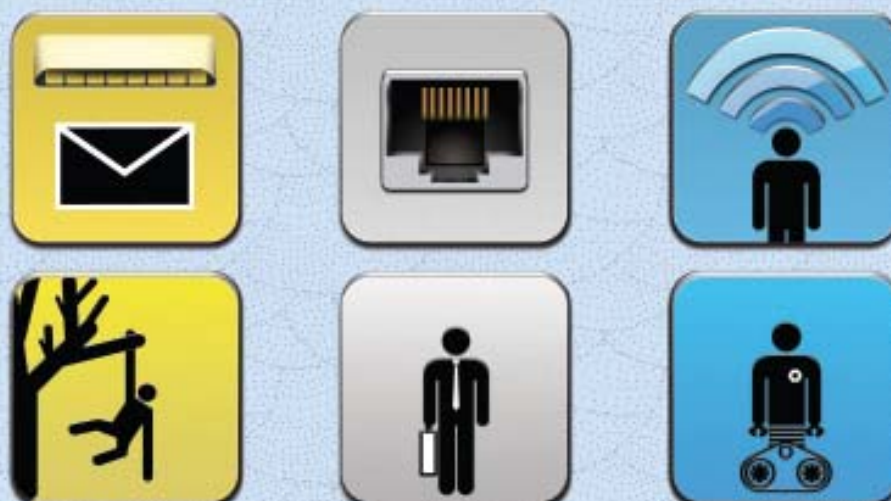


Zbornik 12. mednarodne multikonference
INFORMACIJSKA DRUŽBA - IS 2009

Zvezek A

Proceedings of the 12th International Multiconference
INFORMATION SOCIETY - IS 2009

Volume A



Uredili / Edited by

Marko Bohanec, Matjaž Gams, Vladislav Rajkovič,
Tanja Urbančič, Mojca Bernik, Dunja Mladenič,
Marko Grobelnik, Marjan Heričko,
Urban Kordeš, Olga Markič, Jadran Lenarčič,
Leon Žlajpah, Andrej Gams, Olga S. Fomichova,
Vladimir A. Fomichov, Andrej Brodnik

12. - 16. oktober 2009 / October 12th - 16th, 2009
Ljubljana, Slovenia

Zbornik 12. mednarodne multikonference
INFORMACIJSKA DRUŽBA – IS 2009
Zvezek A

Proceedings of the 12th International Multiconference
INFORMATION SOCIETY – IS 2009
Volume A

Inteligentni sistemi
Vzgoja in izobraževanje v informacijski družbi
Izkopavanje znanja in podatkovna skladišča (SiKDD 2009)
Sodelovanje, programska oprema in storitve v informacijski družbi
Kognitivne znanosti
Robotika
Kognitonika
Druga mini konferenca iz teoretičnega računalništva

Intelligent Systems
Education in Information Society
Data Mining and Data Warehouses (SiKDD 2009)
Collaboration, Software and Services in Information Society
Cognitive Sciences
Robotics
Cognitonics
The Second Mini Conference on Theoretical Computer Science

Uredili / Edited by

Marko Bohanec, Matjaž Gams, Vladislav Rajkovič, Tanja Urbančič, Mojca Bernik,
Dunja Mladenić, Marko Grobelnik, Marjan Heričko, Urban Kordeš, Olga Markič,
Jadran Lenarčič, Leon Žlajpah, Andrej Gams, Olga S. Fomichova, Vladimir A. Fomichov,
Andrej Brodnik

<http://is.ijs.si>

12.–16. oktober 2009 / October 12th–16th, 2009
Ljubljana, Slovenia

Uredniki:

prof. dr. Marko Bohanec
prof. dr. Matjaž Gams

prof. dr. Vladislav Rajkovič
prof. dr. Tanja Urbančič
Mojca Bernik

dr. Dunja Mladenič
Marko Grobelnik

prof. dr. Marjan Heričko

dr. Urban Kordeš
prof. dr. Olga Markič

prof. dr. Jadran Lenarčič
dr. Leon Žlajpah
dr. Andrej Gams

dr. Olga S. Fomichova
dr. Vladimir A. Fomichov

dr. Andrej Brodnik

Založnik: Institut »Jožef Stefan«, Ljubljana
Tisk: Birografika BORI d.o.o.
Priprava zbornika: Mitja Lasič, Jana Krivec
Oblikovanje naslovnice: Ernest Vider - Erc
Tiskano iz predloga avtorjev
Naklada: 150

Ljubljana, oktober 2009

Konferenco IS 2009 sofinancirata
Ministrstvo za visoko šolstvo, znanost in tehnologijo
Institut »Jožef Stefan«

Informacijska družba
ISSN 15819973

CIP - Kataložni zapis o publikaciji
Narodna in univerzitetna knjižnica, Ljubljana

659.2:316.42(082)
659.2:004(082)

MEDNARODNA multikonferenca Informacijska družba (12 ; 2009 ;
Ljubljana)

Zbornik 12. mednarodne multikonference Informacijska družba - IS
2009, 12.-16. oktober 2009 : zvezek A = Proceedings of the 12th
International Multiconference Information Society - IS 2009,
October 12th-16th, 2009, Ljubljana, Slovenia : volume A / uredili,
edited by Marko Bohanec ... [et al.]. - Ljubljana : Institut Jožef
Stefan, 2009. - (Informacijska družba, ISSN 1581-9973)

Vsebina na nasl. str.: Inteligentni sistemi = Intelligent systems ;

Vzgoja in izobraževanje v informacijski družbi = Education in
information society ; Izkopavanje znanja in podatkovna skladišča
(SiKDD 2009) = Data mining and data warehouses (SiKDD 2009) ;
Sodelovanje, programska oprema in storitve v informacijski družbi =
Collaboration, software and services in information society ;
Kognitivne znanosti = Cognitive sciences ; Robotika = Robotics ;
Kognitonika = Cognitonics ; Druga mini konferenca iz teoretičnega
računalništva = The second mini conference on theoretical computer
science

ISBN 978-961-264-010-1

1. Informacijska družba 2. Information society 3. Bohanec, Marko,
1958-
247787264

PREDGOVOR MULTIKONFERENCI INFORMACIJSKA DRUŽBA 2009

V svojem dvanajstem letu ostaja multikonferenca Informacijska družba (<http://is.ijs.si>) ena vodilnih srednjeevropskih konferenc, ki združuje znanstvenike z različnih raziskovalnih področij, povezanih z informacijsko družbo. V letu 2009 smo v multikonferenco povezali rekordnih enajst neodvisnih konferenc. Informacijska družba postaja vedno bolj zapleten socialni, ekonomski in tehnološki sistem, ki je pritegnil pozornost vrste specializiranih konferenc v Sloveniji in Evropi. Naša multikonferenca izstopa po širini in obsegu tem, ki jih obravnava.

Rdeča nit multikonference ostaja sinergija interdisciplinarnih pristopov, ki obravnavajo različne vidike informacijske družbe ter poglobljajo razumevanje informacijskih in komunikacijskih storitev v najširšem pomenu besede. Na multikonferenci predstavljamo, analiziramo in preverjamo nova odkritja in pripravljamo teren za njihovo praktično uporabo, saj je njen osnovni namen promocija raziskovalnih dosežkov in spodbujanje njihovega prenosa v prakso na različnih področjih informacijske družbe tako v Sloveniji kot tujini.

Na multikonferenci bo na vzporednih konferencah predstavljenih 300 referatov, vključevala pa bo tudi okrogle mize in razprave. Referati so objavljeni v zbornikih multikonference, izbrani prispevki pa bodo izšli tudi v posebnih številkah dveh znanstvenih revij, od katerih je ena Informatica, ki se ponaša z 33-letno tradicijo odlične znanstvene revije.

Multikonferenco Informacijska družba 2009 sestavljajo naslednje samostojne konference:

- Inteligentni sistemi
- Kognitivne znanosti
- Kognitonika
- Mondilex
- Robotika
- Rudarjenje podatkov in podatkovna skladišča (SiKDD 2009)
- Sodelovanje, programska oprema in storitve v informacijski družbi
- Soočanje z demografskimi izzivi v Evropi
- Status in vloga tehniških in naravoslovnih poklicev v državi
- Vzgoja in izobraževanje v informacijski družbi
- 2. Minikonferenca iz teoretičnega računalništva 2009

Očitno finančna recesija ni zmanjšala zanimanja za informacijsko družbo; nasprotno, letošnja konferenca je rekordna v več pogledih, recimo glede na število sodelujočih konferenc.

Soorganizatorji in podporniki konference so različne raziskovalne institucije in združenja, med njimi tudi ACM Slovenija. Zahvaljujemo se tudi Ministrstvu za visoko šolstvo, znanost in tehnologijo za njihovo sodelovanje in podporo. V imenu organizatorjev konference pa se želimo posebej zahvaliti udeležencem za njihove dragocene prispevke in priložnost, da z nami delijo svoje izkušnje o informacijski družbi. Zahvaljujemo se tudi recenzentom za njihovo pomoč pri recenziranju.

V letu 2009 sta se programski in organizacijski odbor odločila, da bosta podelila posebno priznanje Slovincu ali Slovenki za izjemen prispevek k razvoju in promociji informacijske družbe v našem okolju. Z večino glasov je letošnje priznanje pripadlo prof. dr. Vladislavu Rajkoviču. Čestitamo!

Franc Solina, predsednik programskega odbora
Matjaž Gams, predsednik organizacijskega odbora

FOREWORD - INFORMATION SOCIETY 2009

In its 12th year, the Information Society Multiconference (<http://is.ijs.si>) continues as one of the leading conferences in Central Europe gathering scientific community with a wide range of research interests in information society. In 2009, we organized record eleven independent conferences forming the Multiconference. Information society displays a complex interplay of social, economic, and technological issues that attract attention of many scientific events around Europe. The broad range of topics makes our event unique among similar conferences. The motto of the Multiconference is synergy of different interdisciplinary approaches dealing with the challenges of information society. The major driving forces of the Multiconference are search and demand for new knowledge related to information, communication, and computer services. We present, analyze, and verify new discoveries in order to prepare the ground for their enrichment and development in practice. The main objective of the Multiconference is presentation and promotion of research results, to encourage their practical application in new ICT products and information services in Slovenia and also broader region.

The Multiconference is running in parallel sessions with 300 presentations of scientific papers. The papers are published in the conference proceedings, and in special issues of two journals. One of them is *Informatica* with its 33 years of tradition in excellent research publications.

The Information Society 2009 Multiconference consists of the following conferences:

- Intelligent Systems
- Cognitive Sciences
- Cognitronics
- Mondilex
- Robotics
- Data Mining and Data Warehouses (SiKDD 2009)
- Collaboration, Software and Services in Information Society
- Demographic Challenges in Europe
- Increasing Interests for Higher Education in Science and Technology
- Education in Information Society
- The Second Mini Conference on Theoretical Computing 2009

Evidently, the economic recession is not affecting Information society; on the contrary, this is a record conference in several terms, e.g. judging from the number of single conferences.

The Multiconference is co-organized and supported by several major research institutions and societies, among them ACM Slovenia, i.e. the Slovenian chapter of the ACM. We would like to express our appreciation to the Slovenian Government for cooperation and support, in particular through the Ministry of Higher Education, Science and Technology.

In 2009, the Programme and Organizing Committees decided to award one Slovenian for his/her outstanding contribution to development and promotion of information society in our country. With the majority of votes, this honor went to Prof. Dr. Vladislav Rajkovič. Congratulations!

On behalf of the conference organizers we would like to thank all participants for their valuable contribution and their interest in this event, and particularly the reviewers for their thorough reviews.

Franc Solina, Programme Committee Chair
Matjaž Gams, Organizing Committee Chair

KONFERENČNI ODBORI

CONFERENCE COMMITTEES

International Programme Committee

Vladimir Bajic, South Africa
Heiner Benking, Germany
Se Woo Cheon, Korea
Howie Firth, UK
Olga Fomichova, Russia
Vladimir Fomichov, Russia
Vesna Hljuz Dobric, Croatia
Alfred Inselberg, Izrael
Jay Liebowitz, USA
Huan Liu, Singapore
Henz Martin, Germany
Marcin Paprzycki, USA
Karl Pribram, USA
Claude Sammut, Australia
Jiri Wiedermann, Czech Republic
Xindong Wu, USA
Yiming Ye, USA
Ning Zhong, USA
Wray Buntine, Finland
Bezalel Gavish, USA
Gal A. Kaminka, Israel
Miklós Krész, Hungary
József Békési, Hungary

Organizing Committee

Matjaž Gams, chair
Mitja Luštrek, co-chair
Lana Jelenkovič
Jana Krivec
Mitja Lasič

Programme Committee

Franc Solina, chair
Viljan Mahnič, co-chair
Cene Bavec, co-chair
Tomaž Kalin, co-chair
Jozsef Györkös, co-chair
Tadej Bajd
Jaroslav Berce
Mojca Bernik
Marko Bohanec
Ivan Bratko
Andrej Brodnik
Dušan Caf
Saša Divjak
Tomaž Erjavec
Bogdan Filipič
Andrej Gams

Matjaž Gams
Marko Grobelnik
Nikola Guid
Marjan Heričko
Borka Jerman Blažič Džonova
Gorazd Kandus
Urban Kordeš
Marjan Krisper
Andrej Kuščer
Jadran Lenarčič
Borut Likar
Janez Malačič
Olga Markič
Dunja Mladenič
Franc Novak
Marjan Pivka
Vladislav Rajkovič

Grega Repovš
Ivan Rozman
Niko Schlamberger
Stanko Strmčnik
Tomaž Šef
Jurij Šilc
Jurij Tasič
Denis Trček
Andrej Ule
Tanja Urbančič
Boštjan Vilfan
David B. Vodušek
Baldomir Zajc
Blaž Zupan
Boris Žemva
Janez Žibert
Leon Žlajpah

KAZALO / TABLE OF CONTENTS

Intelligent Systems	1
PREDGOVOR / PREFACE	3
PROGRAMSKI ODBOR / PROGRAMME COMMITTEE	4
ESTIMATION OF INDIVIDUAL PREDICTION RELIABILITY USING SENSITIVITY ANALYSIS OF REGRESSION MODELS / Zoran Bosnić	7
COGNITIVE COMPLEXITY OF MULTI-CRITERIA GROUP DECISION-MAKING METHODS / Andrej Bregar	11
BEHAVIOUR RANDOMNESS MEASUREMENT AS A PART OF COMPLEX CUSTOMER VALUE INDICATOR / Naděžda Chalupová, Arnošt Motyčka	15
A HYBRID NEURAL NETWORK MODEL FOR SPAM DETECTION / Maria Corduneanu, Carmen Maria Cosoi, Catalin Alexandru Cosoi, Madalin Vlad, Valentin Sgarciu	19
DETECTING ANOMALIES IN SOCIAL NETWORKS USING FRACTAL NETWORKS / Catalin Alexandru Cosoi, Madalin Stefan Vlad, Maria Corduneanu, Carmen Maria Cosoi	22
EVALUATION OF POPULAR FEATURE RANKING ALGORITHMS IN MICROARRAY ANALYSIS / Mario Gorenjak, Mateja Bajgot, Biljana Pejčić, Andrej Sovec, Gregor Štiglic	26
EQUATION-BASED MODELS OF OILSEED RAPE POPULATION DYNAMICS DEVELOPED FROM SIMULATION OUTPUTS OF AN INDIVIDUAL-BASED MODEL / Aneta Ivanovska, Graham Begg, Ljupčo Todorovski, Sašo Džeroski	30
EXPLANATION OF REGRESSION DECISIONS BY ANALOGY WITH THE EXPLANATION IN CLASSIFICATION / Julian Klauser, Igor Kononenko	34
REGRESSION AS COST-SENSITIVE CLASSIFICATION / Egon Kocjan, Igor Kononenko	38
PROBLEM PRED-TESTNE CENILKE NA PRIMERU EQ5D / Marko Ogorevc	42
COMPARISON OF APPROACHES FOR ESTIMATING RELIABILITY OF INDIVIDUAL CLASSIFICATION PREDICTIONS / Darko Pevec, Zoran Bosnić, Igor Kononenko	46
USING STOCHASTIC MODEL FOR IMPROVING HTTP LOG DATA PRE-PROCESSING / Marko Požnenel, Viljan Mahnič, Matjaž Kukar	50
A FUZZY EXPERT SYSTEM TO ENFORCE NETWORK SECURITY POLICY / Bel G. Raggad, Azza Mastouri, Manal Mastouri	54
PORTABILITY OF USER MODELS WITHIN ADAPTIVE WEB-BASED SYSTEMS / Magdalena Raszková, Arnošt Motyčka	58
GENE REGULATORY NETWORKS INFERENCE USING GRAPHICAL GAUSSIAN MODELS / Blagoj Ristevski, Suzana Loskovska	62
TIME SERIES FORECASTING USING MACHINE LEARNING METHODS / Michael Stencl, Ondrej Popelka, Jiri Stastny	66
MACHINE LEARNING FOR OBJECT ESTIMATION USING HIERARCHICAL CRITERIA SYSTEM / Andrey Styskin	70
COLLECTIVE INTELLIGENCE AND ORGANIZATIONS' CONSCIOUSNESS / Viljem Tisnikar	74
APPLYING DATA ENVELOPMENT ANALYSIS FOR INCREASING OPERATIONAL EFFICIENCY IN PROJECT MANAGEMENT / Pavel Tubin	78
POTENTIAL BENEFITS OF USING WEB SERVICES IN CRM SYSTEMS / Jan Turčinek	82
PROJECT SELF-EVALUATION METHODOLOGY: THE HEALTHREATS PROJECT CASE STUDY / Martin Žnidaršič, Marko Bohanec, Nada Lavrač, Bojan Cestnik	85
INTELIGENTNI SISTEM ZA NADZOR OBJEKTOV / Erik Dovgan, Rok Piltaver, Matjaž Gams	89
SE SPLAČA PREMISLITI GLOBLJE? / Matjaž Gams:	93
GLAJENJE TRAJEKTORIJ GIBANJA ČLOVEŠKEGA TELESA ZAJETIH Z RADIJSKO TEHNOLOGIJO / Boštjan Kaluža, Erik Dovgan:	97
ISKANJE VZORCEV V ZAPOREDJU DOGODKOV / Jana Krivec:	101
IZBOLJŠEVANJE PREPOZNAVANJA AKTIVNOSTI IZ POLOŽAJEV ZNAČK / Mitja Luštrek:	105
STROJNA KLASIFIKACIJA SPLETNIH STRANI PO TEMAH / Domen Marinčič:	109
TOWARDS ROBUST RULE ENGINE FOR CLASSIFYING HUMAN POSTURE / Violeta Mircevska:	112
ANALIZA DELOVANJA VIRTUALNEGA SVETOVALCA / Matej Ožek, Matjaž Gams, Jana Krivec:	116
ZAZNAVANJE NENAVIDNEGA OBNAŠANJA S SISTEMOM ZA LOCIRANJE V REALNEM ČASU IN MEHKO LOGIKO / Rok Piltaver:	120
MOVEMENT-BASED AUTOMATIC DISEASE RECOGNITION / Bogdan Pogorelc:	124
IDENTIFIKACIJA GLASOV IN SODNO IZVEDENSTVO V KAZENSKEM POSTOPKU / Tomaž Šef:	128
PATOLOGIJA MINIMIN PREISKOVANJA / Aleš Tavčar:	132

PROBLEM TRANSFORMATION METHODS FOR MULTIGENRE WEB PAGES CLASSIFICATION / Vedrana Vidulin:	136
Education in Information Society.....	141
PREDGOVOR	143
PREFACE.....	144
PROGRAMSKI ODBOR / PROGRAMME COMMITTEE	145
UPORABA IKT PRI POUKU (TUJEGA JEZIKA), NJENE PREDNOSTI IN (MOŽNE) SLABOSTI TER NUJNI POGOJI ZA KAKOVOSTNO DELO Z IKT / Jelka Bajželj.....	147
MODEL ZA OCENO VPLIVA STALNEGA STROKOVNEGA IZOBRAŽEVANJA NA KAKOVOST UČITELJA / Sašo Bizant	148
EVALVACIJA IZOBRAŽEVALNEGA PROCESA S POUČENJEM NA VOJAŠKIH VSEBINAH / Liliana Brožič, Dušan Sušnik	149
UPORABA INTERAKTIVNE TABLE PRI MATEMATIKI V PRVEM TRILETJU OSNOVNE ŠOLE / Urška Bučar	150
MEDOSEBNA VLOGA RAVNATELJA - MANAGERJA V PROCESU DELA IN FUNKCIJI HUMANISTIČNO-ANTROPOCENTRIČNEGA MANAGEMENTA ČLOVEŠKIH VIROV / Bojan Burgar, Jože Florjančič, Mojca Bernik.....	151
POUK MATEMATIKE V OSNOVNI ŠOLI Z UPORABO E-GRADIV / Nevenka Colja	152
GOOGLE APPS - OZADJE, IMPLEMENTACIJA IN UPORABA / Dejan Cvitkovič	153
PROSTO DOSTOPNI IZOBRAŽEVALNI VIRI V E-IZOBRAŽEVANJU / Dejan Dinevski, Samo Fošnarč, Tanja Arh	154
OBRAVNAVA UMETNOSTNEGA BESEDILA – PRAVLJICE S POMOČJO E – GRADIV V 6. RAZREDU / Miroslava Fon	155
IZZIVI NOVIH TEHNOLOGIJ IN ŠOLA BODOČNOSTI / Ivan Gerlič	156
RAZVIJANJE MEHKIH ZNANJ NA TEHNIČNIH FAKULTETAH: IZKUŠNJE S ŠTUDENTSKIM DELOM NA PROJEKTIH / Franc Gider, Tanja Urbančič	157
PREDNOSTI BLOKOVNEGA PROGRAMIRANJA ROBOTOV V OSNOVNI ŠOLI / Milan Hlade	158
NADALJNI KORAKI V RAZVOJU E-IZOBRAŽEVANJA V SLOVENSKEM ŠOLSKEM PROSTORU / Boris Horvat, Matija Lokar, Primož Lukšič, Damijan Omerza, Alen Orbanič	159
NEKATERI STRUKTURNI IN KULTURNI PROBLEMI PRI UVAJANJU E-IZOBRAŽEVANJA. Z NAKAZANIMI REŠITVAMI / Marko Ivanišin.....	160
UPORABA PROGRAMA MICROSOFT WORD PRI TRETJEŠOLCIH / Alenka Kastelic.....	161
UČENJE PREKO IGRE DO SPOZNAVANJA RAČUNALNIKA / Mojca Kogoj	162
ALI UPORABA MULTIMEDIJE IZBOLJŠA UČINKOVITOST IZOBRAŽEVANJA? / Darko Korošec.....	163
ELEKTRONSKI KARIERNI PORTFOLIJ - KONCEPT E-ORODJA, KI PODPIRA KARIERNI RAZVOJ POSAMEZNIKA / Danilo Kozoderc.....	164
UPORABA PROGRAMA ECLIPSECROSSWORD V UČNEM PROCESU PRVEGA VZGOJNO- IZOBRAŽEVALNEGA OBDOBJA / Irena Kresevič.....	165
KAKO OPOGUMITI STAREJŠE OSEBE ZA UPORABO IKT? / Julija Lapuh Bele, Boštjan Jarc, David Rozman	166
KAKO PRIPRAVITI UČNA E-GRADIVA? / Matija Lokar.....	167
PRIMER GRADIVA ZA INTERAKTIVNO TABLO PRI POUKU SLOVENŠČINE / Tatjana Lotrič Komac, Tina Žagar Pernar	168
SLOVENŠČINA NA DALJAVO / Tatjana Lotrič Komac	169
NAUK – NAPREDNE UČNE KOCKE ZA UČITELJE / Primož Lukšič, Matija Lokar, Boris Horvat.....	170
POUČEVANJE METODE SCRUM V SODELOVANJU S PODJETJEM ZA RAZVOJ PROGRAMSKE OPREME / Viljan Mahnič, Strahil Georgiev, Tomo Jarc	171
ODLOČITVENI MODEL ZA IZBIROŠOLSKIH IN OBŠOLSKIH DEJAVNOSTI OTROK / Matea Curkova	172
INFORMATIZACIJA POUKA KLAVIRJA: IZZIV PRIHODNOSTI ALI UTOPIJA? / Lorena Mihelač	173
UPORABA E-GRADIV ZA UČENJE DOMA IN NA DOMU / Jožica Mlakar Broder	174
E-KEMIJA V 8. RAZREDU – IZDELAVA E-GRADIVA / Tomaž Pavlakovič, Sonja Malnarič	175
MERILEC SRČNEGA UTRIPA KOT SREDSTVO IKT PRI ŠPORTNI VZGOJI / Rok Pekolj	175
VLOGA ODLOČITVENEGA MODELA PRI UGOTAVLJANJU VSEBNOSTI TEŽKIH KOVIN V LASEH / Aleksandra Debevec, Marjanca Pograjc Debevec.....	176
ASSURING THE STUDENTS TO WORK INDIVIDUALLY AT HOME USING MOODLE - VIRTUAL LEARNING ENVIRONMENT / Zdenko Potočar	177
INOVATIVNO UČENJE IN POUČEVANJE PRI POUKU GEOGRAFIJE / Andreja Prezelj.....	178

POSTOPNO CELOSTNO UVAJANJE E-IZOBRAŽEVANJA V SPLOŠNI GIMNAZIJI / Tanja Mastnak, Peter Purg, Alenka Budihna	179
OBRAVNAVA PRAVLJICE V PRVEM RAZREDU / Stanka Rakar	180
NAVIGACIJA IN ZUMIRANJE NEVIDNIH POSLOVNIH PROCESOV V MEHATRONSKI INFORMATIKI / Gorazd Rakovec	181
TIMSKO, MEDPREDMETNO POUČEVANJE OB PODPORI IKT / Irena Rakovec Žumer	182
»NOVI« ČLOVEK KOT UČENEC - POMEN PERCEPCIJE IN ZAVESTI ZA UČENJE, PODPRTO Z INFORMACIJSKO TEHNOLOGIJO / Vanda Rebolj	183
UPORABA INFORMACIJSKO-KOMUNIKACIJSKE TEHNOLOGIJE PRI POUČEVANJU TUJEGA JEZIKA V VZGOJNO IZOBRAŽEVALNEM ZAVODU / Ribič Marko	184
MOŽNOST UPORABE PROGRAMA TUX PAINT V DRUGEM RAZREDU / Darja Rijavec	185
PREGLED IN ANALIZA NASTAJANJA, PREIZKUŠANJA IN UPORABE E-GRADIV PRI POUKU / Damjana Šajne, Tanja Urbančič, Iztok Arčon	186
COURSELAB - PREPROSTO ORODJE ZA IZDELAVO E-GRADIV / Peter Škarja, Branislav Šmitek	187
ODLOČITVENI MODEL ZA IZBIRO UČENCA ZA NAGRADO ŠOLE / Magda Slokar Čevdek	188
KAKOVOST ZNANJA PRIDOBLENA Z RAZLIČNIMI NAČINI IZVEDB LABORATORIJSKEGA DELA / Andreja Špernjak, Andrej Šorgo	189
ELEKTRONSKA BAZA PODATKOV O UČENCIH S POSEBNIMI POTREBAMI V OSNOVNI ŠOLI / Amalija Stiplovšek	190
POUČNI RAČUNALNIŠKI PROGRAMI ZA VRTEC / Jelena Stojmenovič	191
INTERAKTIVNA TABLA IN INTERAKTIVNOST PRI POUKU MATEMATIKE NA PREDMETNI STOPNJI OŠ / Jožica Štrajhar	191
AKTIVNE OBLIKE ŠTUDIJA IN VRSTNIŠKO OCENJEVANJE V VISOKEM ŠOLSTVU / Mateja Strnad, Irena Nančovska Šerbec, Jože Rugelj	192
UPORABA E-GRADIV ZA NOVE SREDNJEŠOLSKE UČITELJE / Gašper Strniša	193
UČNA URA Z INTERAKTIVNIMI DEMONSTRACIJAMI / Jože Štrucl	193
POUČEVANJE (SLOVENŠČINE) NA DALJAVO / Polona Tomac Stanojev	194
IKT – MOST MED ŠOLO IN STARŠI / Andreja Vehar Jerman	195
MODEL UGOTAVLJANJA USTVARJALNE UČINKOVITOSTI PODJETIJ / Barbka Vidmar	196
IKT V IZOBRAŽEVANJU ZA TRAJNOSTNI RAZVOJ / Srečo Zakrajšek	197
ALI PRIDOBIVATI ZNANJE S POMOČJO UPORABE IKT ALI S KLASIČNIMI PEDAGOŠKIMI METODAMI IN OBLIKAMI DELA? / Mojca Žepič	198
Data Mining and Data Warehouses (SiKDD 2009)	199
PREDGOVOR / PREFACE	201
ENRYCHER – SERVICE ORIENTED TEXT ENRICHMENT / Tadej Štajner, Delia Rusu, Lorand Dali, Blaž Fortuna, Dunja Mladenič, Marko Grobelnik	203
ENRYCHER – SERVICE ORIENTED TEXT ENRICHMENT / Tadej Štajner, Delia Rusu, Lorand Dali, Blaž Fortuna, Dunja Mladenič, Marko Grobelnik	203
LEARNING EVENT TEMPLATES ON NEWS ARTICLES / Mitja Trampuš, Dunja Mladenic	207
USING ENUMERATIONS FOR WORD CLUSTERING / Lorand Dali, Nada Lavrač	211
SEMI-AUTOMATIC ONTOLOGY EXTENSION USING TEXT MINING / Inna Novalija, Dunja Mladenič	214
CONTEXTUALIZED VISUALIZATION OF ONTOLOGIES AND ONTOLOGY NETWORKS / Boštjan Pajntar, Dunja Mladenič, Marko Grobelnik	218
PROBABILISTIC TEMPORAL PROCESS MODEL FOR KNOWLEDGE PROCESSES: HANDLING A STREAM OF LINKED TEXT / Marko Grobelnik, Dunja Mladenic, Jure Ferlež	222
EXPLORATORY ANALYSIS OF PRESS ARTICLES ON KENYAN ELECTIONS: A DATA MINING APPROACH / Senja Pollak	228
TEXT MINING AND KNOWLEDGE DISCOVERY WITH ONTOGEN 2.0 / Mladen Tomaško	232
AN IMPLEMENTATION OF THE PATHFINDER ALGORITHM FOR SPARSE NETWORKS AND ITS APPLICATION ON TEXT NETWORKS / Anže Vavpetič	236
EXPERIMENTS WITH SATURATION FILTERING FOR NOISE ELIMINATION FROM LABELED DATA / Borut Sluban, Nada Lavrač, Dragan Gamberger, Andrej Bauer	240
Collaboration, Software and Services in Information Society	245
PREFACE	247
PROGRAMSKI ODBOR / PROGRAMME COMMITTEE	248
SIMPLE SOLUTION FOR ONTOLOGY-BASED MAPPING BETWEEN V-MODELL XT AND SELECTED SOFTWARE DEVELOPMENT METHODOLOGIES / Peter Butka	249

MODEL-DRIVEN ENGINEERING AND AN EXAMPLE OF ITS INTRODUCTION / Tomaž Lukman, Giovanni Godena.....	253
A MODEL BASED CODE GENERATION SUPPORT FOR DEVELOPING A PRESENTATION LOGIC / Jan Kryštof, David Procházka, Arnošt Motyčka	257
THE USE OF METAPHORS IN THE DEVELOPMENT OF INFORMATION SYSTEMS / Saša Kuhar, Marjan Heričko	261
ARCHITECTURE FOR SOFTWARE METRICS REPOSITORY / Črt Gerlec, Aleš Živkovič.....	265
SOLUTION REPRESENTATION ANALYSIS FOR THE EVOLUTIONARY APPROACH OF THE ENTITY REFACTORING SET SELECTION PROBLEM / Camelia Chisăliță-Crețu	269
ORGANIZING WEB SERVICES INTERFACES AS A BASIS FOR EFFICIENT SERVICE GOVERNANCE / Aleš Frece	273
INTEGRATION OF FULL-TEXT SEARCH WITH WEB MAPPING SERVICES / David Procházka, Jan Kryštof, Arnošt Motyčka.....	277
THE ARCHITECTURAL DESIGN OF A TOOL FOR TESTING WORKFLOW-BASED APPLICATIONS / Uroš Goljat, Marjan Heričko	281
UTILIZING PROCESS MODELING TO SUPPORT THE COLLABORATIVE COMMUNICATION OF AUTHORITIES IN THE MANAGEMENT OF DISASTER SITUATIONS / Jari Soini, Petri Linna, Hannu Jaakkola.....	285
TOWARDS ADAPTIVE SERVICE-CENTRED APPLICATIONS / Jože Pfeifer	290
DYNAMIC SERVICE BUSINESS MODELS: A PROPOSAL FOR UNIFIED SERVICE PRICING FRAMEWORK / Kristjan Košič, Reinhard Bernsteiner, Marjan Heričko	294
SUCCESS FACTORS AND BARRIERS OF KNOWLEDGE MANAGEMENT – AN EMPIRICAL ANALYSIS OF A SHAREPOINT 2007 IMPLEMENTATION / Michael Amberg, Michael Reinhardt, Jiangping Weng.....	298
THE TRUE VALUE OF AN E-LEARNING SYSTEM THROUGH THE STUDENT'S EYE / Boštjan Šumak, Maja Pušnik, Marjan Heričko.....	303
FUNCTIONAL HORIZONTAL NETWORK MARKETPLACES – A POSSIBLE SOLUTION FOR SERBIAN MARKET / Zoran Jankovič, Mirjana Ivanović, Zoran Budimac	307
USING GEOFENCING TO OVERCOME SECURITY CHALLENGES IN WIRELESS NETWORKS: PROOF OF CONCEPT / Anthony C. Ijeh, Allan J. Brimicombe, David S. Preston, Chris O. Imafidon.....	311
FREE INTERNET ACCESS USING THE ANDROID PLATFORM / Jernej Huber	315
Cognitive Sciences.....	319
PREDGOVOR	321
DEMENCA – RAZPAD UMA IN POGREZ V BLAGODEJNO POZABO / Pirtošek Zvezdan.....	323
RACIONALNO ODLOČANJE IN ČUSTVA / Markič Olga	325
FENOMENOLOGIJA ODLOČANJA / Kordeš Urban.....	329
SE ODLOČAMO GENETSKO ALI PRIVZGOJENO – ANALIZA POSILSTVA? / Gams Matjaž.....	333
PRIDI K MENI: O ODLOČANJU, KAIRÓSU IN TRENUTKIH SREČANJA V PSIHOTERAPIJI / Možina Miran.....	337
OSEBNOSTNI IN KONTEKSTNI DEJAVNIKI TER PARADIGME ODLOČANJA / Tancig Simona	343
SKUPINSKO ODLOČANJE KOT AKTUALIZIRANJE DELOVANJSKIH POTENCIALOV / Ule Andrej.....	347
RAČUNALNIK IN ODLOČANJE: ODLOČITVENI MODELI IN SISTEMI ZA PODPORO PRI ODLOČANJU / Bohanec Marko	351
DO KOD SEŽEJO MATEMATIČNI MDELI V SITUACIJAH ODLOČANJA / Knap Žiga	355
NIKOTINSKA ZASVOJENOST GOSPODA JONESA- Odločanje zgodovinarja v procesu raziskovanja / Ratej Mateja	357
SITUACIJA “ODLOČANJE” IN NJENA REPREZENTACIJA V ZNANSTVENIH TEKSTIH / Bazhenova Elena, Marija Kotyurova	360
VPLIV MAGNETIZMA NA OBČUTLJIVOST OČI - JE V OZADJU MAGNETNI ČUT? / Avbelj Viktor.....	362
AN INTER-DISCIPLINARY SURVEY OF CURRENT STUDIES ON THE NATURE OF CONSCIOUSNESS / Daffern Thomas C.	365
POVEZANOST IMPULZIVNOSTI Z AFEKTIVNIMI DIMENZIJAMI TEMPERAMENTA MED SPOLOMA / Dolenc Barbara, Šprah Lilijana.....	371
NOVI POGLEDI NA DELOVANJE MOŽGANOV: OD SUBMOLEKULSKE DO GLOBALNE RAVNI / Plankar Matej.....	375
VLOGA MOTIVACIJE PRIBLIŽEVANJA IN UMIKA PRI KOGNITIVNI KONTROLI EMOCIONALNIH DRAŽLJAJEV / Šprah Lilijana, Novak Tatjana.....	379
ON CONSCIOUSNESS; INSIGHTS FROM INTUITION, REFLECTION AND DIALOGUE / Žerovnik Eva	383

K VPRAŠANJU O NACIONALNI SLIKI SVETA / Shchukina Irina	387
RAZMERJE DO VODENJA IN IZBRANE SOCIALNO VREDNOTNE ORIENTACIJE PRI RAZLIČNIH KULTURAH V SLOVENIJI / Mihajlović Slađana	390
Robotics.....	395
PREDGOVOR	397
EKSPERIMENTALNA MOBILNA ROBOTSKA PLATFORMA / Peter Čepon, Roman Kamnik, Jernej Kuželički, Tadej Bajd, Marko Munih	399
DESIGN OF A CUSTOM ELECTRONICS DRIVER FOR DIELECTRICS ELASTOMER ACTUATORS / Mitja Babič, Rocco Vertechy, Giovanni Berselli, Jadran Lenarčič.....	403
ROBOTSKE PROGRAMSKE STRUKTURE V PROGRAMIRANJU HUMANOIDNIH ROBOTOV / Andrej Kos.....	407
REGULACIJA TEŽIŠČA ROBOTA PRI VIZUALNO-MOTORIČNEM VODENJU / Blaž Hajdinjak, Jan Babič.....	411
RITMIČNO VODENJE NIHALA Z UPORABO NELINEARNEGA DINAMIČNEGA SISTEMA / Tadej Petrič, Andrej Gams, Leon Žlajpah	415
POSPLOŠEVANJE PERIODIČNIH GIBANJ ZAPISANIH Z NELINEARNIMI DINAMIČNIMI SISTEMI / Andrej Gams, Aleš Ude	419
ROBOTSKA REHABILITACIJA Z NAVIDEZNO RESNIČNOSTJO IN PSIHOFIZIOLOŠKIMI MERITVAMI / Domen Novak, Jaka Zihel, Andrej Olenšek, Janez Podobnik, Matjaž Mihelj, Marko Munih	423
Cognitonics	427
PREFACE / PREDGOVOR	429
EDITORS AND PROGRAM CHAIRS / UREDNIKA	429
COGNITONICS AS AN ANSWER TO THE CHALLENGE OF TIME / Olga S. Fomichova, Vladimir A. Fomichov	431
AN INFORMATION SYSTEM IN SCHOOL FOR A RISK MANAGEMENT OF THE INTERNET: PREVENTING CYBERBULLYING WITHOUT PROHIBITIONS / Hirohiko Yasuda.....	435
A NEW MODEL FOR ONLINE READING COMPREHENSION RESEARCH / S. Ottaviano, A. Chifari, L. Seta, G. Chiazese, G. Merlo, M. Allegra	440
LANGUAGES AND LANGUAGE: THE GREEK CASE / Maria Bontila, Vassilios Dagdilelis.....	444
AESTHETICS AND LOGIC, THE TWO MAIN BRANCHES OF ONE SINGLE TREE / Nicole Szendy	448
AN EXTENDED CONCEPT OF MULTI-MEDIA AND ITS ROLE IN CREATIVITY IN BASIC EDUCATION / Gaba Tsayang, Dimitar M. Totev	452
SCHOOL READINESS: AN ITALIAN TOOL WITH A MULTIFACTORIAL APPROACH FOR ACADEMIC SUCCESS / Daniela Miazza, Maria Assunta Zanetti	456
EXPANDING MENTAL OUTLOOK BY USING CONCEPT MAPS / Dumitru Dan Burdescu, Marian Cristian Mihaescu, Bogdan Logofatu, Costem Marian Ionascu	460
ACADEMIC-SCHOOL READINESS: AN ITALIAN TRAINING / Daniela Miazza, Maria Assunta Zanetti	464
COGNITONICS: A SOPHISTICATED LOOK AT SOCIALIZATION VIA VOGUE / Olga S. Fomichova, Anna V. Molyukova.....	467
The Second Mini Conference on Theoretical Computer Science.....	471
PREDGOVOR	473
PREFACE.....	474
LOWER BOUND ON ON-LINE BIN-PACKING ALGORITHMS / Gábor Galambos, János Balogh, József Békési	475
A FLEXIBLE METHOD FOR DRIVER SCHEDULING IN PUBLIC TRANSPORTATION / Attila Tóth	475
DESIGN SPACE EXPLORATION FOR EMBEDDED PARALLEL SYSTEM-ON-CHIP PLATFORMS USING MODEFRONTIER / C. Kavka, L. Onesti, P. Avasare, G. Vanmeerbeek, M. Wouters, H. Posadas.....	476
IMPROVED ANALYSIS OF AN ALGORITHM FOR THE COUPLED TASK PROBLEM WITH UET JOBS / József Békési, Gábor Galambos, Marcus Oswald, Gerhard Reinelt	477
A FRAMEWORK FOR A FLEXIBLE VEHICLE SCHEDULING SYSTEM / David Paš, József Békési, Miklós Krész, Andrej Brodnik.....	477
MEDIAL AXIS APPROXIMATION OF SIMPLE POLYGONS / Gregor Smogavec, Borut Žalik.....	478
AN EFFICIENT GRAPH REDUCTION METHOD / Miklós Krész, Miklós Bartha SPEAKER TRACKING IN BROADCAST NEWS: A CASE STUDY / Janez Žibert	479
SEZNAM REFERATOV V ŠTUDENTSKI SEKCIJI.....	480

<i>Indeks avtorjev / Author index</i>	481
----------------------------------------------------	------------

Zbornik 12. mednarodne multikonference
INFORMACIJSKA DRUŽBA – IS 2009

Proceedings of the 12th International Multiconference
INFORMATION SOCIETY – IS 2009

Inteligentni sistemi

Intelligent Systems

Uredila / Edited by

Marko Bohanec, Matjaž Gams

<http://is.ijs.si>

14.–15. oktober 2009 / October 14th–15th, 2009
Ljubljana, Slovenia

PREDGOVOR

Inteligentni sistemi imajo jasen trend rasti tako v realnem življenju kot na naši konferenci. Ta trend je nasproten ekonomski recesiji v 2009, saj se je število prispevkov na konferenci občutno povečalo. Stalnica je tudi pozitiven trend, ko inteligentni programi čedalje uspešneje opravljajo naloge inteligentnih pomočnikov in hkrati postajajo tudi bistveno bolj komunikativni v smislu govora in mimike. Inteligentni sistemi postajajo del naše vsakdanjosti.

Konferenca *Inteligentni sistemi* v letu 2009 ostaja mednarodna in vseslovenska hkrati. Prispevki so tako v slovenskem kot angleškem jeziku. Letos posebej izstopajo pristopi, ki temeljijo na izdelavi različnih modelov, analizi podatkov in strojnem učenju. Predstavljene so tudi konkretne aplikacije na različnih področjih, na primer pri upravljanju podjetij, vodenju projektov, na področjih računalniških omrežij in v zdravstvu.. Ponovno so posebej razveseljivi kakovostni prispevki mladih avtorjev. Prispevki dokazujejo uspešnost inteligentnih sistemov pri reševanju zahtevnih praktičnih problemov.

Na letošnji konferenci *Inteligentni sistemi 2009* je predstavljeno skoraj 30 prispevkov kljub poostreni recenziji in posledično večjemu številu zavrnjenih prispevkov. Prispevki so bili recenzirani s strani dveh anonimnih recenzentov. Oblikovne pripombe sva prispevala tudi predsednika konference.

Marko Bohanec in Matjaž Gams, predsednika konference

PREFACE

Contrary to the economic recession in 2009, the area of intelligent systems is still growing and gaining more and more attention. The number of papers increased again this year. Not only that intelligent systems are becoming more and more advanced intelligent assistants, they are improving their communication skills in terms of speech and expression. Intelligent systems are becoming part of our everyday life.

The conference *Intelligent Systems 2009* remains a national and international event and presents papers written in both English and Slovenian languages. In this year, the focus is on approaches based on modeling, data analysis and machine learning. Presented are applications in various problem domains, including the management of companies, projects, computer networks and health care. Particularly promising are high-quality contributions of young authors. The papers confirm the usefulness and effectiveness of intelligent systems in solving and supporting difficult real-life problems.

The Proceedings of Intelligent Systems 2009 include almost 30 papers, in spite of stronger quality criteria and consequently more rejected papers than in the previous years. The submitted papers have been reviewed by two anonymous reviewers. Some additional suggestions for improvements were also provided by the chairmen of the conference.

Marko Bohanec and Matjaž Gams, Conference Chairs

PROGRAMSKI ODBOR / PROGRAMME COMMITTEE

Marko Bohanec, *predsednik*

Institut Jožef Stefan

Matjaž Gams, *predsednik*

Institut Jožef Stefan

Tomaž Banovec

Statistični urad Republike Slovenije

Cene Bavec

Univerza na Primorskem, Fakulteta za Management Koper; IBM

Jaro Berce

Univerza v Ljubljani, Fakulteta za družbene vede

Marko Bonač,

ARNES

Ivan Bratko

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, IJS

Dušan Caf

Telekom Slovenije

Aleš Dobnikar

Center Vlade RS za informatiko

Bogdan Filipič

Institut Jožef Stefan

Nikola Guid

Univerza v Mariboru, Fakulteta za elektrotehniko računalništvo in informatiko

Borka Jerman Blažič

Institut Jožef Stefan

Tomaž Kalin

Ministrstvo za informacijsko družbo

Thiemo Krink

University of Aarhus

Marjan Krisper

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko

Marjan Mernik

Univerza v Mariboru, Fakulteta za elektrotehniko računalništvo in informatiko

Vladislav Rajkovič
Univerza v Mariboru, Fakulteta za organizacijske vede

Ivan Rozman
Univerza v Mariboru, Fakulteta za elektrotehniko računalništvo in informatiko

Niko Schlamberger
Informatika

Tomaž Seljak
Ministrstvo zašolstvo, znanost in šport

Marin Silič
Center Vlade RS za informatiko

Peter Stanovnik
Institut za ekonomska raziskovanja

Peter Tancig
Združanje raziskovalcev Slovenije

Pavle Trdan
Lek

Iztok Valenčič
Nova Kreditna Banka Maribor d.d.

Vasja Vehovar
Univerza v Ljubljani, Fakulteta za družbene vede

Boštjan Vilfan
Univerza v Ljubljani, Fakulteta za računalništvo in informatiko

ESTIMATION OF INDIVIDUAL PREDICTION RELIABILITY USING SENSITIVITY ANALYSIS OF REGRESSION MODELS

doctoral dissertation (extended abstract)

Zoran Bosnić

Laboratory for Cognitive Modeling

Faculty of Computer and Information Science

Tržaška 25, 1000 Ljubljana, Slovenia

Tel: +386 1 4768459; fax: +386 1 468459

e-mail: zoran.bosnic@fri-uni-lj.si

ABSTRACT

This paper is an extended abstract of doctoral dissertation which discusses the estimation of reliability for the individual predictions of regression models. The estimation of reliability of individual predictions, as opposed to the evaluation of the whole predictive model, provides an important aspect of prediction quality which may be strongly beneficial in risk sensitive applications of machine learning. The dissertation compares the performance of 8 such reliability estimates and proposes a methodology to select the best performing estimate for a given domain and predictive model. The performance of the estimates is evaluated on a large number of standard benchmark domains, as well on a real domain. The results have shown that the variance of bagged predictions performs best as reliability estimator, and that the both proposed procedures for automatic selection of the best performing estimate allow achievement of better results than by using each of the individual estimates only.

1 INTRODUCTION

The dissertation [1, 2, 3] discusses the reliability estimation of individual regression predictions in the field of supervised learning. In contrast to the average measures for the evaluation of model accuracy (e.g. mean squared error), the reliability estimates for individual predictions can provide additional information which is beneficial for evaluating the usefulness of the predictions. This additional information can also provide decision support to the users of the prediction systems, based on which they can decide on the corresponding consequential actions (prescribe a therapy, use the autopilot, etc.).

Measuring the expected prediction error is especially important in the risk-sensitive areas where acting upon predictions may have financial or medical consequences (e.g. medical diagnosis, stock market, navigation, control applications). In such areas, appropriate *local* accuracy measures may provide additional necessary information

about the prediction confidence. The difference between traditional approach to the model evaluation and the reliability estimation is illustrated in Table 1. The table also shows an additional advantage of the reliability estimates for individual predictions: They are computed for each particular example in contrast to averaged model estimates which require a separate set of test examples.

Table 1: Comparison of reliability estimates for a model as a whole and reliability estimates for individual predictions

	Purpose	Calculation
<i>Reliability estimate for the whole regression model</i>	one global estimate for a whole model	test examples required for computation
<i>reliability estimates for individual predictions</i>	one reliability estimate for each individual prediction	does not require separate set of test examples

2 CATEGORIZATION OF THE RELATED APPROACHES

In our work, we use term *reliability estimate* to denote any quantity that estimates a quality of a regression prediction. In the related work, the reliability estimates have therefore appeared as either accuracy estimates or error estimates. Depending on how the reliability estimates are implemented, the dissertation separates them into the following two groups:

- (i) model-dependent estimates which exploit the properties of a particular models (e.g. the number of support vectors [4], Lagrange multipliers in the SVM optimization procedure [5], splits in a regression tree, etc.) and
- (ii) model-independent reliability estimates, which exploit the general properties of the supervised learning framework (e.g. changing the learning set [6], etc.).

Besides providing an overview of the related work for the both of the above directions, the dissertation also summarizes and defines various terms which are used in this field (e.g. reliability, sensitivity, stability, confidence, credibility, etc.). The definitions of the terms are systematically shown as a dictionary, which represents an unification of the terminology in the field. The excerpt of this dictionary is shown in Table 2.

Table 2: The dictionary excerpt of the most relevant terms in the area of reliability estimation.

accuracy estimate	One of the aspects of prediction reliability. An estimate which positively correlates with the prediction accuracy or negatively correlates with the prediction error. Similar to confidence, but more general since it does not have probabilistic interpretation (it can take values from an arbitrary real interval and need not be limited to [0, 1]).
confidence	Probabilistically expressed accuracy estimate for a given prediction. Value of prediction confidence therefore represents the probability of its accurateness. It is based on an assumed probability distribution and in classification it can be also defined as $1 - p_2$, where p_2 denotes the probability of the second most probable class [5].
error estimate	One of the aspects of prediction reliability. An estimate which positively correlates to the prediction error. It does not have a probabilistic interpretation and can therefore take values from an arbitrary interval of real numbers. It may be implemented as inverted accuracy estimate.
reliability	A general notion in engineering, denoting the ability of a system or a component to perform its required functions under stated conditions for a specified period of time. In machine learning, we can define reliability as any qualitative property or ability of the system to perform its important task. It is quantitatively estimated with reliability estimate, which can be either positive (accuracy, availability, responsiveness, etc.) or negative indicator (inaccuracy, downtime rate, etc.)
sensitivity	Quantitatively expressed dependence between the changes in system parameters and structure, and the critical aspects of the system operation.
transduction	A term denoting <i>reasoning from particular to particular</i> . The transductive reasoning can be used to construct reliability estimates, which express the probability of how the newly labeled example fits into the distribution of all given examples. Such application may be the estimation of prediction reliability, as in [7].

3 MODEL-INDEPENDENT RELIABILITY ESTIMATES

The main part of the dissertation focuses on developing and comparing new approaches from the group of model-independent reliability estimates. The dissertation proposes 8 new such reliability estimates [8], denoted as SAvar, SABias-s, SABias-a, CNK-a, CNK-s, LCV, BAGV, and DENS. The first three of the listed estimates are developed by adapting the sensitivity analysis [9] approach for the use in the supervised learning. To apply the principles of the sensitivity analysis, we propose a framework for controlled

modification of the input (learning set) and outputs (regression predictions) in the supervised learning setting. By applying minor modifications to the learning set, we exploit the instabilities in predicted values and use them to compose reliability estimates. The other five estimates are adapted from related work based on the following approaches: computing bagging variance, local cross-validation, density estimation, and local error estimation. In the dissertation, the existing estimates are generalized for usage with other regression models.

In the experimental part, the dissertation presents an empirical evaluation of the above reliability estimates using 8 regression models (regression trees, linear regression, neural networks, bagging, support vector regression, locally weighted regression, random forests, generalized additive model) and 28 standard benchmark test domains. The performance of the reliability estimates is measured based on their correlation coefficient to the prediction error of the individual examples. The correlation coefficients are statistically evaluated to confirm whether the reliability estimates significantly estimate the prediction errors. The testing results demonstrated the usefulness of the proposed reliability estimates especially for the use with regression trees, where one of the proposed estimates correlated with the prediction error in 86% of the testing domains (individual results available in the online version of the dissertation [1]). On the average (across all used regression models) the estimate which is based on the bagging variance analysis achieved the highest performance (correlation to the prediction error). The ranking of the performance of tested reliability estimates is shown in Figure 1.

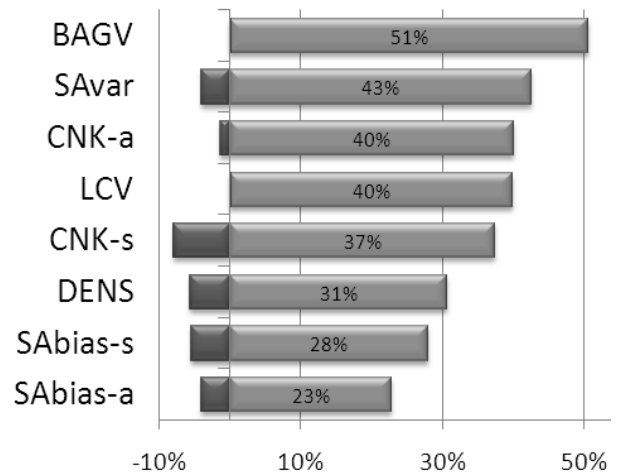


Figure 1: Average performance of tested reliability estimates across all testing domains and models. The figure shows the percentage of experiments exhibiting positive correlation of the estimate to the prediction error (light grey) and the percentage of experiments with the negative correlation to the prediction error (dark grey).

4 AUTOMATIC SELECTION OF THE BEST PERFORMING ESTIMATE

The testing results of the individual reliability estimates revealed that the estimates achieved different performance on different problem domains and using different regression models. Accordingly, in the dissertation we study the problem of the most appropriate reliability estimate selection for a given problem domain and the regression model [10]. We discuss and define two possible solutions of this problem, based on meta-learning and internal cross-validation approach.

In the context of the proposed meta-learning approach we define a meta-problem space for prediction of the best performing reliability estimate. The dissertation presents a possible attribute description of the meta-learning problem and defines it as a classification problem, where each class represents one of the 9 proposed reliability estimates. Using a collection of our 28 testing domains and in combination with 8 different regression models, we construct a meta-learning training set consisting of 224 (28×8) examples. We use this training set to construct a decision tree meta-classifier, which is afterward used to predict the most appropriate reliability estimates for testing domains which do not comprise the meta-training set.

Since decision tree is an interpretable model, we use the constructed meta-classifier to analyze in which cases each particular estimate perform better. The analysis results indicate that the estimates achieve better performance when used with more accurate models (models that achieve lower relative mean squared error on the testing set).

The second approach to the best performing estimate selection is based on the internal cross-validation approach. It is designed to iteratively measure the performance of the reliability estimates on different subsets of the testing domain. The best performing estimate on the average is afterwards used to estimate the reliability of the test examples which were excluded from the estimate selection process.

The testing results have shown that the dynamically selected reliability estimate using both of the above approaches achieves significant correlation to the prediction error in more experiments than any of the individual reliability estimates. Outperforming the most successful individual reliability estimate, which positively correlated with the prediction error in 51% of experiments, the procedures for automatic selection of the best performing estimates performed as follows:

- the *meta learning approach* to automatic selection dynamically selected such estimates that positively correlated with the prediction error in 57% of experiments and negatively correlated with the prediction error in 1% of experiments,
- the *internal cross-validation approach* to automatic selection dynamically selected such estimates that on the average positively correlated with the prediction error in 73% of experiments and negatively in none.

5 IMPLEMENTATION IN A MEDICAL DOMAIN

The individual estimates and the both approaches for automatic selection of the optimal estimate were tested on a real domain from the area of medical prognostics. The data consisted of 1035 breast cancer patients, who had a surgical treatment for cancer between 1983 and 1987 in the Clinical Center in Ljubljana, Slovenia. The goal of the research was to predict the time of possible cancer recurrence after the surgical treatment.

The analysis showed that this is a difficult prediction problem, because the possibility for recurrence is continuously present for almost 20 years after the treatment. Furthermore, the data presents a mixture of two prediction problems, which additionally hinders the learning performance: (i) yes/no classification problem, whether the illness will recur at all, and (ii) the regression problem for the prediction of the recurrence time.

In our study, the bare recurrence predictions were complemented with our reliability estimates. To implement the prediction system, the locally weighted regression was selected for the use with this problem due to its low relative mean squared error (RMSE), compared to the other models. The model was complemented with one of our reliability estimates which was unanimously selected by both of our approaches for the selection of the best performing estimate. A graphical representation of such predicted time of the cancer recurrence, equipped with reliability information is shown in Figure 2. The implemented prediction system helped the the doctors with the additional validation of the predictions' accuracies.

The statistical comparison of reliability estimates to prediction evaluations of the medical experts showed that our reliability estimates correlate to prediction error with statistically equal correlation as the manual evaluations of the experts. This results therefore showed the potential of the proposed methodology in practice.

6 CONCLUSION

Implementation of reliability estimates for the individual predictions can be a helpful tool when using critical decision support systems. In the dissertation, several such reliability estimates are evaluated and proposed. Additionally, two approaches for automatical estimate selection, which increase their consideration for practical use, are proposed and evaluated, as well.

The successful implementation of the proposed methodology in a medical domain indicates the importance and the potential for the use of the reliability estimation in practice. To conclude, the dissertation provides the theoretical time complexities for the computation of the estimates' values. The ideas for the directions of the further work include work on the interpretability of the estimates' values, analysis of the mathematical estimates' properties and best performing reliability estimate selection for an individual example to be predicted.

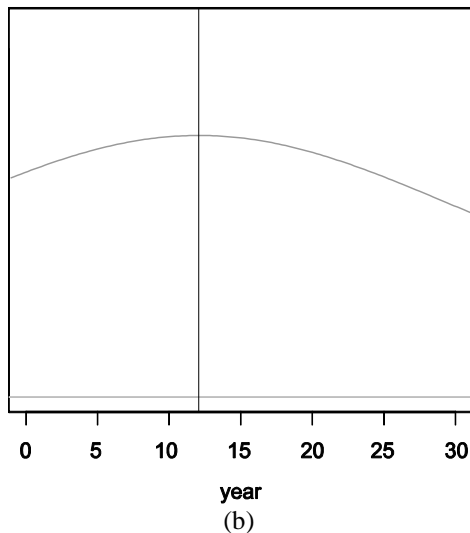
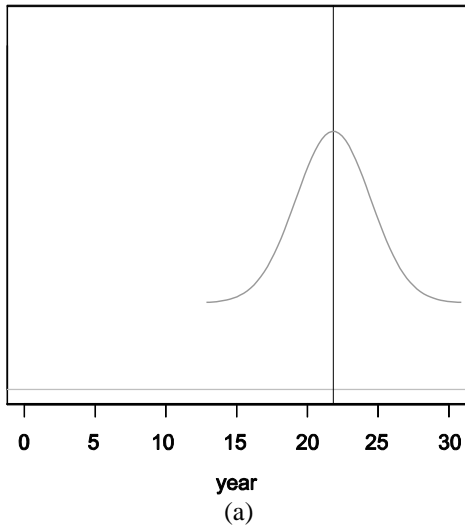


Figure 2: Graphical representation of two recurrence predictions (vertical lines) and their reliability (denoted by the width of the Gaussian, surrounding the vertical line). The figure (a) illustrates an example of the high prediction reliability (the narrow Gaussian), and (b) the low prediction reliability (the wide Gaussian).

References

- [1] Bosnić, Z.: Estimation of individual prediction reliability using sensitivity analysis of regression models (in slovene). PhD Thesis, University of Ljubljana, Faculty of Computer and Information Science, <http://lkm.fri.uni-lj.si/zoranb/dissertation.htm> (2007)
- [2] Bosnić, Z., Kononenko, I.: Estimation of individual prediction reliability using the local sensitivity analysis. *Applied Intelligence* 29(3) (2007) 187–203
- [3] Bosnić, Z., Kononenko, I.: Estimation of regressor reliability. *Journal of Intelligent Systems* 17(1/3) (2008) 297–311
- [4] Gammerman, A., Vovk, V., Vapnik, V.: Learning by transduction. In: *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, Madison, Wisconsin (1998) 148–155
- [5] Saunders, C., Gammerman, A., Vovk, V.: Transduction with confidence and credibility. In: *Proceedings of IJCAI'99*. Volume 2. (1999) 722–726
- [6] Kukar, M.: Quality assessment of individual classifications in machine learning and data mining. *Knowledge and Information Systems* 9(3) (2006) 364–384
- [7] Bosnić, Z., Kononenko, I., Robnik-Šikonja, M., Kukar, M.: Evaluation of prediction reliability in regression using the transduction principle, in: *Proceedings of Eurocon 2003*, B. Zajc and M. Tkalčič, eds. (2003) 99–103.
- [8] Bosnić, Z., Kononenko, I.: Comparison of approaches for estimating reliability of individual regression predictions. *Data & Knowledge Engineering* 67(3) (2008) 504–516
- [9] Bousquet, O., Elisseeff, A.: Stability and generalization. In: *Journal of Machine Learning Research*. Volume 2. (2002) 499–526
- [10] Bosnić, Z., Kononenko, I.: Automatic selection of reliability estimates for individual predictions. *Knowledge Engineering Review* (in press) (2008)

COGNITIVE COMPLEXITY OF MULTI-CRITERIA GROUP DECISION-MAKING METHODS

Andrej Bregar

Informatika d.d., Vetrinjska ulica 2, 2000 Maribor, Slovenia

e-mail: andrej.bregar@informatika.si

ABSTRACT

Several widely used and state-of-the-art multi-criteria methods for group decision-making are analysed and compared with regard to their information complexity and the cognitive load that is imposed on the decision-makers. Five evaluation criteria are considered: total number of preferential parameters, quantity of inputs required for the first iteration of the decision-making process, average number of manual adjustments in each subsequent iteration, amount of data that the decision-makers must observe in each iteration, and complexity of data types. Substantial differences between methods are confirmed according to the defined quality factors. The interactive multi-agent aggregation/disaggregation dichotomic sorting negotiation procedure based on the threshold model is determined to be the most efficient consensus seeking approach.

1 INTRODUCTION

There exist many methods and decision support systems for group decision-making [2, 13]. Their aim is to help potentially conflicting and opposing decision-makers in reaching efficient consensual or compromise solutions. Because they are based on various theories and approaches from the fields of multi-criteria decision analysis, artificial intelligence and operations research, they exhibit different properties with regard to the cognitive load that is imposed on the decision-makers, thoroughness of modelling the problem domain, and rationality of the decision. Several researchers have investigated these properties [8, 14], or defined frameworks for evaluating group decision-making methods [15]. However, no study has been presented in the literature that would systematically compare existing approaches according to their information/cognitive complexity. The purpose of this paper is thus to measure the complexity of some widely used, highly relevant and state-of-the-art group decision-making methods and systems:

1. interactive multi-agent aggregation/disaggregation dichotomic sorting procedure for group consensus seeking based on the threshold model [2, 3],
2. aggregation/disaggregation group decision support system based on the ELECTRE TRI method [5, 7],
3. ELECTRE TRI for groups [6],
4. ELECTRE-GD [11],
5. group PROMETHEE [1],

6. distance based collective preorder inference [10],
7. aggregation and disaggregation of utility function related collective preferences [12],
8. group AHP [17, 18],
9. consensus based group decision-making model integrating various preference structures [4, 9].

The rest of the paper is organized as follows. In Section 2, an overview of evaluated group decision-making methods is provided. These methods are analysed and compared in Section 3. Finally, Section 4 gives some conclusions and directions for further research work.

2 OVERVIEW OF COMPARED METHODS

2.1 Dichotomic Sorting Based Consensus Seeking

The approach is based on the ELECTRE TRI outranking method [16], which is slightly modified so that preferences are modelled in a symmetrically-asymmetrical manner in the neighbourhood of the reference profile b . The purpose of the profile is to divide the set of alternatives into two exclusive categories – all acceptable choices are sorted into the positive class C^+ , while unsatisfactory ones are the members of the negative class C^- . The decision-maker has to provide six preferential parameters for each criterion x_j , including the importance weight w_j , the value of the profile $g_j(b)$, and the thresholds of preference (p_j), indifference (q_j), discordance (u_j) and veto (v_j). Additionally, he can also specify the upper and lower allowed limits of these parameters, which constrain their automatic adjustment in the process of unification with the common opinion of the group. In order to reduce the cognitive load and to enable a rational convergence of individual judgements towards the consensual solution, several mechanisms are applied:

- Preferences may be specified with fuzzy variables or by the holistic assessment of alternatives.
- The most contradictive negotiator is identified by computing the consensus and agreement degrees.
- Several robustness metrics reveal if preferences of an individual are firmly stated.
- The centralized agent negotiation architecture and protocol eliminate the need for a human moderator and minimize the activity of each decision-maker.
- An optimization algorithm is implemented for the purpose of automatic preference unification.
- Weights are derived according to the effect of veto.

2.2 ELECTRE TRI Based Disaggregation Group DSS

The methodology is based on the ELECTRE TRI method and is implemented with the IRIS decision support system. The decision-makers discuss how to sort some exemplary actions into multiple categories, while the IRIS system helps them to iteratively reach an agreement by preserving the consistency of sorting examples both at the individual and the collective level. Some information that may direct the group members is suggested, however the mechanism does not identify the decision-maker who has to conform to the others, leaving this judgement to the moderator.

The decision-maker specifies n criteria-wise evaluations of m compared and several exemplary alternatives, whereby at least one referential example must be provided for each of $k+1$ categories delimited by k profiles. In addition, the upper and lower limits of allowed categories may be set for m sorted alternatives. The decision support system derives for every alternative the acceptability degrees of sorting it into suitable classes, category swaps which are required to attain different permitted classifications, and l constraints that imply the category memberships.

2.3 ELECTRE TRI for Groups

ELECTRE TRI for groups focuses on searching for the best and the worst class into which an alternative may be sorted with respect to the constraints on imprecise parameter values that are set by individual decision-makers. To capture the imprecision, a group member has to manually specify the exact values and the limits of all preferential parameters – n criteria weights, $3 \cdot n$ thresholds p_j , q_j and v_j , and $k \cdot n$ respectively $m \cdot n$ criteria-wise evaluations of referential profiles and alternatives. The veto thresholds can be defined with the linear function $v_j = p_j + \alpha_j / w_j$, where the α_j constant is also subjected to constraints. The method determines the collectively allowed ranges of classes based on the intersections of individual ranges.

2.4 ELECTRE-GD

ELECTRE-GD is an extension of the ELECTRE III method [16]. It derives a group rank-order of alternatives from individual rank-orders and corresponding fuzzy outranking relations. For each pair of alternatives, the degrees $\sigma_k(a_i, a_j)$ and $\sigma_k(a_j, a_i)$ are computed. If $\sigma_k(a_i, a_j) \geq \lambda$, the alternative a_i is at least as good as the alternative a_j , the $\sigma_k(a_i, a_j) \leq \lambda - \beta$ inequality implies that a_i is inferior to a_j , and the case of $\lambda - \beta < \sigma_k(a_i, a_j) < \lambda$ expresses a doubt about outranking. The inequalities are combined with the numbers of alternatives that are worse than a_i and a_j , respectively. Altogether, 45 combinations are possible, from which the global rank-order is induced. Hence, the method is rather complex. Its drawbacks are also that it does not enable iterativeness and that it can assure only a compromise, because the decision-makers must accept the solution recommended according to the rules of the human moderator.

2.5 Group PROMETHEE

PROMETHEE is a family of outranking methods. For a pair of alternatives a_i and a_j , and for each criterion x_k , the $P_k(a_i, a_j)$ function is defined according to the values $g_k(a_i)$ and $g_k(a_j)$, and according to the preference, indifference or Gauss thresholds. This function expresses to what degree a_i outperforms a_j , and can have one of six possible shapes of which the linear is the most common. For every choice a_i , the outranking degrees are aggregated into the positive and negative ranking flows that indicate how much a_i performs better respectively worse than all other alternatives. The inferred flows can be interpreted in two ways – the PROMETHEE I method derives a partial rank-order, while a weak rank-order is obtained with the PROMETHEE II method. The PROMETHEE II net flows are the input into the group setting. It treats each decision-maker as a separate criterion and applies the same aggregation procedure as is used for the case of a single decision-maker.

2.6 Distance Based Inference of Collective Preorder

The procedure for the inference of a weak rank-order of alternatives from partial rank-orders that are suggested by different decision-makers consists of several steps:

1. For the purpose of specifying individual rankings, the decision-makers apply the PROMETHEE I/II and ELECTRE III methods.
2. The decision-makers are assigned their weights by aggregating two types of components. The objective component of the k -th decision-maker according to whom the outranking relation $a_i P^k a_j$ holds for the pair of alternatives (a_i, a_j) is obtained as the ratio between the number of decision-makers preferring a_i over a_j and the total number of decision-makers. The subjective components are derived by ranking group members in a decreasing order of importance with the revised Simos' procedure.
3. Based on the distances between the relations of alternatives and the most/least favourable relations $>$ and $<$, two rankings P^+ and P^- are inferred.
4. By applying the ELECTRE III method, P^+ and P^- are combined into a single weak rank-order.

2.7 Aggregation/Disaggregation of Collective Utilities

This approach combines the concepts of utility theory, consensus seeking and preference disaggregation. Group members specify utilities of m alternatives with respect to n criteria. Based on these utilities, individual rankings of alternatives and criteria weights, a collective rank-order is derived. The decision-makers are then presented with the discrepancies between their own and unified preferences, so that they can express their opinion about the group solution on a five point scale according to four satisfaction criteria. In this way, the necessity to start a new iteration of the unification process is determined.

2.8 Group AHP

Several versions of the group Analytic Hierarchy Process exist. Usually, relative pairwise comparisons that are specified by individual decision-makers on the [1..9] ratio scale are aggregated into reciprocal compromise matrices from which the collective priorities of criteria and alternatives are derived. The aggregated preferences generally do not represent the opinion of the majority of decision-makers, but can instead reflect considerable discrepancies within the group. The consequence is that the Pareto optimality may be violated. To overcome this deficiency, stochastic multiplicative consensus models have been introduced. They define random variables that capture the variation of opinions and result in intervals of potentially overlapping synthesized priorities.

2.9 Consensus Model Integrating Different Preferences

This negotiation model does not require a human moderator as it relies on four types of communicating agents. The decision-makers' preferences may be expressed in various formats, which include the utility function, the rank-order of alternatives, the fuzzy preference relation defined on the [0..1] interval, and the multiplicative AHP based relation. The consensus seeking procedure consists of several steps:

1. Input information is unified by transforming different preference formats into the fuzzy relation.
2. Fuzzy relations of individual decision-makers are aggregated into the collective fuzzy relation with the OWA (*Ordered Weighted Average*) operators.
3. For each alternative, the degree of dominance over all other alternatives is computed.
4. From the dominance degrees, the individual and collective rank-orders of alternatives are derived.
5. The consensus degrees of alternatives and the total consensus degree are obtained with regard to the deviations of the individual rank-orders from the collective rank-order.
6. The decision-makers are ranked according to their distances from the collective choice. The ones that do not reach the concordance threshold must conform to the group. The mechanism suggests up to three suboptimal solutions in terms of rank-order adjustments for each contradictive decision-maker.

3 COMPARISON OF METHODS

It is sensible to measure the complexity of group decision-making methods from the perspectives of two fundamental variables – the initial cognitive load and the cognitive load during the iterative process of consensus or compromise seeking. The initial cognitive load depends on the method's theoretical properties, primarily on the amount of required input data. A crucial factor is also the difficulty of eliciting preferences, since the latter can be specified in the form of holistic assessments, by providing exact or approximate

parameter values, or by using some indirect technique, as the swaps and pairwise comparisons. The cognitive load during the process of group decision-making is reflected in the complexity and quantity of analytic operations that are performed by the decision-maker in every iteration. It can be reduced by incorporating the capability of autonomous interactive guidance and by implementing the disaggregation analysis for the purpose of automated induction of parameter values. The two conceptual variables are operationalized with five quantitative and qualitative criteria:

- the total number of preferential parameters, where a parameter is regarded as a formalized piece of information that is directly or indirectly provided by the decision-maker, such as a criterion weight, a threshold or the evaluation of an alternative (X_1);
- the quantity of inputs that are required for the first iteration of the decision-making process (X_2);
- the number of manual adjustments of preferential parameters in each subsequent iteration (X_3);
- the amount of data analysed in each iteration (X_4);
- the complexity of information types (X_5).

Except for the last criterion, the information complexity can be determined in the form of mathematical expressions consisting of constants and variables, among which the commonly used are the numbers of alternatives (m), criteria (n), classification/sorting categories (k) and decision-makers (o). The inference and detailed explanation of the expressions would exceed the scope of the paper, and can therefore be found in additional literature [2]. The methods are systematically compared in the table.

With regard to information types, there are no considerable differences between the nine evaluated methods. They all model judgements as relations and real numbers. Although some encourage the expression of preferences in a human friendly manner, for example with fuzzy linguistic variables, such information types are merely complementary to quantitative formats and have a partial, limited influence on the decrease of the overall cognitive load.

More interesting is the comparison of methods with regard to the quantity of information that has to be modelled and consequently specified, formalized, analysed and iteratively modified by the decision-makers. The following conclusions can be drawn from the criteria X_1 to X_4 :

- According to all criteria, the ELECTRE TRI based sorting methods are more cognitively demanding than ranking methods. An exception is the dichotomic sorting procedure for consensus seeking. It reduces the amount of processed information by means of two-categorical problem localization and autonomous induction of preferences with the aim of convergent opinion unification.
- Outranking methods that deal with rank-ordering of alternatives – group PROMETHEE, ELECTRE-GD, and the distance based collective preorder inference procedure – require a moderate quantity of

inputs for the first iteration of the decision-making process. However, a large amount of information is analysed in intermediate iterations and after the final round of preference aggregation. These methods also have a deficiency of poorly supporting the iterativeness. Only PROMETHEE allows for the interactive modification of preferences, yet it is limited to weights, which must be adjusted manually, since no induction algorithm is implemented.

- The group application of the utility function has similar properties as the above mentioned ranking methods. It is adequately efficient with regard to the overall amount of preferential parameters and the quantity of initial inputs. However, complex

information structures must be comprehended and manually adjusted in all subsequent iterations.

- Group AHP is subjected to the inherent weakness of the Analytic Hierarchy Process. A high initial cognitive load is imposed on the decision-maker who has to provide for many redundant pairwise comparisons. It also does not enable any iterativeness, but the amount of derived data is acceptable.
- The expressions for the consensus model integrating different preference structures do not take into account multiple criteria. So, most subexpressions should be multiplied by n in order to achieve comparability. In this case, the method would be evaluated as average according to all four criteria.

Table: Comparison of group decision analysis methods with regard to their information complexity.

	X_1	X_2	X_3	X_4	X_5
M_1	maximally $n \cdot (m+18)$	maximally $n \cdot (m+18)$	0 or $6 \cdot n$	$3 \cdot m-1$ to $3 \cdot m+12 \cdot n$	\mathfrak{R} , fuzzy, binary, (P, I)
M_2	$n \cdot (5 \cdot k+m+1)+2 \cdot m$	$(n+1) \cdot (m+k+1)+m$	m or $n \cdot (5 \cdot k+1)+m$	$l+2 \cdot \sum_i C^g(a_i)-C^c(a_i)$	\mathfrak{R} , membership, $(<, >)$
M_3	$3 \cdot n \cdot (m+k+4)$	$3 \cdot n \cdot (m+k+4)$	$2 \cdot n \cdot (m+k+4)$	at least $2 \cdot k \cdot (m+2 \cdot n)$	\mathfrak{R} , $(<, >)$
M_4	$n \cdot (m+4)+2$	$n \cdot (m+4)+2$	no iterativeness	over $2 \cdot m^2 \cdot o$	\mathfrak{R} , binary, (P, I, R)
M_5	maximally $n \cdot (m+4)$	maximally $n \cdot (m+4)$	n	up to $m \cdot (m+3)/2+n$	\mathfrak{R} , 2D visual, (P, I, R)
M_6	$n \cdot (m+4)+o$	$n \cdot (m+4)+l+o-1$	no iterativeness	$m \cdot (m-1)+3 \cdot m+o$	\mathfrak{R} , interval, (P, I, R)
M_7	$n \cdot (m+1)$	$m+m \cdot n+n$	$n \cdot (m+1)+3$	$m \cdot (m+2 \cdot n-1)/2+4$	\mathfrak{R} , qualitative, (P, I)
M_8	$N \cdot m + \sum_i n_i$	$\approx (N \cdot m^2 + \sum_i n_i^2)/2$	no iterativeness	$2 \cdot m$ to $4 \cdot m+1$	\mathfrak{R} , ratio, (P, I)
M_9	m to $m \cdot (m+1)/2+6$	m to $m \cdot (m+1)/2+6$	3 to $m \cdot (m+1)/2$	maximally $5 \cdot m+o$	\mathfrak{R} , ratio, ternary, (P, I)

4 CONCLUSION

The presented comparative study has found significant differences in the cognitive complexity of various existing group decision analysis methods. The dichotomic sorting procedure for consensus seeking has been shown to be the most effective and balanced, while deficiencies have been identified for all other approaches according to at least one evaluation criterion. Within the scope of further research work, additional criteria could be considered, particularly to measure the perceived cognitive load that people have to deal with in real-life problem solving situations.

References

[1] J. P. Brans, C. Macharis, B. Mareschal. *The GDSS PROMETHEE Procedure*. STOOTW/277. Vrije Universiteit Brussel. 1997.

[2] A. Bregar. *An Aggregation/Disaggregation Approach to Integrative Negotiations*. PhD Thesis. University of Maribor, Faculty of Electrical Engineering and Computer Science. 2009. In Slovene.

[3] A. Bregar, J. Györkös, M. B. Jurič. Interactive aggregation/disaggregation dichotomic sorting procedure for group decision analysis based on the threshold model. *Informatica*, vol. 19, no. 2, pp. 161–190. 2008.

[4] A. K. Choudhury, R. Shankar, M. K. Tiwari. Consensus based intelligent group decision-making model for the selection of advanced technology. *Decision Support Systems*, vol. 42, no. 3, pp. 1776–1799. 2006.

[5] S. Damart, L. C. Dias, V. Mousseau. Supporting groups in sorting decisions: Methodology and use of multi-criteria aggregation/disaggregation DSS. *Decision Support Systems*, vol. 43, no. 4, pp. 1464–1475. 2007.

[6] L. C. Dias, J. N. Climaco. ELECTRE TRI for groups with imprecise information on parameter values. *Group Decision and Negotiation*, vol. 9, no. 5, pp. 355–377. 2000.

[7] L. C. Dias, V. Mousseau. IRIS: A DSS for multiple criteria sorting problems. *Journal of Multi-Criteria Decision Analysis*, vol. 12, no. 3, pp. 285–298. 2003.

[8] R. P. Hämmäläinen, A. A. Salo, K. Pöysti. Observations about consensus seeking in a multi-criteria environment. *Proceedings of the 25th Hawaii International Conference on System Sciences*, pp. 190–198. 1992.

[9] E. Herrera-Viedma, F. Herrera, F. Chiclana. A consensus model for multiperson decision-making with different preference structures. *IEEE Transactions on Systems, Man and Cybernetics – Part A: Systems and Humans*, vol. 32, no. 3, pp. 394–402. 2002.

[10] K. Jabeur, J.-M. Martel, S. Ben Khélifa. A distance-based collective preorder integrating the relative importance of the group's members. *Group Decision and Negotiation*, vol. 13, no. 4, pp. 327–349. 2004.

[11] J. C. Leyva-López, E. Fernández-González. A new method for group decision support based on ELECTRE III methodology. *European Journal of Operational Research*, vol. 148, no. 1, pp. 14–27. 2003.

[12] N. F. Matsatsinis, E. Grigoroudis, A. P. Samaras. Aggregation and disaggregation of preferences for collective decision-making. *Group Decision and Negotiation*, vol. 14, no. 3, pp. 217–232. 2005.

[13] N. F. Matsatsinis, A. P. Samaras. MCDA and preference disaggregation in group decision support systems. *European Journal of Operational Research*, vol. 130, no. 2, pp. 414–429. 2001.

[14] S. Paul, W. D. Haseman, K. Ramamurthy. Collective memory support and cognitive-conflict group decision-making: An experimental investigation. *Decision Support Systems*, vol. 36, no. 3, pp. 261–281. 2004.

[15] S. K. Peniwati. Criteria for evaluating group decision-making methods. *Mathematical and Computer Modelling*, vol. 46, no. 7–8, pp. 935–947. 2007.

[16] B. Roy. The outranking approach and the foundations of ELECTRE methods. *Theory and Decision*, vol. 31, no. 1, pp. 49–73. 1991.

[17] T. L. Saaty. *Group decision making and the AHP. The Analytic Hierarchy Process: Applications and Studies*. Springer, New York. 1989.

[18] R. C. Van den Honert. Stochastic group preference modelling in the multiplicative AHP: A model for group consensus. *European Journal of Operational Research*, vol. 110, no. 1, pp. 99–111. 1998.

BEHAVIOUR RANDOMNESS MEASUREMENT AS A PART OF COMPLEX CUSTOMER VALUE INDICATOR

Naděžda Chalupová, Arnošt Motyčka

Department of Informatics

Faculty of Business and Economy, Mendel University in Brno

Zemědělská 1, 613 00 Brno, Czech Republic

Tel: +420 5 4513 22 {49, 00}

e-mail: {nadule, mot}@pef.mendelu.cz

ABSTRACT

This contribution deals with new perspectives of Chaos Theory application into the process of customer behaviour monitoring via a customer value indicator. This indicator can be perceived as a powerful tool implemented into decision support systems. Existing CRM (Customer Relationship Management) and CVM (Customer Value Management) experience reduces the customer value to its profitability only. A new designed model calculates with actions of financial and non-financial customer benefits for an enterprise. The one of acting factors is customer buying behaviour randomness – its monitoring has gained importance recently, especially in connection with irregular development of world economies.

1 INTRODUCTION

The customer value (CV) for the company shall perform a role of indicator of general successfulness of relations with customers. The widely reputable benchmark of this value is the so-called Customer Life-time Value (CLTV). For the company, this value assessment becomes a front support for design of marketing strategies during customer relation management. A unified definition of the customer value still does not exist, which is, of course, highly typical for scopes of marketing or management.

2 CV CALCULATION APPROACHES

Quite a number of simple and also complex models exist for customer value assessment.

2.1 Simple Techniques

As an example of simpler methods one can introduce the ARPU model (Average Revenue Per Unit) or AMPU model (Average Margin Per Unit), which can be very often met at operators in communication and which is instrumental to international benchmarking. Other approaches bearing in mind the fact of the value of particular customers can be different. For example, Lehtinen (2007) characterizes the customer value as the fraction of sales volume to a given customer in the general company sales volume. Ferrell and Hartline (2004) define the customer value as metrics, which

is represented by the present value of cashflows resulting from the relation with a concrete customer in a certain period. Hoffman et al. (2006) approach specifies the customer value on the basis of initial and also forecasted future purchases. According to Hughes (2003), the customer value is the rate of net profit, which can be gained by the company from a concrete customer during the whole relation life cycle.

These indicators are established on a very simple basis and calculate with plain customer return in a certain period only. They are mainly used as indicators of the profit generating ability.

2.2 More Complex Methods

Clow and Baack (2006) introduce a more complex approach: the customer value is based on the average number of visits in a year. It measures the average amount of money spent on a visit and the average duration of relation lifetime with the customer. Further there is the value decreased by customer acquirement and maintenance costs and increased by the value of customer references. The value shall also take the customer loyalty into account. The LTV model (LifeTime Value) or CE model (Customer Equity) and others, which are mentioned in detail e. g. by Kumar and Reinartz (2006), belong among more advanced methods.

Bauer, Hammerschmidt and Braehler (2003) and other authors offer more objective metrics considering a more factors action. Truneček (Vlček, 2004) defines the customer value for the enterprise as present and awaiting future values of a concrete customer discounted on its net present value (the base of this customer value concept calculation are four quantifiable magnitudes: return, costs, investments and risk). The work of Vlček (2004) became the cardinal inspiration for the methodology design of our own.

2.3 Scale LTV Model

A new LTV model was designed according to the definition (Souček, Chalupová, 2008): The customer value represents its financial and non-financial benefits for an enterprise. These are developing over time and its

importance is influenced by concrete situational conditions.

The customer value reflects influences of number of factors that are expressed on the principle of rating scales. This way rated factors are integrated into one complex indicator on the basis of particular actors importance. A unified model would not satisfy all requirements of the whole market.

Following factors were chosen for the customer value (CV) formulation:

- profitability of a customer relationship – p
- duration of a relationship – t
- value for a customer – vc
- reference value – vr
- loyalty of a customer – l
- randomness of customer buying behaviour – r

The basic concept is represented by a multidimensional model in the form of a scalar product

$$CV_i = \sum_{j=1}^n f_{i,j} \cdot w_j$$

whereas applies

$$\sum_{j=1}^n w_j = 1$$

Here the f vector represents rated values of above mentioned factors (p, t, vc, vr, l, r), w are weights representing importance of these factors. The i variable signs monitored customer identification, j is order of a certain factor or its weight. The weights of attributes are set by a competent expert. Further text is focused on methodology of the buying behaviour randomness factor quantification (r).

3 RANDOMNESS OF CUSTOMER BUYING BEHAVIOR

The Chaos theory deals with description and endeavour for better understanding of random events. The theory approaches to time series data analyses in absolutely other way than traditional methods: it allows to find out whether the data has some internal structure or whether it deals with really random data (Hurst exponent) or it can even quantify prediction reliability of a monitored event (Lyapunov exponent). Already this issue stands for the aim of whole computation.

The Chaos theory solves more aspects and uses more methods than mentioned above. A detailed overview offers e. g. Sprott (2003) or others.

3.1 Quantification methods

The time series of customers expenses will be the analysis input, because the development of revenues coming from a customer mostly predicates about customer behaviour (about his/her willingness to expend). The output then will be an indicator quantifying a rate of chaos (presence of complicated memory effect in data) in revenues brought by customer development, which represents the searched rate of randomness in customer buying behaviour (r_i).

Lyapunov exponent that quantifies sensitive dependence on initial conditions, offers the clearly maximal predicative ability. The largest positive exponent from spectrum (each system has the Lyapunov exponents of its own, called Lyapunov spectrum) of Lyapunov exponents is usually meant as the Lyapunov exponent. Its calculation is quite complicated, Wolf et al. (1985) describe the widely used algorithm including Fortran code of this algorithm. The Lyapunov exponent reciprocal indicates directly the *predictability* of watched magnitude.

Of course it is not possible to suppose that at all times (over all customers) at least one positive exponent will be in the resulting Lyapunov spectrum. In case of all exponents being non-positive, it is namely not possible to interpret the largest Lyapunov exponent in other ways than the watched time series is really random (does not contain non-linear dynamics). Therefore there is no point in trying to calculate the predictability (the reciprocal of negative number is the negative number again and the negative value of predictability is a reasonless value in light of interpretation).

Therefore the Lyapunov exponent using in this model cannot be totally disallowed, but it is necessary to manipulate the spurious values in appropriate manner. It means that with the customers, whose time series of revenues development does not exhibit the memory effects presence, the searched quantity (predictability) is zero. That can be written by the formula

$$r_i = \begin{cases} 0 & \text{for } \lambda_i \leq 0 \\ \frac{1}{\lambda_i} & \text{for } \lambda_i > 0 \end{cases}$$

where λ_i is the largest positive Lyapunov exponent from the Lyapunov spectrum.

It is not possible to extrude the situation, when, by the above mentioned method, it will be found out that the predictability of customer behaviour is zero for all or nearly all customers. In this case it is impossible to resolve, which customers behave more chaotic or more randomly, because the fact, that the exponent is negative “more” or “less”, has no practical importance. The negative exponent value (even if the largest from the spectrum) does not predicate about the rate of chaos – it only signs the fact that the monitored customer behaves really randomly (non-predictably).

In this situation the *Hurst exponent* estimation can be considered as applicable for obtaining the value of the customer behaviour randomness indicator. This situation is expressed by the equation

$$r_i = H(\tau_i),$$

where H is the Hurst exponent and τ_i stands for the vector of values representing development of revenues brought by the monitored customer (time-series).

It provides a measure of the relative tendency of a time series to either strongly regress to the mean or continue in its current direction. There is a variety of techniques that exist for estimating the exponent, however assessing the reliability of the estimation can be a complicated issue. Taqqu, Teverovsky and Willinger (1995) have undertaken

empirical comparisons of nine estimators, Rea et al. (2009) provide results of a simulation study, where two of twelve estimators examined performed best. Esposti and Signorini (2006) analyzed the quality of eight methods and then formulated a procedure useful for a reliable H estimation. It works on the basis of an indirect estimation of the stationarity of the data series, and in accordance with this indicator, the procedure recommends the best method for the Hurst exponent estimation.

The Hurst exponent always lies between 0 and 1, and equals 0.5 for processes without underlying trends (random walk). If the Hurst exponent is a different value in $\langle 0, 1 \rangle$, it deals with a long memory process. Higher values indicate a smoother trend, less volatility, and less roughness.

The exponent from $(0.5, 1)$ indicates persistent behaviour of the trend – positive correlation – situation, where the trajectory tends to continue in its current direction (empowered trend) and thus produces enhanced (or anomalous) diffusion. Time series with this characteristic exhibit long-term memory effects, which means that the system is sensitive to infinitesimally changes in initial conditions (what happens today, will influence the future for ever). Opposite to this, values from $\langle 0, 0.5 \rangle$ exhibit antipersistence – negative correlation – a state, where the trajectory tends to return to the point from which it came (the trend is weakened, change of trajectory direction is drawing) and thus the diffusion is suppressed. Antipersistent systems, in contrast to independent systems, travel less, thus they have to change their behaviour more frequently than random process. The trend changes are very common, but unpredictable. The smaller value of the Hurst exponent, the more rugged the series is (it covers more areas, the trajectory direction changes are creating more frequently) and vice versa.

3.2 Factor value rating

The scale rating of certain factors of CV is carried out according to certain rules, which are clearly specified for particular factors. These rules result from the interpretation of real calculated values of p, t, vc, vr, l, r variables.

With the $p, t, vc, vr,$ and l variables, the rating rules are simple: the higher is the calculated variable value, the higher will be the point-like variable representation (f_j). The r variable in its character differs from the others – its value is always not so simply interpretable.

In case of the Lyapunov exponent using for the predictability quantification, the mentioned rule applies. If the Hurst exponent for the behaviour randomness rate quantification would be used, the scale rating is more complicated. Although it deals with the same principle (the better is the situation, the better is its score), the difference is that not only one value, but also a combination of more aspects predicate about the situation advantageousness. The behaviour of monitored time series is the base for potential situation definitions next to particular values of which the Hurst exponent can gather. With respect to these two aspects, these situations can arise:

1. the time series has no memory effect \Rightarrow the behaviour is fully random and therefore unpredictable ($r = 0.5$),
2. the trend is empowered and the time series falls \Rightarrow the depression continues – preferably as small fall as possible is desirable (the lower $r \in (0.5, 1)$ during actual fall of time series values, the better),
3. the trend is weakened and the time series grows \Rightarrow the growth changes itself to fall – preferably as small fall as possible is desirable (the higher $r \in \langle 0, 0.5 \rangle$ during actual grow of time series values, the better),
4. the trend is weakened and the time series falls \Rightarrow the fall ends, the growth begins – preferably as big growth as possible is desirable (the lower $r \in \langle 0, 0.5 \rangle$ during actual fall of time series values, the better),
5. the trend is empowered and the time series grows \Rightarrow the growth continues – preferably as big growth as possible is desirable (the higher $r \in (0.5, 1)$ during actual grow of time series values, the better).

Specified situations in the above mentioned list are ordered with the company view from the worst (the least desirable) one to the best (the most desirable) one. On the base of this order, it is then possible to rate potentially arisen situations on the scale. This system is illustrated on the Figure 1.

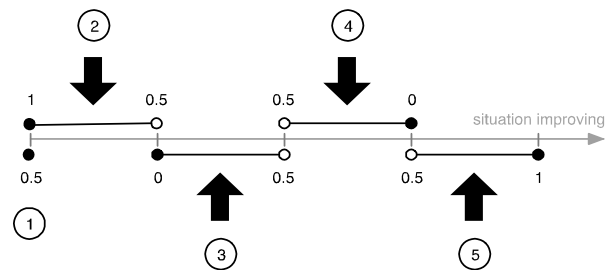


Figure 1: System of situation rating in Hurst exponent using for customer behaviour randomness determination

Here the particular situations are marked by corresponding number in the ring, vertical arrows sign if the time series grow (\uparrow) or fall (\downarrow) and black vectors signalize the intervals in which the r value can be in certain situation found, grey horizontal arrow (\rightarrow) indicates the direction of situation improvement.

4 SOFTWARE IMPLEMENTATION ASPECTS

It is advisable to visualize the CV indicator among a group of key performance indicators collected in managerial cockpit – a group of aesthetic dashboards in analytical systems (for decision support), with views adapted to the interest of each analyst. The application can be realized as a pluggable software component in portal – portlet.

The data required for the buying behaviour randomness factor calculation by both the above mentioned methods have a simple structure. Only the customer identification, date of purchase and total money volume of all purchases of a certain customer on a certain day is needful. The suppose is that data from the business transactional database will be transformed to this structure and then

loaded into data warehouse over which the functionality of the application visualizing the CV indicator will be built.

One of the designed views can be seen in Figure 2. Here the pie chart sectors represent the size of certain customer segments differing from each other on the base of CV range. Excepting the graphic expression, it also represents other information (average CV in the enterprise, if it ascent etc.).

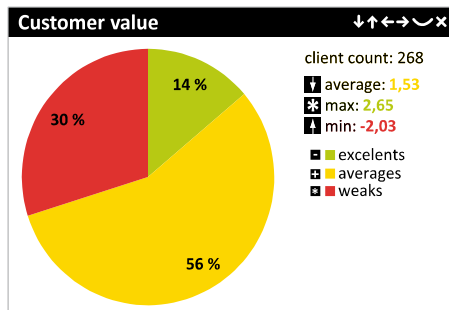


Figure 2: Visualization of customer values in managerial dashboard

The calculations of particular factors are heart of the functionality. Some open source software is available for this factor calculation. The *tshurst* utility from NtropiX (Conover, 2006) and the *lyap_spec* utility, implementing the Wolf algorithm (Wolf et al., 1985), from TISEAN (Hegger, Kantz and Schreiber, 1999) package were experimentally used. The mentioned Espostis and Signorinis procedure for reliable H estimation can be used as improvement alternatively. It is necessary to suggest that the company does not need to have enough robust time series. For this case, it is possible to use a method for the Lyapunov exponent calculation from small data sets (Rosenstein, Collins, De Luca, 1993).

5 CONCLUSION

The new CV model benefit lies in its universal applicability across economy branches. It brings possibilities of better understanding of customer buying behaviour especially in connection with world economics irregular development.

Acknowledgement

This paper is supported by the project IG 190 631 titled "The portal for trading sphere subjects monitoring".

References

[1] Bauer, H. H., Hammerschmidt, M., Braehler, M. The Customer Lifetime Value Concept and its Contribution to Corporate Valuation. In *Yearbook of Marketing and Consumer Research*, Vol. 1, 2003. pp. 47–67. ISSN 1612-9814.

[2] Clow, K. E., Baack, D. E. *Integrated Advertising, Promotion, and Marketing Communications*. 3rd edition. New Jersey: Prentice Hall, 2006. 544 p. ISBN 0-13-186622-2.

[3] Conover, J. *Software For Programmed Trading Of Equities Over The Internet* [online]. last modification:

July 2006 [cit. 25th August 2009]. URL <<http://www.johncon.com/ntropix/utilities.html>>

[4] Esposti, F., Signorini, M. G., *Evaluation of a blind method for the estimation of the Hurst's exponent in time series*. XIV European Signal Processing Conference. Florence, Italy, 2006. URL <<http://www.eurasip.org/Proceedings/Eusipco/Eusipco2006/papers/1568980804.pdf>>.

[5] Ferrel, O. C., Hartline, M. D. *Marketing Strategy*. 3rd edition. USA: South-Western College Publishing, 2004. 648 p. ISBN 0-324-20140-0.

[6] Hegger, R., Kantz, H. Schreiber, T. Practical implementation of nonlinear time series methods: The TISEAN package. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, Vol. 9, Issue 2, 1999. pp. 413–435. ISSN 1054-1500.

[7] Hoffman, K. D. et al. *Marketing Principles and Best Practices*. 3rd edition. USA: South-Western College Publishings, 2006. 598 p. ISBN 0-324-20044-7.

[8] Hughes, A. M. *The Customer Loyalty Solution: What Works (and What Doesn't) in Customer Loyalty Programs*. 1st edition. USA: McGraw-Hill, 2003. 336 p. ISBN 0-07-142904-2.

[9] Kumar, V., Reinartz, W. J. *Customer relationship management: a databased approach*. 1st edition. New York: John Wiley&Sons, 2006. 323 p. ISBN 0-471-27133-0.

[10] Lehtinen, J. R. *Aktivní CRM: Řízení vztahů se zákazníky*. 1st edition. Praha: Grada Publishing, 2007. 160 p. ISBN 978-80-247-1814-9. (In Czech)

[11] Rea, W. et al. *Estimators for Long Range Dependence: An Empirical Study*. 2009. (Submitted to the Electronic Journal of Statistics. ISSN 1935-7524.)

[12] Rosenstein, M. T., Collins, J. J., De Luca, C. J. A practical method for calculating largest Lyapunov exponents from small data sets. *Physica D: Nonlinear Phenomena*, Vol. 65, Issue 1–2, 1993. pp. 117–134. ISSN 0167-2789.

[13] Souček, M., Chalupová, N. Utilization of Chaos Theory in Customer Lifetime Value Management. In *International Journal of Management Cases*. Vol. 10, Issue 3/4, 2008. pp. 73–79. ISSN 1741-6264.

[14] Sprott, J. C. *Chaos and Time-Series Analysis*. 1st edition. New York: Oxford University Press, 2003. 508 p. ISBN 0-19-850840-9.

[15] Taqqu, M. S., Teverovsky, V., Willinger, W. Estimators for long-range dependence: an empirical study. *Fractals: Complex Geometry, Patterns, and Scaling in Nature and Society*, Vol. 3, Issue 4, 1995. pp 785–798. ISSN 0218-348X.

[16] Vlček, R. *Manažerské přístupy podporující vliv zákazníka při řízení firmy* [online]. last modification: March 2004 [cit. 1st August 2009]. URL <http://old.fph.vse.cz/veda/stoly/pdf/ks040310_vlcek.pdf>. 16 p. (In Czech)

[17] Wolf, A. et al. Determining Lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena*, Vol. 16, Issue 3, 1985. pp. 285–317. ISSN 0167-2789.

A HYBRID NEURAL NETWORK MODEL FOR SPAM DETECTION

Maria Corduneanu, Carmen Maria Cosoi, Catalin Alexandru Cosoi, Madalin Vlad, Valentin Sgarciu

Department of Automatic Control and Computer Science

University Politehnica Bucharest

313 Splaiul Independentei Str, Bucharest, Romania

Tel: +4021-402 93 10, +4021-318 10 14

e-mail: maria, carmen, catalin.cosoi, madalinv@ac.pub.ro, vsгарciu@aii.pub.ro

ABSTRACT

Spam has become a global problem. Latest studies estimate that as much as 9 out of 10 emails are spam. Many solutions have been published so far, but every time a suitable solution is found, spam mutates into something new, so new ways to fight it must be found. A good method to fight spam at a proactive level would be the use of neural networks but, as you will see in this paper, applying neural network theory per se is not enough.

1 INTRODUCTION

The currently employed infrastructure for eMail transfer, the simple mail transfer protocol (SMTP), hardly provides any support for detecting or preventing spam. We are also lacking a widely accepted and deployed authentication mechanism for sending emails. Thus, until a new global email infrastructure will be developed so as to allow a better control of this problem, there are two current major approaches that show the greatest potential for coping with the problem: detecting spam based on content filtering or preventing spam to enter our mailboxes by using techniques such as reputation management, white-listing, increasing the costs associated with sending out email messages, and so on.

Current Token-based spam filters (e.g. Signature Filters, Heuristic Filters, Neural Network Filters, Bayesian Filters, Support Vector Machines) distinguish between spam and legitimate email messages based primarily on the tokens found in those messages' text. However this approach has had mixed results. On one hand, many spam messages have token signatures that facilitate filtering. These signatures typically consist of tokens that are invariant for the many variants automatically generated by spammers. On the other hand, spammers can use various techniques to defeat this filters. (Pu et. all, 200X)

Judging by the frequency of their updates, we noticed that token-based filters can be classified in two major categories:

1. Long term filters (updated weekly or monthly, or maybe never)
2. Short term filters (updated hourly or daily)

Each of these filters makes use of a feature/token extraction algorithm in order to have enough information for a good spam vs. legitimate classification. It is also known that long term filters have incredible good detection rates in laboratory conditions, while short term filters very good detection rates in real world conditions.

Although sending billions of email messages advertising ridiculous products that most of us would never in our lives consider buying, what makes spamming profitable is its large volume. According to the New York Times, people click and buy products advertised in pharmaceutical spam emails. Other articles suggest that it costs about 300\$ to send 1 million emails. Assuming that a spammer makes just 25\$ from each sale (which is the lowest profit he can make), it's easy to see that it makes only slightly more than 2 million messages to make an immediate 10 000\$ profit (Beckman, 2007).

Over time, several techniques have been proposed to address this problem, like Bayesian Filtering Techniques (Graham, 2002), URL filtering, heuristic filtering, spam image filtering (Cosoi, 2006) and so on, but each time an acceptable solution was found, spam quickly mutated to something new and harder to catch.

Due to the fact that all the techniques enumerated above are all reactive, the need for a proactive solution is obvious. Heuristic filters look for patterns in the content of an email and match them against a database of known spam characteristics. These characteristics can be in the form of certain words, phrases, punctuation and altered dates. These are strong patterns and they match a single type of spam, offering zero false positives (a legitimate email mistakenly classified as spam), but the process of creating strong patterns is usually insidious and time consuming.

A good way to create strong patterns would be to use a neural network that combines short weaker patterns (if the email has words like "Viagra", "Valium", or if the date of the message is in the future and so on, which individually have a high false positive rate) and to use a neural network in order to combine these into stronger and longer patterns.

2 PROPOSED METHOD

A good neural network type up for this task would be ARTMAP networks (Cosoi, 2006). ARTMAP architectures are neural networks that develop stable recognition codes in real time into response to arbitrary sequences of input patterns. They were designed to solve the stability-plasticity dilemma that every intelligent machine learning system is facing: how to keep learning from new events without forgetting previously learned information. ARTMAP networks were designed to accept binary or fuzzy input patterns (Carpenter & Grossberg, 1991). ARTMAP networks consist of two ART1 networks, ARTa and ARTb, bridged via an inter-ART module, as shown on Fig 1. An ART module has three layers: the input layer (F0), the comparison layer (F1), and the recognition layer (F2).

The neurons, or nodes, in the F2 layer represent input categories. The F1 and F2 layers interact with each other through weighted bottom-up and top-down connections, which are modified when the network learns. There are additional gain control signals in the network that regulate its operation.

In the training phase, the system has to receive a list of features extracted from the email messages and an output category. For example, ARTa will receive an input vector where each field indicates the existence of a certain spam or legitimate characteristic. Also, each input vector will be associated to a label which indicates if the current pattern was extracted from a spam or a legitimate email message, which will be fed to the ARTb module. When the training phase starts, the system will quickly associate inputs and outputs by creating strong patterns for each category.

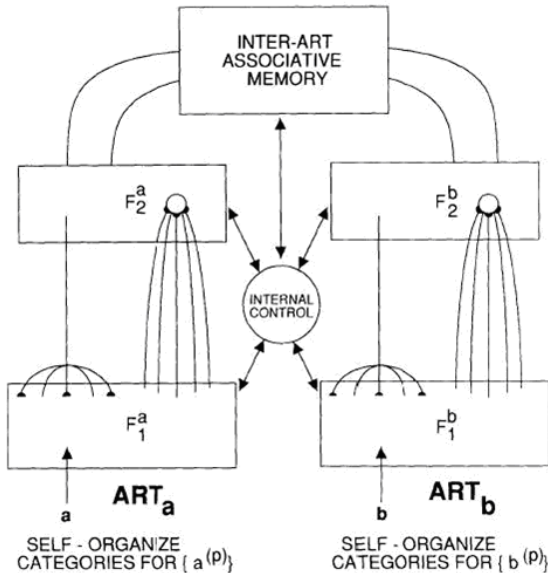


Fig. 1. ARTMAP system diagram (Carpenter & Grossberg, 1991)

The results are very good (Cosoi, 2006), with a false positive rate of almost 1% (which is not exactly the best yet

obtained, but it can be rated among the top 3 AntiSpam filters) and a false positives rate (spam messages mistakenly misclassified as legitimate messages) under 10%. The problem that appears is that since the training phase is performed on a few million legitimate and spam messages samples, and since the individual heuristics are generally weak, the extracted patterns can be quite confusing for the neural network algorithm. For example we can have a situation where important legitimate features and standard weak spam features can determine a mistakenly “this is spam” answer, and vice-versa.

These situations are generally determined by the large corpus of messages on which the neural network has to train in order to achieve an acceptable accuracy. In many situations, in our experiments, the training phase stopped after a fixed number of training iterations was achieved, and not when reaching a pre-established accuracy.

The solution we found to address this problem is to a priori offer a numerical relevance to each individual feature, and also the category (spam or legitimate) for which this feature was created. Our purpose was to create an inhibited connection, in order to stop the neural network to give an answer if the relevance of the pattern was smaller than a pre-established threshold T. Of course, this means that good hits would be eliminated to, but common-sense would say that we can’t actually say an email is a spam message only because it contains the word “Viagra”.

If we consider I and S the relevance for the legitimate heuristics within a subset of a pattern and respectively S the relevance for the spam heuristics, we can combine them in a total relevance for a pattern by using the following simple rule:

$$R = \frac{1 - I + S}{2} \quad (1)$$

Where, I and S are computed as percents of the total sum of the relevancies within a pattern.

By using this result, the neural network can determine if this is an important pattern for the decision process or not. Of course, now this approach is more of a heuristic filter than a neural network. In order to keep all the facilities that a neural network would offer, (and we also chose this type of neural network in order to solve the stability-plasticity dilemma), we had to add a punishment-reward system in the control subsystem of the ARTa module (see Fig. 2). The process we developed is quite simple to explain. Each time the prediction matched the expectation we increased by a small amount the relevance of that pattern. If the prediction and the expectation were different, we decreased the relevance with a small amount. The process can be defined using the following formula.

$$R_{i+1} = (1 - w)R_i + w(R + (-1)^c \cdot \frac{w}{100}) \quad (2)$$

Where $(-1)^c$ has a negative value when the expectation and the prediction are different, and a positive one when the two are the same.

3 RESULTS

Our tests showed that by applying the improvements presented in this paper, the false positives rates dropped radically from an initial 1% to 1 in a million, while the false negative rate reached 13%, compared to an initial value of 10%. Although this method provides a slightly increase of the false negatives rate, it is far more important to prevent tagging as spam a legitimate email message than overlooking a few spam messages.

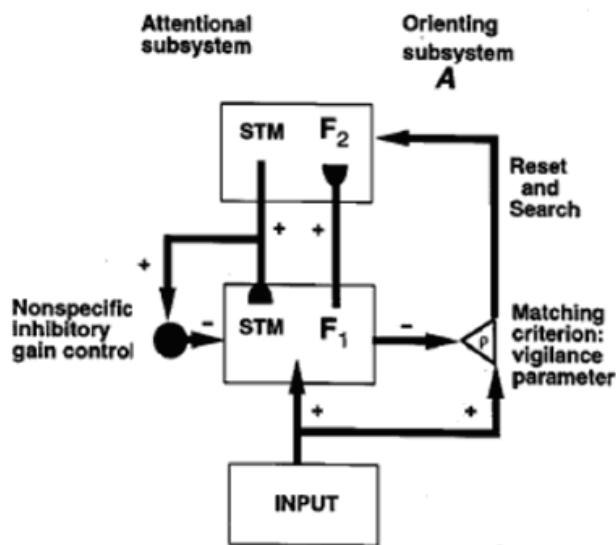


Fig. 2. ARTa system diagram (Carpenter & Grossberg, 1991)

The conditions in which the experiments took place are the following:

- 2.5 million spam messages
- Almost 1 million legitimate email messages
- 75% of the message corpus were used for training the neural network and,
- 25% were used in testing the neural network.

Prosody has great impact on intelligibility and naturalness of speech perception. The proper choice of prosodic parameters, given by phoneme duration and intonation contours, enables natural sounding high quality synthetic speech.

4 CONCLUSION

Our tests showed that by applying the improvements presented in this paper, the false positives rates dropped radically from an initial 1% to 1 in a million, while the false negative rate reached 13%, compared to an initial value of

10%. Although this method provides a slightly increase of the false negatives rate, it is far more important to prevent tagging as spam a legitimate email message than overlooking a few spam messages.

The conditions in which the experiments took place are the following:

- 2.5 million spam messages
- Almost 1 million legitimate email messages
- 75% of the message corpus were used for training the neural network and,
- 25% were used in testing the neural network.

REFERENCES

- [1] Beckman S. (2007). High-Performance Asynchronous IO for SMTP Multiplexing
- [2] Available from: <http://www.spamconference.org> Accessed: 2007-03-31
- [3] Cosoi, A. C. (2006). The medium or the message? Dealing with image spam. Available from: <http://www.virusbtn.com>, Accessed: 2006-12-3
- [4] Cosoi A. C. (2006). An AntiSpam filter based on adaptive neural networks, Available from: <http://www.spamconference.org>, Accessed: 2006-04-15
- [5] Graham P. (2002). A plan for spam, Available from: <http://www.paulgraham.com/spam.html>, Accessed: 2007-05-27
- [6] Carpenter, G. & Grossberg, S. (1991). Supervised real-time learning and classification of nonstationary data by a self-organizing neural network, In: *Pattern recognition by self organizing neural networks*, Carpenter, G. & Grossberg, S., (Ed. MIT press), 501-544, Publisher MIT press, ISBN 0-262-03176-0, Cambridge Massachusetts

DETECTING ANOMALIES IN SOCIAL NETWORKS USING FRACTAL NETWORKS

Catalin Alexandru Cosoi, Madalin Stefan Vlad, Maria Corduneanu, Carmen Maria Cosoi

Department of Automatic Control and Computer Science

University Politehnica Bucharest

313 Splaiul Independentei Str, Bucharest, Romania

Tel: +4021-402 93 10, +4021-318 10 14

e-mail: catalin.cosoi, madalinv, maria, carmen@ac.pub.ro

ABSTRACT

This paper will try to demonstrate that the Romanian blogosphere is a social fractal, a network that scales up and down with equal facility. We will create a network of blogs linked by influence using notions like memes, Internet memes, meme-tracker, and the Yale approach to influence. Fractal geometry provides an effective way to describe the complex property of a 2D map. This paper uses a box-counting method to describe the fractal property of the Romanian Blogosphere.

1 INTRODUCTION

A *meme* is a unit or element of cultural ideas, symbols or practices; such units or elements transmit from one mind to another through speech, gestures, rituals, or other imitable phenomena. The etymology of the term relates to the Greek word *mimema* for mimic. Memes act as cultural analogues to genes in that they self-replicate and respond to selective pressures.

Richard Dawkins coined the word "meme" as a neologism in his book *The Selfish Gene* (1976) to describe how one might extend evolutionary principles to explain the spread of ideas and cultural phenomena. He gave as examples melodies, catch phrases, and beliefs (notably religious belief), clothing/fashion, and the technology of building arches.

Meme-theorists contend that memes evolve by natural selection (in a manner similar to that of biological evolution) through the processes of variation, mutation, competition, and inheritance influencing an individual entity's reproductive success. Memes spread through the behaviors that they generate in their hosts. Theorists point out that memes which replicate the most effectively spread best, and some memes may replicate effectively even when they prove detrimental to the welfare of their hosts.

A field of study called memetics arose in the 1990s exploring the concepts and transmission of memes in terms of an evolutionary model. Criticism from a variety of fronts has challenged the notion that scholarship can examine memes empirically. Some commentators question the idea that one can meaningfully categorize culture in terms of discrete units.

At its most basic, an *Internet meme* is simply the propagation of a digital file or hyperlink from one person to others using methods available through the Internet (for example, email, blogs, social networking sites, instant messaging, et cetera). The content often consists of a saying or joke, a rumor, an altered or original image, a complete website, a video clip or animation, or an offbeat news story, among many other possibilities. An Internet meme may stay the same or may evolve over time, by chance or through commentary, imitations, and parody versions, or even by collecting news accounts about itself. *Internet memes have a tendency to evolve and spread extremely quickly, sometimes going in and out of popularity in a matter of days.* They are spread organically, voluntarily, and peer to peer, rather than by compulsion, predetermined path, or completely automated means.

Blogosphere is a collective term encompassing all blogs and their interconnections. It is the perception that blogs exist together as a connected community (or as a collection of connected communities) or as a social network.

A *meme-tracker* is a tool for studying the migration of memes across a group of people. The term is typically used to describe websites that either analyze blog posts to determine what web pages are being discussed or cited most often on the World Wide Web, or allow users to vote for links to web pages that they find of interest.

The introduction of meme-trackers was instrumental in the rise of blogs as a serious competitor to traditional printed news media. Through automating (or reducing to one click) the effort to spread ideas through word of mouth, it became possible for casual blog readers to focus on the best of the blogosphere rather than having to scan numerous individual blogs. The steady and frequent appearance of citations of or votes for the work of certain popular bloggers also helped create the so-called "A List" of bloggers.

Further on, we must now define what influence is. Alex Mucchielli defined it as an ensemble of manipulation procedures of the cognitive objects which defines the situation.

The Yale approach specifies four kinds of processes that determine the extent to which a person will be persuaded by a communication.

- Attention: One must first get the intended audience to listen to what one has to say.
- Comprehension: The intended audience must understand the argument or message presented.
- Acceptance: The intended audience must accept the arguments or conclusions presented in the communication. This acceptance is based on the rewards presented in the message.
- Retention: The message must be remembered, have staying power.

The Yale approach identifies four variables that influence the acceptance of arguments.

- Source: What characteristics of the speaker affect the persuasive impact?
- Communication: What aspects of the message will have the most impact?
- Audience: How persuadable are the individuals in the audience?
- Audience Reactions: What aspects of the source and communication elicit counter arguing reactions in the audience?

The main distance used in PostRank for top30 is Influence = Acceptance X Retention. (FocusBlog, 2008).

In "The Fractal Blogosphere" at Read/Write Web, Richard MacManus proposes that bloggers not worry too much about the popular/unpopular dichotomy suggested by most common interpretations of the various power laws that govern linking and traffic among blogs, but instead pick a scale that makes sense and judge themselves by their success at the appropriate level.

He proposes an initial idea of five levels, based on audience-size jumps of powers of ten (10 readers, 100, 1000, 10,000, 100,000), calling them "personal," "social," "community," "broadcast," and "celebrity."

Power laws are arguably part of the very nature of links. To explain why, Shirky poses a thought experiment: Imagine that 1,000 people were all picking their favorite ten blogs and posting lists of those links. Alice, the first person, would read a few, pick some favorites, and put up a list of links pointing to them. The next person, Bob, is thus incrementally more likely to pick Alice's favorites and include some of them on his own list. The third person, Carmen, is affected by the choices of the first two, and so on. This repeats until a feedback loop emerges. Those few sites lucky enough to acquire the first linkages grow rapidly off their early success, acquiring more and more visitors in a cascade of popularity. So even if the content among competitors is basically equal, there will still be a tiny few that rise up to form an elite.

The power law is dominant because of a quirk of human behavior: When we are asked to decide among a dizzying array of options, we do not act like dispassionate decision-makers, weighing each option on its own merits. Movie producers pick stars who have already been employed by other producers. Investors give money to entrepreneurs

who are already loaded with cash. Popularity breeds popularity.

As expected, the Romanian Blogosphere will follow the same principles, even though the figures might not have the same magnitude. Each of these blogs will be part of a certain network, linked by blogroll or citation - the number of links that pointed toward each site ("inbound" links, as they're called), because they are the most important and visible measure of a site's popularity.

2 FRACTAL NETWORKS

Scale-free graphs represent a relatively recent investigation topic in the field of complex networks. The concept was introduced by Albert and Barabasi in order to describe the network topologies in which the node connections follow a power law distribution. Common examples of such networks are the living cell (network of chemical substances connected by physical links). Although traditionally large systems were being modeled using the random graph theory developed by Erdos and Renyi (*On random graphs*), during the last few years research has led to the conclusion that a real network's evolution is governed by other laws: regardless of the network's size, the probability $P(k)$ that a node has k connections to other nodes is a power law:

$$P(k) = ck^{-\gamma} \quad (1)$$

This implies that large networks follow a set of rules in order to organize themselves in a scale-free topology. Barabasi and Albert show the two mechanisms that lead to this property of scale invariance: growth (continuously adding new nodes) and preferential attachment (the likelihood of connecting to existing nodes which already have a large number of links). Therefore, scale-free networks are dominated by a small number of highly connected hubs, which on one hand gives them tolerance to accidental failures, but on the other hand *makes them extremely vulnerable to coordinated attacks.* (Barabasi and Albert - *Statistical mechanics of complex networks*)

Based on the remark that random graph-theory does not explain the presence of a power law distribution in scale-free networks, Barabasi and Albert (2002) recommend a growth algorithm that has this property. They show that the assumptions on which the models have been generated up to that point were genuinely false: firstly, considering the number of nodes as being fixed and constant and secondly, the fact that connections were randomly established between the nodes. In fact, real networks are open systems, continuously evolving by adding new nodes. (Ursianu & Sandu, 2007).

As opposed to a random graph, in which all nodes have approximately the same degree, a scale-free graph contains a few so-called hubs (nodes with a great number of links, like the Britney Spears Twitter Profile¹ with 867333

¹ <http://twitter.com/britneyspears>

followers), while the majority of the nodes only have a few connections (50% of the twitter users have an average of 10 connections): this is a power law distribution. In a random network the nodes follow a Poisson distribution with a bell shape, and it is extremely rare to find nodes that have significantly more or fewer links than the average. A power law does not have a peak, as a bell curve does, but it is instead described by a continuously decreasing function. When plotted on a double-logarithmic scale, a power law is a straight line. (Ursianu & Sandu, 2007).

There are two major ways to compute the dimension of this network: box counting method and the cluster growing method.

For the box counting method, let N_B be the number of boxes of linear size l_B , needed to cover the given network. The fractal dimension d_B is then given by

$$N_B \approx l_B^{-d_B} \quad (2)$$

This means that the average number of vertices $\langle M_B(l_B) \rangle$ within a box of size l_B

$$\langle M_B(l_B) \rangle \approx l_B^{d_B} \quad (3)$$

By measuring the distribution of N for different box sizes or by measuring the distribution of $\langle M_B(l_B) \rangle$ for different box sizes, the fractal dimension d_B can be obtained by a power law fit of the distribution.

For the cluster growing method, one seed node is chosen randomly. If the minimum distance l is given, a cluster of nodes separated by at most l from the seed node can be formed. The procedure is repeated by choosing many seeds until the clusters cover the whole network. Then the dimension d_f can be calculated by

$$\langle M_C \rangle \approx l^{d_f} \quad (4)$$

where $\langle M_C \rangle$ is the average mass of the clusters, defined as the average number of nodes in a cluster. These methods are difficult to apply to networks since networks are generally not embedded in another space. In order to measure the fractal dimension of networks we need the concept of renormalization.

In order to investigate self-similarity in networks, we use the box-counting method and renormalization. For each size l_B , boxes are chosen randomly (as in the cluster growing method) until the network is covered. A box consists of nodes separated by a distance $l < l_B$. Then each box is replaced by a node (renormalization). The renormalized nodes are connected if there is at least one link between the un-renormalized boxes. This procedure is repeated until the network collapses to one node. Each of these boxes has an effective mass (the number of nodes in it) which can be used as shown above to measure the fractal dimension of the network.

The fractal properties of the network can be seen in its underlying tree structure. In this view, the network consists of the skeleton and the shortcuts. The skeleton is a special

type of spanning tree, formed by edges the having the highest betweenness centralities, and the remaining edges in the network are shortcuts. If the original network is scale-free, then its skeleton also follows a power-law degree distribution, where the degree can be different from the degree of the original network. For the fractal networks following fractal scaling, each skeleton shows fractal scaling similar to that of the original network. The number of boxes to cover the skeleton is almost the same as the number needed to cover the network. (Ursianu & Sandu, 2007).

3 RESULTS

In order to establish whether these networks are indeed scale-free, we determined the degree-distribution $P(k)$, which is the probability of finding a node with a degree k in the Romanian Blogosphere. The obtained distribution is indeed scale-free and satisfies the power law with the exponential: $\gamma = 2.65$ which satisfies our condition to be between 2 and 3.

$$P(k) \approx ck^{-\gamma} \quad (5)$$

$$\log(P(k)) = (-\gamma)\log(k) + \log(c) \quad (6)$$

$$y = (-\gamma)x + c \quad (7)$$

Using the box number distribution of the network, we obtained the dimension $d_B = 2.72$. This means that this network is indeed a fractal network.

On this level of influence we can find both influent bloggers and also random blogs which hope to increase their visibility by approaching subjects similar to influent bloggers.

In Romania, blogging is still in an incipient state. There are still just a few highly influent blogs, and the approached posts are on similar subjects – news and online businesses.

We believe that this incipient blogosphere will keep expanding and will become an important factor in the news industry. We present, as an example the Amsterdam incident.

The Boeing 737-800, which originated from Istanbul, Turkey, was trying to land at Schiphol when it crashed at about 1040 local time. The plane was carrying about 135 people. The first report on Twitter reportedly came from @nipp, who posted the message "Airplane crash @ Schiphol Airport Amsterdam!!" at 10:42, only 2 minutes after the crash. Barnett said that when CNN saw the image it moved quickly to confirm with Dutch officials that a crash had happened.

And this is just one example of many. Soon, breaking news will be first found on social sites and Citizen Journaling will become more and more influent in the news business. Although, in Romania, this process it is still at the beginning, analyzing the fractal properties of the influence network will become a highly important factor when studying the flow of information in the Romanian society.

4 CONCLUSION

Several fundamental properties of real complex networks, such as the small-world effect, the scale-free degree distribution, and recently discovered topological fractal structure, have presented the possibility of a unique growth mechanism and allow for uncovering universal origins of collective behaviors. However, highly clustered scale-free network, with power-law degree distribution, or small-world network models, with exponential degree distribution, are not self-similarity.

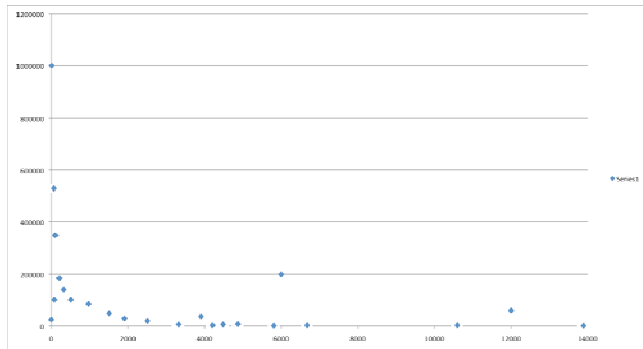


Fig 1. Node degree vs Nr. Of nodes – followers

We believe that analyzing the fractal properties of the Romanian Blogosphere will give us an insight of the future Citizen Journaling and its influence in the local media. Even though at its beginning, influent bloggers are already present and will increase their influence in time and create other Class A bloggers, and this will cause the network to keep expanding.

References

- [1] Barbara, D. (1999). Chaotic Mining Knowledge Discovery Using the Fractal Dimension. In: *1999 ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery (DMKD)*
- [2] Inaoka, H., Ninomiya T., Taniguchi K. (2004). Fractal Network derived from banking transaction. In: *Bank of Japan Working Paper Series*
- [3] Shirky C. (2003). Power Laws, Weblogs, and Inequality. In: *Clay Shirky's Writings About the Internet - Economics & Culture, Media & Community, Open Source*
- [4] Zou, L., Pei W, Li T., He Z., Cheung Y. (2006). Topological fractal networks introduced by mixed degree distribution. In: *Data Analysis, Statistics and Probability*
- [5] Bausch, S., McGiboney, M. Media Alert. Nielsen-Online (see www.nielsen-online.com).
- [6] BBC News, How the Schiphol crash happened. (see <http://news.bbc.co.uk/2/hi/europe/7910215.stm>. Downloaded on 10 april 2009)
- [7] Cohen, N. Moldovans Turn to Twitter to Organize Protests. The New York Times (see

- <http://thelede.blogs.nytimes.com/2009/04/07/moldovans-turn-to-twitter-to-organize-protests/?hp>)
- [8] Bachman, M. Connecting the dots. Nielsen Online (see www.nielsen-online.com).
 - [9] Dawkins, R. The selfish gene. ISBN 0199291144, 9780199291144, Published by Oxford University Press, 2006
 - [10] Yale attitude change program. Persuasive communication theories of persuasion and attitude change . (see www.elcamino.edu/faculty/rwells/PERSUASIVE%20COMMUNICATION.ppt, Downloaded on 1 April 2009).
 - [11] FocusBlog (a Romanian Blogosphere Memetracker, see www.focusblog.ro).
 - [12] MacManus, R. The Fractal Blogosphere. (see http://www.readwriteweb.com/about_readwriteweb.php Downloaded on 14 march 2009).
 - [13] Thompson, C. The Haves and Have-Nots of the Blogging Boom. New York Magazine (see <http://nymag.com/news/media/15967/> Downloaded 14 march 2009).
 - [14] Aberden Group, Research Brief. February 2008, Nielsen Online (see www.nielsen-online.com)
 - [15] Cosoi, A. C., Petre L.G. Workshop on digital social networks. SpamConference 2008, Boston, MIT
 - [16] Albert, R., Barabasi A., Statistical mechanics of complex networks. Review of modern physics 47-97.
 - [17] Ursianu, R., Sandu A., Self-Similarity of scale-free graphs. Proceedings of CSCS 16, Bucharest, Romania, page 121.
 - [18] Erdos, P., Renyi A., On random graphs. Publ Math. Inst. Hung. Acad. Sci, 290-297

EVALUATION OF POPULAR FEATURE RANKING ALGORITHMS IN MICROARRAY ANALYSIS

Mario Gorenjak, Mateja Bajgot, Biljana Pejčić, Andrej Sovec, Gregor Štiglic

Faculty of Health Sciences

University of Maribor

Zitna ulica 15, 2000 Maribor, Slovenia

Tel: +386 2 3004750; fax: +386 2 4747

e-mail: gregor.stiglic@uni-mb.si

ABSTRACT

Selecting the most informative genes from microarray expression data and constructing a reliable set of genes, is an essential part of microarray analysis and it is therefore necessary to explore the most effective feature selection methods. In this study 8 feature selection methods were evaluated with the aim to compare their stability at different numbers of genes in a ranked list. Average overlap of genes selected on different datasets for each method has been calculated, followed by a comparison of overlap using ranks. Furthermore, overlap of all methods between themselves had been compared. Finally we present a novel process for evaluation of obtained lists of genes. In the final experiment, a comparison of resulting gene lists to a top tissue-related gene lists from a text-mining application showed significantly different results, when using particular feature selection techniques.

1 INTRODUCTION

A large amount of genetic data is produced by the development of the DNA microarray technology and because of that it is easier to monitor the expression patterns of thousands of genes simultaneously under particular experimental environments and conditions [1]. Although gene expression microarrays are a popular tool for detecting differences in gene activity across biological samples [2], information from microarrays has not yet been widely used in diagnostic or prognostic decision-support systems, partly due to the diversity of results produced by the different available techniques [3]. It is therefore necessary to explore the most efficient method for selecting discriminative genes from the high dimensional microarray expression data. Considering the huge number of genes, included in the original data set, the impact to the accuracy and speed of classification or prediction systems, data reduction by selecting the most informative genes is very important, as well as constructing a reliable set of genes or gene expression signature for further genetic research. Different methods for selecting genes have been studied in combination with several classification and pre-processing algorithms [3-5]. Pirooznia et al. [6] reported a substantial

influence of chosen feature selection method, number of genes in the gene list and number of cases on the classification success. However, there is a lack of effective comparisons between several gene selection methods. In this study we evaluated 8 different gene selection methods at different number of genes in a ranked gene list by comparing the stability of the ranked gene lists. Furthermore, we compared a resulting gene list to a top disease related genes list provided by a web-based disease-gene text-mining application.

2 DATA AND METHODS

The original data used in this study was obtained from the Expression Project for Oncology (expO) data set from The International Genomics Consortium that was deposited at Gene Expression Omnibus (GEO) repository, accession number GSE2109. Samples from this collection are used in the Gene Expression Machine Learning Repository (GEMLeR), available at <http://gemler.fzv.uni-mb.si/> that represents a collection of 36 data sets derived from the original expO data set. All the samples were based on Affymetrix GeneChip U133 Plus 2.0 arrays.

The first empirical experiment for this study was conducted using all 36 data sets, while the remaining two experiments used only the largest dataset, because of high computational complexity.

Seven widely used feature selection methods that are implemented in Weka [7] machine learning environment were used in this study. Additional to those seven methods we also used an implementation of t-test based feature selection to compare the above mentioned methods to a "classic" gene ranking method that is widely used in bioinformatics. Table 1 represents average computational complexity of feature selection methods that is measured as time needed to rank all genes and generate a list of top 100 genes in the largest dataset.

One might notice an extremely high computational complexity of SVM based method which is caused by the default Weka setting for this feature selection method that eliminates one gene per iteration. By modifying this parameter to eliminate 50% of genes per iteration the time of execution drops to 28.8 seconds. This was the only

modification to default Weka parameters in our experiments. Further information on all feature selection methods can be found in [7].

Table 1: Average time (in seconds) to rank top 100 genes for all feature selection methods

Feature Selection Method	Time (sec)
T-Test (TT)	3.10
Chi Squared (CS)	3.34
Gain Ratio (GR)	3.14
Info Gain (IG)	3.19
OneR (OR)	25.13
ReliefF (RF)	144.22
SVM-RFE (SR)	32234.92
Symmetrical Uncertainty (SU)	3.23

Another tool for additional evaluation of obtained results was used – i.e. Gene Prospector. The Gene prospector is a component of HUGE Navigator [8], an integrated knowledge base for genetics association and human genome epidemiology. The Web-based application selects and prioritizes potential disease-related genes by using a highly curated and updated literature database of genetic association studies [9].

Published literature in human genome epidemiology is selected from PubMed and deposited in the HUGE Navigator database, which contains a curated collection of selected PubMed records from 2001 to the present [10]. The records are retrieved from PubMed weekly, followed by an initial screen of newly added records, which is performed by a text mining program developed by Yu, et al. [11]. A curator reviews the abstracts and manually indexes abstracts that meet the selection criteria with gene symbols, categories and study types. Furthermore, MeSH terms for each article are retrieved from the PubMed database. To facilitate free text search, the meta-thesaurus in the Unified Medical Language System is used as a lookup table for term synonyms, as well as the Entrez gene records from the NCBI Entrez Gene database [12] are used as standards for gene information.

The genes are ranked according to the amount of published literature in human genome epidemiology and published research. The ranked gene list is generated by a heuristic scoring formula, based on the total number of publications in the database for a particular gene-disease combination, with additional weight given to four types of publications: genetic association studies, genome-wide association studies, meta-analyses/pooled analyses and genetic testing articles. Such list of genes ranked by score allows users to find out rapidly which associations have been studied most often and systematically and provides an efficient resource for users seeking to evaluate genetic association.

3 EXPERIMENTAL SETUP

To allow comparison of gene symbols in the last experiment it was needed to convert the probe names into gene

symbols. In case of more probes mapping to the same gene, a maximal expression was used to determine which probe will be mapped to gene symbol. Additionally, to reduce high computational complexity of experiments, a filter based on gene expression variance was used to remove 80% of genes with the lowest variance levels across all samples in 36 datasets. Consequently each of the datasets in our experiments contained 4128 attributes (genes) with an additional class attribute.

The first experiment was used to simulate two separate feature selection studies of the same microarray analysis problem. The following three steps represent the experimental procedure:

- Split the dataset in two halves,
- Execute feature selection (gene ranking) on both halves and calculate overlap of top genes,
- Repeat 100 times using randomized shuffling of samples and calculate average overlap.

Our second experiment compared feature selection methods between themselves to find out the similarities among them. The procedure is similar to the first experiment with a single significant difference – after splitting the data set into two halves, the first one is used as an input to the first feature selection method, while the other half will be used for the second feature selection method. Comparing all of the 8 methods there were 28 pairs of classifiers to be evaluated. Again, the experiment was repeated 100 times for each pair of classifiers.

In our third experiment we observed the similarities of lists of selected genes by 8 compared methods and gene lists generated from text mining process in bioinformatics literature. Eight gene selection methods were applied on the largest dataset, comparing breast cancer tissues against colon cancer tissues. Eight lists of ranked genes were obtained, each with 512 genes included. Each resulting gene list has been compared with two, tissue-related gene lists from the Gene Prospector application, created by the Gene Prospector queries »breast« and »colon«. We observed the overlap score at different number of selected genes, by counting genes displayed in the ranked gene list and at least one of the potentially disease-related genes lists from the Gene Prospector. For each gene selection method the percentage of overlapping genes has been calculated at different number of ranked genes, ranging from 8 to 512.

4 RESULTS

In our initial experiment we calculated average overlap of genes selected on different datasets (obtained by splitting initial data in two datasets) for each of the eight feature selection methods (filters). Observing the average overlap values (Figure 1) we can see that SU and CS feature selection methods have the highest and almost the same values, starting at eight genes, overlap value is approximately 0.5 and climbs to approximately 0.8 at 512 genes. Those values are directly followed by values achieved by IG, GR and OR feature selection methods. The

lowest overlap results were achieved with SR and TT feature selection methods. At 512 genes the TT feature selection method achieved the 4th highest value right behind SU, CS and IG.

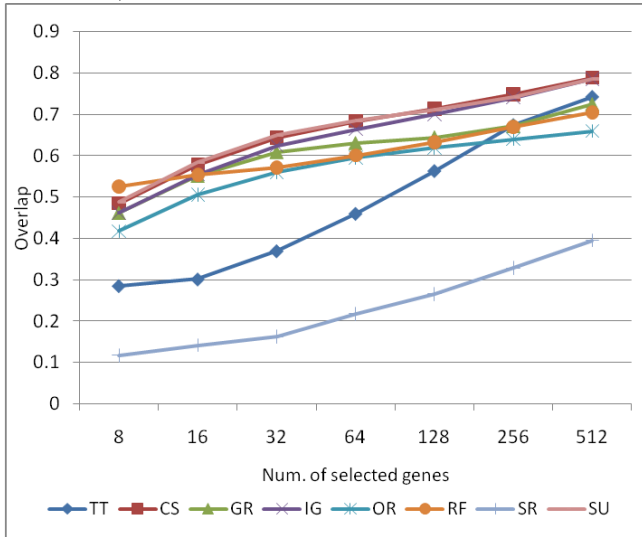


Figure 1: Average overlap of selected genes for different settings.

Additional comparison of overlap was done using ranks of eight compared methods. Microsoft Excel was used to calculate ranks from overlap percentage for all 36 data sets. In formula that calculates ranks the order was equal to 0, which means that values are ranked in decreasing order (the highest number is ranked with number one). Ties were solved by average rank that was calculated in case two or more methods returned the same result.

Figure 2 presents average ranks across all 36 datasets for all compared filters. Best performance was achieved by CS and SU with total average rank values 2.369 for chi squared and 2.258 for SU. Overall, based on ranks, the most stable feature selection methods are CS and SU. Stability of SR feature selection method is the most prone to changes in training data.

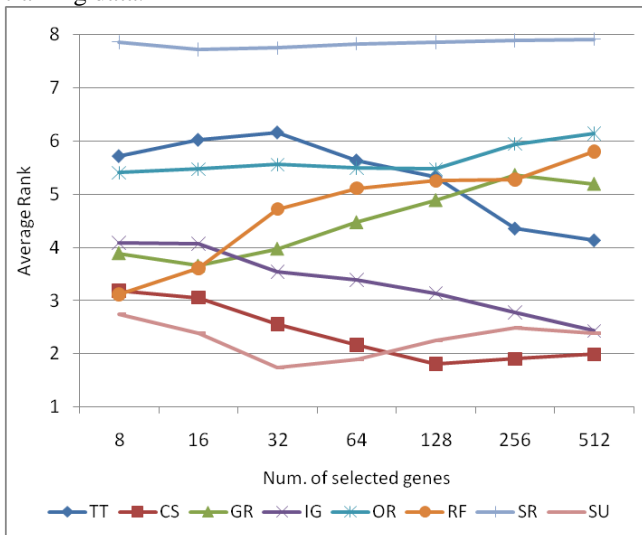


Figure 2: Average ranks of selected genes for different settings (higher overlap means lower rank).

Our second experiment compared results of all feature selection methods between themselves. Therefore all 28 pairs of classifiers were compared for different number of selected genes ranging from 23 to 29. The most outlying results were obtained when TT or SR feature selection methods were compared to the remaining methods. This is demonstrated in Figure 3, where one can notice the weak overlap between the results from both methods when compared to results of CS.

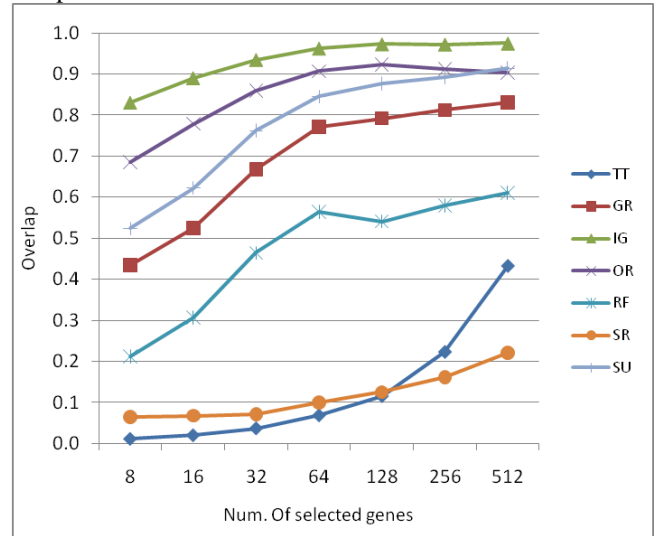


Figure 3: Average overlap between Chi Squared and each of seven other methods.

The final experiment tries to confirm the high instability of TT and SR results that were obtained in the first two experiments. The results of this experiment (Table 2) once again demonstrate that one can get significantly different results with TT or SR compared to the remaining group of feature selection techniques.

Table 2: Number of genes selected by feature selection method and Gene Prospector at the same time.

No. Of Genes	Overlap with Gene Prospector Queries							
	CS	GR	IG	OR	RF	SR	SU	TT
8	1	1	0	1	0	1	1	1
16	2	2	2	2	1	1	2	1
32	3	4	3	3	3	1	4	3
64	6	7	6	5	7	7	6	5
128	10	8	10	10	13	16	9	15
256	21	19	21	20	24	24	20	27
512	45	42	44	44	45	53	43	54

The average percentages of overlapped genes are lower than expected, mainly due to the discordance of gene names from the Affymetrix mappings and the Gene Prospector search application. The TT and SR methods

achieved the highest numbers of overlapped genes with a large increase at 128 genes and upwards. Moreover, the overlap scores at small number of genes, ranging from 8 to 64, for these methods were comparatively low. Generally, the overlap scores are rather irregularly distributed, which could indicate instability of the above-mentioned feature selection methods.

5 CONCLUSIONS

Based on our results achieved in three experiments, it can be concluded that the choice of the feature selection techniques is very important to achieve the most accurate gene selection. In first experiment (Figure 1), the best average overlap results have been achieved with SU and CS feature selection methods with almost the same values trough all gene numbers range. Significant overlap result was also achieved with TT feature selection method, but only for very high number of selected genes. However, in general the worst performing feature selection methods in terms of stability are TT and SR. In our additional comparison of overlap using ranks (Figure 2), best performance with minor difference was also achieved by SU and CS feature selection methods. The difference between SU and CS feature selection method is that SU feature selection method achieved better ranks in range from 8 to 64 genes, while CS feature selection method achieved better ranks in range from 128 to 512 genes. It can be said that SU feature selection method is the best choice for small sets of pre-selected genes, while CS should be used in cases of larger sets of genes.

In second experiment we have discovered the most outlying results when TT or SR feature selection methods were compared to the remaining methods. TT and SR have less common genes in comparison with other feature selection methods. At TT, the acquired data improves at 512 genes, while SR feature selection method constantly returns similarity of chosen genes which is lower than 30 percent. This experiment returned very interesting results, especially considering the high accuracy that can be achieved using SVM for classification.

The third experiment was used to test the biological importance of the results from the previous two experiments. According to Gene prospector TT and SR feature selection methods return the highest rate of genes that have been mentioned in the literature. Based on the results acquired in this experiment, the low overlap of TT and SR against other methods can be explained. However, one should be very cautious using those two methods due to their extremely high instability.

ACKNOWLEDGEMENT

This work was partially supported by Slovenian Research Agency, under grant BI-JP/09-11-002.

References

- [1] Harrington, C.A., Rosenow, C., Retief, J. (2000). Monitoring gene expression using DNA microarrays. *Curr. Opin. Microbiol*, 3: 285-291.
- [2] Song, S., Black, M.A. (2008). Microarray-based gene set analysis: a comparison of current methods. *BMC Bioinformatics*, 9: 502+.
- [3] Zervakis, M., Blazadonakis, M.E., Tsiliki, G., Danilatu, V., Tsiknakis, M., Kafetzopoulos, D. (2009). Outcome prediction based on microarray analysis: a critical perspective on methods. *BMC Bioinformatics*, 10: 53+.
- [4] Cho, S.-B., Won, H.-H. (2003). Data mining for gene expression profiles from dna microarray. *International Journal of Software Engineering and Knowledge Engineering*, 13(6): 593-608.
- [5] Kadota, K., Nakai, Y., Shimizu, K. (2009). Ranking differentially expressed genes from Affimetrix gene expression data: methods with reproducibility, sensitivity, and specificity. *Algorithms for Molecular Biology*, 4: 7+.
- [6] Pirooznia, M., Yang, J.Y., Yang, M.Q., Deng, Y. (2008). A comparative study of different machine learning methods on microarray gene expression data. *BMC Genomics*, 9 (Suppl 1): S13.
- [7] Witten, I. H. and Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition (Morgan Kaufmann Series in Data Management Systems)*. Morgan Kaufmann.
- [8] Yu W., Gwinn M., Clyne M., Yesupriya A., Khoury J.M. (2008). A navigator for human genome epidemiology. *Nat Genet*, 40:124-125.
- [9] Yu W., Wulf A., Liu T., Khoury J.M., Gwinn M. (2008). Gene Prospector: An evidence gateway for evaluating potential susceptibility genes and interacting risk factors for human diseases. *BMC Bioinformatics*, 9:528+.
- [10] Lin B.K., Clyne M., Walsh M., Gomez O., Yu W., Gwinn M., Khoury J.M. (2006). Tracking the epidemiology of human genes in the literature: the HuGE Published Literature Database. *Am J Epidemiol*, 164:1-4.
- [11] Yu W., Clyne M., Dolan S.M., Yesupriya A., Wulf A., Liu T., Khoury M.J., Gwinn M. (2008). GAPscreener: an automatic tool for screening human genetic association literature in PubMed using the support vector machine technique. *BMC Bioinformatics*, 9:205.
- [12] D. Maglott, J. Ostell, K. D. Pruitt, and T. Tatusova, "Entrez gene: gene-centered information at ncbi." *Nucleic Acids Res*, vol. 33, no. Database issue, January 2005.

EQUATION-BASED MODELS OF OILSEED RAPE POPULATION DYNAMICS DEVELOPED FROM SIMULATION OUTPUTS OF AN INDIVIDUAL-BASED MODEL

Aneta Ivanovska¹, Graham Begg², Ljupčo Todorovski³, Sašo Džeroski¹

¹Department of Knowledge Technologies, Jozef Stefan Institute, Ljubljana, Slovenia

²Scottish Crop Research Institute, Invergowrie, Dundee, Scotland

³Faculty of Administration, University of Ljubljana, Ljubljana, Slovenia

e-mail: aneta.ivanovska@ijs.si

ABSTRACT

Individual-based models are becoming increasingly popular in agriculture, where they are used for modeling different types of plant populations. This paper presents a new individual-based model for simulating the dynamics of a transgene within oilseed rape populations. We use the output from this model to develop equation-based models of the oilseed rape (OSR) population dynamics in a single arable field with the equation discovery system LAGRAMGE. We present preliminary results of the analysis of the outputs from individual-based models in agriculture with machine learning.

1 INTRODUCTION

Many different simulation models exist in ecology and agriculture. Most of them are population-based and study the long-term and short-term characteristic properties of a population, such as its density, natality, mortality, age distribution, etc [5]. However, individual-based models (IBMs) are becoming more popular lately, because they capture different aspects of the processes modeled. In IBMs the properties of a system are derived from the properties and interactions among elements of the systems, called *individuals* [2]. Individuals might represent plants and animals in an ecosystem, vehicles in traffic, people in crowds, etc.

In this study we introduce a new individual-based model (IBM-OSR), developed at the Scottish Crop Research Center in Dundee, Scotland, designed to help understand how life history, agronomic and environmental processes determine the persistence of genetically modified (GM) oilseed rape [1]. Encouraged by a positive experience of using machine learning for analyzing outputs from ecological simulation models [3, 4], we applied equation discovery to the output of the IBM-OSR to model the OSR population dynamic. In this paper we give a description of the IBM, the machine learning setting and the equation-discovery experiments carried out, as well as the preliminary results obtained and directions for further work.

2 DESCRIPTION OF THE IBM-OSR MODEL

The model at hand is a stochastic, individual-based model developed to simulate the dynamics of a transgene within oilseed rape populations [1]. The model combines life-history and management processes with environmental drivers to examine the effect of these on the persistence of the transgene, predict the adventitious presence of the transgene in conventional oilseed rape crops, and to test the effectiveness of management strategies in permitting co-existence with conventional oilseed rape.

The model was constructed to represent a population of oilseed rape individuals as a crop and volunteers within a single arable field. The field is defined by three state variables: *soil temperature* and *soil moisture*, which vary with time and soil depth, and *crop cover* which specifies the type of crop being grown at a given time. In addition, the field is divided into a 2-dimensional grid with grid-cells of variable dimension.

Rape individuals in this simulation model can be: *seeds*, *seedlings*, *plants*, and *seeds on plants*. They are characterized by a number of state variables. Of these, three are attributed to individuals of all types: stage, location, and transgenic status. Stage refers to the life-cycle of the individuals, which is separated into seeds present in the seed bank and plants. Location is the position occupied by the individual within the field and is referenced by simple Cartesian co-ordinates.

The population dynamics of the oilseed rape is principally driven by life-history processes which determine the progression of individuals through their life-cycle. The life-history processes modeled are *dormancy*, *germination*, *emergence*, *growth*, *flowering*, *pollination*, *seed production*, and *survival*. Interactions between individuals take place at the plant stage through the processes of growth and pollination. Both processes are spatially explicit: growth is mediated by resource competition with neighbouring individuals, while pollination combines male and female gametes from neighbouring individuals as determined by the out-crossing rate and pollen dispersal.

The model also incorporates a number of management events: *sowing*, *cultivation*, *herbicide application*, and *harvesting*. These generally act to modify the life-history processes. For example, herbicide application reduces plant survival, while cultivation reduces plant survival and alters germination and emergence by repositioning seeds within the seedbank. Top-down constraints are also imposed on the dynamics of the system through the presence of environmental and agronomic drivers. For example, soil temperature and moisture are determinants of dormancy and germination, while the crop type under cultivation influences plant growth rates.

The output of the model is the number and proportion of the GM and non-GM individuals in each stage (seeds, seedlings, plants, seeds on plants). The IBM-OSR is a relatively new model and therefore it is still not validated against empirical data. Validation using empirical data from field trials and sensitivity analyses are planned for further work.

3 OUTPUT FROM THE MODEL

Each simulation of the IBM-OSR model simulates a 10 year crop rotation on a 5m x 5m area of a field. The simulations start with a GM contaminated seedbank. In the 10 years of simulations there are always conventional crops, like winter wheat, oilseed rape and field beans.

The output of the simulation model consists of different types of information about the system:

- **Cultivation techniques** for each year and each crop grown (crop type, cultivation dates and techniques, herbicide application dates, etc.),
- **Life-history parameters**, which differ for each simulation, but are the same for every year within a simulation (death rate, germination window, growth rate, etc.),
- **Environmental parameters** for each day of the 10-year simulations (air and soil temperature, precipitation, wind, sunshine, etc.),
- **Number of individuals** in each stage (seeds, plants and seeds on plants) and each year before harvest.

The main focus of our study was the persistence of GM OSR seeds in a 10-year rotation and the influence of the life history parameters and cultivation techniques on it. The environmental parameters were at this stage omitted.

After careful consultations with domain experts, we filtered the data we had, choosing 21 attributes for further analyses, most of them being life-history parameters and a few cultivation-techniques parameters. The target attributes were the number of individuals in each stage and each year of the simulations. We had 200 simulations, each having 10 years, leaving us with 2000 examples.

4 MACHINE LEARNING SETUP

The goal of analyzing the outputs of IBM model simulations is to learn explanatory models for population dynamics of OSR. To this end, we used equation discovery. We used the ED system LAGRANGE [6, 7], for which we had to define background knowledge and code it into a context free grammar.

The life cycle of the OSR population is structured into 3 different states in which an individual can be found: *sown seed* (C), *seed rain* (yield - Y) and *seedbank* (S), each of which can be GM (G) or conventional (C). The transitions of individuals between these states are defined as functions of life-history characteristics and gene flow.

The population dynamics associated with the life-cycle of OSR can be formalized in the background knowledge as a set of difference equations that relate the state of the system at time $t+1$ to the state of the system at time t :

$$N_{t+1}=AN_t,$$

where A is the transition matrix and its coefficients are interpreted as functions of the life-history characteristics of oilseed rape and gene flow. N is the number of individuals in different stages at a given moment.

We are interested in the OSR population dynamics on the field (seedbank and seed rain), while the dynamics of sown seeds is not important at the moment. Therefore, we created 4 difference equations, for the 2 types of individuals (S and Y) and the two conditions they can be in (GM and conventional), leading to 4 context free grammars for our ED experiments. The grammars defining the population dynamics of GM and conventional seed rain (YC and YG) are almost identical, differing in small details, as well as the grammars defining the population dynamics of GM and conventional seedbank (SC and SG). Due to space limitations, in this paper we will present only the YG and SG grammars.

The life-history parameters that influence the OSR population dynamics are:

- S – annual seedbank survival rate
- G – annual germination rate (different for each type of individuals, G_s , G_y , and G_c)
- R – seed rain
- P – proportion of seeds produced by a conventional plant that are GM
- Q – proportion of seeds produced by a GM plant that are conventional
- M – annual survival rate of plants (important only for seed rain seeds, therefore we have only M_y)
- F – total seed production per plant

The derivation of the detailed functions of the life-history parameters is beyond the scope of this paper.

$$GMseedbankNEXT \rightarrow S \cdot [(1-G_y) \cdot R \cdot YG + (1-G_s) \cdot SG + (1-G_c) \cdot C];$$

$$G_y \rightarrow \frac{HarvCultDelay}{const};$$

$$G_y \rightarrow const;$$

$$HarvCultDelay \rightarrow variable_cultDelay;$$

$$G_s \rightarrow const \cdot (1 - D_{cult})^{const};$$

$$G_s \rightarrow const;$$

$$D_{cult} \rightarrow \frac{DDM}{DDF} \cdot const;$$

$$D_{cult} \rightarrow DDM \cdot \frac{0.2 - DDF}{0.2};$$

$$DDM \rightarrow variable_dormDepthMax;$$

$$DDF \rightarrow variable_dormDeptFifty;$$

$$G_c \rightarrow const;$$

$$S \rightarrow (1 - DR)^{365};$$

$$S \rightarrow (1 - DR)^{const};$$

$$S \rightarrow const;$$

$$DR \rightarrow variable_deathRate;$$

$$R \rightarrow variable_seedLoss;$$

$$YG \rightarrow variable_gmYield;$$

$$SG \rightarrow variable_gmSeedbank;$$

$$CG \rightarrow variable_gmSownSeeds;$$

Table 1: The grammar used to model the GM seedbank in year t as a function of the GM seed rain, seedbank and sown seeds in year $t-1$ using difference equations.

Table 1 presents the grammar for modeling the dynamics of the GM seedbank. $GMseedbankNEXT$ presents the number of GM individuals (seeds) in the seedbank in year t , and is a function of the GM OSR population at time $t-1$ (YG , SG and CG are the numbers of individuals in different life states in year $t-1$) and other life-history parameters.

The grammar modeling the GM seed rain dynamics is presented in Table 2. $GMyieldNEXT$ is the number of GM seed rain individuals in year t , while YC , YG , SC , SG , CC and CG are the number of individuals in all other life states in year $t-1$.

4 RESULTS

Using the 4 different grammars explained in the previous section, we generated equations for each of the stages of individuals: GM seedbank, conventional seedbank, GM seed rain (yield), and conventional seed rain.

The equations describing the OSR seed rain population are very complex due to the extensive grammar we are using to generate them and therefore are not discussed in this paper. The best equations describing the GM and conventional seedbank are presented below:

$$GMseedbankNEXT = (1 - deathRate)^{164.81} \cdot (0.13 \cdot seedLoss \cdot gmYield + 1.09 \cdot gmSeedbank + 0.81 \cdot gmSownSeeds)$$

$$ConSeedbankNEXT = (1 - deathRate)^{117.67} \cdot (0.16 \cdot seedLoss \cdot conYield + 0.82 \cdot conSeedbank)$$

$$GMyieldNEXT \rightarrow F \cdot (1 - M_y) \cdot [P \cdot G_y \cdot R \cdot YC + (1 - Q) \cdot G_y \cdot R \cdot YG + P \cdot G_s \cdot SC + (1 - Q) \cdot G_s \cdot SG + P \cdot G_c \cdot CC + (1 - Q) \cdot G_c \cdot CG];$$

$$F \rightarrow \frac{100 \cdot BM}{e^{const \cdot Dens}};$$

$$F \rightarrow 100 \cdot BM - const \cdot Dens;$$

$$F \rightarrow 100 \cdot BM - e^{const \cdot Dens};$$

$$F \rightarrow const;$$

$$BM \rightarrow variable_maxBiomass;$$

$$P \rightarrow const \cdot OC \cdot PjP;$$

$$P \rightarrow const;$$

$$Q \rightarrow const \cdot OC \cdot PjQ;$$

$$Q \rightarrow const;$$

$$OC \rightarrow variable_outcrossingRate;$$

$$PjP \rightarrow variable_pollenFractionGM;$$

$$PjQ \rightarrow variable_pollenFractionCon;$$

$$M_y \rightarrow const + const \cdot M_{seed} + const \cdot M_d + const \cdot M_c + const \cdot M_{pre} + const \cdot M_{post};$$

$$M_y \rightarrow const;$$

$$M_{seed} \rightarrow 1 - (1 - PDIM)^{const};$$

$$M_d \rightarrow 1 - e^{-const \cdot Dens};$$

$$M_c \rightarrow const;$$

$$M_{pre} \rightarrow 1 - (1 - PreM)^{PreDur};$$

$$M_{post} \rightarrow 1 - (1 - PostM)^{HerbF};$$

$$PDIM \rightarrow variable_pdimMax;$$

$$Dens \rightarrow variable_density;$$

$$PreM \rightarrow variable_preherbMort;$$

$$PostM \rightarrow variable_postherbMort;$$

$$HerbF \rightarrow variable_postherbFreq;$$

$$G_y \rightarrow \frac{HarvCultDelay}{const};$$

$$G_y \rightarrow const;$$

$$HarvCultDelay \rightarrow variable_cultDelay;$$

$$G_s \rightarrow const \cdot (1 - D_{cult})^{const};$$

$$G_s \rightarrow const;$$

$$D_{cult} \rightarrow \frac{DDM}{DDF} \cdot const;$$

$$D_{cult} \rightarrow DDM \cdot \frac{0.2 - DDF}{0.2};$$

$$DDM \rightarrow variable_dormDepthMax;$$

$$DDF \rightarrow variable_dormDeptFifty;$$

$$G_c \rightarrow const;$$

$$S \rightarrow (1 - DR)^{365};$$

$$S \rightarrow (1 - DR)^{const};$$

$$S \rightarrow const;$$

$$DR \rightarrow variable_deathRate;$$

$$R \rightarrow variable_seedLoss;$$

$$YC \rightarrow variable_conYield;$$

$$YG \rightarrow variable_gmYield;$$

$$SC \rightarrow variable_conSeedbank;$$

$$SG \rightarrow variable_gmSeedbank;$$

$$CC \rightarrow variable_conSownSeeds;$$

$$CG \rightarrow variable_gmSownSeeds;$$

Table 2: The grammar used to model the GM seed rain in year t as a function of the conventional seed rain, seedbank and sown seeds, as well as GM seed rain, seedbank and sown seeds in year $t-1$.

From the above equations we can see that the GM (or conventional) seedbank in year t depends on the GM (or conventional, respectively) seed rain (yield), seeds in the seedbank and sown seeds in year $t-1$. The structure of both equations is consistent with the domain expert opinion and is very similar, differing only in the coefficients of the equations.

The survival rate of the seeds in the seedbank is presented by the form $(1 - deathRate)^n$, where $deathRate$ is the daily mortality probability for seeds in the seedbank. Consequently, the proportion of seeds surviving over a year is given by $S = (1 - deathRate)^{365}$. 365 can be replaced by any other constant to give flexibility to the time frame we are taking into account. In this case LAGRAMGE fitted these constants to the data and chose the values 164.31 and 117.67 for the GM and conventional seedbank survival rate respectively.

The parameters that determine how much of the seeds in the seedbank, coming from the seed rain or from the sown seeds, become dormant ($1 - G_{y(s, c)}$) are set to constants. It also appears that the conventional sown seeds in year $t-1$ do not have any influence on the conventional seedbank in year t .

The predictive performance of the equations was obtained on training data, due to the high computational complexity of the equation discovery experiments, and was 0.30 and 0.26 for the GM and conventional seedbank equations respectively, and 0.24 and 0.02 for the GM and conventional seed rain equations. However, this is a first approach in modeling OSR population dynamics from outputs of an individual based model and there are still more modifications and optimizations to be done in order to improve the equation-based models that we obtained.

6 CONCLUSION

In this paper, we presented a new individual-based model, which simulates the dynamics of transgene within OSR populations. We also presented a new approach of modeling the population dynamics of OSR seeds from the output of this individual-based model by using equation discovery.

We used background knowledge encoded in the form of a grammar and applied the equation discovery system LAGRAMGE to build equation-based models. We carried out 4 different equation discovery experiments, for each of the stages the OSR population can be found in.

The structure of the models, although consistent with the domain expertise, is complex and needs further modification and improvements to obtain the needed simplification for interpretation.

Since this is the first attempt to analyze outputs from an IBM with machine learning to generate population dynamics of OSR, the lower predictive performance in terms of correlation coefficients was expected. Further work in improving the predictive performance of the models includes running new simulations with the IBM and

checking the validity of the simulated data. Running the equation discovery experiments with new and improved simulated data with different parameters may prove useful and improve the results.

Another direction for further work is reconsidering the background knowledge used in the equation-discovery process. We can provide a range of complexities of the equations included in the background knowledge from which LAGRAMGE can choose, from letting everything be a constant, to having more complex functional forms.

Finally, the use of equation discovery is a new way of analyzing outputs of individual-based models and building population dynamics models for oilseed rape. Equation discovery is a powerful tool for modeling ecological and environmental systems and combined with strong background knowledge and domain expert involvement can produce very good models.

References

- [1] G.S. Begg, M.J. Elliot, G.R. Squire, J. Copeland. Prediction, sampling and management of GM impurities in fields and harvested yields of oilseed rape. Technical report VS0126, DEFRA, 2006.
- [2] V. Grimm, S.F. Railsback. Individual-based modeling and ecology. Princeton University Press, 2005.
- [3] A. Ivanovska, C. Vens, N. Colbach, M. Debeljak, S. Džeroski. The feasibility of co-existence between conventional and genetically modified crops: Using machine learning to analyse the output of simulation models. *Ecological Modelling* (215), 262-271, 2008.
- [4] A. Ivanovska, L. Todorovski, S. Džeroski. Modelling the outcrossing between genetically-modified and conventional maize with equation discovery. *Ecological Modelling* (220), 1063-1072, 2009.
- [5] S.E. Jorgensen, G. Bendricchio. *Fundamentals of Ecological Modelling*. Elsevier Science, 2001.
- [6] L. Todorovski, S. Džeroski, B. Kompare. Modeling and prediction of phytoplankton growth with equation discovery. *Ecological Modelling* (113), 71-81, 1998.
- [7] L. Todorovski, S. Džeroski. Integrating domain knowledge in equation discovery. In *Computational Discovery of Scientific Knowledge*, 69-97, Springer, Berlin, 2007.

Explanation of regression decisions by analogy with the explanation in classification

Julian Klauser, Igor Kononenko
University of Ljubljana, Faculty of Computer and Information Science,
Ljubljana, Slovenia

ABSTRACT

Explanation of regression based machine learning algorithms makes the understanding and use of those models easier. The presented approach uses the numerical derivative to calculate the individual contribution of each attribute. The method works on a black box model basis, so it doesn't need any information about the model (with one exception that is described below). The demonstration shows that our model explanations follow the prediction patterns and allow comparison of the tested methods.

1 INTRODUCTION

The transparency of prediction is one of the most important requirements for a successful approach. Users often want to not only know what the model's prediction is, but also how the prediction was made. With a good explanation the prediction model can be easier to understand and to know how trustworthy the decision of the model really is.

To tackle this task we set up 5 different regression problem data sets. One tenth of that data is used as a testing set and the rest serves as learning data. We use five different regression models: linear regression, locally weighted regression, support vector machines, regression trees and neural nets.

The explanation consists of the numerical derivative of each instance from the testing set for all of the above mentioned regression algorithms. This means we simply produce a small change in input and observe how the output reacts. Another good feature is that we don't need to rebuild the model for each change and thus the calculation doesn't take up much time. This approach allows us to explain the model decisions from continuous attributes. However this approach can't work on discrete attributes and we use a different explanation for such attributes that don't coincide with the derivative scale.

Besides the discrete attribute problem there is also an issue with regression trees. The root of the problem is the tree structure, because it doesn't allow small changes to impact the prediction result. We'll try to devise a strategy to avoid this via comparing different prediction levels of the regression tree method.

2 BASIC APPROACHES

The data

The data sets are all custom made to better test out the validity of the results. All equations are based on 5 attributes of whom not all contribute to the regression variable (A5 usually doesn't have any defined impact on R). The number of examples of each dataset is 1000. The equation definitions and their expected derivatives are presented in the 'Results' section.

The attribute values were all generated with a uniform RNG. Also the border values (min and max of an attribute) were calculated randomly, ranging from a maximum of -100 to 100 (to prevent numbers growing out of proportion).

The derivative

To calculate how much an attribute contributes to the prediction value we use the numerical derivative. The basic equation to calculate the derivative, if we assume x as entry and y as exit parameter, is

$$d = \frac{\Delta y}{\Delta x}$$

We calculate the prediction for a particular instance 'predR' and when we change the instance's attribute A_i for ϵ (ie: a very small amount) we mark the prediction as 'predR2'. With this in mind our modified derivative equation is

$$d(A) = \frac{predR2 - predR}{\epsilon}$$

The discrete attribute problem

While the derivative of continuous variables works very well, the discrete attributes have to be tackled another way. The major problem is that the value of a discrete attribute can't be added a small number like in the continuous case and we can't measure the distances between different values.

There are many options on how to approach this. First one would be to randomly select a value from the possible discrete value levels. Thus simple, the correctness of this approach is questionable at best. The second option would be to predict all different values of the attribute for a given

instance and select the one, that produces either the biggest or smallest difference. This approach is also debatable, but it seems like it's not the safest bet.

So we decided to go with the third option and that is to calculate the average change in prediction over all possible values of a discrete attribute. Sadly this scale does not match the continuous attribute scale and we had to mark the derivative values differently to point out the difference between them (discrete attributes are marked with a '*').

The regression tree problem

Whilst testing we encountered the problem that all derivatives of the regression trees were exactly zero. The problem is that regression trees predict values very 'step like'. So a small change in the input vector doesn't result in any output change. So we had to add a special version of our derivative method to handle this.

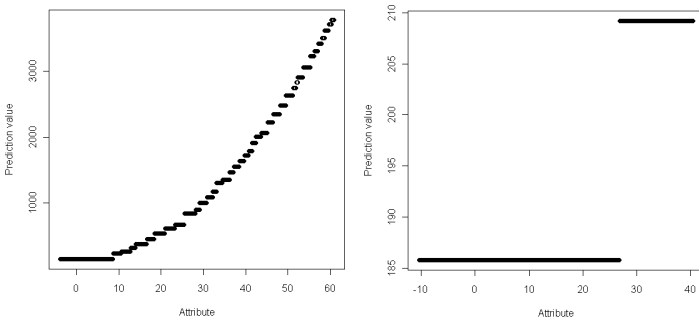


Figure 1. Visualisation of the steps caused by regression trees, when changing an attribute value. On the left: Square function attribute dependency 'steps'. On the right: Linear function attribute dependency 'steps'.

In our method we calculated the borders of the 'step' that the current instance was on. After that we did the same for the neighbouring left or right 'step', depending on what side of the step center our instance was positioned. Followed by a quick calculation between the neighbouring step centers we now got our 'step derivative' (see Fig. 1). Although it's accuracy probably won't be the same as for other models, it does work out well. The only problem was the calculation speed, but we're sure that the method still leaves a lot of room for improvement (bisection, trisection, etc.).

3 RESULTS AND THEIR VISUALIZATIONS

In this section the testing results and its visualizations are presented. The explanation of classification decisions that has been described in [2] and the explanation of decisions of the naïve Bayesian regressor (can be found in [3]) both served as an inspiration on what data sets could be used and how to visualise the derivatives.

For each testing set the equation for R is stated, followed by the expected derivatives. If an attribute does not appear in the equation, then it's derivative should be zero. The

measures RMAE and RMSE (both described in [1]) show the difference between the prediction and the actual values of the testing set. A larger RMAE or RMSE number indicates bad prediction accuracy. The default value for ϵ is 0.00001 (effects of different ϵ values are described in the 'Conclusion' section). After picking a typical instance of the testing set, we visualized it's derivative in a graphical form.

A) $R = A1 + 3 * A2 - 2 * A3$

Observed instance:	Expected derivatives:
A1 = 25.32	$d_{exp}(A1) = 1$
A2 = -19.99	$d_{exp}(A2) = 3$
A3 = -14.52	$d_{exp}(A3) = -2$
A4 = -11.9	$d_{exp}(A4) = 0$
A5 = 10.97	$d_{exp}(A5) = 0$

Model	RMAE	RMSE
Linear Regression	0.0	0.0
Locally weighted regression	0.0	0.0
Support vector machines	0.076	0.005
Regression trees	0.254	0.059
Neural nets	0.199	0.046

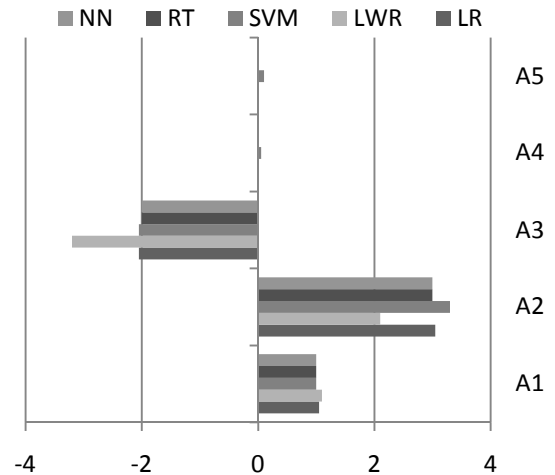


Figure 2: derivatives of example A

B) $R = A12 + 2 * A2 + A3 + 0.2 * A4$

Observed instance:	Expected derivatives:
A1 = 12.8	$d_{exp}(A1) = 2 * A1 = \sim 25.6$
A2 = 48.6	$d_{exp}(A2) = 2$
A3 = -9.08	$d_{exp}(A3) = 1$
A4 = -6.06	$d_{exp}(A4) = 0.2$
A5 = 34.02	$d_{exp}(A5) = 0$

Model	RMAE	RMSE
Linear Regression	0.24	0.06
Locally weighted regression	0.0	0.0
Support vector machines	0.075	0.006
Regression trees	0.102	0.012
Neural nets	0.052	0.003

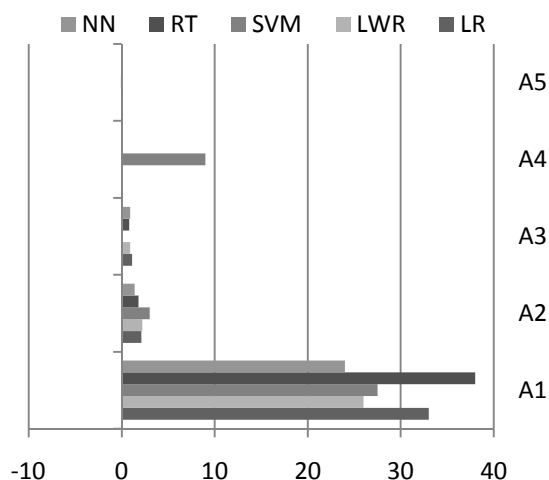


Figure 3: derivatives of example B

C) $R = \min(A1, \max(A2, A3))$

Observed instance:	Expected derivatives:
A1 = 4.97	$d_{exp}(A1) = \sim -1$
A2 = 5.77	$d_{exp}(A2) = \sim 0$
A3 = 32.74	$d_{exp}(A3) = \sim 0$
A4 = -6.12	$d_{exp}(A4) = 0$
A5 = 22.19	$d_{exp}(A5) = 0$

Model	RMAE	RMSE
Linear Regression	0.308	0.142
Locally weighted regression	0.120	0.025
Support vector machines	0.117	0.019
Regression trees	0.104	0.026
Neural nets	0.172	0.049

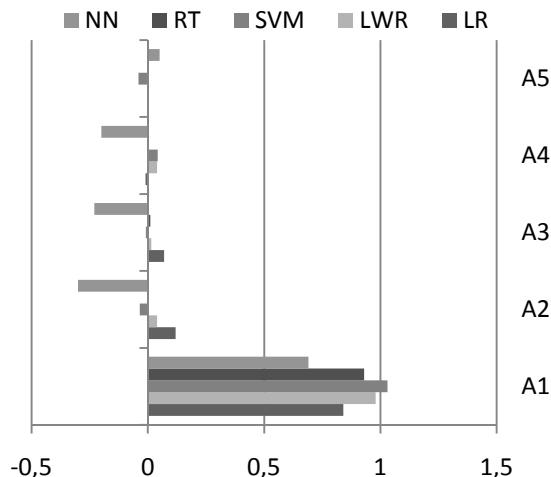


Figure 4: derivatives of example C

D) *Data set with discrete attribute A4. Levels: [A,B,C,D]*

$R = 10 * A1 + 5 * A2 - 5 * A3$; IF A4 = 'A' OR 'B'
 $R = -1 * A1 + 5 * A2 - 5 * A3$; IF A4 = 'C'
 $R = 5 * A2 - 5 * A3$; IF A4 = 'D'

Observed instance:	Expected derivatives:
A1 = 62.57	$d_{exp}(A1) = 10$ (because of A4 = A)
A2 = 7.079	$d_{exp}(A2) = 5$
A3 = 52.08	$d_{exp}(A3) = -5$
A4 = A	$d_{exp}(A4) = ?$
A5 = 38.63	$d_{exp}(A5) = 0$

Model	RMAE	RMSE
Linear Regression	0.42	0.16
Support vector machines	0.082	0.009
Neural nets	0.053	0.004

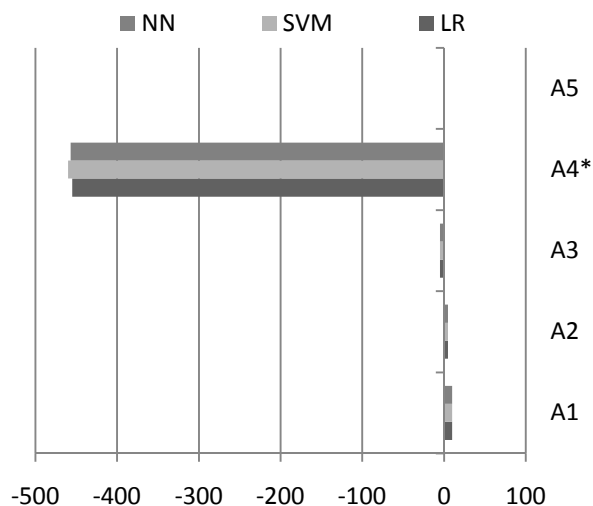


Figure 5: derivatives of example D

E) **Data set with discrete attributes A1, A2 [levels: a,b,c], A3 [levels: a,b,c,d]**

$$R = 1 * A1 + 2 * A2 + 3 * A3$$

where $A_i = 1$ if $A_i = 'A'$, else $A_i = 0$

Observed instance:	Expected derivatives:
A1 = 62.57	$d_{exp}(A1) = \sim 1$
A2 = 7.079	$d_{exp}(A2) = \sim 2$
A3 = 52.084	$d_{exp}(A3) = \sim 3$
A4 = A	$d_{exp}(A4) = 0$
A5 = 38.6	$d_{exp}(A5) = 0$

Model	RMAE	RMSE
Linear Regression	0.0	0.0
Support vector machines	0.067	0.005
Neural nets	0.032	0.001

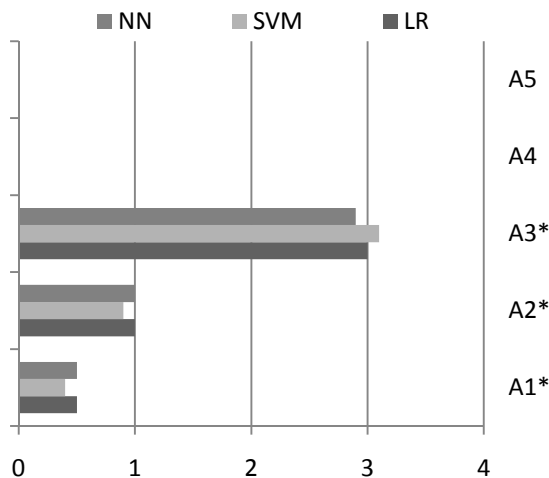


Figure 6: derivatives of example E

6 CONCLUSION

We presented an approach for explanation of regression predictions which generates explanations for individual instances. The approach can be used with all the tested prediction methods used in testing and in theory works for any method. We also noticed that prediction error (in this paper presented with RMAE and RMSE) greatly influences the output of the derivatives, which was to be expected. For future improvement the accuracy of some methods could be increased and would allow better comparison between the methods.

While the methods were tested, many different ϵ values were tried out. The differences were very small, but it turned out that on complex data sets the accuracy is better with a very small ϵ (ie. the smallest possible). Also the approach of trying with $-\epsilon$ didn't show any greater changes in the output, so it suffices to say that the impact on explanations is negligible.

The overall efficiency of the derivative was successful. In

data set 'A' (Fig. 2.) the accuracy of the derivative is determined by the accuracy of the prediction model. Linear regression and locally weighted regression managed a 100% prediction accuracy (and thus giving perfect derivative numbers) due to the simplicity of the problem and the way the methods work.

As the difficulty of the problems increased in data set 'B' (Fig. 3.), linear regression loses a lot of accuracy and thus the derivatives differ from the expected results. But this doesn't mean the explanation of the prediction is bad, the prediction accuracy is at fault here.

Overall the non-contributing attributes (the ones with the expected derivative 0) were correctly assigned by most algorithms.

With increasing differences between the derivatives of attributes in data set 'C' (Fig. 4.), the accuracy of the smaller derivatives is getting lower. If this is a problem of the derivative method or just another side effect of bad prediction accuracy remains unclear.

The discrete 'derivative' works out well when only discrete attributes are used (on simple examples, like data set 'E', Fig. 6.). But when mixed with continuous attributes the scales can't really be compared as the results from data set 'D' prove (Fig. 5.). The influence of the discrete attribute is too big and diminishes the continuous attributes.

Assuming explanations with divided discrete and continuous attribute scales are acceptable, the derivative is successfully providing good regression decision explanation with all tested methods.

References

- [1] I. Kononenko, *Machine Learning (in Slovene)*, 2nd edition, Ljubljana: Faculty of Computer and Information Science, 2005.
- [2] I. Kononenko, R.Šikonja, „Explaining Classifications for Individual instances”, *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 5, pp. 589-600, May 2008.
- [3] N. Zenkovič, “Razlaga predikcij naivnega Bayesovega regresorja” Bachelor Thesis, Ljubljana: University of Ljubljana, Faculty of Computer and Information Science, 2008.

REGRESSION AS COST-SENSITIVE CLASSIFICATION

Egon Kocjan, Igor Kononenko
Faculty of computer and information science
University of Ljubljana
Tržaska 25, 1000 Ljubljana, Slovenia
Tel: +386 1 4768390; fax: +386 1 4264647
e-mail: egon.kocjan@xlab.si,
xaigor@fri.uni-lj.si

ABSTRACT

In this paper we investigate the use of classifiers on regression problems. The methods dependent variable discretization and cost-sensitive classifier learning are used in order to adapt the classifiers to the regression. Cost-sensitive learning is guided by a distance based cost matrix in hope of reducing the effect of information loss caused by discretization. Various classifiers were tested and an improvement of prediction accuracy was noticed for rule based classifiers with cost-sensitive learning.

1 INTRODUCTION

Classifiers are one of the main topics of machine learning. The goal of classification is to construct a model, which can determine the class of the instance (object). Classifiers are built with an automatic process of learning. Classifier learner is an algorithm that takes a set of learning instances (training set) as an input and produces the classifier.

There is a large number of existing classification algorithms. By adapting the classification to regression problems, we reuse and acknowledge the previous work done on classification. The major groups of classification algorithms are:

1. decision trees and rules,
2. Bayes classifiers,
3. nearest neighbor classifiers,
4. discriminant functions,
5. neural networks,
6. hybrid algorithms.

The paper will focus mostly on decision trees and rules, because this group of algorithms proved to be the most successful in the initial performance evaluations. Nevertheless, all the methods described in the paper may be used without any additional adaption on groups 2-6.

A large part of the work presented in the paper is based on existing and related work. The work is split into two major parts.

In the first part, we research the problem of adapting the classification to regression problems. Classifiers expect nominal values for the class, so we need to discretize the

dependent regression variable. If we wish to use the resulting classifier operating on discretized dependent variable as a regression algorithm, a mapping function from the nominal value back to a numeric value is needed. The exact methods were described in the work of Torgo & Gama [1].

In the second part, we introduce a cost-learning scheme in hope of reducing the information loss caused by discretization of the dependent variable. The general method of adapting the various classifiers to cost-sensitive learning is based on work of Domingos [2]. The use of discretized interval distance based cost matrix is evaluated with various classifiers.

2 REGRESSION BY CLASSIFICATION

Regression problems contain a number a samples in the form of predictor variables $x_1 \dots x_n$ and the dependent variable y . The goal of solving the regression problem is to find the relationship between the predictor variables and the dependent variable. The relationship is expressed as a function $y = f(x_1 \dots x_n)$. The dependent variable is numeric in regression, whereas the class attribute is nominal in classification. Regression thus cannot be solved directly by classification.

2.1 Regression by classification method

Regression by classification, described by Torgo & Gama [1], is a method that adapts classifiers to be used on regression problems. The basic idea is to use a discretization on the dependent variable to obtain nominal class variable. As a result, we can use a classifier on the resulting dataset using a function that maps classifier output into the dependent regression variable.

2.2 Discretization

Discretization is a method to split the interval of possible continuous values into a set of intervals that can be used as nominal values. The authors in [1] describe three approaches to discretization, each of them using the number of the output intervals as a parameter (N):

- equally probable intervals (N intervals are created, each containing the same number of elements),
- equal width intervals (N intervals are created, each of them has the same width),
- K-means clustering (N intervals are created, the sum of the distances of each elements is minimized based on the gravity center of the interval; initial set of intervals is defined by Equally probable intervals).

We use the discretization method equal width intervals, because it is simple to understand and it gives satisfactory results, as can be seen later in this paper.

2.3 Mapping classification back into regression

We use the following equation to map the output of the classifier into the dependent variable:

$$y = \frac{\sum_j p_j m_j}{\sum_j p_j} \quad (1)$$

y : dependent variable

p_j : class probability

m_j : interval mean on training set of instances

Single class output is treated as a probability vector: [0, ..., 0, 1, 0, ..., 0].

2.4 Finding the optimal discretization parameters

The authors in [1] describe two possible approaches to adjusting the discretization parameters:

- varying the number of intervals (various number of intervals (N) are tried to find the best performing classifier),
- selective specialization of individual classes (iterative process of splitting intervals, each interval is tested for error estimate. Intervals with an error above calculated threshold are split).

We use the approach varying the number of intervals, because it is simple to understand and it gives satisfactory results, as can be seen later in this paper.

3 COST-SENSITIVE LEARNING

Class misclassification costs provide additional information to the classifier learning algorithm and the classifier. Usually, the class distribution in data sets is not perfectly balanced according to the relative importance of the classes. Therefore, some classes are underrepresented and some are overrepresented. Most classifiers try to maximize the classification accuracy and decrease the complexity of the learned theory. As a result, it may seem worthwhile to the classifier learning algorithm to produce a classifier which simply always classifies the minority class as the majority class.

Most classifiers do not support misclassification cost information directly. There are various methods of pre-filtering the data sets to include misclassification cost information. We use MetaCost, because of its generality.

3.1 MetaCost

MetaCost, described by Domingos [2], is a method to make any classifier cost-sensitive without specifically adapting the classifier learning process to be cost-sensitive. MetaCost is designed to be general: the behavior of the classifier does not need to be known in advance. There are no restrictions regarding the number of classes or cost matrices.

The following equation is used to calculate the conditional risk (of class i) [3]:

$$R(i|x) = \sum_j P(j|x)C(i,j)$$

x : given example

$P(j|x)$: probability of class j

$C(i,j)$: cost of predicting the class to be i instead of j

The algorithm works as follows:

1. create bagging sets of the training set,
2. learn classifier on each bagging set,
3. estimate class probability for each training example using classifiers from step 2,
4. change the training example's class by calculating the minimum conditional risk $R(i|x)$,
5. generate new classifier on the new training set with corrected classes

3.2 Discretization and MetaCost

Discretizing the numeric attribute necessarily reduces information, because multiple values are combined into a single discrete interval. We can observe two particular properties of the numeric attributes:

- ordering,
- magnitude.

It is possible to encode the difference in magnitude with the equation $|r_i - r_j|$. We can use the additional information as a misclassification cost (cost of predicting the class to be i instead of j):

$$C(i,j) = |i - j|$$

A simplified measure $|i - j|$ may be used, because equal width intervals are used.

4 EMPIRICAL EVALUATION

Evaluation of the algorithms was done with 10-fold cross validation. Cross validation was run 10 times for each algorithm. Mean absolute error (MAE) was chosen as the measure of prediction quality, because it describes the actual error and it is widely used. The values of column MAE in evaluation tables were calculated as the average of MAE for all cross validation runs. Standard deviation is marked with the sign \pm .

$$MAE = \frac{1}{N} \sum_{i=1}^N |f(i) - \hat{f}(i)|$$

$f(i)$: the expected value

$\hat{f}(i)$: the predicted value

N : number of values

Relative mean absolute error (RMAE) was calculated to illustrate the prediction accuracy. The values of column RMAE in evaluation tables were calculated as the average of RMAE for all cross validation runs. Standard deviation is marked with the sign \pm .

$$RMAE = \frac{\sum_i^N |f(i) - \hat{f}(i)|}{\sum_i^N |f(i) - \bar{f}|}, \quad \bar{f} = \frac{1}{N} \sum_i^N f(i)$$

$f(i)$: the expected value

$\hat{f}(i)$: the predicted value

N : number of values

4.1 Classification algorithms

Many classification algorithms were tested in the process of evaluation, not all of them are presented in the paper. The algorithms presented in this paper share two common features:

- numeric dependent variable is not supported,
- a noticeable improvement of prediction quality is achieved when using cost-sensitive learning.

All classification algorithms were run in two setups - without cost-sensitive learning (Table 1) and with cost-sensitive learning (Table 2). Thus we can clearly see the improvement gained by cost-sensitivity. Classifiers with cost-sensitive learning are marked with italic *C* in the evaluation tables 3-7. Bold text designates the algorithm variant, that indicates better prediction quality.

The discretization algorithm was run with 10-bins parameter. Several other values from 2 to 50 were tested, but did not provide a significant improvement.

The list of classification algorithms:

- C4.5 [6]: Widely used decision tree classifier. There are many published evaluation results using the C4.5 trees. Evaluation results are easy to verify,
- BFTree [8]: Best first decision tree classifier,
- RIPPER [7]: Repeated Incremental Pruning to Produce Error Reduction - rule based classifier,
- NNG [11]: Nearest neighbour with generalization - rule based classifier,
- Ridor [10]: Ripple-Down Rule Learner.

4.2 Regression algorithms

Two additional and well known algorithms were chosen as a base metric for comparison:

- M5 regression trees [4][12],
- K-nearest neighbours [5].

Bold text designates the algorithm, that indicates better prediction quality.

4.3 Data sets

Algorithms were evaluated on a single synthetic data set and four real-world data sets.

The list of data sets:

- synthetic data set puma8NH [9]: There are 8192 cases, 9 continuous attributes. Not all of the algorithms were evaluated because of time and space reasons. Evaluation results are in table 3,
- real-world data set auto-mpg [9]: The number of cases is 398. There are 3 discrete and 5 continuous attributes. Evaluation results are in table 4,
- real-world data set machine-cpu [9]: The number of cases is 209. There are 6 continuous attributes. Evaluation results are in table 5,
- real-world data set servo [9]: The number of cases is 167. There are 4 discrete and 1 continuous attribute. Evaluation results are in table 6,
- real-world data set housing [9]: The number of cases is 506. There is 1 discrete and 13 continuous attributes. Evaluation results are in table 7.

4.4 Evaluation conclusion

M5 regression tree proved to have the lowest mean absolute error and it clearly has the best prediction quality if we base our decision on MAE.

Decision trees did not indicate a significant amount of improvement when using cost-sensitive learning. C4.5 and BFTree had slightly lower MAE in some test cases.

Distance based cost matrix had a larger effect on rule based classifiers, however. RIPPER improved drastically and consistently in all test cases when using cost-sensitive learning.

5 FUTURE WORK

A simple equal width algorithm for discretization was used. It would be worthwhile to try other methods of discretization. There are several supervised methods for discretization (based on MDL, ReliefF), which might split the continuous attribute range into more appropriate intervals.

We used a simple distance based cost matrix. Further research in more refined cost matrix models as well as other cost-sensitive pre-filters is needed.

Classifier RIPPER reacted drastically to cost-sensitive learning. Further research would be needed to determine the cause.

6 CONCLUSION

We have described the process of identifying the regression problem, conversion of the problem into classification and use of the classifiers to solve the problem. A reduction in mean absolute (MAE) error was noticed, when using cost-sensitive learning with distance based cost matrix on rule based classifiers. The reasons for the reduction in MAE need to be studied further. M5 regression tree had the lowest MAE of all evaluated algorithms.

N : number of bins

D : data set

L : classifier learner

1. DC = discretize data set D with N equal width bins
2. RC = build Regression by discretization classifier by combining L and the mapping from Equation 1.
3. perform cross validation with classifier RC and data set D

Table 1: Cross validation of a classifier without cost-sensitive learning

N : number of bins

D : data set

L : classifier learner

1. DC = discretize data set D with N equal width bins
2. CM = generate cost matrix for N bins: $C(i,j) = |i-j|$
3. MC = build *MetaCost* classifier with learner L and cost matrix CM
4. RC = build Regression by discretization classifier by combining MC and the mapping from Equation 1
5. perform cross validation with classifier RC and data set D

Table 2: Cross validation of a classifier with cost-sensitive learning

Algorithm	MAE	RMAE
M5	2.463 ±0.002	50.62% ±0.04
KNN	3.813 ±0.001	78.35% ±0.04
C4.5	3.131 ±0.014	64.35% ±0.28
C4.5 C	2.942 ±0.006	60.45% ±0.12
RIPPER	4.346 ±0.001	89.32% ±0.03
RIPPER C	3.888 ±0.011	79.90% ±0.23
NNG	3.468 ±0.015	71.28% ±0.30
NNG C	3.462 ±0.038	71.16% ±0.79

Table 3: Evaluation results on synthetic data set puma8NH

Algorithm	MAE	RMAE
M5	2.010 ±0.033	30.69% ±0.50
KNN	2.623 ±0.025	40.05% ±0.36
C4.5	2.601 ±0.036	39.69% ±0.49
C4.5 C	2.641 ±0.118	40.37% ±1.79
BFTree	2.675 ±0.051	40.86% ±0.76
BFTree C	2.564 ±0.058	39.16% ±0.89
RIPPER	4.555 ±0.160	69.65% ±2.44
RIPPER C	3.029 ±0.101	46.17% ±1.58
NNG	2.665 ±0.149	40.70% ±2.30
NNG C	2.630 ±0.052	40.15% ±0.82
Ridor	3.462 ±0.074	52.89% ±1.13
Ridor C	3.275 ±0.056	49.98% ±0.85

Table 4: Evaluation results on real-world data set auto-mpg

Algorithm	MAE	RMAE
M5	30.23 ±1.09	31.43% ±1.13
KNN	32.48 ±0.93	33.71% ±0.93
C4.5	39.95 ±2.55	41.40% ±2.68
C4.5 C	40.70 ±1.79	42.22% ±1.77
BFTree	47.63 ±1.73	49.38% ±1.64
BFTree C	44.10 ± 0.47	45.68% ±0.49
RIPPER	66.25 ±2.52	68.77% ±2.57
RIPPER C	48.54 ±1.33	50.27% ±1.33
NNG	39.70 ±1.30	41.20% ±1.40
NNG C	40.15 ±1.03	41.56% ±1.10
Ridor	48.27 ±1.95	50.05% ±2.04
Ridor C	47.04 ±1.46	48.79% ±1.60

Table 5: Evaluation results on real-world data set machine-cpu

Algorithm	MAE	RMAE
M5	0.303 ±0.010	26.19% ±0.83
KNN	0.524 ±0.011	45.19% ±0.86
C4.5	0.415 ±0.013	35.82% ±1.06
C4.5 C	0.416 ±0.032	35.83% ±2.71
BFTree	0.325 ±0.021	28.00% ±1.81
BFTree C	0.359 ±0.021	31.00% ±1.88
RIPPER	0.519 ±0.049	44.79% ±4.23
RIPPER C	0.409 ±0.027	35.43% ±2.34
NNG	0.420 ±0.039	36.21% ±3.40
NNG C	0.417 ±0.062	35.91% ±5.34
Ridor	0.547 ±0.038	47.17% ±3.39
Ridor C	0.539 ±0.050	46.55% ±4.41

Table 6: Evaluation results on real-world data set servo

Algorithm	MAE	RMAE
M5	2.495 ±0.060	37.45% ±0.85
KNN	3.009 ±0.047	45.19% ±0.72
C4.5	3.014 ±0.115	45.29% ±1.75
C4.5 C	2.994 ±0.071	44.96% ±1.05
BFTree	2.982 ±0.054	44.73% ±0.78
BFTree C	2.919 ±0.062	43.83% ±0.94
RIPPER	3.229 ±0.153	48.43% ±2.29
RIPPER C	3.046 ±0.080	45.70% ±1.18
NNG	3.111 ±0.099	46.73% ±1.47
NNG C	3.079 ±0.074	46.19% ±1.10
Ridor	3.491 ±0.154	52.41% ±2.34
Ridor C	3.206 ±0.087	48.18% ±1.30

Table 7: Evaluation results on real-world data set housing

References

- [1] L. Torgo, J. Gama, "Regression by Classification," *In Proceedings of SBIA'96*, Springer-Verlag, 1996, pp. 51-60
- [2] P. Domingos, "MetaCost: A General Method for Making Classifiers Cost-Sensitive," *In Proceedings of the Fifth International Conference on Knowledge Discovery and Data Mining*, ACM Press, 1999, pp. 155-164
- [3] R. O. Duda. P. E. Hart, *Pattern Classification and Scene Analysis*, Wiley, New York, NY, 1973
- [4] R. J. Quinlan, "Learning with Continuous Classes," *5th Australian Joint Conference on Artificial Intelligence*, Singapore: World Scientific, 1992, pp. 343-348
- [5] D. Aha, D. Kibler, "Instance-based learning algorithms," *Machine Learning*, vol. 6, pp. 37-66, 1991
- [6] R. Quinlan, *C4.5: Programs for Machine Learning*, San Mateo, CA: Morgan Kaufmann Publishers, 1993
- [7] W. W. Cohen, "Fast Effective Rule Induction," *Twelfth International Conference on Machine Learning*, Morgan Kaufmann, 1995, pp. 115-123
- [8] H. Shi, "Best-first decision tree learning," M.S. thesis, University of Waikato, Hamilton, NZ, 2007
- [9] L. Torgo, A jarfile containing 30 regression datasets [Online]: <http://weka.wiki.sourceforge.net/Datasets>
- [10] B. R. Gaines, P. Compton, "Induction of Ripple-Downm Rules Applied to Modeling Large Databases," *J. Intell. Inf. Syst.*, vol. 5, num. 3, pp. 211-228, 1995
- [11] B. Martin, "Instance-Based learning: Nearest Neighbor With Generalization," M. S. thesis, University of Waikato, Hamilton, NZ, 1995
- [12] Y. Wang, I. H. Witten, "Induction of model trees for predicting continuous classes," *Poster papers of the 9th European Conference on Machine Learning*, Springer, 1997

PROBLEM PRED-TESTNE CENILKE NA PRIMERU EQ5D

Marko Ogorevc

Inštitut za ekonomska raziskovanja

Kardeljeva pl. 17, 1109 Ljubljana, Slovenija

Tel: +386 1 5303836; fax: +386 1 5303874

e-mail: ogorevc@ier.si

Povzetek

V članku so predstavljene glavne težave povezane z ocenjevanjem zdravstvenih stanj z instrumentom EQ5D. Problemi nastopijo, ko raziskovalci predhodno ne postavijo hipoteze, za določitev kriterijev oziroma pojasnjevalnih spremenljivk, ki nastopajo v modelu pa uporabijo »stepwise« regresijo. Omenjena metoda je podvržena pristranskosti pred testiranjem in še nekaterim drugim nepravilnostim, kot so vključitev napačne spremenljivke in sprejetje napačnega modela. Kot rešitev problema (velja na splošno kadar imamo opravka z regresijo z nepravilnimi spremenljivkami) je ponujena metoda CART.

1 Uvod

Instrument EQ5D, ki je eden od generičnih instrumentov za merjenje kakovosti življenja, je sestavljen iz dveh delov. V prvem delu osebe ocenijo svoje trenutno zdravstveno stanje na podlagi petih dimenzij, v drugem delu pa ocenijo 16 stanj, ki so prav tako opisana z istimi petimi dimenzijami (*pokretnost, skrb zase, vsakodnevne aktivnosti, bolečina / neugodje in tesnoba / potrtost*). Vsaka izmed dimenzij ima tri enote v zalogi vrednosti, in sicer:

$Z_f = \{ \text{brez težav, nekaj težav, izjemne težave} \}$

Anketiranci (osebe) ocenjujejo zdravstvena stanja tako, da jim dodelijo vrednosti od 0 do 100, kjer 0 pomeni najslabše možno zdravstveno stanje, ki si ga lahko zamislijo, 100 pa opisuje najboljše možno zdravstveno stanje oziroma počutje pri takšnem stanju. V pomoč jim je 20 cm visoka ocenjevalna lestvica (VAS – *angl. Visual Analogue Scale*). Nato se na podlagi določenega vzorca

oseb in stanj izračuna oziroma določijo vsa preostala stanja.

2 Pred-testna cenilka

V literaturi o EQ5D (Cabases et al, 2000, Macran in Kind, 2000, Lubetkin in Gold, 2000) je razvidno, da se za ocenjevanje stanj uporablja pristop z umetnimi spremenljivkami. Napaka, ki so jo storili vsi omenjeni avtorji, je ta, da niso podali hipoteze, ki so jo preverjali. Čeprav trivialna, je nujna za verifikacijo modela. Ker niso na začetku sklepali o obliki modela, so uporabili »stepwise« regresijo, ki je na podlagi določene stopnje tveganja $\alpha=0,05$ v model vključevala spremenljivke in nato testirala njihovo statistično natančnost. V kolikor je bila stopnja tveganja manjša od izbrane (ali absolutna vrednost t-statistike večja od 1,96), je spremenljivka ostala v modelu, sicer v modelu ni bila upoštevana. Tako so, ne da bi se zavedali, vsi po vrsti naredili napako, ki je v literaturi znana pod imenom »pristranskost zaradi predhodnega testiranja« (*angl. Pretest bias*) (Magnus in Durbin, 1999, Magnus 2008).

Predpostavimo, da se odločamo o izbiri med omejenim in neomejenim modelom oziroma, ali naj v model vključimo dodatno spremenljivko (i.e. neomejen model). Oba modela sta linearna in imata naslednjo obliko (Magnus, 2008):

$$\text{Omejen: } Y = \beta X + \varepsilon \quad (2)$$

$$\text{Neomejen: } Y = \beta X + \gamma z + \varepsilon \quad (3)$$

Kjer je Y vektor ($N \times 1$) opazovanj pojasnjene spremenljivke, X matrika ($N \times k$) opazovanj k eksogenih spremenljivk, z je vektor ($N \times 1$) N opazovanj dodatne pojasnjevalne spremenljivke, za katero se odločamo, ali jo

dodamo v model, ε pa je vektor ($N \times 1$) vrednosti inovacij.

Cenilka OLS v omejenem modelu je definirana kot:

$$b_{omejen} = (X'X)^{-1}X'y \quad (4)$$

Če definiramo

$$M = I_n - X(X'X)^{-1}X' \quad (5)$$

$$q = \frac{\sigma}{\sqrt{z'Mz}}(X'X)^{-1}X'z \quad (6)$$

$$\theta = \frac{\gamma}{\sigma/\sqrt{z'Mz}} \quad (7)$$

potem lahko zapišemo cenilki za neomejen model (model, ki ima vključene vse potencialne spremenljivke) kot (Magnus, 2008):

$$b_{neomejen} = b_{omejen} - \hat{\theta}q \quad (8)$$

$$\hat{\gamma} = \frac{z'My}{z'Mz} \quad (9)$$

kjer

$$\hat{\theta} = \frac{\hat{\gamma}}{\sigma/\sqrt{z'Mz}} \sim N(0,1) \quad (10)$$

predstavlja t-statistiko, ki se v tem primeru porazdeljuje normalno (ker predpostavljamo, da je σ znan), z aritmetično sredino 0 in standardnim odklonom 1. θ imenijemo teoretično t-razmerje (t-statistika).

Ker se ne ve, kateri model bi izbrali oziroma ali bi v model vključili dodatno spremenljivko, se raziskovalci običajno poslužujejo predhodnega testiranja – »predtest«. Na podlagi slednjega se dodatna spremenljivka z vključi v model, če je t-statistika velika, in izključi, če je majhna (pri točni stopnji tveganja $\alpha = 0,05$ znaša $\theta = 1,96$). To pa pripelje do pred-testne cenilke, ki ima naslednjo obliko:

$$b = \begin{cases} b_{omejen} & \text{če je } |\hat{\theta}| \leq c, \\ b_{neomejen} & \text{če je } |\hat{\theta}| > c, \end{cases} \quad (11)$$

kjer je c neko pozitivno število (na primer 1,96). Za bolj nazoren prikaz problema lahko cenilko zapišemo tudi kot:

$$b = \lambda b_{omejen} + (1 - \lambda)b_{neomejen} \quad (12)$$

kjer je

$$\lambda = \begin{cases} 0 & \text{če je } |\hat{\theta}| \leq c, \\ 1 & \text{če je } |\hat{\theta}| > c. \end{cases} \quad (13)$$

Omenjena enačba (12) nazorno prikazuje, da je pred-testna cenilka tehtano povprečje vseh cenilk, ki so na voljo pri določenem modelu. Uteži (pri tehtanem povprečju) pa so naključne spremenljivke, ki so odvisne od λ . Pred-testna cenilka je tako zelo komplicirana nelinearna cenilka, ki pa je ne moremo oceniti z OLS (Magnus, 2008). Problem ni toliko v tem, da raziskovalci (ekonomisti in ekonometriki) uporabljajo predhodno testiranje, temveč se kaže v tem, da ne upoštevajo posledic takšnega početja. V praksi (kot tudi v »stepwise« regresiji) se problema lotevajo/mo na naslednji način. Na začetku imamo množico potencialnih modelov, med katerimi s predhodnim testiranjem na podlagi t-statistike in ostalih diagnostik izberemo model, ki nam (oziroma podatkom) najbolj ustreza. V drugem delu poročamo o ocenah koeficientov in njihovi standardnih napakah. Za te pa se po navadi predpostavlja, da so dobljene po standardni OLS metodi, torej da so ocene nepristranske. Ta predpostavka pa je, kot je prikazano zgoraj, napačna. Ocene so pristranske in standardne napake niso rezultat OLS; to je problem predhodnega testiranja oziroma pred-testne cenilke.

3 Problem: »Stepwise« regresija

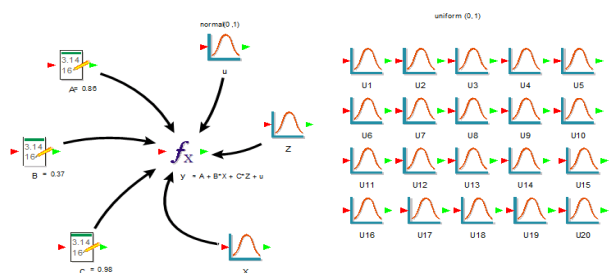
Poleg pristranskosti in nepravilnih poročanj o standardnih napakah, se pri »stepwise« regresiji kažejo napake pri testiranju statistične značilnosti. Če vzamemo stopnjo tveganja $\alpha = 0,05$ in več potencialnih spremenljivk, ki bi jih lahko vključili v model, potem ta stopnja tveganja pomeni, da bo vsaj 1 izmed 20 potencialnih pojasnjevalnih spremenljivk vključena v model, čeprav ne obstaja nikakršna odvisnost med njo in odvisno spremenljivko. Slednje najlažje ponazorimo s primerom.

Če predpostavljamo, da je slučajna spremenljivka Y generirana na način:

$$Y = \alpha + \beta x + \gamma z + \varepsilon, \quad \varepsilon \sim N(0,1) \quad (14)$$

kjer je $\alpha = 0,86$; $\beta = 0,37$; $\gamma = 0,98$; ε je (stohastična) nepojasnjena napaka oziroma ostanki regresije, ki se porazdeljuje normalno, z aritmetično sredino 0 in standardnim odklonom 1, poleg tega pa generiramo še 20 naključnih spremenljivk (U_i) iz enakomerne porazdelitve z mejama 0 in 1. Model, ki je narejen s programskim orodjem GoldSim, je prikazan na Sliki 1.

Slika 1: Prikaz modela (14)



Vir: Lastni vir, GoldSim, 2009

Namen modela je, da generira $N=10.000$ vrednosti za neodvisni spremenljivki x in z ter določi vrednosti odvisne spremenljivke Y , ki poleg konstantnega člena a vsebuje še inovacije ε , ki se porazdeljujejo po standardizirani normalni porazdelitvi. Iz dobljenih podatkov nato v SPSS z metodo »stepwise« regresijo poskušamo določiti prvotni model (14). Hipoteza, ki jo preverjamo, je, ali bo v model vključena vsaj ena slučajna spremenljivka U_i . V nadaljevanju so podani rezultati »stepwise« regresije.

$$H_0: U_i \text{ ni v modelu (14)}$$

$$H_1: \text{Vsaj en } U_i \text{ je v modelu (14); } i = 1, \dots, 20$$

Tabela 1: Izpis iz programskega paketa SPSS

Model	R	R ²	R ² adj	S.N. ocene
1	0,256 ^a	0,065	0,065	1,0071
2	0,280 ^b	0,079	0,078	1,0000
3	0,281 ^c	0,079	0,079	0,9998
4	0,282 ^d	0,080	0,079	0,9996
5	0,283 ^e	0,080	0,079	0,9994

Opombe:

a. Predictors: (Constant), Z

b. Predictors: (Constant), Z, X

c. Predictors: (Constant), Z, X, U9

d. Predictors: (Constant), Z, X, U9, U15

e. Predictors: (Constant), Z, X, U9, U15, U11

Vir: Lastni izračuni, SPSS, 2009

Iz Tabele 1 je razvidno, da so poleg »pravih« pojasnjevalnih spremenljivk z in x , v model (14) lahko vključene tudi spremenljivke U_9, U_{15}, U_{11} . Če bi se odločali po deležu pojasnjene variance, popravljenem za stopinje prostosti, bi izbrali model 3 (4 ali 5), kjer bi poleg »pravih« pojasnjevalnih spremenljivk v model vključili še

slučajno spremenljivko $U_9 (U_{11}, U_{15})$. Torej zavrnemo ničelno domnevo, da med pojasnjevalnimi spremenljivkami modela (14) ne nastopa U_i in sprejmemo sklep, da bo z »stepwise« regresijo pri točni stopnji tveganja $\alpha = 0,05$ in pri 20 dodatnih slučajnih spremenljivkah vsaj ena vključena v model.

4 Rešitev problema

Zaradi omenjenih težav pri ocenjevanju zdravstvenih stanj se pojavi vprašanje, s katerim orodjem in na kakšen način bi lahko ocenili vsa preostala zdravstvena stanja, ne da bi pri tem bili pristranski. Metoda, ki bi zagotovila ustrezno obravnavanje problema, pri katerem ne vemo, ali je linearen niti katere spremenljivke bi bilo potrebno vključiti v model (nekateri raziskovalci, so vključevali še izobraženost posameznika, starost, kadilce,...), je regresijsko drevo.

4.1 Klasifikacijska in regresijska drevesa (CART)

CART je metoda, ki uporablja zgodovinske (pretekle) podatke za izdelavo tako imenovanih odločitvenih dreves. Odločitvena drevesa pa se nato uporabljajo za klasificiranje (razvrščanje v razrede) novih podatkov. Da pa lahko uporabimo CART, moramo v naprej poznati razrede. Metodologijo CART so prvi objavili Breiman, Friedman, Olshen in Stone (1984) v članku z naslovom »Classification and Regression Trees«. Odločitvena drevesa so predstavljena kot niz vprašanj (da/ne), na podlagi katerih se vzorec deli na vedno manjše dele. CART lahko operira tako z numeričnimi kot tudi z kategoričnimi (opisnimi) spremenljivkami.

4.2 Regresijsko drevo

Prej je bilo omenjeno, da je pri uporabi CART nujno predhodno poznati (določiti) razrede. To pri regresijskih drevesih ne drži popolnoma, saj je vektor Y (odvisna spremenljivka) zvezna spremenljivka, in tako predstavlja vrednosti »odgovorov« za vsako izmed opazovanj matrike X (neodvisne spremenljivke) (Timofeev, 2004). Ker regresijska drevesa nimajo v naprej definiranih razredov, ne moremo uporabiti običajnih tehnik razvrščanja (npr. Gini pri klasifikacijskih drevesih). Zato je cepljenje (*angl. Splitting*) napravljeno z algoritmom za minimizacijo kvadratov ostankov, kar pomeni, da algoritem minimizira pričakovano vsoto varianc dveh vozlišč (*angl. Node*):

$$\min[P_l \text{Var}(Y_l) + P_d \text{Var}(Y_d)] \quad (16)$$

kjer je Y vektor odgovorov, črki l in d označujeta levo in desno otroško vozlišče (*angl. Child node*), P pa so

COMPARISON OF APPROACHES FOR ESTIMATING RELIABILITY OF INDIVIDUAL CLASSIFICATION PREDICTIONS

Darko Pevec, Zoran Bosnić, Igor Kononenko

Laboratory for Cognitive Modeling

University of Ljubljana, Faculty of Computer and Information Science

Tržaška cesta 25, 1000 Ljubljana, Slovenia

e-mail: darko.pevec@gmail.com

ABSTRACT

This paper is an extension of previous work on approaches for estimating reliability of individual regression predictions. Here we compare five different methods for reliability estimation of individual predictions applied to classification. Tested on ten domains with seven classification models, our results show interesting potential. Various estimates exhibited varying performance with different models and the same datasets. The best average results were achieved by estimation based on local modeling of prediction error, using the maximal distance.

1 INTRODUCTION

With supervised learning, our goal is to get the best prediction accuracy of new and unknown examples, as possible. Common methods like AUC and alike give an averaged accuracy assessment of models and can be sufficient in most applications. On the contrary, in cases where predictions may have significant consequences, common methods become insufficient as we want to back individual predictions up with a somewhat more credible explanation. In risk-sensitive decision making, for-say in medicine or finances where lives and money can be at stake, having information on single prediction reliability could be of great benefit. Hence, it is quite intuitive to seek for methods for assessment of confidence and/or reliability of individual predictions in a more localized manner.

Various methods have been developed to enable the users of classification and regression models to gain more insight into the reliability of individual predictions [1][2]. We take the model-independent black box approach and also exploit the fact, that we have class probability distributions available with every classification model currently in existence. We adopted four approaches to reliability estimation for individual examples from [1] and tried them with several measures from [2]. They were evaluated on ten testing domains gathered from the UCI Machine Learning Repository [12] using seven classification models. This paper is organized as follows. Section 2 summarizes

previous work from related areas of individual prediction reliability estimation and Section 3 presents the reliability estimates we adopted. We describe our experiments and testing methods, then show the results in Section 4. Last section provides conclusions and ideas for further work.

2 RELATED WORK

An appropriate criterion for differentiating between various approaches is, whether they target a specific predictive model or whether they are model-independent. While the model-specific approaches are less general, they are usually founded on exact mathematical or probabilistic properties. Since the model-independent approaches are general, they cannot exploit parameters, specific to a given predictive model, but rather focus on influencing the parameters that are available in the standard supervised learning framework (e.g. the learning set and attributes). The reliability estimates based on these approaches are defined as metrics over the observed learning parameters. Since the reliability is based on heuristic interpretation of available data, these metrics can take values from an arbitrary interval of numbers. As such, these metrics' values have no probabilistic interpretation [1].

The idea of reliability estimation for individual predictions originated in statistics, where confidence values and intervals are used to express the reliability of estimates. In machine learning, statistical properties of predictive models were used to extend predictions with reliability estimates. It is obvious that model-specific approaches cannot be used with an arbitrary predictive model due to their definition, which is bound to the specific model formalism. In contrast, the methods that are independent of the predictive model are also more generally applicable. These methods utilize approaches such as local modeling of prediction error based on input space properties and local learning [1].

The work presented here compliments and extends work described in [1] by comparing the performance of applicable estimates. They are summarized in the following section.

3 RELIABILITY ESTIMATES

As we are adopting already developed methods for reliability estimation in regression, we need a way to measure differences between predictions in classification. Models for classification are in the given present time all able to give a probabilistic interpretation of their predictions. In hope of finding a method that could give insight of prediction error, we tried seven different ways of evaluating distances between class probabilities as used also in [2]. They were implemented within the methods described in the second part of this section. For further reference, detailed algorithms and figures, the reader is invited to read [1].

3.1 Used measures

1.) Manhattan:

$$\|(P, Q)\|_{man} = \sum_i |p_i - q_i|$$

2.) Euclidean:

$$\|(P, Q)\|_{enc} = \sqrt{\sum_i (p_i - q_i)^2}$$

3.) Maximal distance:

$$\|(P, Q)\|_{max} = \max_i |p_i - q_i|$$

4.) Hellinger distance:

$$\|(P, Q)\|_{hel} = \sum_i (\sqrt{p_i} - \sqrt{q_i})^2$$

5.) Bhattacharyy distance:

$$\|(P, Q)\|_{bha} = \sum_i \sqrt{p_i q_i}$$

6.) symmetric Kullback-Leibler divergence:

$$\|(P, Q)\|_{kl} = \sum_i (p_i - q_i) \log_2 \frac{p_i}{q_i}$$

7.) Cosine between distributions:

$$\|(P, Q)\|_{cos} = \frac{\sum_i p_i q_i}{\sqrt{\sum_i p_i^2} \sqrt{\sum_i q_i^2}}$$

8.) Variance:

$$\|(P, Q)\|_{var} = \sum_i (p_i - q_i)^2$$

Note that our measures have subscripted marks. These are for easy future reference in our tabled results. This excludes *variance*, as it is used only with our *BAGV* estimates.

3.2 Adopted methods for reliability estimation

1.) *Local modeling of prediction error*

Let K be the predictor's class probability distribution for a given unlabeled example $(x, _)$. This approach to local estimation of prediction reliability is based on the nearest neighbors' labels. Given a set of nearest neighbors $N = [(x_1, C_1), \dots, (x_k, C_k)]$, where C_i is the true label of the i -th nearest neighbor, the estimate CNK ($C_{Neighbors} - K$) is for the unlabeled example defined as the average distance between the prediction based on k nearest neighbors and the example's prediction K :

$$CNK = \frac{\sum_{i=1}^k \|(C_i, K)\|}{k}$$

Obviously, CNK is not a suitable reliability measure for the k -nearest neighbors algorithm.

2.) *Density-based reliability estimate*

The density-based estimation of prediction error assumes that error is lower for predictions which are made in denser training problem subspaces, and higher for predictions which are made in sparser training subspaces. Based on this assumption, we trust the prediction with respect to the quantity of information that is available for its computation. A typical use is with decision and regression trees, where we trust each prediction according to the number of learning examples that fall in the same leaf of a tree as the predicted example.

The reliability estimate $DENS$ is a value of the estimated probability density function for a given unlabeled example. To estimate the density, Parzen windows were used, taking the Gaussian kernel. The problem of computing the multidimensional Gaussian kernel was reduced to computing the two-dimensional kernel by using a distance function applied to pairs of example vectors. Given the learning set $L = [(x_1, c_1), \dots, (x_l, c_l)]$, the density estimate for unlabeled example $(x, _)$ is therefore defined as

$$p(x) = \frac{\sum_{e \in L} \kappa(\|(x, e)\|_D)}{l}$$

where D denotes a distance function and κ denotes a kernel function (in our case Gaussian). Therefore the reliability estimate is given by:

$$DENS(x) = \max_{e \in L} (p(e)) - p(x)$$

3.) *Local cross-validation reliability estimate*

The LCV (Local Cross-Validation) reliability estimate is computed using the local leave-one-out procedure.

Suppose that we are given an unlabeled example for which we wish to compute the prediction and the LCV estimate. Focusing on the subspace defined by k nearest neighbors (parameter k is selected in advance), we then generate k local models, each of them excluding one of the k nearest neighbors. Using the generated models, we compute the leave-one-out predictions $K_i, i = 1, \dots, k$ for each of the nearest neighbors. Since the labels $C_i, i = 1, \dots, k$ of the nearest neighbors are given, we are therefore able to calculate the absolute local leave-one-out prediction error $E_i = \|(C_i, K_i)\|$. The LCV estimate is then computed as the average of the nearest neighbors' local errors E_i or weighted by distance. The procedure is schematically illustrated in [1], accompanied by a pseudo-code algorithm.

In experimental work, the algorithm was implemented to be adaptive with respect to the size of the neighborhood, that is to the number of examples in the learning set. The parameter k was therefore assigned to $1/10L$, where L denotes the learning set.

4.) Variance of a bagged model

Since an arbitrary regression model can be used with the bagging technique, the technique was generalized and used as a reliability estimate for use with other regression models [1].

Given a bagged aggregate of m predictive models, where each of the models yields a prediction $B_k, k=1, \dots, m$, the reliability estimate $BAGV$ is defined as the variance (Eq. 8) of prediction's class probability distribution:

$$BAGV = \frac{1}{m} \sum_{k=1}^m \left\| (B_k, K) \right\|_{\text{var}}$$

Due to the fact, that models return probability vectors $B_k = (B_{k1}, B_{k2}, \dots, B_{kc})$, where c denotes the number of classes (the same applies for the prediction $K = (K_1, K_2, \dots, K_c)$), we can also define a reliability estimate $BAGV_{class}$ which focuses solely on the variance of the most probable class c_{max} :

$$BAGV_{class} = \frac{1}{m} \sum_{k=1}^m \left\| (B_{k, c_{max}}, K_{c_{max}}) \right\|_{\text{var}}$$

where $c_{max} = \arg \max_i (K_i)$.

4 EXPERIMENTAL RESULTS

Testing was performed using the leave-one-out cross-validation procedure. For each learning example that was left out in the current iteration, we computed the prediction and all the reliability estimates. The performance of reliability estimates was measured by computing the Spearman's rank correlation coefficient between each reliability estimate and the prediction error (that being the difference between 1 and the predicted probability of the correct class). The significance of correlation coefficients was statistically evaluated using the Welch Two Sample t-test which is an adaptation of Student's t-test intended for use with two samples having possibly unequal variances.

Note that all of the estimates are expected to correlate positively with the prediction error. This means that all the estimates are founded so that higher values represent less reliable predictions and lower values represent more reliable predictions (currently the value 0 represents the reliability of the most reliable prediction).

The performance of reliability estimates was tested using seven classification models implemented in the statistical package R [3]. Here are some key properties of the models used:

Naive Bayes (NB): Naive Bayes Classifier called with `type="raw"`,

k-nearest neighbors (KNN): weighted k -Nearest Neighbor Classifier [4] which returns class probabilities in the value `prob`,

neural networks (NN): three-layered perceptron [5] with five hidden neurons,

support vector machines (SVM): implementation from the library for support vector machines (LIBSVM) [6][7], called with `probability=TRUE`,

decision trees (DT): decision trees [8] with no explicit parameters,

dataset	dataset	A _{discrete}	A _{continuous}
housevotes	435	16	0
wine	178	0	13
parkinsons	195	0	22
zoo	101	16	0
tae	151	4	1
postoperative	90	7	1
monks-3	432	5	0
irisset	150	0	4
glass	214	0	9
hungarian	294	7	6

Table 1: Brief summary of the testing datasets.

random forests (RF): random forests [9][10] with 100 trees,

bagging (BAG): bagging [11] with 50 classification trees.

The aim of our research was to evaluate the reliability estimates with models being treated as black-boxes. Therefore, the focus of our research was not to optimize the above model parameters to improve prediction accuracy, but to evaluate the accuracy of reliability estimates. Each data set is a classification problem, the application domains vary. A brief summary of the data sets is given in Table 1, $A_{discrete}$ and $A_{continuous}$ denote sets of discrete and continuous attributes respectively. We see that there are 3 domains with only discrete attributes, 4 with only continuous attributes and the remaining 3 have a mixture of both.

4.1 Testing of individual estimates

Table 2 sums the number of experiments in which datasets exhibited significant correlation between the reliability

method	nb	knn	nn	svm	dt	rf	bag	%
<i>cnk_{man}</i>	5/1	0/0	6/0	6/0	8/0	6/0	7/0	54/1
<i>cnk_{euc}</i>	5/1	0/0	8/0	6/1	8/1	6/0	7/0	57/4
<i>cnk_{max}</i>	5/1	0/0	8/0	6/0	8/0	6/0	7/0	57/1
<i>cnk_{cos}</i>	4/1	0/1	6/3	4/0	8/0	3/0	4/0	41/7
<i>cnk_{bha}</i>	5/1	0/0	9/0	3/0	7/0	2/0	5/0	44/1
<i>cnk_{hel}</i>	5/1	0/0	8/0	3/0	7/0	2/0	5/0	43/1
<i>cnk_{kl}</i>	5/1	0/0	8/1	3/0	5/0	2/0	1/0	34/3
<i>lcv_{man}</i>	5/0	6/1	2/1	3/2	8/0	6/0	5/0	50/6
<i>lcv_{euc}</i>	5/0	6/1	4/2	3/2	8/0	6/0	5/0	53/7
<i>lcv_{max}</i>	4/0	6/1	4/2	3/2	8/0	6/0	5/0	51/7
<i>lcv_{cos}</i>	5/0	6/1	3/3	2/2	8/0	5/0	5/0	49/9
<i>lcv_{bha}</i>	3/2	2/2	4/1	1/1	4/1	2/1	1/1	24/13
<i>lcv_{hel}</i>	4/0	6/1	4/2	2/2	8/0	6/0	5/0	50/7
<i>lcv_{kl}</i>	4/0	6/1	3/2	2/2	8/0	7/0	5/0	50/7
<i>dens</i>	3/0	2/0	0/2	0/0	2/0	3/0	2/0	17/3
<i>bagv</i>	6/0	6/1	4/0	3/5	8/1	6/1	7/0	57/11
<i>bagv_{class}</i>	4/2	5/2	2/0	4/5	6/1	6/1	5/1	46/17
<i>avg</i>	5/1	3/1	5/1	3/2	7/0	5/0	5/0	

Table 2: Number of experiments exhibiting significant positive/negative correlation between reliability estimates and prediction error / by models.

all models

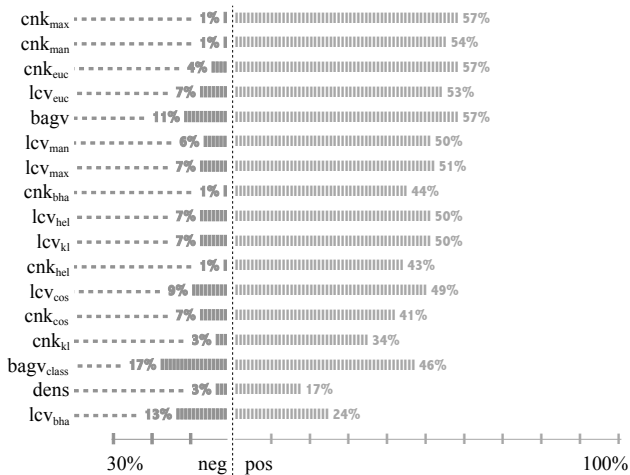


Figure 1: Ranking of reliability estimates by percentage of positive vs. negative correlations with predictor error

estimate and the prediction error throughout used models. Each pair P/N represents the number of positive (P) versus the number of negative (N) correlations. Last row averages these significant correlations for each of our reliability estimates.

It is obvious that our nearest neighbors based estimates *CNK* do not work well with the *k*-nearest neighbors model (even though the *k*-values were 7 for the predictor and 5 in the *CNK* estimate). Secondly, *DENS* estimate did not excel in any given example, as we saw positive correlation with at most three datasets throughout our models and nothing but negative correlation with neural networks. The third most obvious negative deviation of the results is spotted with the *BAGV* and *BAGV_{class}* estimates on the model *SVM*, where negative correlation was statistically significant in half of the experiments.

In Figure 1 we have averaged the results of our reliability estimates through all used models. We see that *CNK_{max}*, *CNK_{man}* and *CNK_{auc}* averaged around 56% positively and 2% negatively correlated experiments. Followers do not seem to obey any specific order. *BAGV* and *BAGV_{class}* took fifth (57 positive / 11 negative experiments) and fifteenth (46 positive / 17 negative) places respectively. *DENS* estimate came second-to-last with only 17% positively and 3% negatively correlated tests and left the last place to *LCV_{bha}* with 24% positive experiments and 13% negative.

5 CONCLUSION

We wanted to see how would our reliability estimates perform on ten various data-sets trying to model a mixture of real-life, synthetic - hard and easy problems. We were able to achieve 57% average positive and 1% negative correlation of the reliability estimate with the prediction error using the estimate *CNK_{max}*. This estimate turned out to be the best choice for nearly half of our models, that being

support vector machines, decision trees, and bagging with decision trees. Still, all reliability estimates' performance varied substantially on different model/domain pairs.

Further work would be foremost to try and confine, that is transform our measures into metrics, as we could then relabel our reliability to *confidence*. We would also like to try other possible measures, as we saw that different measures give remarkably different results. Also, the *BAGV* estimate is going to be further extended and tested with different measures.

Furthermore, all methods exhibit potentials of parametrization, which may improve on their statistic reliability. Most over, statistical package R should be given a wrapper function which unifies probabilistic outputs of available classification models, as there are currently several fringing differences which can cause serious head-scratching to a modest user.

References

- [1] Z. Bosnić, I. Kononenko, *Comparison of approaches for estimating reliability of individual regression predictions*, Data Knowl. Eng. Vol. 67, no. 3, p.504-516 (2008).
- [2] M. Kukar, *Ocenjevanje zanesljivosti klasifikacij in cenovno občutljivo kombiniranje metod strojnega učenja*, Doktorska disertacija, Univerza v Ljubljani (2001).
- [3] R Development Core Team, *A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria (2006).
- [4] Hechenbichler K. and Schliep K.P., *Weighted k-Nearest-Neighbor Techniques and Ordinal Classification*, Discussion Paper 399, SFB 386, Ludwig-Maximilians University Munich (2004).
- [5] Ripley, B. D., *Pattern Recognition and Neural Networks*, Cambridge (1996).
- [6] N. Christiannini, J. Shawe-Taylor, *Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge University Press (2000).
- [7] C. Chang, C. Lin, *LIBSVM: a library for support vector machines*, software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm> (2001).
- [8] Breiman, Friedman, Olshen, and Stone, *Classification and Regression Trees*, Wadsworth (1984).
- [9] L. Breiman, *Random Forests*, Machine Learning 45(1), 5-32, (2001).
- [10] L. Breiman, *Manual On Setting Up, Using, And Understanding Random Forests V3.1*, http://oz.berkeley.edu/users/breiman/Using_random_forests_V3.1.pdf (2002).
- [11] L. Breiman, *Bagging predictors*, Machine Learning 24 (2), 123-140, (1996).
- [12] A. Asuncion D.J. Newman, *UCI machine learning repository*, <http://archive.ics.uci.edu/ml/> (2007).

USING STOCHASTIC MODEL FOR IMPROVING HTTP LOG DATA PRE-PROCESSING

Marko Požnenel, Viljan Mahnič, Matjaž Kukar

University of Ljubljana, Faculty of Computer and Information Science

Tržaška cesta 25, 1000 Ljubljana, Slovenia

Tel: +386 1 4768365; fax: +386 1 4264647

e-mail: marko.pozenel@fri.uni-lj.si, viljan.mahnic@fri.uni-lj.si, matjaz.kukar@fri.uni-lj.si

ABSTRACT

We describe a novel method for interleaved HTTP session reconstruction based on first order Markov model. Interleaved session is generated by a user who is concurrently browsing a web site in two or more web sessions (browser windows). In order to assure data quality for subsequent phases in analyzing user's browsing behavior, such sessions need to be separated in advance. We propose a separating process based on trained first order Markov chains. We developed a testing method based on various measures of reconstructed sessions similarity to original ones. We evaluated the developed method on two real clickstream data sources: web shop and university student records information system. Preliminary results show that method performed well.

1 INTRODUCTION

Data about behaviour of web site visitors have become one of the most important sources of information in most web-aware companies. They play an important part in daily transactions and important business decisions. It is essential to get reliable data analyses, which require both appropriate methods and data. The quality of the the patterns discovered in data analysis depends on the quality of the data on which data mining is performed. The main source of data for the analysis of user behavior represent *clickstream* data [8]. A sequence of clicks that user makes browsing through website is called clickstream. A *user session* is represented by one visit of a user to a web site. For better web usage mining results we need reliable sessions. Clickstream data from a normal website are noisy, page events are often not explicitly linked to page requests. The pre-processing phase is therefore prone to errors [5]. Although many methods for reliable sessions reconstruction have been devised [1, 9], reliable session reconstruction still remains a challenge.

Especially really interested and capable users often browse the same web site with multiple browser windows opened. In each web browser they perform actions to complete a certain task. Typically, users switch between browsing tasks so that they work on a task only for a certain time period. Even if only one user is currently active, we actually have concurrent sessions, each for one web

browser (i.e. task). In a web server log file all concurrent sessions will be seen as single long session. We call such sessions *interleaved sessions*. They cannot be easily separated without some kind of context help. Such sessions have negative effect on data quality so we have to deal with the issue. We have three choices: (i) we neglect the problem, (ii) simply abandon such sessions, (iii) try to separate them. The first choice is bad for data quality since such sessions can affect web usage analysis results. If we abandon such sessions we also abandon useful knowledge about web site usage. Such sessions are usually generated by advanced users whose behaviour could be potentially extremely valuable to us. Therefore we decided to develop a method for separating interleaved sessions.

2 METHODS

2.1 Clickstream

In order to attract more visitors to our web site we have to know, who our visitors are, what they do on our site, and what they would like to be changed. A great aid in achieving this goal is clickstream data. Clickstream data are often large, inadequately structured, and show incomplete picture of users' activity. For example, server side log data do not involve browser and e.g. network caching ('Back' browser actions or requesting pages in intermediate server's cache) [5].

Clickstream data needs to be gathered, preprocessed and cleaned prior to analysis. This step depends on the type and the quality of data. Work done in this phase affects the quality of results of further analyses.

The basic form of clickstream data from a Web server is stateless - no session identifier is logged. Each line in the log file shows an isolated resource retrieval event, but does not provide a link to other events in a user session. Since we are interested in all user actions in a certain period of time, we have to gather all individual events in a user session. The process is called *sessionization*. Without some context help it is hard or impossible to reliably identify complete user session. Berendt et al. [1] report that these sessionization tools are based on heuristic rules and assumptions about the site's usage and are therefore prone to errors.

2.2 Discrete Markov models for clickstream analysis

Markov chain is defined as follows. We have a set of states $S = \{s_1, s_2, \dots, s_N\}$, where N denotes the number of states. The process starts in one of the states and moves forward from one state to another at regularly spaced discrete times. For example, the chain is currently in the state s_i and it moves next to s_j with the transition probability p_{ij} . The starting state is defined by a probability distribution. We denote the steps in which the process changes states as $t = 1, 2, \dots, n$ and the state at time t as q_t . Associated with each state is a set of transition probabilities p_{ij} , where

$$p_{ij} = P(s_i \rightarrow s_j) = P(q_t = s_j | q_{t-1} = s_i)$$

that is, given the present state, the future and the past states are independent. The probability of transition between states in a single step can be written as a matrix T , called the *Transition probability matrix*.

Given a sequence of states (q_1, q_2, \dots, q_k) we can calculate the probability of the sequence by multiplying the probability of the initial state $P(q_1)$ with the probability of transitions to the successive state as follows:

$$P(q_1, q_2, \dots, q_k) = P(q_1) \prod_{i=2}^k P(q_{i-1} \rightarrow q_i)$$

In the first-order Markov chain the next step depends only on current state. If the step depends on the current and previous state, we obtain somewhat more complicated second-order Markov model [4].

3 SEPARATING INTERLEAVED SESSIONS WITH MARKOV MODEL

The process of separating interleaved sessions is one of the phases in data pre-processing. First clickstream data has to be cleaned and sessionized. We refer to sessions, that have been restored without deficiencies, as *clean* sessions. During the sessionization process we detect interleaved sessions which we cannot separate at that time either by using some background knowledge, or by applying the pre-trained MM. Interleaved sessions are separated from clean sessions and are additionally processed. The separation process is based on stochastic methods which have been used to solve some other issues related to clickstream. Because of generality and simplicity we decided to use first-order MM. We build a Markov model and train it with data from clean sessions. We can use last pre-processing clean sessions or clean sessions from last few pre-processings. Trained markov model is then used to separate interleaved sessions. In case of more than two interleaved sessions only the first one is considered as clean, and the second one is submitted to further separation. This results in more reliable pre-processed user behavior data. The last step in a analysis is evaluation of separated sessions with several methods.

We also include site map data as background knowledge. Site map consists of links between pages that are explicitly connected with hyperlinks. A link between pages S_1 and S_2 in a site map means higher prior probability of transition

between these two pages than if there were no link in a site map. Users navigate with higher probability between pages that are linked in a site map. When we train the MM we also use the web site map. Based on links between page sites we calculate initial transition probability between pages $p_{ij}^{(0)}$, where i, j denotes source and target state. Formula for calculating $p_{ij}^{(0)}$:

$$p_{ij}^{(0)} = \frac{1 - P_A^{(ij)}(N - n_t)}{n_t}, n_t \geq 1$$

where j denotes all states that are connected to state i , N denotes number of states, n_t number of outgoing links from state i ¹ and $P_A^{(ij)} = 1/N^2$ default probability transition between any two states. If there is no connection between i and j , probability $P_A^{(ij)}$ is assigned. Parameter $P_A^{(ij)}$ determines the prior probability of transition between arbitrary two pages in the site map. Higher $P_A^{(ij)}$ means higher probability of transition.

We train the model with clean, not interleaved sessions which present users' paths through the web site. Each session is represented as sequence of pages $S = \{q_1, q_2, \dots, q_n\}$ where n denotes length of session. q_1 denotes the entry page and q_n the last page the user visited in this session. For transition from $q_{i-1} = s_j$ to $q_i = s_k$, training data site map data can be combined with *M-estimate* [3]:

$$P(s_j \rightarrow s_k) = p_{jk} = \frac{(s + m * p_{jk}^{(0)})}{n + m}$$

where s denotes number of transitions from state j to k , which we got from training data. n is number of visits of state j . m denotes the weight which presents the ratio between prior (web site map) and posterior knowledge. $p_{jk}^{(0)}$ denotes transition probability based on web site map. Parameter m represents the importance rate of prior knowledge. The higher the m is, the more important the prior knowledge is. If $m = 0$, then we completely neglect the meaning of prior knowledge. In that case *M-estimate* converts to relative frequency $p_{jk} = s/n$.

3.1 The separation process

Separating interleaved session is based on a fact that transition between sites $s_i \rightarrow s_{i+1}$ is more likely to belongs to one of the consisting sessions. If we have interleaved session $S_p = [q_1, q_2, \dots, q_n]$ that consists of two clean sessions length n_1 and n_2 , where $n_1 + n_2 = n$. Let us say that last page of the first session that we already managed to separate is S_{1i} . Similarly for the second session we denote last page as S_{2i} . For each page S_i in interleaved session, which we have not processed yet, we check what is the probability of transition from last page of separated session to current page S_i . If $P(S_{1i} \rightarrow S_i) > P(S_{2i} \rightarrow S_i)$ we add page S_i to the first separated session, otherwise to the second one. Until both of the separated sessions get the first element (entry page), we have to check whether S_i is an entry page

¹ We assume that there is always the reflective transition from s_i to s_j , so n_t is always greater than 0.

for second session. Separating process can be seen on Figure 1.

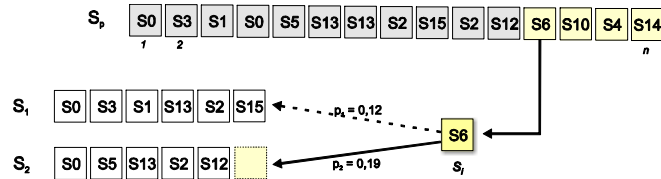


Figure 1: Figure shows simple process of separating interleaved sessions

3.2 Evaluation of separating process

Separated sessions needs to be evaluated to see how successful our method was. Each session is represented as a sequence of pages. Evaluating quality of separated sessions can be viewed as evaluating similarity of symbol sequences [6]. Basically, two sequences are more similar if they have more symbols in common and the symbols' order is similar. There are many methods of measuring similarity between two sequences [7]. We use the following methods since they are appropriate for the evaluation of separating process. *Perfect match* is a simple method where only sequences that perfectly match contribute to the end result.

Alternative approach to measure sequence similarity is based on sequence distance, named *edit distance*. The distance between two sequences is defined as the smallest sum of edit operations' costs that transforms one sequence to another. If we have only three edit operations: inserting, deleting and swapping symbols, and all have the cost of 1, we get *Levenshtein distance*.

A sequence $Z = [z_1, z_2, \dots, z_n]$ is a subsequence of another sequence of sequence $X = [x_1, x_2, \dots, x_m]$ if there exists a strict increasing sequence i_1, i_2, \dots, i_k in X such that for all $j = 1, 2, \dots, k$ we have $x_{i_j} = z_j$ [2]. If we have sequences X and Y , the longest common subsequence of X and Y is a common subsequence with the maximum length. The longer the common subsequence, the more two sessions are similar to each other [6].

We can improve LCS method to differentiate LCS in relation to other elements in the sequence. Chin et al. [7] called this method *weighted LCS* (WLCS). They also propose the use of *F-measure* to estimate the similarity between two sequences X of length m and Y of length n . We decided to use F-measure for presenting end results.

4 MATERIALS

4.1 Synthetic data

First we created a test environment that is similar to real one but is not as complex. We checked what is the average HTTP session length on a local web server. For testing we fixed the number of Web pages to 30. We created an artificial web site map that represented links with higher probability. According to the site map we generated a number of sessions that were used for MM training data

and some of them for creating interleaved sessions. After training MM, we applied the process for separating interleaved sessions and verified the results. About 48% of interleaved sessions were separated 100% correctly, which encouraged us to proceed to real data.

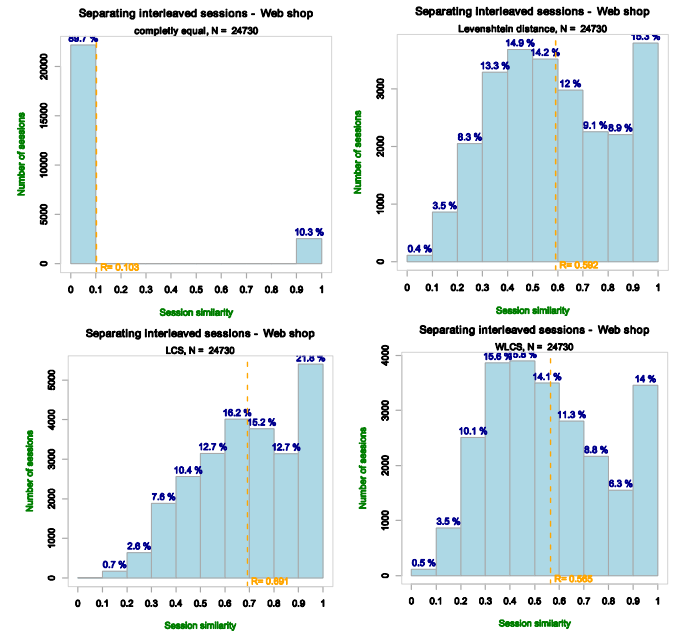


Figure 2: Results of separating for Web shop clickstream

4.2 Real-world data

We applied the interleaved session separating process on two real clickstream sources. The first clickstream originates from log files of university student records information system. It has been used by 16 member institutions. It has approximately 300 different pages. Each state in MM corresponds to an individual page. Typical user paths are well defined. Users have to be logged on in order to use the system. Sometimes they are logged on with different user roles at the same time, and this creates interleaved sessions. Since users have to be logged on we can always determine the session entry point. The Web server log files use the basic CLF format. Clickstream data was taken for 4 months of use, which resulted in 150.000 user sessions.

The second clickstream source is taken from a web shop, which is considerably different from student records information system. Users do not have to sign in (except for buying items), it has many more users and many more pages. We had to cut down number of states of Markov model in order to efficiently use it. Every state of our Markov model represents a group of pages, not an individual page. We transformed the web shop pages to 900 states. Session entry point can be almost any page, which makes separating interleaved sessions harder. The Web shop site map has plenty of links between pages. In fact only few pages are not linked with all others. The web shop generates about 10.000 user sessions a day.

For both clickstreams we took the same steps as with artificially generated data. Initial clean sessions, used for

learning, were generated during the sessionization process of clickstream data. During the sessionization we applied all the necessary steps in order to remove noisy data. We analysed what a typical user session looks like and removed all sessions that did not meet the rules (e.g. to short or too long sessions). 70% of clean sessions were used as a training set for MM, and the rest were used to generate interleaved sessions in order to evaluate separation process. After separating interleaved sessions we evaluated results with evaluation methods that we presented earlier.

we can see that 43% sessions have been separated 100% right. Result is much better in comparison with Web shop.

6 CONCLUSION

We propose a new method for improving the quality of clickstream data in pre-processing phase that is based on a first-order Markov model. We present the motivation that led us to implementation and have applied method on two real data clickstreams. The presented results show that in certain cases method gives promising results. We analysed the domain and detected possible causes of worse results. In order to minimize method deficiencies we plan to work on the issues we presented. First we have to improve the method for detecting interleaved session starting pages. We are also planning to use second-order Markov model and Hidden Markov Model (HMM) for separating process.

References

- [1] B. Berendt, B. Mobasher, M. Nakagawa, and M. Spiliopoulou. The impact of site structure and user environment on session reconstruction in web usage analysis. In *WEBKDD - KDD Workshop on Web Mining and Web Usage Analysis*, pages 159-179, 2002.
- [2] T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms*. The MIT Press and McGraw-Hill Book Company, 1989.
- [3] S. D., B. Cestnik, and I. Petrovski. Using the m-estimate in rule induction. *J. Comput. Inf. Technol.*, 1(1):37-46, 1993.
- [4] M. Deshpande and G. Karypis. Selective markov models for predicting web page accesses. *ACM Trans. Internet Technol.*, 4(2):163-184, 2004.
- [5] R. Kohavi. Mining e-commerce data: The good, the bad, and the ugly. In Foster Provost and Ramakrishnan Srikant, editors, *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 8-13, 2001.
- [6] G. Leusch, N. Ueffing, and H. Ney. A novel string-to-string distance measure with applications to machine translation evaluation. In *In Proceedings of MT Summit IX*, pages 240-247, 2003.
- [7] C-Y. Lin and F. J. Och. Automatic evaluation of machine translation quality using longest common subsequence and skip-bigram statistics. In *ACL '04: Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, page 605, Morristown, NJ, USA, 2004. Association for Computational Linguistics.
- [8] I-H. Ting, C. Kimble, and D. Kudenko. A pattern restore method for restoring missing patterns in server side clickstream data. *Lecture Notes in Computer Science*, 3399:501-512, March 2005.
- [9] J. Zhang and A.A. Ghorbani. The reconstruction of user sessions from a server log using improved time-oriented heuristics. In *Communication Networks and Services Research, 2004. Proceedings. Second Annual Conference on*, pages 315-322, May 2004.

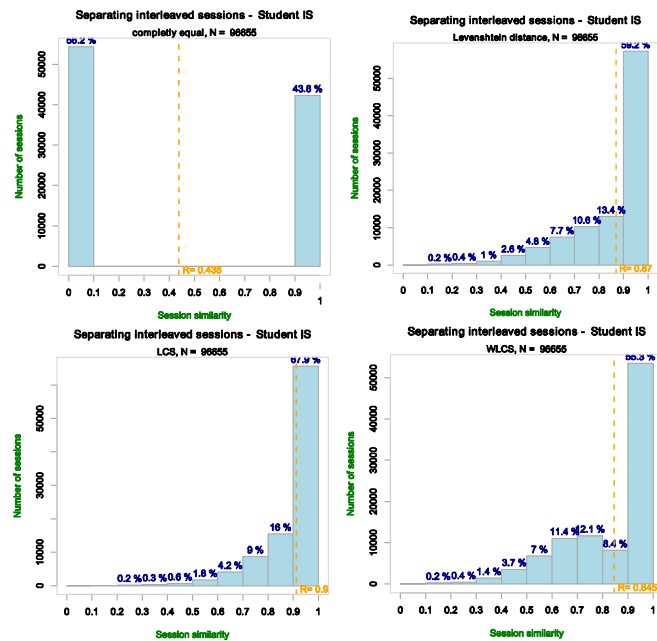


Figure 3: Results of separating for Student records IS clickstream

5 RESULTS

In Figures 2 and 3 we can see graphs for evaluation methods and source of clickstream. Each graph corresponds to one evaluation method. The X axis shows intervals for F-measure based similarity and the Y axis shows number of sessions that fall in that interval.

If we look at Figure 2 we see results for Web shop. 24.730 interleaved sessions have been created and separated. Looking at first graph at that Figure, one sees how many sessions have been separated 100% correctly (session sequence similarity = 1). For web shop this percentage is a little more than 10%, which is quite low. However even 10% is better than throwing away all interleaved sessions. One of the reasons is that grouping pages together affects the results. Since the site map is larger, there may be numerous user paths, what also affects the results. User can enter the web shop at almost any page, so it is harder to detect where the second session in interleaved session starts. Other three graphs on at Figure 2 depict how well the sessions have been separated according to evaluation method. Results on graph that show LCS are better since LCS is less strict method of evaluation than WLCS. Figure 3 reports results for student IS clickstream. At the first graph

A FUZZY EXPERT SYSTEM TO ENFORCE NETWORK SECURITY POLICY

Bel G. Raggad, Seidenberg School of CS & IS, Pace U, New York, braggad@pace.edu
Azza Mastouri, Institut Supérieur de Gestion de Tunis, Tunisia
Manal Mastouri, Institut Supérieur de Gestion de Tunis, Tunisia

ABSTRACT

Security policy of a computing environment is the set of statements defining its acceptable behavior. A computing environment consists of people, activities, data, technology, and network. A security policy may be divided into two parts: nominally auditable policy (NAP), and technically auditable policy (TAP). Even though both components are auditable, the NAP and TAP are not auditable in the same manner. The TAP may be fully translated into security control variables or indicators that can be automatically verified while the NAP cannot be translated into indicators that can be automatically verified. The NAP involves owners' subjective judgment. This article proposes a fuzzy expert system to enforce a security policy (FESP).

1. INTRODUCTION

We define security policy as the set of statements defining the acceptable behavior of the computing environment. The computing environment, as in [21] includes people, activities, data, technology, and network. The security policy may be divided into two parts: nominally auditable policy (NAP), and technically auditable policy (TAP). Even though both components are auditable, the NAP and TAP are not auditable in the same manner. The TAP may be fully translated into security control variables or indicators that can be automatically verified. The NAP, however, cannot be translated into indicators that can be automatically verified and owners' subjective judgment should be trusted.

We propose a fuzzy expert system to enforce a security policy (FESP). The FESP consists of 6 components: 1-Feature selection, 2-Fuzzification, 3-Inference, 4-Composition, 5-Defuzzification, and 6-Response.

Fuzzification is the process of converting crisp input data to fuzzy sets. The linguistic variables in the antecedent part of the rules are evaluated. The corresponding source data are mapped into their membership functions and truth values then fed into the rules. The most commonly used fuzzy inference method is the so-called Max-Min inference method [13], in particular in engineering applications [16]. The Max-Min inference method is applied to the rule set, producing a fuzzy output variable.

The result of the fuzzy inference is a fuzzy set. The defuzzification step produces a representative crisp value as

the final output of the system. There are several defuzzification methods [13] [16]. The most commonly used is the Centroid (Center-of-gravity) defuzzifier which provides a crisp value based on the center-of-gravity of the output fuzzy set.

The theory of fuzzy sets was introduced by Lotfi A. Zadeh, University of California, Berkeley, in the 1960's as a means to model the uncertainty within natural language. The mechanics of fuzzy sets theory was set forth in 1965, based on Zadeh's key notion of graded membership, according to which a set could have members that belong to it only in part. Such fuzzy sets have imprecise boundaries and therefore gradual transition from membership to non-membership of an element in its fuzzy set is observed. The ambition of fuzzy sets is to provide interaction of natural language and numerical models [11].

Fuzzy concepts have been further explored and applied to a diverse range of problems as in flight control, power systems, nuclear reactor control, climate control, etc. [13], [23]. Although human reasoning has been investigated since the inception of fuzzy logic, by far the majority of published work has been concerned with fuzzy control. In addition to control, fuzzy logic, however, fuzzy expert systems have been applied successfully in stock tracking on the Nikkei stock exchange [23], information retrieval [17] and the scheduling of community transport [10].

The most applied methods of implementing fuzzy inferencing are the Mamdani and the Takagi-Sugeno fuzzy expert system [15]. Other references reviewing the fuzzy expert system literature include [4] [5] [6] [7] [8] [9].

2. DEFINITION OF THE FESP

The six components defining the FESP are combined in functional relationship, expressing the system output in terms of its composing processes, as follows

$$z = \delta(\Sigma_{\sigma}(\Pi_{\pi}(\Phi_{|X|}(R))), (1)$$

where

z = crisp output vector defining the security response

δ = Defuzzification process

Σ_{σ} = Composition process using σ

Π_{π} = Inference process using the inference method π .

$\Phi_{[X]}$ = Fuzzification process
 R = Security policy rule base

Let us define the terms in expression (1) in a backward manner.

3. FEATURE SELECTION

The TAP component of the security policy may be written as a set of rules explaining the responsive actions security management need to take if a set of indicators are satisfied. The security control indicators define the conditions that are examined to determine whether or not the current violations of the security policy are of an adversity nature. The set of rules should provide decision support information in terms of the responsive actions security management needs to take.

The feature selection process aims at identifying those features that can provide critical information about any adversity to the security policy. We propose that information owners and security management jointly select the features that are more effective in predicting possible adversity in the current security policy. While low adversity may recommend that security management issue warning to all those agents that are concerned with the detected distress in the security policy, a higher adversity output will trigger immediate response actions to correct the security policy. An output with a moderate adversity recommendation will inform security management who uses subjective judgment to determine the appropriate actions to be taken.

Variable selection refers to the problem of selecting input variables that are most predictive of the adversity of the security policy. Variable selection problems are found in all machine learning tasks, supervised or unsupervised, classification, regression, time series prediction, and so on. Feature selection refers to the selection of an optimum subset of features derived from these input variables [1] [2] [14].

The objective revolves around the efficiency of the inference process and the number of rules to be examined by the inference process for all instances taken by these features. For some applications, intermediate techniques such as variable ranking, variable subset ranking, and search trees are particularly important [2] [3]. These intermediate tools may be combined with other selection criteria elicited from information owners and various security policy stakeholders. They may do so by exploring the tradeoff between the efficiency of the inference process and the size of the feature set.

4. FUZZIFICATION: DEFINING $\Phi_{[X]}(R)$

The expression $\Phi_{[X]}(R)$ represents the set of rules fired by the current input vector X . This set contains all the rules R_i , such that $\alpha(R_i) \geq \alpha_0$. The value α_0 represents the minimal degree

of truth accepted by system owners in satisfying the conditions constituting the premise of a rule. Let X define the system input stream which belongs to the domain $D(X) \subset R^n$. The rule base R consists of the security policy statements represented by the rules $\{R_i\}$, $i=1, m$. Let $I(R_i) \subset N$ denote the set of indexes of the components X_j from the input stream participating in defining the premise of R_i . Also let F denote the set of generic fuzzy subsets F_i , $i=k$, needed to evaluate the input stream, and G denote the fuzzy subsets needed to evaluate system outputs.

The fuzzification process transforms the input stream X into $\mu[R_i](X)$ where μ is a composite membership function defined on the domain $D(X)$ consisting of the vector made of the individual membership function $\mu[F_j]$, $j=1, I(R_i)$. After fuzzifying X , we compute all the alpha values $\alpha(R_i) = \text{Min}\{\mu[F_j](X_j), j \in I(R_i)\}$, $i=1, m$, for all the rules in the security policy rule base R . For any rule R_i , $i=1, m$, if $\alpha(R_i) \geq \alpha_0$, we say that rule R_i is fired. This rule will then participate in the inference process Π .

Let f^Φ denote the number of rules that have been fired at the fuzzification process Φ . We can then denote $\Phi_{[X]}(R) = \{R^{\Phi_1}, \dots, R^{\Phi_{f^\Phi}}\}$. Replacing the latter expression in equation (1), we obtain:

$$z = \delta(\Sigma_\sigma(\Pi_\pi(\Phi_{[X]}(R))) = \delta(\Sigma_\sigma(\Pi_\pi(\{R^{\Phi_1}, \dots, R^{\Phi_{f^\Phi}}\}))) \quad (2)$$

5. INFERENCE PROCESS: DEFINING $\Pi_\pi(\Phi_{[X]}(R))$

In this expression, the term $\Phi_{[X]}(R)$ has been defined as in (2). Let us explain how the process Π_π transforms $\Phi_{[X]}(R)$. We have $\alpha(R^{\Phi_i}) \geq \alpha_0$, $i=1, f^\Phi$. Every rule R^{Φ_i} , $i=1, f^\Phi$ in $\Phi_{[X]}(R)$ produces the fuzzy instance $z[R^{\Phi_i}]$, $i=1, f^\Phi$. The fuzzy process may be then expressed as follows:

$$\begin{aligned} z[R^{\Phi_i}] &= \pi(R^{\Phi_i}), \quad i=1, f^\Phi \text{ such that:} \\ \text{Min-Max Inference Process: } z[R^{\Phi_i}](t) &= \text{Min} \{ \alpha(R^{\Phi_i}), \\ &\mu[R^{\Phi_i}](t) \} \\ \text{Product-Sum Inference Process: } z[R^{\Phi_i}](t) &= \alpha(R^{\Phi_i}) \mu[R^{\Phi_i}](t) \end{aligned}$$

$$\text{We then have: } \Pi_\pi(\Phi_{[X]}(R)) = \{z[R^{\Phi_i}]\}, \quad i=1, f^\Phi. \quad (3)$$

We can rewrite the expression (1) as follows:

$$z = \delta(\Sigma_\sigma(\Pi_\pi(\Phi_{[X]}(R))) = \delta(\Sigma_\sigma(\{z[R^{\Phi_i}]\}, \quad i=1, f^\Phi)). \quad (4)$$

6. COMPOSITION PROCESS

This expression can now be written as follows:

$$\Sigma_\sigma(\Pi_\pi(\Phi_{[X]}(R))) = \Sigma_\sigma(\{z[R^{\Phi_i}]\}, \quad i=1, f^\Phi)$$

The composition process Σ_σ applies σ to the set $\{z[R^{\Phi_i}]\}$, $i=1, f^\Phi$ to produce the fused fuzzy output $\sigma(\{z[R^{\Phi_i}]\}, \quad i=1, f^\Phi)$ as follows:

$$\sigma(\{z[R^{\Phi_i}]\}, \quad i=1, f^\Phi) = z[R^{\Phi_i}] \oplus \dots \oplus z[R^{\Phi_{f^\Phi}}]. \quad (5)$$

where:

$$\text{Min-Max Inference Process: } \sigma(\{z[R^{\Phi_i}]\}, \quad i=1, f^\Phi) = \text{Max} \{z[R^{\Phi_i}]\}, \quad i=1, f^\Phi)$$

$$\text{Product-Sum Inference Process: } \sigma(\{z[R^{\Phi_i}]\}, \quad i=1, f^\Phi) = z[R^{\Phi_i}]$$

+ ... + z[R^φ_{iφ}].

The expression in (1) now becomes: $z = \delta(\Sigma_{\sigma}(\Pi_{\pi}(\varphi_{[X]}(R))))$
 $= \delta(z[R^{\phi}_1] \oplus \dots \oplus z[R^{\phi}_{i\phi}])$. (6)

7. DEFUZZIFICATION PROCESS

Let f denote the system's fused fuzzy output variable. We then have $f = z[R^{\phi}_1] \oplus \dots \oplus z[R^{\phi}_{i\phi}]$. The defuzzification process δ applies the defuzzification function δ to the fused fuzzy output variable to produce its equivalent crisp value z. Sometimes it is useful to just examine the fuzzy output variables produced by the decomposition process to obtain the decision support information the security manager needs, but more often, this fuzzy variable f needs to be converted to a crisp value easy to interpret.

8. RESPONSE PROCESS

Security management may either interpret the fuzzy output and determine the appropriate response strategy given the level of adversity to the current security policy. The fuzzy output variable explains the current conditions of the enforcement state of the security policy. Alternatively, the security management can examine the crisp value representing the fuzzy output variable which should be easier to interpret. Any easy interpretation may be expressed in terms of a decision rule as follows:

If $z \geq z_0^+$, then the situation is critical and there is need to take immediate response actions.

If $z \leq z_0^-$, then security management may be satisfied to issue warning to all stakeholders relevant to the current conditions of the enforcement state of the current security policy.

If z is between z_0^- and z_0^+ then security management may use subjective judgment or negotiate with security policy stakeholders to determine the appropriate correction actions.

9. DEMONSTRATION OF THE FESP

We next provide an example explaining the previous declaration of variables.

$X=(X_1, X_2, X_3) \in D(X)=[0, 100] \times [0, 100] \times [0, 100] \in R^3$.

X_1 = Percentage of users whose password age is higher than the maximum number of days set by the security policy (crisp value)

X_2 = Percentage of users whose password length satisfy the 14 characters required by the security policy (crisp value)

X_3 = Percentage of users whose password string does not satisfy the use of upper and lower case letters, numbers, and special characters as specified in the security policy

R=security policy rule base = { R_1 }

R1: 'If X1 is High and X2 is High then z is High'

R2: 'If X1 is Low and X3 is Low then z is Low'

R3: 'If X1 is Low and X2 is Low then z is Low'

$I(R_1) = \{1, 2\}$

$I(R_2) = \{1, 3\}$

z = Fuzzy output defining the security policy adversity indicator

$F_1 \equiv G_1 = \text{High} =$ Fuzzy subset defined by

$\{\mu_{\text{High}}(t)=t/100, \text{ for } 0 \leq t \leq 100; \text{ and } 0, \text{ elsewhere}\}$

$F_2 \equiv G_2 = \text{Low} =$ Fuzzy subset defined by $\{\mu_{\text{Low}}(t)=1-t/100, \text{ for } 0 \leq t \leq 100; \text{ and } 0, \text{ elsewhere}\}$.

In this example, the security policy rule base consists of two rules, R1: 'If X1 is High and X2 is Low then z is High,' and rule R2: 'If X1 is Low and X3 is Low then z is Low'. The first rule means that if the user has not changed his password as indicated in the password security policy then there is high adversity since the security policy is being violated. There is even more adversity if the security management does not conduct a sufficient number of password audits, expressed by a low X2. The second rule means that if the percentage of users who violate the security policy password length requirements, and the percentage of users violating the security policy requirements for constituting a password, then the adversity indicator of the security policy is low.

Assume $X=(15,40,90)$. We then have $\mu_{\text{High}}(15)=.15$ and $\mu_{\text{High}}(40)=.40$. Let the minimal degree of truth be $\alpha_0=.12$. The alpha value for R_1 equals $\alpha(R_1)=\text{Min}\{.15, .40\}=.15$. Because $\alpha_1(R_1) \geq \alpha_0$, we say that R_1 is fired and will hence be used in the inference process.

We can however see that rule R_2 is not fired by the same input stream $X=(15,40,90)$. In fact, we have $\mu_{\text{Low}}(15)=.85$ and $\mu_{\text{Low}}(90)=.10$. Then, The alpha value for R_2 equals $\alpha(R_2)=\text{Min}\{.75, .10\}=.10$. Because $\alpha(R_2) < \alpha_0=.12$, we say that R_2 is not fired and will not hence participate in the inference process.

On the other hand, if we examine rule R_3 , we find out that this rule is fired. In fact, we have $\mu_{\text{Low}}(15)=.85$ and $\mu_{\text{Low}}(40)=.60$. Then, the alpha value for R_3 equals $\alpha(R_2)=\text{Min}\{.85, .60\}=.60$. Because $\alpha(R_2) \geq \alpha_0=.12$, we say that R_3 is fired and will hence participate in the inference process. We then can have: $\varphi_{[X]}(R) = \{R^{\phi}_1, R^{\phi}_3\} = \{R_1, R_3\}$.

The fuzzification process produces $z =$

$$\delta(\Sigma_{\sigma}(\Pi_{\pi}(\varphi_{[X]}(R)))) = \delta(\Sigma_{\sigma}(\Pi_{\pi}(\{R^{\phi}_1, \dots, R^{\phi}_f\}))) = \delta(\Sigma_{\sigma}(\Pi_{\pi}(\{R_1, R_3\}))).$$

The inference process produces $z = \delta(\Sigma_{\sigma}(\Pi_{\pi}(\varphi_{[X]}(R)))) =$

$$\delta(\Sigma_{\sigma}(\{z[R^{\phi}_i], i=1, f^{\phi}\})) = \delta(\Sigma_{\sigma}(\{z[R_1], z[R_3]\})) \text{ where:}$$

$$z[R_1] = \{t/100 \text{ for } t < 15; .15 \text{ for } 15 \leq t \leq 100; 0 \text{ elsewhere}\}.$$

$$z[R_3] = \{.6 \text{ for } t < 40; 1-t/100 \text{ for } 40 \leq t \leq 100; 0 \text{ elsewhere}\}.$$

The composition process produces the following:

$$z = \delta(\Sigma_{\sigma}(\Pi_{\pi}(\varphi_{[X]}(R)))) = \delta(z[R^{\phi}_1] \oplus \dots \oplus z[R^{\phi}_{i\phi}]) = \delta(z[R_1] \oplus z[R_3])$$

Using the Min-Max inference method we obtain the following fused fuzzy output:

$$f = z[R_1] \oplus z[R_3] = \{.6 \text{ for } t < 40; 1-t/100 \text{ for } 40 \leq t \leq 85; .15 \text{ for } 85 \leq t \leq 100; 0 \text{ elsewhere}\}.$$

Then, after applying the average-of-maximum defuzzification method to the fused fuzzy output, we obtain $\delta(f) = 20$.

Assuming the response decision rule with light adversity if z is less than 30; moderate between 30 and 70; and critical for z higher than 70, the FESP then recommends that warning should be issued to all relevant security policy stakeholders, explaining the current conditions of the enforcement state of the security policy.

9. CONCLUSION

Given its importance in an organization, security policy has to be well written, structured, and continuously revised. The enforcement of corporate security policy is obviously consequential to the organization. This article designed a fuzzy expert system capable of enforcing it by tracking possible violations of the security policy and reporting decision support information to security management. Based on the proposed fuzzy expert system recommendations, the security officer can plan the appropriate response actions towards enforcing the security policy.

References

- [1] D.W Aha, and R.L. Bankert, Feature selection for case-based classification of cloud types: An empirical comparison. In *Proceedings of the 1994 AAAI Workshop on Case-Based Reasoning*, 106-112. Seattle, WA: AAAI Press, 1994.
- [2] H Almuallim, and T.G. Dietterich, Learning boolean concepts in the presence of many irrelevant features. *Artificial Intelligence*, 69(1-2): 279-305, 1994.
- [3] P.S. Bradley, O.L. Mangasarian, and W.N. Street, Feature selection via mathematical programming, *INFORMS Journal on Computing*, 10(2), 209-217, 1998.
- [4] J.C. Bezdek, *Pattern recognition with fuzzy objective function algorithms*, Plenum Press, 1981.
- [5] S.L. Chiu, Fuzzy model identification based on cluster estimation, *Journal of Intelligent and Fuzzy Systems* 2:267-278, 1994.
- [6] O. Cordón, F. Herrera, A three-stage evolutionary process for learning descriptive and approximate fuzzy logic controller knowledge bases from examples, *Int. Journal of Approximate Reasoning* 17(4): 369-407, 1997.
- [7] O. Cordón, F. Herrera, Hybridising genetic algorithms with sharing scheme and evolution strategies for designing approximate fuzzy rule-based systems, *Fuzzy Sets and Systems*, 1999.
- [8] J.C. Dunn, A fuzzy relative of the ISODATA process and its use in detecting compact well separated clusters,

Journal Cybernetics 3:3, 32-57, 1974.

- [9] F. Herrera, M. Lozano, J.L. Verdegay, A learning process for fuzzy control rules using genetic algorithms, *Fuzzy Sets and Systems* 100, 143-158, 1998.
- [10] R.I. John, P. R. Innocent, and M.R. Barnes, "Type 2 fuzzy sets and neurofuzzy clustering of radiographic tibia images," *Proceedings of Third Joint Conference on Information Science*, vol. 1, pp. 58-61, 1997.
- [11] G.J. Klir, and T.A. Folger, *Fuzzy Sets: Uncertainty and Information*. Prentice Hall, Englewood Cliffs, NJ, 1988.
- [12] A.F. Kohn, L.G.M Nakano, and S.M. Oliveira, A class discriminability measure based on feature space partitioning. *Pattern Recognition*, 29, 873-887. 1996.
- [13] C.C. Lee, "Fuzzy logic in control systems: Fuzzy logic controller, part II," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 20, no. 2, pp. 419-435, 1990.
- [14] H. Liu, and R. Setiono, Some issues on scalable feature selection. *Expert Systems with Application*, 15, pp. 333-339. 1998.
- [15] E.H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *Int. Journal of Man-Machine Studies*, vol. 7, pp. 1-13, 1975.
- [16] J.M. Mendel, *Fuzzy Logic Systems for Engineering: A Tutorial*. In Proc. of the IEEE, Special Issues on Engineering Applications of Fuzzy Logic, Vol. 83, No. 3, pp. 345 - 377, 1995.
- [17] K. Nakamura and S. Iwai, "Topological fuzzy sets as a quantitative description of analogical inference and its application to question-answering systems for information retrieval," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-12, pp. 193-204, 1982.
- [18] NIST: Guide for Developing Security Plans for Information Technology Systems <http://csrc.nist.gov/publications/nistpubs/800-18/Planguide.PDF>
- [19] NIST: An Introduction to Computer Security: The NIST Handbook, <http://csrc.nist.gov/publications/nistpubs/800-12>
- [20] NIST: Generally Accepted Principles and Practices for Securing Information Technology Systems <http://csrc.nist.gov/publications/nistpubs/800-14/800-14.pdf>
- [21] B. Raggad, Corporate Vital Defense Strategy: A framework for Information Assurance, The 23rd National IS Security Conference, Baltimore, NSA/NIST, 2000.
- [22] RFC2196, Site Security Handbook, at <ftp://ftp.isi.edu/innotes/rfc2196.txt>.
- [23] T.J. Schwarz, "Fuzzy systems in the real world," *AI Expert*, vol. August, 1990.

PORTABILITY OF USER MODELS WITHIN ADAPTIVE WEB-BASED SYSTEMS

Magdalena Raszková, Arnošt Motyčka

Department of Informatics

Faculty of Business and Economics

Mendel University of Agriculture and Forestry in Brno

Zemedelska 1, 613 00 Brno, Czech Republic

Tel: +420 545 132 237; fax: +420 545 132 245

e-mail: magdalena.raszkova@mendelu.cz, mot@mendelu.cz

ABSTRACT

The adaptivity of a web system represents an ability of a system to respect individuality of a particular user; which can lead to a more effective presentation of information, an improvement of ergonomic quality and accessibility of a system or gaining a competitive advantage in case of commercial system. The adaptation process in adaptive web-based systems proceeds on explicit and implicit feedback from the user. These items of data are saved in so-called user model which predestinates the possible extent and quality of adaptive behaviour of the system. The aim of this paper is to offer a basic view of the issue of user modelling in adaptive web-based systems, portability of the user models and to outline possible future development in this field.

1 INTRODUCTION

The key-stone of adaptive hypermedia systems (AHS) is the respect for diversity of individual users of the system and subsequent presentation (content) adaptation or adaptive navigation related to current knowledge, preferences, aims, abilities, computer literacy of the user, etc.

Presently, the adaptive hypermedia systems are mostly being used in the field of electronic education [1, 2] where it is necessary to adapt the content and navigation to the source knowledge, aims of studies and preferred learning style of the user.

The user model is a crucial point for adaptive behaviour of a system. Under the term *user modelling* in AHS we understand the way of capturing, saving and updating those characteristics of a user which can be used for the needs of the adaptation process.

Changes in the user model in AHS can be initiated in several ways. The more simple way represents direct entries of users made via questionnaires or tests (so-called *explicit feedback*).

However, what makes adaptive systems different from adaptable systems, is the ability to modify the user model automatically on the basis of evaluating his/her interactions with the system (so-called *implicit feedback*). [3]

2 CHARACTERISTICS OF A USER

Generally, any information about the user can serve as his/her user model. A simple log file which records time of access to individual pages or fragments of pages can be considered a very primitive user model.

```
<log>
<user>1234</user>

<record>
<accessdate>Sun Feb 24 07:42:00 CET 2008</accessdate>
<sessionid>58D90A6E2F5FE8FBBBB1F7290E5A1517</sessionid>
<name>File:/tutorial/xml/readme.xhtml</name>
<fragment>false</fragment>
</record>

<record>
<accessdate>Sun Feb 24 07:44:56 CET 2008</accessdate>
<sessionid>58D90A6E2F5FE8FBBBB1F7290E5A1517</sessionid>
<name>File:/tutorial/xml/install.xhtml</name>
<fragment>false</fragment>
</record>
```

Figure 1: Log file in adaptive system AHA! [4]

Within the AHS context the user model represents a set of attributes which describe the user and his/her relation to the model of the domain as concisely as possible. Creating a user model is a complex issue having many forms. It is an essential part of AHS because the method of user modelling predestinates the possible extent and quality of adaptive behaviour of a system. The user model captures characteristics which can be divided into two groups [5]:

2.1 Domain-dependent characteristics

Domain-dependent characteristics extend the definition of domain model concepts with characteristics related to one individual user. In systems supporting education these attributes include for example: time spent on studying a relevant webpage, number of visits of a webpage etc.

2.2 Domain-independent characteristics

Domain-independent characteristics represent those qualities of a user which are valid for more domains, e.g. his/her preferred learning style, favourite colours, knowledge of foreign languages, Internet browser used, limitations (handicap), computer literacy, etc.

2.3 Most frequently modelled characteristics

Characteristics modelled in a particular AHS largely depend on the purpose of the system. We need different characteristics in an educational adaptive system and different in a system for information retrieval. The following list gives most frequently modelled characteristics of a user; these are described in [6] and completed in [7].

Knowledge of the problem domain is one of the most important characteristics modelled in existing adaptive (educational) systems. Approximately one third of adaptation techniques used both for navigation support as well as for adaptive presentation is knowledge-based.

Interests of a user constitute the most important and very often the only part of a user profile in systems which are to retrieve and filter information working with bulk of data. However, even in the other system types the interests of a user are gradually becoming a characteristic of the same importance as knowledge has been so far.

Goals and tasks represent an immediate objective of user's work in the adaptive system. This characteristic of a user is highly variable and the process of analysing goals/tasks of a user is extremely demanding and often inaccurate.

Hyperspace experience generally says what the user knows about the structure of hyperspace (the Internet) and how much he/she is able to be oriented within its structure.

User's background represents a set of all knowledge and experience of a user which, however, are not the topic of the adaptive system itself. These may include the job of the user, knowledge related to the problem domain of a system or language abilities.

Preferences of a user, e.g. favourite colour, font or page layout usually cannot be guessed by the system itself; they are usually gathered by direct questions, or the user is enabled to adjust the system appropriately.

Individual traits is a name for a set of those characteristics of a user which together form his/her individuality; e.g. personality type (extravert/introvert), cognitive abilities, learning style [8].

3 REALIZATION OF A USER MODEL

In AHS the user model is most frequently implemented in the form of structured data in XML format, with the help of languages for description of semantics or in the form of relational database.

3.1 User model in the form of XML structure

XML format or its derivatives appear to be an appropriate format for saving structured data about a user. For example, the AHA! system [4] implements the overlay user model in the form of XML structure (see figure 2) where a concept in the user model exists for each concept of the domain model. In addition, the model contains a pseudoconcept *personal*; the attributes of which save personal information about a user (e.g. login and password).

```
<?xml version="1.0"?>
<!DOCTYPE profile SYSTEM "profile.dtd">
<profile>
  <record>
    <key>personal.name</key>
    <value>Magdalena Raszko</value>
    <firsttimeupdated>>false</firsttimeupdated>
  </record>
  <record>
    <key>personal.course</key>
    <value>ahs</value>
    <firsttimeupdated>>false</firsttimeupdated>
  </record>
  ...
  <record>
    <key>ahs.architektura.visited</key>
    <value>1</value>
    <firsttimeupdated>>false</firsttimeupdated>
  </record>
  <record>
    <key>ahs.architektura.knowledge</key>
    <value>120</value>
    <firsttimeupdated>>false</firsttimeupdated>
  </record>
  <record>
```

Figure 2: The user model in AHA! system in the form of XML structure [4]

3.2 User model in the form of ontology

Ontology can be defined as a formal conceptual model of a certain problem field; conceptions and relations of the model supply semantics for the data. For representations of ontologies the OWL language (Web Ontology Language) is mostly used; it is administered by syndicate W3C [9]. An example of ontology describing characteristics of a user and their relationships is e.g. GUMO (Generalized User Model Ontology) designed by Heckmann et al. [10]. It uses OWL language in combination with RDF (see figure 3).

```
<owl:Class rdf:ID="PhysiologicalState.700016">
  <rdfs:label>
    Physiological State
  </rdfs:label>
  <rdfs:subClassOf>
    rdf:resource="#BasicUserDimensions.700002" />
  <gumo:identifier>
    700016
  </gumo:identifier>
  <gumo:lexicon>
    state of body or bodily functions
  </gumo:lexicon>
  <gumo:privacy>
    high.640033
  </gumo:privacy>
  <gumo:website>
    rdf:resource="#GUMO;concept=700016" />
</owl:Class>
```

Figure 3: Part of GUMO ontology [10]

Ontologies can be a way to overcome the problem of separated models of a user for different applications which prevents their sharing and reuse.

Model sharing is very interesting especially for domain-independent characteristics, but also for domain-dependent characteristics for the same or similar domains. User model represented by ontology can be subsequently accessed via a web service to individual adaptive applications. [11]

3.3 User model in the form of relational database

For realisation of a user model it is also possible to use common relational database systems. For example, the AHA! system [4] enables to save the data either via XML structures or to link the system to MySQL database.

The disadvantage of user models implemented via relational database can be their worse portability between adaptive systems. In contrast to XML structures, their implementation can be platform-dependent.

4 PORTABILITY OF USER PROFILES WITHIN WEB APPLICATIONS


The idea of a unified profile for web and other applications is not new. The project *dataportability.org* [12] deals with the portability of virtual identities within the web. OpenID, one of the main technologies for this unification, has been frequently used for some time now.

OpenID is a decentralized service for identity verification. It is closely linked to the idea of portable digital identity, and respectively it is one of the key elements of so called Identity 2.0. The principle of OpenID can be used in all applications which require login or authorization, e.g. in community webs or analogic applications (discussion forums, services for registered users, etc.). OpenID enables its user to unite all the profiles of his/her own into one, thus creating a complex profile of his/her personality. Several large organizations either issue or accept OpenIDs, including Google, Facebook, Yahoo!, Flickr, MySpace and many more. Instead of filling all required information in the registration, the user can enter only the OpenID address and his/her password and the server downloads all necessary information by itself.

1. CHOOSE YOUR USERNAME

Your OpenID URL is how [sites that accept OpenID](#) know you. You can use you be known by.

Username John Doe, jdoe123

OpenID URL  <http://Magda-Raszko.myopenid.com/>

2. CHOOSE A PASSWORD

You'll use this password to sign in to myOpenID, but you won't have to give it

Password

Password (confirm)

Strength

Figure 4: Registration of OpenID account at myopenid.com

The obligatory data required from the user during his/her registration of OpenID account (e.g. *myopenid.com* [13]) are the username (nickname), password and e-mail.

Some providers of OpenID accounts also enable users to enter optionally their full name, birth date, gender, postal code, country, language and time zone.

If in the future the OpenID format could be extended even with the possibility to enter other characteristics of the user such as preferences, level of computer literacy, language abilities, etc. Subsequently it would be possible to use them for automatic personalization and adaptivity of web applications supporting OpenID.

From the perspective of adaptive web, the APML format (Attention Profiling Mark-Up Language) [14] also seems to be interesting. This XML based format serves for the description of user's interests and their exchange or transmission. In addition, evaluation is added to each interest, i.e. how important a particular interest is for the user. It is also possible to divide user's interests into groups, so it is possible to find out which interests the user has at home and which at work. These preferences could also be used for systems personalization and adaptivity.

5 CONCLUSIONS

The paper dealt with user modelling in adaptive hypermedia systems. Basic characteristics of a user were described as well as approaches to user model implementation. Further, the attention was devoted to the user model portability within the web applications.

The biggest problem in the field of user modelling appears to be an absence of any standardization. Most user models are now created ad-hoc, thus being not portable between individual systems. Data once entered by the user, especially domain-independent data, cannot be used in different systems.

In addition, a user model is always only an estimation of characteristics, therefore it appears convenient to enable the user to edit and specify information which is saved in the system. However, it cannot be expected that the user will upgrade his/her model if this is not portable between individual systems. There should be a tendency to enable at least a transmission of domain-independent characteristics between systems in the future; this will require formalization and standardization of user models. Users could then unite all the profiles of their own from all their web services into one and create an integrated and complex profile of their personalities.

In this respect it would be possible to extend already existing projects such as OpenID or APML. In this context a question arises considering security and protection of personal data, however, this issue goes beyond the scope of this paper.

References

- [1] Raszková, M., Motyčka, A. Adaptivity – the Future of Educational Hypermedia Systems. In *Proceedings of the 11th International Multiconference Information Society*. 2008. vol. A, p. 89–92. ISSN 1581-9973.
- [2] Raszková, M. *Výukové adaptivní hypermediální systémy* [Educational Adaptive Hypermedia Systems]. Diploma thesis. Brno: MZLU, 2008. 75 p.
- [3] Bureš, M., Morávek, A., Jelínek, I. *Nová generace webových technologií* [New generation of web technologies]. vol. 1. Praha: VOX, 2005. 264 p. ISBN 80-86324-46-X.
- [4] De Bra, P. *Publications about the AHA! project and system* [online]. 2007 [cit. 2009-08-23]. <<http://aha.win.tue.nl/>>.
- [5] Bieliková, M. Adaptivna prezentácia hypermedií na webe [Adaptive presentation of hypermedia on the web]. In *Datakon 2003*. vol 1. Brno, 2003, p. 72–91. ISBN 80-210-3215-4.
- [6] Brusilovsky, P. Adaptive Hypermedia. *User Modeling and User-Adapted Interaction*, vol. 11, no. 1–2. Dordrecht: Kluwer Academic Publishers, 2001, p. 87–110. ISSN-0924-1868.
- [7] Brusilovsky, P., Millán, E. *The Adaptive Web: Methods and Strategies of Web Personalization*. Berlin, Heidelberg: Springer Verlag, 2007. ISBN 978-3-540-72079-9. User Models for Adaptive Hypermedia and Adaptive Educational Systems, p. 3–53.
- [8] Malo, R. Diagnostika učebních stylů žáků a návrh jejich podpory v eLearningu [Learning styles diagnostic and suggestion for their support within eLearning]. Brno: MZLU, 2009. Bachelor thesis. <https://is.mendelu.cz/zp/portal_zp.pl?podrobnosti=33203;download_prace=1>.
- [9] W3C. *OWL Web Ontology Language Overview* [online]. 2004 [cit. 2009-09-07]. <<http://www.w3.org/TR/owl-features/>>.
- [10] Heckmann, D. *GUMO – the General User Model Ontology* [online]. 2005 [cit. 2009-09-07]. <http://www.inf.unibz.it/~ricci/ATIS/papers/UM05_Gumo.pdf>.
- [11] Šaloun, P., Bieliková, M. Adaptivní webové systémy pro vzdělávání [Adaptive web-based systems for education]. In *DIVAI 2007*. Nitra: Univerzita Konštantýna Filosofova v Nitre, 2007. p. 1–11. ISBN 978-80-8094-123-9.
- [12] *DataPortability Project* [online]. 2009 [cit. 2009-09-08]. <<http://dataportability.org>>.
- [13] *myOpenID*. First and Largest Independent OpenID Provider [online]. 2009 [cit. 2009-09-08]. <<http://www.myopenid.com>>.
- [14] *APML*. Attention Profiling Mark-up Language: The open standard for Attention Metadata [online]. 2009 [cit. 2009-08-08]. <<http://www.apml.org>>.

Acknowledgements

The research behind this paper is supported by grant agency of MUAF in Brno within a project no. IG190621 called *Domain-Independent Model for Adaptive Web-Based Systems*.

GENE REGULATORY NETWORKS INFERENCE USING GRAPHICAL GAUSSIAN MODELS

Blagoj Ristevski¹, Suzana Loskovska²

¹Department of Information Systems Management
Faculty of Administration and Information Systems Management
St. Kliment Ohridski University – Bitola, Republic of Macedonia
e-mail: blagoj.ristevski@uklo.edu.mk

²Department of Computer Science and Informatics
Faculty of Electrical Engineering and Information Technologies
Ss. Cyril and Methodius University – Skopje, Republic of Macedonia
e-mail: suze@feit.ukim.edu.mk

ABSTRACT

Gene regulatory networks consisting of genes, transcription factors, microRNAs and other components are complex networks which reveal and depict the fundamental gene regulatory mechanisms. In this paper we present the inference of gene regulatory networks based on graphical Gaussian models from gene expression microarray data. The microarray data might be seen as a random sample of a multivariate distribution defined by random variables associated to the genes. Thus, one manner of describing the interaction among genes is by using of conditional dependencies between genes. These independencies can be inferred using structure learning of graphical Gaussian models which are based on multivariate distributed data. As input data we use microarray data and for inference and visualization several R packages. Also we emphasize the importance of including the data obtained from ChIP-chip and ChIP-Seq experiments in gene regulatory networks modeling to obtain more reliable networks and the role which transcription factors and microRNAs have in gene expression regulation.

Keywords – *gene regulatory networks, graphical Gaussian models, inference of gene regulatory networks, structure learning*

1 INTRODUCTION

The high-throughput microarray technology provides observation of gene expression under tens or hundreds of experimental conditions. Currently, several of most challenging tasks in bioinformatics are revealing and describing the interactions among genes, transcription factors, microRNAs and other metabolites in genome level, the learning of large networks' interactions enabled by gene expression microarray data and other data obtained from ChIP-chip and ChIP-Seq experiments.

The microarray data might be seen as a random sample of a multivariate distribution defined by random variables associated to the genes. Each sample corresponds to a vector

of values which quantify the abundance of the messenger RNA produced by each functional genome element under specific experimental conditions.

Hence, one manner of revealing the interaction among genes is by using of conditional independencies which can be inferred using structure learning of graphical Gaussian models.

The remainder of this paper is organized as follows. In the second section we present the graphical Gaussian models and their assumptions and usage in the networks modeling. We depict the partial correlation coefficient, the model selection and network visualization. Network inference and visualization for *Arabidopsis thaliana* time series gene expression data as an input using several R packages and the package *graphviz* is presented in the following section. The concluding remarks and future directions are given in the last section.

2 GRAPHICAL GAUSSIAN MODELS

Besides modeling of gene regulatory networks using Boolean networks, Bayesian networks, dynamic Bayesian networks, Petri networks, state-space models, ordinary differential and difference equations, there are other models which can be used for gene networks reconstruction from gene expression data. The graphical models are very appropriate for bioinformatics modeling and inference. They allow a stochastic description of network association and sophisticated statistical framework for inference. Thus, they are suitable for inference and description of biochemical interactions and regulatory activities [9].

Graphical Gaussian models (GGMs) are often used as a method for gene regulatory networks modeling based on gene expression data and they are very computationally efficient [4]. GGMs as graphical probabilistic models can identify conditional independence relations among the nodes. They assumed that there is a multivariate Gaussian distribution of the input data.

The networks composed of nodes and edges are presented as undirected graphs. Their nodes represent genes, and the edges represent conditional dependence relations between

nodes. There is not an edge between two nodes if the nodes are conditionally independent.

Let Y be the input gene expression data matrix with G columns, corresponding to the number of genes, and with N rows which correspond to the number of samples (time series data points or other experimental conditions) [4]. It is supposed that matrix Y follows a multivariate normal distribution $N_G(\mu, \Sigma)$, where $\mu = (\mu_1, \dots, \mu_G)'$ is the mean vector, and $\Sigma = (\sigma_{ij})_{1 \leq i, j \leq G}$ is the positive definite covariance matrix. $\sigma_{ij} = \sigma_i \sigma_j$ are covariance parameters between genes i and j , and σ_i^2 are related to the variance terms for gene i . The estimation of the covariance matrix of the data distribution is a base for the GGMs inference.

2.1 Partial correlation

First, in the inference of graphical Gaussian model, to make a reliable estimation of the partial correlation matrix $\tilde{P} = (\tilde{\rho}_{ij})$ is required [5]. This matrix is related to the inverse matrix of the covariance matrix Σ . The straightforward estimator is given by the following Eq. (1):

$$\tilde{r}_{ij} = -\frac{\hat{\omega}_{ij}}{\sqrt{\hat{\omega}_{ii} \hat{\omega}_{jj}}} \quad (1)$$

where

$$\hat{\Omega} = (\hat{\omega}_{ij}) = \hat{\Sigma}^{-1} \quad (2)$$

The partial correlation coefficients \tilde{r}_{ij} , which describe the correlation between nodes/genes Y_i and Y_j conditional on all other nodes in the network, are measures of the direct interactions among nodes/genes. Differently from Pearson correlation coefficient which records correlation between gene pairs without regarding to other genes, partial correlation between two genes measures the degree of correlation remaining after removing the effects of the other genes [1] [8].

The above mentioned procedure is appropriate when N is larger than number of genes G , otherwise the covariance matrix is not positive-definite and its inverse matrix cannot be found. In microarray datasets, the sample size N is much smaller than the number of genes G .

For that reason, it is suggested using a shrunk estimate of the covariance matrix. The goal is constructing well conditioned positive definite matrix, so the matrix can be inverted. If λ is a shrinkage coefficient so that $0 \leq \lambda \leq 1$, then shrunk covariance matrix Σ^* is computed by following Eq. 3

$$\Sigma^* = \lambda T + (1 - \lambda)S \quad (3)$$

where \hat{S} is estimated empirical covariance matrix. The shrinkage parameter λ is chosen to minimize the mean-square error and it is determined analytically given by Eq. 4.

$$\lambda^* = \frac{\sum_{i \neq j} \text{var}(r_{ij})}{\sum_{i \neq j} r_{ij}^2} \quad (4)$$

After computing of partial correlation coefficients \tilde{r}_{ij} , the distribution of $|\tilde{r}_{ij}|$ is checked and the edges with significantly small values of $|\tilde{r}_{ij}|$ are removed from the network [3].

2.2 Model selection

The second stage of the inference of GRNs is assigning statistical significance to the edges from the GGM network or in other words model selection.

The usage of the massively parallel structure learning of testing $G(G-1)/2$ edges simultaneously, offers an approach based on Bayesian estimation and inference. A mixture model obtained by Eq. 5 is fitted to the observed partial correlation coefficients \tilde{r}_{ij} .

$$f(\tilde{r}_i) = \eta_0 f_0(\tilde{r}; \kappa) + (1 - \eta_0) f_A(\tilde{r}) \quad (5)$$

where f_0 denotes the distribution under the null hypothesis of disappearance of partial correlation, η_0 denotes the unknown proportion of null edges, and f_A is the distribution of observed partial correlations assigned to existing edges.

The empirical Bayes local false discovery rate (*fdr*) statistics fits with above mixture model and additionally it takes account of the dependencies among the estimated partial correlation coefficients. The Eq. 6 gives the posterior probability that specific edge exists for given \tilde{r}_{ij} .

$$P(\text{non-null edge} | \tilde{r}) = 1 - \text{fdr}(\tilde{r}) = 1 - \frac{\eta_0 f_0(\tilde{r}; \kappa)}{f(\tilde{r})} \quad (6)$$

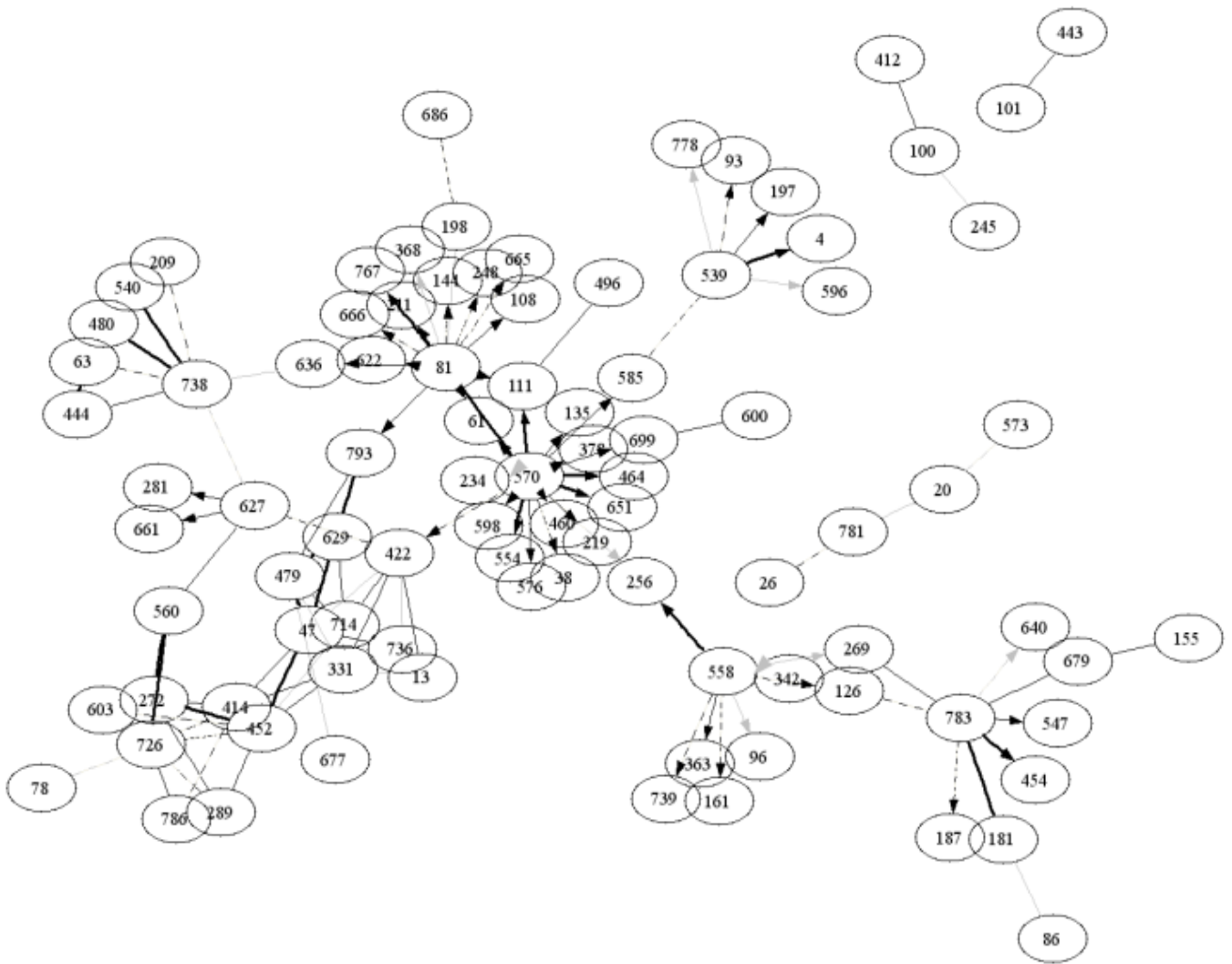
2.3 Network visualization

The following stage is network visualization. The above described inference is performed using GeneNet package and Rgraphviz and graphviz for network plotting. For network inference and visualization other R packages were used from Bioconductor web page (<http://www.bioconductor.org>) and R package archive (<http://cran.r-project.org>).

3 RESULTS

For network inference we use as input dataset of Arabidopsis Thaliana microarray gene expression data. The input data should be properly preprocessed. The dataset is time-series data for 800 genes measured in 11 different time points. The obtained network (Fig. 1) using GeneNet [5] and other R packages consists of 118 most significant edges which present only 0.04% from all 319.600 possible networks.

The Fig. 1 shows the hubs - genes with highest degree. The gene numbered as 570 has 18 edges; genes labeled as 47, 422, 452, 558 and 783 contain 8 edges. Full lines indicate positive regulation, otherwise dotted lines - negative regulation. The intensity of edge lines correspond to magnitude of correlation. Black lines indicate edges with the highest 20% partial correlation values. The gray ones indicate edges at the lower 20% partial correlation values.



Arabidopsis thaliana network

Figure 1. Gene regulatory networks obtained using graphical Gaussian models. The input dataset is Arabidopsis thaliana microarray gene expression data for 800 genes.

4 CONCLUSION AND FUTURE DIRECTIONS

Using GGMs the submodels set of the selected models can identify limited search space where some structure learning algorithm might be applied [2]. To be efficient GGMs structure learning should take into account an available prior knowledge related to interaction among genes [6].

The gene regulatory networks inference based only on microarray gene expression data has several shortcomings regarding to the number of available microarrays, experimental noise, crosshybridization errors, quality of data samples [7]. The further directions for inference of more reliable gene regulatory networks should include the data obtained from ChIP-chip (chromatin immunoprecipitation-chip) and ChIP-Seq (chromatin

immunoprecipitation-Sequencing) experiments, which both include investigation of transcription factors interaction and their corresponding DNA binding sites. Combining of gene expression microarray data and these data offer more biologically relevant inferred networks. These techniques are utilized to study the role of transcription factors in the regulation of gene expression. Besides transcription factors, the further models development should take into account the microRNAs because they also have a significant role in gene expression regulation. MicroRNAs are small non-coding RNAs that control gene expression at transcriptional and posttranscriptional level.

By binding to cis-regulatory elements, both microRNAs and TFs can control several to hundreds of target genes

[10]. TFs and other genes which contain binding sites for other TFs have a high possibility to be targeted by microRNAs. These interactions make microRNAs to be main components in gene regulatory networks, particularly they have key role in several diseases like cancer, cardiovascular, infectious, rheumatologic, metabolic and other diseases [11].

REFERENCES

- [1] N. Kraemer, J. Schaefer, A.-L. Boulesteix, *Regularized Estimation of Large-scale Gene Association Networks Using Graphical Gaussian Models*, Technical Report, Department of Statistics, University of Munich, 2009.
- [2] A. Roverato, R. Castelo, *Structural Learning of Gaussian Graphical Models from Microarray Data with p Larger than n* , *Statistica*, n.4, 2006.
- [3] A. V. Werhli, M. Grzegorzcyk, D. Husmeier, *Comparative Evaluation of Reverse Engineering Gene Regulatory Networks with Relevance Networks, Graphical Gaussian Models and Bayesian Networks*, *Bioinformatics*, Vol.22, no. 20, 2006.
- [4] F. Jaffrezic, G. Tosser-Klopp, *Gene Network Reconstruction from Microarray Data*, *BMC Proceedings*, 2009.
- [5] J. Schaefer, R. Opgen-Rhein, K. Korbinian, *Reverse Engineering Genetic Networks using GeneNet Package*, *R-News* 6/5:50-53, 2006.
- [6] G. Marrelec, H. Benali, *Asymptotic Bayesian Structure Learning Using Graph Supports for Gaussian Graphical Models*, *Journal of Multivariate Analysis*, 2006.
- [7] W. Zhao, E. Serpedin, E. R. Dougherty, *Recovering Genetic Regulatory Networks from Chromatin Immunoprecipitation (ChIP) and Steady-State Microarray Data*, *EURASIP Journal on Bioinformatics and Systems Biology*, 2008.
- [8] S. Ma, Q. Gong, H. J. Bohnert, *An Arabidopsis Gene Network Based on the Graphical Gaussian Model*, *Genome Research*, 2009.
- [9] J. Schaefer, K. Strimmer, *Learning Large-Scale Graphical Gaussian Models from Genomic Data*, *American Institute of Physics*, Vol. 776, 2005.
- [10] O. Hobert, *Gene Regulation by Transcription Factors and MicroRNAs*, *Science*, Vol. 319, no. 5871, 2008.
- [11] A. Drakaki, D. Iliopoulos, *MicroRNA Gene Networks in Oncogenesis*, *Current Genomics*, 2009

TIME SERIES FORECASTING USING MACHINE LEARNING METHODS

Michael Stencl, Ondrej Popelka, Jiri Stastny

Department of Informatics

Mendel University of Agriculture and Forestry in Brno

Zemedelska 1/1665, 613 00 Brno

Tel: +420 545 132 728; fax: +420 545 132 245

e-mail: michael.stencl@mendelu.cz, ondrej.popelka@mendelu.cz, stastny@fme.vutbr.cz

ABSTRACT

In this paper we concentrate on prediction of future values based on the past course of that variable, traditionally these are solved using statistical analysis [16] - first a time-series model is constructed and then statistical prediction algorithms are applied to it in order to obtain future values. This paper describes Radial Basis Functions (RBF) Neural Network and Two-level Grammatical Evolution. Both these methods are applied to solve prediction of simplified numerical time series. Sample dataset includes forty generated observations and the goal is to predict five future values.

1 INTRODUCTION

In this paper we concentrate on prediction of future values based on the past course of that variable. Traditionally prediction problems are solved using statistical analysis [13]. As an alternative to statistical methods of time-series modelling we describe two machine learning methods. These can generally be divided into techniques with analytical approach [2] and techniques inspired by biological processes. Most of the biologically inspired techniques are learning methods. Learning methods accept a set of sample data from which they autonomously learn patterns and trends.

2 Radial Basis function Neural Networks

Unlike statistical methods Artificial Neural Networks do not use a separate algorithm for prediction [10], [13]. Forecasting of future values with artificial neural networks is based solely on learned patterns from the input data.

Radial Basis Function neural networks (RBF-NN) belong to the group of feed-forward models of neural networks. A RBF-NN consists of three layers of nodes (Fig. 1). The first is the input layer which transports the input vector to each of the nodes in the hidden (second) layer. Each node of the input layer is connected to each node of the hidden layer. The hidden layer represents data clusters (so-called areas) which form around a node and contain neighbouring nodes in some perimeter. As described in [14] the local units have the relevant output localized to the point in close neighbourhood defined by its parameters. When an input

vector arrives on some nodes of the hidden layer simultaneously, each node calculates the distance from the input vector to its own centre [14]. The third layer has only one node which sums the outputs of the hidden layer of nodes to yield the decision value [14].

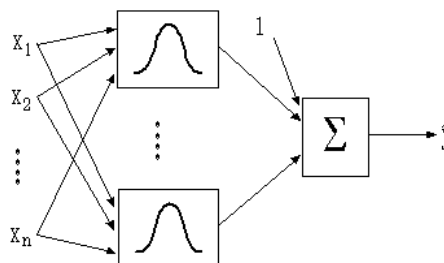


Figure 1: RBF-NN architecture

2.1 RBF-NN learning

In the learning process the values of the training set are gradually fed to the neural network, each time a percept is obtained. Since the hidden layer in this network is represented by so-called areas and the middles of the areas are fast discovered, the larger part of the learning process involves only setting of scales and thresholds of the output layer. Gradient method and Last Mean Square (LMS) methods were tested for defining learning error of the neuron network. These gradient methods use relations derived for outgoing layer using Back-propagation (BP) algorithm. The main difference from Back-propagation is that this method optimizes only scales and thresholds of the outgoing layer [11], [12].

Least Mean Square (LMS) method tries to find optimal scaled vector for general middle quadratic error of the network. This scaled vector is given by normal division:

$w = (H^T H)^{-1} H^T y$, where w is scale vector, H is suggestion matrix $H_{ij} = h_j(x_i)$ and y is vector of outgoing values. This method in contrast to others ones uses transient functions of outgoing neurons layer in place of sigmoid linear function [9], [10].

RBF-NN learning is divided into two stages:

1. **RBF layer neurons learning** (prototypes learning). In the first stage, prototype C and σ are determined for each RBF neuron. For example the algorithms for

cluster analysis can be used. To speed up the first stage, non-adaptive methods can also be used such as uniform or random distribution of RBF neuron centres over the input space.

2. **Output layer neurons learning.** The objective of the second stage of learning is to determine the weights of input neurons, for example the least square method or gradient algorithms can be used.

Prototypes learning. First the number of clusters in input data is estimated, the pertinence function of model m to the cluster is defined and the coordinates of all p vectors C_p being the centres of clusters are estimated [11].

Both stages includes K-Means algorithm used for powerful learning of RBF-NN. Steps of K-Means algorithm:

- Initialize RBF neuron centres C in random.
- Calculate $m()$ for all samples from the training set.
- Calculate new centres C as the average of all samples that pertained to centre k by the pertinence function.
- Terminate if $m()$ does not change, otherwise continue with point 2 [12].

var ::= x_{t-1}	<fnc> ::= --	<expr> ::= <var>
x_{t-2}	ln	<fnc><expr>
x_{t-3}	exp	<fnc><expr><expr>
x_{t-4}	sin	
a	cos	
b		
c		
	<fnc> ::= -	
	+	
	•	
	÷	
		Legend
	fnc	function
	var	variable
	expr	expression

Figure 2: Production rules for arithmetic expressions

The centres (prototypes) of neurons are set so that RBF neurons are represented by model clusters for the best. On the other hand, the learning process could be also described as three stages model. In the first stage, centres c_j are determined for each RBF neuron. The centres c_j are represented by the weights between the input and the hidden layer. For example the cluster analysis algorithms are used. To speed up this stage, non-adaptive methods can also be used such as the uniform or random distribution of RBF neuron centres over the input space. The second stage setups other values of the RBF neurons. The setup values of the RBF neurons (b_j) determine the wideness of the area around the estimated centres of c_j . The objective of the third stage of learning is to determine the weights of the input neurons, for example the least square method or gradient algorithms can be used. In global view, the RBF-NN learning includes unsupervised learning in the first stage. In the second stage the RBF neurons are setup. The typically used function for this setup is the Gaussian Radial Basis Function defined. The RBF neuron determines the important output values in the radial zone with centre in c . The b represents the width of ϕ and determines the size of the radial zone. The setup parameters of the RBF neuron determine the wideness of the controlled area and affect the

generalization capability of the network. If the parameters are smaller means lower generalization capability and on the other hand for wider area the units lost their local mean. In the last stage, the supervised learning is used. The last stage setups the weights w_{sj} . The setup is made by the mineralization process of a typical error function [1].

Description of implementation. Each neuron in the radial basis layer will give the output the value that depends on how close the input vector is to each of the weight vectors of the given neurons. Thus the RBF neurons whose weight vectors are a bit different from input vector p have an output close to zero. On the contrary, the RBF neuron whose weight vector is close to the input vector will have a value close to 1. Individual neuron layers have the form of one-dimensional array. The weight matrix is in the form of two-dimensional array, where the index gives the number of neurons being connected. It is necessary to enter the number of RBF neurons n for one category.

3 Two-level grammatical evolution

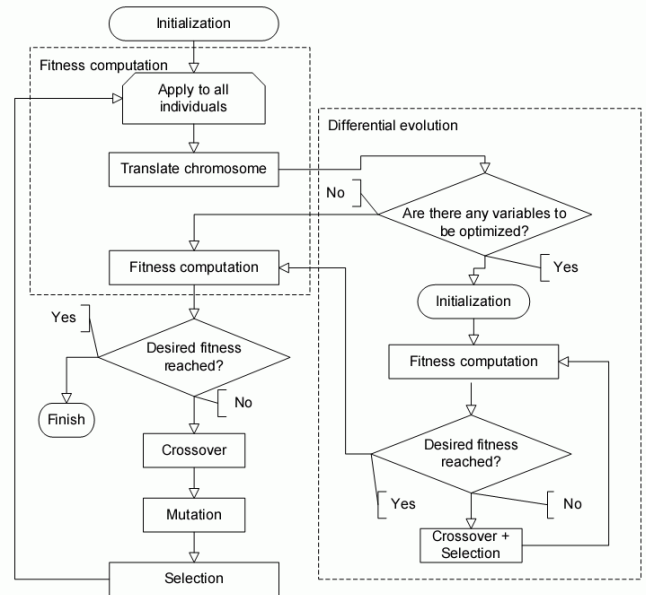


Figure 3: Two-level grammatical evolution flowchart

Two-level grammatical evolution comprises of grammatical evolution with backward processing [4] and differential evolution algorithm. Grammatical evolution is based on a genetic algorithm extended with a translation layer inserted between the chromosome and the actual solution. This layer is formed by a processor of context-free grammar which enables the algorithm to create and optimize generic tree structures and retrieve them in arbitrary reusable format defined in a formal language [5].

Table 1: Definition of grammar used

$$\Pi = \{\text{expr, fnc, var}\}$$

$$\Sigma = \{\text{sin, cos, exp, ln, -, +, *, /, } x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, a, b, c\}$$

$$S = \text{expr}$$

A context-free grammar G is defined as a tuple $G = (\Pi, \Sigma, P, S)$ where Π is set of non-terminals, Σ is set of terminals, S is initial non-terminal and P is table of production rules. We have chosen to represent the time-series using an autocorrelation model. Definition of the grammar is shown in Table 1, for brevity production rules are shown separately in BNF notation on Figure 2.

Non-terminals are items, which can appear in the individuals' body (that is in the solution) only before or during the translation. Terminals are all symbols which appear in the language generated by the grammar, thus they represent the solution. Start symbol is one non-terminal from the non-terminals set, it is used to initialize the

translation process. Production rules define the laws of translation of non-terminals to terminals. The set of non-terminals we used contains arithmetic operators, basic arithmetic functions, variables and constants.

The process of generating a solution is divided into two steps. First an acceptable model is found which represents trends or relations in the data. In the second step the parameters of the model are found and optimized so as to fit the specific dataset. Although grammatical evolution is capable of solving both tasks together [4] it proved to be more flexible to process them separately by using two-level grammatical evolution [6]. For the second step a Differential evolution (DE) algorithm [8] can be used.

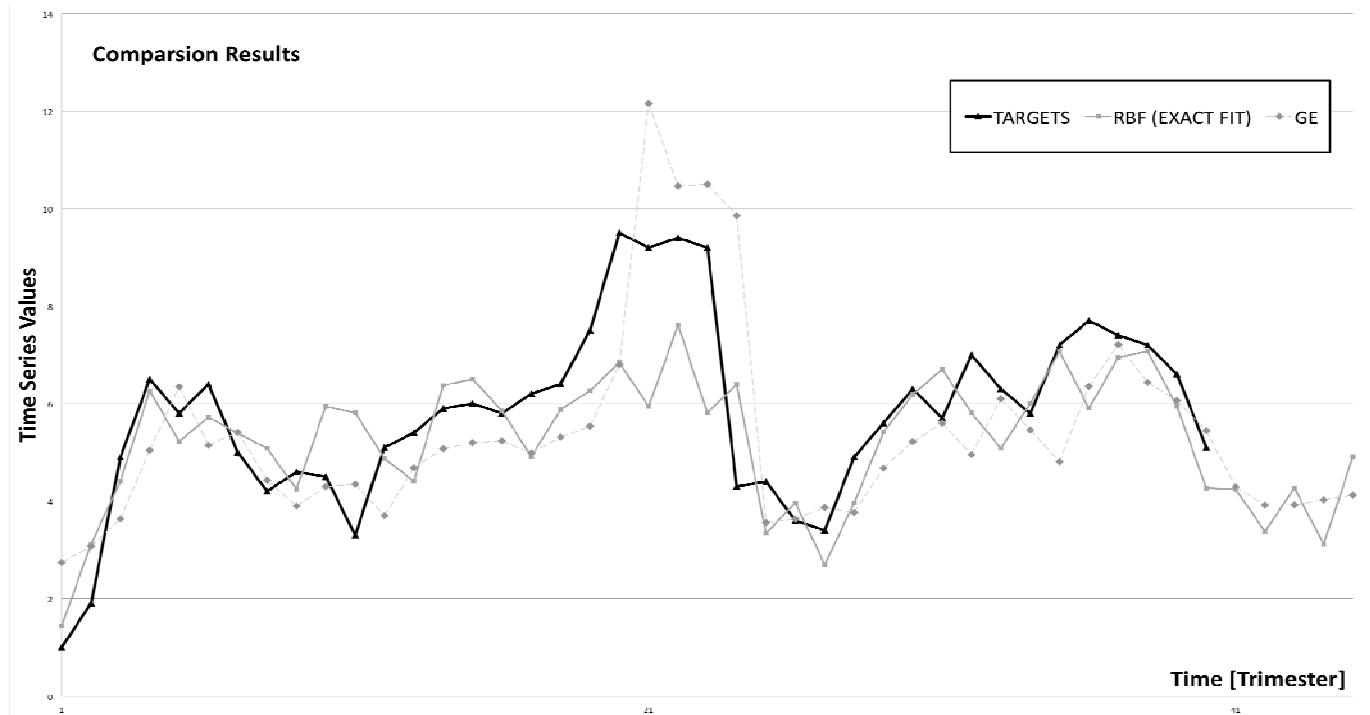


Figure 4: Comparison of trained RBF-NN (Exact Fit) a two-level grammatical evolution on Target Data

Similar to grammatical evolution differential evolution is also a modification of genetic algorithm; it is especially suitable for optimizing of real valued parameters of functions since there is no bias in chromosome encoding [7]. This is very favourable for the crossover and mutation operators.

Each individual in DE is represented with a vector $x_{i,G}$, where $i = 0, 1, \dots, NP - 1$ and NP is number of individuals in the population. The basic scheme (DE/rand/1) involves generation of a perturbed vector of three randomly chosen vectors. The new vector is generated using the formula $v_{i,G+1} = x_{r1,G} + F(x_{r2,G} - x_{r3,G})$, where $x_{r1,G}$, $x_{r2,G}$ and $x_{r3,G}$ are randomly chosen and different individuals and F is amplification factor of the difference. The new vector v_i is inserted in the population in generation $G+1$ only if it has better fitness value then the original vector.

The first level of the optimization is performed using grammatical evolution. According to the grammar (Figure 2) the output can be a function containing variables (x_{t-1} , x_{t-2} , x_{t-3} , x_{t-4}) several symbolic constants (a , b , c). Such function therefore cannot be evaluated and assigned a fitness value. In order to evaluate the generated function a secondary optimization has to be performed to find values for constants.

A flowchart diagram of the two-level optimization is shown on Figure 3. Basically it consists of two nested population loops. The inner loop is using standard differential evolution with DE/rand/1 scheme. The outer loop is a single population of parallel grammatical evolution. This means that in order to evaluate fitness for an individual in GE population it is necessary to create DE population, pass the unknown constants as variable vector to it (each individual can have different number of constants), optimize this vector using DE and substitute optimal values back to GE

individual. The resulting Two-level Grammatical Evolution takes advantage of both the original methods.

6 RESULTS

Sample dataset consists of 40 randomly generated values; values lie in interval $\langle 1, 9.5 \rangle$. First part of the experiment was performed using RBF-NN implementation in Matlab R2007a based on the previously described principles. The dataset was divided into input data and validation data in order to obtain better generalization of the results. Second part of the experiment was realized using our own implementation of two-level grammatical evolution [7]. Grammar used for configuration of the algorithm is described in previous chapter – it is tailored to generate autoregressive formulas of time-series. It is now important to emphasize the difference between both methods. RBF neural network uses fast learning algorithms which adjust the weights in the network so that the network represents trends in the given time-series. Training a RBF neural network is a matter of minutes, but it can provide us only with numeric results. On the other hand Two-level grammatical evolution is a complex algorithm which not only learns the trends in the time-series, it also provides the user with exact description of the time-series. The output of grammatical evolution is both the output data and the formula to obtain them.

Figure 4 shows the input data and two sample runs of both methods. Both methods generally agree on the future values of the time-series. As noted above the output of neural network is only the values displayed, the output of grammatical evolution is both the values displayed and the formula which in this case is

$$\frac{((a/36) + x_{t-3})}{x_{t-1} - 12}, \quad (3)$$

where $a = -1324.739$. Therefore Grammatical evolution provides us with more information, but its training would take much more time than training a neural network. In this simple case it about 10 generations of the underlying genetic algorithm which takes about 40 minutes (about 4 times longer) on the same computer.

Both methods provide comparable results in terms of accuracy. The main difference is in the form in which the results are obtained and in learning performance. Briefly this could be described that it is either possible to obtain numerical results in short time, or use more complicated algorithm to obtain autoregressive formula of the time-series. In future work we would like to use these methods on real-world data with a thorough analysis so that possible inconsistencies in the prediction of both methods can be described and quantified.

Acknowledgements

This work has been supported by the grants: MSM 6215648904/03 – Research design of MUAF in Brno; MSM 0021630529 – Research design of BUT Brno, Research design of MUAF Brno numbers 116/2102

/IG180651 and 116/2102/IG190611; No 102/07/1503 Advanced Optimisation of Communications Systems Design by Means of Neural Networks GACR.

References

- [1] Bishop, C. M., 1991. Improving the generalization properties of radial basis function neural networks. *Neural Computation* 3, pp. 579–588.
- [2] Mitchell, T. M., 1997. *Machine Learning*. 1st edition. McGraw-Hill. ISBN 0-07-042807-7.
- [3] Novak, M., 1998. *Umělé neuronové sítě. Teorie a aplikace*. C. H. Beck, Praha.
- [4] O’Neill, M., Dempsey, I., Brabazon, A., Ryan, C. 2003. Analysis of a Digit Concatenation Approach to Constant Creation In proceedings of the European Conference on Genetic Programming, (EuroGP), Essex, UK, 2003. p. 173-182. ISBN 3-540-00971-X.
- [5] O’Neill, M., Ryan, C. 2003. Grammatical Evolution: Evolutionary Automatic Programming. In an Arbitrary Language, Kluwer, 2003, 160 p, ISBN 1-4020-7444-1.
- [6] Popelka, O. 2007. Two-level Optimization using Parallel Grammatical Evolution and Differential Evolution, Proceedings of MENDEL’2007, Prague, CR, pp. 88-92. 2007. ISBN 978-80-214-3473-8.
- [7] Price, K., 1996. Differential evolution: a fast and simple numerical optimizer. In proceedings of 1996 Biennial Conference of the North American Fuzzy Information Processing Society, NAFIPS, pp. 524-527, IEEE Press, New York, 1996. ISBN:0-7803-3225-3.
- [8] Price, K. V., Storn, R. M., Lampinen, J. A. 2005. *Differential evolution—a practical approach to Global Optimization*. Springer, 543 p., ISBN 3-540-20950-6.
- [9] Ripley, B. D., 1996. *Pattern Recognition and Neural Networks*. Cambridge University Press, Cambridge.
- [10] Sarle, W. S., 1994. *Neural Networks and Statistical Models*. Proceedings of the Nineteenth Annual SAS Users Group International Conference, Cary, NC: SAS Institute, pp 1538–1550.
- [11] Stastny, J., Skorpil, V., 2005. Neural Networks Learning Methods Comparison. *International Journal WSEAS Transactions on Circuits and Systems*, Issue 4, Volume 4, pp. 325–330, ISSN 1109-2734.
- [12] Stastny, J., Skorpil, V., 2007. Genetic Algorithm and Neural Network. *WSEAS Applied Informatics & Communications*, pp. 347–351, ISBN 978-960-8457-96-6, ISSN 1790-5117.
- [13] Tseng, F. M., Yu, H. C., Tzeng, G. H., 2002. Combining neural network model with seasonal time series ARIMA model. *Technological Forecasting and Social Change*, 69, 71–87.
- [14] Wedding D. K., Cios K. J., 1996. Time series forecasting by combining RBF networks, certainty factors, and the Box-Jenkins model. *Neurocomputing*, vol. 10. pp. 149–168. DOI:10.1016/0925-2312(95)00021-6.

MACHINE LEARNING FOR OBJECT ESTIMATION USING HIERARCHICAL CRITERIA SYSTEM

Andrey Styskin

Department of System Analysis

Research Nuclear University Moscow Engineering Physics Institute

115409, Kashirskoe sh. 31, Moscow, Russia

Tel: +7 926 8762705

e-mail: astyskin@gmail.com

ABSTRACT

This paper presents machine learning method to determine parameters of hierarchical criteria system for object estimation based on learning set. The hierarchical criteria model offers knowledge domain in native hierarchical view and therefore is clear to interpret. Machine learning algorithm is iterative. On each iteration it goes through hierarchy using breadth first algorithm and optimize each characteristic independently. This algorithm gives good approximation and works well for small learning set problem.

1 INTRODUCTION

In management, science and commercial field we are usually faced with multi-criteria optimization or choosing best object problems. Multi-criteria decision-analysis (MCDA) is used in such problems. But when there are many criteria the standard MCDA-methods becomes hard to use and interpret results and in this case hierarchical methods are used. In hierarchical methods object criteria are structured in tree/network form and aggregation functions are used to aggregate single criterion to more general ones. When the problem of building hierarchical models becomes regular it is good idea to machine learn the parameters of model: weights and parameters of aggregation functions. We can use expert values of objects or any other kind of ranking function output estimation for learning set objects. The resulting optimized model is well interpreted so a researcher could check the validness himself by comparing the parameter estimations with own fillings about the knowledge domain. And this approach works well for small learning set.

2 HIERARCHICAL MODEL

2.1 Model definition

While building hierarchical model[1] it is required to group and structure measures in form of tree (as we do while the Analytic Hierarchy Process[7]). Lives of such tree are called single criteria. The rest vertexes are called complex criteria.

And the root is called integral criterion. Single and complex criteria differ in units in principle. Single criteria are measured in dimension units and complexes are measured in dimensionless units on interval $[0; 1]$. The values close to 0 indicate the small utility and in opposite values close to 1 indicate big utility. Hierarchical model present both the set of factors for knowledge domain (in form of single and complex criteria) and logical links between them. On each level there are criteria of the same generality level. Figure 1 illustrates automobile pricing problem tree:

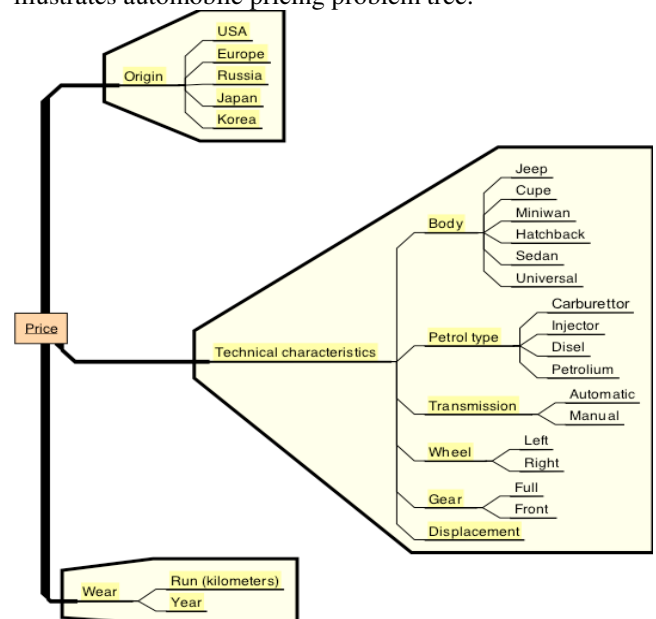


Figure 1: Hierarchical criteria system for automobile pricing problem.

To get estimations of objects with hierarchical model it is also required to define functional links between criteria (where logical links are). So the hierarchical model is:

- Knowledge domain organized in tree view;
- Conversion functions for dimension units that map measures to $[0; 1]$ scale;
- Aggregation functions that aggregate several logical linked measures to more general measure

2.2 Aggregation functions

The aggregation functions are described in [2] in details. To be recurrent the aggregation function for hierarchical model are to meet these properties:

1. Input variables u_1, u_2, \dots, u_m are defined on $[0,1]$ interval. The result of aggregation is also defined on $[0, 1]$;
2. Aggregation function is continuous and monotonous on u_1, u_2, \dots, u_m ;
3. Aggregation function has parameters V_1, \dots, V_m (that mean children criteria weights)

$$\sum_{j=1}^m V_j = 1;$$

4. Edge condition:
 $h(1,1,\dots,1)=1$;
 $h(u_1,1,\dots,1)<1$ while $u_1<1$;
 $h(u_1,0,\dots,0)>0$ while $u_1>0$;

The simplest and common used aggregation function is weighted sum:

$$h(u_1, \dots, u_m) = \sum_{j=1}^m V_j u_j$$

There are 5 distinct classes of aggregate functions:

1. Additive function

$$\text{Mathematically: } h(u_1, \dots, u_m) = \sum_{j=1}^m V_j u_j$$

2. Conjunctive – $h(u_1, \dots, u_m) = 1$, when all $u_i=1$
3. Disjunctive - $h(u_1, \dots, u_m) = 0$, when all $u_i=0$
4. Quasi-conjunctive – means $h(u_1, \dots, u_m)$ is close to 1 only when all u_i are close to 1 and other hand $h(u_1, \dots, u_m)$ is close to 0 when at least one u_i is close to 0.

$$\text{Mathematically: } h(u_1, \dots, u_m) < \sum_{j=1}^m V_j u_j$$

5. Quasi-disjunctive – means $h(u_1, \dots, u_m)$ is close to 0 only when all u_i are close to 0 and other hand $h(u_1, \dots, u_m)$ is close to 1 when at least one u_i is close to 1

$$\text{Mathematically: } h(u_1, \dots, u_m) > \sum_{j=1}^m V_j u_j$$

The problem of constructing aggregate function is described in [3] in details.

For example power function:

$$h(u_1, \dots, u_m) = \left[\sum_{j=1}^m V_j u_j^t \right]^p \text{ where}$$

$$t = \frac{\ln(\lambda_\varphi)}{\ln(1 - \lambda_\varphi)} \text{ and } p = \frac{\ln(\lambda_\psi)}{\ln(1 - \lambda_\psi)}$$

is quasi-conjunctive $\lambda_\psi < 0.5$ and $\lambda_\varphi > \lambda_\psi$ and is quasi-disjunctive $\lambda_\psi > 0.5$ and $\lambda_\varphi < \lambda_\psi$

The parameters of such aggregation functions are easy to interpret and are close to logic “and” and “or” operators that are used to for describing priorities of objects.

3 MACHINE LEARNING

3.1 Problem definition

We have hierarchical model with logically linked criteria, conversion functions for single criteria and output learning set as input. We need to give such aggregation functions that gives the best estimation of output on testing set (we will mathematically define what we suppose to be the best estimation). We will search aggregation functions in form that was introduced in 2.2

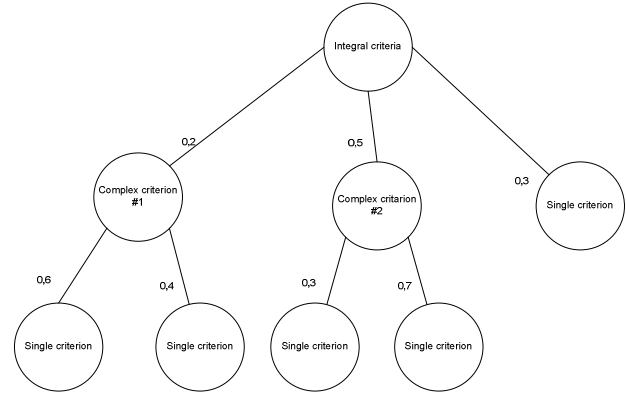


Figure 2: Hierarchical model.

Input:

Objective function:

$$y = f(\vec{x}) = f(\vec{x}, w_{ij}, \lambda_{ij}), \text{ where } w_{ij} -$$

weights of j-th child of i-th complex criterion and λ_{ij} - j-th parameter of i-th complex criterion

Region of acceptability for parameters:

$$\begin{cases} \forall i, j \Rightarrow w_{ij} \in (0;1) \\ \forall i \Rightarrow \sum_j w_{ij} = 1 \\ \forall i, j \Rightarrow \lambda_{ij} \in (0;1) \end{cases}$$

Starting points:

$$w_{ij} = w_{ij}^0$$

$$\lambda_{ij} = \lambda_{ij}^0$$

Learning set:

$$y_i^* = \bar{x}_i^*, i \in [0; n], y_i \in [0;1], x_{il}^* \in [0;1]$$

It is needed to choose such parameters w_{ij} and λ_{ij} that

$$SquareSum = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i^*)^2}{n}} \longrightarrow \min, \text{ where}$$

$$\hat{y}_i = f(\bar{x}_i^*, w_{ij}, \lambda_{ij})$$

3.3 Algorithm

The iterative algorithm on each stage going through each criterion and performs the multidimensional optimization for it. Parameters of other criteria are fixed so dimensionality of optimization is the number of parameters of this particular criterion. Variables are weights and aggregation function parameters.

For particular criteria optimization the algorithm similar to Hook-Jeeves [3] is used:

1. Take initial base point b_1 (current variables values) and step h_1 (values about 10^{-3} works well) for each variable $x_j, j = 1, 2, \dots, n$. Go to the step 2 with $i=1, j = 1$;
2. We calculate $f(b_i+h_j e_j)$, where e_j is a unit vector for x_j .

In case:

$$f(b_i+h_j e_j) < f(b_i) \Rightarrow b_i = b_i+h_j e_j - \text{new base point (we solve minimization problem)}$$

Else

$$f(b_i-h_j e_j) < f(b_i) \Rightarrow b_i = b_i-h_j e_j - \text{new base point.}$$

Otherwise x_j stays as it was.

If $j < n$ then the step 2 is repeated with $j = j + 1$

Otherwise we get new base point b_{i+1} and go to the step 3;

3. If after previous step $b_{i+1} = b_i$, which means that no objective decrease was done, then

$$\text{Set up } \forall i, h_i = \frac{h_i}{2} \text{ and go to the step 2;}$$

Otherwise $b_{i+1} \neq b_i$, and b_{i+1} becomes new base point.

Go to the step 4 with $i=i+1$;

4. The pattern search step. It uses the fact, that the way $(b_i - b_{i-1})$ already gave us objective decrease, so we should move from base point b_i this way. Calculate the value:

$$b_{i+1} = b_i + 2 * (b_i - b_{i-1})$$

In case:

$$f(b_{i+1}) < f(b_i) \Rightarrow \text{repeat the step 4 again with } i=i+1$$

Otherwise:

$$\text{Go to the step 2 with } i = i, j = 1$$

Note: Stop the process after several steps (5-10) to prevent over optimize current particular complex criteria.

Quite good model can be obtained after 50-100 iterations for whole tree.

Breadth first way is used to ensure uniform optimization. It was founded that the order on one level is insignificant.

3.4 Non unimodality problem

It is easy to show that SquareSum function is non unimodal. So optimization algorithm can find only local extremum.

To struggle with this property the idea of crossover (described in [5]-[6]) from genetic algorithm can be used.

We can take two trees from population, split them in the same criteria and generate two more trees made by concatenation of left side from one tree and right side of the other. The new trees should be added to population and optimized by algorithm.

Figure 3 illustrates details of splitting and creating new trees:

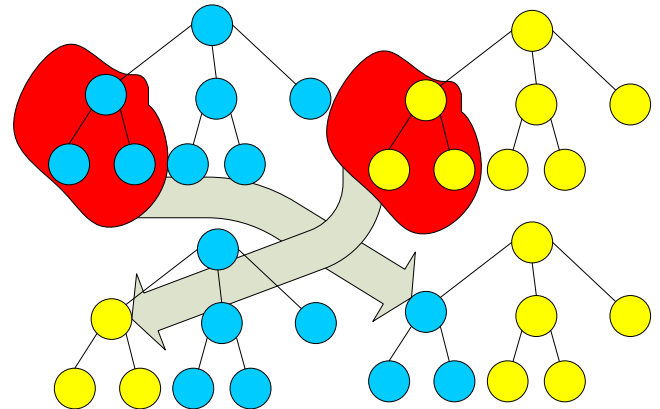


Figure 3: Crossover for two trees.

4 RESULTS

Application that helps building models and optimize their parameters was created.

The optimization is usually takes 100-200 iterations and several minutes for thousands examples with tens criteria in learning set.

The algorithm was tested on second-hand automobile pricing problem. There were about 1000 second-hand cars learning set with 9 criteria: year, run, engine, transmission, wheel, gear, petrol type, body, origin.

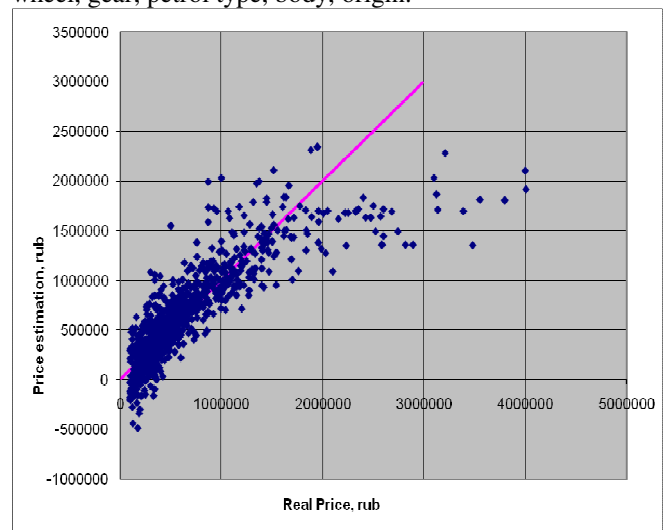


Figure 4: Estimated price to real price correlation graph for classical regression



Figure 5: Estimated price to real price correlation graph for our algorithm

We compare our algorithm with classic polynomial regression (made in Weka[4]). The sum of squared residuals is compared.

There are estimated price to real price correlation graphs for classic polynomial regression on figure 5 and our method on figure 6.

	Classic regression	Our method
Mean square error	54%	33%

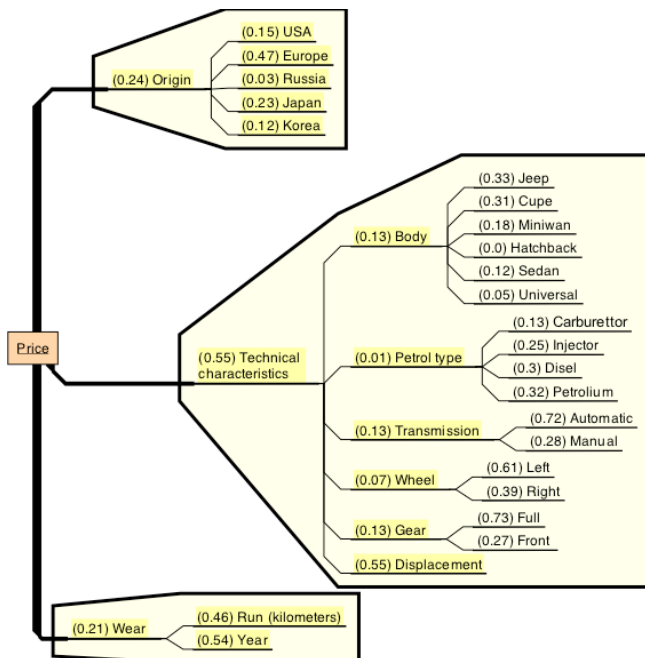


Figure 6: Hierarchical criteria system for automobile pricing problem with parameter estimates.

Figure 6 illustrates the hierarchical model for automobile pricing problem and its parameters after optimization with our algorithm. The weight of particular criteria are reasonable.

5 CONCLUSIONS

In this paper a new method for solving multi-criteria optimization or choosing best object problems was introduced. In comparison with classic data mining algorithms that are either black box which gives good predictions without the way to understand how and what criteria gives the best contribution or number of visualization and asking techniques which cannot be verified, our approach gives easy way both to describe knowledge domain for such problems and to score objects. The results of second-hand automobile pricing problem indicate that the approach is good for scoring objects.

References

- [1] Eltareno E.A., *Scoring and Choice by Multiple Criteria*, MEPhI, Moscow, Russia, 2005 (in Russian)
- [2] Eltareno E.A., *Aggregation Functions*, MEPhI, Moscow, Russia, 2006 (in Russian)
- [3] Hook, R. and Jeeves, T., "Direct Search Solution of Numerical and Statical Problems", J. Assoc. Comput, 1961
- [4] I. H. Witten, E. Frank (2005) "Data Mining: Practical machine learning tools and techniques," 2. edition, Morgan Kaufmann, San Francisco, 2005
- [5] G. Sywerda, "Uniform crossover in genetic algorithms," Proceedings of the 3rd International Conference on Genetic Algorithms, USA, 1989.
- [6] J. Page, P. Poli, and W.B. Langdon, "Smooth uniform crossover with smooth point mutation in genetic programming: A preliminary study," *Genetic Programming*, Proceedings of EuroGP'99, Sweden, 1999.
- [7] Saaty, T.L., "Decision making with the analytic hierarchy process", *Int. J. Services, Sciences*, Vol. 1, No. 1, pp.83-98., 2008

COLLECTIVE INTELLIGENCE AND ORGANIZATIONS' CONSCIOUSNESS

Viljem Tisnikar

Jozef Stefan International Postgraduate School
Jamova 39, 1000 Ljubljana, Slovenia
Tel: +386 41 690619; fax: +386 41 713 118
e-mail: viljem.tisnikar@gmail.com

ABSTRACT

Collective Intelligence is a new paradigm, stemming from increased and novel use of Information and communication technology (ICT) in various human organizations. New approaches are possible for solving traditional hard problems, such as budgeting in organizations. One can learn from automatic plans generation methods from Artificial Intelligence and apply that knowledge within human organizations. We advocate the use of unifying language of mathematical Category Theory in order to compare models from such different fields of research.

1 INTRODUCTION

The key research question of MIT Centre for Collective Intelligence [8] is: "How can people and computers be connected so that — collectively — they act more intelligently than any individuals, groups, or computers have ever done before?" We will refer to any collection of people, working on some common goal using ICT as an Organization. Consciousness is an umbrella term that may refer to a variety of mental phenomena, but for the purpose of this paper we propose the following working definition: "Organization's Consciousness is the ability to be aware of the external environment and the presence of the organization within it". One of the potential advantages of Collective Intelligence is then an increased awareness of more elements of organization's environment and a wider range of potential options for how organizations interact with it. "One of the areas that most scientists and thinkers struggle with in the field of consciousness is explaining how it is created. At this point the best we can figure is that it is an emergent phenomena that arises from the interaction of the comparatively simple elements that underlie it" [9].

In this paper we search for the nature of "comparatively simple elements" that underlie an organization's consciousness and how these elements are structured. In order to describe these elements and their interaction we need some formal language. First, we show analogies between a real life hard problem in organizations – budgeting (Section 2), and some artificial intelligent (AI) formal algorithms (Section 3). In order to compare such different fields we advocate in Section 4 that mathematical

Category Theory language should be used. In Section 5, we shortly present lines of our current research and we conclude in Section 6.

2 HARD PROBLEM: BUDGETING IN ORGANIZATIONS

The current world-wide financial crisis has brought a new emphasis and requirement for intelligent financial planning within various organizations, like businesses or public services. There is no clear and from all agreed upon definition for the term "intelligent" [5]. Nevertheless, in case of efficient financial planning, continuous adaptation to (chaotic) changing environment is needed in order to achieve organization's financial goal, and "goal-directed adaptive behavior" is one of possible definitions of intelligence [11]. From financial perspective the overall goal (for a goal-directed behavior) of an organization is quite simple: cash inflow has to be greater or equal than cash outflow in the same time intervals. On the other hand, the *adaptive* behavior requires continuous recording of the relevant organization's environment and adapting plans accordingly. World's top provider for business software SAP refers to the budgeting in large organizations as "The budgeting nightmare" [12].

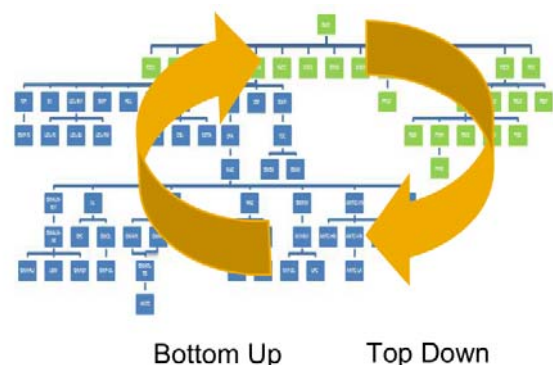


Figure 1: Budgeting in large organizations (from [12])

Budgeting, as depicted in Figure 1, combines the bottom-up synthesis of new information from recorded facts and the top-down analysis of goals into subgoals. There are many perspectives, and each of them can be observed on

different levels of aggregation. The most basic perspectives, inferring hierarchy, are:

- *time intervals*: from several years long strategic level to operational level on a daily basis,
- *places*: from whole world to some specific places, where an organization operates,
- *fields of operation*: from diversified to niched, and
- *financial and accounting* perspective.

Beside these, many additional perspectives can take place. These perspectives and their different levels together make an entangled structure especially due to the fact, that everything is dynamic. Not only yesterday's goals for today have to become facts today, but they also have to be compatible with (derived from) strategic plan defined several years ago. This is not an easy task, so humans in organizations mostly evade such systematic planning and plans follow-through. They use human intelligence and creativity instead, in order to overcome problems due to the poor planning system and consequently the lack of coordination. Although interesting per se, such improvisations bear additional and needless costs of operations. In the field of Artificial Intelligence (AI) however, "automatic and intelligent plans generation" is one of most important open questions. In Section 3, we highlight some AI approaches.

3 AI ALGORITHMS USING TOP-DOWN AND BOTTOM-UP LOGICAL INFERENCES

Since mathematical and logical modeling of human intelligence has proved in the past to be a hard problem [5], AI research is mostly dealing with specific problem-solving algorithms. From these, we shortly present two of them, emphasizing the usage of top-down and bottom-up logical inferences within the same algorithm. Generally, the top-down inference is usually referred to as planning and bottom-up as recalling the past.

3.1 HINT (Hierarchy INduction Tool)

AI Lab from the Faculty of Computer and Information Science Ljubljana developed a computer program HINT, which decomposes the problem into smaller, less complex problems [13]. It is not a typical planning of an artificial agent, but it deals with building of a hierarchical model of a part of the world. The model is then used to classify new data and the classification can be viewed as actions – so the model can also be viewed as a hierarchical plan of actions. A problem in HINT is represented with mathematical *function* and the method for *function decomposition* is presented. To cope with high time complexity of finding an optimal decomposition, authors propose a suboptimal heuristic algorithm. Formal techniques for decompositions of the problem into less complex subproblems can be found also within decision support science/systems (DSS), specifically multi-attribute decomposition method [2]. Researchers compared HINT on one side and hierarchical models derived by human experts with DSS methods on the

other side. The experimental work showed that HINT could discover useful and interpretable concepts. However, authors strongly believe, that such decomposition by computer should not be used alone in the completely automatic fashion, but should rather be incorporated within *interactive* data mining systems to propose new concepts that are then reviewed and interpreted by experts. The authors' believe is compatible with requirements for human intelligence. It is known that within human brains new sensorial data continuously enter and influence "the operating algorithm". The paradigm of an algorithm running in isolation, without continuous sensorial input, is not appropriate for intelligent agents modeling [5,6]. Since the holistic modeling of human intelligence is extremely hard, approaches that combine some computer algorithms and organized interventions from human experts, are valuable. They lead to better understanding of human intelligence.

3.2 Hierarchical Temporal Memory

A kind of more ambitious AI project is presented next. Hierarchical Temporal Memory (HTM) is the technology, which tries to mimic the common algorithm of a human neocortex. Vision, hearing, touch, language, behavior, and most everything else the neocortex does are manifestations of a single algorithm applied to different modalities of sensory input [7].

The main idea is that HTMs, as human brains, automatically derive characteristics of sensed environment through sequences of sensorial records [3]. HTMs are algorithms, a general platform, intended to discover causes in the application specific world. After learning through application specific sensors they can infer causes of novel input and possibly enable making predictions or directing behavior.

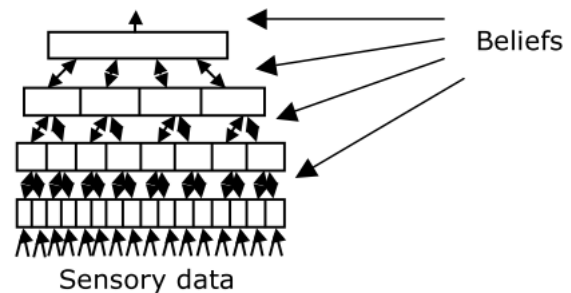


Figure 2: A HTM hierarchy (from [7])

HTMs are dealing with (inverse) trees, where sensory data enter at the bottom and is then aggregated through diverse layers. Through aggregation HTMs calculate and memorize beliefs, similar to Bayesian Networks. Exiting the top is a vector where each element of the vector represents a potential cause of the sensory data. These causes can then be propagated backward (down the inverse tree), affecting new sensory data with pre-existent beliefs. A simple schema is depicted in Figure 2.

HTMs are used for computer vision problems, basically for classifying objects. Authors advocate the potential for application of HTMs for resolving various other hard problems. Modeling such diverse domains requires a universal and powerful enough language. Next, we propose one such language.

4 WHY CATEGORY THEORY?

Mathematical Category Theory (CT) has come to occupy a central position in contemporary mathematics and theoretical computer science, and it can be applied to the study of logical systems [10]. CT is still evolving, its functions are correspondingly developing, expanding and multiplying. At minimum, today CT is a *powerful language*, or conceptual framework, allowing us to see the universal components of a family of structures of a given kind, and how structures of different kinds are interrelated. Categories are algebraic structures with many complementary natures, e.g., geometric, logical, computational, and combinatorial.

Objects play a secondary role in CT and could be entirely omitted from the definition of a category. Since every object in CT has an identity arrow (like identity function in the category of Sets and functions), an arrows-only definition of a category simply recognizes that some arrows have equal domain and codomain. But it is common practice to call such arrows rather objects. However, the arrows (and their composition) are what is important in categories. Identity arrows and (associative) composition of arrows are also the only requirements for various style definitions of a category. Categories were defined in 1945 for reasons of rigor. They were not needed as such, but something like categories had to be defined so that the notions of functors and natural transformations - what the inventors of CT were really interested in and using - could be given an explicit and rigorous presentation. Functors are mappings between categories and they preserve identities and composition of arrows from source category - they *preserve dynamics*. Ordinary set functions can be viewed then as only a degenerate version of a functor - preserving only objects (defined on all objects of domain set), and all possible structure has to be defined in addition with functions. Categorical language can be distinguished from set-theoretic as dynamic and local as opposed to static and absolute.

Almost every known example of a mathematical structure with the appropriate structure-preserving map yields a category. CT treats the notion of structure in a uniform manner and unifies mathematical structures in two different ways. First, a unification provided within a set theoretical environment. Second, and perhaps even more important, once a type of structure has been defined, it is imperative to determine how new structures can be constructed out of the given one. It is also imperative to determine how given structures can be decomposed into more elementary substructures. In both cases, it is necessary to know how structures of a certain kind may combine. The nature of these combinations might appear to be considerably

different when looked at from a purely set theoretical perspective [10].

Category theory also reveals how different kinds of structures are related to one another. These relations can be represented with adjoint functors - a concept of fundamental logical and mathematical importance that is not captured elsewhere in mathematics [1]. Functors can also have the role of objects in new - so called - functor categories. In such a case, natural transformations, a third type of morphisms within CT, are arrows in functor categories.

It is also legitimate to think of a category as an algebraic encoding of a deductive system. This phenomenon is already well-known to logicians, but probably not to its fullest extent [10].

What do "budgeting nightmare", HINT, HTMs, and mathematical Category Theory have in common? They all deal with structures and CT language is the most appropriate language for describing and manipulating with structures and moreover it is formal, exact, and coherent.

5 TOWARDS A CATEGORICAL MODEL OF A UNIVERSAL (INTELLIGENT) AUTONOMOUS AGENT

Using CT language, we envision a model of a universal (intelligent) autonomous agent. Such an agent must have some sensors, the memory, and some actuators. In the memory a model of a relevant environment is stored and actions of the agent are directed upon this model. Like with HTMs and opposed to prevailing practice, the agent can only learn a model of the relevant world through its own sensors. The only role of a designer is to provide that sensorial data are properly summed together - with categorical sum. No intervention of the designer should take place, because the designer does not speak the language of agent's sensors. Like the basic idea with HTMs, temporal sequences of sensorial input play a crucial role for building an agent's hierarchical model. Any such sequence can be viewed as a category. With simple doubling of a sequence and connecting objects of the two sequences (categories) one obtains a new category, a so-called comma-category, which can be interpreted also as a hierarchy (lattice). Doubling of a (double) sequence is a known and important biological mechanism - the DNA replication.

We claim that goals and their decompositions are not simply the hierarchy of perceptions turned around. Goals or signals from memory to actuators are conceptually different than perceptions, even though the hierarchies may resemble each other's opposite. Only recently, the paper abstractly dealing with these subtleties in a systematic manner, was published [4]. Intuitively, maybe this is most clear from "budgeting nightmare" example. Plans are often required to be presented in the same form as reports in order to enable the comparison of the two. But plans "live" in a completely different category than historical data. While historical data

is fixed and certain (corresponding to some mathematical *objects*), plans are uncertain and variable (corresponding to some mathematical *arrows*). Abstractly, the relation between historical data in an agent's memory and plans in the same agent's memory correspond to relation between some sets and some groups (or monoids). Such a relation is the paradigmatic example of adjoint functors in CT.

From these relations between memorized perceptions and from memories generated plans it is possible to build a functor category according to exact mathematical rules. Such a functor category becomes a rich structure even with relatively small number of sensors and actuators and in such a rich structure consciousness (a skeleton subcategory) and intelligence could be explored.

This research was motivated several years ago with an unexpected success of a home-made software of a type known today as "computer supported cooperative work". Basically it was focused on continuous planning and plans follow through in a small business. It was an example of collective intelligence because the organization gained a level of productivity and robustness that was not possible without such software. On this basis we work also on a design for web application software, focused first on cash-flow continuous planning and plan follow through. Such a software is particularly important for start-up businesses, working in turbulent environment. It seems however, that the environment is becoming turbulent also for most other existing organizations.

6 CONCLUSION

With increasing use of ICT in organizations, new terms and questions arise, like Collective Intelligence and Organizations' consciousness. Budgeting or systematic, continuous planning, plans adaptation according to (chaotic) changes in the relevant environment, and plans follow through are traditionally hard problems in organizations. AI research is forced to cope with similar problems formally using mathematical and logical modeling. Learning from diverse experiments with formal problem-solving algorithms, unifying these approaches with use of an appropriate language may lead towards a universal model of an artificial intelligent agent. We presented two such algorithms, namely HINT and HTMs, and advocated the use of unifying language of Category Theory for modeling.

References

- [1] S. Awodey. *Category Theory*. Oxford university Press, 2006.
- [2] M. Bohanec. *Decision making and models* (in slovene). DMFA-založništvo, Ljubljana, 2006.
- [3] G. Dileep. How the brain might work: A hierarchical and temporal model for learning and recognition. *Phd thesis. Department of electrical engineering, Stanford university*. Accessed September 2009 at: <http://www.numenta.com/for-developers/education/DileepThesis.pdf>. 2008.
- [4] D. Ellerman. *The Logic of Partitions: Introduction to the Dual of the Logic of Subsets*. Accessed September 2009 at: <http://www.ellerman.org/Davids-Stuff/Maths/Logic-of-partitions.pdf>. 2009.
- [5] M. Gams. *Weak intelligence: Through the principle and paradox of multiple knowledge*, volume 6 of *Advances in computation: Theory and practice*. Nova science publishers, inc., NY, 2001.
- [6] D. Goldin, P. Wegner. Refuting the strong church-turing thesis: the interactive nature of computing. *Minds and Machines*, 18:1:17–38, March 2008.
- [7] J. Hawkins, D. George. Hierarchical Temporal Memory: Concepts, Theory, and Terminology. *Numenta Inc*. Accessed September 2009 at: http://www.numenta.com/Numenta_HTM_Concepts.pdf. 2007.
- [8] T.W. Malone et al. MIT center for collective intelligence. *Web site*, 2009. URL <http://cci.mit.edu/index.html>. 2009.
- [9] T.W. Malone, R. Laubacher, C.N. Dellarocas. Harnessing crowds: Mapping the genome of collective intelligence. *MIT Sloan Research Paper* No. 4732-09. Available at SSRN: <http://ssrn.com/abstract=1381502>, February 2009.
- [10] J.P. Marquis. Category theory. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/archives/spr2009/entries/category-theory/>, spring 2009 edition, 2009.
- [11] R.J. Sternberg, W. Salter. *Handbook of human intelligence*. Cambridge, UK: Cambridge University Press. 1982.
- [12] L. Yue. SAP BusinessObjects™ Planning & Consolidation: Effectively Manage Performance By Streamlining The Group Budgeting & Consolidation Process. *June 11, 2009 Hong Kong presentation*. Accessed in September 2009 at <http://www36.sap.com/hk/about/events/worldtour09/pdf/T2.4bBridgingStrategyandExecution.pdf>. 2009.
- [13] B. Zupan, M. Bohanec, J. Demsar, I. Bratko. Learning by discovering concept hierarchies. *Artificial Intelligence*, vol. 109, pages 211-242. 1999.

APPLYING DATA ENVELOPMENT ANALYSIS FOR INCREASING OPERATIONAL EFFICIENCY IN PROJECT MANAGEMENT

Pavel Tubin

Department of Economics and Management
Moscow Humanities-Technological Academy (MGTA)
1-st Kotlyakovski per., 1, Russia, Moscow
Tel: +7 926 3380506
e-mail: pavel.tubin@gmail.com

ABSTRACT

A method for identifying managerial effectiveness and efficiency frontiers of project activities and corresponding tools are proposed. The novelty of the method is the ability to dynamically assess managerial efficiency and to receive recommendations for applying adjustments to control actions. The scientific contribution of this work is to propose a new method of performance management and to expand a research field of projects implementation processes.

1. INTRODUCTION

Projects are means of implementing strategy and leveraging performance of organizations [1]. In general, organization performance is evaluated on two basises: means (flexibility, efficiency, etc.) and results (competitiveness, financial performance) [2]. Accordingly, project performance can be viewed from two perspectives: (1) examine the relationship between output and cost (efficiency) and (2) determining the accuracy of the result achieved on the objectives (effectiveness). During the global crisis, when performance of many organizations is unsatisfactory, the focus has shifted to structure optimization and processes reengineering within the organization, i.e. on performance of implementation. However, in practice the impact of many projects is unsatisfactory both in terms of outcome and in terms of process efficiency. Moreover, there is an apparent lack of research of strategic projects process, as well as qualitative metrics to evaluate the efficiency of this phenomenon. Accordingly, development of metrics to assess maximum and in-progress performance of managerial activities as well as methods for increasing efficiency of project implementation are the issues of high relevance nowadays. This paper aims to improve organizations' projects efficiency through assessment of managerial efficiency boundaries during different phases of project implementation, i.e. boundaries not associated with the resourcing of the project but directly with the quality of operational management of the project.

2. STRATEGIC PROCESS: PROJECTS AND EFFICIENCY

Strategic process refers to formulation and implementation activities of strategic objectives. Strategic objective is a desired state of the company from the perspective of key stakeholders. Strategic actions are actions aimed at achieving the strategic objectives, influencing the position of stakeholders. I.B. Gurkov describes the structure of the strategic process as follows [3]:

- Strategic objectives are influenced by external environment, including:
 - ownership structure and company management (corporate governance structures),
 - degree of competitive pressure (market structure).
- During the stage of goals formulation disposable strategic resources of the firm should be taken into account.
- Objectives get translated into certain strategic actions (plans, programs, projects).
- These actions may be attributed to:
 - corporate strategies (restructuring of firm's businesses) and / or
 - competitive strategies (choice of the business community in each market).
- The result of strategic actions implementation determines not only a separate business performance, and the company performance as a whole.
- The results of firm's strategic activities lead to an adjustment of the initial strategic objectives and, consequently, change the structure of strategic assets. In addition, the impact of these actions may lead to changes in market structure and, finally, in ownership structure of firms.

Performance is both input and output in strategic process and it is used not only as a basis for decision-making, but also as a basis for any organizational improvement. Many studies indicated that Performance Measurement Systems

(PMS) play an important role in the effective (in this case - as in the value of the ratio of costs to the result and the importance of achieving the planned goals) the organization management. Many performance measurement systems in terms of architecture are a set of metrics to assess the effectiveness and efficiency. ISO 9000:2000 defines effectiveness as a degree of achievement of planned results and efficiency as a ratio between results and resources invested.

Transition from strategy formulation to specific projects and their implementation unveils problems of two types:

- a problem of effectiveness, related to the achievement of planned results;
- a problem of efficiency, related with the project development process.

Area of interest of this study comprises efficiency problems that arise because of:

1. absence efficiency metrics of project management - the complexity of in-progress (current) monitoring;
2. lack of understanding of "standard" distribution of the project management activities and the associated efficiency frontier.

In practice economic, commercial (financial), budgetary, social, technical and functional efficiencies are marked out. Most of these concepts relates to the effectiveness field being related with the end results, with the important place occupied by cost components of inputs and outputs. Metrics of this type exist in virtually every company and are widely known. It is primarily on financial metrics of performance - from revenue and profits, and ending with EVA and CFROI. With these metrics result of a strategic project can be ex post estimated, as well as forecasted values of the result on the basis of the financial model of the project can be obtained. Metrics of technical efficiency is much less prevalent than financial performance; even rarer are the metrics that are applicable throughout the process of project implementation. In this case there is a metric based on the material resources that cannot gauge pure managerial efficiency of the project. Given a tool that allows evaluation of project activities in its development, opportunities to reduce risk and increase the likelihood of successfully achieving the strategic goal appear.

To maximize an impact of PMS, integration of such systems in the strategic process is essential; that will allow timely adjustments to the control actions on the system based on current measurements of efficiency and effectiveness and help to achieve better results through strategic projects. Accordingly, during the implementation of strategic projects daily routine activities is a key lever impact, affecting both the results themselves, and the trajectory of their achievements. As a part of this work, the in-progress metric of management efficiency is proposed based on the DEA methodology, allowing managers to adjust their impacts during project implementation.

3. DEA

Data envelopment analysis is a nonparametric method in operations research and economics for the estimation of production frontiers. It is used to measure productive efficiency of decision making units (or DMUs). Data Envelopment Analysis (DEA) methodology was reduced by Charnes, Cooper and Rhodes to a linear programming problem [4] and was used to assess the relative efficiency of different industries - from banking to police stations. Consider the mathematical formulation of Data Envelopment Analysis. For each object k ($k=1, \dots, n$) an optimization problem must be solved (total - n tasks). For each object the efficiency index θ should be maximized by variables v_i ($i=1, \dots, m$) and u_r ($r=1, \dots, s$):

$$\theta_k = \frac{u_1 y_{1k} + u_2 y_{2k} + \dots + u_s y_{sk}}{v_1 x_{1k} + v_2 x_{2k} + \dots + v_m x_{mk}} \rightarrow \max$$

$$\frac{u_1 y_{1j} + u_2 y_{2j} + \dots + u_s y_{sj}}{v_1 x_{1j} + v_2 x_{2j} + \dots + v_m x_{mj}} \leq 1 \quad (j = 1, \dots, n)$$

$$v_1, v_2, \dots, v_m \geq 0$$

$$u_1, u_2, \dots, u_s \geq 0$$

where $(x_{1j}, x_{2j}, \dots, x_{mj})$ and

$(y_{1j}, y_{2j}, \dots, y_{sj})$

vectors denoting the input and output.

One of the main issues for future studies on assessing efficiency reads as follows: how to develop a dynamic, not static efficiency measurement system? We argue that various modifications of the DEA model allow to dynamically evaluate efficiency and can be used to solve this problem. In this paper we propose to apply DEA methodology to evaluate project efficiency in different project phases (stages). This is an original approach to this tool, designed originally to determine the efficiency of several DMUs. Based on our study, we propose to assess efficiency of planning, organizing and monitoring activities on a project to build a frontier of management efficiency. Mathematical model used provides an ability to calculate the inefficiency coefficients of the project phases, which form the basis for recommendations for management to achieve the boundary (maximum) efficiency during further project implementation. Scientific novelty of this work comprises development of a system for assessing and improving the projects efficiency based on DEA methodology, which allows to dynamically allocate and assess the managerial component of efficiency, and make it possible to obtain recommendations for adjustments to control actions.

4. APPLYING DEA TO PROJECT MANAGEMENT IN-PROGRESS ASSESSMENT

Ability to implement strategy, not being focused on an individual and involving many members of organization in an interaction, at the lower level is expressed in a system of actions and routine activities of managers responsible

for strategic projects and their teams. To describe this kind of actions there are many different concepts of leadership. An important characteristic of a leadership style (a set of styles) is that this style, on the one hand, is a reflection of operations, managers' routine actions, and, accordingly, his skills in planning, organizing, motivating, and monitoring of staff, directly affecting this activity, combined with situational factors; on the other hand, is a part of organization's ability to implement strategy, affecting a project team and causing variation in the efficiency and effectiveness of strategic project. Regarding the types of activities should be noted that there is no single classification of managers' activities classes, in various studies authors resort to some sets. Nevertheless, there are a number of actions, one way or another occurring in the majority of researches, for example, documents, meetings, negotiations. Based on existing research we can conclude that manager's time distribution is the most important factor affecting the efficiency of the organization as a whole. In this case the existing literature often ignores the nature of managers' activity, thus becomes irrelevant to the majority, many recommendations that appear in this literature are not based on thorough research, are not well-thought and scientific.

Through a set of control actions applied to the manageable inputs of strategy implementation process (projects), management influences results of a project phases. Manager's time budget, i.e. manager's operations timing distribution, is directly linked to control actions. This distribution is affected by external factors, as well as preferred management style, and, accordingly, management cycle ("leader's window") [5]. Based on the research of management cycle for a group we can conclude that, depending on the maturity of a group will change the optimal management style and, accordingly, the distribution of control actions. So, depending on the application of control actions at the examined stage of a project its performance (efficiency) will vary.

Knowing manager's preferred style and group maturity level we are able to analyze phases which showed some inefficiency – was there a shift from manager's preferred style (or from a recommended one for this stage) to the stimulated?

From the perspective of strategic process, control actions and project parameters should be on input in the model, output parameter is the relative efficiency of the tested phase. From the standpoint of the methodology used, the center of the model is Data Envelopment Analysis, for which control actions will serve as inputs and project parameters will serve as outputs. In turn, the result of the methodology application is an efficiency assessment for each of the stages of a project examined and corresponding recommendations for future adjustments based on inefficient stages analysis.

5. MODEL APPROBATION

Testing of developed methodology was carried out in the company, which designs solutions for the banking sector. To test the methodology data was collected on a key project for the company, the strategic nature of which is not in doubt: non-compliance and failure of this project should be fatal for the company. First of all, we reviewed the strategy and past activities of the company to explore strategic projects in the company and key business processes. The model has been adapted for usage in the studied company in terms of input and output data. After processing and aggregation of data it has become possible to analyze the efficiency of control actions using the methodology of Data Envelopment Analysis. Based on the results of analysis that identified stages of effective and ineffective management, evaluated the deviation of efficiency (i.e. potential increase), control actions efficiency dynamics was determined. Measurements were conducted using both BCC and CCR DEA models. In the study for various periods of one and the same project, the difference between the maximum and minimum managerial efficiency, estimated using the model CCR DEA, amounted to 97,3%.

Identified through this analysis sources of inefficiency were correlated with the results of interviews with participants and the project manager.

After that a chain of events that affected efficiency of control actions in the project was reconstituted. Formed recommendations for control actions adjustments were based primarily on efficiency analysis of studied periods, the context and nature of manager's control actions, and address the discrepancy between the techniques and methods of leadership and a situation in the project.

6. CONCLUSION

Application of the proposed method makes it possible to assess the management efficiency during project implementation, with the accuracy of the estimate increases with the number of phases and activities of a project. It is important that the assessment is carried out for several dependent variables simultaneously. Based on the assessment it's possible not only to perform an ongoing correction of control actions, but to build a knowledge base that enables a comprehensive review and further managerial improvement in examined form of projects, which is especially important for IT companies, carrying out a large number of projects similar in structure to develop and implement software. The developed metric can evaluate activities efficiency and allows making adjustments without dividing technical efficiency and project effectiveness, which is achieved through the inclusion of qualitative characteristics of the result in the number of dependent variables in the model.

This work also makes contributions to the scientific research of strategic projects, demonstrating a level of

Pareto-inefficiency of management methods and, consequently, the potential for efficiency gains, and the role of project management as a tool for implementing strategy. In addition, this technique makes it possible to verify a number of theoretical and practical assumptions about project management made on the basis of abstract logical conclusions and expert assessments.

Breakthrough practical results can be obtained by automating procedures for collecting baseline data on projects that, as a positive side effect, lead to the increase of project manager personal efficiency, due to the applied software.

Further research in this field may be associated with improving the model after accumulation of some empirical data. Further practical work include intelligent system software development to fully automate the process of data collection and analysis, repository development for data storage and subsequent analysis to identify common dependencies, the integration of the proposed tools with enterprise project management and enterprise resource planning systems (ERP), i.e. development a system that will help managers to make high-quality decisions and implement projects more effectively.

7. REFERENCES

1. Olivier Furrer, Howard Thomas and Anna Goussevskaia, The structure and evolution of the strategic management field: A content analysis of 26 years of strategic management research - International Journal of Management Reviews, 2008; Volume 10 Issue 1; pp. 1–23.
2. Kit Fai Pun and Anthony Sydney White, A performance measurement paradigm for integrating strategy formulation: A review of systems and frameworks - International Journal of Management Reviews, 2005; Volume 7 Issue 1; pp. 49–71. 23.
3. I.B. Gurkov, Strategic Process in Russian Companies - ENSR, №2, 2009. In Russian.
4. Charnes A., Cooper W.W., Rhodes E., Measuring the efficiency of decision making units - European Journal of Operation Research, 1978; vol. 2: pp. 429–444. 64.
5. John D.W. Beck and Neil M. Yeager, The Leader's Window: Mastering the Four Styles of Leadership to Build High Performing Teams - New York, NY: John Wiley & Sons, 1994.

POTENTIAL BENEFITS OF USING WEB SERVICES IN CRM SYSTEMS

Jan Turčinek

Department of informatics, Faculty of economics
Mendel University in Brno
Zemědělská 1, 613 00 Brno, Czech Republic
Tel: +420 545 132 23
e-mail: turcinek@node.mendelu.cz

ABSTRACT

For better business planning of companies it is important to know utmost available information. This information is handled by CRM systems. CRM as such requires a great number of activities, most of them are nowadays processed with use of IT and information systems. These technologies enable automatic processing of most of CRM processes. SOA enables to build the system tailored to needs of companies. The paper provides an example of web service application in calculation of customer value in order to emphasize the advantages of web services use. As an example we could describe a service of calculation of customer value.

1 INTRODUCTION

When any enterprise wants to reach continual profit, it has to provide continuing sale of its products. For maintenance of sale it is important to hold its current customers and it also is important to find new potential customers. These tasks are substance of branch of management called Customer relationship management (CRM). CRM include a wide scale of communication, data processing and other activities.

The scale and way how these systems are realized depend on requirements of each enterprise. CRM systems are available in many different forms of applications, such as:

- desktop applications,
- part of enterprise information system,
- CRM hosting, etc.

Mostly it means an enclosed package of processes.

During the evolution of CRM systems have many available technologies changed. Older system offered only changeless functionality. Future of CRM systems will be in area of intelligent systems. These systems will be able to adapt their functionality.

Service Oriented Architecture brings a new view to developing of CRM systems.

2 CRM PROCESSES

Customer Relationship Management represents an instrument for formation and administration of longstanding and profitable relationships with customers. It also serves making new profitable relationships. It helps understanding of habits, wishes and needs of customers on base of storage of information of their behavior. It is, thus, a knowledge system and it includes processes, such as

- Data processing
- Use of the data.

The three basic elements of CRM are people, processes and technologies. They are closely interconnected and they are often supplemented with fourth one, formed by following features – people, commercial use of data, technologies and content (Wesling, 2002).

CRM system includes above mentioned elements and the system can be defined as:

- System for management of the entire contact cycle with customers,
- System supporting efficient coordination of linkages to customers,
- System supporting customer care.

These systems are typically solved with information technologies. Wesling (2002) states following benefits of proper use of such system:

- Trouble free process of business processes,
- More individual contacts with customers,
- More time for a customer,
- Differentiation from competition,
- Consolidation of corporate image,
- Access to information in real time,
- Reliable and fast predictions,
- Communication among marketing, sales and services,
- Increase of teamwork efficiency,
- Growth of employee motivation.

In order to achieve these objectives, it is necessary to adjust methodology appropriately to the needs of particular CRM solution and choice of CRM system. The choice depends not only on current company need, but also its economic and personnel conditions. Present market offers number of different solutions based on diverse principles, such as:

- Local application,
- Modules cooperating with office software,
- Online solutions,
- Module of corporate information system.

A new trend today is a use of online solutions based on principles of outsourcing, when a small firm doesn't to own any special resources and it leases particular applications on provider's server.

CRM process set can be decomposed into individual processes, which can then serve for design of services for final information system. CRM as such integrates four primary processes:

- Contact management (uses so called contact center),
- Sales management (for automation of the process there is a SFA functionality – Sales Force Automation, including E-Business),
- Marketing management (uses technologies EMA – Enterprise Marketing Automation) and
- Customer services (uses functionality CSS – customer service and support)

This principle has a lot in common with basic principles of service oriented architecture, which are discussed next.

Service oriented architecture (SOA) belongs to new principles used in corporate systems or in individual applications. It is not pure software architecture. Gala (2006) defines service oriented architecture as a form of technological architecture based on principle of services.

Web Services and Service-Oriented Architecture defines SOA as a set of services, which mutually communicate. This communication can include simple data exchange, or it can include two or more services coordinating certain activity.

The basic structural elements are not processes, as it is for process architecture, but services, just as its name suggests. Misovic (2007) defines corporate service as an activity unit which can be composed from a discrete functionality, or from a certain number of linked processes.

Services as structural units are not extensive, they always represent only on particular functionality. It is then possible to chain them in greater complex, which is easy to adjust them in case of need. It creates flexibility, which is in application architecture hardly achievable.

Service represents certain functionality of real world. The essence of a service, in contrast to processes, is not a process itself, rather its result, which is provided by this process after request.

In SOA we run into “web service”, they are defined as software system designed for remote cooperation of

computers through network. They provide standard tool for cooperation in diverse computer applications working on various platforms (W3C, 2004). Its function is based on sending messages in XML format.

Every web service is characterized by its functionality. As an example of realization of any web service related to CRM could be presented calculation of customer value.

This value shows how big profit could one customer bring to an enterprise. Customer value helps to make better decision in decision making about investments in target marketing (Chalupova, 2009).

Web services, just like functions, return output values or run certain process on basis of given input parameters. Calculation of customer value is a good example of web service application, because there is a number of methods of how to calculate the lifetime customer value. Particular methods can be represented by individual services, which can be easily commuted in case of need for methodology change. On an example of the following formula (1) of one of the methods it is clear that data have character of time series.

$$LTV = \sum_{i=1}^n \pi_i \cdot \left(\frac{1}{1+d} \right)^i$$

where

d represents discount rate,

i represents particular time period,

n stands for length of customer lifecycle,

π_i represents profit collected from a customer in the period I

The obtained value then represents a measure of importance of one customer for a company (Chalupova, 2009). For a description of web services, their functionality and structure of their input and output data, we can use WSDL language, which was designed for these purposes. It describes syntax of service call. Service description for calculation of lifetime customer value could have the following form:

```
<definition name="CustomerValue"
  targetNamespace="http://www.example.cz/hz.wSDL"
  xmlns="http://schemas.xmlsoap.org/wsdl"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  xmlns:ns="http://www.example.cz/hz.wSDL">
  <types>
  ...
  </types>
  ...
  <message>
  ...
  </message>
  ...
</definitions>
```

Structure of input data for value calculation is rather complicated. There is not give a exact number of results due to the reason that we do not know ahead the number of time periods for which it is calculated.

The technology of web services is closely connected with XML languages. WSDL itself describes XML language which transfers input data.

```
<types>
  <schema
    targetnamespace=
      "http://www.example.cz/hz.xsd"
    xmlns=
      "http://www.w3.org/2001/XMLSchema"
  <element name="customer">
    <complexType>
      <element name="period"
        maxOccurs="unbounded">
        <complexType>
          <element name="TRate"
            type="decimal"/>
          <element name="profit"
            type="decimal"/>
        </complexType>
      </element>
    </complexType>
  </element>

  <element name="value"
    type="decimal"/>
</schema>
</types>
```

The return values of web services are then also in form of XML documents. That, for example, in case of use of SVG, provides that the web service returns data, for example in a form of graphs. The service description also includes description of service functionality, description of communication messages, linkage to protocols and other. Call of the service itself is secured by sending a message through SOAP (simple object access protocol) between provider (server) and service consumer (client). These messages include XML document, which include input and return data for the web service which was called.

2 RESULTS AND DISCUSSION

Most of current solutions of CRM systems are offered in a form of complex application, which include constant set of CRM processes. But requirements on functionality of CRM systems can be in each enterprise different.

Enterprises obviously have a possibility of a choice from a wide scope of offered solutions, among which it is possible for find an optimal solution for the moment. If company requirements change in time, it is problematic. Company then needs to invest in new software, or it must adjust its requirements to current solution. Way out of this situation is offered by service oriented architecture, which

doesn't approach a system as an indivisible unit, but as a complex of cooperating services.

CRM process set includes number of processes, which can server as a base for service design, just like in the example of lifetime customer value calculation.

Services can, thus, represent certain partial functionality of CRM system. When a particular functionality is not used, it doesn't need to be included in the system. That leads to system simplification. Such system can provide its uses only those tools, which will be really used. In case of need for system functionality change, it is not necessary to replace the entire system

Extension of functionality is possible through implementation of new services to the existing ones. Company using the system could pay only for those services which are presently used. The approach provided by SOA also brings new questions, which need to be answered before the system is put in operation.

3 CONCLUSIONS

Service oriented architecture brings new approaches to building information system. It provides operators, sometimes also a user, a simple way of how to make a tailored information system which can be very efficient in many cases. A characteristic of applications, built on basis of service oriented architecture, is the possibility of their simple and fast adjustments, according to current needs.

In contrast to other complex solutions, it is not necessary to buy new software, when in need for partial change of a system. However, implementation of new technologies brings new objectives and new questions, which need to be answered.

4 REFERENCES

- L. Gála a kol. *Podniková Informatika*, 1. vyd. Praha: Grada Publishing, 2006. 484 s. ISBN 80-247-1278-4
 - N. Chalupová, *Nové perspektivy uplatnění ICT v oblasti sledování a hodnocení vztahů zákazníka a poskytovatele v procesu obchodování*, Disertační práce, Brno, 2009,
 - M. Mišovič „Současné architektury podnikových informačních systémů“, ACTA UNIVERSITATIS AGRICULTURAE ET SILVICULTURAE MENDELIANEA BRUNENSIS, 2007, roč. 55, č. 6, s. 233-242. ISSN 12-11-85-16
 - M. Mišovič, J. Turčinek „System of Customer Oriented Services“ In BIELIK, P. *Medzinárodné vedecké dni 2008. Konkurenceschopnosť a ekonomický rast. Európske a národné perspektívy*. Nitra: SPU Nitra, 2008, s. 269--273. ISBN 978-80-552-0060-6
 - J. Turčinek „CRM z pohledu servisně orientované architektury“ In *Konkurence -- teoretické a praktické aspekty*. 1. vyd. Jihlava: Vysoká škola polytechnická Jihlava, 2009, ISBN 978-80-87035-23-8
 - H. Wesling, *Aktivní vztah k zákazníkům pomoci CRM*, 1. vyd. Praha: Grada Publishing, 2001. 169 s. ISBN 80-247-05969-9
- W3C, Web Service Architecture [on-line], <http://www.w3c.com/TR/ws-arch>

PROJECT SELF-EVALUATION METHODOLOGY: THE HEALTREATS PROJECT CASE STUDY

Martin Žnidaršič¹, Marko Bohanec^{1,2}, Nada Lavrač^{1,2}, Bojan Cestnik^{3,1}

¹ Jožef Stefan Institute, Jamova cesta 39, Ljubljana, Slovenia

² University of Nova Gorica, Vipavska cesta 13, Nova Gorica, Slovenia

³ Temida d.o.o., Ljubljana, Slovenia

e-mail: martin.znidarsic@ijs.si

Tel: +386 1 477 3366; fax: +386 1 477 3315

ABSTRACT

The paper presents an approach to self-evaluation in collaborative research projects. The approach is taken from a case study of the project Healththreats, where it is used in practice. Aims and focuses of self-evaluation are presented in general and the proposed methodology is described in detail. We conclude with practical experiences and suggestions for improvement.

1 INTRODUCTION

Project-based organization that includes many collaborating partners is very common form of research and development (R&D) consortia. Among the reasons for this are the increasingly interdisciplinary nature of current R&D topics, the financing schemes for R&D (e.g. European Framework Program projects), and the general inclination of the industry towards such kind of work organization.

There are specific good managerial practices that should be followed in order to make collaborative project-based work successful. One of them is project self-evaluation. It is a collection of activities that are conducted in order to provide regular self-assessment of the project's activities and status.

This paper presents a methodology of project self-evaluation that was developed using the methodology of multi-attribute decision support. The developed methodology was used for project self-evaluation in the project Healththreats [5], aimed at the development of an integrated decision support system for health threats and crises management.

Characteristics of collaborative R&D projects, which the method is targeting, are presented in Section 2. The proposed self-evaluation methodology is presented in Section 3 and our experiences in the Healththreats example case are given in Section 4. The paper concludes with Section 5.

2 COLLABORATIVE PROJECTS

A typical setting of R&D collaborative projects is a nationally mixed group of experts from various fields of

research and industry. The work in these projects is usually divided into work packages (WPs). Each WP represents a separate group of activities that are focused towards a common objective. WP has a dedicated WP leader, i.e. a partner that is responsible for its successful operation, but usually many partners are involved carrying out specific WP tasks.

Collaborative projects are managed by project management that coordinates the work, takes care of keeping agreed deadlines and communicates with WP leaders and external entities (regulatory bodies, stakeholders, etc.). It is not uncommon for projects to have a special WP dedicated to managerial tasks.

Self-evaluation is also an organizational activity and is sometimes assigned to a specific WP. The aim of self-evaluation is to monitor the activities and achievements of the project, to present and evaluate them in an objective manner and to provide feedback in form of warnings and recommendations. This helps detecting potential project weaknesses that need timely reaction and improvement, thus helps preventing potential problems before they become real ones. Self-evaluation may also provide useful information for external reviewers of the project.

3 EVALUATION METHODOLOGY

In this section we present a methodology of project self-evaluation, which was developed and used in the project Healththreats, but is applicable to a wide range of similar projects.

3.1 What to evaluate

Collaborative projects are complex systems with many interconnected entities. They consist of WPs, which are on a higher level managed by project management. The WPs have specific tasks and goals, but usually some of the goals of the project depend on the activities that connect the specific achievements of the WPs. Therefore, it is not straightforward and simple to decide what components of the project to evaluate and on which level of detail.

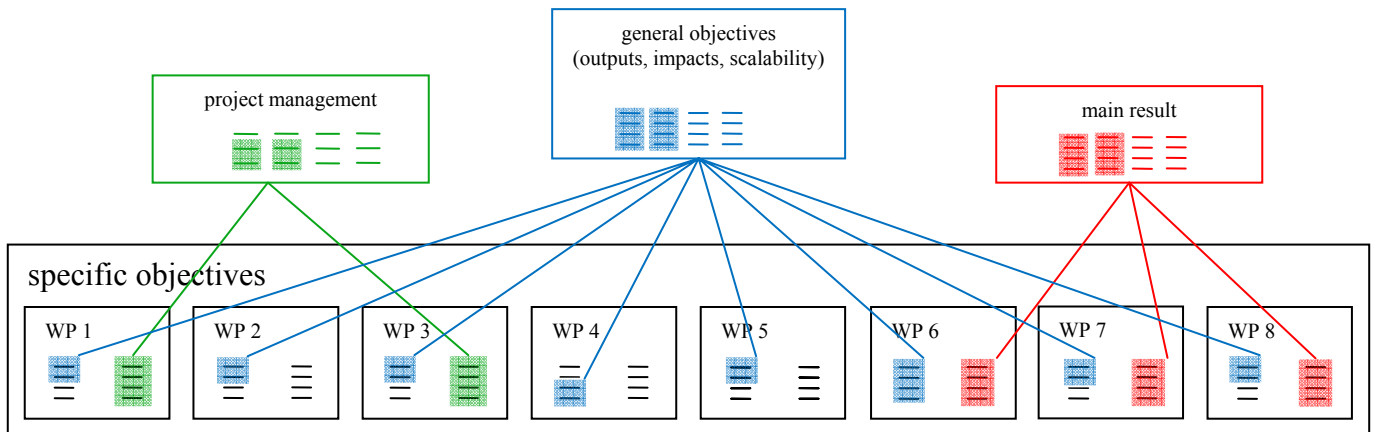


Figure 1: Schema of project self-evaluation aspects.

In our case, we chose to evaluate the project with four focus points that are presented in Figure 1:

- Project management
- General project's objectives
- Main result of the project
- Specific objectives (WP specific goals)

In this figure the small horizontal lines represent selected indicators and the colored/shaded boxes represent the aspects of evaluation. According to this schema, a separate focus point is devoted to *project management* activities. Another focus point is devoted to *general objectives* of the project, which are defined to be general and are monitored and evaluated on a high level of abstraction mainly by descriptive indicators. On the contrary, the *main result* of the project (a particular decision support system for end users) is evaluated at the highest possible detail with very specific, concrete and easily measurable indicators. Somewhat special is the evaluation of WPs, which are evaluated separately, but in a unified way (see section 3.2). All the WP evaluations are gathered under *specific objectives*.

Because of evaluation taking place on different levels of detail and with different contexts, some of the indicators (represented by small colored lines) are used more than once.

3.2 How to evaluate

The evaluation components were evaluated in different level of detail, depending on the focus point. There are different indicators and monitoring/evaluation processes dedicated to each of them. For the purpose of evaluation, the indicators get aggregated into higher level concepts in qualitative hierarchical evaluation models. The modeling and analysis of aggregated criteria follows the multi criteria decision modeling (MADM) approach [6, 3] and use the qualitative DEX methodology [1, 2, 7]. This methodology allows for an

efficient, comprehensible and transparent evaluation of multiple criteria.

PROJECT MANAGEMENT

The management of the project is evaluated according to communication, coordination of work, technical leadership and timeliness. The first three criteria are evaluated on the basis of a yearly questionnaire for project partners, which is provided to all the work package leaders. The indicator of success is the average answer to a question, where the answers can be given using a rating scale from 1 (worst) to 5 (best).

Timeliness is evaluated according to the timeliness of the deliverables and milestones reached. The indicator is a number of deliverables delivered in time, with agreed delay and with an exceptional delay.

GENERAL OBJECTIVES

The general objectives are evaluated through the assessment of:

- Outputs
- Impacts
- Scalability

Outputs

The evaluation of the outputs is the aggregation (a) of the aggregated indicators concerning the evaluation of project web site, publications and presentations, and (b) of the aggregated indicators concerning the developed system. Sketch of a possible aggregation hierarchy for this purpose is shown in Figure 2.

Impacts

The assessments of impacts are intrinsically ex-post assessments. In the Healththreats project the specific impact focus points were selected according to the CDC guidelines [4] and are not mentioned here since they are specific to the project's domain.

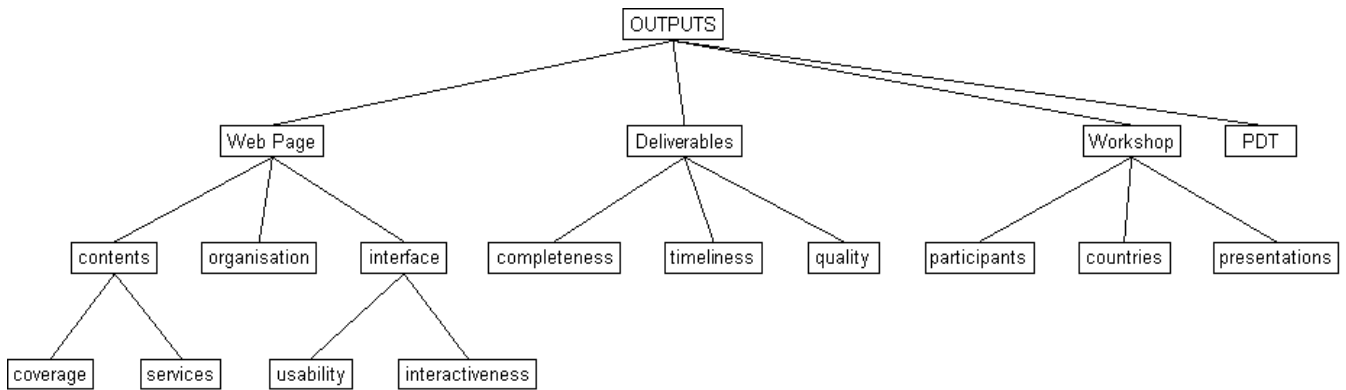


Figure 2: A hierarchy of aggregated indicators for evaluation of project's outputs. PDT is the acronym for the main result of the project in the Healthreals case.

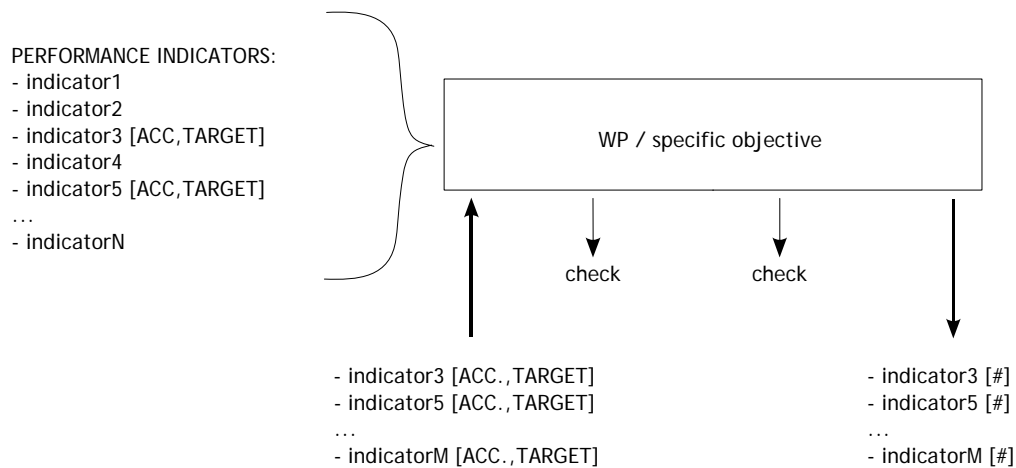


Figure 3: Sketch of procedure for definition, monitoring and evaluation of specific objectives.

The issue of ex-post nature of these assessments was solved by the following procedure: a pre-assessment by the stakeholders is made, according to the expected impacts of the project's results. The assessments are given as ++ (big positive expected impact), + (small positive expected impact), 0 (no impact), - (small negative impact), -- (big negative impact).

The assessments are made in the beginning of the project (before results testing phase) and at the end of the project (after the results testing phase). The final indicator of success is measured as cumulative shift in the positive or negative direction. No shift or a positive shift is considered a successful evaluation, since this would mean that the results of the project support the views and expectations that were present at the start of the project. A negative shift would indicate a possible underachievement or just a starting misinterpretation, thus would need further elaboration and explanation.

Scalability

Scalability of the approach is assessed from two perspectives:

- geographical scalability,
- problem domain scalability.

The geographical scalability describes the ability of the approach to be used in different state regions, states, EU regions. The level of geographical scalability is assessed through the (dis)similarity of issues brought out in the case studies, which will be contributed to the differences caused by geographic (political) factors.

The problem domain scalability describes the ability of the approach to be used in different problem domains. The level of problem domain scalability is determined by the amount of problem-specific features that are incorporated in processes and tools of the project's results.

SPECIFIC OBJECTIVES

Because the specific objectives and the WPs are typically very correlated, we evaluate the success in reaching the specific objectives and the success of the work in WPs together and uniformly for all the work packages.

For each pair, consisting of a work package and its specific objective, a set of indicators has to be chosen, which allows the evaluation of WP's work as well as the level of success in reaching a specific objective. Specific objectives usually share some of the indicators, but most of them are selected specifically for each objective. For some of the indicators, the responsible WP team is able to set its minimum (acceptable) and maximum (target) criteria goals at the beginning of the WP work. The status of all the criteria is checked at the end of the work and can also be monitored in between. A simple sketch of this procedure is presented in Figure 3.

For the indicators that are possible and agreed to monitor during the work, we use a qualitative three level scale for the measurement of completion and current quality for the purpose of monitoring. The values of the scale could be for instance green, yellow and red, where the green would mean a good or normal state of the indicator, the yellow would indicate a minor incoherence with the plans (like a postponed deliverable, or a result of work that needs some agreed further modifications) and the red would indicate a potential problem for the objective (like a deliverable missing a postponed deadline or an improper result).

The WP achievements (the achieved values of individual performance indicators) are given for every 12 months in three columns: acceptable, target and actual. Acceptable column represents minimal acceptable achievements that suffice for the purpose of the work package. Target achievements are set as realistic goals that can be accomplished within a given work package. Actual achievement column includes a list of achieved work package results.

In Healthreats project we monitored the specific objectives results in six month intervals and reported the status in intermediate reports that were made every six months.

4 EXPERIENCE AND SUGGESTIONS FOR IMPROVEMENT

The presented methodology was accepted well by the project partners in the case study and received particularly good feedback from the project officer. It was presented to some related projects as an example of good practice.

From our experience we can now make also some remarks and suggestions. It is very important that the project partners, especially the WP leaders, are involved in the planning of the self-evaluation and that they provide relevant and measurable indicators for their specific tasks. Further, it is important to review and convey the findings of self-evaluation in easily readable (graphical) form and as frequently as possible. The six months interval that we chose in our case study could ideally be shortened to three months to keep the initiative more up-to-date.

5 CONCLUSION

A methodology of self-evaluation for collaborative projects was presented in the paper. Since it is successfully used in practice and received positive remarks from evaluation professionals, it is potentially interesting for any collaborative project consortium that decides to support its work with this activity. There are some lessons presented that were learned in practice, but the methodology might also evolve further or serve as a base that is adapted to specific needs of various types of projects.

Acknowledgment

This work has been financially supported by the Healthreats project, which is sponsored and co-funded by the European Commission's Executive Agency for Health and Consumers (EAHC) and by the Slovenian Research Agency. The authors thank Healthreats' partners for their contribution in formulating specific indicators of the presented methodology.

References

- [1] Bohanec, M., Rajkovič, V. 1990. DEX: An expert system shell for decision support. *Sistemica* Vol. 1, Issue 1, pp. 145–157.
- [2] Bohanec, M. 2008. *DEXi: Program for multi-attribute decision making, User's manual, Version 3.00*. IJS Report DP-9989, Jožef Stefan Institute, Ljubljana, 2008. Available at: <http://kt.ijs.si/MarkoBohanec/pub/DEXiManual30r.pdf>
- [3] Bouyssou, D., Marchant, T., Pirlot, M., Tsoukiàs, A. and Vincke, P.: *Evaluation and decision models with multiple criteria: Stepping stones for the analyst*. International Series in Operations Research and Management Science, Volume 86. Springer, Boston, 1st edition, 2006.
- [4] CDC. *Updated guidelines for evaluating public health surveillance systems: recommendations from the guidelines working group*. MMWR 2001;50(no. RR-13);1-35.
- [5] Healthreats (2007): HEALTHREATS 150107: *Integrated decision support system for health threats and crises management (2007–2010)*. <http://www.healthreats.eu/>
- [6] Keeney, R. L. and Raiffa, H.: *Decisions with multiple objectives: Preferences and value tradeoffs*. Cambridge University Press, 1993.
- [7] Žnidaršič, M., Bohanec, M., Zupan, B. 2008. Modelling impacts of cropping systems: Demands and solutions for DEX methodology, *European Journal of Operational Research*. Vol. 189, Issue 3, pp. 594–608.

INTELIGENTNI SISTEM ZA NADZOR OBJEKTOV

Erik Dovgan, Rok Piltaver, Matjaž Gams

Odsek za inteligentne sisteme

Institut Jožef Stefan

Jamova cesta 39, 1000 Ljubljana, Slovenija

Tel: +386 1 477 3393; fax: +386 1 477 3131

e-pošta: erik.dovgan@ijs.si

POVZETEK

Naraščajoče zahteve po varnosti in zmanjšanju stroškov so privedle do razvoja sofisticiranih načinov za zagotovitev varnosti. Tem zahtevam ustrezajo inteligentne metode, ki omogočajo zagotovitev višje stopnje varnosti za nižjo ceno v primerjavi s tradicionalnim fizičnim varovanjem. Zato smo razvili inteligentni varnostni sistem, ki z uporabo inteligentnih metod za prepoznavo nenavadnih dogodkov omogoča nadzor strogo varovanih objektov, pri čemer se zmanjša potreba po fizičnem nadzoru ter se tako zmanjšajo stroški varovanja. Varnostni sistem je sestavljen iz več inteligentnih modulov: ekspertni sistem, statistika, makro učenje, mehka logika in video modul. Ekspertni sistem sestoji iz množice pravil, ki so vnaprej določena, statistika, makro učenje ter mehka logika ocenjujejo obnašanje ljudi na podlagi preteklega obnašanja, medtem ko video modul opozarja na nedovoljeno gibanje na podlagi posnetkov iz kamer. V članku je predstavljeno delovanje sistema kot celote kot tudi posameznih modulov posebej. Poleg tega je predstavljen tudi način komuniciranja z nadzornikom, ki je ciljni uporabnik predstavljenega sistema. Pri tem je poudarek predvsem na razumljivosti razlag alarmov, ki morajo biti razumljive tudi ljudem, ki niso eksperti na področju umetne inteligence.

1 UVOD

Pomen zagotavljanja varnosti strogo varovanih objektov narašča predvsem zaradi vdorov z uporabo najnovejših tehnologij. Povečevanje fizičnega zagotavljanja varnosti predstavlja predvsem povečevanje stroškov, pri čem pa je učinkovitost vprašljiva zaradi npr. utrujenosti ob spremljanju dogajanja preko posnetkov nadzornih kamer in možnosti podkupovanja. Poleg tega ni zanemarljiva težavnost pri zaposlovanju zaupanja vrednih ljudi, predvsem za varovanje vojaških objektov, ki se zaradi mednarodnih misij nahajajo izven matične države, kjer je treba za varovanje zaposliti lokalne prebivalce. Fizično varovanje v takih primerih očitno ni ustrezno.

Inteligentne metode in postopki že nadomeščajo ljudi na več področjih, predvsem zaradi višje zanesljivosti in natančnosti v primerjavi z ljudmi. Poleg višje zanesljivosti pa uporaba takih metod tudi zmanjša stroške poslovanja. Te

metode je mogoče uporabiti na področju varnosti na več načinov, pri čem se zmanjša potreba po fizičnem varovanju objektov. Eden izmed načinov uporabe je kontrola dostopa, kjer se osebe preverja ob vstopu v objekt [1]. Pri tem se lahko preverja različne fizične lastnosti, kot so prstni odtis, vzorci hoje, obnašanja [2] itd. Lahko pa preverjamo tudi samo obnašanje v objektu. Inteligentne metode delujejo na podlagi naučenega fizičnega vzorca oziroma vzorca obnašanja, ki je lasten določeni osebi, ga je težko posnemati in je zato zadostno zanesljiva mera za identifikacijo, verifikacijo in kontrolo ljudi. Zato uporaba teh metod za varovanje objektov občutljivo poveča varnost za razmeroma nizko ceno, pri čem pa se zmanjša potreba po fizičnem varovanju, kar zmanjša skupne stroške zagotavljanja varnosti.

V tem članku predstavljamo inteligentni sistem za zagotavljanje varnosti v strogo varovanih objektih z nazivom Poveljnikova desna roka (PDR). Vstop in gibanje v teh objektih je dovoljeno samo osebam, ki nosijo identifikacijske značke, katerih pozicija oziroma gibanje sta pridobljena z lokalizacijskim sistemom. Varnost je zagotovljena z nadzorom gibanja, pri čem se vzorec gibanja preverja s petimi inteligentnimi metodami: ekspertna pravila, statistika, mehka logika, video modul ter makro učenje. Vsak izmed modulov stalno nadzira prostor in asinhrono podaja opozorila o anomalijah, ki jih prepozna v obnašanju ljudi v objektu. Pri tem se generira tudi razlaga o anomaliji, ki je razumljiva nadzorniku, ki uporablja sistem. Na podlagi stanja posameznih modulov sistem prikazuje tudi globalno stanje v objektu. Poleg tega sistem tudi shranjuje preteklo dogajanje in tako omogoča pregledovanje le tega. Tako zasnovan varnostni sistem omogoča celovit pregled trenutnega stanja v objektu ter pregledovanje preteklih dogodkov, kar zviša kakovost varovanja objekta.

Članek je sestavljen iz dveh glavnih poglavij. Inteligentni sistem PDR je opisan v 2. poglavju. Na začetku je opisana strojna oprema, ki se uporablja pri tem sistemu, nato pa so opisani inteligentni moduli, ki nadzirajo dogajanje v objektu. Postopek ter rezultati testiranja sistema pa so podani v 3. poglavju.

2 INTELIGENTNI SISTEM PDR

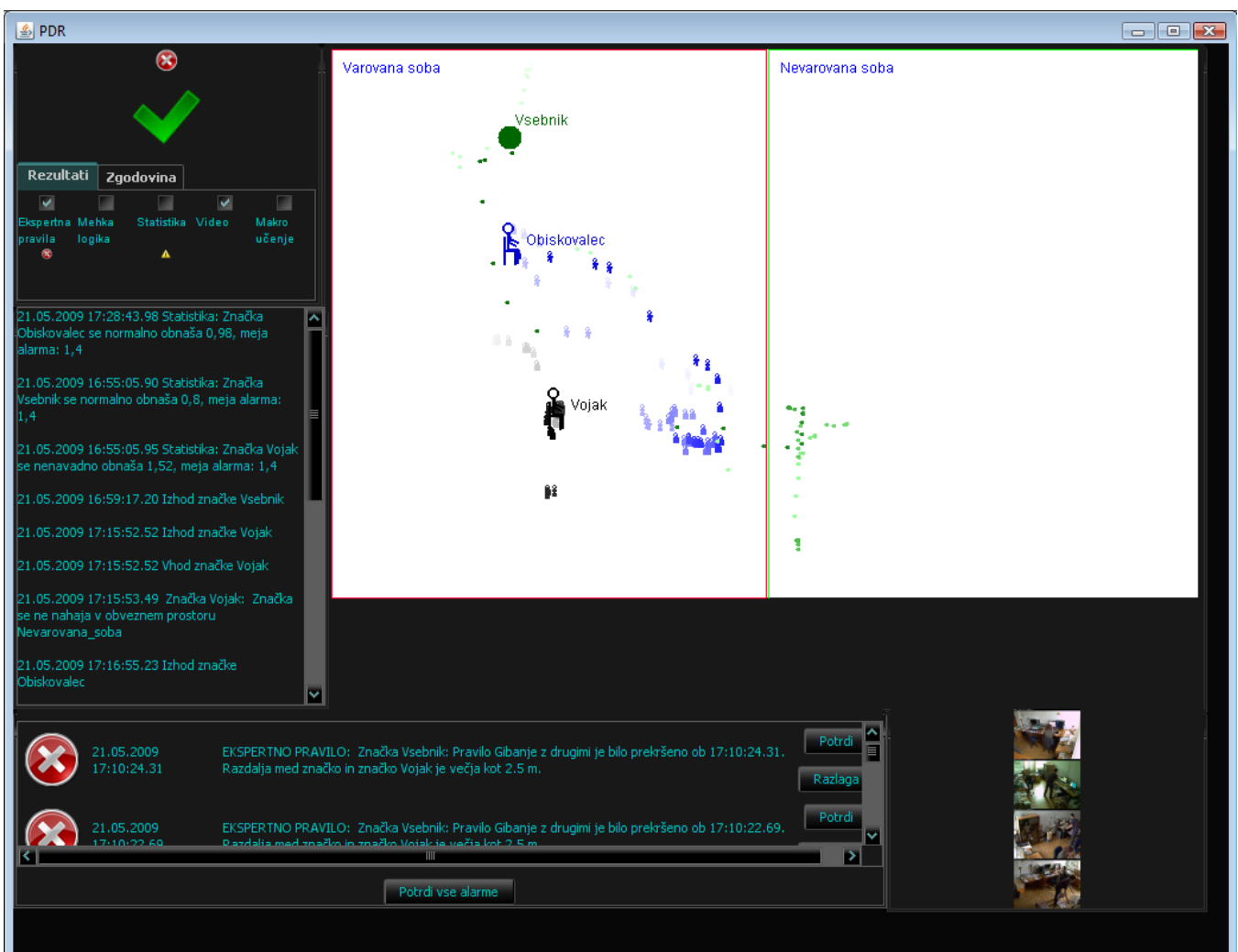
Inteligentni sistem Poveljnikova desna roka (PDR) je sistem, ki nadzira dogajanje v objektu, prepoznava nenavadne dogodke, jih sporoča nadzorniku ter tako zviša nivo varnosti objekta. Sestavljen je iz treh sistemov na nivoju strojne opreme ter iz več modulov na nivoju programske opreme, ki so združeni v integriran inteligentni sistem. V naslednjih poglavjih bomo prvo predstavili delovanje strojnega nivoja, potem pa še delovanje inteligentnega sistema.

2.1 Strojni nivo sistema PDR

Strojna oprema je sestavljena iz lokalizacijskega sistema Ubisense [3], množice kamer ter procesnega strežnika. Ubisense je sistem, ki vrši identifikacijo pozicije oseb s pomočjo radijskih valov [4]. Vsaka oseba, ki vstopi in se

giblje po objektu, mora imeti pripeto identifikacijsko značko, s pomočjo katere zna Ubisense v vsakem trenutku lokalizirati to osebo. Trenutna implementacija Ubisensa ima naslednje lastnosti. Napaka pri določanju pozicije mirujoče značke je v povprečju okoli 5 cm, ki pa naraste, ko se značka premika ter ko se značka nahaja na robnem območju možne lokalizacije, kjer lahko naraste tudi čez 50 cm. Največje število dobljenih pozicij za eno značko je 10 na sekundo. To drži pri pogoju, da spremljamo največ 4 značke. Glede na to, da tehnologija temelji na radijskih valovih, zaznavanje pozicije motijo vsi elementi, ki motijo radijske valove, npr. voda in kovine.

Kamere snemajo prostor iz več kotov. Postavljene morajo biti tako, da pokrijejo celoten prostor. Poleg tega se morajo njihovi vidni koti prekrivati, da lahko vsaki točki v prostoru določimo točno pozicijo s pomočjo slik iz kamer. Procesni strežnik hrani programsko opremo za nadzor



Slika 1: Grafični vmesnik sistema PDR. Zgoraj levo se nahaja stanje celotnega sistema – trenutno ni alarmov, a manjši rdeči križec nakazuje, da so v preteklosti bili še nepotrjeni alarmi. Za vsak modul je prikazano trenutno stanje v obliki alarmov oziroma opozoril ter splošne razlage (levi del okna). Na dnu so prikazani nepotrjeni alarmi ter slike iz kamer. Glavni del okna prikazuje trenutno obnašanje v prostoru, kjer se nahajata dve osebi: vojak in obiskovalec. Obe osebi sedita. Poleg njiju je v prostoru tudi predmet vsebnik, označen s krogcem. Prikazano je tudi nekaj zadnjih pozicij značk.

prostora. Povezan je z Ubisensom in kamerami, pri čem procesira lokacije iz Ubisensa in slike iz kamer v realnem času. Nahaja se v nadzornem prostoru, do katerega ima dostop samo nadzornik, kateri je tudi zadolžen za ukrepanje ob prepoznavi alarma s strani inteligentnega sistema.

2.2 Inteligentni nivo sistema PDR

Inteligentni sistem je sistem, ki nadzira prostor tako, da procesira pozicije pridobljene z Ubisensom in slike pridobljene s kamerami z inteligentnimi metodami ter prepozna nenavadne dogodke, katere sporoči nadzorniku preko uporabniškega vmesnika, prikazanega na sliki 1. Poleg tega prikaže tudi razlago, zakaj je nek dogodek nenavaden. Zgrajen je modularno, kar omogoča enostavno dodajanje novih modulov. Opisan sistem je sestavljen iz petih modulov, ki so predstavljeni v naslednjih podpoglavjih. Na koncu so predstavljeni načini razlage klasifikacij modulov.

2.2.1 Ekspertni sistem

Prvi modul je ekspertni sistem, kateri vsebuje množico generičnih pravil. Z definiranjem vrednosti parametrov pravil dobimo pravila, katera določajo dovoljeno obnašanje v objektu. Vsako kršenje pravil sproži alarm.

2.2.2 Mehka logika

Drugi modul je mehka logika, ki na podlagi preteklega obnašanja določene osebe oceni opazovano obnašanje. Pri tem modul ocenjuje pot, kjer je oseba hodila, hitrost hoje, smer gibanja, višino značke itd. Prostor je razdeljen na območja, kjer se za vsako območje za vse pretekle sprehode shranjuje statistika navedenih atributov. Poti, po katerih se je oseba statistično največ gibala, hitrost, s katero se je največkrat gibala itd. predstavljajo običajno hojo osebe. Če opazovana hoja ni podobna običajni hoji osebe, modul sproži opozorilo, saj predvideva, da je nekaj narobe z osebo oziroma da oseba, ki nosi značko, ni ista kot identificirana oseba.

2.2.3 Video modul

Tretji modul je video modul. Za razliko od ostalih modulov je ta edini modul, ki procesira slike iz kamer. Pri tem prepozna vsako gibanje, ki se zgodi v prostoru. Gibanje je sprememba na določenem področju zaporednih slik določene kamere. Ob zaznavi gibanja se izračuna pozicija gibanja v prostoru ter verjetnost, da se na tej poziciji nahaja oseba. Nato modul preveri, če se na izračunani poziciji nahaja kakšna značka, ki označuje osebo, če je verjetnost, da se giblje oseba, velika. Če tam ni take značke, se sproži alarm, saj je prišlo do vdora v objekt. V nasprotnem primeru, ko je verjetnost, da se giblje oseba, nizka, pa je prišlo do gibanja predmetov. Predmeti se ne morejo sami premikati ter jih ne smejo premikati neavtorizirane osebe, npr. obiskovalci. Zato se preveri, če se v bližini predmeta nahaja značka osebe, katera ima dovoljenje, da premika

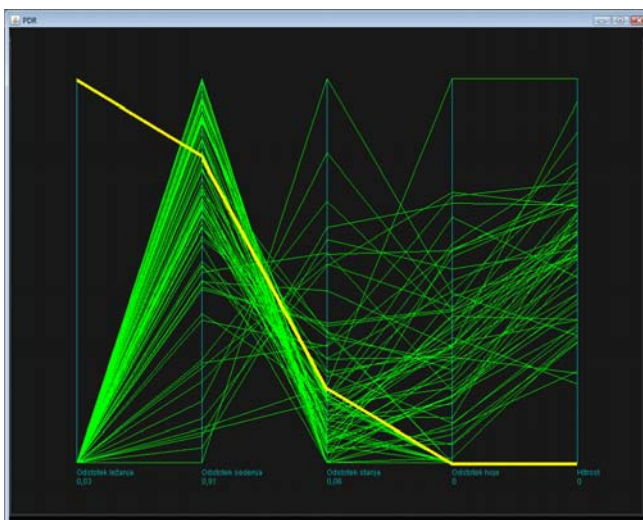
predmete. Če se v bližini ne nahaja takšna oseba, je prišlo do nedovoljenega dogodka npr. kraje in se sproži alarm.

2.2.4 Statistika

Četrti modul je statistika, ki ocenjuje opazovano obnašanje vsake osebe na podlagi preteklih vzorcev obnašanja te osebe. Pri tem za preteklo obnašanje shranjuje podatke o začetnem in končnem času zbiranja podatkov vzorca obnašanja, trajanju zbiranja podatkov vzorca, dnevu, ko je bil vzorec pridobljen, odstotku časa, ko je oseba ležala, hodila, sedela in stala, ter hitrosti gibanja. Vzorce gibanja se shranjujejo preko različnih časovnih intervalov: 30 s, 1 min, 2 min, 4 min itd. Za opazovano obnašanje prav tako izračuna statistični vzorec, ki se primerja z ostalimi vzorci. Pri tem se z algoritmom LOF [5] določi, ali opazovani vzorec odstopa od običajnega vzorca obnašanja [6]. Če pride do odstopanja, se sproži opozorilo, saj je z osebo nekaj narobe oziroma je prišlo do kraje identitete in vdora v objekt.

2.2.5 Makro učenje

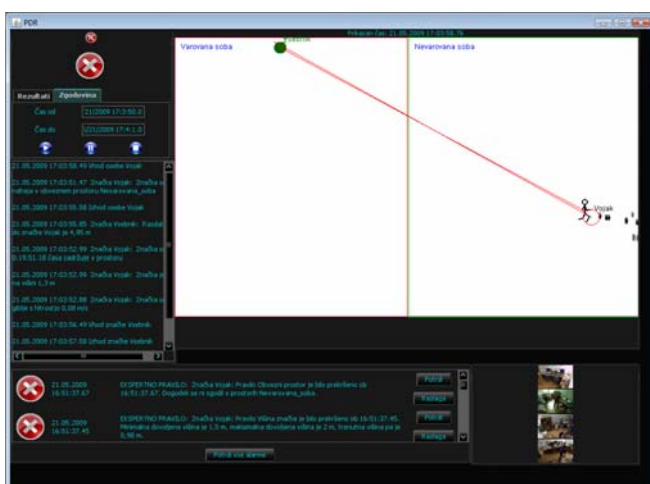
Peti modul je makro učenje, ki deluje podobno kot statistika. Prav tako zbira statistične podatke obnašanja osebe, kjer pa ni osredotočeno na trenutno statistiko obnašanja, ampak obravnava obnašanje na drugem, prostorskem nivoju. Tako se shranjujejo vzorci obnašanja za vsak prostor posebej, saj se osebe v različnih prostorih različno obnašajo. Podatki se ne zbirajo preko različnih časovnih intervalov, ampak preko različnih bivanj v prostorih. Zbirani podatki so: čas vstopa v prostor, čas izstopa iz prostora, trajanje bivanja, dan bivanja, prostor bivanja, odstotek časa, ko je oseba ležala, hodila, sedela in stala, ter hitrost gibanja. Nenavadno bivanje se prav tako kot pri statističnem modulu prepozna z algoritmom LOF.



Slika 2: Razlaga statistike in makro učenja prikazuje dogodek, opisan s petimi atributi. Dogodki, ki prikazujejo preteklo obnašanje osebe, so zelene barve. Opazovani dogodek, ki odstopa od preteklih dogodkov in zato sproži opozorilo, je rumene barve.

2.2.6 Razlaga klasifikacij

Poleg alarmov in opozoril, ki jih moduli vračajo, omogoča inteligentni sistem tudi več načinov prikazovanja razlag o tem, zakaj je bil nek dogodek v določenem trenutku nenavaden. Vsak izmed opisanih modulov poda razlago, zakaj je dogodek nenavaden, ki je razumljiva nadzorniku. Poleg tega statistika in makro učenje podata tudi razlago v obliki slike (kot je prikazano na sliki 2), video modul pa shranjuje in na nadzornikovo zahtevo prikaže video o nenavadnem dogodku. Dodaten mehanizem je tudi predvajanje preteklih dogodkov, kjer se izrisujejo alarmi, ki so bili sproženi z ekspertnimi pravili (kot je prikazano na sliki 3). Tako inteligentni sistem podaja celovit pregled ne samo tekočega ampak tudi preteklega dogajanja.



Slika 3: Prikaz razlage preteklega alarma, ki je označen na glavnem predelu uporabniškega vmesnika z daljico, ki povezuje preveč oddaljeni znački, ki kršita pravilo o obvezni bližini značk, npr. nadzornika in njegovega osebnega računalnika.

3 POSKUSI IN REZULTATI

Testiranje sistema je bilo izvedeno v več fazah. V prvi fazi testiranja je bil testiran ekspertni sistem. Vsako ekspertno pravilo je bilo testirano po predhodno definiranem scenariju. Rezultati testiranja so pokazali, da pravila delujejo, ko Ubisense vrača pozicijo značk z zadostno natančnostjo. Sicer pa lahko pride do določenih anomalij, npr. ko se pozicija neke značke v desetinki sekunde premakne za več kot 1 m. Take Ubisensove napačne zaznave pozicije značk so razmeroma redke, a ko pride do njih, pride do kršenja pravil. Verjetnost takih situacij smo zmanjšali z uporabo različnih filtrov (Kalmanov filter [7], mediana filter [8]), a še vedno se anomalije pojavljajo.

V drugi fazi smo testirali video modul. Pri tem je bila kritična časovna usklajenost med Ubisensom in kamerami, saj vsako odstopanje privede do napačnih alarmov. Na primer zaostajanje ur v kamerah glede na uro v Ubisensu privede do tega, da video modul preverja, če je bila kakšna

značka na poziciji gibanja, ki so jo vrnilo kamere, pred časom, ko se je značka dejansko tam nahajala. Posledično tam značke še ni bilo in sproži se alarm. Glede na to, da je popolna časovna usklajenost praktično nedosegljiva, lahko še vedno pride do napačnih alarmov.

V tretji fazi sta bila testirana modula statistika in makro učenje. Pri tem sta sodelovali dve osebi, ki sta se gibali po dveh nadzorovanimi prostoroma približno eno uro ter naredili vsaka okoli 12 normalnih in 12 nenavadnih vhodov in izhodov. Izvedeno je bilo leave-one-out testiranje, pri čem smo dobili za statistiko 99,3% ter za makro učenje 96,2% točnost prepoznavne normalnih in nenavadnih dogodkov.

V zadnji fazi bomo testirali mehko logiko in celoten sistem.

4 ZAKLJUČEK

Razvili smo inteligentni sistem za nadzor objektov, ki z uporabo inteligentnih metod prepozna nenavadne dogodke ter jih sporoča nadzorniku. Tako se zmanjša potreba po fizičnem varovanju objekta, kar zniža stroške varovanja, poveča pa se varnost, saj je sistem zanesljiv in robusten. Sistem prepozna nenavadne dogodke predvsem na podlagi preteklega obnašanja oseb. Sprememba obnašanja namreč pomeni, da je z osebo nekaj narobe ali pa da je prišlo do kraje identitete. V obeh primerih predstavljen sistem sproži alarm. Takšne situacije smo tudi simulirali med testiranjem, kjer jih je sistem uspešno prepoznal. Tako uspešno stestiran sistem je uporaben v objektih, kjer je zahtevana višja stopnja varnosti, saj pripomore k povešanju varnosti ter izboljšanju preglednosti in nadzora nad objektom s strani nadzornika.

Literatura

- [1] E. Dovgan, M. Gams. Ambient Intelligence for Entry Control. V zborniku 12. WSEAS Mednarodne konference o računalništvu. Heraklion. 2008.
- [2] C. Ramos, J. C. Augusto, D. Shapiro. Ambient Intelligence – the Next Step for Artificial Intelligence. *J. Intelligent Systems*. pp. 15-18. 2008
- [3] Ubisense. <http://www.ubisense.net>.
- [4] RFID. <http://sl.wikipedia.org/wiki/RFID>.
- [5] M. M. Breunig, H. P. Kriegel, J. Sander. LOF: Identifying density-based local outliers. *Proceedings of the International Conference on Management of Data SIGMOD'00*. pp. 93–104. 2000.
- [6] T. Tušar, M. Gams. Outlier detection in an access control system. V zborniku 9. mednarodne multikonference Information Society - IS 2006. vol. A. pp. 136-139. Institut Jožef Stefan. 2006.
- [7] Kalman. http://en.wikipedia.org/wiki/Kalman_filter.
- [8] Mediana. http://en.wikipedia.org/wiki/Median_filter.

SE SPLAČA PREMISLITI GLOBLJE?

Matjaž Gams

Odsek za inteligentne sisteme, Institut Jožef Stefan

Jamova 39, 1000 Ljubljana, Slovenija

Tel: +386 1 4773900; fax: +386 61 1251038

e-mail: matjaz.gams@ijs.si, <http://ai.ijs.si/mezi/matjaz.html>

Povzetek: Po patologiji minimaksa smo se lotili analize minimina in odkrili, da se kvalitativno obnašata podobno. Pri obeh smo odkrili patologije v situacijah, ki se zdijo realistične. Ker se minimaks uporablja v formalnih igrah dveh igralcev, minmin pa v primerih enega igralca, se zdi patologija minmina toliko bolj potencialno nevarna v realnem življenju.

Prispevek okvirno analizira drzno in nepreverjeno hipotezo: Je torej bolje odločati se po nekem svetovnem nazoru, ali je bolje preračunati možnosti in se odločiti glede na najboljši rezultat preiskovanja?

1 UVOD

Več kot tri desetletja je znano, da obstajajo formalne igre dveh igralcev (Nau, 1979), kjer globlje preiskovanje pri določenih pogojih prinese slabše rezultate. To protislovje so poimenovali patologija in glede na tip pregledovanja še patologija pozicije (če je globlja ocena pozicije slabša) in preiskovalna patologija, če je globlje preiskovanje bolj pogosto pripeljalo do slabše poteze kot krajše preiskovanje. Pri tem so bile tako prve kot naslednje raziskave vezane na algoritem minimaks (Smed & Hakonen 2008), ki se uporablja kot najosnovnejši algoritem za preiskovanje dreves iger dveh igralcev.

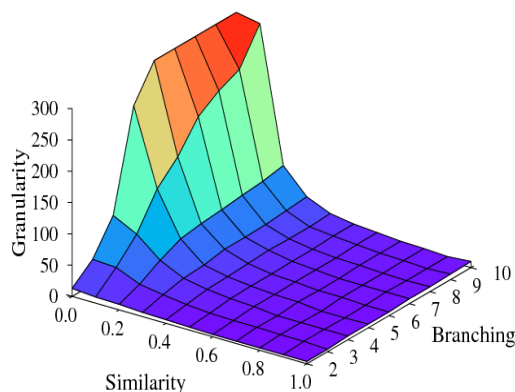
Patologija se seveda ne pojavi, kadar minimaks preišče vse možnosti do konca, tj. kadar je globina preiskovanja globlja kot vsi konci iger, pa tudi pri pomembnih spremembah, recimo pri izgubi pomembne figure v šahu ne. Pač pa se pojavi v igrah, kjer se situacija bistveno ne spremeni v korist enega ali drugega igralca po številu potez, ki je manjše od globine preiskovanja. Primer bi bila neka izenačena pozicija, ki po nekaj potezah ostaja enako izenačena.

V teh razmerah je Beal (Beal, 1980) prvi opisal patologijo na algoritmu minimaks. V naslednjih desetletjih je prišlo do vrste razlag, zakaj se pojavi patologija (Nau, 1979, 1982, 1983; Bratko & Gams, 1982; Pearl, 1983; Scheucher & Kaindl, 1998; Luštrek et al., 2005, 2006; Sadikov et al., 2005), vse te razlage pa so tipično uporabljale en sam mehanizem oz. parameter. Različni medsebojni vplivi skoraj nikoli niso bili raziskovani. Poleg tega so bili patološki primeri v realnih igrah sumljivo odsotni, tako da je bil vtis, da je patologija nekaj izredno eksotičnega.

V Odseku za inteligentne sisteme smo v zadnjih letih razvili modele patologije (Kaluža 2008; Piltaver 2008; Nau idr. 2009), ki so pokazali odvisnost treh ključnih parametrov, tj.

razvejitve b , števila vrednosti hevristične funkcije g in podobnosti s . Ti modeli so pokazali, da je pri velikem b , malem s in malem g zelo velika verjetnost patološkosti. Poleg tega so pokazali, da so igre dejansko lahko patološke na treh realnih igrah: 8 ploščic, šahovske končnice, kalah. Model za omenjene 3 parametre je prikazan na sliki 1.

V tem prispevku bomo analizirali, kaj ti zaključki pomenijo za realno življenje.



Slika 1: Patološka hiperravnina, kjer je preiskovanje v globino 5 enako uspešno kot preiskovanje v globino 1.

2 NAZORNA PRIMERA PATOLOŠKIH DREVES DVEH IN ENEGA IGRALCA

Na slikah 2 in 3 sta dva primera binarnih dreves z globino preiskovanja 2. Na sliki 2 je za igro dveh igralcev in na sliki 3 za igro enega igralca. Korenina drevesa predstavlja osnovno pozicijo, v kateri imamo na izbiro levo ali desno poddrevo. Prave vrednosti vozlišč so navedene v vozlišču in so lahko le 0 ali 1 (izgubljena ali dobljena). Z 1 si običajno predstavljamo pozicijo, v kateri dosežemo svoj cilj, torej zmagamo, v primeru, da smo na potezi. Mi torej želimo maksimirati možen rezultat. Kadar le sami izbiramo poteze, imamo torej sama Max vozlišča. Tako je drevo na sliki 3. Drevo na sliki 2 prikazuje drevo igre, kjer na vsako našo potezo odgovori nasprotni igralec. Nasprotnik nam skuša odvzeti zmago oz. jo spremeniti v naš poraz.

Realna primera omenjenih dreves bi lahko bila: Za igro enega igralca bi lahko dejali, da izbira pot v puščavi, kjer vidi sosednje vrhe puščave. Lahko se odloči glede na videz

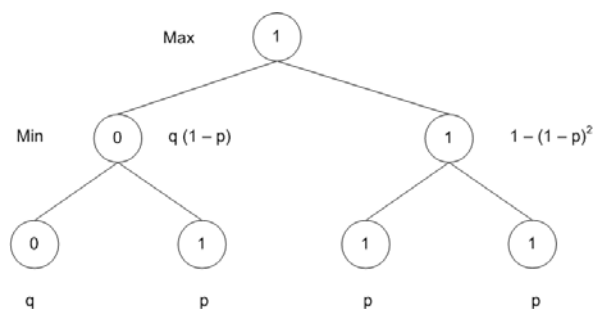
bližnjih sipin, lahko pa vzame daljnogled in se odloči glede na izgled daljnjih sipin..

Kjer pa imamo nasprotnega igralca, si lahko predstavljamo realni primer takole: Lovec lovi antilopo in jo želi ujeti, kar zanj predstavlja 1. Antilopa pa želi ostati živa, zato vleče poteze v tem smislu. Denimo, da lahko antilopa zbeži v travo ali v reko, vendar jo v reki lovec vedno ujame. V travi pa lahko steče v hrib ali v dolino in v hribu jo čaka drugi lovec, medtem ko po dolini uide. Torej jo lovec ulovi v treh od štirih primerov.

Vsi dogodki pa so verjetnostni in pračlovek-lovec lahko vnaprej samo z določeno verjetnostjo oceni situacijo, pa tudi ogled sipin samo z določeno verjetnostjo nakazuje, kje bi bila voda, npr. z nekaj več travnimi bilkami. V teh osnovnih modelih je vedno predpostavljeno, da je verjetnost, da bo 0 narobe ocenjena s »q« in verjetnost, da bo 1 narobe ocenjena s »p«. Pogosto predpostavimo še, da sta p in q enaki, včasih celo, da sta obe 0.1.

Formalna modela ustrezata modelom iger – minimin oziroma maxmax sta igri, kjer igralec na igri skuša doseči najmanjšo ali največjo možno vrednost in igra sam. Kadar igrata dva igralca, eden skuša minimizirati in drugi maksimizirati, od tod izraz minimax. Minimix se uporablja za preiskovanje dreves ALI, medtem ko se minimax uporablja za preiskovanje dreves IN/ALI, kjer se vozlišča zaporedoma izmenjujejo.

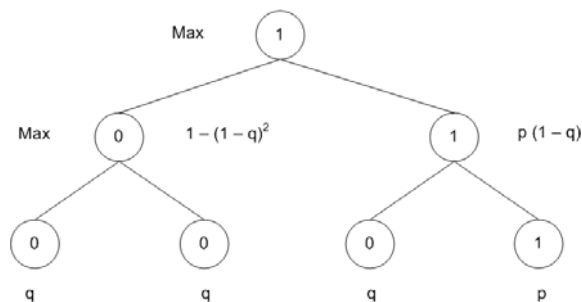
Sedaj lahko preračunamo verjetnosti v drevesih. Če se na nivoju 2, tj. najnižjem nivoju, saj označujemo korenino z nivojem 0, lahko zmotimo z verjetnostjo p ali q odvisno od vrednosti, potem so vrednosti na nivoju 1 označene desno od vozlišča. Za enega igralca sta možnosti $q(1-p)$ za levo vozlišče in za desno vozlišče $1-(1-p)^2$. Izračun je preprost: Da se zmotimo in levo vozlišče, ki ima pravo vrednost 0, ocenimo kot 1, se moramo zmotiti v levem nasledniku (q), medtem ko moramo desno podvozlišče pravilno oceniti ($1-p$). Skupna verjetnost je produkt.



Slika 2: Patološko binarno drevo globine dva za igro dveh igralcev (pračlovek lovi antilopo).

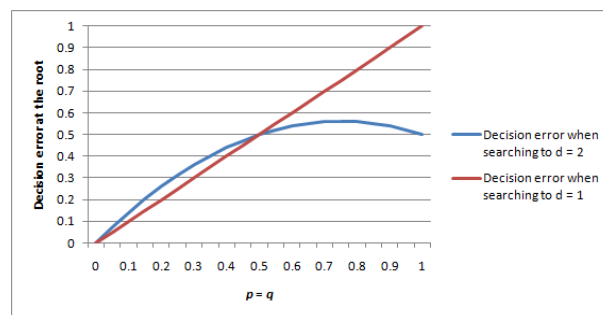
V desnem vozlišču prvega nivoja lahko naredimo napako le, kadar se ne zmotimo v oceni obeh vozlišč hkrati, od tod $1-(1-p)^2$. Ko izračunamo oceno napake v korenini ob dodatni predpostavki, da je $p=q$, dobimo $p(3-2p)/2$. To je torej ocena napake v korenini binarnega drevesa, kjer igra le en igralec.

Pri ocenjevanju verjetnosti napake drevesa, kjer igrata dva igralca, se izračun nekoliko spremeni, zato dobimo vrednosti napake, kot je označena desno pri vozliščih srednjega nivoja. Izračun napake v korenini pri pogoju $p=q$ da enak izraz kot pri prejšnjem drevesu: $p(3-2p)/2$.



Slika 3: Patološko binarno drevo globine dva za igro enega igralca (iščemo oazo v puščavi).

Ker imamo sedaj le en parameter, to je $p=q$, lahko narišemo verjetnost uspešnega preiskovanja, oziroma verjetnost napake $p(3-2p)/2$, v globini 2, v primerjavi z verjetnostno napako v globini 1, tj. p . Slika 4 očitno pokaže, da se preiskovanje v bolj spleča v globini 1 in manj v globini 2 za vse pričakovane vrednosti napak pri oceni pozicije, od 0 do 0,5.



Slika 4: Primerjava izraza $p(3-2p)/2$, tj. vrednost napake pri preiskovanju v globino dva v primerjavi z globino 1, kjer je $p=q$. Obe drevesi na slikah 2 in 3 sta torej patološki za večino smotrnih vrednosti p .

Poskusimo sedaj podati tolmačenje patološkosti v podanih primerih realnega življenja (puščava, lovljenje antilope). Če torej pogledamo iz tekoče sipine do dveh najbližjih in se odločimo na osnovi teh dveh ocen, bomo naredili manjšo napako, kot če gremo pogledati do naslednjih dveh. prvih dveh križišč naprej. Torej se preiskovanje v realnem primeru v večjo globino ne izplača!

Pri tem je bilo seveda še nekaj predpostavk. Prva je bila, da se kvaliteta ocene oz. verjetnost napake ne spremeni z globino. Z drugimi besedami – če pogledamo sipini bolj v daljavi, nič bolj jasno ne vidimo, kje bi bila oaza. Torej mora biti situacija tako nejasna kot prej. Druga predpostavka je bila, da ocenjujemo samo z dvema vrednostma. V realnih igrah, recimo šahu, ima ocenjevalna funkcija nekaj tisoč vrednosti. Vendar se ta hip ob iskanju

oaze ocena z 0 ali 1 ne zdi napačna, saj 1 pomeni lahko samo »bolj verjetno« in 1 »manj verjetno«, ali pa »kaže bolj« in »kaže slabše«.

Preprosti primer puščave torej jasno pokaže, da se v zapleteni situaciji enemu igralcu, ki z globljim preiskovanjem ne dobi boljše ocene, ne splača preiskovati v globino 2 v omenjenem primeru. V (Nau idr. 2009) je pokazano, da se tudi za vse možne kombinacije do globine 5 v takih razmerah ne izplača preiskovati v večjo globino, ob tem da analiza uvaja še nove parametre. Graf je prikazan v naslednjem poglavju. V (Dacar 2009) pa je pokazano, da se za vse primere pri teh pogojih na izplača preiskovati v globino ne glede na obliko binarnega drevesa, oz. porazdelitev vrednosti na drugem nivoju.

Tudi za lovca in antilopo v našem primeru je zaključek podoben: zopet se bolj izplača pomisliti samo en nivo naprej, kot pa dva nivoja naprej. Podobno kot za enega igralca je tudi za dva igralca v (Nau idr. 2009) in (Dacar 2009) pokazano, da so v splošnem razmere patološke.

3 ZAKLJUČEK

Modeli kažejo, da je patologija možna v vsaki domeni, kjer lahko poljubno nastavljamo parametre b , g in s . Pri tem večji s in g zmanjšujeta patologijo, večji b pa jo povečuje.

Praktične igre oziroma praktični problemi bolj ali manj določajo dva parametra – s in b , torej lahko v praksi spreminjamo samo g . Izbrati moramo čim večji g , ki bo imel smiseln pomen. Seveda je pomembna tudi kvaliteta cenilne funkcije, vendar na pojav same patologije nima ključnega pomena.

Kljub temu je prva ugotovitev iz modelov dokaj pomembna in spominja na Goedlov teorem, ki trdi, da v vsakem formalnem sistemu obstajajo resnični stavki, ki se jih ne da formalno dokazati. Analize formalnih modelov minimaxa in minimaksa so pokazale, da je v vsaki domeni možno nastaviti patološke razmere, kjer se preiskovanje v večjo globino ne izplača. Potrebno je le primerno zmanjšati g in s . Torej je patologija potencialno vedno prisotna v preiskovanju, le parametri so v realnih igrah taki, da je ne srečamo.

Drugo pomembno in delno filozofsko vprašanje je, ali torej v zapletenih razmerah ni smotno preiskovati v globino, posebej če ni dobre ocenjevalne funkcije, ki bi prikazala prednosti ene situacije pred drugo.

Odgovor, ki ga nakazujejo modeli, nesporno nakazuje, da se v zapletenih situacijah ne izplača pretirano razmišljati. V realnem življenju je veliko situacij vezanih na časovno komponento, ko je čas za razmišljanje omejen in potrebno reagirati hitro.

Kaj pa v primeru, ko imamo dovolj časa, npr. ko se odločamo za izbiro enega izmed nekaj partnerjev in so vsi primerljivi do te stopnje, da ni možno predvideti bistvenih razlik?

Ta primer je podobno kot veliko realnih družbenih problemov zapleten in ni hitrih jasnih odgovorov. Patologije ni moč vnaprej izločiti. Ni mogoče vnaprej reči, da bo večje

razmišljanje in izbiranje pripeljalo do boljšega rezultata. Konec koncev je stara ljudska modrost: »Kdor izbira, mu izbirek ostane!«

Drzna hipoteza je, da je zaradi potencialno vedno prisotne patološkosti v realnem življenju ljudi morda boljša strategija izbire po vnaprej določenih kriterijih v primerjavi s podrobnim analiziranjem v globino, še posebej, kadar na voljo ni podrobnejše cenilne funkcije.

To se zdi podobno dilemi, ali bi preiskovali sipine v okolici in potem izbrali najbolj perspektivno, ali pa se odločili za npr. sever in ubrali to pot. Pri socialnih, družbenih odločitvah se zdi, da je morda smotrnejša načelna usmeritev, kot pretirano preračunavanje.

Nekoliko bolj nepreverjena se zdi hipoteza, da so bili praljudje uspešnejši z načelnimi usmeritvami, kot npr. ideologije, verovanja, pa celo praznoverja, saj je v teh primerih bolje imeti tudi ne optimalno odločitev na površnem nivoju kot pa preiskovanje v globino. Ta razmišljanja pa bo nujno podrobno preveriti, ali je kaj na tem.

Zahvala:

Zahvaljujemo se ARRS za financiranje raziskovalne programske skupine Umetna inteligenca in inteligentni sistemi. Zahvaljujemo se naslednjim sodelavcem za sodelovanje pri raziskavah: Mitja Luštrek, Boštjan Kaluža, Rok Piltaver, Aleš Tavčar, Aleksander Sadikov, Matej Guid, Ivan Bratko, Dana Nau, Jana Krivec.

Literatura:

Beal DF. (1980) An Analysis of Minimax. In *Advances in Computer Chess 2*, ed Clarke MRB (Edinburgh University Press, Edinburgh), pp 103–109.

Bratko I., Gams M. (1982) Error Analysis of the Minimax Principle. In *Advances in Computer Chess 3*, ed Clarke MRB (Pergamon Press, Oxford), pp 1–15.

Kaluža B. (2008) Analysis of pathological minimax models and Pearl's game (Slovene: Analiza patoloških modelov minimaksa in Pearllove igre). BSc thesis, University of Ljubljana.

Luštrek M., Bratko I., Gams M. (2005) Why minimax works: An alternative explanation. In *Proc IJCAI*, pp. 212–217.

Luštrek M., Gams M., Bratko I. (2006) Is real-valued minimax pathological. *Artificial Intelligence* 170(6–7): 620–642.

Nau DS. (1979) Quality of Decision Versus Depth of Search on Game Trees. PhD thesis, Duke University.

Nau DS. (1982) An investigation of the causes of pathology in games. *Artificial Intelligence* 19(3): 257–278.

Nau DS. (1983) On game graph structure and its influence on pathology. *International Journal of Computer and Information Sciences* 12(6): 367–383.

Pearl J. (1983) On the nature of pathology in game searching. *Artificial Intelligence* 20(4): 472–453.

Piltaver R. (2008) Search pathology in eight-puzzle (Slovene: Patologija preiskovanja v igri osmih ploščic). BSc thesis, University of Ljubljana.

Sadikov A., Bratko I., Kononenko I. (2005) Bias and pathology in minimax search. *Theoretical Computer Science* 349(2): 261–281.

Scheucher A., Kaindl H. (1998) Benefits of using multivalued functions for minimaxing. *Artificial Intelligence* 99(2): 187–208.

Smed J., Hakonen H. (2006) *Algorithms and Networking for Computer Games*. Wiley.

GLAJENJE TRAJEKTORIJ GIBANJA ČLOVEŠKEGA TELESA ZAJETIH Z RADIJSKO TEHNOLOGIJO

Boštjan Kaluža, Erik Dovgan
Odsek za inteligentne sisteme, Institut "Jožef Stefan"
Jamova cesta 39, 1000 Ljubljana, Slovenija
{*bostjan.kaluza,erik.dovgan*}@ijs.si

POVZETEK

Z radijsko tehnologijo (UWB) je mogoče spremljati gibanje človeka pri vsakdanjih opravilih na cenovno ugoden način, a za ceno slabše kakovosti zajetih signalov. V prispevku analiziramo šum pri tako zajetih signalih in predstavimo postopek za zmanjšanje učinka šuma, ki temelji na upoštevanju narave šuma, anatomije telesa in porazdelitve šuma. Predhodni rezultati kažejo izboljšanje ocene pozicije in zanesljivejše ocenjevanje hitrosti.

1 UVOD

Prepoznavanje obnašanja ljudi je mogoče le, če se naprave zavedajo okolice, kar pomeni, da morajo biti sposobne pridobiti podatke iz okolice. Eden izmed načinov pridobivanja podatkov je tudi nadzor okolice oziroma ljudi s sensorji. Pri tem so pomembni predvsem lokalizacijski sensorji, preko katerih naprave pridobijo informacijo, kje se ljudje in objekti nahajajo. Obetajoča tehnologija na tem področju je tehnologija ultraširokega spektra (UWB - ultra-wideband [10]), ki omogoča lokalizacijo brezžičnih sensorjev ob razmeroma majhni porabi električne energije ter ob manjši ceni v primerjavi s sorodnimi tehnologijami. Njena slabost pa je predvsem šumnost podatkov. Natančnost pozicij je kritičnega pomena za potrebe kasnejše obdelave podatkov, kot so npr. prepoznavanje aktivnosti, obnašanja ter odkrivanje znakov bolezni, kar pripomore k natančnejšemu diagnosticiranju fizičnega stanja uporabnika.

V članku predstavimo metodo za glajenje signalov, zajetih z UWB tehnologijo. Signale gladimo v več korakih. Prvi korak je odstranitev izjemnih vrednosti z mediana filtrom. Sledi uporaba domenskega znanja, pri čem upoštevamo, da so vhodni podatki pozicije sensorjev, pritrjenih na človeško telo. Le-to ima določene omejitve (npr. hrbtenica se ne more prepogniti kot noga), ki jih upoštevamo pri ocenjevanju vrednosti pozicije. Zadnji korak je uporaba Kalmanovega filtra [3], s katerim poleg zanesljivejše ocene pozicije ocenimo tudi podatke, ki se uporabijo pri nadaljnji analizi podatkov, npr. hitrosti značk.

Prispevek ima naslednjo strukturo. V drugi sekciji pregledamo sorodna dela na področju odstranjevanja iz-

jemnih vrednosti, upoštevanju anatomskih omejitev in ocenjevanja parametrov na podlagi izmerjenih vrednosti. V tretji sekciji analiziramo šum in kakovost vhodnih signalov. Celoten postopek filtriranja in posamezni filtri (mediana, algebrajske omejitve in Kalmanov filter) so predstavljeni v četrti sekciji, medtem ko predhodne rezultate uporabe filtrov na realnih in sintetičnih podatkih predstavimo v peti sekciji. Prispevek se zaključi v šesti sekciji s povzetkom in diskusijo.

2 SORODNA DELA

Na področju obdelave signalov je prisotnih veliko metod, ki uspešno zmanjšujejo šumnost signalov. V tem pregledu bomo pregledali nekaj del s področja odstranjevanja izjemnih vrednosti, upoštevanja omejitev in ocenjevanja vrednosti.

Qui idr. [6] so opisali primerjavo različnih filtrov pri odstranjevanju impulzivnega šuma na podatkih, izmerjenih s sensorji v motorjih letal. Pri počasnem spreminjanju signala se je izkazalo jedrno glajenje, pri sledenju nenadnim spremembam pa je bil najboljši kaskadni rekurzivni mediana filter. Do podobnih zaključkov so prišli tudi Verma idr. [9] pri odstranjevanju izjemnih vrednosti v signalu, ki simulira tok goriva pri letalu. Pokazali so, da z mediana filtrom uspešno odstranijo izjemne vrednosti.

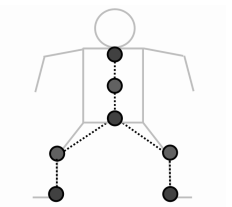
Sul idr. [7] so predstavili postopek za upoštevanje telesnih omejitev z uporabo Kalmanovega filtra. Osredotočili so se na podatke o gibanju človeškega telesa z več omejitvami, npr. translacijske in rotacijske. Te omejitve so dodane Kalmanovemu filtru kot funkcije napake, ki jih je potrebno minimizirati. S filtrom so uspešno združili različne odseke gibanja, pri čem je filter zgladil vmesne nenadne premike. Musi idr. [4] so predstavili razširjen Kalmanov filter za filtriranje podatkov iz nizkocenovnih sensorjev pri modeliranju vstajanja človeka. Za opis vstajanja so definirali kinematične enačbe, ki so združene s Kalmanovim filtrom v Razširjen Kalmanov filter. Model uspešno filtrira kote, kotne hitrosti, pospeške in momente med opazovanimi sklepi.

3 PODATKI

Vhodni podatki so pozicije množice značk, ki so nameščene na človeškem telesu. Pridobimo jih s loka-

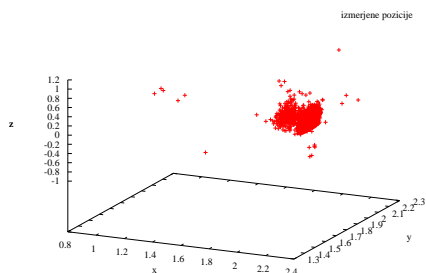
lizacijskim sistemom Ubisense [7], ki deluje na podlagi UWB tehnologije. Najvišja frekvenca zajemanja signalov je 10 Hz, pri čemer lahko istočasno zajema k večjemu signale štirih značk. Zajeti podatki v posameznem trenutku vsebujejo absolutne koordinate x , y in z ter čas lokalizacije.

Značke so pritrjene na telo na naslednjih pozicijah: levi gleženj, desni gleženj, pas in prsi, kot je prikazano na sliki 1. Z potrebe filtrov izračunamo še vmesne točke na kolenih in na trebuhu (več o tem v sekciji 4.2).



Slika 1: Pozicije značk na telesu.

Slika 2 prikazuje izmerjene koordinate x , y in z ene značke med mirovanjem na poziciji (2, 2, 0). Za potrebe kasnejše obdelave pozicij je tovrstni šum moteč.

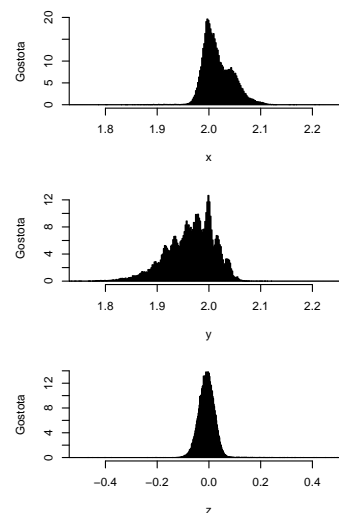


Slika 2: Izmerjene pozicije značke, ki miruje na koordinatah (2, 2, 0).

Histogrami koordinat na sliki 3 prikazujejo, da šum v podatkih ni porazdeljen z normalno porazdelitvijo. To smo pokazali tudi s Kolmogorov-Smirnov testom[2], pri čem smo primerjali vsakega izmed signalov značk s signali, pridobljenimi z normalno porazdelitvijo. Izračunana statistika D je za koordinate je $D_x = 0,2094$, $D_y = 0,2776$ in $D_z = 0,1428$, pri čem pa je mejna vrednost med $1/(65536)^{0.5} = 0,00390625 < D < 2/(65536)^{0.5} = 0,0078125$, kateri pa se nobena izmed koordinat ne približa. Ker je pri vseh koordinatah D večji od mejne vrednosti, zaključimo, da koordinate ne vsebujejo Gaussovega šuma. Dodatna analiza šuma pozicij z avtokorelacijo je pokazala, da je šum skoraj naključen oziroma ni nobenega nenaključnega procesa, ki bi vnašal šum v zajete koordinate.

4 GLAJENJE

Izsledki analize šuma dajejo slutiti, da bo potrebno kombinirati več metod za glajenje signala. V naslednjih



Slika 3: Histogrami koordinat mirujoče značke na poziciji (2, 2, 0).

podsekcijah predstavimo filtre za odstranjevanje impulzivnega šuma, upoštevanje algebrajskih omejitev med značkami in glajenje ter ocenjevanje s Kalmanovim filtrom. Predstavljeni filtri so vezani zaporedno.

4.1 Mediana

Mediana filter [5] spada v skupino nelinearnih filtrov in se pri analizi signalov in slik običajno uporablja za odstranjevanje šuma tipa *sol in paper*. Filter je sestavljen iz okna lihe dolžine $2n + 1$ zaporednih vzorcev signala. Vzorce v oknu razporedi po velikosti in kot rezultat y_k vrne srednjo vrednost – mediano:

$$y_k = \text{median}(x_{k-n}, x_{k-n+1}, \dots, x_k, \dots, x_{k+n}) \quad (1)$$

Izhod filtra ima n vzorcev zamude. Mediana filter ohranja ostre spremembe signala, učinkovito odstranjuje impulziven šum, slabo pa odpravlja Gaussov šum.

4.2 Algebrajske omejitve

Pri izločanju vpliva šuma si lahko pomagamo z anatomskimi omejitvami sklepov, kar pomeni, da morajo značke v vsakem trenutku zadovoljiti te omejitve. Postopek je sestavljen iz treh delov: (i) projekcija izmerjenih vrednosti v veljavno območje, (ii) reševanje omejitev z relaksiranjem in (iii) upoštevanje skeleta.

V prvem koraku preslikamo izmerjene vrednosti vseh značk v veljavno območje, torej območje, ki je fizično mogoče, npr. predpostavimo, da se vse meritve nahajajo v območju kvadra (npr. soba) omejenega s točkama $t_1 = (x_1, y_1, z_1)$ in $t_2 = (x_2, y_2, z_2)$. Da bi obdržali vse vrednosti znotraj veljavnega območja, je potrebno izmerjeno pozicijo p preslikati v notranjost kvadra:

$$p' = \min(\max(p, t_1), t_2) \quad (2)$$

V nadaljevanju predstavimo omejitve, ki morajo veljati med posameznimi značkami. Vzemimo znački p_1 ter

p_2 in zahtevajmo, da je razdalja med njima d , s čimer dobimo omejitev:

$$|p_1 - p_2| < d \quad (3)$$

Če omejitev (3) ni zadovoljena, je potrebno znački premakniti bodisi bližje bodisi bolj narazen v smeri premice, ki jo znački določata. Novo pozicijo značk izračunamo z izrazoma:

$$p'_1 = p_1 + \frac{1}{2}(p_2 - p_1) \left(\frac{\|p_2 - p_1\| - d}{\|p_2 - p_1\|} \right) \quad (4)$$

$$p'_2 = p_2 - \frac{1}{2}(p_2 - p_1) \left(\frac{\|p_2 - p_1\| - d}{\|p_2 - p_1\|} \right) \quad (5)$$

Človeško telo lahko obravnavamo kot skupino povezanih togih teles. Dve telesi, npr. gleženj in koleno, lahko povežemo z daljico. V našem primeru razpolagamo s štirimi značkami - dve na gležnjih, ena na pasu in ena na vratu. Takšen nabor značk ne omogoča sestave skeleta s statičnimi omejitvami (razdalja med pasom in kolonom se namreč spreminja, medtem ko je razdalja med gležnjem in kolonom konstantna), zato smo določili še tri vmesne točke - dve kolena in trebuh, kot je prikazano na sliki 1.

Določanje vmesnih točk kolena poteka po naslednjem postopku. Vzemimo pozicijo točke p_1 za središče prve krogle k_1 z radijem r_1 in pozicijo točke p_2 za središče druge krogle k_2 z radijem r_2 . Vmesno točko p_3 izračunamo kot presečišče krogel k_1 in k_2 tako, da premaknemo koordinatni sistem v središče ene izmed krogel in ga poravnamo z osjo, ki jo določata točki p_1 in p_2 . Krogli sta potem oddaljeni za razdaljo d . Enačbi krogel odštejemo in izrazimo koordinate. Rezultat se nahaja na krožnici:

$$x = \frac{d^2 - r_2^2 + r_1^2}{2d} \quad (6)$$

$$y^2 + z^2 = r_1^2 - \left(\frac{d^2 - r_2^2 + r_1^2}{2d} \right)^2 \quad (7)$$

Točna pozicija kolena nas ne zanima, temveč želimo le posredno vplivati na omejitve med točkama p_1 in p_2 . Za rešitev lahko izberemo poljubno točko na krožnici in jo preslikamo v prvotni koordinatni sistem.

Vmesno točko na trebuhu izračunamo podobno kot kolensko točko, le da za središči krogel vzamemo točki na pasu in vratu ter radije r_3 in r_4 .

Sedaj lahko predstavimo artikulirano telo kot zvezo omejitev med posameznimi sklepi, kot so s črtkanimi črtami povezani na sliki 1. Konkretno vrednosti omejitvev so odvisne od človeka do človeka, dobre približke pa se lahko pridobi s poznavanjem višine človeka. Za povezave med sklepi na sliki 1 sestavimo sistem omejitev po zgledu enačbe (3). Zadovoljitev ene omejitve povzroči kršitev druge omejitve, a teorem Gauss-Siedlove iteracije [1] nam zagotavlja, da iterativno zadovoljevanje omejitev konvergira k rešitvi. Vse omejitve zapišemo kot

linearen sistem enačb (8), matriko A pa lahko razcepimo na zgornje in spodnje trikotno matriko U in L (9) in uporabimo dekompozicijo za rešitev linearnega sistema enačb z uporabo prejšnje znane vrednosti:

$$Ax = B \quad (8)$$

$$(L + U)x = B \quad (9)$$

$$x_{t+1} = L^{-1}(B - Ux_t) \quad (10)$$

Postopek konvergira k rešitvi, če je matrika A pozitivno definitna in diagonalno dominantna. Pogoja sta zadoštna, kar pomeni, da postopek lahko konvergira tudi v primerih, ko pogoja nista izpolnjena. Število potrebnih iteracij je običajno majhno, saj metoda hitro konvergira k rešitvi že po nekaj iteracijah. Postopek je še posebej primeren za izvajanje v realnem času, saj je rešitev na voljo ob vsaki iteraciji.

Poleg omejitev med sklepi je potrebno upoštevati fizikalne omejitve udov, kot sta npr. pospešek in hitrost, zato za vsak sklep vpeljemo omejitev maksimalnega premika. Vzemimo, da je pospešek gležnja lahko največ $a \text{ m/s}^2$, kar predstavlja največji premik $x = at^2$ v časovnem intervalu t , kjer je $1/t$ frekvenca zajemanja signala. Potemtakem je pozicija premika sklepa omejena s kroglo z radijem x . V primeru, ko razdalja med prejšnjo in trenutno pozicijo presega radij x , novo pozicijo predstavimo na rob krogle v smeri nove pozicije. Za pohitritev izračunov lahko namesto krogle vzamemo očrtano kocko in uporabimo enačbo (2). Napaka je v tem primeru zanemarljiva, saj predstavlja k večjemu 15 % odstopanje (kar npr. pri $x = 10 \text{ cm}$ pomeni 1,5 cm).

4.3 Kalmanov filter

Kalmanov filter je optimalni rekurzivni linearni algoritem za pridobitev ocene stanja sistema ali iskanih količin. Vključuje vsa razpoložljiva znanja o dinamiki sistema in merskih naprav, statističnih lastnosti merskih napak, napakah sistema in dinamiki modela ter začetnih vrednostih iskanih količin. Kalmanov filter predpostavlja, da so šumi sistema in meritev beli in Gaussovi. Rešuje probleme v zvezi z manjkajočo, nepopolno ali moteno informacijo.

Treutno vrednost stanja x_k oceni glede na prejšnje stanje sistema:

$$x_k = F_k x_{k-1} + B_k u_k + w_k, \quad (11)$$

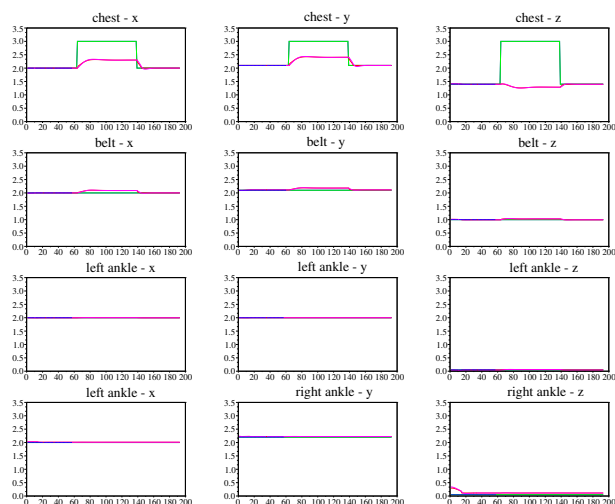
kjer je F_k matrika prehodov, B_k kontrolna matrika in w_k matrika šuma. V našem primeru želimo popraviti izmerjeno lokacijo značke (x, y, z) in zanesljivo oceniti hitrost (v_x, v_y, v_z) . Matrike B_k v našem primeru ne potrebujemo, matriko w_k ocenimo iz podatkov o mirujoči znački, pri matriki F_k pa upoštevamo enostaven dinamičen sistem:

$$x(t) = x(t-1) + dx(t-1) = x(t-1) + vx * dt \quad (12)$$

$$vx(t) = vx(t-1) = dx(t-1) \quad (13)$$

5 MERITVE

Delovanje filtrov smo preizkusili na dveh signalih. Najprej smo generirali sintetični signal za štiri mirujoče značke na gležnjih, pasu in vratu. V nekem trenutku smo signal pokvarili tako, da smo značko na vratu premaknili v izjemno pozicijo. Na sliki 4 so prikazane x , y in z koordinate vseh značk. Z modro krivuljo je označen sintetični signal, izhod po filtru mediana je zelene barve, izhod po upoštevanju algebraskih omejitev je rdeče barve in korekcija po Kalmanovem filtru je rožnata. Iz slike je razvidno, da filter mediana zaradi daljšega intervala sledi intenzivni spremembi signala, medtem ko upoštevanje algebraskih omejitev konzervativno sledi premiku točke na vratu, posledično pa premakne tudi značko na pasu. S tem smo vnesli nekoliko šuma v signal značke na vratu, a učinek šuma signala značke na vratu je bil precej zmanjšan. Na znački na gležnjih ni zaznati vpliva.

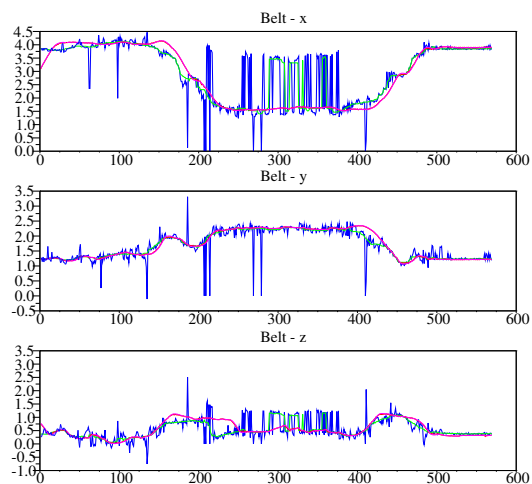


Slika 4: Odziv filtrov na motnjo v sintetičnem signalu. Značka na pasu se je v nekem trenutku impulzivno premaknila na novo pozicijo in tam ostala.

Za drugi poskus smo zajeli realne signale gibanja človeka pri opravljanju naslednjih aktivnosti: hoja, usedanj, leganje, padec, vstajanje in sedenje. Na sliki 5 so prikazane x , y in z komponente signala značke na pasu. Z modro krivuljo je označen zajeti signal, z zeleno izhod po mediani, z rdečo izhod po upoštevanju algebraskih omejitev in z rožnato korekcija s Kalmanovim filtrom. Iz slike 5 je razvidno, da filter mediana dobro odstrani impulziven šum, če le-tega ni preveč. V slednjem primeru se izkaže upoštevanje algebraskih omejitev, saj kljub zelo šumnemu signalu ne sledi izjemnim vrednostim. Slabost slednjega je opazna zamuda glede na filter mediana. Manjše oscilacije so na koncu povsem zglajene s Kalmanovim filtrom.

6 ZAKLJUČEK

V prispevku smo v treh korakih predstavili postopek za glajenje trajektorij značk. V prvem koraku s filtrom me-



Slika 5: Glajenje realnega signala značke na pasu.

diana odstranimo izjemne vrednosti, v drugem koraku zadovoljimo algebrasko omejitve med značkami in v tretjem koraku s Kalmanovim filtrom ocenimo pozicijo in hitrost. Kombinacija filtrov uspešno zmanjša šumnost signalov v predhodnih meritvah.

Tako algebrasko omejitve, ki pomagajo izboljšati model telesa, kot tudi mediana in Kalmanov filter, odstranijo detajle v signalu, ki so morda pomembne pri nadaljnji obdelavi. Telo je še vedno v veljavnih pozicijah, vendar je gibanje zglajeno. Pri iskanju parametrov je zato potrebno poiskati kompromis med stopnjo detajlov in količino šuma.

Literatura

- [1] R. Barrett et al. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*, 2nd edition, Philadelphia, 1994.
- [2] M. Chakravarti, R. G. Laha, J. Roy. *Handbook of Methods of Applied Statistics (1)*. John Wiley and Sons. 1967.
- [3] R. E. Kalman. *A New Approach to Linear Filtering and Prediction Problems*. Transaction of the ASME - Journal of Basic Engineering, št. 35, str. 35–45, 1960.
- [4] J. Musič, R. Kamnik, M. Munih. *Model based inertial sensing of human body motion kinematics in sit-to-stand movement*. Simulation Modelling Practice and Theory, vol. 16, str. 933–944, 2008.
- [5] R. K. Pearson. *Mining Imperfect Data: Dealing with Contamination and Incomplete Records*. Philadelphia, 2005.
- [6] H. Qiu, N. Eklund, N. Iyer, X. Hu. *Evaluation of Filtering Techniques for Aircraft Engine Condition Monitoring and Diagnostics*. Proc. of the International Conference on Prognostics and Health Management, str. 1–8, 2008.
- [7] C. W. Sul, S. K. Jung, K. Wohn. *Synthesis of Human Motion using Kalman Filter*. Proc. of the international Workshop on Modelling and Motion Capture Techniques For Virtual Environments, vol. 1537, str. 100–112, 1998.
- [8] Ubisens. *Real-time Location System*. <http://www.ubisense.net>, 12-09-2009.
- [9] R. Verma, R. Ganguli. *Denoising Jet Engine Gas Path Measurements Using Nonlinear Filters*. IEEE/ASME Transactions on Mechatronics, 10:(4), str. 461–464, 2005.
- [10] Wikipedia. *Ultra-wideband*. <http://en.wikipedia.org/wiki/Ultra-wideband>, 12-09-2009.

ISKANJE VZORCEV V ZAPOREDJU DOGODKOV

Jana Krivec

Odsek za inteligentne sisteme

Jozef Stefan Institute

Jamova 39, 1000 Ljubljana, Slovenia

e-mail: jana.krivec@ijs.si

ABSTRAKT

V življenju imamo velikokrat opraviti z učenjem zaporednih dogodkov. Procesiranje zaporednih dogodkov je eden izmed osnovnih procesov vključen v mnoge človekove kognitivne sposobnosti, ki jih potrebujemo za uspešno interakcijo z neprestano spreminjajočim se okoljem. Zaradi lažjega zapomnjevanja in znajdenja v okolju si človek zaporedne informacije združuje v smiselne vzorce. V prispevku so predstavljene metode, s katerimi si lahko pomagamo pri obdelavi podatkov zaporednega procesiranja in odkrivamo človeku smiselne, tipične vzorce v gradivu smiselnih zaporednih dogodkov. Predstavljen je primer identifikacije vzorcev v zaporednih potezah šahovske partije.

1 UVOD

Nenehna interakcija z dinamičnim okoljem zahteva od človekovega kognitivnega sistema posebno občutljivost za zaporedno strukturo dražljajev zunanjega sveta. Ena bistvenih spretnosti človekovega procesiranja zaporednih informacij je združevanje posameznih informacij v smiselno povezane enote, ki so med seboj povezane in se hkrati razlikujejo od drugih informacij (skupke). Ta spretnost se lepo kaže pri uporabi jezika, kjer osnovno asociativno sekvenčno učenje prihaja do velikega izraza. Saffran in drugi [6] so denimo pokazali, da izpostavljanje umetnemu, jeziku podobnemu zvoku (npr. *bupadapatubitutu...*) pri poslušalcih vzpodbudi, da zvoke oblikujejo v nekakšne umetne besede (e.g. *bupada*, *patubi*, etc.). Na osnovi teh podatkov je Saffran et al. [6] predvideval, da se sposobnost besedne členitve razvije s pomočjo mehanizmov, ki uporabljajo statistične zakonitosti prisotne v zaporedju dogodkov, kot npr. da je verjetnost zaporednih zlogov višja znotraj posamezne besede kot med besedami. Cleeremans and Jiménez [1] sta zaključila, da je ključni proces vpleten v učenje zaporednih dražljajev, vpeljevanje asociativnih učnih procesov, ki privedejo do izboljšane občutljivosti statističnih značilnosti, ki jih učni material vključuje. V tem kontekstu je večno zanimivo vprašanje, kaj je tisto, kar se ljudje naučijo, ko so izpostavljeni učenju materiala z zaporedno strukturiranimi dražljaji?

V našem primeru smo ugotavljali, ali obstajajo tipični vzorci oz. statistične zakonitosti v zaporedju šahovskih potez oz. variant (skupkov) in jih poskušali identificirati.

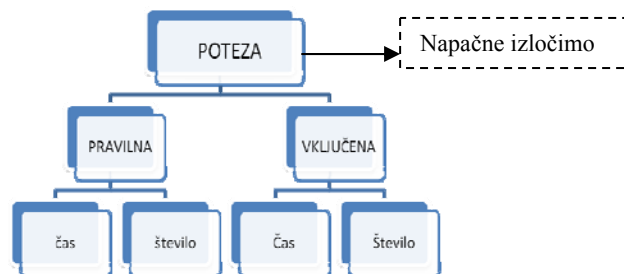
Predstavili bomo metode, ki smo jih uporabili za identifikacijo takšnih vzorcev.

2 NALOGA REKONSTRUKCIJE

Za dokazovanje in odkrivanje vzorcev v zaporedju potez smo uporabili naloge rekonstrukcije petih šahovskih variant. Iz izbrane osnovne pozicije, smo časovnim razmakom 2 sekund med potezami, prikazali 16 potez (32 polpotez) dolgo šahovsko varianto. Udeleženci so morali po prezentaciji varianto iz osnovne pozicije čim bolje in čim hitreje rekonstruirati. Ob tem smo merili čas med rekonstrukcijo zaporednih potez. V psihologiji se človekove kognitivne procese velikokrat meri in raziskuje s pomočjo nalog reakcijskih časov. Metodo rekonstrukcije smo izbrali, ker temelji na priklicu informacij iz dolgoročnega spomina (zajame proceduralno znanje). V raziskavi je sodelovalo 52 šahistov, starih od 14 do 56 let ($M=14$). Njihova povprečna šahovska moč (merjena z mednarodnim »ELO« sistemom) je bila 2275 ELO točk (varirala je od 1600, ki predstavlja nivo novinca, pa do 2670 ELO, ki je rating najmočnejših velemejstrov).

3 OBDELAVA PODATKOV

V obdelavo rezultatov smo vključili podatke o pravilnosti in časovnem intervalu rekonstrukcije posamezne poteze (glej Sliko 1). Pravilna poteza je tista, ki je rekonstruirana povsem točno (v tem primeru je število (Slika 1) =1, sicer pa 0), vključena pa tista, ki je sicer rekonstruirana, a na napačnem mestu ali po predhodni napaki v rekonstrukciji. Čas predstavlja interval med rekonstrukcijo dveh zaporednih potez.



Slika 1. Podatki, ki jih imamo za vsako rekonstruirano potezo.

Pri identifikaciji skupkov iz šahovskih variant smo na zgoraj omenjenih podatkih uporabili različne metode.

3.1 Metode s časovnim kriterijem

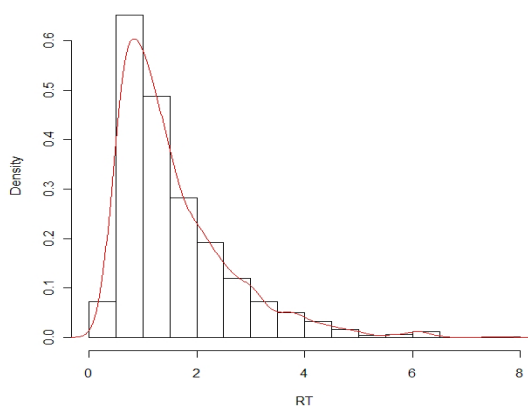
Najprej smo si pomagali s principom metode, ki sta jo razvila Chase in Simon [7] za izolacijo vzorcev v šahovskih poziciji. Ta temelji na predpostavki, da so daljši časovni premori v reakcijskih časih pri rekonstrukciji smiselnega gradiva (to je gradiva o katerem imamo določeno predznanje), povezani s priklicem nove strukture (skupka) iz spomina. V našem primeru smo dodatno predvidevali, da krajši reakcijski časi predstavljajo najbolj tipične predstavnike oz. poteze v vzorcu.

V reakcijskih časih odgovorov oz. rekonstruiranih potez smo iskali dve vrsti izstopajočih vrednosti:

A* - pri posameznih udeležencih smo iskali tiste čase, ki so pomembno daljši ali krajši od večine časovnih intervalov, ki jih je posameznik za svoje odgovore porabil. S tem smo pri vsakem posamezniku ugotovili, katera je tista poteza, ki zanj predstavlja nov skupek. Nadalje smo izvajali t.i. kolektivno rekonstrukcijo, kjer smo glede na frekvenčno porazdelitev pogostosti poteze ki predstavlja začetek novega vzorca pri vseh udeležencih, ugotavljali katere poteze najverjetneje veljajo za začetek novega skupka.

B** - za posamezne poteze smo iskali tiste povprečne čase vseh udeležencev, ki pomembno odstopajo od povprečnih časov, ki so jih udeleženci v povprečju porabili za ostale poteze.

Reakcijski časi imajo svoje tipične lastnosti, ki jih je treba upoštevati pri njihovi obdelavi. Surovi reakcijski časi so v povprečju skoraj vedno nagnjeni k zgornjemu delu zaradi posameznih daljših časov (outlierjev). Krivulja porazdelitve časov je zato navadno desno asimetrična (Glej Slika 2). To pomeni da distribucija najprej strmo naraste, nato pa počasi pada.



Slika 2. Histogram tipične razoreditve časovnih intervalov odgovorov.

Desno nagnjeno razporeditev reakcijskih časov pri človeku je zabeležilo več empiričnih raziskav (Heathcoate, Popiel & Mewhort, 1991; Hockley, 1982, 1984; Hohle, 1965; Ratcliff, 1978, 1979, 1981, 1988a, b; Ratcliff & Murdock, 1976; Ulrich & Miller, 1992, po [4]). Takšno razporejanje podatkov

najbolje opišejo eksponentne krivulje, med katere spadajo: Ex-Gausova (e.g., Balota et al., 2008; Hohle, 1965; McGill, 1963; Ratcli_ & Murdock, 1976; Ratcli_, 1978), the Ex-Wald (Schwarz, 2001), the Inverzna Gausova (e.g., Lamming, 1968; Stone, 1960), Gamma (e.g., Christie, 1952; Luce, 1960; McGill, 1963) (po [4]). To moramo upoštevati predvsem pri računanju srednjih vrednosti podatkov in njihovih variabilnosti.

Detekcija odstopajočih reakcijskih časov (ang. »outlierjev«) je izredno težavna naloga.

Obstaja več metod detekcije reakcijskih časov:

- Metode »Cenzuriranja podatkov«: rezanje od določenih vrednosti naprej, porezanje določenega procenta odgovorov, rezanje glede na standardni odklon (+1 SD, +2 SD),
- Transformacijske metode: logaritemska transformacija, inverzna transformacija

Vedno pa naj bi obstajala vrednost optimalnega rezanja. Lokacija le-te je odvisna od spreminjanja oblike distribucije, glede na spreminjanje povprečnega reakcijskega časa. Ko je variabilnost okrog mediane določene osebe nizka v primerjavi s standardnim odklonom celotne distribucije, potem je vrednost inverzne transformacije (1/R) blizu optimalne točke rezanja. V primeru, ko je variabilnost med medianami oseb velika, se je kot učinkovita metoda rezanja oz. identificiranja odklonskih vrednosti izkazalo rezanje po standardnih odklonih (+1 ali 2 SD) [5]. V našem primeru gre za veliko variabilnost med medianami posameznih oseb ali medianami posameznih potez, zato smo se odločili za detekcijo s pomočjo rezanja po standardnih odklonih.

Postopek je naslednji:

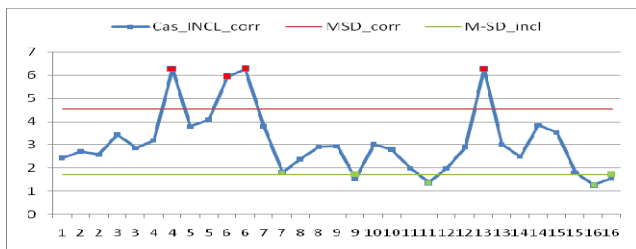
*- Najprej smo za vsakega posameznika izračunali srednjo vrednost niza podatkov. Teoretične raziskave so potrdile [3],[4], da je v primeru, ko podatki vsebujejo izstopajoče vrednosti najbolj »odporna« mera mediana, ki smo jo uporabli tudi mi. Nadalje smo izračunali standardno deviacijo časov posameznih oseb. Vse čase, ki so bili višji od $Me+1SD$ smo vzeli kot indikatorje novega vzorca ali težke poteze, vse čase, ki so bili nižji od $Me-1SD$ pa kot indikatorje tipičnih/lahkih potez. Po teoretičnih predpostavkah je proti izstopajočim vrednostim sicer najbolj odporna mera razpršenosti »kvartil« in »Pearson 2« kot mera nagnjenosti distribucije [5]. Morda bi bilo torej smoterno namesto pogoja $\pm 1SD$ upoštevati pogoj $Q1, Q3$. Vsaka identificirana izstopajoča vrednost je dodala pripadajoči potezi vrednost 1. Sledil je prikaz frekvenčnih porazdelitev omenjenih vrednosti po vseh preizkušancih. Zgornjih 25% potez ($>Q3$) se smatra za poteze, ki dejansko predstavljajo nov skupek oz. zelo težko potezo.

** - Pri iskanju izstopajočih časov rekonstrukcije posameznih potez je postopek nekoliko drugačen. Zaradi različnega osnovnega tempa rekonstrukcije različnih udeležencev najprej izračunamo delež svojega celotnega časa rekonstrukcije, ki ga za določeno potezo porabil

posamezen udeleženec (relativen reakcijski čas), po naslednji formuli:

$$f(x) = (t(x) / t(\text{skupen})) * \text{število potez}$$

Nato za vasko potezo izračunamo povprečni relati ven reakcijski čas vseh udeležencev (mediano). Mediano poteze vzamemo kot nov podatek pri računanju mediane povprečnih časov vseh potez v varianti (mediano median potez). Ponovno sledimo metodi rezanja po SD (+1) da ugotovimo katere so izstopajoče vrednosti oz. izstopajoče poteze.

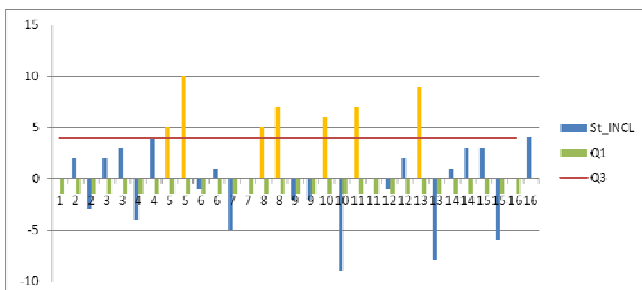


Slika 3. Primer razporejanja povprečnih reakcijskih časov in identifikacije izstopajočih vrednosti z metodo rezanja po standardnih odklonih (1SD).

Is Slike 3 vidimo, da imajo poteze 4.črnega, 6.belega in črnega in 13.belega v povprečju pomembno daljše reakcijske čase od ostalih potez, kar pomeni da začenjajo nov skupek ali da so precej težje poteze od ostalih. Nasprotno imajo poteze 7.črnega, 9.črnega 11.črnega in 16.belega in črnega pomembno krajše čase in predstavljajo tipične poteze v skupku oz. varianti.

3.2 Metoda s kriterijem uspešnosti

V naslednji metodi smo uporabili podatke o pravilnosti/vključenosti potez. Ta metoda bazira na predpostavki, da večji upad v frekvenci pravilno rekonstruiranih ali vključenih potez predstavlja začetek novega skupka oz. zelo težko potezo, velik porast pa zelo tipično ali lahko potezo. Postopek je naslednji: Pri vseh udeležencih skupaj za vsako potezo izračunamo frekvenco pravilno rekonstruiranih in vključenih potez. Nato izračunamo spremembo v frekvencah med zaporednima potezama. Iz sprememb izračunamo Q1 in Q3. Poteze katerih frekvence so višje od Q3 so začetek novega skupka ali težje, tiste, ki so nižje od Q1 pa so tipične ali lahke poteze.



Slika 4. Primer frekvenčne distribucije rekonstruiranih vključenih potez in identifikacije izstopajočih vrednosti z metodo rezanja po kvartilih (Q1 in Q3).

Za končno odločitev o skupku na osnovi reakcijskih časov in uspešnosti rekonstrukcije smo dodali zahtevo ujemanja vsaj pri dveh kombinacijah pogojev PRAVILNA/VKLJUČENA in ČAS/ŠTEVILO.

3.3 Metoda z asociacijskimi pravili

Naslednja metoda odkrivanja vzorcev v zaporednih dogodkih je osnovana na principu asociacijskih oz. povezovalnih pravil. Asociacijska pravila oblike $X \rightarrow Y$, kjer sta X in Y množici postavk (items) in predstavljajo algoritem strojnega učenja za odkrivanje podatkov oz. stvari, ki se navadno zgodijo/pojavljajo vzajemno. V našem primeru torej iščemo poteze, ki so navadno rekonstruirane vzajemno (pravilne ali vključene).

Osnovna baza na kateri se gradi pravila izgleda takole:

Tabela1. Primer podatkovne tabela pri analizi z metodo asociacijskih pravil.

Osoba	1Lg3	1Le7	2Ld3	20-0	38d2	3Te8	40-0	4Sh5	5Sc4	5Lf6	6Se3	6Sf8	7de5	7de5
1	2	2	2	2	2	2	2	2	2	2	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	1	1	1	1
3	2	1	1	1	1	0	1	1	1	1	1	0	1	1
4	2	2	1	1	1	1	1	1	0	0	0	0	1	1
5	2	2	2	2	2	2	2	1	1	0	1	0	1	1
6	2	2	2	2	2	0	1	0	0	0	0	0	0	0
7	2	2	2	2	1	1	1	1	0	0	0	0	0	0

Legenda: 2...poteza je pravilno rekonstruirana
1...poteza je "vključena"
0...poteza je napačno rekonstruirana

Vsaka vrstica v Tabeli 1 predstavlja rekonstrukcijsko uspešnost posameznega udeleženca pri določeni potezi (stolpec). Kar nas zanima je, katere poteze je večina udeležencev navadno vzajemno pravilno rekonstruirala.

V ta namen smo na podatkih izvedli Ridor algoritem [8], pri maksimalnem številu pravil 1000. Kot prvi rezultat smo vzeli najdaljše pravilo. V primeru, da sta bili dve pravili enako dolgi smo vzeli tisto, ki je imelo večje zaupanje in večjo pokritost (število primerov za katere pravilo velja). Iz pravila smo izločili poteze (če so), ki niso bile pravilno rekonstruirane. Zaporedne poteze v pravilu predstavljajo skupek. Nato smo iz baze podatkov izločili poteze, ki smo jih izluščili iz prvega pravila in ponovili postopek na preostalih podatkih. Postopek smo ponavljali dokler nismo izločili vseh potez.

Primer:

4Kc5=2 136 ==> 4Th6=2 5Tg6=2 5Th4=2 6Tg8=2 6Tg4=2 136 [conf:\(1\)](#)
62. 4Kc5=2 136 ==> 5Tg6=2 5Th4=2 6Tg8=2 6Tg4=2 136 [conf:\(1\)](#)
-> SKUPEK: Th6, Kc5, Tg6, Th4, Tg8, Tg4

Problem te metode je lahko v tem, da lahko slabo prepozna meje med dvema zaporednima skupkoma.

3.4 Metoda z uspešnostjo rekonstrukcije pozicij

Ta metoda je še v idejni fazi. Predpostavlja, da so skupki poteze od ene do druge ključne pozicije. Ključne pozicije dobimo z naslednjim postopkom: vsaki poziciji v originalni/prkazani variant, ki jo posameznik obnovi v svoji rekonstrukciji variante dodelimo 1 točko. Tako

dobimo frekvenčno poradelitev posameznih pozicij v varianti. Pozicije, ki imajo frekvenco rekonstrukcij višjo od tretjega kvartila (zgornjih 25% pozicij) smaramo kot "ključne pozicije", poteze od ene do druge ključne pozicije pa kot skupek.

4 PRIMER IDENTIFICIRANEGA SKUPKA

Ker je raziskava še v teku vam predstavimo le preliminaren primer identificiranega skupka.

V eni izmed variant smo po opisanih metodah z reakcijskimi časi in uspešnostjo rekonstrukcije izločili izstopajoče poteze. Dobili smo naslednje kombinacije, kjer je vsaj v dveh kriterijih hkrati prisotna določena poteza:

(CORR-cas \wedge CORR-st) \vee (INCL-cas \wedge INCL-st): **Th6, Kb6 v Ke5**

(CORR-cas \wedge INCL-cas) \vee (CORR-st \wedge INCL-st): **Th6, Kd3, Tg5 v Kb6, Ke2, Ke5**

(CORR-cas \wedge INCL-st) \vee (INCL-cas \wedge CORR-st): **Kb6, Th6, Ke5**

V končne meje skupkov vzamemo naslednje poteze:

Th6, Kb6, Ke5, Kd3, Tg5, Ke2

Prvi skupek je tako do poteze Th6, drugi od vključno poteze Th6 do poteze Kb6 (Glej Sliko 5)....



Slika 5. Primer tipičnega vzorca v proceduralnem skupku.

Na sliki 3 lahko vidimo kako dejansko izgleda skupek v šahovski varianti: Poteze 1.Th6, Kc5 2.Tg6 Th4 2.Tg8 Tg4 3.g6 (1-7 na Sliki 5) naj bi bile med seboj močnejše (tudi pomensko) povezane in tvorijo skupek oz.v šahovskemu žargonu motiv.

Asociacijska pravila so v tem primeru podala podobne meje skupka (glej Primer v odseku 3.3): **SKUPEK=1.Th6, Kc5 2.Tg6 Th4 2.Tg8 Tg4**. Iz skupka so izpustila potezo 3.g6.

5 IDEJE ZA PRIHODNJE DELO IN IZBOLJŠAVE

5.1 Korekcije

• Na osnovi raziskav v okviru Template teorije [2] ter preliminarne študije predvidevamo, da obstajajo razlike v velikosti in številu skupkov pri različno dobrih šahistih, zato

nameravamo preveriti ali razlike dejansko obstajajo in moramo zato udeležence najprej razdeliti po skupinah glede na ELO in šele nato izvajati opisane metode.

• Preveriti bi bilo potrebno, ali se reakcijski časi pomembno razlikujejo med različnimi deli partije/ med različnimi variantami. V primeru da se to potrdi bi bilo dobro posamezne dela partije posebej obravnavati, saj bi s tem zmanjšali vpliv daljših časov (predvidoma ob koncu variante), ki bi lahko vplivali na to, da se izstopajoče vrednosti v sicer hitrejših reakcijskih časih prvega dela ne bi identificirale.

• V zadnjih delih variante se pogosto zgodi, da samo še par preizkušancev pravilno ali sploh rekonstruira poteze, zato so lahko reakcijski časi precej popačeni. Morda bi bilo potrebno izločiti poteze od takrat naprej, ko sta ostala samo dva udeleženca.

5.2 Ideje za prihodnje delo:

• V šahovskih variantah je včasih težko ločiti ali je skupek povezava belih in črnih potez ali je značilen skupek le za poteze figur določene barve. V prihodnjem delu želimo gledati skupke še posebej za bele in posebej za črne. Na ta način bi lahko primerjali kateri skupek je močnejši (bolj povezan), "mešani" ali "ločen" za posamezne figure.

• Preveriti želimo, kakšne so korelacije med rezultati različnih metod.

• Preverili bomo ali se reakcijski časi in skupki razlikujejo po spolu in starosti?

Literatura

- [1] Cleeremans, A. & Jiménez, L. (1998). Implicit sequence learning: The truth is in the details. In M.A. Stadler & P.A. Frensch (Eds.), *Handbook of Implicit Learning*. Thousand Oaks, CA: Sage Publications.
- [2] Gobet, F., & Jackson, S. (2002). In search of templates. *Cognitive Systems Research*, vol.3, pp. 35-44.
- [3] Miller, J. (1988). A warning about median reaction time. *Journal of Experimental Psychology: Human Perception and Performance*, vol. 14, pp. 539-543.
- [4] Moscoso del Prado Martin, Dr Fermin (2008) *A Theory of Reaction Time Distributions*. (Unpublished) : <http://cogprints.org/6310/>
- [5] Ratcliff, R. (1993)"Methods for dealing with reaction time outliers". *Psychological Bulletin*, vol. 114, pp.510-532.
- [6] Saffran, J.R., Newport, E.L., Aslin, R.N., Tunick, R.A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, 8, 101-105.
- [7] Simon, H. A., Chase, W. G. (1973). "Perception in chess" *Cognitive Psychology*, vol.4, pp.55-81.
- [8] I. H. Witten, E., Frank (2005). "Data Mining – Practical Machine Learning Tools and Techniques (sec. ed.)," Morgan Kaufmann.

IZBOLJŠEVANJE PREPOZNAVANJA AKTIVNOSTI IZ POLOŽAJEV ZNAČK

Mitja Luštrek

Odsek za inteligentne sisteme

Institut Jožef Stefan

Jamova cesta 39, 1000 Ljubljana, Slovenija

Telefon: +386 1 4773380; telefaks: +386 1 4773131

E-pošta: mitja.lustrek@ijs.si

POVZETEK

Iz koordinat štirih značk, pritrjenih na telo, smo s strojnim učenjem prepoznavali aktivnosti. Koordinate smo pridobili s sistemom za določanje položaja Ubisense. Ugotovili smo, da je najprimernejši algoritem za strojno učenje naključni gozd in da je atributni vektor najbolje sestaviti iz niza atributov, ki pripadajo desetim zaporednim zajemom koordinat značk. Glavni prispevek so štirje načini za izboljševanje prepoznavanja aktivnosti, ki upoštevajo časovno zaporedje prepoznav. Z njimi se je klasifikacijska točnost povečala za 1,66 odstotne točke.

1 UVOD

Prebivalstvo v Evropi in razvitem svetu nasploh se naglo stara. To pomeni, da je starejših, ki so potrebni skrbi, vedno več, mladih, ki bi zanje lahko skrbeli, pa vedno manj. Zato je veliko raziskav usmerjenih v iskanje tehnoloških pripomočkov za skrb za starejše. Eden izmed takih pripomočkov se razvija v projektu Confidence [1].

Cilj projekta Confidence je starejše osebe opremiti z radijskimi značkami, s pomočjo katerih bo moč zaznavati položaj telesa in na podlagi tega ugotavljati zdravstvene težave, posebej padce. V prispevku se ukvarjamo s prepoznavanjem aktivnosti, ki je predpogoj za nadaljnjo analizo obnašanja uporabnika.

Ker so značke in senzorji, ki bodo uporabljeni v projektu, še v razvoju, smo namesto njih uporabili sistem za določanje položaja Ubisense [7]. Ta sistem uporablja radijsko tehnologijo in je najboljši komercialni približek strojni opremi, ki bo razvita v projektu. V preteklih poizkusih [3][4][5] smo uporabljali infrardeči sistem Smart [2], katerega izmerkom smo zaradi bistveno večje natančnosti od pričakovane v Confidence dodajali Gaussov šum. Takšni podatki so bili manj realistični od Ubisensovih.

Aktivnosti smo prepoznavali s strojnim učenjem. V prispevku najprej primerjamo klasifikacijsko točnost različnih algoritmov za strojno učenje na Ubisensovih podatkih in načine sestavljanja atributnega vektorja. Nato se posvetimo glavni temi prispevka, to je načinom za izboljševanje prepoznavanja aktivnosti, ki prepoznane aktivnosti popravlja na podlagi njihovega zaporedja v času.

2 UČNI IN TESTNI PODATKI

Za učne in testne podatke smo uporabili 25 posnetkov obnašanja petih oseb, po pet za osebo. Posnetki so sestavljeni iz 32 aktivnosti v osmih sklopih. Prepoznavali smo osem različnih aktivnosti: hoja, normalno sedenje, sedenje na tleh, ležanje, na vseh štirih, normalno spuščanje (sedanje in leganje), dviganje (iz sedenja in ležanje) ter padanje. Sestava posnetkov je prikazana v Tabeli 1.

Sklop	Aktivnosti
Hoja	hoja
Normalno sedanje	spuščanje, sedenje, dviganje
	hoja
Spotikanje	padanje, ležanje
	dviganje, hoja
Normalno leganje	spuščanje, ležanje, dviganje
	hoja
Omedlavanje stoje	padanje, ležanje
	dviganje, hoja
Hitro sedanje	spuščanje, sedenje
Zdrs s stola	padanje, sedenje na tleh
	dviganje, hoja
Hitro leganje	spuščanje, ležanje, dviganje
	hoja
Iskanje predmeta pod mizo	spuščanje, na vseh štirih, ležanje
	dviganje, hoja

Tabela 1. Sestava učnih/testnih posnetkov.

Osebe na posnetkih so bile opremljene z štirimi značkami: na obeh stopalih, na pasu in za vratom. Koordinate smo zajemali s frekvenco 9,25 Hz. Predprocesirane so bile z medianinim in Kalmanovim filtrom [6].

3 OSNOVNO PREPOZNAVANJE AKTIVNOSTI

Za prepoznavanje aktivnosti s strojnim učenjem moramo iz koordinat značk izračunati atributni vektor. Ob vsakem zajemu koordinat iz njih izračunamo nabor atributov, nakar več zaporednih naborov združimo v atributni vektor. To storimo z metodo drsečega okna, kar pomeni, da se vsak atributni vektor od prejšnjega razlikuje za en nabor.

V preteklosti [5] smo uporabljali več različnih vrst naborov atributov: attribute v referenčnem koordinatnem sistemu, attribute v štirih različnih telesnih koordinatnih sistemih in kote med deli telesa. A glede na to, da so se referenčni atributi obnesli najboljše in da je ostale z le štirimi značkami težko izračunati (v preteklosti smo uporabljali do 12 značk), smo se tokrat omejili na referenčne. Ti so bili:

- koordinate z vseh značk,
- hitrosti vseh značk,
- hitrosti vseh značk v smeri z,
- razdalje med vsemi pari značk,
- razdalje med vsemi pari značk v smeri z,
- razdalje med vsemi pari značk v ravnini xy.

Koordinat x in y nismo uporabili, ker so odvisne od mesta v prostoru, kjer se aktivnost dogaja, to pa nas ne zanima.

3.1 Algoritmi za strojno učenje

Najprej smo za Ubisensove podatke ugotovili, kateri algoritem za strojno učenje je za prepoznavanju aktivnosti najprimernejši. Pri tem smo se oprli na dognanja iz poizkusov s Smartom [5]. Attribute smo v atributne vektorje združevali tako, da smo deset zaporednih naborov atributov nanizali enega za drugega. Od algoritmov za strojno učenje smo primerjali tistih pet, ki so se najboljše obnesli na Smartovih podatkih. Poizkuse smo pognali v paketu za strojno učenje Weka [8] s privzetimi nastavitvami. Klasifikacijsko točnost smo dobili z desetkratnim prečnim preverjanjem na podatkih, opisanih v razdelku 2. Rezultati so prikazani v Tabeli 2. Vidimo lahko, da največjo točnost da naključni gozd, ostale metode z izjemo podpornih vektorjev pa so blizu. Zanimivo je, da so podporni vektorji, ki so se s Smartovimi podatki obnesli najboljše, tu najslabši.

Algoritem	Klasifikacijska točnost
3-najbližji sosedi	95,5
podporni vektorji	87,4
naključni gozd	96,7
“bagging”	94,2
“boosting” Adaboost M1	94,1

Tabela 2. Klasifikacijska točnost [%] z različnimi algoritmi za strojno učenje.

Ker je cilj projekta Confidence izdelati sistem, ki bo deloval na novem uporabniku brez prehodnega učenja, smo izračunali še klasifikacijsko točnost na način »izpusti eno osebo«. To pomeni, da smo za učne podatke uporabili posnetke štirih oseb, za testne pa posnetke pete, kar smo ponovili za vseh pet oseb. Klasifikacijska točnost z algoritmom naključni gozd je bila v tem primeru 81,5 %. Razlog za tolikšno zmanjšanje v primerjavi s prečnim preverjanjem je, da pri slednjem odseki iste aktivnosti (ki so si zelo podobni) pridejo med učne in testne podatke.

3.2 Združevanje atributov v atributne vektorje

Pri združevanju naborov atributov znotraj drsečega okna v atributne vektorje je mogoče spreminjati število naborov in

način združevanja. Uporabili smo od enega do deset naborov. Več ni smiselno, saj deset naborov obsega že 1,08 s, nekatere aktivnosti pa so krajše od ene sekunde.

Uporabili smo tri načine združevanja naborov atributov v atributne vektorje. Prvi način je že omenjeno nizanje. Drugi je povprečenje vsakega atributa čez vse nabore znotraj drsečega okna. Tretji pa je izračun naklona linearne interpolacije vsakega atributa čez vse nabore znotraj okna; tega smo uporabili le skupaj s povprečenjem.

Primerjava različnih števil naborov atributov in načinov združevanja je prikazana v tabeli 3. Kot algoritem za strojno učenje je bil uporabljen naključni gozd, rezultati pa so dobljeni z desetkratnim prečnim preverjanjem. Vidimo, da se več naborov vedno obnese bolje kot manj in da je najboljši način združevanja najpreprostejši: nizanje.

Združevanje \ št. naborov	1	3	5	10
Nizanje	96,0	96,1	96,3	96,7
Povprečenje		96,1	96,2	96,5
Povprečenje + naklon	/	95,3	96,0	96,5

Tabela 3. Klasifikacijska točnost [%] z različnimi načini združevanja atributov v atributne vektorje.

Ker so si klasifikacijske točnosti vseh treh načinov združevanja podobne, smo jih pri desetih naborih atributov primerjali še na način »izpusti eno osebo«. Pri nizanju smo dobili točnost 81,5 %, povprečenje in povprečenje + naklon pa sta bila z 81,0 % in 80,9 % tudi tu slabša.

4 IZBOLJŠANO PREPOZNAVANJE AKTIVNOSTI

Zamislimo si, da klasifikator za prepoznavanje aktivnosti kot aktivnost desetkrat zapored prepozna hojo. Če temu enkrat sledi spuščanje, nismo zelo gotovi, da je do spuščanja res prišlo. Če mu sledi več spuščanj, smo pa dokaj prepričani, da se uporabnik res spušča. A če po drugi strani nemu spuščanju sledi več hoj, se zdi bolj verjetno, da je bilo spuščanje prepoznano napačno. Takisto moramo posumiti, da je prišlo do napake, če hoji sledi ležanje, saj bi vmes moralo biti spuščanje ali padanje. Zato si želimo informacije o zaporednih prepoznanih aktivnosti izkoristiti za izboljšanje prepoznavanja aktivnosti.

4.1 Opis izboljšav

Prepoznane aktivnosti opazujemo v oknu dolžine W . Naj bodo $[a_1^p, \dots, a_W^p]$ prepoznane aktivnosti in $[a_1, \dots, a_W]$ pripadajoče prave aktivnosti. Privzemimo, da v vsakem oknu lahko nastopata največ dve različni pravi aktivnosti: leva a_L in desna a_R . To pomeni, da za vsak par pravih aktivnosti obstaja W različnih oken (indeks pomeni število pojavitev desne aktivnosti):

$$w_0(a_L, a_R) = [a_L, a_L, \dots, a_L, a_L, a_L]$$

$$w_1(a_L, a_R) = [a_L, a_L, \dots, a_L, a_L, a_R]$$

$$w_2(a_L, a_R) = [a_L, a_L, \dots, a_L, a_R, a_R]$$

...

$$w_{W-1}(a_L, a_R) = [a_L, a_R, \dots, a_R, a_R, a_R]$$

Če obstaja A različnih aktivnosti, imamo tako $A(A-1)W$ različnih oken pravih aktivnosti. Za vsako tako okno lahko določimo verjetnostno porazdelitev pravih aktivnosti. To s pomočjo učenih podatkov storimo na štiri načine.

Prvi način za vsako okno pravih aktivnosti prešteje vsa okna prepoznanih aktivnosti, ki se pri tistem oknu pravih aktivnosti pojavijo v učenih podatkih. To pomeni, da če npr. pri oknu pravih aktivnostih [hoja, spuščanje, spuščanje] prepoznamo [hoja, hoja, padanje], za ena povečamo števec takih primerov. Če je W velik, je možnih oken prepoznanih aktivnosti ogromno, zato večine v učenih podatkih ni. Da ublažimo razliko med nikoli in redko videnimi okni prepoznanih aktivnosti, uporabimo Laplaceovo glajenje: na začetku števec vseh oken prepoznanih aktivnosti nastavimo na 1. Takšno glajenje uporabimo tudi pri ostalih načinih.

Prvi način tako določi verjetnost vsakega okna prepoznanih aktivnosti pri vsakem oknu pravih aktivnosti:

$$P([a_1^P, \dots, a_W^P] | [a_1, \dots, a_W])$$

Ko želimo določiti okno pravih aktivnosti za neko okno prepoznanih aktivnosti $[a_1^P, \dots, a_W^P]$, za vsa okna pravih aktivnosti $w_i(a_L, a_R)$ preverimo, kolikšna je verjetnost $P([a_1^P, \dots, a_W^P] | w_i(a_L, a_R))$. Za tisto okno pravih aktivnosti, za katero je ta verjetnost največja, odločimo, da ustreza dejanskim aktivnostim.

Drugi način je milejši od prvega in za vsako okno pravih aktivnosti prešteje le pojavitve prepoznanih aktivnostih na posamičnih mestih okna. To pomeni, da če pri oknu pravih aktivnostih [hoja, spuščanje, spuščanje] prepoznamo [hoja, hoja, padanje], za ena povečamo števec hoj na prvem in na drugem mestu tega okna pravih aktivnosti ter padanj na tretjem mestu. To pomeni, da določimo verjetnosti:

$$P(\alpha_i^P = \hat{a}_j | [a_1, \dots, a_W]); i = 1 \dots W, j = 1 \dots A$$

Z α_i označimo aktivnosti na i -tem mestu v oknu, z \hat{a}_j pa j -to vrednost, ki jo ta aktivnost lahko zavzame.

Ko želimo določiti okno pravih aktivnosti za neko okno prepoznanih aktivnosti $[a_1^P, \dots, a_W^P]$, za vsa okna pravih aktivnosti $w_i(a_L, a_R)$ preverimo, kolikšna je verjetnost:

$$P(\alpha_1^P = a_1^P | w_i(a_L, a_R)) \cdot \dots \cdot P(\alpha_W^P = a_W^P | w_i(a_L, a_R))$$

Za tisto okno pravih aktivnosti, za katero je ta verjetnost največja, odločimo, da ustreza dejanskim aktivnostim.

Tretji način je še milejši in za vsako okno pravih aktivnosti prešteje le pojavitve prepoznanih aktivnostih na mestih, ki pripadajo levi in desni pravi aktivnosti. To pomeni, da če pri oknu pravih aktivnostih [hoja, spuščanje, spuščanje] prepoznamo [hoja, hoja, padanje], za ena povečamo števec hoj pri levi pravi aktivnosti (hoji) ter števec hoj in padanj pri desni pravi aktivnosti (spuščanju). To pomeni, da določimo verjetnosti:

$$P(\alpha_L^P = \hat{a}_j | [a_1, \dots, a_W]); j = 1 \dots A$$

$$P(\alpha_R^P = \hat{a}_j | [a_1, \dots, a_W]); j = 1 \dots A$$

Ko želimo določiti okno pravih aktivnosti za neko okno prepoznanih aktivnosti $[a_1^P, \dots, a_W^P]$, za vsa okna pravih aktivnosti $w_i(a_L, a_R)$ preverimo, kolikšna je verjetnost

$$P(\alpha_L^P = a_1^P | w_i(a_L, a_R)) \cdot \dots \cdot P(\alpha_L^P = a_{W-i}^P | w_i(a_L, a_R)) \cdot P(\alpha_R^P = a_{W-i+1}^P | w_i(a_L, a_R)) \cdot \dots \cdot P(\alpha_W^P = a_W^P | w_i(a_L, a_R))$$

Za tisto okno pravih aktivnosti, za katero je ta verjetnost največja, odločimo, da ustreza dejanskim aktivnostim.

Četrty način je enak tretjemu, le da je še milejši. Za neko okno pravih aktivnosti pojavitve prepoznanih aktivnosti ne štejemo le, če nastopijo pri tistem oknu pravih aktivnostih, ampak če nastopijo kjerkoli pri pravih aktivnostih iz okna.

Združevanje več načinov izboljševanja prepoznavanja aktivnosti je enostavno. Naj bo $P_k([a_1^P, \dots, a_W^P] | w_i(a_L, a_R))$ verjetnost, ki jo za neko okno prepoznanih aktivnosti in okno pravih aktivnosti $w_i(a_L, a_R)$ določi k -ti način. Najprej za vsak način te verjetnosti za vsa okna pravih aktivnosti $i = 0 \dots W-1, a_L, a_R = \hat{a}_1 \dots \hat{a}_A$ normaliziramo na interval $[0, 1]$, da načine izenačimo. Nato pa za vsako okno pravih aktivnosti verjetnosti za vse načine $k = 1 \dots 4$ seštejemo. Za okno pravih aktivnosti z največjim seštevkom odločimo, da ustreza dejanskim aktivnostim.

Na koncu je treba biti pozoren še na nekaj podrobnosti. Prepoznavanje aktivnosti izboljšujemo z drsečim oknom. To pomeni, da vsako aktivnost popravljamo W -krat: od takrat, ko se pojavi na desnem robu okna, do takrat, ko pride do levega roba. Popravki se lahko razlikujejo. A največ vemo o aktivnosti, ko pripotuje do levega roba okna, zato jo dokončno popravimo šele takrat.

Če privzamemo, da je dokončno popravljena aktivnost pravilna, se mora okno, ki ji sledi, začeti s to aktivnostjo ali pa more vsebovati same enake aktivnosti. V nasprotnem primeru namreč kršimo predpostavko, da znotraj okna dolžine W nastopata največ dve različni aktivnosti. Brez te omejitve bi se lahko npr. zgodilo, da bi za prave aktivnosti določili [sedenje, sedenje, sedenje], okno pomaknili za eno aktivnost naprej (podčrtana aktivnost bi se zdaj štela za pravilno), za prave aktivnosti določili [spuščanje, sedenje, sedenje], okno spet pomaknili naprej in za prave aktivnosti določili [dviganje, sedenje, sedenje]. Rezultat bi bil zaporedje dokončno popravljanih aktivnosti sedenje, spuščanje, dviganje, v resnici pa bi uporabnik ves čas sedel. Z omejitvijo bi v drugem in tretjem koraku v poštev prišla samo okna pravih aktivnosti, ki se začenejo s sedenjem ali vsebujejo same enake aktivnosti. Ker bi bila okna iz treh spuščanj in dviganj premalo verjetna, bi bilo zaporedje dokončno popravljanih aktivnosti sestavljeno iz treh sedenj.

4.2 Poizkusi

Za preizkus načinov za izboljševanje prepoznavanja aktivnosti sta potrebni dve testni množici: ena za učenje osnovnega klasifikatorja za prepoznavanje aktivnosti in druga za učenje mehanizma za izboljševanje. Če bi namreč tako osnovni klasifikator kot mehanizem za izboljševanje učili na istih podatkih, bi klasifikator na njih delal predobro,

da bi se mehanizem sploh lahko česa naučil. Uporabili smo podatke, opisane v razdelku 2. Kot učno množico za klasifikator smo uporabili prvi in četrti posnetek vsake osebe, kot učno množico za izboljševanje drugi in peti posnetek, kot testno množico pa tretji posnetek.

Najprej smo preverili, kolikšno izboljšavo prinese vsak izmed štirih načinov za izboljševanje glede na klasifikacijsko točnost brez izboljševanja. Osnovni klasifikator za prepoznavanje aktivnosti smo zgradili z algoritmom naključni gozd in atributnimi vektorji iz desetih nanizanih naborov atributov. Za dolžino okna za izboljševanje smo izbrali $W = 10$. Rezultati so prikazani v tabeli 4. Vidimo, da se najbolje obnese način 2 in vsi načini skupaj. Način 2 očitno najbolje ujame ravnotežje med natančnostjo opisa napak, ki jih dela osnovni klasifikator, in zahtevo po dovolj učnih podatkih. Najslabši je način 1, ki ga pesti pomanjkanje učnih podatkov, slabša pa sta tudi načina 3 in 4, ki najbrž premalo natančno opisujeta napake, ki jih dela osnovni klasifikator. Odločili smo se, da bomo v prihodnje uporabljali vse načine skupaj. Kot prvo niso slabši od nobenega posamičnega načina, kot drugo pa je moč pričakovati, da se bodo obnesli pri količini podatkov, ki najbolj ustreza kateremukoli izmed načinov.

Izboljševanje	Klasifikacijska točnost
Brez	83,4
Način 1	77,5
Način 2	84,7
Način 3	84,3
Način 4	84,6
Vsi načini skupaj	84,7

Tabela 4. *Klasifikacijska točnost [%] z različnimi načini izboljševanja.*

V naslednjem koraku smo preverili, kakšen je vpliv velikosti okna W na izboljševanje. V levi polovici tabele 4 vidimo, da je najprimernejše okno velikosti 20, s katerim se klasifikacijska točnost s 83,4 % dvigne na 85,0 %. Bržkone gre tudi v tem primeru za najboljše ravnotežje med natančnostjo opisa napak, ki jih dela osnovni klasifikator za prepoznavanje aktivnosti, in zahtevo po učnih podatkih.

Okno	Izboljšanje	Št. naborov	Izboljšanje
5	0,62	1	1,74
10	1,30	3	2,08
15	1,47	5	2,01
20	1,66	10	1,66
25	1,41		

Tabela 5. *Izboljšanje klasifikacijske točnosti [v odstotnih točkah] z različnimi velikostmi okna za izboljševanje in različnimi števili naborov atributov v atributnem vektorju.*

Na koncu smo pri oknu za izboljševanje velikosti $W = 20$ preverili še, kakšen je vpliv števila naborov atributov v atributnem vektorju. Domnevali smo, da bo izboljšanje pri manj naborih večje, saj krajši atributni vektorji sami po sebi zajamejo manj informacij o dogajanju v času. V desni

polovici tabele 4 vidimo, da je bila domneva pravilna. A ker klasifikacijska točnost osnovnega klasifikatorja pri manj naborih pade hitreje, kot jo izboljševanje poveča, je še vedno najbolje uporabiti deset naborov.

5 ZAKLJUČEK

V prispevku smo določili najprimernejši algoritem za strojno učenje in attribute za prepoznavanje aktivnosti iz Ubisensovih podatkov. Ugotovili smo, da naključni gozd, ki kot atributni vektor uporablja niz desetih zaporednih naborov atributov posamičnih položajev telesa, doseže klasifikacijsko točnost 96,7 % pri desetkratnem prečnem preverjanju in 81,5 % na način »izpusti eno osebo«.

Razvili smo štiri načine izboljševanja prepoznavanja aktivnosti, ki upoštevajo informacije o časovnem zaporedju prepoznav. Z vsemi štirimi skupaj se je klasifikacijska točnost dvignila za 1,66 odstotne točke. Ta dvig se zdi majhen, a glede na to, da večina posegov v prepoznavanje prinaša podobno velike spremembe, je možno, da brez radikalno drugačnega pristopa več ni mogoče doseči.

V prihodnje nameravamo mehanizem za izboljševanje temeljiteje preizkusiti, denimo na način »izpusti eno osebo«. Poleg tega razmišljamo o mehanizmu za izboljševanje, ki bi poleg časovnega zaporedja upošteval tudi podobnosti med aktivnostmi.

Zahvala

Raziskave, opisane v prispevku, je financirala Evropska skupnost s Sedmim okvirnim programom pod številko projekta 214986. Znatno del programja, uporabljenega pri delu, opisanem v prispevku, je napisal Boštjan Kaluža.

Literatura

- [1] Confidence. <http://www.confidence-eu.org/> [2008-09-15]
- [2] eMotion. Smart motion capture system. <http://www.emotion3d.com/smart/smart.html> [2008-09-15]
- [3] Luštrek, M., Gams, M., in Veléz, I. (2009). Posture and movement monitoring for ambient assisted living. IST-Africa 2009.
- [4] Luštrek, M., in Gams, M. (2008). Prepoznavna položaja telesa s strojnim učenjem. Informacijska družba 2009, str. 30-33.
- [5] Luštrek, M., in Kaluža, B. (2009). Fall Detection and Activity Recognition with Machine Learning. Informatica 33 (2), 205–212.
- [6] Maybeck, P. S. (1979). Stochastic models, estimation, and control. Mathematics in Science and Engineering 141.
- [7] Ubisense. <http://www.ubisense.net/> [2008-09-15]
- [8] Witten, I. H., in Frank, E. (2005). Data Mining: Practical machine learning tools and techniques, 2nd Edition. Morgan Kaufmann, San Francisco, ZDA.

STROJNA KLASIFIKACIJA SPLETNIH STRANI PO TEMAH

Domen Marinčič
Odsek za inteligentne sisteme
Institut "Jožef Stefan"
Jamova 39, 1000 Ljubljana, Slovenija
domen.marincic@ijs.si

Povzetek

Z naraščajočo velikostjo svetovnega spleta postaja potreba po kategorizaciji spletnih strani vedno večja. Kategorizacija je neobhodno potrebna za mnogo nalog, kot je iskanje dokumentov v spletu, izdelava in vzdrževanje spletnih imenikov in podobno. V članku opisujemo algoritem za kategorizacijo spletnih strani po temah, ki stavi na hitrost in enostavnost delovanja, saj za kategorizacijo uporablja predvsem informacije dostopne v spletnih straneh samih. Kot učna in testna množica za algoritem je služil Yahoojev spletni imenik.

1 Uvod

Klasifikacija spletnih strani igra pomembno vlogo pri mnogih opravilih upravljanja z informacijami. Na spletu je klasifikacija vsebine spletnih strani odločilna za razvoj spletnih imenikov, analizo spletnih povezav, analizo tematske strukture spleta ipd. Klasifikacija spletnih strani tudi izboljša kakovost spletnega iskanja.

Klasifikacija spletnih strani, poznana tudi kot spletna kategorizacija, je postopek določanja kategorije spletni strani. Pri tem običajno obstaja zaključen seznam kategorij, ki je vnaprej določen. Kategorizacijo se pogosto predstavi kot problem nadzorovanega učenja, kjer zbirka spletnih strani označenih s kategorijami služi kot učna množica za strojne klasifikatorje. Glede na organizacijo kategorij lahko problem obravnavamo kot enonivojsko ali hierarhično klasifikacijo. Pri prvem primeru kategorije obravnavamo kot vzporedne, kjer ni relacij podrejenosti/nadrejenosti med kategorijami. V drugem primeru so kategorije organizirane v hierarhijo običajno predstavljeno v obliki drevesa.

Uporaba spletne kategorizacije je pomembna na mnogo področjih. V primeru spletnih imenikov gre običajno za ročno uvrščanje strani, delo ki ga morajo opraviti uredniki, kar zahteva precej človeškega napa. Predkupno klasificiranje strani v kategorije [1, 2] z naknadnim pregledovanjem in popravki s strani urednikov lahko precej zmanjša količino potrebnega dela. Drugo področje je izboljšanje kakovosti rezultatov iskanja po spletu. Ključne besede pri iskanju običajno

ne določajo vrste zadetkov na nedvoumen način. Če ima uporabnik poleg ključnih besed na voljo še izbiro kategorije strani, lahko število zadetkov zožamo - s pogojem, da imamo na voljo učinkovit način avtomatske kategorizacije [3]. Prav tako je možno izboljšati delovanje sistemov za odgovarjanje na vprašanja, saj lahko z določanjem teme vprašanja in kategorije strani omejimo število virov kjer je možno priti do relevantnih informacij [4, 5].

Spletna klasifikacija je podproblem splošnega problema klasifikacije samostojnega besedila. Od klasifikacije samostojnega besedila se spletna klasifikacija razlikuje v sledečih pogledih. Klasifikacija samostojnega besedila običajno poteka na dobro definiranih zbirkah besedila z enotnim stilo in strukturo, kar pri spletnih straneh ni res. Po drugi strani je prednost spletnih strani, da so izrecno strukturirane s pomočjo HTML oznak, kar daje več informacij o samem besedilu. Poleg tega pri spletnih straneh obstajajo še povezave na druge spletne strani, kar je možno uporabiti kot dodaten kontekst pri avtomatski kategorizaciji.

Za nadzorovano učenje strojnih klasifikatorjev se spletne strani opiše s formalnimi atributi. Pri kategorizaciji spletnih strani razlikujemo attribute, ki so vsebovani na sami straneh, ter attribute, ki jih dobimo iz povezanih strani. Atributi znotraj strani so tisti, ki so najboljše definirani in do katerih je najlažji dostop. A ker je vsebina le delno strukturirana, uporaba vreče besed kot atributov ne daje optimalnih rezultatov. Potrebno je vsaj filtrirati najmanj pogoste besede oziroma n-grame [6]. Drug način, kako strukturirati attribute znotraj strani, je uporaba HTML oznak, kar dokazano izboljšuje točnost kategorizacije [7]. Pri njihovem delu se razlikuje med besedilo, ki izhaja iz naslova strani, od naslova poglavij in telesa strani. Najboljše rezultate s doseže z uteženo kombinacijo atributov različnih tipov. Najhitrejša in enostavna, pa zato tudi manj učinkovita, je kategorizacija na podlagi URL naslova strani [8, 9], ki pride prav tudi takrat, ko vsebina strani ni dostopna.

Drugo vrsto atributov predstavljajo atributi sosednjih spletnih strani, kjer sosednost lahko definiramo s pomočjo hiperpovezav. V splošnem lahko predpostavimo, da če strani a in b pripadata isti kategoriji, da imajo sosednje strani podobne karakteristike, kar lahko

s pridom izkoriščamo pri kategorizaciji strani po temah. Naprimer: prisotnost mnogih športnih strani med sosedi nakazuje, da je lahko tudi kategorizirana stran s področja športa. Poskusi kažejo, da uporaba takih atributov najboljše dela pri splošnih temah višjega nivoja, medtem ko klasifikacija na bolj podrobno razmejene teme ne daje več optimalnih rezultatov. Drugo vprašanje je še določanje oddaljenosti sosedov in vrste sosedov. Obstoječi poskusi se osredotočajo na sosede, ki niso oddaljeni več kot dva skoka, torej starši, otroci, brati, stari starši in vnuki. Poskusi tudi kažejo, da so izmed teh najbolj koristni brati [10, 11].

2 Algoritem za klasifikacijo

Razvili smo algoritem za enonivojsko kategorizacijo spletnih strani po temah. Za seznam tem smo uporabili zgornji nivo Yahoojevega spletnega imenika:

- Animals
- Business & Finance
- Computers & Internet
- Culture & Community
- Entertainment & Arts
- Family & Home
- Games
- Government & Politics
- Health & Wellness
- Hobbies & Crafts
- Music
- Regional
- Religion & Beliefs
- Romance & Relationships
- Schools & Education
- Science
- Sports & Recreation

2.1 Učenje

Za učenje smo uporabili strani iz spletnega imenika Yahoo. Iz vsake od izbranih kategorij smo izbrali naključno določenih 1000 spletnih strani. Le-te smo pridobili s pomočjo lastno razvitega pajka, ki je Yahoojevo spletišče obiskoval s frekvenco eno stran na pet

sekund. Pri pajkih so vedno potrebni daljši časovni intervali med obiski istega spletišča, saj bi sicer spletne strežnike preveč obremenili.

Vsaka spletna stran je opisana z atributi, ki izvirajo iz spletne strani same. Attribute smo razdelili glede na vir:

- naslov spletne strani,
- naslov poglavja v spletni strani,
- telo spletne strani.

Strukturiranje atributov glede na njihov izvor omogočajo HTML oznake. Iz vsakega od naštetih virov smo upoštevali besede odprtega tipa (glagoli, samostalniki, pridevniki, prislovi), medtem ko besede zaprtega tipa (vezniki, medmeti, členi ipd.) niso bile upoštevane pri določanju atributov.

Za vsako kategorijo so bili zgrajeni tri atributni vektorji, eden za attribute iz naslova spletne strani, drugi za besede iz naslovov poglavij, tretji iz telesa strani. Vsaka beseda je v vektorju predstavljena kot ena izmed dimenzij, velikost dimenzije pa določa število pojavitev besede v učnih spletnih straneh. Pri tem smo v vsakem vektorju obdržali le zgornjih 1000 dimenzij glede na velikost.

2.2 Avtomatska klasifikacija

Za klasifikacijo strani je uporabljena metoda skalarnega produkta. Za vsako klasificirano stran se izdelajo tri atributne vektorje po enakem postopku, kot so bili pridobljeni atributni vektorji v postopku učenja. Nato za vsak tip vektorja izračunamo skalarni produkt med vektorjem strani in vektorjem kategorije. Vse tri skalarne produkte se nato sešteje, pri čemer je vektor naslova obtežen z utežjo 5, vektor naslovov podpoglavij z utežjo 3 in vektor telesa spletne strani z utežjo 1. Uteži so bile določene empirično.

Končni rezultat klasifikacije je rang kategorij, pri čemer je na prvem mestu kategorija spletne strani, ki ji klasifikacijski algoritem priredi največjo oceno.

3 Ovrednotenje

Za ovrednotenje smo opravili poskus na testni množici spletnih strani, ki je vsebovala po 100 strani za vsako kategorijo. Nobena od spletnih strani v testni množici se ni nahajala v učni množici. Rezultate kategorizacije spletnih strani navajamo v spodnji tabeli, pri čemer navajamo delež primerov, ko je prava kategorija dobila:

- prvi rang (prvi stolpec v tabeli),
- enega izmed prvih treh rangov (drugi stolpec v tabeli),
- enega izmed prvih petih rangov (tretji stolpec v tabeli).

Rang	Delež pravilnih klasifikacij
Rang 1	19,2%
Rang 1-3	25,3%
Rang 1-5	31,4%

Poskusi kažejo, da se v najmanj primerih prava kategorija pojavi z najvišjim rangom, medtem ko pravi kategoriji kategorizacijski algoritem večkrat dodeli enega izmed prvih treh oziroma enega izmed prvih petih rangov.

4 Zaključek

Iz rezultatov poskusov lahko sklepamo, da relativno enostaven algoritem za klasifikacijo daje rezultate z zmerno točnostjo klasifikacije. Tak algoritem je uporaben predvsem, če imamo na voljo omejene učne vire, ter mora delovati zelo hitro. Za boljše rezultate bi bilo vsekakor treba vključiti bolj izdelano filtriranje dimenzij atributnih vektorjev. Še en korak naprej pa bi bila vključitev sosednjih spletnih strani v atributni opis - kar pa zahteva učne podatke o spletnih povezavah, ki so na voljo le največjim podjetjem, ki se ukvarjajo s spletnim iskanjem.

Literatura

- [1] Huang, C.-C., S.-L. Chuang, and L.-F. Chien (2004a). Liveclassifier: Creating hierarchical text classifiers through web corpora. In *Proceedings of the 13th International Conference on World Wide Web*, New York, NY, pp. 184192. ACM Press.
- [2] Huang, C.-C., S.-L. Chuang, and L.-F. Chien (2004b). Using a web-based categorization approach to generate thematic metadata from texts. *ACM Transactions on Asian Language Information Processing (TALIP)* 3(3), 190212.
- [3] Chekuri, C., M. Goldwasser, P. Raghavan, and E. Upfal (1997). Web search using automated classification. In *Proceedings of the Sixth International World Wide Web Conference*, Santa Clara, CA. Poster POS725.
- [4] Yang, H. and T.-S. Chua (2004a). Effectiveness of web page classification on finding list answers. In *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, New York, NY, pp. 522523. ACM Press.
- [5] Yang, H. and T.-S. Chua (2004b). Web-based list question answering. In *Proceedings of the 20th international conference on Computational Linguistics*, Morristown, NJ, pp. 1277. Association for Computational Linguistics.
- [6] Mladenic, D. (1998). Turning Yahoo into an automatic web-page classifier. In *Proceedings of the European Conference on Artificial Intelligence*, pp. 473474. Mladenic, D. (1999, Jul/Aug). Text-learning and related intelligent agents: A survey.
- [7] Golub, K. and A. Ardo (2005, September). Importance of HTML structural elements and metadata in automated subject classification. In *Proceedings of the 9th European Conference on Research and Advanced Technology for Digital Libraries*, Volume 3652 of LNCS, Berlin, pp. 368378. Springer.
- [8] Kan, M.-Y. (2004). Web page classification without the web page. In *Proceedings of the 13th International World Wide Web Conference Alternate Track Papers & Posters*, New York, NY, pp. 262263. ACM Press.
- [9] Kan, M.-Y. and H. O. N. Thi (2005). Fast webpage classification using URL features. In *Proceedings of the 14th ACM International Conference on Information and Knowledge Management*, New York, NY, pp. 325326. ACM Press.
- [10] Chakrabarti, S., B. E. Dom, and P. Indyk (1998). Enhanced hypertext categorization hyperlinks. In *Proceedings of the ACM SIGMOD International on Management of Data*, New York, NY, pp. 307318. ACM Press.
- [11] Qi, X. and B. D. Davison (2006). Knowing a web page by the company it keeps. In *Proceedings of the 15th ACM International Conference on Information and Knowledge Management (CIKM)*, New York, NY, pp. 228237. ACM Press.

TOWARDS ROBUST RULE ENGINE FOR CLASSIFYING HUMAN POSTURE

Violeta Mirčevska, Matjaž Gams
Department of Intelligent Systems
Jožef Stefan Institute
Jamova 39, 1000 Ljubljana, Slovenia
e-mail: violeta.mircevska@ijs.si

ABSTRACT

This paper presents a procedure for developing rule engine by combining machine learning with expert knowledge. This procedure was applied in the domain of classifying human posture based on information of location of body parts. The procedure overcomes in certain measure the problem of over-fitting to non-representative training dataset. Tests show improvement in accuracy of the developed rule engine in comparison to machine learning techniques. Possibilities for improvement and automation of the procedure for rule engine development are discussed at the end of the paper. The presented procedure can be applied to any problem domain in which representative training dataset is not available or is difficult to obtain.

1 INTRODUCTION

The European FP7 project CONFIDENCE – Ubiquitous Care System to Support Independent Living [1] aims at developing a system that will monitor the health conditions of the elderly in real time in a non-intrusive way [2] [3]. The reasoning of the system is based on measurements of location of user's body parts. Positions are measured solely with the use of radio sensors.

As part of the work in the CONFIDENCE project, we are researching techniques for classifying human activity based on measured location of body parts. Tests show that machine learning techniques tend to over-fit to training data. Main reason for this is the difficulty for obtaining representative training dataset for posture classification due to the wide variety of possible body constitutions that need to be recorded. Therefore, we considered combining machine learning techniques with expert knowledge. Since humans are good at imagining body structures and postures not represented in the training data, it is reasonable to expect that application of expert knowledge will improve the accuracy and robustness of the classifier.

This paper presents the procedure for developing a rule engine by combining machine learning and expert knowledge in the domain of classifying human posture. More precisely, we present the structure of the rule engine, the methodology according to which it was developed and the achieved results. Detailed description of posture classification with the use of machine learning techniques

can be found in [4] [5] [6]. Additionally, since the obtained results are promising, we discuss possible improvements and possibilities for automation of the procedure for rule engine development.

The paper is organized as follows. Section 2 presents the architecture of the rule engine and the methodology for its development. In Section 3 we present evaluation of the rule engine and compare it with other machine learning techniques. Section 4 discusses possibilities for improvements and automation of the procedure for rule engine development. Section 5 concludes the paper.

2 RULE ENGINE FOR CLASSIFYING HUMAN ACTIVITY

The proposed rule engine (Figure 1) has hierarchical architecture and is composed of three levels of reasoning – strict rules, weak rules and default rule. More detailed explanation of these reasoning levels will be presented on the domain of classifying human posture.

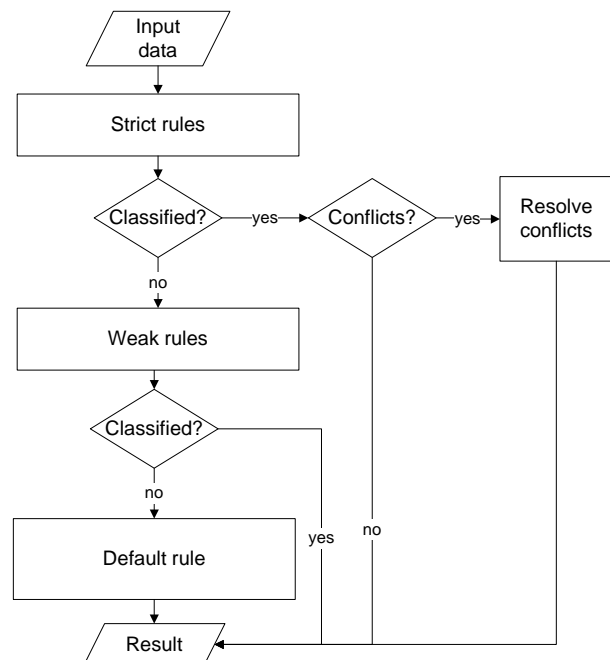


Figure 1: Proposed rule engine architecture

In the domain of classifying human posture we examine six postures of interest - standing, sitting, lying, falling, moving downwards (normally) and moving upwards. The classification is done based on information about the position of the neck and the ankles of the user. More precisely, only the Z-coordinate of the neck and ankles is considered. The X- and Y- coordinates are not relevant, because they refer to the place in the room where the user is. Being one of the highest tags of the body, the height of the neck with respect to the ankles is important for distinguishing among standing, sitting and lying. Moreover, the neck has the highest velocity in the Z direction, making it suitable for distinguishing among falls, moving downwards and upwards.

The strict rules contain precise definitions of the body configuration in each of the postures of interest. They are created by enriching rules extracted from decision tree models with expert knowledge (Section 2.1). Each instance is first processed by the strict posture rules. If it is covered by set of strict posture rules describing only one posture class, this class is assigned to it. Conflicts when a particular instance is covered by rules of more than one posture are resolved as presented in Table 1. In the domain of human posture classification conflicts appear between rules for adjacent classes, e.g. standing and going down. Since in this particular application, the rules for standing, sitting, lying and falling were constructed in a way that only pure postures are captured, they are chosen when there is a conflict with a rule for moving downwards/upwards.

Table 1: Resolution of conflicts among the posture rules

Conflict	Result
Standing and moving downwards/upwards	Standing
Sitting and moving downwards/upwards	Sitting
Lying and moving downwards/upwards	Lying
Falling and moving downwards/upwards	Falling

The weak posture rules specify the most probable class according to the user's neck-ankle distance. The weak posture rules were created by using expert knowledge. Each instance which is not covered by any of the strict posture rules is processed by the weak posture rules.

Finally, the default rule is used to assign a class to an instance that is not covered by both the strict and the weak posture rules. Since the current posture of a person is highly correlated with the posture he/she had in the previous time interval, the default rule assigns the class of the previous time interval to the instance in the current time interval.

2.1 Strict rule extraction procedure

The strict rules were extracted in the following way:

1. Create one-against-all dataset for a particular posture
2. Create decision trees
3. Extract rules with high precision and possibly high recall
4. Modify the extracted rules by expert knowledge.

Similar procedure for extraction of rules is presented in [7].

In order to concentrate on differences between a particular posture and all other postures, one-against-all datasets were created for each posture of interest. Decision trees were generated with the purpose of identifying attributes that best separate the examples of concrete posture from all other postures.

We constructed several decision trees using different attribute sets in order to find all relevant hypotheses for the postures of interest. The decision tree induction technique performs general-to-specific hill-climbing search through the space of possible hypotheses. Because of this, relevant information may be hidden behind the best hypothesis. Therefore, a decision tree was first built with all attributes. Then, the procedure was repeated by removing the attribute at the root node or attribute near the root node with the aim of finding relevant hidden hypotheses, until the classification accuracy of the resultant tree significantly dropped.

From the decision trees, rules with high precision and, possibly, high recall were extracted and modified by expert knowledge. The conditions of the extracted rules were made stricter, especially for the classes standing, sitting, lying and falling, improving their precision at the expense of recall. The aim was to correctly classifying pure postures, neglecting the borderline cases.

The rules were added to the set of rules for the particular class.

3 EXPERIMENTS

The performance of the rule engine was tested and compared with four machine learning techniques (support vector machines, random forest, bagging and decision trees) on two datasets obtained by recording human activity in two different circumstances.

3.1 Data

The evaluation of classifiers was done with the use of two datasets obtained by recording human activity with the use of Smart infrared motion capture system [8]. The first recording, containing 135 sequences of behavior of three persons, includes examples of standing/walking, lying down, sitting down, and falling. The second recording, which contains 775 sequences of behavior of five people, includes the basic activities recorded in the first set, examples of several kinds of falls and, based on discussions with physicians, examples of walking and lying of people with different health problems, such as Parkinson's disease, hemiplegia etc.

3.2 Results

The performance of the rule engine was compared with the performance of four machine learning techniques –support vector machines, random forest, bagging and decision trees (Table 3). The support vector machines, random forest and bagging classifier were evaluated in Weka [9] with default settings, whereas the decision tree classifier was induced with the minimal number of instances per leaf set to 2% of the training dataset. The performance of the machine learning techniques was evaluated with 10-fold cross validation on the data from both phases and with three separate training and test set scenarios. The performance of the rule engine is presented with its accuracy on the test dataset for each separate training and test set scenario.

Examination of the classification accuracy of the machine learning techniques in the different evaluation scenarios suggests a certain degree of over-fitting. The accuracy of the classifiers is highest when they are evaluated with 10-fold cross validation. The random selection of training and test dataset in 10-fold cross validation leads data about the behavior of concrete person in a concrete phase to be present in both the training and test dataset. Therefore, over-fitting is most likely to be present in this evaluation scenario. The classification accuracy falls when the classifiers are induced on data about two persons and tested on data of a third person. In this case, the training dataset does not contain data about the behavior of the person on which the model is tested. However, since all persons were instructed to behave in the same way in both phases of recordings and they were able to observe and copy each other, the models induced in this evaluation scenario are likely over-fitted to this particular behavior of the persons. The most significant drop in accuracy happens when the classifiers are induced on one phase of recordings and tested on the other. In this case, the training and test dataset contain different behavior and there are persons for which recordings were only made in the second phase. The fall of classification accuracy in this scenario confirms that the models induced with machine learning get over-fitted to the persons and behavior present in the training dataset.

As seen in Table 3, the classification accuracy of the

decision trees techniques is 10 percentage points lower than the classification accuracy of the rule engine when the classifier is trained on one phase of recordings and tested on the other. This shows that the incorporation of expert knowledge in the knowledge obtained by inducing decision trees increased the generality of the posture classifier.

The accuracy of the other three machine learning techniques is also smaller when training is done on one phase of recordings and testing on the other. The difference is especially significant for the random forest and bagging classifier, where the drop in accuracy is more than 12% when the classifier is induced on the data from the first phase of recordings and tested on the data from the second phase. The difference in classification accuracy between the machine learning techniques and the rule engine is not significant when the classifiers are trained on data about two persons and tested on a third person. Nevertheless, the higher classification accuracy of the rule engine still suggests that the incorporation of common sense improved the generality of the classifier.

4 POSSIBILITIES FOR ENHANCEMENT OF THE PROCEDURE FOR RULE ENGINE DEVELOPMENT

The developed rule engine is robust, can easily be adapted to particular user and is computationally cheap. The presented procedure for rule engine development principally determines:

1. which attributes are important for concrete class of interest?
2. what are the limits for the attributes in the conditions of rules for concrete class?

One drawback of the presented procedure for rule engine development is its work intensiveness. In this section, we discuss possibilities for enhancement of the procedure for rule engine development by:

- providing the expert with information concerning important aspects of the addressed problem domain
- automating parts of the procedure for rule engine development

Table 2: Comparison of the classification accuracy among the rule engine and four machine learning

Evaluation		Support vector machines	Random forest	Bagging	Decision trees	Rule engine
Training dataset	Test dataset					
First phase	Second phase	84.68%	74.01%	74.76%	79.92%	87.00%
Second phase	First phase	77.96%	79.63%	79.25%	70.78%	81.76%
Two persons	Third person	86.89%	85.99%	86.52%	87.24%	89.42%
10-fold cross validation		89.51%	95.85%	94.91%	87.66%	

4.1 Determination of attributes important for concrete class

The determination of attributes important for concrete class will be more accurate if the expert is presented with information concerning important aspects of the addressed problem domain by means of ontologies, statistics and genetic programming results.

Ontologies define problem domains and information related to them and, therefore, provide important information in the process of determining attributes important for concrete class. Today, there is active research in the field of ontologies and definition of wide variety of domains by means of ontology can be expected in near future. Examination of ontologies defining concrete class of interest being examined will be a must.

The expert may be assisted in the process of determining attributes important for concrete class by providing her with information about the statistical importance of the examined attributes. Attributes may be ranked according to measures of statistical importance, such as relief, information gain, gain ratio and similar. The ranking would clearly present which attributes are relevant to the examined class.

Finally, combination of attributes important for concrete class of interest can be obtained by genetic programming. The stochastic nature of genetic programming may overcome the obstacle of over-fitting to non-representative problem domain. By this, the genetic programming results would present one more aspect concerning significant combinations of attributes for the examined class.

4.2 Determination of the limits for the attributes in the conditions of rules for concrete class

The determination of the limits for the attributes in the conditions of rules for concrete class may be automated by means of evolutionary algorithms. By training the previously presented rule engine on the first set of recordings using evolutionary algorithms, the classification accuracy on this set of recordings increased by more than 4% and the accuracy on the data from the second set remained in the same range. In the applied approach, one individual contained all limits in all of the rules present in the rule engine. One population consisted of 40 individuals. The starting population contained the limits determined with the previously expressed approach of combining machine learning and expert knowledge. The 39 rest individuals were created by randomly modifying these limits by 50%. The population was evolved 400 times.

6 CONCLUSION

This paper presented a procedure for developing rule engine by combining machine learning and expert knowledge. Incorporation of expert knowledge overcomes in certain measure over-fitting to non-representative training dataset, problem occurring with machine learning techniques.

The presented procedure was applied in the domain of classifying human posture based on measured location of

body parts. Representative training dataset is difficult to obtain for this domain due to the wide variety of body constitutions in each body posture. The improvement in accuracy, as compared to machine learning techniques, indicates that the developed rule engine is robust. Additionally, the developed rule engine can easily be adapted to particular user and is computationally cheap.

Possibilities for improvement of the procedure for rule engine development need to be considered. One aspect relates to providing the expert with information encompassing different aspects of the examined problem domain in order to assist her in the determination of attributes important for classifying concrete class. Other aspect is examination of possibilities for automatically determining the limits for attributes in the conditions of rules for concrete class by means of evolutionary algorithms.

Acknowledgement

The research leading to these results has received funding from the European Community's Framework Programme FP7/2007-2013 under grant agreement n° 214986.

References

- [1] CONFIDENCE, <http://www.confidence-eu.org/>. [2009-09-05]
- [2] C. Quemada, I. Val, M. Pietrzyk, T. Von der Grün, N. González, I. Vélez, A Smart Care System for Elderly People to Support Independent Living: CONFIDENCE, *Smart Systems Integration* (2009).
- [3] N. González, I. Vélez, CONFIDENCE: Support To Older People's Independent Living, *Proceedings of Psychologia* (2008).
- [4] M. Luštrek, B. Kaluža, Fall detection and activity recognition with machine learning, *Informatika* 33 2 (2009).
- [5] M. Luštrek, M. Gams, Posture And Movement Recognition From Locations Of Body Tags, Ami-08 (European Conference on Ambient Intelligence), 19-22 November 2008, Nürnberg (Germany).
- [6] M. Luštrek, M. Gams, I.Vélez, Posture and movement monitoring for ambient assisted living, IST-Africa 2009, 6-8 May 2009, Kampala (Uganda).
- [7] M. Ožek, M. Gams, Use of Data Mining Techniques for Process Analysis on Small Databases, *Dynamic and Advanced Data Mining for Progressing Technological Development*, IGI Global, USA (accepted).
- [8] eMotion, Smart motion capture system, <http://www.emotion3d.com/smart/smart.html>. [2009-09-05]
- [9] I. H. Witten, E. Frank, *Data Mining: Practical machine learning tools and techniques*, 2nd Edition, Morgan Kaufmann, San Francisco, USA, 2005.

ANALIZA DELOVANJA VIRTUALNEGA SVETOVALCA

Matej Ožek, Matjaž Gams, Jana Krivec

Odsek za Inteligentne sisteme

Institute »Jožef Stefan«, Jamova 39, 1000 Ljubljana, Slovenija

e-mail: matej.ozek@ijs.si; matjaz.gams@ijs.si; jana.krivec@ijs.si

Povzetek

Na spletni strani Instituta »Jožef Stefan« že eno leto deluje virtualni svetovalec Robi. V prispevku bomo analizirali vprašanja in odgovore. Sistem je uporabljalo 1200 oseb in zastavilo preko 11000 vprašanj, kar je glede na skromno povezavo na internetni strani IJS odličen dosežek.

1. UVOD

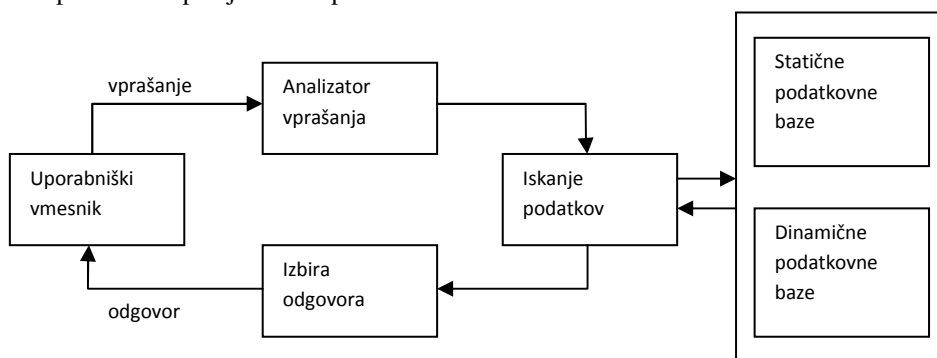
Zaradi vedno bolj zmogljivih računalnikov in vedno hitrejšega prenosa informacij nastaja vedno več spletnih aplikacij. Na ta način prodirajo na splet tudi inteligentna orodja za pomoč pri iskanju informacij. En tak primer je prva aplikacija Virtualnega svetovalca, ki se imenuje Vida. Razvilo jo je tuje podjetje za spletne strani davčne uprave, odsek za inteligentne sisteme pa je pomagal pri prenosu mednarodnega sistema v slovenščino. V letu 2008 pa so se pojavile mnoge podobne aplikacije in danes lahko vidimo Virtualnega svetovalca na spletnih straneh mnogih občin in podjetij (npr: Občina Maribor, Univerza v Mariboru, SiOL, Institut »Jožef Stefan«).

Na sliki 1 vidimo shemo delovanja Virtualnega svetovalca. Glavne tri komponente so [2]: uporabniški vmesnik, modul za obdelavo vprašanj in baza znanja. Ponavadi je baza znanja najpomembnejši del sistema, saj vsebuje urejene odgovore in algoritme, s katerimi se odgovori generirajo. Skoraj enako pomembni so programi za obdelavo vprašanj, saj je od njih odvisno, ali bo računalnik vprašanje sploh razumel. S stališča uporabnika pa je zelo pomemben

uporabniški vmesnik oz. »avatar«. Ta je ponavadi animiran lik, ki se deloma odziva na vprašanja uporabnika. Animacija spodbuja uporabnika, ki je navajen iskanja po ključnih besedah, k oblikovanju slovnično pravilnih vprašanj.

Za konstrukcijo Virtualnega svetovalca se večinoma uporabljata dva pristopa. Prvi je izdelava čim boljšega programa za analizo odgovora, ki vsebuje [1] lematizatorje, ontologije in uporablja metode umetne inteligence za to, da modelira temo pogovora, kar olajša izdelavo najprimernejšega odgovora. Dobra stran tega pristopa je, da so odgovori, ki jih sistem vrača, zelo dobri. Slaba stran pa je, da za slovenščino še ne obstajajo programi za analizo besedila, ki so po kvaliteti primerljivi z velikimi jeziki (npr. angleščina). Potrebno je tudi paziti na težave, na katere so opozorili avtorji pogovornega robota Bildgesmythe (zmagovalec tekmovanja pogovornih robotov CBC 2008). Po uporabi metode strojnega učenja naivni Bayes za izbiro odgovorov, se je kvaliteta odgovorov zmanjšala in ne povečala. Drug pristop izdelave Virtualnih asistentov gradi v nasprotno smer. Izbiramo le tiste funkcije, ki so nujno potrebne. Tako vprašanje povežemo z odgovorom preko ključnih besed. Pri tem pristopu je daleč najbolj pomembna dobro urejena baza znanja.

Predvidevamo, da bo dolgoročno obveljal prvi pristop. Vendar pa bomo v nadaljevanju predstavili, da trenutno deluje drugi pristop enako dobro, ali celo bolje kot prvi, poleg tega pa potrebuje za izdelavo veliko manj sredstev.



Slika 1: Struktura Virtualnega svetovalca

2. DELOVANJE VIRTUALNEGA SVETOVALCA ROBIJA

Na spletni strani IJS je septembra 2008 začel delovati Virtualni svetovalec Robi. Cilj raziskave sistema Robi je ugotoviti, ali je mogoče z minimalnim vložkom narediti uspešen sistem za odgovarjanje na vprašanja.

Najpomembnejši del Virtualnega asistenta Robi je baza znanja. Razdeljena je na del v naprej pripravljenih odgovorov in del odgovorov, ki se ustvarjajo sproti. Med slednje spadajo odgovori, za katere se dobi informacije sproti iz spletnih strani instituta, in tipizirani odgovori, ki se črpajo iz drugih baz, npr. iz telefonskega imenika. K vsakemu odgovoru spada seznam ključnih besed ob katerih se določen odgovor prikaže uporabniku.

Program za obdelavo vprašanj z regularnimi izrazi išče ključne besede po vprašanju. Vsak nabor ključnih besed je povezan z odgovorom, ki ga Robi vrne. Tak program deluje hitro in ga ni težko narediti tudi za delo z večjim številom uporabnikov.

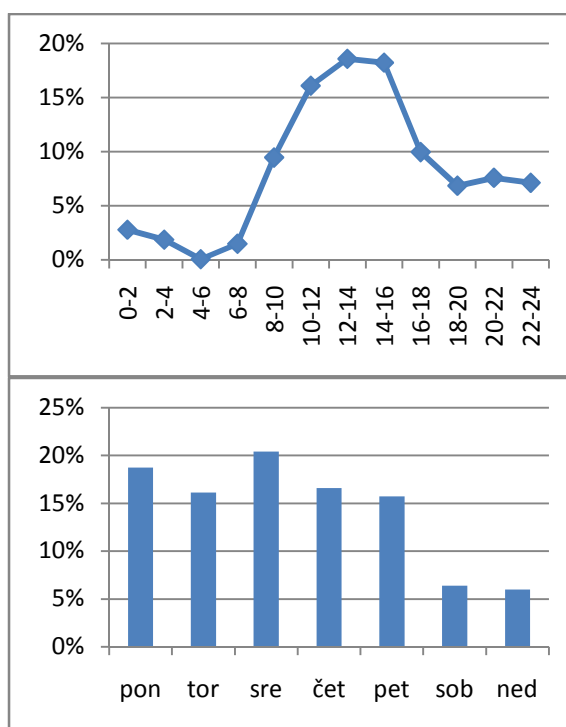
V prvem letu delovanja smo Robiju štirikrat zamenjali uporabniški vmesnik. Gre tako za menjavo vizualne podobe, kot za prilagoditev uporabnikom, ki dostopa z mobilnega telefona.

Vsi pogovori z Robijem se beležijo. Za analizo imamo datum in uro pogovora, uporabnikov IP, vprašanje in odgovor. Primer izpisa enega prvih uspešnih pogovorov:

```
463472533 :: 2008-08-31 17:05:47.446000 :: 90.157.x.x
Q: dober dan
A: Lepo pozdravljeni!
176562268 :: 2008-08-31 17:05:56.821000 :: 90.157.x.x
Q: kje je IJS
A: Naslov Instituta Jožef Stefan je Jamova cesta 39, 1000 Ljubljana.
Najdite nas na <a href='http://zemljevid.najdi.si/search_maps.jsp?q=Jamova+cesta+39&tab=maps' target='_blank'>zemljevidu.</a>
1908578208 :: 2008-08-31 17:06:06.368000 :: 90.157.x.x
Q: s čim se ukvarjate na ijs
A: Institut Jožef Stefan je vrhunska raziskovalna ustanova.
230050785 :: 2008-08-31 17:06:12.290000 :: 90.157.x.x
Q: kaj raziskujete
A: Na Institutu Jožef Stefan se ukvarjamo z raziskavami in razvojem
tehnologij, kot so nanotehnologije, novi materiali, biotehnologije,
tehnologije vodenja in proizvodnje, komunikacijske tehnologije,
računalniške tehnologije in tehnologije znanja, okoljske tehnologije in
reaktorske tehnologije.
820474647 :: 2008-08-31 17:06:20.149000 :: 90.157.x.x
Q: ti is faca
A: Res? Hvala.
```

3. ANALIZA PODATKOV

Robi je v prvem letu delovanja odgovoril na 11312 vprašanj (od septembra 2008 do septembra 2009). Vprašanja so prišla iz 1194 IP naslovov, kar se preslika v približno 1000 ljudi, ki so postavljali vprašanja. Po IP sklepamo, da je več kot pol vprašanj prišlo iz Inštituta. Iz slike 2 vidimo, da je največ vprašanj zastavljenih od 12. do 14. ure, v ponedeljek in sredo.



Slika 2: Vprašanja po urah v dnevnu (interval je dolg 2 uri) in glede na dan v tednu.

Vprašanja so zelo različna. Tudi najpogostejša vprašanja se pojavijo le krog 30-krat. Najpogostejša »vprašanja« so pozdravi. Če gledamo le prava vprašanja in najbolj podobna združimo dobimo naslednji vrstni red:

1. Kdo je XY? Povej telefonsko številko od XY?
2. Kaj znaš/delaš? Kdo si?
3. Kdo je direktor instituta? Informacije o direktorju?
4. Kaj je IJS? S čim se ukvarjate na IJS?

Na peto mesto pride kramljanje. Za kramljanjem pa ni več drugih izrazitih tem. Torej le malo ljudi sprašuje o IJS kaj več kot splošno vprašanje o tem, kaj delamo. S temi podatki lahko potrdimo, da je bila izbira pravilna, da Virtualni svetovalec ni le odgovarjalec na vprašanja, ampak ima tudi modul za klepet.

V tabeli 1 je statistična analiza delovanja virtualnega asistenta. Razvidno je, da je Robi dobil približno 1000 vprašanj na mesec. Odgovoril na 71% vprašanj. Vsak uporabnik je v povprečju zastavil 10 vprašanj. Ob temeljitem pregledu baze se izkaže, da je deviacija od povprečja velika, saj je bilo kar nekaj uporabnikov, ki so zastavili preko 100 vprašanj (vprašanja razvijalcev Robija niso všteta v spodnjo statistiko) in mnogo uporabnikov je zastavilo le eno ali dve vprašanji. Ker se nekateri odgovori delajo po obrazcu, je število različnih odgovorov, ki jih je sistem dal, večje od števila odgovorov v bazi (bistveno različnih odgovorov).

vseh vprašanj	11312
Vprašanja, na katera sistem ni znal odgovoriti	3236
iskanje informacij o osebi (kdo je ...)	1220
število različnih oseb, o katerih se je spraševalo	525
število IP, od koder so vprašanja prišla	1194
število različnih odgovorov, ki so bili dani	1265
število bistveno različnih odgovorov	583

Tabela 1: Virtualni asistent v številkah.

Dnevnike, ki se avtomatsko beležijo, smo obdelali tudi z metodami strojnega učenja.

Za analizo smo podatke uredili v obliko, ki je primerna za obdelavo v Weki. Za to smo 1265 odgovorov, ki so se pojavili v dnevnikih, združili v 583 bistveno različnih odgovorov. Za attribute smo vzeli IP in 583 odgovorov. Vrednosti atributov bi bile lahko enake številu, kolikokrat je določen IP dobil ta odgovor. Vendar v našem primeru regresijske metode, ki bi jih morali uporabljati, niso primerne, če pa attribute diskretiziramo, metode strojnega učenja vračajo slabe rezultate zaradi premnogih atributov, ki imajo različne vrednosti. Zato smo se odločili, da bo atribut imel vrednost d , če je IP dobil ta odgovor in n , če ga ni dobil. Na ta način smo dobili bazo podatkov s 1194 primeri in 584 atributi. Baza podatkov je zelo razpršena, kar pomeni, da imajo primeri večino atributov n in le nekaj d . V povprečju ima vsak primer približno 90% atributov z vrednostjo n .

Večina algoritmov za asociacijska pravila in grupiranje deluje slabše na razpršenih podatkih. Zato odstranimo attribute, ki imajo vrednost d le v redkih primerih. Na ta način smo dobili 568 atributov.

Zaradi razpršenih podatkov nam metode za izbiro najpomembnejših atributov niso dale dobrih rezultatov, zato

pa je grupiranje takoj razdelilo uporabnike v tri skupine. Za kriterij, po katerem se uporabniki razvrstijo v tri skupine, sta le dva atributa: spraševanje po človeku in odgovor »ne vem«. V podatkih se torej pojavijo tri močne skupine:

1. Uporabniki, ki so iskali podatke o osebah zaposlenih na IJS in so včasih dobili odgovor »ne vem«. (277)
2. Uporabniki, ki niso iskali podatkov o osebah na IJS in so včasih dobili odgovor »ne vem«. (419)
3. Uporabniki, ki niso iskali podatkov o osebah na IJS in nikoli niso dobili odgovora »ne vem«. (549)

Če pregledamo odgovore, vidimo da so v prvi in drugi skupini aktivnejši uporabniki, ki so postavljali več vprašanj. Ta so bila bolj specifična za področje, na katerem delajo. Uporabniki prve skupine so bili iz IJS, uporabniki iz druge skupine pa se pretežno nahajajo zunaj instituta. V tretji skupini so večinoma uporabniki, ki so malo poklepetali z Virtualnim svetovalcem in niso vrtali z vprašanji. Ta skupina je največja (46%).

ZAKLJUČEK

Virtualni asistent Robi je dokaz, da je možno z enostavnim sistemom doseči kvaliteto primerljivo z najdražjimi sistemi za odgovarjanje v slovenskem jeziku. Glede na množico vprašanj, ki jih je dobil in veliko število uporabnikov smo prepričani, da se bodo virtualni svetovalci obdržali na spletnih straneh in da se bo njihovo število še povečalo.

Razvoj pogovornih robotov ima že dolgo zgodovino [3], prav tako poznamo že mnoge uspešne sisteme za odgovarjanje na vprašanja [2], vendar danes Virtualni svetovalci presegajo priljubljenost obeh. Gotovo zato, ker združujejo prednosti obeh sistemov.

V članku [5] Minker piše, da je še mnogo izzivov za Virtualne asistente. Prvi preskok napoveduje v kratkem, ko bo omogočena interakcija z govorom. Ljudje govorno komunikacijo zelo povezujemo z inteligentnostjo. Na ta korak se tudi pri nas že pripravljamo s sistemom za sintezo govora – Govorec.

Na žalost kvaliteta odgovorov slovenskih virtualnih svetovalcev ne dosega najboljših v svetovnem merilu. Kljub temu pa poteka mnogo projektov, ki bodo to razdaljo zmanjšali.

LITERATURA

- [1] D. Verdonik, M. Sepesy Maučec. Pogovor z računalnikom. *Življenje in tehnika*, 6, junij 2008.
- [2] M.A. Liljenback, ContextQ:Experiments in Interactive Restricted-Domain Question Answering. Master Thesis. Computer Science department, Faculty of San Diego State University, fall 2007.
- [3] A. Turing. 1950. Computing machinery and intelligence, *Mind* 50: str. 433-460
- [4] J. Krivec, M. Ožek, M. Gams. Računalniški sistemi za avtomatski dialog. Zbornik multikonference Informacijska družba, 2008
- [5] W. Minker, R. Lopez-Cozar, M. McTear. The role of spoken language dialogue interaction in intelligent environments. *Journal of Ambient Intelligence and Smart Environments*, 1, 2009

ZAZNAVANJE NENAVADNEGA OBNAŠANJA S SISTEMOM ZA LOCIRANJE V REALNEM ČASU IN MEHKO LOGIKO

Rok Piltaver

Odsek za inteligentne sisteme

Institut Jožef Stefan

Jamova 39, 1000 Ljubljana, Slovenija

Tel: +386 1 477 3230

e-mail: rok.piltaver@ijs.si

Povzetek

Varnost v sodobnem okolju postaja vse bolj pomembna. Da bi znižali ceno zagotavljanja varnosti in dvignili nivo kvalitete smo razvili prototip inteligentnega sistema za nadzor gibanja ljudi in opreme, ki deluje v zaprtem območju. Programska oprema sistem je zgrajen modularno, pričujoči prispevek pa predstavlja enega izmed modulov, ki ga imenujemo Modul mehka logika. Prispevek predstavi osnovne principe delovanja, privzete predpostavke in omejitve ter se posveča tehničnim podrobnostim implementacije predlaganega algoritma. Poudarek je predvsem na prostorski in časovni učinkovitosti, ki modulu omogočata delovanje v realnem času in odpornosti na šum v vhodnih podatkih.

1 UVOD

V sodobnem času postaja varnost ljudi, lastnine in podatkov iz dneva v dan bolj pomembna. Običajne metode varovanja vključujejo fizično prisotnost varnostnega osebja in tehnične pripomočke, kot so alarmne naprave in senzorji gibanja. V obeh primerih je dejanska varnost odvisna predvsem od usposobljenosti, motiviranosti in psihofizičnih sposobnosti varnostnikov. Kot vsi ljudje so tudi varnostniki podvrženi utrujenosti, zmotljivosti in pristranskosti. Znano je na primer, da ljudje ob spremljanju posnetkov iz več video kamer zaradi monotoni naloge hitro postanejo utrujeni in zato pogosto spregledajo varnostne grožnje. Kot alternativa monotonemu človeškemu delu se ponuja možnost uporabe sodobne informacijske tehnologije in metod umetne inteligence, ki lahko zmanjša ali celo odpravi številne človeške pomanjkljivosti in obenem zniža stroške varovanja.

Opisana dejstva so nas motivirala za razvoj prototipa inteligentnega varnostnega sistema Poveljnikova desna roka (PDR). Podrobnejši opis sistema je objavljen v [7], zato bomo tu podali le najpomembnejše lastnosti sistema, ki so potrebne za razumevanje okvira predstavljenega dela. Namen sistema PDR je samodejno zaznavanje nenavadnega gibanja oseb in predmetov v stavbi z nekaj sobami ter opozarjanje operaterja na varnostne grožnje. Sistem je posebej prilagojen za varovanje arhivov tajnih

podatkov ter je sestavljen iz posebne strojne in programske opreme.

Podatke o gibanju oseb in pomembnih predmetov sistem pridobi s pomočjo sistema za lociranje v realnem času (ang. Real-time location system). Lociranje omogočimo tako, da na predmete in ljudi, ki jim želimo slediti, pritrdimo posebne značke, ki oddajajo signale (npr. radijske ali ultra-zvočne). Oddane signale sprejmejo senzorji postavljeni v nadzorovanem prostoru, specializirana programska oprema nato obdela izmerjene signale in izračuna 3D lokacije značk. Poleg sistema za lociranje v realnem času sistem PDR uporablja tudi več IP video kamer, ki skupaj s programsko opremo razvito na Fakulteti za elektrotehniko Univerze v Ljubljani omogočajo detekcijo in lociranje gibajočih se oseb in predmetov. Sistem video kamer je namenjen predvsem zaznavanju oseb brez značk, ker v sistemu PDR zahtevamo, da so vse osebe, ki se gibljejo v varovanem prostoru, označene z značko za sledenje v realnem času. Strojna oprema sistema PDR vključuje tudi mrežno infrastrukturo in strežnik, ki procesira zbrane podatke ter uporabniku pošilja opozorila in alarme o nenavadnem dogajanju v opazovanem prostoru.

Programska oprema, ki teče na strežniku, je razdeljena v več modulov. Vsak modul uporablja drugačno metodo umetne inteligence, ki omogoča zaznavanje skupine nenavadnih dogodkov. Video modul skrbi za detekcijo oseb brez značk. Modul Ekspertna pravila [6] omogoča uporabniku ekspliciten vnos prepovedanih dogodkov v obliki pravil, kar omogoča samodejno zaznavanje kršenja le-teh. Poleg opisanih modulov sistem PDR vsebuje še tri module, ki so se sposobni samodejno učiti vzorcev običajnega obnašanja in jih ločiti od neobičajnih vzorcev. V tem prispevku opisujemo Modul mehka logika, ki je eden od treh učečih modulov in je poimenovan po matematičnem modelu [10], ki ga uporablja. Preostala dva učeča modula sta Modul statistika, ki se uči statističnih lastnosti gibanja oseb in predmetov, ter Makro modul, ki se uči makroskopskih lastnosti obnašanja, kot so npr. običajen čas vstopa določene osebe v določen prostor, število obiskov prostora na dan ipd.

V naslednjem razdelku smo opisali ključne ideje in predpostavke predlagane metode za detekcijo neobičajnega obnašanja s pomočjo mehke logike in sistema za lociranje v realnem času. V razdelku tri smo opisali pomembne

tehnične detajle implementacije algoritma. V četrtem razdelku smo predstavili metodo za vizualizacijo znanja Modula mehka logika, v zadnjem razdelku pa smo podali zaključke, do katerih smo prišli v okviru opisanega dela.

2 KLJUČNE IDEJE PREDLAGANE METODE

Pri razvoju algoritma za zaznavanje nenavadnega obnašanja na podlagi podatkov o gibanju po prostoru smo sprejeli naslednji predpostavki:

1. Pogosti dogodki so običajni in zato nezanimivi.
2. Redki dogodki so zanimivi in verjetno niso dovoljeni.

O resničnosti predpostavk govori splošno znana teorija informacij, zato bomo na tem mestu našli posledice za uporabnika sistema PDR, ki izhajajo iz predpostavk. Uporabnika sistema ne smemo obveščati o pogostih dogodkih, ker mu takšna obvestila prinesejo le malo ali celo nič informacije in ga zato po nepotrebnem obremenjujejo. Pogosti dogodki so v okolju, v katerem večino časa ni nezaželenih oz. nedovoljenih dogodkov, dovoljeni in glede varnosti nezanimivi. Za primer pogledajmo obnašanje uradnika, ki dela v pisarni za svojo mizo. To počne zelo pogosto, zato sklepamo, da je sedenje za mizo njegova običajna aktivnost in da je dovoljena oz. celo zaželena. Po drugi strani so dogodki, ki se zgodijo le redko, zanimivi, saj je pri le-teh zelo verjetno, da gre za nezaželene ali prepovedane dogodke. Redko se na primer zgodi, da delavec omedli in pade na tla ter obleži. Za varovanje je odkrivanje takšnih in podobnih dogodkov zelo pomembno. Z avtomatskim opozarjanjem na redke dogodke je mogoče zagotoviti višjo stopnjo varnosti. Očitno je, da vsi redki dogodki niso nedovoljeni, so pa gotovo zanimivi vsaj za boljše razumevanje dogajanja v opazovanem prostoru.

Iz sklepanja v prejšnjem odstavku sledi najpomembnejša ideja predlagane metode za detekcijo nenavadnega obnašanja: za vsak opaženi dogodek je potrebno preveriti kako pogost je, če je redek je o tem potrebno obvestiti uporabnika varnostnega sistema PDR. Ideja naše metode je torej v tem, da »štejemo«, kako pogosto se zgodijo določeni dogodki, in da na podlagi števila podobnih dogodkov v preteklosti sklepamo o zanimivosti trenutnega dogajanja.

Dogodke v opazovanem prostoru delimo v dve skupini: dogodki, ki so posledica gibanja, in tisti, ki so posledica mirovanja. V gibanju je dogodek določen z lokacijo v prostoru ter smerjo in hitrostjo gibanja. V mirovanju je dogodek določen z lokacijo, stanjem opazovane osebe (leži, sedi, stoji) in trajanjem dogodka. Pri beleženju pogostosti dogodkov uporabljamo mehko diskretizacijo dogodkov, ki nudi številne prednosti. Za posamezen trenutek gibanja s pomočjo sistema za lociranje v realnem času izmerimo lokacijo v smereh x in y (v ravnini tal), ter ocenimo hitrosti v smereh x in y . Lokacijo v vsaki smeri diskretiziramo v razrede velikosti 50 cm. Opazovani prostor zato vidimo kot mrežo s kvadrati velikosti 50×50 cm. Iz hitrosti v

posamezni smeri izračunamo smer gibanja v ravnini tal ter jo diskretiziramo v eno od osmih možnih smeri (S, SV, V, JV, J, JZ, Z, SZ). Skalarno vrednost hitrosti diskretiziramo v štiri razrede (zelo počasno, počasno, običajno in hitro gibanje). Mehka diskretizacija posameznih lastnosti dogodkov ima naslednje štiri prednosti. Prva prednost je, da za shranjevanje podatkov o številu ponovitev posameznega dogodka potrebuje manj pomnilnika, natančnost pa se zaradi razmeroma fine diskretizacije bistveno ne zmanjša. Druga prednost je, da je časovna zahtevnost štetja zabeleženih dogodkov, ki so podobni danemu dogodku, zaradi diskretizacije konstantna in ni odvisna od števila zabeleženih dogodkov. Tretja prednost je, da mehka diskretizacija v sistem samodejno vnese linearno interpolacijo in tako omogoči natančnejše ocenjevanje pogostosti redkih dogodkov. Četrta prednost pa je, možnost hitrega posodabljanja števecv dogodkov, ki ima konstantno časovno zahtevnost.

Poleg predpostavk, ki govorijo o tem, da so redki dogodki bolj zanimivi od pogostih, je pomembna še tretja predpostavka:

3. Nenavadnost zaporedja dogodkov narašča s številom nenavadnih dogodkov v zaporedju oz. s trajanjem zaporedja.

Veljavnost predpostavke bomo ilustrirali na primeru: če se uradnik iz prejšnjega primera splazi pod pisalno mizo za kratek čas (ker npr. pobira svinčnik, ki mu je padel), je to manj nenavadno, kot če se pod mizo nahaja dlje časa. Stopnjo nenavadnosti opazovanega dogajanja zato računamo s pomočjo drsečega okna znotraj katerega računamo povprečno nenavadnost dogodkov. Z računanjem povprečja nenavadnosti po času se izognemo lažnim alarmom, ki bi jih sicer povzročile napake v meritvah položajev značk in kratka zaporedja redkih dogodkov. Zaporedje dogodkov razglasimo za nenavadno, če je v njem dovolj nenavadnih dogodkov oz. če so dogodki v oknu v povprečju dovolj redki.

Nenavadnost posameznega dogodka določimo tako, da primerjamo njegovo pogostost s pogostostjo ostalih zabeleženih dogodkov iz preteklosti. Uporabnik sistema v ta namen določi dve meji. Spodnja meja določa delež najbolj redkih dogodkov, ki jih obravnavamo kot popolnoma nenavadne in jim pripišemo maksimalno oceno nenavadnosti. Obratno določa zgornja meja delež najbolj pogostih dogodkov, ki jih obravnavamo kot povsem običajne in zaželene ter jim zato pripišemo nenavadnost enako 0. Nenavadnost dogodkov, ki so po pogostosti med obema mejama, linearno pada z naraščajočim deležem dogodkov, ki so redkejši od opazovanega dogodka.

3 IMPLEMENTACIJA ALGORITMA

Ker uporabljen sistem za lociranje v realnem času izmeri le lokacijo v 3D prostoru, je za delovanje predlagane metode potrebno iz zaporednih izmerjenih lokacij izračunati hitrosti v posameznih smereh, ugotoviti ali se človek, ki nosi značko premika ali miruje in v kakšnem stanju se nahaja:

leži, sedi ali stoji. V ta namen smo razvili tri primitivne rutine, ki jih uporabljajo tudi ostali moduli sistema PDR.

Prva primitivna rutina računa hitrosti z uporabo Kalmanovega filtra. Za stanje filtra smo izbrali vektor s šestimi komponentami: oddaljenosti od koordinatnega izhodišča v vsaki od treh smeri in hitrost premikanja vzdolž vsake osi. Novo stanje se iz prejšnjega izračuna kot vsota prejšnje lokacije z produktom hitrosti in časa med zaporednima meritvama za vsako smer posebej. Kovariančno matriko šuma meritev Kalmanovega filtra smo določili glede na specifikacijo natančnosti sistema za lociranje v realnem času, parametre šuma procesa, ki generira vhodne podatke za Kalmanov filter, pa smo uglasili eksperimentalno. Poleg ocenjevanja hitrosti je Kalmanov filter primeren tudi za glajenje šuma sistema za lociranje v realnem času.

Druga primitivna rutina je binarni klasifikator za ločevanje med gibanjem in mirovanjem, ki smo ga zgradili s pomočjo strojnega učenja. Razred klasifikatorja je gibanje ali mirovanje, atributi pa so izračunani znotraj drsečega okna, ki vsebuje po 20 zaporednih meritev položaja. Za attribute smo uporabili povprečno hitrost gibanja v oknu, dolžino prepotovane poti, razdalje med zaporednimi lokacijami značke in povprečno spremembo smeri gibanja. Za učenje smo uporabili približno eno uro ročno označenih posnetkov gibanja in mirovanja. Zaradi šuma v meritvah pozicij je klasifikacijska točnost za posamezen izmerjen položaj približno 95 %. Več o samem klasifikatorju je zapisano v [8].

Tretja primitivna rutina oceni stanja osebe (leži, sedi, stoji) iz zaporedja oddaljenosti značke od tal. Algoritem za določanje stanja ima 3 parametre. Prva dva parametra sta meji t_{lo} in t_{hi} med posameznimi stanji, tretji parameter pa je toleranca d . Algoritem v spominu hrani stanje v prejšnjem trenutku in glede nanj dinamično določi mejo med posameznimi stanji, tako da mejama t_{lo} in t_{hi} ustrezno prišteje ali odšteje vrednost parametra d . Če je trenutno stanje pod določeno mejo algoritem poveča mejo za d , sicer pa jo zmanjša za d . Kot trenutno stanje se določi tisto, ki je v opazovanem oknu zaporednih višin opazovane značke glede na dinamično postavljene meje najbolj pogosto. S primerno nastavitvijo parametra d , preprečimo preskakovanje med sosednjima stanjema, ki bi se sicer zgodilo kot posledica šuma v meritvah višine značke. Meji t_{lo} in t_{hi} smo določili s pomočjo strojnega učenja, tako da smo ju odčitali iz odločitvenega drevesa, ki glede na višino značke klasificira stanje osebe v ležanje, sedenje ali stanje. V ta namen smo posneli in označili pol ure ležanja, sedenja in stanja.

Vse tri primitivne rutine so robustne na šum v meritvah lokacij značk in so prilagojeni sistemu za lociranje v realnem času, ki ga uporabljamo v prototipu sistema PDR. Za določanje parametrov primitivnih rutin smo uporabili meritve šuma uporabljenega sistema za lociranje in strojno učenje v programskih paketih Orange [3] in Weka [9]. Drseča okna smo uporabili za dodatno zmanjšanje vpliva šuma in za glajenje zaporedij klasifikacij. Če bi imeli na

voljo natančnejši sistem za lociranje (npr. na ± 1 cm), bi lahko bile vse tri primitivne rutine bistveno enostavnejše in natančnejše. Kljub omejeni natančnosti uporabljenega sistema za lociranje (± 20 cm) so opisane primitivne rutine uporabne, ker sistem omogoča dovolj hitro osveževanje – pozicije posameznih značk zajema s frekvenco 9 Hz. Če sistem za lociranje ne bi bil dovolj natančen in bi imel poleg tega še prenizko frekvenco osveževanja, bi predlagane primitivne rutine odpovedale. Zato je kvaliteta sistema za lociranje ključnega pomena za dobro delovanje Modula mehka logika.

Poleg primitivnih rutin je za delovanje Modula mehka logika v realnem času pomembna tudi časovna in prostorska zahtevnost uporabljenih algoritmov. V nadaljevanju sledita opisa dveh najbolj pomembnih detajlov prototipa Modula mehka logika.

Za shranjevanje števecv dogodkov lahko uporabimo dve štiri dimenzionalni tabeli. Prva tabela vsebuje števce dogodkov v gibanju druga pa je namenjena števcem dogodkov v mirovanju. Vsaka od dimenzij tabele ustreza eni lastnosti dogodka (položaj v x ali y smeri, hitrost, smer, stanje, trajanje dogodka). Posodabljanje števecv v večdimenzionalnih tabelah ima konstantno časovno zahtevnost in je hitro, če so lahko celotne tabele za vsako osebo shranjene v notranjem pomnilniku računalnika. Če so tabele prevelike in jih je preveč postane delo z njimi zaradi omejene velikosti hitrega pomnilnika počasno. Druga slabost tabele je, da porabijo preveč pomnilnika: za vsako stanje namenijo svoj števec, tudi če je ta enak nič. V praksi se mnogi dogodki nikoli ne zgodijo in imajo zato njihovi števci vedno vrednost nič. Tabela je torej redka (sparse) in zato pomnilniško neučinkovita.

Druga možnost je, da namesto tabele uporabljamo npr. rdeče-črno [2, 4] ali AVL drevo [1]. V drevesu so shranjeni le števci, ki so večji od nič. Na ta način je mogoče prihraniti pri velikosti potrebnega pomnilnika. Za ključ v drevesu uporabimo indeks iz opisane štiri dimenzionalne tabele, za vrednost pa števec. Iskanje, posodabljanje vrednosti in vstavljanje novih števecv v drevo ima logaritemsko časovno zahtevnost v odvisnosti od števila vstavljenih števecv, kar je nekoliko počasneje od operacij v večdimenzionalni tabeli, vendar je v praksi dovolj hitro za delo v realnem času. Odločitev med tabelo in drevesom je torej odvisna predvsem od števila opazovanih oseb, velikosti opazovanega prostora in od deleža predvidenih dogodkov, ki se nikoli ne zgodijo.

Drugi pomemben detajl implementacije Modula mehka logika je povezan z učinkovitostjo računanja nenavadnosti posameznih dogodkov. Ker je v prvi fazi izračuna potrebno po velikosti sortirati vse števce pogostosti posameznih dogodkov, je pomembna predvsem izbira algoritma za sortiranje. V ta namen smo uporabili algoritem quicksort [5], ki je v praksi hitrejši od ostalih algoritmov za sortiranje s časovno zahtevnostjo $\Theta(n \log n)$. Ker se vrstni red števecv spreminja počasi, je potrebno sortiranje pognati le občasno; med posameznimi sortiranjmi privzamemo

približno sortiran vrstni red števecv kot dovolj dober približek za ocenjevanje pogostosti posameznih dogodkov.

V drugi fazi računanja nenavadnosti posameznih dogodkov je potrebno s pomočjo sortirane tabele ugotoviti, kakšen delež dogodkov ima manjšo pogostost od danega dogodka. Za iskanje po urejeni tabeli smo uporabili algoritem binarnega iskanja, ki ima časovno zahtevnost $O(\log n)$, kjer je n število elementov v tabeli.

4 VIZUALIZACIJA

Ko Modul mehka logika zazna nenavadno gibanje o tem obvesti uporabnika z opozorilom. Ker opozorilo tipa: "Janez Novak se giblje nenavadno," za uporabnika ni dovolj informativno smo razvili metodo za vizualizacijo znanja, ki ga ima Modul mehka logika o opazovanih dogodkih.

V vsak kvadrat mreže, s katero razdelimo opazovani prostor ob mehki diskretizaciji, narišemo graf v obliki tarče. Vsak kos tarče predstavlja posamezen dogodek gibanja. Koncentrični krogi "tarče" predstavljajo hitrost gibanja ob dogodku: kolobarji z manjšim premerom pomenijo manjšo hitrost, tisti z večjim polmerom pa večjo hitrost. Trikotni izrezi "tarče" pa predstavljajo smer gibanja. Vsak kos tarče pobarvamo s sivino, ki je odvisna od pogostosti dogodka, ki ga posamezen kos opisuje. Bela barva pomeni, da se dogodek še nikoli ni zgodil, črna barva pa pomeni, da je dogodek najbolj pogost izmed vseh opaženih dogodkov. Dogodke, ki so v trenutnem drsečem oknu za izračun nenavadnosti, pobarvamo podobno, a uporabimo barvno skalo od rdeče (za nenavadne dogodke) do zelene (za običajne dogodke), da izbrane dogodke ločimo od ostalih.

Za dogodke v mirovanju namesto grafov v obliki tarče uporabimo grafe v obliki tabel. Vsaka vrstica tabele predstavlja po eno stanje (ležanje, sedenje, stanje), stolpci pa predstavljajo trajanje dogodka. Levi stolpec pomeni trajanje ene sekunde, vsak naslednji stolpec pa predstavlja dvakrat daljši dogodek.

Iz takega grafa lahko uporabnik z malo prakse hitro razbere v katerem delu prostora je prišlo do nenavadnega dogodka in kakšen je ta dogodek bil. Poleg razlage opozoril je vizualizacija primerna tudi za splošno analizo dogajanja v opazovanem prostoru.

5 ZAKLJUČEK

V prispevku smo predstavili enega izmed učečih modulov inteligentnega varnostnega sistema Poveljnikova desna roka, ki je namenjen nadzoru nad gibanjem ljudi in pomembne opreme v omejenem prostoru. Modul bazira na predpostavki, da so pogosti dogodki običajni in nezanimivi, med tem ko so redki dogodki zanimivi, ker so z veliko verjetnostjo nedovoljeni ali vsaj nezaželeni. Modul Mehka logika deluje tako, da šteje število ponovitev podobnih si dogodkov in glede na število opaženih podobnih dogodkov v preteklosti izračuna nenavadnost krajšega zaporedja opazovanih dogodkov.

Modul mehka logika določi tip dogodka iz zaporedja 3D lokacij značke, ki jo nosi opazovana oseba. Za klasificiranje dogodkov smo razvili tri primitivne rutine, ki ocenijo kako hitro se giblje značka, ali se premika ali miruje in ali oseba, ki nosi značko, leži, sedi ali stoji. Za shranjevanje števecv dogodkov smo predlagali dve podatkovni strukturi, ki smo ju primerjali po prostorski in časovni zahtevnosti. Za računanje nenavadnosti dogodkov smo predlagali poseben algoritem in razložili vpliv implementacije predlaganega algoritma na sposobnost delovanja modula v realnem času. Za potrebe razlage opozoril, ki jih opisani modul pošlje uporabniku, in analizo dogajanja v opazovanem prostoru smo razvili in opisali metodo za vizualizacijo znanja, ki ga modul hrani o preteklem dogajanju v opazovanem prostoru. Opisani modul je sposoben zaznati veliko število nezaželenih dogodkov in o njih v razumljivi obliki opozoriti uporabnika sistema Poveljnikova desna roka.

Literatura:

- [1] G. Adelson-Velskii, E. M. Landis. "An algorithm for the organization of information", *Proceedings of the USSR Academy of Sciences* 146, str. 263–266, 1962, (rusko), angleški prevod: M. J. Ricci, *Soviet Math. Doklady*, 3 str. 1259–1263, 1962.
- [2] R. Bayer. "Symmetric Binary B-Trees: Data Structures and Maintenance Algorithms", *Acta Informatica*, 1, str. 290–306, 1972.
- [3] J. Demšar, B. Zupan, G. Leban. "Orange: From Experimental Machine Learning to Interactive Data Mining", White Paper (www.aillab.si/orange), Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, 2004.
- [4] L. J. Guibas, R. Sedgewick. "A Dichromatic Framework for Balanced Trees", *Proceedings of the 19th Annual Symposium on Foundations of Computer Science*, str. 8–21, IEEE Computer Society, 1978.
- [5] C. A. R. Hoare. "Quicksort", *Computer Journal* 5 (1), str. 10–15, 1962.
- [6] R. Piltaver, G. Matjaž. "Expert system as a part of intelligent surveillance system", *Zbornik osemnajste mednarodne Elektrotehniške in računalniške konference ERK 2009*, Zv. B, str. 191–194, 2009.
- [7] R. Piltaver, G. Matjaž. "Poveljnikova desna roka", *Zbornik sedemnajste mednarodne Elektrotehniške in računalniške konference ERK 2008*, Zv. B, str. 141–144, 2008.
- [8] R. Piltaver. "Strojno učenje pri načrtovanju algoritmov za razpoznavanje tipov gibanja", *Zbornik 11. mednarodne multikonference Informacijska družba - IS 2008*, str. 13–17, 2008.
- [9] I. H. Witten, E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques* (2nd edition), Morgan Kaufmann, 2005.
- [10] L. A. Zadeh. "Fuzzy sets", *Information and Control* 8 (3), str. 338–353, 1965.

AVTOMATSKO RAZPOZNAVANJE BOLEZNI IZ GIBANJA

Bogdan Pogorelc^{1,2}

¹Odsek za inteligentne sisteme
Institut Jožef Stefan
Jamova c. 39, 1000 Ljubljana, Slovenija
Tel: +386 1 477 32 30
e-mail: bogdan.pogorelc@ijs.si

²Špica international d.o.o.
Pot k sejmišču 33
1231 Ljubljana
Slovenija

POVZETEK

Prispevek predstavlja detekcijo bolezni, ki se odražajo na gibanju. Gibanje je zajeto s sistemom za zajem gibanja, ki je nameščen na telesu uporabnika v obliki značk. Izhod sistema je inteligentni sistem, ki ločuje med 5 oblikami hoje: s hemiplegio, s Parkinsonovo boleznijo, z bolečo ного, z bolečim hrbtom, normalno. Klasifikacijske točnosti od 85-95% so dosežene. Tudi študija vpliva postavitve značk sistema za zajem gibanja in stopnje šuma na klasifikacijsko točnost detekcije zdravstvenih težav je predstavljena. Lahko je uporabljena za pomoč v prihodnjih študijah vedno bolj pomembnega področja ambientalno podprtega bivanja.

1 UVOD

Razvite države se soočajo s hitro rastjo svojega prebivalstva. Do leta 2050 naj bi se število ljudi nad 60 let v Evropi podvojilo na 40% celotnega prebivalstva oz. na 60 % aktivnega prebivalstva [1]. Posledično bi to pomenilo, da mlajši ne bi mogli več zadostno skrbeti za ostarele. V izogib temu problemu se po vsem svetu išče nove tehnične rešitve, ki bi omogočale starejšim daljše samostojno življenje ob minimalni podpori delovno aktivnega prebivalstva. To je tudi cilj dela, ki je predstavljeno v tem prispevku.

V prispevku predlagamo inteligentni sistem za spremljanje starejših z namenom da razpoznamo nekaj najbolj pogostih in kritičnih zdravstvenih težav starejših, ki se odražajo na gibanju. Te so: hemiplegia (navadno po možganski kapi), Parkinsonova bolezen, bolečina v nogi in bolečina v hrbtu.

Sistem je zasnovan tako, da uporabnik na sebi nosi določeno število značk. Le te so razpoznane s senzorji, kar omogoča razpoznavo potencialnih bolezni. Cilj predstavljene raziskave je dvojen. Prvič, da ugotovimo, do kake mere je možna avtomatska razpoznavna bolezni, in drugič, da raziščemo kakšne klasifikacijske točnosti dobimo ob variiranju števila značk in dodajanju šuma. Postavitev značk je kompromis med uporabnostjo in točnostjo – uporabniki si želijo nositi čim manj značk, a

premalo značk ne more zagotoviti primerne točnosti razpoznavanja.

2 SORODNO DELO

Za avtomatsko razpoznavo bolezni moramo najprej zajeti gibanje. Za ta namen se lahko uporabijo inercialni senzorji, ki so ponavadi v obliki MEMS (Mikro-Elektro-Mehanski Sistemi) sistemov, ki jih sestavljajo pospeškometri in žiroskopi [2]. Za rekonstrukcijo človekovega gibanja se prav tako uporablja procesiranje videa. Če ti sistemi vključujejo značke za namestitev na telo za lažje lociranje, se slika zajame z infrardečimi (IR) kamerami in po računalniški obdelavi signalov dobimo pozicije značk [10].

Obstajajo pa tudi bolj specifični sistemi, npr. za zaznavanje tremorja (tresenja) pri Parkinsonovi bolezni se uporabljajo senzorji za zaznavo upogiba sklepov [3] in elektromiografija.

Za namen naše raziskave je bil uporabljen sistem IR-kamer z značkami nameščenimi po telesu [10]. Iz časovnih vrst pozicij značk smo s pomočjo metod strojnega učenja klasificirali ljudi v zdrave oz. v ljudi z eno od 4 bolezni.

Nismo se ukvarjali z razpoznavo vsakodnevnih aktivnostih, ker je bil slednji problem že velikokrat rešen, npr. v [6, 12] ampak smo se soočili z nekoliko drugačnim izzivom.

Pregled sorodnega dela je pokazal, da se prepoznave bolezni navadno lotevajo s podobnim sistemom za zajem gibanja, kot smo ga uporabili mi, za razliko od našega pristopa pa izrisane trajektorije posameznih delov telesa medicinsko usposobljeni ljudje pregledujejo in ocenjujejo ročno. Prav tako tudi v primeru [16], ki predstavlja sistem za dolgoročno spremljanje pacienta s Parkinsonovo boleznijo (PD) s pomočjo pospeškometra.

Uporaba podobnega sistema za zajem gibanja je v primeru [11] uporabljena za avtomatsko razpoznavo bolezni kot so hemiplegia in diplegia. Klasifikacijska točnost 92.5% je bila dosežena s pomočjo valjčne transformacije in SOM (Self-Organizing Maps) nevronske mreže. Kljub manjšemu številu značk in večjemu dodanemu šumu (in verjetno tudi manjši frekvenci

zajemanja našega sistema – ni podatka za njihovega) se je naša metoda odrezala primerljivo.

Pomemben del raziskave, predstavljene v tem prispevku, je študija vpliva števila in postavitve značk ter dodanega šuma na klasifikacijsko točnost. Vir [5], ki obravnava razpoznavo padcev, predstavlja tudi študijo optimalne postavitve pospeškometrov na telo. Avtorji ugotavljajo, da postavitev na glavo izboljša klasifikacijsko točnost, a ker je nepraktična, predlagajo postavitev za pas.

3. SISTEM ZA RAZPOZNAVO BOLEZNI IZ GIBANJA

3.1 Izbor bolezni

Osredotočili smo se na 4 zdravstvene težave po nasvetu medicinskega eksperta. Specifične zdravstvene težave smo izbrali glede na pogostost v starosti nad 65 let, medicinsko pomembnost in zmožnost razpoznave iz gibanja. Tako smo dobili:

- **Parkinsonova bolezen:** degenerativna bolezen možganov (centralnega živčnega sistema), ki pogosto prizadene motorične sposobnosti, govor in druge funkcije. Pogost simptom je tremor (tresenje), togost in posturalna nestabilnost. Tremor je prisoten na udih, ki v tistem trenutku niso obremenjeni.
- **Hemiplegia:** je paraliza roke, noge in trupa po eni strani telesa. Ponavadi je rezultat možganske kapi, a tudi bolezni možganov ali hrbtenjače lahko povzročijo podobno stanje. Ker vpliva na gibanje, posebno na hojo, povzroča nestabilnost in je lahko vzrok za padce.
- **Bolečina v nogi:** vzorec hoje je nekoliko podoben vzorcu hoje po kapi, saj je prisotno šepanje. Pri starejših gre ponavadi za bolečino v kolenu ali boku.
- **Bolečina v hrbtu:** prisotna je podobnost s hemiplegio in bolečo nogo, saj gre za nesimetričnost korakov zaradi nagnjenosti hrbta v eno stran in bolečine pri obremenitvi. Nesimetričnost korakov kljub temu ni tako izrazita kot v prejšnjih dveh primerih zdravstvenih težav.

3.2 Konstrukcija značilk

Zdravnik ponavadi diagnosticira obravnavane zdravstvene probleme iz opazovanja hoje [7]. Ker so si vzorci podobni, mora biti pozoren na veliko detajlov, ki smo jih poskušali zapisati z merljivimi spremenljivkami. Za nalogo avtomatske prepoznave bolezni smo predlagali in testirali uporabo naslednjih 13 značilk (predvidevajo, da je oseba prizadeta po desni strani, ker če bi bila po levi, bi se strani zamenjale):

1. Absolutna razlika povprečnih razdalj desni komolec – desni bok in desno zapestje – levi bok.

2. Povprečni kot desnega komolca.
3. Kvocien med maksimalnim kotom levega in maksimalnim kotom desnega kolena.
4. Razlika med maksimalnim in minimalnim kotom desnega kolena.
5. Razlika med maksimalno in minimalno višino levega ramena.
6. Razlika med maksimalno in minimalno višino desnega ramena.
7. kvocien {razlike med maksimalno in minimalno višino levega in maksimalno in minimalno višino desnega }
8. Absolutna razlika {razlike maksimalne in minimalne hitrosti levega in razlike maksimalne in minimalne hitrosti desnega gležnja}
9. Absolutna razlika povprečne razdalje desna rama – levi komolec in leva rama – desno zapestje.
10. Povprečna hitrost desnega zapestja.
11. Pogostost desnega komolca prečkanja svoje povprečne vrednosti kota.
12. Povprečni kot med vektorjem {desna rama – desni bok} in vektorjem {desna rama – desno zapestje}
13. Absolutna razlika povprečnih višin desnega in levega ramena.

4 MODELIRANJE BOLEZNI

Uporabili smo 122 posnetkov 5 oseb, ki je izvajalo naslednje aktivnosti:

- 25 posnetkov normalne hoje.
- 25 posnetkov hoje s hemiplegio.
- 25 posnetkov hoje s Parkinsonovo boleznijo.
- 25 posnetkov hoje z bolečinami v nogi.
- 22 posnetkov hoje z bolečinami v hrbtu.

Zaradi nedosegljivosti ljudi z navedenimi zdravstvenimi težavami so poskuse izvedli zdravi prostovoljci, ki so imitirali paciente, tako, da so posnemali hojo, ki jim jo je pokazal zdravnik.

Posnetki so nastali z 12 značkami, pritrjenimi na ramenih, komolcih, zapestjih, bokih, kolenih in gležnjih s frekvenco zajema 10 Hz. Za zajem je bil uporabljen IR sistem za zajem gibanja Smart [10] s standardno deviacijo šuma 0.5 mm.

Da bi testirali robustnost predlaganega pristopa smo dodajali spremenljivo stopnjo Gaussovega šuma h koordinatam, ki so izhod sistema Smart. Da bi bil dodatek šuma pristnejši, smo v drugem poskusu dodajali Gaussov šum s standardno deviacijo kot naj bi jo imel Ubisense UWB sistem za lociranje v realnem času [14], to je 4.36 cm vodoravno in 5.44 navpično. Tako se v nadaljevanju sklicujemo na večkratnike Ubisenseovega šuma. V procesu predobdelave smo zašumljene podatke filtrirali s Kalmanovim filtrom [13].

Naloga strojnega učenja je bilo klasificirati hojo v 5 razredov: 4 razredi hoje z boleznijo in 1 razred normalna hoja. Klasifikator je bil naučen na posnetkih, ki so bili označeni s pravkar omenjenimi razredi. Za vsak posnetek

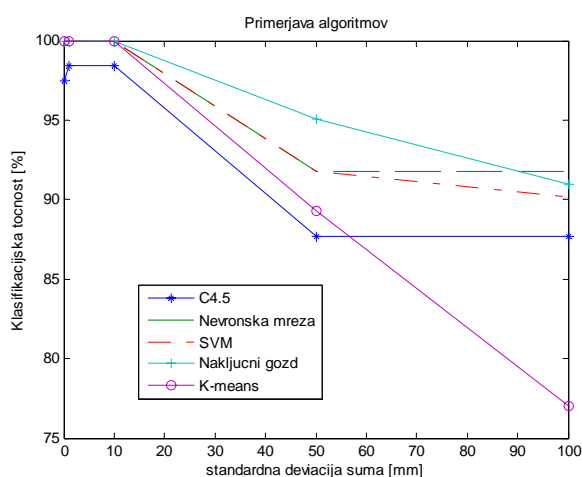
smo dobili 13-dimenzionalni vektor značilke, ki je bil uporabljen z večimi algoritmi strojnega učenja [4], od katerih se je metoda podpornih vektorjev (SVM) [15] najbolje izkazala. Ker je nošenje vseh 12 značk nekoliko moteče, smo iskali načine, kako bi njihovo število zmanjšali. Zanimalo nas je tudi variranje šuma hkrati s številom značk.

Rezultati prvega poskusa so bili narejeni z 10-kratnim prečnim preverjanjem, v drugem poskusu pa smo uporabili metodo "spusti enega človeka", kjer smo naučili model na štirih in testirali na petem. Razlog za to je bil, da preprečimo pretirano prileganje modelov (klasifikatorjev) učnim podatkom, se pravi, da rezultati kažejo pričakovano točnost na še nevidnih ljudeh. V praktični uporabi bi človek začel nositi sistem, ko bi bil zdrav, a sistem bi bil naučen na ostalih bolnih ljudeh.

V prvem poskusu smo dodajali Gaussov šum s standardno deviacijo od 0-100 mm. Preizkusili smo 4 algoritme nadzorovanega učenja in enega za nenadzorovano učenje v Weki. Rezultate prikazuje Tabela 1 in Slika 1.

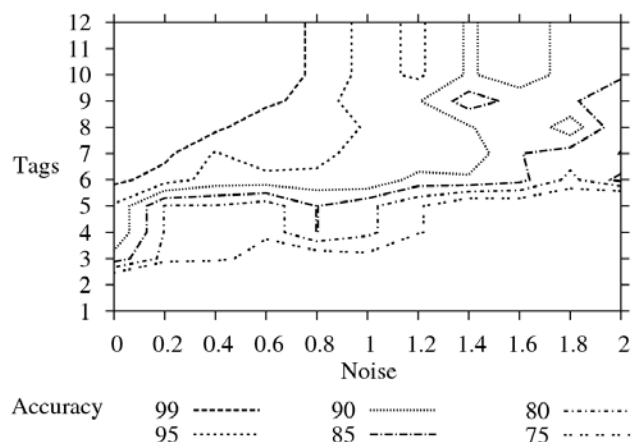
		Standardna deviacija šuma				
		0mm	1mm	1cm	5cm	10cm
Algoritmi	C 4.5	97.5	98.4	98.4	87.7	87.7
	Nevronska mreža	100	100	100	91.8	91.8
	SVM	100	100	100	91.8	90.2
	Naključni gozd	100	100	100	95.1	91.0
	K-means	100	100	100	89.3	77

Tabela 1: Klasifikacijske točnosti [%] in odstotek pravilno razporejenih vzorcev hoje za 4 nadzorovane in nenadzorovano metodo strojnega učenja in z dodatkom različnih stopenj Gaussovega šuma.



Slika 1. Primerjava algoritmov.

V drugem poskusu smo varirali št. značk in količino dodanega šuma. Količino šuma smo spreminjali od 0 do $U_{\text{bisense}} \times 2$ v inkrementih $U_{\text{bisense}} \times 0.2$. Slika 2 prikazuje klasifikacijsko točnost za najboljšo kombinacijo 12, 11,...1 značk in za spreminjanje šuma.



Slika 2. Klasifikacijska točnost razpoznavne bolezni [%] glede na število značk in dodan šum.

Slika 2 prikazuje, da je pri Ubisensovem šumu klasifikacijska točnost 95 % nedosegljiva in da za klasifikacijsko točnost 90% rabimo vsaj 6 značk. V zgornjem levem kotu grafa je področje z izjemno visoko klasifikacijsko točnostjo.

5 ZAKLJUČEK

Predstavili smo sistem za avtomatsko prepoznavo bolezni, ki se odražajo v gibalnem vzorcu, starejših. Uporabljen je bil infrardeči sistem za zajem gibanja. Trajektorije zaznanih pozicij značk so bile pretvorjene v obliko za metode strojnega učenja, pri čemer smo konstruirali primerne značilke.

Avtomatska detekcija zdravstvenih težav je redko uporabljena, a naši rezultati so pokazali, da se da doseči klasifikacijske točnosti med 85–95 % z določeno stopnjo šuma.

Preizkusili smo tudi vpliv postavitve značk in stopnje šuma na klasifikacijsko točnost detekcije zdravstvenih težav. V splošnem je več šuma oz. manj značk povzročilo manjšo klasifikacijsko točnost. Rezultati so uporabni tudi v vse bolj pomembnem področju ambientalno podprtega bivanja.

ZAHVALA

Operacijo delno financira EU, in sicer iz Evropskega socialnega sklada. Avtor se zahvaljuje Martinu Tomšiču, Bojanu Nemcu in Leonu Žlajpahu in drugim za pomoč pri pripravi podatkov, Antonu Gradišku za medicinsko

ekspertizo in Matjažu Gamsu ter Zoranu Bosniću za diskusijo.

VIRI

1. S. Toyne, Ageing: Europe's growing problem. BBC News. <http://news.bbc.co.uk/2/hi/business/2248531.stm>, 2009-01-19.
2. D. Strle, V. Kempe, "MEMS-based inertial systems", Informacije MIDEM 37(2007)4 Ljubljana, pp. 199-209.
3. S. Ribarič, J. Rozman, "Sensors for measurement of tremor type joint movements", Informacije MIDEM 37(2007)2, Ljubljana, pp. 98-104.
4. Kononenko, I., Kukar, M. Machine Learning and Data Mining: Introduction to Principles and Algorithms. Horwood publ., 2007.
5. Bourke, A.K., Scanaill, C.N., Culhane, K.M., O'Brien, J.V., and Lyons, G.M. An optimum accelerometer configuration and simple algorithm for accurately detecting falls. In *Proc. BioMed 2006* (2006), 156–160.
6. Confidence: Ubiquitous Care System to Support Independent Living. <http://www.confidence-eu.org>.
7. Craik R., and Oatis C. *Gait Analysis: Theory and Application*. Mosby-Year Book (1995).
8. Harrison RD. *Harrison's principles of internal medicine*. New York: McGraw-Hill, 14th edition, 1998.
9. Perry J. *Gait Analysis: Normal and Pathological Function*. McGraw-Hill, Inc., 1992.
10. eMotion. Smart motion capture system. <http://www.emotion3d.com/smart/smart.html>.
11. Lakany, H. Extracting a diagnostic gait signature. *Pattern recognition 41* (2008), 1627–1637.
12. Luštrek, M., and Kaluža, B. Fall detection and activity recognition with machine learning. *Informatica 33*, 2 (2009), accepted in publication.
13. Maybeck, P.S. Stochastic models, estimation, and control. *Mathematics in Science and Engineering 141* (1979).
14. UbiSense. <http://www.ubisense.net>.
15. Witten, I.H., and Frank, E. *Data Mining: Practical Machine Learning Tools and Techniques (2nd edition)*. Morgan Kaufmann (2005).
16. Moore ST, et al., Long-term monitoring of gait in Parkinson's disease, *Gait Posture* (2006), doi:10.1016/j.gaitpost.2006.09.01

IDENTIFIKACIJA GLASOV IN SODNO IZVEDENSTVO V KAZENSKEM POSTOPKU

Tomaž Šef

Odsek za inteligentne sisteme, Institut "Jožef Stefan"
Jamova cesta 39, 1000 Ljubljana, Slovenija
tomaz.sef@ijs.si

POVZETEK

Kriminaliteta lahko zaradi številnih negativnih socialnih, ekonomskih in družbeno-moralnih posledic prizadene pomemben delež prebivalstva ter ogrozi nacionalno varnost v celoti. Zato je potreben odločen boj zoper njo. Identifikacija glasov lahko pomaga v katerikoli fazi postopka; bodisi že na samem začetku preiskave (predkazenski postopek) ali pa v fazi sojenja (kazenski postopek).

Različne metode identifikacije glasov so lahko bolj ali manj subjektivne oz. objektivne. **V forenzični praksi te metode vedno povezujemo**, s čimer dobimo kombinirano oceno zanesljivosti dokaznega gradiva.

Na Institutu Jožef Stefan trenutno gradimo novo **specializirano govorno bazo** za potrebe razpoznavanja govorcev v forenzične namene. Na njeni osnovi želimo podrobneje ovrednotiti razlike v snemalnih pogojih, ki nastopajo pri posnetkih iz sodne prakse, kompenzirati njihove učinke ter kvantitativno opredeliti negotovost oz. nezanesljivost, ki je posledica teh dejavnikov.

1 UVOD

Soočenje s kriminaliteto v nacionalnem in mednarodnem okviru **zahteva celovit, načrten in sistematičen odziv državnih institucij, civilne družbe in vseh državljanov** na nacionalni, regionalni in lokalni ravni. Kriminaliteta ima strukturne učinke. Hromi zaupanje v delovanje državnih institucij ter spodbuja družbene konflikte, ki lahko vodijo v ekstremitetizem.

Kriminaliteta po drugi strani predstavlja resno grožnjo demokraciji tudi v primeru poenostavljenih rešitev na področju njenega preprečevanja in zatiranja, kar lahko omogoči kršitev človekovih pravic in svoboščin ter dostojanstva.

Zato je **potrebno nenehno iskanje ravnovesja med učinkovitostjo boja zoper kriminaliteto in pravili, ki jih zahteva vladavina prava ter spoštovanja človekovih pravic in svoboščin**. Navedeno pa moramo upoštevati tudi v zvezi z razpoznavanjem govorcev v kazenskem postopku.

2 SPLOŠNO O SODNEM IZVEDENSTVU V KAZENSKEM POSTOPKU

Določbe o sodnem izvedenstvu v Republiki Sloveniji vsebuje Zakon o kazenskem postopku (v nadaljevanju: ZKP); v členih: 248 do 267. Kadar je za ugotovitev ali za

presojo kakšnega pomembnega dejstva potrebno dobiti izvid in mnenje nekoga, ki ima potrebno strokovno znanje, se odredi izvedenstvo (248. člen ZKP).

Izvedenec (expertus) je torej oseba določena s strani procesnega organa, ki s svojim strokovnim znanjem pomaga procesnemu organu pri ugotavljanju pomembnih dejstev v kazenskem postopku.

Izvedenstvo (expertisis) je procesno dejanje s katerim se dobi poseben dokaz (v materialnem smislu): izpovedba izvedenca. Izpovedba izvedenca je vir iz katerega črpa procesni organ svoje znanje o dejstvih. Izvedenstvo pomeni tako tudi pomoč sodišču pri izvrševanju njegove funkcije. Izvedenstvo se nanaša na dejanska in ne na pravna vprašanja.

Skupna značilnost izpovedbe izvedenca in izpovedbe priče je, da gre v obeh primerih za zaznavanje dejstva s strani subjekta, ki je različen od procesnega organa. **Bistvena razlika med njima** [1] pa je v tem, da:

- priča izpove o dejstvih preteklosti, izvedenec pa o dejstvih sedanjosti;
- priča izpove o dejstvih, ki jih je zaznala izven postopka, izvedenec pa o dejstvih v postopku in po pisni odredbi organa, ki vodi postopek. Ker priča izpove o dejstvih, ki jih je sama zaznala je zaradi tega nezamenljiva. Ravno nasprotno pa praviloma velja za izvedenca. Le izjemoma se za izvedenca sukcesivno vzame oseba, ki je bila zaslišana kot priča (3. odst. 251. člena ZKP);
- priča načeloma le reproducira svoja zapažanja, izvedenec pa daje tudi mnenje (čeprav se tudi v izpovedi priče odraža njeno mnenje, ki si ga nedvomno ustvari o zadevi, le da je to mnenje v ozadju v primerjavi z njenimi zapažanji dejstev);
- da se za izvedenca (za razliko od priče) zahteva strokovna in pravna sposobnost.

Zahtevana strokovna sposobnost je odvisna od konkretnega izvedenstva, ki ga je potrebno opraviti. Sicer pa pomeni strokovna sposobnost, da je izvedenec za ugotavljanje in presojo dejstev strokovno usposobljen oziroma ima potrebno strokovno znanje.

Opravljanje izvedenstva je državljanska dolžnost, kar pomeni, da je izvedenec dolžan sprejeti izvedenstvo. **Obveznosti izvedenca** so:

- a) odzvati se vabilu organa (1. odst. 250. člena ZKP),
- b) podati svoj izvid in mnenje,
- c) govoriti resnico in priseči (1. in 2. odst. 252. člena ZKP).

Kršitev te državljanske dolžnosti ima lahko za posledico:

- a) denarno kazen (določeno v 1. odst. 78. člena ZKP),
- b) prisilno privedbo (2. odst. 250. člena ZKP) ali
- c) kaznivo dejanje krive izpovedbe (1. odst. 252. člena ZKP v zvezi z 289. členom KZ).

Izvedenec opravi izvedensko delo tako, da predmet izvedenstva (ki ga določa sodišče s svojimi vprašanji) skrbno pregleda, natančno navede vse kar opazi in dožene in poda svoje mnenje nepristransko in **v skladu s pravili znanosti ali strokovnega znanja** (1. odst. 252. člena ZKP). Izvedenec lahko predlaga, naj se izvedejo dokazi ali priskrbijo predmeti ali podatki, ki so pomembni za izvid in mnenje. ZKP določa procesni vidik izvedenstva. Vsebinski vidik izvedenstva pa določajo pravila stroke [1].

Izvid (visum repertum) je oporoka ali zapis tega, kar je izvedenec pri ogledu z opazovanjem dognal. Gre za ugotavljanje dejstev za katera je potrebno posebno strokovno znanje.

Mnenje (parere) je izvedenčeva strokovna presoja tega, kar je ugotovil z izvidom, in obsega tudi razloge (utemeljitev) za tako presojo.

V zapisniku o izvedenskem delu (v katerega se vpišeta izvid in mnenje) ali v pisnem izvidu ali mnenju je potrebno navesti, kdo je to delo opravil, ter njegov poklic, strokovno izobrazbo in specialnost (256. člen ZKP).

Zaradi zagotovitve pravilnosti izvedenstva in nadzora nad izvedenstvom, lahko organ pred katerim teče kazenski postopek, stranka in zagovornik (neposredno) postavljajo vprašanja izvedencem (7. odst. 178. člena in 1. odst. 334. člena ZKP) in dajejo pripombe na njihov izvid in/ali mnenje. Če so v mnenju izvedencev nasprotja ali pomanjkljivosti ali če nastane utemeljen dvom o pravilnosti danega mnenja, pa se te pomanjkljivosti ali dvom ne dajo odpraviti z novim zaslišanjem, se zahteva mnenja drugih izvedencev (258. člen ZKP) [1].

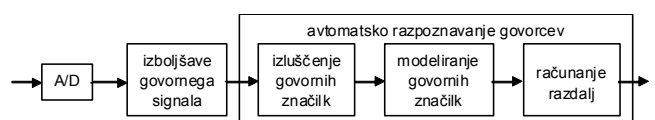
Glede ocene dokazne vrednosti izpovedbe (izvida ali mnenja) izvedenca velja enako, kot glede ocene kateregakoli drugega dokaza, torej »načelo proste presoje dokazov«, kar pomeni, da sodišče pri presoji ni vezano na nobena formalna pravna pravila o vrednosti nekega dokaza. Je pa ta »prosta presoja« izpovedbe izvedenca omejena s splošnimi pravili človekovega mišljenja in z izkustvenimi pravili: sodišče mora tako oceniti izpoved izvedenca, da je ta ocena razumljiva in sprejemljiva za vsakega nepristranskega in (pravno) poučenega človeka. V zvezi s tem je sodišče v svoji »prosti presoji« omejeno tudi z dolžnostjo obrazložitve svoje presoje. Kadar se sodišče z izvedenčevim izpovedbo (mnenjem) ne strinja, ga ne more modificirati ali nadomestiti s svojim mnenjem ali s svojimi ugotovitvami, temveč mora za odpravo nasprotij ali pomanjkljivosti v mnenju ali dvomov v pravilnost danega mnenja, najprej opraviti novo zaslihanje že postavljenega izvedenca, nato pa odrediti drugega izvedenca. Če se sodišče ne strinja z nobenim izmed izvedenskih mnenj, šteje dejstvo, ki se je ugotavljalo z izvedencem, za nedokazano, in ravna po načelu »in dubio pro reo« (v dvomu v korist obdolženca) [1].

3 METODE IDENTIFIKACIJE GLASOV

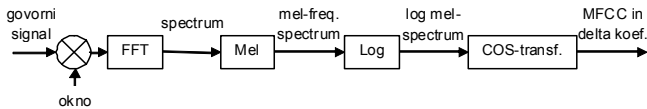
Različne metode identifikacije govorcev so lahko bolj ali manj subjektivne oz. objektivne. Tudi pri objektivnih metodah imamo opraviti z določenim vplivom človeka; npr. računalnik je sprogramiran, rezultati pa so interpretirani s strani eksperta. Najbolj subjektivna metoda identifikacije govorcev v forenzične namene je slušno-zaznavna metoda oz. slušna analiza. Nekoliko bolj objektivna je slušno-instrumentalna metoda. Med najbolj objektivne štejemo polavtomatske in avtomatske metode identifikacije govorcev.

Slušno-zaznavna metoda (angl. »aural-perceptual approach«) oz. slušna analiza (angl. »auditory analysis«) v osnovi temelji na pozornem poslušanju posnetkov s strani izkušenega fonetika, pri čemer se zaznane razlike v govoru uporabijo za ocenjevanje stopnje podobnosti med glasovi. Prvi pristopi k omenjeni metodi so vključevali pozorno poslušanje narečnih posebnosti, govornih napak in kvalitete glasov, vendar je bila primernost teh postopkov vseskozi vprašljiva. Čeprav so se te metode izkazale za neprimerne pri razpoznavanju govorcev v forenzične namene, **koristijo pri iskanju značilnosti govorcev in njihovem razvrščanju v skupine.** Poleg prej navedenih razlik se govorci razlikujejo še po hitrosti govora, trajanju premorov, izgovarjavi glasov in slogu govorjenja. Pri razpoznavanju govorcev so pomembne tudi višjenivojske govorne karakteristike, kot npr. narečne in jezikoslovne posebnosti ter prozodične lastnosti. Z upoštevanjem vseh navedenih značilnosti lahko podamo subjektivno oceno verjetnosti glede podobnosti dveh glasov. **Slušna analiza ima svoje omejitve** in se pri običajni fonetični analizi uporablja predvsem za izluščanje zanimivih lastnosti in parametrov, ki jih nato podrobneje analiziramo s slušno-instrumentalno metodo [1].

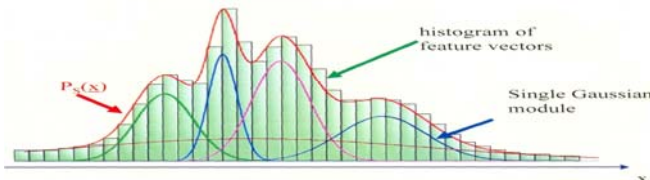
Slušno-instrumentalna metoda (angl. »auditory instrumental approach«) vključuje meritve različnih parametrov, kot so npr. osnovna frekvenca (F0), hitrost govora, potek osnovnega tona, razne spektralne karakteristike govornega signala itd. Parametri se nato medsebojno primerjajo po srednjih ali povprečnih vrednostih in variancah. Pri **računalniški akustični analizi (angl. »computerised acoustic analysis«)** dobimo numerične vrednosti različnih govornih parametrov s pomočjo posebne programske opreme. Pri tem je vloga eksperta še vedno zelo pomembna, saj se je potrebno odločiti, kateri govorni vzorci so dovolj dobre kvalitete za analizo. Poleg tega je potrebno izbrati oz. določiti primerljive dele govornih vzorcev, ki bodo analizirani, in ovrednotiti dobljene rezultate. Parametri pri akustično-fonetični analizi večinoma izvirajo iz lingvistično-fonetičnih raziskav in so neposredno povezani s slišnimi fonetičnimi značilnostmi [2].



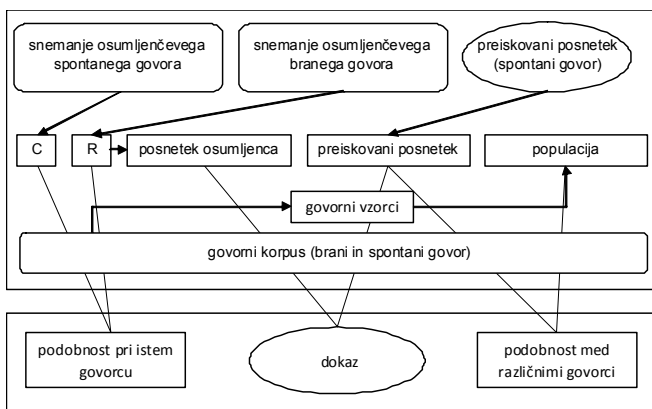
Slika 1: Postopek avtomatskega razpoznavanja govorcev



Slika 2: Modeliranje značil: MFCC koeficienti



Slika 3: Modeliranje govornih značil z GMM



Slika 4: Računanje razdalj

Polavtomatsko (angl. »forensic semiautomatic speaker recognition«) in avtomatsko (angl. »forensic automatic speaker recognition«) razpoznavanje govorcev v forenzične namene je uveljavljen termin za metode (pol)avtomatskega razpoznavanja govorcev, ki so prilagojene za uporabo v forenzične namene. Pri polavtomatskih metodah prihaja med preiskavo do interakcije eksperta in računalnika. Pri avtomatskem razpoznavanju govorcev (slika 1) pa se medsebojno primerjajo statistični modeli akustičnih parametrov (slika2) glasov znanih govorcev (iz govorne baze) s statističnim modelom akustičnih parametrov nepoznane osebe, ki jo želimo identificirati. Na podlagi te primerjave izračunamo kvantitativno oceno podobnosti med (od govorca odvisnimi) parametri glasu nepoznane osebe na posnetku in parametri obdolženca s čimer ocenimo prepričljivost dokaza (slika 4). Pri avtomatskem razpoznavanju govorcev v forenzične namene je prepričljivost dokaza odvisna od relativne verjetnosti, da opazimo neke značilnosti nepoznanega glasu v statističnem modelu akustičnih parametrov obdolženca in v statističnih modelih glasov potencialne populacije. **Metode razpoznavanja govorcev, ki temeljijo na tehnikah statističnega modeliranja**, kot npr. Gaussov

mešani model (angl. Gaussian Mixture Model, GMM; slika 3), imajo to dobro lastnost, da neposredno vrnejo verjetnost, ali posamezna izgovorjava lahko pripada statističnemu modelu govorca [3]. Zadnje čase temeljijo sistemi za (pol)avtomatsko razpoznavanje govorcev v forenzične namene na **oceni kvocienta verjetnosti (angl. »likelihood ratio«) dveh medsebojno konkurenčnih hipotez** (obdolženec je vir preiskovanega (vprašljivega) posnetka ali obdolženec ni vir preiskovanega posnetka), ki podaja stopnjo zanesljivosti dokaza. Samodejni parametri, kot so »kepstrum« ali »delta kepstrum« delujejo dobro pri avtomatski identifikaciji in verifikaciji govorcev v komercialne namene. Zelo malo pa je bilo opravljenih raziskav, kako se ti samodejni parametri obnesejo pri forenzičnih »stvarnih pogojih«. **Avtomatski sistemi za razpoznavanje govorcev se ne smejo uporabljati samostojno pač pa le kot dopolnitev drugih metod**, sicer obstaja možnost napačne identifikacije [4].

Rezultate slušno-zaznavne metode (rezultati temeljijo na subjektivni oceni), slušno-instrumentalne metode (rezultati temeljijo na subjektivni oceni in statističnih verjetnostih) in avtomatskih metod (rezultati temeljijo zgolj na statističnih verjetnostih) v praksi preučujemo povezano s čimer dobimo **kombinirano oceno zanesljivosti** dokaznega gradiva.

3 RAZLIKE MED »STVARNIMI« IN »LABORATORIJSKIMI« POGOJI

Razpoznavanje govorca je zahtevno opravilo tako pri komercialnih kot forenzičnih aplikacijah. Medtem ko pri nekaterih aplikacijah lahko ocenimo, privzamemo ali predvidimo delovne pogoje, pri skoraj vseh primerih razpoznavanja govorcev v forenzične namene to ni možno. Pri razpoznavanju govorcev v forenzične namene imamo opravka s spornimi posnetki izgovorjav, ki predstavljajo dokazno gradivo in so posneti v »**stvarnih pogojih**« med samim izvajanjem kaznivih dejanj.

Pojem »stvarni pogoji« uporabljamo kot nasprotje »**laboratorijskim pogojem**«, ko ne moremo nadzirati, pričakovati ali predvidevati pogojev v katerih se bodo pridobili posamezni govorni posnetki. Celo več; obtoženec ponavadi ne želi korektno sodelovati in skuša ovirati ali pre-prečiti pridobitev kakršnikoli zanj obremenilnih informacij.

Zaradi »stvarnih pogojev« pridobivanja posnetkov je govorni signal bolj spremenljiv oz. variabilen. Vire variabilnosti govornega signala lahko razvrstimo v naslednje kategorije:

- (i) *svojske variabilnosti govornih signalov istega govorca*: vrsta govora, staranje, časovni presledek med dvema posnetkoma, narečje, žargon, socialni status, čustveno stanje, vpliv omamnih sredstev itd.,
- (ii) *izsiljene oz. umetne variabilnosti govornih signalov istega govorca*: »Lombardov« učinek, stres zaradi zunanega vpliva, »cocktail-party« učinek itd.,

- (iii) *zunanja variabilnost odvisna od kanala*: tip telefona ali mikrofona, fiksna/mobilna telefonija, komunikacijski kanal, pasovna širina, dinamični obseg oz. razpon, električni in akustični šum, odmev, popačenje itd.

Forenzični pogoji so doseženi, ko se dejavniki variabilnosti, ki predstavljajo t.i. »stvarne pogoje«, pojavljajo brez kakršnegakoli principa, pravila ali norme. Lahko so konstantni preko celotnega klica ali pa se hipoma pojavijo ali izginejo; na celoten proces vplivajo povsem nepredvidljivo.

Uporaba metod za razpoznavanje govorcev v razmerah, ko obstaja nevarnost, da je **prestopnik poizkušal prikriti oz. popačiti svoj glas**, v splošnem ni mogoča. Vendar je uporaba govornih značilnosti, ki so kombinacija večih različnih govornih značilk, bolj neodvisna od govorca in manj občutljiva na poizkuse posnemanja. V normalnih razgovorih storilci kaznivih dejanj govora ne pačijo ali kako drugače zavestno spreminjajo, če ne sumijo na prisluškovanje. Prav tako je skoraj nemogoče potvarjati govor med obsežnejšim zasliševanjem na policijski postaji.

4 TEKOČE RAZISKAVE S PODROČJA RAZPOZNAVANJA GOVORCEV V FORENZIČNE NAMENE V SLOVENIJI

Poglavitni cilj dejavnosti s področja razpoznavanja govorcev v forenzične namene, ki trenutno potekajo na Institutu Jožef Stefan, je **izgradnja nove specializirane govorne baze** (v slovenskem jeziku) za namene določanja, analize in meritve učinkov oz. posledic najpomembnejših virov variabilnosti, ki jih najdemo v (dejanskih) komercialnih in forenzičnih aplikacijah, in proučevanja njihovega vpliva na sisteme avtomatskega razpoznavanja. S tem se želi doseči nadzorovane pogoje za razpoznavanje govorcev:

- Govorna baza mora biti posneta v večih delih, vključno z direktnim snemanjem in snemanjem preko telefona, branjem besedil z različnimi hitrostmi, branim govorom nasproti spontanemu govoru, uporabi različnih mikrofonov in telefonskih aparatov, različna narečja govorcev (lahko se spreminja že pri istem govorniku glede na to ali besedilo bere ali pa spontano govori), stalne izgovarjave besedil za vse govorce preko vseh delov posnetkov napram specifičnim izgovarjavam vsakega govorca znotraj posameznega posnetka.
- Govorna baza mora biti fonološko in zlogovno uravnotežena tako, da posnema frekvenco pojavljanja fonemov in zlogov v govornem slovenskem jeziku.
- Starostna porazdelitev govorcev v podatkovni bazi mora upoštevati sociološke dejavnike povezane z uporabo različnih tehnologij.
- Variabilnost parametrov med posameznimi posnetki je pomemben dejavnik, ki ga je potrebno upoštevati. Med posameznimi snemanji mora biti vsaj 14 dnevni zamik.

Na osnovi opisane govorne baze želimo ovrednotiti razlike v snemalnih pogojih, ki nastopajo pri posnetkih iz sodne prakse, skušamo kompenzirati njihove učinke ter

kvantitativno opredeliti negotovost oz. nezanesljivost, ki je posledica teh dejavnikov. Snemalni pogoji vplivajo tako na slušno kot instrumentalno in avtomatsko razpoznavanje govorcev.

5 SKLEP

Napredek na področju računalništva in digitalnega procesiranja signalov je prinesel številne izboljšave tako pri hitrosti kot samih zmožnostih različnih akustičnih analiz govornega signala in statistične obdelave podatkov. Ob vsem tem se nam pogosto zastavlja vprašanje, ali se bo vse to pokazalo tudi pri izboljšani identifikaciji govorcev v forenzične namene?

Vsekakor nam bo tehnologija omogočala, da bomo to delo opravljali vedno hitreje, s čimer bodo ti postopki čedalje bolj dosegljivi. Le redko kdo pa zaenkrat verjame, da bodo postopki identifikacije govorcev v forenzične namene postali popolnoma avtomatizirani (kot je to npr. prikazano v ameriških kriminalkah). Med strokovnjaki s tega področja prevladuje mnenje, da zaradi same narave človeškega glasu in zvočnih lastnosti posameznega jezika ter pomanjkanja nadzora nad forenzičnimi govornimi vzorci, nikdar ne bomo imeli popolnoma avtomatskega sistema identifikacije govorcev v forenzične namene in to neglede na razvoj same tehnologije. Pred akustično analizo govornega signala, ki bo navadno selektivna in ne avtomatska, bo vedno potrebno izvesti tudi skrbno slušno analizo (npr. za ugotavljanje raznih lingvističnih posebnosti).

Iz vsega navedenega lahko razberemo, da si zgolj zaradi tehnološkega razvoja ne smemo preveč obetati. **Pravi napredek pri identifikaciji govorcev v forenzične namene nam lahko prinesejo le dodatna znanja**: boljše poznavanje ozadja podatkov ter kako jih pravilno obdelati, boljše razumevanje in poznavanje individualnosti vsebovane v govoru, kateri so najboljši parametri, kje naj jih iščemo in pod kakšnimi »stvarnimi pogoji« so primerljivi [5].

Literatura

- [1] T. Šef, P. Baucon, Sodno izvedenstvo in razpoznavanje (identifikacija) govorcev v kazenskem postopku, *Pravosodni bilten*, Ljubljana, 2/2007.
- [2] P. Rose, Technical forensic speaker recognition: evaluation, types and testing of evidence, *Computer Speech and Language*, 2005.
- [3] A. Alexander, Forensic Automatic Speaker Recognition Using Bayesian Interpretation and Statistical Compensation for Mismatched Conditions, *doktorska disertacija*, Lausanne, EPFL, 2005.
- [4] M. Jessen, *Some Experiences from Tests of an Automatic Speaker Recognition System under Forensic Conditions*, Bundeskriminalamt, EAFS, 2007.
- [5] P. Rose, *Forensic speaker Identification*, Taylor & Francis, 2002.

PATOLOGIJA MINIMIN PREISKOVANJA

Aleš Tavčar, Mitja Luštrek, Matjaz Gams
Odsek za inteligentne sisteme
Institut "Jožef Stefan"
Jamova cesta 39, 1000 Ljubljana, Slovenija
{ales.tavcar,mitja.lustrek,matjaz.gams}@ijs.si

POVZETEK

Do preiskovalne patologije naletimo, ko globlje preiskovanje dreves iger, v določenih primerih, daje slabše rezultate, kot plitvejše. Na ta pojav so raziskovalci naleteli pri proučevanju algoritma minimaksa, ki se uporablja pri igranju iger. V tem prispevku se ukvarjamo s patološkimi modeli enoagentnega (minimin) preiskovanja, ter obravnavamo dejavnike, ki vplivajo na patologijo: vejitve, zrnatost in odvisnost.

1 UVOD

Z enoagentnim preiskovanjem se srečujemo v primerih, ko agentu pri iskanju najboljše rešitve nihče ne nasprotuje, tako kot je to pri igrah za dva igralca. Pri iskanju najkrajše poti se igralec neovirano premika npr. po zemljevidu in išče poteze, ki ga bodo privedle do iskanega cilja.

Podobno kot pri igrah z dvema igralcema, lahko tudi pri tej vrsti preiskovanja uporabimo predstavitev z drevesom igre, kjer vozlišča predstavljajo položaje, povezave pa možne poteze. V listih pa se nahajajo vrednosti, ki označujejo kakovost rešitve in ne le zmago ali poraz. Bistvena razlika med enoagentnim preiskovanjem in igrami za dva igralca pa je način prenosa vrednosti v listih do korena drevesa. Te vrednosti prenašamo z miniminom, torej vrednost nekega vozlišča določimo tako, da izberemo najmanjšo vrednost v sinovih tega vozlišča.

Pri igranju iger običajno pregledamo drevo igre od trenutnega položaja do neke globine, stanja tam hevristično ocenimo in te ocene prenesemo nazaj do korena drevesa, ter se na njihovi podlagi odločimo za ustrezno potezo. Opis daje slutiti, da globlje preiskovanje daje boljše rezultate, kar se tudi pokaže v praksi. Matematične analize modelov [1, 2, 3] pa so pokazale ravno nasprotno: v nekaterih primerih globlje preiskovanje daje slabše odločitve. Ta pojav se imenuje *preiskovalna patologija*.

Patologija minimaksa je že dodobra raziskana, istega pa ne moremo trditi za patologijo minimina. Začetne raziskave enoagentnega preiskovanja so le potrdile možnost patologije [4] in na umetnih drevesih podale razloge za njen nastanek [5]. Naša naloga je zgraditi

splošen model enoagentnega preiskovanja, ki bi omogočal preučevanje vpliva na patologijo treh dejavnikov: vejitve, zrnatosti in podobnosti. S spreminjanjem teh količin pregledamo celoten prostor vrednosti parametrov in iz rezultatov ugotovimo iskane značilnosti modela.

2 SORODNA DELA

2.1 Patologija minimaksa

Odkritje patologije povezujemo s prizadevanji raziskovalcev o poizkusu formalne razlage, zakaj globlje preiskovanje z minimaksom v realnih igrah daje boljše rezultate. Pri tem sta Dana S. Nau leta 1979 [1] in Donald F. Beal leta 1980 [2] neodvisno odkrila zanimiv pojav, ki so ga poimenovali patologija minimaksa.

Razhajanja med teoretičnimi modeli in uporabo minimaksa v praksi je povzročilo veliko zanimanje med raziskovalci. Začeli so iskati lastnosti realnih iger, ki prispevajo k uspešnosti minimaksa. Med najpomembnejšimi so vejitve, število vrednosti v vozliščih in odvisnost bližnjih vozlišč. Kaluža idr. [6] so sestavili minimaks model, ki upošteva vse našete parametre in preveri njihov vpliv na pojav patologije. **V tej nalogi bomo naredil analizo, ali podobne relacije veljajo tudi v modelu minimin.**

2.2 Patologija minimina

Patologijo enoagentnega preiskovanja so odkrili Bultko idr. [4] šele leta 2003, zato je tudi slabše raziskana od patologije minimaksa. Patologijo je pokazal na enostavnem binarnem drevesu, ni pa ponudil razlage, zakaj do nje pride.

Podrobnejšo razlago pojava poda Luštrek [5, 7], kjer razloge razdeli v dve skupini:

- domena problema, na katero težko vplivamo
- uporabljena hevristična funkcija, na katero imamo vpliv.

Bultko je kasneje [8] ugotovil, da se patologija pojavi tudi izven sintetičnih preiskovalnih dreves, saj je tudi pri proučevanju igre osmih ploščic z uporabo hevristične funkcije predstavljene z umetno nevronske mrežo naletel na patologijo.

Piltaver idr. [9] so proučeval patologijo v igri osmih ploščic. V igro so dodali nove poteze in tako omogočil spreminjanje vejitve. Pri različicah igre so opazili 13 skupin iger s faktorji vejitve med 1,56 in 4,44. Omenjeni faktorji vejitve so povprečni, saj nobena izmed iger nima uniformnega faktorja vejitve. Do jasnih zaključkov glede vpliva faktorja vejitve avtor ni prišel. Možno je zaznati, da večja vejitev povzroči večjo patologijo. Naslednji parameter, ki ga je moč spreminjati, je zrnatost hevrstične funkcije. Izkaže se, da večja zrnatost hevrstične funkcije zavira patologijo. Odvisnost bratskih vozlišč pa ni bilo mogoče neposredno spreminjati, saj je odvisnost vozlišč lastnost problema. Vpliv odvisnosti bratskih vozlišč so obravnavali na vseh različicah igre hkrati in tako dobil več izmerjenih vrednosti faktorja grozdenja za posamezne igre. Dobljene vrednosti so uredili po naraščajočem faktorju grozdenja in iz grafa razbrali, da večja odvisnost vozlišč povzroči manjšo patologijo preiskovanja.

S patologijo v igri osmih ploščic sta se ukvarjala tudi Sadikov in Bratko [10], kjer sta ugotovila, da je pesimistična hevrstična funkcije učinkovitejša od optimistične, saj patologijo oslabi.

3 NEODVISNI MODEL

Gradnja neodvisnega modela se zgleduje po modelih opisanih v [6, 3], s to bistveno razliko, da je prirejen minimin preiskovanju in nekaterim dodatnim omejitvam domene. Model omogoča uporabo različnega števila pravih in hevrstičnih vrednosti v vozliščih drevesa.

Pri računanju patologije se bomo osredotočili na napako poteze. Zanimal nas bo delež izbranih potez v korenu, ki ne vodijo v sina z najboljšo pravo vrednostjo. Verjetnost take napake, pri določeni globini preiskovanja d in hevrstični funkciji h , označimo z $E(h,d)$. Sedaj lahko definiramo patološkost p z naslednjim izrazom:

$$p = \frac{E(h, d_2)}{E(h, d_1)}, 1 \leq d_1 < d_2 \leq d_{max} \quad (1)$$

Iz 1 lahko sklepamo, da vrednost $p > 1$ pomeni prisotnost patologije. Nasprotno pa je patologija odsotna, ko je $p < 1$, saj je v tem primeru napaka pri globljem preiskovanju manjša kot pri plitvejšem.

Ugotavljanje patologije v modelu začnemo z gradnjo sintetičnega preiskovalnega drevesa. Zgrajeno drevo je podane vejitve b in globine d_{max} . Patološkost smo opazovali pri izbranih globinah preiskovanja $d_1 = 1$ in $d_2 = 5$. S prirejenim algoritmom minimin, ki na vsakem nivoju vrne maksimalno vrednost, se sprehodimo čez drevo in ko pridemo do končne globine d_{max} , listom dodelimo eno izmed g naključnih realnih vrednosti z intervala $[0, 1]$. Ob vračanju algoritma pa se izračunajo tudi prave vrednosti notranjih vozlišč. Pri tem si zapomnemo b pravih vrednosti na prvem nivoju drevesa, s katerimi bomo v nadaljevanju ugotavljali napako poteze. Statično napako hevrstične funkcije simuliramo tako, da pri preisko-

vanju do določene globine d pravih vrednostim vozlišč na nivoju d dodamo šum in s tem dobimo hevrstične vrednosti. Za dodajanje šuma uporabimo Gaussov šum s standardnim odklonom $\sigma = 0, 1$.

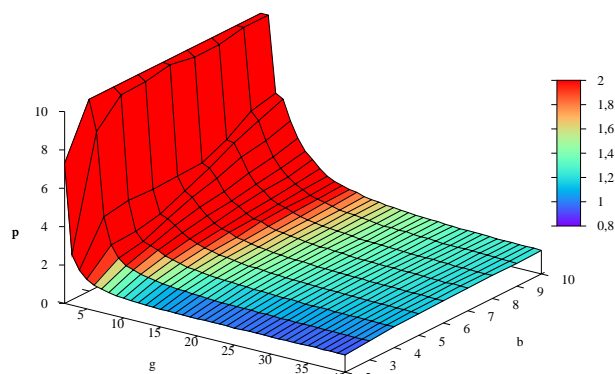
Pomemben parameter v modelu je zrnatost vozlišč preiskovalnega drevesa. Ta lastnost se je izkazala za zelo pomembno pri analizi minimaks patologije, zato smo jo na podoben način vpeljali tudi v našo simulacijo. Interval $[0,1]$ smo razdelili na g podintervalov enake dolžine. Posamezna vrednost je predstavljena s srednjo vrednostjo podintervala. Nato mejo zadnjega intervala premaknemo na vrednost p_{leaves} in ustrezno popravimo tudi meje ostalih podintervalov. p_{leaves} definiramo z izrazom 2:

$$p_{leaves} = \sqrt[N]{1 - p_1}, N = b^{d_{max}-1} \quad (2)$$

Poglejmo še razmislek, ki je privedel do tega izraza. Preveč maksimalnih vrednosti v sinovih korena prepreči pojav patologije, zato definiramo mero p_1 , ki predstavlja delež maksimalnih vrednosti na prvem nivoju drevesa. Iz tega podatka želimo izračunati kolikšen mora biti delež nemaksimalnih vrednosti v listih (p_{leaves}), da dosežemo v korenu delež p_1 maksimalnih vrednosti. Vrednosti v listih sedaj umestimo v ustrezen podinterval in tako dobimo v listih g različnih vrednosti.

Vsaka meritev je pridobljena s simulacijo Monte Carlo z 30000 drevesi. Vsakič smo izračunali količine, ki nas zanimajo, in rezultate povprečili. Tako smo dobili povprečene vrednosti za raven patologije in se izognili nihanjem, ki jih vnaša gradnja naključnih sintetičnih dreves.

Sedaj si pogledjmo rezultate simulacije s celotnim naborom parametrov. Na sliki 1 vidimo grafični prikaz izračunane stopnje patološkosti neodvisnih dreves ($s = 0.0$). Ta je na grafu prikazana v odvisnosti od zrnatosti g in vejitve b . Navpično os grafa smo omejili na vrednost $p = 10$, saj tako lahko natančneje ugotovimo smernice pri spreminjanju posameznih parametrov.



Slika 1: Prikaz stopnje patološkosti neodvisnih dreves. Prikazana je v odvisnosti od zrnatosti g in vejitve b

Pri konstantni zrnatosti se s povečevanjem vejitve povečuje stopnja patološkosti. To lahko razberemo iz naraščajočih krivulj, ki so prečne na g os. Povečevanje zrnatosti blagodejno vpliva na odpravo patološkosti, saj relativno hitro pada in se približuje proti $p = 1$. S povečanjem vejitve se število pravih vrednosti, potrebnih za odpravo patologije, zelo hitro povečuje.

4 ODVISNI MODELI

V prejšnjem poglavju smo opisali gradnjo naključnih preiskovalnih dreves. Kot je bilo pričakovano so taka drevesa patološka. Naslednja naloga je najti primeren način gradnje odvisnih ter delno odvisnih dreves. Taka drevesa so v praksi namreč bolj pogosta, saj so pri realnih igrah sosednja vozlišča precej podobna. Z eno odigrano potezo se namreč položaji bistveno ne spremenijo.

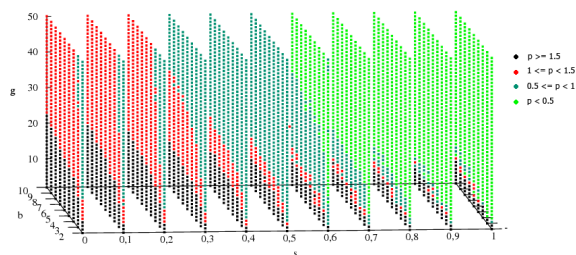
Uporabljeni pristop temelji na urejanju naključnih vrednosti. Tak model se je uporabilo tudi pri analizi patologije minimaksa [6] in zato bo uporaba tega načina omogočila bolj pošteno primerjavo rezultatov obeh modelov. Gradnja popolnoma odvisnih dreves se pri izbranem modelu začne od spodaj, torej pri listih drevesa. Vsem listom dodelimo pomožne naključne vrednosti, ki jih nato sortiramo po velikosti. Sedaj imajo listi od leve proti desni naraščajoče pomožne vrednosti. Dobljene vrednosti nato še preslikamo v ustrezne razrede zrnatosti na enak način, kot smo to storili v prejšnjem razdelku. Tako smo zgradili popolnoma odvisno drevo, kateremu pripišemo podobnost $s = 1$.

Sedaj imamo na razpolago neodvisna in popolnoma odvisna drevesa in nam ostane le še kombiniranje le-teh v delno odvisna drevesa. Preizkusili smo več načinov sestavljanja delno odvisnih dreves. Vsem pa je skupno to, da vsebujejo delež s_0 listov odvisnega in delež $1 - s_0$ listov neodvisnega drevesa. Pri vseh načinih gradnje pa dobimo kvalitativno enake rezultate.

Na sliki 2 je prikazan prostor parametrov vejitve, odvisnosti in zrnatosti ter izmerjena stopnja patologije za posamezne vrednosti. Na sliki je viden padec patološkosti pri povečevanju odvisnosti. Pojav je najbolj izrazit pri večjih vejitvah, kjer je jasno viden prehod iz črne barve v sivo. Naslednja značilnost je povečevanje stopnje patologije s povečevanjem vejitve, saj se z večevanjem le-te črno obarvana območja širijo proti večjim zrnatostim. Če opazujemo le en stolpec na grafu opazimo še eno značilnost: s povečevanjem zrnatosti stopnja patologije pada.

Na sliki 3b so prikazane mejne vrednosti g , pri katerih patologija izgine ($p < 1$): s povečevanjem podobnosti meja patologije pada. Opazimo tudi, da je za odpravo patologije pri večjih vejitvah potrebna večja zrnatost. Model je kvalitativno primerljiv z minimaks modelom, le da v našem primeru meja patologije pri majhni stopnji podobnosti ne pade tako hitro, kot pri minimaks modelu.

Zavoljo primerljivosti s prejšnjimi raziskavami in umestitev realnih iger v zgrajeni model smo stopnjo



Slika 2: Stopnja patologije v minimin modelu preiskovanja.

podobnosti izmeriti s faktorjem grozdenja f (angl. clustering factor), ki je povzet po Sadikovu [11]. Definirana je kot količnik med povprečnim standardnim odklonom sinov in standardnim odklonom celotnega drevesa. Z izračunom je mogoče zaznati, kdaj so si bližnja vozlišča nadpovprečno podobna. Formula za izračun faktorja grozdenja je podana z izrazom 3.

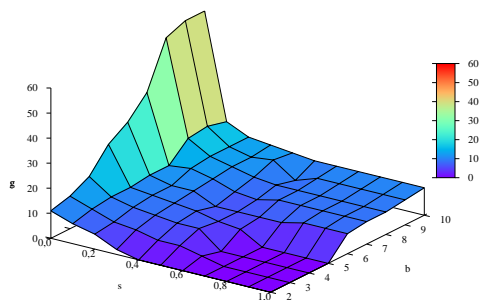
$$f = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^b (v_{ij} - \bar{v}_i)^2}}{\sqrt{\frac{1}{N} \sum_{i=1}^N (v_i - \bar{v})^2}} \quad (3)$$

Na sliki 3 so prikazane mejne vrednosti zrnatosti g za oba modela. Na obeh slikah opazimo podobne značilnosti modela.

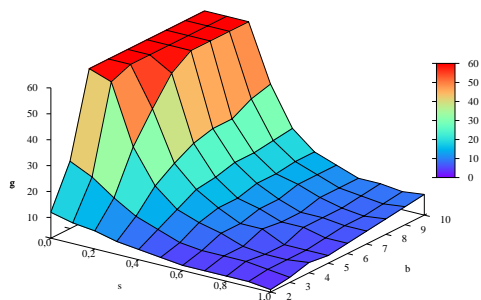
1. meja zrnatosti g , pri kateri patološkost doseže $p = 1$ z vejitvijo narašča, s povečevanjem podobnosti pa pada,
2. patološkost je v modelu minimin bolj izrazita, saj je pri manjši odvisnosti meja zrnatosti, kjer patologija izgine, precej višja kot v minimaks modelu.

4.1 Primerjava z igro osmih ploščic

Poglejmo še primerjavo poteka patološkosti pri igri osmih ploščic in minimin modela. Za igro osmih ploščic so na voljo [9] podatki o povprečni vejitvi, stopnji podobnosti merjeni s faktorjem grozdenja in izračunano stopnjo patološkosti za večje število zrnatosti. Nad našim modelom smo pognali simulacijo z enakimi naborom parametrov. Na sliki 4 lahko vidimo potek patološkosti pri igri osmih ploščic in minimin modelu. Igra uporablja dve hevristični funkciji, za oceno položajev. Patološkost igre pri uporabi obeh funkcij je prikazana na grafu. Iz slike je razvidno, da model relativno dobro modelira potek stopnje patološkosti v igri, saj je potek vseh krivulj kvalitativno podoben. Le pri manjši zrnatosti je napaka modela večja.



(a) Minimaks model z urejanjem naključnih vrednosti



(b) Minimin model

Slika 3: Mejne vrednosti g , pri katerih je $p = 1$

5 ZAKLJUČEK

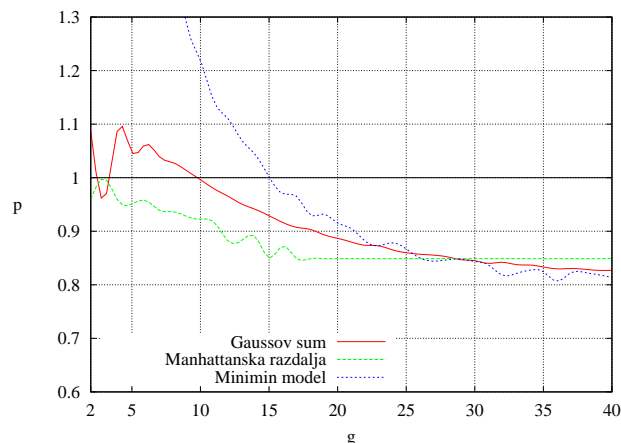
V članku smo ugotavljali, ali relacije v minimaks modelu veljajo tudi v modelu minimin. **Ugotovimo, da za oba modela kvalitativno veljajo enake relacije.**

Pri minimaksu igra vejitev pomembno vlogo, saj se s povečevanjem vejitve patološkost okrepi. Podobno obnašanje smo zaznali tudi pri analiziranju minimin modela.

Naslednji pomemben dejavnik je število pravih vrednosti v vozliščih drevesa. Povečevanje zrnatosti stopnjo patološkosti oslabi. Pri neodvisnem modelu opazimo, da se število vrednosti, ki so potrebne za odpravo patologije, z naraščanjem vejitve eksponentno povečuje. Tu je mogoče zaznati pomembno razliko med minimaks in minimin modeloma. Pri slednjem je namreč zadostno število vrednosti za odpravo patologije bistveno večje, kot pri minimaks modelu. Ta razlika je najbolj opazna pri večjih vejitvah.

Zadnji in eden najpomembnejših dejavnikov pa je odvisnost vozlišč v drevesu. Povečevanje odvisnosti pozitivno vpliva na patološkost v modelu, saj jo ublaži.

Opravili smo tudi primerjavo minimin modela z igro osmih ploščic in ugotovili, da model precej dobro modelira stopnjo patologije v igri, le pri manjši zrnatosti je napaka modela večja.



Slika 4: Primerjava minimin modela z igro osmih ploščic.

Literatura

- [1] Dana S. Nau. *Quality of decision versus depth of search on game trees*. Doktorska disertacija, Duke University, 1979.
- [2] Donald F. Beal. An analysis of minimax. V zborniku *Advances in Computer Chess 2*, str. 103–109. Edinburgh University Press, 1980.
- [3] Mitja Luštrek, Matjaž Gams in Ivan Bratko. Is real-valued minimax pathological? *Artificial Intelligence*, 170(6):620–642, 2006.
- [4] V. Bulitko, Lihong Li, R. Greiner, I. Levner. Lookahead pathologies for single agent search. V zborniku *18th International Joint Conference on Artificial Intelligence*, str. 1531–1533, 2003.
- [5] M. Lustrek. Pathology in single agent search. V zborniku *8. mednarodne multikonference Informacijska družba*, str. 345–348, 2005.
- [6] Boštjan Kaluža, Mitja Luštrek, Matjaž Gams in Aleš Tavčar. Pathology in minimax searching. V zborniku *Proceedings of the Sixteenth International Electrotechnical and Computer Science Conference (ERK 2007)*, zv. B, str. 107–110, 2007.
- [7] Mitja Lustrek, Vadim Bulitko. Thinking Too Much: Pathology in Pathfinding. V zborniku *ECAI 2008: 18th European Conference on Artificial Intelligence*, str. 899–900, 2008.
- [8] V. Bulitko. Lookahead Pathologies and Meta-level Control in Real-time Heuristic Search.. V zborniku *15th Euro-micro Conference on Real-Time Systems*, str. 113–16, 2003.
- [9] Rok Piltaver, Mitja Luštrek, Matjaž Gams. Search pathology of 8-puzzle. V zborniku *Proceedings of the 10th International Multiconference Information Society (IS 2007)*, zv. A, str. 8–12, 2007.
- [10] A. Sadikov, I. Bratko. Pessimistic heuristics beat optimistic ones in real-time search. V zborniku *European Conference on Artificial Intelligence (ECAI)*, str. 148–152, 2006.
- [11] Aleksander Sadikov. *Širjenje napak hevrističnih ocen v preiskovanju grafov iger*. Doktorska disertacija, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, 2005.

PROBLEM TRANSFORMATION METHODS FOR MULTI-GENRE WEB PAGES CLASSIFICATION

Vedrana Vidulin
Department of Intelligent Systems
Jožef Stefan Institute
Jamova cesta 39, 1000 Ljubljana, Slovenia
vedrana.vidulin@ijs.si

ABSTRACT

Availability of tools that simplify creation of composite web pages resulted in increased appearance of multi-genre pages containing several parts, each of different genre. To properly address the genre interplay, a recent proposal in automatic web genre identification is multi-label classification. The dominant approach to such classification is by binary relevance problem transformation method where a multi-label machine learning problem is transformed into sub-problems of learning binary single-label classifiers, one for each genre. In this paper we explore label power-set problem transformation method, where each genre combination is represented with a single distinct label, and one multi-class classifier is learned. The approaches are compared to determine which better captures the correlations between genres appearing together in multi-genre pages. Experimental results show that both of the approaches failed to properly capture the correlations since obtained differences were the result of improved classification of one-genre pages.

1 INTRODUCTION

From the communication aspect the goal of a web page is to send a message to a reader. The message has a topic (e.g. the life of elephants), which is communicated in a specific way depending on the purpose of a web page. For example, scientific web page will present to a zoologist a high number of objective facts about elephants, and when wishing to entertain, it will present pictures and video material. In the light of the previous explanation, genre can be described as intentional styling of a web page with the objective to communicate the topic in a specific manner. For the purpose of automatic web genre classification, a genre can simply be defined as the style of a web page [4].

Automatic web genre identification (AWGI) is useful in information retrieval and digital library systems, which combine genres with keywords to obtain more precise hits. In the field of information extraction predictable structure of detected genre facilitates extraction of desired information from the text.

The problem that recent line of research deals with is how to model composite, multi-genre web pages [7, 8]. State-of-

the-art web genre classifiers still mostly use single-label classification scheme [5, 6], which attribute to a web page one genre label from the set of predefined labels. Recent line of research [7, 8], however, showed that multi-label classification scheme is more suitable for capturing this complexity. In our study we follow this scheme with some modifications.

The model based on zero-to-multi genre assignment was implemented by Santini [7]. It assigned multiple labels to multi-genre web pages, while the absence of clear genre conventions resulted in zero label assignment. The model was implemented in two steps. First, text types were inferred by odds-likelihood or subjective Bayesian method. Second, if-then rules were created to identify genres from the combination of text types and features as linguistic and HTML.

Another model of Stubbe et al. [9] was partially concerned with the issue of multi-label approach to AWGI. They built multiple genre-specific classifiers, one per genre, by combining features provided by the genre expert into rules. They pointed that the classifiers can be combined into a scheme which can assign several genre labels to a web page. An improved classifier's precision was observed. However, they did not explore this issue in more detail since their main task was to exploit interdependencies between genre-specific classifiers to improve the precision of a single label assignment.

In contrast to presented approaches based on expert knowledge, our goal is to induce a multi-label classifier from the example web pages with supervised machine learning (ML) methods. We achieve this through the problem transformation approach where a multi-label ML problem is transformed into one or more single label ML sub-problems, depending on the type of transformation. We explore two transformation methods: binary relevance (BR) and label power-set methods (LPS).

While the BR method shares several characteristics with the approaches of Santini [7] and Stubbe et al. [9], LPS method differs from previous research. Its main advantage is the induction of a single classifier, for which we assume that it can better learn the correlation between different genres, and consequently better recognize web pages containing multiple genres. In this paper, we compare the two transformations to understand their suitability for learning a web genre model from example web pages.

The rest of the paper is organized as follows. In Section 2 we present the problem transformation methods. Section 3 presents the experimental setup by describing evaluation procedure, multi-label data we experimented with, and ML algorithms used. Section 4 presents experimental results with discussion, and Section 5 concludes the paper.

2 PROBLEM TRANSFORMATION METHODS

The data mining task of multi-label classification associates an example x with a subset of labels $Y \subseteq L$, where L represents previously known finite set of labels. There are two approaches to multi-label classification: by problem transformation and by algorithm adaptation [10]. We have chosen problem transformation approach since it can be combined with existing state-of-the-art ML algorithms. Two transformations are explored in this paper: BR and LPS methods.

The BR transformation method assumes learning $|L|$ binary sub-classifiers $F_l : X \rightarrow \{l, \neg l\}$, one for each label $l \in L$. A new example is classified by all of the $|L|$ sub-classifiers and the class is a union of labels that received a positive answer. The main disadvantage of this approach is that it does not encode correlations between genres. Because of the loss of information, predicted label sets are likely to contain either too many or too few labels, or labels that would never co-occur in practice.

The label LPS transformation method assumes treatment of different combinations of labels as distinct single labels. Therefore, the goal is to learn a multi-class classifier $F : X \rightarrow P(L)$, where X represents example and $P(L)$ the power set of L . This transformation encodes label correlations, however, with the negative side-effect of producing high number of categories. In some cases, newly obtained categories were represented with only one example. To properly train and test a classifier, we removed all the examples labeled with the categories not represented with at least one example in the both train and test sets.

3 EXPERIMENTAL SETUP

3.1 Evaluation

The performance of multi-label classifiers was evaluated using stratified 3-fold cross-validation. Stratification can be approached in different manners in multi-label setting. One approach is to do the problem transformations first and than to separate the data into folds. It results in better balance of classes on the level of individual ML sub-problems. The second approach is to separate the folds before problem transformations and to stratify in a manner to obtain equal distribution of single genre categories over the folds. Since our goal was to obtain the same train-test splits to allow comparisons between induced classifiers we used the second approach. Considering that the number of examples per category decreases after LPS transformation, we used three instead of ten folds to increase the chance of obtaining more test examples per class.

The performance of classifiers was evaluated with several measures: exact match ratio, micro-averaged precision, recall and F-measure, and macro-averaged precision, recall and F-measure.

Exact match ratio (EX) counts exact matches between the predicted and the actual labels (Eq. 1). This measure is, in a way, similar to accuracy in the case of the single-label classification. However, it does not account for e.g. two out of three correctly predicted labels which is fairly good success in the multi-label setup.

$$EX = \frac{\sum_{i=1}^M I[Y_i^{\text{predicted}} = Y_i^{\text{actual}}]}{M} \quad (1)$$

$I[S]$ is 1 if the statement S is true and 0 otherwise, and M represents the number of classified examples.

Micro-averaged measures weight all the web pages equally, representing the averages over all the (web page, genre category) pairs. They tend to be dominated by the classifier's performance on common categories. Micro-averaged precision ($\pi(\text{micro})$) represents the ratio of web pages correctly classified as l ($TP = \text{true positives}$), and all the pages correctly and incorrectly ($FP = \text{false positives}$) classified as l (Eq. 2). Micro-averaged recall ($\rho(\text{micro})$) represents the ratio of web pages correctly classified as l , and all the pages actually pertaining to the class l ($FN = \text{false negatives}$) (Eq. 3). Micro-averaged F-measure ($F(\text{micro})$) represents a harmonic mean of $\pi(\text{micro})$ and $\rho(\text{micro})$ (Eq. 4). $|L|$ represents the number of categories.

$$\pi(\text{micro}) = \frac{\sum_{l=1}^{|L|} TP_l}{\sum_{l=1}^{|L|} (TP_l + FP_l)} \quad (2)$$

$$\rho(\text{micro}) = \frac{\sum_{l=1}^{|L|} TP_l}{\sum_{l=1}^{|L|} (TP_l + FN_l)} \quad (3)$$

$$F(\text{micro}) = \frac{2 \times \pi(\text{micro}) \times \rho(\text{micro})}{\pi(\text{micro}) + \rho(\text{micro})} \quad (4)$$

Macro-averaged measures weight equally all the genre categories, regardless of their frequencies. They tend to be dominated by the classifier's performance on rare categories. Macro-averaged precision ($\pi(\text{macro})$) is computed firstly by computing the precision for each category separately, and then by averaging over all categories (Eq. 5). The same procedure is used for computing the macro-averaged recall ($\rho(\text{macro})$) (Eq. 6), and macro-averaged F-measure ($F(\text{macro})$) (Eq. 7).

$$\pi_l = \frac{TP_l}{TP_l + FP_l}, \quad \pi(\text{macro}) = \frac{\sum_{l=1}^{|L|} \pi_l}{|L|} \quad (5)$$

$$\rho_l = \frac{TP_l}{TP_l + FN_l}, \quad \rho(\text{macro}) = \frac{\sum_{l=1}^{|L|} \rho_l}{|L|} \quad (6)$$

$$F_l = \frac{2 \times \pi_l \times \rho_l}{\pi_l + \rho_l}, \quad F(\text{macro}) = \frac{\sum_{l=1}^{|L|} F_l}{|L|} \quad (7)$$

An issue raised during evaluation is what to treat as a label set L after LPS transformation. Since each combination of individual labels can be treated as new separate label, we can obtain two different micro and macro-averages, depending on our view what constitutes a label set. For the purpose of comparison with BR approach, we considered a set of individual labels present in a data set as a label set L .

3.2 Data

The data set was obtained by extracting 2,491 features from 1,539 web pages in the 20-Genre-Collection corpus (<http://dis.ijs.si/MitjaL/genre/>), containing genres: *Blog, Childrens', Commercial/promotional, Community, Content delivery, Entertainment, Error message, FAQ, Gateway, Index, Informative, Journalistic, Official, Personal, Poetry, Pornographic, Prose fiction, Scientific, Shopping, User input*.

We considered broad spectar of features generally divided into four types: surface, structural, presentation and context. All features except those pertaining to URL were expressed as ratios. Since it is more probable that a certain feature would appear more frequently in longer pages, expressing features as ratios eliminates the influence of page length.

From 1,539 web pages, 1,059 are labeled with one, 438 with two, 39 with three and 3 with four labels. On average, there are 1.34 labels per web page.

More details about the data set can be found in Vidulin et al. [11, 12].

3.3 Algorithms

On the transformed data we applied LIBSVM [1] and ADABOOST [2] to learn the classifiers.

LIBSVM has built-in problem transformation functionalities, and is good at handling high number of features and sparse data. In the process of tuning the algorithm, we followed the recommendations of Hsu et al. [3]. First, we scaled feature values to fall into the $[0, 1]$ interval. Then, we applied the RBF kernel

($K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0$). Finally, we selected the parameters C and γ with the grid search in the space of models induced with exponentially growing sequences of parameters C and γ . The quality of parameters was evaluated with the 3-fold cross-validation on the training set. The parameters of a model with the best cross-validation accuracy were picked. In the case of BR transformation the choice of the parameters was separately done for each sub-classifier.

ADABOOST is a meta-learning algorithm. In our previous research [11] we boosted J48 decision trees [13], and obtained the best performance among five algorithms (sequential minimal optimization, Naive Bayes, J48 decision trees, Random Forrest and ADABOOST).

4 RESULTS AND DISCUSSION

The performances of the four classifiers induced with LIBSVM and ADABOOST on the LPS and BR data sets are presented in Table 1 and Table 2. Both LPS classifiers correctly

classified higher number of examples than the BR classifiers, which can be seen from higher exact match ratio. Further, LPS classifiers have higher recall, while BR classifiers are more precise. Higher precision of BR classifiers is due to their ability to assign N/A answer to a web page, while LPS classifiers must choose between one of the categories from the fixed list. When BR classifier states N/A this is accounted as FN for each of the actual classes, increasing the number of FN and consequently influencing only the recall. In contrast, when LPS classifier makes wrong decision it is accounted as FN when actual class is considered and FP when incorrectly predicted class is considered. Therefore, it influences both precision and recall, and consequently F-Measure.

Table 1: The performances of the two classifiers induced with LIBSVM (M.=measure, T.=transformation, C.=comparison)

LIBSVM			
M./T.	LPS	BR	C.
EX	38%	29%	▲
π (micro)	0.47	0.55	▼
ρ (micro)	0.40	0.29	▲
F(micro)	0.43	0.38	▲
π (macro)	0.58	0.57	▲
ρ (macro)	0.40	0.33	▲
F(macro)	0.44	0.38	▲

Table 2: The performances of the two classifiers induced with ADABOOST (M.=measure, T.=transformation, C.=comparison)

ADABOOST			
M./T.	LPS	BR	C.
EX	38%	29%	▲
π (micro)	0.45	0.68	▼
ρ (micro)	0.41	0.32	▲
F(micro)	0.43	0.43	=
π (macro)	0.48	0.71	▼
ρ (macro)	0.44	0.34	▲
F(macro)	0.45	0.44	▲

To understand which classifier better recognizes multi-genre web pages, we broke down the correct predictions into categories according the number of actual labels and the number of correctly predicted labels (Tables 3 and 4). To allow the comparisons between the LPS and the BR classifiers, we transformed the numbers of correctly predicted labels into ratios. For example, in the case of two-genre pages there were 128 examples. The LIBSVM LPS classifier correctly predicted one of the two genres for 45 examples or the 35% of the two-genre cases, and two of the two genres for 14 examples or the 11% of the two genre cases. As can be seen from the Tables 3 and 4, the removal of the examples labeled with

improperly represented categories in LPS setting (cf. Evaluation section), resulted in different number of web pages per the number of labels category.

From the presented results it can be seen that higher EX of the LPS classifiers in comparison to the BR classifiers is due to better recognition of single-genre web pages (on average 10 percentage points improvement). In the case of two-genre web pages the quality of recognition was on average the same - around 10%. Because of the small number of the three-genre and four-genre web pages, we cannot make proper conclusions for more than two genres per page.

Table 3: LPS transformation: Correctly classified examples having one, two, three and four labels (ACT.=actual labels, PRED.=predicted labels)

NO-LABELS		LPS			
ACT.	PRED.	LIBSVM		ADABOOST	
1	1	169/352	48%	165/352	47%
2	1	45/128	35%	44/128	34%
2	2	14/128	11%	11/128	9%
3	1	2/4	50%	1/4	25%
3	2	0/4	0%	1/4	25%
3	3	0/4	0%	0/4	0%

Table 4: BR transformation: Correctly classified examples having one, two, three and four labels (ACT.=actual labels, PRED.=predicted labels)

NO-LABELS		BR			
ACT.	PRED.	LIBSVM		ADABOOST	
1	1	133/353	38%	132/353	37%
2	1	54/146	37%	54/146	37%
2	2	14/146	10%	11/146	8%
3	1	5/13	38%	5/13	38%
3	2	2/13	15%	2/13	15%
3	3	0/13	0%	0/13	0%
4	1	0/1	0%	0/1	0%
4	2	0/1	0%	0/1	0%
4	3	0/1	0%	0/1	0%
4	4	0/1	0%	0/1	0%

5 CONCLUSION AND FUTURE WORK

In this paper we compared two problem transformation methods for multi-label web genre classification - LPS and BR - to understand which one can better capture the correlations between genres that appear together in multi-genre web pages. To this end four multi-label classifiers were induced, two LPS and two BR. Overall performances of both LPS and BR classifiers are relatively low. For example, the

exact match ratio is around 38% for LPS and around 29% for BR classifiers. A potential reason is the high number of features (2,491) in comparison to the number of examples (1,539), an aspect which we intend to address in further experiments through feature selection.

BR classifiers outperformed LPS classifiers in precision due to N/A answer, while LSP classifiers performed better considering other criteria such as recall and exact match ratio. Further evidence showed that the differences between the two approaches were largely in different ability to correctly classify single-genre web pages. Therefore, we can conclude that under presented circumstances both approaches fail to address the issue of correct recognition of multi-genre web pages.

As part of the future work, the approaches could be tested on another data set, preferably larger and with more multi-label examples. Several other ML algorithms could be applied. This would rule out the influences of the specific corpus and the specific ML algorithms.

References

- [1] R.E. Fan, P.H. Chen, and C.J. Lin. Working Set Selection Using Second Order Information for Training Support Vector Machines. *The Journal of Machine Learning Research*, 6:1889–1918, 2005.
- [2] Y. Freund and R.E. Schapire. Experiments with a New Boosting Algorithm. In *Proceedings of the 13th International Conference on Machine Learning*, pages 148–156. Morgan Kaufmann, 1996.
- [3] C.W. Hsu, C.C. Chang, and C.J. Lin. A Practical Guide to Support Vector Classification. <http://www.csie.ntu.edu.tw/~cjlin>, 2008.
- [4] J. Karlgren. *Stylistic Experiments for Information Retrieval*. PhD thesis, Stockholm University, 2000.
- [5] J. Karlgren and D. Cutting. Recognizing Text Genres with Simple Metrics Using Discriminant Analysis. In *Proceedings of the 15th International Conference on Computational Linguistics*, pages 1071–1075. Association for Computational Linguistics, 1994.
- [6] C.S. Lim, K.J. Lee, and G.C. Kim. Multiple Sets of Features for Automatic Genre Classification of Web Documents. *Information Processing and Management*, 41(5):1263–1276, 2005.
- [7] M. Santini. *Automatic Identification of Genre in Web Pages*. PhD thesis, University of Brighton, 2007.
- [8] M. Santini. Zero, Single, or Multi? Genre of Web Pages Through the Users’ Perspective. *Information Processing and Management*, 44(2):702–737, 2008.
- [9] A. Stubbe, C. Ringlstetter, and K.U. Schulz. Genre as Noise: Noise in Genre. *International Journal on Document Analysis and Recognition*, 10(3):199–209, 2007.
- [10] G. Tsoumakas and I. Katakis. Multi-Label Classification: An Overview. *International Journal of Data Warehousing and Mining*, 3(3):1–13, 2007.
- [11] V. Vidulin, M. Luštrek, and M. Gams. Using Genres to Improve Search Engines. In *Proceedings of International Workshop: Towards Genre-Enabled Search Engines: The Impact of Natural Language Processing*, pages 45–51, 2007.
- [12] V. Vidulin, M. Luštrek, and M. Gams. Multi-Label Approaches to Web Genre Identification. *Journal for Language Technology and Computational Linguistics*, 24(1):93–110, 2009.
- [13] I.H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques*. Elsevier, 2005.

Zbornik 12. mednarodne multikonference
INFORMACIJSKA DRUŽBA – IS 2009

Proceedings of the 12th International Multiconference
INFORMATION SOCIETY – IS 2009

Vzgoja in izobraževanje v informacijski družbi

Education in Information Society

Uredili / Edited by

Vladislav Rajkovič, Tanja Urbančič, Mojca Bernik

<http://is.ijs.si>
<http://lopes1.fov.uni-mb.si>

16. oktober 2009 / October 16th, 2009
Ljubljana, Slovenia

PREDGOVOR

Kako iz podatkov izluščiti informacijo, kako informacijo pretvoriti v znanje, kako priti od znanja k spoznanjem in modrosti? Odgovore na ta vprašanja išče človek že od prazgodovinskih časov. O tem pričajo npr. tudi stvaritve jamskega človeka na stenah Altamire, Lascaux-ja in drugih jam, pa tudi bogastvo knjižnega gradiva vse do potencialne neskončnosti interneta.

V teh prizadevanjih smo si ustvarili različne tehnologije in metodologije, ki pa same po sebi niso dale in ne morejo dati zadovoljivih odgovorov. Odgovor lahko da skupaj s tehnologijo in metodologijo le človek, ki pozna vsebino. Šele harmonizacija vsebine, tehnologije in metodologije lahko rojeva novo vrednost, človeško vrednost.

Osvežimo, kaj nam na poti od podatka do znanja prinaša računalnik? Računalnik že zdavnaj ni več le stroj za računanje, ampak je stroj za obravnavo simbolov. Znanje v njem je predstavljeno s simboli in je za razliko od knjige, kjer gre (po Einsteinu) za mrtvo znanje, lahko tudi živo. To pomeni, da deluje in rešuje probleme. Je pa bistvena razlika med človekom in strojem. Človek je v svojem reševanju problemov kreativen. Se znajde v novih nepredvidljivih situacijah. Računalnik pa žal, ali pa na srečo, praktično ne, kljub prizadevanjem na področju umetne inteligence. Pa vendar računalnik uvrščamo med temeljna tehnološka odkritja. Zaradi računalništva in informatike ustvarjamo nova znanja in spoznanja, kar vodi v nova odkritja, do katerih brez tega nebi prišli.

Nova znanja in spoznanja ter odkritja so možna in nujna tudi v vzgoji in izobraževanju. Tehnologija nam omogoča, da je obstoječe znanje lahko dosegljivo vsakomur in vedno. To dejstvo korenito spreminja šole vseh vrst. Iskanje pravih sprememb ni in ne more biti zaključena zgodba. Vsak udeleženec procesa šolanja od učencev preko učiteljev do staršev se mora zavedati svoje dolžnosti v iskanju novih rešitev problemov, ki vodijo tudi v nove, bolj človeške odnose med ljudmi v tem procesu in izven njega. Ne gre le za iskanje »receptov«, ki narekujejo, kako se stvari počne bolj ali manj uspešno, a pri tem ne osmislijo samega početja. Strokovnjak pa je tisti, ki rešuje probleme tako, da sproti zajema iz teorije področja svojega delovanja in s tem lažje pride ustvarjalno do novih rešitev. Prav takim novim rešitvam pa je namenjena tudi letošnja že 12. konferenca »Vzgoja in izobraževanje v informacijski družbi«, ki poteka v sklopu multikonference »Informacijska Družba«.

Vladislav Rajkovič, Mojca Bernik, Dejan Dinevski, Tanja Urbančič

PREFACE

How to extract information from data, how to transform information into knowledge and how to make the transition from knowledge to cognition and wisdom? Humans have been looking for answers to these questions ever since the prehistoric times. As witnesses to this search stand among others creations of the cavemen on the walls of Altamira, Lascaux and other caves, the wealth of literary heritage as well as Internet with its potential infinity.

In these endeavors humans have created various technologies and methodologies which alone have not led and could not have led to satisfactory answers. The answer can be provided only together with the human who is familiar with the content. Only when the content, technology and methodology are in sync, new value can be created, that is the human value.

Let us quickly remind ourselves what is it that the computer contributes on the path from data to knowledge? For quite some time now the computer has no longer been considered only a machine for merely performing calculations but a machine for handling symbols. The knowledge inside the computer is presented in the form of symbols. In comparison to a book, inside which according to Einstein the knowledge is 'dead', the knowledge inside a computer can also be 'alive'. This means that the computer operates and solves problems. However, there is one key difference between a human and a computer. A human takes a creative approach to problem solving and finds a way in new unpredictable situations. The computer, fortunately or unfortunately, practically cannot do this despite the endeavors in the field of artificial intelligence. Still, computers are regarded as a basic technological invention. Due to computer science and informatics we create new knowledge what leads to new inventions that would otherwise not be within our reach.

New knowledge as well as inventions are possible and necessary also in education. Technology enables us to make existing knowledge available to everyone at all times. It is this fact that is profoundly changing schools of all kinds. The search for appropriate changes is a never-ending story. Every stakeholder in the process of education from students to teachers and parents has to fulfill their duty for finding new solutions to problems. This brings about new more humane relationships among people within and outside the educational process. It is not a search for 'recipes' on how to do something but a matter of knowing what is the purpose of doing this in the first place. A professional is the one who puts to use theory from her/his field of interest and delivers new solutions through a creative process. To such new professional solutions is also dedicated this year's 12th conference »Education in information society« that is held under the framework of the multi-conference »Information society«.

Vladislav Rajkovič, Mojca Bernik, Dejan Dinevski, Tanja Urbančič

PROGRAMSKI ODBOR / PROGRAMME COMMITTEE

Predsednik

Vladislav Rajkovič

Univerza v Mariboru, Fakulteta za organizacijske vede in Institut Jožef Stefan

Alenka Adamič

Zavod MIRK

Vladimir Batagelj

Univerza v Ljubljani, Fakulteta za matematiko in fiziko

Igor Bernik

Univerza v Mariboru, Fakulteta za organizacijske vede

Mojca Bernik

Univerza v Mariboru, Fakulteta za organizacijske vede

Janez Bešter

Univerza v Ljubljani, Fakulteta za elektrotehniko

Lea Bregar

Univerza v Ljubljani, Ekonomska fakulteta

Borut Čampelj

Ministrstvo za šolstvo in šport RS

Dejan Dinevski

Univerza v Mariboru, Medicinska fakulteta

Saša Divjak

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko

Tomi Dolenc

ARNES

Marko Ferjan

Univerza v Mariboru, Fakulteta za organizacijske vede

Ivan Gerlič

Univerza v Mariboru, Pedagoška fakulteta

József Györkös

Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko

Eva Jereb

Univerza v Mariboru, Fakulteta za organizacijske vede

Mitja Jermol

Institut Jožef Stefan

Peter Kokol

Univerza v Mariboru, Fakulteta za zdravstvene vede

Alenka Krapež

Gimnazija Vič

Nives Kreuh

Zavod RS za šolstvo

Tanja Logar

Center za poklicno izobraževanje

Robert Marinšek

Center za mobilnost in evropske programe izobraževanja in usposabljanja

Gregor Mohorčič

Zavod RS za šolstvo

Mojca Orel

Gimnazija Moste

Jože Rugelj

Univerza v Ljubljani, Pedagoška fakulteta in Institut Jožef Stefan

Niko Schlamberger

Slovensko društvo Informatika

Tomaž Skulj

Zveza za tehnično kulturo Slovenije

Branislav Šmitek

Univerza v Mariboru, Fakulteta za organizacijske vede

Olga Šušteršič

Univerza v Ljubljani, Zdravstvena fakulteta

Tanja Urbančič

Univerza v Novi Gorici

Margareta Vrtačnik

Univerza v Ljubljani, Naravoslovno tehniška fakulteta

Rado Wechtersbach

Zavod RS za šolstvo

Uporaba IKT pri pouku (tujega jezika), njene prednosti in (možne) slabosti ter nujni pogoji za kakovostno delo z IKT

The use of information communication technology (ICT) in foreign language teaching, its advantages and (possible) disadvantages and the essential conditions for a quality ICT work

Jelka Bajželj

ESIC Kranj, Gimnazija
jelka.bajzelj@guest.arnes.si

Povzetek:

V prispevku avtorica predstavi konkretne možnosti uporabe IKT pri pouku (tujega jezika), številne prednosti tovrstnega dela in tudi možne pomanjkljivosti oz. slabosti. Glavna prednost vključevanja IKT v učenje in poučevanje je zagotovo višja motivacija, a učenci in dijaki pogosto niso vajeni samostojnega dela z računalnikom, zato je pri razvijanju ustreznih učnih strategij dijakov potrebna stalna podpora učitelja. Avtorica spregovori tudi o nujno potrebnih pogojih za kakovostno delo z IKT, pri čemer je (med drugim) prisotna problematika (ne)pridobivanja specialnih didaktičnih znanj za uporabo IKT pri pouku znotraj dodiplomskega študija bodočih učiteljev. Uporaba novih informacijsko komunikacijskih tehnologij za učenje in poučevanje pravzaprav na nek način postavlja pod vprašaj vsaj nekatere oblike institucionaliziranega učenja.

Ključne besede: informacijsko komunikacijska tehnologija, pouk tujih jezikov, kakovostno učenje in poučevanje, samostojno učenje, komunikacija, motivacija

Abstract:

In the article the authoress presents the possibilities of the use of ICT in foreign language teaching, as well as its advantages and disadvantages. The main advantage of ICT in learning and teaching is by all means the motivation although the students very often are not capable to develop the learning strategies independently using a computer and therefore they constantly need the help and advice of a teacher. The essential conditions needed to use ICT are also discussed. The (not) acquiring of specialized didactic knowledge about ICT during the study of the future teachers is also mentioned. The use of ICT in learning and teaching somehow calls some institutionalized ways of teaching into question.

Keywords: Information communication technology ICT, foreign language teaching, quality learning and teaching, independent learning, communication, motivation

Model za oceno vpliva stalnega strokovnega izobraževanja na kakovost učitelja

A Model of Effect of Permanent Education on Teachers Quality

Sašo Bizant

TŠC Kranj, Slovenija

saso.bizant@guest.arnes.si

Povzetek

Namen prispevka je izdelati model za oceno vpliva stalnega strokovnega izobraževanja na kakovost učitelja, ki bo pripomogel h kvalitetnejši izvedbi izobraževalnega procesa saj se bodo učitelji na osnovi modela lahko samovrednotili in izboljšali kritične točke svojega dela. Izdelan je model, in sicer iz variante stalnega strokovnega izobraževanja in variante kakovosti učiteljevega dela.

Z modelom za oceno vpliva stalnega strokovnega izobraževanja na kakovost učitelja smo interpretirali dobljene rezultate analize vprašalnika, ki je potekala v mesecu marcu 2009. Na osnovi analize vprašalnikov smo strukturirali kriterije, jih ovrednotili, določili zaloge vrednosti in funkcijske koristnosti, opisali variante, ovrednotili variante in jih s pomočjo računalniškega programa za večparametrsko odločanje DEXi analizirali.

Ključne besede: večparametrsko odločanje, stalno strokovno izobraževanje, kakovost učiteljevega dela, računalniško orodje DEXi.

Abstract

The purpose of the article is to draw up a model to assess the impact of continuous professional education on teachers' daily work. The model could contribute to better quality of education as teachers will be able to use it for self-assessment purposes and thus improve the main components of their work. The model is based on the variant of permanent professional education and training and the variant of quality.

By using the model, I evaluated and interpreted the information collected from a questionnaire devised for this purpose (the answers were provided during March 2009). I then structured the different criteria, analysed it, defined the values in terms of their practicality, described and assessed the variants and analysed them by using the programme for multiparameter decision-making DEXi.

Keywords: multiparameter decision-making, continuous professional education and training, quality of teachers, computer tool DEXi.

Evalvacija izobraževalnega procesa s poudarkom na vojaških vsebinah

Education Process Evaluation with Emphasis on Military Contents

Liliana Brožič, Dušan Sušnik,

Ministrstvo za obrambo, Slovenska vojska, Vojkova cesta 55a, Ljubljana.

liliana.brozic@mors.si, dusan.susnik@mors.si

Povzetek

Evalvacija kot proces vrednotenja je eden izmed elementov izobraževanja, ki pomembno vplivajo na njegovo kakovost ne glede na število njegovih ponovitev. Celovitost vrednotenja prinaša dodano vrednost v obliki novih usmeritev za celoten izobraževalni proces, ki jih lahko z vsako ponovitvijo implementiramo in sproti nadgrajujemo.

V Slovenski vojski se zavedamo pomena kakovosti v izobraževalnih procesih. Zanimiv izziv predstavlja organizacija izobraževanja Višje štabnega šolanja v sodelovanju z javnimi izobraževalnimi institucijami.

O razlogih za organizacijo takega izobraževanja, njegovem pomenu za Slovensko vojsko in evalvaciji izobraževalnega procesa bo predstavljeno v prispevku.

Ključne besede: evalvacija, vojaško izobraževanje in usposabljanje, evalvacija vojaških vsebin, Višje štabno šolanje.

Abstract

Evaluation as a process of validity is one of the education elements with high important influence of quality no matter of the number of its repetition. Entirely of validity brings new extra value as a form of aim towards to complete education process. They can be implemented each time process is on again and upgrade them at same period.

The significant of quality in education process is something we are very aware of in Slovenian Armed Forces. Senior Staff Cours with cooperation with public education institutions offers interesting challenge.

Article presents reasons for organizing military education process as mentioned, its importance for Slovenian Armed Forces and evaluation of a process .

Keywords: Evaluation, Military Education and Training, Military Contents Evaluation, Senior Staff Course.

Uporaba interaktivne table pri matematiki v prvem triletju osnovne šole

Use of an Interactive Board with Mathematics in the First Trimester of Primary School

Urška Bučar
OŠ Dolenjske Toplice
urska.bucar@guest.arnes.si

Povzetek

V prispevku opisujemo uporabo INTERAKTIVNE TABLE pri pouku matematike v prvem triletju osnovne šole. Danes smo preplavljeni z moderno tehnologijo. Včasih se nam je zdelo vsakdanje to, da smo imeli televizijo, radio, videorekorder, danes pa otroci živijo v svetu računalnikov in interneta in nanje so navajeni. Nova računalniška tehnologija v razred prinaša podobe sveta, zvočne in slikovne efekte, nov način učenja in poučevanja. Interaktivna tabla za svoje delovanje potrebuje projektor in računalnik ter svoj program, ki nudi številne možnosti, s katerimi popestrimo vsakdanji pouk. V sklopu menija najdemo slikovne pripomočke, zvočne posnetke, možnosti dodajanja videa in najpomembnejše – strani lahko pripravimo interaktivne, s čimer zadovoljimo potrebe učnih stilov, ki se v razredu pojavljajo. Šole se opremljajo z interaktivnimi tablam, učitelji pa se bomo morali za njeno uporabo tehnološko in didaktično pripraviti. V okviru svojega podiplomskega študija se že nekaj ukvarjamo s pripravo IKT didaktičnega gradiva za poučevanje matematike na razredni stopnji. Interaktivna tabla nam je omogočila pripravo didaktičnega gradiva, ki bo zadovoljilo različne učne stile in bo tako otrokom zanimivo in privlačno.

Ključne besede: matematika, interaktivna tabla, orodja, delo v razredu, didaktični pripomoček

Abstract

In the lecture, we describe the use of the INTERACTIVE WHITE BOARD for mathematics lessons in the first triennium of primary school.

Today we are flooded with modern technology. We used to take for granted the fact that we had a television, radio, video recorder, but nowadays children live in a world of computers and internet and they are accustomed to them. New computer technology brings images of the world into the classroom, sound and image effects and new ways of learning and teaching.

To work, an interactive board needs a projector, a computer and its supporting program with which we can liven up everyday lessons. Under the tab Tools, we can find image aids, sound recordings, options to add video and most importantly -- we can make interactive pages with

which we can satisfy the needs of all the learning styles that are present in the classroom.

In the future, schools will be equipped with interactive whiteboards, and teachers will have to be technically and didactically prepared to use them.

During postgraduate studies, we have prepared an outline of IKT didactic material for teaching mathematics in the classroom. The interactive board has enabled us to prepare didactical lesson plans, which will satisfy all learning styles and will be both interesting and attractive to pupils.

Keywords: Mathematics, interactive board, tools, classroom work, didactical aids

Medosebna vloga ravnatelja - managerja v procesu dela in funkciji humanistično-antropocentričnega managementa človeških virov

Principal Manager's Interacting Role in the Process of Work and the Function of Humanistic-Antrophocentric Human Resources Management

Bojan Burgar

Osnovna šola Ormož, Slovenija
os.ormoz@guest.arnes.si

Jože Florjančič, Mojca Bernik

Univerza v Mariboru, Fakulteta za organizacijske vede, Slovenija
joze.florjancic@fov.uni-mb.si, mojca.bernik@fov.uni-mb.si

Povzetek

Medosebne vloge ravnatelja v procesu dela in funkcijah humanistično antropocentričnega managementa zahtevajo njegovo temeljno osredotočenost na zaposlene in uporabnike storitev ter kreativnost posameznika kot strokovnjaka in osebnosti. To je usmerjenost medosebne vloge kreativne obzirnosti in odprtega stila vodenja.

Ključne besede: človeški viri, medsebojni odnosi, medosebna vloga, kompetenca, humanistično-antropocentrični management, kreativna obzirnost, vodenje

Abstract

Principal's interacting role in the process of work and in functions of humanistic – anthropocentric management demand his fundamental concentration on employees and users of services as well as individual's creativity as an expert and personality. This is a guidance of the interacting role of creative tactfulness and the open styl of leadership.

Keywords: Human resources, mutual relationships, interacting role, competence, humanistic – anthropocentric management, creative tactfulness, leadership.

Pouk matematike v osnovni šoli z uporabo e-gradiv

Mathematical Lessons and E-Learning Materials in Primary Education

Nevenka Colja
OŠ Miren, Slovenija
nevenka.colja@gmail.com

Povzetek

Hiter razvoj sodobne informacijsko-komunikacijske tehnologije prinaša velike spremembe tudi na področje šolstva, v proces izobraževanja. Sodobna informacijsko-komunikacijska tehnologija omogoča uvajanje novih oblik učenja, poučevanja ter posredovanja znanj. Učitelji morajo svoje oblike in metode dela prilagajati potrebam in zahtevam sodobnega časa. S pomočjo uporabe računalnika in različnih didaktičnih gradiv za učenje in poučevanje, ki jih najdemo na internetu, postane pouk bolj zanimiv, saj so učenci bolj motivirani za delo.

Prispevek se osredotoča na uporabo e-gradiv pri pouku matematike v osnovni šoli. Ideja za poglobljeno delo z e-gradivi se je porodila ob izvajanju projekta »Mesec širjenja uporabe e-gradiv«, katerega cilj je bil vzpodbuditi preizkušanje in uporabo e-gradiv pri pouku, pridobiti mnenja učiteljev in učencev o njihovi uporabni vrednosti in usmeritve za nadaljnji razvoj in nadgradnjo e-gradiv. Tudi po zaključku projekta smo pri pouku matematike še naprej uporabljali e-gradiva v šoli, kot tudi za domače delo. Na podlagi ankete med učenci in na podlagi mnenja učiteljice so v prispevku prikazane nekatere prednosti in pomanjkljivosti, ki se pojavljajo pri delu z e-gradivi.

Ključne besede: informacijsko-komunikacijska tehnologija, poučevanje, izobraževanje, e-gradiva

Abstract

Fast development of modern information and communication technology brings great changes to the process of education. The modern technology enables the introduction of new learning and teaching methods as well as knowledge transfer. Teachers' working methods have to be adjusted to the needs and demands of the present time. Computers and different learning and teaching materials, which can be found on the internet, motivate pupils and therefore make classes more interesting.

This paper focuses on the use of e-learning materials in mathematical lessons in primary schools. The idea for more detailed work with e-learning materials developed during the project "Mesec širjenja uporabe e-gradiv" (Increasing the Use of E-Learning Materials Month). The aim of this project was to encourage the testing and use of e-learning materials in teaching; to get teacher's and pupils' opinion on their useful value and to seek guidance for further development and upgrading of such materials. When the project was concluded, the teacher continued to use e-learning materials in class and for homework. This paper shows some advantages and disadvantages of the use of e-learning materials, based on teacher's and pupils' opinion.

Keywords: information and communication technology, teaching, education, e-learning materials

Google Apps - ozadje, implementacija in uporaba

Google Apps - background, implementation and use

Dejan Cvitkovič

Srednja šola za gostinstvo in turizem v Ljubljani, Slovenija

dejan.cvitkovic@ssgtlj.si

Povzetek

Google Apps je ena izmed storitev, ki jo organizacijam ponuja Google. Storitve temelji na računalništvu v oblaku, katerega osnovna značilnost je shranjevanje programov in podatkov na internetu, dostop do njih pa je možen iz poljubne, na internet povezane lokacije in naprave. Kot tak se Google Apps uvršča tudi med programsko opremo kot storitev (angl. Software as a Service – SaaS), katerega poglobljena značilnost je prenos bremena vzdrževanja programske opreme na ponudnika, saj je le-ta shranjena na enem mestu in ne na vsakem računalniku posebej. V prispevku je poleg teoretičnih opredelitev zgoraj omenjenih pojmov opisan tudi primer implementacije Google Apps v organizacijo. Prav tako so predstavljeni konkretni primeri uporabe.

Ključne besede: Google Apps, računalništvo v oblaku, programska oprema kot storitev

Abstract

Google Apps is one of the services for organisations from Google. It is based on cloud computing, whose primary characteristic is storing applications and data on the internet. Access to applications and data is possible from any internet connected location or device. As such Google Apps is one of Software as a Service (SaaS) solutions, whose primary characteristic is shifting the burden of software maintenance on the provider since that is the place where all applications run. The article describes theoretical backgrounds of the above mentioned technologies and guidelines for implementation of Google Apps into an organisation. Some practical examples of use are also described.

Keywords: Google Apps, Cloud computing, Software as a Service

Prosto dostopni izobraževalni viri v e-izobraževanju

Open Educational Resources in e-Learning

Dejan Dinevski¹, Samo Fošnarič¹, Tanja Arh²

¹University of Maribor, Faculty of Education, Koroška c. 160, Maribor, Slovenia

²Institute Jožef Stefan, Jamova cesta 39, 1000 Ljubljana, Slovenia

Povzetek

Prosto dostopni učni viri v e-izobraževanju so vir informacij prihodnosti za udeležence vseživljenjskega izobraževanja. Oprta koda in odprti standardi predstavljajo temelj »Gibanja prosto dostopnih učnih virov«, ki se v zadnjem desetletju oblikuje na globalni ravni. V članku so raziskane karakteristike odprte kode v relaciji z e-izobraževanjem, obstoječimi in novimi pedagoškimi principi ter problemi avtorske zaščite. Predstavljene so dobre prakse, ideje in obstoječe iniciative ter vizija prihodnosti prosto dostopnih učnih virov.

Ključne besede: e-izobraževanje, odprta koda, prosto dostopni učni viri

Abstract

Open educational resources in e-learning are the future source of information for lifelong learners. Open source and open standards are defined as the basis of the "Open educational resource movement" that is beginning to form on a global level in the last decade. The characteristics of the OS are investigated in the relation to e-Learning, existing and new pedagogical principles and copyright issues. Several good practices, ideas and existing initiatives are presented and the vision of the future of open educational resources is introduced.

Keywords: e-Learning, Open Source, Open educational resources

Obravnavava umetnostnega besedila – pravljice s pomočjo E – gradiv v 6. Razredu

Treatment of Art Text – Fairytale with a Help of E-Materials in 6. Class

Miroslava Fon

Osnovna šola Ivana Roba Šempeter pri Gorici
miroslava.fon@gmail.com

Povzetek

V 6. razredu osnovne šole učenci spoznavajo umetnostna in neumetnostna besedila. Med umetnostnimi besedil spoznajo pravljico in njene značilnosti, potem pa to znanje nadgradijo tudi s pisanjem lastne avtorske pravljice.

Pri tem delu pa je lahko tako učitelju, kot učencu v veliko pomoč računalnik.

Različni računalniški programi nudijo veliko možnosti za usvajanje, utrjevanje in poglobljanje novih znanj. Poznavanje pravljice so učenci nadgradili tako, da so najprej napisali lastno avtorsko pravljico in jo ilustrirali. Delo so nato nadaljevali s pomočjo nekaterih e- gradiv.

Ključne besede: pravljica, računalnik, računalniški program, e- gradiva.

Abstract

In 6. class of elementary school pupils get in touch with art and non art texts. Among the art texts they get introduced with fairytales and their characteristics and then they upgrade this knowledge with writing of their own fairytale.

At this work a computer is a big help to a teacher, as to a pupil.

Different computer programs offers a lot of possibilities for creation, consolidation and deepening of new notions. The pupils have upgraded the knowledge of fairytales with writing of their own fairytale and they also illustrated it. Then they continued the work with several e-materials.

Keywords: fairytale, computer, computer program, e-material

Izzivi novih tehnologij in šola bodočnosti

Challenges of Advanced Technologies and School of the Future

Ivan Gerlič

Univerza v Mariboru, Fakulteta za naravoslovje in matematiko
ivan.gerlic@uni-mb.si

Povzetek

Doba napredne tehnologije zahteva drugačnega posameznika; takšnega, ki bo lahko nemoteno deloval v dobi digitalizacije in bo lahko s pomočjo znanj in prednosti napredne informacijsko komunikacijske tehnologije bolje spoznal in varoval okolje. Ni boljšega časa za učenje in konkretiziranje takih vrednot kot v šoli. Današnji učenci, učenci digitalne dobe, so vse bolj veščji v uporabi sodobnih medijev in mrež, s čimer zahtevajo spreminjanje razmer v šoli kot instituciji in tudi spremenjeno vlogo učiteljev. Bodo nove informacijsko in multimedijsko oz. hipermedijsko podprte učilnice in nove didaktične ideje izboljšale učne rezultate in spodbudile inovativnost in večje veselje do učenja in znanja? To se sprašujemo in na to bomo skušali v našem prispevku odgovoriti.

Ključne besede: tradicionalna šola, šola prihodnosti, računalniška učilnica, multimedijska učilnica, hipermedijska učilnica, izobraževalna omrežja

Abstract

The era of advanced technology claims a different individual. The individual that would undisturbed function in the era of digitalization and would better perceive and protect our environment with the help of accomplishments and benefits of advanced information and communication technology (ICT). There is no better time for learning and concretization of these values as in school. Nowadays pupils, the pupils of digital era, are more and more experienced in the usage of contemporary media and networks. That is why they require the modification of conditions and teachers' role in the school. Will new information and multimedia or hypermedia support for classrooms and new didactical ideas improve learning results, and stimulate innovation and greater pleasure towards learning and knowledge? All these questions will be answered in our contribution.

Keywords: traditional school, school of the future, computer classroom, multimedia classroom, hypermedia classroom, education networks.

Razvijanje mehkih znanj na tehničnih fakultetah: izkušnje s študentskim delom na projektih

Developing Soft Skills for Engineering: Experience With Student Team Projects

Franc Gider^{1,2}, Tanja Urbančič¹

¹University of Nova Gorica, Vipavska 13, 5000 Nova Gorica, Slovenia

²Slovenian Technology Agency, Dunajska 22, 1511 Ljubljana, Slovenia

E-mail: franc.gider@tia.si

Povzetek

Medtem ko si sodobne tehnologije počasi utirajo pot v izobraževalne programe na vseh nivojih, ne smemo pozabiti tudi na razvijanje sposobnosti za uspešno uporabo teh tehnologij za reševanje kompleksnih problemov v sodobnem svetu. Članek predstavlja obetaven način pridobivanja kompetenc za delo v interdisciplinarnih projektih skupinah. V predmetu »Skupinski projekt« na bolonjskem magistrskem programu Gospodarski inženiring študentje pridobivajo izkušnje in razvijajo sposobnosti timskega dela, koordiniranja, projektnega vodenja, raziskovanja, reševanja konkretnih problemov, predstavitev rezultatov javnosti in upravljanja s časom. Predmet je na programu v prvem letniku kot obvezna sestavina študijskega programa. Študentom nudi konkretno izkušnjo projektnega dela, podobnega tistemu v realnih poslovnih okoljih. Preizkusijo se v interdisciplinarnem timskem delu, pri čemer morajo premagovati negotovosti in težave, ki se pojavljajo pri tovrstnih projektih. Po prvih izvedbah se je nabralo že več izkušenj, s pomočjo katerih izboljšujemo izvajanje predmeta. Kljub temu še vedno predstavlja precejšen izziv tako za študente kot za profesorje.

Ključne besede: skupinsko delo, študentski projekt, gospodarski inženiring

Abstract

While advanced technologies are entering schools at all levels, we claim that other skills needed to deal with them in a complex nowadays world should not be neglected in the education system. In the article a promising approach to the interdisciplinary post graduate education is described. The idea of the »Team project« course is to give the students opportunity to gain experience with an extensive project where different skills are needed, e.g. team work, coordination of work, project management, research, problem solving, public presentations, time management. The course is obligatory and held in the first year of the master study program at the School of Engineering and Management of the University of Nova Gorica. The main advantage of the course is to give the students hands-on experience of the work on a project which is very close to the reality. Students are forced to gain the experience in team work, and to overcome the uncertainties and obstacles which arise in such projects. Since the launch of the course several improvements were made with the aim to improve the performance of the course. Nevertheless, the course remains a challenge for students as well as for teachers.

Keywords: team work, student project, engineering and management

Prednosti blokovnega programiranja robotov v osnovni šoli

Advantages of Function Block Programming

Milan Hlade

OŠ Koroška Bela Cesta talcev 2 4270 Jesenice

milan.hlade@gmail.com

Povzetek

Prispevek opisuje primer uporabne povezave računalnika, vmesnika in stroja. Učenci osvojijo osnove programiranja. Utrdijo pridobljeno znanje o elektriki, računalništvu, tehniki, fiziki, kemiji, biologiji.... Spoznajo dobre in slabe strani naprav, avtomatov. Spoznajo tudi področja uporabe robotskih naprav in avtomatiziranih strojev. Znajo uporabiti pridobljeno znanje na različnih področjih in s tem tudi znajo povezovati različna znanja.

Predmet navaja učence na koristno uporabo računalnika. Spoznajo, da je računalnik mogoče uporabiti za krmiljenje povsem vsakdanjih naprav. Učno uspešnejšim učencem nudi omenjeni program še mnogo več. Vse možnosti in funkcije programa niso opisane zaradi pomanjkanja prostora.

Ključne besede: računalnik, vmesnik, mobilni robot, blokovno programiranje, medpredmetno povezovanje

Abstract

In this contribution I describe an example of useful connection between computer, interface and machine. Pupils get familiar with the basis of programming. They consolidate knowledge about electricity, computer science, technology, physics, chemistry, biology... They get to know good and bad sides of devices, machines. They know how to use the knowledge on different areas and consequently connect different knowledges.

This subject aim is useful use of computer. Pupils learn that the computer can be used for steering entirely common devices. This program offers much more especially to very successful pupils. All possibilities and functions are not described due to lack of space.

Keywords: computer, interface, mobile robot, block programming, intersubject connection

Nadaljni koraki v razvoju e-izobraževanja v slovenskem šolskem prostoru

Further steps in the development of better quality e-learning in Slovenian educational environment

Boris Horvat¹, Matija Lokar¹, Primož Lukšič¹, Damijan Omerza², Alen Orbanic¹

¹IMFM in FMF UL, Jadranska 19, 1000 Ljubljana, Slovenija

²Hruška.si, Ilirska 21, 1000 Ljubljana, Slovenija

boris.horvat@fmf.uni-lj.si, matija.lokar@fmf.uni-lj.si, primoz.luksic@fmf.uni-lj.si
damijan@hruska.si, alen.orbanic@fmf.uni-lj.si

Povzetek

V tem prispevku bomo predstavili stanje na področju e-izobraževanja v slovenskem osnovnošolskem in srednješolskem prostoru v prvi polovici leta 2009 in predlagali potrebne korake, ki jih moramo narediti, da bi stanje na tem področju izboljšali. Osredotočili se bomo predvsem na kvaliteto e-gradiv, na orodja, ki jih potrebujejo učitelji ter na storitve, ki jih za kvalitetno izvajanje e-izobraževanj potrebuje širša skupnost v Sloveniji.

Ključne besede: e-izobraževanje, e-gradiva, analiza stanja, orodja, storitve, praksa, Slovenija

Abstract

The paper presents the current state of e-learning in Slovenian (primary and secondary level) education as well as the necessary steps that need to be taken to improve it. We focus mostly on the quality of existing e-learning content, on tools needed to produce content of good quality and on services needed for an ever-expanding community in Slovenia that uses e-learning in practice.

Keywords: e-learning, e-learning content, current state, tools, services, practice, Slovenia

Nekateri strukturni in kulturni problemi pri uvajanju e-izobraževanja. Z nakazanimi rešitvami.

Structural and Cultural Problems in Implementing E-Learning. With Solutions' Outline.

Marko Ivanišin
UM, FERI
marko.ivanisin@uni-mb.si

Povzetek

E-izobraževanje postaja realnost, ki se zaradi vključevanja tehnologij v poučevanje zdi bolj pisana na kožo učencem kot učiteljem. V pričujočem prispevku si bomo pogledali, na katerih mestih se pri vpeljevanju e-izobraževanja v izobraževalnih ustanovah srečujemo z ovirami, katere so te in po kakšnih mehanizmih delujejo. Z odgovorom na slednje (kulturno) vprašanje si bomo tudi pomagali pri reševanju problemov, saj je napačno verjeti, da se lahko z enostavnimi (strukturnimi) rešitvami lotimo kompleksnih (kulturnih) problemov. Tako npr. za slabšo (izobraževalno) tehnološko pismenost učiteljev oz. »starejše generacije« in njihov odklonilen odnos do novih tehnologij (izobraževanja) ni krivo samo pomanjkanje opreme in razpoložljivega časa za učenje pač pa spremembe v »logiki« novih tehnologij, ki sovpada z načinom razmišljanja tehnološko veliko bolj večjih učencev oz. mlajših generacij.

Ključne besede: e-učenje, tehnologija, paradigme

Abstract

E-education is becoming a reality that seems, because of use of technology for teaching and learning, to be more real for learners than for teachers. In this contribution we will take a look at which stages of e-education implementation at educational institutions we face obstacles, which are they and what mechanisms they work by. By answering the latest (cultural) question we will try to solve complex (cultural) questions that many falsely believe can be solved by simple (structural) solutions. Following that e.g. weak (educational) technical literacy of teachers and »older generation« and their negative attitude towards new technologies (in education) is not merely based on lack of equipment and on lack on time devoted to learning but it is primarily based on changes in technological paradigms that correspond to ways of thinking of technically much more literate learners and younger generation.

Keywords: e-learning, technology, paradigms

Uporaba programa microsoft word pri tretješolcih

The Use of the Microsoft Word Programme with the Third Grade Pupils

Alenka Kastelic

OŠ Rudolfa Ukoviča Podgrad, Podgrad 99b, 6244 Podgrad
alenka.kastelic@guest.arnes.si

Povzetek

Prispevek govori o spoznavanju učencev 3. razreda s programom Microsoft Word.

Program Microsoft Word so spoznali skozi različna besedila, vaje. Najprej so se mimogrede naučili upravljanja s tipkovnico in računalnikom. Spoznali so osnovne gumbe v programu Microsoft Word (velikost besedila, zapis velikih in malih črk, vrsta in barva pisave, različne poravnave, oštevilčenje, podčrtovanje, krepko, ležeče, orodji uveljavi in razveljavi, Word Art, vstavljanje slike), se jih naučili smiselno uporabljati ter dokument tudi shraniti. Kot končni izdelek so si učenci samostojno naredili pohvalo za opravljen računalniški tečaj. Pri izdelavi končne pohvale so učenci uporabili tudi svetovni splet – internet, s pomočjo katerega so iskali različne fotografije, smešne in zanimive figure, ki so jih vstavili v pohvalo. Pohvalo so učenci tudi samostojno natisnili. Učenci so si z usmerjanjem učiteljice ustvarili tudi elektronsko pošto na Gmailu ter se naučili poslati preprost email.

Ključne besede: tipkovnica, Microsoft Word, Word Art, osnovni gumbi, pohvala, samostojno delo, računalniški tečaj, učenci, elektronska pošta

Abstract

My article deals with the third grade pupils learning to use the Microsoft Word programme.

They learnt about this programme by using various texts and exercises. First they acquired by the way how to manage the keyboard and computer. They were acquainted with the basic keys in the Microsoft Word programme (the text font, writing of capital and small letters, type and colour of letters, different settlements, numbering, underlining, bold, italic, different orders, exercise and invalidate, Word Art, insertion of picture), then they learnt how to use them meaningfully and to save a document as well. Their final product was an independently made certificate for the accomplished computer course. At this stage the pupils used the Internet, too – they were searching for different photos, funny and interesting characters later put into the certificate. They even printed it on their own. The teacher led them to create an e-mail box on the Gmail and they were taught how to send a simple e-mail.

Keywords: keyboard, Microsoft Word, Word Art, basic keys, independent work, computer course, pupils, e-mail.

Učenje preko igre do spoznavanja računalnika

Learning Thought Play to Computer

Mojca Kogoj
mojca.kogoj@siol.net

Igra je način, kako otrok spoznava zunanji svet,
igra je za otroka učenje,
igra je zanj resna oblika vzgoje,
igra je nujno potrebna otroškemu organizmu,
ki raste.

(Krupska)

Povzetek

Otroci se praviloma vse, vsaj v zgodnjem otroštvu, učijo le na podlagi igre. Igra ima neke vrste posebnosti, po katerih se razlikuje od vseh drugih dejavnosti. Igra je svobodna spontana dejavnost, ki poteka brez katerekoli zunanje nujnosti. Igra poteka po notranji nujnosti, ta ima svojo logiko, ki je drugačna od logike zunanje realnosti, kakor jo dojemajo odrasli. Logika igre se kaže ravno v tem, da je v njej realnost predelana, in sicer v skladu z otrokovim doživljanjem.

Pravilno vključiti igro v vzgojni proces lahko tudi pomeni, da pozna vzgojni pomen igre le vzgojitelj, učitelj, medtem ko se ga otrok, ki se igra, niti ne zaveda, niti ni to njegov poglavitni cilj.

Uporabo didaktičnih iger pa je bolj ali manj optimalna, če igre ponudimo v ustreznem didaktičnem sistemu in ne le naključno. V pomoč nam je lahko več skupin didaktičnih iger tudi pri začetnem učenju računalnika. Izmed devetnajstih skupin did. iger se da izbrati najprimernejše. Izkušnje kažejo, da jih imajo otroci zelo radi. V prispevku je predstavljeno, kako se da preko did. iger mlajšim otrokom približala računalnik.

Ključne besede: otroci, didaktične igre, igre, računalnik

Abstract

Children learn most things in the early childhood during playing, which is spontaneous and without outside pressure. Playing has its own logic that is different from an adult reality. The reality during the play is reformed in accordance with child's comprehension. It is important to include playing into an educational process in such way that its intent is visible only to a teacher and not to a child. Children benefit from didactical playing most if it is part of an appropriate didactical system. Didactical games can also be used by learning basic computer skills. My experience show that children love didactical games. In my contribution I will explain how I introduced basic computer skills to children by using didactical games.

Keywords: children, didactical games, games, computer

Ali uporaba multimedije izboljša učinkovitost izobraževanja?

Does the Use of Multimedia Increase the Efficiency of Education?

Darko Korošec

Srednja šola Slovenska Bistrica

darko.korosec@guest.arnes.si

Povzetek

Multimedija združuje v celoto različne medije in deluje na več čutil hkrati. Z uporabo multimedije lahko povečamo kvaliteto izobraževanja, vendar moramo poznati omejitve dijakov in učiteljev ter omejitve tehnologije. Spremembe v načinih poučevanja so nujne in uporaba multimedije je ena od možnosti. Sama uporaba multimedije ne izboljša pouka, če pa je multimedija uporabljena neprimerno, ga lahko celo poslabša, po nekaterih raziskavah pa celo podaljša čas učenja. Opisani so primeri uporabe v srednji šoli in podani rezultati raziskave o mnenju dijakov o uporabi multimedije v šoli. Podani so zaključki in predlagane rešitve za učinkovito uporabo multimedije pri pouku.

Ključne besede: multimedija, učenje, učinkovito izobraževanje, informacijska in komunikacijska tehnologija (IKT)

Abstract

Multimedia integrates different media into a system which affects different human senses at the same time. By using multimedia the quality of teaching can be increased, but the limitations of the teacher and the students as well as the limitations of the technology must be taken into consideration. Changes of teaching methods are necessary and the use of multimedia is one of the possibilities. The use of multimedia in class does not increase the quality of teaching by itself. If it is used in an inappropriate way it can even deteriorate the quality of teaching, and some researches show that it can even prolong the learning time. In this article the practical experiences with using multimedia at secondary school are described and the results of the research how students feel about using multimedia at school are presented with conclusions and proposed solutions for using multimedia at school effectively.

Keywords: multimedia, learning, efficient education, information and communication technologies (ICT)

Elektronski karierni portfolij - koncept e-orodja, ki podpira karierni razvoj posameznika

Electronic career portfolio - The concept of e-tools that support the career development of individuals

Danilo Kozoderc
danilo@vizer.si

Povzetek

Članek predstavlja koncept elektronskega portfolija, ki je orodje za beleženje in vrednotenje znanja, veščin in kompetenc. Članek predstavi zahteve in strukturo takšnega orodja. E-portfolij mora omogočati vnos izkušenj, izobraževalnih dogodkov, izdelkov oz. prototipov ter samovrednotenje le tega. Na ta način oblikovano orodje predstavlja osebno mapo, ki opisuje karierno pot posameznika in pomeni osebno bazo znanja, ki jo lahko zelo hitro vključimo v bazo znanja podjetja.

Članek predstavlja tudi dva koncepta: koncept portfolija kot unikata in koncept obstoječega okvira, ki imetnika usmerja glede vsebine in oblike. Članek podpira ta drugi koncept, ki omogoča primerljivost izdelkov – portfolijev in z usmerjanjem vodi posameznika k odkrivanju lastnih potencialov.

Ključne besede: kariera, karierna orientacija, elektronski karierni portfolij, učenje, baza znanja, razvoj človeških virov

Abstract:

This article presents the concept of an electronic portfolio, which is a tool for recording and evaluation of knowledge, skills and competencies. The requirements and structure of such a tool are presented. E-portfolio should allow the introduction of experience, educational events and products. prototypes and selfevaluation of those elements. A tool designed in this way represents a personal folder, which describes the career path of individuals. This means the personal knowledge base that can be very quickly integrated into the knowledge base of the company.

This article presents two concepts: the concept of portfolio as a unique form and concept of the existing framework, which directs the holder about the content and the form. This article supports the second concept, which allows comparability of products - portfolios and with guidance leads the individual to identify their own potential.

Keywords: career, career orientation, electronic career portfolio, learning, knowledge base, human resource development

Uporaba programa EclipseCrossword v učnem procesu prvega vzgojno-izobraževalnega obdobja

The use of the programme EclipseCrossword in the first educational triade process

Irena Kresevič

OŠ Rudolfa Ukoviča Podgrad

irena.kresevic@guest.arnes.si

Povzetek

V prispevku želim predstaviti možnosti uporabe računalniškega programa EclipseCrossword pri pouku v prvem vzgojno-izobraževalnem obdobju. Ta program omogoča oblikovanje in reševanje lažjih pa tudi bolj zahtevnih križank z različno tematiko. Učenci lahko samostojno oblikujejo in rešujejo križanke z različnih učnih področij. Da nastane križanka, mora učenec vpisati geslo in opis gesla v za to pripravljeno tabelo. S takimi križankami lahko učenci preverjajo svoje predznanje, usvajajo nove učne vsebine ali preverjajo na novo pridobljeno znanje. Možno jih je objaviti na spletu ali natisniti v obliki učnih listov.

Ključne besede: EclipseCrossword, sestavljanje križank, reševanje križank, geslo, aktivno učenje, znanje

Abstract

My contribution to the seminar presents some possibilities of the use of the computer programme EclipseCrossword in the first educational triade process. The programme enables forming and solving easy crosswords, and even more difficult ones in various themes. The pupils can independently form and solve the crosswords of different educational fields/subjects. To prepare a crossword, a pupil has to put the password and its description in an already made graph. The pupils can use such crosswords to check their pre-knowledge or newly obtained knowledge, or to appropriate new educational themes. They can be published on the Internet or printed in a form of learning papers.

Keywords: EclipseCrossword, crossword forming, crossword solving, passwords, active learning, knowledge

Kako opogumiti starejše osebe za uporabo IKT?

How to Encourage Older People to Use ICT Skills?

Julija Lapuh Bele, Boštjan Jarc, David Rozman

B2 d.o.o., Slovenija

julija.bele@b2.eu, bostjan.jarc@b2.eu, david.rozman@b2.eu

Povzetek

Delež uporabnikov IKT in interneta je v populaciji prebivalcev Slovenije, starejših od 65 let, najnižji glede na druge starostne skupine in tudi nižji kot je povprečje v EU. Razlogi so različni, predvsem pa starejšim manjka znanja in poguma, da bi to znanje pridobili. Rezultati projekta računalniškega opismenjevanja starejših od 65 let, ki ga predstavljamo v tem članku, kažejo, da je populacijo starejših mogoče motivirati in ustrezno usposobiti za uporabo IKT in interneta. Učinkovito usposabljanje izvedemo s pomočjo kombinacije metod klasičnega in e-izobraževanja. Model je vreden posnemanja, saj omogoča nižje stroške izvedbe in premošča številne ovire, ki jih imajo starejši občani pri soočanju s tehnologijo in učenjem. V članku opisujemo teoretične podlage za izbrani model in rezultate, ki smo jih dobili na podlagi prvih dveh izvedb.

Ključne besede: e-izobraževanje, kombinirano izobraževanje, spletno učenje, vseživljenjsko učenje

Abstract

The population of Slovenians' aged over 65 is the group with the lowest portion of ICT and Internet users relative to other age groups and this statistic is also lower to the EU average. This is due to different reasons, but the main factor seems to be that the elderly lack the skills and courage to acquire this knowledge. The results of the project Computer literacy of people over 65 years, presented in this article, show that the population of older people can be motivated and properly trained to use ICT and the Internet. Effective training is carried out through a combination of classical methods and e-learning. The model is worth reproducing, as it lowers cost of implementation and bridges a number of obstacles present with the elderly in the face of technology and learning.

The article describes the theoretical basis of the selected model and the results obtained on the basis of the first two performances.

Keywords: e-learning, blended learning, web-based learning, lifelong learning

Kako pripraviti učna e-gradiva?

How to prepare e-materials for teaching?

Matija Lokar

Univerza v Ljubljani, Fakulteta za matematiko in fiziko
matija.lokar@fmf.uni-lj.si

Povzetek

Kako pripraviti učna e-gradiva?

Avtorji učnih e-gradiv, namenjenih učiteljem, vse prevečkrat pripravijo le-te v obliki monolitnih blokov, sestavljenih na način, kot so običajni tekstovni delovni zvezki. To zahteva, da jih učitelj uporabi kot celoto, na točno predpisan način. A ker je učitelj posrednik med učnim gradivom in učencem, je on tisti, ki mora narediti izbor gradiv in jih kombinirati na ustrezen način. E-gradiva naj bodo zato pripravljena na fleksibilen način. Zakaj ne izrabiti možnosti, ki jih ponujajo nove tehnologije in učitelju vsaj ponuditi možnost, da gradiva pripravi ustrezno svojim in učenčevim potrebam.

Ključne besede: e-izobraževanje, e-gradiva, tehnologija, atomarna gradiva, učitelj

Abstract

The authors of e-materials, meant for the use of teachers, all too often prepare them as monolithic blocks, constructed in the way an ordinary workbook would be. This demands that the teacher takes them as a whole, precisely in the order they were written in. As the teacher usually serves as an intermediary between the teaching materials and the student, he/she should make choices about the materials and how to combine them into new ways. E-materials should be designed in a flexible way. Why not use the possibilities that new technologies offer and at the very least give teachers the chance to adapt the materials to their own and their students' needs.

Keywords: e-learning, e-learning content, technology, atomized content, teacher

Primer gradiva za interaktivno tablo pri pouku slovenščine

Learning material for Slovene lesson for an interactive whiteboard

Tatjana Lotrič Komac, Tina Žagar Pernar

OŠ Naklo, Slovenija

tatjana.lotric@guest.arnes.si, tina.zagar1@guest.arnes.si

Povzetek

Avdio-, vizualna in multimedijaska tehnologija postaja v zadnjem času vse pomembnejša sestavina sodobnega pouka. Zavedajoč se tega dejstva, se učitelji, ki imamo možnost uporabe interaktivne table, trudimo le-to orodje uporabiti čim pogosteje. Primernost uporabe upravičujejo tudi pretekla spoznanja o možnostih različnih gradiv za različne učne stile in ravni učenja.

Za pouk slovenščine je na voljo le malo oz. skoraj nič obstoječih e-gradiv oz. gradiv za interaktivno tablo, zato smo učitelji slovenščine primorani pripravljati lastno interaktivno učno gradivo. V prispevku je tako prikazan primer učnega gradiva za uporabo z elektronsko tablo, in sicer ob obravnavi prve slovenske komedije, nastale v 17. stoletju. Vključuje že obstoječe spletne povezave ter Smartboardova učna orodja, glavnina gradiva pa predstavlja dejavnosti, ki so plod lastne inovativnosti.

Ključne besede: interaktivna tabla, slovenščina, obstoječa interaktivna učna gradiva, lastna interaktivna učna gradiva

Abstract

Undoubtedly, audio, visual and multimedia technologies are becoming increasingly embedded components of classroom activities. Therefore teachers decided to integrate use of interactive whiteboards as often as possible. During the previous learning process we became aware of different ways of preparing resources for individual learning styles and different levels of learning.

Since almost no existing software for teaching Slovene language and literature is available, teachers are forced to produce their own interactive learning resources. This article represents an example of learning resource for Slovene literature. It is focused on the first Slovenian comedy, dated in the 17th century. It implements some existing links from the Internet and from the Smartboard's lesson activity toolkit, but mostly it focuses on the activities, prepared especially for this specific lesson.

Keywords: interactive whiteboard, Slovene language, existing interactive learning resources, prepared interactive learning resources

Slovenščina na daljavo

Distance learning Slovene

Tatjana Lotrič Komac
tatjana.lotric@guest.arnes.si

Povzetek

Naraščajoči mobilnosti posameznikov in družin ter dinamičnemu razvoju novih tehnologij v zadnjih letih vse bolj skladno sledijo tudi vzgojno-izobraževalni procesi. Učitelji tako lahko svojim (časovno in krajevno ločenim) učencem iz leta v leto ponujamo kakovostnejše programe izobraževanja na daljavo. Čeprav poučevanje in učenje na daljavo še vedno ne more povsem nadomestiti klasičnega dela v razredu, pa lahko pouk slovenščine za slovenske otroke v tujini, ki so sicer vpeti v tuje izobraževalne sisteme, pomeni pristen stik z maternim jezikom in kulturo ter domovino. Kombinacija individualiziranega samostojnega dela in različnih konferenčnih sistemov tako omogoča slovenskim otrokom po svetu, da skoraj povsem vzporedno sledijo izobraževalnim in funkcionalnim ciljem, ki jih sočasno osvajajo vrstniki v domovini.

V prispevku predstavljam izkušnje in primere prakse v odprtokodnem učnem okolju Moodle pri poučevanju slovenščine na daljavo za otroke slovenskih diplomatsko-konzularnih, gospodarskih, tiskovnih idr. predstavnikov RS po svetu.

Ključne besede: slovenščina, poučevanje na daljavo, učenje na daljavo, mednarodne šole, Moodle, primeri dobre prakse

Abstract

Due to increasing mobility of individuals and families and dynamic development of new technologies also education processes are expanding. Each year teachers can offer better programs and forms of distance education for pupils who are separated in space and time. Although distance education and teaching are not able to entirely replace traditional teaching methods in the classroom, we believe that Slovene children living abroad and learning according to the local educational system are by distance learning encouraged to be in touch with their mother tongue, culture and country. The combination of individualized distance learning program and different conference systems enables the Slovene children in foreign countries to learn and achieve all the educational and functional goals, their coevals learn in Slovenia.

In this issue I focus on the experiences and examples, connected with distance teaching of Slovene language in open source e-learning software platform Moodle for children of diplomatic, consular, business, press and other representatives of Republic of Slovenia all over the world.

Keywords: Slovene language, distance teaching, distance learning, international schools, Moodle, examples of good practice

NAUK – Napredne Učne Kocke za učitelje

NAUK – Advanced Learning Blocks for Teachers

Primož Lukšič, Matija Lokar, Boris Horvat

Inštitut za matematiko, fizik in mehaniko, Jadranska 19, 1000 Ljubljana
ter

Univerza v Ljubljani, Fakulteta za matematiko in fiziko, Jadranska 19, 1000 Ljubljana

primoz.luksic@fmf.uni-lj.si, matija.lokar@fmf.uni-lj.si, boris.horvat@fmf.uni-lj.si

Povzetek

V procesu izdelave izobraževalnih e-gradiv, ki je v Sloveniji v zadnjih letih močno poraslo zaradi različnih razpisov na to tematiko, se je vedno bolj zapostavljalo učitelja in vpeljevalo dvosmerno relacijo med avtorjem gradiva ter uporabnikom – učencem. Cilj projekta NAUK je zato izgradnja naprednih učnih vsebin, ki učitelju dajejo možnost prilagajanja, saj le-ta najbolj ve kaj potrebuje pri pouku. V prispevku tako srečamo izkušnje iz preteklih let, predstavimo možne scenarije uporabe repozitorijev učnih gradiv ter na teh osnovah zgrajenega portala nauk.si. Upamo, da bo s tem ponovno vzpostavljena zveza avtor-učitelj-učenec.

Ključne besede: e-izobraževanje, učni gradniki, učitelj

Abstract

In the process of creating educational learning content, which has become increasingly popular in Slovenia due to the large number of grants on this topic, the teacher has been more and more left behind while trying to implement a two-way relation between the content author and the user – student. The goal of NAUK project is therefore the creation of advanced learning content that gives the teacher the means for adapting it since he or she knows most of all what he or she needs in the classroom. The article presents our experiences from the previous years, the possible scenarios of use of the content repositories and on that basis the creation of nauk.si portal. We hope that this will again establish the connection author-teacher-student.

Keywords: e-learning, learning blocks, teacher

Poučevanje metode Scrum v sodelovanju s podjetjem za razvoj programske opreme

Teaching the Scrum method in cooperation with a software development company

Viljan Mahnič¹, Strahil Georgiev², Tomo Jarc²

¹Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, Tržaška 25, Ljubljana
viljan.mahnic@fri.uni-lj.si

²SRC Sistemske integracije, d.o.o., Tržaška 116, Ljubljana
strahil.georgiev@src.si, tomo.jarc@src.si

Povzetek

Vse večja uporaba agilnih metodologij za razvoj programske opreme zahteva, da učenje teh metodologij postane sestavni del izobraževanja bodočih inženirjev računalništva in informatike. Po drugi strani pa je možno skozi poučevanje teh metodologij preveriti tudi posamezne agilne koncepte in poiskati natančnejše odgovore na vprašanja o njihovi učinkovitosti. Zato se kot najprimernejša oblika poučevanja pogosto uporablja delo na projektih, ki omogočajo, da študenti v praksi spoznajo značilnosti agilnega pristopa, obenem pa služijo kot študije primera za ovrednotenje posameznih agilnih konceptov. V članku opisujemo, kako smo v sklopu predmeta Tehnologija programske opreme izpeljali učenje agilne metode Scrum v sodelovanju s podjetjem za razvoj programske opreme. Učenje je potekalo ob delu na realnem projektu, za katerega je seznam zahtev posredovalo podjetje, sodelavec tega podjetja pa je ves čas sodeloval s študenti kot predstavnik naročnika. Študenti so pri svojem delu dosledno uporabljali metodo Scrum in na koncu vsake iteracije s pomočjo ankete ocenili svoje izkušnje. V članku je najprej na kratko predstavljena metoda Scrum, nato sledi opis poteka dela na projektu, na koncu pa so predstavljeni rezultati ankete.

Ključne besede: agilne metodologije, Scrum, razvoj programske opreme, izobraževanje inženirjev računalništva, sodelovanje univerze z gospodarstvom

Abstract

The increasing use of agile methods for software development creates need for these methods to become part of education of future computer and information science engineers. On the other hand, teaching these methods gives us an opportunity to verify individual agile concepts and their effectiveness. For that reason, project work is appropriate and frequently used form of teaching that enables students to get acquainted with agile methods and, at the same time, provides case studies for evaluating individual agile concepts. We describe our approach to teaching the Scrum agile method, within the software technology course, in cooperation with a software development company. Students were taught through work on a real project for which a list of requirements was submitted by the company. A co-worker of this company participated throughout the teaching period playing the role of customer's representative. During their work, students consistently used the Scrum method and at the end of each iteration they evaluated their experience by means of a questionnaire. In the article the Scrum method is presented first, then a description of work on the project is given and finally the results of the survey are described.

Keywords: agile methods, Scrum, software development, computer engineering education, university-industry co-operation

Odločitveni model za izbiro šolskih in obšolskih dejavnosti otrok

Decision model for selecting curricular and extracurricular activity

Matea Curkova

OŠ Ferdo Vesel, Šentvid pri Stični

matea.curkova@guest.arnes.si

Povzetek

Aktivnosti, s katerimi se ukvarjamo iz lastnega veselja, nam nudijo sprostitve, pomagajo razvijati različne spretnosti, širijo znanje in pomagajo pri pridobivanju novih izkušenj. Hkrati spoznavamo ljudi s podobnimi interesi in razvijamo svojo osebnost. S tem namenom šole in druge organizacije ponujajo čedalje več raznovrstnih interesnih dejavnosti, starši in otroci pa se morajo odločiti in izbranim dejavnostim nameniti svojo energijo, čas in denar. Odločitev o izbiri interesne dejavnosti ni enostavna, saj temelji na presoji številnih kriterijev. Poleg tega je odločevalcev več, in ti imajo lahko povsem diametralne interese. Pri tako kompleksnih odločitvah nam lahko pomaga strokovno pripravljene večparametrski odločitveni model. Uporaba le-tega omogoča kvalitetnejše svetovanje otrokom in staršem pri izbiri najprimernejših interesnih področij za otroka v okviru šolskih in obšolskih dejavnosti in nadaljno kvalitetno izbiro konkretne dejavnosti, ki se je bo otrok rad udeleževal ter bo vsaj delno v interesu staršev in v okviru njihovih finančnih in časovnih zmožnosti.

Ključne besede: osnovne šole, interesne dejavnosti, sistem za podporo odločanju, DEXi.

Abstract

The activities, that we endeavour out of our own pleasure offer us relaxation, help us develop skills, widen our knowledge and give us new experiences. At the same time we meet people with similar interests and develop our personality. With those benefits in mind, schools and external youth organizations offer increasing number of extracurricular activities. It is up to the parents and their children to decide in which activity or activities they will invest their energy, time and money. This decision may not be easy as it is based on simultaneous judgement based on multiple criteria. Besides there may be multiple parties of interest involved and their desires may be diametrically different. Competent multiple attribute decision making model may help with complex decisions like that thus helping parents and their children to recognize suitable area of interest and based on this, select the right activity. The activity that will bring joy to the child and be in accordance with parents financial and time constrains.

Keywords: elementary schools, supplementary activity, decision-making support system, DEXi.

Informatizacija pouka klavirja: izziv prihodnosti ali utopija?

ICT in teaching piano: A challenge for the future or just an utopia?

Lorena Mihelač

lorena.mihelac@guest.arnes.si

Povzetek

Članek poskuša odgovoriti na vprašanje, če moramo na informatizacijo pouka klavirja gledati kot na utopistično dejanje ali kot na izziv vsem udeležencem v tem procesu v bližnji prihodnosti. Z ugotavljanjem načina poučevanja klavirja v preteklosti in v sedanjosti se usmerja na nove sodobne načine poučevanja z uporabo IKT-ja. Pri tem gleda na informatizacijo pouka klavirja ne samo z vidika učitelja klavirja, ampak tudi z vidika samega učenca klavirja oz. njegovih starših. Članek izpostavlja vprašanje, če je informatizacija pouka klavirja smiselna, ker je poučevanje klavirja oz. učenje/vadenje klavirja zelo specifično področje v vzgoji in izobraževanju. Članek predstavlja (vsaj za slovenske razmere) najnovejše dosežke s področja IKT-ja v poučevanju klavirja. Članek predstavlja tudi različne možnosti uporabe IKT-ja v procesu poučevanja klavirja. Z že uveljavljenimi modeli poučevanja/učenja klavirja pojasnjuje možnosti prepletanja tradicije in inovativnih modelov poučevanja z uporabo IKT-ja. Članek pojasnjuje, kako bi uporaba IKT-ja sam proces poučevanja tako kot učenja/vadenja klavirja bistveno lahko olajšala in pospešila z vidika učiteljev kot z vidika učencev. Članek izpostavlja nazadnje odprte možnosti za uporabo IKT-ja pri diagnosticiranju in preventivi različnih obolenj, ki so sestavni del pianističnega sveta.

Ključne besede: pouk klavirja, IKT in učitelj klavirja, IKT in učenec klavirja, IKT in diagnosticiranje poškodb pri igranju klavirja.

Abstract

The article is trying to answer the main question, if we have to look on the use of ICT in teaching piano as on an utopistic act or as on a challenge for all subjects involved in the proces of teaching the piano. It exposes the teaching of piano in the past as nowadays and it focus itself on new models in teaching piano using ICT. The article is pointing out the use of ICT from the viewpoint of the teacher of piano as the piano pupil and his parents as well. The article is trying also to answer the question if it is reasonable to use ICT in teaching piano, as the proces of teaching piano is very specific among other subjects in education. It is representing (at least for the Republic of Slovenia) the newest inventions in the field of ICT and different possibilities of using ICT in teaching piano. It tries to give an idea, how it is possible to combine a traditional model with the newest ICT technologies in teaching piano. It gives also explanations of the benefits of using ICT in teaching piano and how it is possible to improve the process of learning and studying piano. The article is pointing out the use of ICT in diagnostic and preventing different stress injuries, which are very often in the pianistic world.

Keywords: piano education, ICT and the piano teacher, ICT and the piano pupil, ICT and diagnostic in piano injuries.

Uporaba e-gradiv za učenje doma in na domu

Learning and Teaching with E-materials at Home

Jožica Mlakar Broder

OŠ Matije Valjavca Preddvor, Šolska ulica 9, Preddvor
jozica.mlakar-broder@guest.arnes.si

Povzetek

Klasičen način poučevanja matematike je potrebno občasno popestriti, da povečamo motiviranost učencev za delo. V mesecu novembru smo se tudi pri pouku matematike seznanili z uporabo e-gradiv. Z njimi smo nadaljevali in jih uporabili za utrjevanje in preverjanje znanja, za pridobivanje dodatnih informacij in samostojno učenje ter kot samostojno domače delo. Učencem je bilo tako delo všeč predvsem zaradi reševanja nalog na drugačen način (manj zapisovanja, naloge v obliki odgovorov DA/NE oziroma prav/narobe ...). Učenci se pri rednem pouku včasih dolgočasijo, ko pa uporabljajo računalnik, delo lahko prilagodijo svojim sposobnostim. Takoj dobijo tudi povratno informacijo o svojem delu - učenje zato postane zanimivejše. Učenci menijo, tisti s slabim in z dobrim uspehom oziroma učenec na domu, da je tako delo bolj aktivno in zanimivo ter dober pripomoček za utrjevanje že znane snovi ali pa spoznavanje nove.

Ključne besede: e-gradiva, samostojno učenje, domače delo, učenje na domu

Abstrast

Classic method of teaching mathematics needs an occasional refreshment to increase students' learning motivation. We introduced learning with e-materials in maths lessons in November. They were used for revising the subjects, for additional information and autonomous learning as well as autonomous homework. Students liked such work, particularly the different manner of exercise solving (less writing, Yes / NO answers, T / F statements...). Students are sometimes bored during the lessons, but using computers they can adapt their work to their abilities. Feedback information is given to them instantly – learning becomes more interesting. In all students' opinion that kind of learning is more active and interesting and at the same time an appropriate tool for revising the subjects or introducing the subjects.

Keywords: e-materials, autonomous learning, homework, teaching at home

E-kemija v 8. razredu – izdelava e-gradiva

E - Chemistry in 8th grade primary school – producing E - materials

Tomaž Pavlakovič, Sonja Malnarič

OŠ Belokranjskega odreda Semič, Šolska ulica 1, 8333 Semič

Povzetek

V prispevku bomo predstavili tehnično stran izdelave e-gradiva, ki zajema pripravo slikovnega gradiva (fotografije, sheme, ilustracije), video gradiva, animacij, izdelavo interaktivni elementov in postavitev gradiva na ustrezni strežnik. Predstavili bomo tudi programsko in strojno opremo, ki smo jo uporabljali pri izdelavi elektronskega gradiva za 70 ur kemije v 8. razredu. Poleg tega e-gradivo vsebuje še vsebine iz naravoslovja za 7. razred in dodatne vsebine iz kemije, ki presegajo učni načrt.

Ključne besede: kemija, e-gradiva

Abstract

In the following we will present the technical way of making E – material which contains making photo and picture materials: (photos, schemes, and illustrations), video materials, animations and making interactive elements. It also contains setting this E- material on the internet server. We will also present programme and computer equipment which we used at making E – materials for 70 lessons of Chemistry in 8th grade primary school. The E materials also contain contents from Science for 7th grade and extra contents which are not in national lesson plan for Chemistry.

Keywords: Chemistry, E- materials

Merilec srčnega utripa kot sredstvo IKT pri športni vzgoji

Heart rate monitor as a part of ICT at physical education

Rok Pekoļj

OŠ Koroška Bela

rok.pekoļj@telesat.si

Povzetek

Učenci so s pomočjo IKT spoznali, kako se odziva srce in posledično celoten organizem na fizično aktivnost. V proces dela smo zaradi lažjega razumevanja in povezovanja posameznih vedenj vključili medpredmetno povezovanje z biologijo (delovanje srca, ožilja, mišic in fizioloških procesov), matematiko (vnašanje parametrov v tabelo in koordinatni sistem, izris diagrama na podlagi podatkov) in računalništvom (delo s programom za urejanje preglednic in risanje diagramov - Microsoft Excel). Sodelujoči so s primerjanjem natančnosti samostojnega merjenja v primerjavi z merilcem srčnega utripa dokazali smiselnost uporabe merilcev srčnega utripa za ponazarjanje intenzivnosti dela med osnovnošolsko populacijo.

Ključne besede: Srce, obremenitev, frekvenca srčnih utripov, merjenje, merilec srčnega utripa.

Abstract

With the help of ICT the pupils are capable to understand how the human heart is responding during physical activity. Because of easier understanding of the problem we incorporated an interdisciplinary communication with subjects Biology (cardiovascular work, muscle work and physiologic processes) and Mathematics (inserting parameters in the table and system of coordinates, drawing the diagram) and Computer Science (using the Microsoft Excel program for the table and diagram analysing). The pupils compared the manual to the electronic measuring the heart rate and proved usenes of the heart rate monitors for the explanation intensity of exercising among the school population.

Keywords: Heart, effort, heart rate, measuring, heart rate monitor.

Vloga odločitvenega modela pri ugotavljanju vsebnosti težkih kovin v laseh

Decision making model for identificatin of heavy metal in hair samples

Aleksandra Debevec

Univerza v Ljubljani, Medicinska fakulteta
aleksandra.debevec@guest.arnes.si

Marjanca Pograjc Debevec

Gimnazija in ekonomska srednja šola Trbovlje
marjana.pograjc@gess.si

Povzetek

Ljudje večkrat zmotno mislimo, da živimo v okolju prijaznem prostoru. Nikoli se ne vprašamo po škodljivih snoveh, ki jih ne vidimo, a vplivajo na nas. Zlasti to velja za kadmij, živo srebro in svinec, ki se dolgotrajni izpostavljenosti v telesu akumulirajo in povzročijo nepopravljivo škodo. V članku želimo predstaviti odločitveni model, s katerim hitro in enostavno ugotovimo vsebnost škodljivih snovi v hrani, zraku, vodi in ostalih elementih, ki smo jim izpostavljeni. Oblikovan in v praksi preizkušen je prototip sistema za podporo večparametrskemu odločanju, ki temelji na uporabi lupine ekspertnega sistema DEXi, s katerim je možno oceniti kakovost življenjskega prostora.

Ključne besede: nevidne škodljive snovi, ugotavljanje vsebnosti škodljivih snovi v okolju, računalniško podprto odločanje, DEX lupina ekspertnega sistema.

Abstract

People are often wrong and think, that we life in an environment friendly place. We never ask ourselves about the dangerous substances that we can't see, but have effect on us. That is especially true for Cd, Hg and Pb, which accumulate in our bodies if we are exposed to them for a long time and that's what causes damage, which cannot be repaired. In this article we want to present decisive model which can help us to find out fast and easy if food, air, water and other elements, which we are exposed to, contain harmful substances. The prototype of the system which supports multi-attribute decision making is formed and tested in practice. It is based on the use of Shell of the Expert System DEX. That system makes it possible to estimate the quality of our living space.

Keywords: invisible harmful substances, determine the harmful substances in the environment, computer supported decision making, DEX Shell of the Expert System.

Assuring the students to work individually at home using Moodle - virtual learning environment

Zagotavljanje samostojnega dela študentov s pomočjo spletne učilnice Moodle

Zdenko Potočar

Ekonomska šola Novo mesto, Višja strokovna šola
zdenko.potocar@guest.arnes.si

Povzetek

V okviru bolonjske prenovе študijskih programov mora vsak študent opraviti določen obseg samostojnega dela doma. Pojavlja se problem, na kakšen način ugotavljati količino vložnega dela študenta in kako zagotoviti raznolikost aktivnih metod dela. V prispevku so predstavljene možne rešitve zagotavljanja samostojnega dela študentov višješolskega študija s pomočjo uvedbe spletne učilnice Moodle in ugotavljanje rezultatov dela študentov na Ekonomski šoli Novo mesto, Višji strokovni šoli. Rezultati analize anket študentov in analize statističnih podatkov dela v spletni učilnici nam kažejo, da uvedba LMS omogoča ugotavljanje študentove aktivnosti doma ob uporabi različnih aktivnih metod učenja.

Ključne besede: upravljanje učenja, LMS, LCMS, e-izobraževanje, spletna učilnica

Abstract

According to the Bologna process every student needs to do a portion of work at home individually. However, there is a problem of supervision the amount of time each student spends working at home and a problem of how to assure diversity of active working methods.

This text introduces possible solutions to enable students to work individually at home using Moodle web classroom and solutions to evaluate students' work. These solutions were tested among students of School of economics Novo mesto, Vocational College for Business Assistants and Economists Novo mesto.

Results of students' surveys and statistical data analysis lead us to a conclusion, that implementation of LMS enables us to successfully evaluate (lahko tudi monitor) students' activities at home using different active methods of learning.

Keywords: learning management, LMS, LCMS, e-learning, virtual learning environment

Inovativno učenje in poučevanje pri pouku geografije

Creative Learning and Teaching at the Geography Lessons

Andreja Prezelj
OŠ Koroška Bela
andreja.prezelj@telesat.si

Povzetek

V članku sem prikazala uporabo e-gradiv pri geografiji v šestem razredu devetletke. Izbrala sem tematiko o gibanju Zemlje, ki je za šestošolce težje razumljiva in predstavljiva. Na enem mestu sem zbrala e-gradiva o gibanju Zemlje, dodala pa sem tudi nekatera nova, ki so rezultat mojega dela z učenci. Prikazala sem uporabo programa hot potatoes, ki omogoča utrjevanje in preverjanje znanja v multimedijiski učilnici. Njegova prednost je v tem, da učenci dobijo sprotno povratno informacijo o pravilnosti odgovorov, svoje znanje pa lahko v takšni obliki utrjujejo tudi doma. Primerjala sem posamezne faze učnega procesa v multimedijiski učilnici ter pri klasični učni uri v razredu. Motivacijo sem izpustila namenoma, saj je uporaba informacijsko komunikacijske tehnologije pri pouku za učence že sama po sebi dovolj velika motivacija.

S pomočjo te primerjave sem tudi utemeljila svoje prepričanje o tem, zakaj je uporaba e-gradiv pri geografiji dobra izbira. Uporaba e-gradiv je pokazala precejšnje prednosti uporabe IKT pri pouku: višja motivacija, aktivno sodelovanje učencev, samostojno iskanje podatkov, ki vodijo do pravih rešitev, boljše predstave o procesih in pojavih ter lažje razumevanje le teh, utrjevanje znanja tudi doma, poučno preživljanje prostega časa. Delo učencev sem predstavila tudi v powerpointu in moviemaker

Ključne besede: IKT, gibanje Zemlje, internetna gradiva powerpoint, moviemaker, hot potatoes, usvajanje nove učne vsebine, preverjanje znanja

Abstract

In my article I have presented the usage of e-materials in the 6th grade of primary school. I have chosen the theme of **Earth** Revolution and Rotation because I consider it more difficult for the pupils to understand. I have chosen suitable e-materials connected to **Earth** Revolution and Rotation on the internet and I have also added my personal interactive exercises that are the result of my work with pupils. By using Hot potatoes program I have presented the possibility of revising and checking pupils' knowledge. The program itself offers the immediate feedback and pupils can also use it at home. In my article I have compared the classic Geography lesson to the lesson that has been carried out in the multimedia **classroom**. **I have checked advantages and disadvantages of the lessons as well as the methods and the** separate stages of the teaching/learning process. I have come to the conclusion that e-materials at Geography should be used as an **everyday teaching** method. There are several advantages of e-learning materials: higher motivation, active participation by pupils in the learning process, better understanding of complicated processes ... I have presented the pupils' projects (PowerPoint presentation) and I have made a workshop about the **Earth** Revolution and Rotation (moviemaker).

Keywords: rotation, revolution, e-materials, PowerPoint, moviemaker, hot potatoes, revise and check pupils' knowledge, creative teaching

Postopno celostno uvajanje e-izobraževanja v splošni gimnaziji

Gradual Holistic Introduction of E-Learning in a General Gymnasium High School

Tanja Mastnak¹, Peter Purg², Alenka Budihna³

¹Gimnazija Bežigrad, Slovenija

²IAM – Inštitut in akademija za multimedije, Leskoškova 12, 1000 Ljubljana, Slovenija

³Gimnazija Bežigrad, Slovenija

Povzetek

Prispevek predstavlja in diskutira izziv postopnega celovitega uvajanja e-izobraževanja v program splošne gimnazije na primeru Gimnazije Bežigrad. V skladu s sodobnimi pedagoškimi smernicami in glede na naraščajočo prisotnost e-izobraževanja na višješolskem ter predvsem visokošolskem nivoju je tudi v Sloveniji vedno več srednjih šol, ki v svoje programe uvajajo e-izobraževanje na ravni celotnega programa. Primer Gimnazije Bežigrad predstavlja inovativen način e-izobraževanja po metodi snežne kepe, ki predvideva celoletno kombinirano izvedbo z uporabo Moodle pri določenem predmetu šele po predhodni pilotski izvedbi posamezne učne enote, ob ustreznem spremljanju, vrednotenju in (kolektivni) refleksiji. Tako je znotraj učiteljskega kolektiva zagotovljeno spontano naraščanje interesa za nove pilotske in kasneje polne letne izvedbe v kombiniranem načinu (živa izvedba z online podporo). Prispevek predstavlja rezultate prve faze uvajanja e-izobraževanja s skupino petih nosilcev pilotskih izvedb učnih enot, ki so svoje izkušnje predstavili tudi drugim zainteresiranim učiteljem in jih tako s pomočjo pozitivnega vzora ter odprtih diskusij o problemih motivirali za aktivno in odgovorno vključevanje v proces postopnega celostnega uvajanja e-izobraževanja kot normalnega modela kurikularne izvedbe na Gimnaziji Bežigrad. Prva faza uvajanja je predstavljena z rezultati opravljenih anket ter SWOT analizo na podlagi vrednotenja mnenj vseh udeležencev pilotskih izvedb, tako dijakov kot učiteljev.

Ključne besede: e-izobraževanje, Moodle, kombinirano izobraževanje, splošna gimnazija, metoda snežne kepe, celostno

Abstract

The article presents and discusses the challenge of a gradual but holistic introduction of e-learning in the case of Gimnazija Bežigrad general gymnasium high school programme. Contemporary pedagogical guidelines and the increasing presence of e-learning in university programmes call for more and more high schools to integrate e-learning into their course runs on entire program level. The case of Gimnazija Bežigrad is an example of innovative implementation of e-education according to the snow ball method that foresees full annual-cycle blended learning use of Moodle within a certain course only after a preceding piloting of a singular teaching unit, and with appropriate coaching, evaluation and (collective) reflection. The latter assures a spontaneous increase of interest for new piloting runs and for later full annual-cycle blended course runs, within the teacher collective. The article describes the first steps taken based on a group of five piloted teaching units whose teachers have presented their experience to other interested teachers. This method of positive examples and open discussions of problems motivated other teachers from Gimnazija Bežigrad to take an active and responsible part in this gradual integration of e-learning as part of a normal teaching model. The first phase is introduced by presenting results of questionnaires, SWOT analysis and evaluations of all the participants – piloting teachers, students and other teachers involved.

Keywords: e-learning, Moodle, blended learning, general gymnasium, snowball method, holistic

Obravnavna pravljice v prvem razredu

Treating Fairy Tales in the First Class

Stanka Rakar

Osnovna šola Koroška Bela Jesenice, Cesta Talcev 2, 4270 Jesenice
stanka.zupancic@telesat.si

Povzetek

Uporaba IKT pri obravnavi pravljice v prvem razredu je nekoliko omejena v primerjavi z uporabo pri starejših učencih. Pri obravnavi pravljice smo uporabili računalnik, projektor in interaktivno tablo.

Obravnavna pravljice z uporabo IKT v primerjavi s klasično obravnavo se je izkazala za dobro: tak način dela za najmlajše otroke je zanimiv, motivira učence za spremljanje pravljice in tudi za nadaljnje naloge. Prednosti take metode so tudi v dinamičnem vzdušju, otroci so uporabili več čutov kot pri klasični obravnavi, hkrati so lahko spremljali besedilo in sliko, ki je bila velika in dobro vidna vsem.

Otroci so bili nad uporabo IKT navdušeni in si želijo še več takega pouka.

Ključne besede: pravljica, informacijsko komunikacijska tehnologija, prvi razred, klasična obravnavna

Abstract

The use of ICT in dealing with a fairy tale in the first class is somewhat limited compared to the use with older pupils. In treating a fairy tale we used the computer, a projector and an interactive board. Treating the fairy tale by means of ICT in comparison with the conventional treatment proved to be good. This kind of work is interesting for the youngest children it motivates pupils to follow the fairy tale and also for further tasks. We see the advantages of such methods in a dynamic atmosphere in the classroom. The children used more senses as in classical treatment. At the same time they were able to follow the text and the pictures, which were big and clearly visible to all. The children were enthusiastic over the use of ICT and they want more of such teaching.

Keywords: a fairy tale, ICT, the first class, classical treatment

Navigacija in zumiranje nevidnih poslovnih procesov v mehatronski informatiki

Navigation and zooming of nonvisible business processes in mechatronic informatics

Gorazd Rakovec
gorazd.rakovec@siol.net

Povzetek

Dejstvo je, da so pri modernih informacijskih sistemih eden od najšibkejših členov ročno vneseni podatki nastali po subjektivnih zaznavah oseb. Taki podatki so vneseni s tipkarskimi napakami, s precejšnjimi zakasnitvami in so dostikrat drugačni od objektivnega dejanskega stanja procesov. Drugi aktualen problem pa je, da je večina poslovnih procesov človeku nevidnih, zato o njih človek ne more dati podatkov. Obravnavani so človeku nevidni procesi, ki jih zaznavajo samo ustrezni senzori z merilniki v obliki bralnik terminalov, ki izdelane elektronske podatke posredujejo na strežnik v informacijski sistem. Od tu naprej se avtomatsko izdelani podatki lahko obdelujejo kot vsi ostali v informacijskem sistemu. Obravnavani so procesi v proizvodnih podjetjih. Za postavitev senzorjev v proizvodnji je treba poznati tudi stroje s tehnološkimi procesi in vse skupaj tvori mehatronski sistem, katerega arhitektura je podrobneje opisana. V prispevku so popisane glavne faze mehatronizacije informacijskih sistemov in glavni gradniki mehatronskih sistemov za vsak nivo posebej od navigacijskega sistema proizvodnih procesov, do nadzornega sistema strojnih procesov in v najvišji obliki tudi z nadzorom tehnoloških procesov. Opisan je didaktični sistem za poučevanje proizvodnih, strojnih in tehnoloških mehatronskih procesov.

Ključne besede: Mehatronika, informatika, tehnološki, strojni, proizvodni, sistem, proces, dejavnik, navigacija, nadzorni sistem

Abstract

Manual data input is one of the highest imperfection of information systems today. Most data has a subjective nature. Such data has many typing faults. They are inputted with much delay and the content are many times different than real state of the processes. The majority of business processes are nonvisible for the human, so the man can not give any realistic data about them. The nonvisible processes can be perceptioned only with propriate sensors with measuring systems named Data Acquisition terminals which generate electronic data and send them to the server of information system. When the automatic generated electronic data are in the data base, they can be processed like all other information system data. The processes of industrial companies are treated in this article. The sensors should be mounted on the machines in the production with technological processes and everything together is a mechatronic system. The architecture of mechatronic system is described. The main phases of mechatronisation of information systems with the main elements from navigation systems of production processes to machine process monitoring and finally technological processes monitoring are described. The didactic system for the production, machine and technological mechatronic processes lecturing

Keywords: mechatronics, informatics, technological, mechanical, manufacturing, system, process, factor, navigation, monitoring

Timsko, medpredmetno poučevanje ob podpori ikt

Ict Supported By Cooperating, Team Teaching And Connections Between Different Subjects

Irena Rakovec Žumer
ESIC Kranj, Gimnazija
irena.rakovec-zumer@guest.arnes.si

Povzetek

Projekt posodobitev gimnazij je prinesel mnoge kurikularne obogatitve, med drugim tudi sodelovalno, timsko poučevanje in medpredmetno povezovanje. S kolegico nemcistko sva timsko, medpredmetno poučevali slovensko in nemško književnost v športnem oddelku ob podpori IKT. IKT nam je lahko v veliko pomoč pri različnih oblikah poučevanja – tudi pri timskem, medpredmetnem poučevanju. Da se z uporabo IKT lahko izboljšajo rezultati izobraževanja, nam je dokazal projekt Kakovost za prihodnost vzgoje in izobraževanja.

Ekonomška šola Kranj (od 1. 1. 2009, ESIC Kranj) je druga srednja šola v Sloveniji, ki je prejela certifikat Kakovost za prihodnost vzgoje in izobraževanja (SIQ KzP-012). Mnogo dela za vzpostavitev sistema smo opravili z IKT. Kazalniki kakovosti so pokazali, da so se nam rezultati izobraževanja in zadovoljstvo udeležencev vzgojno-izobraževalnega procesa hitreje izboljšali, ko smo na področju IKT naredili vrsto izboljšav. Uvedli smo obvezno komuniciranje strokovnih delavcev preko e-pošte, e-redovalnico, vzpostavili internetni portal (<http://go.to/ekonomska-sola>), spletno učilnico (<http://239.gvs.arnes.si/ucilnica-ekonomc>). IKT uporabljamo vsakodnevno pri pouku, kajti na šoli imamo veliko športnikov, ki jim je IKT v pomoč, ko so odsotni.

Ključne besede: timsko, medpredmetno poučevanje, certifikat Kakovost za prihodnost vzgoje in izobraževanja, kazalniki kakovosti, IKT

Abstract

The Secondary school updating project has brought many innovations, among which are also cooperating, team teaching and connections between different subjects. Me and my colleague, the teacher of German, have worked together teaching the Slovene and German literature to the sports class by the help of ICT. ICT can be of great help at various teaching methods and at the team and interdisciplinary teaching methods as well. The project Quality for the future of upbringing and education has shown us that the results of education can improve a great deal also by the use of ICT.

The Secondary School of Economics Kranj (since 1st of January, 2009, ESIC Kranj) is the second secondary school in Slovenia that has won the certificate Quality for the future of upbringing and education. (SIQ KzP-012). A lot of work, necessary for setting up the system, has been carried out by the help of ICT. The indicators of quality have showed us that the results and the satisfaction of the participants with the educational process are improving more quickly as soon as we have introduced some improvements in the field of ICT. The communication via e-mail between the staff members has become obligatory; further on the e-markbook has been introduced, the internet portal (<http://go.to/ekonomska-sola>) and a web classroom (<http://239.gvs.arnes.si/ucilnica-ekonomc>) have been set up. ICT is used daily in our classes. As we have many sportspeople enrolled, the attendance of whom is very low, IC technology is of great help to them to meet the school requirements.

Keywords: cooperating, team teaching, certificate Quality for the future of upbringing and education, indicators of quality, ICT

»Novi« človek kot učenec - Pomen percepcije in zavesti za učenje, podprto z informacijsko tehnologijo

»New« Man As a Student The Importance of Perception and Consciousness for E-Learning

Vanda Rebolj
vanda.rebolj@siol.net

Povzetek

V članku so predstavljeni rezultati raziskave o percepciji med učenjem odraslih v virtualnem okolju, ki je bila opravljena z namenom raziskati nekatere vzroke za upočasnitev napredka pri uvajanju e-učenja pri nas. Percepcija in njeni učinki so pomembni na zavedni in nezavedni ravni, vendar jim ne posvečamo toliko pozornosti kot v bolj raziskanem klasičnem okolju. Raziskava je namenjena odkrivanju nekaterih rešitev za učno prakso. Hkrati pa raziskava potrjuje tezo, da ne moremo vedno zaključevati na podlagi nekdanjih resnic o učencih, saj se tudi ti spreminjajo pod vplivom tehnologije. Pomanjkanje strokovnih argumentov in dobre prakse sproža pedagoški konzervativizem, ta pa učni proces pelje v nasprotno smer, kot jo zahtevajo poslovni procesi v organizacijah, iz katerih prihajajo in v katere po šolanju odhajajo odrasli učenci.

Ključne besede: e-učenje, e-izobraževanje, virtualno učno okolje, percepcija, odrasli, pomnenje

Abstract

The article presents the results of a research on perception during the learning process of adults in a virtual environment. The aim of the research was to determine why the process of e-learning introduction in Slovenia has been slowed down. Perception and its effects upon learning are important on the conscious as well as on the subconscious level but they have not been given as much attention as in the classical learning environment. Disturbed perception which results from the lack of expertise in preparation of the e-environment is a serious obstacle for learning. The objective of the research was to find solutions for the actual teaching practice but at the same time the research emphasizes that conclusions cannot always be made on the basis of former facts about students. We have to bear in mind that the impact of technology changes the students as well. Lack of professional arguments and of good practice leads to pedagogical conservatism which can lead the school's progress, also in the area of adult education, in the opposite direction from the one required by business processes in the organizations in which the adult students come from or in which the students are employed after they finish their education.

Keywords: e-learning, e-education, virtual learning environment, perception, adults, remembrance

Uporaba informacijsko-komunikacijske tehnologije pri poučevanju tujega jezika v vzgojno izobraževalnem zavodu

Use of Information technology at the teaching of foreign language in educational institution

Ribič Marko
ribic23@yahoo.com

Povzetek

V šolah je v današnjem času informacijska in komunikacijska tehnologija že prepletena v vzgojno izobraževalni sistem. Prispevek obravnava vpletenost te tehnologije, predvsem uporabo računalnika in interneta, v poučevanje tujega jezika v vzgojno-izobraževalnem zavodu. Predstavljen je pomen uporabe računalnika in računalniških programov oziroma izobraževalnih spletnih strani za popestritev in motivacijo pri poučevanju angleščine za otroke in mladostnike s posebnimi potrebami, natančneje z učnimi in vedenjskimi težavami. V prispevku smo opredelili nekatere spletne izobraževalne vsebine, ki jih uspešno uporabljamo pri učenju in poučevanju angleškega jezika v vzgojno-izobraževalnem zavodu.

Ključne besede: informacijska tehnologija, računalnik pri pouku, učenje angleškega jezika, spletne izobraževalne vsebine

Abstract

Nowadays IT (Information technology) is strongly incorporated into educational system. The article deals with the incorporation of IT, especially the use of computer and Internet at the teaching of foreign languages at Institution of education Višnja Gora. The relevance of computer use, software and educational web sites for enrichment and motivation at English teaching is presented. The article deals with teaching of English to students with special needs; especially students with learning and behaviour problems. Certain internet educational contents which are used with success in English teaching are presented.

Keywords: IT, computer in teaching, teaching of English, internet educational contents

Možnost uporabe programa Tux Paint v drugem razredu

The Use of Tux Paint Program in Second (2nd) Grade

Darja Rijavec

darja.rijavec@gmail.com

Povzetek

Tux Paint je brezplačen didaktični program, namenjen predšolskim otrokom in učencem prvega in drugega triletja. Omogoča številne dejavnosti, s pomočjo katerih učenci osvajajo novo znanje in istočasno spoznavajo strojno in programsko opremo računalnika. Članek predstavlja natančen opis programa Tux Paint in preglednico dejavnosti, izhajajočih iz učnih načrtov za drugi razred osnovne šole, ki jih je avtorica izvedla z učenci med podaljšanim bivanjem, v času namenjenem samostojnemu učenju in ustvarjalnemu preživljanju prostega časa. Avtorica ugotavlja, da so bili učenci optimalno učno aktivni in za delo zelo motivirani. Izražali so pozitivna čustva: zadovoljstva, uspešnosti, veselja..., kar je predpogoj, da bodo tudi kasneje samoiniciativno opravljali domače delo. Opazila je, da je računalnik pomembno motiviral učence, saj so samoiniciativno ustvarjali s pomočjo tega programa tudi doma.

Ključne besede: Tux Paint, učni načrt, dejavnosti, drugi razred.

Abstract

Tux Paint is a free didactic software program designed for preschool children and pupils in the first and second three year cycle in primary school. It enables children to develop new knowledge through different activities and at the same time to get to know the machine and software equipment. This article presents a detailed Tux Paint program description with a table of activities deriving from 2nd grade primary school curriculum. The author of the article performed those activities during child care time which is intended for individual learning and the creative use of free time. The author ascertains that the pupils were very active and highly motivated for work. They expressed positive feelings, such as satisfaction, successfulness, happiness, etc. It is undoubtedly a prerequisite that enables the pupils to do their homework independently in the future. The author also noticed that the computer significantly motivated the pupils because they created with the program at home on their own initiative.

Keywords: Tux Paint, curriculum, activity, 2nd grade.

Pregled in analiza nastajanja, preizkušanja in uporabe e-gradiv pri pouku

A Review and Analysis of the Formation, Testing and Usage of E-materials in Teaching

Damjana Šajne¹, Tanja Urbančič², Iztok Arčon³

¹OŠ Dragomirja Benčiča Brkina Hrpelje

²Poslovno-tehniška fakulteta, Univerza v Novi Gorici

³Poslovno-tehniška fakulteta, Univerza v Novi Gorici

Povzetek

V članku obravnavamo nastajanje, preizkušanje in uporabo e-gradiv pri pouku. Predstavljene so nekatere smernice za pripravo e-učnih gradiv, ki smo jih zbrali s pomočjo literature oz. spletnih strani. Opisan je sistem za upravljanje L(C)MS ter standard SCORM. Sistem LCMS omogoča kreiranje in posredovanja e-učne vsebine. Z upoštevanjem standarda SCORM je e-učne vsebine mogoče hitro ter enostavno prilagajati in posodabljati. Pri orodju za izdelavo e-učnih gradiv je večji poudarek na programu eXe, ki je bil uporabljen v sklopu projekta e-um za izdelavo e-učnega gradiva za matematiko. Posebna pozornost je namenjena preizkušanju in uporabi e-gradiv pri pouku, kjer predstavljamo mnenje učitelja matematike in učencev. Poudarek je tudi na ocenjevanju kakovosti e-gradiv.

Ključne besede: e-izobraževanje, e-učna gradiva, sistemi za upravljanje, orodje exe, standard SCORM, e-um

Abstract

The following article deals with creating, testing and using e-materials in different lessons. Some guidelines for preparing e-materials have been collected from the available literature and websites. Further on, administration system L(C)MS and standard SCORM are presented. LCMS systems make creating and mediating possible. E-materials can be quickly and easily adopted and updated by using SCORM standard. When creating e-materials, the focus is on eXe programme. This tool was used in project called "e-um"(e-sense). The focus was on testing and using e-materials. On these bases we present teacher's and learners' opinion. We also pay some attention on quality estimation of e-materials.

Keywords: e-learning, e-teaching materials, administration systems, exe programme tools, SCORM standard, »e-um« (e-sense)

CourseLab - preprosto orodje za izdelavo e-gradiv

CourseLab – a simple tool for making e-learning materials

Peter Škarja, Branislav Šmitek

Univerza v Mariboru, Fakulteta za organizacijske vede, Slovenija
peter.skarja@siol.net , brane.smitek@fov.uni-mb.si

Povzetek

Z razvojem in uvajanjem novih programov v izobraževanje, so izvajalci izobraževanj primorani poučevati drugače. S povečano uporabo odprtokodnih avtorskih orodij se spodbuja izdelava in uporaba e-gradiv. Z odprtokodnim programom CourseLab smo dobili preprosto avtorsko orodje za izdelavo e-gradiv. V današnjem času z razvojem informacijsko-komunikacijske tehnologije, bi moral vsak učenec oziroma študent imeti izbiro med klasičnim načinom učenja in učenja s pomočjo e-gradiv. Učenje s pomočjo le-teh je bolj razgibano in interaktivno, zato si učenci lažje zapomnijo snov. CourseLab pa omogoča tudi tistim učiteljem, ki nimajo veliko računalniškega znanja, da izdelajo preprosta e-gradiva, ki so jim v pomoč pri pouku.

Ključne besede: e-gradiva, CourseLab, animacija, avtorska orodja, e-izobraževanje

Abstract

With the development and introducing new programmes into education, the performers of education are forced to teach differently. The increasing use of open source authoring tools stimulates the making and use of e-learning materials. Open source programme CourseLab is a simple authoring tool for making e-learning materials. Nowadays, with the development of information and communication technology, every pupil or student should have the choice whether to learn by a classical method or by means of e-learning materials. The latter is more stirring and interactive, therefore students memorize the material better. CourseLab also enables teachers who do not have a lot of computer knowledge to make simple e-learning materials which help them at their lessons.

Keywords: e-learning materials, CourseLab, animation, authoring tools, e-education

Odločitveni model za izbiro učenca za nagrado šole

The decision-making model for selecting students for school awards

Magda Slokar Čevdek
Osnovna šola Solkan, Slovenija

Povzetek

V prispevku je predstavljen Odločitveni model izbire učenca za nagrado šole. Model je oblikovan s pomočjo programa za večparametrsko odločanje DEXi.

V prispevku najprej predstavimo, kaj je nagrada šole in osvetlimo kompleksnost izbora za tako nagrado. Sledi kratka opredelitev odločanja s predstavitvijo vseh faz in opis programa Dexi. V nadaljevanju podrobneje opišemo Odločitveni model izbire učenca za nagrado šole po vseh fazah odločanja, kot so: opredelitev problema, oblikovanje in strukturiranje spiska kriterijev, določanje zaloga vrednosti, določanje funkcije koristnosti, opis variant, vrednotenje variant, njihova analiza ter analiza tipa »kaj-če«. Predstavljeni model je eden izmed možnih strokovnih pristopov, ki lahko pomaga pri oblikovanju predlogov za podelitev nagrade šole, saj je zaradi pestrosti odločitvene skupine, večjega števila variant kandidatov za nagrado šole in vrste kriterijev, proces odločanja kompleksen.

Prispevek zaključujemo z razmislekom o prednostih in pomanjkljivostih opisanega modela.

Ključne besede: nagrada šole, odločanje, sistemi za podporo odločanju, DEXi,

Abstract

The following paper presents The decision-making model for selecting students to be awarded by the school. The model is supported and modelled by the program for multi-parameter decision making called Dexi.

The presentation first illustrates what the school award is and highlights the complexity of selection for such an award. Next there is a short definition of decision-making with description of all its phases and the program Dexi. After that it describes in details the decision-making model for selecting students to be awarded by the school presenting all the decision-making phases such as: the definition of the problem, shaping and structuring the list of criteria, defining the stores of values, the function of usefulness, the description and evaluation of possible variants and their analysis and the what-if type analysis.

The presented model is one of the possible expert methods that can support the modelling of the proposals for the school award as the process of decision is complex, due to the variety of the decision-making group, a huge number of candidate variants and types of criteria.

In the conclusion the paper takes into consideration the advantages and disadvantages of the presented model.

Keywords: School award, decision-making, systems supporting decision-making, DEXi.

Kakovost znanja pridobljena z različnimi načini izvedb laboratorijskega dela

The knowledge quality acquired with different methods of laboratory work

Andreja Špernjak, Andrej Šorgo

Fakulteta za naravoslovje in matematiko UM, Koroška cesta 160, 2000 Maribor
andreja.spernjak@uni-mb.si, andrej.sorgo@uni-mb.si

Povzetek

V biologiji in naravoslovnih vedah nasploh je eksperimentalno in laboratorijsko delo eno izmed najučinkovitejših metod za pridobivanje znanja. Poskusi v šolskem laboratoriju so lahko izvedeni na realen ali virtualen način. Med realne poskuse uvrščamo računalniško podprt laboratorij in klasično izvedene poskuse, medtem ko so pri virtualnih poskusih eksperimenti izvedeni s pomočjo računalniške simulacije ali animacije. V raziskavi so učenci od petega do devetega razreda osnovne šole izvajali tri biološke poskuse (aktivnost kvasovk, pljučno dihanje in srčni utrip) s pomočjo klasične izvedbe poskusov, računalniško podprtega poskusa in računalniške simulacije.

Učenci so na starostni stopnji od 10 – 14 let sposobni izvesti vse tri načine laboratorijskega dela, vendar je kakovost pridobljenega znanja z vsakim posameznim načinom dela drugačna, čeprav razlike niso velike. Prispevek h kakovosti znanja pa ni odvisen le od načine izvedbe, temveč tudi od vsebine, kjer je nek način bil uporabljen.

Ključne besede: biologija, izobraževanje, klasičen laboratorijski pristop dela, načini laboratorijskega dela, računalniško podprt laboratorij, računalniške simulacije

Abstract

In Science teaching laboratory work is recognized as one of the cornerstones. In school science laboratory work computers can be used as computer supported laboratory (real) and as virtual laboratory. In the first case “real” laboratories involve bench top experiments utilizing data acquisition systems while “virtual” laboratory entails interactive simulations and animations. Lower secondary school students in age between 10 and 14 performed three laboratory exercises (Activity of yeast, Gas exchange in breathing, Heart rate) as classic, computer supported and virtual laboratory. Pupils at the different age stage get different quality of knowledge.

Keywords: biology, classical method of laboratory work, computerised laboratory, computer’s simulations, education, methods of laboratory work.

Elektronska baza podatkov o učencih s posebnimi potrebami v osnovni šoli

Electronic database about children with special needs at Primary school.

Amalija Stiplovšek

Osnovna šola Koseze, Ljubljana, Slovenija
amalija.stiplovsek@guest.arnes.si

Povzetek

Ob vključitvi otrok s posebnimi potrebami v osnovno šolo, je nastalo veliko dodatnega dela za učitelje. Za te učence je potrebno vsako šolsko leto pripraviti individualiziran program in vsak učitelj, ki poučuje takega učenca, mora sodelovati. Učitelj mora pri svojem predmetu pripraviti cilje, ki naj bi jih učenec s posebnimi potrebami usvojil. Učitelji se vsako leto menjajo, zato bi bilo smiselno pripraviti bazo podatkov o učencih s posebnimi potrebami. Baza podatkov bi morala vsebovati ključne podatke o učencu, katere cilje je pri predmetih dosegel preteklo šolsko leto, katere strategije za poučevanje so uspešne, katera so njegova šibka področja, kje potrebuje pomoč. Ti podatki bi se zbirali v času osnovnošolskega izobraževanja učenca in nato z zaključno evalvacijo poslani na srednjo šolo, kjer bo učenec nadaljeval šolanje. Taka baza podatkov bi bila na spletni strani Ministrstva za šolstvo in šport. Osnovne šole bi imele dostop do nje z geslom samo za svoje učence. Podatki o učencih bi bili dostopni v vsakem trenutku in učiteljem ne bi bilo potrebno odkrivati strategij poučevanja, ki so bile že odkrite. Tako bi učenci s posebnimi potrebami imeli več možnosti, da uspešno končajo šolanje.

Ključne besede: elektronska baza podatkov, učenci s posebnimi potrebami, osnovna šola

Abstract

In Slovenia, children with special needs have been included to primary school since 1998. There has been a lot of work to be done for teachers. Teachers have to prepare the programe for teaching for each child with special needs every year. Each teacher has to prepare goals for his subject that each child with special needs has to achieve. Pupils have different teachers every year. We think of making a database on the web, which includes the main information about these children. This database should contain the goals that the pupil achieved last year, which strategy of teaching is the best for her/him and where she/he needs help. Teachers write in this database every year of pupils learning at primary school. Finally evaluation can than be sent to the secondary school. This database should be on the web of Ministry of Education and Sport. This database is available to teachers every moment, of course by using a password. Information would be available to teachers in secondary schools and they would just use the strategies that were used in the primary school. In this way pupils with special needs would have better conditions for successful education.

Keywords: electronic database, children with special needs, Primary school

Poučni računalniški programi za vrtec

Educational computer programs for kindergarten

Jelena Stojmenovič
jelena_s21@yahoo.co.uk

Povzetek

Uporaba računalnika v vrtcu kot obogatitveni vzgojno-izobraževalni medij je vedno bolj pogosta. S primerno uporabo izbranih poučnih računalniških programov otroci raziskujejo in so aktivno udeleženi v procesu učenja. Prispevek predstavlja vlogo vzgojitelja pri delu z računalnikom, namen poučnih računalniških programov ter opredeli nekatere cilje, ki jih lahko uresničujemo z uporabo računalnika pri izvajanju dejavnosti v vrtcu. Predstavljenih je nekaj programov za otroke, ki se najpogosteje uporabljajo v vrtcih.

Ključne besede: računalnik v vrtcu, računalniški programi za otroke, vloga vzgojitelja.

Abstract

Use of computer in kindergarten as enrichment educationally-educational medium is more and more frequent. With suitable use of selected educational computer programs children are investigating and are involved actively in process of learning. The article presents the role of educator at work with computer, purpose of educational computer programs and defines some goals we can realize with use of computer at performing of activity in kindergarten. Presented are some programs for children that are the most often used in kindergarten.

Keywords: computer in kindergarten, computer programs for children, role of educator.

Interaktivna tabla in interaktivnost pri pouku matematike na predmetni stopnji OŠ

Interactive Whiteboards and the Use of Interactive Whiteboards During Math Lessons that Are Attended by Pupils from the 6th to 9th Grade at Primary Schools

Jožica Štrajhar
jozica.strajhar@amis.net

Povzetek

V prispevku je razmislek o uvedbi interaktivnih tabel v OŠ: kaj je to, zakaj naj bi jih uvajali v OŠ, organizacijska vprašanja, kdo in kako naj bi jih uporabljal, izobraževanje za uporabo. Avtorica predstavlja interaktivnost pri pouku matematike na predmetni stopnji OŠ kot posledica uporabe interaktivne table in njenih orodij, e-gradiv, matematičnih programov Geogebra in Graph.

Ključne besede: interaktivna tabla, e-gradiva, Geogebra, Graph.

Abstract

This contribution wants to take into consideration the introduction of interactive whiteboards into Primary schools: what are interactive whiteboards, why should we introduce them into our Primary schools, some questions concerning organizational matters, who should use them and in which way they could be used, as well clarify the question of training and instructions in order to use this whiteboards. The author introduces the use of interactive whiteboards during Math lessons that are attended by pupils from the 6th to 9th grade at Primary schools as a consequence or result of the usage of interactive whiteboard and its tools, e-materials, mathematical programmes Geogebra and Graph.

Keywords: interactive whiteboard, e-materials, Geogebra, Graph.

Aktivne oblike študija in vrstniško ocenjevanje v visokem šolstvu

Active learning methods and peer assessment in higher education

Mateja Strnad, Irena Nančovska Šerbec, Jože Rugelj

Univerza v Ljubljani, Pedagoška fakulteta, Kardeljeva ploščad 16, 1000 Ljubljana
mateja.strnad@pef.uni-lj.si, irena.nancovska@pef.uni-lj.si, joze.rugelj@pef.uni-lj.si

Povzetek

V prispevku predstavljamo aktivne in sodelovalne oblike učenja pri študentih začetnikih, ki potekajo v okolju za mešano učenje, kjer se prepletajo tradicionalne in z IKT podprte oblike dela. Pri raziskavi smo osredotočeni na rezultate učenja, ki jih opazujemo skozi prizmo Bloomove digitalne taksonomije. Za doseganje učnih ciljev na višjih kognitivnih stopnjah smo uporabili različne aktivne oblike učenja, kot so problemsko učenje, podprto z mini nastopom in ocenjevanjem v delavnici, učenje iz video vsebin z uporabo foruma vprašanj in odgovorov, sodelovalno učenje z vrstniškim ocenjevanjem z uporabo Wikija in programiranje v paru. Predstavili bomo vrstniško ocenjevanje, ki smo ga vključili v učni proces. Podali in analizirali bomo tudi rezultate anket, s katerimi smo ugotavljali odnos študentov do e-učenja na začetku in ob koncu semestra, mnenje o vrstniškem ocenjevanju ter odziv študentov na programiranje v paru.

Ključne besede: aktivne oblike učenja, sodelovalno učenje, vrstniško ocenjevanje, učenje z IKT

Abstract

Active and collaborative methods of learning for novice students in the environment for blended learning are presented in the paper. We focus on learning outcomes, which are observed in the framework of Bloom's digital taxonomy. Different methods of active learning, such as problem based learning, learning from video contents supported by forum for group communications, collaborative learning with peer assessment supported by Wiki, and pair programming, were used to achieve learning goals. Peer assessment is presented as integral part of learning activities. The results of survey about students' attitude towards e-learning, about their relation to peer assessment, and about their response to peer programming are presented and analyzed.

Keywords: active learning methods, collaborative learning, peer assessment, ICT supported learning

Uporaba e-gradiv za nove srednješolske učitelje

Usage of e-material for new high school teachers

Gašper Strniša

Tehniški šolski center Kranj, Kidričeva 55, 4000 Kranj,
e-pošta: strnisa.tsckr@gmail.com

Povzetek

Novi srednješolski učitelji so zaradi zaposlitvenega pogoja, kateri zahteva visoko izobrazbo dobro podkovani na strokovnem področju. Glede na to, da pa za določene predmete ne obstaja predpisanih učbenikov, morajo snov katero želijo predstaviti dijakom, poiskati v več knjigah, jo smiselno povezati in prilagoditi nivoju razumevanja dijakov. To jim vzame ogromno časa, zaradi česar morajo pogosto žrtvovati svoj prosti čas, ali pa kakovost predavanj. Odlično rešitev predstavljajo e-gradiva objavljena na spletnih straneh Ministrstva za šolstvo in šport, katero preko raznih projektov skrbi za njihovo izdelavo in uporabo. V tem prispevku je predstavljen tudi primer uporabe e-gradiva iz področja računalništva, katerega so tako dijaki kot učitelj odlično sprejeli.

Ključne besede: e-gradiva, novi učitelj, vejitve, strokovna gimnazija, licenca.

Abstract

The new high school teachers have to be highly educated and have expert knowledge because of the job application criteria. Due to the fact that there are no textbooks for certain subjects at school, they have to find the material in different books and adapt it in a way that the students will understand it. That takes a lot of time and the teachers have to sacrifice a lot of their free time or quality of their lectures. An excellent solution to this problem is e-material published on the web sites of the Ministry of Education and Sport which is also responsible for the making and usage of e-material. This article presents an example of usage of e-material in the computer science field, which was well adapted by both students and teachers.

Keywords: e-material, new teacher, branching, technical gymnasium, license.

Učna ura z interaktivnimi demonstracijami

A class with interactive demonstrations

Jože Štrucl

joze.strucl@guest.arnes.si

Povzetek

Na vzorčnem primeru učne ure z interaktivnimi demonstracijami bomo predstavili sodobno zasnovano e-gradivo za podporo poučevanja osnov digitalne tehnike. Opisali bomo osnovne lastnosti e-gradiva in izpostavili njegove glavne prednosti. Na primeru vzorčne ure bomo prikazali, kako vključimo e-gradivo v delu v razredu. Našteli bomo različne možne načine uporabe. E-gradivo je na voljo v obliki spletišča in v obliki za učno okolje Moodle. Opisali bomo prve izkušnje uporabnikov.

Ključne besede: e-gradivo, simulacije, zvok, interaktivnost, demonstracije.

Abstract

On an example of a class with interactive demonstrations we will present an up to date designed e-learning material that supports teaching in introductory digital logic classes. We will describe basic properties of e-learning material and expose its main benefits. On the example of a class we will show how to integrate the material into the classroom problematic. We will go through different possibilities of use. E-learning material exists in a web form and in a Moodle oriented form. We will describe the first experiences of users.

Keywords: e-learning material, simulations, sound, interactivity, demonstrations.

Poučevanje (slovenščine) na daljavo

Distance Teaching (of Slovenian)

Polona Tomac Stanojev

ESIC Kranj, Ekonomsko-trgovska šola
polona.tomac@gmail.com

Povzetek

Prihodnost poučevanja slovenščine je podobno, kot velja tudi za druge splošnoizobraževalne in strokovne predmete osnovnošolskega in srednješolskega izobraževanja, pogojena tudi z delom in izobraževanjem na daljavo, za kar je še posebej primerna spletna učilnica oz. aplikacija, ki jo omogoča, tj. npr. Moodle. Takšno sodelovanje med udeleženci izobraževanja, ki predvideva predvsem delo na daljavo, je še posebej koristno in pomembno za določene ciljne skupine izobraževanja, kot so, denimo, dijaki s posebnimi potrebami, vrhunski športniki idr., ki zaradi različnih razlogov ali osebnih okoliščin ne (z)morejo redno obiskovati pouka in tako niso v vsakodnevnem stiku niti s svojimi sošolci in niti z učitelji. Zanje torej pomeni delo na daljavo učinkovito komunikacijsko bližnjico med njimi in ostalimi udeleženci izobraževanja, s čimer jim je lahko znatno olajšano tako sporazumevanje s sošolci, vrstniki in učitelji kakor tudi samo pridobivanje znanja, saj jim delo v e-učilnici oz. Moodlu zagotavlja in omogoča oboje.

Slovenščina je sicer nekoliko specifičen učni predmet, kjer se izmenično ali povezano obravnavajo ter nadgrajujejo jezikovna in literarna poglavja ter teorija in praksa, a ker je tudi zasnovanost spletne učilnice takšna, da z dejavnostmi, ki jih ponuja, dopušča učinkovito in dejavno obravnavo ter povezovanje vsega navedenega, so ustvarjeni pogoji tako za uspešno poučevanje kakor tudi učenje.

Ključne besede: e-učilnica, Moodle, pouk na daljavo, slovenščina

Abstract

Like other core and specialist curriculum subjects in primary and secondary education, the future of teaching Slovenian lies in distance education, for which e-learning software platforms such as Moodle seem to be particularly appropriate. Such cooperation between participants of distance education seems to be very useful and significant for special target groups of students, such as students with special-needs and top athletes, who cannot attend classes regularly due to a variety of reasons and personal circumstances and are therefore not able to be in everyday contact with their classmates and teachers. For them, distance learning represents a communication short-cut between themselves and other participants in education. Thus communications between classmates, colleagues and teachers, as well as knowledge acquisition, are facilitated as the e-classroom or Moodle makes both possible.

Slovenian as a school subject has some specific characteristics as its grammar and literature segments, as well as theory and usage, are discussed intermittently or successively and built upon. The e-classroom is designed to enable the previously mentioned activities to be carried out effectively, ensuring the conditions for successful teaching and learning are equally fulfilled.

Keywords: e-classroom, Moodle, e-learning, Slovenian

IKT – most med šolo in starši

ICT - connection between school and parents

Andreja Vehar Jerman

andreja.vehar-jerman@guest.arnes.si

Povzetek

Za uspešno delo dijakov je nujno potrebna povezava med starši dijakov in šolo. V času, ko je odsotnost staršev od doma, zaradi zahtev službe, vedno daljša, je smiselno povečati sodelovanje s starši z uporabo informacijsko komunikacijske tehnologije. E-redovalnica je zaenkrat edino orodje, ki staršem omogoča, neodvisno od časa, da se seznanijo z učno problematiko svojih otrok.

V prispevku so predstavljeni razlogi za razširitev sodelovanja med starši in razrednikom z uporabo elektronske pošte. Razrednik lahko starše obvešča o učnih in/ali vzgojnih težavah otroka, na ta način jim podaja tudi splošne informacije. Na drugi strani starši lahko vnaprej opravičijo napovedano odsotnost, se pozanimajo o napovedanih dejavnostih v šoli, se seznanijo s spremembami urnika, morebitnimi nadomeščanji.

Sama IKT vsekakor ne more nadomestiti pristnega človeškega srečanja na govornih urah, lahko pa z njeno pomočjo dosežemo, da so starši vsaj minimalno obveščeni o delu in napredku svojih otrok.

Ključne besede: Sodelovanje s starši, uporaba IKT, govornih ur.

Abstract

Due to increased working activities in the present times, the absence of the parents from home becomes critical. Because of this reason the cooperation between student's parents and school teachers becomes more sensitive. To compensate a time shortage for consultations between parents and teacher, modern electronic communication technologies can be used. The use of e-gradebook, which can be accessed at any time of day, is nowadays the only way to get the information about student's situation in the school.

The paper presents the main reasons for use an e-mail as an alternative way of communication between parents and teachers. In such way a teacher can inform the parents about student's educational problems and/or can share useful general information. On the other hand the parents can announce the absence of the student or can check the current and future important activities at the school, timetable changes, etc.

Modern electronic communication technologies can not replace personal consultations between parents and teachers, but can help to obtain a certain level of quality information about student's educational progress.

Keywords: Cooperation with parents, a use of modern electronic communication technologies, personal consultations between parents and teachers.

Model ugotavljanja ustvarjalne učinkovitosti podjetij

The Model for Establishing Creativeness in Companies

Barbka Vidmar
Srednja šola Zagorje
barbi_vidmar@yahoo.com

Povzetek

Na podlagi raziskave je bil razvit model za ugotavljanje ustvarjalnosti v zasavskih podjetjih s pomočjo lupine ekspertnega sistema DEXi. Ustvarjalnost v podjetju se ugotavlja s pomočjo ključnih, pomembnih in zaželenih pogojev, ki morajo biti prisotni v podjetju. Glavni namen raziskave je bil, da smo z njeno pomočjo ugotovili ključne dejavnike, ki zavirajo ustvarjalnost v zasavskih podjetjih in predlagali ukrepe za izboljšanje stanja.

Model je bil preizkušen na vzorčnih podatkih 82 vodilnih managerjev v zasavskih podjetjih, omogoča enostavno analizo stanja, na podlagi katerih lahko predlagamo rešitve. Model je uporaben tudi za ugotavljanje ustvarjalnosti podjetij v drugih regijah.

Ključne besede: Model za ugotavljanje ustvarjalnosti, ekspertni sistem, lupina ekspertnega sistema DEXi, ključni, pomembni in zaželeni pogoji za ustvarjalnost, regija, ustvarjalnost.

Abstract

The subject of this master work is the research which developed the model for establishing creativeness in companies in Zasavje using the shell of the expert system DEXi. Creativeness is established with the help of essential, significant and desired conditions that must be present in a company. The mayor purpose of this research is to find out the crucial factors which obstruct creativeness in the companies of Zasavje region to suggest the measures for improving the situation.

That model was tested on sample data of 82 leading managers in companies in Zasavje. It enables simple analyses of situation and on the basis of these analyses we are able to suggest the solutions.

The model is also applicable for establishing creativeness in companies in other regions.

Keywords: Model for establishing creativeness expert system, shell of the expert system DEXi, essential, significant and desired conditions for creativeness, region, creativeness.

IKT v izobraževanju za trajnostni razvoj

ICT in education for sustainable development

Srečo Zakrajšek

IAM (Inštitut in akademija za multimedije, Ljubljana)

sreco.zakrajsek@iam.si

Povzetek

Ekonomska kriza, ki je povezana s preobremenjevanjem in onesnaževanjem okolja na eni strani, in na drugi strani z globalizacijskimi procesi, ki so v veliki meri tudi posledica razvoja sodobnih tehnologij in medijev, postavlja velike zahteve pred družbo in še posebej pred izobraževalne sisteme. Sodobne tehnologije omogočajo nove didaktično-metodične pristope in rešitve, zahtevajo pa tudi bistveno večja vlaganja v opremo in izobraževanje učiteljev.

V prispevku je prikazan potek uveljavljanja izobraževanja za trajnostni razvoj v EU in možnosti, ki jih pri tem ponujajo oziroma omogočajo IKT (informacijsko komunikacijske tehnologije). Podrobneje je prikazano stanje v slovenskem izobraževalnem sistemu.

Ključne besede: IKT, informacijsko komunikacijske tehnologije, trajnostni razvoj, izobraževanje

Abstract

The economic crisis that is connected with burdening and polluting the environment; with the globalization processes which are in a great degree a consequence of development of modern technologies and media, sets high demands on society and especially on the educational system. Modern technologies allow new didactic approaches and solutions, but they also require significantly more investments in equipment and the education of teachers.

This article presents a course for the provision of education on sustainable development in the EU, and the opportunities that are provided or can be made possible by ICT (informational communication technologies). More specifically, it presents the situation in the Slovenian educational system.

Keywords: ICT, Information and Communication Technologies, sustainable development, education

Ali pridobivati znanje s pomočjo uporabe IKT ali s klasičnimi pedagoškimi metodami in oblikami dela?

How to Acquire the Knowledge? With ICT Technology or With Classic Pedagogical Methods and Forms?

Mojca Žepič

OŠ Matije Valjavca Preddvor, Šolska ulica 9, Preddvor
mojca.zepic@guest.arnes.si

Povzetek

V današnjem času je potreba po širjenju ter lažji dostopnosti potrebnih znanj vse večja.

IKT tehnologija ponuja veliko priložnost za izboljšanje procesa učenja in poučevanja, saj motivira učence. Njena uporaba v učnem procesu spreminja obliko obravnavanja učne snovi in tudi obliko učenja.

V prispevku bom prikazala analizo primerjave rezultatov pridobljenega znanja dveh različnih učnih enot s pomočjo e-grafov oz. s klasičnimi metodami in oblikami dela.

Ključne besede: IKT, znanje, motivacija

Abstract

Today's world requires a need for spreading and easy-accessing of necessary knowledge and skills which is increasing rapidly. ICT Technology offers a great chance to improve the learning and teaching process since it is motivating for students.

The use of ICT during the learning process transforms the form of the teaching themes and topics and also the form of learning.

A comparative analysis of knowledge results gained with two different learning units – with the help of e-materials and classic teaching methods and forms - will be dealt in the article.

Keywords: ICT Technology, knowledge, motivation

Zbornik 12. mednarodne multikonference
INFORMACIJSKA DRUŽBA – IS 2009

Proceedings of the 12th International Multiconference
INFORMATION SOCIETY – IS 2009

Izkopavanje znanja in podatkovna skladišča

Data Mining and Data Warehouses

Uredila / Edited by

Dunja Mladenić, Marko Grobelnik

<http://is.ijs.si>

16. oktober 2009 / October 16th, 2009
Ljubljana, Slovenia

Preface/Predgovor

Data Mining and Data Warehouses (SiKDD 2009)

Data handling technologies have significantly progressed in the 90's. The first phases mainly dealing with storing and efficiently accessing the data, resulted in the development of industry delivering tools for handling large databases, standardization of related processes, queering languages, etc. When the data storage was not a primary problem any more the need for improving the database organization resulted in the databases supporting not only transactions but also analytical views of the data. At this point data warehousing with OLAP (On-Line-Analytical-Processing) entered as a usual part of a company information system. The OLAP paradigm still requires from the user to set well defined questions which is not always easy and possible. This led to the development of Data Mining offering automatic data analysis trying to obtain some new information from the existing data and enabling the user some new insights in the data. The Slovenian KDD conference covers a broad area including Statistical Data Analysis, Data/Text and Web Mining, Semantic Web, Link Detection and Link Analysis, Data Warehouses.

Izkopavanje znanja in podatkovna skladišča

Tehnologije, ki se ukvarjajo s podatki so v zadnjem desetletju (devetdeseta leta) močno napredovale. Iz prve faze, kjer je šlo predvsem zato kako podatke shraniti in kako do njih učinkovito dostopati, se je razvila industrija za izdelavo orodij za delo s podatkovnimi bazami, prišlo je do standardizacije procesov, povpraševalnih jezikov itd. Ko shranjevanje podatkov ni bil več poseben problem se je pojavila potreba po bolj urejenih podatkovnih bazah, ki bi služile ne le transakcijskem procesiranju ampak tudi bolj analitskim pogledom v podatke – pojavilo se je skladiščenje podatkov (data warehousing), ki postaja vse bolj standarden del informacijskih sistemov v podjetjih. Paradigma OLAP (On-Line-Analytical-Processing) zahteva od uporabnika, da še vedno sam postavlja sistemu vprašanja in dobiva nanje odgovore in na vizualen način preverja in išče izstopajoče situacije. Ker seveda to vedno ni mogoče, se je pojavila potreba po avtomatski analizi teh podatkov oz. z drugimi besedami to, da tehnologija sama pove, kaj bi utegnilo biti zanimivo za človeka – to prinašajo tehnike izkopavanja znanja (data mining), ki iz podatkov, ki že nekje obstajajo, skušajo pridobiti novo znanje, ki uporabniku ponudi novo razumevanje svojih lastnih procesov.

Slovenska KDD konferenca ponuja nekaj predstavitev, ki se ukvarjajo z modernejšimi pogledi na delo s podatki – predvsem poslovno analitske poglede: pristope, orodja, probleme in rešitve.

Editors and Program Chairs /Urednika

- Marko Grobelnik
- Dunja Mladenić

ENRYCHER – SERVICE ORIENTED TEXT ENRICHMENT

Tadej Štajner, Delia Rusu, Lorand Dali, Blaž Fortuna, Dunja Mladenić, Marko Grobelnik

Department of Knowledge Technologies

Jozef Stefan Institute

Jamova 39, 1000 Ljubljana, Slovenia

Tel: +386 1 4773419; fax: +386 1 4251038

e-mail: tadej.stajner@ijs.si

ABSTRACT

This paper describes a natural language processing and information extraction framework and illustrates several use cases where the service-oriented approach has proven to be useful. We also describe an abstract extensible schema model for representing document enrichments.

1 INTRODUCTION

In our experience, many knowledge extraction scenarios generally consist of multiple steps, starting with natural language processing, which are in turn used in higher level annotations, either as entities or document-level annotations. This in turn yields a rather complex dependency scheme between separate components. Such complexity growth is a common scenario in general information systems development. Therefore, we decided to mitigate this by applying a service-oriented approach to integration of a knowledge extraction component stack. The motivation behind Enrycher[17] is to have a single web service endpoint that could perform several of these steps, which we refer to as 'enrichments', without requiring the user to bother with setting up pre-processing infrastructure himself.

The following chapters will describe the specific components, integration details and some of the use cases that motivated this experiment of integration.

2 RELATED WORK

There are various existing systems and tools that tackle either named entity extraction and resolution, identification of facts, document summarization. The OpenCalais system [15], for example, creates semantic metadata for user submitted documents. This metadata is in the form of named entities, facts and events. In the case of our system, named entities and facts represent the starting point; we identify named entities within the document, determine the subject - verb - object triplets, and refine them by applying co-reference resolution, anaphora resolution and semantic entity resolution. As opposed to OpenCalais, we continue the pipeline to extract assertions from text, which represent newly identified relationships, present in text. This process enables the construction of a semantic description of the

document in the form of a semantic directed graph where the nodes are the subject and object triplet elements, and the link between a pair of entities is determined by the verb (predicate triplet element). The initial document, its associated triplets and semantic graph are then employed to automatically generate a document summary. The resulting triplets are then in turn used to construct a semantic graph, an effective and concise representation of document content [12].

3 ARCHITECTURE

The process consists of several phases, each depending on the output of the previous one. The dependencies between components can be illustrated by the following chart (Figure 1):

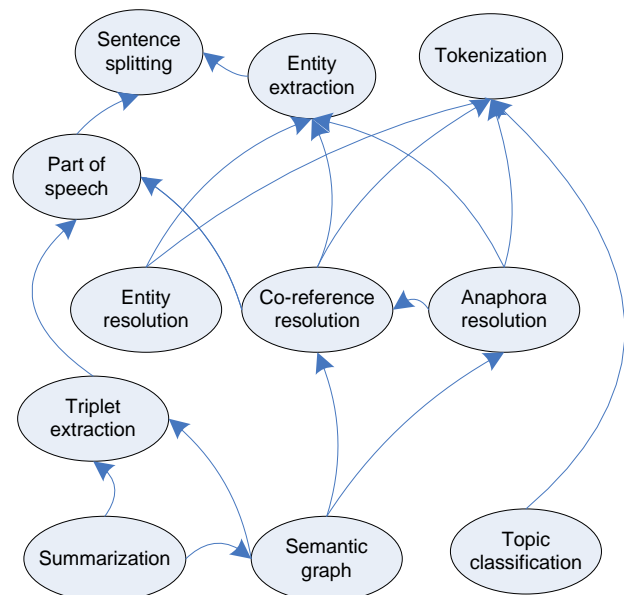


Figure 1: Component dependency graph

3.1 LANGUAGE-LEVEL PROCESSING

While language-level features are usually not explicitly stated as a requirement in most use cases, they are instrumental to most of the further enrichments that are required in those use cases:

- Sentence splitting
- Tokenization
- Part of speech tagging
- Entity extraction

3.2 ENTITY-LEVEL PROCESSING

Whereas the language-level processing step identified possible entities, the purpose of this phase is to consolidate the identified entities. This is done with anaphora resolution, where pronoun mentions are merged with literal mentions, co-reference resolution that merges similar literal mentions and entity resolution, which links the in-text entities to ontology concepts.

3.2.1 NAMED ENTITY EXTRACTION

We gather named entities in text using two distinct approaches to named entity extraction, a pattern-based one [9] and a supervised one [10].

3.2.1 ANAPHORA RESOLUTION

Anaphora resolution is performed for a subset of pronouns: {I, he, she, it, they}, and their objective, reflexive and possessive forms, as well as the relative pronoun who. A search is done throughout the document for possible candidates (named entities) to replace these pronouns. The candidates receive scores, based on a series of antecedent indicators (or preferences): givenness, lexical reiteration, referential distance, indicating verbs and collocation pattern preference [1].

3.2.2 CO-REFERENCE RESOLUTION

Co-reference resolution is achieved through heuristics that consolidate named entities, using text analysis and matching methods. We match entities where one surface form is completely included in the other, one surface form is the abbreviation of the other, or there is a combination of the two situations described [1].

3.2.3 SEMANTIC ENTITY RESOLUTION

Rather than just extracting information from text itself, the motivation behind entity resolution is to integrate text with an ontology. This consists of matching previously extracted named entities to ontology concepts. Since named entities are often ambiguous, especially in multi-domain ontologies, such as DBpedia [13], we have to employ sophisticated methods to determine the correct corresponding semantic concept of a named entity. The underlying algorithm uses ontology entity descriptions as well as the ontology relationship structure to determine which are the most likely meanings of the named entities, appearing in the input text. Because the approach is collective, it does not treat distinct entity resolution decisions as independent. This means that it can successfully exploit relational

similarity between ontology entities, which means that entities, which are more related to each other, tend to appear more often together. This is explored in further detail in [11], with concrete implementation details in [6].

3.3 ENTITY GRAPH PROCESSING

3.3.1 TRIPLET EXTRACTION

The triplet is a semantic structure composed of a subject, a verb and an object. This structure is meant to capture the meaning of a sentence. We try to extract one or more triplets from each sentence independently. Two approaches to triplet extraction have been tried, both of which take as input a sentence with tokens tagged with their part of speech.

In the first approach the sentence is parsed, and then the triplets are extracted based on the shape of the parse tree obtained. The rules of triplet extraction from a parse tree are explained in detail in [5].

In order to avoid the performance bottleneck introduced by deep parsing, we tried another approach where instead of parsing, we only do noun phrase chunking on the input sentence. The result of chunking is a sequence of tags on which pattern matching rules are applied in order to find the triplets which must be extracted. This pattern matching rules are similar to regular expressions applied on text. The difference is that as opposed to regular expressions which have as the processing unit a character, the triplet extraction rules recognise the tags as the smallest units which can be matched.

The second approach brings an important speedup to the triplet extraction process. However, due to the sequential structure of the chunked sentence, it loses some of the representational power when compared to the richer structure of a parse tree. This is why it is more difficult, if not impossible, to find some of the triplets in a chunked sentence than finding them in a parsed sentence. Another advantage of the chunked approach is that the pattern matching rules are easier to understand and extend.

3.4 DOCUMENT-LEVEL PROCESSING

While the language-level processing operates on the token and phrase domain and the entity-level processing operates on the in-text entities and concepts, the document-level processing uses the preceding enrichments to annotate the document as a whole.

3.4.1 SEMANTIC GRAPH VISUALIZATION

The semantic representation of text is achieved through linking triplet elements together, where the nodes are represented by the subject and object elements, and the relationship between them is linked with the corresponding verb. The yielded graph is a directed one, from the subject element to the object one.

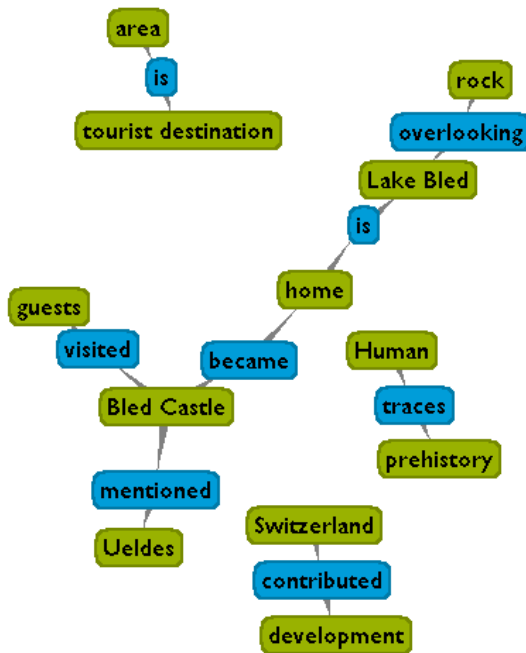


Figure 2: Example of a semantic graph visualization: Wikipedia article on Bled

Thus we can represent plain-text in a more compact manner that enables visual analysis, highlighting the most important concepts and the relations among them. An example is illustrated in Figure 2.

3.4.2 TAXONOMY CATEGORIZATION

A common use case in working with documents is classifying them in categories. This component annotates the input text with a hierarchical classifier which chooses relevant categories based on word and phrase similarity [14]. The current on-line implementation uses the Open Directory as an example of a taxonomy.

3.4.3 CONTENT SUMMARIZATION

The document's semantic graph is a starting point for automatically generating the document summary. The model for summary generation is obtained by machine learning, where the features are extracted from the semantic graph structure and content [1].

3.5 MODEL SCHEMA

The schema that is used in the inter-service communication is abstracted to the point that it is able to represent:

- **Document-wide metadata:** identifier, document-wide semantic attributes (e.g. categories, summary),
- **Text:** sentences, tokens, part of speech tags,
- **Annotations:** entities and assertion nodes, identified in the article with all identified instances,

possibly also with semantic attributes (e.g. named entities, semantic entities)

- **Assertions:** identified $\langle \text{subject}, \text{predicate}, \text{object} \rangle$ triplets, where subjects, predicates and object themselves are annotations.)

4 USAGE

The system abstracts the setup and workflow from the user by exposing only a single web service endpoint, which in turn pipelines the request through other web services. All communication is done with REST-like XML-over-HTTP requests.

5 USE CASES

5.1. VISUAL ANALYTICS

Visual analysis of documents based on the semantic representation of text in the form of a semantic graph can aid data mining tasks, such as exploratory data analysis, data description and summarization. Users can thus get an overview of the data, without the need to entirely read it. This kind of concept overview offers straightforward data visualization by listing the main facts (the triplets), linking them in a way that is meaningful for the user (the semantic graph), as well as providing a document summary. [4].

5.2 SEMANTIC INTEGRATION OF TEXT AND ONTOLOGIES

An important part of information systems integration is providing interoperability of data. This is a major issue when dealing with plain text, because it is inherently unstructured. On the other hand, one of the most pragmatic approaches is representing knowledge in a common ontology. Therefore, we designed our system to not only identify and consolidate named entities in text, but uses the semantic entity resolution component to match it with ontology concepts, which enables us to represent nodes in the graph as semantic concepts.

5.3 QUESTION ANSWERING

Document enrichment techniques such as triplet extraction and semantic graphs have been applied to build a question answering system [3]. The use case is that the answer to a natural language question is searched in a collection of documents from which triplets have been previously extracted. Triplets, possibly incomplete, are also extracted from the question, and they are matched against the triplets extracted from the documents to find the answers.

5.4 STORY LINK DETECTION

A task, related to news mining and analysis is story link detection [7], where the objective is to identify links between distinct articles that form a coherent story. [2] shows that enriching the text with entity extraction and

resolution improves story link detection performance. This indicates that such enrichment on documents may also be beneficial for other topic detection and tracking or semantic search tasks.

6 DISCUSSION AND FUTURE WORK

A use case for Enrycher in a related domain of computational linguistics is evaluating local discourse coherence of text. This is an intrinsic measure that indicates readability of text. Since it is automatic, it is also convenient for large-scale evaluation of automatically generated text. The concrete method is based on detecting rough shifts in entity mentions and short entity topics as indicators of poor coherence. As Enrycher supplies grammar roles and entities in triplets, we can match them to the sentences they have been extracted from and evaluate discourse coherence.

Another interesting research area that we are currently tackling is extracting knowledge from large-scale document collections, such as news corpora, where we are exploring possible usability and visualization improvements. Since we extract triplets and possibly resolve their nodes to semantic concepts, we can create new ontologies from corpora of text automatically. Since we are able to do semantic entity resolution, we can also perform alignment of newly extracted ontologies with other ontologies.

As of writing, we are developing additional applications that use Enrycher at their cores. One such example is a mobile RSS news reader, which leverages Enrycher to perform text summarization on news items to make them more suitable to consume on a screen space constrained mobile device.

7 CONCLUSION

We show that Enrycher offers a user-friendly way to qualitatively enhance text from unstructured documents to semi-structured graphs with additional annotations. Since the system offers a full knowledge extraction stack, it makes the system simpler to use than having the user to implement and configure several processing steps that are usually required in knowledge extraction tasks. We described various use cases in both research and applied tasks which we were able to solve with the use of Enrycher as infrastructure.

8 ACKNOWLEDGMENTS

This work was supported by the Slovenian Research Agency and the IST Programme of the EC under ACTIVE (IST-2008-215040) and PASCAL2 (IST-NoE-216886).

References

[1] D. Rusu, B. Fortuna, M. Grobelnik and D. Mladenić: Semantic Graphs Derived From Triplets With Application. In *Document Summarization. Informatica Journal*, 2009

- [2] T. Štajner, M. Grobelnik: Story Link Detection with Entity Resolution. In *Proceedings of Semantic Search Workshop at WWW2009*, Madrid, Spain, 2009
- [3] L. Dali, D. Rusu, B. Fortuna, D. Mladenić, M. Grobelnik: Question Answering Based on Semantic Graphs. In *Proceedings of Semantic Search at WWW2009*, Madrid, Spain, 2009
- [4] D. Rusu, B. Fortuna, D. Mladenić, M. Grobelnik and R. Sipoš, *Proceedings of Visual Analysis of Documents with Semantic Graphs Workshop VAKD '09 at KDD-09*
- [5] D. Rusu, L. Dali, B. Fortuna, M. Grobelnik, D. Mladenić. Triplet Extraction from Sentences. Ljubljana: 2007. In *Proceedings of the 10th International Multiconference "Information Society - IS 2007"*. Vol. A, pp. 218 - 222.
- [6] T. Štajner, From unstructured to linked data: entity extraction and disambiguation by collective similarity maximization. In *Proceedings of Identity and reference in web-based knowledge representation Workshop at IJCAI 2009*
- [7] J. Allan. Introduction to Topic Detection and Tracking. *Kluwer Academic Publishers, Massachusetts, 2002, pp. 1-16.*
- [8] J.J. Thomas and K.A. Cook. A Visual Analytics Agenda. *IEEE Comput. Graph. Appl.* 26, 1 (Jan. 2006), 10-13.
- [9] H. Cunningham, GATE, a general architecture for text engineering, *Computers and the Humanities, 2002*
- [10] Jenny Rose Finkel, Trond Grenager, and Christopher Manning. Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling. *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL 2005)*, pp. 363-370.
- [11] X. Li, P. Morie, and D. Roth, Semantic integration in text: From ambiguous names to identifiable entities," *AI Magazine. Special Issue on Semantic Integration*, vol. 26, no. 1, pp. 45-58, 2005.
- [12] I. Herman, G. Melançon, M.S. Marshal: Graph visualization and navigation in information visualization: A survey. *IEEE Transactions on Visualization and Computer Graphics*, 2000.
- [13] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives, Dbpedia: A nucleus for a web of open data, *Lecture Notes in Computer Science*, vol. 4825, p. 722, 2007.
- [14] M. Grobelnik, D. Mladenić. Simple classification into large topic ontology of Web documents. In *Proceedings of the 27th International Conference on Information Technology Interfaces, 20-24 June, Cavtat, Croatia, 2005.*
- [15] OpenCalais, <http://www.opencalais.com/>
- [16] R. Barzilay, M. Lapata. Modeling Local Coherence: An Entity-Based Approach. In *Computational Linguistics*, Vol. 34, No. 1, Pages 1-34, 2008
- [17] Enrycher, <http://enrycher.ijs.si>

Learning Event Templates on News Articles

Mitja Trampus, Dunja Mladenić

Department of Knowledge Technologies

Jožef Stefan Institute

Jamova 39, 1000 Ljubljana, Slovenia

E-mail: {mitja.trampus, dunja.mladenic}@ijs.si

ABSTRACT

We propose a pipeline for learning event templates from a large corpus of textual news articles. An event template is a machine-usable semantic data structure, in our case a graph, describing a certain event type. For instance, most earthquake news reports mention something in direction of "x people dead" or "town y shook at time z". Such templates can be used as an input for information extraction tasks or automated ontology extension. We present preliminary results of applying the proposed pipeline on a subset of News articles.

1 INTRODUCTION

Given the large amount of information encoded in written English and present on the web and elsewhere, there is a clear and long-understood need for machines to canonicalize that information as autonomously as possible in order to be able to use its inherent value.

One of the main approaches toward this end is (high-level) *information extraction*, where an algorithm is developed to fill a structured template (e.g. a database table row or a small ontology subgraph) with information extracted from unstructured text. The templates and the corresponding learning examples (tagged text), however, have to be prepared manually. In this work, we propose a step towards learning (automatically identifying) such templates prominent in a collection of news articles. Newswire is a particularly suitable domain for this task because many articles get written about each separate event, enabling us to exploit redundancy when determining the importance of pieces of information.

2 RELATED WORK

Automatic construction of templates for information extraction is already relatively well-researched (e.g. [6, 8]). However, the goal of existing approaches is to obtain *syntactic* templates for detecting words or phrases of a certain type (e.g. book titles). Our goal is to construct *semantic* templates (in the form of graphs) describing whole events; the templates do not act on the raw article text, but rather on semantic graphs describing separate events. We also aim to obtain templates that are useful in themselves, for ontology extension, not only information extraction. Furthermore, we learn the templates in a completely unsupervised manner as opposed to existing weakly supervised approaches.

Graph-based templates are also used in [7] in a context similar to ours, though the semantics are shallower. Also, the authors focus on information extraction and do not attempt to generalize the templates. Identification of templates in textual product descriptions is addressed in [10] in form of identifying product attributes and their values.

3 OVERVIEW

We propose an approach based on a pipeline for constructing abovementioned event templates in the form of small semantic graphs. Nodes represent actors or objects (nouns) and the links between them represent actions (verbs); see Figure 3 for an example of an automatically constructed template. Additionally, each node is rich with statistics about the context within separate articles it appears in, which will in future hopefully be a good starting point for training information extraction methods.

To test the proposed approach, we have used the Google News portal (although any news aggregation service would do). At this stage, we have limited ourselves to processing 7132 news articles from all topical categories, mostly published in March 2009.

4 THE PIPELINE

Each of the pipeline phases is described through an illustrative example. Consider the subset of articles reporting on various bombing attacks: in the next subsections, we will follow the information they convey and the form this information takes as it passes through the pipeline.

To avoid confusion, let us first detail some terminology: an *article* is a single web page which is assumed to report on a single *story*. A story is an event that is covered by one or more articles. Each story may fit some *event template*.

For example, the *event template* describing bombings in general may be supported by a *story* of a suicide bomber¹ in Baghdad and a *story* of NATO bombing Kabul. The story on Baghdad is in turn covered by a hundred or so web *articles* which are no longer an abstract concept but chunks of HTML code. Schematic overview of the pipeline is in Figure 1.

4.1 Data acquisition and preprocessing

We first need to obtain the data; to that end, we crawl <http://news.google.com> approximately every 40 minutes to obtain links to articles as well as a grouping of articles into

¹ We apologize in advance for such a morbid example; sadly, it is exactly topics of this kind that get terrific news coverage and are therefore both familiar to everyone and convenient to analyze.

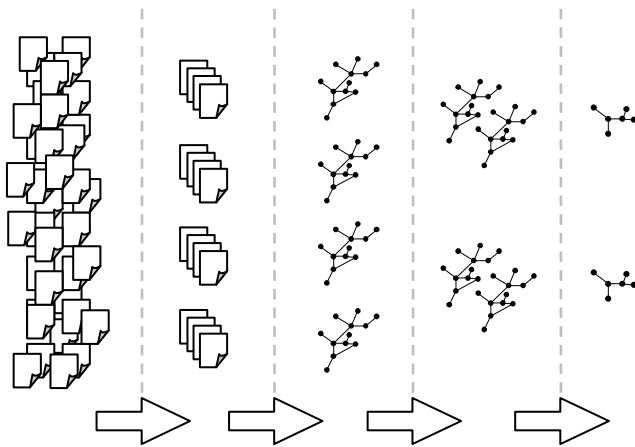


Figure 1. *The five main stages of the pipeline.* Cleaned articles (1) are grouped (2) according to the story they cover. A semantic graph is constructed for each story (3). Topically related story graphs are clustered (4); the largest subgraph common to most of the graphs in each cluster (5) is the event template.

stories. Each article is then downloaded from the publisher's website and cleaned of all HTML markup, advertisements, navigation and similar. We have developed a heuristic algorithm for **identifying the content part** of most any news article; the basic idea is to traverse the DOM tree and extract the first block-level element (TD or DIV) containing a lot of text and very little of anything else, particularly links and images. This approach successfully identifies the title and the body of an article with accuracy of about 90%.

In the end, some additional cleanup is performed like encoding, whitespace and punctuation normalization.

4.2 Data annotation

Next, we enrich the text with semantic annotations of several types as follows. Using the ANNIE tool from the GATE [1] framework, we **detect named entities** and tag them as person, location or organization. Following that, we use the Stanford parser [2] to **extract triplets** (subject-predicate-object); the authors report the precision and recall of this stage to be about 85%. As a last step, we use the web service by Rusu [3] to perform **coreference and pronoun resolutions** ("Mr. Obama", "President Barack Obama" and "he" might all refer to the same entity within an article).

4.3 Story graph construction

Starting from a group of annotated articles on a single story, we want to construct a semantic graph relaying the gist of that story. This is similar to the classic problem of multi-document summarization; however, we have stronger assumptions about inter-document coherence (assumed to be high as all documents report on the same story) and we want to present the summary in the form of a semantic graph.

First we have to **identify the stories**, i.e. clusters of articles with high topical and temporal similarity. As already mentioned, we currently simply use existing Google's clustering results. Once a story has been identified, we once more perform coreference resolution on all of its articles simultaneously (since all mentions of e.g. Obama might have gotten mapped to "Mr. President" in one article and to "Barack Obama" in another).

We now have to **identify the important triplets**. Since each story is typically represented by at least 20 articles, typically 50-200, we can rely relatively heavily on statistics: the important triplets are those that appear many times throughout the articles. However, care must be exercised: in their attempt to meet the deadlines, journalists often copy-paste whole paragraphs from another source. Unfortunately, such plagiarism cannot be detected by string matching in its simplest form because short fragments of copied paragraphs often do get altered. Writers sometimes even creatively merge paragraphs from two or more sources. In any case, much of the text is repeated verbatim which would cause triplets from those passages to be rated too high. To mitigate the problem, we **compute paragraph similarities** based on character 4-gram overlap and weight paragraphs with $1/d_{sim}$ where d_{sim} is the number of paragraphs "very similar" to current one. The method, while simple, gives results with accuracy on par with what humans can do in such a loosely defined problem.

At this point, for the purposes of the algorithm, we discard the full article text and only keep the (weighted) triplets. The weight of a triplet is defined to be the sum of weights of all paragraphs it appears in, multiplied by "position score" (triplets that appear at the beginning of an article get a higher position score). Further, triplets with verbs like "report", "tell" suggest they are the result of sentences of the form "eyewitnesses told the police that ..." and therefore uninformative; their overall weight is decreased drastically.

Triplet scores are further improved by making pairs of **similar triplets** increase each other's score. Similar triplets are identified using WordNet; the actual similarity score between two triplets is a product of experimentally set factors. The factors describe the number of words in which triplets overlap, the type of overlap (exact string match or via WordNet) and the position of overlap (e.g., it turns out that matching objects are more indicative of similar triplets than matching subjects). As WordNet does not provide uniform coverage of all topics, we have to compensate for that: triplets that appear similar to an extraordinary high number of other triplets are reduced in weight as its numerous similarities are most likely due to (too) rich synsets in that portion of WordNet. We also tried adjusting the similarity score in reverse proportion with the a priori probability of overlapping words, but that seemed not to affect performance noticeably (although evaluation was only informal). We do, however, employ a list of stopwords.

Finally, the scored triplets are viewed as tiny graphs; each graph has two weighted nodes (the scored subject and object) with a directed, weighted, labeled edge connecting them (label being the verb). Nodes are consolidated wherever possible, effectively creating a single connected component from most of two-node graphs.

We refer to the result as a *story graph*; an example can be seen in Figure 2. The central node in that graph is the subject "suicide bomb", involved in several triplets including "target camp" (the top right heavily linked node), "killed people", "blow mosque". We prune the graph from several hundred to about 100 nodes; only the several most important ones are shown in the figure.

We are currently working on a method to measure the quality of constructed semantic graphs. Both constructing a

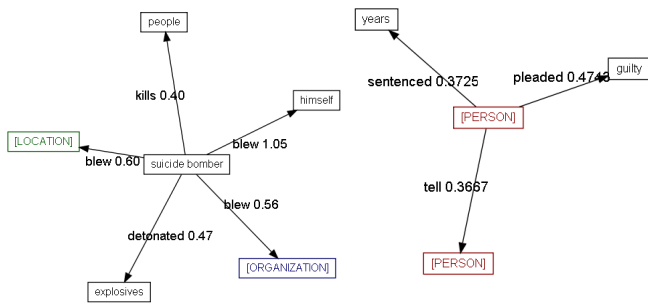


Figure 3. **The end result.** Two event templates as output by the algorithm. The left graph attempts to provide a template for stories on bombings, the left one for stories on court sentencings.

time should be invested into it. Until then, sample outputs of the algorithm will have to speak for themselves. One of them, the bombing, has already been presented in Section 4. Figure 3 shows the final output of another story cluster, this time on the topic of court sentencings. The template graphs in the figure were extracted from about 10 story graphs each.

6 DISCUSSION AND FUTURE WORK

The results, although sketchy, show promise for using the template graphs in ontology extension. Had we used some ontology other than WordNet in the last step, we would essentially get information encoded in the terms of that ontology. While mapping English words to ontology concepts is in general hard, this problem is mitigated by the high redundancy of information found in a collection of news like ours. Each hypernode of our template graph is represented by a whole set of words and therefore easier to interpret in an automated fashion.

In a similar vein, information extraction based on such templates should be feasible as well, since each hypernode is again equipped with a context and a list of words which we can think of as positive examples.

In the future, we hope to be able to verify these claims; in the short run, however, the focus will be on increasing the performance of each pipeline phase.

In the data annotation phase, the use of a faster triplet tagger is a mandatory improvement as the rate of tagging is currently about 2 articles per minute. For named entities we plan to replace ANNIE with a disambiguator proposed in [4] which uses public knowledge sources including DBpedia and GeoNames to tag entities with higher accuracy and using globally consistent IDs.

The clustering of articles into stories will probably be left in Google's domain as its performance is not problematic, although we do have an equivalent in-house solution in store. When scoring triplets at the story level, we might try to exploit the local topology of each article's semantic graph as demonstrated in [5], although statistics alone currently seem to suffice. All in all, the added structure carried by the graphs (as opposed to plain words) will have to be better exploited on all fronts. At this point our assumption that nodes and links of semantic graphs correspond directly to subject-verb-object triplets in English language may prove to be too strong. Indeed, this is not at all always true: for example, for the sentence "neighbors have reported to have seen the car crash into building", parsers would return

"neighbors reported car" or similar. The real information, "car crash building", remains hidden deep within the parse tree. With intransitive verbs, even improving the parser would not help: e.g., for "Michael Jackson died quickly", sensible graph representations like "MJ—become—dead", "death—happen—quickly" have no foundation in triplets as there are no triplets at all in the sentence. Both problems are mitigated extensively by redundancy: it is highly probable that some article will use a phrase that the pipeline can recognize, like "Michael Jackson suffered a stroke". If this proves not to be enough, there is interesting work by [9] which aims to syntactically break problematic sentences like the ones above into more parser-friendly but equivalent sentences.

We are also considering altogether dropping the phase of story clustering and trying to mine frequent subgraphs in all the stories. Computational complexity is an obvious issue here, especially because the subgraph support can be fuzzy.

Finally, the most obvious shortcoming of our work so far is the absence of efficiency measures. As the speed and accuracy of the pipeline increase, it will also become feasible to execute larger and more structured tests to properly evaluate its performance.

7 ACKNOWLEDGMENTS

This work was supported by the Slovenian Research Agency and the IST Programme of the EC under PASCAL2 (IST-NoE-216886), ACTIVE (IST-2008-215040) and VIDII (EP-08-01-014).

References

- [1] H. Cunningham, K. Humphreys, R. Gaizauskas, Y. Wilks, "GATE: a General Architecture for Text Engineering," Proceedings of the 16th conference on Computational linguistics, 1996.
- [2] D. Klein, C. D. Manning, "Accurate Unlexicalized Parsing," Proceedings of the 41st Meeting of the Association for Computational Linguistics, 2003.
- [3] D. Rusu, B. Fortuna, M. Grobelnik, D. Mladenic, "Semantic Graphs Derived from Triplets with Application in Document Summarization," Proceedings of the 11th International Multiconference "Information Society - IS 2008", SiKDD 2008.
- [4] T. Štajner, M. Grobelnik, "Story Link Detection with Entity Resolution", presented at WWW 2009 Workshop on Semantic Search, 2009.
- [5] J. Leskovec, M. Grobelnik, N. Milic-Frayling, "Learning Substructures of Document Semantic Graphs for Document Summarization," Proceedings of the 7th International Multi-Conference Information Society, 2004.
- [6] A. Arasu, H. Garcia-Molina, "Extracting structured data from Web pages," Proceedings of the 2003 ACM SIGMOD conference on Management of data, 2003.
- [7] H. Tanev, B. Magnini, "Weakly Supervised Approaches for Ontology Population," Proc. of EACL-2006, 2006.
- [8] S. Brin, "Extracting Patterns and Relations from the World Wide Web," in Lecture Notes in Computer Science, 1998.
- [9] A. Hickl, "Weakly Supervised Approaches for Ontology Population," Proc. of 22nd Conference on Computational Linguistics Coling-2008, pp. 337-344, 2008.
- [10] R. Ghani, K. Probst, Y. Liu, M. Krema, A. Fano, Text Mining to Extract Product Attributes, SIGKDD Explorations 2006.

USING ENUMERATIONS FOR WORD CLUSTERING

Lorand Dali, Nada Lavrač

Department of Knowledge Technologies

Jozef Stefan Institute

Jamova 39, 1000 Ljubljana, Slovenia

Tel: +386 1 477 3144;

e-mail: lorand.dali@ijs.si, nada.lavrac@ijs.si

ABSTRACT

This paper presents an application of exploratory data analysis consisting in a method of finding related nouns and clustering them using the information inherent in enumerations occurring in text. The results obtained are presented in a demo, and access to the functionality is published as RESTful web services.

1 INTRODUCTION

The basic idea on which the rest of this paper relies is that people tend to use related nouns in an enumeration. For instance such enumerations would be: zebra, giraffe, elephant, lion; or train, bus, car, plane; or Germany, France, Spain, Italy. One will rarely (if ever) see an enumeration of totally unrelated nouns like: zebra, helicopter, February, Germany.

The aim of the work presented in the paper is to make use of the information inherent in enumerations. Possible uses could be:

- Given a noun find the nouns which are most related to it
- Determine the extent to which two nouns are related
- Grouping of similar nouns
- Ordering of nouns from general to specific
- Document classification

This paper addresses the problem of finding related nouns and clustering nouns in groups based on relatedness.

The following sections will describe the data, the algorithms used on it, and the results obtained.

2 THE DATA

The data consists of enumerations of nouns as they occur in the Reuters RCV1 news article corpus, which contains over 12 million sentences. 500 000 enumerations of at least 3 nouns were extracted, and in these enumerations occur about 220 000 distinct nouns. The decision to take into account only enumerations with length at least 3 came from the wish to keep the data to a manageable size. And also I think that the longer an enumeration is, the more related the terms in it are. More data about the data is

shown in Table 1. In the left half of the table is shown how many sentences contain a certain number of enumerations (of at least 3 nouns).

Enumerations per sentence		Enumeration lengths	
0	12 119 657	2	3 270 921
1	460 587	3	380 012
2	16 550	4	77 070
3	1 098	5	22 530
4	749	6	8 482
5	20	7	4 612
6	12	8	3 004
7	3	9	1 075
8	3	10	938

Table 1 Statistics about the data

It can be seen that most (almost all) of the sentences do not contain any suitable enumerations at all. The right half of the table shows how many enumerations are of a certain length. It becomes clear that the vast majority of the enumerations are of length 2, of which none were taken into account for the experiments done. This means that I have used only about 10% of the available data.

2.1 Data Preprocessing

The question which is answered here is how to obtain, from a passage of text which contains an enumeration, a list of nouns appearing in the enumeration. In order to achieve this two preprocessing steps were done: part of speech tagging, and noun phrase chunking. For example the text ‘*milk and toast and honey*’ contains an enumeration which we would like to extract. The part of speech tagger will transform this text into `<noun><conj><noun><conj><noun>`, and the nounphrase chunker to `<NP><conj><NP><conj><NP>`. Now it is easy to recognise sequences of noun phrases separated by conjunctions or commas to obtain the list of nouns i.e. [milk, toast, honey]

Other preprocessing steps are converting all words to lower case, and trying to bring nouns to their base form using WordNet[1].

For the part of speech tagging CRFTagger[2], and for the nounphrase chunking CRFChunker[3] have been used.

2.2 Graph Representation

The enumeration data is stored as an undirected graph where the nouns are the nodes. Two nodes are connected with an edge if they were enumerated together at least once. The weight of the edge is how many times the two nouns have been enumerated together. Figure 1 shows an example of graph which was obtained from the following enumerations:

```
[brandy, whiskey, rum, gin]
[chips, coke]
[whiskey, pizza, beer, chips]
[cheese, beer, pizza]
[chips, beer, coke, popcorn]
```

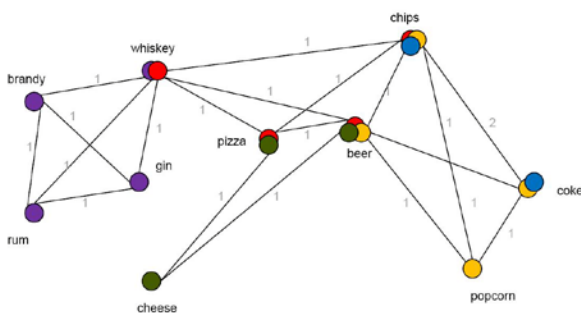


Figure 1 Graph representation

3 ALGORITHMS USED

3.1 Random Walk

Random walk is a process in which at each step we are at a node in a graph, and we randomly choose another node from the graph to be in at the next step. The node chosen for the next step has to be a neighbor of the node we are in at the current step, and the probability of it being chosen is determined by the weight of the edge which connects it. At each step there is also a probability of stopping i.e. ending the process of random walk.

Using random walk we can compute for each node the probability of ending in any other node if starting from there. This probability tells us how related, or how associated, a word is to another. If from a certain word there is a high probability of reaching another given word after a random walk, then this means that the first word is highly related to the second one. It has to be noted that in general it is not true that if a word is strongly related to another then the other one will also be strongly related to the first one. So for each word pair the relation is described by two numbers, one for each direction.

For the graph described in this paper random walk has been computed using Monte Carlo sampling. From each

node a large number of walks (2000 in this case) have been started, and the probabilities were estimated using the relative frequencies of the obtained destinations.

3.2 Clustering

Clustering is a way of grouping similar entities together. The similarity between two entities is expressed by a distance. The smaller the distance, the more similar the two entities are. In clustering each entity belongs to exactly one cluster, and the goal is that members of the same cluster have a small distance to each other while the clusters are far from each other.

For being able to compute distances, the entities have to be given as data points (vectors), or alternatively a distance matrix, showing the distance from each entity to each entity, can be directly given.

In the clustering application presented here the entities to be clustered are nouns (the nodes in the graph). The distance between two nouns is computed as the average between the two associations between the words. The goal is not to cluster the entire set of words, but rather a given subset. The clustering method of choice was the hierarchical agglomerative clustering which has as output a tree with the leaves the elements which must be clustered. It works like such that at the beginning each entity is a cluster of its own and at each step the two clusters which are nearest to each other are merged, until only one cluster remains. The inter-cluster distance has been computed as the average of distances between the distances of all pairs of elements where one element is from one cluster and one is from the other.

For example if we cluster the words *apple*, *train*, *doctor*, *helicopter*, *lawyer* and *orange* we obtain the tree shown in Figure 2

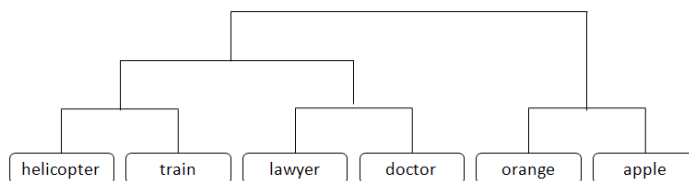


Figure 2 Hierarchical clustering

It can be seen that fruit (apple and orange) professions (lawyer and doctor) and vehicles (helicopter and train) are clustered together. Moreover from the resulting three clusters the one with vehicles and professions are closest to each other.

The clustering has been done using the C Clustering Library[4] through a Python interface called Pycluster[5].

4 DEMO

The results of the work are presented in a demo in which the user can do three things:

He can explore for any given word what words it has been enumerated with, and to what words it is mostly related.
He can check the relations between any pair of words.
He can give a set of words to be clustered hierarchically like shown in Figure 2.
The functionality used in the demo is also published as RESTful web services.

5 CONCLUSIONS AND FUTURE WORK

Time has come to draw the conclusions. A method of finding words related to a given word and to group related words together has been tried. Although no evaluation has been done we can intuitively conclude that the method works quite well, and that indeed there is much information in enumerations.

A point which is worth discussing is the choice of algorithm used to compute relatedness. The random walk probabilities estimation through random sampling needs very little memory, no matter how big the data. The runtime is long if a large number of samples is taken, but it turns out that just 2000 samples give quite satisfactory results. Moreover the algorithm presented is very suitable for parallel computing as it can be easily divided into independent jobs (one for each node) which can be executed in any order. Also the method is scalable as it allows easy addition of new nodes after all probabilities have been estimated.

The other alternatives would be to compute the random walk probabilities with matrix multiplications or with singular value decomposition. With both of these methods the difficulties caused by the limited memory must be surpassed. A parallel implementation is not as straightforward, and also online addition of new nodes is more tricky in these cases.

As future work document classification based on enumeration data could be tried. I expect that such a classifier would be more accurate, or at least would need less training data.

Also, using the asymmetry of the relatedness measure computed by the random walk, an ordering of words from general to particular could be tried, as specific words tend to be more strongly associated with general words than the other way around.

ACKNOWLEDGEMENTS

This work was supported by the Slovenian Research Agency and the IST Programme of the EC under NeOn (IST-4-027595-IP) and ACTIVE (IST-2008-215040).

References

- [1] WordNet: An Electronic Lexical Database
- [2] Xuan-Hieu Phan, "CRFTagger: CRF English POS Tagger", <http://crftagger.sourceforge.net/>, 2006.
- [3] Xuan-Hieu Phan, "CRFChunker: CRF English Phrase Chunker", <http://crfchunker.sourceforge.net/>, 2006.
- [4] The C Clustering Library, The University of Tokyo, Institute of Medical Science, Human Genome Center
- [5] Pycluster, <http://bonsai.ims.u-tokyo.ac.jp/~mdehoon/software/cluster/software.htm>.

SEMI-AUTOMATIC ONTOLOGY EXTENSION USING TEXT MINING

Inna Novalija, Dunja Mladenić
Department of Knowledge Technologies
Jozef Stefan Institute
Jamova 39, 1000 Ljubljana, Slovenia
Tel: +386 1 4773144
e-mail: inna.koval@ijs.si, dunja.mladenic@ijs.si

ABSTRACT

This paper addresses the process of the ontology extension for a selected domain of interest which is defined by keywords and possibly a glossary of relevant terms. A new methodology for semi-automatic ontology extension, aggregating the elements of text mining and user-dialog approaches for ontology extension, is proposed and evaluated. We conduct a set of ranking, tagging and illustrative question answering experiments using Cyc ontology and business news collection. The experiments show that the precision of business news tagging increases from 61% to 87% and the corresponded recall increases from 46% to 81% after the ontology extension with concepts extracted from business news.

1 INTRODUCTION

This paper explores the process of the ontology extension motivated by usage of the extended ontology for business news analysis. The main contribution of this paper is in proposing a new methodology for semi-automatic ontology extension based on text mining and user-dialogue approaches. Our research also contributes to the analysis of business news by the means of semantic technologies. The new methodology for the semi-automatic ontology extension, aggregating the elements of text mining and user-dialog approaches for ontology extension, is suggested and used for inserting the new financial knowledge into Cyc [10], which maintains one of the most extensive common-sense knowledge bases worldwide.

The experiments on ranking, business news annotation and simple question answering show that the extended financial ontology allows for a better financial news analysis.

The evaluation of the methodology of the ontology extension shows its ability to fasten the ontology extension process.

The paper is structured as follows: Section 2 presents the information about the existing approaches of ontology extension; the new methodology of ontology extension is discussed in Section 3, Sections 4 and 5 describe the experiments and the results, the discussion and future work are covered in Section 6.

2 EXISTING APPROACHES OF ONTOLOGY EXTENSION

The automatic and semi-automatic ontology extension processes are usually composed of several phases. Most approaches include defining the set of the relevant ontology extension sources, preprocessing the input material, ontology augmentation according to the chosen methodology and ontology evaluating and revision phases. The notable approaches of ontology extension include natural language processing based approach [2] [7], networks/graphs based approach [5] [6], user-dialogue based approach [8] and pattern based approach [1].

3 METHODOLOGY

As a part of the research, we propose a new methodology for semi-automatic ontology extension, which combines text mining methods with user-oriented approach and supports the extension of multi-domain ontologies. The proposed methodology embodies three main modules: the Domain Information Module (DIM), the Domain Subset Extraction Module (DSEM) and the Ontology Extension Module (OEM) and deals with an extendable multi-domain ontology.

The proposed methodology for semi-automatic ontology extension accounts for the following phases, displayed with numbers in Figure 1:

1. *Domain information identification.* The domain experts identify the appropriate Domain Keywords. As well, in DIM a Domain Relevant Glossary, containing concepts with descriptions is determined.
2. *Extraction of the relevant domain ontology subset from multi-domain ontology.* In DSEM the Keywords are used by the Upper-Level Domain Extractor to restrict the multi-domain Ontology to the specific domains of interest. Afterwards, the Domain Knowledge Extractor uses a Domain Relevant Glossary to obtain the Ontology Subset for the particular domains of interest.
3. *Domain relevant information preprocessing.* The information from the Domain Relevant Glossary and the extracted relevant Ontology Subset are linguistically preprocessed in OEM. The preprocessing phase

includes tokenization, stop-word removal and stemming.

4. *Composing the list of potential concepts and relationships for ontology extension.* The ranked list of the relevant concepts and possible relationships suitable for ontology extension is composed in OEM.
5. *User validation.* Furthermore, in OEM the user validates the initial results and the final list of ontology extension concepts and relationships in the relevant format is created.
6. *Ontology extension.* The Ontology extension is taking place in the Ontology Extension Module. It represents adding the new concepts and relationships between concepts into the Ontology.
7. *Ontology reuse* As a part of the new extension process, we reuse the previously extended Ontology in DSEM and in OEM.

We have adapted the methodology in order to obtain an exhaustive specific methodology for Cyc knowledge base extension. The main adaptations are based on microtheories (Mt) that Cyc is using to represent thematic subsets of the ontology. Usage of Knowledge Entry (KE) concept templates helps to incorporate the user feedback. Since our motivation is in business news annotation, we have chosen Business and Finances as the domains of primary interest. Given the fact that Cyc knowledge base contains common sense knowledge [4], we assume that Cyc KB includes some financial knowledge - a financial knowledge base (Cyc FKB).

4 EXPERIMENTS

In order to evaluate the proposed methodology we conducted a series of ranking, news tagging and illustrative question answering experiments on the data sources, described below.

For the data evaluation we have used the RSS feeds data from CNN [9], Reuters [14], Forbes [12], Financial Times (FT) [13] and Yahoo! Finance [11] websites.

The news collection used in the current research accounts for 1455 Reuters news, 4584 CNN news, 5812 Yahoo! Finance news, 15374 Forbes news and 34311 Financial Times news.

Following the first phase of the proposed methodology, domain knowledge identification should be made in the initial phase. For these purposes we have selected the Harvey [3] financial glossary which contains around 6000 hyperlinked financial terms.

In order to evaluate the suggested methodology, we have conducted ranking experiments on the subset of 500 random Yahoo! Finance news. The most frequent financial terms have been extracted and 100 random financial terms have been chosen. Cyc Financial Knowledge Base is then extended, using the proposed methodology, with concepts corresponding to the chosen financial terms. The efficiency of the automatic concept ranking is measured afterwards.

Tagging experiments show how the business news tagging improves after ontology extension with the domain relevant glossary. The tagging/annotation experiments provide testing on a random subset of 100 Yahoo! Finance news. We have identified the financial terms, occurring most frequently in the selected news, tagged the terms with Cyc Concept Tagger and checked the precision and recall of news tagging. Furthermore, we have added the simplest assertions about the missing financial terms into Cyc and again found the precision and recall of news tagging.

Illustrative question answering demonstrates the capacity of Cyc to answer simple financial questions before and after the extension of Cyc Financial Knowledge Base. Let us assume that we have a simple question and we want to get an answer using an unextended and extended Cyc Knowledge Base. The experimental results are given in Section 5.

5 RESULTS

The results of the experiments suggest that the financial ontology extension leads to better business news annotation and confirm the applicability of the suggested methodology for ontology extension to Cyc Knowledge Base augmentation.

5.1 Ranking

The proposed methodology results by the user getting a ranked list of relevant Cyc financial concepts for each glossary concept and confirms the relationships between the glossary and Cyc concepts. The comparison of the automated and manual rankings shows that in 60 out of 100 cases the correct concept can be found among the top 10 automatically suggested concepts for each glossary term. More precisely, for 23 out of 100 glossary terms the automatically suggested concept occupied the first position in manual ranking. It means that using the proposed methodology the user is able to compare Cyc and glossary concepts and establish the equivalence and parent-child relationships much faster than just using the manual search for the relevant concepts in Cyc.

5.2 Business News Tagging

We have found 231 financial terms in the random sample of 100 Yahoo! Finance news. The precision of business news tagging increases from 61% to 87% and the corresponded recall - from 46% to 81% after adding the simplest assertions about the missing terms into Cyc. This is confirming our hypothesis that the Cyc ontology has still space for extension as for financial domain with terms relevant for financial news analysis.

5.3 Illustrative Question Answering

In this section we illustrate relevance of the proposed Cyc ontology extension for question answering in the financial domain.

For the research purposes we have selected the following simple question:

“Which stock exchanges exist in Western Europe?”
which might be translated in CycL query as:

```
(#$and
  ($isa ?X $$StockExchange)
  ($residenceOfOrganization ?X
  $$WesternEurope))
```

Using an unextended Cyc KB we get no appropriate answers because of the insufficient representation of stock exchange instances and residence relationships in Cyc. After we extended Cyc KB, using the proposed methodology, with the following new assertions about Frankfurt stock exchange:

```
Constant: FrankfurtStockExchange.
In Mt: OrganizationDataMt.
isa: StockExchange.
In Mt: OrganizationDataMt.
residenceOfOrganization:
CityOfFrankfurtGermany.
```

Extension of Cyc Knowledge Base according to the proposed methodology with European stock exchange instances allows the user to provide Cyc with new stock exchange instance and get the following answer for the asked question:

```
[Explain #0] FrankfurtStockExchange
```

6 CONCLUSION

In this paper the aspects of ontology extension and business news analysis have been explored. The new methodology of ontology extension, combining text mining methods and user-based approach, has been proposed and exposed to the preliminary evaluation.

In contrast with many other methodologies for ontology extension, our methodology deals with ontologies and knowledge bases, covering more than one domain. However, it allows restricting the area of ontology extension to a specific domain. Furthermore, the user validation helps to avoid adding to the ontology irrelevant concepts and relationships.

The future work should include further extension of Cyc Knowledge Base and using it for more sophisticated news analysis. Furthermore, the proposed methodology for ontology extension should be tested on other domains. In addition, a particular attention should be given to the problem of the disambiguation of the glossary terms and terms extracted from news sources.

7 ACKNOWLEDGMENTS

This work was supported by the Slovenian Research Agency and the IST Programme of the EC under NeOn (IST-4-027595-IP) and ACTIVE (IST-2008-215040).

References

- [1] Blomqvist, E.: Pattern-based Ontology Construction. In: *KWEPSY*. 2007.
- [2] Burkhardt, F., Gulla, J. A., Liu, J., Weiss, C., Zhou, J.: Semi Automatic Ontology Engineering in Business Applications. *Workshop Applications of Semantic Technologies*, INFORMATIK. 2008.
- [3] Harvey, C.R.: Yahoo Financial Glossary, Fuqua School of Business, Duke University. 2003.
- [4] Lenat, D.: Cyc: A Large-Scale Investment in Knowledge Infrastructure. *Communic. of the ACM* 38 (11). 1995.
- [5] Liu, W., Weichselbraun, A., Scharl, A., Chang, E.: Semi-Automatic Ontology Extension Using Spreading Activation. *Journal of Universal Knowledge Management*, No. 1, pp. 50 – 58. 2005.
- [6] McDonald, J., Plate, T., Schvaneveldt, R.: Using pathfinder to extract semantic information from text. In: *Schvaneveldt*, pp. 149–164. 1990.
- [7] Sabrina T., Rosni A., Enyakong T.: Extending Ontology Tree Using NLP Technique. In: *Proceedings of National Conference on Research & Development in Computer Science REDECS 2001*. 2001.
- [8] Witbrock, M., Baxter, D., Curtis, J., Schneider, D., Kahlert, R., Miraglia, P., Wagner, P., Panton, K., Matthews, G., Vizedom, A.: An Interactive Dialogue System for Knowledge Acquisition in Cyc. In: *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence*. 2003.
- [9] CNN News, <http://www.cnn.com>
- [10] Cycorp, Inc., <http://www.cyc.com>
- [11] Yahoo! Finance, <http://finance.yahoo.com>
- [12] Forbes News, <http://www.forbes.com>
- [13] Financial Times News, <http://www.ft.com/home/europe>
- [14] Reuters News, <http://www.reuters.com>

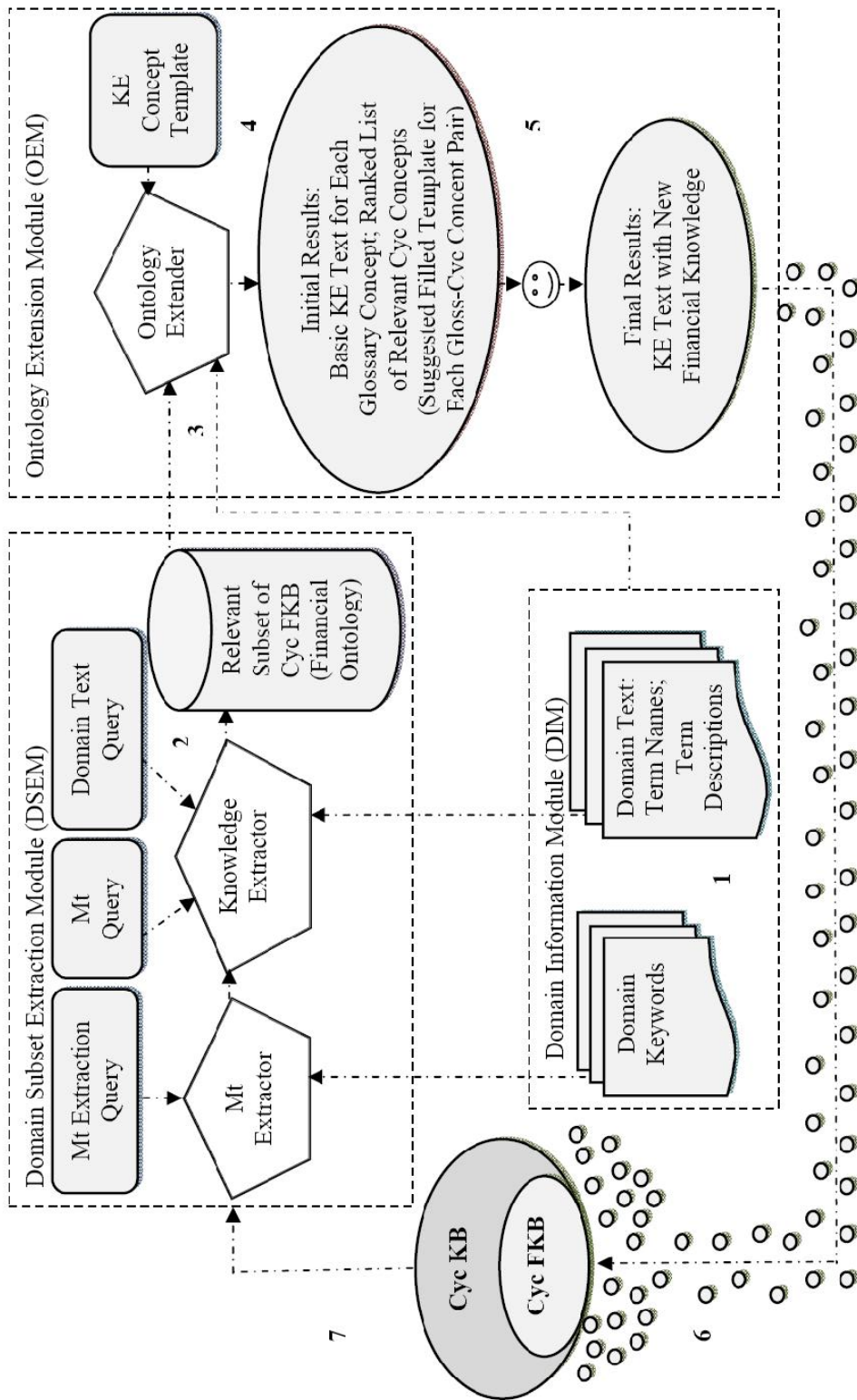


Figure 1: Methodology for Semi-Automatic Ontology Extension (Cyc Adaptation).

CONTEXTUALIZED VISUALIZATION OF ONTOLOGIES AND ONTOLOGY NETWORKS

Boštjan Pajntar, Dunja Mladenić, Marko Grobelnik

Department of Knowledge Technologies

Jozef Stefan Institute

Jamova 39, 1000 Ljubljana, Slovenia

Tel: +386 1 4773419; fax: +386 1 4251038

e-mail: {bostjan.pajntar, dunja.mladenic, marko.grobelnik}@ijs.si

ABSTRACT

Contextualized visualization of ontologies and ontology networks is addressed here via visualization of ontology and ontology mappings. The idea is to enable the user to browse through an ontology inside a context of related networked ontologies, where the context is defined through the mappings between the ontologies. The proposed visualization consists of two parts. First, a pair of ontologies is visualized in an intelligent way and second, alignment of these two ontologies is visualized as links between the related concepts. The proposed approach is implemented in cONTOext as a Flash application integrated within the Facebook plug-in architecture. The extended approach is implemented in OntoConto as plug-in to a NeOn Toolkit platform where it is seamlessly integrated with the other technologies. Namely, integration inside the existing NeOn Toolkit provides an easy access to the loaded ontologies, a possibility to edit them and, enables the user to access several different alignments between the ontologies.

1 INTRODUCTION

Networked ontologies in general can be a very complex structure. There are many ways how the knowledge stored inside them can be obtained. One possible way is to focus on a part of the networked ontologies i.e. a pair of related ontologies and in this setting try to understand the context that one ontology provides for the other.

The proposed approach is, with some variation, implemented in two demos: cONTOext plug-in to Facebook social networking platform [1] and, OntoConto plug-in to NeOn Toolkit platform for networked ontologies [2].

cONTOext emphasizes contextual visualization of ontologies to support collaborative editing of network ontologies. The main idea is to leverage efforts of multiple individuals editing similar domains (in the form of ontologies) at the same time via sharing knowledge by introducing mappings between ontologies. The obvious

practical benefit is in the situation when other users provides context through the ontologies they have created.

OntoConto emphasizes contextual visualization of ontologies by providing context for a selected ontology via showing the related parts of another ontology (here also referred to as context ontology). This is achieved by visualizing and manipulating the target ontology and context ontology side by side, with mappings connecting nodes of the two ontologies. Namely, for each concept in the chosen ontology the context is defined by an aligned concept in the context ontology and its neighbourhood of concepts. The user can simply switch contexts by either change the context ontology or by changing the alignment (i.e., selecting another method for the alignment) connecting the two ontologies.

The rest of this paper first describes the proposed approach and related work (Section 2) and next it demonstrates OntoConto, the NeOn Toolkit plugin (Section 3).

2 APPROACH DESCRIPTION

The main idea behind the proposed approach is to provide a context of a concept in one ontology by providing a part of the related ontology that is similar/related to this concept. This is knowledge that is inherently present in every network of ontologies and harnessing it can be beneficial in several ways. Ontology creator/editor gains help by finding out how certain parts of an ontology are presented in a different ontology and this can help with making editing decisions. An ontology user on the other hand, is presented with a tool that can help in understanding an ontology or a collection of ontologies. For example if the user has a task of annotating some data, the presented approach and the developed plug-in can help to understand each specific concept in the ontology by providing a context in another, maybe more familiar ontology.

Notion of context is defined here similar as in [3] by relating individual concepts of one ontology to the concepts of another ontology. Due to our emphases on visualization we have extended that original notion of context by including concept neighbourhood as follows. If a concept A in the first ontology is aligned to the concept B in the second, then since the "near" concepts of B are in some

way related or similar to B, this provides context for A. To define the "near" concepts of B, different approaches could be used. For example any type of a relation in side of an ontology could be used to define a unique metric between concepts, by counting how many hops over specified relation are needed to get from one concept to another. Every such metric would provide a specific neighbourhood defined by the relation used. This notion of context can be also illustrated by the idea of presenting an ontology through "semantic lenses" defined by the other ontology that is giving the context [4]. We are using the most basic SubConceptOf relation of the ontologies to define neighbourhood of the concept, as it is present in most if not all ontologies. Apart from this, SubConceptOf relation defines an underlying tree structure which can be very easily and intuitively visualized.

In our approach the two selected ontologies are visualized side by side, and the aligned concepts are visually linked across the ontologies. This supports the user in several different tasks. Most notably the user can edit ontologies inside an automatically generated context of another ontology. Apart from that, visualization also provides means that help with the evaluation and manipulation of the mappings. For an example see Figure 1.

The visualization used in the proposed approach can be broken into three parts: visualization of an ontology, visualization of the mappings and user interface.

2.1 Visualization of an ontology

Visualization of an ontology as proposed here is based on visualizing concepts ordered by an underlying tree structure stemming from SubConceptOf relation. Due to the demand to visualize a context for a selected concept and also due to the limitation of how many concepts can be distinctly visualized at the same time our approach visualizes only four layers of the ontology.

Often visualizing four layers at one time still proves too much information to visualize distinctly, thus additional visualization techniques are used. First concepts on the same level are vertically displaced in order to enhance the visibility of a single concept. Next we propose a visualization technique that is somewhat similar to using lens. Visual entities near the centre of the screen are enlarged, on account of reduced entities that are closer to the borders. This technique enables the user to clearly see the detail of interest, while retaining the big picture.

2.2 Visualization of the mappings

For the visualization of the mappings two selected ontologies are visualized side by side. This enables the user to acquire contextual information in networked ontologies across the two selected ontologies. The basic idea is rather simple: a simple link is visualized spanning the space between the mapped concepts.

Important functionality is provided by allowing the users to obtain third party auto-generated alignments. Currently our

approach is integrated with OntoLight (in cONTOext) and with the NeOn Toolkit Alignment Server (in OntoConto). Other such services could be added easily. Moreover, in OntoConto it is possible to store and load alignments inside the Neon Toolkit's workspace. This enables interoperability with other plug-ins.

2.3 Graphical User Interface

The proposed approach was developed with great care to simplify user interface with very simple actions to select and browse ontologies and alignments. The user is not required to understand ontology and mapping notation nor existing ontology editing software. This is inspired by the usage scenario of United Nations Food and Agriculture Organization (UN FAO) [5], where a group of users named "subject experts" is defined. Subject experts know about the domain to be modelled, but usually know little or nothing about ontology design issues or software for ontologies.

The user is enabled to easily load and then browse through two ontologies. Clicking on any concept selects it as the selected node, and the whole ontology gets dynamically redrawn in to the new position, with selected node on the top.

Editing of ontologies is supported either by the plug-in (cONTOext) or by the platform (OntoConto). Editing of the mappings is performed by selecting a mapping via clicking on it and then, editing its weight, the two concepts it links or deleting it altogether. It is also possible to create a new mapping by clicking two concepts to be connected while holding the key "m".

3 ONTOCONTO DEMO DESCRIPTION

OntoConto is a plugin inside the NeOn Toolkit application / platform. Primarily it functions as one of the graphical user interfaces of the Toolkit, with several advanced visualization techniques and this is why it was implemented with Adobe's Flash application. Flash is one of the leading dynamic visualization development applications, allowing easy development of advanced visualization techniques.

Integration is done by first implementing a regular java plugin which connects to the Toolkit's data model and extracts all the information of the ontologies and stored alignments. Second, this plugin consumes the Alignment Server and could also easily consume any other alignment providing plugin or webservice. Last, it implements a browser which in turn includes the OntoConto's visualization part written in Flash. Non-trivial communication is required in order to fully integrate the Java plugin and Flash applet.

Alignment Server serves several different methods that automatically generate alignments [6]. To consume it, it must be called with two OWL ontologies and a method. It returns the alignment in the RDF format. This makes it a perfect tool to be used by OntoConto.

In this way, it is possible to visualize any two ontologies loaded inside the NeOn Toolkit and request different

alignments for them. The Ontology of interest can be explored and edited inside the context of another. To change the context it is enough to either change the context ontology or change the alignment linking them. Finally it is also possible to make and store changes to the current alignment.

The OntoConto plugin is available at <http://kameleon.ijs.si/ontoconto>. To install it, one should have NeOn Toolkit installed. The binary version of OntoConto then needs to be copied to the plugins directory. No additional plugins are required.

4 USAGE SCENARIOS

The proposed approach was developed with great care to simplify user interface with very simple actions to select and browse ontologies and alignments. The user is not required to understand ontology and mapping notation nor existing ontology editing software. This is inspired by the usage scenario defined by UN FAO, where several groups of users are defined one being “subject experts”, who will be in charge of the everyday editing and maintenance work of the networked ontologies. Their duties also involve the creation of multilingual versions of ontologies. Subject experts know about the domain to be modelled, but usually know little or nothing about ontology design issues or software for ontologies. In user requirements [5, page 27] it is stated: “Subject experts should be provided with more intuitive interfaces than those available to ontology experts and application developers; in particular interfaces for subject experts should conceal much of the purely ontological and engineering decisions.”

The used combo selections, one click user interface and no ontology notation fits perfectly to serve this user group. Apart from the clear need of the proposed contextual visualization for users unfamiliar with ontologies, the proposed approach and demo also provides for expert users. Visualization in general provides easy exploration, understanding of any complex model, which an ontology network definitely is.

The first simple scenario this demo fulfils [5, page 40] deals with visualizing an ontology. Ontologies are visualized distinctly with the cut of the section of interest and simple one way to navigate to other sections.

The main scenario this demo fulfils is about System-supported creation of mappings between concepts. [5, page 46]. Actual manipulation of the mappings should always be done in context to the neighbourhood of both concepts it spans, for only so can the mapping be truly understood.

Apart from the pre-thought scenarios, this demo has already proved to serve other needs as well. An example is in evaluating the algorithms for ontology mapping [6].

The experiments are usually designed to answer how good these algorithms detect alignments in comparison with an expert of the domain. The expert defines the matchings between concepts of the two respective ontologies. With this

control set precision and recall are calculated for the alignments detected by the different algorithms.

To be able to detect concepts which represent the same aspect of the real world it is necessary to take the whole context into account. Concepts labelled with the same name do not necessarily have to describe exactly the same element; one concept might be more general or might describe a different part. To detect such ambiguities one has to check the concept with its neighbourhood, and that is enabled by the proposed demo.

The proposed demo also simplifies the qualitative analysis of the mapping algorithms. Some ontology mapping algorithms proposes semantic relations and the user has to choose the correct ones. For algorithms of this kind the strict quantitative analysis calculating recall and precision is not sufficient. The proposed context visualization enables to check the matching proposals and compare the different possibilities for each mapped concept. Moreover it helps to get a quick first impression of the available algorithms.

5 ACKNOWLEDGEMENTS

This work was supported by the Slovenian Research Agency and the IST Programme of the EC PASCAL2 (IST-NoE-216886), NeOn (IST-4-027595-IP) and ACTIVE (IST-2008-215040).

References

- [1] Grobelnik, M., Pajntar, B., Mladenić, D.: Reasoning with context. deliverable 3.2.2, NeOn(2008)
- [2] Pajntar, B., Mladenić, D., Grobelnik, M., Kohler, S., Dzbor, M.: Neon toolkit plugin realizing revised and extended techniques for contextualized visualization of ontologies and ontology networks. deliverable 4.5.2, NeOn (2009)
- [3] Grobelnik, M., Brank, J., Fortuna, B., Mozetič, I.: Contextualizing ontologies with ontolight: A pragmatic approach. *Journal Informatica* 32 (2008) 79-84
- [4] Fortuna, B., Mladenić, D., Grobelnik, M.: Visualization of Temporal Semantic Spaces. In: *Semantic Knowledge Management: Integrating Ontology Management, Knowledge Discovery, and Human Language Technologies*. Springer (2009)
- [5] Iglesias M., Caracciolo C., Jaques Y., Sini M., Calderini F., Keizer J., Hunte Ward F., Nissim M., Gangemi A.: User requirements. deliverable 7.1.1, NeOn (2006)
- [6] Le Duc, C., d'Aquin, M., Barrasa, J., David, J., Euzenat, J., Palma, R., Plaza, R., Sabou, M., Villazon-Terrazas, B.: Matching ontologies for context: The neon alignment plug-in. deliverable 3.3.2, NeOn (2008)

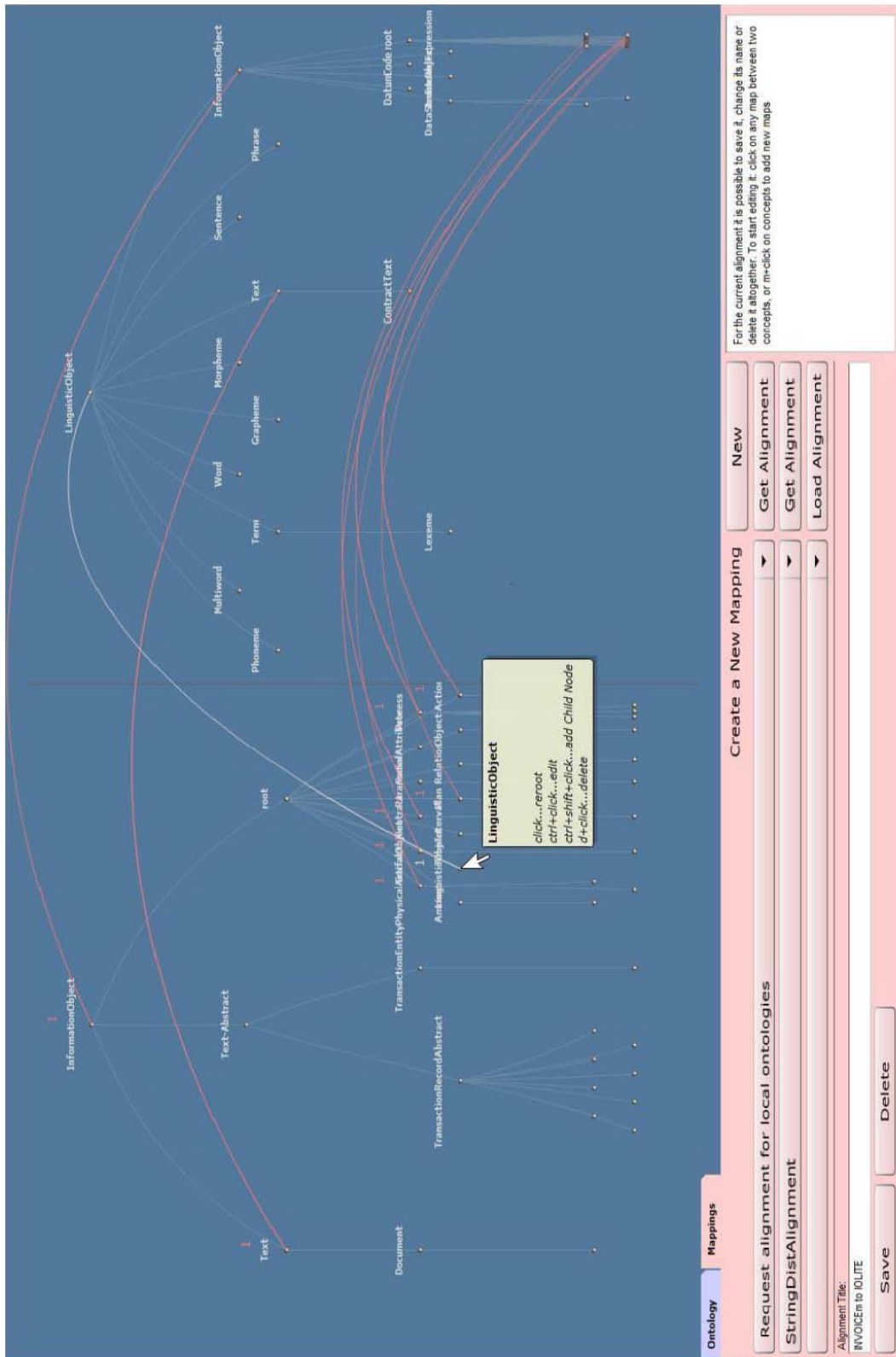


Figure 1. Two ontologies are loaded and visualized side by side. StringDistAlignment method was called for the alignment. Ontology of interest can be studied in the context ontology. There the concept “LinguisticObject” lays in between “InformationObject” and “Text”. It might make sense to redesign the corresponding concept in the starting ontology.

PROBABILISTIC TEMPORAL PROCESS MODEL FOR KNOWLEDGE PROCESSES: HANDLING A STREAM OF LINKED TEXT

Marko Grobelnik, Dunja Mladenić, Jure Ferlež

Department of Knowledge Technologies

Jožef Stefan Institute

Jamova 39, 1000 Ljubljana, Slovenia

E-mail: {marko.grobelnik, dunja.mladenic, jure.ferlez}@ijs.si

ABSTRACT

The paper presents an approach to modelling the data obtained from an observed environment driven by knowledge processes. It is based on the proposed a formalism for presenting probabilistic temporal process model consisting of three major components: (1) background knowledge (in the form of ontologies), (2) observed data (in the form of a stream of data items represented in different data modalities and possibly enriched with background knowledge) and, (3) objectives to optimize (providing guidelines for analytic techniques). The goal is to enable maintaining a data structure - to store, summarize and respond to a wide variety of queries about the observed low level data and about information and knowledge derived from the process. The formalism is realized in software components. Its functioning is illustrated on three scenarios: personal email, corporate email and document collections. The resulting platform is called TNT (Text-Network-Time) according to the main data modalities being addressed within the software.

1 INTRODUCTION

Communication today is greatly supported by electronic media meaning that there is potentially a trace of that communication to be found in electronic form. In we focus on communication inside organizations, the usual practice in organisations is to have more or less well defined formal business processes, informal side of an enterprise life is far from definitions which would allow formal approaches. In a way, we can say that the informal part of the enterprise's life is everything that is not regulated by formal business processes. This means that we can potentially have difficulties even describing what the informal part is, not to mention difficulties when defining, monitoring and analyzing processes within an organization beyond the formal rules.

We can capture the informal part of an enterprise: (a) in a top-down manner where informalities could be described by some kind of rules and logic formalisms, and (b) in a bottom-up data-driven manner, where the goal is to model the processes by observing a life of an enterprise from the side where people leave their traces (in the form of some kind of log files such as, communication records, emails, documents, search logs, etc.). In this paper we focus on the

bottom-up/data-driven type of modelling where the goal is to perform automatic discovery of regularities in a functioning of an enterprise based on the available observed data in addition to background knowledge.

Probabilistic temporal process model for knowledge processes is defined here to cover different levels of data complexity where time is an inherent property of the data regardless its complexity. By data complexity we mean here data ranging from simple record of action over records equipped with additional textual description to network of records (where both records and links can have additional textual description). The resulting class of models (modelling temporal processes) is based on an underlying data structure that enables capturing information provided as text, network and time (TNT). In the basic setting the data structure is a timeline of simple records where the main goal is to find patterns leading to identification of tasks that the user is performing and using the identified tasks to predict/suggest to the user the next action (system and/or information source that the user might like to access next). However the timeline of records can be a timeline of documents where the underlying data structure should enable compact storage and representation of text to support search and visualization. Moreover, the records timeline can contain links between the records based on some record properties forming a dynamic graph/network. In that case, mining can be performed on nodes of the graph and/or links of the graph (e.g., taking into consideration neighbourhood of a node).

The main contribution of this paper is the proposed approach based on a unifying formalism being able without significant modifications to deal with different complex scenarios within enterprises. The formalism is designed in a way which allows reusing of appropriate analytic methods from the areas like machine learning, data mining, and social network analysis while preserving enough expressive power of generated structures for modeling the targeted real life situations. The rest of this paper is organized as follows. Section 2 gives a brief description of related work, while Section 3 presents the proposed approach. Implementation of the proposed approach is given in Section 4, while Section 5 and Section 6 give details via describing the data structures and process modelling. The paper concludes with discussion in Section 7.

2 RELATED WORK

Modelling knowledge processes as addressed here assumes that the user is a central actor in the knowledge process accessing some systems/tools and some information sources (data). Probabilistic temporal process model for knowledge processes will be generated from (a) a stream of captured actions captured and (b) relevant background knowledge.

Stream of actions can be seen as a sequence of actions that are potentially connected in more complex actions forming tasks. The area of research dealing in particular with these kinds of scenarios is a subfield of Data Mining called Stream Mining [1, 5], the area got popular as the amount of data increased and the usual scenario where that data was analysed in batch mode was not appropriate anymore. Analogous to Data Mining [4, 8] Action Mining can be used to describe the observed data or for prediction. Probabilistic temporal process model, as defined here, enables both descriptive and predictive mining. In descriptive mining, actions (described by some properties capturing nature of the observed processes) can be described in a more general or coherent way, causal analysis can be performed on them, and anomaly detection can be applied. In predictive mining based on frequent sequences of actions possibly useful patterns can be first proposed to the user to guide identification of tasks. Then each of the identified tasks can be modelled (e.g., by a task template) and applied in the future activity of the user to predict the next user's action. In the above scenarios scalability is an important issue which needs to be addressed separately – the methods dealing with streams are typically able to take data transactions from a stream and change the resulting model incrementally with a small cost per data transaction. Generally we distinguish methods maintaining one global model of the stream data and the ones with many local models (which could act as one big model as well).

Depending on a specific scenario there may be some additional relevant knowledge (commonly referred to as background knowledge) available. Background knowledge can be of rather general nature (such as, an ontology that is describing possible actions) or specific (such as, description of experts skills or some relation between the experts e.g., being frequently in the same project team). The process of action mining may however require generation of background knowledge from some related data. For instance, automatically suggesting experts for specific topics based on authorship of internal enterprise documents on the topics. The main function of background knowledge in this work is to enrich the measured data from the target environment. Namely, the data we are receiving need some extra explanation which is not a part of the measurements – additional information/knowledge in the form of databases, rules, ontologies etc. should therefore serve mainly as data interpretation facility and a way how to bring semantics into the measurements. For example, this can mean extending the existing data records coming from measurements with extra features (based on background knowledge) which will be used for building better statistical models.

3 APPROACH DESCRIPTION

The proposed approach to modeling complex dynamic systems uses the data obtained from an observed environment driven by knowledge processes. The main dimensions along which the approach is being developed are the type of input data the approach is able to deal with, scalability issues and background knowledge for interpretation of the observed data. The main data modalities being used are structured content (e.g., relational bases), graphs extended to networks (e.g., social networks), and textual content (document databases). The data items are coming into the system through time opening the temporal dimension when dealing with the data. For additional interpretation (e.g., data enrichment or introducing semantics) of the data we introduce databases or ontologies which further enrich the observed data. Scalability is achieved by a careful selection of appropriate analytic methods which assure appropriate time and space complexity – the approaches is designed to cover everything from online scenarios (mining streams of data) to batch style processing. Architecture of the proposed approach is shown in Figure 1. The top most object is an environment (“System”) including one or more observable and measurable entities (“Entity”) whose activities are generalized (“Process Model”).

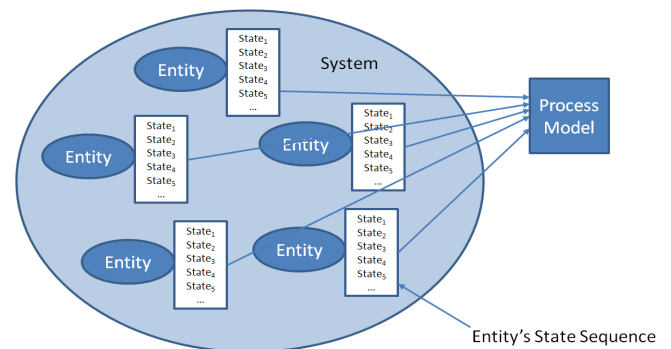


Figure 1 Architecture of the proposed process modelling.

The environment “**System**” can be in general anything with some observable internal dynamics. This includes environments like personal computer desktop, e-mail client, document or web server, news forums, web 2.0 portals, social relationships of any kind, agent based systems, corporate business processes, etc. It is not necessary to have a complete understanding and observation of the environment – we should acknowledge that there can always be some hidden parts having influence on the dynamics within the environment.

Within the environment we expect to have one or more (potentially a very large number of) entities (of type “**Entity**”) which interact between themselves, with a hidden or visible internal architecture and can have interaction with the world outside the observable environment. Each entity should have a well defined state which captures relevant measurable information about the activity – state of an entity in general cannot capture everything needed for complete understanding of the dynamics, but rather what is practical to measure and analyse. A state in general consists of a set of

variables representing different aspects of the measured entity. Each entity has a log of past activities represented as a sequence of past states which serves for further analytic purposes. Examples of entity logs are mouse movements, activated applications, incoming/outgoing emails, web server log files, news forum postings, instance messenger logs, etc. All the entity logs include lots of information which is unmanageable without proper abstraction. The goal of the presented architecture is to construct abstractions (“**Process Model**”) in the form of processes describing entity logs in a compressed form understandable for human reading and reusable for various machine applications. The idea is to analyse the data generated by entities and construct a model in the form of patterns, regularities and other kinds of rules for different kinds of tasks like descriptive & exploratory analysis, causal analysis, anomaly detection, prediction etc. The discovered models can then be further used in higher level applications for e.g. optimising personal efficiency on a desktop, for efficient communication over email or instance messenger, for competence discovery within an enterprise, for understanding of informal aspects of an enterprise etc.

4 SYSTEM ARCHITECTURE

The implemented system called TNT (Text-Network-Time), is designed as a set of components maintaining server side functionality. It consists from three major chunks of the functionality: creating TNT database, maintaining the TNT database, querying the TNT database. At the core of the TNT design is a specialized data structure supporting "dynamic networks" enriched with "dynamic content". This structure is able to respond to queries with temporal, network and content constraints. The main functionality of the TNT system is to deal with event processing, events analysis and extraction of abstract event patterns out of the data in a way, which enables accommodating the size and the nature for the streaming data which could be expected within corporate setting. It is expected this to be in the range of ten of events (e.g. emails) and queries per second.

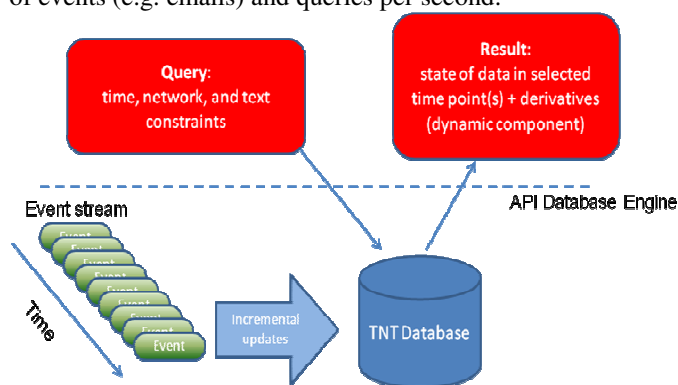


Figure 2 Architecture of the TNT system implementing the proposed approach.

The usual relational databases (such as SQL Server) are not suitable for these kind of tasks because of their transaction processing nature which does not efficiently support the large semi structured data (such as networks and text). Closer to the data structures we are using in TNT database are the inverted index databases for textual information retrieval. However, these have a problem of not

supporting temporal queries and networks (graphs), but only support indexing over text. In other words - in the same way as the specialized databases use customized data structures (like textual search, image search, OLAP/data warehouses, triple stores, ontologies, etc.) we use specialized data structures for dealing with the type of operations we will need for mining and searching the kind of data we have in focus (dynamic social networks, dynamic content).

The TNT database will process the events originating in the knowledge workers desktop environment. The TNT database is directly using data structures defined in Section 5. The most important are the dynamic networks enriched with content/textual part with limited queue data structures containing the temporal information. The system is designed as a set of components maintaining server side functionality. It consists from three major chunks of the functionality (see Figure 2): (1) **TNT database** – a data structure allowing other components to insert and query the event data coming into the system. (2) **Maintaining the TNT database** – this includes creation of the TNT database and incremental updates of the TNT data structure for each event coming into the system (Figure 2). It is important to note that the efficient implementation of the incremental updating is the key for the scalability of the system and directly influences efficiency of the throughput of the events. Amortized complexity [Napaka! Vira sklicevanja ni bilo mogoče najti.] of the updates within our system is constant in time and linear with upper limit (based on the parameters of the system) in space. Amortized complexity means averaging the time and space complexity over the worst case internal infrastructural operations (such as resizing etc). In TNT updates are happening with constant time amortized complexity, while space amortized complexity is linear with the respect to the system parameter which limits the time window of data being monitored. Internally the TNT database maintains the information on a social network between actors. Every actor is described with information on the content, social network and time of the events associated with the actor. Additionally, the content and temporal information on communication between any two actors is maintained in the links of the social network. The TNT is designed to consume one dense stream of events which affect the internal data structures. For instance, TNT is able to consume email events between people inside an organisation in a centralized manner. Each individual event triggers update of the right part of the content, social network, and temporal internal structures of TNT, which describe the actors, the content of communication and time constraints of the email event. We tested the TNT database on a stream of email events generated out of the ENRON dataset [2]. The TNT system running on 2.8GHz machine was able to consume the whole stream of 500 000 email events between 150 people in approximately 15 minutes. (3) **Querying the TNT database** – querying is conceptually similar to other client-server types of architectures, the main difference is in the structure of the query consisting from three major parts: temporal part, network part, and textual part. Therefore, each query isolates network part and content part within the data and shows dynamics of the data within the specified temporal

constraints (examples will be shown in the examples section). The resulting output of a query consists from state (network and textual part) of data in different time point (based on the temporal part of the query) and corresponding derivatives showing dynamics (trends) of data.

5 DATA PRELIMINARIES

The key element for modelling knowledge processes is the availability of data within enterprises (where project case studies are just a good sample of situations on the market). In the ideal case, the data would capture all the activities on different levels of an organisation which would allow reconstruction of most processes and other relevant concepts to understand how the organization is functioning. But unfortunately, this is impossible to achieve for many practical reasons. After some review of what kind of data is realistic to retrieve (like availability) and what type of data is manageable (like data complexity, scale, granularity), from technical point of view, we concluded to use some of the fundamental data types on which we have define characteristic operations.

The data can be of different types (modalities). The most typical data types are the following: structured records (e.g. databases), textual data (text documents with some or without structure), and networks (set of connected objects and their relationships). Each of the above data modalities allows characteristic set of operations which could be used for extracting higher level constructs from data. Each of the data types has inherent temporal nature, which means that structured data records and documents are coming in streams and networks have dynamic nature. We have defined the data model for the structures we expect to be dealing with in a way which allows relatively straightforward implementation in programming languages.

Any data can be of atomic or structures types: **DataType = AtomicDataType | StructuredDataType**. Atomic data types are the ones known from programming languages – in the description below we won't go into the details of lexical representations for each of the atomic data types but rather stay on the level of informal description: **AtomicDataType = Void | Boolean | Integer | Float | String**. Each of the atomic data types has defined set of operations which are well known from algebra and programming language practice. Structured data types are composed from the atomic data types and other structured data types. **StructuredDataType = Pair | KeyVal | Tuple | Vector | Stack | Queue | Set | Map**. All of the structured data types require additional subtypes which need to be specified when the data type is getting instantiated.

Complex data types consist of basic data types and implement the key data structures for monitoring processes within enterprises. We tried to restrict the selection to the minimal possible set of data types covering typical data modalities one can find in enterprise environments. We identified three major types of data which we use in a static and stream context (meaning either we have data available as a static dataset or the data is coming in a stream).

Structured records as found in databases represent the most basic way of representing data in an environment.

Typically all the standard data of an enterprise would be represented in this way – this corresponds to the data stored in relational databases, indexed in various ways and managed by DBMS type of systems.

Textual data with potentially some additional structure (in a form of e.g. documents, web pages, presentations) represent most of the unstructured data retrievable from enterprise document servers. This type of data represents core of the less formalised part of the enterprise data. By extracting data from text and putting it into other structured forms (like structured records or networks) we make steps towards manageability of the less formal information.

Networks are the most generic formalism to describe structure in the data. The basic underlying structure for a network is a graph with additional information attached to the vertices and edges of a graph. By structuring the information in different ways on a network we can represent very complex situations in an environment. In addition to the standard definition of a static network we will define also a dynamic version of it, which changes through time which will represent basis for monitoring processes within an enterprise.

6 PROCESS MODELING

Process modeling involves some process and modeling methods applied on the process. This Section defines both of them in the light of the proposed approach.

6.1 Process definitions

A “process” is often used as an abstraction for a sequence of events. Processes with some particular meaning would often get a special name denoting the effect the characteristic sequence of events has on the environment. For different contexts where the notion of process has its meaning see en.wikipedia.org/wiki/Process. A process always happens to an entity **Entity** which operates within a system **System**. **Entity** is described by its state **State** which is usually described by a set of variables of a type **Tuple** denoting various characteristics of the observer **Entity.State = Tuple**, where variables within a Tuple describe state of an Entity. A process is a sequence of states in which the observed entity appears. In the case of discrete time (being reasonable assumption in our case) we can say that the sequence of states is a **Vector** of the type **State**. **StateSequence<State> = Vector<State>**, where the states are order in time Each state in the sequence is derived by the following formula: **StateSequence[t=0] = initial state** where state variables get initial assignment of values. **StateSequence[t+1] = NextState(Process[t], Action, BKnowledge)**. Where an **Action** can be either an empty value or of a type **Tuple** generated by an external source within the **System**. Typically, an **Action** could be an asynchronous event which influences the state of the observed **Entity**. **Action = Tuple**. Background knowledge represented by the parameter **BKnowledge** serves as an additional (mainly static) resource for interpretation of the **Process**, **State** and **Action** within the function **NextState**. **BKnowledge** can appear in the form of a database, set of rules or even as an ontology providing semantics for the data being involved in the description of

states within sequences and further on within processes. In this formalization we will avoid specifics of the background knowledge – each example of the **NextState** function in the next sections will provide its own specific type of additional information needed to describe the states and processes. Operationally **BKnowledge** can be understood as a generic query function returning values on the requested (domain specific) queries. Having the data collected in the form of one or several **StateSequence** structures (e.g. one for each **Entity** within the **System**) we can build a model of the data by using analytic techniques such as statistics, machine learning, data mining etc. The model is expressed in a formalism (hypothesis language) supported by the selected analytic method (like rules, decision trees, linear or non-linear functions, etc.). The purpose of the model is to abstract or summarize the analysed data into a shorter description which typically probabilistically resembles the structure and properties of the data. Specific type of modelling and selection of the analytic methods depends on the task (described in the next section). From **StateSequence** data construct **ProcessModel** via modelling using **BKnowledge** and **Target** value: $\text{ProcessModel}_{\text{Target}} \leftarrow \text{Modelling}(\text{AnalyticAlgorithm}, \text{Vector}\langle\text{StateSequence}\rangle, \text{BKnowledge}, \text{Target})$. **ProcessModel** is further used in any situations as independent functions getting on the input data from a **StateSequence** and returning the **Target** value which were modelled within the **Modelling** phase: $\text{Target} \leftarrow \text{ProcessModel}_{\text{Target}}(\text{StateSequence}, \text{BKnowledge})$.

6.2 Process Mining

Analysis on the top of the event sequences can be done in various ways. Many areas of science are actively working in subfields of analysis of processes which are mainly determined by the structure and properties of the **State** description of the observed **Entity** and by the language used to describe **NextState** function.

First example of process modelling is from mathematical analysis. Differential equations are well known mechanism for describing processes for the cases where the **State** is described with continuous variables and the function **NextState** is described in analytical language. In the same way, but on the other side of the spectrum within the areas of data mining there is a family of algorithms on “Mining Frequent Episodes” [11] which operate on a set of binary variables and the language for describing **NextState** function is “association rules formalism” (if-then rule).

An important class of processes, which will be of particular interest for our work, are situations where the **State** encodes a network (of a type **Network** or **DynamicNetwork**) and by the change of state (function **NextState**), the observed **Entity** moves from one node to another. Apart from encoding a **Network**, a **State** can include in the rest of its variables also other characteristics of the modelled **Entity**. This is entirely domain dependent and can cover wide range of situations. The importance of this class of processes is relation to process diagrams which are widely used within business process modelling. Diagrams, which can be easily represented by an instance of a type **Network** and

corresponding **NextState** function can describe broad class of situations which we can envisage in process mining.

Two typical scenarios for analysing processes are: (1) Analysing the data appearing within the **StateSequence** sequence. This allows finding regularities within the sequences without any specific assumptions on the structure of data etc. This kind of scenario allows mainly understanding of the process event sequences but not other tasks. (2) Identification of the **NextState** function is the most common approach to the analysis of the process data. The goal is to capture specifics of the process within the model by which we describe the function. For certain classes of models (defined typically by the language used for the description of the **NextState** function) we can efficiently calculate the function while for some other classes it can be much more difficult or almost impossible. Usually we are trading between the computational efficiency of the reconstruction of the function versus expressivity of the language for describing **NextState**. Main tasks that we are addressing by process modelling are process description (identify main regularities within the processes and describe them in a formal language), causal modelling (reconstruct causal chains of state sequences including a subtask of root cause analysis), anomaly detection (identify unusual situations) and prediction (model to predict future unseen situations).

On the top of the structures defined in the previous sections like **Process**, **StateSequence**, **DynamicNetwork**, **Network**, **Document**, etc. one can use a broad class of methods from research fields like machine learning, data/text/web mining, statistics, time series analysis, social network analysis, computational linguistics, etc. Following the structure from the previous section, we can assign some more relevant analytic methods to solve each of the tasks. The assumption here is to have on the input the structure **StateSequence** (possibly enriched with **BKnowledge**) upon which we execute some of the analytic methods to solve the task. Most of the tasks deal with the reconstruction of the function **NextState** or with analysis of **StateSequence** data.

In process description we would like to explain and understand the functioning of the function **NextState** operating in the typical situations described in the **StateSequence** data structure. Main classes of methods for decomposing the **StateSequence** data into some kind of structure are so called unsupervised methods like clustering and other eigenvector decompositions. These kinds of methods provide insights into the structure of the data and together with additional domain knowledge being used for exploratory data analysis enable deeper understanding on what is going on within the data. These methods are often an input for data visualization methods which typically generate graphical summaries of the data. An example of such descriptive analysis would be analysis of a corporate email server an approximate organigram of the analysed organisation based on the sub-communities discovered in the analysis [7].

In causal analysis we typically need to construct inverse of the function **NextState**, where the goal is to predict which state was preceding the current state. First, we need to

construct (either from data or domain knowledge) set of discrete states in which the process can appear. Further, with the analysis of the past data of process behaviour we construct models of relationships between the discrete states which enable causal analysis. Typical scenario we try to solve is given an event, causing the process to appear in a certain state, to generate list of other states which most likely preceded our current event to happen. The key here are the methods to detect relationships between the states – this can be done in various ways (depending on the problem we are solving). The most obvious one is correlation between states and events that happen where correlation represent probabilistic relationship between the states. In the presence of additional information we can build local predictive models for “predicting” which path to follow from a state backwards to the cause. Another option is also to build global model of relationships via Markov models where we can test how likely a certain sequence of actions could have happened [10]. Example of causal relationships are probabilistic Markov-network style if-then rules which establish probabilistic relationship of preceding events based on the later sequence of events. **If** Device-123 failed, **then possible causes are:** failure of Device-456 (with probability 0.7), failure of Device-567 (with probability 0.2), failure of Device-789 (with probability 0.1).

In anomaly detection we generally apply statistical measures for measuring surprise or probability of a certain sequence or an event to happen. In a general case, we compare a particular probabilistic distribution with a general one and if it is statistically significantly different then we report an anomaly. With an additional domain knowledge which provides clearer model of behaviour of the observed system we can detect anomalies in a more accurate way. Example of anomalies in the scenario of monitoring computer’s desktop activities of the user could be unusual usage of the resources based on the running applications. Unusually high usage of certain resources could be an anomaly caused by a virus or malfunction of the application. Anomalies are often reported as comparison between what is usual and in what way the anomalous situation differs from that. In the application of observing corporate business processes (being typically in the form of process diagrams) an unusual increase of activities in the part of the diagrams where intensity is not as high. This kind of information could be a signal for management to react in a proper way. In the application of observing corporate e-mail server, it could be a change in the intensity of communication between some project group members which can have a cause in problems in their relationships which can further lead to problems on the observed project.

In prediction we predict which state will be the next one in the sequence of states within the **StateSequence** structure. In this task we typically take a window of recent states and try to predict most probable next state. The models we build for prediction are typically built locally for each state in the vocabulary of states. To solve the task we have available a broad range of predictive modelling methods from machine learning and data mining. The methods being used the most in the recent years for predictive tasks are Support Vector

Machines (SVM) [3] and Conditional Markov Fields (CRF) [9]. An example of prediction can be within the application on personal e-mails where the goal is to predict when we might expect a response to our e-mail sent to a certain person.

7 DISCUSSION

We have described an approach to modelling complex dynamic systems by modelling the data obtained from an observed environment that is driven by knowledge processes. It is based on the proposed a formalism for presenting probabilistic temporal process. The main functionality of the developed TNT system implementing the proposed approach is to deal with event processing, events analysis and extraction of abstract event patterns out of the data in a way, which enables accommodating the size and the nature for the streaming data which could be expected within corporate setting. The TNT system pipeline currently consist of utilities which enable the system to process the email events’ data and metadata, to store it inside the TNT database, and use queries which report on temporal, content and social network aspects of the data and meta data about email stored inside the TNT system.

8 ACKNOWLEDGEMENTS

This work was supported by the Slovenian Research Agency and the IST Programme of the EC PASCAL2 (IST-NoE-216886) and ACTIVE (IST-2008-215040).

References

1. Aggarwal, C., Data Streams: Models and Algorithms. In Advances in Database Systems, 2006.
2. R. Bekkerman, A. McCallum, and G. Huang. Automatic Categorization of Email into Folders: Benchmark Experiments on Enron and SRI Corpora. CIIR Technical Report IR-418 2004
3. Cristianini, N., J Shawe-Taylor, J. (1999) An introduction to Support Vector Machines: and other kernel-based learning, Cambridge University Press
4. Fayyad, U., Piatetski-Shapiro, G., Smith, P., and Uthurusamy R. (eds.) (1996) Advances in Knowledge Discovery and Data Mining. MIT Press, Cambridge, MA, 1996.
5. Gama J., Gaber, M.M., (2007) Learning from Data Streams: Processing Techniques in Sensor Networks
6. Grčar, Miha, Mladenčić, Dunja, Grobelnik, Marko. User profiling for interest-focused browsing history. In Proceedings of the 8th international multi-conference on Information Society IS 2005. Ljubljana: Institut "Jožef Stefan", 2005, pp. 182-185.
7. Grobelnik, M., Mladenčić, D., Fortuna, B., Ontology Generation from Social Networks. In Semantic Knowledge Management: Integrating Ontology Management, Knowledge Discovery, and Human Language Technologies by Davies, Grobelnik, Mladenic (eds.), Springer, 2009
8. Hand, D.J., Mannila, H., Smyth, P. (2001) Principles of Data Mining (Adaptive Computation and Machine Learning), MIT Press.
9. Lafferty, J.D., McCallum, A., Pereira, F.C.N. (2001) Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data, Proceedings of the Eighteenth International Conference on Machine Learning.
10. Manning, C. Schütze, H. (1999) Foundations of Statistical Natural Language Processing - MIT Press. Cambridge, MA.
11. Mannila, H., Toivonen, H., Verkamo A. (1995) Discovering frequent episodes in sequences. In Proceedings of the First International Conference on KDD.

EXPLORATORY ANALYSIS OF PRESS ARTICLES ON KENYAN ELECTIONS: A DATA MINING APPROACH

Senja Pollak

Jožef Stefan Institute

Jamova 39, 1000 Ljubljana, Slovenia

e-mail: senja.pollak@ijs.si

ABSTRACT

This paper investigates the utility of applying data mining techniques to media analysis, more specifically, to the analysis of a corpus of articles covering the 2007 Kenyan election and post-election crisis, aimed at capturing the differences between local (Kenyan) and Western (British) newspaper articles. Having formulated this task as a binary classification problem, we have succeeded to reveal interesting phenomena in the data using data/text classification methods and class association rules, opting for techniques where interpretability of results prevails over their accuracy.

1 INTRODUCTION

Media analysis has been a topic of several studies, including a recent media analysis study performed by Fortuna et al. (2008). Text classification methods prove to be a useful vehicle e.g., for different newspaper article classification tasks, such as article genre, topic or author classification.

One of the starting points of this paper is the theory of pragmatics, which pays a lot of attention to choice-making in language use (Verschuere, 1999). Lexical, syntactic or discursive choices are significant: the use of words, syntactic structures, modality markers, etc. as well as absence of their use is always meaningful.

Our approach to text classification is formulated as a binary classification task, aimed at distinguishing between the articles from Kenyan newspaper *Daily Nation* and British newspaper *The Independent* (a forthcoming, more extensive experimental study will take a larger set of articles from different Kenyan and European/American newspapers, and will – like in this study – aim at distinguishing between two selected classes: *local* and *Western*). The starting hypothesis of this work is that news coverage of Kenyan events is not the same in the local and in the international (Western) media. Using data mining/machine learning techniques, the main goal of our work is to explore how some of the lexical choices differ in the Western and local media.

The structure of this paper is as follows. In Section 2 we present the data. Section 3 outlines the text classification and machine learning techniques used in our analysis. Section 4 presents selected results of the analysis. Section 5 presents the conclusions and plans for further work on a larger selection of 464 newspaper articles.

2 DATA

This section presents the data (presented in Section 2.1), data cleaning (2.2) and data representation (2.3).

2.1 Data description

Originally, the corpus was collected as part of the project *Intertextuality and Flows of Information* in the field of pragmatics. The collected corpus consists of articles from six different newspapers (Kenyan, British and American) in English language, covering the Kenyan election and post-election crisis between December 2007 and April 2008.

The Kenyan presidential and parliamentary elections were held on the 27th of December 2007. The two main candidates were the incumbent President Mwai Kibaki and the opposition presidential candidate Raila Odinga. Kibaki is a member of the traditionally dominant Kikuyu ethnic group and Odinga is a member of the Luo ethnic group. Kibaki was declared the winner and sworn in despite the opposition leader's claims of victory. The election was followed by violence and conflicts.

For the first experiments, which are the focus of this paper, we had only a limited set of articles at our disposal. Our data set consists of 72 articles from only two newspapers: 36 articles from *The Independent* for Western media (WE) and 36 from *Daily Nation* for local media (LO). This subset was used in the reported study and will be followed by experiments on a selection of a larger number of articles.

2.2 Cleaning the data

Since our aim is to better understand the way of reporting on the same event by two different newspapers, we had to remove all information that could be distinctive for the two classes, but not important for our work. To illustrate, newspapers have normally only few journalists covering Kenya events, so if not removed, the author's name could easily be selected as a distinguishing feature. Therefore, we removed meta-information such as newspaper source, authors of articles, dates of publication, photographers, mails of authors, types of articles, etc. For this purpose, we made scripts in Perl and used only the remaining relevant data for document classification (titles, text and photo descriptions).

2.3 Data representation

Each of the two classes, LO (local) and WE (Western), contains 36 instances (articles). The class is a nominal attribute.

For data representation we had to select the attributes, and represent the articles with feature vectors. We selected word unigrams (W1) and word bigrams (W2) as attributes. The attributes have a numeric value, calculated on the basis of term frequency (*tf*). Since having all the attributes would result in too large feature vectors, attribute selection of best 500 attributes ranked by *chi square* values was first performed. This feature selection and transformation to feature vectors was done in TACTiCS¹. We also used a binary representation of this selection of attributes (1 - word is present in the document, and 0 - word is not present in the document) as well as concatenation of word unigrams and word bigrams. To summarize, we experimented with the following feature sets: W1 (word unigrams weighted by term frequency), W2 (word bigrams weighted by term frequency), W1-bin (word unigrams transformed to binary representation), W2-bin (word bigrams transformed to binary representation), W1W2 (concatenated W1 and W2 feature sets: 500 selected word unigrams followed by 500 selected word bigrams).

3 TEXT CLASSIFICATION AND MACHINE LEARNING ALGORITHMS USED

This section starts with a theoretical definition of text classification and situates our task within this framework in Section 3.1, while Section 3.2 presents the tools and motivates our choice of machine learning techniques used.

3.1 Text classification

Text classification can be defined as “automatic assignment of documents to a predefined set of categories” (Sebastiani, 2002), and learning of text classifiers (classification models) can be performed by supervised machine learning. Text categorization is the task of assigning a binary value (T or F) to each pair $\langle d_j, c_i \rangle \in D \times C$, where D is a domain of documents and $C = \{c_1, \dots, c_n\}$ is a set of predefined categories. In our case $C = \{WE, LO\}$, WE meaning Western media and LO meaning local Kenyan media.

Since we want exactly one category to be assigned to each $d_j \in D$, and we have only two complementary categories $|C|=2$, our case is an example of binary text classification, a special case of a general N class single-label (or non-overlapping categories) case.

3.2 Algorithms used

The choice of symbolic data mining algorithms from the Weka 3.6.0 environment² was motivated by the need for ensuring the interpretability of results, possibly at a cost of not achieving the highest possible accuracy. For this reason, we used decision tree and decision rule classifiers and not better performing algorithms such as support vector machines or nearest neighbours algorithms. We used the following algorithms available in Weka: J48 for learning decision trees (Quinlan, 1986 and 1993), 1R for decision rules (Holte, 1993), JRip for decision rules (Cohen, 1995), PRISM for decision rules (Cendrowska, 1987), and PART

for decision list learning (Frank and Witten, 1998). An experiment was done also by learning of class association rules with Predictive Apriori (Scheffer, 2001).

4 RESULTS, EVALUTATION AND INTERPRETATION

This section is divided into two parts. Section 4.1 presents some results of algorithms applied to the entire data set (72 instances) without performing accuracy evaluation. In this setting we used no pruning, trying to build models and patterns that describe the given data set. This part is called exploratory data analysis (descriptive analysis). In Section 4.2 we present the results of predictive/classificatory data analysis on which evaluation was done with 10-fold cross-validation.

4.1 Exploratory analysis

We present the construction of models induced from the whole data set. The classification accuracy of these models, if evaluated on the training data itself, can be interpreted as an upper bound for the model’s performance of new, unseen data (nearly always 100%) (Witten and Frank, 2005).

Results of JRip are presented in Table 1. We see that we get some interesting descriptions of the corpus. We present the rules that cover the Western articles (where local articles are treated as ‘else’). We must keep in mind for the interpretation that in learning of these rules, the examples covered by the currently constructed rule are excluded from the data in the next rule construction iteration, hence each rule below covers only the examples not yet covered by previous rules. We present the analysis of two rules, which show how the information obtained with text mining tools could be useful for further discourse analysis. We present the results on the binary W1-bin feature set.

Table 1: JRip rules induced from the W1-bin feature set for class WE (minimum number of object=1, without pruning)³.

(raila = 1) and (tribe = 1) => class=WE (11.0/0.0)
(raila = 1) and (go = 1) and (next= 0) => class=WE (6.0/0.0)
(may = 1) and (new = 1) => class=WE (4.0/0.0)
(union = 1) and (john = 1) => class=WE (3.0/0.0)
(national = 0) and (major = 1) and (nations = 0) => class=WE (4.0/0.0)
(rather = 1) and (real = 0) => class=WE (2.0/0.0)
(by = 0) and (its = 0) => class=WE (2.0/0.0)
(could = 1) and (raila = 0) and (calm = 0) and (running = 0) => class=WE (3.0/0.0)
(alone = 1) and (emergency = 1) => class=WE (1.0/0.0)
=> class=LO (36.0/0.0)

Rule 1 that covers 11 out of 36 Western articles says that if words *Raila* and *tribe* are used, the article belongs to the Western class. *Raila Odinga* is the Kenyan presidential candidate that lost the election. The choice of word *tribe* is very interesting for the analysis. *Tribe* is a very ideologically marked word. If we talk about tribes, tribal wars and conflicts, it can be a pejorative, ‘savage’

³ The first number in the parenthesis after the class means the number of examples correctly covered by the rule; the second number means false positive examples and is in our case always 0.0.

¹ TACTiCS: <http://www.cnts.ua.ac.be/stylometry/demo.html>.

² Weka: Data Mining Software in Java <http://www.cs.waikato.ac.nz/ml/weka/>

description of African situation. Talking about the presidential candidate and at the same time of tribal division exclusively in the Western media should be further analysed. As explained earlier, the two main candidates *Raila Odinga* and *Kibaki* are from two different ethnic groups, but emphasizing the tribal aspect is an ideological choice of Western media.

Rule 4 is also interesting for the analysis, where we must keep in mind that the rules cover the examples that have not yet been covered by the preceding rules. If we check the articles where words *union* and *John* appear at the same time, we see that Rule 4 covers the articles, which refer to *John Kufuor*.

- *John Kufuor, the head of the African Union...*

- *John Kufuor, the President of Ghana who is the current chairman of the African Union,...*

The next set of rules was obtained with the PRISM algorithm. PRISM generates only rules with 100% accuracy. We obtained a large set of rules perfectly covering the entire corpus. In Table 2 we present just first three rules describing the *Western* and the *local* class. Since PRISM selects only 100% accurate rules, it presents a different view on our data and the rules are much more specific.

Table 2: Selection of PRISM rules on the W1-bin feature set.

If briefly = 1 then WE
If tribe = 1 and rest = 0 then WE
If challenger = 1 and continue = 0 then WE
If running = 1 and national = 0 then LO
If raila = 0 and go = 1 then LO
If kericho = 1 and challenger = 0 then LO

We made experiments also with association rules (Predictive Apriori). For applying association rules, we first performed feature set reduction (choosing a smaller number of attributes from initially 500 attributes selected). We performed attribute filtering with different filtering methods and obtained new feature sets of binary word unigrams. We present the association rules on the feature set obtained by previous selection of 10 attributes by Relief (with kNN set to 10). The selected features are: *emergency, tour, described, tribe, raila, sell, sharing, cancel, union, could*. With Predictive Apriori we can mine class association rules and the measure used is predictive accuracy, combining confidence and support into a single measure (Scheffer, 2001). We searched for the best five rules. For the feature set selected on W1-bin by Relief, the induced class association rules are presented in Table 3.

Table 3: First five class association rules on the W1-bin feature set.

1. tribe=1 raila=1 11 ==> class=WE 11 acc:(0.9923)
2. union=1 tribe=1 4 ==> class=WE 4 acc:(0.96755)
3. tribe=1 described=1 4 ==> class=WE 4 acc:(0.96755)
4. tribe=1 emergency=1 2 ==> class=WE 2 acc:(0.92507)
5. tribe=1 12 ==> class=WE 11 acc:(0.9175)

We see again that *tribe* is the most characteristic word appearing in class Western in combination with words like *raila, union, emergency* or appearing alone.

4.2 Predictive analysis

In this section we present the classification models on which we measure the accuracy. The evaluation criterion was the percentage of correctly classified instances. For model testing we used 10-fold cross-validation, where nine folds are used for building the classifier, one fold is used for testing, and the average accuracy is computed by repeating this action 10 times.

We made a set of experiments using different feature sets. In the majority of cases the results show quite a low accuracy and a high standard deviation. We only present the most interesting results, so not all the algorithms or feature sets are covered. Results of the experiments are presented in Table 5.

1R selects only the most important rule. Surprisingly, this is one of the best performing algorithms on our data set. We present the result on the word unigrams feature set. For W1 the accuracy is 69.64% and the chosen feature is *Raila*. For interpretation purposes the result on W1-bin is better/simpler, even if the accuracy is lower (61.43%):

Raila:
0 -> LO
1 -> WE

The next set of experiments was done with J48. Among the parameters that we can choose in J48 for decision tree pruning we used the minimal number of objects in the leaves. Without additional feature selection, the best results were obtained by the tree built on the concatenated W1W2 feature set and the minimal number of objects set to 4 (accuracy 58.9% but high standard variation). We can improve the obtained results by first applying a wrapper feature subset selection method. Wrapper approaches are being tuned to the learning algorithm being used.

We chose the same settings for the wrapper feature selection and for classifier learning. The minimum number of objects was set to two. The classifier has achieved 80.56% accuracy (this is the best accuracy achieved in our experiments on the small dataset of 72 documents) and the model is presented in Table 4.

Table 4: Decision tree for the W1W2 feature set, with previous wrapper feature subset selection (accuracy: 80.56%).

```

tribe <= 0
| political/leaders <= 0
| | what <= 0.001479
| | | pledged<= 0
| | | | forces <= 0
| | | | | what <= 0.000516
| | | | | it/had<= 0
| | | | | | union <= 0.002273: LO (30.0/7.0)
| | | | | | union > 0.002273: WE (4.0)
| | | | | | it/had> 0: WE (3.0)
| | | | | what > 0.000516: WE (5.0)
| | | | forces > 0: WE (3.0)
| | | pledged> 0: WE (3.0)
| | what > 0.001479: LO (8.0)
| political/leaders > 0: LO (4.0)
tribe > 0: WE (12.0/1.0)

```

At the first node we find the word *tribe* like in the majority of the above presented examples. The presence of *tribe* results in leaf WE, while its absence leads us to the

Table 5: Results of 10-fold cross-validation model testing.

Algorithms	One R	J48-M2	J48-M4	JRip-N2	JRip-N4	.PART '-M 2	.PART '-M 4
Feature set	Classification accuracy (\pm st.dev.)						
W1	69.64(16.60)	47.86(25.83)	43.75(17.84)	66.96(25.49)	65.18(21.58)	36.25(18.82)	42.32(18.48)
W1_BIN	61.43(13.13)	37.14(18.71)	47.32(15.37)	59.11(24.66)	59.11(24.66)	31.96(16.42)	57.50(21.46)
W2	68.04(11.56)	27.86(13.15)	33.57(18.09)	60.00(19.76)	24.11(15.27)	60.00(19.76)	38.93(14.79)
W2_BIN	63.93(11.69)	29.64(17.48)	49.82(12.38)	62.50(13.36)	63.93(11.69)	24.82(14.93)	36.25(19.38)
W1W2	62.32(12.04)	36.96(21.72)	58.93(23.88)	56.96(19.28)	57.14(22.08)	31.25(14.99)	51.79(18.39)
W1W2_wrapperJ48	64.11(15.80)	80.54(11.92)	70.71(12.69)	50.36(22.87)	51.79(23.76)	71.07(16.01)	70.71(12.69)
W1W2_wrapperPART	69.82(15.28)	70.77(15.35)	70.59(14.36)	60.48(16.42)	62.93(16.62)	72.36(14.28)	70.89(13.74)

second node with an interesting word bigram (*political leaders*). The presence of this bigram is the indicator for a Kenyan newspaper.

Another model (presented in Table 6) was obtained with PART with previous wrapper feature selection.

Table 6: PART decision list for the W1W2 feature set, with previous wrapper feature subset selection (accuracy: 86.1%).

```

tribe <= 0 AND real<= 0.001582 AND running <= 0 AND emergency <=
0.000557 AND and/the > 0 AND and/the <= 0.003215: WE (24.0/4.0)
tribe <= 0: LO (36.0/5.0)
: WE (12.0/1.0)

```

5 CONCLUSIONS AND FUTURE WORK

We can see that given a 50%-50% prior class distribution, the used machine learning algorithms do not result in very high classification accuracy on a small dataset of 72 articles. As indicated by the initial experiments on a larger data set (464 instances, accuracy around 90%), low accuracy is mainly due to a small data set available for this study. Nevertheless, machine learning algorithms do provide new insights leading to improved understanding of lexical choices.

We utilized data mining techniques in order to find out which word unigrams and word bigrams can be interpreted as the distinguishing words between the *Daily Nation* Kenyan newspaper and *The Independent*. We observed that better results are obtained on numeric feature sets (based on term frequency) than on binary ones. However, binary representation enables much easier interpretation. The best results were obtained when classification was combined with wrapper feature selection.

From the content point of view, the main selected feature was *tribe*, that appears in Western newspaper *The Independent* and not in Kenyan *Daily Nation*. This is an important finding, because *tribe* is not an ideologically neutral term, but is frequently used for stereotyping the African situation (Ray, 2008) and promoting “a myth of primitive African timelessness” (Lowe et al. 2008).

In further work we will repeat and extend the experiments by analysing 232 articles for each class (local and Western) from six different newspapers (where initial experiments indicate substantially increased accuracies of induced models). We plan to experiment also with lemmatized or stemmed feature sets, with previous stopword removing and using also syntactic features.

ACKNOWLEDGMENTS. The data was obtained from the Antwerp Center of Pragmatics, University of Antwerp.

The corpus was collected by L. Michiels and R. Coesemans as a part of the project Intertextuality and Flows of Information, led by Prof. J. Verschueren. I am grateful to my supervisor, prof. W. Daelemans, as well as to Kim Luyckx for her help in transforming the data into feature vector format using system TACTiCS, developed at CNTS.

REFERENCES

- Cendrowska, J. (1987). PRISM: An algorithm for inducing modular rules. *International Journal of Man-Machine Studies*, 27(4), p. 349-370.
- Cohen, W. (1995). Fast effective rule induction. In *Proceedings of the 12th International Conference on Machine Learning*, p.115–123.
- Fortuna, B., Galleguillos, C. and Cristianini, N. (2008). Detecting the bias in media with statistical methods. In *Text Mining: Theory and Applications*, Taylor and Francis Publisher, London.
- Frank, E. and Witten, I. H. (1998). Generating accurate rule sets without global optimization. In *Proceedings of the 15th International Conference on Machine Learning*, p. 144-151.
- Holte, R.C. (1993). Very simple classification rules perform well on most commonly used datasets. *Machine Learning*, 6 (11), p. 63-91.
- Lowe C., Brimah, T., Marsh P.-A., Minter, W. and Muyangwa, M. (1997, updated 2008). Talking about "tribe" moving from stereotypes to analysis. http://www.africaaction.org/bp/documents/TalkingaboutTribeFeb2008Update_001.pdf, last accessed 18. 8. 2009.
- Quinlan, J. (1986). Induction of decision trees. *Machine Learning*, 1, p. 81–106.
- Quinlan, J. (1993). *C4.5: Programs for Machine Learning*, Morgan Kaufmann, San Francisco.
- Ray, C. (2008). How the word 'tribe' stereotypes Africa. *New African* 471. p. 8-9.
- Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys*, 34 (1), p. 1-47.
- Scheffer, T. (2001). Finding association rules that trade support optimally against confidence. In *Proceedings of the 5th European Conference on Principles of Data Mining and Knowledge Discovery*, p. 424–435.
- Verschueren, J. (1999). *Understanding Pragmatics*, Arnold (Understanding Language Series), London.
- Witten, I. H. and Frank, E. (2005). *Data Mining Practical Machine Learning Tools and Techniques*, 2nd ed., Elsevier, San Francisco.

TEXT MINING AND KNOWLEDGE DISCOVERY WITH ONTOGEN 2.0

Mladen Tomaško

Police Academy, Police College and Jožef Stefan International postgraduate school

e-mail: mladen.tomasko@gmail.com

ABSTRACT

This paper presents a text mining technique used for extracting knowledge from two databases containing the same set of documents. The first database contains only abstracts, while the second one contains full text documents. Data preprocessing and processing is described, followed by a description of problems and encountered doubts. Special attention was given to the comparison of results taken from these databases. Some remarks on the tool used for data mining were given and for a conclusion some future researches in this field were proposed.

1 INTRODUCTION

In the last decade or two, a large amount of data was collected all over the world. Handling it is almost an impossible task. When we search for some specific topic, we are lost because we find numerous irrelevant documents (“data deluge”).

Text mining technique can help us to extract important data more efficiently. This tool sorts our documents by searching similar words or phrases in documents. It also gives us a possibility to interfere on search process and to actively improve accuracy. Text mining is a relatively “young” technique; it is a part of data mining or – in wider context - a part of data analysis. The roots of this tool lie in researching learning computer to understand ontology, documents concept and implementation of this for document classification, document clustering and document visualization. It is extremely useful for web crawling and extracting information we are interested in.

There are numerous tools in the market for text mining, some specialized for certain fields (e. g. medicine, law etc.) and some more general. The main point of good text mining tools is concept understanding and most efforts today are dedicated to improve that property.

Spending less time searching for relevant information gives us more time for researching and working. A short overview how data mining can help us is explained in this article through a practical example.

2 TEXT MINING

2.1 Data description

For practical work we had two different text files, properly organized and prepared for deeper analysis. Preparing and collecting data can take a while and can require some special skills, but it is an important step to obtain good final results.

We had two datasets; one with only short description (abstract) of documents and the other with full text. These two files give us a good base for explaining how efficient text mining will be. Since we are searching for new knowledge in these files, we will use different methods to make our work fast and accurate.

There are 4571 documents about chemistry. Each document is represented with title and short abstract (in the first file) or with title and full text (in the second file). The first file has about 670 Kb and the larger one has about 12,5 Mb. For the real world these are relatively small databases, but representative enough that the result could be generalized.

2.2 Data preparation

Data were taken from different resources, mainly from the web. The use of different characters coding tables brought about inconsistencies in the final file. The first task was to eliminate these characters, which follows the wrong documents classification.

Then we tried to start with data mining process, but there was another type of noise. In numerous documents there were messages from web pages about the incompatibility between browser and frames these web sites use.

To avoid this problem we prepared a set of words and add it to the stop word list. The result was much better concept suggestions.

2.3 Tool used for text mining

For text mining we used OntoGen 2.0 tool [1, 3], that is a semiautomatic tool which allows us ontology and concept visualization and a lot of ways to improve concept definitions. Despite some initial difficulties described in details later, the overall user experience using OntoGen was fair.

information to accept right decision with which database to go further. One another thing is also obvious. In larger database documents are just in size of normal abstract and not like ordinary articles or documents which are usually at least four or five pages long. This can follow us in trap to take a wrong conclusion because of insufficient information.

After first steps taxonomy trees look like on the two figures below. The first one is much better defined due to sufficient text in the documents and data mining programme has enough information to do a good classification during the first step.

With appropriate methods we can improve similarities and accuracy in our concept.

2.5 Improving accuracy and similarity of concept

In OntoGen we can work in supervised and unsupervised way. Supervised means that we exactly know how our data must be organized. We used unsupervised way, but we tried to improve accuracy (similarity) in certain branches. We can join branches together, divide it into more sub-branches or just exclude some documents that are not directly connected with our concept. Another possibility is using sim graph in Concept's documents tab to include or exclude documents. Using all this we finished with a relatively small taxonomy tree with branches that can be divided into more sub-branches if necessary.

Figure 4 shows the final tree is shown. Branches have good similarity and after random checking and evaluation we realized that accuracy is good enough. Working this way is time consuming (because of user interaction), but in the end we get far better results than results given by programme itself.

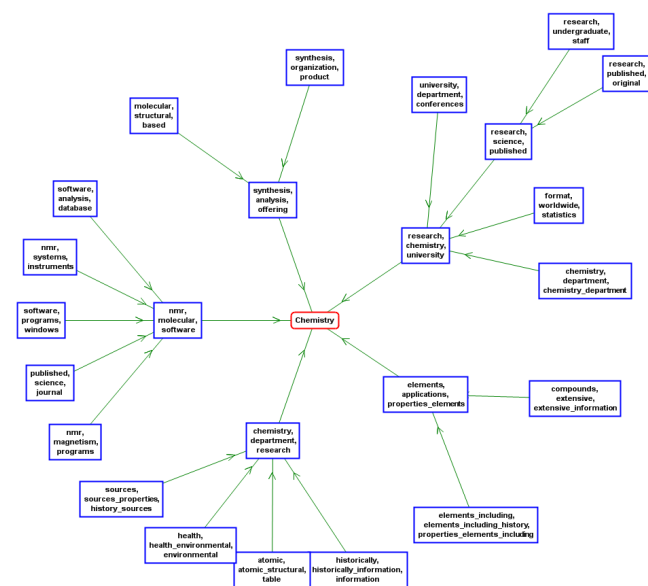


Figure 4: Ontology visualization (abstracts)

If we compare this tree with the next one in the Figure 5. we can see they are very similar despite the fact that the first one is produced from abstracts and the second one from full text documents.

As mentioned before, the main difference is the amount of time needed to obtain similar results. With full text documents with no intensive human work we can get almost the same results. However, for large databases it is important that we can obtain such abstract from documents. That can save us a lot of computing time and simplify building of a taxonomy tree.

If we need a model with high accuracy, help of an expert is needed. He can help us to improve our ontology concept manually. In our example with reasonably small database, it was easier to experiment with data. On large databases a proper strategy is needed which saves time and money.

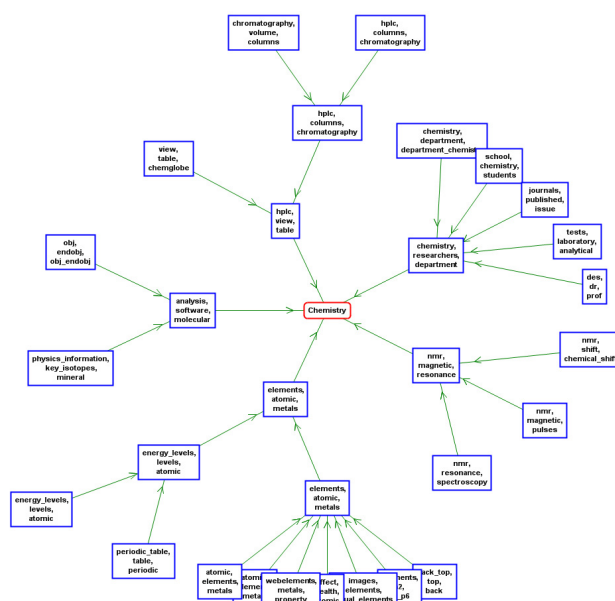


Figure 5: Ontology visualization (full text)

The tree in Figure 5 shows the selected final ontology. Accuracy is good enough in almost all branches, whole process to obtain this result was straightforward and no special knowledge about chemistry was needed. There is enough difference between branches and similarity inside branch is good.

We didn't change names what programme proposed, even though calculation of SVM keywords gives slightly different results. We noticed that if user manually change the name of node then all words are in the same line and there is no visible way how change it. That follows in unclear graph and we decided to stay with programme's proposed names. The hardest part of building tree is proper division of nodes who have relatively low similarity. This part usually demands lot of manual work.

As an example how visualization of concept with good similarity looks, there is one of node from our taxonomy tree in Figure 6.

An implementation of the Pathfinder algorithm for sparse networks and its application on text networks

Anže Vavpetič

Faculty of Computer and Information Science,

Vladimir Batagelj

Faculty of Mathematics and Physics, Department of Mathematics,

Vid Podpečan

Jožef Stefan Institute

Ljubljana

Abstract

The Pathfinder algorithm is extensively used for pruning weighted networks. It is particularly useful in the analysis of co-citation networks. The present paper reviews two versions of the algorithm. A version with improved time and space complexities, called the Binary Pathfinder, and a version optimized for sparse networks, called Sparse Pathfinder. These two algorithms were implemented and tested on text networks. The obtained results show that the algorithm optimized for sparse networks works notably faster on such data.

1 Introduction

In larger (at least some hundreds of nodes) weighted networks the visual inspection can't be used anymore for identifying essential parts of the network. An approach to this problem are the pruning algorithms. They are used to remove less significant links, allowing the more salient links to be found. An example of a network pruning algorithm is the Pathfinder algorithm, developed in cognitive science to determine the most important links in a network [6]. The output defined by the Pathfinder algorithm is known as a Pathfinder network or PFnet. Initially, Pathfinder networks were used exclusively to represent relationships between concepts or keywords, but later works have extended its use to many other fields of application, for example co-citation networks.

1.1 Basic idea of the Pathfinder algorithm

Let $\mathcal{N} = (V, E, w)$ be a network. V is the set of nodes, E is the set of links, and $w : E \rightarrow \mathbb{R}_0^+$ the weight. We denote $n = |V|$ and $m = |E|$. Assuming that the weight represents a distance, the pruning idea of the Pathfinder algorithm is based on the triangle inequality, which states that the direct distance between two points must be less than or equal to the distance between those two points going through an intermediate point. It can be easily extended to all paths: the direct distance between two nodes must be less than or equal to the dist-length (sum of all

weights) of every path of between these two nodes; therefore also less than or equal to the length of geodesic path (i.e. the shortest path). The algorithm eliminates the links which violate the extended triangle inequality.

The Pathfinder algorithm preserves the geodesic distance between every pair of nodes, while simplifying the network and so clarifying it for the subsequent analysis [7].

The extended triangle inequality gives rise to the first parameter of the algorithm, i.e. the link-length – maximum number of intermediate links that will be considered, usually named q . Since the link-length of geodesic can not exceed $n - 1$ the maximum possible value of q is $n - 1$.

To calculate the distance between two nodes along a path the Pathfinder algorithm uses the Minkowski operation

$$a \odot_r b = \sqrt[r]{a^r + b^r} \quad (1)$$

r is the second parameter of the Pathfinder procedure. For different values of r we get: $a \odot_1 b = a + b$; $a \odot_2 b = \sqrt{a^2 + b^2}$; $a \odot_\infty b = \max(a, b)$. The operation \odot is associative. This means that for a path π with links with weights w_1, w_2, \dots, w_k , we calculate its length as $d(\pi) = w_1 \odot w_2 \odot \dots \odot w_k$.

2 Implementation

2.1 The Binary Pathfinder

The first version of the Pathfinder algorithm considered in this paper is the Binary Pathfinder. The original Pathfinder has space complexity $O(qn^2)$ and time complexity $O(qn^3)$. The Binary Pathfinder [5] is an improvement to the original algorithm, reducing its time complexity to $O(n^3 \log q)$ and the space complexity to $O(n^2)$.

Before we describe the Binary Pathfinder, we list some definitions used by [5] and [6]:

- The Pathfinder network $\text{PFnet}(r, q) = (V, E, w)$ is a sub-network of network \mathcal{N} .
- The weight of the link from node u to node v is denoted by w_{uv} . They are collected in a $n \times n$ matrix W .

- $W^{i+1} = W \odot W^i$ is computed as follows:

$$w_{uv}^{i+1} = \min\{w_{ut} \odot w_{tv}^i : t \in V\}$$

- The minimum-distance matrix for paths not exceeding i links is denoted D^i and its elements are computed as follows:

$$d_{uv}^i = \min(w_{uv}^1, w_{uv}^2, \dots, w_{uv}^i) \text{ for } u \neq v$$

$$\text{and } d_{vv}^i = 0.$$

The crux of the improvements to the original Pathfinder made by [5] lies in the calculation of matrices D^i . They have pointed out that for determining the PFnet we only need the matrix D^q for the comparison with the initial weight matrix. It is unnecessary to generate all of the D^i s. They have shown that $D^{i+j} = D^i \odot D^j$:

$$d_{uv}^{i+j} = \min\{d_{uv}^i, d_{uv}^j, w_{ut}^i \odot w_{tv}^j : t \in V\}$$

where $d_{uv}^1 = w_{uv}$. Thus we can make larger steps in computing the distance matrices. The Binary Pathfinder algorithm is based on matrices $D^1, D^2, D^4, D^8, \dots$ [5]:

$$i = 1; nq = 0$$

$$\text{Generate } D^1 = W; D^q = \infty$$

$$\text{If } (q \bmod 2) = 1 :$$

$$\text{Compute } D^q = D^q \odot D^1$$

$$nq = 1$$

$$\text{While } (2 * i) \leq q :$$

$$\text{Compute } D^{2i} = D^i \odot D^i$$

$$\text{If } ((q - nq) \bmod (4 * i)) > 0 :$$

$$\text{Compute } D^q = D^q \odot D^{2i}$$

$$nq = nq + (2 * i)$$

$$i = 2 * i$$

Comparing elements of D^q and W , wherever $d_{uv} = w_{uv}$, add (u, v) as a link to the PFnet.

The improvement in time complexity seems only minute, but as larger networks are dealt with, the difference grows immensely, as can be seen in the results section of [5].

2.2 The Pathfinder optimized for sparse networks

The second algorithm tested by this study is the Pathfinder procedure optimized for sparse networks. It is based on the idea presented in [1] by Vladimir Batagelj. The motivation for this algorithm is the fact that most real life networks are sparse and that in sparse networks the matrix D^q can be computed faster using an adapted version of Dijkstra's algorithm [3] (for $q = n - 1$) or an adapted breadth-first search (BFS) algorithm (for $q < n - 1$). To speed-up the algorithm we represent the network with a graph data structure using adjacency lists, where each node has a list of its neighbor nodes, resulting in a neighbor retrieval query with a time complexity of $O(1)$. To efficiently calculate the matrix D^q for $q = n - 1$ we run the Dijkstra's algorithm based on Minkowski operation once for every

node as a source node, thus producing a corresponding row of the resulting matrix. For $q < n - 1$ the Dijkstra's algorithm can not be easily adapted. We replace it with an adapted shortest path algorithm based on BFS search. The last step of this algorithm is the same as in other implementations of the Pathfinder algorithm, i.e. comparing the elements of D^q and W and wherever $d_{uv} = w_{uv}$, we add (u, v) as a link in the resulting PFnet.

2.2.1 Specifics

As mentioned, the Dijkstra's algorithm is modified so that in calculating dist-lengths, we use Minkowski's operation \odot instead of the addition. With the BFS algorithm we (1) limit the search depth to q and (2) skip all the paths starting from node v that would yield a path length $d > \max\{w(v, u) : u \text{ is a successor of } v\}$.

Sparse Pathfinder($r, q, \mathcal{N} = \langle V, E, w \rangle$) :

If $q \geq n - 1$: Dijkstra(r, q, \mathcal{N})

Else: BFS(r, q, \mathcal{N})

Comparing elements of D^q and W , wherever $D^q[u, v] = w_{uv}$, add (u, v) as a link to the PFnet.

End Sparse Pathfinder

Dijkstra($r, q, \mathcal{N} = \langle V, E, w \rangle$) :

$pq := \{\} \dots$ priority queue sorted by distance

For each $v \in V$:

For each $u \in V$: $dist[u] := \infty$; Mark u as unvisited

$dist[v] := 0$;

Mark v as visited and insert it into pq .

While $pq \neq \{\}$:

$t := \min(pq)$; delete_min(pq)

For each successor z of t :

$new_dist := dist[t] \odot_r w_{tz}$

If z not visited:

$dist[z] := new_dist$

Mark z as visited and insert it into pq .

Else if $new_dist < dist[z]$:

decrease_key(pq, z, new_dist)

For each $u \in V$: $D^q[v, u] := dist[u]$

End Dijkstra

BFS($r, q, \mathcal{N} = \langle V, E, w \rangle$) :

$Q := \{\} \dots$ FIFO queue

For each $v \in V$:

For each $u \in V$: $dist[u] := \infty$

$dMax := \max\{w_{vu} : u \text{ is a successor of } v\}$

putLast($Q, v, 0, 0$)

$dist[v] := 0$

While $p \neq \{\}$:

$(u, d, l) := \text{firstFrom}(Q)$

$l := l + 1$

For each neighbor t of u :

$new_dist := d \odot_r w_{ut}$

If $new_dist \leq dMax$ and $new_dist < dist[t]$:

$dist[t] := new_dist$

If $l < q$: putLast(Q, t, new_dist, l)

For each $u \in V : D^q[v, u] := dist[u]$

End BFS

In order to improve the efficiency of Dijkstra, the priority queue was implemented as a minimum binary heap with the following time complexities: $O(\log n)$ for insertion, deleting the minimum element, decreasing a key, and $O(1)$ for testing whether the priority queue is empty.

For such an implementation the time complexity of Dijkstra's algorithm is $O(m \log n + n \log n)$, which is dominated by $(m \log n)$, giving the total time complexity of the Sparse Pathfinder algorithm of $O(nm \log n)$. For dense networks, $m = O(n^2)$, we get the same time complexity as for the Binary Pathfinder. The space complexity remains the same as for the Binary Pathfinder, $O(n^2)$. In theory the complexity could be further improved to $O(nm)$ by using Fibonacci heap. But we considered the 'Bottom line' of [8, slide 26] which says: Fibonacci heap is best in theory, but not worth implementing.

The BFS algorithm makes a complete search of all possible paths originating in node v , of link-length at most q , and dist-length at most $dMax$. Its efficiency strongly depends on the properties of network (average degree, distribution of weights, parameter r) and it is very difficult to analyze analytically.

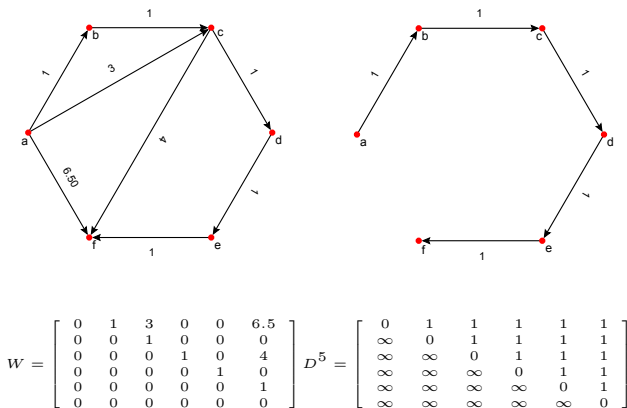


Figure 1: A graph before and after we apply the Pathfinder procedure ($r = \infty, q = 5$) with the initial weight matrix W and the target matrix D^5

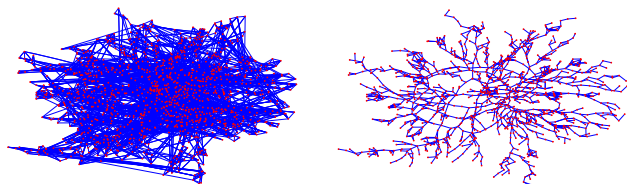


Figure 2: Network of terms generated from 100 documents and its PFnet ($r = \infty, q = n - 1$)

3 Experimental results

Both algorithms were implemented in C++ compiled with gcc 4.3.2 and integrated into the machine learning and data mining framework Orange [4] and can thus be easily imported as a Python module. The tests were done on a Intel Core 2 Duo 3GHz machine with 2GB of RAM running Ubuntu Linux. Input files were given in Pajek's format [2]; output networks were written back in the same format as well, and finally also drawn with Pajek.

The given input graphs can be interpreted as follows. Each link represents a co-occurrence, which means that both words appear together in at least one document, whereas the link's weight represents the normalized number of co-occurrences through all documents. This weight is a similarity measure. In order to transform the weights into dissimilarities required by Pathfinder, we applied the formula $w' = \frac{1}{w}$, where w is the original weight.

First we present the comparison of the performance of both algorithms on the input graphs. The statistics gathered can be seen in Tables 1 and 2. On these networks, the optimized Pathfinder algorithm works much faster than the Binary Pathfinder. The reason for this is that m is much smaller than n^2 . In Table 1 the parameters r and q were set to $r = \infty$ and $q = n - 1$ as these values are most commonly used in practice; in Table 2 we provide some time measurements for smaller values of q , where the adapted BFS algorithm is used. We observe that the Sparse Pathfinder can be efficiently applied on such networks, as its computing time increases much more slowly than the Binary Pathfinder's, thus allowing to produce PFnets from larger networks in a reasonable time frame.

Figure 1 illustrates how the network changes when the Pathfinder procedure is applied. Figure 2 presents the effect of the Pathfinder on the network. On the left side the original network is presented, and on the right side the corresponding PFnet.

4 Conclusion

As noted by [5] the Pathfinder networks are of great interest in the study of different types of weighted networks. They are found to be particularly useful in scientometrics in studying advancing frontiers of research, disciplines, profiles of authors, etc.

Since the original algorithm has severe practical limitations, rising from its time and space complexity, we have implemented two improved versions: the Binary Pathfinder, a version presented in [5], and an algorithm optimized for sparse networks.

Both algorithms have been applied to several text networks, generated from a various number of documents in order to check the applicability to such networks. We have found that especially the algorithm for sparse networks has potential to be used for pruning of such networks, as its computational time rises much slower than the Binary Pathfinder's with the number of nodes and links, thus allowing to produce PFnets from

Input (with $r = \infty$ and $q = n - 1$)			Binary PF	Sparse PF	Output
Network	n	m	$t(s)$	$t(s)$	n
stem+cell_10docs.net	64	127	0.187	0.006	71
epilepsy+migraine_50docs.net	517	1115	56.951	1.278	536
stem+cell_100docs.net	1215	2828	1035.302	17.606	1419
epilepsy+migraine_100docs.net	1322	3021	1183.712	21.145	1480
migraine+protein_100docs.net	1322	3021	1200.313	21.168	1480
2ksparse.net	2622	5244	11069.679	75.938	2674
2kdense.net	2622	39330	10913.843	390.118	3800
5ksparse.net	5355	10710	$\approx 97,000$ ($\approx 27h$)	593.982	5277
5kdense.net	5355	80325	$\approx 97,000$ ($\approx 27h$)	3425.378	6919

Table 1: Algorithm performance, $q = n - 1$ / Dijkstra

Network	Binary PF $t(s)$			Sparse PF $t(s)$		
	$q = 3$	$q = 5$	$q = 10$	$q = 3$	$q = 5$	$q = 10$
$r = \infty$						
stem+cell_10docs.net	0.064	0.088	0.103	0.001	0.002	0.006
epilepsy+migraine_50docs.net	18.841	27.486	33.784	0.052	0.136	1.524
stem+cell_100docs.net	217.067	299.909	366.942	0.410	0.988	25.963
epilepsy+migraine_100docs.net	288.932	376.673	491.588	0.369	1.269	21.753
migraine+protein_100docs.net	302.332	395.689	486.578	0.342	1.142	20.046
2ksparse.net	1953.135	2512.451	3249.126	2.264	22.020	86.208

Table 2: Algorithm performance, small q / BFS

larger inputs in reasonable time.

Acknowledgments

The work presented in this paper was supported by the Slovenian Research Agency grant Knowledge Technologies, and by the grant of the European Commission under the 7th Framework Programme FP7-ICT-2007-C FET-Open, contract no. BISON-211898.

References

- [1] V. Batagelj. Fast pathfinder algorithm for large sparse networks (unpublished). Notes of the talk presented at the 1172-th Sredin seminar, Ljubljana, February 11, 2009.
- [2] V. Batagelj and A. Mrvar. Pajek – analysis and visualization of large networks. In M. Jünger and P. Mutzel, editors, *Graph Drawing Software*. Springer, 2003.
- [3] T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms*. McGraw-Hill, 2001.
- [4] J. Demšar, B. Zupan, and G. Leban. Orange: From experimental machine learning to interactive data mining. *White Paper*, 2004.
- [5] V. P. Guerrero-Bote, F. Zapico-Alonso, M. E. Espinosa-Calvo, R. G. Crisóstomo, and F. de Moya-Anegón. Binary pathfinder: An improvement to the pathfinder algorithm. *Information Processing & Management*, 2006.
- [6] R. W. Schvaneveldt. Pathfinder associative networks. *Norwood, NJ: Ablex*, 1990.
- [7] R. W. Schvaneveldt, D. W. Dearholt, and F.T. Durso. Graph theoretic foundations of pathfinder networks. *Computers and Mathematics with Applications*, 1988.
- [8] R. Sedgewick and K. Wayne. Lectures 15: Shortest paths, 2009. <http://www.cs.princeton.edu/courses/archive/spr09/cos226/lectures/>.

EXPERIMENTS WITH SATURATION FILTERING FOR NOISE ELIMINATION FROM LABELED DATA

Borut Sluban (1), Nada Lavrač (1), Dragan Gamberger (2), Andrej Bauer (3)

(1) Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia

(2) Rudjer Bošković Institute, Bijenička 54, 10000 Zagreb, Croatia

(3) Faculty of Mathematics and Physics, Jadranska 21, 1000 Ljubljana, Slovenia

e-mail: borut.sluban@gmail.com

ABSTRACT

For classification tasks, data filtering is aimed at improving prediction accuracy. This paper presents some upgrades of the existing saturation filter, which is used for the elimination of noisy examples from labeled data. The main contribution of this paper is the exhaustive experimental evaluation of saturation filtering. The filter was applied on 12 different datasets and its performance was tested with four different learning algorithms. The evaluation showed the improvement of prediction accuracy on most of the datasets and statistical testing proved that the performance of a learning algorithm was significantly better with the use of the saturation filter than without it.

1 INTRODUCTION

In predictive data mining learning methods are used to induce models or theories from class-labeled data. The induced models are used for classification or prediction. In a classification task, data is usually formed from examples which are labeled by the class to which they belong. The task is to find a model that will enable a newly encountered instance to be classified.

A common problem affecting the prediction accuracy of an induced model is noise in the training data. Noise usually means outliers and random errors in training examples (erroneous attribute values and/or erroneous classification). Therefore appropriate noise handling procedures are essential to ensure the prediction accuracy and applicability of induced models.

One of the approaches to noise handling is filtering of noisy examples from the training dataset. In this way the model induced from the filtered data will be less complex and more accurate when classifying unseen examples. The upgrade of a filtering algorithm, as well as the experimental work presented in this paper are based on the *saturation filter* introduced in [3].

The paper is structured as follows. In Section 2 the theoretical background and the idea of the saturation filter is briefly described. Section 3 describes the reimplementation of the saturation filter. An overview of the testing results is given in Section 4. The paper concludes with some

directions for further work in Section 5.

2 SATURATION FILTER

The idea of the saturation filter, introduced in [3], follows the Occam's razor principle, which suggests that among all the hypotheses (models) that are correct for all the training examples one should select the simplest hypothesis. This principle is implicitly used for noise elimination, since by choosing a simpler hypothesis we try to avoid overfitting a noisy training dataset.

A *target concept* of a given problem domain is defined as the source of all possible correct examples of the concept. The task of inductive learning is to find a good representation of the target concept in a selected hypothesis language. This representation is called a *target theory*. By following the Occam's razor principle and selecting the simplest hypothesis which correctly represents the target concept we get the so-called *target hypothesis*. Elimination of noisy examples from the training set helps in the induction of the target hypothesis, since a hypothesis induced from noiseless data will be less complex and more accurate when classifying unseen examples.

The name for the filter was derived from the saturation property of training data, meaning that a training dataset is saturated, if it can be used to induce a target theory. The approach to obtain a saturated training set was based on the observation that the elimination of noisy examples, in contrast to the elimination of examples for which the target theory is correct, reduces the *CLCH* value of the training set, where *CLCH* stands for the *Complexity of the Least Complex correct Hypothesis*.

Suppose that a complexity measure c for a hypothesis language is defined and that for any hypothesis H its complexity $c(H)$ can be determined. Then for a training set E one can determine the complexity of the least complex hypothesis which is correct for all the examples in E : this complexity, denoted by $g(E)$ is called the *CLCH* value.

In [3] it was shown that if E is noiseless and saturated (containing enough training examples to be able to induce a correct target hypothesis from it), then – if a noisy example e_n is added to E – it follows that $g(E) < g(E_n)$, where $E_n = E \cup \{e_n\}$ and e_n is a noisy example for which the target hypothesis is not correct. The property $g(E) < g(E_n)$ means

that a noisy example in E_n can be detected as the one that enables *CLCH* value reduction. The approach in an iterative form is applicable also when E_n includes more than one noisy example.

In this way noisy examples can be detected and eliminated from the training set. But it must be noted that although the saturation property of a training set is the main theoretical condition for the mentioned filter, in practice it is usually not possible to satisfy the saturation condition. Despite that, the filtering algorithm is still applicable since there still may be enough training examples that form a saturated subset for some subconcept of the target concept.

3 IMPLEMENTATION

The reimplementaion of the saturation filter and the testing of its performance were made with the help of the open source software for data mining called *Orange* [2]. It is software for data mining through visual programming or Python scripting, it includes many modules and tools for data preprocessing, modeling and knowledge discovery in databases.

For the elimination of noisy training examples with the saturation filter we had to choose a complexity measure that would distinguish between models induced from training data. The idea was to build an *unpruned* decision tree with machine learning tools available in Orange. Since an unpruned decision tree is a classification model which is correct for all the training examples, we chose the number of nodes of the unpruned decision tree as the complexity measure of the model.

To construct the saturation filter we needed two methods. The first one was a *saturation test*. At first it computes the complexity of the classification model for the given training set, then in each step it excludes only one training example and computes the complexity of a classification model induced from the rest of the training examples. If the complexity of the new model is smaller than the one computed in the beginning, then the excluded example is marked as potentially noisy. However at this point it is not

yet finally excluded, it is returned to the training set and the same procedure is repeated for each training example. The examples which had the most effect in reducing the complexity of the classification model with their exclusion are labeled as the *most noisy* and are passed on to the second method. If there is no such example, then the training set is considered to be saturated.

The second method, the *filter*, randomly chooses one among the most noisy examples and finally excludes it from the training dataset, while the others are returned back. This is repeated as long as the saturation test finds noisy examples, meaning that a saturated subset has not yet been obtained.

While testing the filter on different datasets, in some cases the number of examples suggested as being noisy by the saturation test was quite high (relative to the size of the training set). In this case the filter excluded them all one by one. To avoid excluding a subset of examples that represent a small subconcept of the target concept, an addition to the saturation filter was introduced as the tolerance level parameter t . If the parameter is specified the filtering process stops when the size of the set of potentially noisy instances in an iteration of the saturation test exceeds the given percentage t of the size of the training set.

4 EVALUATION OF RESULTS

The experimental work done with the previously described reimplementaion of the saturation filter and the results obtained from the testing are presented in this section.

4.1 Datasets used for testing

The datasets used for testing the performance of the saturation filter were obtained from the Orange website and are mostly from the *UCI Machine Learning Repository* [1]. For simplicity only 2-class datasets were chosen. The first eight datasets are real-life datasets containing noise, whereas the last four are artificially generated (precise representations of a concept or a set of all possible configurations of a concept) and do not contain any noise (Table 1).

Dataset	No. of examples	No. of attributes	Class ratio	Description
Breast Cancer LJ	286	9	70:30	Breast cancer of patients in Ljubljana, 1988.
Breast Cancer WI	683	9	65:35	Breast cancer of patients in Wisconsin, USA, 1991.
BUPA	365	6	58:42	Liver disorder, male patients.
Credit Approval	690	15	56:44	Approval of credit applications.
Heart Disease	303	13	54:46	Presence of heart disease in the patient.
MONK-3	554	6	52:48	Target concept: ($A_5 = 3$ and $A_4 = 1$) or ($A_5 \neq 4$ and $A_2 \neq 3$).
SA - Heart	462	9	65:35	Coronary heart disease, male patients, SAR
Voting	435	16	61:39	Congressional voting records. USA, 1984.
MONK-1	556	6	50:50	Target concept: $A_1 = A_2$ or $A_5 = 1$.
MONK-2	601	6	66:34	Target concept: exactly two of the attributes have the value 1.
Promoters	106	57	50:50	Promoter DNA sequences of the bacteria E. Coli.
Tic Tac Toe	985	9	65:35	All possible endgame configurations of the game Tic Tac Toe.

Table 1: Datasets used for testing the performance of the saturation filter.

4.2 Algorithms used in the experimental evaluation

For testing the performance of the saturation filter, four learning algorithms from the Orange library were used for the induction of classification models. These are:

- Decision tree learner (unpruned)
- Decision tree learner (pruned)
- Naïve Bayes classifier
- Rule learning algorithm (CN2)

4.3 Classification accuracies of the basic saturation filter

Classification accuracies were computed using 10-fold cross-validation for all the four classifiers induced from non-filtered and filtered datasets.

The testing results are presented in Table 2. Among the first eight datasets filtering with the saturation filter showed to achieve better classification results on:

- four datasets with all four learning algorithms
- one dataset with three learning algorithms
- two dataset with two learning algorithms
- one dataset with one learning algorithm

However on two domains (Breast cancer LJ and SA-heart) no learning algorithm except the Naïve Bayes classifier managed to outperform the default classifier.

The classification accuracies on the last four datasets which were artificially generated, were expected to be worse if filtering were applied (since removing non-noisy examples would result in the induction of a less accurate classification model). These expectations were fully confirmed on two datasets, but on the other two dataset there were (a bit

surprisingly) some improvements with one or two learning algorithms. However, the Naïve Bayes classifier did not reach the classification accuracy of the default classifier on the MONK-2 dataset.

It is also interesting to compare the third and the fourth column in Table 2, where we can see that the combination of the saturation filter and the unpruned decision tree learner performs comparable to (in some cases even better than) the decision tree learner which has its own built-in noise handling procedure, called pruning. However, according to statistical *t-test*, the differences are due to large standard deviations in most cases insignificant.

4.4 Improvements with tolerance level parameter t

In addition to the reimplementing of the original saturation filter described in [3], we have proposed its improvement using tolerance level parameter t , whose goal is to prevent the exclusion of certain subsets which could possibly represent a small subconcept of the target concept. The application of the parameter makes sense only in the case of datasets where the saturation test seems to find “larger” sets of potentially noisy examples.

Parameter values used for testing were 0.015, 0.02, 0.025 and 0.03, which means that the filtering process stops if the size of the set of potentially noisy examples in an iteration of the saturation test exceeds the given percentage t of the size of the training set. In Table 3 the improved classification accuracies are shown in bold. The values listed are the highest classification accuracies achieved, mostly achieved by setting t value to 0.015 or 0.02.

Dataset	Unpruned decision tree		Pruned decision tree		Naïve Bayes Classifier		Rule learning alg. (CN2)	
	non-filtered	filtered	non-filtered	filtered	non-filtered	filtered	non-filtered	filtered
BreastCancerLJ	61.9 (±11.4)	67.1 (±6.74)	67.2 (±9.36)	67.1 (±6.74)	72.7 (±3.06)	73.1 (±4.81)	67.9 (±6.05)	67.2 (±5.11)
BreastCancerWI	94.3 (±2.99)	93.8 (±2.78)	94.3 (±2.98)	93.7 (±2.64)	96.8 (±2.16)	96.5 (±1.89)	94.7 (±2.10)	95.9 (±2.16)
BUPA	60.5 (±6.54)	63.4 (±6.80)	60.5 (±6.54)	63.4 (±6.80)	67.0 (±5.76)	67.3 (±6.70)	60.5 (±9.45)	66.1 (±6.32)
Credit Approval	83.8 (±3.60)	85.1 (±2.83)	85.4 (±3.07)	85.5 (±3.11)	85.7 (±2.78)	85.8 (±3.09)	75.2 (±3.46)	73.8 (±7.05)
Heart Disease	65.4 (±7.13)	70.0 (±10.8)	65.7 (±7.56)	69.3 (±10.9)	82.8 (±8.30)	83.2 (±7.48)	68.6 (±11.5)	69.3 (±8.02)
MONK-3	96.6 (±2.84)	97.8 (±2.26)	98.9 (±2.18)	98.9 (±2.18)	96.4 (±3.04)	96.4 (±3.04)	88.6 (±8.89)	91.0 (±9.56)
SA - Heart	58.6 (±8.53)	60.4 (±6.66)	58.6 (±8.53)	60.4 (±6.66)	69.2 (±7.58)	69.5 (±7.19)	56.7 (±9.76)	60.0 (±7.91)
Voting	83.9 (±3.67)	96.1 (±3.90)	96.1 (±4.30)	96.3 (±3.78)	90.1 (±4.78)	93.3 (±4.87)	93.6 (±3.94)	95.4 (±3.72)
MONK-1	98.4 (±2.03)	100.0 (±0.00)	98.4 (±2.03)	100.0 (±0.00)	74.6 (±6.07)	74.6 (±6.07)	90.5 (±7.65)	86.3 (±8.68)
MONK-2	76.0 (±4.12)	66.2 (±5.83)	73.2 (±3.67)	66.0 (±5.79)	62.4 (±3.06)	61.9 (±3.23)	92.0 (±2.76)	86.3 (±5.11)
Promoters	81.3 (±6.84)	83.7 (±8.99)	83.1 (±6.74)	82.7 (±10.8)	85.9 (±10.5)	85.8 (±9.79)	71.7 (±13.7)	72.0 (±13.1)
Tic Tac Toe	86.5 (±5.14)	85.7 (±3.54)	86.4 (±5.24)	85.7 (±3.54)	70.3 (±5.75)	71.1 (±5.47)	86.5 (±3.13)	84.5 (±4.90)

Table 2: Classification accuracies of four different learning algorithms on 12 different datasets

Dataset	Unpruned decision tree		Pruned decision tree		Naïve Bayes Classifier		Rule learning alg. (CN2)	
	non-filtered	filtered	non-filtered	filtered	non-filtered	filtered	non-filtered	filtered
BreastCancerLJ	61.9 (±11.4)	67.1 (±6.74)	67.2 (±9.36)	68.9 (±8.69)	72.7 (±3.06)	73.1 (±4.81)	67.9 (±6.05)	68.9 (±3.41)
BreastCancerWI	94.3 (±2.99)	95.2 (±3.43)	94.3 (±2.98)	95.2 (±3.36)	96.8 (±2.16)	96.5 (±1.89)	94.7 (±2.10)	96.1 (±2.16)
SA - Heart	58.6 (±8.53)	60.4 (±6.66)	58.6 (±8.53)	60.4 (±6.66)	69.2 (±7.58)	69.7 (±6.80)	56.7 (±9.76)	60.8 (±9.65)
Promoters	81.3 (±6.84)	88.5 (±5.91)	83.1 (±6.74)	86.6 (±6.51)	85.9 (±10.5)	85.8 (±9.79)	71.7 (±13.7)	72.8 (±7.29)

Table 3: Improved classification accuracies (bold) by the use of tolerance lever parameter t .

4.5 Statistical evaluation of the results

To compare the performance of two learning algorithms on several datasets the *Wilcoxon signed-rank test* is used. In this case we wanted to see if the classification accuracy results obtained by the combinations of saturation filter and learning algorithm were statistically significantly better than the classification accuracy results obtained from only the learning algorithm.

First, the Wilcoxon signed-rank test was performed for the two algorithms on all the 12 datasets presented in Table 1, in order to get an idea how the filter performs on representative set of datasets. But since we did not expect improvements of the classification accuracy when applying the filter on the last four datasets which do not contain noise, we made the Wilcoxon test also separately for the first eight datasets which contain noise. At the end the test was made once more on these same eight datasets, but with the classification accuracies obtained by using the saturation filter with tolerance level parameter t .

The results from these three Wilcoxon signed-rank tests made on two different sets of datasets are presented in Table 4. The table has to be understood in the following way: the algorithm combining the saturation filter and the learning algorithm is better than the algorithm without the filter, with probability p stated in the table.

Datasets	Unpruned decision tree learner	Pruned decision tree learner	Naïve Bayes classifier	Rule learning algorithm (CN2)
All (8+4)	0.0499	0.6101	0.1579	0.8753
Only noisy (8)	0.0173	0.1415	0.0499	0.0929
Noisy (8) (with param. t)	0.0117	0.0117	0.0423	0.0423

Table 4: p -values obtained from the Wilcoxon signed-rank test (in the case where the combination of the saturation filter and the learning algorithm is better than only the learning algorithm).

By choosing statistical significance level $\alpha = 0.05$, we see from Table 4, that the application of the saturation filter on all 12 datasets yielded statistically significantly better results only while using the unpruned decision tree learner (bold in Table 4). If the Wilcoxon test is made only on the eight noisy datasets, then the filtering showed to be statistically significantly better with the use of the unpruned decision tree learner and the naïve Bayes classifier.

The most interesting Wilcoxon signed-rank test results however, for statistical significance level $\alpha = 0.05$, were obtained in the case where the saturation filter with tolerance level parameter t was used. The combination of the saturation filter with tolerance level parameter t and all four learning algorithm showed to be statistically significantly better than only the learning algorithms with no filtering applied.

5 CONCLUSION

Considering the experimental test results we can conclude that our implementation of the saturation filter combined with all four different learning algorithms results in improving the classification accuracy by previous filtering of

training data. The use of the filter shows improvements on most datasets, and with the additional use of tolerance level parameter t the improvement on even few more datasets can be observed. Finally, the Wilcoxon signed-rang test shows statistical significance of the improvement obtained by the use of the saturation filter.

At the negative side, however, the implemented algorithm is rather slow, due to its iterative process of searching for noisy training examples. Therefore, the current implementation is practically suitable only for datasets of the size up to 1000 instances (or a few 1000s). Another shortcoming of the current implementation is the complexity measure used in the filtering process: it might be better to use a more sensitive complexity measure that could better distinguish between noisy and non-noisy training examples.

Improvements of these two shortcomings, along with testing on more datasets and testing with more learning algorithms, is the subject of further work, which could make the saturation filter more widely applicable and make it perform even better.

Acknowledgements

The authors are grateful to Petra Kralj Novak for helpful advice on the implementation in Orange, and Dragi Kocev for help with the statistical evaluation of the results.

References

- [1] A. Asuncion, D.J. Newman, UCI Machine Learning Repository, <http://www.ics.uci.edu/~mllearn/MLRepository.html>, 2007.
- [2] J. Demšar, B. Zupan, G. Leban, Orange: From Experimental Machine Learning to Interactive Data Mining, <http://www.aillab.si/orange>, 2004, Faculty of computer and information science, Univerza v Ljubljani.
- [3] D. Gamberger, N. Lavrač, Conditions for Occam's Razor Applicability and Noise Elimination, Lecture Notes in Artificial Intelligence: Machine Learning: ECML-97 (M. Van Someren, G. Widmer, ur.), vol. 1224, Springer-Verlag, 1997, str. 108-123.
- [4] B. Sluban, Saturation filter for noise elimination from labeled data, Diploma Thesis, Faculty of Mathematics and Physics, University of Ljubljana, 2009.

Zbornik 12. mednarodne multikonference
INFORMACIJSKA DRUŽBA – IS 2009

Proceedings of the 12th International Multiconference
INFORMATION SOCIETY – IS 2009

**Sodelovanje, programska oprema in storitve v informacijski
družbi**

Collaboration, Software and Services in Information Society

Uredil / Edited by

Marjan Heričko

<http://is.ijs.si>

14. oktober 2009 / October 14th, 2009
Ljubljana, Slovenia

PREFACE

The transition into an information and knowledge society is being accompanied by a redefinition of business processes and a renewal of information solutions and services. There is significant growth in the size and complexity of these systems. Therefore, the main challenge is how to successfully manage and control the complexity of software and services in order to enable efficient collaboration between individuals, organizations and/or systems. The CSS'2009 proceedings bring together papers that address different aspects of software and service engineering, including model-driven development, process models, metric repository, workflow testing, database quality and refactoring, ontology-based approaches, searching techniques, business process modeling, security, knowledge management portals, wireless networks and mobile platforms, adaptive service-based applications as well as service governance and new service business models. The papers cover a broad spectrum of research as well as practical experience.

Special thanks go to the members of the Program Committee and additional reviewers as well as to the authors who have contributed to the success of this sub-conference. We hope that these proceedings will be beneficial for your reference and that the information in this volume will be useful for further advancements in research and industry.

prof. dr. Marjan Heričko

CSS 2009 – Collaboration, Software and Services in Information Society Conference Chairman

PROGRAMSKI ODBOR / PROGRAM COMMITTEE

Dr. Marjan Heričko

University of Maribor, Faculty of Electrical Engineering and Computer Science

Dr. Ivan Rozman

University of Maribor, Faculty of Electrical Engineering and Computer Science

Dr. Kimmo Salmenjoki

Seinäjoki University of Applied Sciences

Dr. Lorna Uden

Staffordshire University, Faculty of Computing, Engineering and Technology

Dr. Gabriele Gianini

University of Milano, Faculty of Mathematical, Physical and Natural Sciences

Dr. Hannu Jaakkola

Tampere University of Technology Information Technology (Pori)

Dr. Mirjana Ivanović

Department of Mathematics and Informatics, Faculty of Science, University of Novi Sad

Dr. Zoltán Porkoláb

Eötvös Loránd University, Faculty of Informatics

Dr. Adel Taweel

University of Manchester, Department of Computer Science

Dr. Aleš Živkovič

University of Maribor, Faculty of Electrical Engineering and Computer Science

Mag. Urška Kežmah

University of Maribor, Faculty of Law

Dodatni recenzenti / Additional Reviewers

Dr. Vili Podgorelec, Dr. Gregor Polančič, Dr. Marko Hölbl, Dr. Luka Pavlič

University of Maribor, Faculty of Electrical Engineering and Computer Science

V sodelovanju z / In collaboration with



SIMPLE SOLUTION FOR ONTOLOGY-BASED MAPPING BETWEEN V-MODELL XT AND SELECTED SOFTWARE DEVELOPMENT METHODOLOGIES

Peter Butka

Gerstner Laboratory, Department of Cybernetics
Faculty of Electrical Engineering, Czech Technical University
Karlovo namesti 13, 121 35 Prague, Czech Republic
e-mail: peter.butka@gmail.com

ABSTRACT

Development activities within organizations usually follow well-known methodology for software development processes by defining the sequence of steps in particular development activity. Many methodologies lack the precise formal description and it is not easy to understand projects already instantiated in one methodology in the context of another approach. Knowledge-based representation of software development processes could be a meeting point for (at least partial) mapping or capturing of relationships between different methodologies. In this paper we will present simple knowledge representation used for V-Modell XT methodology based on its documentation, where table-based mapping from several methodologies is available. Our simple ontology solution based on this documentation is ‘proof-of-concept’ for such mapping and could be re-used with other approaches, e.g. with metamodels as explicit mapping.

1 INTRODUCTION

Software engineering focuses on the processes and methods for design and construction of software products and their attributes. Therefore construction of high-quality software products requires systematic processes to define the sequence of steps during the project course, a set of suitable methods for construction, and set of quality assurance activities to verify and validate the deliverables against the specification and the customer requirements. Projects and development activities within organizations usually follow some well-know software development methodology (SDM). Many methodologies lacks of precise formal description and it is not easy to understand projects instantiated in one methodology in the context of another approach. It is because of low level of interoperability between them (relationships are weak or hard to explain). Weakness of interoperability between methodologies leads to design of knowledge representation for software development processes. This is the space for mapping or capturing of relationships between different methodologies.

Ontologies in software engineering domain (according to its usage) have very general objectives [6]:

1. *Clarify the knowledge structure* – during the ontological analysis the domain concepts and relationships between them are defined in such a way that the adequate execution of this step eases the clear specification of the nature of the concepts and terms being used, with respect to the body of knowledge that is to be constructed;
2. *Reduce conceptual and terminological ambiguity* – ontological analysis provides a framework for the unification between people (and/or agents-systems) with differing necessities and/or points of view, depending on their particular context;
3. *Allow the sharing of knowledge* – it is possible to achieve a set of conceptualizations of a specific domain, and the set of terms which support it. With an adequate syntax, these conceptualizations and the relationships between them are expressed and codified in an ontology, which can be shared by agents (person or system) having similar needs for the same domain.

Several approaches are available for the knowledge representation of methodologies, basically separated into two groups. In first case approaches create representation for direct mapping between methodologies from analysis of the similarities (and therefore achieve better interoperability). In second case abstracted common concepts from the structure of existing methodologies are incorporated into meta-model standards and re-used for creation of methodologies in constructive way (with extensions).

In [2] (or [3] for more detailed report) we have already provided more information regarding this problem, together with comprehensive analysis of the applicability for V-Modell XT methodology. According to this analysis we have found that metamodel solution based on some standard, e.g. one called Semantic Engineering Metamodel for Development Methodologies (SEMMDM, also known as international standard ISO/IEC 24744) described by the authors in [4], is suitable for a wide scope of applications. SEMMDM-based knowledge representation is suitable to be

used as *metamodel* (extended with concepts from V-Modell XT) for new project instances (metamodel for instantiation of projects), *vocabulary* for understanding the domain of SDM (basic concepts from metamodel and well-known semantic shared between SDM, any extension can be understandable using well-known metamodel concepts), *mapping* (fully modelled V-Modell XT can be mapped to any methodology, which is also modelled using this standard) in intentional way or by some tool (shared metamodel concepts and structures), and *support (guidance) tools* (SEMDM approach is agile and flexible, we can imagine useful applications, where created ontology are used for e.g. automatic tailoring and dynamic combination with other steps from different methodologies).

Other knowledge representation approaches are not able to fulfil all aspects and potential applications. The only problem is “intentional” way of mapping – we need to have another methodology modeled using SEMDM in order to have mapping available due to usage of standard elements. One way how to make clear connection between approaches is to provide it explicitly. With V-Modell XT in mind, it is possible to re-use table-based mapping from their documentation (part Mapping to Standards), where selected methodologies are mapped to V-Modell XT elements. We have modeled this using ontology and such explicit mapping then should be connected to the metamodel-based ontology of V-Modell XT in order to achieve model with full application potential and clear explicit mapping.

In next chapter introduction to SDM is provided, then proposed solution for mapping based on the V-Modell XT documentation is presented followed by the conclusions.

2 SELECTED SOFTWARE DEVELOPMENT METHODOLOGIES

One of the most important issues of an appropriate software process is development of software projects. A software processes typically defines the sequence of steps, products (deliverables, artefacts) and time plan (milestones, time periods, deadlines) during the software development life-cycle. In industrial practice a wide range of common SDM exist, which meet individual application requirements regarding project size, duration, effort, involved engineers, application domain, etc. Thus, the selection of a software process depends on the project scope and the project context. Most of the SDM approaches are based on common life-cycle models. Basic life-cycle approaches include the major steps in software development, which are common to typical projects in software engineering. Many different SDM exists (e.g. waterfall, spiral model, Unified Process Framework, several agile approaches, etc.), but we will shortly describe only those important for the paper.

CMMI (Capability Maturity Model Integration) [1] was developed to be able to carry out interdisciplinary developments, in particular software and system development projects faster and better, and to obtain at the

same time high quality products. Important concepts here are process areas and maturity levels. Process Area defines related practices in an area like Requirements Management, Project Planning, etc. Maturity level is degree of process improvement across a predefined set of process areas. Specific goals and practices are defined for accomplishment of process area and needed maturity level. *ISO15288* [5] (Life Cycle Management - System Life Cycle Processes) provides a framework of processes that cover the whole life cycle of a system. Compared to the V-Modell, the descriptions of the activities are much less detailed. It is limited mainly to a listing of the activities without examining how to proceed during the individual activities. ISO 15288 contains 25 processes that form the following four process groups: Agreement Processes, Enterprise Processes, Project Processes and Technical Processes.

V-Modell XT [6] is a mandatory framework for planning and execution of public domain software processes in Germany. It is flexible and product-centric software process model approach due a modular design, process customization (company and domain-specific requirements) and tailoring (project specific requirements). Central elements are Project types, Process Modules and Project Execution Strategies based on Decision Gates (comparable to milestones). The selection of a project scope (building software, hardware or complex systems) is the first decision. A project type focuses on the view of the project from customer (acquirer) perspective, from developer perspective, etc. Products and deliverables are important for the model and are organized in so-called “process modules”, which encapsulates products (deliverables), roles and product responsibilities, and activities to construct the product. Additionally, decision gates (comparable to project milestones) are used for project progress decision. If the product does not meet the expected requirements, additional actions have to be done to meet decision gate requirements. For process tailoring, the V-Modell XT allows to select an appropriate project execution strategy (Incremental System Development, System Maintenance, etc.). The V-Modell XT does not rely on a specific modelling technology or on specific tools, and main contribution is clear vocabulary and its comprehensive documentation.

3 SIMPLE MAPPING SOLUTION BASED ON THE DOCUMENTATION OF V-MODELL XT

In V-Modell XT documentation [7] problem of mapping is already analysed in very simple way (but sufficient, if application is used according to their assumptions). They deal with several conventions whose most important contents are related to elements of the V-Modell XT. Thus it indicates to what extent the V-Modell XT covers these conventions or to what extent it is compatible with them. It is good for those familiar with another standard, who want to classify this standard in relation with the V-Modell XT or seek a quick lead-in from this standard to the V-Modell

XT, and also should be interesting for retraining persons, who are familiar with one of the described conventions, to the V-Modell XT. In their reference following standards are available: AQAP 150 (Allied Quality Assurance Publication, NATO Quality Assurance Requirements for Software Development), CMMI (Capability Maturity Model Integration, v.1.1), ISO 15288 (Life Cycle Management - System Life Cycle Processes), ISO 9001:2000 (Quality Management Systems - Requirements), V-Modell 97 (Development Standard for German Federal IT Systems, Process Model as of June 1997). In the actual representation of the mapping the topics of the standard to be represented are listed on the left side and opposite to them on the right side elements of the V-Modell XT that cover these topics (see example in Table 1).

Table 1: Example of the mapping table from CMMI to V-Modell XT (for part Project Monitoring and Control).

Element of the Standard (CMMI)	Is Fulfilled By (V-Modell XT)
Monitor Project Against Plan	Chapter: Decision Gates ProcessModule: Configuration Management Product: Project Status Report Product: Risk List Product: Project Progress Decision
Manage Corrective Actions to Closure	Product: Project Status Report

From Table 1 we can see that Mapping to Standards part of the documentation is suitable to inform any methodologist, project manager or developer, about the all necessary elements from V-Modell XT that he/she needs to understand in order to fulfil element of other standard (e.g. CMMI). Analysis of the documentation chapter in V-Modell XT gives us all elements of this standard (direct modeling elements, documentation elements like chapters in documentation book, etc.). Summarization of the usage of different V-Modell XT elements (model or documentation related) for the mapping to V-Modell XT is in Table 2. Some of the elements are used for every mapping, while some of them are used only for one of them. In case of ISO15288 standard only four different elements are used, all other have a wide usage of more elements. If we want to cope with these mappings, then simple solution is to model elements of the other standards in the ontology and make a connection to the V-Modell XT concepts. For that reason we have modelled part of the V-Modell XT using OWL (Ontology Web Language) standard in Protégé editor. Then for the modelling of mapping from other standards we only need to model part related to elements of other standards. The elements from some standard are hierarchically organized under their corresponding parts and connected using mapping relation. As a proof of concept we only modelled two of them, we choose CMMI and ISO15288 (they are quite common and well-known standards with a quite different number of V-

Modell XT elements used in mapping). That is why we need Element concept with all elements of CMMI and ISO15288 in it. Therefore, we have subclasses CMMI_Element and ISO15288_Element. Also VMXT_Element is there, which is used for its elements (both concepts directly from model and documentation-like elements). Then basic structure consists of all process areas within CMMI and ISO15288, V-Modell XT part consists of V-Modell XT elements used in mapping tables. Concrete data from tables is then imported into ontology as instances. This could be little confusing, because of different “domain” of elements (on a left side process elements, on the right side process elements, products, chapters from documentation, etc.), but for the purpose of the proposed mapping it is suitable, and any connection to more complex model could be easily added. Most important relation for the mapping is to connect CMMI and ISO15288 elements with V-Modell XT elements, e.g. let say we have CMMI Project Monitoring and Control process area, according to documentation it is completely covered by the V-Modell XT using elements from Table 1. Then Figure 1 presents ontology part related to this process area. Relation between elements of CMMI and V-Modell XT is called *isFulfilledBy*.

Table 2: Example of the mapping table from CMMI to V-Modell XT (for part Project Planning).

Element/ Standard	AQAP 150	CMMI	ISO 15288	ISO 9001	V-Modell 97
Process Module	X	X	X	X	X
Product	X	X		X	X
Product Group	X	X		X	X
Activity	X	X	X	X	X
Activity Group	X	X	X		X
Sub-Activity		X			
Role	X			X	
Project Execution Strategy			X		X
Chapter	X	X		X	X
Topic	X	X		X	
Section	X	X		X	
Method Reference		X			

Only necessary subclasses and instances are shown, whole ontology consists of more elements (classes and instances) from CMMI, ISO15288 and V-Modell XT, and also more

relations (objective properties) are used for the description of domain (especially V-Modell XT part) like *isFulfilledBy* (CMMI_Element or ISO15288_Element is mapped to VMXT_Element, already mentioned mapping relation from tables), *isFromProductGroup* (VMXT_product is in direct relation with its VMXT_ProductGroup), *isFromActivityGroup* (same as for products, but between VMXT_Activity and VMXT_ActivityGroup), *isPartOfActivity* (Sub-activity is part of some Activity), *isPartOfProduct* (Topic is part of some product, like some scheme in documentation, or some subsection, so VMXT_Topic is in this relation to VMXT_Product), etc. All relations should also have inverse relation (property) that can be used for the inverse queries. Ontology is relatively small with 40 concepts under main Element concept, mapping and other basic relations between instances (inverse of them also available, if needed) and more than 300 instances, plus datatype/annotation properties for detailed description of concepts/instances.

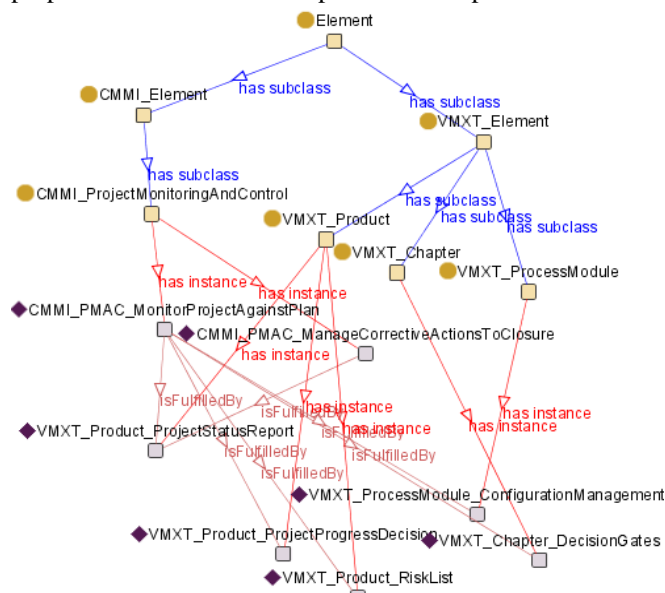


Figure 1: *Project Monitoring and Control part of ontology.*

Because structure is simple and purpose-oriented (mapping to standard based on the table approach, where elements from standard X are connected to elements of V-Modell XT), ontology is not suitable as vocabulary application or metamodel for general development process. On the other side, it should be still helpful for understanding specific part of SDM domain (part related to involved standards). Motivation for such solution is simple. Main purpose is mapping tool between V-Modell XT and several standards, useful also for re-use with ontology solution based on metamodels. Similarly, we can imagine simple help tool (guidance tool), which can help developer or project manager transform his/her knowledge about e.g. CMMI to V-Modell XT. Extension of such application can use models to produce some products, like documentation parts or diagrams, with a help of such knowledge-based tool. Such simple solution could be helpful also like training material

for people in organizations. And important motivation is that model is easily extensible and already contains relationships not available in table-based mapping, e.g. connections between product and product groups, activity and activity groups, etc. Therefore solution is useful not only as another table-based mapping, but as knowledge-based representation of V-Modell XT and other methodologies in practice.

4 CONCLUSION

In this paper we presented our simple mapping solution between V-Modell XT and other selected software development methodologies, based on the ontology modeling of table-based mapping proposal from the V-Modell XT documentation. As a proof of concept CMMI and ISO15288 were used in modeling. Such solution should be re-used for the mapping and supporting tools in development processes (where mentioned methodologies are involved) by the developers or project managers, if they want to understand instantiation of their projects in new standard. For the future it is also interesting that this solution should be connected to metamodel solution and directly used as explicitly provided mapping part of combined ontology.

ACKNOWLEDGMENT

This research work has been supported by a Marie Curie Transfer of Knowledge Fellowship of the European Community's 6th Framework Programme under the contract MTKD-CT-2005-029755: CzechVMXT.

References

- [1] CMMI Product Team, Capability Maturity Model Integration (CMMI), Version 1.1, 2002.
- [2] P. Butka. Different knowledge representations of software development methodologies used for V-Modell XT. *Proceeding of 30th International Conference on Information Systems, Architecture & Technology (ISAT 2009)*, Szklarska Poreba, Poland, 2009.
- [3] P. Butka. Knowledge-based Representation of Software Development Processes – Existing Approaches and Applicability for V-Modell XT. Technical Research Report, Deliverable for MTKD-CT-2005-029755: CzechVMXT project, 2009.
- [4] C. Gonzalez-Perez, B. Henderson-Sellers. An ontology for software development methodologies and Endeavours, In: *Ontologies in Software Engineering and Software Technology*, Springer, 2006, pp. 123-152.
- [5] ISO/IEC 15288: *Life Cycle Management - System Life Cycle Processes*. ISO/IEC Technical Report (TR), International Standards Organization, 2004.
- [6] F. Ruiz, J.R. Hiler. Using Ontologies in Software Engineering and Technology. In: *Ontologies in Software Engineering and Software Technology*, Springer, 2006, pp. 49-102.
- [7] V-Modell XT framework, <http://www.v-modell-xt.de>, last accessed, March 2009.

MODEL-DRIVEN ENGINEERING AND AN EXAMPLE OF ITS INTRODUCTION

Tomaž Lukman, Giovanni Godena
Department of Systems and Control
Jožef Stefan Institute
Jamova 39, 1000 Ljubljana, Slovenia
e-mail: tomaz.lukman@ijs.si

ABSTRACT

This paper discusses Model-Driven Engineering (MDE), an approach to software engineering that promises to increase software quality and productivity. This is achieved through the employment of Domain-Specific Modeling Languages (DSML) and model transformations. As MDE captures more and more attention of the industry, many organizations are facing the question: How to introduce MDE into the domain(s) for which we develop software? This and additional questions may be addressed by the example of introducing MDE into the domain of industrial process control, which is the main focus and contribution of our paper.

1 INTRODUCTION

Two of the recognized challenges that software engineering is facing today are how to alleviate the constantly increasing complexity of software development [1] and how to sustainably increase the productivity of developing software systems [2].

2 MODEL-DRIVEN ENGINEERING

Model-Driven Engineering¹ (MDE) is a software development paradigm that has the potential to sustainably raise the development productivity and to alleviate the complexity of software engineering. It offers an environment that ensures the systematic and disciplined use of models throughout the development process of software systems. The essential idea of MDE is to shift the attention from program code to models. This way models become the primary development artifacts [3]. The two pillars of MDE are Domain Specific Modeling Languages (DSMLs) and model transformations. In order to achieve all the potential benefits of MDE both of these pillars have to be supported by sophisticated software tools.

2.1 Domain-Specific Modeling Languages

Many software development organizations are specialized in the development of software within one or only a few

application domains. Project after project they gather valuable expert knowledge about those domains. This domain-specific knowledge is often the prime intellectually property of an organization [2]. To shield the organization against losing this property due to personnel fluctuation this knowledge should be made explicit and one of the best ways to do so is to codify it into a DSML.

A DSML formalizes the application structure, behavior, and requirements within a particular application domain [1]. Because of that we actually reuse this formalized knowledge every time we create a model with that particular DSML.

In practice two major ways of defining a DSML have been established. The first approach is based on the extension the Unified Modeling Language (UML), which is a general purpose modeling language (GPML). This extension, which actually specializes UML for a specific domain is achieved through the definition of a UML profile that consists of a set of stereotypes, tagged values and constraints. The specific instance of MDE that was proposed when working with UML based DSMLs is the Model-Driven Architecture (MDA). The second approach is to define a DSML “from scratch”, where there are no limitations that UML is forcing on the semantics and the syntax of a developed DSML.

2.2 Model transformations

MDE aims to automate many of the complex but routine development tasks which still have to be done manually today [4] with model transformations. This is possible because of the employment of formally defined and precise DSMLs. Examples of tasks that are automatable with model transformations are the generation of lower-level models from higher-level models and the refactoring of models. A strong benefit of using DSMLs is that a large portion of the code (up to 100%) can be automatically generated, which is not possible with GPMLs, like UML.

2.3 Software tools

As already mentioned, sophisticated tools have to be employed to achieve all the benefits of MDE in practice. In the case that a DSML is realized through a UML profile there are plenty of tools already available, which support modeling and model transformations with such DSMLs. In

¹ also known as Model-Driven Development (MDD)

some cases however custom tools have to be developed or adapted to enable specific features (e.g., verification). In the other case when the developed DSML is realized independently of UML all the needed development tools have to be developed. This is a challenging task, which requires a large amount of resources. Based on the literature [5] and our own experience the effort needed to develop a basic set of tools (graphical model editor and code generator) without specialized software tools is measured in several man-years. Fortunately in the last years more and more specialized tools that enable a cheaper, easier and faster development of tools needed to realize MDE with a specific DSML are appearing. These tools that are sometimes named meta-tools automatically generate (a part of) the tool support based on a formal definition of the DSML and/or transformations.

3 EXAMPLE OF INTRODUCING MODEL-DRIVEN ENGINEERING

In order to offer a glimpse at some of the challenges one may encounter when introducing MDE into a specific domain, the following subsections present a real-life example of introducing MDE into the domain of industrial process control.

3.1 Motivation for introducing a DSML

Systems in the domain of industrial process control are responsible for controlling a particular industrial process, which transforms the process inputs into the desired outputs. Software is a very important part of this system, because it must synchronize the operation of the process equipment in the system and generate proper reactions based on the process equipment signals and operator commands. When software fails in this domain it will mostly cause great material damage and sometimes even endanger human lives or the environment. Therefore this software has to be more reliable and fail-safe than software in many other domains e.g., business information systems. The motivation of introducing a DSML into this domain was the state-of-practice in the 90's of the previous century, which made it hard to construct reliable software in the given financial and time constraints. One of the major issues was the low level of abstraction in the early as well as in the later stages of the software lifecycle. This made it hard to manage the growing size and complexity of the process control systems under development. In the early development stages (if they were even carried out) the software was usually specified by hundreds or thousands of lines of pseudo code, which were hard to validate and to comprehend on a global scale. Even if graphical models were used (e.g., Statechart models), that did not necessary lead to an optimal decomposition of the system, mainly because the modeling language with which they were constructed lacked of the appropriate elements and domain constraints. Additionally none of the 5 programming languages for the implementation platform, which was defined by the IEC 61131-3 standard, was object-oriented. Therefore the application of object-oriented methods and

modeling languages in the domain of process-control software was difficult because of the big semantic distance between the object-oriented abstractions and the available programming languages [6].

3.2 The ProcGraph DSML

The potential solution for the mentioned problems was the development of a DSML, which had the appropriate high-level and domain-specific abstractions. This language was named ProcGraph and more information about it can be found in [6] since its diagram types and elements will not be presented here due to the given space limitation. The ProcGraph language has been developed based on the experience that researchers and engineers have gathered over several years working on complex real-life projects. Until today it has been successfully used for the specification of over ten real-life projects ranging from a glue-production plant to a TiO₂ production plant.

3.3 Motivation for introducing tool support

Over the years that the ProcGraph DSML was used in practice the transformation from the models build in it to the target programming language(s) have been made manually by developers. The reason that the code generation had to be done manually was the absence of tool support for the ProcGraph DSML. The idea arose to automate this transformation to gain on productivity and avoid the errors introduced by the manual transformation. Although this idea appeared already in the previous decade it became more feasible with the appearance of MDE. The preferred type of tool support for ProcGraph is an integrated development environment (IDE) that integrates all the required features into one tool. The initial version of this IDE, which is schematically shown on Figure 1, should have the following features/components:

- A visual ProcGraph editor that enables the construction of ProcGraph models in an easy and straightforward way. The editor should enable only the construction of valid (well-formed) models, which conform to the domain constraints that embody valuable knowledge about the industrial process control domain.
- A model repository that holds the structured data model i.e., an internal representation of ProcGraph models.
- A code generator, which generates code from the models in the model repository. For our initial version of the IDE only one code generator is being developed. This generator will generate program code for the platforms provided by Mitsubishi.

3.4 To MDE through the development of tool support

At the beginning of the development we were determined to use as many meta-tools as needed to achieve an efficient (economical and rapid) development process. Therefore we systematically selected and procured the most appropriate tools according to our requirements. The meta-tools that were selected for our ProcGraph IDE were the Graphical Modeling Framework (GMF) and openArchitectureWare.

With them it is possible to develop an Eclipse based IDE for an arbitrary DSML.

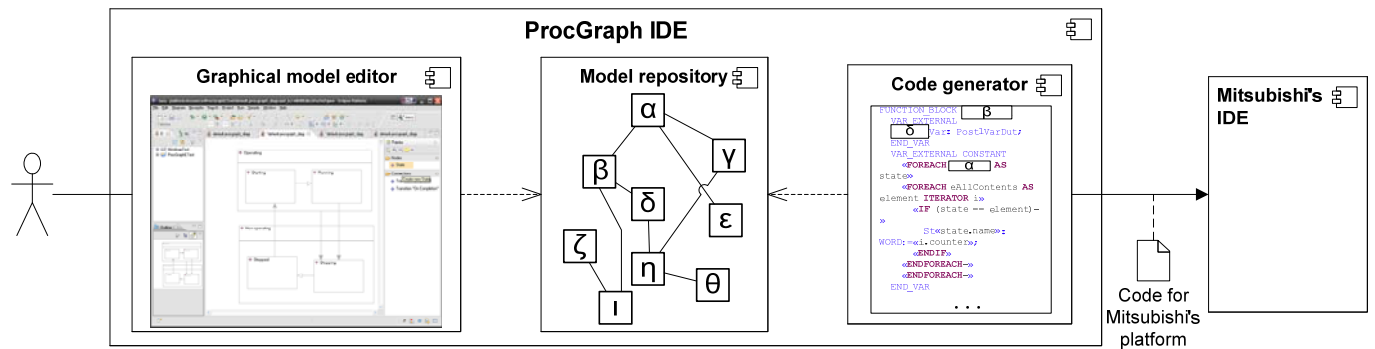


Figure 1: The scheme of the ProcGraph IDE structure and its interfaces.

GMF is used for the development of graphical model editor and the corresponding model repository. The GMF framework/tool consists of two components, the generative and the runtime component. The generative component enables the automatic generation of the basic graphical model editor based on various models of the selected DSML. The GMF runtime component includes all the classes that are useful at the implementation of graphical model editors. openArchitectureWare is a framework for MDE that has many features, including a powerful code generation engine.

The major activities that are being undertaken in the process of developing the IDE for ProcGraph are depicted in Figure 2. In the following subsections these activities will be described through the perspective of the IDE features presented in section 3.3.

3.4.1 Developing the model repository

The first activity was the definition of the domain model for our industrial process control domain. The domain model should contain all the relevant domain entities, relationships among these entities and constrains. The domain model, as it is called in GMF, is actually the metamodel of the selected DSML. To enable the modeling of a domain GMF uses the Eclipse Modeling Framework (EMF) tool. EMF has all the necessary features that facilitate the development of a structured data model repository for arbitrary models. The domain model is defined (instantiated) with ECore, which is EMF's own meta-metamodel. Ecore is based on the standardized EMOF (Essential Meta-Object Facility). We had no difficulties in defining the metamodel for ProcGraph, because most of its elements were already known. After the ProcGraph metamodel was defined in Ecore (more specific in a graphical editor for Ecore, which is very similar to a UML class diagram editor), EMF was used to automatically generate the model repository and a simple hierarchical editor for this repository. The editor allowed the construction and manipulation of ProcGraph models (just the data model without the appropriate graphical representation i.e., the concrete syntax), which enabled the independent and parallel development of the two remaining components.

3.4.2 Developing the graphical model editor

The next model that is needed by GMF to generate an editor is the notation model for the selected DSML. The notation model defines the concrete syntax with which the users will be modeling. The graphical elements (symbols) of the DSML can be defined by composing primitive figures (e.g., circles and squares) into a compound figure. For this task GMF provides a simple hierarchical editor or an experimental graphical editor. An alternative is to define the graphical elements (nodes and relationships) programmatically with the help of the Draw2D library and then import them into the notation model. The modeling of the ProcGraph concrete syntax was relatively easy with the exception of some specific symbols, which require dynamic changes of their appearance based on the model content. To construct these specific symbols a deep knowledge of Draw2D is required, so some of them were left out for the next versions.

The next activity was the definition of the tools model. This model defines all the tools with which the user can manipulate the content of a specific diagram. This was an easy task for the ProcGraph DSML.

The next model which we had to define was the mapping model. The purpose of it is to relate all the previously defined models, which are initially independent from each other. We experienced no larger complications at the definition of the mapping model for the ProcGraph DSML. The last modeling activity is the tuning of the model editor, through the generator model which decorates all the previous models with additional information about the editor. With the appropriate parameters of this model the end user experience can be significantly improved. The problem we had was that a lot of experimentation had to be done because of the myriad of parameters that are poorly documented. After the generator model was prepared the automatic generation of the graphical model was carried out. It is necessary to be aware that even in the case that all the models perfectly reflect the developers intent, the generated editor is far from good enough for professional development with the selected DSML. Therefore the editor has to be specialized and extended through manual coding.

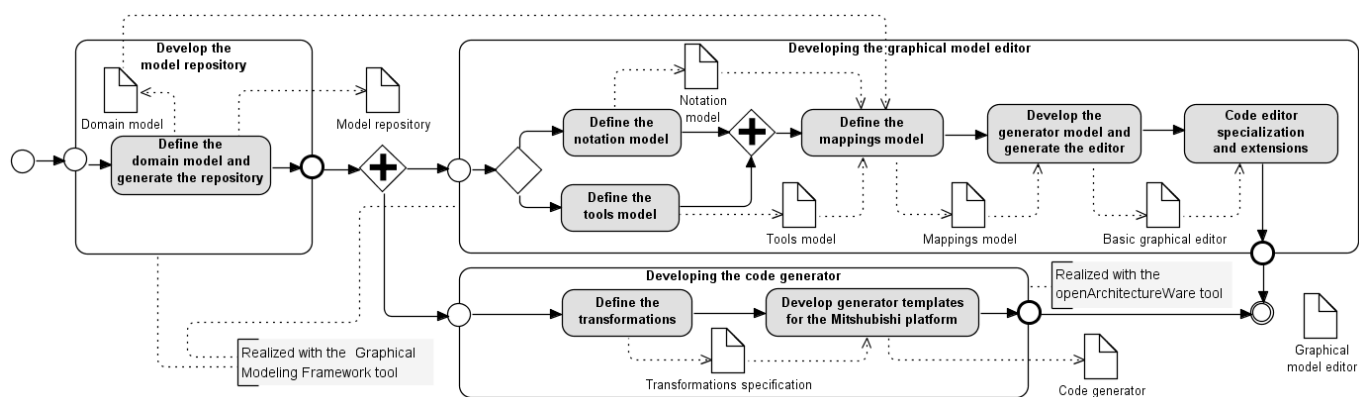


Figure 2: The process of developing the ProcGraph IDE, which will realize MDE in the domain of industrial process control.

The manual specialization and extension with which we are still occupied is the most demanding and effort consuming activity. In the case of ProcGraph this activity has already consumed 2 man-months and will consume several more before the editor can be tested by potential customers.

3.4.3 Developing the code generator

At the beginning of the code generator development we decided that we will begin with the formal specification of the model-to-code transformations. We are currently searching for the most appropriate formalism to define this transformation from the ProcGraph language into the Function Block Diagram graphical programming language. This specification will serve as the basis for implementing the platform specific generators.

The next activity will be the implementation of the code generator for the Mitsubishi platform. In openArchitectureWare the code generator is implemented through templates, which are coded in a special scripting language (template language) called Xpand. A template is a textual file composed of a mixture of static and dynamic parts. The static parts of the text are the same as the ones in the generated code and usually describe the part of the code that does not change through different applications. The dynamic parts are written in Xpand and are also called metacode. The metacode accesses parts of the source model, which is located in the model repository. These parts of information are retrieved and then inserted into the template. This way the code is generated from the current source model (which abstracts a specific application). Implementing the generator for the Mitsubishi platform will be challenging, because of the odd format in which the Function Block Diagram programming language has to be described so that the Mitsubishi IDE can import it and then to translate it into executable code.

3.5 Expected benefits

As implicitly indicated the development of the ProcGraph IDE is not yet complete and therefore several man-months of effort are still required to finish the first useful version. Therefore the benefits of MDE in this domain can only be

predicted based on the experiments done with the incomplete tool and our experience. These benefits will be:

- Increasing software quality – Because only well-formed models can be constructed there will be less design flaws, which will increase the quality of software. Additionally the automatic code generation will reduce coding errors.
- Improved productivity – due to the 100% code generation the productivity will sustainably increase.

4 CONCLUSION

In this article a short overview of the MDE approach to software engineering has been given. Moreover an example of the ongoing introduction of MDE into the domain of industrial process control was presented. This example may be useful for organizations who are considering the MDE approach to developing software in one or some of their domains, but have little idea of where and how to start. Also this example gives interesting insight into the domain of industrial process control, which differs from many other application domains, where software systems are being used.

References

- [1] D. C. Schmidt, "Model-Driven Engineering," *IEEE Computer*, vol. 39, no. 2, pp. 25-31, 2006.
- [2] B. Selic, "The pragmatics of model-driven development," *IEEE Software*, vol. 20, no. 5, pp. 19-25, 2003.
- [3] E. Seidewitz, "What models mean," *IEEE Software*, vol. 20, no. 5, pp. 26-32, 2003.
- [4] C. Atkinson, and T. Kuhne, "Model-driven development: a metamodeling foundation," *IEEE Software*, vol. 20, no. 5, pp. 36-41, 2003.
- [5] J.-P. Tolvanen, "MetaEdit+: domain-specific modeling for full code generation demonstrated," in 19th OOPSLA Conference, 2004, pp. 39-40.
- [6] G. Godena, "ProcGraph: a procedure-oriented graphical notation for process-control software specification," *Control Engineering Practice*, vol. 12, no. 1, pp. 99-111, 2004.

A MODEL BASED CODE GENERATION SUPPORT FOR DEVELOPING A PRESENTATION LOGIC

Jan Kryštof, David Procházka, Arnošt Motyčka

Department of Informatics

Mendel University of Agriculture and Forestry in Brno,

Faculty of Business and Economics

Zemědělská 1, 613 00 Brno, Czech Republic

Tel: +420 545 132 241; fax: +420 545 13 27 27

e-mail: jan.krystof@mendelu.cz

ABSTRACT

This article deals with issue of user interaction with graphical user interface of software application with focus on tight involving of interaction objects in UML activity diagrams. We propose an approach which uses macro-based activities, stereotypes of interaction objects from UML profile and other specific objects. The purpose of our approach is to provide better modeling facilities for specifying user-interface interaction and provide more detailed models for tasks of code generation such as code generation or documentation.

1 INTRODUCTION

Rapid development of software systems providing graphical user interface emerged a need for precious modeling of user-interface interaction in order to provide a support for various aspects of issue of development of graphical user interfaces. Phases of analysis, preparation, development and evaluation need to have proper way of description and specification. Models can provide the right solution for this purpose and bring independent and clear artifacts for the wide spectrum of involved participants: developers, analysts, HCI specialists, testers and regular users meet during the phases of the development process and collaborate on designing, development, testing and evaluation of the user-interface interaction.

In order to provide a support for model-based development, there was designed a concept proposing to use specialized tools enabling modeling and development of user interfaces: MB-UIDE (Model-Based User Interface Development Environment) [1]. The purpose of these IDEs is to enable definition of declarative models, which present appropriate artifacts for diverse tasks within realization of user interfaces. Advantages of these tools are especially [1]:

- Providing a platform independent approach for modeling user interfaces and related user-interface interaction. MB-UIDEs do not use any concrete data types when dealing with interaction objects, but they use

an abstraction instead, such as an inputter, action invoker etc.

- Providing a methodology controlling the whole process of the UI development.
- Providing a functionality for support of specific tasks related to the UI development such as road maps generation or menu definition.

We can mention the WebRatio product as a representative of commercial products and the ArgoUwe as a representative of open-source products. The WebRatio is a product designed to support WebMI [4] methodology and the ArgoUwe supports UWE [3] methodology. Both methodologies share similar characteristics since they are derived from former methodologies such as OO-HDM or OO-H [2]. Their purpose is to guide developers through a development process consisting of working out a set of model artifacts containing the requirement model, domain model, task model, navigation model and presentation model. Both methodologies provide modeling facilities which are mostly based on UML and there also exist UML profiles for the WebMI and Uwe. Furthermore, ArgoUwe and WebRatio use the MDA concept and provide features for semi-automatic code generation for Java EE platform. The methodologies provide rich logical frameworks with guidance and development environment thus they have a high chance to be adopted in the commercial sphere. In spite of it, we can identify several drawbacks, which we can find in both of them. The paper [3] identified a gap with the respect to usability issues, particularly UWE does not provide a way to model usability guidelines at an abstract and also navigational model concentrates too much on navigation link between pages rather than on links between related information. We also miss facilities for modeling interaction object in more concrete level as we deal in the paper [6]. Furthermore both methodologies are focused on web-based application and do not provide better support for independent applications. This drawback becomes very obvious when we deal with the navigation model. The main drawback we see in low support of presentation logic. For instance, the UWE focuses on modeling the business logic but does not make any deeper relation between the logic and

interaction objects. The lack of this kind of modeling then disables a possibility to provide code generation support for presentation logic in the implementation phase. We perceive this lack as a crucial disadvantage and we focus on bridging the gap in this paper.

2 MODELING FACILITIES

First of all, we will specify the term presentation logic. The presentation logic is a part of the whole functionality which is presented in the presentation layer and which is related to a change of an interaction unit. This change can happen during user-interface interaction or during an asynchronous message from the underlying application layer.

After introduction of the presentation logic we need to argue, which kind of modeling facilities we will use. In this case we will refer to UWE and WebMI again. The UWE approach extends the UML meta-model. WebMI approach uses its own meta-model but its facilities have similar notation to UML. However both of them provide corresponding UML profiles [4, 5], thus they enable also other modeling tools being involved.

2.1 Modeling interaction objects

Our approach is based on using a profile, which we have already designed [6]. The profile represents a solution for modeling five domains which we identified in the presentation layer. The profile provides also facilities for detail modeling structural domain beside others. The structural domain in our profile includes stereotypes for assembling the visible part of the GUI which are basically Containers, Presentation units and Controlling units [6], which we consider in our approach as interaction objects. However in practical applications, we do not use these objects on such a high level of abstraction, but we use concrete objects instead such as a button, drop down box or dialog window. We deal with practical application of this profile in the paper [7].

Furthermore we defined a set of layout-oriented relationships in the profile in order to enable model layout of particular windows or containers as we deal in the paper [12]. We are also able to transform the structural part of the particular container into a source code according to the MDA concept. Our approach of transformation is described in the paper [13].

2.2 Modeling interaction

As we have already mentioned we designed a UML profile which provides facilities to model some scopes of the presentation layer. The profile is mainly focused on structural scope modeling but it also provides facilities to model a functionality which is related to interaction objects and can help us with modeling the interaction. However the functionality is defined in a corresponding meta-model which does not have a behavioral character and can not be used in the sense of activity diagrams in UMLi. The meta-model was created in order to model static relationships between entities, such as an interaction object, an action or

a procedure of the presentation layer thus we need to find and adopt another approach for interaction modeling.

There have been many papers written trying to provide a way for modeling user-interface interaction with respect to navigation [8] or business logic [9]. From our point of view, the largest work in this scope has been done by the UMLi [10] and related research. In all mentioned references, there were UML activity diagrams used and we use them in our approach as well.

The UMLi uses activity diagrams and a set of high abstract objects in order to model user-interface interaction objects. The set consists of six constructors which are Inputters, Editors, Displayers, ActionInvokers, Containers and FreeContainers. UMLi also provides a set of abstraction actions in order to express semantics of interaction. These actions are represented by stereotypes interacts, presents, confirms, cancels and activates. An Activity diagram in UMLi is realized as a composition of activities, which are connected with other activities by control flow or object flow. If the relationship of an object-flow is marked by one of interaction stereotypes, the client object represents an interaction object and is also marked by an appropriate stereotype. The activity connected as the supplier object in the relationship is marked by a text label informing more in detail about its responsibility. UMLi also proposes to use macro-activities and uses them to model behavioral categories usually observed in user-interfaces: optional, order-independent and repeatable behavioral. With the use of these we can better model options and responsibilities of a user.

The extended activity diagrams bring a way how to model user-interface interaction, provide semantics and propagate interaction units. Especially the propagation of interaction objects is the corner-stone feature for modeling presentation logic. UMLi bridges a gap which is in the case of the UWE and WebMI since UMLi provides a systematic way how to involve interaction objects into modeling. On the other hand it is worth mentioning, that modeling interaction objects on too high level of abstraction does not meet needs for tasks like code generation.

UMLi is realized as an extension of UML meta-model which implies a compatibility with the standard UML. Unfortunately this also means that all features of UMLi can be usually used only with some special modeling environments specially tailored for the UMLi. During our investigating we did not find appropriate UML profile of UMLi which would bring features of UMLi into those major modeling concepts - UWE and WebMI.

3 APPLICATION

We decided to adopt the UMLi way of interaction modeling and bring concrete interaction objects defined in our profile. This integration of the profile with the UMLi way of interaction modeling and adding a set of several specific objects set a good infrastructure for modeling presentation logic and generation of a source code. We will explain the basics of our approach in this section.

3.1 Modeling infrastructure

In order to model the presentation logic, we use UML activity diagrams, which:

- contain the basic activity macros and additional macros such as a zeroOrMore and oneOrMore macros. The zeroOrMore macro models that performing of related activities is conditional. When the condition is satisfied, the activities can be performed as many times as it is specified, including none. The zeroOrMore macro is an extension of “repeatable” behavioral from UMLi. The oneOrMore macro specifies mandatory set of activities which must be performed at least once.
- contain actions marked by stereotypes “fill”, “extract”, “activate” and “change”. These actions model data flow between interaction object and user or objects of the presentation logic. The stereotype “activate” represents actions which execute other actions and stereotype “change” represents a change of an object property. To discern if the action was invoked by a user or the presentation logic, we mark all actions performed by the user with the “interact” stereotype.
- Contain concrete interaction objects from the structural scope of our profile, can be used by the user or the presentation logic.
- Contain objects marked by “Accessor” stereotype. The Accessor objects are designed to manipulate data conveyed by interaction objects or to change their properties such as color, background or border style. Every concrete object in the profile has it's implementation of an Accessor class.
- Contain objects with “Result”stereotype used for reading all data, which are presented by interaction objects.

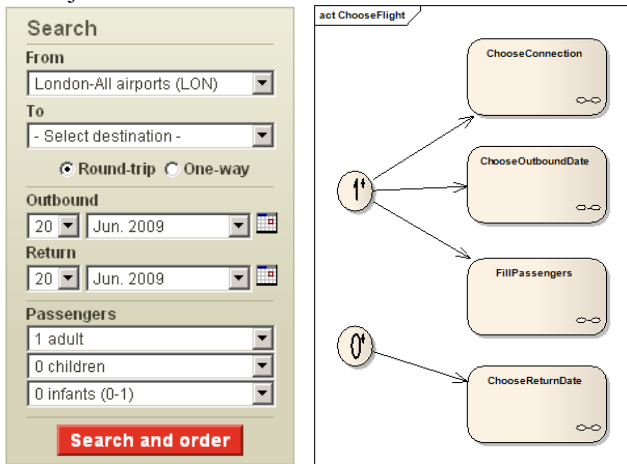


Figure 1: Screenshot and task-model for booking flights use-case

The common task in the user-interface interaction of client-server systems is to validate data provided by a user before they are processed by the application logic. The validation can be performed on the server side as it is in the case of Jakarta Struts framework or on the client side by the

presentation logic. In the following, we provide an example of a task where a user of a booking flight system provides information about the return flight.

3.2 Modeling

Use case: [...] User sets arrival and departure destinations by filling fields “From” and “To”. If he does not choose “One-way” option, he also needs to set information about return flight, which means providing the date of departure by filling fields “day” and “month - year ” with the use of a calendar component or by direct using of these fields. Using the calendar will fill both those fields automatically. [...].

The use case can be split into tasks, which are shown on the figure 1. The diagram contains task s “ChooseConnection” and “ChooseOutboundDate”, which are mandatory and can be repeatable performed, since they are associated with with “oneOrMore” behavior. The task “ChooseReturnDate” is conditional and is modeled on the Figure 2.

The condition is satisfied in the case that the radio button “roundTripCB” is selected by the user. The condition is modeled by a decision node which controls execution of a task “SettingReturn”. By performing the “SettingReturn” task, the user provides information about departure by setting two drop down boxes “dayOutboundDBox” and “monthYearOutboundDBox”. In order to fill both these fields, it is also possible to use the calendar component, which fills them automatically. The result of the “SettingReturn” activity is an object of the Result stereotype, which holds all data from affected fields.

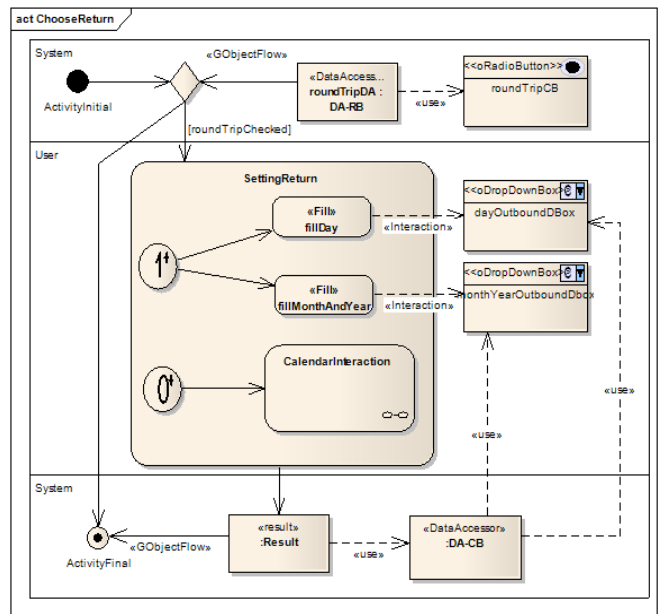


Figure 2: Detailed activity diagram

3.3 Tasks related to model-text transformation

Activity diagrams can be processed automatically and read data can be used for transformation into source code. The transformation of tasks is mainly related to sections, which

are maintained by the presentation logic and which are located by swimlanes in the "System section" within diagrams. With the respect to the figure 2, there is the decision node and the result object maintained by the presentation logic. In the scope of model-text transformations, we focus on tasks such as testing completeness of provided data, checking state of controlling units and related conditional tasks, providing contextual help. The process of transformation is implemented as a sequence of model-model transformation, which refines model to optimal state which is taken as an input for the final model-text transformation and which produces a source code. The source code does not provide a complete solution of modeled tasks but produces reasonable fragments which represents a guidance for the start of the implementation phase and can speed up the work. Easily maintainable templates can be also tailored in order to support internal rules of developer teams (annotations, form of identifiers).

3 CONCLUSION AND FUTURE WORK

The paper focuses on user-interface interaction modeling in order to provide better modeling facilities and support development part of work. In this phase of our investigating we specify modeling facilities and form an infrastructure for modeling with the respect to one of our goals: source code generation. Modeling of the presentation logic is a part of our research in the scope of model driven realization of the present layer of software applications. The goal of our research is to design UML profile providing facilities to model structure, navigation and functionality independently on the fact, whether the application is run independently on a web-browser or is run within a web-browser. By using the profile we can create a set of declarative models of the presentation layer and transform them into a source-code for target platform or to transform them into technical or user documentation. We do not provide any environment as it is in the case of MB-UIDE but we propose to create an additional user-interface layer over a current modeling tool which is used for manipulating UML models. The interface accesses underlying UML models and is capable to call methods of our libraries providing UML data processing (e.g. generation of source code or documentation). In this issue we have investigated into modeling tools Sparx Enterprise Architect and Visual Paradigm.

At this moment, we are working on detail specification of our methodology in order to start the evaluation phase. One of main goals is to find out, if the set of our declarative models is capable to describe relevant aspects of real-life applications with GUI and enhance the entire development process. Until this phase is done, we can mention similar projects with evaluations in literature [15, 16].

Acknowledgement

The research behind this paper is supported by grant agency of MUAFF in Brno within a project no. IG190661 - Advanced Modeling of Graphical User Interfaces.

References

- [1] P. Silva, User interface declarative models and development environments: a survey, In Proceedings of DSVIS 2000, 2000, pp. 207–226.
- [2] C. Cachero, J. Gómez, A. Párraga, O. Pastor: „The OO-H Method“, IWWOST 2001, Valencia, Spain, 2001. On line: <http://www.dsic.upv.es/~west2001/>
- [3] Atterer, R. (2005), WhereWeb Engineering Tool Support Ends: Building Usable Websites, in ‘SAC ’05: Proceedings of the 2005 ACM symposium on Applied computing’, ACM Press, NY, USA, pp. 1684–1688.
- [4] Nathalie Moreno , Piero Fraternali , Antonio Vallecillo, A UML 2.0 profile for WebML modeling, Workshop proceedings of the sixth international conference on Web engineering, July 10-14, 2006, Palo Alto, California
- [5] Koch, N., and Kraus, A., 2002, The expressive power of UML-based Web engineering. In D. Schwabe, O. Pastor, G. Rossi, and L. Olsina, eds., Proceedings Second International Workshop on Web-Oriented Software Technology (IWWOST’02), pp. 105–119.
- [6] J. Kryštof, A. Motyčka, Metamodel for presentation layer. Proceedings of the 11th International Multiconference Information Society - IS 2008. 2008. č. 11, pp. 270-273. ISSN 1581-9973.
- [7] J. Kryštof, N. Chalupová, Prerequisites for new GUI modeling approach. In JANECH, J. Objekty 2008. Žilina: Žilinská univerzita v Žilinie, 2008, pp. 127-136. ISBN 978-80-8070-927-3.
- [8] Lieberman, B. UML Activity Diagrams: Detailing User Interface Navigation. [online], 2004. Dostupné z: <http://www-128.ibm.com/developerworks/rational/library/4697.html>
- [9] Koch, N., Kraus, A., Cachero, C., Meliá, S. Modeling Web Business Processes with OO-H and UWE. 3rd International Workshop on Web Oriented Software Technology (IWWOST03), Oviedo, Spain, 2003.
- [10] P. Pinheiro da Silva and N.W. Paton. "User Interface Modelling in UMLi " IEEE Softw. 20. pp. 62-69.
- [11] F. Valverde, J. I. Panach, and Ó. Pastor, "An Abstract Interaction Model for a MDA Software Production Method," in 26th International Conference on Conceptual Modeling (Posters). Auckland, New Zeland, 2007.
- [12] J. Kryštof, Formal description of a layout of the GUI. In 11th International Conference MEKON 2009. 1. vyd. Ostrava: VŠB - TUO, Faculty of Economics, 2009, ISBN 978-80-248-2013-2.
- [13] J. Kryštof, J. An automated platform independent realization of GUI with use of UML. In Gaudeamus. IMEA 2009. Hradec Králové: Univerzity Hradec Králové, 2009, pp. 14-18. ISBN 978-80-7041-851-2.
- [14] K. Stirewalt. Automatic Generation of Interactive Systems from Declarative Models. PhD thesis, Georgia Institute of Technology, December 1997.
- [15] C. Wiecha, W. Bennett, S. Boies, J. Gould, and S. Green. ITS: A Tool for Rapidly Developing Interactive Applications. ACM Transactions on Information Systems, pp. 204-236, July 1990.

THE USE OF METAPHORS IN THE DEVELOPMENT OF INFORMATION SYSTEMS

Saša Kuhar, Marjan Heričko

University of Maribor, Faculty of Electrical Engineering and Computer Science

Smetanova ulica 17, 2000 Maribor, Slovenia

Tel: +386 2 220 7359; fax: +386 2 220 7272

e-mail: sasa.kuhar@uni-mb.si, marjan.hericko@uni-mb.si

ABSTRACT

This paper introduces the theories of metaphors in general, in both user interfaces and in software engineering. Since the choice of metaphor is crucial for success, the guidelines for choosing an appropriate metaphor are listed. We also conducted a research of the identification of visual metaphors by software developers.

1 INTRODUCTION

Metaphors help us understand new and abstract concepts and for this reason have been implemented in nearly all computer systems. Ever since their first appearance in graphic user interfaces and because of some inappropriate implementations, metaphors have been the subject of diverse opinions about the benefits they supposedly provide in human-computer interaction. Metaphors are supposed to be a helpful tool in extreme programming and object-oriented programming, as well, but, despite the claimed benefits, are rarely used.

The paper is organized in the following manner: the theory of metaphor is explained in section 2, followed by description of metaphors in user interfaces in section 3 and in software engineering in section 4. The guidelines for choosing an appropriate metaphor are listed in section 5. A research on the identification of system functionality with the use of metaphors is introduced in section 6.

2 THE THEORY OF METAPHOR

Metaphors are constantly present, not only in verbal language, but also in our minds, even if we are usually not aware of them. Aristotle already emphasized their benefits for understanding and learning about new things and concepts. He wrote that a "metaphor consists in giving a thing a name that belongs to something else" [1].

Contemporary theories [2], [3], have provided new knowledge about our mental system, which is supposedly fundamentally metaphorical in nature. The metaphor thus does not exist only in language but rather in our mind "in the way we conceptualize one domain in terms of the other" [2]. Lakoff and Johnson divided the metaphor into two parts, where the explaining concept of a metaphor is the vehicle, or the source domain, and the concept to be explained is the tenor or the target domain. If we take a look at Shakespeare's

metaphor "Juliet is the sun", the sun is the source domain and Juliet is the target domain (and is for that reason warm, bright, the center of the solar system, etc).

The Oxford Companion to Philosophy [4] states that metaphors "are interpreted and they are interpreted differently by different readers and hearers". When a metaphor is processed in one's mental system, it is influenced by the individual's former knowledge, experiences, abstraction system and cultural context. The interpretation of a metaphor can for that reason differ significantly from person to person.

According to Antti Pirhonen [5], metaphors have a lifecycle with five stages: birth, insight, full metaphorical power, commonness and death. Soon after its birth, the metaphor reaches its full power. When its use becomes frequent, the strength fades and gradually the metaphor turns into idiom or everyday expression.

Lakoff and Johnson [2] stated that the metaphor is only partial. Not all of the properties of source domain are transferred to the target domain. In the case of Shakespeare's metaphor, only some properties of the sun are transferred to Juliet (she is not huge or burning as the sun is).

To sum up, the metaphor is a process where the concept of one thing is transferred to another (usually new or abstract) thing. This allows people to gain knowledge and insight into unknown things. Everything is interpreted in the context of an individual's mental system, as influenced by their former experiences. Because the metaphor starts in our minds, it is not merely a linguistic expression, but can exist in different forms, e.g. picture, sound, gesture.

3 METAPHORS IN USER INTERFACES

The basic characteristic of metaphors in user interfaces (UI) is the same as in the general description, i.e. the metaphor is an instrument for explaining one thing by identifying it with another. The source domain is usually an object or a concept from the real world, whereas its characteristics represent system functionality. The main benefit of metaphors in UI is (as Neale and Carroll state [6]) that "metaphors help establish user expectations and encourage predictions about system behavior".

Metaphors in UIs appeared with graphical user interfaces in the late 60's at the Xerox Palo Alto Research Center (PARC) with the computers Xerox Star and Xerox Alto [7]. After the

first commercial success with Apple Macintosh, the developers widely and sometimes irrationally (Microsoft BOB and General Magic's Magic Cap) implemented metaphors in UIs. The irrational use of metaphors led to early research and criticism. Nevertheless, metaphors kept their place in computer UIs and have also spread to nearly all devices that have graphical user interfaces: on the web, mobile phones, navigation devices, on DVD and MP3 players, home entertainment systems, video game consoles, etc.

According to Lakoff and Johnson, metaphors can be classified into 3 groups: orientational, ontological and structural. We will expand this classification to UI metaphors.

3.1 The classification of UI metaphors

Oriental metaphors, according to Lakoff and Johnson [2], "give a concept a spatial orientation". Since the UIs exist in two-and-a-half dimensional screens (due to the capability of windows overlapping), the developers can implement orientational metaphors quite frequently.

In general, orientational metaphors serve two purposes in UIs: navigation and quantification. The buttons "next" and "back" are mostly positioned next to each other, "back" being on the left and "next" on its right side. Often they are accompanied with arrows for left and right, as a visual aid for the user. Thus, the concepts of back and next have been transformed into the spatial orientation terms "left" and "right".

The "slider" and the "status bar" are examples of quantificational orientational metaphors. They provide the information of a quantity by visualizing it either in an up-down or left-right orientation. The "slider" furthermore enables users to set the amount themselves.

Ontological metaphors use the concepts of entities, substances and containers, enabling us to categorize, group, quantify and refer to experiences or activities - and thereby reason about them [2].

Since programs are generally intangible abstract concepts, the ontological metaphors in UIs can be very useful, mainly for referring, quantifying and explaining the cause. The documents can thus have a name, location and size. One can copy and paste them or throw them away. Programs can have errors in them and the system can provide information about causes for different operations.

In structural metaphors, the source is a well-defined object or concept from the real world, giving its target a specific structural meaning [2].

When the above definition is explained within the concept of UIs, the structural metaphors use real-life objects and with its structure they define a certain abstract system concept. Because they are well defined, they are the most recognizable UI metaphors: "desktop", "window", "trashcan", "menu", etc.

3.2 Criticism of the use of metaphors in UI

Criticisms on the use of metaphors in UIs can be condensed into 5 points:

1. UI metaphors can never transfer all the meaning of a concept from a real-world object to a system concept. The user can therefore expect the same features in the system as in its realistic representative, which may cause confusion [8].
2. The target domain can have extra features that its source domain does not. As a result, the users might not be aware of all of the system functionalities.
3. Metaphors cannot grow or shrink. The features of a system might well correlate with the current metaphor, but new features might not [6].
4. Metaphors lose power and meaning through time and turn into idiom.
5. The implementation of metaphors into UIs is too frequent. Even though they might be useful for novice users, they can hinder expert users [6].

As Johnson, Lakoff and others [2], [3] have stated, metaphors are a part of our mental system. We use them often and would find it difficult to express our (abstract) thoughts without them. Because computers are mainly abstract concepts, metaphors are a natural tool for explaining system functionalities.

We could presume that users would generate their own metaphors for system functionalities, if they were not any implemented already. The question is only if the user's metaphors would be as appropriate as the developer's, since users are not field experts [9].

Alty et al. [10] reject the criticism of inequality between the source and domain properties by turning the differences into a tool for measuring a metaphor's adequacy.

Pirhonen [5] believes that idioms should not be considered problematic. Novice users can benefit from dead metaphors (as they still provide them with knowledge on abstract concepts), while expert users can simply see them as idioms and are not explicitly hindered by them.

Despite all the criticisms, metaphors are still considered to be a useful tool in human-computer interaction and are for that reason implemented in nearly all user interfaces.

4 METAPHORS IN SOFTWARE ENGINEERING

In software engineering, metaphors can either provide names for programming principles and problems (the traveling salesman problem, evolutionary algorithms, genetic programming, neuron networks, decision trees, data mining and data warehousing) or they come as methods or practices (extreme programming and object-oriented programming). In this article, we will focus on metaphors as methods, since the naming of principles is merely a language issue and not considered to be a tool for the development of information systems.

4.1 Metaphors in extreme programming

One of the twelve basic practices of extreme programming is the system metaphor. It is a story that everyone - customers, programmers and managers - can tell about how the system works [11]. The system metaphor simplifies communication during development, as it suggests a common system of names, objects and inter relationships. Furthermore, it

suggests the key structure for how the problem and the solution are perceived; it shapes the systems by identifying the key objects and sometimes suggests new ideas about the system [12].

Studies of the extreme programming project have shown that the chosen metaphors were poor and underutilized during development. The metaphors are either too weak (lacking architectural structure or vocabulary) or too strong (forcing the system to adapt their features). The metaphor might also be misleading, unfamiliar or could limit the conception of the system [13].

4.2 Metaphors in Object-oriented programming

Ideas about how the development of a system is conceived and what kind of support is expected is according to Heinz Züllighoven [14], called a guiding metaphor. A guiding metaphor - or leading motif - is a theme or other coherent idea, clearly defined and named, whose purpose is to represent or symbolize a person, object, place, idea, or state of mind. It provides a common orientation for all participating groups throughout the development process and can be, for that reason, compared to the system metaphor in extreme programming.

5 GUIDELINES FOR CHOOSING AN APPROPRIATE METAPHOR

We found several articles with guidelines for choosing an appropriate metaphor in UIs. As stated above, the metaphors in UIs do not merely relate to visualization, but reveal system functions. Thus, the guidelines could be generalized and applied to software engineering as well.

Carroll et al. [6] suggest three steps: (1) think of possible candidates for metaphors (by analyzing existing programs, user observations or the invention of a new metaphor) (2) measure the appropriateness of the candidates (by defining the mismatches between the source and target domain) and (3) solve the defined mismatches.

Erickson [15], on the other hand, suggests a different approach, where the developers can find a metaphor for each system part where the functionality problem occurs. For the appropriateness measures, he suggests defining five attributes: the amount of structure, the applicability of structure, representability, audience suitability and extensibility.

Alty et al. [10] defined six steps in the user interface planning: identification of system functionality, generation and description of potential metaphors, analysis of metaphor-system pairings, implementation, evaluation and feedback on design. In step two, the generation and description of potential metaphors, several techniques are suggested: design metaphors, extending existing metaphors, brainstorming, market feedback and work-place studies.

Madsen introduced a large collection of heuristic methods for the shaping of appropriate system metaphors. For generating metaphors, he suggests listening to users, building on existing metaphors, using predecessor artifacts as metaphors, noting metaphors already implicit in the problem description and looking for real-world events

exhibiting key aspects. In the evaluation process, the developers should find a metaphor with a rich structure, evaluate the applicability of the structure and choose a metaphor suitable for the audience, with a well-understood literal meaning. The metaphor should also have a conceptual distance between the source and the target meaning, but at least one bridging concept as well. During the development of the system, he proposes the use of a key aspect based on a metaphor, the restructuring of the perception of reality, the elaboration of assumptions, the identification of the unused parts of the metaphor and the generation of conflicting parts, as well as telling the metaphors story [16].

6 VISUALIZATION OF METAPHORS

The success of the metaphor relies on its ability to communicate the right information to the user. For the UI metaphors, being mainly visual, Alty et al. point out three main aspects: realism, consistency, appropriateness.

6.1 Realism

The user should be capable of identifying system functionality quickly and easily as possible, while the more realistic a metaphor the better it is. But this is not always the case, as some metaphors represent only a definite number of real-world object properties. Thus, the realism in the visualization of a metaphor should be considered in correlation with the similarity of the source and the target domain properties.

6.2 Consistency

According to Alty et al., there are two types of consistency. Within-metaphor consistency applies to the same metaphors in different UIs, where they should carry the same system functionality (the "trashcan", should always have the same functionalities applied to it). Between-metaphor relates to different metaphors in different interfaces, which are in correlation with one another and should include the same logical behavior (the "copy-paste" metaphor should enable copying from one UI and pasting into another).

6.3 Appropriateness

An inappropriate metaphor has (usually cultural) unsuitable connotations, which interferes with the end users. It is therefore important that the developers find a metaphor in harmony with the intended user population [10].

7 IDENTIFYING SYSTEM FUNCTIONALITY WITH THE USE OF METAPHORS

In this article, have repeatedly stated that metaphors carry information about system functionality and are, because of that, a presumably helpful tool in human-computer interaction. Since there were also many criticisms leveled against their benefits, we conducted research where the goal was to find how useful metaphors in UIs really are.

For that purpose, a survey for the identification of system functionalities with the help of metaphors was conducted. We used a questionnaire with 20 metaphors and asked

software developers to identify the functionalities the metaphors implement.

The chosen metaphors were: home*, trashcan, link, F1 function key, slider, save, options, new window, rating, new comment, RSS, shopping cart*, bookmarks*, favorites*, top sites*, page menu*, show images*, browser trashcan*, dodge/burn (Photoshop) and status (MSN messenger). [Note: * - connotes metaphors in web browsers.]

7.1 Results

As we can see from Figure 1, most of the users that identified the system functionality correctly had seen the given metaphor before. Only three of the interviewees recognized the function, even though they had not seen the specific metaphor before. We can also note a large amount of incorrect answers by interviewees that had seen the given metaphor before.

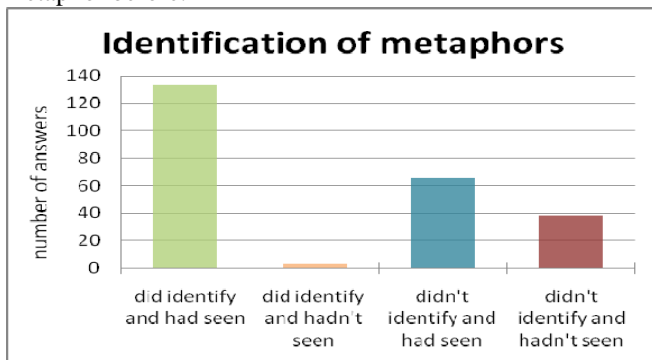


Figure 1: The amount of answers in correlation with former experiences with them

The results show the metaphor is not strong enough by itself for the user to recognize system functionality. The users did, however, identify some of system functionalities, mainly with the help of previously known metaphors.

8 CONCLUSION

Metaphors occur in almost every user interface as well as in software engineering, hence they should not be ignored during system development. Some theories emphasize their contributions in human-computer interaction, but as we have shown in this research they are not powerful enough to be used on their own for system functionality recognition purposes.

Usually the systems provide extra help and tips that help users easily comprehend and master interaction with the program. Erickson has pointed out the importance of enabling free exploration of the program, as this helps the user identify possible mismatches between real-world objects and system functionalities.

For future research, we suggest a study of functionality recognition by metaphors with a focus on user behavior while operating the system. This kind of research would unveil more detailed information on the role metaphors play in human-computer interaction.

References

- [1] L. Honeycutt, Aristotle's Rhetoric: <http://www.public.iastate.edu/~honey1/Rhetoric/rhet3-10.html>, visited: September 2009.
- [2] G. Lakoff, M. Johnson, *Metaphors we live by*, The university of Chicago Press, 2003.
- [3] M. J. Reddy, *The conduit metaphor: A case of frame conflict in our language about language*, in: A. Ortony, *Metaphor and Thought*, 2. edition, Cambridge University Press, 1993, pages: 164 - 202.
- [4] T. Honderich, *The Oxford Companion to Philosophy*. Oxford University Press, 1995.
- [5] A. Pirhonen, *To Simulate or to Stimulate? In search of the Power of Metaphor in Design*, in: A. Pirhonen, P. Saariluoma, *Future Interaction Design*, Springer London, 2005, pages: 105 - 123.
- [6] J. M. Carroll, D. C. Neale, *The Role of Metaphors in User Interface Design*, in: M. Helander, T. K. Landauer, P. V. Prabhu, *Handbook of Human-Computer Interaction*, 2. edition, Elsevier, New York, 1997, pages: 441 - 462.
- [7] J. Reimer, *A history of the GUI*, 2005: <http://arstechnica.com/old/content/2005/05/gui.ars>, visited: September 2009.
- [8] D. Saffer, *The Role of Metaphor in Interaction Design*, A thesis submitted for candidacy for the degree of master of Design in Interaction Design, The School of Design Carnegie Mellon University, Pittsburgh, Pennsylvania, 2005.
- [9] Alan Cooper, *The Myth of Metaphor*, in: *Visual Basic Programmer's Journal*, 1995.
- [10] J. L. Alty, R. P. Knott, B. Anderson, M. Smyth, *A framework for engineering metaphor at the user interface*, in: D.M. Murray, *Interacting with Computers 13*, Elsevier, 2000, pages: 301 - 322.
- [11] K. Beck, *Extreme Programming explained*, Addison-Wesley Professional, Boston, 2005.
- [12] W. C. Wake, *The system Metaphor*: <http://www.xp123.com/xplor/xp0004/index.shtml>, visited September 2009.
- [13] R. Khaled, P. Barr, J. Noble, R. Boddle, *Extreme Programming System Metaphor: A Semiotic Approach*, in: M. Billinghamurst, *Proceedings of the Sixth Australasian conference on User interface*, Newcastle, Australia, 30. 1. - 3. 2. 2005, ACM Press, 2005, pages: 109 - 117.
- [14] H. Züllighoven, *Object-oriented construction handbook: developing application-oriented software with the tools & materials approach*, Morgan Kaufmann Publishers, Heidelberg, 2005.
- [15] T. D. Erickson, *Working With Interface Metaphors*, in: B. Laurel, *The Art of Human-Computer Interface Design*, Addison-Wesley Publishing Company, Reading, 1990, pages: 65-73.
- [16] K. H. Madsen, *A Guide to Metaphorical Design*, v J. Cohen, *Communications of the ACM*, 1994, pages: 57 - 62.

ARCHITECTURE FOR SOFTWARE METRICS REPOSITORY

Črt Gerlec, Aleš Živkovič

University of Maribor

Faculty of Electrical Engineering and Computer Science

Smetanova ulica 17, 2000 Maribor, Slovenia

Tel: +386 2 220 7355; fax: +386 2 220 7272

e-mail: crt.gerlec@uni-mb.si

ABSTRACT

Software systems are becoming larger and more complex. Consequently, they are hard to maintain and upgrade. In order to better understand these systems, engineers have started to collect metrics data using special storage spaces, known as metrics repositories. With the existence of appropriate tools, large amounts of data can be processed and visually presented in order to improve design and code quality. This paper presents the design and implementation of the metrics repository for collecting and analyzing product metrics.

1 INTRODUCTION

Reviewing the quality of software products is a key factor for companies' long-term success. With a good system that collects fundamental programming data (e.g. change in source code), project managers or team leaders can assess code quality as well as recognize the change in quality of a software product as a whole. The system is known as metrics repository and its aim is to collect and store product metrics collected from software development and project management tools. The data is manually extracted from different tools that collect data from different sources (e.g. integrated development environments, test results, schedule data). Based on this data, different metrics are calculated that help a project team preserve a better overview of the project. When these metrics data are stored in a repository, we are talking about a software metrics repository [1].

2 RELATED WORK

Different types of repositories have often been described and analyzed in scientific papers. Sillitti et al. [5] presented the software metrics repository and tools for automated data extraction from different sources. The tool collects, integrates and analyzes software metrics and personal software process data.

Yap et al. [6] defined the generic data model to handle measurement data. The aim was to make software metrics definitions more robust, definable and structured.

3 PRODUCT METRICS

In every engineering discipline, measurements are an important source of information. The collected data allows engineers to understand how things work and how to make changes in order to improve the results. However, we can identify some difficulties in these measurements [2][3]:

- Collecting metrics data is time consuming. Projects are often late and there is no time to manually analyze the collected data. Collecting data in such a manner does not provide immediate benefits.
- Manual data collection produces too many errors or leads to missing data. Usually, the analysis process is not sufficient enough. The errors are mostly produced in critical periods (e.g. stressful working periods, before timelines).

Product metrics measure facts, products or document some results of the software development process [4]. Each fact or document produced during the software life cycle can be measured. In general, measurement data brings important information about code quality, processes and changes in a software product. However, there are still some open issues regarding the use of product metrics that need to be addressed [5]:

- Different metric definitions are an important issue. Even "simple" metrics (i.e. the number of methods) can be treated in different ways. For instance, how to classify constructors, inherited, overloaded, overridden methods, properties etc.
- Data integrity has been a long-term concern for software engineering. Tools are not precise enough and they cannot guarantee the correctness, completeness and self-consistency of data. This can lead to potential data ignorance. For instance, how to treat Java's anonymous inner class when calculating cyclomatic complexity of the methods that contain them.
- Tools usually generate enormous amount of data. To understand and interpret collected data, the

appropriate software metrics visualization techniques are required.

- Some methods require calibration to the target environment. The calibration usually requires a substantial amount of data.

4.1 Quality index

Besides well-known object oriented metrics (e.g. lines of code, cyclomatic complexity, maintainability index, Halstead metrics, lack of cohesion in methods, weighted methods per class, etc.), a metric for measuring software product quality, called the Quality index, has been proposed [7].

The first step in calculating a quality index is to estimate project size and effort based on the UML models and project characteristics. After the implementation phase, a code analysis is performed. The acknowledge effort, based on the code analyses and actual effort, is defined as:

$$E_{ack} = E_a * (RE_1 + (QI * RE_2))$$

where E_{ack} is the acknowledged effort (expressed in hours), E_a is actual effort, RE_1 in RE_2 are reward factors and QI is the quality index.

The Quality Index is defined as:

$$QI = \frac{\sum_{i=1}^n PMQR_i}{n}$$

$$PMQR \in \{1..5\}$$

$$PMQR_i = f_i(mv_i),$$

where $PMQR$ is the product metric quality rating, n is the number of code metrics used in the calculation, mv is the code metric value and f is the function that transforms the metric to the product metric quality rating.

The Quality index is composed of n software product metrics. The number of metrics and their type depend on the project and environment characteristics. Because all metrics have their own threshold values, these thresholds need to be estimated considering the following attributes:

- development team,
- development environment,
- domain,
- customer,
- development type.

4 SOFTWARE METRICS REPOSITORY ARCHITECTURE

To ensure software product quality and recognize code quality changes, we have decided to implement a specific software metrics repository. Besides the core functionality (collecting metric values), its purpose is to define quality index. Once it is defined, we can conduct further analysis and discover the dependability of $PMQR$ transformation functions on the final quality index value.

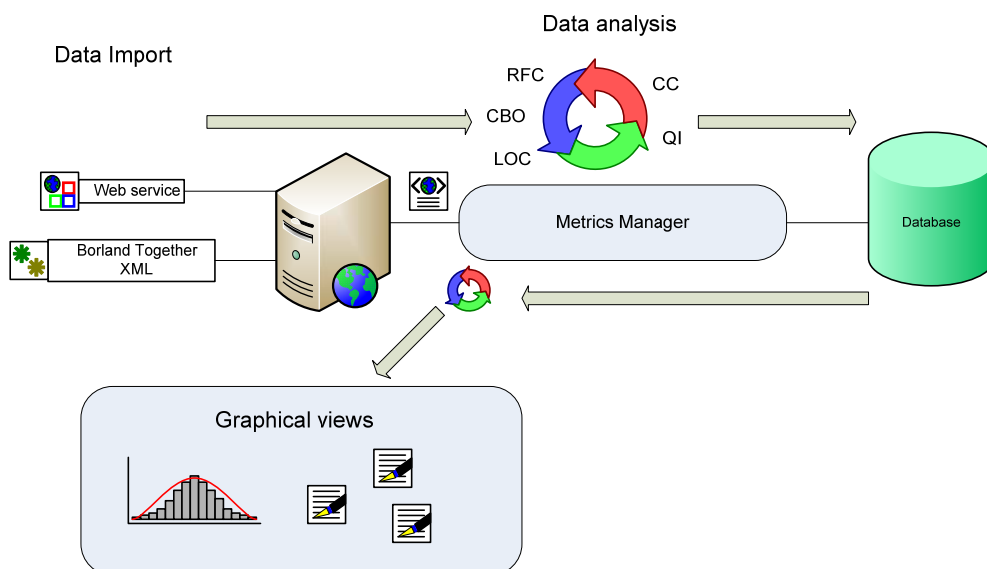


Figure 1: Software metrics repository architecture.

The metrics repository consists of different modules (Figure 1):

- metrics list
 - basic matrices
 - composed metrics
- data import
- defining quality index
 - defining PMQR functions
 - interactive quality index analysis
- presentation of metrics values through versions

4.1 Metrics list

Before the import of the metric to the repository, we have to mark each metric with a unique identifier. If the metric is not defined, it is skipped during the data import process. The repository allows the use of basic (lines of code, number of classes, number of methods) and composed metrics (MI). When C is a composed metric and A and B are the basic metrics, the repository enables custom calculations using simple mathematical operations. Currently, the repository supports all arithmetic operations (e.g. addition, subtraction, multiplication, divisions) and some advanced functions (e.g. logarithm). All composed metrics are calculated automatically during the data import process.

4.2 Defining Quality Index

The quality index plays a special role in the repository. The aim of building the repository was to enable further research related to the quality index. The quality index consists of various metrics that require transformation functions, namely the Product Metric Quality Rating (PMQR). In other words, for each metric, we have to define quality intervals. For instance, quality intervals for the maintainability index are shown in Table 1.

Maintainability index	PMQR function value
20 – 40	1
40 – 60	2
60 – 70	3
70 - 80	4
80 – 100	5

Table 1: Defining the PMQR function and Quality index.

In this case, the higher maintainability index has a higher PMQR function value (quality). However, some PMQR functions, like the one for DIT, are different.

4.3 Interactive Quality Index analyses

When all metrics and PMQR functions are defined, the tool calculates a quality index for all projects versions. However, at this point, we do not have a clear understanding about the quality index changes and how the changes in the PMQR

function affect the quality index performance. Therefore, the repository has a built-in mechanism that allows for an interactive definition of the PMQR.

To use this mechanism, the quality index has to be defined. Depending on the intervals of the PMQR function, sliders (interval changes) are dynamically generated.

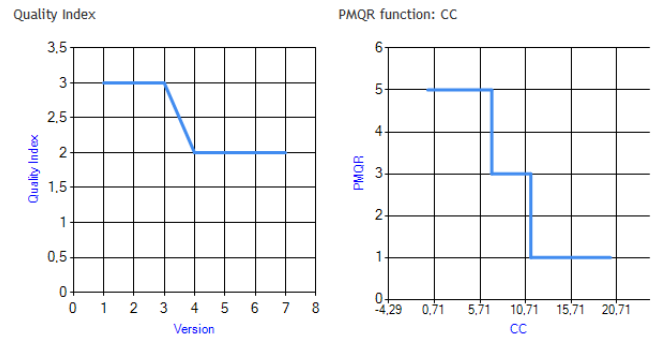


Figure 2: Interactive Quality Index analyses.

Figure 2 shows the graphical representation of the PMQR. By changing the sliders (intervals), the calculation of the final quality index is changed. Researchers and project managers are able to analyse the impact of the PMQR functions on the calculated quality index.

4.4 Data import

The repository’s primary functionality is data import. Our solution enables two options for importing data into the repository. The first option is by using web services and the second is with the results file generated by Borland Together 2005.

A unique xml schema has been designed for importing data. It consists of the following elements:

- Elements
 - ElementId
 - Name
 - Type (e.g. project, class, method)
 - ChildElements
 - MetricValues
 - MetricIdentifier (unique identifier)
 - Value

The data import process (from different sources) has a special mechanism, which transforms the original structure into the internal (xml notation). For instance, when the data import by Borland Together 2005 is performed, the unique results structure is converted to internal. However, such a procedure guarantees that the data is always imported in the same way.

4.5 Metrics value history

The repository supports the storage of software metrics for different project versions. Based on the version, the metric values are then generated and shown on the graph. Researchers and project managers can analyze the metrics history and view the metrics change between project versions. For instance, if the rise of the cyclomatic complexity is recognized, additional actions may be required to lower the complexity in the next version.

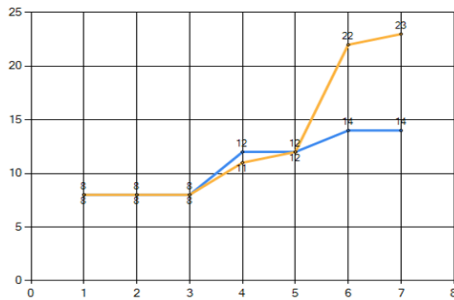


Figure 3: Metrics value history.

Figure 3 shows the metric changes collected during different project versions.

5 CONCLUSION

To manage the quality in a complex software system, the appropriate methods, metrics, repositories and tools are necessary. To ensure positive long-term business success, project managers have to maintain systems in a way that provides consistently good quality. In other words, they have to recognize potential bad-smells as well as system failures and take appropriate actions in order to avoid them.

To address all these issues, we have implemented a software metrics repository with additional features that support the analysis of the software products based on a quality index. Our further research work will be directed towards data collection and analysis based on open-source projects.

References

- [1] Warren Harrison. A flexible method for maintaining software metrics data: a universal metrics repository Towards. *Journal of Systems and Software*, Vol 72, Issue 2. 2004.
- [2] P. M. Johnson, A. M. Disney. A Critical Analysis of PSP Data Quality: Results from a Case Study. *Empirical Software Engineering*. Vol 4, Num 4. 1999.
- [3] P. M. Johnson, A. M. Disney. Investigating Data Quality Problems in the PSP. *6th ACM SIGSOFT international symposium on Foundations of software engineering*. Florida, USA.
- [4] Norman Fenton, Shari Lawrence Pfleeger. *Software Metrics: A Rigorous and Practical Approach*. Thomson Computer Press. 1994.

- [5] Marco Scottoa, Alberto Sillittib, Giancarlo Succib, Tullio Vernazza. *A non-invasive approach to product metrics collection*. *Journal of Systems Architecture*. Volume 52. Issue 11. 2006
- [6] Ng Keng Yap, Abdul Azim Abdul Ghani, Ali Mamat, Hazura Zulzalil. *The Robust Software Metric Data Model Defined in XML*. *International Journal of Computer Science and Network Security*. Vol. 8 No. 2. 2008.
- [7] Heričko, Marjan, Živkovič, Aleš, Porkoláb, Zoltán. A method for calculating acknowledged project effort using a quality index. *Informatica*. Vol. 31. No. 4. 2007.

SOLUTION REPRESENTATION ANALYSIS FOR THE EVOLUTIONARY APPROACH OF THE ENTITY REFACTORING SET SELECTION PROBLEM

Camelia Chisăliță-Crețu

Department of Computer Science, Babeș-Bolyai University
1, M. Kogalniceanu Street, RO-400084 Cluj-Napoca, Romania
Tel: +40-264-405300/5240
e-mail: cretu@cs.ubbcluj.ro

ABSTRACT

Refactoring is a commonly accepted technique that improves the structure of object oriented software. The paper presents a multi-objective approach to the Entity Refactoring Set Selection Problem (ERSSP) by treating the cost constraint as an objective and combining it with the effect objective. The results of a new solution representation for an already proposed weighted objective genetic algorithm on a experimental didactic case study are presented and compared with other previous results.

1 INTRODUCTION

Software systems continually change as they evolve to reflect new requirements, but their internal structure tends to decay. Refactoring is a commonly accepted technique to improve the structure of object oriented software [4]. Its aim is to reverse the decaying process in software quality by applying a series of small and behaviour-preserving transformations, each improving a certain aspect of the system [4]. The ERSSP is the identification problem of the optimal set of refactorings that may be applied to software entities, such that several objectives are kept or improved. The paper proposes a new solution representation for the Multi-Objective Entity Refactoring Set Selection Problem (MOERSSP) and performs the weighted objective genetic algorithm on the Local Area Network (LAN) simulation source code. Obtained results for our case study are presented and compared with already obtained results.

The rest of the paper is organized as follows: Section 2 gives the definition of the Multi-Objective Optimization Problem (MOOP), while Section 3 shortly reminds the formal definition of the MOERSSP. The proposed approach and several details related to the genetic algorithm applied to the LAN simulation are described in Section 4. The results of the experiments are presented and compared in Section 5.

2 MOOP MODEL

MOOP is defined in [8] as the problem of finding a decision vector $\vec{x}=(x_1, \dots, x_n)$, which optimizes a vector of M objective functions $f_i(\vec{x}), 1 \leq i \leq M$, that are subject to inequality constraints $g_j(\vec{x}) \geq 0, 1 \leq j \leq J$ and equality constraints $h_k(\vec{x}) = 0, 1 \leq k \leq K$. A MOOP may be defined as:

$$\text{maximize } \{F(\vec{x})\} = \text{maximize } \{f_1(\vec{x}), \dots, f_M(\vec{x})\},$$
with $g_j(\vec{x}) \geq 0, 1 \leq j \leq J$ and $h_k(\vec{x}) = 0, 1 \leq k \leq K$ where \vec{x} is the vector of decision variables and $f_i(\vec{x})$ is the i -th objective function; and $\vec{g}(\vec{x})$ and $\vec{h}(\vec{x})$ are constraint vectors.

There are several ways to deal with a multi-objective optimization problem. In this paper the weighted sum method [5] is used.

Let us consider the objective functions f_1, f_2, \dots, f_M . This method takes each objective function and multiplies it by a fraction of one, the "weighting coefficient" which is represented by $w_i, 1 \leq i \leq M$. The modified functions are then added together to obtain a single fitness function, which can easily be solved using any method which can be applied for single objective optimization. Mathematically, the new mapping may be written as:

$$F(\vec{x}) = \sum_{i=1}^M w_i \cdot f_i(\vec{x}), 0 \leq w_i \leq 1, \sum_{i=1}^M w_i = 1.$$

3 MOERSSP FORMULATION

The complete definition for the ERSSP is presented in [1]. In order to understand the problem a brief summary is given here. Let $SE = \{e_1, \dots, e_m\}$ be a set of software entities, e.g., a class, an attribute from a class, a method

from a class, a formal parameter from a method or a local variable declared in the implementation of a method. The weight associated with each software entity $e_i, 1 \leq i \leq m$ is kept by the set $Weight = \{w_1, \dots, w_m\}$, where $w_i \in [0, 1]$ and $\sum_{i=1}^m w_i = 1$. A set of possible relevant chosen refactorings [1] that may be applied to different types of software entities of SE is gathered up through $SR = \{r_1, \dots, r_t\}$. There are various dependencies between such transformations when they are applied to the same software entity, a mapping emphasizing them being defined by: $rd : SR \times SR \times SE \rightarrow \{Before, After, AlwaysBefore, AlwaysAfter, Never, Whenever\}$.

The effort involved by each transformation is converted to cost, described by the function $rc : SR \times SE \rightarrow N$. Changes made to each software entity $e_i, i = \overline{1, m}$ by applying the refactoring $r_l, 1 \leq l \leq t$ are stated and a mapping is defined: $effect : SR \times SE \rightarrow Z$. The overall effect of applying a refactoring $r_l, 1 \leq l \leq t$ to each software entity $e_i, i = \overline{1, m}$ is defined by the mapping $res : SR \rightarrow Z$.

The purpose is to find a subset of entities $ESet_l$ for each refactoring $r_l \in SR, l = \overline{1, t}$ such that the cost objective is minimized and the overall effect objective is maximized. The ERSSP is the identification problem of the refactorings that may be applied to each software entity such that the proposed objectives are kept or improved.

Multi-objective optimization often means compromising conflicting goals. For our MOERSSP formulation there are two objectives taken into consideration in order to maximize refactorings effect upon software entities. The first objective function minimizes the total cost. In order to have an maximized objective, it was subtracted from MAX , the biggest possible total cost, as:

$$maximize \{f_1(\vec{r})\} = maximize \left\{ MAX - \sum_{l=1}^t \sum_{i=1}^m rc(r_l, e_i) \right\}, \vec{r} = (r_1, \dots, r_t)$$

The second objective function maximizes the total effect of applying refactorings upon software entities, considering the weight of the software entities in the overall system, like:

$$maximize \{f_2(\vec{r})\} = maximize \left\{ \sum_{l=1}^t res(r_l) \right\}, \vec{r} = (r_1, \dots, r_t)$$

The goal is to identify those solutions that compromise the refactorings costs and the overall impact on transformed entities. The final fitness function for MOERSSP is defined by aggregating the two objectives and may be written as:

$$F(\vec{r}) = \alpha \cdot f_1(\vec{r}) + (1 - \alpha) \cdot f_2(\vec{r}), \quad 0 \leq \alpha \leq 1.$$

4 PROPOSED APPROACH DESCRIPTION

The decision vector (chromosome) $\vec{S} = (S_1, \dots, S_m)$, $S_i \subseteq SR \cup \phi, 1 \leq i \leq m$ determines the refactorings that may be applied in order to transform the proposed set of software entities SE . The item S_i on the i -th position of the solution vector represents a set of refactorings that may be applied to the i -th software entity from SE , where each entity $e_u \in SE_i, S_i \subseteq SR \cup \phi, 1 \leq u \leq q, 1 \leq q \leq m, 1 \leq i \leq t$.

This means it is possible to apply more than once the same refactoring to different software entities, i.e., distinct gene values from the chromosome may contain the same refactoring. Besides the solution representation proposed in [1], a new solution representation is presented here. A steady-state evolutionary algorithm was applied here, a single individual from the population being changed at a time. Both rc and res mappings are normalized to fit between 0.0 and 1.0; decimal scaling for the rc mapping and min-max normalization for the value of the res function have been applied here.

The parameters used by the evolutionary approach are as follows: mutation probability 0.7 and crossover probability 0.7. Different number of generations and of individuals are used: number of generations 10, 50, 100, 200 and number of individuals 20, 50, 100, 200. The value of α parameter used while aggregating the objectives was set to 0.5 which gives the same importance to both objectives.

A simple one point crossover scheme is used. A crossover point is randomly chosen. All data beyond that point in either parent string is swapped between the two parents. Mutation operator used here exchanges the value of a gene with another value from the allowed set. In other words, mutation of i -th gene consists of adding or removing a software entity from the set that denotes the i -th gene. We have used 11 mutations for each chromosome.

The algorithm proposed was applied on a simplified version of the Local Area Network (LAN) simulation source code that was presented in [2]. Figure 1 shows the class diagram of the studied source code. It contains 5 classes with 5 attributes and 13 methods, constructors included.

Thus, for the studied problem the software entity set is defined as: $SE = \{c_1, \dots, c_5, a_1, \dots, a_5, m_1, \dots, m_{13}\}$. The chosen refactorings that may be applied are: *renameMethod*, *extractSuperClass*, *pullUpMethod*, *moveMethod*, *encapsulateField*, *addParameter*, denoted by the set $SR = \{r_1, \dots, r_6\}$ in the following. The dependency relationship between refactorings is defined as follows: $\{(r_1, r_3) = B, (r_1, r_6) = AA, (r_2, r_3) = B, (r_3, r_1) = A, (r_6, r_1) = AB, (r_3, r_2) = A, (r_1, r_1) = N, (r_2, r_2) = N, (r_3, r_3) = N, (r_4, r_4) = N, (r_5, r_5) = N, (r_6, r_6) = N\}$.

The values of the final effect were computed for each refactoring, but using the weight for each existing and possible affected software entity, as it was mentioned in Section 3. Therefore, the values of the res function for each refactoring are: 0.4, 0.49, 0.63, 0.56, 0.8, 0.2.

The current version of the source code lacks of hiding information for attributes since they are directly accessed by clients. The abstraction level and clarity may be increased by creating a new superclass for `PrintServer` and `FileServer` classes, and populate it by moving up methods in the class hierarchy.

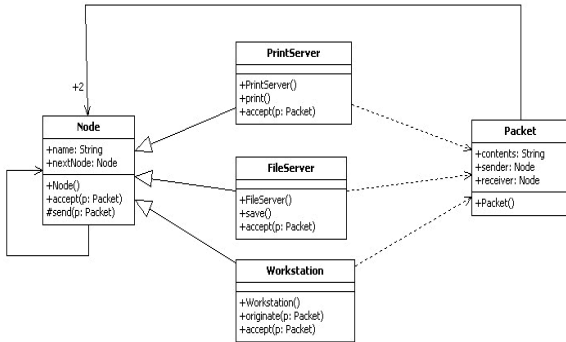
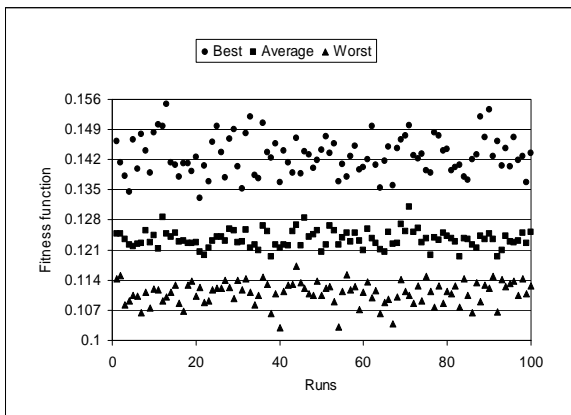
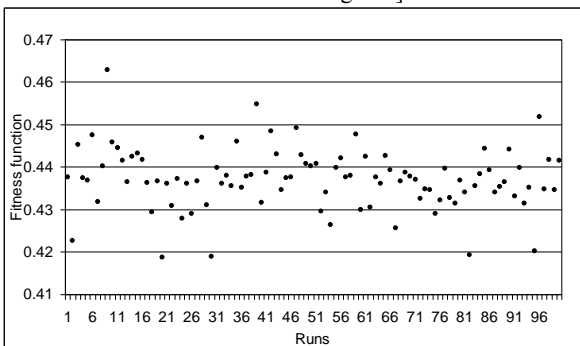


Figure 1: Class diagram for LAN simulation

5 OBTAINED RESULTS BY THE EVOLUTIONARY APPROACH



(a) [Experiment with 10 generations and 20 individuals with eleven mutated genes]



(b) [Experiment with 200 generations and 100 individuals]
Figure 2: The evolution of fitness function (best, worse and average) and the best individual evolution

The algorithm was run 100 times and the best, worse and average fitness values were recorded. The experiment proposes equal weights, i.e., $\alpha = 0.5$, for the studied fitness function, as: $F(r) = 0.5f_1(r) + 0.5f_2(r), r = (r_1, \dots, r_m)$.

Figure 2a presents the evolution of the fitness function (best, worse and average) recorded for each run within 20 chromosomes populations and 10 generations. The evolution of the fitness function (recorded for the best individual in each run) for 200 generations and 100 individuals is depicted in Figure 2b.

The experiment has shown there are small differences among the results obtained in all the 100 runs. In the 200 generations experiments for 100 chromosomes populations the greatest value of the fitness function was 0.16095 (with 96 individuals with the fitness >0.15) while in the 200 evolutions experiments for 200 individuals populations the best fitness value was not more than 0.15995 (82 individuals with the fitness >0.15). But the best chromosome was found in the experiment with 200 generations and 20 individuals having the value 0.17345 (all individuals with the fitness 0.15). Figure 3 shows the evolutions of the number of chromosomes for different number of populations and individuals.

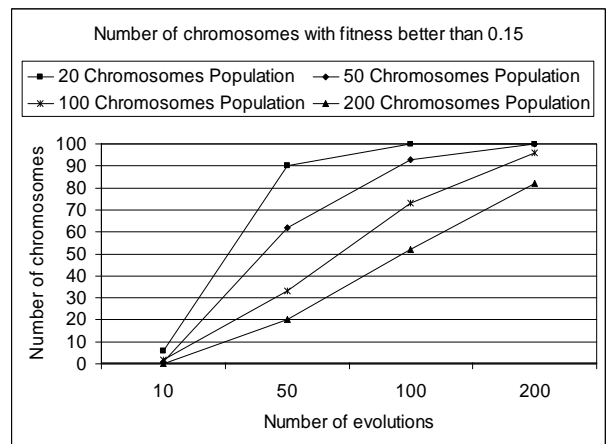


Figure 3: The evolution of the number of chromosomes with fitness better than 0.439 for the 20, 50, 100 and 200 individuals populations, with $\alpha = 0.5$

The best individual obtained by this solution representation suggests several improvements that may be applied in order to improve the structure of the class hierarchy. But, they do not shape a coherent refactoring strategy, suggesting just some refactorings that cannot be connected one to another. Thus, in order to enhance polymorphic behaviour, for the `PrintServer` and `FileServer` classes a new base class called `Server` is created. More, the signature of the `print` method from the `PrintServer` class should be changed and then renamed in order to allow the `accept` method to be pulled up to the new base class. The `save` method signature from the `FileServer` class should be

changed and then renamed too. But, the studied best chromosome suggests only the pulling up methods aspect and no rename method. The correct access to the class fields is accomplished by encapsulating them within their classes. Thought, the current solution representation suggests the `encapsulateField` refactoring for two of the five `public` class attributes of the studied source code.

5.1 Discussion

Another solution representation for the current problem is discussed thoroughly in [1]. The chromosome structure of the solution proposed in this paper is an *entity based* one, having a large number of genes. The chromosome presented in [1] being similar to the former, is a *refactoring based* one but having a lower number of genes. Therefore, the recommended refactorings proposed here may not shape a relevant strategy compared to the other solution representation. Thus, in [1] the best individual was obtained for 200 generations for 20 chromosomes populations, with a fitness value of 0.4793. For the current solution representation, for the recorded experiments the best chromosome was obtained for 200 generations and 20 individuals, with a fitness value of 0.17345. The best chromosomes studied may be transposed from a representation to another, which means their fitness values may be compared and their efficiency evaluated. Smaller individual populations produce better individuals (as number and quality) than larger ones, due to the poor populations diversity itself. But, the large number of genes of the individual structure induces poor quality to the current solution representation. Further analysis and experiments with different values for the genetic algorithm parameters may be done, e.g., $\alpha = 0.3$ or $\alpha = 0.7$.

5.2 Related Work

Fatiregun et al. [3] applied genetic algorithms to identify transformation sequences for a simple source code, with 5 transformation array, whilst we have applied 6 distinct refactorings to 23 entities. Seng et al. [7] apply a weighted multi-objective search, in which metrics are combined into a single objective function. An heterogeneous weighed approach is applied here, since the weight of software entities in the overall system and refactorings cost are studied.

Mens et al. [6] propose the techniques to detect the implicit dependencies between refactorings. Their analysis helped to identify which refactorings are most suitable to LAN simulation case study. Our approach considers all relevant applying of the studied refactorings to all entities.

6 CONCLUSIONS AND FUTURE WORK

The paper discusses a new solution representation of the ERSSP. The results for an weighted approach ($\alpha = 0.5$) are presented and compared with another decision vector representation studied in [1]. Some related work are

mentioned too. The weighted multi-objective optimization is discussed here, but the Pareto approach may prove to be more suitable when it is difficult to combine heterogenous fitness functions into a single overall objective function. Thus, a further step would be to apply the Pareto front approach in order to prove or deny the superiority of the second possibility. Here, the cost is described as an objective, but it can be interpreted as a constraint, with the further consequences.

References

- [1] C. Chisăliță-Crețu. A Multi-Objective Approach for Entity Refactoring Set Selection Problem, in *Proceedings of the 2nd International Conference on the Applications of Digital Information and Web Technologies (ICADIWT 2009)*, London, UK, August 4-6, 2009, *accepted paper*.
- [2] S. Demeyer, D. Janssens, T. Mens. Simulation of a LAN, *Electronic Notes in Theoretical Computer Science*, 72 (2002), pp. 34-56.
- [3] D. Fatiregun, M. Harman, R. Hierons. Evolving transformation sequences using genetic algorithms, in *4th International Workshop on Source Code Analysis and Manipulation (SCAM04)*, Los Alamitos, California, USA, IEEE Computer Society Press, 2004, pp. 65-74.
- [4] M. Fowler. Refactoring: Improving the Design of Existing Software. Addison Wesley, 1999.
- [5] Y. Kim, O.L. deWeck. Adaptive weighted-sum method for bi-objective optimization: Pareto front generation, in *Structural and Multidisciplinary Optimization*, MIT Strategic Engineering Publications, 29(2), 2005, pp. 149-158.
- [6] T. Mens, G. Taentzer, O. Runge, Analysing refactoring dependencies using graph transformation, *Software and System Modeling*, 6(3), 2007, pp. 269-285.
- [7] O. Seng, J. Stammel, D. Burkhart, Search-based determination of refactorings for improving the class structure of objectoriented systems, in *Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation*, M. Keijzer, M. Cattolico, eds., vol. 2, ACM Press, Seattle, Washington, USA, 2006, pp. 1909-1916.
- [8] E. Zitzler, M. Laumanns, L. Thiele, SPEA2: Improving the Strength Pareto Evolutionary Algorithm, *Computer Engineering and Networks Laboratory*, Technical Report, 103(2001), pp. 5-30.

ORGANIZING WEB SERVICES INTERFACES AS A BASIS FOR EFFICIENT SERVICE GOVERNANCE

Aleš Frece

University of Maribor, Faculty of Electrical Engineering and Computer Science
Smetanova 17, 2000 Maribor, Slovenia; e-mail: ales.frece@uni-mb.si

ABSTRACT

Core building blocks in service oriented environment are services. On one side service providers provide services and on the other side service consumers consume them. Junction point between these two is service interface. Service interface represents agreement between service provider and service consumer on what functionality service provides, in what form should consumer pass its requests for service to perform requested functionality and in what form it will accept expected result. This agreement has to be stable for service consumers to communicate with consumed service in an undisturbed manner. But in practise this is not always the case. So, how to efficiently handle these cases?

As any IT environment service oriented environment is also a subject of constant changes because services have to change often on account of emerging software requirements. After changes are introduced to a service it is necessary to inform all parties that introduced change concerns, that is all relevant service consumers. They have to assess change impact on them and adapt accordingly if necessary. The biggest challenges here are the way to define interfaces to minimize effort for service consumers to adapt to as broad range of changes as possible and to manage interface change introduction (part of service governance) for minimal impact on service consumers.

1 INTRODUCTION

With a right service interface definition and with taking into account on agreed service interface change introduction approach it is possible to automate change management. This means less human effort on service change introduction. For a good communication between service interface architect, implementators of services and implementators of service consumers it is necessary for a service interface to tell as much as possible about itself and about a service it describes. This goal can be achieved with bringing in interface elements naming and other conventions.

Service changes can influence consumers directly. This reflects through service interface change. Influence can also be indirect (service interface stays intact but service implementation changes, e.g. service functional change

requires consumer to semantically use service in a different way) or there can be no impact on consumers (e.g. only service implementation flaw is fixed with no impact on service interface or service semantics). In service oriented architecture services are implemented with means of a web service. In this case their interface is described through WSDL (Web Service Description Language) document.

2 TYPES OF SERVICE INTERFACE CHANGES

There are, roughly speaking, two types of changes that can be introduced into WSDL document that will not have impact on service consumers [1]. First is adding a WSDL operation (if a consumer does not need this new operation, it does not have to adapt) and the second adding new XML (eXtensible Markup Language) types that will not be contained in the existing XML types. This means that these new types do not even have impact on existing operations (e.g. new types are used only in newly added operation). These changes are called backwards-compatible changes [2]. However, Bechara [3] mentions that adding an operation can cause compatibility with some types of clients to break. These clients check congruity of interface they are able to call with the interface they are about to call each time they invoke a service to prevent semantically false invocations. In practise these clients are more an exception than a rule because of performance reasons [3]. All others changes to WSDL document are called non-backwards-compatible changes [2]. They have impact on service consumers in some way, so consumers have to adapt to these changes if they want to use modified service. Non-backwards-compatible changes include deleting and/or renaming an operation, changing operation parameters (their order or type) and changing one or more types that are used by one or more existing operations [2]. These non-backwards-compatible changes demand efficient data type or XML schema versioning [4], but this issue is too complex to be addressed in this article. However, Bechara [3] identifies one special case of XML schema alternation that does not break compatibility with service consumer. It is about a case where only optional elements are added into the XML schema.

Regardless of the service interface changes impact on service consumers it is necessary to mark these changes in some way so it is evident that they were made. Procedure of marking changes to a WSDL document is called WSDL

document versioning. It is possible to version WSDL documents inside the document itself with a help of XML namespaces (targetNamespace attribute) or to store WSDL document in a repository and mark it with version data as WSDL document's metadata (this leaves targetNamespace attribute intact).

3 EXISTENT WSDL DOCUMENT VERSIONING APPROACHES

Bechara [3] suggests major-minor version approach. Minor version is of two types. It can tell whether a service flaw was fixed or service efficiency was improved. This type of change does not imply any service interface change. On the other hand minor version can tell whether a new operation was added which as a rule does not have impact on service consumers. It is possible to distinguish between these types of minor versions with version mark. For example, when a first type of change is introduced it is possible to increase second version decimal place (e.g. from 1.02 to 1.03) and for the second type first version decimal place is increased (e.g. from 1.02 to 1.12). Major version is reserved for non-backwards-compatible changes. Major release is also a service with intact interface which implements a (semantically) different functionality. Consumers without adaptation cannot use this kind of services semantically correct any more. Bechara [3] speaks in favour of versioning exclusively with a help of a service repository like UDDI (Universal Description Discovery and Integration) registry.

Likewise Trenaman [5] suggests major-minor versioning but does not distinguish between different minor versions and does not use service repository. Trenaman uses WSDL elements names and targetNamespace attribute instead. Next minor version includes whole WSDL document of former version and at the same time it introduces all changes that do not break backwards compatibility. Major version is put into WSDL's targetNamespace attribute. XML schema's targetNamespace attribute holds not only major but also minor version. Trenaman [5] advises inclusion of both versions into wsdl:portType name and consequently into wsdl:service name (it suggests for every wsdl:portType there should be one wsdl:service). Like Bechara [3], Trenaman [5] proposes new major version if at least one existing operation is semantically changed.

A combination of some sort is proposed by Brown [2] that recommends repository usage only for backwards-compatible changes (repository increases WSDL document minor version in its metadata). At the same time Brown [2] recommends minor version inclusion into a WSDL's XML comment. Of course this kind of comment is not necessary but it could help implementators to identify service version form WSDL document that was not acquired directly from repository. For marking non-backwards-compatible changes Brown [2] advises usage of unique XML namespace because it is send with every input or output service message. In this way service knows what to do with

each version of request it receives. For version denotation Brown [2] suggests usage of number (e.g. X/Y) or date (e.g. year/month/day) mark as a suffix to WSDL's targetNamespace attribute.

4 BEST WSDL DOCUMENT VERSIONING APPROACH SELECTION AND ITS IMPROVEMENT

From the presented WSDL document versioning approaches one will be selected as a basis for improvement. Selected approach will be improved with better solutions from non-selected ones and some of new suggestions will be introduced to get a new improved approach.

4.1 Best WSDL document versioning approach selection

Generally the most appropriate and complete approach is Brown's approach [2] because it merges strengths of both Bechara [3] and Trenaman [5]. So the new approach will base o Brown [2]. His approach allows number and date as version mark but Bechara [3] and Trenman [5] suggest only number mark because the date mark is as a W3C recommendation reserved for XML schema versioning. So numbers for major version mark in WSDL's namespace for non-backwards-compatible changes will be selected. Repository and its metadata will be selected to mark backwards-compatible changes as minor version.

4.2 WSDL document versioning approach improvement

In practice it is possible to notice that service endpoints are a subject of most often changes, especially when staging environments are used within service development. Endpoint change reflects wsdl:port element change, particularly its soap:address sub-element (in most common cases SOAP over HTTP is used as a binding mechanism). In this way through service development one new WSDL document per transition between staging environments would emerge. These WSDL documents would describe the same service but each document in a separate environment (development, test, production). This is why it makes sense to figure out which WSDL definition elements are environment independent (lets call these definitions abstract part of the WSDL) and which are specific to a single environment (lets call them concrete part of the WSDL).

It is possible to state that in abstract part it is possible to describe (wsdl:definitions) an abstract web service interface (wsdl:portType) that includes operations (wsdl:operation). Operations have inputs (wsdl:input) and they can also have ouptus (wsdl:output). Inputs and outputs are realized with messages (wsdl:message) that can be composed of one or more parts (wsdl:part). Each part is realized with one of XML elements or XML data types defined (xsd:element) in or imported (xsd:type) into the data definition segment (wsdl:types/xsd:schema).

In the concrete WSDL part there is an approach missing: in which way input and output messages will be encoded for a transportation “on the wire” (wsdl:binding), web service (wsdl:service) that implements (wsdl:port) in abstract WSDL part defined interface (wsdl:portType) and that will be available on some endpoint address (soap:address).

This division of WSDL (abstract and concrete part) can be considered in the manner of WSDL namespace versioning for non-backwards-compatible changes. This means each part will be versioned in a separate WSDL document with a help of their targetNamespace attributes:

- Abstract part that is recognized by containing wsdl:portType (and at the same time does not contain wsdl:binding nor wsdl:service) has to include major version (mark on the end of namespace in a form of a string »_vX« where X is major version, e.g. »_v1«).
- Concrete part that is recognized by containing wsdl:binding and/or wsdl:service (and at the same time does not contain wsdl:portType) has to include not only major but also minor version (mark on the end of namespace in a form of a string »_vX_Y« where X is major version and Y minor version, e.g. »_v1_0«).

Therefore if a WSDL document contains wsdl:portType it means it is an abstract part and has to include only major version but if it contains wsdl:binding and/or wsdl:service it means it is a concrete part and has to include not only major but also minor version. In this way wsdl:portType cannot be in the same WSDL document as wsdl:binding and/or wsdl:service because this document would have to respect both version rules (it should include only major and at the same time major-minor version) which it obviously cannot. Conditions for abstract (only »_vX«) and concrete (only »_vX_Y«) namespace suffix are evidently exclusive, which means that with namespaces conditions abstract-concrete separation rule is automatically respected. Each part of a WSDL document is then as a separate WSDL document put into repository where it can additionally be versioned with repository’s metadata for backwards-compatible changes.

4.3 Inclusion of naming conventions

With correct WSDL documents versioning their unique recognition and self-explanation is enabled. Similarly it can be achieved with the elements they contain. For naming the contained elements there are many naming conventions propositions – for each purpose a different proposition (e.g. Ananiev [6] focuses on web service implementation in Java so it applies Java naming conventions to WSDL elements). Moreover, it stays each organization’s responsibility to define its own naming rules taking into account previously used naming conventions. Proposed approach will focus only on basic rules for WSDL elements naming that will assure self-explanation and at the same time they will take object-oriented naming conventions in [7] into consideration. Let us define the following naming rules:

- Names of wsdl:portType, wsdl:binding, wsdl:service and wsdl:port have to begin with a capital letter because they define a static entity (e.g. after WSDL

transformation into object-oriented language implementation this is a class name),

- Name of wsdl:operation has to begin with a small letter because it defines a dynamic entity, an action (after WSDL transformation into object-oriented language implementation this is a method name).

Let us add some rules for a better self-explanation of the element names through which it will be possible to distinguish between them (e.g. so it is possible to distinguish what is a wsdl:portType and what is a wsdl:service only by the name of the element):

- wsdl:portType name should have as meaningful name as possible,
- wsdl:binding, wsdl:service and wsdl:port names have to arise from a whole wsdl:portType name they relate to and at the end they have to append an appropriate suffix (Binding, Service, Port) that tells us what is the element behind the name (e.g. if wsdl:portType is OneInterface, then the rest are OneInterfaceBinding, OneInterfaceService and OneInterfacePort),
- wsdl:message, wsdl:input and wsdl:output have to arise from a whole wsdl:operation name that uses them and at the end they have to append an appropriate suffix (RequestMessage, ResponseMessage, Request, Response) that indicates in which direction they are sent (e.g. wsdl:operation is operName, input wsdl:message is operNameRequestMessage and output wsdl:message is operNameResponseMessage, wsdl:input and wsdl:output are operNameRequest and operNameResponse).

With enforcing these rules it is possible to achieve the goal of a better WSDL element self-explanation.

5 DEFINED APPROACH AND RULES REALIZATION WITH OPEN STANDARDS

As early as web service identification this service must be led to governance lifecycle. Before this step is taken this web service has to have an interface defined and it is very important to check its definition correctness. Interface definition correctness includes WSDL document separation into abstract-concrete parts and conformance with not only naming convention but also namespace versioning rules.

5.1 XML schema limitations

WSDL that defines web service interface is actually an XML dictionary. XML dictionary restrictions are expressed with XML schemas (also called XSDs). XML schema focuses on dictionary’s grammar. It defines where in the tree structure elements should reside, which types they should include etc. For realization of the restrictions in the WSDL documents the article proposes a powerful expressiveness is needed not in structure that is already defined within the WSDL specification but in the naming of the elements. However, XML schemas with XPath usage enable us to express patterns within strings but these restrictions have to be expressed inside WSDL elements

definition. This is apparently not possible without extending the WSDL specification but by doing this the compatibility with WSDL implementation engines is lost. That is why an alternative language in a separate document is needed to express restrictions on WSDL element names.

5.2 Schematron as an alternative to XML schema

To realize defined WSDL document separation into abstract-concrete parts, naming convention and namespace versioning rules it is possible to use another XML dictionary called Schematron. Schematron is a language for making restriction assertions on patterns found in XML files [8]. It is a part of ISO open standard called DSDL [9]. DSDL joins and enables cooperation of many XML dictionaries that are focused on XML document validation. Schematron offers some open source implementations. One offered implementation is declarative in a form of an XSLT file and another two are procedural in Python and Perl programming languages. With a Schematron's approach through pattern search it is possible to express restrictions on XML documents independent of their structure. Additionally there is no need to learn new skills to use Schematron comparing to XSDs if knowing of Schematron's relatively simple schema is omitted. After Schematron schema definition standard's implementation searches for nodes within the target XML document for pattern match (expressed with XPath). Implementation then checks if found nodes meet the rules (also in XPath) and informs us about any irregularities. Rules for WSDL document namespace versioning defined earlier can be expressed with Schematron and XPath 1.0 or 2.0 like in Figure 1.

```
<?xml version="1.0" encoding="UTF-8"?>
<schema xmlns="http://purl.oclc.org/dsdl/schematron">
  <ns uri="http://schemas.xmlsoap.org/wsdl/" prefix="wsdl"/>
  <pattern name="ServiceVersioning">
    <rule context="//wsdl:portType/..">
      <assert id="abstract-WSDL-part-version"
        test="number(substring-after(@targetNamespace, '_v'))">
        Major version (like "_v1") has to be at the end of
        targetNamespace of this WSDL document
        (<value-of select="@targetNamespace"/>),
        if WSDL includes portType
        (<value-of select="wsdl:portType/@name"/>).
      </assert>
      ...
    </rule>
  </pattern>
</schema>
```

Figure 1: Versioning rule enforcement with Schematron.

In a similar way it is possible to express earlier defined naming conventions (see Figure 2). This Schematron file fragment shows only a rule for when to use capital letters (first rule) and an example of keyword checking (second rule for checking that wsdl:service contains string »Service« at the end of the element name).

6 CONCLUSION

Before starting with web service governance WSDL documents that define interfaces of the web services have to meet some conditions. Only in this way service governance can be efficient. Article proposes some of these conditions

which are a basis of a new approach to efficient web service development management and governance. This basis of a new approach includes separation of WSDL document into abstract-concrete parts, namespace naming rules for a good versioning approach and WSDL elements naming conventions for good self-explanation of their names. WSDL document separation is implicitly included into proposed versioning rules. Article also proposes a new approach where all these rules are expressed with an open standard Schematron.

```
<?xml version="1.0" encoding="UTF-8"?>
<schema xmlns="http://purl.oclc.org/dsdl/schematron">
  <ns uri="http://schemas.xmlsoap.org/wsdl/" prefix="wsdl"/>
  <pattern name="ServiceNaming">
    <rule context="//wsdl:portType|//wsdl:binding|//wsdl:service|//wsdl:port">
      <report id="capital-letter-in-name" test="fn:matches(@name, '^[A-Z]')">
        Name of portType, binding, service and port (<value-of select="@name"/>)
        must start with capital letter.
      </report>
    </rule>
    ...
    <rule context="//wsdl:service">
      <report id="keyword-at-the-end-of-service" test="contains(@name, 'Service')
        and string-length(substring-after(@name, 'Service')) = 0">
        Service's name (<value-of select="@name"/>) must end with keyword "Service"
      </report>
    </rule>
    ...
  </pattern>
</schema>
```

Figure 2: Naming convention enforcement with Schematron.

In this way article addresses some of the biggest web service change management and governance challenges. With rules in Schematron form the proposed new approach provides a basis for web services interface governance automation. At the same time proposed approach brings some of the best practices into governance procedure. By doing all that it is possible to reduce a necessary web services change introduction effort and improve communication between web service interface architects and web service's implementors on one side and web service client's implementors on the other side.

References

- [1] N. Bieberstein, *Service-Oriented Architecture Compass: Business Value, Planning, and Enterprise Roadmap*, FT Prentice Hall, 2005, ISBN: 0131870025
- [2] K. Brown, M. Ellis: *Best practices for Web services versioning*, IBM, January 2004
- [3] G. Bechara: *Web Services Versioning*, Oracle
- [4] C. Peltz, A. A. Subbarao: *Design Strategies for Web Services Versioning*, SOA World Magazine, April 2004
- [5] A. Trenaman: *WSDL Versioning Best Practise*, IONA Technologies, April 2007
- [6] A. Ananiev: *WSDL Naming Conventions*, MyArch Solutions, April 2007
- [7] *Naming Conventions – Naming Scheme 1*, <http://www.atomicobject.com/pages/Naming+Conventions>
- [8] *Schematron: A language for making assertions about patterns found in XML documents*, <http://www.schematron.com>
- [9] *DSDL: Document Schema Definition Languages*, <http://www.dsdl.org>

INTEGRATION OF FULL-TEXT SEARCH WITH WEB MAPPING SERVICES

David Procházka, Jan Kryštof, Arnošt Motyčka

Department of Informatics

Faculty of Business and Economics

Mendel University of Agriculture and Forestry in Brno

Zemědělská 1, 61300 Brno, Czech Republic

Tel: +420 545132240; fax: +420 545132245

e-mail: david.prochazka@mendelu.cz

ABSTRACT

There is a debate about the future shape of the *Web Map Service* standard. Different projects aimed on searching in map services are of interest. A still unsolved issue is the interconnection method of the searching and the map service. There are three main approaches: searching could be integrated in the *Web Map Service* standard, it could be a separate standard or there could be a searching engine without any connection to any standard. Each approach has its benefits and also its flaws. This paper compares the mentioned perspectives and proposes a possible solution.

1 INTRODUCTION

The amount of web map services and map layers provided by them is constantly growing. Frequently there is a problem to find an appropriate map service or even to locate a required layer within the service. One of the causes is the architecture of the *Web Map Service* standard (WMS).

There is no searching engine that allows the user to locate a map service with required map layers. The user is forced to browse through different catalogs with many obsolete records and web pages of institutions providing the geodata. Moreover, in the WMS there is only one request providing the information about the map service (*GetCapabilities*). This request is without any options and returns all information about the map service at one dash. Then the user is forced to parse a significant amount of data to locate the required layer. This situation is becoming unbearable with a growing number of map layers.

It is possible to say that there is a significant need for a searching infrastructure that enables us to locate required map layers. Such an infrastructure is also one of the priorities of the European Union project INSPIRE (<http://inspire.jrc.ec.europa.eu/>).

2 RELATED WORK

We could recognize several approaches in geodata retrieval. The first one is development of different catalogs. These catalogs usually contain own metadata describing free web

mapping services or paid geodata sources. Currently, one of the most promising projects is the *INSPIRE Geoportal* for European Union (<http://www.inspire-geoportal.eu>).

Another approach to aggregation of WMS is presented by *Geospatial Information Database (GIDB) Portal System* of the *Naval Research Laboratory's (NRL) Mapping Branch* [1]. The key part of their system is an agent searching in the Internet for the map services. A discovered service is indexed and it is created a WMS interface to this map service. A similar project is the *GeoBrain* [2].

Our research is partially similar to the *GIDB* project. We developed a virtual mapserver providing a WMS compliant interface to aggregated map services (see [3], where you can find a review of mentioned projects as well). The current research is focused on methods of WMS full-text search [4].

3 WEB MAP SERVICE

There are many web services focused on distribution of geodata. Probably most spread are the *Web Map Service* (WMS) and *Web Feature Service* (WFS) standards.

The basic purpose of the *Web Map Service* (WMS) is to provide raster images with geodata on a specified location (see *Web Map Server Specification* [5]). Requests are usually sent through HTTP protocol.

WFS distributes the geodata in vector form. WFS is a generally more complex standard. It allows to update data on a server or to process provided data according to the requirements of the client [6]. For the purpose of this article we will focus on WMS. WMS supports three basic requests.

GetCapabilities – This request is mandatory (must be supported by all WMS compliant mapservers). The response on this request is an XML file called a capabilities file. In the first part of the XML file is described the mapserver itself – supported output formats, URLs for sending requests, etc. The second part describes all map layers – their coverage, name, content, scale, etc. Example of the request sent through GET method of HTTP protocol follows.

```
http://server.com/cgi-bin/mapservice?  
service=WMS&version=1.1.1&request=GetCapabilities&
```

The second mandatory request is *GetMap*. The response is an image file with a map layer or layers. The image is given by the layers themselves and by the requested geodetic system, bounding box and finally properties of the target image file (format and resolution).

The last and optional request is *GetFeatureInfo*. If the *queriable* attribute is set to "true", this request is supported. In this case, the request returns information about the feature in the map layer specified in URL on the given location.

4 WEB SERVICES STANDARDS

As written in W3C description, WSDL is an XML format for describing network services as a set of endpoints operating on messages containing either document-oriented or procedure-oriented information. WSDL is a general format, however, usually is used in conjunction with SOAP, HTTP and MIME. For further details see specification on [7].

SOAP (*Simple Object Access Protocol*) is a protocol for exchanging information between the web services. SOAP envelopes transported data in an XML message. Such messages are in a standardized machine-parseable format, therefore they could be easily processed by any service. Specification is available on [8].

Standards mentioned above are the generally accepted framework for implementation of the web services. Nevertheless, they are not widely supported by mapping services. One of major reasons is the age of mapping services. They originated before the boom of web services and proliferation of related standards. Currently there is a pressure on adoption of these standards (also by the INSPIRE activity).

5 SEARCHING METHODS

It is necessary to define clearly two completely different searching problems. The first one is searching for a map layer in a given map service – we have a map service and we are trying to find a required map layer or group of layers. The second problem is the method of finding an appropriate map service on the Internet. In this article we will focus especially on the first problem – identification of group of layers in a given service. But it is obvious that both problems are partially overlapping.

We distinguish three main approaches to searching in WMS. The first approach could be designated as an extension of the WMS standard. The second one is development of an independent standard that will be connected to the WMS. The third one is development of an independent searching engine (as is *Google* for web pages).

5.1 Extension of Web Map Service Standard

This approach could lead to very promising results from the user point point-of-view. It is very straightforward and simple. Let us assume that we need a new WMS request that provides us information about some specific group of layers

instead all possible layers (as provides the *GetCapabilities* request).

Such request is named *FindLayer*. It is used instead of the *GetCapabilities* request. The response on this request is again a capabilities file, but with just a specific part of map layers fulfilling given conditions. Its mandatory parameter is *words*. It is used to specify search keywords. Its optional parameter is *operator* that allows to set the operator between the keywords.

Because we are dealing with geospatial data, except the search for specific keywords, the user could be searching for layers from a specific location. For fulfilling such request, there is possible to use parameter *bounding box (bbox)*.

To get better results it is useful to allow the user to specify further the search priorities: e.g. keywords found in *title* have higher priority than the results from *abstract* of the map layer. This surely goes hand in hand with the used search method. For such setting it is possible to use parameters with names of map layer entities given by the *GetCapabilities* specification – *name*, *title*, *keywords*, *abstract*).

An even more straightforward approach is adding mentioned parameters directly to the *GetCapabilities* request. In this case every user is able to receive a complete capabilities file or a file with specific layers.

Parameter	Required	Description
REQUEST=GetCapabilities	Mandatory	Request
SERVICE=WMS	Mandatory	Mapping service
VERSION=int	Mandatory	Service version
WORDS=string,...	Mandatory	Searching word
OPERATOR=and or	Optional	Keywords operator
BBOX=int,int,int,int	Optional	Search area using WGS-84 coord.
NAME=int	Optional	Priority of the element
ABSTRACT=int	Optional	
KEYWORDS=int	Optional	
TITLE=int	Optional	

Table 1: Extended *GetCapabilities* request

Example of the request – find layers with roads in Slovenia:
`http://server.com/cgi-bin/mapservice?service=WMS&version=1.1.1&request=GetCapabilities&words=roads&bbox=47.4,16.2,45.17,13.4&`

Let us summarize this approach. We formulated a completely new request that could be part of the WMS standard. Such request is able to work as an extended version of a *GetCapabilities* request or standalone request. The advantage of this approach is clearly its simplicity. The result of searching – a capabilities file – could be directly used in contemporary geographical information systems or transformed into XHTML or any other XML based format. However, this approach has four significant drawbacks.

1. **Extensibility:** Presented searching method is associated with a single map service, hence it could not be in future extended to work with more services.
2. **Different results:** There are many algorithms for text data searching. Ranking the relevance depends completely on the used algorithm. Therefore the order of returned layers and even their appearance in the result capabilities file could differ.
3. **Implementation:** Problems could be also with the implementation of this request into map servers. Advanced searching algorithms are complex and what looks as a simple extension of the standard could lead to a significant implementation task.
4. **Complication of the standard:** Such extending could in time complicate the standard.

5.2 Development of a query language

The second possibility is development of an independent query language. This approach could be compared for example with *XQuery* [9]. *XQuery* is independent on a processed XML file. It allows to find elements with given properties. *XQuery* itself is not suitable for searching as been defined but it is the delineation between data storage and data processing which is important. The basic idea is very similar to the object oriented design: every entity should have just one purpose and for fulfilling complex tasks more entities have to cooperate. In the object-oriented design this entity is a class or its instance, or in this case a web service.

From this point of view it is obvious that the searching should be independent as it is for example processing the geodata using *Web Processing Services* [10]. WMS should be focused just on geodata distribution. Consequently also the implementation of WMS would not be complicated. This goes hand in hand with the trend of web services orchestration.

The query language could be similar to the *FindLayer* request presented above. The key question is the format of those responses on a given request.

Parameter	Required	Description
REQUEST=FindMap	Mandatory	Request
SERVICE=MSS	Mandatory	Map Search Service
VERSION=int	Mandatory	Service version
WORDS=string,...	Mandatory	Searching word
OPERATOR=and or	Optional	Keywords operator
BBOX=int,int,int,int	Optional	Search area using WGS-84 coord.
NAME=int	Optional	Priority of the element
ABSTRACT=int	Optional	
KEYWORDS=int	Optional	
TITLE=int	Optional	

Table 2: *FindMap* from *Map Search Service*

Example of the request – find layers with roads in Slovenia:

```
http://server.com/cgi-bin/searchservice?
service=MSS&version=0.0.1&request=FindMap&
words=roads ,streets&bbox=47.4,16.2,45.17,13.4&
operator=or&
```

The first possibility is using the capabilities file. This allows simple integration with GIS software. The user is usually asked for the URL of the capabilities file, therefore the address itself could contain the search request. The capabilities file with response could redirect the client to an appropriate WMS service.

However, a significant problem from the long-term point-of-view is the architecture of the capabilities file. The capabilities file in the WMS standard provides all possible information at one time. This approach is not effective. Providing all information at a time has not more sense than a method returning values of all attributes of an object.

A step forward will be adoption of the completely new WSDL/SOAP interface. There could be a set of specific requests: get the general information about the service, get the content of the service (list of map layers), describe a map layer. The responses could be the well-known parts of the capabilities file appropriately enveloped into a SOAP message.

However, also this approach has significant drawbacks:

1. It will be necessary to implement a completely new standard in GIS applications.
2. Orchestration of web map services is still in unsatisfactory condition. None of the major web mapping services is really using WSDL – the key standard in this field. Even SOAP is not supported by all standards including namely the WMS.
3. Enveloping binary data provided by WMS into SOAP messages will lead to significant rise of the messages size. This is especially problematic in many implementations of SOAP, where messages are buffered in the memory before are sent. This could have significant influence on the overall performance of the system.

If we summarize this approach, we could say, that designing the query language has many problems. However, for many reasons it is still more suitable than the extension of the WMS. Most important arguments are that both standards will be single-purpose and implementation of the searching would not complicate the implementation of the WMS.

5.3 Development of an independent engine or catalog

This approach is currently most popular. The reason is obvious. There is still no generally accepted approach to searching in map services, therefore all projects are still more or less experiments.

There could be an important argument that it is not possible to compare an independent engine or catalog and a standard. That is certainly true, but there is an important question: is it necessary to have any standard for searching the map layers? There is no query language or similar standard for web pages, users are working with *Google*, *Yahoo!* and similar services and there is generally no need for any special standards.

To answer this question we must focus on search result processing. In case of the web pages it is usually the user who is processing them. The user himself is seeking the information.

In case of the web services this situation is more complicated. Currently, main usage of the web mapping services is in GIS applications (importing layers for analysis or map creation). In this case automatic processing of queries by the software should be a contribution, but it is not necessary.

However, the importance of mapping services is still rising as a data source for other web services such a Location Based Services, navigation, crisis management and so on. These applications must provide an interface to information from many (map) sources. Therefore it is not possible to rely on some manual processing of map layers. The service itself must process the data sources and present just the required final information (e.g. yes, there is a danger of flood in this area).

Moreover, in manual processing it is relatively easy to change the provider. Users of *Yahoo!* could switch to *Google* or other company. It is in fact just the question of custom. But applications must be re-implemented to the new API. This leads to high expenses and also debatable reliability of the solution.

6 CONCLUSION

For the reasons presented above, we are convinced that it is necessary to design a searching standard that could be implemented by vendors of such services. Such key applications as e.g. previously mentioned crisis management could not rely on an API given by a single vendor.

A discussion could be about the method of searching and especially about the standard used for communication. We are convinced that it is not possible to rebuild currently used standards from the scratch to support the web service standards. Such act is not possible especially from the developers' point-of-view, because all open standards are based on broad agreement that is against radical changes [11]. Furthermore, before acceptance of such technology, there must be an objective consideration of all consequences and implementation problems (e.g. mentioned problem with enveloping of binary data) because a badly designed application programming interfaces are plague of the software engineering [12]. However, this acceptance is practically inevitable. Therefore the searching standard must be designed with respect to the future support of the WSDL/SOAP communication.

On the basis of our experimental implementation, we propose to use an independent query language that allows integration with existing web mapping services as been presented in section 5.2. Interconnection with one particular standard could be complication for further development in future and extended standard could also loose its simplicity and clearness.

The experimental searching service, as well as the source codes and further information are available on address <http://echo.mendelu.cz>.

References

- [1] Sample, J. et al. Enhancing the US Navy's GIDB Portal with Web Services. In *Internet Computing*, IEEE. Sept.-Oct. 2006, 10, 5, pp. 53–60.
- [2] Zhao, P. – Di, L. Semantic Web Service Based Geospatial Knowledge Discovery. In *IEEE International Conference on Geoscience and Remote Sensing Symposium 2006*. 2006, pp. 3490–3493.
- [3] Procházka, D. Motyčka, A. Geospatial search service. *Proceedings of the 11th International Multiconference Information Society - IS 2008*. Ljubljana, Slovenia, 2008. vol. A, pp. 227-230. ISSN 1581-9973.
- [4] Procházka, D. Procházková, J. Moebius: An interface to web map services. In *Geoinformatics FCE CTU*. Prague, Czech Republic, April, 2009. vol. 3, pp. 39-50. ISSN 1802-2669.
- [5] de la Beaujardiere, J. *OpenGIS Web Map Server Specification Implementation*, Open Geospatial Consortium, Inc., on-line: <http://www.opengeospatial.org/standards/wms>. 2009.
- [6] Vretanos, P. A. *Web Feature Service Implementation Specification*, Open Geospatial Consortium, Inc., on-line: <http://www.opengeospatial.org/standards/wfs>. 2009.
- [7] Cristensen, E. et all *Web Services Description Language (WSDL) 2.0: Core Language*, W3C, on-line: <http://www.w3.org/TR/wsdl20>, 2007 .
- [8] Gudgin, M. et all *SOAP Version 1.2 Part 1: Messaging Framework (Second Edition)*, W3C, on-line: <http://www.w3.org/TR/soap12-part1/>, 2009.
- [9] Boag, S. et all *XQuery 1.0: An XML Query Language*, W3C, on-line: <http://www.w3.org/TR/xquery/>, 2007.
- [10] Schut, P. *OpenGIS® Web Processing Service*, Open Geospatial Consortium, Inc., on-line: <http://www.opengeospatial.org/standards/wps>, 2009
- [11] Lake, R. Technology and Social Perception, on-line: <http://www.galdosinc.com/archives/576>, 2008.
- [12] Henning, M. API Design Matters. In *Communications of the ACM*, ACM, vol. 52, no. 5, pp. 46-57. May, 2009.

Acknowledgement

This article was written in context of VZ MSM 6215648904/03/03/01.

THE ARCHITECTURAL DESIGN OF A TOOL FOR TESTING WORKFLOW-BASED APPLICATIONS

Uroš Goljat, Marjan Heričko

University of Maribor, Faculty of Electrical Engineering and Computer Science

Smetanova ulica 17, 2000 Maribor, Slovenia

e-mail: {uros.goljat, marjan.hericko}@uni-mb.si

ABSTRACT

In this paper we will present some of the problems that arise when testing workflows created with the Windows Workflow Foundation (WF). Based on that, we will present the architecture behind a tool that enables the structured testing of workflows in isolation. That means that workflow is tested separately from the application in which it will later be used. The presented architecture will serve as the base for a theoretical implementation of the tool named WF-TESTER.

1 INTRODUCTION

In the last few years, companies and governmental institutions have increasingly been paying attention to automation and the management of business processes through the use of IT technologies. This means that the majority of applications in use today have the same goal: to support business processes. Some of these business processes are fully automated and communicate only with other business processes. Others, which are in the majority, include communication with people that initiate the process, confirm documents used in processes, or handle exceptional cases. In either case, it is possible to specify the series of steps that are included in the business process as a workflow. To specify business processes in the form of a workflow, workflow management systems, like the *Windows Workflow Foundation (WF)*, are used.

Because workflows present the main components of workflow-based applications, they need to be tested appropriately with the use of correct testing techniques and testing tools in order to allow automated and repeatable testing. These testing techniques ensure the overall quality of implemented workflows and, consequently, for the application in which the workflows reside in. That is to say, if the workflow, which is a main part of the workflow-based application, works correctly, then this means that the entire application will also boast better quality.

2 WORKFLOW-BASED APPLICATIONS WITH WF

Building workflows and running them inside software opens up a lot of new challenges. One particular challenge is that the execution of a workflow can take hours, days or even weeks to complete. This raises the question of how an application can and should keep track of the workflows'

state over an extended period of time. Of course, this is only one such challenge. Equally important challenges can be found in [2].

The WF framework provides a programming model, an engine and tools that allow developers to build and deploy workflow-based applications on the .NET Framework. It was first released in November 2006 as part of .NET 3.0 [1].

2.1 The key concepts of WF

Prior to building workflow-based applications using WF, developers need to get familiar with some key concepts that are fundamental for the development of WF workflows and applications [2] (the relationship between the described concepts is shown in Figure 1):

- **Workflow Designer:** a graphical tool to visually design and model workflows. Workflows are designed inside Visual Studio.
- **Activities:** these are the basic building blocks of every workflow built with WF. They are standalone pieces of functionality that can be reused across multiple workflows. The work that an activity implements can be very simple (e.g. a send e-mail activity) or quite complex (e.g. a composite activity that executes nested activities in a transaction).
- **Workflow:** a group of activities that represent the implemented business process or its parts. The type of activities contained in a workflow defines the type.
- **Base activity library (BAL):** a set of activities that range from the most basic workflow control to more complex activities, such as invoking WCF services. BAL includes nearly 30 different activities [1].
- **Runtime engine:** a WF component that is responsible for executing workflows. It also manages the execution of runtime services that are vital for workflows to properly execute.
- **Host process and run-time services:** The host process is needed to host and manage the runtime engine. It is also responsible for conducting runtime services that are responsible for providing services such as transactions and persistence to the run-time engine. A host process can be any type of .NET application.

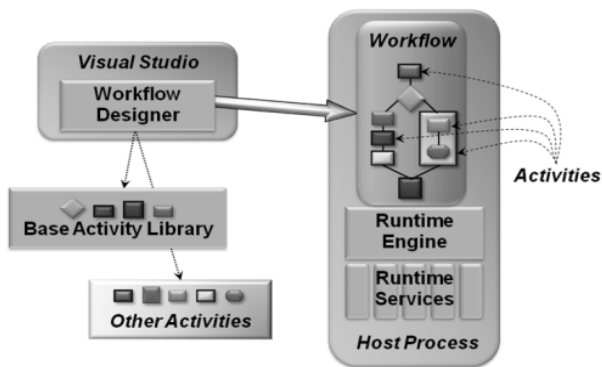


Figure 1: WF fundamental components [2]

2.2 Workflow types

WF allows three ways of modeling workflows through different workflow types. Workflows built using WF can be modeled as: finite-state machines (“state-machine workflows”), a sequences of activities (“sequential workflows”), or data-driven workflows.

State-machine workflows are presented as a set of states and application events. Application events control the transition between states and occur when a workflow is in a specific state. The lifecycle of a state machine workflow is determined by the initial state in which the lifecycle begins and by one or more terminal states in which the lifecycle ends. In most cases, workflows modeled as state machines are non-deterministic. This type of workflow is ideal for modeling business processes where the workflow itself includes a lot of user interactions.

Sequential workflows are presented as a set of predefined steps (activities) that have a prescribed order, and the path through the workflow is deterministic. The control flow within the workflow is defined through well-known constructs such as branching (*if-else*) and loops (*while*). Modeling workflows as a series of predefined steps is ideal for modeling business processes where little or no user interaction is needed.

Data-driven workflows are presented as a special type of sequential workflows. In this case, the control flow of activities is determined by conditional expressions that are presented by rules. The rules check external data to determine the execution path of a workflow.

3 WORKFLOW TESTING CHALLENGES

For testing workflows, some traditional testing approaches can be used (e.g. unit testing). However, because workflows are always part of a workflow-based application, testing workflows raises some unique challenges.

The main challenge is how to test a workflow outside of an application (i.e. in isolation) without modifying the definition of a workflow for testing purposes. To allow such a testing approach, we need to ensure mock objects for two types of workflow interaction partners:

- A mock for hosting application.
- A mock for interacting partners accessible through Web Services.

Based on this, we then need a generic test tool that will allow for the testing of workflows in isolation and that also intercepts communication between the workflow and its interaction partners. To enable the testing of workflows without prior modifications, the test tool needs to be designed as a standalone application, which accepts workflow definitions from assembly files. To enable the testing of workflows in isolation, they need to be properly designed so that they can be tested outside of the application. Two such design guidelines are:

- All communication between the workflow and host application need to be modeled through appropriate activities i.e. for handling communication from the host application to the workflow, a *HandleExternalEvent* activity needs to be used.
- Web Service calls need to be modeled through activities that enable the use of the Windows Communication Foundation (WCF) framework.

Due to a lack of time, poor design or suboptimal implementation of a WF workflow, testing is often difficult and left to the last phases of the development lifecycle. Or the structured testing of workflows is skipped. In cases where workflows are badly designed and it is not possible to test them in isolation, the testing of the workflow is done through simple Windows Forms applications with multiple buttons, where each button is used to test a part of a workflow. This approach does not allow the testing of the workflow as a whole and is used to test the workflow quickly and simultaneously during the development phase. This kind of ad-hoc testing approach is not good in the long term because it cannot deal with changes and with the ad-hoc approach we cannot assure repeatable testing. A key requirement to structured testing in the case of WF workflows is their proper architectural design.

3.1 Structured testing approach

When testing traditional software, we choose between two testing approaches; the black box approach and the white box approach. With the black box testing approach we only have an external view of the testing object (in this case the testing object is a workflow under test) with no knowledge of its internal structure. The behavior is observed from the outside, and the execution path is controlled through changing input values. On the other hand, the white box testing approach allows for structured testing because we have full access to the internal structure (e.g. data and control flow) of the testing object. With regard to both approaches, the white box testing approach for testing workflows is more suitable. However, because the test tool does not have access to the workflow’s source code, a gray box testing technique has to be used [3]. In addition to the

structured testing approach for testing WF workflows, we also need some metrics that provide results about testing. The most important testing metric in our case is test coverage.

3.1.1 Test coverage for sequential workflows

For sequential workflows, test coverage metrics used by white box testing technique [4] can be used. Those coverage metrics are: code coverage, branch coverage and path coverage. Because the smallest unit of testing a workflow is an activity, the code coverage metric has to be defined as activity coverage, which shows the percentage of executed activities during the execution of a test case. Because sequential workflows are already represented as directed graphs, cyclomatic complexity [5] can also be used to give a measure of the minimum number of test cases needed.

3.1.2 Test coverage for state-machine workflows

For state-machine workflows, test coverage metrics for sequential workflows can be used in addition with test coverage metrics used for testing software based on finite state machines [6][7]. Several coverage criteria have been defined. One of them is Transition Coverage [6], which measures the percentage of state machine transitions made during the execution of a test case.

4 TEST TOOL ARCHITECTURE

In this section we will present the architecture for a test tool that enables the testing of WF workflows in isolation, i.e. apart from the application in which they are included. The architecture is based on four layers that build upon one another. The four layers are depicted in Figure 3. The implementation of separate layers can be done in a few ways, which will be briefly described in the description of each layer.



Figure 2: *WF Test Tool Architecture*

The test architecture was adopted from [8] where it was originally used for the design of a test framework to test BPEL processes. Since the concepts of BPEL and WF are

similar -- both are used to model business processes -- this architecture is also suitable for testing WF workflows.

The Test Case Specification layer is responsible for the definition of test cases. Two approaches can be used [8]: the data-centric approach, where the test case specification is done through files that the tool can interpret and the logic-centric approach, where the test case specification is done through a programming language.

The Test Case Organization layer is responsible for grouping test cases defined in test case specifications into test suites and adds meta-data that connects test cases with a specific, tested workflow. As in [8], again two approaches are possible: to integrate the test case organization with test case specification (e.g. xUnit) or to formulate the test case organization separately and only include links to the test case specification.

The Test Execution layer is responsible for the execution of test suites and/or test cases. Two approaches can be considered [8]. The first, which is more suitable for testing WF workflows, is simulated testing where all interacting partners are simulated through a special layer which intercepts interactions and handles them locally based on a test case specification. The second approach is real-life testing, where all interaction partners are replaced by mock implementations and therefore the workflow has to be changed for testing purposes. Since we do not want to change workflows before testing, this approach is not suitable in our case.

The Test Results layer is responsible for gathering and displaying test results after the testing is finished. The test results can be presented as a pass/fail result and through test metrics. To calculate test metrics, the test execution layer needs to ensure a mechanism that enables the tracking of executed workflow activities. The test results layer then uses this data. A key requirement is that test results are given in a way that is independent from the environment in which they are later used.

5 WF-TESTER TOOL

The WF-TESTER tool presents a theoretical implementation of a test tool based on the four-layer architecture presented in Section 4. The tool is presented as a standalone application that accepts the compiled workflow definitions through assembly files.

The Test Case Specification layer is implemented via a data-centric approach. Separate test cases are defined through XML documents and its structure is defined through the XML Schema. This approach allows for a simple way of defining test cases and also provides good support for automated test execution. From a testing point of view, sequential and state-machine workflows are quite different. Hence, the test case specification needs to be formulated separately for each type of workflow (e.g. two different XML Schemas).

The Test Case Organization layer is also implemented through the XML document and its accompanying XML Schema. This is due to the tight coupling between test case

specifications and organization. The test case specification is logically separated from test case organization however they are still contained in one file. The test suite specification file includes all the data that is necessary for the proper execution of a tested workflow and a set of test cases as defined by the test case specification layer. Minimally, the following data must be provided through the test case organization layer and is necessary for the proper execution of a workflow during the execution of a test case:

- The path to the assembly, with a compiled definition of the tested workflow.
- A list of referenced assemblies used by the tested workflow.
- A list of Web Services used by the workflow with the location of their mock implementations.

The WF-TESTER also provides a more intuitive approach for defining test suites and test cases (not to edit XML documents by hand). This allows the tester to focus on testing only. The resulting test suite specification file needs to be saved to a persistent medium, to ensure repeatable testing.

The Test Execution layer is implemented as a simulated testing approach as proposed in Section 4. The core of the execution layer of the WF-TESTER tool consists of three components:

- *The WorkflowObserver (WO)* component is responsible for the interception of local communication between the workflow and host application in both directions. The WO is implemented through the use of Tracking Services, which are a part of the WF framework [9]. Based on the test case specification and the current executing activity, the WO determines if the data needs to be accepted from the workflow or if the data should be sent back to the workflow. This approach allows for the testing of the workflow without prior modifications.
- *The WorkflowWebServicesObserver (WSO)* component is responsible for the interception of communication with outer interaction partners, e.g. Web Services. Behind the scenes, WSO uses the ChannelManagerService [10] service that is used by the Send activity to resolve the web service endpoint. Before the execution of the test, the WSO configures the ChannelManagerService with the endpoint addresses of mock web services specified in the test suite specification. This kind of approach also allows for the real-life testing approach without changing the workflow definition for testing purposes.
- *The WorkflowExecutionObserver (WEO)* component is responsible for the detailed tracking of workflows' execution when running test cases in the sense of "through which path the workflow executed," and "which activities were executed." This is done through Tracking Services, which are configured so that all information from every executed activity (e.g. execution time, status, execution results) is recorded.

The acquired data need to be stored to a persistent medium (e.g. SQL Server Database) so that they can be used for later queries and analysis in the test results layer.

The Test Results layer is responsible for displaying test results to the tester after the testing is finished. The WF-TESTER shows the pass/fail type of results as a green/red light besides a test case. The more important test results are provided through test coverage metrics. Test coverage metrics are computed from the data collected by the WEO component during test execution. For example, the Activity Coverage metric is computed from the number of activities defined in the workflow, divided by the number of executed activities.

6 CONCLUSION

In this paper, we presented the WF framework for modeling the business processes as workflows. Since workflows constitute the core of workflow-enabled applications, they demand special attention when it comes to testing. Workflow testing brings with it many new challenges, which can be resolved with the proper design of workflows and a structured testing approach.

Based on the testing challenges, we have presented the architecture for a tool that enables the structured testing of WF workflows. Later, the presented architecture was used for the theoretical implementation of a WF test tool called WF-TESTER, which enables the structured testing of WF workflows.

References

- [1] Kitta T: Professional Windows Workflow Foundation, Wiley Publishing Inc., Indianapolis, USA, 2007
- [2] Chappel D: Introducing Windows Workflow Foundation, available at http://download.microsoft.com/download/f/3/2/f32ff4c6-174f-4a2f-a58f-ed28437d7b1e/Introducing_WF_in_NET_Framework_35_v1.doc, Accessed: August 2009
- [3] Software Testing Fundamentals: Gray Box Testing, available at <http://www.softwaretestingfundamentals.com/2009/03/gray-box-testing.html>, Accessed: August 2009
- [4] Zhu, H., Hall, P. A., and May, J. H: »Software unit test coverage and adequacy«, ACM Comput. Surv. 29, 4 (Dec. 1997), 366-427
- [5] NIST: Cyclomatic Complexity, available at: <http://hissa.nist.gov/HHRFdata/Artifacts/ITLdoc/235/chapter2.htm>, Accessed: August 2009
- [6] IPL: Testing State Machine with AdaTEST and CANTATA, available at: <http://www.iplbath.com/pdf/p0827.pdf>, accessed: August 2009
- [7] Chow, T. S: Testing Software Design Modeled by Finite-State Machines, IEEE Trans. Softw. Eng. 4, 3 (May. 1978), 178-187
- [8] Mayer, P. and Lübke, D: Towards a BPEL unit testing framework, In Proceedings of the 2006 Workshop on Testing, Analysis, and Verification of Web Services and Applications (Portland, Maine, July 17 - 17, 2006). TAV-WEB '06. ACM, New York, NY, 33-42
- [9] Gristwood, D: Windows Workflow Foundation: Tracking Services Introduction, available at: [http://msdn.microsoft.com/en-us/library/bb264459\(vs.80\).aspx](http://msdn.microsoft.com/en-us/library/bb264459(vs.80).aspx), accessed: August 2009
- [10] Microsoft, ChannelManagerService, available at: <http://msdn.microsoft.com/en-us/library/system.workflow.runtime.hosting.channelmanagerservice.aspx>, Accessed: August 2009

UTILIZING PROCESS MODELING TO SUPPORT THE COLLABORATIVE COMMUNICATION OF AUTHORITIES IN THE MANAGEMENT OF DISASTER SITUATIONS

Jari Soini, Petri Linna, Hannu Jaakkola
Department of Software Engineering
Tampere University of Technology – Pori
P.O. Box 300, FIN-28101 Pori, Finland
e-mail: jari.o.soini@tut.fi

ABSTRACT

The paper studies how process modeling could be adapted in a multi-authority field to assist collaborative communication in disaster situations. In this context there are various authorities involved and this time-critical operational environment requires effective and seamless collaboration related to information transmission for all participants and also the optimized, integrated use of management systems and resources. Based on preliminary study results carried out by Finnish authorities, there seems to be a need for technologies which can integrate communication and visualization aids in order to build a collaborative understanding of what is happening and who is doing what. The current research is focusing on study how to exploit process modeling in this context and to clarify how formal process notation can be utilized. This research was based on the assumption that visual process descriptions are more expressive, illustrative, and easier to understand than textual descriptions. Graphical descriptions and links between tasks and roles inside the organization and also between the participant organizations could be one tool for clarifying and enhancing co-operation. The paper discusses how and why process modeling should be utilized to support the collaborative communication of authorities in the disaster management.

1 INTRODUCTION

A disaster brings together a team of people often representing different organizations, resources and roles. In this context there are various authorities involved and this time-critical environment requires effective and seamless collaboration regarding information transmission for all participants and also the optimized, integrated use of management systems and resources. The participant authorities have to respect the maxims of collaborative communication in order to efficiently manage the complexity inherent to the disaster context. Effective

management of this multi-authority, time-critical and hectic situation requires the assistance of supportive information technology. However, previous studies [1],[2],[3] have shown that the deployment of collaborative decision support systems technologies within the context of emergency management is not yet common. There is a clear need for technologies which can integrate communication and visualization aids in order to build a collaborative understanding of what is happening and who is doing what and when [4],[5]. This scenario shows the need for research into this topic. This paper presents one approach for solving this issue.

An ongoing research project *Seamless Services and Mobile Connectivity in Distributed Disaster Knowledge Management (SSMC/DDKM)* [6] is studying how the Business Process Modeling (BPM) approach, can be adapted to the benefit of communication transfer and sharing in this environment. We are studying how this communication link-up may be implemented by utilizing process modeling and using a graphical description of the information flows between the different tasks or roles related to emergency management. Our research is based on the assumption that the graphical description and links between the tasks and roles inside the organization and also between the participant organizations could be a useful tool for clarifying and enhancing the collaboration of authorities in the management of disaster situations. In general, the purpose is to clarify how well formal process modeling notation is suitable for utilization in this context. Several related methods developed for this purpose are also available such as the ARIS (Sheers Architecture of Integrated Information Systems) methodology and notation [7] and the WfMS (Workflow Management system) model [3]. This paper deals with our research findings and experiences during the current research work.

2 BACKGROUND

This study is based on the ongoing SSMC/DDKM research project being carried out at the Tampere University of

Technology (TUT) [8]. Research has been started into the conceivable methods and techniques that may be beneficial for enhancing information sharing and interoperability between participant organizations in emergency situations. One special task of this two-year project is to study whether process modeling could be utilized for clarifying information transfer and communication between authorities in emergency situations. The project will develop the technology required for modeling accident and disaster knowledge and related processes to support improved information and knowledge management in disasters and accidents, to determine standardized description solutions, and to evaluate the practicability of the tools needed in modeling. The aim is to support interaction between the different parties by means of process modeling. Moreover, process models support the transfer of seamless services between information systems toward SOA-based interoperability (e.g. [9],[10]), which is one of the main research focuses in the SSMC/DDKM project. As the concept of SOA is based on cumulative knowledge in software engineering and software technology, it is our primary candidate for solving the issues presented in this study in the later stages of our project [11].

In Finland, the Finnish Government has published a recommendation [12] for describing public administration processes, with the aim of standardizing and clarifying the description of said processes. For example, Business Process Modeling Notation (BPMN) [13] based terms and concepts have been used in relation to process modeling and this was a strong justification for the selection of BPMN as the model for study in our research project.

3 CHALLENGES IN COLLABORATIVE COMMUNICATION IN DISASTER SITUATIONS

Previously, studies have shown that the primary challenges in responding to disaster situations are generally related to breakdowns in information, communication, and coordination [14],[15]. These recognized difficulties relate merely to deficient communication and coordination among organizations [16]. Sophisticated technologies can establish technical inter-organizational connections, but do not ensure a comprehensive connection between participant organizations during dynamic emergency response situations [17]. In this kind of cross-organizational communication, differences in the participants' organization structure (e.g. hierarchical vs. ad-hoc organizations) and also the lack of a common vocabulary between response organizations add to the problems. The organizational aspects must be taken into account when developing parallel and interconnected processes between different parties. The types and nature of the information needed by an organization involved in disaster and catastrophe management and the position of the organization's processes in the emergency management life

cycle introduce a very wide landscape of different knowledge and application needs.

Earlier studies [3],[18] have shown that emergency plans are very similar to business processes. Therefore, it can be assumed that these plans can be modeled with the notations used for business modeling. In practice, these plans are currently textual documents of different levels of detail and formality. These emergency plans usually exist in the form of multi-page printed documents and there is a lack of suitable approaches to support the modeling, execution and management of these plans [19]. Based on this, we decided to investigate formal modeling notations and modeling principles such as icon languages for use between response organizations. The common assumption is that a "picture" can be more easily understood than textual descriptions only (see e.g. [19]).

4 PROCESS MODELING – A TOOL FOR IMPROVING EMERGENCY MANAGEMENT

The processes defined during disaster responses are based on various kinds of human interactions in disaster scenarios. It is important to study and optimize these interactions because humans, not technologies, play the key role in disaster response [20]. Business process models are a suitable means to describe the temporal and logical order of tasks for achieving a given goal. They can be used with the help of visuals to describe complex links and tasks on many different levels and with the appropriate precision [21],[22]. Process models (for example: UML - Unified Modeling Language or BPMN - Business Process Modeling Notation) are generally described in a visual language (for example: XPDL - XML Process Definition Language, BPML - Business Process Modeling Language, PetriNets or Unified Modeling Language Activity Diagram), so that the majority of the data content of the model is presented with graphic symbols and links [23]. In any event, the process should be facilitated by guiding people on what to do - on how to divide and coordinate the work - and by ensuring effective communication [13]. In this case the applicability of these languages is being investigated with regard to whether they enhance the coordination of intervention actions to recover from a disaster event.

In this research the BPMN notation [24] was chosen as an experimental tool because it is the recommended graphical notation for process description in Finnish agencies [13], which are the target organizations in the SSMC/DDKM project. Briefly, BPMN is a graphical notation that enables the visual description of business processes. One of the main ideas of this approach was to develop a model that is readily understandable by all users from developers to people in the organization. BPMN is quite a new standard developed by the Business Process Management Initiative (BPMI), which nowadays belongs to the Object Management Group (OMG) [25]. BPMN is widely supported by service suppliers and its greatest benefit is its

connectivity to an SOA environment. The basics of BPMN notation are described as “the types of graphical objects that comprise the notation and how they work together as part of a Business Process Diagram (BPD)”. Figure 1 below presents the core set of BPMN elements.

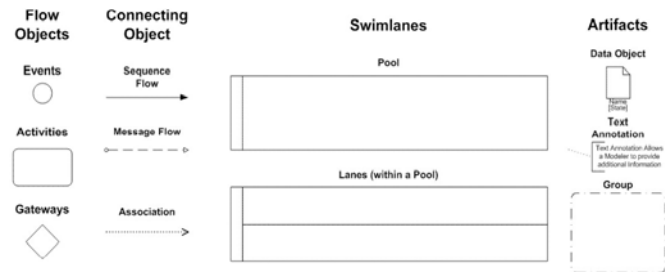


Figure 1: Core set of BPMN elements [26].

A BPD is made up of a set of graphical elements, which enable the easy development of simple diagrams (see Figure 1 above). Below, an example of a process diagram is described in Figure 2, consisting of processes related to the Finnish Emergency Response Center (ERC) response plan modeled with BPMN notation.

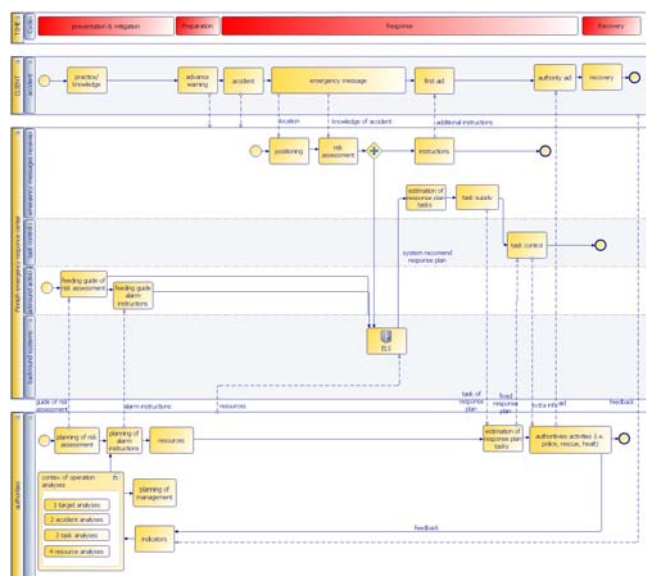


Figure 2: An example process diagram of several tasks related to ERC response plans modeled using BPMN notation [27].

This model was made in close co-operation with Finnish emergency authorities during the SSMC/DDKM project. An example process diagram (Figure 2 above), describing the usage of a response plan (RP), was made based on interviews with emergency experts [28],[29]. Literature sources (e.g. guides, operating procedures, manuals) were also used in order to become familiar with ERC processes [30],[31],[32].

At present, there are multiple modeling tools available which support the execution of BPMN (see e.g. [33],[34]). These tools are used when it is necessary to describe service interfaces to other organizations. For example, Business Process Execution Language for Web Services (BPEL) [35] provides the possibility to visualize the execution of business processes. BPEL is an XML-based language written by OMG (Object Management Group) designed to execute web services. It enables data transfer and compatibility between different applications and systems [36]. The position of various people in relation to BPMN and BPEL has been put in graphical form below in Figure 3.

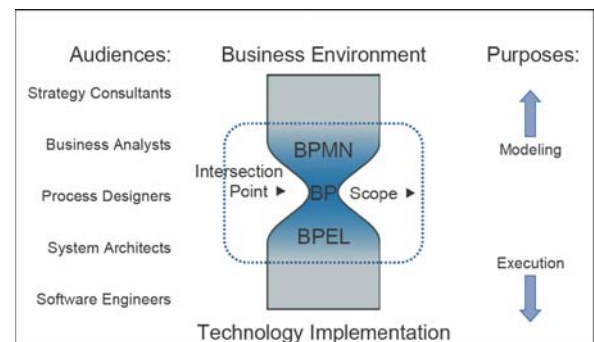


Figure 3: BPMN modeling - BPEL programming [37].

Figure 3 describes on which level each user is. On the top level are corporate management consultants, who make process modeling in a graphic form using BPMN and below there are the software engineers, who write software code using BPEL from modeled process diagrams. Our ongoing research work on the SSMC/DDKM project relates precisely to investigating this modification from BPMN to BPEL.

To sum up, some observations on the focal issues in this study are given below. Firstly, before starting the modeling work it must be clearly understood why we are modeling and after that the goals of the modeling must be carefully defined. The appropriate accuracy and level of the modeling can and should be defined based on this knowledge. Secondly, it must be noted that all operations/processes are not worth modeling, because in practice certain matters can best be presented with textual description. Thirdly, disaster situations, which are characterized by their random and unpredictable behavior, and in which multi-authority co-operation and cross-organizational communication are key, are quite challenging to model. The interdependence between the tasks, and non-formal tasks, which are difficult to analyze, and the differing interpretations of how actions should be conducted cause additional challenges from the process modeling perspective [38]. However, based on our research experience from the SSMC/DDKM project, it seems that process modeling as such can also be adapted for the emergency management context. The main challenges

observed in the utilization of process modeling in this environment seem to be related to the current practices and behavioral norms existing in the various authority organizations. There is a lot of work to be done on this issue. This task should be started by harmonizing the vocabulary and concepts used in emergency situations between the participant authorities involved. This is one of the basic and also most fundamental aspects from the process modeling point of view.

5. CONCLUSION

This paper discussed process modeling utilization in disaster situations. In these types of emergency situations, various authorities are involved and an effective response management requires seamless collaboration in terms of information transfer and communication between all the various participants. The research focused on studying how to exploit process modeling in disaster situations and to clarify how a formal process notation can be utilized in this context. The research was based on the assumption that visual process descriptions are more expressive, illustrative, and easier to understand than textual descriptions. The aim of this study was to research the usefulness of Business Process Modeling Notation (BPMN) for a multi-authority field and to determine whether BPMN is suitable for the demands of emergency authorities. The results show that process modeling can be utilized and also justify some reasons why it should be utilized.

References

- [1] B. Van de Walle, M. Turoff. Decision support for emergency situations. *Information Systems and E-Business Management*, Vol. 6. No. 3. pp. 295–316. 2008.
- [2] S. Mehrotra, C. Butts, D. Kalashnikov, N. Venkatasubramanian, et al. CAMAS: A Citizen Awareness System for Crisis Mitigation. *In Proceedings of the 2004 ACM SIGMOD international conference on Management of data*. pp. 955-956. Paris. France. 2004.
- [3] C. Shell, I. Braun. Using a Workflow Management System to Manage Emergency Plans. *In Proceedings of the 6th International ISCRAM 09 Conference*. Gothenburg. Sweden. 2009.
- [4] M. Castrén, S. Ekman, M. Martikainen, T. Sahi, J. Söder. *Suuronnettomuusopas*. 1. painos. Kustannus Oy Duodecim. Jyväskylä. Finland. 2006.
- [5] E. Dykstra. Toward an international System Model in Emergency Management. *In Proceedings of an international System Model in Emergency Management*. Paper presented at the Public Entity Risk Institute Symposium. 2003.
- [6] H. Jaakkola, J. Soini, J. Leppäniemi. *Seamless Solutions and Mobile Connectivity – Towards location independent and flexible system architectures*. Tampere University Press. In Publication No. 12. pp. 1-17. Finland. 2008.
- [7] A. W. Scheer. *ARIS – Business Process Frameworks*. Springer. Heidelberg. 1999.
- [8] Tampere University of Technology (TUT), Retrieved 14/8/09 World Wide Web, <http://www.tut.fi/public>
- [9] M. Papazoglou, D. Georgakopoulos. Service-Oriented Computing. *Communications of the ACM*. Vol. 46. No. 10. pp. 24-28. 2003.
- [10] OASIS. Organization for the Advancement of Structured Information Standards. Retrieved 03/07/09 World Wide Web, http://www.oasisopen.org/committees/tc_home.php?wg_abbrev=soa-rm
- [11] J. Leppäniemi, J. Soini, P. Linna, H. Jaakkola. Towards a flexible service-oriented reference architecture for situational awareness systems in distributed disaster knowledge management. *In Proceedings of the International Conference on Management of Engineering and Technology PICMET'09*. pp. 959-965. Portland. Oregon. 2009.
- [12] JHS 152 Process modeling. Retrieved 11/5/09 World Wide Web, <http://www.jhs-suositukset.fi/suomi/jhs152>
- [13] S. White. Introduction to BPMN. Retrieved 2/8/09 World Wide Web, [http://www.bpmn.org/Documents/Introduction to BPMN.pdf](http://www.bpmn.org/Documents/Introduction%20to%20BPMN.pdf)
- [14] B. S. Manoj, B. A. Hubenko. Communication challenges in emergency response. *Communications of the ACM*. Vol. 50. No. 3. pp. 51-53. 2007.
- [15] L. Palen, S. R. Hiltz, S. B. Liu. Online forums supporting grassroots participation in emergency preparedness and response. *Communications of the ACM*. Vol. 50. Issue 3. 2007.
- [16] A. Meissner, T. Luckenbach, T. Risse, T. Kirste, H. Kirchner. Design Challenges for an Integrated Disaster Management Communication and Information System. *In papers presented at The First IEEE Workshop on Disaster Recovery Networks DIREN 2002*. New York. USA. 2002.
- [17] M. Turoff, M. Chumer, B. Van de Walle, X. Yao. The design of a dynamic emergency response management information system. *Journal of Information Technology Theory and Application*. Vol. 5. No. 4. pp. 1–36. 2004.
- [18] M. Zalewski, P. Sztandera, M. Ludzia, M. Modeling and Analyzing Disaster Recovery Plans as Business. Springer Berlin. Heidelberg. Vol. 5219. 2008.
- [19] H. Eriksson, P. Magnus. *UML*. 3. painos. IT Press. Jyväskylä. Finland. 2002.
- [20] T. Catarci, M. de Leoni, A. Marrella, M. Mecella, B. Salvatore, G. Vetere, S. Dustdar, L. Juszcyk, A. Manzoor, H.-L. Truong. Pervasive software environments for supporting disaster responses. *IEEE Internet Computing*. Vol. 12. No. 1. pp. 26– 37. 2008.
- [21] C. Plesums. *Introduction to Workflow*. Workflow Handbook 2002. Eds. by Layna Fischer. Published by Future Strategies Inc. Lighthouse Point. FL. USA. 2002.

- [22] E. Siwaraman, M. Kamath. On the Use of Petri Nets for Business Modeling. School of Industrial Engineering & Management. Stillwater 74078. Oklahoma State University. USA. 1999.
- [23] C. Prior. *Workflow and Process Management*. Workflow Handbook 2003. Eds. by Layna Fischer. Published by Future Strategies Inc. Lighthouse Point. FL. USA. 2003.
- [24] M. Lindvall, I. Rus. Process Diversity in Software Development. IEEE Software. Vol. 17. No. 4. pp. 14-18. 2000.
- [25] Business Process Management Initiative. Object Management Group. Retrieved 5/6/09 World Wide Web <http://bmi.omg.org/>
- [26] OMG. Object Management Group. Retrieved 25/7/09 World Wide Web <http://www.BPMN.org/Samples/Elements/Core%20BPMN%20Elements.htm>
- [27] P. Linna. Requirements engineering of Emergency center's response plan. Tampere University of Technology. *MSc Thesis*. 2009.
- [28] Emergency Response Centers' authorities. Pori. Finland. Interviews in autumn 2008.
- [29] Emergency Service Collage's authorities. Kuopio. Finland. Interview in 10/6/09.
- [30] M. Castrén, S. Ekman, M. Martikainen, T. Sahi, J. Söder. *Guide of catastrophe*. Jyväskylä. Finland. 2006.
- [31] T. Raivio, Y. Gilbert, H. Lonka, M. Hjelt. *Estimation of Emergency Responce Center*. Gaia Consulting Oy. Helsinki. Finland. 2007
- [32] J. Salonen, T. Okkolin. *Guide to formulating emergency guidelines*. Emergency Service Collage. Kuopio. Finland. 2008.
- [33] BPMN Implementors and Quotes. Object Management Group (OMG). Retrieved 5/6/09 World Wide Web http://www.bpmn.org/BPMN_Supporters.htm
- [34] N. Russell, A. H. Hofstede, W. M. van der Aalst, N. Mulyar. Workflow Control-Flow Patterns: A Revised View. *BPM Center Report BPM-06-22*. BPM center.org. 2006.
- [35] J. Zwiers, B. de Vos. Guide to BPEL. Retrieved 25/7/09 World Wide Web <http://www.radikalfx.com/bpel/usage.html>
- [36] M. B. Juric, B. Mathew, P. Sarang. *Business Process Execution Language for Web Services*. Second edition. Packet Publishing. Birmingham. UK. 2006.
- [37] Introduction to BPMN. Object Management Group (OMG). Retrieved 25/7/09 World Wide Web <http://www.bpmn.org/Documents/OMG%20BPMN%20Tutorial.pdf>
- [38] G. Giaglis, R. J. Paul, G. I. Doukidis. Simulation for Intra- and Inter- Organizational Business Process Modeling. In *Proceedings of the 28th conference on Winter Simulation*. pp. 108-117. Coronado. California. USA. 1996.

TOWARDS ADAPTIVE SERVICE-CENTRED APPLICATIONS

Jože Pfeifer

Hermes SoftLab d.o.o., a ComTrade company
Zagrebška cesta 104, 2000 Maribor, Slovenia
Tel: +386 2 4508858; fax: +386 2 4508816
E-mail: joze.pfeifer@hermes-softlab.com

ABSTRACT

This paper presents an introduction and comparison of some of the most promising development towards adaptive service-centred applications. Such applications can be seen as the next logical step in the evolution of services, by leveraging QoS users will be able to differentiate services not only by their given functionality, but also by their non-functional aspects, such as performance, reliability and execution price. By introducing semantics for services through OWL-S we will also be able to automate the discovery, invocation, composition and interoperation of web services.

1 INTRODUCTION

Service-centred applications are applications that rely on external services (such as web services) for normal operation. If one of the used services should become unavailable or fail during the application's execution, the whole application can become unusable or even crash.

Given the fact that nowadays there are a lot of services available, which offer the same functionality, but offer different non-functional properties, users should be able to choose their preferred services not only by their given functionality, but also by their non-functional properties.

With the introduction of adaptation in such applications the failed service would be replaced by a functionally equivalent service. This means that the application would, despite the failure of the former service, continue to work without problems by utilizing the latter service, thus making the application more reliable and valuable for the end user.

To implement such a system we would have to agree upon some non-functional aspects, with which we could measure and compare services with each other. To achieve this, we could define some Quality of Service (QoS) requirements, such as performance, reliability and execution price.

The next logical step would be to introduce semantics into the field of services. The current proposal is to use Web Ontology Language for Services (OWL-S) in the pursuit to automate the discovery, invocation, composition and interoperation of services.

2 QUALITY OF SERVICE

Quality of Service (QoS) is becoming increasingly important to service providers. By implementing QoS, providers would be able to differentiate themselves from their competition not only by their offered functionality, but also by the non-functional properties of their services.

With the introduction and support of standards such as SOAP, UDDI and WSDL from major web service providers it is nowadays possible to use web services for a wide variety of purposes – financial services, high-tech and media and entertainment. As most of the services are going to need to establish and adhere to standards, QoS will become an important selling and differentiating point of these services.

With the help of [1, 2] we identified 15 QoS requirements for web services, which include:

- Performance
- Reliability
- Scalability
- Capacity
- Robustness
- Exception Handling
- Accuracy
- Integrity
- Accessibility
- Availability
- Interoperability
- Security
- Network-Related QoS Requirements
- Regulatory
- Execution price

Some of them, such as performance and capacity are quite straightforward and can be understood without much explanation – performance is measured in milliseconds and means the time it takes the service to process our request and return the result. Capacity on the other hand is the maximum number of concurrent users that the service can handle without them affecting its average performance levels. But to measure and compare requirements like interoperability and security, certain standards should be introduced. With the help of those standards (e.g. an ISO standard) we would then be able to define, measure and

compare services by their non-functional quality parameters.

3 WEB ONTOLOGY LANGUAGE FOR SERVICES

Web Ontology Language for Services (OWL-S) represents the next step in the evolution of web services. By simplifying the discovery, invocation, composition and interoperation of web services it enables us to more easily search for, execute and compose similar services into something useful [3].

Amongst the most useful resources on the web are those, which offer us some kind of services. With that we don't mean web pages with static information, but services, which enable their user to affect some action or change in the world, such as selling some product or the control of some physical device. The main idea is that the semantic web should enable users to automatically locate, select, use, compose and monitor web services.

For the usage of web services the software agent needs a description of the service in a format he understands and the way, in which it can be accessed. An important goal for semantic web markup languages is the establishment of a framework, within which these descriptions are made and shared. Web sites should be able to employ a standard ontology, which is composed from a number of basic classes and properties and is used for declaring and describing services. At the same time the ontological structural mechanisms of the OWL language offer an acceptable, web-compatible representation language framework within which this is possible to achieve.

Researchers from many organisations have defined the OWL-S ontology (formally known as DAML-S), which represents a language for describing services, because it gives us a standard vocabulary, which we can use, together with other aspects of the OWL description language, for creating descriptions of services. The ontology is useful for atomic and composed services.

An atomic service is a simple service, like inputting your phone number and getting your home address in return. Composed services are, as the name suggests, composed of smaller, atomic services. They were the main reason for many of the defined elements of the ontology.

OWL-S doesn't forbid the declaration of simple, static web sites as services, but the main motivation for OWL-S were more complex tasks, which are as follows:

1) Automatic web service discovery

It represents the automated process of locating web services, which can provide a particular class of service capabilities, while adhering to some client-specified constraints.

For example, the user may want to find a service that sells airline tickets between two given cities and accepts a particular credit card. Nowadays this would mean that the user would have to manually search for a suitable service (e.g. via Google).

With OWL-S, the information necessary for web service discovery could be specified as computer-interpretable

semantic web markup at the service web sites and a service registry or ontology-enhanced search engine could be used to locate the services automatically.

2) Automatic web service invocation

A computer program or agent automatically executes a web service, given only a declarative description of that service, as opposed to when the agent has been pre-programmed to be able to call that particular service.

This means that a user can request the purchase, from a site found by searching and then selected by that user, of an airline ticket on a particular flight.

OWL-S provides a declarative, computer-interpretable API that includes the semantics of the arguments to be specified when executing these calls, and the semantics of that is returned in messages when the services succeed or fail. A software agent should be able to interpret this markup to understand what input is necessary to invoke the service and what information will be returned.

3) Automatic web service composition and interoperation

This task includes the automatic selection, composition and interoperation of web services to perform some complex task, given a high-level description of an objective.

For example, the user may want to make all the required travel arrangements for a trip to a conference. Currently, the user must select all required web services and specify the composition manually. He must also make sure, that any software needed for the interoperation of services, that must share information, is created manually.

With OWL-S the information necessary to select and compose services will be encoded at the service web sites. OWL-S provides declarative specifications of the prerequisites and consequences of application of individual services and a language for describing service compositions and data flow interactions.

The structure of an OWL-S document can be seen in figure 1. The top level of the service ontology is composed of three elements, which describe what the service does, how it can be accessed and how it works.

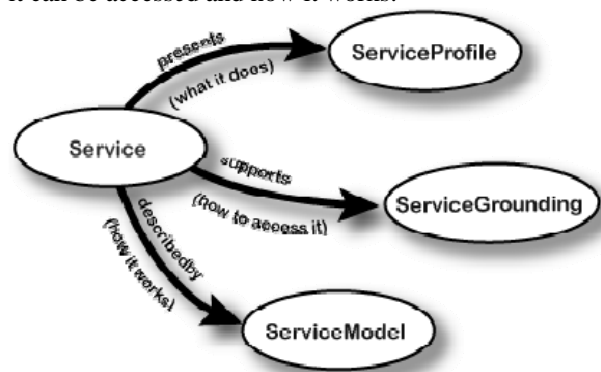


Figure 1: Top level of the service ontology.

4 ADAPTATION

Because web services are prone to failure, the introduction of adaptation should be considered. With it we can implement a failsafe mechanism, whereby we replace faulty services with working ones. This can be achieved with some kind of agent, which would, in the event of service failure, look for an alternative one and execute it instead.

Adaptation of services can be achieved with the help of both aforementioned technologies. QoS enables us to compare services by their non-functional properties, such as performance, reliability and execution price, whereas OWL-S enables us to automate the process of discovery, execution, composition and interoperation of web services. If we would combine those two we would be able to add an additional layer of abstraction to services.

To trigger adaptation, we would first have to feel the need to trigger it. There are two approaches in determining the need to trigger adaptation:

1) Monitoring of services

It represents monitoring of changes or deviations after they happened. Therefore adaptation based on monitoring is reactive and often too late, because the changes or deviations already caused some unwanted consequences.

Despite the fact, that this approach is easier to implement, it has certain weaknesses, which can cause some serious side effects, such as loss of money, unhappy users, longer execution times or even the inability to trigger the adaptation, because the application exited in an inconsistent state.

2) Proactive adaptation

Proactive adaptation represents the solution for the weaknesses of the first approach, because, in the ideal case, the system detects the need to adapt and executes it even before the failure occurs. The goal of testing is the systematic execution of services, so we can detect service failures (e.g. deviations from the actual functionality or QoS from the accepted levels). This approach uses online testing for triggering adaptation, which means that the testing activities are executed while the service applications are running.

5 COMPARISON OF APPROACHES

In our work we looked at several different approaches and decided on describing and comparing the following three, because they all include something novel and useful for introducing adaptation to service-centred applications.

In [4] Zeng et al. describe quality driver composition of web services, where we, based on QoS criteria, select service compositions during execution of composed services. Two properties of this approach are:

- Multi-dimensional quality model for web services, that is extensible and represents non-functional properties that are typical for web services;

- Quality driven service selection, which enables us to circumvent local service selection and use a global planning approach.

With global selection of services and linear programming the paper introduces two important concepts, whereby the former is used with global constraints (e.g. limiting the execution time for the whole process) and the latter at implementing an optimal execution plan in foreseeable time. The architecture of the prototype system can be seen in figure 2.

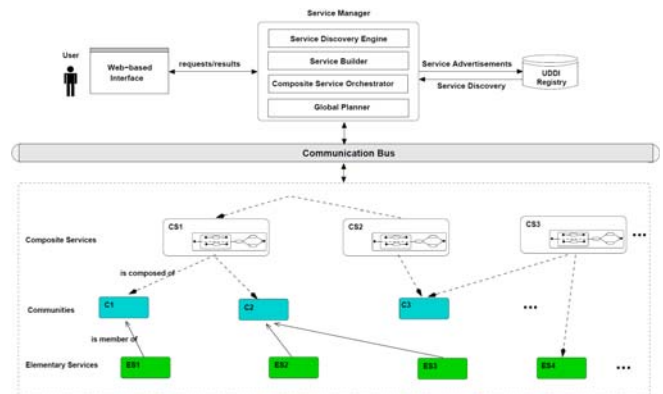


Figure 2: Architecture of the prototype.

In [5] Hielscher et. al. describe proactive self-adaptation (PROSA) and list the bad sides of reactive adaptation, because adaptation is oftentimes triggered too late. They introduce their implementation of such an adaptive system; in the end they even describe the usefulness of proactive adaptation in case of various failures. The architecture of the framework can be seen in figure 3, where four major activities are prescribed:

- 1) **Test initiation:** determines the need to initiate online tests;
- 2) **Test case generation/selection:** this activity determines the test cases to be executed during online testing;
- 3) **Test execution:** test cases from activity 2 are executed;
- 4) **Adaptation triggering:** analysis of the test results provides information on whether to adapt the service-based application and thus to create adaptation requests.

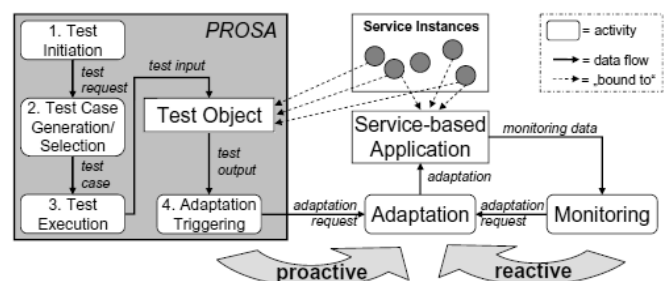


Figure 3: The PROSA framework.

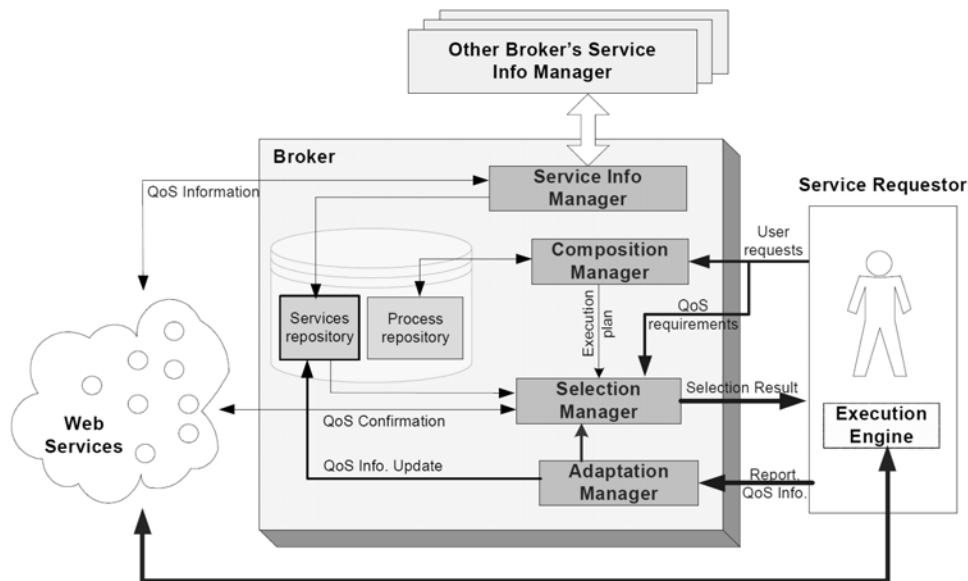


Figure 4: QoS Broker Architecture.

In [6] Yu and Lin introduce an agent based framework (figure 4) for the composition of web services based on QoS. They introduce a solution that, based on user requests, returns a composed business process, which is then executed with the help of an execution language such as BPEL (Business Process Execution Language).

The main functions of a QoS broker include:

- **Service tracking:** A broker has a service repository to record all feasible web services it is aware of
- **Dynamic service composition:** A broker maintains some predefined business process plans in the process repository. New plans can be built or existing ones updated based on user requirements and newly discovered services;
- **Dynamic service selection:** This is the key function of the proposed service broker. It selects services to execute a business process so that the user-defined utility is maximized, and user's QoS requirements are satisfied;
- **Dynamic service adaptation:** In case of individual service failure during a business process execution, a broker needs to reconstruct the business process to ensure a good performance.

6 CONCLUSION

In this paper we introduced potential technologies, which could represent a logical evolution, rather than a revolution of services. By introducing QoS and OWL-S to web service discovery, invocation, composition and interoperation, web services could be simplified and abstracted to the next level, where we wouldn't search for a particular service, but only specify the type of service we want to execute and our QoS requirements. A service agent, which would compare services by the given QoS, would perform the search for the optimum service and with the help of OWL-S, it could easily execute any of them.

The current state of both mentioned technologies is rather disappointing, because after more than 5 years of their first mentions they are still considered to be in their early adoption phases, particularly QoS, which doesn't even have a standard set of requirements. The current definition of those requirements isn't enough, because we would still need a standard that defines them, specifies how to measure them and compare different services with them. Only then would we be able to fully leverage OWL-S and all of its advantages.

The described approaches solve adaptation in their own way, but if a convergence could be achieved from all relevant researches we would be able to standardize the result and pave a new way for service flexibility and reliability.

References

- [1] IBM, »Understanding quality of service for Web services«, <http://www.ibm.com/developerworks/library/ws-quality.html>
- [2] W3C, »QoS for Web Services«, <http://www.w3c.or.kr/kr-office/TR/2003/ws-qos/>
- [3] W3, »OWL-S«, <http://www.w3.org/Submission/OWL-S/>
- [4] L. Zeng, B. Benatallah, M. Dumas, J. Kalagnanam, Q. Z. Sheng, Quality Driven Web Services Composition, conference *World Wide Web*, Budapest, 2003
- [5] J. Hielscher, R. Kazhamiakin, A. Metzger, A Framework for Proactive Self-adaptation of Service-Based Applications Based on Online Testing, conference *Towards a Service-Based Internet*, Madrid, 2008
- [6] T. Yu, K. Lin, A Broker-Based Framework for QoS-Aware Web Service Composition, conference *e-Technology, e-Commerce and e-Service*, Hong Kong, 2005

DYNAMIC SERVICE BUSINESS MODELS: A PROPOSAL FOR UNIFIED SERVICE PRICING FRAMEWORK

Kristjan Košič¹, Reinhard Bernsteiner², Marjan Heričko¹

*¹University of Maribor, Faculty of Electrical Engineering and Computer Science
Smetanova ulica 17, 2000 Maribor, Slovenia*

kristjan.kosic@uni-mb.si, marjan.hericko@uni-mb.si

*²Management, Communication & IT, MCI - Management Center Innsbruck
Universitaetsstrasse 15, 6020 Innsbruck, Austria
reinhard.bernsteiner@mci.edu*

ABSTRACT

In this paper we report our research work on dynamic service business models. Various service pricing strategies are presented. Based upon current research findings and pricing models in the use today we propose a unified service pricing framework with its emphasis on the dynamics. The drivers of dynamics pricing mechanism are service-oriented computing and service level agreements, which importance is stressed in this paper. At the end we present future research ideas and still open research areas.

1 INTRODUCTION

Software industry is transitioning to service industry. Organizations are realizing that software can and must be treated as a value delivery process, a view that brings many new opportunities and opens new market areas. Nowadays, service consumers mostly buy results, rather than applications [1], and therefore appropriate business models must be adopted in order to succeed and sustain competitiveness. As mentioned by Vassiliadis in [1], the traditional e-commerce paradigm is giving way to e-service paradigm, and service-oriented computing (SOC) is the driver for this change.

Free internet access has fueled the growth of software systems and software products in a form of service evolution. Software rental or result buying is based upon pay-per-use basis or a subscription model [2]. Thin client concept has become ignited again by the service-oriented computing and e-services paradigm. Thin clients operate by downloading or accessing functionalities (applications or services) by per need basis. They depend on service providers, which are also known as “Computer Utility Companies” or “Application Service Providers (ASP)”.

However, nowadays when everything is becoming interconnected, integrated and therefore depended on each other, we need new pricing mechanisms that will provide satisfaction on both sides (service consumers, and service providers).

Most of the research on the service pricing mechanisms was made in the area of optimal resource allocation and

realization of single service functionalities. However there was no research that was addressing depended service pricing mechanisms. Our paper complements and supplements several other papers [3,4,5,6] addressing service pricing issues and dynamic service pricing models. It proposes a unified service pricing framework.

1.1 Outline of the paper

This paper is organized in the following manner. Section two motivates and presents basic service pricing models and the role of Service Level Agreement (SLA) in the model. We present a literature review of dynamic pricing models. In section three, we outline a foundation for service pricing mechanism. Architecture is based on a web-service realization. We conclude the paper in section four where we also present future research ideas.

2 SERVICE PROVISIONING AND DYNAMIC SERVICE PRICING

As mentioned in the introduction, our research was focused on analysis of dynamic service pricing models. Most of the existing service classifications and categorizations [6,7,8] that we found were resource based. Models that defined service resource allocation pricing and reservations were mostly mathematical and defined from economical view. The missing part was a clear definition of service roles in the process and last, but not least, Service Level Agreements (SLA) that define the service contracts.

However, our idea expands service pricing, not just on resource pricing and allocation, but on “dynamic” service provisioning. Under dynamic we understand dynamic service behavior, which is driven by clients (service consumer) decisions and selections in the business process or just parts of the business process.

Customers should be able to control total cost of technology ownership through scheduled payment schemes, and service-oriented computing [1] is the realization technology for that. SOC has enabled new flexible models, merely focused on-line, on machine to machine provisioning. The main difference between Application Service Providers (ASP) and SOC is that ASP is focused on a provision of a solution or a tool, and SOC provides entire management

and a solution [1]. There are two types of service providers [1]:

- *Application Services*: they were initially focused on transaction processing (e.g. payroll services). This application services also include web services.
- *Infrastructure Services*: also known as utility service (Utility Service Computing). The user runs its own services on a vendor provided information system. Today these services are hidden under buzzwords such as “grid and cloud computing”. An example would be Amazon Services Platform.

There is a variety of charging algorithm for services. These algorithms use different realization technologies (game theory, machine learning, decision models and others). Some charge a yearly fee for a defined number of services, other charge per user or per month. There are also billing models that charge by pay-as-you-go (based upon time metered basis). As already mentioned in this paper, our goal is the realization of dynamic service provisioning framework. Service Level Agreements are the means that will help us to realize this promise. In next section we will present service level agreements their role and the formalization needed for realization of dynamic pricing systems.

2.1 The role of Service Level Agreements in the service pricing process

Service level agreement (SLA) is defined as a formal contract between the service provider and the user who is consuming offered services (Figure 1).

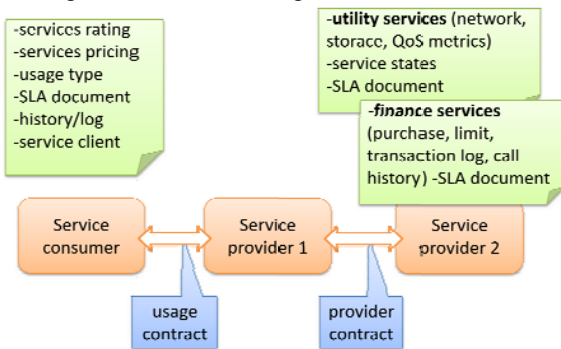


Figure 1 Service contract relationships [5]

Services have to be acceptable by the user. SLA realization can be measured with metrics (at least from the quality of service-QoS point of view). Clear-well defined SLA is a must for a disputed free information system. A typical SLA duration is one to three years [1]. SLA becomes more complex and complicated depending on the services we use and the type of business and goals of the user. Computing technology industry association [9] defined four types of SLA, which are still in use today:

- *Network SLA*: defines network connection between client and service provider.

- *Hosting SLA*: server availability, disaster recovery, physical server security.
- *Application/Service*: measurement of application service performance. It includes definitions of metrics and penalties for failure to meet agreed service levels including service availability, performance, upgrades.
- *Help desk SLA*: technical and nontechnical support for customers (customer satisfaction, responsiveness).

Although SLAs are very well formally defined, their realization lacks in practice. It is very difficult to define complete agreements, especially because of the multi-level agreements. Another thing missing in realization of SLA is insufficient cost/pricing information. As mentioned by Vassiliadis in [1], existing SLA methodologies need to be improved to enable unified, business-oriented approaches for fulfilling commitments. SLA must be adaptable as the service to which they are measured [10] (Figure 2).

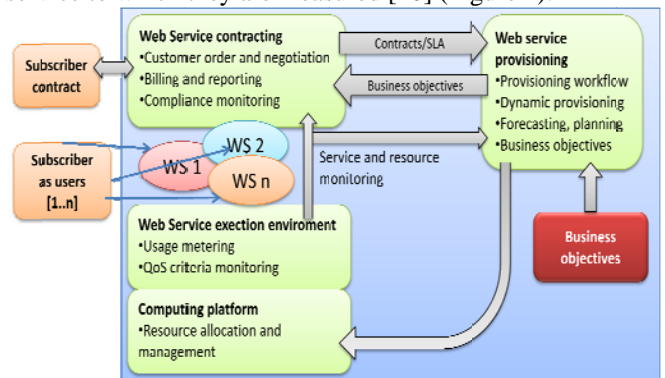


Figure 2 WSLA computing environment [5]

IBM proposed first formal specification of SLA in 2004 [5]. The product was called SAM (SLA Action Manager). SAM is distributed SLA management system that can be incorporated in utility computing environment. This technology was then incorporated into IBM ETTK [5,11] (Emerging Technologies Toolkit) and is commonly known as a WSLA framework. The main driver, for a successful realization of this framework is Web Service Level Agreement language-WSLA [11].

SLAs are essential, they need to be defined at the right time and with all the necessary information (pricing, performance, reliability and others). Without SLA a successful realization of service-oriented offerings is hard to realize.

3 UNIFIED SERVICE PRICING FRAMEWORK PROPOSAL

In section one and two we became familiar with underlying service pricing concepts (pricing categories and strategies, SLA, SOC) that define and make-up service pricing process. In order to realize dynamic service provisioning framework we propose a new framework, which will be realized with web-services technology stack. Our architecture proposal

extends already known service pricing models (intranet Resource Management Unit (iRMU), VPN Traffic Pricing Experiment System (VTPES) and Stochastic General Economic Equilibrium (SGEE)) defined by Lin and Zhangxi in [8], by providing additional pricing information and simulation options. Most of the known models were defined for specific domains. Suggested architecture is based upon conceptual components of iRMU and VTPES[8].

Modules of the pricing system [8] are related to the layers in web-services technologies stack as shown in table 1[8].

Table 1 Pricing system modules

Module	Technology stack
Service Billing	BPEL4WS/UDDI
Service Admission	BPEL4WS/UDDI
Service Access Utilities	BPEL4WS/UDDI
Traffic Measuring	SOAP/NETWORK
Resource Scheduling	SOAP/NETWORK
Resource Pricing	UDDI/WSDL

Service billing module bills service access as defined in UDDI repository. Resource pricing modules updates the prices for resource usage, which is measured by traffic measuring module. Service access utilities help clients to set up their payment accounts.

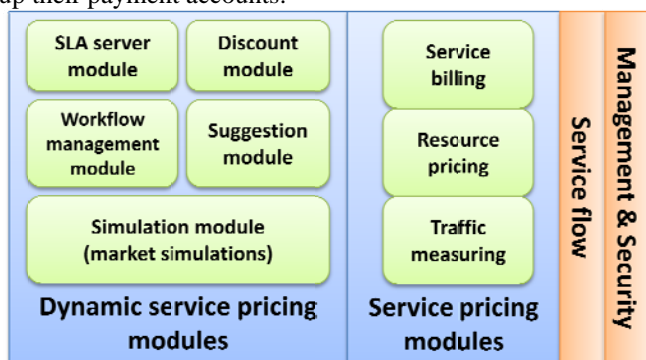


Figure 3 Service pricing system modules

Above picture shows some the modules of web pricing system as proposed by Lin and Zhangxi [8] and dynamic modules recommended in our framework. Architecture proposed by Lin and Zhangxi [8] lacks the dynamics, although the user can select services, and view the transaction or usage price at the end. The dynamics should give clients the ability to cooperate in the pricing process, to see which services are better/cheaper and what is the best service workflow. This dynamics can be realized with IBM's WSLA framework [5,11]. Our proposed architecture extends existing architecture [8] with the following modules (Figure 3): (i) SLA Server Module (SSM), (ii) Workflow Calculation Module (WCM), (iii) Service Discount Module (SDM), (iv) Service Suggestion Module (SGM) and (v) Simulation Pricing Module (SPM).

SLA server module, with the help of WSLA service specifications, has access to all levels of the SLA contract (levels are layered upon different contracts, resources and

providers). Module has all information related to service and resource pricing.

Workflow calculation module takes all selected services from client business process into calculation. Based on the semantic description of the services, the module searches for additional services with similar functionalities and suggests them to the client. In this way client has full control, over the service selection process, even more, the semantic agents help him to find lower priced services or service workflows and combine them into his process.

Service discount modules main function is to take into account services, with special discounts or offerings. Module works in combination with service suggestion module. The main functionality of this module is the ability to suggest better alternatives for the business process selected by the user and from the marketing point of view, to offer and promote new services.

One of the main differences from previous pricing systems (iRMU, VTPES, SGEE and Lin and Zhangxi [8]) is that business process consists from one or more service requests, which brings us to composite services and transactional behavior. Service pricing is no longer static based, but it takes into account service workflow defined by the user and it dynamically suggests better offerings from the service offering module. Table two summarizes new proposed modules.

Table 2 Dynamic service pricing modules

New modules	Functionalities
SLA server module	-Pricing information -Usage contract realization -Provider contract realization -QoS specification -WSLA compliance monitoring
Workflow module	-Business process realization -SLA contract verification and validation
Discount module	-Pricing offerings -Special offer calculation -Workflow suggestion
Suggestion module	-Semantic search for similar services -Service suggestion (price, functionalities)
Simulation module	-Service providers can simulate pricing situation based on the QoS metrics collected by the traffic measuring engine -Optimization of pricing criteria -Market simulations

From a client point of view, the workflow of the proposed architecture would consist of the following steps:

- Service offerings are registered in the main repository by service providers.
- Client is selecting services for business process realization.

- During service selection client is informed of similar services with better offerings.
- Semantic module is also monitoring service selection process and suggests similar processes with special offerings.
- Client confirms service selection – a new usage contract is defined.
- Client confirms usage contract and enters additional payment information. Usage contract defines payment strategies and billing cycle.
- Client can start using requested services.

But not all service offerings in the main repository are from the same service provider. That is where provider contract realization comes into account. All the services user selected for the realization of the business process are contained in different SLA layers between various service providers.

From the service provider point of view, service publishing in the main repository would consist of the following steps:

- Service provider publishes service in the main repository.
- Client can consume the service if it is not composed in the business process.
- During publishing process service providers offer their starter pricing offerings on the server. A template SLA for the provider is created.
- Other providers negotiate and define their provider contracts based upon provided template.
- Service can be consumed in various business processes; the only condition is the existence of intermediate SLA agreements.

Service providers can also simulate and test pricing options defined in SLA in the simulation module. Based upon QoS metrics, pricing options can be optimized for all parties (end-users and service providers). SLA's are by nature adaptable and configurable.

4 CONCLUSIONS AND FUTURE WORK

Information goods and service pricing is a critical topic in the new future internet era. Advances in the internet and e-business technologies have opened rich opportunities where dynamic pricing models can be used. Companies resorting to dynamic pricing strategies are increasing in numbers steadily [12]. Increasingly complex dynamic strategies are being researched from various aspects like machine learning, game theory, expert systems and genetic algorithms. In this paper we have covered the following topics:

- We presented current service pricing business models.
- We categorized service pricing methods.
- We presented a shift from the ASP to SOC and utility services (service outsourcing).
- We stressed out the importance of SLA in the dynamic pricing decision process.
- Based on the existing service pricing models, we proposed a unified framework for service pricing.

The main message of this paper is that there is a growing need for dynamic and adaptable business models, especially in the services (e-services) domain. For future research we plan to implement the proposed unified service pricing framework and use it as a simulation tool for market simulations (clients, consumer, buying and selling behavior, dynamic pricing strategies).

REFERENCES

- [1] B. Vassiliadis, A. Stefani, J. Tsaknakis, and A. Tsakalidis. From application service provision to service-oriented computing: A study of the IT outsourcing evolution, *Journal of Telematics and Informatics* 23(4), 271–293, 2006
- [2] H.K. Cheng, G.J. Koehler. Optimal pricing policies of web-enabled application services, *Decision Support Systems* 35(3), 259–272, 2003
- [3] W. Chang, S. Yuan. IPrice: A collaborative pricing system for e-service bundle delivery. *Int. J. of Electronic Business*, 5:623–642, December 30 2007. ISSN 1741-5063.
- [4] E. Julian, D. Ertogrul, A. Papageorgiou, N. Repp, and R. Steinmetz. The impact of service pricing models on service selection. *International Conference on Internet and Web Applications and Services*, 316–321, 2009.
- [5] A. Dan, D. Davis, R. Kearney, A. Keller, R. King, D. Kuebler, H. Ludwig, M. Polan, M. Spreitzer, and A. Youssef. Web services on demand: WSLA-driven automated management. *IBM System. Journal*, 43(1):136–158, 2004. ISSN 0018-8670
- [6] T. Kern, J. Kreijger, and L. Willcocks. Exploring asp as sourcing strategy: Theoretical perspectives, propositions for practice. *The Journal of Strategic Information Systems*, 11(2):153–177, 2002. ISSN 0963-8687.
- [7] D. Gouscos, M. Kalikakis and P. Georgiadis. An Approach to Modeling Web Service QoS and Provision Price, *Fourth International Conference on Web Information Systems Engineering Workshops (WISEW'03)*, IEEE Computer Society, 121–130, 2003.
- [8] L. Zhangxi, R. Sathya and Z. Huimin: Usage-based dynamic pricing of Web services for optimizing resource allocation, *Information Systems and E-Business Management* 3(3), 221–242, 2005.
- [9] CompTIA, www.comptia.org/visited in 2009.
- [10] T.O. Bangemann. *Shared services in finance and accounting*, Gower Publishing, ISBN: 978-0-566-08607-6, 2007.
- [11] WSLA, <http://www.research.ibm.com/wsla/>, visited in 2009.
- [12] Y.R. Narahari, K. Ravikumar, and S. Shah. *Dynamic pricing models for electronic business*, Indian Academy of Sciences, April 2005.

Acknowledgement

This work is partly supported by the Slovenian Research Agency and Austrian Office for Academic Co-operation and Mobility, Grant BI-AT/09-10-004.

SUCCESS FACTORS AND BARRIERS OF KNOWLEDGE MANAGEMENT – AN EMPIRICAL ANALYSIS OF A SHAREPOINT 2007 IMPLEMENTATION

Michael Amberg, School of Business and Economics Friedrich-Alexander University Erlangen-Nuremberg, Nuremberg, Germany, Michael.Amberg@Wiso.Uni-Erlangen.de

Michael Reinhardt, School of Business and Economics Friedrich-Alexander University Erlangen-Nuremberg, Nuremberg, Germany, Michael.Reinhardt@Wiso.Uni-Erlangen.de

Jiangping Weng, School of Business and Economics Friedrich-Alexander University Erlangen-Nuremberg, Nuremberg, Germany, jianping.weng@Wiso.stud.uni-erlangen.de

ABSTRACT

This paper gives insights into the success factors and barriers of Knowledge Management in theory and practice. Organizational, conceptual and technical conditions are analyzed in regard to their role for a successful Knowledge Management implementation in enterprises. A real world, SharePoint Server 2007 based enterprise Knowledge Management system is described and evaluated via a structured user survey. The results and discussion of our findings are aimed to help practitioners on their way for a successful Knowledge Management implementation.

1. INTRODUCTION

Industrialized countries experience an increasing shift from an industrial society towards knowledge society (cp. [Willke et al. 2001]). In this context a huge challenge arises for enterprises to manage and enhance their knowledge in an effective and efficient way. These challenges can be found on an individual, social and corporate level, as the amount of information increases steadily while the maximum usability period of knowledge steadily shrinks (cp. [Fücks2008], [Falk2007]). Furthermore, new information technology (IT) provides steady potentials improved workflows (cp. [Bartel et al. 2007]).

Hence, the ability to manage knowledge efficiently and with a systematic approach becomes a crucial success factor for enterprises in the information age. This paper tries to outline practical experiences to answer the question: Which barriers and success factors are perceived for Knowledge Management, based on an implementation of a SharePoint IT system?

First, this paper gives a theoretical background on Knowledge Management (KM) definitions and concepts. Subsequently, the KM project environment and IT system of a major German electronic-company is described, as it aims to foster the KM, using a Microsoft Office SharePoint Server 2007 implementation. Subsequently the research approach is described. Based on an online questionnaire survey, the success factors and barriers of KM are discussed and improvement potentials are shown.

2. THEORETICAL BACKGROUND

The terms data, information and knowledge are related, but frequently used synonymously or mistakenly. To be clear about the concepts, we draw on [North2005], who suggest the following stairway model for these terms: “Data” usually refers to an order of “characters” or “symbols”, e.g. letters or numbers without evident meaning, but using already certain syntax. Carrying certain meaning data forms “information”. This single information, connected with other information, e.g. in a certain context or experiences, forms knowledge. As knowledge shall enable a goal oriented “action”, motivation and application of the knowledge are important. If this “action” is carried out right, [North2005] speaks of competence. Finally, a unique or “better than others” competence enables a company to gain a competitive advantage (cp. [Porter et al 1985], [Porter2008]).

2.1 Knowledge Management

Knowledge management was defined by [Takeuchi1995] as “the process of continuously creating new knowledge, disseminating it widely through the organization, and embodying it quickly in new products/services, technologies and systems”. This concept consequentially also addresses three major components of knowledge management namely people (cp. [Arthur et al. 2008]), organization (cp. [Mayasandra et al. 2004]) and technology (cp. [Bartel et al. 2007]). They should be equally taken into account, when considering the purpose of knowledge management. Further specific conditions of a company lie in a firm’s structure (cp. [Teece2000]), its coworker motivation (cp. [Kulkarni et al 2007]) and the employee qualification (cp. [DeTienne2004]). A practical approach to implement KM is suggested by [Probst et al. 1999], classifying the knowledge management actions which are to be taken within an organization. Thus Knowledge Management is divided into six core categories, as visualized in Figure 1.

The first process of *knowledge identification* shall foster internal and external transparency for existing knowledge. In our case it addresses the approach of an integrated system and mechanisms to retrieve knowledge within this system.

The goal of *Knowledge acquisition* is to identify abilities that are to be purchased externally. An organisation has

external sources and relationships to customers and suppliers as well as competitors. Thus an organisation has the possibility to internalize knowledge by acquiring experts from innovative enterprises and therefore acquire new knowledge that it probably could hardly create with its own resources. An integrated system with suitable access control addresses this point.

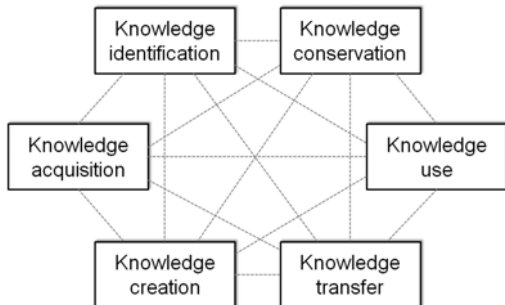


Figure 2: Knowledge Management Bricks, according to [Probst et al. 1999].

Generating and establishing (new) knowledge within an organization is the task for the *knowledge creation*: Based on a good usability the system shall enable the creation and development of (new) knowledge, based on the cross-linking of existing information.

Knowledge transfer aims to bring knowledge to the right place. This process is obligatory to make isolated information and experience usable for the whole organisation. Further, implicit knowledge shall be made explicit and transferable.

Knowledge use addresses the (situated) utilization of knowledge. Given the increasing information load, the effective usage of all available knowledge shall be ensured.

Finally, the *knowledge conservation's* ambition is to protect a company from knowledge loss. A major challenge within the analyzed project was the prevention of knowledge loss when elderly employees leave the company. Due to demographic change, this challenge is likely to increase over the next years.

Of course, these knowledge management bricks can not be handled independently, as they have correlation effects to each other. Consequently, a system dealing with these processes must form “an integrated and integrative technology architecture [...]”, as a Knowledge Management system seems “to require a variety of technological tools in three areas: database and database management, communication and messaging, and browsing and retrieval” [Alavi et al. 1999].

2.2 SharePoint Server 2007 Implementation

The implementation of the Knowledge Management system is based on a Microsoft SharePoint 2007. It provides a web-based interface for the department-wide collaboration and the management of information and documents (cp. [Microsoft2009]). In April 2007, the decision was taken to support the internal KM with

SharePoint, the project kick-off was in May 2007 and the roll-out was finished after 8 months at the end of 2007. The project team was composed of domain experts and technical support of the IT department. The core functionalities that have been realized with the analyzed solution are the following: *Navigation and access* via customized menu structures and hierarchies; *access management* for folders, documents, list elements and users; *document management* supporting storage, meta data management, versioning, etc; a *wiki platform* to link knowledge; *customizable messaging* in case of updates on elements; comprehensive *search-functionality*.

3. RESEARCH METHOD

To gain practical insights into the Knowledge Management success factors and barriers, the Microsoft Office SharePoint Server 2007 based KM-system of a major German electronic-company was evaluated. Therefore, an online questionnaire (cp. Chapter 4) was designed and implemented, encompassing 10 questions, which were derived from theoretical research and practical insights, in form of interviews and workshops. A five-step scale has been given for most questions to the interviewees. The answer categories at multiple choice answers were weighted with the mean (+2 for “totally agree”, +1 for “agree”, 0 for “neutral”, -1 for “disagree”, -2 for “totally disagree”).

The questionnaire could be filled in for six weeks during spring 2008. About 250 employees were invited for participation, of which 57 answered to the questionnaire. The goal of the survey was to establish crucial factors for a successful Knowledge Management.

4. RESULTS AND DISCUSSION

The first question to the participants was, *if they actually know the Knowledge Management Portal*. Responding to that question, 88 percent are already familiar with the application, while 12 percent did not get to know it so far. Of course, information events and newsletters were launched previously during the project, in which the portal has been announced. Hence, a positive side effect of the survey was, to draw their attention towards the application or to remind them about the completion. Only 28 percent of the respondents assessed the *current use of the department's internal knowledge* with “very good” or “good”, while 34 percent perceived the knowledge usage as “insufficient” or “very insufficient”. 39 percent took a neutral attitude or gave no statement (cp. Figure 2).

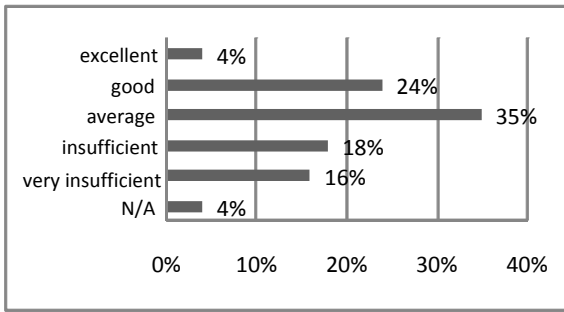


Figure 2: How do you rate the current use of existing knowledge in your department?

On the other hand 57 percent of the respondents expect Knowledge Management to be “important” or even “very important” for the department in the future. Only 7 percent think of Knowledge Management to be “unimportant” (cp. Figure3). This importance correlates with our expectations in the beginning and also indicates that there is still potential for improvement left.

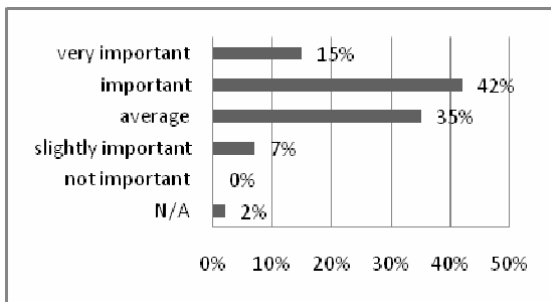


Figure 3: How important is the Knowledge Management with the KM-system in the future?

Based on this question, we see that there is a strong demand for a new approach in Knowledge Management. The following questions try to find out which are most crucial fields of action and potential starting-points. Hence, it is very important to be clear about the goals, which shall be reached with a Knowledge Management initiative. These goals form the foundation of further roll-out and communication initiatives.

As Figure 4 illustrates, about 89 percent of the respondents expect primarily improvements in regard to the collection and documentation of experiential knowledge. Furthermore, the development and exploitation of knowledge are seen as important goals. In fact this is a key point, as it enables the department to act according to existing knowledge and also stay competitive and adapt itself to new situations (cp. [Tushman1986]). Another crucial factor, which shall be supported with the means of an IT-system, is the transfer of knowledge. The enabling of such knowledge flows supports the creation of new capabilities and the efficient usage of knowledge (cp. [Davenport et al. 1998]).

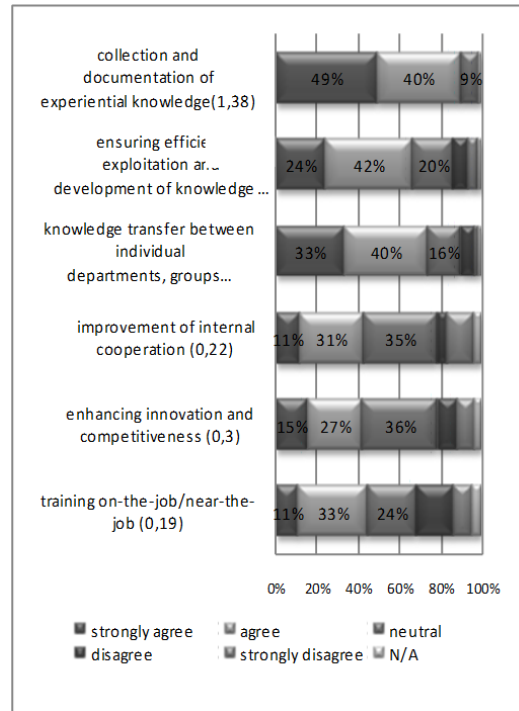


Figure 4: Which goals shall be achieved by KM-system?

Being asked for the required skills to use the KM system, 58 percent perceived internet skills as “important” or “very important”. 71 percent named the domain knowledge as at least important and 62 percent thought their practical experience to be at least important (cp. Figure 5).

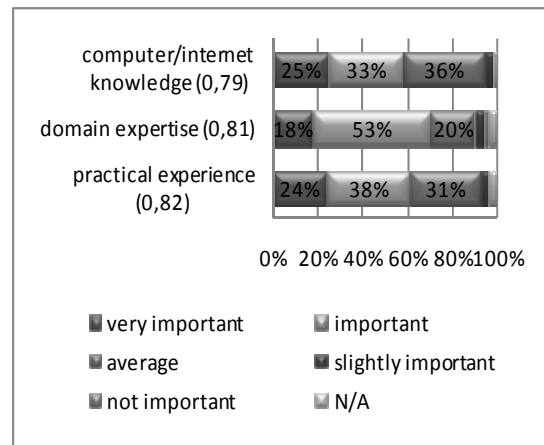


Figure 5: How do you evaluate the following knowledge as precondition for the usage of the KM-system?

This expectation underlines that Knowledge Management is much more than providing just a tool and that the crucial factors lie in the conceptual framework of an application. Hence, effort must be taken to define structures (e.g. categories, links), in which the knowledge can be stored analogical to naturally used categories. Furthermore, navigational elements and search mechanism should support the finding of the right knowledge at the right time. It is

obvious that the perceived *success factors* are very important for an effective KM (cp. Figure 6).

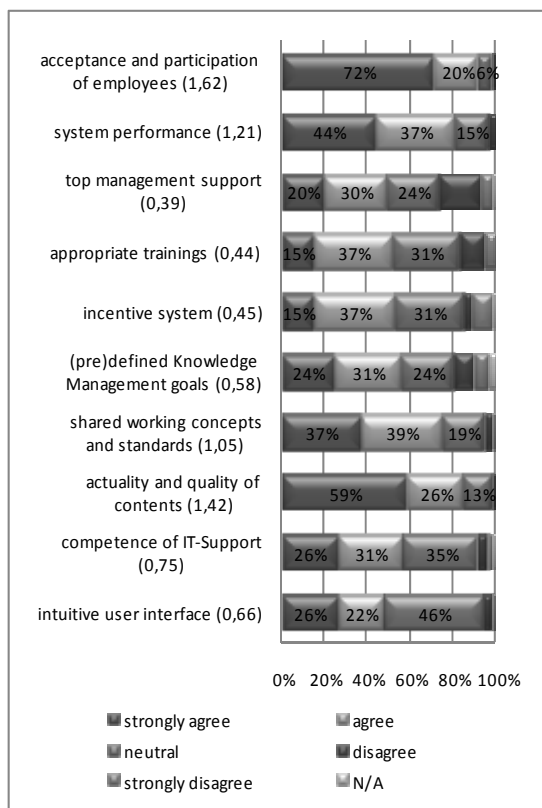


Figure 6: Which factors are particularly important for a successful KM-system?

The crucial success factors are seen in areas, where crucial KM-requirements meet the daily business. Especially the acceptance and participation of the employees, the system performance and the actuality and quality of the contents can be named at this place. These factors are important to provide an additional value to the users and to reduce the additional overhead as far as possible. Software, tools and services, which are able to gain a high user acceptance, usually show a very good usability (cp. [Courage et al. 2004]). Hence it is important to meet the functional and domain specific requirements and parallel and provide a tool, with a high ease of use (cp. [Courage et al. 2004]). Further, the value of a KM system increases with the number of active users. A sustainable and effective development of the KM-system requires an active and problem solving oriented approach. An interesting question is how the system can be integrated into the daily business and the underlying (editorial) processes.

Going one step further on the technical side, we asked for *crucial technical functions* for the KM system (cp. Figure 7).

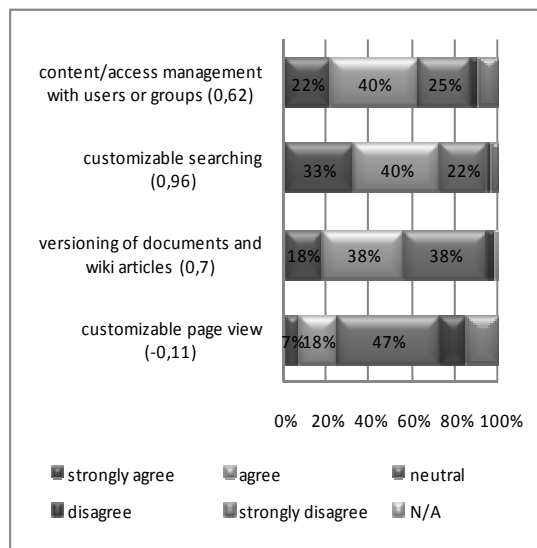


Figure 7: Which of the following functions are particularly important for the KM-system?

Significantly, the search functionality was rated to be most important on the functional side, as it enables the access of knowledge in regard of finding relevant information. Another important point was the access management, as it cares for the rights and roles and addresses questions of confidentiality.

Finally, we asked for the *current barriers in regard to a more extensive system usage* (cp. Figure 8). Interestingly, there are only four barriers with a positive assessment, which indicates that only these barriers are perceived as real barriers by the respondents. The biggest barrier is seen in the time shortage. People do not seem to have enough time to use the system effectively – for example to edit articles or browse and discover information. Other critical barriers, which should be actively worked on, are the outdated information provided, the missing usage and the missing expert support.

These barriers point out related issues than prior research, that “effective knowledge management systems involve far more than just technology, encompassing broad cultural and organizational issues” [Alavi et al. 1999].

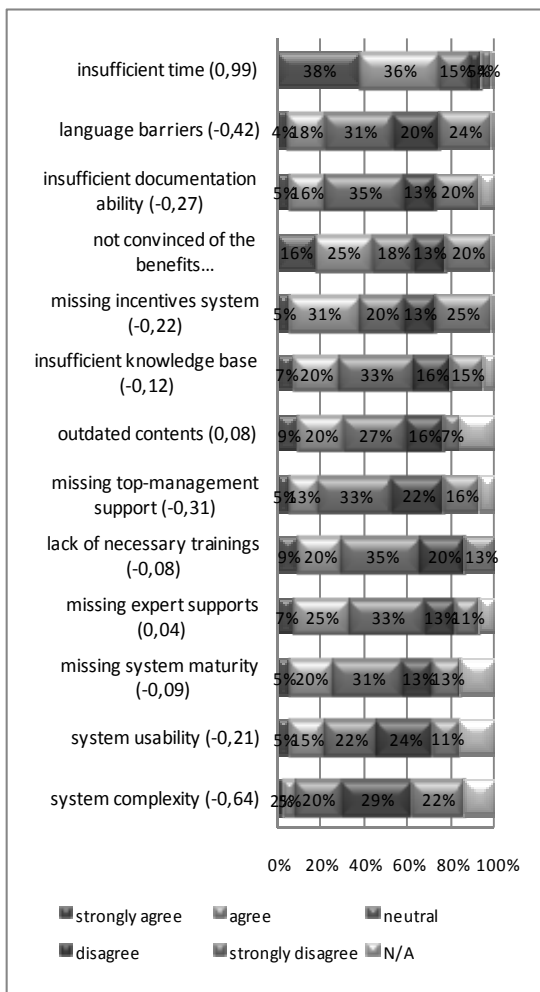


Figure 8: Which of the following barriers is preventing currently a successful implementation of the KM-system?

References

[Alavi et al. 1999] Alavi, M., Leidner, D. (1999). Knowledge Management Systems: Issues, Challenges and Benefits. Communications of the Association for Information Systems, Vol. 1, Art. 7.

[Arthur et al. 2008] Arthur, M., Defillippi, R., Lindsay, V. (2008). On Being a Knowledge Worker. Organizational Dynamics, Vol. 37, No. 4, pp. 365–377, Elsevier.

[Bartel et al. 2007] Bartel, A., Ichniowski, C., Shaw, K (2007): How does Information effect Productivity? Plant-Level Comparisons of Product Innovation. Process Improvement, and Worker Skills. In: The Quarterly Journal of Economics, November 2007.

[Courage et al. 2004] Catherine Courage & Kathy Baxter (2004): Understanding Your Users – A Practical Guide to User Requirements Methods, Tools, and Techniques. Morgan Kaufmann, 2004.

[Davenport et al.1998] Davenport, T. de Long, D.W. and Beers, M. (1998): Successful knowledge management projects. Sloan Management Review, 43-57.

[DeTienne2004] DeTienne, K., Dyer, G., Hoopes, C., Harris, S. (2004). Toward a Model of Effective Knowledge Management and Directions for Future Research: Culture, Leadership, and CKOs. Journal of Leadership & Organizational Studies (Baker College), Vol. 10 Issue 4, p26-43, London: SAGE.

[Falk2007] Falk, S. (2007). Personalentwicklung, Wissensmanagement und Lernende Organisation in der Praxis: Zusammenhänge - Synergien - Gestaltungsempfehlungen (2. Edition., Vol. 2). München, Hampp.

[Fücks2008] Fücks, R. (2008): Wissen ist Macht. Ein Plädoyer für eine neue Wissenspolitik. Retrieved May 05, 2008, from Heinrich Böll Stiftung: <http://www.boell.de/alt/downloads/zukunftsgruen/zg4fuecks.pdf>.

[Kulkarni et al 2007] Kulkarni, U., Ravindran, S., Freeze, R. (2007). A Knowledge Management Success Model: Theoretical Development and Empirical Validation. Journal of Management Information Systems, Vol. 23 Issue 3, p309-347, New York: M.E. Sharp Inc.

[Mayasandra et al. 2004] Mayasandra, R., Pan, S. (2004): Consequences of Implementing Knowledge Management Initiatives in Different Organizational Subcultures. European Conference on Information Systems (ECIS).

[Microsoft2009] Microsoft SharePoint Website (2009). <http://sharepoint.microsoft.com/Pages/Default.aspx>.

[North2005] North, Klaus (2005): Wissensorientierte Unternehmensführung: Wertschöpfung durch Wissen. 4th. Edition, Gabler, Wiesbaden.

[Porter2008] Porter, M.E. (2008): The Five Competitive Forces That Shape Strategy, In: Harvard Business Review, Januar 2008, S. 78-03.

[Porter et al 1985] Porter, M.E. & Millar, V.E. (1985): How information gives you competitive advantage, In: Harvard Business Review, Juli/August 1985, S. 149-174.

[Probst et al. 1999] Probst, Gilbert., Raub, Steffen, Romhardt, Kai (1999): Wissen Managen: Wie Unternehmen ihre wertvollste Ressource optimal nutzen. 3. Edition. Frankfurt/Main, Gabler.

[Takeuchi 1995] Takeuchi, H. (1995). The knowledge creating company. How Japanese companies create the dynamics of innovation. New York: Oxford Univ. Press.

[Teece2000] Teece, D. (2000). Strategies for Managing Knowledge Assets: the Role of Firm Structure and Industrial Context. Long Range Planning, Vol. 33 Issue 1, p35-54, Elsevier.

[Tushman1986] Tushman, M. L. / Anderson, P. (1986): Technological Discontinuities and Organizational Environments, In: Administrative Science Quarterly, 31. Jg. (1986), Nr. 3, S. 439.

[Willke et al. 2001] Willke, H., Krück, C., & Mingers, S. (2001). Systemisches Wissensmanagement. (2. Edition). Stuttgart, Lucius und Lucius; Lucius & Lucius.

THE TRUE VALUE OF AN E-LEARNING SYSTEM THROUGH THE STUDENT'S EYE

Boštjan Šumak, Maja Pušnik, Marjan Heričko

University of Maribor, Faculty of Electrical Engineering and Computer Science

Smetanova ulica 17, 2000 Maribor, Slovenia

Tel: +386 2 220 7378; fax: +386 2 220 7272

e-mail: bostjan.sumak@uni-mb.si

ABSTRACT

With the advent of Web 2.0, e-learning systems are being extended with new features, such as wikis, blogs, RSS and 3D virtual learning spaces. This study examined the importance of different e-learning system services from a student's point of view. Quantitative research was performed to answer questions about the importance of several e-learning system services and the frequency of actual use of individual e-learning system services. Statistical tests (t-test, ANOVA) were performed during the search for differences for various independent variables like gender, age, internet experience, e-learning experiences, number of courses where the e-learning system is used and the frequency of e-learning system use. Although students find that an e-learning system should provide Web 2.0 technologies and services, these are not fully exploited. This paper concludes with a discussion and directions for future research activities.

1 INTRODUCTION

The e-learning market in 2001 was worth \$5.2 billion and was expected to reach \$23.7 billion in 2006, according to predictions by the International Data Corporation (IDC). Another IDC estimate predicted that the value of the e-learning market worth will be between \$21 billion and \$28 billion by 2008 [1]. According to Cortona Consulting, corporate e-learning in the US and Europe will reach \$50 billion by 2010 [2]. For several years, e-learning systems have been growing in different parts of the industry, especially in education. Despite initial resistance, e-learning systems have managed to come forward and have consequently, thanks to user feedback, become more user-friendly. E-learning systems have developed and are starting to satisfy all the needs of a modern student. With the advent of Web 2.0 technologies and services, e-learning systems are being extended with new features, such as wikis, blogs, RSS and 3D virtual learning spaces. Whenever a new learning environment is presented, it needs to be adapted, where the adaption process has to deploy different learning styles [3]. Developers, designers, and institutional purchasers of e-learning systems must carefully consider the needs and values of system users and

ensure that the system in question effectively meets their demands [4]. E-learning 2.0 is about collaboration and places an emphasis on social learning and the use of social software such as blogs, wikis, podcasts, etc. Web 2.0 technologies and services play a crucial role in the building of new and innovative e-learning environments, which are providing interoperable social and learning services. Through these learning environments, content is not only being composed, but syndicated, as in the case of blog posts and podcasts. Learners were used to being passive actors. New technically and socially open Web 2.0 services are encouraging learners to actively participate in the creation of learning content. It is expected that with the advent of E-learning 2.0, content is more likely to be produced by learners than teachers. Many teachers and subject specialists already use blogs to publish up-to-date information and commentary on their subject areas. Podcasts are also a very popular technology, used for pushing learning content to learners. RSS readers and similar applications can be used to aggregate content. Media sharing sites like Flickr, Youtube, Slideshare, etc. can be used for publishing and providing multimedia resources to learners or wider audience. Through the above-mentioned Web technologies and services, students can not only consume content but can also very easily tag and discuss with authors and other participants.

The present study aims to investigate students' perceptions about the importance of different e-learning system services. The study also aims to reveal the frequency of the students' use of individual e-learning system services. The paper is organized as follows: in the next chapter, a definition of an e-learning system is given together with the research questions that the study hopes to answer. In the chapter that follows, the research methodology is presented. Chapter four gives descriptive statistics and the results of statistics methods are given to further explain the influence of different factors (age, gender, e-learning experience, e-system experience, the number of courses where e-learning system is being used and frequency of use) on students' perceptions. The last chapter concludes the paper together with a discussion about the implications and limitations of the study. Directions for future research activities are also given.

2 THE VALUE OF AN E-LEARNING SYSTEM

An e-learning system is a system that provides necessary services for handling all aspects of a course through a single, intuitive and consistent web interface. Course content management, communication, the uploading of content, the return of students' work, peer assessment, student administration, the administration of student groups, the collection and organization of students' grades, questionnaires, quizzes, and tracking tools are all examples of some of the tools that are usually provided by an e-learning system. This study was performed using an open-source web based e-learning system – Moodle. Moodle has been successfully deployed and used at our faculty for the last five years. Via a unified web environment implemented through Moodle, students can enroll in courses, download learning materials, communicate with other participants by using forums, write blogs, contribute in content creation using wikis, communicate with professors and teaching assistants through the built in messaging system, finish their activities and upload files, check grades, etc. Professors and teaching assistants use Moodle to manage learning content materials, manage students and their grades, check the uploaded students' work, prepare quizzes, create content using Wikis, aggregate news from different RSS feeds, etc.

When students with different learning habits and technology backgrounds are presented with Moodle, it is important that Moodle be successfully accepted. The adoption process can be influenced by different individual factors like age, gender, computer anxiety, internet experience, etc. Students that currently attend the faculty, are generally well-experienced internet users and usually do not have problems with acceptance and the use of new internet technologies. However, we wanted to know how students of different ages, gender, internet experience, and e-learning experience feel about different e-learning system services. These were the research questions:

RQ1: *From the students' point of view, which are the most important Moodle services?*

RQ2: *How often do students use individual Moodle services?*

RQ3: *Do different factors like gender, age, internet experience, e-learning experience, number of courses in Moodle and frequency of Moodle use influence students' perceptions about the importance and actual use of individual Moodle services?*

3 RESEARCH METODOLOGY

To answer the research questions stated in the previous chapter, quantitative research was conducted in the form of an online survey. Students answered questions about the importance of the inclusion of various Moodle services and about the frequency of actual use of a particular service within Moodle.

3.1 Instrument development

The questionnaire was composed of two parts. In the first part, students gave their opinion on the importance of an individual e-learning system service. Students' perceptions about the importance of 14 features/services were assessed by a five-point scale (very important, important, undecided, not so important, not important), where the respondents gave the extent to which they believed that a certain service is important or not important at all. In the second part of the questionnaire, students had to mark how often they used a certain e-learning system service. For assessing the frequency of actual use of the 15 individual e-learning systems services, a five-point scale (very often, often, rare, very rare, never) was used.

3.2 Sample and data collection

The study participants were comprised of undergraduate students at the Faculty of Electrical Engineering and Computer Science in Maribor, Slovenia. At the Institute of Informatics, a common Moodle system is established, where 115 courses were registered at the time of research. Through the news system in Moodle, a request form with a link to the online survey was dispatched to 609 students, currently enrolled in different undergraduate courses. 284 online surveys were started, of which 235 were finished and 49 returned incomplete. The response rate was therefore 39%.

3.3 Statistical Analysis

To answer the research questions, SPSS software was used for statistical analysis. With SPSS software, the frequency means and standard deviations were estimated. To get an answer to the third research question, the independent-samples t-test and one-way ANOVA tests were performed to test whether several means for different population groups are equal across individual dependent variables. Prior to the statistical tests, adequate normality checks (skewness and kurtosis) for individual independent variables were performed in order to check the normality of the distribution.

4 DATA ANALYSIS AND RESULTS

4.1 Sample characteristics

Table 1: *Sample characteristics*

Demographic characteristics	Frequency	Percentage
<i>Gender</i>		
Male	188	80
Female	47	20
<i>Age</i>		
18 – 20 years	45	19.1
21 – 22 years	124	52.8
23 – 24 years	48	20.4
25 – 26 years	15	6.4
more than 26 years	3	1.3
<i>Years of education</i>		
0 – 1 year	69	29.4
1 – 2 years	64	27.2
2 – 3 years	45	19.1

3 – 4 years	36	15.3
4 – 5 years	15	6.4
more than 5 years	6	2.6
<i>Internet Experience</i>		
Some experience	2	0.9
Experienced	117	49.8
Very experienced	116	49.4
<i>E-learning system experience</i>		
No experience	1	0.4
Some experience	33	14.0
Experienced	164	69.8
Very experienced	37	15.7
<i>Number of courses where Moodle is used (for the present year)</i>		
1	4	1.7
2	24	10.2
3-5	127	54.0
5-8	61	26.0
8-13	16	6.8
13-21	3	1.3
<i>Frequency of Moodle use</i>		
A couple times a year	2	0.9
A couple times a month	4	1.7
Weekly	81	34.5
Daily	148	63.0

4.2 The descriptive statistics

Table 2: *The importance of e-learning services*

Feature/ Service	Very imp. (5)	Imp. (4)	Can't decide (3)	Not so imp. (2)	Not imp. (1)	Mean / Std. Dev.
FA1	190 80.9%	40 17%	4 1.7%	1 0.4%	0	4.78 / 0.480
FA2	169 71.9%	54 23%	9 3.8%	3 1.3%	0	4.66 / 0.480
FA3	74 31.5%	103 43.8%	33 14%	21 8.9%	4 1.7%	3.94 / 0.983
FA4	65 27.7%	100 42.6%	29 12.3%	35 14.9%	6 2.6%	3.78 / 1.087
FA5	116 49.4%	92 39.1%	16 6.8%	10 4.3%	1 0.4%	4.33 / 0.816
FA6	113 48.1%	103 43.8%	10 4.3%	8 3.4%	1 0.4%	4.36 / 0.757
FA7	137 58.3%	74 31.5%	10 4.3%	8 3.4%	6 2.6%	4.40 / 0.916
FA8	112 47.7%	85 36.2%	20 8.5%	16 6.8%	2 0.9%	4.23 / 0.928
FA9	143 60.9%	78 33.2%	10 4.3%	4 1.7%	0	4.53 / 0.662
FA10	49 20.9%	103 43.8%	44 18.7%	35 14.9%	4 1.7%	3.67 / 1.021
FA11	172 73.2%	56 23.8%	6 2.6%	1 0.4%	0	4.70 / 0.537
FA12	127 54%	90 38.3%	10 4.3%	7 3%	1 0.4%	4.43 / 0.749
FA13	144 61.3%	79 33.6%	8 3.4%	4 1.7%	0	4.54 / 0.648
FA14	113 48.1%	86 36.6%	25 10.6%	10 4.3%	1 0.4%	4.28 / 0.850

Table 2 presents the frequencies of answers and mean and standard deviation statistics for the importance of the following Moodle services/features: access to e-learning materials (FA1), electronic posting of homework (FA2),

collaboration with other students (FA3), communication with other students (FA4), communication with professors (FA5), unified Web interface (FA6), Single-Sign-On (FA7), automatic notification about grades (FA8), a complete overview about grades and success for a particular course (FA9), the inclusion of Web 2.0 technologies and services like blogs, wikis, RSS, etc. (FA10), security for private data (FA11), on-line exam subscriptions (FA12), E-calendars (FA13) and on-line quizzes (FA14).

Table 3: *The actual use of e-learning services*

Feature/ Service	Very often (5)	Often (4)	Rare (3)	Very rare (2)	Never (1)	Mean / Std. Dev.
FU1	149 63.4%	70 29.8%	15 6.4%	1 0.4	0	4.56 / 0.633
FU2	74 31.5%	120 51.1%	29 12.3%	9 3.8%	3 1.3%	4.08 / 0.839
FU3	41 17.4%	124 52.8%	51 21.7%	13 5.5%	6 2.6%	3.77 / 0.891
FU4	84 35.7%	127 54%	18 7.7%	4 1.7%	2 0.9%	4.22 / 0.729
FU5	33 14%	56 23.8%	78 33.2%	50 21.3%	18 7.7%	3.15 / 1.141
FU6	4 1.7%	18 7.7%	88 37.4%	91 38.7%	34 14.5%	2.43 / 0.891
FU7	9 3.8	34 14.5%	83 35.3%	77 32.8%	32 14.6%	2.62 / 1.015
FU8	85 36.2%	114 48.5%	31 13.2%	3 1.3%	2 0.9%	4.18 / 0.769
FU9	61 26%	102 43.4%	39 16.6%	20 8.5%	13 5.5%	3.76 / 1.100
FU10	23 9.8%	61 26%	76 32.3%	32 13.6%	43 18.3%	2.95 / 1.234
FU11	2 0.9%	15 6.4%	70 29.8%	111 47.2%	37 15.7%	2.29 / 0.839
FU12	0 0%	17 7.2%	49 19.1%	58 24.7%	115 48.9%	1.85 / 0.975
FU13	0 0%	4 1.7%	18 7.7%	20 8.5%	193 82.1%	1.29 / 0.680
FU14	1 0.4%	12 5.1%	29 12.3%	52 22.1%	141 60%	1.64 / 0.916
FU15	24 10.2%	39 16.6%	37 15.7%	41 17.4%	94 40%	2.40 / 1.411

The second part of the questionnaire was composed of questions about how often students use a particular service in Moodle for learning and completing their studying tasks. In this part, they were asked about the frequency of use for the following services: electronic posting of homework (FU1), search for learning materials (FU2), on-line learning and browsing through e-materials (FU3), downloading of learning content (FU4), an e-calendar for a complete overview of events, dates for submission of homework, etc. (FU5), communication with professors through synchronous/asynchronous communication mechanisms (FU6), communication with other students through synchronous/asynchronous communication mechanisms (FU7), a complete grades overview (FU8), on-line quizzes (FU9), on-line subscription to exams (FU10), personal

profile updating (FU11), collaboration in learning content creation using a wiki (FU12), writing a blog (FU13), reading news through the RSS module (FU14), and searching for definitions in the e-dictionary (FU15).

4.4 Factors influencing students' perceptions

In the case of students' perceptions about the importance of individual services to be provided by an e-learning system, the skewness and kurtosis estimates showed that there were no normal distributions, except in the case of importance for Web 2.0 technologies and services (FA10). For FA10, ANOVA and t-tests (gender, internet experience) were performed. For other independent variables about importance, the Kruskal-Wallis and Mann-Whitney (gender, internet experience) tests were carried out. The Kruskal-Wallis test showed that there was a significant difference in the case of students' communication with professors and in the case of using an e-calendar according to students' ages. Students with different degrees of internet experience can have different opinions about the importance of the possibility to communicate with professors. The Mann-Whitney test showed that for different genders there are significant differences in the means for students' belief about the importance of providing access to learning materials and about providing a complete overview about grades.

According to the results, students with a different level of e-learning experience have varying opinions about the importance of services for communication with other students and online subscriptions to exams. Students' perceptions about the importance of individual e-learning system functionalities were not influenced by the frequency with which students use Moodle or the number of subjects where Moodle was actively used. The ANOVA test results showed no significant differences in students' opinions about the importance of Web 2.0 being enabled by an e-learning system. In the case of the actual use of individual services, the skewness and kurtosis tests found a normal distribution for FU5, FU6, FU7, FU10, FU11 and FU15. For these dependent variables, ANOVA and t-tests (for two population groups) were performed. For others dependent variables (FU1, FU2, FU3, FU4, FU8, FU9, FU12, FU13, FU14) the Kruskal-Wallis and Mann-Whitney (gender, internet experience) tests were conducted.

Differences were found in the frequency of use of on-line subscriptions to exam service for students of different genders and for students enrolled in a different numbers of subjects in Moodle. Age and e-learning experience can influence the students' frequencies of using the e-dictionary. Students with more e-learning experience and students that used Moodle on a daily basis, used Moodle more often to communicate with professors. The same students also used Moodle more frequently to communicate with other students. The statistical tests also revealed that students who used Moodle daily, more often used Moodle to download learning materials and to solve on-line

quizzes. The frequency of use of a complete grades overview tended to be different for students of a different gender, e-learning experience, and for students with different frequencies of Moodle use. The differences in the frequency of use for the electronic posting of homework and search capabilities for learning materials were also found for students of different gender and students that accessed Moodle more frequently.

5 CONCLUSIONS AND FUTURE RESEARCH

The presented study contributes to the existing body of research about e-learning, resulting in empirical evidence about students' beliefs about the importance of different services within an e-learning system. The findings are important for different e-learning stakeholders. As discovered in this research, students are finding Web 2.0 technologies and services important, although they are not using them within an e-learning system.

New technologies and services with social openness may lead to a social revolution, where learners actively participate in learning content creation and the learning process. E-learning system developers should therefore consider a way of integrating an e-learning system with existing Web 2.0 technologies and services (wikis, blogging systems, RSS, Twitter, Youtube, Flickr, Slideshare, Facebook, MySpace, etc.). Future internet technologies and services are going to increase the accessibility of training, teaching, and learning. To succeed, new technologies and services will have to have a positive impact on the capabilities of learners. E-learning system developers must consider how to integrate the e-learning environment with new services and technologies in an appropriate way.

References

- [1] Brown, A. 2006. Learning from a distance. *Journal of Property Management*, 71, (4) 49-57
- [2] Chiu, C.M., Sun, S.Y., Sun, P.C., & Ju, T.L. 2007. An empirical analysis of the antecedents of web-based learning continuance. *Computers & Education*, 49, (4) 1224-1245
- [3] Lee, M.G. 2001. Profiling students' adaptation styles in Web-based learning. *Computers & Education*, 36, (2) 121-132
- [4] Pituch, K.A. & Lee, Y.K. 2006. The influence of system characteristics on e-learning use. *Computers & Education*, 47, (2) 222-244

FUNCTIONAL HORIZONTAL NETWORK MARKETPLACES – A POSSIBLE SOLUTION FOR SERBIAN MARKET

Zoran Janković, Mirjana Ivanović, Zoran Budimac

Department of Mathematics and Informatics

Faculty of Sciences, Trg D. Obradovića 4, 21000 Novi Sad, Serbia

E-mail: bpzoran@yahoo.com, mira@dmi.uns.ac.rs, zjb@dmi.uns.ac.rs

ABSTRACT

During an evolution of B2B Network Marketplaces (NMPs), two main types have emerged: vertical NMPs and horizontal NMPs. Whereas vertical marketplaces cover particular markets and industries, horizontal marketplaces include a huge number of different enterprises from dislocated markets and various industries. Horizontal NMPs are more affordable for most of small-sized and medium-sized enterprises, but they mainly have much less functionalities to offer. In this paper, the possibilities of improving functionalities of horizontal marketplaces are discussed, with an example of a B2B portal that comprises suggested improvements.

1. INTRODUCTION

Business-to-Business (B2B) describes computer-aided commerce between enterprises, supported by the Internet and other computer and network technologies. This term does not include digital transactions inside enterprises' boundaries, for example the transactions between two departments of a single enterprise.

Depending on a number of participants, organizational structure and relations between participants, there are two main types of B2B systems: Network marketplaces and Private industrial networks [1]. Private industrial networks gather a limited number of strategic business partners that use a B2B system in order to improve their supply chain. These systems may be organized by some single large enterprise that involves its business partners, or it may be in possession of several companies that have same privileges in system use.

Net marketplaces bring together a large number of suppliers and buyers, who can share information, trade and cooperate in many ways. This type of a B2B system usually means free entrance for all enterprises that belong to market which is supported by this system.

2. EXISTING SOLUTIONS

There is no easy way to classify Net marketplaces. One criterion for classifying is a modality of cooperation between enterprises. There is spot purchasing and long-term cooperation. Spot purchasing means that suppliers and buyers create individual transactions based on current needs and possibilities. Long-term cooperation is just opposite – participants cooperate over many different transactions that

may take months or even years. Another criterion for classifying NMP is alternation of priority in communication. In this respect, there are two main types of Net Marketplaces: Transaction-oriented marketplaces and collaboration-oriented marketplaces [2]. While transaction-oriented marketplaces are characterized by catalogs, auctions or exchanges and support for negotiated pricing, collaboration marketplaces are characterized by planning capabilities such as continuous planning, forecasting, and replenishment or product life-cycle management.

Collaboration marketplaces have different benefits than transaction-oriented marketplaces and different adoption considerations.

There are many similar portals, which all have certain common features. They are all wide spread, with a huge number of members covering many different industries. This type of B2B systems is called "Horizontal marketplaces". On the other side, vertical B2B systems are anticipated as smaller systems for individual markets or industries.

Horizontal marketplaces are primarily profiled for spot purchasing, without intention for long-term cooperation. Likewise, as product catalogs are a main purpose of horizontal marketplaces, they are mostly defined as transactional-oriented. Namely, considering the development of horizontal marketplaces, main problem is how to adjust its functionality for many different enterprises, different markets or regions. The solution can be obtained by reducing the functionality to base transactional processes: catalogs, messaging and order management. Enterprises create catalogs of their goods and services, demands/offers and send messages to each other, which can be in different forms: an enquiry for products, an answer to offer/demand, a query about delivery status etc. These forms of cooperation avoid all individual business processes and make a B2B system acceptable for enterprises in most of industries, markets and regions. However, still there is a problem of internationalization and localization. The localization of all business terms to any specific language is a less significant problem in comparison with the adoption of individual information of portal members in another regions and markets. Certain attributes of enterprises have different importance from country to country. That is the reason why many of B2B systems are localized for many specific countries or regions.

Another problem of horizontal marketplaces is how to group products, services, offers and demands. In most cases, they are grouped into hierarchical groups of goods/services. Users have to click on links that represent groups until the

last element is chosen. There are two main problems in this approach: 1. A user has to click as many times as many sub-groups exist in order to find desired elements. Every click mainly means a new HTTP request, so it spends additional time for navigation. 2. Users have an overview of a current group structure, but they lose an indication about the groups which are not chosen, so if they want to make a different choice, they have to move back to previously chosen groups or to a root element.

One of the most usual ways for paying membership on NMPs is a transaction fee system. Members have to pay a fee for every committed transaction.

As markets become more competitive, transaction fees will most likely decrease or even disappear [3]. Apart from other implications, this means that providers will need to shift from dealing with transactions to offering more comprehensive solutions to business needs. For example, products can be bundled with related information and services in order to forge customer loyalty and long-lasting relationships. Furthermore, providers need to be able to offer separate products, in a form of traditional catalogs, as well as group offers of products, which may include some benefits for buyers, some special offers and so on. Hence, suppliers need more flexibility in their business. In addition, suppliers need a possibility to search for buyers in the easiest possible way, and vice versa.

3. RELATED WORK

“Alibaba” [9] is the name of a typical transactional-oriented B2B portal based on spot purchasing. This portal was founded in China and has been spread all around the world. Enterprises from the entire world share information about products and services using this portal. Any company can join and make requests about other enterprises’ offers and demands for products and services. Every member can be verified by a third party in order to provide trusted identification to other portal members. Members may have various privileges on this site depending on a purchased membership version.

Another large B2B portal is “Kompass” [10]. This portal was founded in the USA and has also been spread all around the world. Membership is also free, but with a limited functionality. For a more serious use (a contact with another enterprises, gathering detailed information etc.) a membership fee is required. This system is also intended for spot purchasing and individual messaging.

Nowadays, the most commonly used way to develop a B2B portal is by using Content Management Systems (CMS) specialized for B2B portals [11, 12]. These products are intended for easy development of horizontal B2B systems that contain described functionalities – catalogs of goods and services sorted into offers and demands, and, on the other hand, messaging between portal members. This way of deployment is easy, fast and productive. A B2B system like this can be localized to many languages and adjusted for particular needs.

The main disadvantage of such systems is the fact that communication between members and the portal itself is

much more developed than communication between members. Another disadvantage is a reduced possibility of adding new functionalities and changing and improving existing ones.

The most famous B2B CMS systems are B2B “Portal Script” [12] and “Alibaba Clone” [11]. They have similar properties and make it possible for everyone to deploy a functional B2B system.

An alternative way to deploy a B2B system is by deploying a completely new project from the beginning, respecting all the standards of application deployment. [8]. Such applications are much more flexible and extensible for new functionalities. In addition, these solutions can be oriented to communication between users, rather than to communication between the portal and users. The most important limitation of this approach is a longer development period, a need for standardization of certain business processes and installation of these standards into an application.

Solutions in Serbia are mainly deployed by using some of B2B CMS systems and partially localized to the Serbian language. Disadvantages of these portals are related to described weaknesses of B2B systems generally deployed in this way. An example of widely used B2B portals is “Serbian Industry” [14]. This portal has been deployed using one of B2B Content Management Systems. It offers possibilities for presenting enterprises’ information, publishing product information, offers, demands and messaging between enterprises. However, this portal does not include advanced possibilities related to collaboration between enterprises. For example, it does not include automated negotiate system. Moreover, navigation for choosing product categories is based on simple navigation as it has been already described. Functionality of this portal is mainly limited to possibilities embedded in CMS used for deployment.

Other B2B portals are much less used on Serbian market. However, some companies use other types of B2B systems, like private industrial networks. These companies are generally well situated and they use B2B system to improve communication with existing business partners, without a possibility for making new ones. However, small and medium-sized enterprises cannot afford such systems. For that reason, there is a need for a system, which can serve all enterprises on Serbian market, providing a wide range of functionalities.

4. PROPOSED SOLUTION – MARKET4NET B2B PORTAL

Hereafter, our research is about to present how all listed disadvantages can be overcome by making an independent project of a horizontal network marketplace intended for Serbian market, concerning all the described limitations. It means choosing suitable attributes for every object in application, and pick appropriate terms for making business and harmonization with internal law and business rules.

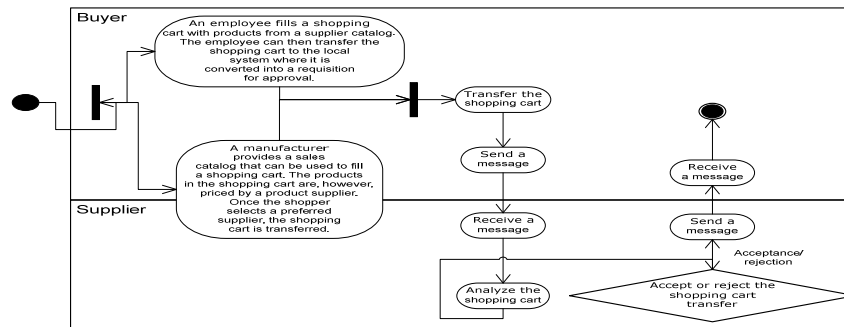


Figure 1: RosettaNet process "Request Shopping Cart Transfer"

Market4Net (www.market4net.com) [13] is an independently deployed Net Marketplace solution intended for Serbian and regional markets. This solution tends to overcome difficulties of deploying a horizontal marketplace with maximal functionality. This application has been deployed using following methodologies and standards:

- an object-oriented methodology using the UML diagrams for system description
- RossetaNet standards [4] for defining global business processes that rely on previously defined goals. RossetaNet represents a business protocol and a set of standards for fundamentals of cooperation between business partners over the Internet. In our research, standards were adapted to an object-oriented methodology and the UML language
- The J2EE [15] platform for technical implementation, using Spring Framework [5], Hibernate ORM [6] and JBoss application server [7].

Usage of RossetaNet standards [4], its combination with object-oriented methodology with the UML language and its adaptation with local business rules, is presented through a description of business process called "Request Shopping Cart Transfer" (Figure 1).

This process is presented as an activity diagram, using the process definition from RosettaNet documentation [4].

This process implies an existence of the object called "Shopping Cart". Nonetheless, this term is acceptable on markets with widely used e-payment systems, which is not the case in Serbia. Many small-sized and medium-sized enterprises in Serbia do not have that possibility. For this reason, terms in this diagram have to be adapted to the Serbian market. The term "Shopping Cart" is about to be translated into "Inquiry" – a term that is not tied to payment. In that case, a buyer sends only a query about some group of products. This is in accordance with one of main purposes of "Market4Net" business portal – collaboration, rather than transactions.

Some of main functionalities of this portal are:

- **Set up structured and non-structured information for every member – enterprise.** Structured information is adapted to Serbian and regional terminology and includes general and contact information. Non-structured information includes a self-made description of every enterprise, without text size limits and with possibilities of formatting text, unlimited number of pictures and

unlimited number of news and actualities for every enterprise.

- **Make on-line partnerships between enterprises.** This is one of the main advantages comparing with portals based on B2B CMS systems. Every enterprise chooses its enterprises for cooperation. In that way, they are protected from unwanted messages ("spam") from other portal members. Any two enterprises can make partnership, with a common agreement. On the other hand, all members may decide not to limit their cooperation to partners only. Therefore, the entire cooperation over this portal is based on trust.
- **Publish product catalogs and make queries about products among enterprises.** Enterprises can show inquiries for products or services of their partners, and, depending on information attached to the query, those inquiries can be treated as orders. This information can include a simple message for the closer explanation of a query, attachment, and delivery type. An answer to the query (Figure 2) can include all invoice elements: costs of every product, delivery costs, other costs, total costs, additional information and attachments, which may include a formal invoice. The terminology for all of these elements is adapted, so it represents a communication and information exchange, without implicit transactions. Enterprises have freedom to interpret this information as they decide, depending on relationship between other collaborative enterprises.
- **Exchange information about offers and demands.** Every enterprise can publish an unlimited number of demands and offers and answer to demands and offers of other enterprises. A portal member receives answers to its offers and demands and chooses the most appropriate ones. In addition, this functionality has installed a partnership element – an enterprise may choose to receive answers to offers or demands from partners only.
- **Simple messaging between enterprises.** This way of communication may include a range of possibilities that are not included in other functionalities – for example query about enterprises' news and actualities, transaction canceling, notification of invoice rejection etc.

Obviously, this portal does not mean simple transactions between enterprises. Thus, it does not include transaction fees and there is no limitation for cooperation between enterprises.

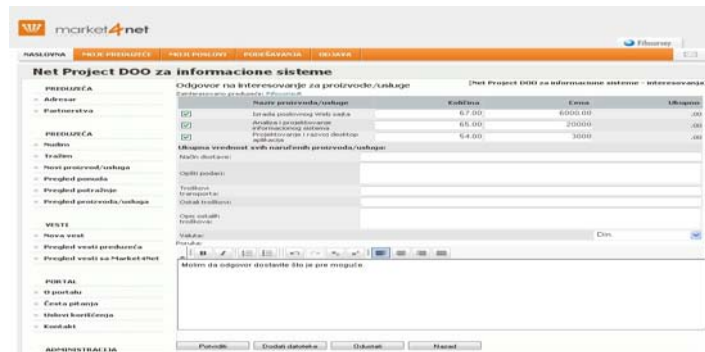


Figure 2: Market4Net B2B Portal appearance –an answer to the inquiry query

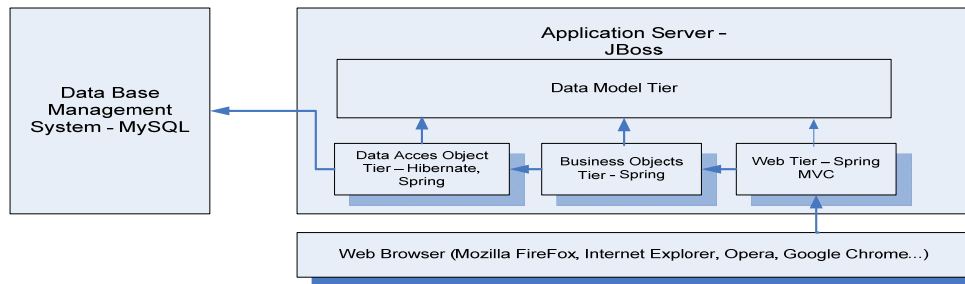


Figure 3: Modular structure of the application

Every product, service, offer and demand is related to a certain group of goods/services. Groups are arranged in a tree structure, so users can navigate it more easily. Choosing groups does not require a new request, which makes it faster and more intuitive. Users have an overview of the tree structure in every moment, so they can easily return to any other group at any moment. This feature solves problems of navigation when searching for adequate goods/services, offers and demands.

One of the most important goals while deploying this application is the flexibility for improving existing functionalities and extendibility for adding new ones. In that sense, non CMS or similar system, which would reduce the flexibility, is used for deployment.

Moreover, entire, application is built modularly, using J2EE components, with possibility to change easily almost any part of application (Figure 3).

This way developed application can change and improve its functionality, as well as particular technologies in application tiers. For example, for data access, Hibernate Object Relational Mapping (ORM) framework is used. However, the implementation of this tier is completely independent from other tiers, so it can be changed without a need for changing other tiers.

5. CONCLUSION

There is not any particular solution to address all issues about the deployment and the use of horizontal marketplaces. This paper covers some of usual aspects of the problem, focusing on some specific problems, like transactions vs. cooperation, product grouping, localization etc. Proposed solutions to these problems are included into Market4Net B2B portal, which is a functional and ready-to-use application for the Serbian market [8, 13]. Nevertheless, described problems and solutions can be a base for further

improvements of existing horizontal network marketplaces and creating new ones, based on market demands.

REFERENCES

- [1] Kenneth C. Laudon, Carol Guercio Traver. *E-Commerce: Business, technology, society*. Addison Wesley. 2002.
- [2] M. Lynne Markus, and others. Adoption and Impacts of B2B Marketplaces: Transaction Versus Collaboration Marketplaces. *European Conference on Information Systems (ECIS)*. 2003.
- [3] Ruth A. Wienclaw. *B2B Business Models*. EBSCO Publishing Inc. 2008.
- [4] © RossetaNet. RosettaNet Program Office Overview: Clusters, segments, and pips, Version 02.07.00. 2009.
- [5] Craig Walls, Ryan Breidenbach. *Spring in Action - Second Edition*. Manning Publication. 2008.
- [6] Hibernate Reference Documentation. 2008. (<http://www.hibernate.org>)
- [7] JBoss Application Server 4.2 Getting Started Guide. *JBoss Community Documentation Project*. 2006.
- [8] Zoran Janković. Development and Implementation of B2B E-commerce systems (in Serbian, MSc thesis - draft version). *Faculty of sciences*. Novi Sad. 2009.
- [9] Alibaba B2B Marketplace - <http://www.alibaba.com>
- [10] Kompas Business Search - <http://www.kompass.com>
- [11] Alibaba Clone - <http://www.alibabacclone.com>
- [12] B2B Portal Script – http://www.i-netsolution.com/b2b_portal_script.html
- [13] Market4Net B2B Portal – www.market4net.com
- [14] Serbian Industry (<http://www.serbianindustry.com>)
- [15] Jennifer Ball, Debbie Carson, Ian Evans, Scott Fordin, Kim Haase, Eric Jendrock. *The Java™ EE 5 Tutorial*. Sun Microsystems. 2006.

USING GEOFENCING TO OVERCOME SECURITY CHALLENGES IN WIRELESS NETWORKS: PROOF OF CONCEPT

*Anthony .C. Ijeh¹, Allan .J. Brimicombe², David S. Preston¹, Chris .O. Imafidon¹,
Annette O. Uwaechie³*

School of Computing and Technology, University of East London, UK

² Centre for Geo Information Studies, University of East London, UK

³ Zenith Bank PLC, Nigeria

Tel: +44 2082237778; Fax: +44 2082232963

E-mail: a.ijeh@uel.ac.uk

ABSTRACT

Using a Proof of Concept approach, this paper examines Geofencing impact on Wireless Network Security. Results indicate that Geofencing technology can function effectively in a small-to-medium hardware environment. Also the majority of the simulations recorded reasonable read rates even though the simulations were set up over a short period of time without a great deal of fine tuning. Moreover, Geofencing could have a positive impact on the wireless network security challenges within organisations by being used as an access control mechanism in conjunction with authentication based on a user location. Despite these encouraging findings in relation to using Geofencing for wireless network security purposes within small-to-medium sized organisations; it is recommended that the Prototype for which £13,852 of funding has been secured through the London Mayors office and the University of East London, Project Reference No: 2M-046, be tested before committing to the commercial implementation of Geofencing as Security Strategy

KEYWORDS: Geofencing technology, wireless networks, proof of concept, simulation

1 INTRODUCTION

Geofencing is the application of location based services and the delivery of data and information services where the content of those services is tailored to the current or some projected location and context of a mobile user [1]. Recent research has shown that the wireless network protocols are vulnerable even with the creation of wireless protocol access (WPA2) and suggested the use of message integrity

checks [2]. Against this backdrop' research for a solution to wireless protocols is greater than ever and some research has made great strides into finding a permanent solution. Place lab is one of such research organisations. "A key observation here is that many developed areas have wireless hotspot coverage so dense that cells overlap. By consulting the Place Lab directories, which will map wireless hotspot MAC addresses to physical locations, mobile computers and PDAs equipped with Wi-Fi can determine their location to within a city block" [3]

2 RESEARCH PURPOSE AND APPROACH

Currently the security of all Information systems is based on the Open Systems Interconnection Reference Model (OSI Reference Model or OSI Model) it was developed as part of the Open Systems Interconnection (OSI) initiative. In its most basic form, it divides network architecture into seven layers which, from top to bottom, are the Application, Presentation, Session, Transport, Network, Data-Link, and Physical Layers. The highest security that can be offered as at today is based on the layer 2 OSI Model which is the data link layer where security is provided through routers and protocol. The security solution will use specially programmed technology to locate a wireless device. The objective is for the wireless device to only function within a defined parameter. This is so that the parameter can be used to control the acts of the wireless device when it communicates with a designated database. **Figure 1** shows an RFID Tag (blue icon) being used to monitor a wireless Laptop Red (red icon). **Figure 2** is the test bed and walked line measurement which had been predefined prior to the exercise

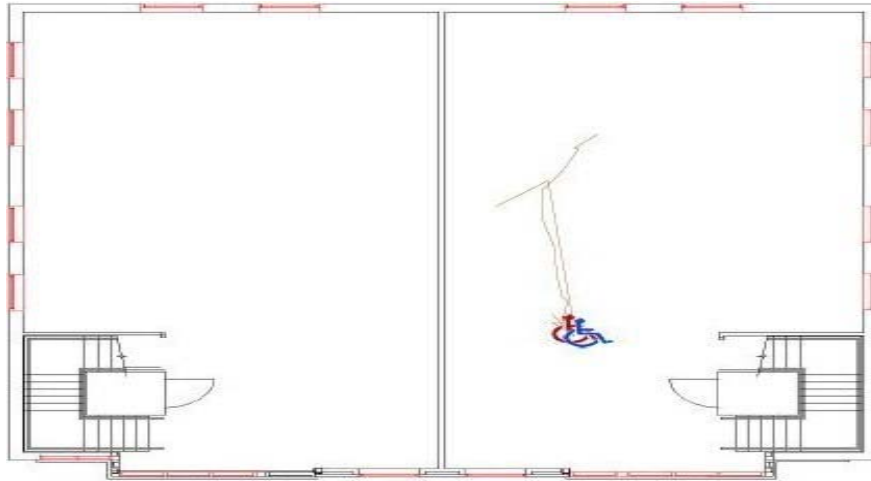


Figure 1: Activity within the defined parameters of a test bed

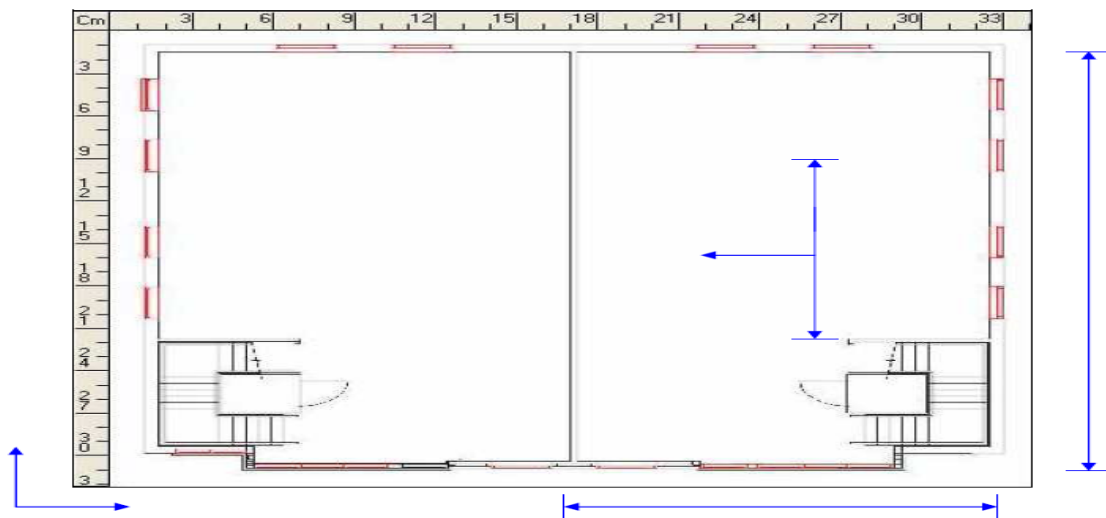


Figure 2: Pre-measured test bed and marked route

3 RELATED WORKS

With the technology of Airetrak Wi-Fi Tracking solutions at its disposal this study proposes that on pinpointing the location of the user using Airetrak technology to the most accurate point possible; varying levels of user access should then be used as a security measure depending on the location of the individual. So say for example if an individual works in the finance department, and they try to log into the network from the HR department then their access to the network will be blocked because they are not logging in from where they are meant to be gaining access to the system [4]

However some researchers suggested there are issues with using the Open System Interconnection (OSI) model in developing Information Security Trust Models. They claim that “The presentation of the layer-by-layer issues has

focused on elements within each layer. However, the purpose of the seven-layer model is communication, and communication between layers is where many key issues lie” [5]. In support of this some researchers have backed up this claim stating that “The seven-layer model is more than adequate for network purposes, but when used in the context of developing information security Trust Models there are organisational information security Trust Model concepts that sit outside of the conventional OSI Layer network model” [6]. In addition to these other researchers have suggested that the ways in which data is collected has been the major impact on Information Security Trust Models, as Small Medium Enterprises do not have the resources of their larger counterparts and as such cannot support the OSI Layer Models guidelines [7]

This paper supports these separate arguments and will suggest an Information Security Trust model that

accommodates all three arguments to be used by both Small Medium and Large Enterprises in adherence to the Open System Interconnection (OSI) 7 Layer Model

4 METHOD

The SSM will be created and tested using a wireless location appliance (WLA), which will enable the researcher to monitor which wireless devices access the network within the geographical test bed. A defined geographical location has been chosen because the access points (AP) within the test bed will enable a more effective monitoring of the wireless devices used within it. Once a user steps out of the geographical test bed, their access will automatically be cut off and once they step back into the test bed area they will have to log on again. Thus by using location-based technology which operates within a defined area, the SSM will be able to prevent unauthorised users from gaining access to a wireless network through sniffing or war driving; [8, 9]. The results will then be analysed using statistical software (MATLAB) and (SPSS) in order to identify the relationships between the dependent variables e.g. test bed, security strategy model and the independent variables e.g. wireless apparatus, and also to determine the precise geographic location of the wireless laptop when it performs any or all of the following the acts: If the wireless apparatus was logged onto within the geo-fenced area; If the wireless apparatus was logged onto outside the geo-fenced area; If

the wireless apparatus was able to stay logged on outside the geo-fenced area, if taken from within to the outside of the geo-fenced area; If the wireless apparatus can be logged onto on the boundary of the geo-fenced and non geo-fenced area. This study will not be looking at network security or positioning technology or trying to improve either of them.

5 ELEMENTS OF THE INVESTIGATION THAT ARE NOVEL

In using Airetrak Wi-Fi Tracking Solutions technology, which can pinpoint a user's location to the nearest possible inch, the author believes by varying access levels of security depending on the user's pinpointed location the study has uncovered a new area of wireless security and possibly a new protocol. By using a holistic approach to understanding the development and management of protocols for wireless security and privacy locations, the study ascertained how the location of key data transmitted over the wireless network can be restricted to defined areas in order to enhance security. **Figure 3** shows RFID tags all located within the predefined test bed, because of their location access can be granted to the devices which they are attached to. Also located are the icons used in our experiment which are located within the test bed (blue and red) and to which based on their location access was given to our wireless laptop

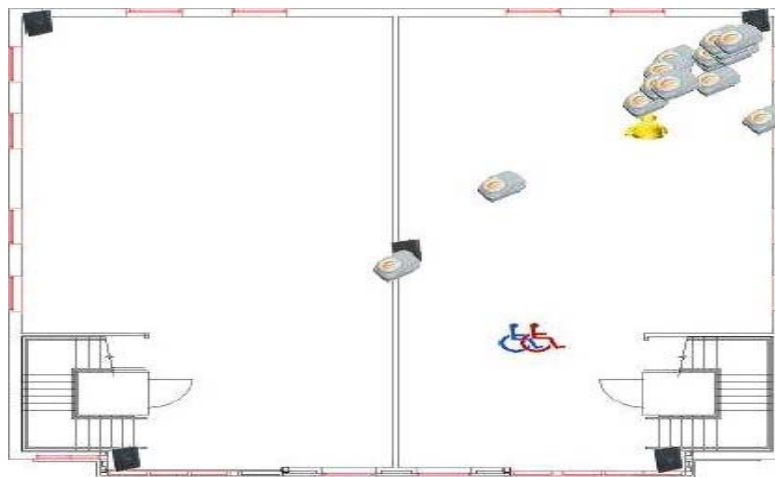


Figure 3: Using Geofencing to grant access within a pre-defined area

7 DISCUSSION AND CONCLUSION

This experimental study was based in a laboratory and adopted a lot of work from previous researchers in order to focus on its key objective “Using Geofencing to Overcome Security Challenges in Wireless Networks: Proof of Concept”. The primary contribution of the research was the

design of the Security Strategy Model and the development of a supporting Application Framework (Security Solution) **Figure 4** shows the Geofencing Trust Model for wireless Security

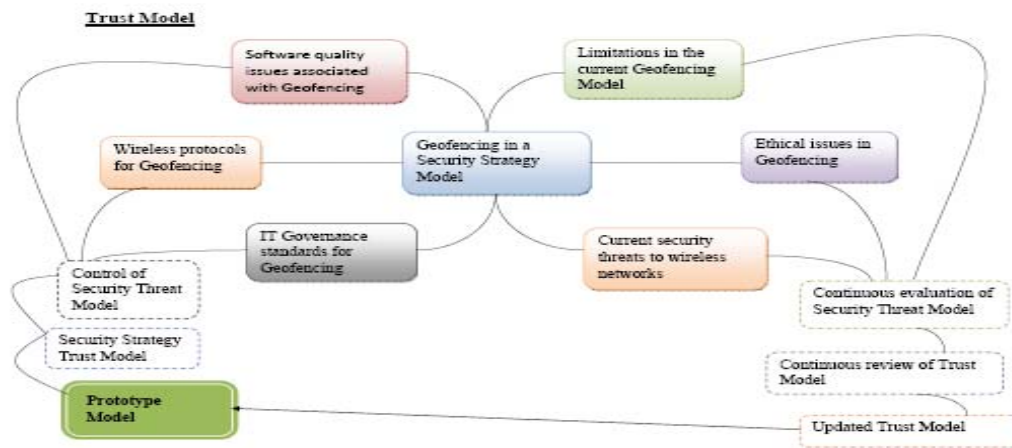


Figure 4: Geofencing Trust Model for wireless Security

REFERENCES

- [1] Brimicombe, A.J and Li, C (2009) Location Based Services and Geo-Information Engineering, Wiley-Blackwell (2009) Page 1 – 20
- [2] Hori, Y. and Sakurai, K (2006) ‘Security Analysis of MIS Protocol on Wireless LAN comparison with IEEE802.11i’ in *wireless mobile networks Mobility 06*, Oct. 25–27, (2006), Bangkok, Thailand; Page1-5
- [3] Hong I Jason (2003) Privacy and Security in the Location-enhanced World Wide Web http://www.cs.berkeley.edu/~tygar/papers/Privacy_security_location-enhanced_web.pdf (Accessed on: 09/01/08)
- [4] Hightower, J. et al (2006) Practical Lessons from Place Lab IEEE Pervasive Computing, vol. 5, no. 3, 2006. <http://www.placelab.org/publications/>
- [5] Reed, D. (2004) Applying the OSI Seven Layer Network Model to Information Security IN “SANS GIAC GSEC Practical Assignment” version 1.4b Option One. Assessed from: http://www.sans.org/reading_room/whitepapers/protocols/applying_the_osi_seven_layer_network_model_to_information_security_1309?show=1309.php&cat=protocols (Assessed on: 23-3-09)
- [6] Surman, G. (2002) Understanding security using the OSI model IN “SANS GIAC GSEC Practical Assignment” version 1.3. Assessed from: http://www.sans.org/reading_room/whitepapers/protocols/understanding_security_using_the_osi_model_377?show=377.php&cat=protocols Assessed on: 23-3-09
- [7] Beachboard, J. (2008) Improving Information Security Risk Analysis Practices for Small - and Medium-Sized Enterprises: A Research Agenda IN *Issues in Informing Science and Information Technology* Volume 5, Issue 5; Pg 73 – 85. Assessed from: <http://proceedings.informingscience.org/InSITE2008/ISITv5p073-085Beach415.pdf> Assessed on: 23-3-09
- [8] Ijeh, A.C. et al(2009) “Geofencing in a Security Strategy Model” In the Proceedings of the International Global Security, Safety, and Sustainability 5th International Conference, ICGS3 2009, London, UK, September 1-2, 2009. Pages 104 – 111
- [9] Ijeh, A.C. et al (2009) “The Significance of Security in Transmitting Clinical Data” In the Proceedings of the International Global Security, Safety, and Sustainability 5th International Conference, ICGS3 2009, London, UK, September 1-2, 2009. Pages 122 - 131

FREE INTERNET ACCESS USING THE ANDROID PLATFORM

Jernej Huber

University of Maribor, Faculty of Electrical Engineering and Computer Science

Smetanova 17, 2000 Maribor, Slovenia

tel: +386 31 355 918

e-mail: jernej.huber@uni-mb.si

ABSTRACT

This paper presents a vision of ubiquitous computing and a brief history of the mobile phone application programming. In addition it shortly describes the Android Platform and presents an implemented solution for gaining a free internet access using the Android platform and Web 2.0 technologies.

The implemented solution includes a REST web service for the Wi-Fi hotspot database management, the hotspot submission web site and the application for the Android Platform, which shows us a list of nearby Wi-Fi hotspots, as well as other locations like hotels, restaurants and museums, and gives us a navigation details of the chosen location, when walking towards it with the Android-enabled Smartphone in our hands.

1 INTRODUCTION

We are facing the true ubiquity of the computing in all the aspects of our lives. Because the ubiquity of computing includes a wide range of aspects, we present only a few.

We also present the evolution of mobile platforms, which nowadays for the first time allow everyday developers to contribute their knowledge of building innovative application. These applications are starting to change the tempo of our lives.

The announcement of the Android Platform has changed a many trends in the mobile operating systems market. All of the important mobile platforms are now opening their development tools and programming stacks and starting to simplify the application development to give everyday high level programmers a chance to contribute their share in the constantly growing mobile application market.

Because of the high prices of the international data transfers on the mobile network we provided a solution for locating and finding the nearest free Wi-Fi hotspots through which we can access the internet on our Android-enabled mobile phone for free.

2 UBIQUITY OF COMPUTING

The ubiquity of computing contains a broad range of technologies. We present the tree aspects – the vanishing computers, an ambient intelligence and the internet of things.

2.1 Vanishing computers

The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it [1], with other words – the vanishing computers.

2.2 Ambient intelligence

Today we are constantly plied with new, noisy and disruptive electronic devices. Consequently, we are spending too much time in learning all the new technologies instead of dealing with the vital problems. For this reason, the experts in the development departments of large enterprises, such as Philips, started thinking about so-called ambient intelligence, a shift in thinking, which would allow the development of more human-oriented technologies.

The idea of ambient intelligence is in the moving of the most information and communication systems to the background while the users move to the forefront of the technology-enhanced environment management. Ambient intelligence is generalized to the wider residential environment, i.e. the work, service, manufacturing and public spaces and the entire transport system, including the means of transport. It requires the entirely new ways of communication between people and the ubiquitous computing systems, which go beyond today's relatively inefficient methods of using the desktop and laptop computers, and mobile phones. It is not the forecast of the future, as it should be a serious vision, offered by advisory groups within the international research programs [2].

2.3 Internet of things

The research, taken out by the European Commission forecasts that in the future every object will have a unique way of identification – the unique address, which will

constitute a addressable continuum of computers, sensors, actuators and mobile phones – almost all the things or objects that surround us. Once these things have the ability of mutual addressing and verification of their identities, they can exchange information among themselves and, if necessary, actively process information against pre-defined deterministic and non-deterministic schemas [3].

EU Research predict that the internet of things will become a reality in the next 20 years – with omnipresent smart devices that will wirelessly communicate over hybrid and ad-hoc networks of devices, sensors and actuators, working in synergy to improve the quality of our lives and consistently reducing the ecological impact of mankind on the planet [3]. The EU Action plan on the future of Internet things has stressed that internet of things is not yet a tangible reality, but rather a prospective vision of a number of technologies that, combined together, could in the coming 15 years drastically modify the way our societies function [4].

3 UBIQUITY OF MOBILE PLATFORMS

Technologies, which are surrounding the vision of ubiquitous and pervasive computing, for example haptic computing (interaction with users on the basis of their senses), sentient computing, context-aware pervasive systems, augmented reality, localization using GPS, identification using bar codes and RFID tags etc., are now available in the almost every contemporary mobile platform (including the Android platform), which enable the construction of advanced and innovative applications for the smartphones. What seemed like a science fiction around a decade ago is now literally available in our pocket.

3.1 Evolution of mobile platforms

Cell phones have same roots with embedded device programming. Most embedded devices ran (and in some cases still run) proprietary operating systems. Reason for choosing proprietary systems was because simple devices did not need very robust and optimized operating systems. The next step as product of device evolution for many of the more complex embedded devices was to move to more standardized operating system platforms. This shift

happened about six years ago, as small-footprint operating systems such as Linux, or even an embedded version of Microsoft Windows, have become more prevalent on many embedded devices.

Another reason for using proprietary systems was the fact that cell phone manufacturers typically used hardware that was completely developed in-house, or at least hardware that was specifically developed for the purposes of running cell phone equipment.

The downside to proprietary systems is that anyone who wanted to develop applications for cell phones needed to have inside knowledge of the proprietary environment. The solution was to purchase expensive development tools directly from the manufacturer. This isolated many of the potential developers.

Another thing that kept cell phone development out of the hands of the everyday developer was a lack of device memory and storage capacity. Not until few years ago the price of memory has reached very low levels that allow manufacturers to put on the cell phones enough memory for running many applications at the same time.

In the following years device manufacturers still closely guarded the operating systems that run on their devices. While a few had opened up to the point where they have allowed some Java-based applications to run within a small environment on the phone, many did not allow this.

Even the systems that did allow some Java apps to run did not allow the kind of access to the underlying system that standard desktop developers were accustomed to having.

This barrier to application development began to crumble in November of 2007 when Google, under the Open Handset Alliance, released Android and forced the rest of the mobile platform vendors to open up their technologies and simplify the application development [5].

3.2 Mobile operating system market

Smartphone sales were strong during the second quarter of 2009, with sales of 40.9 million units in line with Gartner's forecast of 27 per cent year-on-year sales growth for 2009 [6]. Figure 1 shows the market share of the most important mobile platforms.

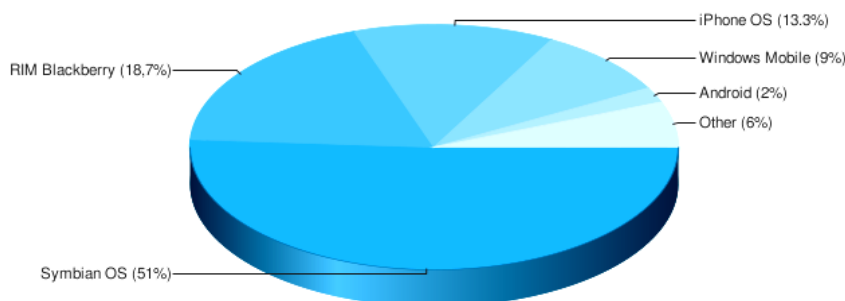


Figure 1: Market share of the most popular mobile platforms.

Given the higher margins (the selling prices minus the cost of production), the smartphones offer the biggest opportunity for manufacturers. It is the fastest-growing market segment and the most resistant to declining average selling prices [6].

4 THE ANDROID PLATFORM

Because of its wide range of capabilities, we can easily confuse Android with a desktop operating system.

4.1 Features

Figure 2 shows a simplified architecture of the Android platform [7].

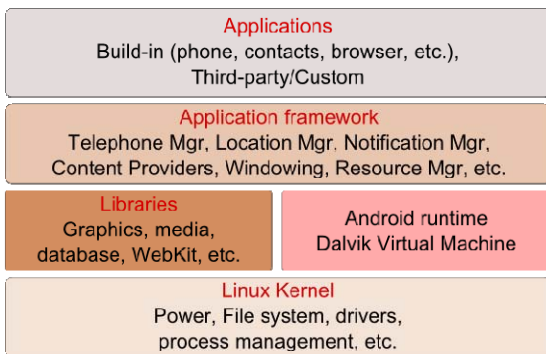


Figure 2: Simplified architecture of the Android platform.

Android is a layered environment built upon a foundation of the Linux kernel. The UI subsystem includes the Layouts of Viewgroups or Views, and the Widgets, which are a View objects that serve as an interface for interaction with the user, for example buttons, checkboxes, and text-entry fields [7].

Connectivity options include WiFi, Bluetooth and wireless data over a cellular connection (GPRS, EDGE, and 3G). The MapView UI element allows us to link Google Maps to display an address directly within the applications. Support for location-based services such as GPS and accelerometers is also available, though Android devices are obliged to be equipped with the mentioned hardware. There is also vast video/photo/audio support. The data-storage challenge is eased with the use of open source SQLite database [8].

Android includes an embeddable browser built upon WebKit, the open source browser engine powering the many of the popular web browsers (Safari, Chrome).

4.2 Application architecture

Android applications are written in the Java programming language (the Apache harmony open source implementation of Java) and they run within a virtual machine (VM), which is not a JVM, but the Dalvik Virtual Machine, an open source technology, optimized for mobile devices. Each Android application runs within an instance of the Dalvik VM, which resides within a Linux-kernel managed process. An Android application consists of one or more of the following concepts [9]:

- **Activities** are presentation layer of applications (similar to forms in desktop computing).
- **Services** run in background and are used when applications need to be persistent for a long time (for example, update checker).
- **Content providers** are the means of exposing the data from various sources to various applications. We can create our own content providers or use the included in the Android platform, like the content provider for accessing the data from our contacts list.
- **Broadcast receivers** are used to build event driven applications, because they can respond to a various events, such as the receipt of a text message.

The necessary configuration and deployment information are stored in manifest file called AndroidManifest.xml. The manifest includes the required component info and types of events the application is able to process, and the required permissions the application needs to run (most common example is an access to the network). Such a declarative security helps to maintain security, because the users see which permissions are required when installing applications [9].

5 IMPLEMENTED SOLUTION

Because of the high cost of a data transfer when roaming, we came to an idea to try to minimize the expenses. Therefore we developed a solution for gaining a free internet access using the Android platform and Web 2.0 technologies, as seen in Figure 3.

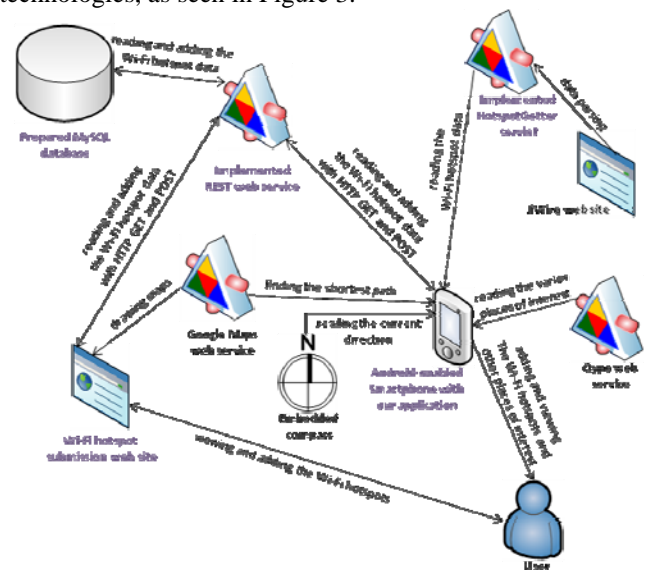


Figure 3: Schema of the implemented solution.

Figure 4 shows a screenshot of implemented application - navigation to the chosen destination - application gives us a distance and direction of the chosen destination when walking towards it with the Android-enabled Smartphone in our hands.



Figure 4: Navigating to the chosen destination with our application.

Figure 5 shows the comparison of the estimated daily expenses in € when using our application and direct web access to the JiWire website when roaming in EU countries on Slovenian mobile operators.

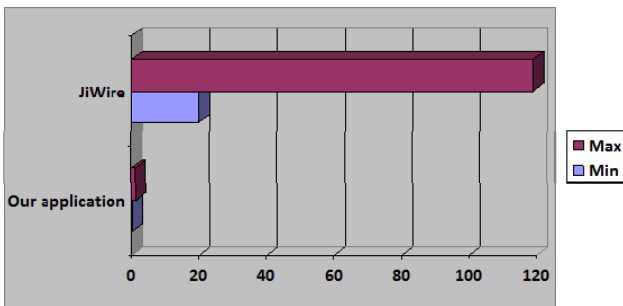


Figure 5: Comparison of the estimated daily expenses in € when using our application and direct web access

We used the following services and protocols (among others):

- **JiWire** is a worldwide free hotspot finder with listings for over 270000 hotspots in 140 countries [10]. We developed a servlet which parses the Wi-Fi hotspot data and represents it as a XML document, to minimize transferred data and consequently, the cost of the data transfer.
- **Qype** is the social networking web site, which allows user to create local reviews of hotels, restaurants, museums, clubs and similar locations with the monthly total of about 6.3 million unique visitors to Europe [11]. We used Qype API to get information about interesting surrounding locations.
- **OAuth** is a protocol which allows a user to grant access to his private information on one site (the Service Provider), to another site (called Consumer), without sharing his identity. For consumer developers, OAuth is a method to publish and interact with protected data. For service provider developers, OAuth gives users access to

their data while protecting their account credentials [12]. OAuth protocol was used when accessing the Qype API.

6 CONCLUSION

We have presented the vision of ubiquitous computing and a brief history of the mobile phone application programming and shortly described the Android Platform, which allows everyday developers to simply build innovative applications. With the use of the Android platform, we tried to minimize the expenses when roaming on the foreign mobile networks and developed the application, which helps us locate the surrounding Wi-Fi hotspots on which we can freely access the internet.

References

- [1] M. Weiser. *The Computer for the 21st Century*. Scientific American (Special Issue: Communications, Computers, and Networks). Vol. 265 (No. 3). pp. 94–104. 1991.
- [2] S. Dobrišek. *Vseprisotni služabnik v ozadju: ambientalna inteligenca*. Moj mikro. Vol. 24 (No. 6). pp. 76-77. 2008. In Slovenian.
- [3] Commission of the European Communities, European Technology Platform on Smart Systems Integration. *Internet of Things in 2020, Roadmap for the future*. 2008.
- [4] Commission of the European Communities. *Internet of Things - An action plan for Europe*. 2009.
- [5] J. DiMarzio. *Android™: A programmer's Guide*. McGraw-Hill. 2008.
- [6] *Gartner Says Worldwide Mobile Phone Sales Declined 6 Per Cent and Smartphones Grew 27 Per Cent in Second Quarter of 2009*. Gartner Newsroom. <http://www.gartner.com/it/page.jsp?id=1126812>. Last visit in September of 2009.
- [7] R. Meier. *Professional Android™ Application Development*. Wiley Publishing, Inc. 2009.
- [8] *Android developers*. Android | Official Website. <http://developer.android.com>. Last visit in September of 2009.
- [9] R. Rogers, J. Lombardo. *Android Application Development: Programming with the Google SDK*. O'Reilly Media, Inc. 2009.
- [10] *JiWire. The Leading Mobile Audience Media Channel & Wi-Fi Advertising Network*. JiWire. <http://www.jiwire.com>. Last visit in September of 2009.
- [11] J. Kincaid. *Yelp's European Counterpart Qype Continues Global Expansion*. TechCrunch. <http://www.techcrunch.com/2008/11/26/yelps-european-counterpart-qype-continues-global-expansion>. Last visit in September of 2009.
- [12] *OAuth: introduction*. OAuth. <http://oauth.net/about>. Last visit in September of 2009.

Zbornik 12. mednarodne multikonference
INFORMACIJSKA DRUŽBA – IS 2009

Proceedings of the 12th International Multiconference
INFORMATION SOCIETY – IS 2009

Kognitivne znanosti

Cognitive Sciences

Uredili / Edited by

Urban Kordeš, Matjaž Gams, Olga Markič

<http://is.ijs.si>

15. oktober 2009 / October 15th, 2009
Ljubljana, Slovenia

PREGOVOR

Letos že šestnajstič nadaljujemo tradicijo vsakoletnih srečanj kognitivnih znanstvenikov v okviru Slovenskega društva za kognitivne znanosti in že dvanaajstič pod okriljem multikonference »Informacijska družba«.

Interdisciplinarno raziskovalno področje kognitivnih znanosti vključuje discipline s celotnega znanstvenega spektra: nevrologijo, psihologijo, računalništvo, lingvistiko, filozofijo in družbene vede (predvsem antropologijo in sociologijo). Področje raziskovanja se je v zadnjih letih razširilo od raziskovanja miselnih in višjih duševnih procesov tudi na čustva in na drugi strani na družbeno "kognitivno" delovanje. V delo kognitivnih znanosti se vključujejo tudi znanstveniki s področja fizike, kemije, biologije, ekonomije itd.

Temeljno vprašanje kognitivnih znanosti kot "dežnika" vseh disciplin, ki se v svojem raziskovanju dotikajo duševnih procesov, je, kako integrirati raznolike pristope, saj se vsaka od disciplin problemov loteva s svojega zornega kota in uporablja svoj strokovni jezik in svoje metode.

Uvedba ustreznega načina povezovanja spoznanj znanstvenih disciplin s spektra kognitivnih znanosti bi že sama po sebi rešila nekaj perečih problemov, na čelu s tako imenovanim "težkim problemom" (povezavo med duševnim in telesnim). Ravno tovrstno povezovanje je osnovni cilj kognitivne konference, ki jo prirejamo v okviru multikonference Informacijska družba.

Z vzpodbujanjem sodelovanja bi radi vzpodbudili vznik interdisciplinarne raziskovalne paradigme, ki bo združevala in nadgrajevala raziskovalne pristope, ki so se uveljavili v dosedanjem razvoju kognitivnih znanostih.

Zaradi tega vsako leto - in letos ni bila nobena izjema - poskušamo zbrati čim več raziskovalcev s širokega področja raziskovanja kognicije. Zdi se, da je letos ta projekt dobro uspel. V zborniku so zbrani raznoliki prispevki, ki dokazujejo, da so kognitivne znanosti v Sloveniji še kako žive in aktualne.

Še posebno pozornost pa smo letos namenili fenomenu odločanja. Odločanje kot kognitivni proces je v sodobni kognitivni znanosti ena pomembnejših in najbolj raziskovanih tem. K spoznanjem prispevajo najrazličnejše discipline, od nevrologije do ekonomije; vsaka od njih razvija svoje specifične raziskovalne pristope, svoj znanstveni jezik in temu primerne formulacije spoznanj. Zaradi raznolikosti pristopov je vse bolj jasna potreba po sintezi spoznanj. Članki, predstavljeni na letošnji konferenci do neke mere izpolnjujejo upe po takšni sintezi. Zbral se je širok spekter raznolikih področij, ki ponujajo svoj prispevek k razumevanju tega večplastnega pojava: nevroznanost, filozofija, fenomenološko raziskovanje, psihologija, umetna inteligenca, zgodovina in drugi. Gre za odlične prispevke k interdisciplinarnemu dialogu, ki nam - ko jih preberemo kot celoto - omogočijo širši, polnejši pogled tega procesa, ki ga po eni strani poznamo intimno, po drugi strani pa je še velika uganka.

Na kognitivni konferenci vzpodbujamo interdisciplinarno debato med znanstveniki, ki se tako ali drugače ukvarjajo z raziskovanjem zavesti. Pričujoči teksti vsebujejo poročila o raziskavah, pa tudi razmišljanja o predpostavkah, na katerih temeljijo posamezne raziskave (ki so znotraj posameznega raziskovalnega polja včasih tako samoumevne, da se jih sploh ne zavedamo).

Urban Kordeš
Matjaž Gams
Olga Markič

DEMENCA – RAZPAD UMA IN POGREZ V BLAGODEJNO POZABO

Prof. dr. Zvezdan Pirtošek

Nevrološka klinika UKC Ljubljana, predsednik Slovenskega sveta za možgane

Danes so dementni naši očetje, jutri bomo mi in pojutrišnjim naši otroci.

Ko je Alzheimer davnega leta 1907 prvič opisal demenco, mu nihče ni verjel, da takšna bolezen obstaja, saj takrat ni bilo veliko starih ljudi. Če je bilo to pred sto leti še razumljivo, je pa težko doumeti, da te bolezni ne razumemo danes in da nanjo nismo pripravljeni. Danes je v Sloveniji 27.000 naših očetov, mater, stricev in včasih tudi otrok, ki imajo motnje spomina in jim lahko postavimo diagnozo demence. V naslednjem desetletju se bo to število podvojilo in po črnih scenarijih naj bi bilo leta 2040 imelo nevrodegenerativne bolezni možganov kar štirikrat toliko prebivalcev Slovenije. Vsak tretji med nami bo dementen, če nam bo usoda tako (ne)prijazna, da bomo dočakali svoje 80. leto. Vse več bo starih ljudi in vse dlje bomo živeli. Deklica, ki se v tem trenutku rojeva v Tokiu, ima 50% možnosti, da bo dočakala 101 rojstni dan. Pričakovati je torej pravo eksplozijo nevrodegenerativnih bolezni možganov, poleg demence še zlasti Parkinsonove bolezni. To je alarmantno. In v Sloveniji na takšno stanje nismo pripravljeni.

Stopam v četrto desetletje sestankov in pogovorov z ljudmi in ustanovami, ki bi morale skrbnikom in strokovnjakom pri skrbi za bolne starejše ljudi pomagati. In moram povedati, da sem kot zdravnik in kot človek utrujen. Prehrojena pot je dolga, a dediščina teh sestankov borna. V treh desetletjih - ko sta se izmenjala dva družbena sistema in ne-vem-koliko vlad - sem se pogovarjal s politikami, z uslužbenci na ministrstvih in v različnih organizacijah. Ljudje srednjih let, v lepih oblekah in kravatah. Ali pa tisti starejši, že nekoliko naveličani in cinični birokrati in 'strokovnjaki' ki vnaprej ocenijo, da sredstev za te bolezni in za stare tako ali tako ne bo in zanje niti ne bodo zaprosili. Pogovori, v katerih dolžne empatije, vzpodbudnega trepljanja in vljudnosti ne manjka, a tudi neizrečena sporočila, da je izkušnja starosti in bolezni nekaj oddaljenega, tujega, nekaj, kar se dogaja drugim, ne meni, ne v našem družbenem sistemu, ne pod našo vlado. Osupljiva nezmožnost videti, da je starost nekaj, kar bo privilegij ali pa prekletstvo vsakega posebej.

V teh letih so svetlo izjemo predstavljali splošni in družinski zdravniki. Predvsem nanje se je zgrnilo silno breme porasta nevrodegenerativnih bolezni starostnikov, kljub temu, da jim zavarovalniška politika ni naklonjena. A razvoj stroke jih je prehitel: ob čudovitih diagnostičnih možnostih zadnjih dveh desetletij in ob novih načinih zdravljenja se pri obravnavi bolnika z demenco samokritično vse bolj zavedajo omejenega dometa splošne ambulante ali obiska na domu. Vse bolj potrebujejo pomoč specialistov.

In v tujini so ustrezno podporo tudi dobili. Zlasti potem, ko je ameriški predsednik Reagan takoj po svojem drugem mandatu stopil pred kamere in dejal: "Imam Alzheimerjevo demenco". Pri nas se je precej zgodaj potrebam stroke odzvala psihiatrija, a tudi tam dolgo in bolj ko ne kot 'one man show', ko je dr. Aleš Kogoj utrjeval temelje današnji psihiatrični obravnavi bolnikov z demenco. Interna medicina in nevrologija sta bolnike z demenco obravnavali stihijsko, brez prave doktrine. Slovenska medicina žalostno ostaja ena redkih v Evropi, ki nima specialistov za obravnavo starostnikov (geriatrije). In nevrologija bi morala ob dejstvu, da se z demenco (za migreno drugo najpogostejšo nevrološko boleznijo) do lani ni usmerjeno ukvarjal niti en sam nevrolog, že zdavnaj uravnovesiti svojo nenavadno (neevropsko)notranjo strukturo. Šele potem, ko sta nas na tem področju krepko prehiteli tudi Hrvatska in Srbija in nas pustili v društvu Makedonije in Albanije, je stroka lani končno ustanovila Center za kognitivne motnje. V njem naj bi bolnikom nudili multidisciplinarno obravnavo; žal tega ni mogoče, saj je bila dejavnosti dodeljena le desetina (!) potrebnih sredstev, na pregled pa lahko pride tako malo bolnikov, da je to le kapljica v morje. In mnogim dementnim bi lahko učinkovito pomagali, če bi jih zajeli v zgodnejši fazi bolezni.

Novo upanje pa prihaja iz Bruslja. Po tem, ko je francosko predsedstvo EU lani - v odmevnem govoru Sarkozyja - za svojo prioriteto na področju zdravstva opredelilo demenco, je junija letos tudi komisarka za zdravje Androulla Vassiliou države članice dramatično opozorila na demografska gibanja in zahtevala, da uskladijo raziskovalne aktivnosti in klinično doktrino obravnave nevrodegenerativnih bolezni možganov.

Zakaj družba bolj ne prisluhne težavam starejših? Enega izmed razlogov vidim v tem, da so demenca in možganske bolezni nasploh zelo povezane s stigmo. Malokdo je sposoben zase ali za svojega bližnjega izreči - glejte, demenco ima(m). Največkrat bomo doživeli zgolj molk, morda prtajeno šepetanje, sram pri družini in zadrego pri sodelavcih. Naše bolne dedke, babice, očete in matere družine dostikrat skrivajo in se njihove bolezni, kot mnogih možganskih bolezni, sramujejo. Gre torej za stigmo, proti kateri se moramo boriti.

Veliko lažje je bolnikom, ki lahko artikulirano povedo, da imajo diabetes, osteoporozo ali kaj tretjega. Na žalost tega privilegija bolniki, ki jim razpada spomin in um in počasi zapadajo v konec koncev blagodejno stanje popolne pozabe, nimajo. Pravi bolniki pa vse bolj postajajo njihovi izčrpani otroci, njihovi možje in žene, ki se sami, brez veliko pomoči družbe in medicine nekako borijo, da bi svojim ljubljenim kot poslednje darilo izborili dostojno kakovost življenja. Včasih jim to uspeva, dostikrat le s težavo, in včasih ob rednih službah sploh ne. Ob nezadržnem in dolgem slovesu dostikrat ostaja le še težko breme krivde.

Kot rečeno, bolniki z demenco se sami zase ne morejo boriti, kot se zase ne morejo boriti mali otroci, starostniki ali psihotični bolniki. Moralna dolžnost zdravnikov, drugih strokovnjakov in skrbnikov je, da v njihovem imenu naslovimo na tiste, ki odločajo, apel, da zanje poskrbijo. Bolnik z demenco in njegov skrbnik ponavadi izgubljata bitko s časom. S tem, ko se Slovenija ne pripravlja na skorajšnje eksplozijo nevrodegenerativnih možganskih bolezni, izgubljammo bitko s časom vsi.

In tudi Sarkozy in Androulla Vassiliou nam kmalu ne bosta mogla več pomagati.

LITERATURA

- [1] Alexander GE, DeLong MR, Strick PL 1986. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Ann Rev Neurosci* 9; 357-81.
- [2] Baxter LR, Schwarty JM, et al. 1992. Caudate glucose metabolic rate changes with both drug and behavior therapy for obsessive-compulsive disorder. *Arch of General Psychiatry*, 49: 681-9.
- [3] Claxton G 1999. Whodunnit? Unpicking the 'Seems' of Free Will. *J of Consciousness Studies*, 6; 8-9: 99-113.
- [4] Fisher CM 1983. Honored guest presentation: Abulia minor vs. Agitated behavior. *Clin Neurosurg*, 31: 9 – 31.
- [5] Fisher CM 2000. Alien hand phenomena: A review with the addition of six personal cases. *An J Neurol Sci*, 27: 192 – 203.
- [6] Frith C,D, Friston K, Liddle PE, and Frackowiack RSJ. (1991). Willed action and the prefrontal cortex in man. A study with PET. *Proc R Soc Lond (B)*, 244: 241-6.
- [7] Grey Walter W 1963. *Presentation to the Osler Society*, Oxford University; quoted in *Consciousness Explained*, D Dennett (Boston, MA: Little Brown, 1991)
- [8] Hume D. 1739. *A Treatise of Human Nature*. Edited by L.A. Selby-Bigge, 1888 and 1896. Oxford: Oxford University Press
- [9] Ingvar DH and Philipson (1977). Distribution of cerebral blood flow in the dominant hemisphere during motor ideation and motor performance. *Ann Neurol*, 2 (3): 230-7.
- [10] Jankovic J 1992. Diagnosis and classification of tics and Tourette's syndrome. In Chase T, Friedhoff A, Cohen DJ (eds): *Tourette's syndrome. Advances in Neurology*, vol. 58, New York, Raven Press, pp 7 – 14.
- [11] Jog MS, Kubota Y, Connolly CL, et al 1999. Building neural representations of habits. *Science*, 286: 1745-1749.
- [12] Kane R. 1996. *The significance of free will*. (New York: Oxford University Press)
- [13] Kant I. 1788. *Critique of Practical Reason*, ed. and tr. LW Beck (Macmillan)
- [14] Koller WC, Biary NM 1989. Volitional control of involuntary movements. *Mov Disord* 4: 153 – 156.
- [15] Kornhuber HH and Deecke L 1965. Hirnpotentialänderungen bei Willkurbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential and reafferente Potentiale. *Pfuegers Arch. Ges. Physiol.* 284: 1-17.
- [16] Libet B, Gleason CA, Wright Jr. EW, Gleason CA, Pearl DK 1983. a. Time of conscious intention to act in relation to onset of cerebral activity (readiness potential). *Brain* 106: 623-42.
- [17] Libet B, Wright Jr. EW, Gleason CA 1983. b. Preparation- or intention-to-act, in relation to pre-event potentials recorded at the vertex. *Electroencephalography and Clinical neurophysiology*, 56: 367-72
- [18] Obeso JA, Rothwell JC, Marsden CD 1981. Simple tics in Gilles de la Tourette's syndrome are not prefaced by a normal premovement potential. *J Neurol Neurosurg Psychiatry*, 44: 735-738.
- [19] Papa SM, Artieda J, Obeso JA 1991. Cortical activity preceding self-initiated and externally triggered voluntary movement. *Mov Disord* 6: 217-224.
- [20] Posner MI and Raichle ME (1994). *Images of Mind*. New York: WH Freeman and Company.
- [21] Rauch SL, Jenike MA, et al. 1994. Regional cerebral blood flow measured during symptoms provocation in obsessive-compulsive disorder using 15O-labeled CO2 and positron emission tomography. *Arch of General Psychiatry*, 51: 62-70.

RACIONALNO ODLOČANJE IN ČUSTVA

Olga Markič

Filozofska fakulteta, Univerza v Ljubljani
Aškerčeva 2, 1000 Ljubljana, Slovenia
e-mail: olga.markic@guest.arnes.si

POVZETEK

V članku najprej na kratko predstavim različne pomene pojma racionalnost. Poleg kategoričnega in normativnega pojmovanja, s katerima so se filozofi in znanstveniki v preteklosti največ ukvarjali, je z razmahom kognitivne znanosti postalo zelo zanimivo preučevanje praktične racionalnosti in odločanja v konkretnih, vsakdanjih situacijah. Raziskave kažejo, da igrajo pri takem odločanju pomembno vlogo čustva, ki pomagajo omejiti prostor izbire in pogosto vodijo do hitrejših intuitivnih odločitev.

1 RACIONALNOST

Pojem racionalnosti lahko razumemo na različne načine, zato najprej sledi kratka pojmovna analiza, nekakšen zemljevid področja. V splošnem lahko razlikujemo med kategoričnim in normativnim pomenom. Pri prvem dobimo dvojico racionalen / neracionalen, kjer je kriterij za uvrstitev v kategorijo vedenje na podlagi razlogov. Uporabljamo razlago s sklicevanjem na cilje, norme ali vrednote. Kategorični pomen lahko navežemo na znano Aristotelovo misel, da je človek racionalna žival. Kadar pa imamo v mislih normativni pomen, gre pri uvršljanju v dvojico racionalen /iracionalen za ugotovitev, ali je (ni) prepričanje ali vedenje ustrezno utemeljeno z določenimi razlogi, normami ali vrednotami.

Naslednje razlikovanje se tiče uporabe pojma racionalnosti na mislih ali na vedenju. Kadar ga uporabljamo za misli, gre za proces od detektiranja dražljaja do sposobnosti predstavljanja predmeta, kadar pa gre za vedenje, se osredotočamo na proces od avtomatičnega vedenjskega odgovora do sposobnosti oblikovati zamisel in delovati na podlagi želja in namer (intencionalnost). V prvem primeru govorimo o epistemski racionalnosti, kjer je vprašanje ujemanje predstav z objektivnim svetom. Pri tem sta izpostavljena dva cilja, t.j. doseči resnico in izogniti se napačnemu. S tem področjem se ukvarjata predvsem logika in epistemologija. Druga oblika racionalnosti pa je strateška racionalnost. Tu gre predvsem za doseganje ujemanja sveta in ciljev relativno glede na dejavnikove izbire in želje. S temi vprašanji se ukvarjata predvsem teorija odločanja in teorija iger.

Vpeljemo pa lahko še tretjo obliko razlikovanja, ki poteka po ločnici teoretično nasproti praktičnemu. Med teoretične pristope uvrščamo vprašanja znanosti, kot so vprašanja razlage in napovedi. Zanje je značilen neoseben pristop in raziskovanje je v principu dostopno vsakomur. V praktičnem razmišljanju pa se vprašanje nanaša na osebno ravnanje, na primer, kaj moram storiti oziroma kaj je najbolje storiti (tehtanje razlogov za različne alternative). Ljudje za to uporabljamo praktične silogizme in si pomagamo s psihologijo želja in prepričanj. Gre torej za prvoosebno gledišče. Ker pa včasih prihaja do nasprotij med našimi željami in prepričanji na eni ter delovanjem na drugi strani, so filozofi že od grških časov dalje ukvarjali tudi s vprašanjem šibke volje (akrasia).

V tem prispevku se bomo omejili predvsem na vprašanje, kako na racionalno odločitev vplivajo čustva. Ali so čustva pomemben dejavnik, ki imam veljavo vzporedno z bolj razumskim pristopom, ali pa so čustva le moteča in jih je treba pri odločanju čim bolj omejiti. Lahko bi rekli, da je med filozofi prevladovalo stališče, da imajo čustva pomembno vlogo pri postavljanju ciljev, vrednot in motiviranju, vendar pa so si za doseg praktičnih odločitev razum in čustva v opoziciji. O pomembnosti vloge čustev priča znana Humova opredelitev, da je "razum zgolj suženj strasti". A po drugi strani se je za sam proces razumskega odločanja zdelo, da so čustva pri tem le v nadlogo. Instrumentalna racionalnost, ki gleda na razum kot na neke vrste "preračunavanja", naj jih za doseganje cilja ne bi potrebovala. Za tak pogled se je uveljavil izraz "Negativni pogled na čustva", zasledimo pa ga lahko že pri Platonu.

Vendar se je izkazalo, da pri praktičnem odločanju v situacijah v vsakdanjem življenju, ki vsebujejo nejasnosti in tveganja, instrumentalni razum ni dovolj. Kažejo se težave, kako oceniti, katere od možnih posledic je treba vzeti v obzir. To se je zelo očitno pokazalo pri razvijanju programov umetne inteligence, kjer so se računalničarji znašli pred zagato, kako prepoznati relevantne informacije (problem okvirja). V zadnjem času so zato mnogi filozofi, nevroznanstveniki in psihologi opozorili na pomembno vlogo čustev tudi v samem procesu odločanja in ne zgolj v postavitvi ciljev. Tako filozof de Sousa pravi, da si pri problemu iskanja lahko uspešno pomagamo s hipotezo o vlogi čustev, ki omogočijo, da postane relevanten samo

majhen delež možnih alternativ in dejstev (de Sousa 2007). Podobno z vidika nevroznanosti in psihologije ugotavlja tudi Antonio Damasio. Njegovo opredelitev čustev in občutkov in njihovo vlogo pri odločanju si bomo pogledali v nadaljevanju (več o Damasijevi opredelitvi čustev in občutkov v Markič 2008).

2 DAMASIJEV POGLED NA VLOGO ČUSTEV IN OBČUTKOV PRI ODLOČANJU

Antonio Damasio je eden vodilnih raziskovalcev na področju nevrobiologije duševnih procesov. Poleg številnih znanstvenih člankov je objavil tri zelo odmevne knjige, v katerih obravnava čustva: *Descartes's Error: Emotion, Reason, and the Human Brain* (1994), *The Feeling of What Happens* (1999) in *Iskanje Spinoze: Veselje, žalost in čuteči možgani* (2008). V prvi knjigi *Descartesova zmota* je poudarek na vlogi čustev in občutkov pri odločanju. Kot lahko razberemo iz naslova, je po njegovem mnenju Descartes napačno umestil racionalno odločanje kot netelesno in ločeno od čustvovanja.

Damasio je v knjigi opisal med nevrologi zelo znan primer Phineasa Gagea, ameriškega delavca, ki se je leta 1848 poškodoval pri razstreljevanju skal. Pri nesreči mu je sunek eksplozije porinil železni drog skozi glavo. Na veliko presenečenje je Gage kljub hudi poškodbi ostal živ in si kasneje fizično skoraj povsem opomogel. A iz skrbnega, delovnega in urejenega moža se je po nesreči spremenil v nezanesljivega, nepotrpežljivega in nesposobnega slediti načrtom, postal je povsem drug človek. Rekonstrukcija Gageve poškodbe je pokazala, da sta bila poškodovana frontalna (čelna) režnja, predvsem levi.



Slika 1: Rekonstrukcija poškodbe Phineasa Gagea (avtor John M. Harlow)
Vir: Wikipedia
(http://en.wikipedia.org/wiki/File:Phineas_gage_-_1868_skull_diagram.jpg)

Damasio je raziskoval podoben primer, pacienta Elliota, ki je imel po odstranitvi tumorja poškodovano ventromedialno

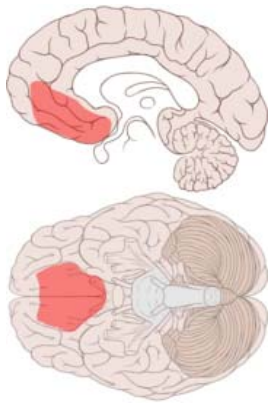
prefrontalno skorjo. Tudi Elliotov značaj se je po operaciji spremenil, njegovo vedenje je postalo neprimerno, v službi je sistematično izkazoval zmanjšane sposobnosti presoje. Čeprav na standardnih testih inteligence in spomina niso opazili večjih sprememb, se je v vsakdanjem življenju vedel nerazumno. Še posebej velike težave je imel pri sistematičnem načrtovanju. Zdelo se je, da natančno proučuje vsako podrobnost glede prihodnjega dejanja, do katerega pa sploh ni prišlo. Damasio opisuje dogodek, ko je v ambulanti Elliota povprašal, kateri od dveh predlaganih datumov za naslednje srečanje bi mu bolj ustrežal. Običajno bi človek po krajšem premisleku hitro prišel do ustreznega odgovora, pri Elliotu pa se je proces odločanja neverjetno podaljšal, saj je pretresal vse možne razloge za ali proti določenemu datumu, dokler ga po pol ure ni Damasio prekinil in mu sam predlagal datum. Elliot je odločitev mirno sprejel in s tem zaključil preračunavanje (Damasio 1994, str. 193–194).

Tudi pri delu z drugimi pacienti z okvarjeno ventromedialno prefrontalno skorjo je Damasio opazil, da imajo vsi težave pri odločanju v vsakdanjem življenju. Tu je, za razliko od reševanja miselnih nalog v laboratoriju, kjer so problemi omejeni in točno določeni, vedno prisotna precejšnja stopnja negotovosti in nejasnosti. Toda kako pojasniti njihovo neracionalno vedenje, če pa so njihove sposobnosti razumskega računanja neokrnjene? V prevladujočih teorijah odločanja so tradicionalno poudarjali predvsem razumski vidik, saj je bilo prav razumsko odločanje tista sposobnost, ki je odlikovala človeka nasproti drugim bitjem. Kot smo omenili v predhodnem razdelku, je v zgodovini filozofije prevladoval "Negativni pogled na čustva" in je veliko filozofov (npr. Platon, Descartes, Kant) zagovarjalo stališče, da je pri racionalnem odločanju vpletenost čustev predvsem ovira jasnemu razmišljanju.

Vendar so prav praktične izkušnje s pacienti Damasia spodbudile, da je začel raziskovati vlogo čustev pri odločanju. S sodelavci je pregledal in testiral še več bolnikov z okvarjeno ventromedialno prefrontalno skorjo. Ugotovil je, da pri odločanju v vsakdanjem življenju oziroma v laboratorijskih pogojih, ki skušajo simulirati take okoliščine, čustveni odzivi igrajo pomembno vlogo. Ker pri omenjenih bolnikih čustvenih odzivov niso zaznali, je zato sklepal, da je prav njihova odsotnost vir težav pri odločanju.

Hipoteza somatskih označevalcev

Zakaj imajo bolniki z okvarjeno ventromedialno prefrontalno skorjo težave pri odločanju v negotovih okoliščinah in pogosto ne delujejo v skladu z družbenimi pravili? Damasio se je reševanja tega vprašanja lotil tako, da je začel raziskovati mehanizem, s katerim lahko čustveni procesi vodijo ali usmerjajo vedenje, predvsem odločanje. Svoj pristop je poimenoval *hipoteza somatskih označevalcev*.



Slika 2: Somatski označevalci so verjetno shranjeni v ventromedialni prefrontalni skorji.

Vir: wikipedia

(http://en.wikipedia.org/wiki/Somatic_markers_hypothesis)

Ugotovil je, da pacientom v situacijah, ko se soočajo z različnimi možnostmi delovanja, ne uspe aktivirati s čustvi povezanega spomina, ki bi jim pomagal izbrati zanje najugodnejšo različico. Izkazalo se je, da so še posebno težavne situacije, v katerih se morajo odločati med izrazito nasprotujočimi se odločitvami in negotovimi izidi, kot so na primer izbira poklicne poti ali poslovne odločitve. Spoznali pa smo, da je tudi na videz tako enostavna odločitev, kot je izbira datuma za naslednji zmenek, za pacienta Elliota predstavljala velik problem.

Damasio ugotavlja, da imajo pri odločanju pomembno vlogo primerjave s čustvi in občutki, ki so človeka spremljali v podobnih preteklih dogodkih. Prav tako je vključena tudi ocena rezultata, do katerega so ti pretekli dogodki privedli, in morebitna pohvala ali kazen, ki jo je oseba dobila. Tak postopek človeku omogoča, da na osnovi preteklih izkušenj simulira možne izide v prihodnosti in se potem odloči za potezo, ki ga bo pripeljala do najbolj ugodne rešitve. Ljudje začnemo sčasoma situacije, ki jih izkusimo pod vplivom družbenih čustev in čustev sreče in žalosti, ki jih sprožita nagrajevanje in kaznovanje, razvrščati v pojmovne kategorije. Te kategorije, ki jih tvorimo na duševni in ustrezni živčni ravni, pa nato povezujemo z možganskim aparatom, ki sodeluje pri sprožanju čustev. Na tak način se v novih situacijah, ki ustrezajo določeni kategoriji, hitro in avtomatično pojavijo ustrezna čustva. Na nevrološki ravni Damasio takole opiše mehanizem: »Ko omrežja v zadnji senzorični skorji in v temporalnih ter parietalnih predelih obdelujejo neko situacijo, ki sodi k dani pojmovni kategoriji, se aktivirajo prefrontalna vezja, ki hranijo tej kategoriji pripadajoče zapise. V naslednjem koraku se aktivirajo predeli, kakršen je ventromedialna prefrontalna skorja, ki s pomočjo pridobljene vezi med dano kategorijo dogodka in preteklimi čustvenimi-občutenimi odzivi sprožijo ustrezne čustvene signale. Takšna ureditev nam omogoča, da kategorije družbenega znanja – najsi bodo pridobljene ali obogatene z osebnim izkustvom – povežemo

s prirojenim, gensko danim aparatom družbenih čustev in naknadnih občutkov.« (Damasio 2008: 139)

Čustvenih signalov se človek lahko zaveda in ponovno podoživlja občutke, lahko pa je delovanje prikrito in poteka avtomatično. V tem primeru čustveni signali možnosti in izide odločitev označijo s pozitivnim ali negativnim predznakom in tako pripomorejo k odločitvi za dejanje, ki je bolj v skladu s preteklimi izkušnjami. Ker pridemo do odločitve razmeroma hitro in brez zavestnega razmišljanja, za tako vrsto odločanja pogosto rečemo, da je intuitivno. Podobno kot de Sousa tudi Damasio meni, da čustva zožujejo prostor odločanja, poleg tega pa še »povečujejo verjetnost, da se bo dejanje skladalo s preteklimi izkušnjami« (Damasio 2008: 140). Prav ta zožitev prostora odločanja pa je pri omenjenih bolnikih manjkala.

Damasio je s sodelavci opravil vrsto testiranj, pri katerih so simulirali odločanje v pogojih negotovosti. Uporabili so psihološki test, ki ga je zasnoval Bechara s sodelavci z Univerze v Iowi in ga zato pogosto imenuje kar »Iowa test hazarderske igre« (Iowa Gambling Test). Testiranje poteka tako, da imajo udeleženci pred seboj 4 kupe kart. Povedo jim, da vsaka karta, ki jo izberejo, prinese določeno vsoto denarja, vendar občasno izbrana karta lahko pomeni tudi izgubo. Izvlečena karta tako lahko včasih prinese nagrado, včasih pa kazen. Slabe karte so po kupih različno razporejene, tako da so nekateri kupi »dobri«, saj na dolgi rok prinesejo dobiček, drugi pa so »slabi«, ker na dolgi rok prinesejo izgubo. Cilj naloge je pridobiti čim več denarja. Rezultati, ki so jih dobili Damasio in sodelavci, so pokazali, da so zdrave osebe po določenem številu izvlečenih kart obračale le še oba »dobra« kupa, medtem ko so se pacienti še naprej odločali tudi za »slaba« kupa kart, čeprav so včasih že vedeli, da s tem izgubljajo denar. Vsem sodelujočim so merili tudi prevodnost kože, ki je eden izmed fizioloških pokazateljev čustvene reakcije. Pokazalo se je, da so pri zdravih osebah zaznali stresne reakcije že mnogo prej, preden so zavestno spoznali, da je kup slab. V nasprotju s tem pa pacienti z okvarjeno ventromedialno prefrontalno skorjo teh reakcij na »slabe« rezultate niso imeli. Damasio je zato sklepal, da ti pacienti zaradi okvare niso razvili fiziološke reakcije na grožnjo kaznovanja. Prišlo je torej do oslabitve s čustvom povezanega signala in s tem do težav pri izbiri odločitve.

3 ZAKLJUČEK

Raziskovanja nezavednih dejavnikov, ki vodijo do odločitev na podlagi intuicije namesto zavestnega tehtanja alternativ, kažejo na precenjevanje vloge razumskega preračunavanja pri odločanju v vsakdanjih situacijah. Idealizirani racionalni modeli odločanja, ki so poudarjali predvsem uporabo logike in verjetnostnega računa, so se izkazali kot precej nerealen opis procesa odločanja. (Gigerenzer 2007). Damasio je s hipotezo somatskih označevalcev poudaril vlogo čustev in občutkov pri odločanju ter podal nevrobiološko in

evolucijskoo razlago. Ne glede na to, ali bo Damasijeva opredelitvi čustev vzdržala kritke, ali pa bo potrebno njegovo teorijo dopolniti ali celo bistveno spremeniti, je s svojim raziskovanjem spodbudil k nadaljnjemu raziskovanju vloge čustev pri odločanju in k ponovnim premislekom o racionalnosti.

LITERATURA

Bechara, A., Damásio, A. R., Damásio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex, *Cognition*, 50, str. 7-15.

Damasio, A. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: G.P. Putnam's Sons.

Damasio, A. (1999). *The Feeling of what Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt Brace and Co.

Damasio, A. (2008). *Iskanje Spinoze: Veselje, žalost in čuteči možgani*. Ljubljana: Krtina.

De Sousa, R. (2007). *Why think? Evolution and the Rational Mind*. Oxford, New York: Oxford University Press.

Gigerenzer, G. (2007). *Gut Feelings*. Allen Lane, Penguin Books.

Markič, O. (2008). "Nevrobiologija čustev in občutkov", v Damasio, A. (2008). *Iskanje Spinoze: Veselje, žalost in čuteči možgani*. Ljubljana: Krtina, str. 287 – 302..

FENOMENOLOGIJA ODLOČANJA

Urban Kordeš

Pedagoška Fakulteta Univerze v Ljubljani
Kardeljeva ploščad 16, 1000 Ljubljana, Slovenia

Tel: +386 1 589 2200

e-mail: urban.kordes@guest.arnes.si

POVZETEK

Na področjih, ki ne dovoljujejo pogleda na ljudi kot na trivialne stroje, ki delujejo deterministično oziroma brez lastne avtonomije, nepoznavanje njihove (naše) izkustvene pokrajine pomeni izgubo zelo pomembnega, morda celo bistvenega elementa. Zato se v zadnjem času odpira novo raziskovalno področje: poskus sistematičnega opazovanja izkustva. Gre za t.i. fenomenološko inspirirane raziskovalne pristope.

Del članka je namenjen predstavitvi tega mladega raziskovalnega področja – prikazane so skupne osnove pristopov in nekatere značilne metodološke izpeljanke. Vse to pa je navezano na možnosti raziskovanja odločanja s prvoosebnega vidika, torej odločanja kot doživljajskega fenomena (in ne kot nevrološkega procesa).

Fenomenološko raziskovanje

Pregled začnimo z opisom temeljev na katerih stoji raziskovanje izkustva. En od glavnih je spoznanje, da je človekovo izkustvo oziroma doživljanje primarno. Ne le da je primarno: to je vse kar imam na voljo – ničesar ne morem spoznati izven polja svojega doživljanja. Misli, pomeni, opisi, vizije, slike, občutki, čustva, ... nič od tega ni izven zavestnega doživljanja (oziroma mojega živega izkustva). Doživljajski svet seveda lahko urejam (npr. z razdelitvijo na čustva, misli itd. ali pa z znanstveno razlago opaženega), ne morem pa spoznati ničesar kar bi bilo izven mojega izkustva.

S fenomenologijo inspirirano oziroma krajše fenomenološko raziskovanje se torej osredotoča na opazovanje živega, neposrednega izkustva. Pojem "fenomenološko" je povzelo po filozofski smeri, ki se je začela okrog leta 1900 in si je zgornje ugotovitve postavila za svoj osnovni epistemološki kredo. Naj naštejemo nekaj osnovnih metodoloških smernic, kot jih je načrtoval oče te filozofske smeri, matematik in filozof Edmund Husserl¹:

¹ Smernice povzemam po poenostavljenem, a odličnem pregledu fenomenologije v Kotnik (2003, str. 102) – od tod so vse dobesedne navedbe, ki delno sledi Ihdeju (1977, str. 32 do 45).

- Odredotočenje na fenomene (stvari kot se kažejo v našem izkustvu) in epoche – postavljanje v oklepaje. Husserl predlaga "puščanje običajnih privzetkov o stvareh ob strani" (oz. dajanje le-teh v "oklepaje") in fenomenološko redukcijo – redukcijo opazovanega na fenomene "kot tisto edino, kar je dano v izkustvu in gotovo".
- Pravilo: "Ne razlagaj, ampak opisuj!" To je najpomembnejša metodološka smernica za fenomenološko raziskovanje. Morda izgleda navodilo na prvi pogled preprosto, vendar je njegova izvedba zelo kompleksna in zahteva dobršno mero refleksije in spretnosti. Šele, ko v praksi poskusimo samo opisovati doživljanje, ne da bi ga kakorkoli klasificirali, umeščali v teoretske okvire, pojasnjevali ipd., se zavemo kako globoka je naša potreba po razlagi in kako težko se ji izognemo.
- Vzdržimo se prepričanj oziroma ocenjevanj "realnosti" opaženih fenomenov. Res je, da brez intersubjektivnega preverjanja (ki je značilno za znanstveno metodo) ne moremo razločiti med "iluzijo" in "resničnostjo". Toda z vidika primarnosti neposrednega izkustva je takšno razlikovanje le en način urejanja doživljajskega sveta (in torej nima višje vrednosti od drugih možnih razlikovanj), zaradi tega Husserl priporoča, da se celo to – na videz tako primarno sodbo – postavimo v oklepaj in opazujemo polje izkustva kot se kaže, brez presojanja.

Ker se je Husserlu zdelo, da bi samo z zgornjimi navodili ostali "izgubljeni med fenomeni" (prav tam) je dodal še četrto smernico, ki priporoča iskanje strukture in konstantnih značilnosti opazovanih fenomenov. Husserl je verjel, da je na tak način mogoče ustvariti "transcendentalno" znanost – torej sistem, ki presega samo enkratno minljivo konkretno izkustvo, ampak iz njega izloči bistvene (transcendentalne) elemente.

Opisana epistemološko-metodološka drža je tisto, kar fenomenološko raziskovanje (kot kvalitativno-metodološka smer) povzema po fenomenologiji kot filozofski disciplini. Ne zgleduje pa se po filozofski odmaknjenosti od empiričnega raziskovanja ampak poskuša kar se da velik del svojih trditev utemeljiti na empiričnem raziskovanju doživljanja. Kot bo razvidno v nadaljevanju teksta, so metode, kako priti do teh podatkov raznolike in še ne

povsem dodelane, vsem pa je skupno to, da se izogibajo t.i. "armchair introspection", torej mnenju, da že tako vemo vse kar se da vedeti o našem doživljanju. Fenomenološko raziskovanje lahko torej v splošnem definiramo kot metodološko kategorijo, ki zajema vse empirične raziskovalne metode, pristope, načine zbiranja in/ali analize gradiva, ki temeljijo na zgoraj opisanih (fenomenoloških) epistemoloških temeljih ter smernicah.

Fenomenološko raziskovanje in odločanje

Danes smo priča mnogim različnim, vzporednim poskusom raziskovanja doživljanja. K razvoju modernega fenomenološkega raziskovanja kot empirične znanstvene discipline, predvsem pa k njegovi uveljavitvi v kognitivni znanosti, je bistveno prispeval Francisco Varela s svojim nevrofenomenološkim projektom, katerega začetek zaznamuje novo dobo znanstvenega zanimanja za izkustvo.

Že v temeljnem članku nevrofenomenologije (Varela, 1996) je Varela opazil, da je raziskovanje volicije in posledično odločanja ena od pomembnih nalog in hkrati potencialov (nevro)fenomenološkega polja: "The nature of will as expressed in the initiation of a voluntary action is inseparable from consciousness and its examination. Recent studies give an important role to neural correlates which precede and prepare voluntary action, and the role of imagination in the constitution of a voluntary act (Libet, 1985; Jeannerod, 1994). Yet voluntary action is preeminently a lived experience which has been thoroughly discussed in the phenomenology literature, most specifically in the role of embodiment as lived body (corps propre, Merleau-Ponty, 1945), and the interdependence between lived body and its world (Leibhaftigkeit)." (Varela, 1996; vključno s citiranimi deli)

Kljub temu, da so pojmi kot volicija, odločanje, svobodna volja itd. očitno primarno doživljajski (torej opisi doživljanja in ne teoretski koncepti), do zdaj ni bilo večjih raziskav na to temo. Ob koncu pričujočega teksta bodo predstavljeni delni rezultati fenomenološke študije odločanja, ki je bila izvedena v okviru interdisciplinarnega raziskovalnega projekta "Metodološki vidiki raziskovanja kognitivnih procesov – učenje in odločanje", že takoj pa lahko povzamemo nekaj ugotovitev, ki so vzniknile kot filozofska refleksija po podrobnejšem pregledu doživljajskih poročil, povezanih z odločanjem.

Odločanje lahko obstaja le znotraj mišljenja

Prva stvar, ki jo opazimo, ko na odločanje pogledamo s fenomenološke perspektive, je, da v večini razprav o tem pojavu (procesu) ni jasno definirano, ali je odločanje vedenjska kategorija (pogled s tretjeosebne perspektive) ali doživljajska (prvoosebna perspektiva). Raziskovalci, ki se ukvarjajo z odločanjem, vse preradi gledajo na ta proces kot na nekaj, kar je "tam zunaj", neodvisno od opazovalca in neodvisno od tega kako subjekt doživlja. Torej na

podoben način, kot gledamo na primer fizikalne pojave (gravitacija je "tam zunaj", če jo opazim ali ne).

S perspektive fenomenološkega raziskovanja je takšen pogled nesmiseln. Očitno je, da si moramo najprej biti na jasnem na kaj mislimo, ko govorimo o odločanju (enako velja seveda za vse ostale kognitivne fenomene). Zdi se, da obstajajo znotraj sodobne kognitivne znanosti trije možni pogledi, vsak določen z epistemološkim ozadjem pripadajoče vede, vsak znotraj svojega področja legitimen:

Psihološki: Kategorijo (npr. "odločanje") določi ekspert na podlagi opazovanja vedenja. Opazovanje lahko vključuje tudi vprašanja (navadno v obliki vprašalnika), ki pa jih ekspert interpretira, glede na ostale zbrane podatke. Zadnjo besedo pri uvrstitvi ima torej ekspert, ne subjekt.

Nevroznanstveni: Zaenkrat nevroznanost še ni dovolj precizna, da bi lahko samo s pogledom na fiziologijo z gotovostjo trdila, katera kognitivna funkcija se pravkar vrši, vendar je prizadevanje te discipline seveda usmerjeno v to smer.

Fenomenološki: Fenomenološko raziskovanje se ne sprašuje o vzrokih za določeno doživljanje (kot na primer psihoanaliza), niti po njegovih fizioloških korelatih. Zanima ga samo kaj je subjekt doživljal in on sam je edini ekspert za to. Fenomenološki raziskovalec lahko poročevalcu pomaga izostriti pogled na doživljanje, lahko mu pomaga raziskovati njegovo lastno doživljajsko pokrajino, ne more pa soditi katero poročilo je resnično in katero ne. Tisti, ki doživlja ("subjekt" je dokaj neprimeren izraz zanj, vendar se ga bom v nadaljevanju vseeno držal) je edini ekspert za svoje doživljanje. In z vidika fenomenološkega raziskovanja lahko samo on oceni katere vrste izkustvo je doživel.

Iz zgornjega sledi, da je – s fenomenološkega gledišča – odločanje samo tisto doživljanje, ki ga subjekt spozna za odločanje.

Mnoga vedenja, ki bi jih lahko od zunaj označili kot odločanje, so – z doživljajskega vidika – samo delovanje v skladu z okoliščinami (oziroma: delovanje, primerno okoliščinam). V psihološki literaturi na temo odločanja lahko najdemo mnogo tovrstnih primerov. Tvorci teorije t.i. naravnega odločanja (Klein, Rasmussen in drugi) na primer večkrat omenjajo svoje raziskave gasilcev, ki so se v kritičnih situacijah (tako ali drugače) "odločili", ne da bi sami svoje vedenje videli kot odločitev. Gasilec, ki se je "odločil" zapustiti sobo v katero je pravkar prišel je v tem trenutku doživljal to vedenje kot edino možno in ne kot enega od možnih izborov. (Kasneje ga je šele racionaliziral in odkril vzroke/razloge zanj.)

V fenomenoloških poročilih, ki smo jih analizirali, se je pojavilo mnogo primerov, ki so od zunaj verjetno izgledali kot odločanje, poročevalci pa svojega doživljanja niso videli kot takega. Npr. v primeru nakupa v pekarni ali izbora jedila v restavraciji: obkraj je poročevalec, kljub

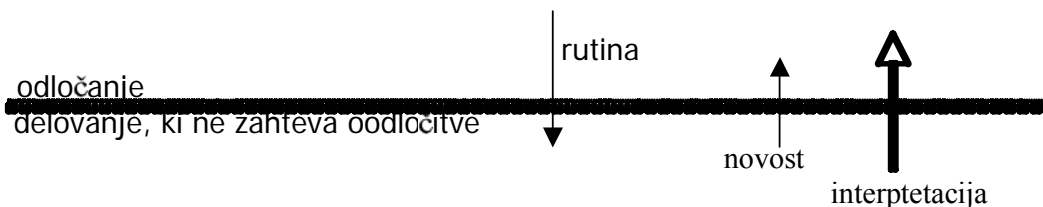
izbiri, ki se je ponujala, "začutil" kaj hoče in naročila ni videl kot odločitve.

V fenomenoloških poročilih pa lahko najdemo tudi nekaj nasprotnih primerov: včasih doživljamo, da se odločamo v trenutkih, ko naše vedenje ne kaže na to. Primer je poročilo moškega s Tourettovim sindromom, ki se mu je po končanem tuširanju vsilila misel, da bo, če stopi ven iz kabine, moral ubiti očeta. Z vidika zunanjega opazovalca bi verjetno lahko zaznali obotavljanje pri zapuščenju kabine, težko pa bi temu rekli "odločanje", saj ne obstajajo alternativne možnosti.

Zaključimo lahko, da nas v odločanje ne sili narava sveta. Izven človeške resničnosti oziroma boljše povedano, izven človeške racionalne interpretacije, ni odločanja. Ideja izbire in posledično koncept odločanja je racionalna iznajdba. Odločanje je samo v jeziku oziroma v mišljenju. Glede na raziskave Russela Hultburta (1992) se sicer velikokrat pojavlja v obliki »neformaliziranega mišljenja«, torej razmišljanja brez jasno izraženih besed, vendar vseeno razmišljanja.

Vsekakor ni tako, da bi nas življenjske situacije silile v odločitev. V odločitev nas sili interpretacija življenjske situacije.

Ob tem postane zanimivo vprašanje za fenomenološki razmislek: kaj je tisto, kar napravi določene doživljajske situacije "zrele" za odločanje. Zakaj se v določeni situaciji začnemo "odločati", v kakšni drugi – na videz povsem analogni – pa ne? Odgovor seveda presega ozke meje tega teksta, lahko pa bi grafično predstavili nekaj dejavnikov, ki vplivajo na to, ali subjekt zazna svoje doživljanje kot odločanje ali kot neprekinjen "tok":



Naslednje, še bolj pomembno vprašanje pa je: kako pride do odločitve. Vse kaže, da je čisti racionalni proces zelo redko, če sploh kdaj, sposoben prevesti tehtnico na eno stran. Fenomenološke raziskave se na tej točki strinjajo z nevroznanstvenimi ugotovitvami, da je potreben nekakšen energetski sunek, ki prekine stisko negotovosti. V tej luči bi lahko odločanje videli kot čakanje oziroma aktivno iskanje impulza (energetizacije, ki da eni od možnosti dovolj teže, da prevlada nad ostalimi). Miselni proces se tako kaže bolj kot ovira ali pa samo kot krinka za nekaj drugega, ne (toliko) racionalnega.

A propos zgoraj napisanemu bi fenomenološka poročila lahko grobo razdelili na dve kategoriji: na poročila, ki opisujejo proces odločanja kot (več ali manj) mučno

nihanje med različnimi idejami in na tista, ki govorijo o čakanju na impulz (ki mu včasih predhodi racionalen razmislek o možnostih). Če vzamem nevrološke opise samo kot metaforo, bi lahko rekel, da gre enkrat za tekmujoče se zanke (recimo, da se frontalni korteks bori proti neki drugi zanki/recimo odvisnosti/ ali pa se borita med sabo dve ali več neokortikalnih zank) in drugič za sodelovanje (kjer neokorteks ojači ali vsaj dopusti /"pojasni"/ nek drug impulz). Enkrat torej gre za boj oz. konkurenco oz. reševanje frustracije, drugič pa za usklajeno delovanje.

Ker okvir pričujočega pisanja ne dopušča resne analize zgornjih primerov, naj samo še povzamem.

Jasno je, da gre tudi s fenomenološkega vidika (podobno kot s filozofskega) za zelo široko kategorijo, morda boljše povedano – za družino doživljanj (procesov) – in da je veliko vprašanje, koliko skupnega lahko najdemo v tej družini. Poleg zgoraj naštetih primerov lahko opazimo tudi: odločanje da, odločanje za, odločanje med itd. Potem so tu seveda trenutki, kjer samo spoznamo nujnost (tudi, če je dogodek pozneje okarakteriziran kot pomembna življenjska odločitev).

Že resen razmislek o pojavu pokaže, da se preprosti psihološki eksperimenti ne morejo dotakniti doživljajskega bistva odločanja in da je zato potreben (in tudi etično nujen) dialoški pristop, kjer se oba – subjekt in raziskovalec – urita v občutljivosti za drug drugega. Pri fenomenološkem raziskovanju se postavlja tudi bolj temeljno epistemološko vprašanje: kdo je sploh raziskovalec. Je možno raziskovati doživljanje drugega, če le-ta nima iskrenega in globoko eksistenčnega interesa po opazovanju svojega doživljanja?

Literatura

- [1] Hulburt, R. (1992): *Sampling Normal and Schizophrenic Inner Experience*. Plenum Press, New York
- [2] Varela, F (1996): *Neurophenomenology – remedy for a hard problem*. *Journal of consciousness studies*, 3, 1, 330-350
- [3] Kotnik, R. (2003): *Pouk filozofije kot osebno doživljanje: izkustveno poučevanje filozofije kot aplikacija načel Gestalt terapije*, (Knjižna zbirka Učbeniki, 5), Pedagoška fakulteta v Mariboru

SE ODLOČAMO GENETSKO ALI PRIVZGOJENO – ANALIZA POSILSTVA?

Matjaž Gams

Odsek za inteligentne sisteme

Institut Jožef Stefan, Jamova 39, 1000 Ljubljana, Slovenia

matjaz.gams@ijs.s

POVZETEK

Odločanje pri posilstvu je dolgoletna tema raziskav, saj zadeva analizo smotrnosti takšnega početja v smislu, ali je to dejanje genetsko – evolucijsko pogojeno, ali pa je privzgojeno oz. svobodna volja. Tema posilstva je bila izbrana zato, ker se je v svetu ponovno razživila debata v tej temi, ko je nekaj znanstvenic z novimi orodji in podatki analiziralo stara spoznanja, tehnike in podatke. Pokazalo se je, da je bila večina desetletja starih analiz narejena na z današnjega stališča staromodno in ne najboljši način, ter da ne dopušča tako preprostega sklepanja. Po drugi strani skuša novi rod znanstvenic zanikati genetske in evolucijske osnove teorij odločanja, kar se zdi strokovno vprašljivo. Morda bi lahko celo rekli, da se vračamo v stari »komunistični način« razmišljanja, po katerem se človek odloča ne glede na genetske predispozicije, samo tokrat strokovno vsaj delno podprto. Avtor tega prispevka pa meni, da je sklepanje pri in o posilstvu zelo zapleten pojav, kjer sodelujejo vsi omenjeni faktorji, hkrati pa ostaja nejasno, kolikšen del tega dogodka je genetsko pogojen in kolikšen del privzgojen. Boljše analize in boljše razumevanje še niso prinesle dokončnega odgovora.

1. UVOD

Ali se pri odločanju ljudje odločajo samostojno, ali pa je njihovo mnenje pogojeno z genetiko, vzgojo ... in v končni konsekvenci sploh nimajo dejansko svobodne volje?

Wikipedija (Fre will) navaja, da je prvo vprašanje o **determinizmu**. Po determinizmu je vsako odločanje vnaprej determinirano z okoliščinami, zgodovino itd.

Če je torej vse determinirano, nimamo svobodne volje, kar po Wikipediji odgovarja trdemu determinizmu. Če pa lahko izbiramo po lastni volji, potem lahko vsaj občasno izberemo svojo odločitev ne glede na okoliščine. Ta smer se angleško imenuje Compatibilizem. Za kognitivne znanosti je pomembno, da nekatere smeri zagovarjajo svobodno voljo preko interdisciplinarnih kognitivnih znanosti, recimo preko kvantne mehanike in naključnih dogodkov skušajo vpeljati nedeterminističnost. Preko dualizma skušajo drugi znanstveniki vpeljati svobodo misli, odmaknjeno od fizikalnih okvirov. Nekatere filozofske smeri se zdijo avtorju tega prispevka nekoliko nenavadne, kot tista, da je vse determinirano ali pa vse naključno.

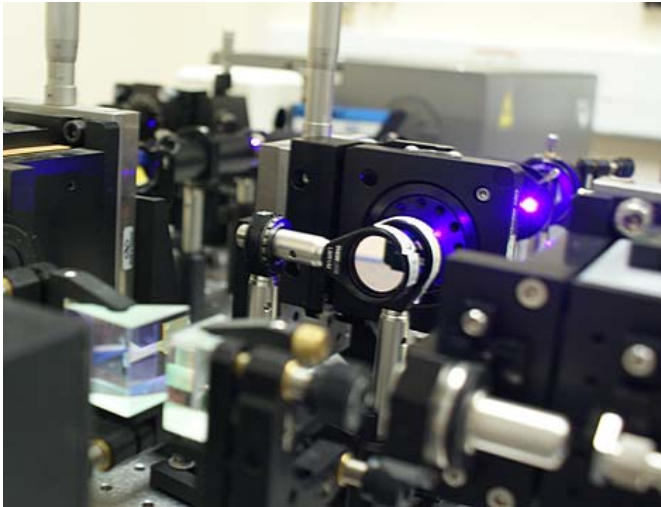
Najbrž lahko s precejšnjo gotovostjo po zdravi pameti zatrdimo, da so določene odločitve vezane na pogoje in omejitve, prav tako pa so določene odločitve po svobodni volji oziroma celo naključne. Taka osnovna vprašanja – **ali sploh obstaja prosta volja**, niso del raziskav, povezanih z umetno inteligenco, in ostajajo del raziskav drugih znanosti. Za kognitivno informatiko, inteligentne sisteme in umetno inteligenco (AI Magazin 2008) je ključno vprašanje konstrukcije programov, ki se bodo obnašali podobno kot ljudje – torej **bodo izkazovali neko lastnost, podobno lastni volji, ne glede na filozofsko utemeljitev, ali gre za navidezno ali resnično** (Brooks 2008, Gams 2003).

2. RAČUNALNIŠKI (NE)DETERMINIZEM

Večina obstoječih programov je determinističnih, saj načeloma le izvajajo zaporedje formalnih ukazov nad formalnimi simboli. Sta pa vsaj dve pomembni področji, kjer **determinističnosti ni več**: agenti in kvantni računalniki.

Medtem ko so **kvantni računalniki** v principu nedeterministični, pa v realnosti delujejo dokaj deterministično, gledano navzven. Zadnje leto so raziskovalci uspeli zgraditi nekaj čipov oz. vezij, ki zmorejo opravljati določene kvantne naloge, kjer so kvantni algoritmi boljši kot klasični. Na konferenci IJCAI (Gams 2009) so predstavili kvantno vezje v velikosti noge, ki je računalo optimume na osnovi tunelskega efekta – namesto preiskovanja vseh lokalnih optimumov, kar je NP težko, preiskovalna točka enostavno preskoči na boljši sosednji optimum. Te mesece poročajo tudi o kvantnem čipu za dekodiranje, kjer na kvantni način lahko izjemno hitreje in uspešneje dešifrirajo zakodirana sporočila (IEEE Spectrum 2009). Iz članka: »Modern cryptography relies on the extreme difficulty computers have in factoring huge numbers, but an algorithm that works only on a quantum computer finds factors easily. Today in Science, researchers at the University of Bristol, in England, report the first factoring using this method—called Shor’s algorithm—on a chip-scale quantum computer, bringing the field a tiny step closer to realizing practical quantum computation and code cracking. Quantum computers are based on the quantum bit, or qubit. A bit in an ordinary computer can be either a 1 or a 0, but a qubit can be 1, 0, or a ”superposition” of both at the same time. That makes solving certain problems—like factoring—exponentially faster, because it lets the computer try many more solutions at once. The race is on to find the ideal

quantum computer architecture, with qubit contenders that include ions, electrons, superconducting circuits, and in the University of Bristol's case, photons.«



Slika 1: Kvantni sistemi so načeloma nedeterministični. Slika prikazuje sistem za kvantno dešifriranje bristolske univerze.

Drugo pomembno področje nedeterminističnih programov so **inteligentni agenti**. Ker smo o njih že večkrat pisali (Gams 2001), na kratko povzemimo: Inteligentni agenti imajo v smislu računalniških programov edini lastno voljo, ker se morajo odločati v okoliščinah, ki niso vnaprej predvidljive. Poglavitne lastnosti so avtonomnost (lastnost svobodne volje), socialnost (reševanje nalog v sodelovanju z drugimi akterji) in racionalnost (pričakovanje primerne reakcije v neznanih okoliščinah). S filozofskega stališča so inteligentni agenti seveda še vedno programi, saj so napisani kot formalni sistemi, ki obdelujejo formalne simbole. Po mnenju tega avtorja (Gams 2001, 2003) in nekaterih drugih pa imajo agenti zaradi svojih posebnih lastnosti, recimo interakcije (Wegner 199??), večje zmožnosti kot univerzalni Turingovi stroji in ti programi, vsaj v smislu definicij inteligentnih sistemov, imajo »**inženirsko lastno voljo**«.

3. ANALIZA ODLOČANJA PRI POSILSTVU

Ne glede na to, da so osnovne računalniške teorije na videz dokaj ločene od dogajanja v človeškem svetu, se kognitivna informatika stalno zgleduje po ugotovitvah o ljudeh. V zadnjih letih se je po svetu razživel debata o odločanju pri posilstvu, zato jo tu navajamo in analiziramo.

Zakaj so storilci posilstva skoraj izključno moški vseh ras, ideologij, izobrazb, premoženj? Npr. izraelski predsednik Moshe Katsav je bil obtožen posilstva leta 2007 in je osramočeno odstopil. Po mnenju feministk gre predvsem za škodljivo in nasilno dokazovanje premoči moškega spola nad ženskim, po mnenju evolucionističnih psihologov pa gre za evolucionistično zakodirano obnašanje. Moški, ki so genetsko nagnjeni k posilstvu, naj bi imeli evolucionistično prednost, saj je večja verjetnost, da bodo imeli več potomcev. Slej ko prej

torej tovrstni gen prevlada in zato imajo danes vsi moški po svetu gene za posilstvo. In tu pridemo do

- a) vprašanja evolucionistične »ekonomike« posilstva – ali so genetsko k posilstvu nagnjeni moški res evolucionistično uspešnejši? In še bolj relevantno za temo odločanja –
- b) ali je torej odločanje evolucionistično pogojeno?

Vprašanje a) lahko v veliki meri odloči o b), saj je v primeru evolucionistične uspešnosti posiljevalcev večja verjetnost, da je to zakodirano v genih, torej je svoboda razmišljanja toliko manjša.

Evolucionistični psihologi so zadnja desetletja v razvitem svetu veljali za znanstveno verodostojne razlagalce osnov posilstva, medtem ko različna ideološko usmerjena gibanja v strokovnih debatah ne uživajo ravno velikega zaupanja. Če poenostavimo, evolucionistični psihologi obravnavajo moške v smislu evolucionističnega boja – močnejši, agresivnejši in uspešnejši (bogatejši) so, več uspeha imajo pri lepšem spolu, podobno kot pri drugih sesalcih, ker je večja verjetnost, da bodo njihovi otroci uspeli in prenesli gene naprej. Evolucionistično pogojeno naj bi moški imeli genetske module tako za posilstvo kot npr. za agresivnost do otrok, ki niso njihovi potomci. Statistika potrjuje tovrstne teorije – po starejših analizah očimi nekaj desetkrat bolj verjetno poškodujejo pastorko kot svoje otroke. Primer iz živalskega sveta - levi po prevzemu krdela pobijejo mladiče, mlajše od nekaj mesecev, medtem ko hrano raje delijo s svojimi mladiči kot z levinjami. Nesporno imajo levi tovrstno obnašanje zakodirano v DNK. Ljudje tudi?

Statistike nesporno dokazujejo, da določeni posamezniki kljub psihičnemu zdravljenju in uporabi kemičnih zdravil ostanejo nepopravljivi, recimo večji del pedofilov. Sodna praksa sicer te ljudi obravnava ne kot bolnike, ki se ne morejo sami odločati, ampak kot ljudi z lastno voljo.

Evolucionistične razlage odnosov med spoloma gredo še dlje – po tej teoriji naj bi npr. moški bolj zamerili svoji ženski seks z drugim moškim, saj to pomeni večjo nevarnost, da bo imela otroka z drugim, kot pa če se zaljubi v drugega. Nasprotno pa naj bi ženske bolj zamerile svojemu moškemu, če bi se emotivno navezala na drugo žensko, kot pa če bi seksal z njo, ker bi zaljubljenost lahko pomenila, da bi za vedno izgubila podporo svojega moškega, seks z drugo žensko pa ne bi ogrozil podpore njenemu potomstvu. Skratka, vedno in povsod naj bi odločala »ekonomika« evolucionističnih prednosti, ki naj bi bila večinoma zakodirana v genih.

V zadnjih nekaj letih pa evolucionistična psihologija doživlja temeljito prevetritev. Pri podrobnem preiskovanju evolucionističnih razlag posilstva, izbire partnerja, prešuštva ali odnosa očimov se je razkrila vrsta dvomov (Sharon Begley, Newsweek, 29.6.2009).

Kar nekaj objav iz ZDA in Kanade je pokazalo, da so bile prvotne študije o nekaj desetkrat večji verjetnosti poškodbe

pastorka opravljene nesistematično. Švedska študija iz leta 2005 je pokazala kompleksnost tovrstnih dogodkov – pogosto so bili storilci tovrstnih dejanj nasploh nagnjeni k nasilju, tudi mentalno nezdravi. Njihova razlaga je, da so očimi, tj. poročeni z ločenko z otroki, pogosto problematični ljudje, npr. pretirano nagnjeni k alkoholu. Namesto verjetnosti 1 proti 30 med poškodbo svojega otroka ali pastorka novejše študije navajajo vrednosti 1 proti 2 do 5.

Ta številka ni zamajala prepričanja avtorjev teorije o mentalno-genetskih modulih za agresivno obnašanje do pastorkov, Martina Dailyja in Margo Wilson (*The Truth about Cinderella: A Darwinian View of Parental Love*), ki trdita, da je celo ob takšnih razmerjih, kot jih navajajo novejše študije, neznanstveno zanikati signifikantnost relacije. Poleg tega je Pepelka zgodba o hudobni mačehi in ne očimu in nakazuje, da so enako tudi ženske izrazito evolucijsko preračunljive, kar so tako ali tako vedno trdili evolucijski psihologi – da so razlike med moškimi in ženskami le zaradi različne biološke vloge, v bistvu pa sta oba spola najmočneje podvržena istemu genetskemu in evolucijskemu boju. Ne glede na odgovore imajo kritiki prav vsaj toliko, da so bile prvotne preiskave nekoliko površne, saj niso upoštevale vseh odvisnosti. Zanimiva je tudi primerjava med vzhodno in zahodno prevladujočo razlago. Recimo na vходу so bile v veljavi družboslovne in ne genetsko utemeljene raziskave, zato evolucijskih razlag družbenih pojavov skoraj ni bilo. Lahko bi se tudi vprašali, ali na slovenskih fakultetah že poučujejo te razlage ali ne.

A poglobljenih študij je čedalje več in so čedalje bolj zapletene. Stephen Beckerman v *Proceedings of the National Academy of Sciences* leta 2009 poroča o analizah plemena Waorani (amazonških indijancev), kjer se še vedno bojujejo med seboj. Izkazuje se, da so super bojevniki slabši očetje kot bolj umirjeni posamezniki, saj je verjetnost uspeha njihovih potomcev zaradi večje verjetnosti maščevanja sovražnikov manjša. Podobno so študije smotnosti oz. »ekonomije« posilstva (Thornhill, Palmer, *A Natural History of Rape: Biological Bases of Sexual Coercion*, MIT Press, 2001) s kvantitativnimi modeli razmer v paragvajskem plemenu Ache z ocenami verjetnosti uspeha dogodka in negativnih posledic pokazale negativen izračun. Od tu so nasprotniki evolucijskega vpliva sklepali, da osnovna evolucijska profitabilnost posilstva ne drži, še manj torej gen za posilstvo.

Pri izbiri partnerjev so donedavne študije večinoma pokazala, da imajo moški raje mlade ženske z vitkim pasom v primerjavi z debelejšimi postavami, medtem ko so ženske uspešnost in položaj moškega postavljale nad izgledom. Oboje naj bi bilo povezano z verjetnostjo uspešnosti potomcev. Novejše analize kritikov pa so pokazale, da je ocenjevanje odvisno od socialnih razmer – dobro situirane ženske so pogosto izbirale moške tudi ali celo predvsem glede na videz. V modernih družbah zahodne Evrope, kjer je dosežena velika stopnja enakopravnosti spolov, moški

pogosto izbirajo družice glede na finančno uspešnost in manj na izgled.

Če si ogledamo še statistike o prešuštvu v zahodni civilizaciji – ameriški študentje bi dvakrat bolj kot študentke prizadelo, če bi partnerka storila prešuštvo, kar je skladno s teorijo evolucijske psihologije. Toda presenetljivo je polovica študentov bolj kot fizično ocenila nesprejemljivo mentalno prešuštvo, kar postavlja donedavne teorije v čudno luč. V Nemčiji in na Nizozemskem pa je tovrstni odstotek le 28 oziroma 23. In nasprotno – le 13% ameriških, 12 nizozemskih in 8% nemških žensk bi bolj motila partnerjeva čustvena zveza z drugo žensko v primerjavi s kratkotrajnim skokom čez plot. Očitno je ocena odvisna od družbenih razmer – v sodobni enakopravni in spolno odprti družbi hiter ženski seks po statistiki sodeč ni več tako neprijeten kot emotivna odtujenost. Hiter seks tudi ne vodi več do otrok, kar bistveno spremeni »izračun«.

Po mnenju kritikov evolucijsko podprtih teorij je nauk teh analiz, da je naivno sklepati o enotni človeški naravi in okolju na osnovi genetsko zakodiranih modulov – ljudje smo prilagodljiva bitja in živimo v raznovrstnih bioloških in družbenih okoljih. Spreminja se okolje, spreminjajo se ljudje in celo geni. Ljudje se bistveno razlikujemo od živih bitij v tem, da se razumsko odločamo racionalno in ne predvsem glede na prirojeno. Z drugimi besedami – če se neko ustaljeno obnašanje izkaže za neracionalno, ga lahko ljudje spremenimo v relativno kratkem času, medtem ko živali za to potrebujejo genetsko spremembo, kar lahko traja tudi tisoče let. Primer je iberijski ris, kjer se v leglu mladičev sproži genetsko zakodiran spopad med mladički, dokler ne ostane samo eden; podoben pojav je pri nekaterih pticah. Danes, ko ogrožene iberijske rise umetno vzgajajo, jih morajo umetno ločevati, da preprečijo v sedanjih razmerah povsem neracionalno potezo.

Pa vseeno – zakaj še vedno posiljujejo tako amazonški Indijanci kot Eskimi? Zakaj je na Japonskem število posilstev od leta 1997 do 2001 poskočilo za 40%, število obscenih dejanj pa se je potrojilo od leta 1991 do 2001? Najbrž kritiki evolucijske psihologije utemeljeno opozarjajo, da so se razmere tudi med spoloma v zadnjih desetletjih tako spremenile, da so nekateri pojavi ali »računice« drugačni kot pred stoletji. Vprašanje pa je, v kolikšni meri to velja za posilstvo. Praljudje so posiljevalce, če so jih ujeli, pokončali. Danes jih pošljemo v ječo. Se je pa bistveno spremenila evolucijska ekonomika, saj posilstvo ne pripelje več do otroka, če ženska noče.

Moderni modeli posilstva podobno kažejo, da je zgodovinsko gledano posilstvo v splošnem za povprečnega moškega verjetnostno negativen pojav, da pa je za določene posameznike in v določenih razmerah evolucijsko uspešno. Torej gen za posilstvo lahko ali celo verjetno obstaja, sproži pa se glede na razmere.

Evolucija je pač tako osnoven zakon za vsa živa bitja in tudi ljudi, kot so fizikalni zakoni za fizikalna dogajanja. Če se spremenijo razmere, se spremenijo tudi živa bitja, oz. uspešnejša izpodrinejo manj uspešne. Ljudje smo izredno hitro odzivni ne samo glede obnašanja, ampak tudi glede spremembe genov. Današnji ljudje imamo vrsto poskupin genov in se torej genetsko dokaj razlikujemo med seboj, nekatere spremembe pa so stare le nekaj tisoč ali deset tisoč let. Stvari očitni niso preproste.

Ne glede na to, ali se danes obnašamo v skladu z evolucijskimi principi, ali pa je moderno obnašanje bolj ali manj iracionalno ali naključno, se osnovni »izračuni ekonomike evolucije« niso bistveno spremenili – dolgoročno preživijo uspešnejši in to je gibalno napredka. Če bi večino potomcev imeli manj uspešni, bi to pripeljalo do nazadovanja človeštva. Moderni inteligentni agentni modeli pa kažejo tudi pomembnost medčloveških odnosov - združba primerno sodelujočih posameznikov je uspešnejša kot npr. družba popolnih altruistov ali egoistov.

4. ZAKLJUČKI

Kakor se je že ponovno razživela debata o odločanju pri posilstvu in je mnogo obetala, se zdi, da ni prinesla bistvenih novosti, predvsem s stališča kognitivne informatike oz. inteligentnih sistemov. Zdi se, da je del odločanja pogojen oz. determiniran, pa naj bo to evolucijsko-genetsko ali praktično s stališča interesov eksistence akterja. Del pa ostaja tisto nejasno, čemur v kognitivni informatiki rečemo »inženirska svobodna volja«.

Literatura

AI Magazin (2008). Achieving Human-Level Intelligence through Integrated Systems and Research, AI Magazine, vol. 27, no. 2, summer 2006.

Brooks, M. (2008). I, Rodney Brooks, Am a Robot, IEEE Spectrum Online, <http://www.spectrum.ieee.org/jun08/6307>

Free will, Wikipedia, http://en.wikipedia.org/wiki/Free_will

Gams, M. (2001). Weak intelligence : through the principle and paradox of multiple knowledge, (Advances in computation, vol. 6). Huntington: Nova Science, XIX, 245 str., graf. prikazi, ilustr. ISBN 1-56072-898-1.

Gams, M. (2003). Are minds stronger than Turing machines?. V: ZAJC, Baldomir (ur.). Zbornik ERK

Gams, M. (2009). V postelji z robotom : umetna inteligenca leta 2009. Delo (Ljubl.), 12. sept. 2009, l. 51, št. 211, str. 36.

IEEE Spectrum, september 2009, <http://spectrum.ieee.org/computing/hardware/chip-does-part-of-codecracking-quantum-algorithm>

PRIDI K MENI: O ODLOČANJU, KAIRÓSU IN TRENUTKIH SREČANJA V PSIHOTERAPIJI

Mag. Miran Možina, dr. med., psihiater in psihoterapevt
Slovenski inštitut za psihoterapijo
Trg MDB 5, 1000 Ljubljana, Slovenija, Tel: 386 41 748812
e-mail: miran.mozina@guest.arnes.si

IZVLEČEK

Trenutek odločitve je lahko večji ali manjši kairos, to je ugoden trenutek, ko se porodi nekaj novega. V »zdaj trenutkih« nenadoma začutimo, da se znani vzorci vedenja, znano medosebno okolje oz. odnos lahko spremeni. Ti trenutki zahtevajo osredotočeno pozornost in nas postavijo pred izbiro in odločitev - ali ostati v utečenem vzorcu ali ne. V okviru ti. generičnih načel vključuje sinergetika tudi kairotični vidik in nam omogoča, da lažje razumemo diskontinuirane prehode, preskoke, ki so tipični za kaotične, samoorganizirajoče procese v kompleksnih sistemih. Za ponazoritev procesa odločanja iz vidika sinergetskih generičnih načel in »trenutkov srečanja«, to je dialoških kairotičnih trenutkov, prikažem psihoterapevtsko seanso.

1 ODLOČANJE IN KAIRÓS

Najprej pogledimo na odločanje iz vidika kairosa. Stern (2004) nas namreč opozarja, da je ta vidik v psihoterapiji zanemarjen. Z zgodbami, ki jih pripovedujemo o sebi, svojem življenju (tudi z zgodbami o odločanju), umeščamo naša življenja v resničnost kronosa, to je linijskega časa, ki ga lahko mislimo kot črto, ki teče iz preteklosti v prihodnost. In sedanost iz vidika kronosa lahko umestimo na to črto. Tega smo vajeni. Kronos nam daje vtis objektivnosti. Lahko pa pogledamo nase, so rekli stari Grki, se vzremo tudi iz vidika kairosa. To pa je zanje bil subjektivni čas, ki ima svoje lastne meje in se izmuzne minevanju linijskega časa oz. ga preseže. Čeprav vsebuje tudi preteklost. Je subjektivni oklepaj v kronosu. Kairos je tako subjektivna kot tudi psihološka časovna enota.

V psihologiji in psihoterapiji je dolgo prevladoval (in še prevladuje) pogled na sedanost iz vidika kronosa. Vendar se naša vsakdanja izkušnja – naš subjektivni občutek živosti, dogajanja, kot ga živimo iz trenutka v trenutek lokalno – ne ujema z idejo, da sedanost nima časovne debeline oz. gostote. Doživljanje, ko poslušamo glasbo, gledamo ples ali ko smo z nekom v interakciji, zahteva sedanost, ki traja. Enako velja za življenje na lokalni ravni.

Psihološki trenutki sedanosti morajo po eni strani trajati, da se stvari lahko dogajajo in se morajo po drugi strani odviti v enem, subjektivnem zdaj. Vzemimo primer iz glasbe: kratka glasbena fraza (običajno v okviru dveh do osem sekund) je temeljna enota v procesu doživljanja poslušanja glasbe. Fraza je glasbena analogija za trenutek sedanosti v

vsakdanjem življenju. Glasbeno frazo intuitivno dojamemo kot temeljno enoto z mejami. Njeno trajanje začutimo. In kar je najbolj zanimivo, glasbeno frazo, kot jo slišimo, doživimo v trenutkih, ki so daljši od hipa, hkrati pa teh trenutkov ne doživimo ločeno kot posamezne bite napisanih not. Doživimo jo kot povezano, analogno, tekočo celoto v sedanosti, v zdaj. Ponavadi se v zdaj ne zavedamo minevanja časa, hkrati pa minevanje časa zaznavamo izven zavesti.

Grki so psihološki trenutek sedanosti poimenovali Kairos, zanje je to bil milostiv, dobrotljiv, naklonjen, ugoden trenutek, ko se nekaj porodi. V kairosu se na nov način povežejo prej nepovezani elementi. Nastajanje novih povezav prispeva k temu, da take trenutke doživimo kot posebne in osvežujoče. Eden od izvornih pomenov besede je povezan s pastirji, ki so opazovali zvezde. Ko se noči in se prižigajo zvezde, se zdi, kot da se na nebu dvigajo in nato spet spuščajo. Trenutek, ko zvezda doseže svoj višek in ko se začne spuščati, je kairos.

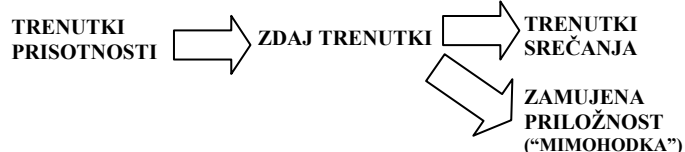
Ena od številnih legend o Kairosu, v kateri nastopa kot poosebljenje sreče oz. srečne priložnosti, pa pravi, da je bil zelo lep mladenič. Vendar je bil v nečem poseben: bil je popolnoma nag, namazan z olivnim oljem, lase pa je imel povezane v čop, ki mu je visel čez sredo čela. Poleg tega je Kairos stalno hitro tekkel po svojih poteh, večkrat tudi mimo ljudi. Mnogi so ga želeli ujeti. Ker pa je bil zaradi olja izmuzljiv, so ga lahko ujeli samo, če so ga zgrabili za čop, ki je bil kot konjski rep sredi glave. Če so bili prehitri, se jim je izognil, če so bili prepočasni, so ga zgrabili za roko ali kakšen drug telesa, ki je bil namazan z oljem in Kairos se je izmuznil. Sporočilo oz. nauk te legende je torej, da morajo ljudje aktivno zaznati in izkoristiti pravi trenutek, da bi uresničili svojo željo in dosegli srečo. Tudi beseda sreča izhaja iz besede srečanje – srečanje želje in tiste priložnosti v realnem svetu, ki omogoča izpolnitev želje. (Možina, 2007)

Da bi se v ugodnem, kairotičnem trenutku srečala naša želja in priložnost v realnem svetu, pa se moramo pogosto ODLOČITI med dvema ali več možnostmi. Torej Kairos velikokrat pomeni odločitev. V določenem trenutku lahko začutimo – zdaj ali nikoli. Odločiti se moram ZDAJ, ali pa bo šla priložnost mimo. Kairos je trenutek edinstvene priložnosti, ki jo lahko ujamemo ali pa se nam izmuzne, saj dogodki zahtevajo dejanje ali so ugodni za neko dejanje. Življenjske okoliščine ali dogodki so taki, da se zavemo, da

moramo ukrepati, nekaj narediti, in tudi tvegati, ZDAJ, da bi spremenili svojo usodo – bodisi za naslednjo minuto ali za celo življenje. Če ne bomo ukrepali, se bo naša usoda prav tako spremenila, a na drugačen način, ker nismo ukrepali, ker nečesa nismo naredili, ker se za nekaj nismo odločili.

Kairos je majhno okno, priložnost v našem bivanju, postajanju, nastajanju. Stern (2004) je oblikoval model treh vrst kairotičnih trenutkov: »trenutki prisotnosti« (»present moments«) lahko prerastejo v »zdaj trenutke« (»now moments«) in ti v »trenutke srečanja« (»moments of meeting«) ali pa le te zamudimo (»mimohodka«) (glej sliko 1).

Slika 1: Prisotnost lahko preraste v trenutke, ko smo osredotočeni tu in zdaj, ti pa v trenutke srečanja



V »zdaj trenutkih« nenadoma začutimo (lahko tudi kot nevarnost), da se znani vzorci vedenja, znano medosebno okolje oz. odnos lahko spremeni. Ti trenutki zahtevajo okrepljeno navzočnost, osredotočeno pozornost in nas postavijo pred izbiro in odločitve - ali ostati v utečenem vzorcu ali ne. In če ne, kaj naj naredimo? Znajdemo se na razpotju. Silijo nas v nerutinski odziv, ki je nov glede na znani vzorec. Nepričakovani in neznani so v svoji obliki in času in se nam zato lahko zazdijo čudni, nenavadni. Pogosto povzročijo zmedo, saj nenadoma ne vemo, kaj se dogaja in kaj narediti. Vezani so na neznano prihodnost in jih lahko občutimo kot zastoj ali kot priložnost. So izrazito osebni, kot nekakšni trenutki resnice. Pogosto jih spremljajo pričakovanja ali strah zaradi nujnosti izbire oz. odločitve. Ne moremo narediti takojšnjega načrta akcije ali najti takojšnje razlage. V teh trenutkih se znajdemo na čistini. Ne moremo se izogniti tveganju.

O izzivu čistine poje pesnik Tone Pavček (1978) takole:

<p>»Na mestu stojim. Sam. Na zagatni čistini. In nisem edini, ki ne vem, kam.</p>	<p>In ne moreš uteči iz sebe, iz svojih dni in ne gesla izreči, ki te osvobodi.</p>
---------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------

<p>Je res vseeno ali na levo ali na desno, naprej ali nazaj, ali stati na mestu, na tej goli čistini dočakati kraj...</p>	<p>Na mestu stojim. Na zagatni čistini. Se prebijem k dolini? Ali tu obležim?«</p>
---------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------

Vse je odprto, golo.
Človek sam pred seboj.
Kot bi stotisoč nožev bodlo
in ni kritja za boj.

Običajno poskušamo razumeti kairos iz vidika ene osebe, a ko ga pogledamo iz vidika dveh, iz dialoškega vidika, se nam razkrije še bolj bogat in pester. Npr. ko smo zaljubljeni, smo posebej dojemljivi za prisotnost druge osebe, ki v nas vzbuja intenzivna čutenja. Ob njej se spontano ojača naša prisotnost v sedanjem trenutku. Zlahka se osredotočamo. In če imamo srečo, naša zaljubljenost lahko preraste v trenutke intimnosti, ko se čas na protisloven način ustavi v intenzivnem zazrtju drugega in sebe v njegovih očeh. To je kairos v dvoje, kairos v trenutkih globokega medosebnega stika, v »trenutkih srečanja«.

Po Sternu bi torej lahko trenutke, ko se odločamo v medosebnem prostoru, razdelili v tri faze:

- pripravljalna faza z občutki neizbežnosti glede odločitve, ki nas čaka;
- faza zmedenosti, ko posameznik ugotovi, da je vstopil v neznan in nepričakovan medosebni prostor;
- faza odločanja, ko se posameznik odloča, ali bo ta trenutek zagrabil ali ne (ali se bo npr. odločil za približevanje ali oddaljevanje od pomembne osebe). Če ga, to vodi v trenutek srečanja. V nasprotnem primeru pa to pomeni zamujeno priložnost oz. "mimohodko".

2 KAIROS IN ODLOČANJE V PSIHOTERAPEVTSKEM SOPOTNIŠKEM DIALOGU

Kairos ima lahko tudi zdravilno moč in je zato pomemben za vse oblike zdravljenj, tudi za psihoterapijo. Je kot izvir sveže, žive studenčnice, ki na čudežen način odplakne tudi najhujše bolečine, je obliž za rane. Je zdravilo, ki pospešuje procese celjenja. Je žarek, ki osvetli temo duševnega trpljenja, vrača smisel, vero in upanje.

Kot obstaja širok spekter kairosov, od zelo redkih in izjemnih (kairos z velikim K), ko se preokrene širok zavoj reke našega življenja, do takih, ki nimajo skoraj nobenih posledic, obstajajo Odločitve z velikim O in tiste z malim o. Odvisno od konkretnih okoliščin in od tega, za kaj se gre, jih različno poimenujemo: »trenutek v času«, »trenutna odločitve«, »trenutek brezčasa oz. večnosti«, »trenutek odločitve« (ko se npr. odločamo na ozadju ambivalence), »odločilni trenutek« (ko npr. pritisnemo na sprožilec fotoaparata), »bistveni trenutek« (npr. ko ujamemo bistvo situacije), »trenutek resnice« (npr. ko se izpovemo), čuden, »zajeban« trenutek (npr. v mladostniškem žargonu) ali »trenutek tu in zdaj« (npr. v psihoterapiji). Obstajajo torej skrajno banalni trenutki sedanjosti (mikro-kairosi in mikro-odločitve), ki skoraj neopazno spreminjajo tok življenja, saj trenutek za trenutekom puščajo majhno, a prepoznavno sled. To je snov, to so delčki, iz katerih je zgrajeno naše neprestano doživljanje in kar je za nas, psihoterapevte, še posebej važno, to so delčki, ki omogočajo terapevtske spremembe.

Psihoterapije brez kairosa in neprestanega odločanja za kairotični pristop ni. Psihoterapevtski proces (kot tudi naše

življenje) obrodi sadove samo, če so naši klienti deležni teh milostnih trenutkov – od tistih z manjšim k do tistih z velikim K, do katerih tudi sem in tja pride in ki pustijo globoko sled v spominu naših klientov, pa tudi nas, terapevtov. Ki ojačajo medsebojno čustveno vez med terapevtom in pacientom. Saj psihoterapija v svojem bistvu pomeni za pacienta iskanje prehodov, utiranje poti v nove življenjske pokrajine, širjenje obzorij, odpiranje novih perspektiv, tveganje novih odločitev. Psihoterapija je kot „sopotništvo“ (Yalom, 2002) in „dialogoterapija“ (Barnes, 2008) plodna avantura samo, če se lahko tako klient kot terapevt odpreta za nepričakovano, novo, za kairoso in če stalno tvegata nove odločitve, ki ohranjajo svežino in enkratnost vsakega terapevtskega srečanja in vsakega terapevtskega procesa. V dialogoterapiji terapevt kot sopotnik za vsakega pacienta odkriva novo terapijo. Vsaka terapija je unikaten izdelek. Za sopotništvo in dialogoterapijo mora terapevt vedno znova, pred vsakim srečanjem s klientom sprejeti zavestno odločitev za izstopanje iz rutin, mehničnega ponavljanja vzorcev, za stalno opuščanje varnega zavetja tehnik, miselnih in vedenjskih navad. Psihoterapija ni „štancanje“, ampak stalna kreacija.

Psihoterapija torej ni enkratna avantura samo za paciente, ampak tudi za terapevte. Če je terapevt odprt za kairoso - in pomemben (če ne bistven) del njegovega usposabljanja naj bi bilo prav večanje občutljivosti za ta vidik psihoterapije – bo s svojim zgledom in ne le z besedami pritegnil pacienta v kraljestvo tega skrivnostnega boga, ki mu je ljuba avantura. Zato se tik pred terapevtsko seanso rad vsaj za nekaj minut ustavim, upočasnim, se usedem na stol, na katerem bo sedel pacient in se osredotočim na svoje dihanje in svojo zaznavo – kaj vidim, slišim in čutim tu in zdaj. Da se odprem za stisk roke ob pozdravu in za prvi očesni stik. Za vljudnostne besede ob pozdravu. Da bom pacienta v skladu z Buberjevim priporočilom zagledal in zaslišal, kot bi se prvič srečal z njim: »Kljub vsem podobnostim, ki jih vsaka življenjska situacija ima, tako kot novorojeni otrok, kaže tudi obraz, ki ga še nismo srečali in ga nikoli več ne bomo. To zahteva od tebe odziv, ki ga nisi mogel vnaprej pripraviti. Ne zahteva ničesar, kar je preteklost. Zahteva prisotnost, odgovornost; zahteva tebe.« (Buber, 1977)

In potem, ko terapevt zapre vrata svoje ordinacije in nanj obesi – NE MOTI – ko izključi telefon in ko se popolnoma posveti svojemu pacientu ali pacientom, ustvarja na nek način svet, posvečen prostor in čas in s tem pogoje za obisk boga Kairosa. Saj se neprestano uglašuje na svojega pacienta, na njegove občutke, ki se iz sekunde v sekundo spreminjajo. Vključi nevidni mikroskop, ki povečuje minimalne gibe, vzgibe, drobnarije v načinu izrečenega, v mimiki... Kot izkušeni plesalec začne prisluškovati klientovim ritmom in intenziteti, crescendom in decrescendom njegovih čutenj, in poskuša ujeti njegov korak, na njegovih frekvencah. Diha v ritmu pacientovega dihanja. Subtilno in diskretno oponaša pacienta, kot bi bil njegovo ogledalo. Da s tem pospeši procese naravnega zrcaljenja, spodbudi aktivnost zrcalnih nevronov (Bauer, 2005), s katerimi nas je v evoluciji opremila narava, da smo

izjemno povečali zmožnosti za medosebni stik. In ko se med terapevtom in pacientom vse bolj razvija skupno gibanje, ki vključuje seveda tudi veliko neuglašanih plesnih korakov, prehitavanja, zaostajanja, nesporazumov, motenj, lahko pride do srečnih naključij, ko se prisotnost obeh istočasno ojača v trenutkih srečanja, v »kairoso à deux«. Takrat se zagledata v obeh drug drugega in se uzreta skozi oči drugega brez mask, kot osebi, kot človeka na istem čolnu, kot sopotnika v veliki avanturi nepredvidljivega bivanja. Dve zavesti povezani v kairoso intersubjektivne, medosebne zavesti.

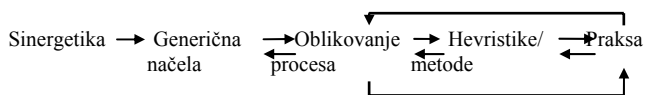
3 KAKO LAHKO PSIHOTERAPEVT Z UPOŠTEVANJEM SINERGETIČNIH GENERIČNIH NAČEL POMAGA PACIENTU K POMEMBNIM ODLOČITVAM

Sinergetika definira psihoterapijo kot destabilizacijo disfunkcionalnih in porajanje novih odnosnih vzorcev, novih vzorcev vedenja, čustvovanja in mišljenja v okviru stabilnega, zaupnega in varnega odnosa med terapevtom in pacientom. Sinergetika, teorije nelinearnih sistemov in teorija kaosa omogočajo matematično formalizacijo in simulacijo telesnih, psiholoških in socialnih samoorganizirajočih procesov v kompleksnih sistemih, kar na široko odpira vrata raziskovanju (Haken, Schiepek, 2006). Za psihoterapevtsko prakso pa sinergetika ponuja osem ti. generičnih načel (Schiepek et al, 2005), s pomočjo katerih terapevt lažje razume diskontinuirane, samoorganizirane prehode, seveda tudi nepričakovane lastne in pacientove odločitve, do katerih pride v obliki nenadnih preskokov (glej sliko 2):

1. *Ustvarjanje pogojev za stabilnost* (terapevtski postopki za povečevanje varnosti, zaupanja v odnosu med terapevtom in pacientom, podpiranje pacientovega občutka lastne vrednosti).
2. *Prepoznavanje vzorcev relevantnega sistema* (prepoznavanje tistih disfunkcionalnih, problemskih odnosnih vzorcev, vzorcev mišljenja, čustvovanja in vedenja, ki jih pacient želi spremeniti, ob hkratnem prepoznavanju vzorcev rešitve; opisovanje in analiza problemskih vzorcev oz. sistemskih procesov, kolikor je le-to potrebno za definiranje terapevtskih ciljev in odkrivanje rešitev).
3. *Upoštevanje smiselnosti* (za pacienta morajo biti naloge in koraki, ki omogočajo napredovanje in vodijo k ciljem, ves čas smiselni zanj; terapevt daje priznanja za pacientov proces spreminjanja in upošteva pacientov življenjski slog in osebne razvojne naloge tudi če se to ne ujema z njegovim življenjskim slogom; cilji morajo biti po meri pacienta, v prvi vrsti smiselni zanj in ne terapevta).
4. *Prepoznavanje parametrov nadzora / Omogočanje energiziranja* (ustvarjanje pogojev, ki spodbujajo pacientovo motivacijo; prebujanje pacientovih virov; upoštevanje pacientovih ciljev in skrbi).

5. *Destabiliziranje / Krepitev nihanj* (eksperimenti in paradoksi; prekinjanje disfunkcionalnih vzorcev odnosov, mišljenja, čustvovanja in vedenja; vnašanje novih razlikovanj; izjeme – zavestne in spontane; neobičajno, novo vedenje itd.)
6. *Upoštevanje »kairosa«/resonance/sinhronizacije* (časovna usklajenost in koordinacija med terapevtskimi postopki, terapevtskim komunikacijskim stilom ter pacientovimi psihičnimi in socialnimi procesi / ritmi; skupno gibanje, skupni »ples«).
7. *Prekinjanje simetrije* (s „simetrijo“ je tu mišljeno stanje, v katerem so poleg utečenih v enaki meri možni oz. prisotni različni (novi) odnosni vzorci in vzorci mišljenja, čustvovanja, vedenja; za prekinjanje simetrije terapevt usmerja pacienta na njegove cilje, pričakovanja in uresničevanje načrtovanih novih vzorcev).
8. *Ponovno stabiliziranje* (ukrepi za stabiliziranje in integracijo novih miselnih, čustvenih, vedenjskih vzorcev in odnosnih vzorcev).

Slika 2 (po Schiepek et al, 2005): Znanje o generičnih načelih samoorganizirajočih se procesov nam lahko omogoča, da terapevtsko ravnanje organiziramo, poenostavimo in utemeljimo.



Pomembne odločitve (v obliki diskontinuiranih prehodov) so pogosto povezane z destabilizacijami, ki vzbudijo strahove. V fazah kritične nestabilnosti pride do izrazitih fluktuacij, ker dotakratni vzorci za pacienta izgubijo svojo „privlačnost“ ali mu postanejo celo neprijetni, in ker lahko istočasno pacient le težko predvideva, kako bo potekal nadaljnji razvoj. Te strahove in vznemirjenosti vsled negotovosti kot tudi z njimi povezana „vzratna gibanja“ (npr. ti. „recidivi“) k znanim vedenjskim vzorcem večinoma tolmačimo kot „odpore“, vendar pa jih nikakor ne smemo interpretirati kot pacientovo pomanjkljivo pripravljenost za sodelovanje. „Recidivi“ in „napake“ v fazah kritične nestabilnosti so še kako smiselni in ljudje ob „recidivih“ in „napakah“ potrebujejo od terapevtov in drugih ljudi podporo, tudi preko drža, ki kaže „prijateljski odnos do napak“.

V fazah nestabilnosti se tudi poveča pacientova občutljivost za minimalne znake (lahko je npr. to že ena terapevtova beseda ali droben neverbalni znak), ki ga lahko spodbudijo k pomembnim življenjskim odločitvam in s tem močno oblikujejo nadaljnji razvoj. V teoriji kaosa je to poznano kot „metuljev fenomen“, kjer zamah metuljevih kril na Japonskem lahko sproži serije in kaskade v kompleksnem sistemu vremena, tako da se razvije orkan v Ameriki. V teh fazah lahko pacientom pomagamo pri zahtevnih odločitvah tudi z metaforami, slikami in verbalnimi sugestijami na različnih ravneh zavesti in na različnih senzornih kanalih. Tako npr. Keeney (1987) v svojem kibernetičnem modelu terapevtske spremembe, ki se mi vedno znova obnese,

predlaga, da lahko terapevt pomaga pacientu pri pomembnih odločitvah tako, da najprej reče JA, NAREDI to in to in argumentira z razlogi za, nato pa nadaljuje z NE NAREDI tega in argumentira z razlogi proti (lahko seveda tudi začne z ne naredi in nadaljuje z naredi; važno pa je, da ne reče JA in NE istočasno, ampak da ujame dober timing in preseneti pacienta, s čimer mu pomaga izstopiti iz običajnih vzorcev v kreativni proces; ko pacient že misli, da se je terapevt opredelil za ali proti, ga terapevt preseneti in odpre še drugo možnost in morda še kakšno), nato pa pove zgodbo ali zgodbe, ki kot Rorschach omogočijo pacientu, da projicira vanje tisto, kar je zanj pomembno, da razširi svoje asociacijsko polje, spodbudi svoj ustvarjalni proces, kar mu pomaga pri odločanju. V nadaljevanju bom to prikazal na konkretnem primeru iz svoje prakse.

4 PRIDI K MENI: PRIKAZ PRIMERA ODLOČANJA NA PSIHOTERAPEVTSKI SEANSI

Za ponazoritev procesa odločanja iz vidika kibernetike, sinergetskih generičnih načel in kairotičnih trenutkov srečanja bom prikazal psihoterapevtsko seanso, ki sem jo imel pred leti z 39-letnim Janezom, dobro stoječim obrtnikom z nekaj zaposlenimi. Pred tem srečanjem je hodil k meni na pogovore okoli osem let v intervalih. Eden glavnih fokusov terapije so bili njegovi odnosi z ženskami. V obdobju osmih let, ko je hodil k meni, je imel resnejše odnose s petimi ženskami. Ko sem ga spoznal, je po šestih letih pristal, kot je sam rekel, v „dolgočasnem“ odnosu z dve leti mlajšo parterko. Ni več čutil pravega naboja in privlačnosti, občasno jo je tudi na skrivaj varal, vendar tudi ni mogel zaključiti odnosa, čeprav sta to že dvakrat poizkusila. Potem se je prvič v življenju zares močno zaljubil v poročeno žensko, ki ga je znala ravznemati, a ga je hkrati držala na distanci, tako da je po nekaj mesecih obupal. Mu je pa ta zaljubljenost pomagala, da je zaključil »dolgočasno« zvezo. V naslednjem resnejšem odnosu je po nekaj mesecih ugotovil, da ga je varala in je po odkritju njene dvojne igre odnos takoj prekinil. Nadaljeval je z vdovo, ki je imela enajst let staro hčerko, je pa živela v več kot sto kilometrov oddaljenem kraju. Ta vikend odnos se je izpel v okoli dveh letih. V času seanse, ki jo bom opisal, je bil že leto in pol v odnosu z deset let mlajšo Lidijo, frizerko, ki je tudi živela v več kot sto kilometrov oddaljenem kraju in s katero se je tudi v glavnem videval preko vikendov.

STABILIZACIJA TERAPEVTSKEGA ODNOSA, UTRJEVANJE ZAUPANJA IN VARNOSTI: Ko se je vsedel, so bile ob nekaj globokih vzdihih njegove prve besede, da je potr in utrujen in tudi zgledal je brezvoljen, saj se je kar pogreznil v stol, njegov glas pa je bil monoton, govor upočasnen. Ko je poročal, kaj je novega, odkar se nisva videla, se je izogibal temi Lidije, zato sem ga po nekaj minutah direktno vprašal. Povedal je, da se še naprej vleče to, da tehtata, ali naj se preseli k njemu v njegovo hišo ali ne. To je bila ena glavnih tem najinih pogovorov zadnje mesece in na različne načine sem ga spodbujal, da ne bo mogel nikoli vedeti, ali bi z Lidijo lahko živela skupaj, če

tega ne bosta poskusila. Vikend odnos je bil dober okvir na začetku, zdaj pa že lep čas ne več. Ona je vse bolj pritiskala, on pa je vse bolj omahoval.

PREPOZNAVANJE DISFUNKCIONALNIH, PROBLEMSKIH VZORCEV IN VZORCEV REŠITVE: V vseh odnosih z ženskami, tudi z Lidijo, je doživljal močno ambivalenco ("Rad bi imel družino AMPAK nisem siguran, če je ona prava..."), v enem delu sebe je bil oklevajoč, ziheraški in pretirano kritičen. Po eni strani je bil v odnosu lahko nežen, igriv in včasih celo preveč gentelmski, po drugi strani pa je v posebnem zvezku vodil »črno kroniko«, kjer je skrbno beležil dogodke, v katerih so se njegove partnerke pokazale kot problematične in v katerih se je omajalo njegovo zaupanje, da bi bile lahko primerne za njegov družinski projekt.

Kot običajno je tudi od najinega zadnjega srečanja pisal črno kroniko, ki jo je prinesel s seboj. Najbolj ga je motil njen "dramatični" talent, kot je sam rekel. V izražanju čustev, tako pozitivnih kot negativnih, je bila precej bolj neposredna kot on, prav tako je bila precej bolj nepremišljena v dejanjih (npr. pri seksu ne bi pazila, bi rada čimprej zanosila, saj si je zelo želela otroka, on pa niti pod razno ne bi hotel imeti otroka z njo, dokler ne bi bil res siguran, da je TA PRAVA). In vedno bolj pogosto sta se zapletala v znano stopnjevanje po principu »želve in nevihte«. Bolj kot je ona grmela in bliskala, bolj se je on zapiral v svoj oklep negativnih misli. In bolj kot se je on zapiral, bolj je bila ona negotova glede njunega odnosa in prihodnosti, tako da mu je delala vse več »scen«, v katerih je tudi z nedvornim dramskim talentom preverjala, ali jo sploh še ima rad. In ko ga je enkrat v navalu jeze celo udarila, je bil zanj to skoraj kronski dokaz, da mora biti s tako žensko še bolj previden. Ali naj se poroči s tako, ki ga bo tepla?

Po drugi strani je že od samega začetka najinih srečanj sprejel moje vabilo, da odpre tudi »belo kroniko«, v kateri se je nabralo marsikaj pozitivnega v odnosu z Lidijo. Na vsakem srečanju sem ga vprašal, s čim je zadovoljen, kaj so viri in rešitve, ki jih odkriva. Tako je zabeležil konkretne opise raznih dogodkov in mi o njih redno poročal: sproščeno vzdušje na dopustih, veselje nad potovanji, znala sta biti razigrana, otroško razposajena, njena bolj odprta čustvenost, spontanost se je večkrat dobro dopolnjevala z njegovo večjo obvladanostjo, previdnostjo, razvijal je sposobnost za konstruktivni konflikt, ljubeči odmik ob soočanjih, dobro ujemanje v spolnosti, skupna želja po otroku in družini, itn.

ENERGIZIRANJE, SPODBUJANJE PACIENTOVE MOTIVACIJE: Nadaljeval sem tako, da sem mu povedal zgodbo o pacientu, s katerim sem delal pred leti zaradi hude krize in depresivne reakcije po drugi ločitvi. Ta pacient se je pred desetimi leti ločil in zapustil ženo z dvema otrokoma, starima sedem in pet let. Zgodila se mu je namreč »velika« ljubezen. Strastno se je zaljubil v osemnajst let mlajšo žensko. Potem ko se je z njo drugič poročil in sta po desetih letih skupnega življenja načrtovala otroka, pa se je nenadoma zaljubila v drugega moškega in odšla. Glavno

sporočilo zgodbe, ki sem mu ga želel posredovati, je bilo: "Ljubiti pomeni TVEGATI! V ljubezni ni garancije!"

SMISELNOST: Nato sem nadaljeval s povzetkom njegovih ljubezenskih odnosov, tako da sem spodbujajoče poudarjal, kako se je v vsakem nečesa naučil, napredoval in osebno zorel. Od prve resne zveze, kjer sta bila oba s partnerko preveč prilagojena, je doživel neke vrste iniciacijo, ko se je prvič močno zaljubil, vendar je bilo to žal enostransko. V naslednjih dveh zvezah se je postopno še bolj odpiral in predajal, tako da je v odnosu z Lidijo doživljal največ pozitivnega od vseh dosedanjih zvez. Hkrati pa se je v odnosu z njo kot tudi v vseh odnosih z ženskami srečeval tudi z ambivalenco – po eni strani bi se rad vezal in osnoval družino, po drugi strani pa je bil nezaupljiv, saj se je stalno spraševal, ali ga ne čaka nekje še boljša družica. Želel si je gotovosti, garancije kot pri nakupu avtomobila ali stroja za svojo obrt. Leta pa so minevala in tako je hkrati v njem rasla zavest, da zna ob takem »garancijskem« pristopu ostati praznih rok, sam, brez otrok in družine.

DESTABILIZACIJA NA OSNOVI KIBERNETSKEGA MODELA SPREMEMBE »NAREDI – NE NAREDI – ZGODBE« IN ZDAJ TRENUTKI: Ne vem točno, kdaj sem se v toku pogovora spomnil na eno od Yalomovih (2002) priporočil, da se pogosto obnese, če terapevt spodbudi pacienta, da naj na sami seansi tudi kaj novega naredi, poizkusi ne samo govori. Da bi ga spodbujal, da naj se pogovori z Lidijo in ji predlaga poskus skupnega življenja za določeno obdobje, se mi nenadoma ni več zdelo plodno, saj sem mu to že večkrat predlagal, pa je vedno znova okleval. Tudi midva sva se glede pat položaja, v katerem se je znašel, že nekaj časa vrtela v začaranem krogu.

Tako sem se odločil, da bom tvegaj in mu predlagal, da naredi korak iz svojega vzorca cincanja pred mano, na sami seansi. Dobra ideja, vendar kako jo izpeljati z dobrim timingom, v dobrem stiku z njim, v skupnem plesu. Naredil sem nekaj vdihov in izdihov v ritmu njegovega dihanja, medtem ko mi je nekaj razlagal. Skupno dihanje je eden od mojih običajnih načinov, kako intenziviram svojo prisotnost, ko poglobim svoj hipnotski trans z odprtimi očmi, da lažje odkrijem trenutek, ko izstopim na čistino. Nenadoma sem se nagnil naprej in s pomočjo svojega dramatičnega talenta rekel s posebnimi poudarki in pomenljivimi pavzami: "Janez, mislim, da je napočil trenutek odločitve... Nimava več kaj odlašati... Poklical jo boš ZDAJ.... (presenečeno me je pogledal in otrpnil) in ji rekel: 'Odločil sem se – pridi živet k meni...'" »Kako to misliš?« me je vprašal s široko odprtimi očmi. »Tako kot sem rekel!« Nastala je tišina, da se je slišalo le najino dihanje. Tega ni pričakoval, napravil sem nekaj novega, svežega, nepričakovanega. Nato sem malo zmehčal: »Ali pa ji boš rekel: 'Narediva pavzo!« Še malo sem počakal in dodal: »Ali pa bosta imela navaden 'površinski klepet'.« Še vedno je bil brez besed. "Kaj lahko izgubiš?... Nič. V zadnjih mesecih se odnos z Lidijo postopno slabša... Če bosta tako nadaljevala, se bosta razšla... Ujela sta se v začarani krog – bolj kot ti oklevaš, več drame iz njene strani. Več Lidijine drame, bolj te je

strah, bolj oklevaš itn... Pravzaprav s tem, da jo zdaj pokličeš, ne moreš nič več izgubiti, saj si na sigurni poti h koncu vajinega odnosa. Najhuje, kar se lahko zgodi, če jo pokličeš, da bo še slabše, kar bo pa itak. Lahko pa se ti oz. vama odpre na bolje. Hkrati doživljaš v odnosu z njo največ pozitivnega do zdaj...« Našel sem nekaj konkretnih dobrih dogodkov, ki mi jih je opisal na prejšnjih seansah (kako se dobro ujameta v spolnosti, na dopustih, kako znata biti razigrana, otroško razposajena itn.). Leta se nabirajo... Vzemi telefon v roke in pokliči. Prepusti se trenutku, občutku. Ko boš zaslišal njen glas, ti bo že nekaj reklo, kako naprej. Saj lahko samo malo poklepetata. Ni treba, da domišliš, kako bo. Prepusti se. Zaplavaj... Kaj ti bo prinesla prihodnost, ne veš, lahko pa boš začutil tu in zdaj, ali je v tebi še dovolj upanja in pozitivnih občutkov za to, da jo povabiš k sebi.«

Da bi mu dal še nekaj časa, ki ga je rabil za »usodno« odločitev, sem mu povedal še zgodbo o filmu Himalaya. Pripovedovanje zgodbe je moj priljubljen indirektni način terapevtske komunikacije. V filmu oče, vaški poglavar, po tragični smrti enega sina, vodiča karavan, povabi drugega, lamo, ki živi v dnevu hoda oddaljenem svetišču, da naj mu priskoči na pomoč pri vodenju karavane čez visok prelaz. Od uspeha te poti je odvisno preživetje vaše skupnosti. Ta najprej očeta zavrne, nato pa spremeni odločitev in se nepričakovano pojavi na vratih očetove hiše. Ko ga oče prijetno presenečen vpraša, zakaj si je premislil, mu sin odvrne: »Vprašal sem svojega predstojnika za nasvet. In mi je odvrnil: 'Pot Bude je, da med dvema možnostima vedno izbere težjo.'«

Da bi mu olajšal globok stik s seboj, sem vdihoval v ritmu njegovih vdihov, pripovedovanje zgodbe pa sem povezal z njegovimi izdihmi. Poslušal me je negiben, s pogledom, fokusiranim na meni, vendar je še vedno okleval. V sebi sem se že sprijaznil, da ga morda moja ideja ne bo premaknila, vendar se sam nisem hotel povleči nazaj, vzdrževal sem napetost in zaupal, da se bo sam odločil po svoji meri. V čakanju na njegov jasni da ali ne, sem nato podelil še svojo ljubezensko izkušnjo. V letih najinega srečevanja sem mu že večkrat razkril kakšen del svojih osebnih izkušenj, kar je vedno imelo dober učinek in zdaj se mi je zdel spet primeren trenutek. Ko sem podelil svojo zgodbo, sem dodal, da se v mojem iskanju ljubezni kaže obratna značilnost kot pri njem – tvegat sem znal, večkrat celo preveč. In vendar kljub temu da sem tvegat, v trenutku najinega pogovora nisem živel v izpolnjujočem odnosu, ampak sem bil prav tako v krizi. Sporočiti sem mu hotel, da sva glede ljubezni sopotnika, da tudi sam nisem našel, ampak da še iščem. Na lastnem primeru sem mu pokazal, da ne morem biti drugega kot skromen in da ga lahko razumem in sprejem tudi, če se NE bo odločil za klic.

»Ali veš, kaj je rekel Oscar Wilde?« sem ga vprašal. »Ne,« je odgovoril. »Za ljubezen do mladega, ekscentričnega plemiča je plačal najvišjo ceno, svojo svobodo, zdravje in na koncu življenje, saj je moral zaradi vplivnega in maščevalnega plemičevega očeta, ki je zaradi homoseksualnosti proti Wildu vložil tožbo na sodišču, po

krivični obsodbi za dve leti na prisilno delo v ječo. Tam je izgubil svoje zdravje, premoženje, zapustila pa ga je tudi žena z otroki, ki niso smeli imeti več stika z njim. Zaradi tega je imel tako hude občutke krivde, da se je javno odrekel razmerju s plemičem. Pa vendar se je po vrnitvi iz ječe pred svojo smrtjo spet za nekaj časa vrnil k razvajenemu in sadističnemu ljubimcu. Ko sem vse to pred kratkim gledal v biografskem filmu o Wildu, pa me je najbolj pretresla zaključna Wildova misel v zadnjem kadru filma: 'V življenju sta možni dve vrsti ljubezenske nesreče: ena, če se ti želja ne izpolni in druga, če se ti izpolni.'« Po kratkem premoru sem dodal: »Eno je v ljubezni sigurno, na to lahko dobiš ali dobimo garancijo: ni receptov, ampak vedno znova zabredemo v njene skrivnostne vode. In ljubezen ni muc, ampak je tiger. Ne moremo je udomačiti.«

Janez je nato odvrnil: »Prav. Jo bom poklical, vendar pojdi ven iz sobe. Te bom poklical po pogovoru.« Kakšnih dvajset minut sem nato čakal na hodniku.

TRENTKI SREČANJA, NELINEAREN, DISKONTINUIRAN PRESKOK V NOV VZOREC: Ko me je poklical nazaj, je sijal od zadovoljstva. Opisal mi je pogovor, ki je na njegovo veliko presenečenje potekal spontano, tekoče, tudi z veliko humorja in igrivosti. Takoj ko jo je zaslišal, mu je bilo prijetno in pogovor je stekel. Nekako naravno je prišel trenutek, ko jo je brez posebnega napora povabil: »Odločil sem se, da te povabim k sebi. Pridi živet k meni!« Bila je prijetno presenečena. Dogovorila sta se, da bosta čez vikend nadaljevala s pogovorom. Vendar ni bilo samo besedno poročilo tisto, ki me je prepričalo, da je prišlo do »kvantnega skoka«, še bolj prepričljiva je bila sprememba vzdušja: Janez je imel jasen obraz, bil je navihan, igriv, nasmejan, zadovoljen in hvaležen. Ko je odhajal, me je spontano objel. Tudi sam sem se počutil živ, prisoten, vesel, da sva doživela močan trenutek srečanja. Čez dva meseca se je Lidija preselila k njemu. Zadnji stik z njim sem imel leto in pol po opisanem srečanju, ko sta še vedno živela skupaj.

LITERATURA

- Barnes, G. (2008). An Introduction to Dialogotherapy. *Kairos – Slovenska revija za psihoterapijo*, letn. 2, št. 3/4, str. 25-60.
- Bauer, J. (2005). Warum ich fühle, was du fühlst: Intuitive Kommunikation und das Geheimnis der Spiegelneurone. Hamburg: Hoffmann und Campe.
- Buber, M. (1977). Ja i ti. Beograd: Vuk Karadžić.
- Haken, H., Schiepek, G. (2006). Synergetik in der Psychologie. Selbstorganisation verstehen und gestalten. Göttingen: Hogrefe.
- Keeney, B.P. (1987). Aesthetik des Wandels. Hamburg: Isko-Press.
- Yalom, I. (2002). The Gift of Therapy: An Open Letter to a New Generation of Therapists and Their Patients. New York: HarperCollins.
- Možina, M. (2007). O Kairosu. *Kairos – Slovenska revija za psihoterapijo*, letn. 1, št. 1/2, str. 3-5.
- Pavček, T. (1978). Pesmi. Ljubljana: DZS.
- Schiepek, G. (1999). *Die Grundlagen der Systemischen Therapie*. Göttingen: Vandenhoeck & Ruprecht.
- Schiepek, G., Ludwig-Becker, F., Helde, A., Jagdfeld, F., Petzold, E.R., Kröger, F. (2005). Sinergetika za prakso. V: Bohak, J., Možina, M. (ur.). Četrtri študijski dnevi Slovenske krovnice zveze za psihoterapijo, 3. in 4. junij 2005, Rogla. *Sodobni tokovi v psihoterapiji : od patogeneze k salutogenezi : [zbornik prispevkov]*. Maribor: Slovenska krovna zveza za psihoterapijo: 25-33.
- Stern, D.N. (2004). *The present moment in psychotherapy and everyday life*. W.W. Norton & Company: New York, London.

OSEBNOSTNI IN KONTEKSTNI DEJAVNIKI TER PARADIGME ODLOČANJA

Simona Tancig
Pedagoška fakulteta
Univerza v Ljubljani
Kardeljeva pl. 16, 1000 Ljubljana, Slovenia
Tel: +386 1 5892 200; fax: +386 1 5892 233
e-mail: simona.tancig@guest.arnes.si

POVZETEK

V prispevku so predstavljeni nekateri pogledi na odločanje skozi prizmo vloge čustev, motivov in zavesti (nezavedno, zavedno). Prispevek obravnava tudi odločanje ekspertov, njihove sposobnosti in kontekstualne dejavnike odločanja. Obe tematiki se povezuje s paradigmatškimi pogledi na odločanje. Predstavljene so 3 temeljne paradigme odločanja: formalno-empiricistična, racionalistična in naturalistična.

1 POGLEDI NA ODLOČANJE: VLOGA ČUSTEV, MOTIVOV IN ZAVESTI

Človekovo odločanje je zelo prilagodljivo in nanj sistematično vpliva množica dejavnikov, ki jih normative teorije izbire ne upoštevajo. Človekove preference so v precejšnji meri pod vplivom percepcije rizika in vrednosti, težnja izogniti se odločitvenemu konfliktu, različnih identitetnih vlog posameznika, mentalnih modelov kot tudi emocij in motivov ali potreb.

Novejše raziskave odločanja skušajo razjasniti osebnostne kot tudi kontekstne dejavnike odločanja. Med smerni, ki jih zastopajo nekateri raziskovalci, je vsekakor večja osredotočenost na vlogo čustev kot pomembne silnice, ki vpliva na odločanje (Hsee, Kunreuther, 2000). Kot pomembno področje raziskav je tudi pojav konsistentne nepovezanosti med *odločitveno koristjo*, ki jo implicira neka odločitev, ter *izkušnjo koristnosti*, t.j. *hedonistično izkušnjo*, ki jo dejansko uresniči ta izbrana opcija. Podroben opis procesa odločanja zahteva vključitev teh in še nekaterih drugih dejavnikov. Uspešen deskriptivni model mora upoštevati normative kriterije, prav tako pa mora vključiti vrsto psihičnih procesov, ki lahko vplivajo na izbiro oz. odločitev. Proces izpopolnjevanja deskriptivnih teorij se pospešeno razvija in rezultati kažejo, da nastajajo teorije oz. modeli, ki so precej drugačni od elegantnega in optimalnega normativnega postopka. Ta odmik od normativnih teorij pa ne pomeni nujno, da sta njihov vpliv in moč oslABLJENA. Sicer pa so te teorije že same po sebi empirični projekt, ker predstavljajo odločanje, ki ga na nek način pojmuje kot idealno, hkrati pa se z razumevanjem, kako v resnici poteka

proces odločanja, odpirajo nove možnosti oblikovanja postopkov odločanja, ki vodijo odločevalce ob upoštevanju vseh dejavnikov, da se približajo nekemu idealnemu postopku.

Glede na to, da je odločanja tesno povezano z različnimi vidiki vsakdanjega življenja, poklicne dejavnosti, politike, ekonomije ipd., dajejo te raziskave pomembne ugotovitve, ki pojasnjujejo, zakaj ljudje mislijo, delujejo in se odločajo, kot se.

Tako je bila narejena cela vrsta raziskav, ki skušajo odkriti vpliv emocij in motivov tudi v povezavi z različnimi nivoji zavestnosti (zavedno in nezavedno).

Ugotovljeno je bilo, da so nekatere odločitve odvisne celo od trenutnega emocionalnega stanja, ki pogojuje trenutno evalvacijo opcij, in tako tudi izbiro. To so t.i. *miopične* (kratkovidne) *odločitve*, ki se pojavijo, ko se za trenutek sproži nek miselni okvir, ki poveča vrednost in željo, ki morda ne odražata bolj globalne preference odločevalca. Emocije vplivajo na asociacije ali predstave, ki pridejo v zavest pri odločanju. Ker se predstave obravnavajo (pregledajo) hitro in brez napora, so raziskovalci za tako čustveno ocenjevanje opcij, ki pogojuje odločitve predlagali izraz *afektivna heuristika* (Slovic, Finucane, Peters, MacGregor, 2002). Ugotovljeno je bilo, da celo t.i. *anticipatorne emocije* (npr. predvidevanje emocionalne reakcije v rizični situaciji) lahko vplivajo na kognitivno oceno situacije odločanja in tako vplivajo na izbiro (Loewenstein et al., 2001).

Podobno tudi potrebe in motivi vplivajo na rezoniranje in odločanje. Pogosto se zaradi tega pojavijo nekonsistentnosti v odločanju, ker osebe ne prepoznajo, da so bile njihove preference spremenjene s trenutno induciranimi sentimentami, ki so pogojeni s situacijo. Kaže se, da celo v primeru, ko se osebe zavedajo svoje »ujetosti« v prehodne potrebe, motive ali čustva, se ne bodo rešile njihovega vpliva.

V nedavnih raziskavah so se osredotočili na identificiranje principov, ki opisujejo vmesnik med motivacijo in kognicijo ter njegov vpliv na mišljenje, rezoniranje in presojanje (Higgins, Molden, 2003).

Obstajata tudi dva razreda motivacijskega vpliva. Prvi vključuje željo po *doseganju* določenega *tipa rezultata* v presojanju in odločanju; drugi vključuje človekovo željo

uporabiti določen tip strategij pri presojanju. Sicer pa motivacija vpliva na mišljenje in rezoniranje tako, da usmerja človekove kognitivne procese (npr. priklic in iskanje informacij) k doseganju želenih ciljev.

Pomembne so tudi ugotovitve, obsežnih raziskav, ki temeljijo na sodobnih konceptih zavesti in pojasnjujejo kako nezavedna (subliminalna) aktivacija in nezavedni procesi regulacije čustev vplivajo na presojanje in odločanje. Pri omenjenih raziskavah je pomembna tudi vloga nevroznanosti, ki skuša s slikanjem možganske aktivnosti identificirati možganska področja, ki so vključena v omenjene procese.

Vrsta raziskav se je osredotočala na implicitni in eksplicitni vpliv emocij na presojanje in odločanje na različnih področjih., npr. v politiki, pravosodju, vsakodnevnem življenju ipd.

Najbolj široko sprejeti modeli presojanja in odločanja v različnih vedah od kognitivne znanosti, psihologije do političnih ved so skoraj-racionalni modeli. Ti modeli predvidevajo, da so sicer ljudje v precejšnji meri racionalni, vendar lahko odstopijo od racionalne izbire, npr. uporabijo bližnjice in ne pretehtajo vsako posamezno opcijo glede koristi in vrednosti ter verjetnosti, do z dosegajo cilj. Nekateri teoretiki, ki so razvili skoraj-racionalne modele, domnevajo, da osebe bolj izčrpno pregledajo opcije in uporabijo več kognitivnih procesov za probleme, ki so za njih pomembni, pri manj pomembnih glede posledic pa uporabljajo kognitivne bližnjice (Gigerenzer, Goldstein, 1996).

Čeprav nekateri raziskovalci (Mellers, 2000) poudarjajo večjo vlogo čustev v odločanju, še vedno prevladujejo modeli presojanja, odločanja in reševanja problemov v kognitivnih psihologiji, političnih vedah in ekonomiji, ki postavljajo kognicijo nad čustvi v optimalnih in suboptimalnih presojah in odločitvah.

Tako kot mnogi raziskovalci čustev (Plutchik, 1980; Panksepp, 2005) tudi Westen (1994) zagovarja veliko bolj pomembno vlogo emocij in regulacije emocij v presojanju in odločanju. Sodobni pogledi na motivacijo poudarjajo mehanizme »približevanja« in »izogibanja«, ki so pogojeni s pozitivnimi in negativnimi čustvi (Carver, 2001), Westen (1985) je predlagal, da je komponenta »vrednosti« v teoriji pričakovane vrednosti primarno čustvena, če se osebe implicitno ali eksplicitno odzivajo s približevanjem ali izogibanjem na različne alternativne opcije na osnovi priklicanih, asociiranih ali anticipiranih čustev. To pomeni, da se ljudje pri presojanju in odločanju približujejo ali odmikajo alternativnim opcijam glede na emocionalne posledice, čeprav so te posledice lahko mediirane izrazito kognitivno. S tem v zvezi je pomembna Westenova (1985) trditev, da ti procesi izogibanja in približevanja, ki temeljijo na čustvih in ki motivirajo »racionalno« odločanje, pogosto izkrivijo sklepe, do katerih osebe pridejo, podatke, ki jih pregledajo, in načine, kako jih tehtajo.

V poskusih modeliranja kompleksnih procesov kognicije in regulacije emocij, so se Westen in sodelavci

(Westen, Gabbard, 2002) oprli na psihoanalitični konstrukt »kompromisna formacija«, ki se nanaša na proces, v katerem oseba generira kompromisne rešitve pri soočanju z mnogokratnimi, pogosto tekmujočimi motivi. V kognitivni znanosti bi temu našli analogijo v *konekcionističnih mrežah*, kjer se odvijajo procesi uravnoteženja k rešitvi, ki optimizira stopnjo ujemanja z obstoječimi podatki. Westen in sodelavci so namesto zaporednega obravnavanja (motivi prihajajo v zavest posamično in vplivajo na vedenje) predpostavljali, da motivi lahko delujejo simultano in vplivajo na mišljenje ali dejavnosti izven zavesti, kar vodi h kompromisnim rešitvam. Predlagali so model regulacije emocij, presojanja in odločanja, ki vključuje *konekcionistično osredotočenost* na omejitve, ki jih narekujejo podatki, kot tudi omejitve, ki jih narekujejo hedonistične posledice različnih rešitev. Po tem modelu presojanje o emocionalno pomembnih zadevah (ki v vsakdanjem življenju vključujejo večino presoj in odločitev) odražajo sočasno zadovoljitev dveh množic omejitev: *kognitivne omejitve*, ki jih narekujejo podatki in njihova logična vsebovanost, ter *emocionalne omejitve*, ki jih narekujejo emocionalne asociacije in anticipirane emocije.

2 PARADIGME ODLOČANJA: OD KLASIČNE DO NATURALISTIČNE

Tradicionalne raziskave odločanja niso pomembno prispevale k razumevanju profesionalnega presojanja in odločanja v praksi. To je spodbudilo raziskovalce k spremembi prevladujoče normativne klasične paradigme, imenovane *presojanje in odločanje*, s številnimi modeli in pristopi, ki so nastali v tem okviru (Hoffman, 1995). Nov pogled na odločanje so sprožili raziskovalci, ki so raziskovali presojanje in odločanje v različnih poklicih in v naravnih okoljih. Pomembno je bilo zgodnje odkritje raziskave, da poklicno oz. ekspertno presojanje v največji meri temelji na procesih ocene situacije (ugotavljanje narave problema) in nato evalvacije posamezne opcije z mentalno simulacijo, ki je pripeljala do zadovoljivega odgovora ali akcije.

Kvalitativne analize poklicnega presojanja, ki se dogaja v naravnem okolju pogosto tudi pod pritiski in oteževalnimi okoliščinami, so pomembno prispevale v odmiku od prevladujočih modelov odločanja, npr. po teoriji koristnosti je za »dobro« odločanje potrebno, da oseba pregleda vse alternativne možnosti in vsako od teh evalvira glede koristi oz. izgube. Uspešno poklicno in ekspertno odločanje, ki poteka v praksi, je radikalno drugačno od predpisanih procesov »dobrega« odločanja, ki jih ponujajo različni normativni modeli odločanja. Raziskave odločanja v naravnih situacijah so bile nosilci nove paradigme odločanja – t.i. naravne (naturalistične) teorije odločanja. Paradigma naravnega odločanja je postavila eksperta v središče raziskovanja in tako je ekspertiza postala bistvo raziskav odločanja v naravnih situacijah.

Cohen (1993) je prikazal razvoj paradigem odločanja in opisal tri vrste paradigem raziskovanja

odločanja : formalno-empiricistično paradigmo (klasična teorija odločanja), racionalistično paradigmo in naturalistično paradigmo.

Formalno-empiricistična paradigma je bila najbolj zgodnja in je bila dominantna do 60-ih let prejšnjega stoletja. Ta paradigma je slonela na normativnem predpisovalnem modelu racionalnega vedenja odločevalcev, ki je predvideval, da osebe izbirajo med večjim številom različnih možnosti. Za to paradigmo so značilni usmerjenost na vhod-izhod (*input-output*), zelo obsežno preiskovanje informacij kot tudi formalni razvoj abstraktnega modela, ki je neodvisen (prost) od konteksta in zavezan kvantitativnemu testiranju (Lipshitz, 2001). Omenjena paradigma se je osredotočala na vedenjsko testiranje formalnih modelov, ne pa na razumevanje kognitivnih procesov, ki so udeleženi v procesu odločanja.

Racionalistična paradigma je prehitela klasično paradigmo odločanja, vendar je obdržala nekatere njene glavne značilnosti, kot so normativni (predpisovalni) modeli odločanja, ki so standardni za evalvacijo kvalitete odločanja. Racionalistična paradigma vključuje koncept napak pri odločanju zaradi odločanja brez ustreznih podpor. Po racionalističnem pojmovanju je nastala diskrepanca v odločanju zaradi napake, ki jo naredi odločevalec in ki ne izvira iz modela odločanja, za razliko od pristopov formalno-empiricistične paradigme, ki so v takih primerih iskali modifikacije modela. To kaže na očiten premik k kognitivnim procesom in njihovim omejitvam kot tudi k odkrivanju psihičnih procesov v odločanju – med drugim tudi vloge intuicije.

Tretja paradigma je *naturalistična* in je usmerjena v odločanje v naravnih situacijah ter umešča eksperta v središče raziskovanja. Namesto obsežnega procesa izbiranja je v ospredju skladnost situacije in akcije, osredotočenost na vhod-izhod je nadomestila usmerjenost na procese in kontekstno proste formalne modele se nadomestili modeli, vezani na kontekst. Raziskovalci so se v okviru te paradigme lotili sestavljanja deskriptivnih modelov, značilnih za ekspertne odločevalce v naravnih okoljih, ne da bi se naslanjali na normativne modele (Lipshitz, 2001). Zanimivo je, da se ta paradigma ni bistveno spremenila od časa svoje pojavitve. Še vedno jo označujejo odločevalci, ki so posebno uspešni na nekem ožjem poklicnem področju, ali eksperti, pravila odločanja glede na ujemanje situacije in akcije, kontekstno vezano neformalno modeliranje ter usmerjenost na procese in predpisovanje, ki temeljijo na empiričnih ugotovitvah (Lipshitz, 2001). Čeprav Lipshitz (2001) hkrati ugotavlja, da je v nekaterih pogledih prišlo do sprememb in sicer naj bi v zgodnejših delih bil poudarek na značilnostih poklicnih kontekstov in je bila ekspertiza drugotnega pomena, so kasneje ekspertni odločevalci in ekspertiza postali središče raziskav odločanja v naravnih poklicnih situacijah.

Razlike in podobnosti paradigem raziskovanja odločanja bi lahko povzeli, kot kaže **tabela**.

FORMALNO EMPIRICISTIČNA PARADIGMA ODLOČANJA	RACIONALISTIČNA PARADIGMA ODLOČANJA	NATURALISTIČNA PARADIGMA ODLOČANJA
normativni modeli racionalnega vedenja	normativni modeli	neformalni modeli, vezani na kontekst odločanja
formalni, abstraktni modeli prosti konteksta in kognicij odločevalcev	modeli so standardi kvalitete odločanja	konstrukcije modelov ekspertnega odločanja v naravnih okoljih
napake v odločanju se pripisujejo pomanjkljivostim modelov	napake v odločanju se pripisujejo odločevalcu in ne modelu	
osredotočenost na vhode in izhode procesa odločanja	procesna usmeritev: osredotočenost na kognitivne procese in njihove omejitve v odločanju	procesna in kontekstna usmeritev: osredotočenost na kognitivne procese ekspertov in naravne kontekste odločanja
obsežna (celovita) preiskava informacij		ujemanje (skladnost) situacije in akcije
izbiranje med alternativami		mentalna simulacija

Tabela. Značilnosti paradigem odločanja

3 SPOSOBNOSTI EKSPERTOV, KONTEKST IN ODLOČANJE

Pri raziskovanju ekspertov, za katere velja, da so dosegli izjemne spretnosti in znanje na nekem ožjem področju, se je raziskovanje s pozicij naturalistične paradigme osredotočilo predvsem na raziskovanje procesov razvoja in uporabe ekspertize v kontekstu. Raziskovalci so identificirali številne variable ekspertnosti, ki so pomembne za raziskovalce odločanja v naravnih situacijah (Pliske, McCloskey, Klein, 2001):

- Za eksperte so značilne *perceptivne sposobnosti* fine diskriminacije. V situaciji vidijo več kot novinci predvsem z opažanjem pomembnih »ključev«.
- Zanje so značilne tudi bogate notranje reprezentacije o delovanju elementov nekega konkretnega področja prakse. Ti *mentalni modeli* jim omogočajo, da hitreje razumejo situacijo.
- Imajo obsežen repertoar kompleksnih vzorcev. Hitro prepoznajo, kaj je *tipično* v neki situaciji kot tudi kdaj se stvari ne odvijajo v pričakovani smeri.

- Imajo širok repertoar *taktik*. Vedo, kako je treba narediti določene stvari. Imajo bogato proceduralno znanje.
- Eksperti poznajo več dejstev in podrobnosti – *deklarativno znanje*. Imajo več *tacitnega znanja* kot novinci. Tacitno znanje, ki je operativno znanje, ni dostopno zavesti, ni verbalno kodirano, niti ga ni mogoče zlahka artikulirati.
- Eksperti imajo dobro razvito sposobnost *mentalne simulacije*, ki je pomembna pri preverjanju in izpopolnjevanju poteka neke akcije oz. pri razumevanju, kako je prišlo do neke situacije.
- Eksperti v primerjavi z novinci posvetijo več pozornosti in več časa razumevanju dinamične situacije oz. *oceni situacije*. Novinci več čas posvetijo razmišljanju o *poteku akcije*.
- Posedujejo več strategij za *upravljanje z negotovostjo* na nekem področju delovanja.
- Imajo dobro poznavanje in razumevanje svojih lastnih močnih in šibkih točk. So boljši pri spremljanju, nadzorovanju in samouravnavanju lastnih kognitivnih procesov – *metakognicija*.

Omenjene značilnosti ekspertov omogočajo ustrezno ravnanje v oteževalnih okoliščinah. Potreba po razumevanju odločanja v kontekstu časovnih pritiskov, negotovosti, slabo definiranih ciljev in visokih osebnih interesov (vložkov) je bila glavna spodbuda za odkrivanje odločanja v naravnih situacijah. S publikacijo *Odločanje v akciji: modeli in metode* (Klein et al., 1993) se je odločanje v naravnih situacijah združilo v posebno področje študija, ki je med drugim omogočilo opredeliti kontekstualne faktorje, značilne na odločanje v naravnih situacijah:

- Slabo definirani ali odprti problemi kot nasprotje umetnih in dobro strukturiranih.
- Negotovo, nezanesljivo dinamično okolje v nasprotju s statičnimi, simulacijskimi situacijami.
- Akcija – povratna zanka in ne odločitev v prvem poskusu.
- Časovna omejenost in stres v nasprotju z imeti dovolj časa za nalogo.
- Visoko osebni interesi (vložki) v nasprotju s situacijami, kjer za odločevalce ni pravih posledic odločanja.
- Vključenost večih posameznikov kot nasprotje individualnega odločanja.
- Organizacijski cilji in norme kot nasprotje odločanja v vakuumu.

VIRI

- Carver, C.S. (2001). Affect and the functional base of behavior: On the dimensional structure of affective experience. *Personality and Social Psychology Review*, 5.
- Cohen, M.S. (1993). Three paradigms for viewing decision biases. V G.A. Klein, J. Orasanu, R. Calderwood, C.E.

- Zsombok (ur.), *Decision making in action: Models and methods*. Norwood, N.j: Ablex.
- Gigerenzer, G. Goldstein, D.G. (1996). Reasoning the fast and the frugal way: Models of bounded rationality. *Psychological Review*, 103.
- Higgins, E. T., & Molden, D. C. (2003). How strategies for making judgments and decisions affect cognition: Motivated cognition revisited. V G.V. Bodenhausen, & A. J. Lambert (ur.) *Foundations of social cognition* (pp. 211-236). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hoffman, R.R. (1995). *A review of naturalistic decision making research on the critical decision method of knowledge elicitation and the recognition priming model of decision-making with a focus on implications for military proficiency*. (pripravljeno za Univerzo v Nottinghamu pod pokroviteljstvom Defense Research Agency-Farnborough and Epistemics, Ltd.)
- Hsee, C.K., Kunreuther, H.C. (2000). The affection effect in insurance decisions. *Journal of Risk and Uncertainty*, 20.
- Klein, G.A., Orasanu, J., Calderwood, R., Zsombok, C.E. (ur.) (1993). *Decision making in action: Models and methods*. Norwood, NJ: Ablex
- Lipshitz, R. (2001). Puzzle seeking and model-building on the fire ground: A discussion of Karl Weik's keynote address. V E. Slas, G. Klein (ur.), *Expertise and naturalistic decision making*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Loewenstein, G.F., Weber, E.U., Hsee, C.K., Walch, N. (2001). Risk as feelings. *Psychological Bulletin*, 127.
- Mellers, B. (2000). Choice and the relative pleasure of consequences. *Psychological Bulletin*, 126.
- Panksepp, J. (2005). Affective consciousness: Core emotional feelings in animals and humans. *Consciousness & Cognition: An International Journal*, 14.
- Pliske, R.M., McCloskey, M.J., Klein, G. (2001). Decision skills training: Facilitating learning from experience. V E. Salas, G.A. Klein (ur.), *Linking expertise and naturalistic decision making*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Plutchik, R. (1980). *Emotions: A psychoevolutionary synthesis*. New York: Harper & Row.
- Slovic, P., Finucane, M., Peters, E., MacGregor, D.G. (2002). The affect heuristics. V T. Gilovich, D. Griffin, D. Kahneman (ur.), *Heuristics and Biases: The psychology of intuitive judgement*. New York: Cambridge University Press.
- Westen, D. (1985). *Self and society: Narcissism, collectivism, and the development of morals*. New York: Cambridge University Press.
- Westen, D. (1994). Toward an integrative model of affect regulation: Applications to social-psychological research. *Journal of Personality*, 62.
- Westen, D., Gabbard, G.O. (2002). Developments in cognitive neuroscience: I. Conflicts, compromise, and connectionism. *Journal of the American Psychoanalytic Association*, 50.

SKUPINSKO ODLOČANJE KOT AKTUALIZIRANJE DELOVANJSKIH POTENCIALOV

Andrej Ule

Filozofska fakulteta

Aškerčeva 2, 1000 Ljubljana, Slovenija

Tel. +386 1 2411102, fax: +386 1 2411495

e-mail: Andrej.ule@guest.arnes.si

ABSTRAKT

Predstavim nekaj značilnosti kolektivnega odločanja in dilem racionalnega skupinskega odločanja in uvedem zamisel o kristalizaciji odločitev kot aktualizacije skupinskih potencialov.

ABSTRACT

Some traits of collective decision and dilemmas of group rational decision making and the idea of cristalization of decisions as actualization of the group potentials are presented.

Odločanje in racionalnost

Odločanje razumem kot proces in dejanje izbire akterjev (posameznikov, skupin, institucij) med več možnimi smermi razmišljanja, čustvovanja, vrednotenja ali delovanja v določeni situaciji. Odločanje kot dejanje je lahko povsem »mentalno« ali tudi fizično, tj. kaže se v zunanjem svetu. Odločanje nam pomaga zmanjšati kompleksnost situacije in z njo povezane dvome in negotovosti na obvladljivo mero. Je že neko dejanje, ki (običajno) vodi k drugim dejanjem, npr. k izvrševanju ene od alternativ delovanja. Le malo odločitev se stori z maksimalno gotovostjo, zato večina odločitev le zmanjša kompleksnost situacije (problema), ne pa, da jo povsem eliminira ali povsem reši problem. To že zato, ker akter(ji) le redko posedujejo popolno znanje o vseh alternativah in možnih rezultatih, prav tako nimajo povsem jasnih in enoličnih preferenc (želja, interesov, potreb). Zato odločanje normalno vsebuje elemente negotovosti in tveganja. Tako izkušnja kot znanstvene raziskave pa nam kažejo, da je teh negotovosti manj, če naši razmisleki in odločitve potekajo na podlagi »dobrih razlogov«, ne pa *ad hoc*, slučajno, zgolj pod pritiskom razmer itd. V tem smislu so izjemno pomembne *racionalne odločitve*, torej odločitve na podlagi dobrih, ustreznih, zadostnih razlogov. Kakšni so ti razlogi in kako pridemo do njih in do racionalnih odločitev, pa je še težko in nerešeno vprašanje.

Odločanje se torej povezuje z *racionalnostjo*. To sta pojma, ki ju pogosto uporabljamo, a le redko natančno definiramo. Sta konceptualno povezana že v svoji osnovi, kajti vsaj velik del tega, čemur pravimo odločanje, je izbiranje med alternativami na podlagi določenih razlogov, vzorčni, paradigmatični primeri racionalnosti pa so ravno odločitve z razlogi. Zaradi večplastnosti in nejasnosti obeh pojmov se omejujem na določene segmente pojavnosti in problemov, ki

ju običajno zajemamo s tema dvema pojmom, namreč na tiste, ki se tičejo kolektivnega, skupinskega razmišljanja in delovanja, še bolj konkretno, kreativnega skupinskega odločanja.

V tem tekstu razumem *racionalnost kot zmožnost (ability) mišljenja ali delovanja na podlagi razlogov (reasons)*, ki je lahko *podvrženo kritiki*.¹ Pod »mišljenjem« zajemam obsežno območje človeških kognitivnih procesov, ki jim lahko določimo propozicionalno vsebino, torej vsebino, podano v obliki povednih stavkov, pod »delovanjem« pa vedenje, ki mu lahko določimo kak smoter, cilj (teleološkost).

Sposobnost za (dobro) odločanje zahteva od akterjev celo vrsto zmožnosti in kompetenc, ki se jih lahko naučimo v vsakdanji praksi odločanja ali skozi študij in vežbanje različnih odločitvenih situacij. Ključne zmožnosti oz. kompetence posameznikov za učinkovito odločanje so: *razumevanje* situacije, *ocenjevanje* pomena odločitve, *razmislek* o odločitveni situaciji, o možnih izbirah, rezultatih, koristi in stroških, o možnosti uresničenja alternativ, *izbor* (ene izmed alternativ), *sledenje vrednotam* (sledenje temu, kar je za nas oz. v dani okoliščini dobro). Predvsem slednja zmožnost pogosto zajema tudi čustvovanje in druge zunaj-racionalne mentalne procese.

Raziskovanje racionalnosti odločitev je pomembno zato, ker lahko izboljša ti. intuitivno odločanje posameznikov ali skupin. Še zlasti intuitivno skupinsko odločanje podlega mnogim psihološkim pritiskom, npr. pritisku po čim večjem ujemanju in soglasju v skupini (npr. ti. »group think« pojav), spregledovanje kritičnih informacij ipd. (Janis, Mann, 1977).

Ocena o racionalnosti odločanja je pogosto precej kompleksna, ker moramo upoštevati celo vrsto kognitivnih in kontekstualnih dejavnikov, zato jo v praksi pogosto poenostavimo, npr. tako, da upoštevamo kot normo racionalnosti določene idealizirane modele racionalnega odločanja. Med takšne modele sodijo npr. logični modeli

¹ Ta opredelitev je blizu Toulminovi opredelitvi racionalnosti prepričanosti kot »odprtosti za argument« (Toulmin, Rieke, Janik, 1979: 13) in Habermasovi opredelitvi racionalnosti kot sposobnost nekega prepričanja, izjave ali delovanja (Äusserung), da se podverže transsubjektivnemu procesu upravičenja in kritike (Habermas, 1981: 27), vendar ne implicira niti Toulminove teorije argumentacije niti Habermasove teorije diskurza kot medija komunikacijske racionalnosti.

reševanja problemov, klasični model racionalne izbire, modeli teorije iger in modeli praktičnega sklepanja. Ti modeli vsebujejo nekaj temeljnih idealizacij, npr. predpostavko o logični koherenci prepričanj akterjev, predpostavko, da v svojih odločitvah vedno težijo k maksimiziranju pričakovanih koristi, predpostavko o popolnosti primerjav med možnimi izbirami (za poljubni izbiri velja, da lahko rečemo, da je ena bolj privlačna od druge ali da ne dajemo prednost nobeni od njiju), tranzitivnost izbir (če akter preferira izbiro a pred izbiro b in preferira izbiro b pred izbiro c, potem preferira a pred c), predpostavko, da racionalni akterji sledijo osnovnim načelom logike in teorije verjetnosti ipd. (Elster, 1986, Hastle, Dawes, 2001: 261-277). Ti modeli dalje predpostavljajo, da se ljudje odločajo glede na pričakovane koristi pri posameznih delovanskih izbirah in glede na preference med izbirami, da so sposobni »izračunati« koristi (in stroške) za izvedbo vsake od izbir ter pri tem maksimizirajo svoje koristi.

V kolikor gre za izbire v socialnih situacijah, kjer je naša izbira odvisna od izbir drugih oseb, potem moramo uporabiti koncept *strateške igre*, tj. praviloma (ne pa nujno) skupinske situacije, kjer udeleženci na podlagi poznavanja svojih preferenc, domnev o preferencah svojih »soigralcev«, predpostavk o racionalnosti in znanju soigralcev, izbirajo strategijo svojega delovanja in to tako, da bi s tem kar najverjetneje dosegli maksimalni rezultat zase (tj. uresničili čim več od svojih preferenc in eo ipso, minimizirali morebitne »stroške« ali »izgube«).

Znano je, da idealizirani modeli in metode racionalnega odločanja niso suho zlato, kajti empirično raziskovanje odločanja kaže, da postavljajo previsoka merila racionalnosti, saj ljudje pogosto in odstopamo od njih, celo če gre za razmeroma elementarne oblike logičnega ali verjetnostnega sklepanja. Naše preference so redko tako koherentne, kot to zahteva klasični model racionalnega odločanja, poleg tega pa se praviloma ne obnašamo kot »racionalni egoisti«, tj. kot osebe, ki bi maksimirali (le) svoje koristi. Na ta dejstva so opozarjali mnogi raziskovalci, tako da so ta spoznanja postala splošno sprejeta v sodobnih raziskavah racionalnosti (Evans, 1987, Tversky, Kahneman, 1988, Evans, Over, 1996, Elio, 2002a).

Herman Simon, sicer tudi eden od tvorcev kognitivne znanosti, je že v petdesetih letih predlagal model ti. omejene (bounded) racionalnosti, ki naj bi vključil vsaj nekatera standardna odstopanja dejanskega mišljenja in odločanja ljudi od klasičnih predpostavk racionalnosti (Simon, 1956, 1982). Po tem konceptu ljudje zaradi naravnih in kulturnih omejitev svojega kognicijskega aparata preprosto ne moremo doseči klasičnih standardov racionalnosti, zato se pogosto zadovoljimo s ti. »zadoščanjem« (satisficing), tj. s sklepanji in odločitvami, ki sicer niso optimalne, vendar za nas še vedno »dovolj dobre«.

V okviru razprav o povezavah vsakdanjega razmišljanja in racionalnosti so se razvili koncepti otelesene (embodied) racionalnosti, ki skušajo ugotoviti optimalne (ne pa idealno maksimalne) načine reševanja problemov in odločanja v določenih razmerah, pri čemer upoštevajo značilnosti okolja, probleme, ki jih ljudje rešujejo in človeško kognitivno arhitekturo (Anderson, 1991, Oaksford, Chater, 2002). S tega ozira je možna kritika določenih razmišljanj, odločitev in dejanj kot ne- oz. celo kot iracionalnih, a obenem je možna tudi razlaga tega početja, ki ne pomeni posrednega upravičenja početja.

Nekoliko drugačna smer razmišljanja je tista, ki uvaja dve oz. več vrst racionalnosti oz. sistemov razmišljanja, da vseh napak v razmišljanju in odločanju ne moremo pojasniti s slučajnimi napakami v razmišljanju ali z komputacijskimi omejitvami v človeškem kognitivnem aparatu. Zdi se razumno sklepati na sistemske izvore teh napak, ki pa ne ukinjajo racionalnosti na sploh.²

Skupinsko odločanje

Ko govorim o skupinskem odločanju se omejujem na razpravo o manjših, neformalnih skupinah, izpuščam pa razpravo o racionalnosti večjih kolektivov ali institucij. To omejitev sprejemam prvič zaradi omejenega obsega razprave in drugič zato, ker je analiza odločanja v manjših, neformalnih skupinah bistvenega pomena za vse kolektivno odločanje, saj se mnoge značilnosti kolektivne kognicije pri manjših skupinah »dedujejo« tudi pri večjih in bolj formalno konstituiranih skupinah. Prav tako se omejujem le na tiste značilnosti kognitivnih procesov v skupinah, ki jih lahko zajamemo z repertorajem sodobne kognitivne znanosti, npr. jim lahko najdemo vsaj hevristične računalniške modele.

Racionalno kolektivno odločanje predstavlja odločanje skupin, ki temelji na razumni presoji situacije. Kot v primeru posameznika ni treba, da so te presoje vedno zavestne, sklepalne, skupinsko vodene in sprejete, pač pa so v danih okoliščinah racionalne za skupino. Ugotovitev, da je neka skupina v dani situaciji razmišljala ali delovala racionalno, pomeni, da so člani skupine (ali vsaj reprezentativen del članstva skupine) presojali o tem, kaj je za vse njih bolje oz. najbolje v dani situaciji in se podlagi teh presoj odločali in delovali.

Na videz lahko domnevamo, da bo interakcija posameznikov, ki težijo k racionalnim odločitvam, praviloma privedla k racionalnim skupnim odločitvam. Čeprav je temu pogosto tako (npr. delovanje trga v razmerah, ko niso prisotne velike perturbacije v tokovih dela,

² Razni avtorji so predlagali delitev načinov razmišljanja na bolj avtomatsko potekajočo, spontano, implicitno, nezavedno razmišljanje ljudi v interakcijskih situacijah in za bolj zavestno, reflektivno, normativno vodeno, eksplicitno razmišljanje in podobno za dve vrsti racionalnosti (Evans, Over, 1996, Sloman, 1996, Pollock, 1993, Stanovich, 1999).

blaga in denarja), nam je tako iz vsakdanje prakse kot iz znanstvenega opazovanja delovanja skupin znano, da temu ni vedno tako. Nekateri tovrstni primeri so naravnost standardne dileme individualnega odločanja v kolektivnem kontekstu in se pogosto formalno obravnavajo s pomočjo teorije iger. Znamenite so npr. *tragedija skupnega, lovska preža, zaporniška dilema*, idr. V prvi gre za dilemo, kako izkoriščati kako za vse ljudi potrebno skupno, a omejeno dobrin. V drugi dilemi gre za situacijo, ko se mora posameznik odločiti, ali bo sledil neredil potezo, ki koristi vsej skupini ali potezo, ki močno koristi njemu samemu, a škodi skupini. V tretji dilemi pa se akterji odločajo med sodelovanjem ali nesodelovanjem, pri čemer vsi nekaj pridobijo, če sodelujejo, vendar če nekdo ne sodeluje, drugi pa sodelujejo, lahko dobi več. Če pa noben ne sodeluje, vsi izgubijo (ali dobijo zelo malo).

Ti problemi sodijo v širši sklop *problemov skupinskega delovanja*. John Elster ugotavlja, da ima skupina problem skupinskega delovanja tedaj, ko je bolje za vse, če nekateri njeni člani sodelujejo v delovanju, kot če ne sodeluje nihče, vendar pa je za vsakega posameznika bolje, če ne sodeluje (Elster, 2007: 126). Prej navedeni paradoksi odločanja so elementarne oblike teh problemov. Pogosto jih znamo brez večjih zapletov Tudi poskusi na realnih osebah z realnimi situacijami podobne vrste so pokazali, da so ljudje bolj kooperativni in bolj uspešni v sodelovanju kot kažejo razne teorije individualne racionalnosti. Strogo vzeto, po teh teorijah do sodelovanja med ljudmi pogosto sploh ne more priti-to je preprosto neracionalno početje. Poskusi in računalniške simulacije so pokazale, da se sodelovanje poveča, če se »igre« nadaljujejo (in ni določeno, kdaj se ponovitve končajo), ali če akterji poznajo svoje partnerje in so z njimi že sodelovali v drugih primerih. Grobo vzeto velja: kolikor več sodelovanja je bilo prej, toliko več ga lahko pričakujemo v nadalje. Mancur Olson je v svojem znanem delu »The Logic of Collective Action« (1965) ugotavljal, da ljudje v splošnem sodelujejo pri zagotavljanju neke skupne dobrine, če gre za *manjše skupine, kjer člani pričakujejo očitno korist od sodelovanja*, ali če so člani prisiljeni sodelovati oz. obstajajo kakšne druge *posebne vzpodbude za sodelovanje*. Organiziranost skupine lahko pripomore k večjemu sodelovanju članov, a ne nujno. Organiziranost je nujno potrebna v velikih skupinah, kjer prispevki posameznih članov na splošno niso opazni za druge člane (»anonimnost sodelovanja«). Elster meni, da je napačno domnevati, da sodelujemo zgolj zaradi lastnih koristi, ali pa, da nas vedno vodi kak poseben motiv. V sodelovanju se povezujejo različne, tako racionalne kot neracionalne motivacije posameznikov, ki krepijo ena drugo (2007: 397).

V mnogih primerih sodelovanja je prisotna potreba po medsebojni usklajenosti (koordinaciji) dejanj več posameznikov. Zato odločanje za sodelovanje poteka pod tiho ali očitno domnevo, da so drugi pripravljene koordinirati svoje delovanje z drugimi posamezniki. Takšno je npr. odločanje za skupen izlet, za ples s soplesalko ali soplesalcem, za igranje muzike v ansamblu, za sodelovanje

kake skupinske igre, pa celo tako preproste stvari, kot je odločanje za skupno hojo po parku. Tu se ljudje zanašajo na *skupno* domnevo sodelujočih, da se bo vsakdo od njih trudil korektno opravljati svoj del skupne dejavnosti. Vse nadaljne odločitve članov skupine potem potekajo ob tej tihi predpostavki. Če bi skušali postaviti razmišljanje akterjev kot neke vrste praktični sklep, potem bi morali narediti *dkupen* praktični sklep za vse udeležence z več premisami, ki povzemajo njihove partikularne, medsebojno koordinirane smotre in sredstva delovanja in *skupno pravilo* racionalnosti, ki ob danih premisah terja nujno izpeljavo sklepa o medsebojno usklajenem delovanju vseh akterjev. To pa presega horizont ti. metodološkega individualizma v razlagah kolektivnih dejanj in odločitev.

Kristaliziranje skupinskih odločitev

Zdi se, da moramo predpostaviti proces postopne ali nagle kristalizacije medsebojno osklajenih racionalnih odločitev posameznih akterjev v kolektivnem odločanju, ki se ne ujema s shemami racionalnega odločanja posameznikov (bodisi s »popolno« ali »omejeno« racionalnostjo). Pojem postopne kristalizacije racionalnih skupnih rešitev, pravil, norm spominja na pojem *žariščnih točk* (focal points), o katerem veliko razpravljajo teoretiki iger in racionalnosti. Gre za razlago ugotovitev Thomasa Schellinga, da ljudje znajo koordinirati svoje vedenje tako, da je to vsem v prid, čeprav si ne delijo nobenih izrecnih skupnih prepričanj ali skupnih percepcij, ki bi jih vodile v koordinaciji (Schelling, 1960). Pač pa obstajajo določena implicitna vodila za koordinacijo, ki so v dani situaciji nekako poudarjena (salient). Tem vodilom pravimo žariščne točke koordinacije. Pomembno vlogo igrajo domneve posameznikov o tem, kaj naj storijo in kaj nameravajo storiti drugi člani skupine ob predpostavki, da vsi ali večina nameravajo sodelovati pri kakem skupnem početju. Ni treba, da so ti premisleki učinkovito izpeljani, lahko so zgolj naznačeni v mentalnem prostoru posameznikov ali nepopolno navrženi v pogovoru. Tu gre za pomembne in zapletene procese prehoda implicitnega v eksplicitno znanje, ki jih tu ne morem obravnavati. Domnevam, da ta proces ni le socialno-psihološko kavzalen, temveč tudi socialno racionalen, tj. argumentativen in vsaj delno miselno reflektiran v zavesti udeležencev.

Domnevam, da *obseg in kvaliteta razlogov in argumentov* za različne trditve ali odločitve, ki jih dajejo v »obtok« člani skupine ter medsebojno navezovanje teh razlogov pomembno prispevajo k kristalizaciji skupnih odločitev. Te odločitve se kristalizirajo okrog (implicitnih) žariščnih točk razprave in delovanja članov skupine. Prehodi od implicitnega k eksplicitnemu znanju in kristaliziranje kolektivnih odločitev niso le intelektualni, retorični, komunikacijski procesi, temveč delujejo s specifično močjo, ki *aktualizira določene potenciale za odločanje* in delovanje posameznikov ali skupin. Govorim o »energetizaciji« razmišljanja in odločanja. Pri posamezniku so močan vir energetizacije želje in hotenja ter čustva (tako pozitivna kot negativna), v skupinah pa občutek skupinske kohezije,

stopnja identifikacije s skupino ter čustvena in normativna zavezanost skupno sprejetim ciljem pri članih skupine. V situacijah, kjer se mora več ljudi hitro odločiti o tem, če in kako bodo sodelovali, se naslanjajo na utečene vzorce delovanja (reagiranja), na žariščne točke v iskanju ustreznih izbir, na svoje čustva, občutke in hotenja, le malo pa na morebitne dodatne razloge ali argumente, ki bi si jih delili v komunikaciji. V situacijah pa, kjer se člani neke skupine znajdejo pred kompleksnejšim problemom, ki terja dodatno razpravo in medsebojno usklajevanje stališč, pa pričakujemo bolj ali manj obsežno in intenzivno opiranje na osebne in skupinske »poole« razlogov in argumentov, s katerimi pride najprej do usklajevanja medsebojnih razlik in potem do kristalizacije skupinske odločitve. Lahko bi dejali, da je skupinska razprava in identifikacija s skupino *aktualizirala* do tedaj potencialne za odločanje posameznih članov v obliki *kolektivne odločitve* za kooperativno delovanje. Nekateri argumenti v toku razprave prispevajo bolj k usmerjanju razprave k morebitni rešitvi problema, drugi pa izražajo »stanje duha« in dajejo razpravi posebno občutje, recimo temu, da »energetizirajo« razpravo. Za uspešno odločitev. zaporedje odločitev, ki vodijo k rešitvi problema, je potrebno oboje, tako »prava smer« kot tudi ustrežna »energija«.

Ta proces bi lahko primerjali s kolapsom valovne funkcije v kvantni mehaniki, ko se pod vplivom merjenja prostor potencialnih aktivacij fizikalnega sistema zreducira na eno od možnosti. Prostoru potencialnosti ustreza struktura implicitnih vrednotenj različnih možnosti (alternativ), vendar te alternative ne nastopajo kot elementi logične disjunkcije, temveč v neke vrste »prepletenosti«, ki jo raplete šele proces odločanja oz. točneje, stališča in argumenti v razpravi in energetiziranje razprave. Po tej analogiji odločitev igra podobno vlogo kot eksperiment in opazovanje v kvantni fiziki, nanreč razvezovanja prepletenosti v določeno »opazljivo vrednost« (v določeno odločitev oz. dejanje). Pomembna razlika je v tem, da je fizikalno opazovanje običajno enkratno, skorajda trenutno poseg v mikroproces, ki v enem koraku kolapsira prepletenost, odločanje pa je pogosto rezultat razprave v skupini, kar je nek proces, ki teče skozi več faz, v katerem se (običajno) postopoma aktualizira izbira. Obstaja nekaj matematičnih modelov odločanja po kvantno-mehanskem modelu, vendar noben ni zgrajen za skupinsko odločanje (Busenmayer idr., 2006, Khrennikov, 2009, Zak, 2000), a postavlja se nam vprašanje, kakšne so možnosti za njihovo nadaljevanje v analizo skupinskega odločanja.

Literatura:

Anderson, J. R. (1991): Is human cognition adaptive? *Behavioral and Brain Sciences*, 14: 471-485.

Busenmayer, J. R., Wang, Z., Townsend, J. T. (2006): Quantum dynamics of human decision-making. *Journal of Mathematical Psychology*, 50: 220-241.

Elio, R. (2002b): Introduction. V: R. Elio (ur.) (2002a): *Common Sense, Reasoning, and Rationality*. Oxford University Press, Oxford: 1- 52.

Elster, J. (1986): Introduction. V: J. Elster (ur.) (1986): *Rational Choice*. Blackwell, Oxford: 1-33.

Elster, J. (2007): *Explaining Social Behavior. More Nuts and Bolts for the Social Science.*. Cambridge Univ. Press.

Evans, J. (1987): *Bias in Human Reasoning: Causes and Consequences*. Erlbaum, Hillsdale.

Evans, J. Over, D. E. (1996): *Rationality and Reasoning*. Psychology Press, Hove.

Hastle, R., Daves, R. M. (2001): *Rational Choice in an Uncertain World*. Sage, London.

Janis, I., Mann, L. (1977): *Decision Making: A Psychological Analysis of Conflict, Choice, and Commitment*. Free Press, 1977, New York.

Habermas, J. (1981): *Theorie des kommunikativen Handelns*. Bd. 1. Suhrkamp, Frankfurt/M.

Khrennikov, A. (2009): Quantum-like model of cognitive decision making and information processing. *Biosystems*, 95: 179-187.

Oaksford, M., Chater, N. (2002): Commonsense reasoning, logic, and human rationality. V: R. Elio (ur.) (2002a): *Common Sense, Reasoning, and Rationality*. Oxford University Press, Oxford: 174-214.

Olson, M. (1965): *The Logic of Collective Action. Public Goods and the Theory of Groups*. Harvard Univ. Press.

Pollock, J. L. (1993): The phlogany of rationality. *Cognitive Science*, 17: 563-588.

Schelling, T. (1960): *The Strategy of Conflict*. Harvard University Press, Cambridge/M..

Simon, H. (1956): Rational choice and the structure of the environment. *Psychological Review*, 63: 101-116.

Simon, H. (1982): *Models of Bounded Rationality*. MIT, Cambridge/M.

Slovan, S. A. (1996): The empirical case for two systems of reasoning. *Psychological Bulletin*, 119: 3-22.

Stanovich, K. E. (1999): *Who is Rational? Studies of Individual differences in Reasoning*. Erlbaum, Mahwah/NJ.

Toulmin, St., Rieke, R., Janik, A. (1979): *An Introduction to Reasoning*. Prentice Hall, New York 1979.

Tversky, A., Kahneman, D. (1988): Rational choice and the framing of decisions. V: D. E. Bell, H. Raiffa, A. Tversky (ur.) (1988), *Decision Making: Descriptive, Normative, and Prescriptive Interactions*. Cambridge Univ. Press, Cambridge: 167- 192.

Zak, M. (2000): Quantum decision-maker. *Information Science*, 128: 199-215.

RAČUNALNIK IN ODLOČANJE: ODLOČITVENI MODELI IN SISTEMI ZA PODPORO PRI ODLOČANJU

Marko Bohanec

Institut Jožef Stefan, Odsek za tehnologije znanja

Jamova 39, SI-1000 Ljubljana, Slovenija

Tel: +386 1 477 3309; fax: +386 1 477 3315

e-mail: marko.bohanec@ijs.si

POVZETEK

V prispevku predstavljamo nekaj pogledov na odločanje, proces odločanja in reševanje odločitvenih problemov s strani računalništva. Poudarek je predvsem na uporabi računalnika pri odločanju, ki ga izvaja posameznik ali skupina ljudi. V prispevku predstavimo discipline, ki se ukvarjajo s podporo pri odločanju in definiramo proces odločanja in njegove lastnosti, ki nas zanimajo v tem okviru. Podrobneje predstavimo odločitveno analizo in sisteme za podporo pri odločanju.

1 UVOD

Pri razvoju računalniških aplikacij je eno najpomembnejših vprašanj vprašanje funkcionalnosti: kaj lahko računalnik ponudi uporabniku in kako mu lahko pomaga pri njegovem delu? Zato računalnikarje tudi pri odločanju zanima, kako bi lahko z računalnikom in ustreznimi računalniškimi programi ljudem pomagali pri odločanju, da bi se odločali lažje, hitreje, bolj učinkovito in nasploh bolje. Govorimo o programih, sistemih, metodah in tehnikah za *podporo pri odločanju* (angl. Decision Support) [7,9].

Pri tem se srečujemo s številnimi vprašanji. Kaj sploh je odločanje, kako poteka in kako široko naj ga zajamemo? Kakšne vrste odločitev poznamo? Od česa so odvisne naše odločitve? Kaj predstavlja naše *vhodne* podatke in kaj so pričakovani *izhodi*? Katere vidike odločanja lahko učinkovito podpremo in katerih ne? Kaj sploh je *dobra* odločitev? V prispevku opisujemo, kako se s temi vprašanji soočajo discipline, ki se ukvarjajo z metodološkimi in računalniškimi vidiki podpore pri odločanju.

2 OSNOVNE DISCIPLINE

Prvo vprašanje, ki ga v zvezi z odločanjem postavi računalnikar, je, ali gre za človeško ali strojno odločanje? Pri podpori odločanja želimo seveda pomagati človeku, torej nas zanima človeško odločanje. Toda dejstvo je, da je računalnik izredno prilagodljiv stroj, ki se je na nek način sposoben odločati tudi sam. Za to, da računalnik lahko izvaja neko nalogo, ga je potrebno programirati, torej definirati zaporedje ukazov, ki naj jih izvaja. Pri izvajanju pogosto nastopajo različne situacije, na katere se mora

računalnik odzvati različno. Na osnovi podatkov, ki jih ima, se mora *odločiti*, kako nadaljevati z izvajanjem naloge. Zato je ena temeljnih značilnosti računalniških programov prav ta, da se vejijo – vsebujejo ukaze, ki “preklaplajo” med različnimi zaporedji ukazov. Navzven se to bolj ali manj uspešno kaže kot sposobnost računalnika, da se odloča sam. Pri tem danes računalniki postajajo vse boljši in so se že sposobni odločati tudi v nepredvidljivih situacijah. Za primer omenimo računalniško krmiljene avtomobile, ki so bili v okviru izziva DARPA (<http://www.darpa.mil/grandchallenge/index.asp>) sposobni povsem samostojno prevoziti okrog 100 km poti v urbanem okolju.

S podporo človeškega odločanja se ukvarja cela vrsta disciplin, ki jih pogosto označimo s skupnim imenom *odločitvene znanosti* (angl. Decision Sciences). Gre predvsem za tri skupine disciplin [1]:

1. Prva skupina se uvarja z vprašanjem *racionalnega odločanja*: kako in po kakšnih pravilih naj se človek odloča pametno, smiselno, racionalno? Gre za *normativni* (kako bi bilo treba) in pretežno teoretični pristop, ki ga srečamo na primer pri odločitveni teoriji (angl. Decision Theory) [4,6].
2. Druga skupina opazuje, kako se ljudje dejansko odločamo. Govorimo o *deskriptivnem* ali *opisnem* pristopu, ki je značilen za večino kognitivnih znanosti.
3. Tretja skupina pa se dejansko ukvarja s *podporo pri odločanju*: kako udejaniti spoznanja normativnih in opisnih disciplin ter na ta način pomagati ljudem, da bi se odločali lažje in bolje? Sem sodijo na primer operacijske raziskave (Operational Research), odločitvena analiza (Decision Analysis) in sistemi za podporo pri odločanju (Decision Support Systems).

3 ODLOČANJE

V tehničnih disziplinah, ki se ukvarjajo s podporo pri odločanju, *odločanje* običajno definiramo kot proces izbire ene izmed več variant, alternativ, možnosti oziroma različic. Izbrati želimo tisto alternativo, ki najbolj ustreza ciljem odločevalca (ali skupine odločevalcev). Bistveni komponenti odločanja sta torej dve: množica *alternativ*,

med katerimi izbiramo, in *cilji*, ki jih želimo uresničiti z odločitvijo.

Odločanje je *proces*. To pomeni, da lahko traja dalj časa in je v splošnem sestavljen iz več aktivnosti, kot so:

- spoznavanje odločitvenega problema;
- zbiranje in preverjanje informacij;
- identifikacija alternativ;
- predvidevanje posledic odločitev;
- *odločitev* (izbira alternative),
- realizacija odločitve,
- obveščanje o odločitvi in razlogih zanjo;
- vrednotenje odločitve.

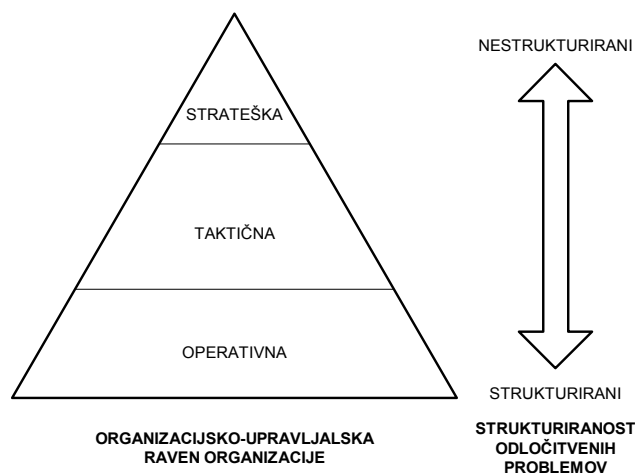
Ključni korak odločanja (kot procesa) je *odločitev*, to je izbira alternative na osnovi presoje, osnovane na zbranih informacijah. Z izbiro dajemo prednost določenim rešitvam in sicer tistim, za katere predvidevamo, da nas bodo najhitreje, najlažje ali najbolj zanesljivo pripeljale do cilja (ali vsaj čim bliže). Odločitev je rezultat zavestne kognitivne aktivnosti, zato je odločevalec odgovoren zanjo in za njene posledice. *Realizacija odločitve* je običajno povezana s porabo virov, kot so čas, energija in denar, kar pomeni, da so odločitve nepovratne in se njihovih posledic ne da povsem izničiti [8].

Paleta odločitvenih problemov je obsežna in raznolika. Na eni strani imamo vsakdanje odločitve, ki jih rešujemo zlahka in skoraj nezavedno (ali naj kupim štruco črnega ali belega kruha?), na drugi pa na primer probleme, ki zahtevajo ogromne vire, zadevajo mnogo ljudi in imajo pomembne družbene posledice (kako dolgoročno urediti železniški promet v Sloveniji?). Nekje vmes so pomembni življenjski problemi posameznikov (kaj naj študiram?), družin (kje in kako stanovati?) in organizacij (kako rešiti proizvodnjo, ki usiha?). Pri podpori odločanja nas običajno zanimajo le dovolj "zahtevni" odločitveni problemi, se pravi takšni, ki se jih "splača" lotiti na nek sistematičen način, in ki imajo dovolj "pomembne" posledice, da se o njih splača pridobiti čim več informacij, jih temeljito premisliti, se o njih posvetovati in jih v splošnem podpreti z neko metodo, računalniškim programom ali informacijskim sistemom. Hkrati velja, da lahko primerno podpremo le odločitveni problem, ki ga dovolj dobro razumemo in poznamo: vedeti moramo, o čem točno se odločamo, kaj hočemo doseči z odločitvijo, kakšne posledice ima lahko odločitev, vsaj približno moramo poznati alternative in njihove lastnosti, se zavedati glavnih virov negotovosti in podobno.

V odvisnosti od teh lastnosti potem različne odločitvene probleme delimo na več načinov:

- Po *zahtevnosti* delimo odločitve predvsem na *rutinske* in *tiste*, ki to niso.
- Po *pogostosti* ločimo *enkratne* in *ponavljajoče se* odločitve.
- Glede na *čas in zaporedje* ločimo med *enostopenjskimi* in *večstopenjskimi* odločitvami, slednje pa so lahko *vzporedne* ali *zaporedne*.

- Po *številu kriterijev*, ki jih upoštevamo pri izbiri najboljše alternative, delimo odločanje na *enokriterijsko* (ali *enoparametrsko*) in *večkriterijsko* (*večparametrsko*).
- Glede na *negotovost* razlikujemo med odločanjem v *gotovosti*, s *tveganjem* in v *negotovosti*.
- Po številu udeležencev so odločitve *individualne* ali *skupinske*.
- S stališča podpore pri odločanju je zelo pomembna delitev glede na *strukturiranost* (slika **Error! Reference source not found.**): pri *strukturiranih* odločitvah je postopek odločanja znan in dobro definiran, pri *nestrukturiranih* pa ne; prve lahko vsaj deloma avtomatiziramo, drugih pa praktično ne.
- Glede na *organizacijsko raven*, na kateri poteka odločanje v gospodarskih in drugih organizacijah (slika 1), ter na vpliv teh odločitev na organizacijo, lahko le-te razdelimo v *strateške* (so dolgoročne in vplivajo na celotno organizacijo), *taktične* (srednjeročni vpliv na del organizacije) in *operativne* (kratkoročni vpliv pri izvajanju operativnih nalog).



Slika 1: Ravnosti odločanja in strukturiranost odločitev

Tabela 1: Lastnosti informacij glede na raven odločanja v organizaciji.

Lastnost informacije	Operativne odločitve	Strateške odločitve
Točnost	visoka	nizka
Podrobnost	podrobna	izpeljana (agregirana)
Časovni horizont	sedanjost	prihodnost
Pogostost uporabe	pogosto	redko
Viri	notranji	zunanji
Pogled	ozek	širok
Tip informacije	kvantitativna	kvalitativna

Poznavanje lastnosti odločitvenih problemov je pomembno zato, ker opredeljuje lastnosti informacij, ki jih imamo na voljo pri podpori odločanja, ter lastnosti posameznih nalog, ki jih je potrebno podpreti. Na primer, informacije, s

katerimi imamo opravka na različnih ravneh odločanja v organizaciji, se med seboj bistveno razlikujejo; njihove najpomembnejše značilnosti povzema tabela 1.

4 ODLOČITVENA ANALIZA

Odločitveno analizo [3,8] imenujejo tudi “uporabna odločitvena teorija”. Ponuja vrsto metod in tehnik za reševanje in analizo zahtevnih odločitvenih problemov. Probleme poskuša strukturirati in jih razgraditi na manjše, bolj obvladljive podprobleme. Pri tem upošteva elemente, kot so alternative, med katerimi izbiramo, dostopno informacijo o odločitvenem problemu, znanje in zahteve odločevalca ter poskuša oceniti negotovost in tveganje pri odločitvah. Na ta način dobljene odločitve naj bi bile vsaj v načelu racionalne, vendar se odločitvena analiza ob pomanjkanju popolnih informacij in drugih težavnih pogojih pogosto zadovolji tudi z odločitvami, ki niso povsem optimalne, pač pa le “dovolj dobre” oziroma “zadovoljive”.

Odločitvena analiza predlaga sistematičen pristop k reševanju odločitvenih problemov. Odločitveni proces razdeli na posamezne faze, kot so:

1. identifikacija odločitvenega problema,
2. identifikacija alternativ,
3. razgradnja problema in modeliranje,
4. vrednotenje, analiza in izbira alternativ,
5. realizacija odločitve.

V osnovi si faze sledijo zaporedno, vendar se lahko tudi prepletajo in ponavljajo. V njih poskušamo celotni odločitveni problem čim bolj natančno opredeliti (kaj je predmet odločanja, kaj so cilji in posledice odločitve, kakšne in katere alternative pridejo v poštev, kje so viri negotovosti in tveganja, itd.), ob tem pa poskušamo strukturirati in razgraditi odločitveni problema na manjše in lažje obvladljive podprobleme.

Za odločitveno analizo je predvsem značilna tretja faza, v kateri na osnovi zbranih informacij izdelamo enega ali več modelov, ki jih v četrti fazi uporabimo za vrednotenje alternativ ter različne analize in simulacije, kot sta analiza *kaj-če* in simulacija Monte Carlo, ter na tej osnovi izberemo najboljšo alternativo.

Za primer si oglejmo dva preprosta odločitvena modela za izbiro zaposlitve. Mladi ekonomist je za svojo prvo zaposlitev dobil štiri ponudbe, imenujemo jih A, B, C in D. Tabela 2 prikazuje eno najbolj osnovnih metod odločitvene analize: metodo *primerjave alternativ po parih*. V tabeli opredelimo preferenčne relacije med alternativami. Številka 1 pomeni, da imamo alternativo, ki je napisana v levem stolpcu, raje od alternative, napisane v glavi tabele. Številka -1 označuje ravno obratno. Številka 0 pomeni indiferenco, torej da sta obe alternativni enakovredni. Iz tabele lahko z nekaj logičnega sklepanja razberemo, da je odločevalcu najbolj všeč ponudba B, sledita enakovredni ponudbi A in D, najslabša pa je ponudba C.

Tabela 2: Matrika primerjav med ponodbami za službo.

Ponudba	A	B	C	D
A		-1	1	0
B			1	1
C				-1
D				

Drugi primer (tabela 3) pa prikazuje model vrednotenja ponudb, ki poteka na osnovi več parametrov. (uporabili smo metodo Kepner-Tregoe [5]). Vrednosti alternativ so izražene s točkami med 0 in 10, kjer 0 pomeni zelo slabo, 10 pa zelo dobro oziroma idealno vrednost. Podobno so določene uteži, kjer 10 označuje najbolj pomemben parameter, 0 pa takega, ki sploh ne vpliva na odločitev. Končna ocena alternativ je vsota zmnožkov med utežmi in vrednostmi parametrov. Tudi v tem primeru je najboljša ponudba B.

Tabela 3: Večparametrsko vrednotenje ponudb.

Utež	Alternativa Parameter	A	B	C	D
		10	Plača	8	10
7	Zanimivost dela	4	8	2	6
5	Lokacija	4	2	9	1
5	Varnost zaposlitve	4	6	9	2
4	Ugled	8	9	7	7
3	Napredovanje	6	4	8	10
3	Sodelavci	2	0	4	8
	Ocena	204	244	228	189

Takšne večparametrške modele lahko zelo koristno uporabimo za različne analize, na primer tako, da variramo vrednosti v tabeli in med seboj primerjamo rezultate. V ta namen lahko uporabimo tudi računalniške programe, ki so bodisi splošnonamenski ali specializirani. Splošnonamenski programi so elektronske preglednice ali matematični programi, v katerih uporabnik bodisi sam definira potrebne postopke in metode za delo z odločitvenimi modeli bodisi uporabi ustrezne že realizirane predloge ali dodatke. V ta okvir spadajo programi Microsoft Excel, OpenOffice.org, MATLAB in Mathematica. Specializirani programi v splošnem podpirajo: (1) zajemanje, oblikovanje in preoblikovanje modela vrednotenja in njegovih komponent, (2) zajemanje podatkov o alternativah, (3) vrednotenje in analizo alternativ ter (4) pripravo grafičnih prikazov in poročil. Znani programi te vrste so na primer [2] TreeAge Pro, DPL, Analytica, Hiview, ELECTRE, Expert Choice, Criterium Decision Plus, HIPRE, Web-HIPRE in domači DEXi.

5 SISTEMI ZA PODORO PRI ODLOČANJU

Sistemi za podporo odločanja (angl. Decision Support Systems, DSS) so interaktivni računalniški programi, ki pomagajo odločevalcem pri uporabi podatkov in modelov

za spoznavanje in reševanje odločitvenih problemov. Glavne značilnosti takšnih sistemov so:

- vsebujejo podatke in modele;
- namenjeni so predvsem za pomoč menedžerjem pri reševanju delno strukturiranih in nestrukturiranih odločitvenih problemov;
- podpirajo odločevalca: mu pomagajo z izbiro in prikazom informacij, vendar ga ne nadomeščajo;
- bolj kot učinkovitosti so namenjeni izboljšanju kvalitete in uspešnosti odločitev.

Takšni programi lahko pomagajo odločevalcem na najrazličnejše načine. Že to, da hranijo podatke in so jih sposobni poiskati, ko jih uporabnik potrebuje, lahko pomaga pri odločanju. S podatki lahko tudi računajo, na primer združujejo (agregirajo) v različne statistične kazalnike, kot so vsota, srednja vrednost in trend. Ponudijo lahko številne metode in tehnike za pregledovanje in analiziranje podatkov. Podatke in izračune lahko uredijo in prikažejo v obliki poročil, vrtilnih tabel ali grafikonov. V podatkih lahko iščejo vzorce ali zakonitosti in jih izrazijo z različnimi modeli. Omogočajo uporabo različnih modelov za razvrščanje podatkov in vrednotenje alternativ. Rešitve odločitvenih problemov pomagajo iskati s pomočjo optimizacijskih ali simulacijskih modelov ter z logičnim sklepanjem na osnovi podanih odločitvenih pravil. Omogočajo povezovanje različnih informacijskih virov in z ustreznimi komunikacijskimi sredstvi podpirajo sodelovanje odločevalcev in odločitvenih skupin. Gre torej za zelo široko in raznovrstno paletu računalniških orodij in pristopov. Ti lahko vključujejo tudi metode odločitvene analize.

Sisteme za podporo odločanja v splošnem delimo glede na prevladujoče področje oziroma komponento, s katero s ukvarjajo: podatki, modeli, procesi ali komunikacije [7]. Poleg tega obstajajo tudi tako imenovani generatorji, ki omogočajo razvoj namenskih sistemov za podporo odločanja. Med sisteme za podporo odločanja uvrščamo tudi razne bolj ali manj specializirane informacijske sisteme [9], kot so upravljalski informacijski sistemi (angl. Management Information Systems, MIS), direktorski informacijski sistemi (Executive Information Systems, EIS), geografski informacijski sistemi (GIS), sistemi tipa OLAP (On-Line Analytical Processing), sistemi za odkrivanje znanja (Knowledge Discovery Systems) in nekateri tipi ekspertnih sistemov (Expert Systems, ES).

V skupino sistemov za podporo odločanja sodijo tudi sistemi za podporo skupinskega odločanja (angl. Group DSS) in sistemi za podporo skupinskega dela (Group Support Systems, GSS). To so interaktivni računalniški programi, ki omogočajo reševanje nestrukturiranih odločitvenih problemov, pri katerih sodeluje več odločevalcev, ki delujejo v skupini. Takšni sistemi pomagajo skupinam pri analizi odločitvenega problema in iskanju možnih rešitev. Pri tem morajo poleg podatkov in modelov upoštevati in podpreti tudi dinamiko skupinskega

odločanja. Za to ponujajo programe, ki omogočajo medsebojno sodelovanje, usklajevanje in komunikacijo med udeleženci, na primer prek elektronske pošte, sistemov za posredovanje sporočil ali videokonferenc.

6 ZAKLJUČEK

Računalnik je zmogljivo orodje, s katerim si lahko pomagamo tudi pri reševanju bolj ali manj zahtevnih odločitvenih problemov. Pomoč pa je učinkovita le, če dobro poznamo konkretni odločitveni problem in njegove lastnosti ter se zavedamo zmožnosti, ki jih ponujajo sodobni sistemi za podporo pri odločanju. V tem prispevku smo predstavili pogled na odločanje, ki ga ima nanj večina tehnično-računalniških disciplin. Te razumejo odločanje kot celovit proces pridobivanja, preverjanja, modeliranja in uporabe znanja, ki vodi do odločitve in naprej do njene realizacije. Odločitveni problemi se med seboj razlikujejo po številnih lastnostih, ki jih je potrebno upoštevati pri njihovem reševanju in računalniški podpori.

Odločitvena analiza temelji na izdelavi odločitvenih modelov, s katerimi vrednotimo, razvrščamo in analiziramo odločitvene alternative. Zahteva pazljivo preučevanje in razgradnjo odločitvenega problema, obenem pa izboljšuje dokumentiranost in preglednost odločitvenega postopka.

Sistemi za podporo pri odločanju so specializirani informacijski sistemi, ki pomagajo uporabnikom pri reševanju odločitvenih problemov na različne načine: s podatki, grafikoni, različnimi analizami. Poleg podatkov praviloma vsebujejo tudi odločitvene modele in spodbujajo komunikacijo odločevalcev med seboj in z drugimi udeleženci odločitvenega procesa.

LITERATURA

- [1] Bohanec, M. (2006): *Modeli in odločanje*. DMFA – založništvo, Ljubljana, 2006.
- [2] Bohanec, M.: *IJS decision support resources*. Spletna stran <http://kt.ijs.si/MarkoBohanec/mare.html>
- [3] Clemen, R. T. (1996): *Making Hard Decisions: An Introduction to Decision Analysis*. Duxbury Press.
- [4] French, S. (1986): *Decision Theory: An Introduction to the Mathematics of Rationality*. Ellis Horwood Halsted Press.
- [5] Kepner, C. H., Tregoe, B. B. (1981): *The New Rational Manager*. Princeton Research Press.
- [6] Omladič, V. (2002): *Matematika in odločanje*. DMFA – založništvo, Ljubljana.
- [7] Power, D.J. (2002). *Decision support systems: concepts and resources for managers*. Westport, Conn., Quorum Books.
- [8] Skinner, D. C. (1999): *Introduction to Decision Analysis*. Probabilistic Publishing.
- [9] Turban, E., Aronson, J., Liang, T.-P. (2005): *Decision Support Systems and Intelligent Systems*. Seventh Edition, Prentice Hall.

DO KOD SEŽEJO MATEMATIČNI MDELI V SITUACIJAH ODLOČANJA

Žiga Knap

Oddelek za filozofijo, Filozofska fakulteta, Univerza v Ljubljani

Aškerčeva ul.2, SI-1000 Ljubljana, Slovenija

Tel: +386 1 241 11 02;

Prijateljjeva 11, SI-1000Ljubljana, Slovenija

Tel: +386 1 252 10 49

e-mail: ziga.knap@guest.arnes.si

IZVLEČEK

Predstavljen je poskus utemeljitve ocene, do kod sežejo matematične metode, matematični modeli pri sprejemanju odločitev (=odločanju). Poskus odgovora na principialno vprašanje, ali se vsaka situacija odločanja lahko modelira z adekvatnim matematičnim modelom, poskušamo najti v analogijah s Churchovo tezo in osnovno tezo o upravljavskih (kibernetskih) sistemih (Yablonskij).

ADEKVATNOST MATEMATIČNEGA MODELA. V trditvi, da lahko vsak problem odločanja (odločitev) izrazimo z adekvatnim matematičnim modelom je vsa težava skrita v besedici »adekvatno«. Ta beseda je v tem kontekstu povezana s semantiko, med tem ko lahko matematični model smatramo za formaliziran, torej se v celoti podreja sintaktičnim pravilom in je konstruiran strogo samo z matematičnimi sredstvi in pojmi.

V matematiki so bili in še so problemi, za katere se je zdelo, da se dajo rešiti s pomočjo primerne algoritma. (Klasičen primer je Evklidov algoritem, ki je hkrati tudi ostal kot primer in vzorec, kaj pomeni algoritem). Ko se konec 19. in v začetku 20. stoletja nekaterih problemov nikakor ni dalo rešiti s pomočjo algoritma, se je oblikovala zamisel, da algoritmično niso rešljivi. Znikanje algoritmične rešljivosti je zahtevalo natančno definicijo algoritma, svojevrsten prehod od intuitivnega razumevanja nekega matematičnega pojma k strogemu matematičnemu opisu tega pojma. V prvi polovici 20. stoletja je bilo predloženih nekaj konkretnih algoritmov, kot so Turingov stroj, algoritem Posta, normalni algoritmi Markova, pozneje se je pokazalo, da so vsi ti konkretni algoritmi enakovredni med sabo. In tu se srečamo s tezo Churcha (ki se sicer nanaša na rekurzivne funkcije, ampak pravzaprav velja za algoritme) in ki nekoliko ohlapne rečeno trdi, da vsak intuitivno razumljen algoritem mere narediti samo to, kar lahko naredijo naštetih konkretni algoritmi. Sredi 20. stoletja, ko je bila na pohodu kibernetika, je Yablonskij formuliral analog tezi Churcha za upravljavske (kibernetske) sisteme, ki pravi, da lahko vsak fizični upravljavski (kibernetski) sistem adekvatno predstavimo s pomočjo matematičnega upravljavadekvatno (kibernetskega) sistema. Matematični kibernetki sistem je

bil konkretno konstruiran s pomočjo strogo matematičnih pojmov. V obeh primerih teze ne moremo dokazati, oziroma »dokaz« gre per negationem, do zdaj še ni bil najden noben primer ne algoritma, ne kibernetkega sistema, ki bi ovrgel eno ali drugo tezo. Ne glede na enako filozofijo obeh tez, pa se teza Churcha nanaša ne matematične objekte, medtem ko kibernetki sistemi modelirajo dele stvarnosti, dele tako imenovanega realnega sveta. In tu vidim tesno povezavo s problemom odločanja, saj je odločanje povezano z nekim delom stvarnosti, pa naj bodo to situacije v ekonomiji, politologiji, sociologiji, biologiji, psihologiji, v vseh konfliktnih ali manj konfliktnih situacijah v katerih se znajde posameznik ali skupina.

Ko poskušam najti odgovor na vprašanje, ali matematični modeli lahko da odgovor na problem odločanja, se mi zdi, postaviti tezo, ki bi bila analogna opisnima dvema tezama problematično. Morda bo naredila sliko jasnejšo obravnava poteka, poti in korakov vključitve matematičnega modela v problem odločanja.

POT DO MATEMATIČNE REŠITVE PROBLEMA ODLOČANJA. Sam proces konstrukcije matematičnega modela vidimo v določenem procesu, ki ga skušamo prezentirati s sledečo shemo. Naj bo intuitivno dana **odločevalska situacija (1)**, ki je opisana v nekem jeziku, poimenujmo ta opis kot **opis situacije (2)** odločanja. Iz opisa situacije odločanja formuliramo **problem odločanja (3)**, ki ga oklestimo in formuliramo **racionalni problem odločanja (4)** iz katerega poskušamo dobiti **formaliziran problem (5)** (formalizacija), ki ga poskušamo predstaviti z **matematičnim modelom (6)** problema odločanja iz katerega dobimo **matematični problem (7)**, ki ga je treba rešiti in v splošnem lahko je rešljiv, lahko pa tudi ni rešljiv in če smo ga uspešno rešili, dobimo **rešitev matematičnega problema (8)**, ki spet v splošnem lahko je (9), lahko pa tudi ni rešitev (10) intuitivno dane odločevalske situacije oziroma problema odločanja.

Predlagana shema predstavlja linearno serijo blokov, vendar iz vsakega bloka izhajata dve »puščici«, ena se zapiči v naslednji blok, druga pa izzveni v prazno, kar shematsko dopušča, da se proces racionalnega modela sprejemanja odločitev na tistem mestu prekine. (Problem validnosti in zanesljivosti modela v samem prelaganem postopku/ shemi

ni razviden, ne glede na to, da je prisoten, saj na koncu postopka ugotavljamo, koliko je matematična rešitev adekvatna rešitvi odločanja.) (1) – intuitivno dana odločevalska situacija se nanaša na situacijo v kateri je potrebno sprejeti odločitev, neizdiferencirano odločanje, ki še ni povezano zo predeljeno jezikovno formulacijo. Zdi se, da je potrebno predvsem za neke vrste ontološko ozadje, saj v kakršnikoli formulaciji problema odločanja imamo že opravka z jezikovno formulacijo. (2) – v opisu odločevalske situacije imamo opravka z nekim tekstom, navadno v jeziku, ki je blizu naravnemu, vsakdanjemu jeziku, čeprav je to lahko že zelo ozek strokovni jezik, na primer, ko zadamo neko matematično nalogo. (Zenonove aporije so zadane v te sorte jeziku. Staroveški problem odločitve razdelitve dediščine 15 ovac med 4 sinove, tako da dobi najstarejši polovico, drugi sin polovico manj kot najstarejši itd. jezik ni matematičen, čeprav je sam problem prof. Vidav predstavil kot primer tipično matematičnega problema.) Povezava med(1) in (2) je odvisna od semantike, semantičnega bloka. Čeprav je v tem koraku že lahko viden matematični problem, to ni nujno. (3) – opis odločevalske situacije je tista osnova iz katere izločimo problem odločanja, ki je seveda odvisen od namenske funkcije (funkcije koristnosti, koristi). Tudi v tem koraku je poudarjena formulacija problema v nekem standardnem, sprejemljivem jeziku komuniciranja, ki pa ni nujno niti racionalen in tudi ne nujno matematičen. (4) – racionalen problem odločanja (racionalnost, glej Newton-Smith (1981)) pomeni poskus iz opisa problema odločanja izluščiti racionalen opis, formulacijo. (5) – iz racionalnega opisa problema odločanja poskušamo problem formalizirati, formaliziran opis problema bo osnova za konstrukcijo matematičnega modela. (6) – konstrukcija matematičnega modela problema odločanja. Še drugače povedano, to pomeni formulirati problem odločanja v matematičnem jeziku. Kakšen je ta jezik, je odvisno od samega problema odločanja in videnja, kako tak problem potopiti v matematiko. Na razpolago je vsa matematika in seveda tudi možnost postavitve novega matematičnega problema. (7) – matematičen problem- po poti vseh korakov je bila v ozadju misel, da se pri transformaciji v prehodu od enega koraka k drugemu ohranja bistvo problema odločanja, kar je lahko res ali pa tudi ne. (8) – matematični problem lahko znam rešiti ali pa tudi ne. Lahko pa sploh nima rešitve. Ali v slučaju, da matematični problem nima rešitve, tudi odločevalski problem nima rešitve? Najbrže je to dokaz, da ne eksistira racionalne rešitve odločevalskega problema. Pa tudi, ko matematični problem znam rešiti. S tem še ni rečeno, da je res rešen tudi problem odločanja. Ugotoviti moram, ali je rešitev matematičnega problema tudi rešitev odločevalskega problema. Imam opravka s problemom vlidnosti celotnega postopka odločanja. Vračam se k bloku (1) in semantičnem bloku, ki morata podpreti ali ovreči validnost rezultata. Naj zaključim s citatom iz knjige Hanssona: *Decision Theori* (Hansson (2005)). » Teorija odločanja je teorija o odločanju. Predmet ni zelo poenoten. Nasprotno, mnogo je poti poti

teoretiziranja o rešitvah in zato tudi mnogo različnih tradicij...soraj vse, kar človek počne, vključuje odločanje«

Literatura

- Hansson Sven Ove (2005) *Decision Theory. A Brief Introduction*, Royal Instituta of Technology (KTH), Stockholm
- Kleene Stephen Cole (1952) *Introduction to Metamathematics*, New York
- Neumann J. von, Morgenstern O. (1944) *Theory of games and economic behavior*, Princeton University press
- Newton-Smith W. H. (1981) *The rationality of Science*, Boston, etc.
- Omladič Vesna (2003) *Matematika in odločanje*, Ljubljana
- Yabonskij Sergej (1959) *Osnovnye ponyatiya kibernetiki, Problemy kibernetiki 2*, (str. 7-38) Moskva

NIKOTINSKA ZASVOJENOST GOSPODA JONESA

Odločanje zgodovinarja v procesu raziskovanja

Dr. Mateja Ratej

Sekcija za interdisciplinarno raziskovanje ZRC SAZU

Odločanje je z vidika zgodovinarskega poklica nadvse pomembna, bržkone celo ključna miselna operacija. Tako imenovana dejstva o preteklosti je namreč potrebno med seboj povezati, povezave pa so zmeraj plod zgodovinarjeve domišljije, osebnostnih značilnosti, intimne naklonjenosti zgodovinskim akterjem in v končni fazi – odločitve o tem, kaj, kako, zakaj se je zgodilo. Zadeva je sila kočljiva, saj želi biti zgodovina eksaktna znanost, katere dognanja temeljijo na uveljavljenih metodoloških postopkih, kjer skratka za arbitrarno intimo zgodovinarja ni mesta. Zatorej so vprašanja o tem, v kakšni meri je rekonstrukcija posameznih preteklih dogodkov plod zgodovinarjeve subjektivne oz. objektivne sodbe, vselej aktualna. No ja, roko na srce, večini resnih zgodovinarjev je že zelo dolgo jasno, da objektivna zgodovina ne obstaja, saj so ljudje umeščeni v prostor in čas, iz katerega izhajajo, do te mere, da jim spoznanje absolutne resnice ni dostopno. S tem seveda ne želijo trditi, da resnica ne obstaja, temveč da lahko zgodovinarji napišemo le delne in nepopolne približke resnici. Prepričanje o nevtralnosti dejstev in vera v vzgojne učinke zgodovine pa izhajata iz racionalizma preteklih stoletij. Če je tako sedanjost določene skupnosti do neke mere resda rezultat preteklega dogajanja, je še toliko bolj res, da sedanjost nenehno prilagaja preteklost svojim potrebam.

Edward H. Carr, ki mu deloma pripada zgornje razumevanje zgodovine, je bil eden najprodornejših zgodovinarjev 20. stoletja. Zgodovine ni nikoli študiral ali poučeval, temveč se je po diplomu iz klasičnih ved zaposlil na britanskem Foreign Officeu. Da se kot raziskovalec ni oblikoval v zaprtem in sterilnem strokovnem diskurzu, mu je več kot koristilo pri razumevanju umeščenosti historične vede v javni (tj. politični) prostor. In vprašanje zgodovinarjeve objektivnosti ga je resnično mučilo. Mučilo tako zelo, da si je izmislil hipotetično zgodbo in z njo za nekaj desetletij dodobra stresel svetovno zgodovinarsko sceno, da je nemimo valovila nekoliko v smeri navdušenega pritrjevanja, predvsem pa v smeri vzvišenega zavračanja. Takole gre: *»Po neki zabavi, kjer je zvrnil nekaj kozarčkov čez mero, je gospod Jones sedel v svoj avto s pokvarjenimi zavorami in na nekem križišču do smrti povozil gospoda Robinsona, ki je šel v bližnjo trafiko kupit cigarete. Kaj je povzročilo nesrečo? Vinjenost gospoda Jonesa, pokvarjene zavore, slaba vidljivost ali nikotinska zasvojenost gospoda Jonesa?«*

Slovenski javni prostor je poseben tudi po tem, da pogosto problematizira intimne ideološke pozicije zgodovinarjev. Medtem se prvovrstni italijanski in svetovno uveljavljen zgodovinar Giovanni Levi samo nasmehe, ko mu omeniš

bajko o nevtralnih (v teh krajih se rado reče nepristranskih) miselnih pozicijah zgodovinarjev. Osebna arheologija zgodovinarjev Leviju še malo ne dela preglavic. Nasprotno, prepričan je, da je njihov (in naš) skupni prostor polje (miselnega) spopada in ne polje konsenza. Iskanje odgovora na vprašanja, kaj imamo kot skupnost skupnega v kulturnem smislu, kaj je legitimno, kaj nas mora povezati v skupni temelj vrednot, je pač predmet neprekinjene vojne reprezentacij preteklosti. Ko me je Levi pred nekaj leti pričakal s cigaro med zobmi v Benetkah, od koder je doma, in me nato v labirintu ozkih ulic pripeljal v mično in tihotno kavarno, je ponovno nasprotoval avtoritarnemu slogu pisanja, ki ga zavzemajo zgodovinarji, ko v glavnem pripovedujejo o preteklih dogodkih tako, kakor bi opisovali objektivno realnost. Kajti zgodovinarji le redko napišejo kaj o težavnem procesu raziskovanja, kar bi dalo njihovim zaključkom pridih hipotetičnosti in nezaključenosti. In medtem ko zgodovina dobesedno ignorira sodobne družboslovne kritike konceptov racionalizacije, so druge humanistične vede že zdavnaj zrele vizirale avtoriteto avtorja.

A četudi izgleda položaj za zgodovinsko vedo v začetku 21. stoletja naravnost brezupen, je takšen le na prvi pogled. Izhodnih strategij je na polju teorije zgodovinarstva več kot dovolj. Giovanni Levi, na primer, predlaga zgodovinarjem, naj raje kot da vlagajo miselne napore v vzročno posledično nizanje dejstev, preverijo, v kolikšni meri je nek dogodek (predvsem) stvar splošnega interesa, ki zadeva obče človeške potrebe in strahove v skupnosti. Kajti dogodek ni nujno nekaj, kar se je zgodilo, temveč nekaj, kar so ljudje pripravljene sprejeti kot takšno. Spraševanje o tem, kako so se stvari dejansko zgodile, zato Levi razume kot postopek, ki naj bi obvladal siceršnji nepregleden in pogosto nerazumljiv kaos bivanja. Če zdaj do kraja postrgamo kožo z zgodovinarjevega hrbta in ga dokončno razgalimo, lahko rečemo, da je zgodovina nepretregano reinterpretiranje nečesa, kar je že znano. Temelj zgodovinarjevega dela še zdaleč ni vračanje nečesa pozabljenega v zavest ljudi, pri čemer je Levi celo prepričan, da so zgodovinarska odkritja slejkoprej stvar naključja. Zgodovinska resnica pa je bolj ali manj znana oz. se poraja iz novosti v vprašanih, ki si jih zastavlja vsakokratna generacija. Ker zgodovina razvršča dogodke od bolj do manj pomembnih, ker izbira pravilne razlage in s tem ustvarja parcialno podobo preteklosti, zbuja videz relativnega reda in skupnega kulturnega prostora. Kar ni malo.

Johan Gustav Droysen (1808–1884) je v svojem orisu historične metode zapisal, da je zgodovina edina znanost, ki uživa to negotovo srečo, da je znanost in umetnost hkrati. Pa poglejmo torej, kaj se je v 19. stoletju zgodilo z zgodovinsko vedo, da je pričela obsedeno braniti status dejstev in ekskluzivne resnice. Zgodovinopisje se je v 19. stoletju razvijalo pod vplivom industrijske revolucije, nagle tehnične modernizacije, velikih znanstvenih dosežkov in splošnega preoblikovanja evropskega političnega prostora po zedinjenju Italije in Nemčije – Evropa je v tistem času vse bolj dobivala podobo zbirke nacionalnih enot. V prvi polovici stoletja je trg preplavila množica zgodovinskih del, ki so zgodovino pričela dojemati kot razvojni proces, ki ga določa notranja logika in ni zgolj zapovrstno nastevanje dogodkov. Deloma se je takšna strast po zbiranju ostankov preteklosti porodila iz prizadevanja, da bi preteklost zavarovali pred napadom mehanizirane sedanosti, glavni vir zbirateljske strasti pa je bil – NACIONALNOST. Pri ljudeh, namreč, ki do takrat še niso doživela prebuje nacionalne zavesti, so postali avtorji besednjakov in zbiralci ljudskih pesmi dejanski utemeljitelji narodne zavesti.

Če nosi HERODOT naslov očeta zgodovine, VOLTAIRE pa ime očeta moderne zgodovine, smo zdaj prispeli do tretjega nosilca prestižne titule. Leopold Von Ranke (1795–1886) alias: oče znanstvene zgodovine. Ranke je bil leta 1841 imenovan za zgodovinarja pruske države, leta 1865 pa je bil povzdignjen v plemiški stan. Poleg tega, da se je ukvarjal z zgodovino papežev, germanskih in romanskih ljudstev, z nemško zgodovino v času reformacije, s sorodnostmi oz. razlikami med zgodovino in politiko ipd., je učil, da je potrebno zgodovino pisati samo na podlagi poročil prič in neposrednih dokumentov, saj nam lahko samo ti razkrijejo univerzalne odnose. Zaradi njih je zgodovinska veda univerzalna znanost, ki jo sestavljajo trije temeljni elementi: KRITIČNA METODA, OBJEKTIVNO PROUČEVANJE in OBLIKOVANJE SINTEZE. Tu ni prostora za sodbe, tudi ne za nasvete sedanosti, ki naj bi koristili prihodnosti. »Moje delo želi pokazati samo, kako je dejansko bilo,« je njegova največkrat citirana predpostavka. WIE ES EIGENTLICH GEWESEN! Kot temeljne vire je Ranke definiral: dnevnike, pisma, diplomatska poročila, originalne pripovedi prič itd. Menil je, da morajo zgodovinarji poleg vernega opisa dejstev izluščiti splošne tendence in vodilne ideje, ki so posamezniku, ustanovi, državi, obdobju dale značaj.

Spoznanja zgodovinarja morajo služiti predvsem NARODU, je bil prepričan – v tem pogledu je bil Ranke človek Herderjeve definicije, po kateri je narod vir vse resnice in vseh vrednot, ki lahko kot takšen sam presodi, kaj je dobro zanj. V nasprotju z razsvetljenci, ki so verjeli v razum, so historicisti (ime se jih je prijelo zaradi stalnega sklicevanja na zgodovino) edino narodu pripisali relativno stabilnost. Leopold von Ranke je bil zavoljo svojih stališč in mestoma njihove napačne interpretacije še pred sredino 20. stoletja močno kritiziran zgodovinar, v drugi polovici 20. stoletja pa je izraz *rankejansko zgodovinopisje* postal sinonim za

zastarelo, suhoparno in naivno interpretacijo preteklosti. Ne glede na to sta bila njegov IDEAL OBJEKTIVNOSTI (zgodovinar mora opisati vse tako, kot se je zgodilo, pri tem pa ga ne smejo bremeniti lastna stališča) in KRITIČNA METODA še daleč v 20. stoletje temelj modernega zgodovinopisja. Ne glede na to moramo ugotoviti, da 19. stoletje kljub preboju rankejevskega koncepta zgodovinopisja ni izzvenelo v nacionalistični evforiji, temveč bolj v pozivu k emancipaciji, k odpiranju novih in novih vprašanj, ki so zgodovino močno približala družboslovju, saj je tudi zgodovinopisje pričelo gledati na človeško preteklost kot na preteklost družbenih procesov.

Na zgodovinopisje je tako na prelomu stoletja 20. stoletja močno vplival sociolog Max Weber (1864–1920), ki je videl osnovno nalogo zgodovine v tem, da dogodke in procese predstavlja in interpretira z vidika njihovega kulturnega pomena; izjemno skeptičen je bil do zgodovinarskega prisvajanja resnice, saj je bil prepričan, da predstavljeni resničnosti zmeraj dodamo tudi celoto našega 'nomološkega' izkustvenega znanja. Z drugimi besedami: OB DOSLEDNEM UPOŠTEVANJU KRITIČNE ANALIZE ZGODOVINSKEGA GRADIVA JE ZGODOVINARJEV SKLEP VERJETNA PODOBA PRETEKLEGA DOGAJANJA.

V drugi polovici 20. stoletja pa je Rankejevemu empirizmu najvidneje nasprotoval ravno E. H. Carr, ki ga je označil kot naivnega, dolgočasnega in zastarelega. Ena najbolj razvpih debat na to temo v okviru angleškega zgodovinopisja se je odvijala med Carrom in Geoffreyem Rudolphom Eltonom in jo je iniciralo Carrovo delo *What is History?*, ki je prvič izšlo leta 1961. V njem je Carr med drugim zagovarjal tezo, da so toliko opevana dejstva pravzaprav preprosto tista, ki jih zgodovinarji izberejo in definirajo kot takšna. Milijoni so prečkali Rubikon, razlaga Carr, vendar zgodovinarji samo Cezarjevo prečkanje slavne reke interpretirajo kot signifikantno. Vsa zgodovinska dejstva so samo rezultat interpretativne izbire zgodovinarjev, na katere vplivajo norme, vrednote, pravila itd. dobe, iz katere zgodovinarji izhajajo.

Pomembno je tudi, da so zgodovinarji v osemdesetih letih prejšnjega stoletja generalno pričeli poudarjati, da so pretekle izkušnje neposredno vpete v JEZIK. Povedano drugače: zgodovinarji so se pričeli ob tematizaciji dejanskega stanja in izpraševanju empirično preverljivih pogojev dejanske zgodovine posvečati tudi jezikovnim zmožnostim zgodovinarjevega opisa. Ugotavljali so, na kak način zgodovinar oblikuje svojo zgodbo (zgodovino) in na kakšen način jo posreduje bralcem. Odgovorili so na nekatera ključna vprašanja zgodovinopisja: kakšna je struktura historične zavesti, na čem temeljijo najpogostejše oblike historične predstavitve, kakšen spoznavnoteoretski status imajo specifične historične razlage v primerjavi z alternativnimi vrstami razlag, s katerimi je prav tako mogoče interpretirati historično gradivo. Nekje pred koncem 20. stoletja je bilo večini zgodovinarjev jasno tudi, da

vsakokratnih oblik historične interpretacije ne gre razumeti kot nekaj dokončnega, temveč jih moramo obravnavati kot dinamični (dialogični ali retorični) niz jezikovnih protokolov, kot mrežo komunikacijskih besedil.

A obenem so se prav tedaj zgodovinarji ponovno začeli v velikem številu zatekati k izključujočim sodbam, namesto da bi historični predstavitvi dokončno odrekli status objektivnega prikaza in se posvetili mnogoterim pogledom človeka na realnost – samo tako bi se namreč lahko izvili iz zahteve po izdelavi občeveljavnih dejstev. Potem ko je razpravo o statusu historične predstavitve dolgo časa obvladovala filozofija zgodovine, so začeli proti koncu 20. stoletja tudi tisti zgodovinarji, ki so dotlej trdili, da je epistemološka samorefleksija zgodovinopisja stvar filozofije, znova tematizirati razmerje med preteklostjo in zgodovino. Ena njihovih temeljnih ugotovitev je bila, da je potrebno preteklost obravnavati kot besedilo, oz. da je zgodovina REPREZENTACIJA, tj. predstavitev ne le obdobja, temveč tudi tistega, ki o njem piše in obdobja, iz katerega piše. Strinjali so se tudi, da je reprezentacija v globokem razmerju z jezikom kot medijem, skozi katerega je šele skonstruirana realnost, ki je razumljiva točno zato, ker je skonstruirana.

Če nadaljujemo s tezo sodobne historične antropologije, da je imaginarno prav tako imaginativno kot realno (da je potemtakem fikcija možna tudi znotraj resnice) in da znakovne kode obvladujejo prav vse reprezentacije, zgodovinopisje zlahka pripeljemo do lastnega razpada. Tako ni čudno, da se zgodovinarji (v teh krajih premalokrat) sprašujejo, ali so nam instrumentalne uporabe jezika sploh zmožne posredovati védenje o zgodovini. Ali je sploh mogoče govoriti o zgodovinski resnici? Dekonstruktivistični tezi, da onkraj besedila ni ničesar in je vsak zapis odprt interpretaciji, se je, kot kaže, učinkovito postavila po robu ena najvidnejših ameriških zgodovinarok Gabrielle M. Spiegel. Na tem mestu se ne bomo podrobneje spuščali v zgodovinarokino teoretsko pozicijo, zaključimo le z njenim bistvom, da je ob jasnem protokoliranju jezikovnih modifikacij, ki konstruirajo vsakokratni pomen besed, tj. s pridobitvijo osnov kredibilnosti, možno do neke mere verodostojno rekonstruirati preteklost.

Literatura

- E. H. Carr, *What is History*, London: Penguin Books, 1990.
- Oto Luthar, Marjeta Šašel Kos, Nada Grošelj, Gregor Pobežin, *Zgodovina historične misli; od Homerja do začetka 21. stoletja*, Ljubljana: Založba ZRC, ZRC SAZU 2006.
- Giovanni Levi, *Nematerialna dediščina; življenjska pot piemontskega eksorcista iz XVII. stoletja*, Ljubljana: ŠKUC, Filozofska fakulteta, 1995.
- Giovanni Levi, *O zgodovinski resnici*, v: *Časopis za kritiko znanosti*, št. 207–208/2002.
- Keith Jenkins, *Re-thinking History*, London – New York: Routledge, 2003.
- Igor Grdina, *Poti v zgodovino*, Založba ZRC, ZRC SAZU, 2003.

SITUACIJA “ODLOČANJE” IN NJENA REPREZENTACIJA V ZNANSTVENIH TEKSTIH

Elena Bazhenova, Marija Kotyurova

Kafedra ruskega jezika i stilistiki, Filologičeskij fakultet

Permskij gosudarstvennyj universitet

Bukireva, 15, 614990 Perm, Ruska federacija

Tel: +7 342 2 396 256

Fax: (7)(342) 237 16 36

e-mail: bazhenova_e2000@mail.ru; kotyurova@yandex.ru

ABSTRACT

“Decision-Making” Situation and its Representation in Scientific Texts The “decision-making” situation is regarded as cognitive and communicative component, which is realized in scientific texts. This situation reflects the main scientist’s cognitive intentions of research. The “decision-making” is expressed in a text by special words (mostly mental verbs).

V središču zanimanja kognitivne lingvistike je jezik kot sredstvo kodiranja in reprezentacije misli: mentalne osnove produkcije in razumevanja govora, udeleženos jezika v predelavi in prenosu informacije in drugo. Za raziskovalce znanstvenega teksta je kognitivna lingvistika pomembna kot instrument raziskovanja spoznavne dejavnosti in njenega proizvoda, racionalnega znanja, ki je realizirano v govornem proizvodu.

V zavesti znanstvenika je vsebina spoznavne dejavnosti prezentirana v obliki teoretično osmišljenega vzorca stvarnosti, to je fragmenta znanstvene slike sveta znanstvenika. V tekstu se vsebina zavesti preoblikuje v skadu z linearno strukturo govora, kot niz jezikovnih enot. Ta zakonitost je veljavna tudi v situaciji “odločanja”, ki spremlja spoznavno dejavnost v znanosti - v vsej njeni razsežnosti od znanstvenikove formulacije problema, konstrukcije hipotez do opredelitve zaključkov.

Znanje je možno predstaviti v statističnem in dinamičnem aspektu. Tako je prezentirano v tekstu. Situacija “odločanje” zaradi metonimičnosti nominacije, kot mi vidimo, zavzema dvojni položaj, saj odlikava **process sprejemanja** dlčitev in tudi rezultat tega procesa, **sprejeto odločitev**. Poleg tega, ker je znanje v dinamičnem aspektu v tesnem odnosu s spoznavno dejavnostjo ni to le akt enkratnega značaja, ampak spremlja vse ključne momente te dejavnosti: izbor predmeta raziskovanja, določitev cilja in nalog, izbiro snovi, nahajanje bistva problema, formuliranje hipotez, “pritegnitev” že znanega teoretičnega znanja, ki podpira hipoteze, spisek metod raziskovanja, razdelavo prijemov analize področja raziskovanja in, na kraju, sprejemanje odločitve o izpeljvi določenega eksperimenta.

Vsi ti, tako splošni kot posebni momenti spoznavne dejavnosti, ki jih družijo sprejemanja odločitve, se manifestirajo v znanstvenem tekstu. Zanimivo je, da so reprezentirani ali kot **strateško** pomembni v procesu spoznavne dejavnosti (ti postopki bodo šele realizirani; glej naprimer: *določimo pojem ..., formulirajmo zakon ...*) ali kot aktualni, **taktično** pomembni za bralca. V tem zadnjem slučaju avtor samo konstatira zanj jasne in sprejemljive trditve; glej naprimer: *jezikovne kategorije običajno razumemo kot ..., ta količina je konstanta ...* in podobno.

V znanstvenem tekstu je situacija odločanja tesno povezana z **novim** avtorskim znanjem (to je znanjem, ki ga je avtor osebno pridobil). Hkrati pa je očitno, da znanstveni tekst ne nastopa samo kot nositelj znanja, ampak tudi kot sredstvo znanstvenega komuniciranja. To pa pomeni, da tekst izpolnjuje svojo komunikativno funkcijo samo v tem primeru, če je razloženo znanje adekvatno sprejeto s strani bralca. Upoštevajoč to, avtor “gradi” stvaritev in se pri tem orientira na določen nivo poučenosti potencialnega interpretatorja, in pri tem poskuša v taki ali drugačni meri predvideti njegove možne reakcije na tekst. Z drugimi besedami, tekst se ustvarja z upoštevanjem **deja naslovnika**, ki igra pomembno vlogo pri nastajanju teksta.

Torej, tudi v primeru, ko je avtor prepričan v svoj prav, ko se je torej “odločil”, on prezentira znanje, ki ga je pridobil v obliki **strateško** možnega ali kot nujnega. O tem se prpričamo, naprimer, ko avtor v tekstu uporablja polfunkcionalne glagole v obliki prihodnjega časa naslednjega tipa: *definirajmo, študirajmo, klasificirajmo, podčrtajmo* in mnogo drugih, ki nedvomno vključujejo tudi tak odtenek smisla, kot “odločitev”.

Tako se s pomočjo glagolov *opozorimo o, spomnimo se, vrnimo se k* avtor sklicuje na spomin bralca; z *dopustimo, predstavimo, predpostavimo* pa se avtor obrača na predstavno sposobnost bralca, s *preštudirajmo, preidimo k, pokažimo* pa avtor apelira na brlčvo pozornost. Z glagoli *primerjajmo, soočimo, dokažimo, izračunajmo, razdelimo, izmerimo* pa avtor vzpodbuja bralca k mislenim opracijam; *naredimo sklep, strnimo povedano, posplošimo* pa so glagoli, ki potegnejo bralca v logično dejavnost. Vendar, v

bistvu, vsi glagoli tega tipa eksplicirajo lastno spoznavno dejavnost avtorja.

Opozorimo, da navedeni glagoli niso uporabljeni samo kot izraz dinamike spoznavne dejavnosti, ki vključuje tudi "odločanje", ampak je njihov namen tudi pritegnitev bralca v **skupno** iskanje renice. Poglemo, kako to izgleda v realnem tekstu: **Перейдем** к вопросу о... В дальнейшем **мы будем придерживаться** следующих терминов... Такое разграничение **будет соблюдаться нами** в настоящей работе... Именно на эти важнейшие характеристики **мы будем опираться** при классификации... **Обратим** внимание на... **Поясним** сказанное примером... **Разберем** подробнее этот пример... Далее **рассмотрим**... **Переходя** от... к..., **отметим**... Теперь **обсудим** вопрос о... **Проанализируем** одно из явлений... **Остановимся** более подробно на... **Подчеркнем** значимость... **Обобщая** вышесказанное, **подведем итоги**... Ugotovimo lahko, da naslovljenec tako dobi možnost ne samo skoncentrirati se na pomembne momente razvijanja teme teksta, ampak skupaj z avtorjem prehodi vso pot od ne-znanja k znanju. Še več, pri bralcu vznikne občutek, da se kot udeleženec znanstvene komunikacije ne seznanja samo z rezultati raziskave, ampak jih v nekem smislu sam dosega, to je, sam **odloča**.

Podobni signali za aktiviranje bračeve pozornosti so posebej namenjeni za markacijo »korakov« presojanja, zato njihovo maksimalno gostoto nahajamo v tistih fragmentih teksta, kjer je poudarek na **procesu** iskanja in argumentacije novega znanja (teoretično in empirično fundiranje predlagane teze, dokazovanje hipoteze, rešitev naloge ali enačbe).

V sodobni kognitivistiki, se hkrati s tezo o objektivnem znanstvenem znanju, prav tako poudarja tudi nemožnost izločitve spoznavajočega subjekta iz procesa ustvarjanja znanstvene like sveta: znanstvenik ne nastopa samo kot vodnik logike faktov, ampak tudi kot mislec, ustvarjalec, ki aktivno potrjuje svojo pozicijo v znanosti. Izhajajoč iz tega, v tekstu reprezentacijo situacije »odločanja« neizbežno spremljajo jezikovna sredstva, ki izražajo avtorjevo reflektiranje realizacije njegove spoznavne dejavnosti. Avtorjeva refleksija se sooča tako z modulom dvoma (*verjetno, možno, kaže da* in drugo) kakor tudi z modulom gotovosti, prepričanosti (*brezpogojno, nesporno, nedvomno* in drugo). Podobne jezikovne enote formirajo modalni okvir odločanja. Tako, prvič predstavljeno novo znanje običajno spremlja dvom, skepsa; vendar s poglobljanjem dokazov za pravilnost hipoteze se modalnost menja v smeri večje gotovosti, uverjenosti avtorja v njegov prav; izvorno znanje je, praviloma, »okrašeno« s prepričanjem v njegovo pravilnost.

Menimo, da je na sodobni stopnji razvoja lingvistike teksta pojasnjevanje govorne specifike znanstvenega dela nemogoče brez upoštevanja osnovnih trditev kognitivne lingvistike. Ta znanstvena smer orientira raziskovalca na nahajanje globinskih povezav med formalno-jezikovno stranjo teksta in specifičnostjo spoznavne dejavnosti znanstvenika.

References

- [1] Баженова Е.А. Научный текст в аспекте политекстуальности. Пермь, Изд-во Перм. ун-та, 2001.
- [2] Котюрова М.П. Стилистика научной речи. М, Academia, 2009.
- [3] Кубрякова Е.С., Демьянков В.З. Когнитивная лингвистика. In Краткий словарь когнитивных терминов. М., МГУ, с. 53–55, 1996.
- [4] Рябцева Н.К. Ментальные перформативы в научном дискурсе. In Вопросы языкознания, № 4, с. 12–27, 1992.

VPLIV MAGNETIZMA NA OBČUTLJIVOST OČI - JE V OZADJU MAGNETNI ČUT?

Viktor Avbelj

Odsek za komunikacijske sisteme

Institut Jožef Stefan

Jamova 39, 1000 Ljubljana, Slovenija

Tel: +386 1 4773824; fax: +386 1 4773111

e-pošta: viktor.avbelj@ijs.si

ABSTRAKT

Veliko živalskih vrst si pri orientaciji v prostoru pomaga z magnetnim poljem Zemlje. Med njimi so nekatere ptice, ribe, žuželke in sesalci. Rezultati nedavnih eksperimentov kažejo, da magnetno polje vpliva na zaznavanje svetlobe pri ljudeh, kar se je pokazalo kot možnost zaznavanja geomagnetizma že v predhodnih študijah vedenja ptic. Enega teh eksperimentov na ljudeh smo z manjšimi spremembami ponovili, da bi ugotovili, ali lahko potrdimo objavljene rezultate. Posebej smo bili pozorni na možnost, da bi zaznavanje neželenih sopojavov z drugimi čuti lahko vplivalo na rezultate. Trenutni izsledki kažejo, da vpliva geomagnetnega polja na zaznavanje svetlobe pri človeku v izbranih eksperimentalnih pogojih ne moremo potrditi.

1 UVOD

Organizmi spreminjajo svoje vedenje pod vplivom magnetnega polja. Je človek pri tem izjema [1]? Čeprav so v mnogih študijah pokazali, da magnetizem vpliva na vedenje živali, dokončnih odgovorov na vprašanja o mehanizmu zaznavanja magnetizma in lokaciji receptorjev še ni [2]. V ospredju sta dve teoriji, pri prvi gre za receptorje, pri katerih kristali magnetita (Fe_2O_3) mehansko vplivajo na nevrone [3], pri drugi gre za svetlobno-odvisno kemično reakcijo na katero vpliva tudi magnetizem [4]. Ker je pri tem nujna svetloba, domnevajo, da so senzorne celice v očeh [5, 6], medtem ko so pri prvi teoriji senzorne celice lahko kjerkoli. Nikakor ni izključeno, da bi ptiči ne imeli obeh mehanizmov, enega v očeh in drugega v kljunu [7].

Nedavno so z opazovanjem goveda na prostem s pomočjo satelitskih slik ugotovili, da se živali postavijo v smeri silnic zemeljskega magnetnega polja [8] in da se to spremeni v bližini daljnovodov [9].

S poskusi na ljudeh je Thoss s sodelavci ugotovil, da je zaznavanje svetlobe odvisno od jakosti in smeri magnetnega polja [10, 11]. Ta njihov poskus [10], ki predpostavlja zaznavanje magnetizma z očmi, smo z manjšimi spremembami ponovili. V nadaljevanju opisujemo postavitev našega eksperimenta, potek in rezultate.

2 METODE

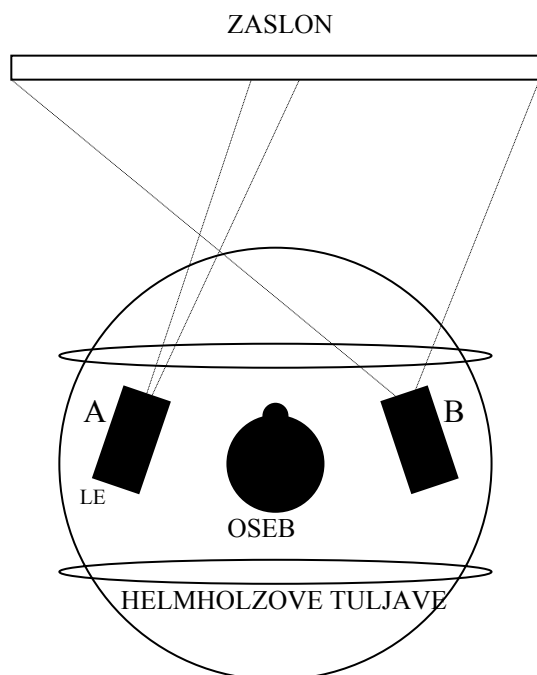
2.1 Osebe

V poskusu so sodelovali študenti elektrotehnike, ki so meritve izvajali v okviru vaj, kjer so imeli možnost izbrati udeležbo v poskusu ali drugačno vajo. Prostovoljci, ki so se odločili za sodelovanje v poskusu, so poskus opravili v dveh skupinah. V prvi je bilo 8 oseb in v drugi 11. Vsi so dobili enaka navodila v pisni obliki, kjer je bil pojasnjen cilj eksperimenta in okoliščine. Posebej je bilo poudarjeno, da ne gre za kakšno tekmovanje ali primerjave med osebami. Niso bili seznanjeni s podrobnostmi eksperimenta, kot je na primer jakost in smer magnetnega polja ali kdaj in kako se v poskusu to spremeni. V navodilu je bilo napisano, da v poskusu jakost magnetnega polja ne preseže jakosti zemeljskega magnetnega polja. Enake informacije je dobila tudi oseba, ki je vodila eksperiment (eksperimentator). S tem smo zadostili pogoj za dvojno slep poskus. Naloga eksperimentatorja je bila, da poskusni osebi praktično pokaže, kako naj ravna v poskusu in potem samo še zažene vnaprej pripravljen postopek, ki je pod nadzorom računalnika tekel samodejno do konca.

2.2 Merilne naprave in potek poskusa

Postavitev naprav je prikazana na sliki 1. Oseba sedi na stolu in v vodoravni smeri opazuje zaslon, ki je oddaljen 1,2 metra. Zaslon je osvetljen s stalno svetlobo svetilke (B). V središče zaslona sveti še utripajoče svetilo (A) - svetlobna dioda z valovno dolžino 590 nm. Utripanje ima periodo 1 Hz in to tako, da je polovico časa svetilo prižgano, polovico pa ugasnjeno. Svetlost svetila določa program v računalniku tako, da je v začetku svetlost 100 enot, potem se povečuje s korakom po 1000 enot, vse dokler poskusna oseba ne opazi utripanja in pritisne na tipko ter s tem spremeni svetlost spet nazaj na 100 enot. Oseba ponavlja vse to dokler ne mine 300 sekund, ko se poskus samodejno konča. Glava osebe je v srednjem delu dveh parov Helmholtzovih tuljav, s katerimi je mogoče spreminjati jakost in smer magnetnega polja. Naprave so v zvočno izoliranem prostoru orientirane tako, da je pogled osebe usmerjen proti jugu. Orientacijo smo določili s kompasom z magnetno iglo.

V začetku poskusa so tokovi v tuljavah nastavljeni tako, da je vertikalna komponenta magnetnega polja enaka nič, horizontalna pa je enako velika kot je magnetno polje Zemlje. Po 150 sekundah je naključno izbrano, da se polje spremeni, ali pa ostane nespremenjeno. Če se spremeni, potem postane še horizontalna komponenta polja enaka nič. Vsaka oseba sodeluje v dveh poskusih, kjer se v enem polje spremeni, v drugem pa ne. Pri vsaki osebi je naključno izbrano ali se to zgodi pri prvem ali pri drugem poskusu.



Slika 1: Postavitev eksperimenta z utripajočim svetilom A in svetilom B, ki osvetljuje celoten zaslon (pogled od zgoraj).

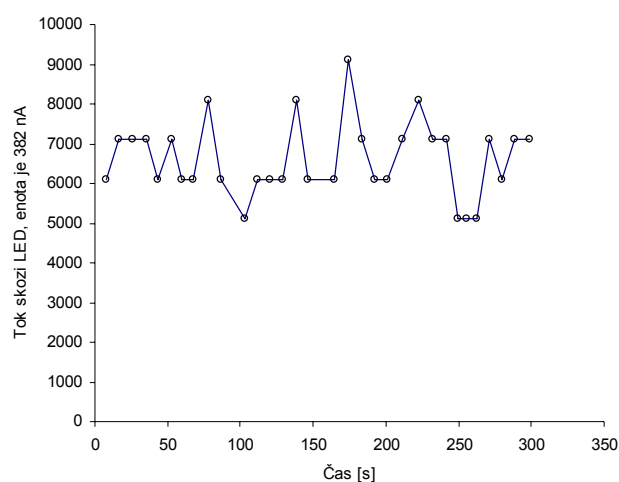
Kadar pride do spremembe tokov v tuljavah, se spremeni tudi toplota, ki jo te oddajajo. Ta toplota je v paru tuljav, ki določa vertikalno komponento polja, enaka ves čas poskusa in znaša 2,63 W. V paru tuljav, ki določa horizontalno komponento, je v začetku poskusa moč 1,22 W, v primerih, ko se jakost polja spremeni, je moč 0,36 W. Tuljave smo navili v plastične cevi z debelo izoliranimi vodniki tako, da bi poskusna oseba ne mogla začutiti razlike zaradi spremembe temperature niti da bi pri tem prihajalo do slišnih zvokov, ki bi lahko zmotili osebo v zvočno izoliranem prostoru. Prav sluh je čutilo z visoko občutljivostjo, ki je komaj izven termičnega šuma molekul v ušesih.

Po izvedbi poskusov prve skupine oseb smo upoštevali pripombe nekaterih, da se je med meritvijo slišalo škrebjanje diska, ki je sproti zapisoval podatke o jakosti tokov v tuljavah in jakosti toka skozi utripajočo lučko ter ostale parametre. Vse krmilne naprave skupaj z

računalnikom smo preselili v sosednjo sobo, kjer je bil v času meritev oseb iz druge skupine tudi eksperimentator.

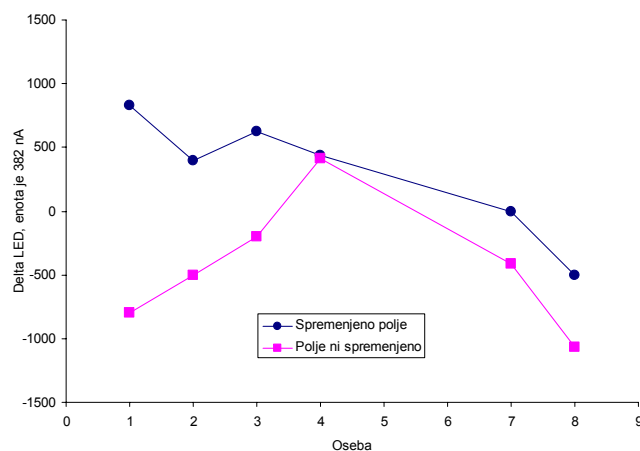
3 REZULTATI

Od osmih oseb iz prve skupine se ena ni udeležila drugega poskusa, pri eni osebi pa je bila prva meritev neregularna. Tako so ostale za nadaljnjo obdelavo še meritve šestih oseb iz prve skupine. Na sliki 2 je prikazan primer izmerjenih tokov, pri katerih je oseba zaznala utripajoče svetilo (LED).

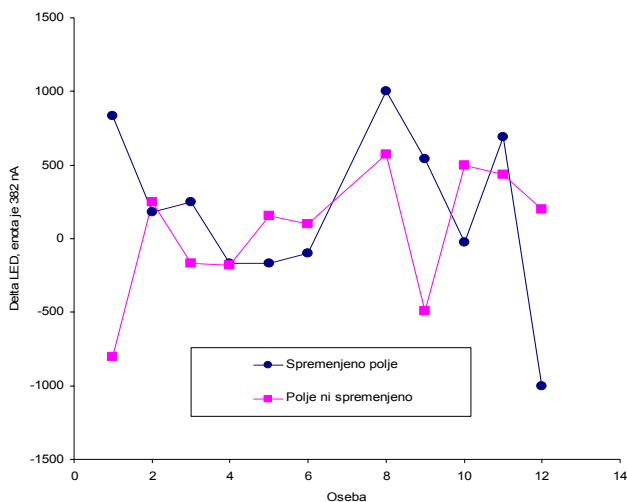


Slika 2: Pet-minutna meritev praga svetlosti, ki jo določa tok skozi utripajoče svetilo (LED). Magnetno polje je 150 s po začetku spremenilo vrednost z začetne, ki je bila enaka zemeljskemu magnetnemu polju na vrednost nič.

Ker pri nekaterih osebah prag svetlosti v začetku meritve ni bil stabilen, smo pri vseh izločili prvih 100 sekund posnetka. Izračunali smo povprečje tokov za interval 150-200 s in za interval 100-150 s. Razlika teh dveh povprečij (Delta LED) je bilo merilo, s katerim smo določali spremembo praga zaznavanja. Rezultati prve skupine oseb so na sliki 3, rezultati druge skupine (11 oseb) so na sliki 4.



Slika 3: Sprememba toka utripajočega svetila (Delta LED) glede na pogoje eksperimenta za prvo skupino oseb.



Slika 4: Sprememba toka utripajočega svetila (Delta LED) glede na pogoje eksperimenta za drugo skupino oseb.

Pri drugi skupini ni statistično značilne razlike v spremembi praga zaznavanja svetlosti (Delta LED) pri poskusih, kjer je bilo magnetno polje spremenjeno. V prvi skupini je premalo podatkov za tak izračun.

4 ZAKLJUČEK

Ob spremenjenih pogojih smo ponovili poskus, katerega rezultate je objavil Thoss s soavtorji leta 2007 [10]. V naših eksperimentalnih pogojih ne moremo potrditi, da sprememba magnetnega polja vpliva na prag zaznavanja svetlobe.

Bistvene razlike med našim in njihovim poskusom so:

- pri nas je bilo vključenih 17 oseb, pri njih 55
- naš poskus je trajal 5 minut, njihov 31,5 minut
- mi smo poskus začeli z magnetnim poljem enakim

Zemljinemu, oni s poljem nič

Iz naših rezultatov je razvidno, da je korak, s katerim povečujemo svetlost utripajočega svetila, večji, kot pa so razlike v spremembi zaznavanja svetlosti, merjene preko toka skozi utripajoče svetilo (Delta LED). Zato je potrebno veliko število posameznih meritev, da bi se pokazala statistično značilna razlika. Zaradi premajhnega števila oseb in krajšega časa merjenja tega pogoja nismo izpolnili. Kvantizacija merjene količine je pregroba. Zato predlagamo spremenjen eksperiment, kjer bi oseba sama nastavljala svetlost utripajočega svetila na mejno vrednost, kjer utripanje ravno še zazna. S tem bi zmanjšali korak kvantizacije in verjetno tudi povečali število meritev v istem času. Pri naših eksperimentih se je prag svetlosti, ki je bil pri nekaterih v začetku nestabilen, hitro stabiliziral - prej kot v eni minuti, pri Thoss-u in soavtorjih [10] se je to zgodilo šele po nekaj minutah. Je vzrok tega pojava dejstvo, da so bile pri njih osebe že takoj v začetku eksperimenta v spremenjenih magnetnih pogojih (polje ni nič), pri nas pa v magnetnem polju z jakostjo enako zemeljskemu? Eden od udeležencev v poskusih je opazil, da je svetlost zaslona

modulirana z njegovim srčnim utripom. Ker sta frekvenci utripanja svetila in utripanja srca blizu skupaj, bi bilo v nadaljnjih eksperimentih potrebno upoštevati tudi vpliv kardiovaskularnega sistema na zaznavanje svetlobe. Pričakujemo, da bi vključitev ljudi v magnetne eksperimente, v primeru, da je pri ljudeh v dolgi dobi evolucije še kaj ostalo od magnetnega čuta, ki ga ima in uporablja veliko živali, lahko hitreje pripeljala do odgovorov na vprašanja o lokaciji in naravi magnetoreceptorjev ter o orientaciji v prostoru.

Zahvala

Zahvaljujem se profesorju Dušanu Feferju in Samu Begušu s Fakultete za elektrotehniko Univerze v Ljubljani, ki sta omogočila, da je eksperiment lahko potekal v laboratoriju fakultete, za njuno pomoč in nasvete ter vsem študentom prostovoljcem in eksperimentatorjem.

Reference

- [1] Kirschvink JL. Magnetoreception: Homing in on vertebrates. *Nature*. 1997 Nov 27;390:339-340.
- [2] Johnsen S, Lohmann KJ. The physics and neurobiology of magnetoreception. *Nat Rev Neurosci*. 2005 Sep;6(9):703-12.
- [3] Diebel CE, Proksch R, Green CR, Neilson P, Walker MM. Magnetite defines a vertebrate magnetoreceptor. *Nature*. 2000 Jul 20;406(6793):299-302.
- [4] Wiltschko R, Stapput K, Bischof HJ, Wiltschko W. Light-dependent magnetoreception in birds: increasing intensity of monochromatic light changes the nature of the response. *Front Zool*. 2007 Feb 15;4:5.
- [5] Wegner RE, Begall S, Burda H. Magnetic compass in the cornea: local anaesthesia impairs orientation in a mammal. *J Exp Biol*. 2006 Dec;209(Pt 23):4747-50.
- [6] Ritz T, Adem S, Schulten K. A model for photoreceptor-based magnetoreception in birds. *Biophys J*. 2000 Feb;78(2):707-18.
- [7] Wiltschko W, Wiltschko R. Magnetic orientation and magnetoreception in birds and other animals. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol*. 2005 Aug;191(8):675-93.
- [8] Begall S, Cervený J, Neef J, Vojtech O, Burda H. Magnetic alignment in grazing and resting cattle and deer. *Proc Natl Acad Sci U S A*. 2008 Sep 9;105(36):13451-5. Epub 2008 Aug 25. Erratum in: *Proc Natl Acad Sci U S A*. 2008 Nov 4;105(44):17206.
- [9] Burda H, Begall S, Cervený J, Neef J, Nemeč P. Extremely low-frequency electromagnetic fields disrupt magnetic alignment of ruminants. *Proc Natl Acad Sci U S A*. 2009 Apr 7;106(14):5708-13.
- [10] Thoss F, Bartsch B. The geomagnetic field influences the sensitivity of our eyes. *Vision Res*. 2007 Apr;47(8):1036-41.
- [11] Thoss F, Bartsch B, Telschaff D, Thoss M. The light sensitivity of the human visual system depends on the direction of view. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol*. 2002 Apr;188(3):235-7.

AN INTER-DISCIPLINARY SURVEY OF CURRENT STUDIES ON THE NATURE OF CONSCIOUSNESS

Thomas C. Daffern

International Institute of Peace Studies and Global Philosophy - IIPSGP
Chithurst Manor, The Old School House, Chithurst, Near Petersfield, West Sussex, U.K
Tel 01730 716496 Mobile: 07951 600959
e-mail: iipsgp@educationaid.net website www.educationaid.net

ABSTRACT:

The current state of research into consciousness in the natural and social sciences, and other forms of cognition, is presented. The author explores how art, literature and religion can also shine scientific light on the nature of consciousness and must not be overlooked in scientific colloquia. Some ideas are also given as to how one could use consciousness to arrive at better solutions for contemporary societal and human problems. The phenomenon of consciousness is discussed in the light of transpersonal psychology and philosophy. This is an abridged text shortened from the full paper

Consciousness and the natural sciences:

Consciousness Studies is a new interdisciplinary term which draws heavily on the work of natural scientists. Their peer reviewed *Journal of Consciousness Studies* is published by the University of Arizona. The Centre for Consciousness Studies is also based at the same University and is a multidisciplinary centre focusing on all aspects of the subject. Another parallel journal, *Consciousness and Cognition*, is the journal of the Association for the Scientific Study of Consciousness. This body, which is arguably the leading such scientific network in the world, founded in 1994 by a group of scientists working on different aspects of brain research, neurosciences, cognitive studies and philosophy of science, organises regular international conferences and publishes an online journal PSYCHE, as well as the above mentioned *Consciousness and Cognition*. The work is international and includes scientists from many countries worldwide, and from different cultural and intellectual standpoints. The Society has also published specifically three monographs on neurobiology and consciousness studies. [1]. Finally, the Society also awards an annual William James Prize for work by the most promising young scholar active in the field of consciousness studies [2]. With all this work going on, hasn't the location of consciousness been achieved? Hasn't science figured out exactly what is going on? One hypothesis is that of **Stuart R. Hameroff** [3], who locates the subjective experience of

consciousness as a phenomenon located in the microtubules inside the brain's neuronal system which is where the electric impulses occur that constitutes the literal rush of brain activity that Hameroff associates with "thinking" and "brain waves". Many other consciousness researchers have been working in recent years on different aspects of the neurophysiology of consciousness, from different perspectives and angles. **Itzhak Bentov** in his important early study *Stalking the Wild Pendulum: On the Mechanics of Consciousness* (1977) proposed that consciousness is a kind of eternal nowness which provides the pleroma in which all energy and matter rotate, swirl, transform and intergenerate – that consciousness, like Sunyata in Buddhism, is the kind of ultimate field of nothingness in which manifestation is itself manifesting. He also advocated the parallel idea that a kind of steady state, or eternally recycling sustainable universe, is the most appropriate cosmology to adopt (also known as the Continuous Big Bang Model). In 2000 *A Brief Tour of Higher Consciousness: A Cosmic Book on the Mechanics of Creation*, Inner Traditions - Bear and Company, was also published some 20 years after his tragic death in an airline crash. One of his key ideas, which is again very consistent with Buddhist and other Eastern metaphysical teachings, is that the sum total of all knowledge that has ever been generated (or ever will be) is somehow available to our minds now, if only we know how and where to tune in. This is similar to the idea in Yogachara Buddhism, that there is an alayavijnana, a storehouse-consciousness, which individual minds can tap into, but which is collectively far vaster and more ancient than our own individual knowings can ever be. **Joseph E. Bogen** (1923-2006) was another neurophysiologist who specialised in consciousness research, particularly interested in the phenomenon known as split brain research. As a Professor of Neurosurgery, Bogen pioneered the research on whose findings the later consensus that right and left hemispheres of the brain control different aspects of consciousness was based (intuitive = right; rational = left). Tragically he was working on a grande oeuvre to sum up his lifetime of research into the physiological nature of consciousness at the time of his death. He argued that consciousness is such a subjective phenomenon, that we will never be able to point to a purely physical or material correlate and say "that is consciousness" since by definition it is a subjective phenomenon. What we can say about the

subjective experience of being conscious, however, is that it seems to involve a relationship between a centre (a distributive node) and periphery (sensory data, “out-there experiences” to be conscious of). He argued that the physiological correlates in brain structures able to reproduce this centre-periphery relationship are where we ought to be looking for the seat of this subjective experience in itself. Another natural scientist who explored the nature of consciousness in some depth was the physicist **David Bohm** (1917 – 1992) who at one time worked with Einstein on relativity and quantum theory. Bohm ended up working at the University of Bristol and then Birkbeck College. A close friend and colleague of Krishnamurti, in his later years he recorded many dialogues and conversations with Krishnamurti in which they discussed the nature of consciousness, the one from the perspective of a philosopher and the other as a physicist. Bohm’s technical argument was that there is a kind of “implicate order” behind manifest reality, which manifests at the physical level as a degree of ordered arrangements between quanta of energy, particles, waves etc. which conform to certain laws and norms of behaviour, and which enables the physical universe to take the form and structure that it does. In the world of mind, he speculated, there must be some kind of similar invisible quanta of thought, whose interactions and dynamic interfusings generate a similar universe of mental discourse. Could the internal architectonic of the two universes, the mental and the physical, ever be harmonised ? Could the same formulae or governing general laws be discovered that determine both inner and outer realities ? Bohm felt this was possible and furthermore that it was in the same general area that Einstein had been searching in his quest for a Unified Field Theory. Another natural scientist to work on consciousness was **Rodney Cotterill** (1933-2007) a British educated scientist who taught in Denmark for many years, and ended up researching the innermost structures of the brain and how they can account for the subjective experience of consciousness itself. He felt sure that a purely physical explanation could eventually be arrived at once the material sciences become sophisticated enough to analyse what is going on exactly, and where. He felt that computer activity is also itself a key and a guide to how the brain itself operates and generates consciousness. A further natural scientist involved in this work is Prof **Antonio Damasio** (1944-) who worked for many years as head of neurology at the Iowa state hospital system, and more recently has moved to the University of Southern California to direct the *Brain and Creativity Institute* there. For many years Damasio has explored the scientific basis of emotion, of states of subjectivity and feeling that humans attach special ontological importance to, arguing that they have physical correlates in internal brain function, and that part of what makes us uniquely human is the ability to process and galvanise these complex feeling-nets in successful social interactions, which in turn also influence our decision-making abilities. What happens when these internal skills become impaired? What happens when the physical

structure of the brain becomes damaged ? Do our personalities, our emotional subjectivities change ? Do our ethical principles, according to which we live our lives, depend on the successful functioning of our physiological cerebral architecture ? Where exactly in our neural nets are our memories stored away ? These are some of the important questions Damasio has been asking over many years, as explored in his various publications, including: *Descartes’ Error: Emotion, Reason and the Human Brain* (1994), *The Somatic marker hypothesis and the possible functions of the prefrontal cortex*, (1996), *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*, (1999), *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*, (2003). Another natural scientist research aspects of consciousness has been **Dr. Lüder Deecke** who specialised in medical neurology and undertook doctoral research in Germany on the crucial moments before making a conscious decision to undertake a certain motor action (e.g. raising a finger) which can be measured exactly by scientific equipment attached to the skin and also to the brain itself. They discovered that there is a strange phenomenon which they termed *Bereitschaftspotential* (translated into English as “readiness potential”) and which indicates that the physical motor units of the brain and nervous system are gearing up for the mechanistic correlates of an action, prior to the subject actually performing that action. In other words, consciousness somehow seems to be in a kind of prior time lag with the physical mechanisms that it apparently controls. Deecke and his supervisor found proof, so to speak, for the existence of the mental precedent of motor actions, as being recorded in the minute fluctuations in the subject’s brain waves. Deecke ended up as Head of the Department of Clinical Neurology at the University of Vienna and has remained actively interested in the way that consciousness manifests its presence in terms of physical brain structure. **Sir John Eccles** (1903-1997) was another important neurologist who won a (shared) Nobel Prize for Medicine in 1963 for his researches on the nature of the synapse governing relations between neurons. Australian in origin, Eccles worked in the UK, Australia and the USA, and worked out the details of how nerve impulses are carried across synapses, realising eventually that they are conducted by neurotransmitters such as acetylcholine. Interestingly, Eccles remained a Christian theist and a believer in a higher spiritual dynamic underlying the human world of experience. In a series of remarkable books he explored with great intellectual integrity how the further researches of scientific knowledge in his field eventually petered out into trails in the upper reaches of the unknown from where perhaps only the mystic or the theologian might dare to take the ascent any further. His publications included: *The Physiology of Nerve Cells* (1957) *The Physiology of Synapses*. (1963), *The brain and the unity of conscious experience* (1965) *Facing reality: Philosophical Adventures by a Brain Scientist* (1970), *The Understanding of the Brain* (1973), *The Self and Its Brain* (1977) this later being co-written with Sir Karl Popper; *The human mystery* (1979), *The Human Psyche* (1980), *Evolution Of The Brain* :

Creation Of The Self (1989) and lastly *How the Self Controls the Brain* (1994). Another contemporary researcher into the nature of the brain and consciousness is **Susan Greenfield**, who is the Director of the Royal Institution of Great Britain and Professor of Synaptic Pharmacology at the University of Oxford where she has set up and directs a new *Institute for the Future of the Mind*. The focus of this interdisciplinary research team is interesting – not content to answer the ancient questions about the nature of consciousness, they are also exploring the general parenters of the effects of rapid social and technological change on the way that the brain and mind work. How will consciousness cope with the increasing challenges and shocks that might affect us in our increasingly inter-dependent world ? (see www.futuremind.ox.ac.uk for further details). It is to this increasingly inter-disciplinary approach to the study of consciousness that we must now turn.

Consciousness and the social sciences – the notion of science is itself a key to the study of consciousness, since both words come from the same root, Latin scio, I know. Science means literally knowledge, and consciousness, means literally “that with which we know” or “knowing with” (the exact meaning is ambiguous and encompasses both meanings in English). Consciousness, this mysterious x, is thus at the root of all our knowing – without it, there can be neither science, nor knowledge. It is the observer, the experimenter, the framer of scientific hypotheses, the calculator of variables. The way that formal scientific method was articulated by Sir Francis Bacon, Descartes, Newton, Galileo and other pioneering minds in the early periods of the scientific revolution, was designed to exclude subjectivity from the calculations. They took as their model of truth the harsh rigour of mathematical calculations. Later, the idea of scientific rigour was applied to the study of society, as part of the work of the 18th century enlightenment, where Adam Smith analysed the social interactions that generated wealth in societies and progressed what later became the social science of economics; Hobbes had earlier studied how political societies interact and tried to mould the method of his mentor Bacon to this unfamiliar terrain. Vico asked how societies change and progress over time, and what forces and strata of society carry with them wisdom or understanding from epoch to epoch, as their context changes and evolves. Destutt de Tracy asked how ideas interact with social groupings, and how ideology arises, as the articulated consciousness of particular factions, interests and groupings, and how peace can arise from their interactions. Disgusted by the excesses of the French revolution with its reign of terror, Destutt retired to the countryside to develop his intellectual approach to social science. Comte thought that the evolution of mankind is a tableau of gradual progress, from darkness to enlightenment, in which truth, positive knowledge, scientific knowledge, gradually replaces barbarism, superstition, irrationality. Comte’s saints had test

tubes and measuring compasses, phials and weighing machines, and their advent presaged the arrival of the common good. St Simon took these ideas and argued that consciousness is a class phenomenon, that class differentiation has caused the oppressive disfigurement of consciousness from class to class, and that only socialism can produce a new kind of society, in which working for the general good becomes the norm, rather than selfish self interest. Robert Owen also explored these ideas, and furthermore became very interested in the study of transpersonal conscioiusness towards the end of his life, and a leading figure in the spiritualist movement. Marx and Engels took over all these ideas, but paradoxically, being at such pains to make them “scientific” they banished the study of consciousness as such from their ideal academy, and seem instead to have regarded consciousness as a given, something so self evident it didn’t need further analysis. This approach might have been inspired partly by their readings of Hegel, but in fact Hegel’s own teachings on the matter had been far more complex, and he had devoted a very considerable amount of study and thought to how consciousness manifests in institutors, structures, societies, religions, political ideas, judicial concepts, philosophical ideas etc. In fact, it could be said that the great secret of Hegel and Hegelianism, and its power, was precisely that it put the study of consciousness firmly on the map of philosophy for the first time in a scientific manner. Yet immediately afterwards, this effort had been subverted, paradoxically, by Marx, who sought to turn Hegel upside down. ! This crucial debate, which was fought out first in the classrooms and later on the battlefields, became known as the struggle between materialism and idealism. Yet the actual scientific study of consciousness and its social manifestations got lost amidst the din of competing factions. Is it perhaps now long overdue that we recommence its study, remembering the ancient philosophical teaching – *know thyself* ! Without question, then, we can say that consciousness is partly a social phenomenon – that the means of communicating consciousness, whether by word or by gesture or body language, is a socially acquired characteristic. Consciousness is in some sense also the way whereby we “know together.” Is consciousness present wherever communication is occurring ? This would seem close to self evident. Consciousness occurs where apparently separate things interconnect and commune and communicate. These can be nerves, communicating through the rich pharmacopoeia of their synapses, or people, meeting in a park at dusk and beginning a love affair... We all know what it feels like to be conscious, and how it fluctuates and changes, yet it is ironic how little studied consciousness actually is, even now. We have agreed it’s the root and foundation of all other science, all our other knowing – yet we have ignored and neglected its study hitherto. Perhaps we can begin to remedy this woeful situation. What kinds of fields of the social sciences would be involved in its rigorous study, exactly ? Firstly, the **economic sciences**: – exactly how wealth, poverty, greed, work, labour, make people feel; how does the quality of employment, how do worker-management

relations, impact on people's subjective experience of work ? How do new ideas come to be adopted as successful economic enterprises ? How do companies with vision, with creative knowledge and products, make it to the fore, while others, with outmoded ideas and old fashioned understandings of knowledge, come to lag behind ? How do companies such as Google, for example, which operate on a lateral, non-hierarchical framework, manage to succeed so well ? How does consciousness affect management science and leadership ? Surely, the sense of vision, of purpose, that successful institutions embody, has something to do with breadth and depth of consciousness ? What happens where workers or staff feel under-utilised – when their “minds wander” and when they feel that their consciousness is not excited, not actualised, not drawn on ? Is this not the real meaning of what Marxist's call alienation – the feeling of being other than, divorced from, and separated from, one's working reality. How then can we as a world community create working conditions for people in which alienation is a thing of the past, in which workers feel a sense of belonging and engagement and enthusiasm for their work ? What role does consciousness play in this ? How does ownership per se affect consciousness ? If I own my workplace, in some sort of share scheme, or via the state owning it, or not, how does this variable affect the quality of my working experiences ? What exactly is ownership ? Is it false consciousness per se, or is it a valid mode of experience, useful to demarcate civilised life from sheer undifferentiated barbarism – is the notion of private property a victory or a defeat for consciousness ? Is money itself something that affects consciousness ? Why do we feel elated when rich and fearful when poor ? Is the buoyant ego associated with wealth and the sad and depressed ego associated with poverty a natural extension of the economy of Libido / Thanatos that Freud proposed, and that Darwin also elaborated on scientifically within the framework of evolutionary theory (Darwin parted with A.R. Wallace however exactly on the problematic of the phenomenology of consciousness and its role in evolution). Is the competition for resources, also a competition for consciousness, as in the Nietzschean acceptance of the will to power ? Is that really, at bottom a will to consciousness ? Or is there another way than sheer brutal competition ? What would an economy based on love look like, neither a coercive communism, nor a coercive capitalism ? Has it ever been tried ? An economic system based on the free cooperation of friendly and loving individuals ? Finally, we must also ask – why do we seem to be so far from this still ? The recent economic crashes of 2008 and 2009, the current world recession – how far are they impacting on and a result of certain particular modes of consciousness ? The fear of liquidation, redundancy, homelessness, unemployment, these are pressing and vital concerns for millions of people worldwide – and they are an affront to consciousness itself. Somehow our economics and politicians have to factor back consciousness itself, feeling, sentience, into economic theory. Above all economists should realise that the

generation of knowledge is a crucial factor in wealth creation, and new knowledge requires new forms of dissemination and sharing. The exponential growth of information technology, the whole rise of the computer industry and the internet etc. are a witness to the importance of the knowledge economy. The challenge that faces us now however is how to transform the knowledge economy into a wisdom economy – and in this task, the question of actually identifying what consciousness is and how it can be generated, maintained, supported, validated, esteemed and enhanced – is crucial. What kinds of economic activities, structures and enterprises will generate new ideas, new consciousness ? How will they satisfy people's innermost needs best ? These are questions for the economics of the future, one which will lead to peace, rather than the current economics of violence, warfare, greed, hatred, guilt, debt, and competition. On 16 September 2009 the Archbishop of Canterbury took part in a TV discussion with economists and other thinkers concerning the spiritual dimension of the current global financial recession, arguing that it had been produced by a sense of unreality generated by the cult of expertise on the part of economists who had donned the mantle of expertise which had lulled ordinary people to assume that the financial industry knew what it was actually doing. He also stated some of the spiritual costs involved, and called for a degree of humility and repentance on the part of the world's financial elites, whose models of reality had so long denied ethics and spirituality from their worldview. Other social sciences that have much to say about consciousness include **sociology**, where Pitirim Sorokin long ago pioneered the role that creative altruism plays in vanguard social developments, and who ran a centre at Harvard University dedicated to its study. Since the work of Durkheim and Weber the role that consciousness plays in forming social groups and structures has been a long and constant matter of inquiry for the sociological imagination. **Anthropology** has also for many decades pioneered the analysis of the roles of consciousness in social systems found around the world, including the way that myths, legends, symbols and stories play in bonding and cementing group and tribal life. The detailed study of initiation rituals underway worldwide is another example of how the transference of consciousness is taken seriously in primal human cultures. Joseph Campbell and other mythographers have tried to map these stories into overall patterns of meaning, following on from the pioneering work of Sir James Frazer in the early 20th century. Some mythographers have announced that all the legends and myths of the world amount to just 7 basic narratives, and on further reflection, it will be apparent that these 7 stories all represent different experiences that occur to consciousness in the process of birth, maturation, love and death. Anthropology thus can hold up the mirror of our own minds enacted in societies worldwide. Mircea Eliade in his pioneering of Shamanism has also discovered something of the same grand narrative as did Campbell and Jung. Physical anthropology has likewise made enormous strides in studying the evolution of the human brain in Palaeolithic times, and in comparing the various primitive proto-human forms for

evidence of thinking capacity, speech, thumb/finger juxtapositioning, upright bipedalism, female sexual activity not restricted to oestrus, capacity for tool making, all of which are factors involved in the evolution of consciousness. Recent (2009) discoveries in Georgia (Caucasus Mountains) have shone further light on the antiquity of an advanced hominid presence in Europe, along with tool making and other signs of early consciousness development. One other interesting fact of early so called primitive man, is that we believed that consciousness was a phenomenon-at-large, and that animals and even plants and rocks and river and mountains, to some extent shared in an animistic world in which everything, somehow, participated in a sea of conscious experience. The function of mankind was not to monopolise consciousness, but rather to act as a steward, servant or harmoniser of the different aspects of consciousness that exist at large in society, from stellar beings to aquatic, or aerial or land beings. All of nature is somehow alive and our function is to celebrate, and indeed to worship this Great Mystery that lies at the heart of existence. Modern quantum physics and advanced metaphysical thought have in many ways come back to realising that our so called primitive ancestors were not necessarily so primitive after all. David Peat's study of *Blackfoot Physics* (1996) is a good example of this realisation. The current author has also analysed in detail the American Indian ethical and philosophical worldview, in the *Encyclopaedia of Applied Ethics*. (Academic Press, 1998, new edition forthcoming 2010)

Political aspects of consciousness studies: here is another huge field, for the history of political thought is just that – a history of the way that different consciousness have grasped reality and framed optimal or realistic ways for political groupings to coexist. Aristotle in fact started the scientific study of politics by arguing that it was the quest for happiness, and that man's being is inevitably a collective phenomena, whether in the family, the tribe, the clan, the city state or the wider political community. Politics then is partly how one can harmonise the demands of competing and contesting consciousness, each with their own narratives of truth, against those of other competing narratives. Nations are also collective conscious entities, to some extent, with a life larger than the sum total of the individual who comprise them. Values, beliefs, codes, ideologies are transmitted subtly from generation to generation –and yet, as they shift, new nations emerge from the old – nothing is constant, nothing stays the same, and yet everything does... But who has actually studied consciousness and its role in political theory and practice ? Which political philosophers have actually talked about it ? Which politicians ? Has our political discourse become so impoverished that we are not able to talk about such things without appearing unduly naïve or opportunistic ? In the Declaration of Independence of 1776 of the fledgling United States of America that Enlightenment project deriving from the 18th century, are found the words, penned by Thomas Jefferson: “*We hold*

these truths to be self-evident, that all men are created equal, that they are endowed by their Creator with certain unalienable rights, , that among these are life, liberty and the Pursuit of Happiness. To secure these rights governments are instituted among men, deriving their just power from the consent of the governed.” There is a huge question here about equality, in relation to consciousness; as a teacher of many years, I can absolutely state that children are NOT born equal, and do not present themselves with equal aptitude for all subjects. They may have this potentially, as human beings (ontological equality) but most manifestly do not possess it in actuality (social equality). To argue otherwise is I am afraid to simply deny the actual evidence of life. But we are all ontologically equal, and in some sense one can say we are “created equal” i.e. through the same biological and perhaps metaphysical processes. *Difference* however is equally important, as an aspect of human life, as sameness; indeed, politics is still struggling how to accommodate successfully the question of difference. The multiple differences that affect us as human beings are the very meat and drink of politics – culture, ethnicity, religion, language, class, intelligence, gender, sexuality, spirituality, wealth, knowledge, education – perhaps the first thing to strike a visitor from another world about human beings would be how different we all are from one another. Political systems struggle and strain to cope with all this diversity – they forge nations, constitutions, political parties, governments, political ideas – yet the fundamental philosophical dilemmas remain. We re similar, yet different, equal yet unequal. And over this truism countries fall apart, engage in civil wars, endure endless agonies of disagreements and infighting... Yet could consciousness itself be part of the solution ? Could the scientific study of the politics of consciousness itself be some part of a solution towards this dilemma into which mankind has plummeted ? one thinks of the marvellous speech of Mandela in which he speaks of healing relations with the other; one thinks of the Truth and Reconciliation Commission for South Africa. One thinks of Obama's Presidential acceptance speech in which he spoke of healing the war within the mind and soul of America as a nation. One thinks of Gandhi being prepared to fast to the death in order to prevent attacks by Hindu fanatics on his Muslim brothers and sisters All these are pointers towards the recognition that consciousness is indeed part of the key, part of the solution, to the political problems affecting mankind. But the study of the role of consciousness and its potential contribution to a more positive politics is indeed in its infancy. [4]

Consciousness and history: towards a transpersonal historiography [5-7] – And what of history ? Is history itself, the whole of time, but a play of classes, and battles, and kings and wars, or power struggles, of catastrophes, or plagues and conquests and explorations ? Or is there something else, some hidden drum whose beat has long been obscured ? What would history be like if consciousness were put back into it – both in its study and in its re-creation ? Since Freud came and went, psychohistorians have speculated that the inner life of

children might inform their later maturity as adults, and the dawning of the age of psychohistory has been a long one, full of trials and tribulations, in which the cruelty of parents to children has been duly and comprehensively noted (by De Mause etc.). In the current author's own doctoral work however he advanced the notion that it is time to develop a transpersonal psychohistory, which takes account of the further reaches of human nature, as defined by Maslow, Jung and the host of transpersonal psychologists, many of whose names are included in note 18 below. In the authors view, to study psychohistory without exploring the transpersonal dimensions of thought and psychic life, would be like studying consciousness without being conscious, which would in fact surely be a logical impossibility. This is a new field for historians however, and being the most conservative of intellectuals, historians are unlikely to agree to its proposal all at once. One doesn't have to be a prophet however, to realise that if mankind is to generate any kind of acceptable future for ourselves over the next few thousand years, we are going to have to get some rigorous and scientific thinking about transpersonal history, and at least agree to differ in our major narratives about the meaning of time, religion and culture, so the major civilisations on offer can learn to share this planet with one another in harmony. Alternatively, the reader might care to cut out this essay and to bury it deep down in a glass bottle somewhere for future generations to discover, and to lament our failure of consciousness, even at this eleventh hour, pregnant with such potentiality and possibility.....

References

- [1] Steven Laureys, ed (2005). *Progress in Brain Research, The boundaries of consciousness: neurobiology and neuropathology*. Elsevier.; Axel Cleeremans (Ed.), ed (2003). *The Unity of Consciousness: Binding, Integration, and Dissociation*. Oxford University Press.; and Thomas Metzinger, ed (2000). *The Neural Correlates of Consciousness: Empirical and Conceptual Questions*. MIT Press.
- [2] See <http://www.theassc.org/> for more details
- [3] *Toward a Science of Consciousness III: The Third Tucson Discussions and Debates* (1999). Stuart R. Hameroff, Alfred W. Kaszniak and David J. Chalmers (Editors). The MIT Press.
- [4] The current author has attempted to take this approach forward in a number of theoretical and practical approaches, including the work of the Truth and Reconciliation Commission for the Middle East, and the proposed European Union Mediation Service (EUMS). See www.lulu.com/iipsgp for publications about these ongoing projects. See also the important study *Spiritual Politics: Changing the world from Inside Out* by Connie McLaughlin and Gordon Davidson (Ballantine Books, New York, 1994)
- [5] Daffern, Thomas Clough *Towards A Transpersonal History Of The Search For Peace 1945-2001* (IIPSGP Publications, USA, 2009);
- [6] Daffern, Thomas Clough *Transpersonal History And Peace Historiography – Exploring The Implications For The Philosophy Of History: Towards A Definition Of The Field* (IIPSGP Publications, USA, 2009) *Revisioning Transpersonal Theory: a participatory Vision of Human Spirituality*.
- [7] See for a good overview of the field, *Transpersonal Knowing: Exploring The Horizons of Consciousness*, by Tobin Hart (SUNY Press, 2000) and *Revisioning transpersonal theory: a participatory vision of human spirituality* by J.N. Ferrer
- [8] Tim Bayne, Axel Cleeremans and Patrick Wilken *The Oxford Companion to Consciousness* ed. (OUP, 2009) has recently appeared and is an excellent overview of many the fields of research indicated in this paper.

POVEZANOST IMPULZIVNOSTI Z AFEKTIVNIMI DIMENZIJAMI TEMPERAMENTA MED SPOLOMA

Barbara Dolenc, Lilijana Šprah

Druženomedicinski inštitut, Znanstvenoraziskovalni center SAZU

Novi trg 2, Ljubljana

tel: 01 470 64 39; faks: 01 426 14 93

e-mail: dolenc.barbara@gmail.com

POVZETEK

Prispevek predstavlja raziskavo, kjer smo proučevali razlike v izraženosti afektivnih dimenzij temperamenta in različnih aspektov impulzivnosti med spoloma. Hkrati nas je zanimalo, ali se dimenzije medsebojno povezujejo in, kakšno vlogo igra pri tem spol. V raziskavi je sodelovalo 1096 študentov in študentk, ki so izpolnili lestvico impulzivnosti BIS-11 in vprašalnik afektivnih temperamentov TEMPS-A. Rezultati so pokazali, da imajo ženske bolj izražene depresivne, ciklotimne in anksiozne lastnosti, moški pa razdražljive in hipertimne značilnosti. Moški so imeli tudi bolj poudarjeno impulzivnost kot splošno lastnost ter impulzivnost pri načrtovanju. Le ciklotimna dimenzija se je povezovala z vsemi aspekti impulzivnosti pri obeh spolih medtem, ko so imele ostale afektive dimenzije temperamenta različne trende povezanosti, tako z impulzivnostjo kot s spolom.

1.0 UVOD

Impulzivno vedenje definiramo kot hitro, spontano in nenačrtovano akcijo, ki odraža nesposobnost odlaganja gratifikacije (Arce in Santisteban, 2006). Dosedanje raziskave na področju impulzivnosti in razlik med spoloma v izraženosti le-te kažejo, da se moški in ženske ne razlikujejo pomembno (Patton, Stanford in Barratt, 1995; Stanford in dr., 2009).

Medtem ko se različni aspekti impulzivnosti že vrsto let nahajajo pod budnim očesom raziskovalcev, pa se šele v zadnjih letih ponovno pojavlja koncept afektivnega temperamenta, ki sicer izhaja že iz grško-romanske in osrednjeevropske psihiatrije. Afektivni temperament se nanaša na »specifične konstitucijsko-zasnovane afektivne dispozicije« (Perugi, 2005, str. 260). Gre za dimenzionalni konstrukt, ki ga lahko v njegovih ekstremih razumemo kot odklonskega v statističnem in morda tudi v kliničnem smislu. Afektivni temperament predstavlja adaptativno mešanico lastnosti, ki v določenih ekstremih, lahko vodi do bolezni ali modificiranega izražanja nekaterih afektivnih stanj (Akiskal in Akiskal, 2005). Raziskave kažejo, da med spoloma obstajajo določene razlike v izraženosti posameznih dimenzij afektivnih temperamentov. Moški naj

bi imeli pomembno bolj izraženo dimenzijo hipertimnega (Pompili in dr., 2008; Vahip in dr., 2005; Vázquez in dr., 2007) in razdražljivega temperamenta (Vázquez in dr., 2007), medtem ko naj bi pri ženskah prevladovala lastnosti depresivne, ciklotimne (Erfurth in dr., 2005; Figueira in dr., 2008; Rózsa in dr., 2008), pa tudi anksiozne dimenzije (Vahip in dr., 2005; Vázquez in dr., 2007).

Čeprav tako impulzivnost, kot tudi afektivne dimenzije temperamenta predstavljajo osebnostne lastnosti, na podlagi katerih bi lahko sklepali na morebitne odklonske simptome, predvsem v smeri razpoloženjskih motenj (Akiskal in Akiskal, 2005; Swann, Steinberg, Lijffijt in Moeller, 2008), o sami neposredni povezanosti impulzivnosti z nekaterimi simptomi afektivnosti ni veliko znanega.

Signoretta, Maremmani, Liguori, Perugi in Akiskal (2005) so na primer v svoji raziskavi ugotovili povezanost ciklotimnih dispozicij z impulzivnostjo. Avtorji navajajo, da je ciklotimna dimenzija temperamenta od vseh najbolj patološka in se povezuje z emocionalnimi in vedenjskimi težavami ter z osebnostno in socialno nestabilnostjo. Stanford, Greve in Dickens Jr. (1995) poročajo tudi o visoki povezanosti impulzivnosti z nekaterimi lastnostmi razdražljivega temperamenta. Impulzivni posamezniki so namreč pogosto poročali o nakopičeni napetosti, jezi, o občutku, da bodo »izbruhnili«, prav tako pa so večkrat kazali tudi določene eksplozivne, nekontrolirane akcije, kar so tipične lastnosti razdražljive dimenzije temperamenta (Akiskal in Akiskal, 2005). Avtorji zaključujejo, da impulzivnost in razdražljivost tvorita dimenzijo vedenjske kontrole, ki je vpletena tudi v inhibicijo agresivnih izbruhov (Stanford, Greve in Dickens Jr., 1995). Nekateri raziskave odkrivajo tudi povezanost depresivne dimenzije temperamenta z impulzivnostjo pri načrtovanju, saj naj bi se slednja povezovala z brezupom in anhedonijo (Swann in dr., 2008) ter pomanjkanjem občutka za prihodnost (Patton, Stanford in Barratt, 1995).

2.0 NAMEN RAZISKAVE

Namen raziskave je bil proučiti, ali se med spoloma pojavljajo razlike v izraženosti afektivnih dimenzij temperamenta ter v različnih aspektih impulzivnosti in, ali se omenjene dimenzije medsebojno povezujejo podobno pri moških in ženskah.

3.0 METODA

3.1 Udeleženci

Pri raziskavi je sodelovalo 1096 študentov iz različnih slovenskih fakultet, od tega je bilo 416 moških in 680 žensk. Povprečna starost vseh udeležencev je bila 20,55 let.

3.2 Pripomočki

Barrattova Lestvica impulzivnosti BIS-11 (Patton, Stanford in Barratt, 1995): namenjena merjenju impulzivnosti, kot osebnostne poteze in se lahko uporablja tako v klinične namene kot pri zdravih osebah. Vsebuje 30 postavk, ki se nanašajo na tri aspekte impulzivnosti (sekundarni faktorji), in sicer *pozornost* (sposobnost osredotočiti se na trenutne naloge), *motorika* (spontano in nepremišljeno reagiranje) ter *načrtovanje* (trenutna orientacija oziroma pomanjkanje vpogleda v prihodnost).

Vprašalnik TEMPS-A (Temperament Evaluation of Memphis, Pisa, Paris and San Diego, Akiskal in Akiskal, 2005): vprašalnik afektivnih temperamentov, ki je bil razvit z namenom ocene afektivnih dimenzij temperamenta tako v klinični praksi, kot tudi pri zdravih posameznikih. Vprašalnik vsebuje 110 postavk za ženske in 109 za moške. Postavke odražajo lastnosti dimenzij afektivnega temperamenta, ki naj bi bile stalne, prisotne skozi celotno življenje posameznika. Te lastnosti se izražajo v petih različnih dimenzijah, in sicer v depresivni, ciklotimni, hipertimni, razdražljivi in anksiozni dimenziji temperamenta.

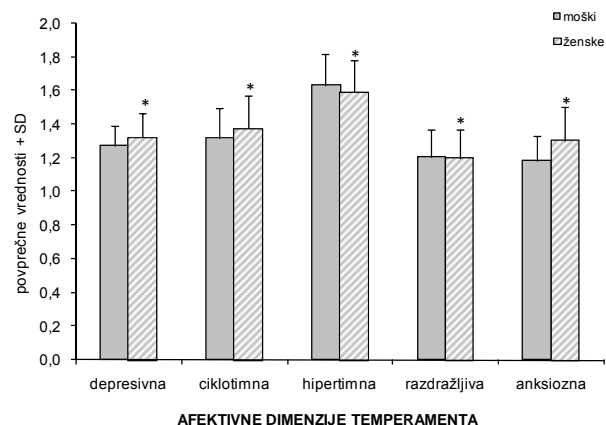
3.3 Postopek

V raziskavo smo vključili študente iz različnih slovenskih fakultet Univerze v Ljubljani in Mariboru. Udeleženci so skupinsko, prostovoljno ter anonimno izpolnili Barrattovo lestvico impulzivnosti BIS-11 in vprašalnik afektivnih dimenzij temperamenta TEMPS-A (Temperament Evaluation of Memphis, Pisa, Paris and San Diego, Akiskal in Akiskal, 2005). Rezultate smo analizirali s statističnim paketom SPSS 14.0, kot mero statistične pomembnosti smo vzeli $p < 0,05$ oz. $p < 0,01$.

4.0 REZULTATI

4.1 Afektivne dimenzije temperamenta

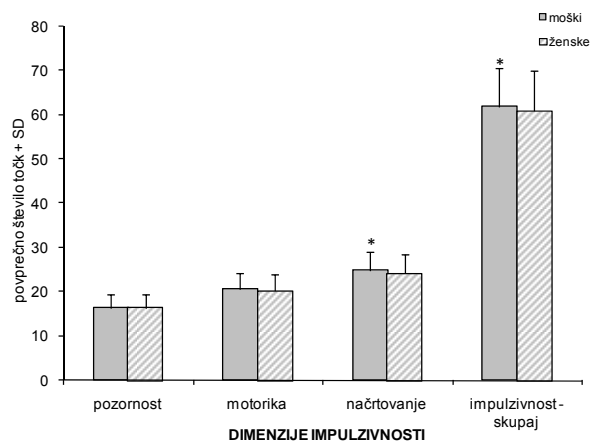
Ker je sama struktura vprašalnika TEMPS-A za oba spola različna, saj moški odgovarjajo na postavko manj, rezultate v nadaljevanju prikazujemo ločeno glede na spol. Rezultati so pokazali statistično pomembno bolj izraženo depresivno ($Z = -5,24$; $p = 0,000$), ciklotimno ($Z = -4,344$; $p = 0,000$) in anksiozno dimenzijo ($Z = -10,681$; $p = 0,000$) pri ženskah, medtem ko so pri moških pomembno bolj izražene lastnosti hipertimne ($Z = -3,377$; $p = 0,001$) in razdražljive dimenzije temperamenta ($Z = -2,004$; $p = 0,045$) (slika 1).



Slika 1. Razlika v izraženi posameznih afektivnih dimenzij temperamenta, ocenjenih s TEMPS-A, med moškimi in ženskami (* - statistično pomembna razlika med spoloma v dimenziji).

4.2 Dimenzije impulzivnosti

Primerjava med spoloma na posameznih dimenzijah impulzivnosti je pokazala, da moški kažejo pomembno višjo impulzivnost na področju načrtovanja ($Z = -1,994$; $p = 0,046$), statistično pomembne pa so bile tudi razlike v skupnem skoru impulzivnosti ($Z = -2,299$; $p = 0,021$) (slika 2).



Slika 2. Razlika v izraženi posameznih dimenzij impulzivnosti, ocenjenih z BIS-11, med moškimi in ženskami (* - statistično pomembna razlika med spoloma v dimenziji).

4.3 Povezanost afektivnih dimenzij temperamenta in impulzivnosti

Analiza povezanosti med posameznimi dimenzijami afektivnega temperamenta in impulzivnostjo je prikazana v tabeli 1. Iz nje je razvidno, da se impulzivnost v načrtovanju povezuje z vsemi dimenzijami, s hipertimno negativno, z ostalimi štirimi dimenzijami pa so korelacije pozitivne. S pozornostjo korelirata le ciklotimna in razdražljiva dimenzija, pri moških pa tudi anksiozna, medtem ko so

korelacije afektivnih dimenzij temperamenta z motorično impulzivnostjo pri moških skoraj nepomembne (z izjemo ciklotimne dimenzije), pri ženskah pa motorična impulzivnost pozitivno korelira s ciklotimno in razdražljivo ter negativno z depresivno in anksiozno dimenzijo temperamenta. Skupna impulzivnost pomembno pozitivno korelira s ciklotimno in razdražljivo dimenzijo, pri ženskah tudi z anksiozno, ter negativno s hipertimno dimenzijo.

Tabela 1

Spearmanovi koeficienti korelacije med posameznimi dimenzijami afektivnega temperamenta in dimenzijami impulzivnosti.

	Impulzivnost			
	Pozornost	Motorika	Načrtovanje	Skupaj
<i>Depresivna</i>				
moški	0,020	-0,049	0,124*	0,048
ženske	-0,070	-0,125**	0,166**	0,007
<i>Ciklotimna</i>				
moški	0,284**	0,152**	0,325**	0,318**
ženske	0,254**	0,179**	0,362**	0,341**
<i>Hipertimna</i>				
moški	-0,019	-0,009	-0,296**	-0,162**
ženske	-0,007	0,014	-0,256**	-0,130**
<i>Razdražljiva</i>				
moški	0,229**	0,062	0,319**	0,270**
ženske	0,214**	0,086*	0,367**	0,283**
<i>Anksiozna</i>				
moški	0,102*	-0,075	0,204**	0,109*
ženske	-0,028	-0,099**	0,132**	0,015

* statistična pomembnost na 5% nivoju tveganja

** statistična pomembnost na 1% nivoju tveganja

5. 0 RAZPRAVA

V prezentirani raziskavi smo proučevali, ali pri izraženosti afektivnih dimenzij temperamenta in impulzivnosti prihaja do razlik med moškimi in ženskami ter, ali med obema dimenzijama obstajajo pomembne povezave in kako, če sploh, se le-te razlikujejo med spoloma.

Rezultati raziskave so pokazali pomembne razlike med moškimi in ženskami tako v izraženosti afektivnih dimenzij temperamenta, kot tudi v izraženosti splošne impulzivnosti ter impulzivnosti v načrtovanju. Ženske imajo bolj izražene lastnosti kot so omejenost z rutino, samoobtoževanje, bolj so občutljive na kritiko in čustveno labilne, njihovo razpoloženje in energija pogosto nihata, so tudi bolj zaskrbljene in imajo nekoliko pogostejše težave, ki so povezane s somatiko. Moški pa so za razliko bolj optimistični, z večjo željo po zabavanju, so bolj šaljivi, zgovorni in samozavestni, nagnjeni k tveganju in včasih tudi bolj vsiljivi, hkrati pa so nekoliko bolj kritični, nezadovoljni, jezavi, razdražljivi in bolj nagnjeni k pritoževanju in agresivnosti v primerjavih z ženskami. Rezultati se ujemajo tudi s preteklimi študijami (Erfurth in dr., 2005; Figueira in dr., 2008; Pompili in dr., 2008; Rózsa in dr., 2008; Vahip in dr., 2005; Vázquez in dr., 2007), kjer se je pokazal podoben trend. Poleg razlik v samem

temperamentu pa se moški in ženske razlikujejo tudi v impulzivnosti, čeprav razlike na tem področju niso tako zelo velike. Moški so v primerjavi z ženskami nekoliko bolj orientirani v sedanjosti in nimajo toliko vpogleda v prihodnost, pa tudi sama impulzivnost kot poteza je pri njih bolj izražena. V tem oziru naši rezultati niso popolnoma v skladu s primerljivimi raziskavami, ki navajajo, da med spoloma ne obstaja pomembnih razlik v impulzivnem vedenju (Patton, Stanford in Barratt, 1995; Stanford in dr., 2009).

Naša raziskava je odkrila nekatere pomembne povezave med različnimi aspekti impulzivnosti in afektivnimi dimenzijami temperamenta. Najbolj izstopajo relativno visoke pozitivne korelacije ciklotimne dimenzije s prav vsemi aspekti impulzivnosti. Kar navaja na zaključek, da so tisti posamezniki, ki kažejo več lastnosti nihanja razpoloženja in energije in so bolj čustveno labilni, hkrati tudi bolj impulzivni. Takšne osebe se težje osredotočijo na trenutne naloge, imajo manj kognitivne vztrajnosti, reagirajo spontano, hitro in nepremišljeno ter so osredotočeni bolj na sedanjost, s slabšim vpogledom v prihodnost. To se ujema tudi z ugotovitvami Signorette in sodelavcev (2005), ki definirajo ciklotimno dimenzijo temperamenta kot najbolj odklonsko in povezano z impulzivnostjo ter nekaterimi vedenjskimi in emocionalnimi težavami.

Med posameznimi aspekti impulzivnosti smo opazili najbolj izstopajoče povezave pri impulzivnosti v načrtovanju, kjer se pojavijo razmeroma visoke korelacije z vsemi dimenzijami temperamenta. Zanimivo je, da visoka izraženost hipertimne dimenzije temperamenta korelira negativno z impulzivnostjo pri načrtovanju, kar lahko pomeni, da so tisti posamezniki, ki so veseli, optimistični, šaljivi, neutrudljivi, nagnjeni k tveganju in vsiljivi, tudi bolj orientirani v prihodnost, z višjo samokontrolo in bolj sposobni kognitivne kompleksnosti. Ta podatek ponovno kaže na nejasno vlogo samega hipertimnega temperamenta, na kar opozarjajo tudi nekateri drugi raziskovalci (Cavanagh, 2004; Koukopoulos in Ghaemi, 2009). Lastnosti hipertimne dimenzije temperamenta bi namreč lahko predstavljale konstelacijo adaptativnih lastnosti, ki so, če le niso izražene v pretirani meri, osnova zdravi osebnosti. Meja med zdravim in odklonskim pri tej dimenziji je torej precej nejasna, saj po drugi strani nekatere raziskave kažejo pomembnost hipertimne dimenzije za razvoj same klasične bipolarni motnje (Angst in Marneros, 2001). Tudi naši rezultati nakazujejo bolj adaptativno vlogo hipertimnega temperamenta. Pri posameznikih, ki imajo več depresivnih, ciklotimnih, razdražljivih ali anksioznih lastnosti lahko opazimo ravno obraten trend – pomanjkanje vpogleda v prihodnost, nizko samokontrolo in nezmožnost kognitivne kompleksnosti. To se v določeni meri sklada tudi z nekaterimi preteklimi raziskavami, kjer so ugotovili povezanost načrtovanja z predvsem depresivnimi lastnostmi (Patton, Stanford in Barratt, 1995; Swann in dr., 2008).

Dokaj visoke korelacije lahko opazimo tudi med razdražljivo dimenzijo afektivnega temperamenta in impulzivnostjo pri pozornosti in načrtovanju, pri ženskah pa

tudi z motorično impulzivnostjo. Kritični, nezadovoljni, razdražljivi, jezavi posamezniki, nagnjeni k stalnemu pritoževanju, k nasilju in ljubosumju, kažejo tudi več odkrenljivosti, se težje osredotočajo na nalogo, manjka jim samokontrole in vpogleda v prihodnost. Ženske s tovrstnimi razdražljivimi lastnostmi pa hkrati kažejo več spontanega in nepremišljenega reagiranja kar potrjujejo tudi izsledki Stanforda, Grevea in Dickensa Jr. (1995), kjer so razkrili povezanost impulzivnosti z lastnostmi razdražljivega temperamenta. Avtorji so menili, da sta impulzivnost in razdražljivost sestavni del vedenjske kontrole, ki je vpletena tudi v samo inhibicijo agresivnih izbruhov. Podoben trend lahko opazimo tudi v naši raziskavi.

Pri ženskah smo opazili negativno povezanost depresivne dimenzije z motorično impulzivnostjo, kar pomeni, da bodo posameznice, ki so bolj omejene z rutino, samokritične, občutljive in nagnjene k samoobtoževanju, hkrati kazale manj spontanega in nepremišljenega reagiranja, medtem ko pri moških tega trenda nismo opazili. Prav tako so zanimive razlike v povezavah z anksiozno dimenzijo temperamenta, kjer smo pri moških opazili pozitivno povezanost anksioznih lastnosti z odkrenljivostjo in nesposobnostjo osredotočiti se na neko nalogo, medtem ko pri ženskah te povezave niso pomembne. Posameznice, ki so torej bolj zaskrbljene in anksiozne, kažejo tudi manj spontanega reagiranja.

6. ZAKLJUČEK

V raziskavi smo proučevali, ali med moškimi in ženskami obstajajo razlike v izraženosti afektivnih dimenzij temperamenta in impulzivnosti ter kakšna je medsebojna povezanost omenjenih dimenzij. Rezultati naše študije so pokazali, da se pojavljajo pomembne razlike med spoloma tako v okviru afektivnih dimenzij temperamenta, kakor tudi pri različnih aspektih impulzivnosti. Tako so bile depresivne, ciklotimne in anksiozne lastnosti bolj izražene pri ženskah, razdražljive in hipertimne pa pri moških. Moški so imeli tudi na splošno bolj izraženo dimenzijo impulzivnosti. Afektivne dimenzije temperamenta so se povezovale z različnimi vidiki impulzivnosti. Edino ciklotimna dimenzija temperamenta je korelirala z vsemi aspekti impulzivnosti pri obeh spolih medtem, ko so se ostale dimenzije povezovale z impulzivnostjo zelo različno in tudi v odvisnosti od spola. Številne raziskave so potrdile, da ima TEMPS-A dobro prognostično moč predvsem glede diferenciacije med različnimi skupinami oseb z bipolarno motnjo in pri oceni tveganj za razvoj motenj razpoloženja pri ogroženih posameznikih. Njegovo prognostično moč potrjuje tudi naša raziskava, saj se je ciklotimna dimenzija, ki predstavlja pomemben prediktor razvoja bipolarnih motenj razpoloženja, pomembno povezovala pri obeh spolih z impulzivnostjo, torej z vedenjskim vzorcem, ki je značilen za osebe z bipolarno motnjo razpoloženja.

*** Študija je bila izvedena v sklopu raziskovalnega projekta L3-9698-2485-06, ki sta ga sofinancirala ARRS in Astrazeneca.*

6. LITERATURA

- [1] Akiskal, H. S. in Akiskal, K. K. (2005). Epilogue. The Renaissance of the Ancient Concept of Temperament (with a Focus on Affective Temperaments). V M. Maj, H. S. Akiskal, J. E. Mezzich in A. Okasha (ur.), *Personality disorders* (str. 479 – 500). Chichester: John Wiley & Sons Ltd.
- [2] Angst, J. in Marneros, A. (2001). Bipolarity from ancient to modern times: conception, birth and rebirth. *Journal of Affective Disorders*, 67 (1-3), str. 3 – 19.
- [3] Arce, E. in Santisteban, C. (2006). Impulsivity: a review. *Psicothema*, 18 (2), str. 213 – 220.
- [4] Cavanagh, J. (2004). Epidemiology and Classification of Bipolar Disorder. V M. Power (ur.), *Mood Disorders: A Handbook of Science and Practice* (str. 1 – 52). Chichester: John Wiley & Sons Ltd.
- [5] Erfurth, A., Gerlach, A. L., Michael, N., Boenigk, I., Hellweg, I., Signoretta, S., Akiskal, K. K. in Akiskal, H. S. (2005). Distribution and gender effects of the subscales of a German version of the temperament autoquestionnaire briefTEMPS-M in a university student population. *Journal of Affective Disorders*, 85 (1-2), str. 71 – 76.
- [6] Figueira, M. L., Caeiro, L., Ferro, A., Severino, L., Mateus, P., Abreu, M., Akiskal, H. S. in Akiskal, K. K. (2008). Validation of the Temperament Evaluation of Memphis, Pisa, Paris and San Diego (TEMPS-A): Portuguese-Lisbon version. *Journal of Affective Disorders*, 111 (2-3), str. 193 – 203.
- [7] Koukopoulos, A. in Ghaemi, S. N. (2009). The primacy of mania: A reconsideration of mood disorders. *European Psychiatry*, 24 (2), str. 125 – 134.
- [8] Patton, J. H., Stanford, M. S. in Barratt, E. S. (1995). Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology*, 51 (6), str. 768 – 774.
- [9] Perugi, G. (2005). Are Cyclothymic Temperament and Borderline and Histrionic Personality Related Concepts?. V M. Maj, H. S. Akiskal, J. E. Mezzich in A. Okasha (ur.), *Personality disorders* (str. 260 – 263). Chichester: John Wiley & Sons Ltd.
- [10] Pompili, M., Girardi, P., Tatarelli, R., Iliceto, P., De Pisa, E., Tondo, L., Akiskal, K. K. in Akiskal, H. S. (2008). TEMPS-A (Rome): Psychometric validation of affective temperaments in clinically well subjects in mid- and south Italy. *Journal of Affective Disorders*, 107 (1-3), str. 63 – 75.
- [11] Rózsa, S., Rihmer, Z., Gonda, X., Szili, I., Rihmer, A., Kó, N., Németh, A., Pestaloty, P., Bagdy, G., Alhassoon, O., Akiskal, K. K. in Akiskal, H. S. (2008). A study of affective temperaments in Hungary: Internal consistency and concurrent validity of the TEMPS-A against the TCI and NEO-PI-R. *Journal of Affective Disorders*, 106 (1-2), str. 45 – 53.
- [11] Signoretta, S., Marenmani, I., Liguori, A., Perugi, G. in Akiskal, H. S. (2005). Affective temperament traits measured by TEMPS-I and emotional behavioral problems in clinically-well children, adolescents, and young adults. *Journal of Affective Disorders*, 85 (1-2), str. 169 – 180.
- [13] Stanford, M. S., Greve, K. W. in Dickens Jr., T. J. (1995). Irritability and impulsiveness: relationship to self-reported impulsive aggression. *Personality and Individual Differences*, 19 (5), str. 757 – 760.
- [14] Stanford, M. S., Mathias, C. W., Dougherty, D. M., Lake, S. L., Anderson, N. E. in Patton, J. H. (2009). Fifty years of the Barratt Impulsiveness Scale: An update and review. *Personality and Individual Differences*, 47 (5), str. 385 – 395.
- [15] Swann, A. C., Steinberg, J. L., Lijffijt, M. in Moeller, F. G. (2008). Impulsivity: Differential relationship to depression and mania in bipolar disorder. *Journal of Affective Disorders*, 106 (3), str. 241 – 248.
- [16] Vahip, S., Kesebir, S., Alkan, M., Yazici, O., Akiskal, K. K. in Akiskal, H. S. (2005). Affective temperaments in clinically-well subjects in Turkey: initial psychometric data on the TEMPS-A. *Journal of Affective Disorders*, 85 (1-2), str. 113 – 125.
- [17] Vázquez, G. H., Nasetta, S., Mercado, B., Romero, E., Tifner, S., Ramón, M. L., Garelli, V., Bonifacio, A., Akiskal, K. K. in Akiskal, H. S. (2007). Validation of the TEMPS-A Buenos Aires: Spanish psychometric validation of affective temperaments in a population study of Argentina. *Journal of Affective Disorders*, 100 (1-3), str. 23 – 29.

NOVI POGLEDI NA DELOVANJE MOŽGANOV: OD SUBMOLEKULSKE DO GLOBALNE RAVNI

Matej Plankar

Inštitut BION

Stegne 21, 1000 Ljubljana

Tel: 01 513 11 46; email: matej.plankar@bion.si

IZVLEČEK

Vse bolj jasno postaja, da teorija klasične kemične sinapse ne zmore pojasniti spomina in prav tako ne zadošča za razlago kompleksnega informacijskega procesiranja možganih. V biofiziki se vse bolj uveljavlja koncept biološke koherence, ki omogoča neprimerno bogatejšo procesualne zmožnosti. Citoskelet pridobiva ključno vlogo kot materialni substrat kognicije.

1 UVOD

Klasična nevrološka doktrina, ki velja še iz časov nevrofiziologa Cajala (ki je v začetku 20. stol. prispeval ključna nevroanatomska spoznanja) je, da informacija v obliki spremembe transmembranskega potenciala potuje od aksonskega vhoda preko kemične sinapse do dendrita ali some drugega nevrona (klasična kemijska sinapsa). Kadar postsinaptični nevron doseže vzdražni prag (z nabiranjem ekscitacijskih in/ali preprečevanjem inhibicijskih membranskih potencialov), pride do sprožitve serije akcijskih potencialov (AP), katerih frekvenca oziroma časovni vzorci odgovarjajo jakosti dražljaja in/ali vedenja. Zaradi kompleksne integracije mnogih vhodov in enega samega izhoda, ki spominja na računalniško procesiranje, veljajo AP oziroma njihovi vzorci še danes za nevrološki kod informacijskega procesiranja in s tem tudi same kognicije (AP kot nevrološki ekvivalent informacijskemu bitu) (Hameroff 2006). Prepustnost kemične sinapse je odvisna od sinaptične moči, ki jo primarno modulirajo živčni prenašalci skupaj z lastnimi nabori receptorjev, sekundarnih obveščevalcev in drugih regulatornih molekul. Zmožnost spreminjanja sinaptične moči imenujemo sinaptična plastičnost in velja za biokemijski temelj učenja in spomina.

2 NEZADOSTNOST KEMIČNIH SINAPS

Danes vemo, da je povezljivost klasičnih kemičnih sinaps (akson-dendrit) v sesalčjih možganih visoka le v področjih iste modalnosti (npr. centri za voh, sluh), medtem ko je intermodalna povezljivost majhna ali nična (Woolf 2006). Poleg tega povezave, vključno s kortiko-kortikalnimi, večinoma ohranjajo topografijo, kar zmanjšuje njihovo razpršenost in s tem povezljivost. Kako lahko na ta način pride do vezave različnih modalnosti in enovitega, sklenjenega doživljanja?

Čedalje bolj jasno postaja, da klasična sinapsa ni edini način informacijskega procesiranja med nevroni. Nekateri kortikalni nevroni namreč nimajo aksonov, dendriti komunicirajo z drugimi dendriti preko električnih sinaps, ki so pravzaprav presledkovni celični stiki in tvorijo citoplazemski kontinuum. Vloga električnih sinaps je danes zreducirana na zgolj embrionalno fazo zaradi precejšnjega upada njihovega števila kasneje v razvoju, a v novjšem času znova pridobivajo na pomenu, tokrat v novi luči biofizikalnih spoznanj o izraziti elektromagnetni naravi bioloških procesov (razdelek 3). Pri tovrstnih prenosih membranskih potencialov ni nujno, da pride do nadpraznega vzdraženja, zato jih v splošnem označujejo kot šum, torej epifenomen. Vendar pa so prav ti podpražni signali zmožni koherentnih oscilacij preko makroskopskih možganskih območij z izredno majhnim faznim zamikom (klasične Hebbove nevrnske skupnosti niti ne morejo doseči tovrstne koherence zaradi omejene hitrosti AP). Natančnejše študije so pokazale, da se koherentne oscilacije pojavijo na dokaj majhem območju nevronov, nato poniknejo in se znova pojavijo na drugem območju. Gama-frekvenčna koherenca elektroencefalograma (EEG) pri 40-80Hz velja za najbližji nevrološki približek pozornosti in zanj so odgovorne prav električne sinapse oziroma presledkovni stiki (Hameroff 2006). S poskusi na opicah so pokazali, da se sinhronizacija distalnih nevronov, vpletenih v različne percepcijske modalnosti (npr. barva, oblika, gibanje), zgodi šele, kadar je zunanji dražljaj opicam predstavljen tako, da ga zaznajo kot enoten dogodek (McFadden 2006).

Presledkovni stiki se pojavljajo med dendriti, aksoni in drugimi aksoni, med nevroni in glia celicami in med glia¹ celicami. Hameroff (2006) imenuje tovrstne globalne skupnosti povezanih celic hipernevroni. Zaradi različnih sinaps lahko nevroni delujejo večplastno (npr. inhibitorni interneuroni imajo mnogo električnih sinaps, ki posredujejo koherentne oscilacije, medtem ko njihovi aksoni tvorijo inhibitorno GABA sinapso). En nevron ima lahko številne

¹ Glia celice, ki jih je sicer mnogo vrst, po številu prekašajo nevrone pribl. za faktor 10. Samo prostornina astrocitov (prevladujoča vrsta) lahko znaša do 50% možganske prostornine. Pogledi na delovanje glia celic so se močno spremenili, sprva so menili, da so zgolj metabolna opora nevronom, danes pa je znano, da so aktivno udeležene v modulaciji sinaptične plastičnosti. Spoznanja o njihovi vključenosti v hipernevronske mreže globalnih koherentnih oscilacij pa njihovo vlogo še razširjajo (Kreft 2008, Wikipedia: Glia cell, 2008).

presledkovne stike, tako da bi teoretično lahko bile vse možganske celice povezane v en sam hiperneuron, vendar pa bi previsoka sinhronizacija zmanjšala procesualno kapaciteto (nenavadno visoka sinhronizacija je npr. med epileptičnimi napadi), tako da so v danem trenutku samo nekateri izmed njih odprti. Poskusi kažejo, da je njihovo delovanje regulirano zelo kompleksno (mikrotubuli, G proteini); modulacija električnih sinaps je na ta način domnevno vsaj tako dinamična kot plastičnost kemičnih (Hameroff 2006).

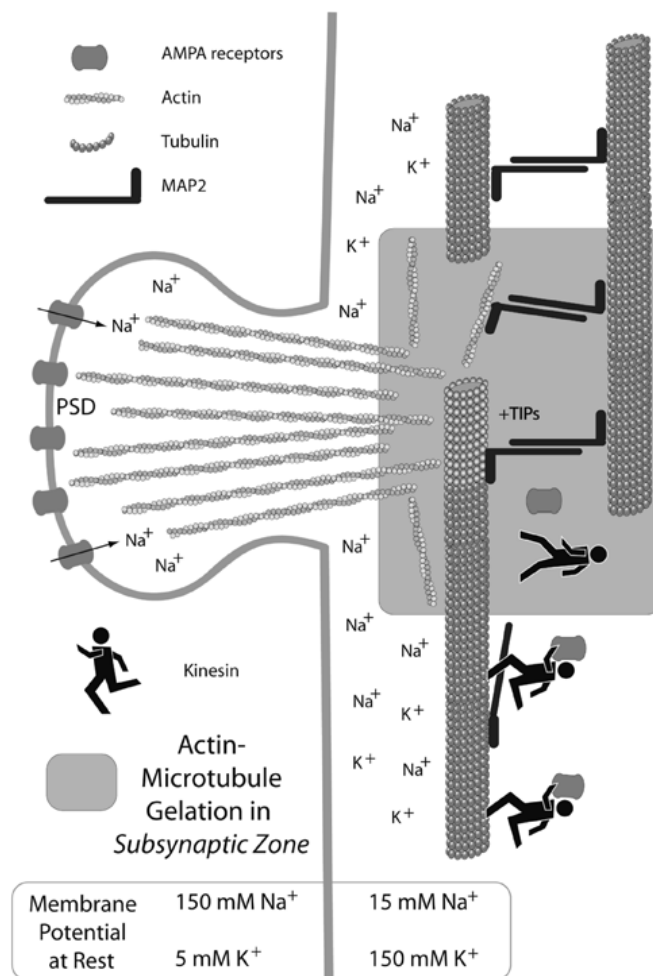
Drugi argument proti kemičnim sinapsam se nanaša na spomin. Mnogi poskusi namreč kažejo, da se po nekaj urah do enega dneva po intenzivnem učenju število sinaps sicer poveča, vendar pa že po enem do treh dneh pade na raven pred učenjem. Sinapse so podvržene dinamičnemu toku stalnega preurejanja, s tem da se njihovo število oziroma gostota dolgoročno praktično ne spreminja. Sinaptični »pretok« je torej podvržen strogim kontrolnim mehanizmom, ki število sinaps vzdržujejo znotraj določenih mej. Ta dejstva nakazujejo, da vsaj za dolgoročni spomin, torej kot mesto za shranjevanje informacij, klasične sinapse niso primeren kandidat (Woolf 2006).

Boljši kandidat za fizikalno osnovo spomina je podsinaptični prostor (Slika 1). To je območje goste mreže citoskeletnih elementov: MT, aktinski filamenti (AF) in povezovalne MAP2 molekule ter molekularnih motorjev in drugih strukturnih in regulatornih molekul, ki povezuje postsinaptično zgoščino² z notranjostjo dendrita. Po vezavi živčnega prenašalca sproščeni ioni (ki se prenašajo vzdolž aktinskih filamentov v postsinaptični zgoščini podobno kot po električnem vodniku in torej ne difundirajo) oziroma sekundarni obveščevalci posredno vplivajo na stopnjo povezljivosti in polimerizacije aktinskih filamentov in MT v podsinaptičnem prostoru ter s tem na kratkoročno (npr. povezljivost s kanalčki) in dolgoročno (npr. delovanje molekularnih motorjev, ki prenašajo strukturne in regulatorne elemente za delovanje sinapse) modulacijo prepustnosti sinapse. Morda še pomembnejši učinek pa je možnost modulacije povezljivosti celotnega citoskeletnega omrežja v nevronu. Ena ključnih strukturno-funkcionalnih molekul pri tem je MAP2, katere dimeri se vežejo na MT in jih povezujejo v stabilno omrežje, obenem pa vežejo tudi aktinske filamente v postsinaptični zgoščini³ (Woolf 2006). Na ta način so praktično vse sinapse na nivoju nevrona, ob

² Gosto omrežje aktinskih filamentov v sinaptičnem gumbku (synaptic spine), ki funkcionalno povezuje pretok ionov med ionskimi kanali in podsinaptičnim prostorom.

³ O osrednji strukturno-regulatorni vlogi MAP2 govori tudi dejstvo, da ima kar 40 (ali več) mest za fosforilacijo, kar pomeni, da je njeno delovanje zelo kompleksno regulirano (v smislu molekularnih mrež predstavlja pomembno vozlišče), in je obenem največji porabnik metabolne energije v dendritu. Znane so tudi nekatere nevrodegenerativne bolezni zaradi motenj delovanja MAP2 (tauopatije) kot npr. Alzheimerjeva bolezen (Hameroff 2006).

upoštevanju elektromagnetne komunikacije preko električnih sinaps pa tudi na globalnem (internevronske) nivoju, povezane v enotno mrežo z možnostjo natančnega usmerjanja povezljivosti njenih elementov, kar implicira njen informacijsko-procesualni značaj.



Slika 1: Postsinaptična zgoščina (levo) in podsinaptični prostor (desno)

Podsinaptični prostor je torej vmesnik, ki združuje lokalno in globalno kontrolo, najbolj neposredno določa lokacijo, velikost in prepustnost kemične kot tudi električne sinapse, procesi v njem pa so odvisni tako od sinapse same kot tudi od globalnega omrežja citoskeletnih elementov oz. njihove elektrodinamike. Na sinapse lahko torej bolj kot na procesualne enote gledamo z vidika komunikacijskih kanalov, kot neke vrste radijskih sprejemnikov, ki šele na najvišji hierarhični ravni usmerjajo subtilnejše in predvsem neprimerno bogatejši informacijske tokove na globalni, internevronske ravni. Če so torej akcijski potenciali in sinaptična plastičnost le vrh ledene gore in kot taki niso primarna substanca kognitivnih procesov, je naslednje vprašanje, kaj je torej materialna oz. biofizikalna osnova informacijskih tokov v možganih, ki omogočajo kognicijo ter kako nastanejo. Tudi v kognitivni znanosti se vse bolj

uveljavlja koncept biološke koherence, ki pri tem igra ključno vlogo.

3 BIOLOŠKA KOHERENCA

3.1 Nastanek in pomen koherence

Klasična molekularna znanstvena paradigma, ki je epistemološko redukcionistična, ne razloži nekaterih bazičnih bioloških problemov, npr. časovno-prostorske razlike med mestom sprostitve ter mestom porabe metabolne energije, hitrosti biokemijskih reakcij (z uporabo statistične mehanike oziroma Brownovega gibanja) (Vitiello 2001), večje porabe energije za vzdrževanje transmembranskega potenciala, kot jo zmore dovesti celični metabolizem (Jaeken 2007), izkoristka pretvorbe metabolne energije v mehansko (Del Giudice 2005) itd. Vsem problemom je skupno, da termalizacija metabolne energije ne razloži visoke učinkovitosti bioloških procesov. Iz tega sledi, da drugi zakon termodinamike takorekoč vsiljuje kolektivno dinamiko bioloških molekul, kjer so vse komponente sistema podvržene isti dinamiki s skupno fazo in frekvenco, kar pomeni, da se ne termalizirajo (niso podvržene Maxwell-Boltzmannovi statistični porazdelitvi in ne zvišujejo temperature sistema), pač pa se pretvarjajo iz ene v drugo obliko kolektivno organizirane elektrokemične energije, zmožne visoko specifičnega prepoznavanja in interakcij (Del Giudice in sod. 2005, 2009). Prvi, ki je uvidel, kako tovrsten red nastane v bioloških sistemih, je bil angleški biofizik Fröhlich (1975, 1977), po katerem sproži koherentni režim nadkritičen dotok metabolne v povezan sistem dipolov (npr. protein) ob dovolj visoki polarizacijski gostoti. Dotok energije povzroči preusmeritev (entrainment) oziroma fazni prehod sprva termično vzbujenih (difuznih) oscilacij dipolov v enoten nihajni način ali polarizacijski režim, kjer celoten sistem dipolov oscilira v isti fazi in z isto osnovno frekvenco oziroma harmoniki. Pomen koherence je v kolektivni amplitudi elementov koherentnega sistema (dipolov), kar si lahko predstavljamo kot skupino vojakov, ki sinhrono koraka čez most; če se frekvenca korakov resonančno ujame z naravno frekvenco mostu, se lahko most poruši. Isti princip velja za delovanje biomolekul, kar implicira visoko specifično resonančno interakcijo molekul, kjer se te lahko prepoznavajo in interagirajo brez fizičnega stika, preko interferenčnih vzorcev EM polja (Vitiello 2001).

Italijanska skupina (Del Giudice, Vitiello, Preparata, Doglia in Milani) je Fröhlichov model nadgradila v okviru kvantne teorije polja. Po tej teoriji za nastanek koherence ni več potrebna specifična metabolna energija kot npr. ATP, pač pa koherentni režim inducirajo spontane vakuumske fluktuacije vektorskega potenciala elektromagnetnega polja. Med atomi oziroma molekulami (katerih gostota mora presegati določen prag, zato velja samo za kondenzirne snovi), ki so v dosegu fluktuacije, se vzpostavijo fazne korelacije dolgega dosega. EM fluktuacija (fotoni) se na ta način ojača in stabilizira, zaradi efektivne mase pa tudi »ujame« v sistem – koherentno domeno, ki tako ne more več

oddati energije navzven, zato ima tak sistem obenem tudi najnižjo energijo. Sistem izgubi rotacijsko invarianco (simetrijo), saj so komponente sistema vezane v kolektivno dinamiko z omejenimi (kolektivnimi) stopnjami prostosti in se torej ne morejo več gibati vsaka zase, temveč na točno določen način, čemur pravimo zlom simetrije. Prav skupna (omejena) dinamika omogoča najnižjo energijo in stabilizira sistem kot osnovno stanje (Del Giudice in sod. 1988, 2005, 2009a b, Preparata 1995, Vitiello 2001).

3.2 Vloga citoskeleta pri informacijskem procesiranju

Model italijanske skupine je še nadalje izpopolnila ter ga aplicirala tudi na kognitivno področje širša mednarodna skupina raziskovalcev, katerih glavni predstavniki so Mershin, Mavromatos, Nanopoulos, Hameroff, Penrose, Tuszynski, Pribram, Jibu, Yasue idr. Nevroni so dobesedno napolnjeni z mikrotubuli, ki so kot osrednji element citoskeleta tudi daleč najbolj raziskani; spoznanja o njihovi integrativni funkciji pri uravnavanju delovanja sinaps in spomina vse bolj odpirajo tudi vprašanja o njihovi zmožnosti prenosa ter procesiranja informacij. Njihova klasična vloga, zreducirana zgolj na bolj ali manj pasivne strukturne elemente oziroma na »tračnice«, vzdolž katerih molekularni motorji prevažajo celični tovor, je tako že zdavnaj presežena.

Za izrazite elektrodinamične lastnosti MT sta ključni fizikalni lastnosti feroelektričnost (spontana ureditev dipolov ob nadpražni aplikaciji zunanega električnega polja) in zmožnost preklapljanja dipolov v polimeriziranem stanju. Konformacija tubulinov (gradnikov MT⁴) je namreč zelo labilna, zato lahko že zelo majhna sprememba energije okolice povzroči preklon dipolnega momenta in s tem spremembo konformacije MT (Mershin in sod. 2006). Zaradi povezanosti s sosednjimi tubulini (MT lahko obravnavamo kot nagibno heksagonalno mrežo) lahko prihaja do kompleksnih medsebojnih interakcij – preračunavanj oz. integracije vhodnih in izhodnih stanj. Tovrstno klasično procesiranje so sprva računalniško modelirali kot celični avtomat, podoben igri življenja, kjer mreža povezanih celic z vsako generacijo dobiva nove oblike, ki se lahko ohranjajo in celo razvijajo (Hameroff 2006). V kolikor predpostavimo, da so stanja tubulinov v kvantni superpoziciji (koherenci), se hitrost interakcij drastično poveča, saj superpozicija omogoča paralelno procesiranje poljubnega števila kvantnih stanj. MT na ta način lahko deluje kot kvantni računalnik, katerega

⁴ Mikrotubuli so votle supramolekulske cevi z zunanjim premerom 25 in notranjim 14 nm. Gradbeni element MT je dimerni protein tubulin, sestavljen iz dveh skoraj identičnih, 50 kDa globularnih monomerov α - in β -tubulina. Pri fiziološkem pH je tubulin negativno nabit zaradi C₁₅ ostanka (»rep«) na zunanji površini tubulina. Večina negativnega naboja je skoncentrirana na α -tubulinu, tako da vsaka tubulinska molekula predstavlja permanentni električni dipol, po ocenah med 1200 in 1700 Debye.

delovanje lahko simuliramo s pomočjo kvantne Hopfieldove mreže (Behrman 2006).

Na osnovi omenjenega principa razvijata Penrose in Hameroff lasten model, ki mu pravita orkestrirana objektivna redukcija (Orch OR). Po tem modelu citomatriks (citoskelet skupaj z vodo in ioni) deluje kot globalni orkestrirni aparat za objektivno redukcijo valovne funkcije tubulinov oz. MT, ki na ta način kontinuirano nihajo med superponiranim, kvantno koherentnim stanjem, ki zajema celoten, globalno zastopan hiperneuron (ki naj bi po izračunih glede na energijo sistema obsegal okoli 10^4 celic) in klasičnim stanjem, pri čemer prvo omogoča informacijsko procesiranje (kot termično izoliran kvantni sistem), slednje pa integracijo (zapis informacije) v materialno strukturo možganskega aparata in s tem med drugim nastanek spominov.

Vlogo MT kot informacijsko-procesualnega substrata kognitivnih procesov še nadalje razširjajo spoznanja o njihovi predpostavljeni funkciji kvantno-elektrodinamskih votlin. Notranjost MT je po ocenah raziskovalcev dovolj izolirana pred termičnim šumom, da vodne molekule v njej ohranjajo koherentno elektromagnetno dinamiko dipolnih momentov, kar omogoča emisijo koherentnih fotonov (superradianca), ki lahko vzajemno interagirajo s tubulinskimi monomeri. Emisija koherentnih fotonov odpira povsem nove možnosti za kvantno procesiranje, saj fotoni medij "vidijo" kot popolnoma transparenten (fenomen samoinducirane transparence), kar jim omogoča neomejeno, tj. brez izgub, širjenje vzdolž filameta. MT na ta način delujejo kot valovni vodniki za koherentno elektromagnetno polje (podobno kot multimodalni laser, le da z mnogo manjšo amplitudo), ki s kompleksnimi interferenčnimi vzorci integrira dipolne interakcije tubulinov. Eksperimentalna spoznanja o vlogi električnih sinaps pri pozornosti oz. globalne koherentne oscilacije pa to funkcijo razširjajo tudi na internevronske raven (Jibu in sod. 1994, Vitiello 2001).

Omenjeni načini informacijskega procesiranja v možganih so le nekatere izmed številnih fascinantnih, predvsem pa neprimerno bogatejših možnosti, ki jih odpira nov koncept organizacije biološke materije, tj. biološke koherence. Pri tem bo eno ključnih biofizikalnih vprašanj, ali so biološki sistemi dovolj izolirani pred termičnim šumom, da lahko koherentne domene dosežejo makroskopsko, torej biološko relevantno prostorsko razsežnost ter časovno stabilnost, primerljivo z biološkimi procesi. Številna teoretična, vse bolj pa tudi eksperimentalna spoznanja, so zelo obetavna.

4 LITERATURA

[1] Behrman E., Gaddam K., Skinner S.R. 2006. Microtubules as a Quantum Hopfield Network. V: Tuszynski, J. (ur.): The Emerging Physics of Consciousness. Springer-Verlag Berlin.

[2] Del Giudice E., Doglia S., Milani M. in Vitiello G. 1988. Coherence of electromagnetic radiation in biological systems. *Cell Biophysics* 13:221-224.

[3] Del Giudice E., De Ninno A., Fleischmann M., Mengoli N., Milani M., Talpo G. in Vitiello G. 2005. Coherent Quantum Electrodynamics in Living Matter. *Electromagnetic Biology and Medicine* 24: 199–210.

[4] Del Giudice E. in Tedeschi A. 2009. Water and Autocatalysis in Living Matter. *Electromagnetic Biology and Medicine* 28: 46–52.

[5] Del Giudice E., Pulselli R.M. in Tiezzi E. 2009. Thermodynamics of irreversible processes and quantum field theory: An interplay for the understanding of ecosystem dynamics. *Ecological Modelling* 220:1874-1879.

[6] Fröhlich H. 1975. The extraordinary dielectric properties of biological materials and the action of enzymes. *Proc. Nat. Acad. Sci.* 72(11): 4211-4215.

[7] Fröhlich H. 1978. Coherent Electric Vibrations in Biological Systems and the Cancer Problem. *IEEE Transactions on Microwave Theory and Techniques* 26(8): 613-617.

[8] Hameroff S. 2006. Consciousness, Neurobiology and Quantum Mechanics: The Case for a Connection. V: Tuszynski, J. (ur.): The Emerging Physics of Consciousness. Springer-Verlag Berlin Heidelberg.

[9] Jaeken L. 2007. A New List of Functions of the Cytoskeleton. *IUBMB* 3: 127-133.

[10] Jibu M., Hagan S., Hameroff S., Pribram K. in Yasue K. 1994. Quantum optical coherence in cytoskeletal microtubules: implications for brain dynamics. *BioSystems* 32:195-209.

[11] Kreft M. 2008. Novi pogledi na astrocitate. *Interno gradivo*.

[12] McFadden N. 2006. The CEMI FIELD Theory: Seven Clues to the Nature of Consciousness. V: Tuszynski, J. (ur.): The Emerging Physics of Consciousness. Springer-Verlag Berlin Heidelberg.

[13] Mershin A., Sanabria H., Miller J.H., Nawarathna D., Skoulakis E.M.C., Mavromatos N.E., Kolomenskii A.A., Schuessler H.A., Luduena R.F., Nanopoulos D.V. 2006. Towards Experimental Tests of Quantum Effects in Cytoskeletal Proteins. V: Tuszynski, J. (ur.): The Emerging Physics of Consciousness. Springer-Verlag Berlin.

[14] Preparata G. 1995. QED Coherence in Matter. World Scientific Publishing Company.

[15] Vitiello G. 2001. My Double Unveiled: The Dissipative Quantum Model of Brain (Advances in Consciousness Research). John Benjamins Publishing Co.

[16] Woolf N. 2006. Microtubules in the Cerebral Cortex: Role in Memory and Consciousness. V: Tuszynski, J. (ur.): The Emerging Physics of Consciousness. Springer-Verlag Berlin Heidelberg.

VLOGA MOTIVACIJE PRIBLIŽEVANJA IN UMIKA PRI KOGNITIVNI KONTROLI EMOCIONALNIH DRAŽLJAJEV

Lilijana Šprah, Tatjana Novak

Družbenomedicinski inštitut, Znanstvenoraziskovalni center SAZU

Novi trg 2, 1000 Ljubljana, Slovenija

Tel: +386 1 470 6439; fax: +386 1 426 1493

e-mail: tatjana.novak3@guest.arnes.si

POVZETEK

V predstavljeni študiji smo proučevali značilnosti motivacijskih sistemov približevanja in umika ter kognitivno inhibicijo emocionalnih dražljajev pri zdravih osebah in pri osebah z bipolarno motnjo razpoloženja. Rezultati so pokazali, da lahko nekatere tendence motivacijskih sistemov približevanja in umika povežemo tudi z učinkovitostjo kognitivnega procesiranja relevantnih in motečih emocionalnih dražljajev. Izsledki raziskave govorijo v prid nekaterim nevrofiziološkimi modelom kognitivne inhibicije, ki predpostavljajo, da se deficiti v nekaterih nevroanatomskih področjih, povezanih z motivacijskimi dimenzijami, lahko odražajo v okrnjenosti kognitivnega procesiranja informacij z emocionalno valenco.

1.0 UVOD

Posameznik je vsakdanjem življenju soočen z različnimi značilnostmi dražljajev, tako v smislu različnih modalitet in intenzitete kot tudi njihove relevantnosti. Sposobnost njihove inhibicije, kadar so irelevantni, je bistvena za adaptivno vedenje. Vedenjski odziv na dražljaje je pogojen z vrsto kognitivnih procesov in z nekaterimi motivacijskimi vidiki posameznikovega funkcioniranja. Z razvojem novejših metod in tehnik, ki omogočajo identifikacijo nevrobioloških sistemov, vpletenih v čustveno in kognitivnega odzivanje, so raziskovalci oblikovali vedenjski model s katerim skušajo pojasniti povezanost med nagnjenji posameznika k določenim oblikam vedenja in med nekaterimi aspekti kognicije.

Depue in Iacono (1989) ter Gray (1990; 1991) so med prvimi opisali t.i. aktivacijski in inhibicijski vedenjski model, ki je vključeval biološke in okoljske dejavnike. Predpostavljali so obstoj dveh osnovnih motivacijskih sistemov, ki se odzivata na dejavnike okolja v smeri umika in približevanja. Omenjena motivacijska sistema predstavljata podlago vedenjskim in čustvenim reakcijam ter celo nekaterim osebnostnim dimenzijam in imata določljiv nevrobiološki substrat. To sta: motivacijski sistem umika (Behavioural Inhibition System - BIS) in nagrajevalni motivacijski sistem (Behavioural Activation System - BAS) (Carver in White, 1994; Amodio in sod., 2008). BAS in

BIS uravnava adaptivno vedenje. Medtem ko BAS usmerja akcijo glede na signale nagrade in nekaznovanja ter je povezan s pozitivnim čustvenim nabojem (vznesenost, ugodje, hrepenenje), pa se BIS odziva na vse dražljaje, ki povzročajo konfliktno odgovorne težnje kot je motivacija po hkratnem približevanju in umiku dražljaju, in je povezan z negativnim čustvenim nabojem (strah, tesnoba, neugodje). Odzivanje BIS-a se lahko odraža v prekinitvi vedenja v teku (vedenjska inhibicija), v preusmeritvi pozornosti ali v povečanem vznurjenju. Osebe s povečano občutljivostjo BIS in BAS raziskovalci povezujejo z širokim spektrom motenj razpoloženja in razvojem zasvojenosti (Depue in Iacono, 1989; Franken in sod., 2006; Alloy in sod., 2008). Model hiperaktivnega BAS, predpostavlja povečano odzivnost na nagrajujoče dražljaje in poudarjeno impulzivnost, kar lahko vodi tudi v stanja manjše in povečane dovzetnosti za zlorabo drog. Konsistentno z Grayjevo teorijo raziskave tudi potrjujejo povezanost med bolj občutljivim BIS-om in bipolarnimi depresivnimi fazami ter anksioznimi stanji (Johnson in sod., 2003, Campbell-Sills in sod., 2004).

Različne študije so potrdile vpletenost BIS in BAS v uravnavanje nekaterih kognitivnih procesov kot so pozornost (Putman in sod., 2004), delovni spomin (Gray in Braver, 2002) in vedenjska inhibicija (Amodio in sod., 2008). Slednja je še posebej zanimiva z vidika razumevanja funkcije kognitivne kontrole inhibicije, saj uravnava presojo o tem, kateri dražljaji imajo prednost in odzivanje nanje ter odpornost na moteče učinke nerelevantnih dražljajev. Omogoča procesiranje in odzivanje na relevantne dražljaje ob hkratnem inhibiranju in izločanju nerelevantnih dražljajev (Blasi in sod., 2006). Dražljaji z emocionalnim nabojem lahko še posebno ovirajo ali olajšajo kognitivno kontrolo inhibicije.

2.0 NAMEN RAZISKAVE

Namen raziskave je bil proučiti osnovne značilnosti in povezanost med občutljivostjo motivacijskih sistemov BIS in BAS ter kognitivno kontrolo inhibicije emocionalnih dražljajev pri zdravih osebah in pri osebah z bipolarno motnjo razpoloženja (BMR) v evtimični fazi.

3.0. METODA

3.1. Udeleženci

Sodelovalo je 64 zdravih prostovoljcev (24 moških in 40 žensk; starost: $36,23 \pm 11,95$ let; $14,59 \pm 2,82$ let šolanja) in 62 oseb z BMR v fazi eutimije (27 moških in 35 žensk; starost: $39,02 \pm 11,24$ let; $13,76 \pm 2,25$ let šolanja).

3.2. Uporabljeni instrumenti

BIS/BAS lestvica (Carver in White, 1994) je samoocenjevalna 4-stopenjska lestvica, ki meri individualne razlike v občutljivost BIS in BAS. BAS lestvico sestavljajo podlestvica *Težnje*, ki meri motivacijo in vztrajnost slediti zastavljenim ciljem, podlestvica *Odzivanje na nagrado*, ki meri tendenco odzivanja s povečano energijo in pozitivnimi čustvi v situacijah, kjer se pričakuje ali zgodi želeni dogodek, in podlestvica *Iskanje zabave*, ki meri impulzivno iskanje zabavnih in prijetnih situacij. BIS lestvica vsebuje nagnjenost k negativnemu čustvenemu odzivanju, anksioznosti in odzivanju s strahom v potencialno nevarnih situacijah. Več doseženih točk na posameznih podlestvicah odraža večjo občutljivost BIS/BAS sistema.

Emocionalni Go/NoGo test je instrument za oceno kognitivne inhibicije pri procesiranju dražljajev z emocionalno valenco, ki jih predstavljajo neprijetne, prijetne in nevtralne slike iz baze International Affective Picture System. Test se uporablja tudi kot mera kognitivnega konflikta med odgovoroma na tarčni in moteči dražljaj, ki se simultano aktivirata (Amodio in sod., 2008). Test vsebuje 6 blokov z dvema dražljajema, ki imata različno emocionalno valenco (npr. *prijetni – neprijetni*). Naloga udeleženca je bila reagirati na v naprej določene tarčne dražljaje (Go) s pritiskom na tipko in inhibirati odgovor na moteče dražljaje-distraktorje (NoGo) tako, da se vzdrži pritiska na tipko. Beležili so se reakcijski časi za pravilne odgovore na Go dražljaj in število pravilnih odgovorov glede na emocionalno valenco tarče oz. distraktorja.

3.3. Postopek

Pred vključitvijo v raziskavo so bili vsi udeleženci seznanjeni z njenim potekom in cilji ter so pisno privolili v sodelovanje. Testiranje je potekalo individualno, vsak udeleženec je najprej izpolnil BIS/BAS lestvico in se nato še po predhodni vaji udeležil Go/NoGo testa.

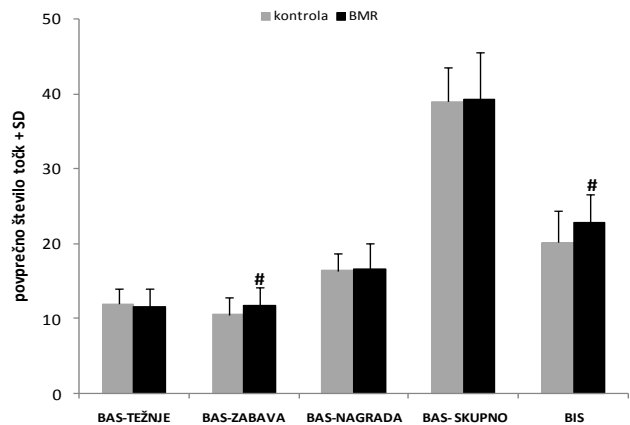
Rezultate smo analizirali s statističnim paketom SPSS 14.0, kot mero statistične pomembnosti smo vzeli $p < 0,05$ oz. $p < 0,01$. Pri preverjanju učinkov tarčnih in motečih dražljajev na reakcijske čase odgovarjanja, smo uporabili ANOVO s ponovljenimi merjenji. Po teoriji detekcije signalov smo vpeljali mero zaznavne občutljivost za emocionalno valenco tarčnega in motečega dražljaja - d' (Stanislaw in Todorov, 1999). S Pearsonovim korelacijskim koeficientom smo preverjali povezanost med posameznimi dimenzijami

BIS/BAS lestvice in reakcijskimi časi odzivanja na tarčne oz. moteče dražljaje ter mero zaznavne občutljivosti (d').

4.0 REZULTATI

4.1. BIS/BAS občutljivost

Osebe z BMR so imele v primerjavi z zdravimi osebami pomembno bolj izraženo komponento BAS-zabava ($t(123) = -2,66$, $p = 0,01$) in BIS ($t(123) = -3,83$, $p = 0,00$) (slika 1).



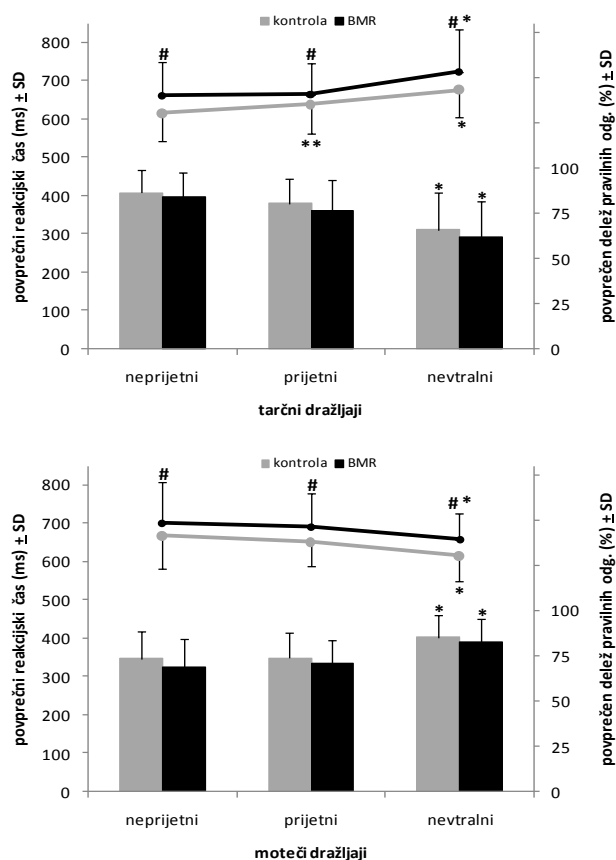
Slika 1. Primerjava dimenzij BIS in BAS občutljivosti med skupinama zdravih prostovoljcev in oseb z BMR.

4.2. Kognitivna kontrola inhibicije emocionalnih dražljajev

Osebe z BMR so imele glede na zdrave osebe pomembno pomembno okrnjeno funkcijo kognitivne inhibicije emocionalnih dražljajev. Odražala se je predvsem v daljših reakcijskih časih pri vseh emocionalnih valencah dražljajev ne glede na to, ali so se pojavljali v vlogi tarčnega dražljaja (Go) ali kot moteči dražljaj (NoGo) (glej # v grafikonih slike 2). Pravilnost odgovorov pa se med skupinama ni pomembno razlikovala. Kontrolna skupina je bila bolj zaznavno občutljiva na prijetne ($t(124) = 2,56$, $p = 0,01$) in nevtralne tarče ($t(124) = 2,03$, $p = 0,044$) in so jih manj motili nevtralni distraktorji ($t(124) = 2,06$, $p = 0,04$). Razlike med skupinama so se približale statistični pomembnosti ($p = 0,07$) tudi pri večji občutljivosti oseb z BMR na neprijetne in prijetne moteče dražljaje. V obeh skupinah se je izkazalo, da emocionalna valenca in interakcija med valenco in tipom dražljaja vpliva tako na reakcijske čase ($F(2,120) = 3,94$, $p = 0,02$ za valenco, $F(2,120) = 18,31$, $p = 0,00$ za interakcijo) kot na pravilnost odgovorov ($F(2,248) = 17,87$, $p = 0,00$ za valenco, $F(2,248) = 73,85$, $p = 0,00$ za interakcijo).

Analiza odzivov na tarčne dražljaje je pokazala, da sta imeli obe skupini daljše reakcijske čase pri ocenjevanju nevtralnih tarč v primerjavi s prijetnimi in neprijetnimi tarčami ($p = 0,00$ za obe razliki) (glej * grafikon zgoraj v sliki 2). Medtem ko so osebe z BMR primerljivo hitro odgovarjale na prijetne in neprijetne tarče, pa so kontrole hitreje odreagirale na neprijetne tarče v primerjavi s prijetnimi ($t(62) = -2,55$, $p = 0,02$) (glej ** grafikon zgoraj v

slike 2). Neprijetne tarče so bile najbolj pravilno prepoznane (delež pravilnih odgovorov: 86 ± 13 za kontrole in 84 ± 14 za osebe z BMR, $p=0,00$ za obe skupini), nevtralne tarče pa najmanj pravilno prepoznane (delež pravilnih odgovorov 66 ± 21 za kontrole in 62 ± 20 , $p=0,00$ za obe skupini) v obeh skupinah (glej * v grafikonu zgoraj slike 2).



Slika 2. Primerjava odzivanja na tarčne (grafikon zgoraj) in na moteče (grafikon spodaj) dražljaje med skupinama zdravih oseb in oseb z BMR. Vrstični grafikon predstavlja reakcijske čase odzivanja na različne tarčne/moteče dražljaje, stolpci v grafikonu pa predstavljajo delež pravilno prepoznanih tarčnih/motečih dražljajev (*- stat. pomembna razlika med različnimi tarčnimi/motečimi dražljaji znotraj posamezne skupine; # - stat. pomembna razlika med različnimi tarčnimi/motečimi dražljaji med skupinama).

Tabela 1. Korelacijski koeficienti med dimenzijami BIS/BAS in zaznavno občutljivostjo (d') za emocionalne valence posameznih tarčnih in motečih dražljajev (* $p<0,05$; ** $p<0,01$).

	BAS težnje		BAS zabava		BAS nagrada		BAS skupno		BIS		
	kontrola	BMR	kontrola	BMR	kontrola	BMR	kontrola	BMR	kontrola	BMR	
Tarčni dražljaji	Neprijetni	-0,142	0,086	0,039	-0,056	0,264*	0,116	0,088	0,095	-0,060	-0,155
	Prijetni	0,088	0,045	0,187	0,051	0,302*	0,242	0,282*	0,141	0,034	-0,052
	Nevtralni	0,017	0,158	0,292*	0,079	0,287*	0,113	0,294*	0,237	0,027	0,093
Moteči dražljaji	Neprijetni	-0,033	0,087	0,209	0,076	0,241	0,145	0,208	0,269*	0,039	0,117
	Prijetni	-0,031	0,171	0,166	0,057	0,279*	0,147	0,208	0,138	-0,155	-0,127
	Nevtralni	-0,010	0,055	0,220	-0,029	0,480**	0,211	0,344**	0,061	0,184	-0,042

V analizi inhibicije motečih dražljajev smo odkrili, da sta obe skupini porabili pomembno več časa za inhibicijo distraktorjev z emocionalno valenco v primerjavi z nevtralnimi distraktorji ($p=0,00$ za obe razliki) (glej * v grafikonu spodaj slika 2). Reakcijski časi za inhibicijo prijetnih in neprijetnih distraktorjev pa se niso bistveno razlikovali v nobeni od skupin. Skupini sta moteče dražljaje točneje prepoznavali v primeru nevtralnih distraktorjev kot v primeru prijetnih ali neprijetnih distraktorjev (delež pravilnih odgovorov: 85 ± 12 za kontrole in 83 ± 13 za osebe z bipolarno motnjo, $p=0,00$ za obe skupini) (glej * v grafikonu spodaj slika 2). Pozitivna in negativna emocionalna valenca motečih dražljajev ni pomembno vplivala na pravilnost odgovorov v nobeni skupini.

4.3. Povezanost med tendencami BIS/BAS in kognitivno inhibicijo emocionalnih dražljajev

Analiza povezanosti med hitrostjo odzivanja na tarčne in moteče dražljaje ter posameznimi BIS/BAS dimenzijami je pokazala, da je pri zdravih osebah edino dimenzija BAS-zabava pomembno ($p<0,05$) korelirala z odzivanjem na prijetne tarčne (-0,25) in neprijetne (-0,29) ter prijetne moteče dražljaje (-0,26). Medtem, ko se je pri osebah z BMR pojavila pomembna korelacija le med BIS in prijetnim distraktorjem (0,26). Prav tako se je pri zdravih osebah le dimenzija BAS (predvsem BAS-nagrada in BAS-skupni skor) pomembno povezovala z zaznavno občutljivostjo do tačnih in motečih dražljajev. Pri osebah z BMR so bile korelacije v tem oziru nižje in nepomembne z izjemo BAS-skupno in zaznavno občutljivostjo za neprijetne distraktorje (tabela 1). Dimenzija BIS se ni v nobeni od skupin pomembno povezovala z zaznavno občutljivostjo do emocionalne valence tarče oz. distraktorja.

5.0 DISKUSIJA

V predstavljeni raziskavi smo proučevali osnovne razlike med zdravimi osebami in osebami z bipolarno motnjo razpoloženja (BMR) v dimenzijah občutljivosti BIS in BAS ter v učinkovitosti procesiranja tarčnih in motečih dražljajev v okviru kognitivne inhibicije dražljajev z emocionalno valenco.

Rezultati naše študije so pokazali, da imajo osebe z BMR v primerjavi z zdravimi osebami, bolj izraženi dimenziji BAS-zabava in BIS, ki tudi sicer odslikavata nekatere značilnosti te motnje (npr. impulzivnost). Rezultati se skladajo z nekaterimi zaključki raziskav (Johnson in sod., 2003; Alloy in sod., 2008), ki spremembe v občutljivosti BAS-a in BIS-a povezujejo z značilnostmi manije in depresije pri bipolarni motnji. Pri osebah z BMR smo podobno, kot raziskovalci v primerljivih študijah (Murphy in sod., 1999; Swann in sod., 2009) identificirali

slabšo kognitivno kontrolo inhibicije emocionalnih dražljajev glede na zdrave osebe. Odražala se je predvsem v počasnejšem odzivanju na dražljaje, ne glede na emocionalno valenco tarče in distraktorja in v slabši zaznavni občutljivosti za prijetne oz. nevtralne tarčne dražljaje ter večji občutljivosti za neprijetne oz. prijetne moteče dražljaje. Obe skupini sta razvili vzorec t.i. emocionalne facilitacije, in sicer v obliki hitrejših odzivov na neprijetne in prijetne tarčne dražljaje v primerjavi z nevtralnimi. Nasprotno, se je vzorec t.i. emocionalne inhibicije pojavil pri počasnejšem odzivanju na neprijetne in prijetne moteče dražljaje. Emocionalna facilitacija in inhibicija sta bili bolj izraženi pri osebah BMR kot pri zdravih osebah, zato bi lahko bil ta fenomen tudi posledica okrnjene funkcije selektivne pozornosti in nekaterih nevrofizioloških korelatov, ki so značilni za motnje razpoloženja (Putman in sod., 2004).

Glede na to da raziskave, ki preučujejo nevrofiziološki substrat kognitivne kontrole (Hewig, 2005; Amodio in sod., 2008) izpostavljajo nekatera nevroanatomska področja, ki hkrati predstavljajo tudi nevrobiološko osnovo motivacijskim tendencam približevanja in umika, smo v naši študiji preverili, ali obstaja povezava med motivacijskim sistemom umika (BIS) oz. nagrajevalnim motivacijskim sistemom (BAS) in med nekaterimi vidiki kognitivne inhibicije emocionalnih dražljajev (hitrost odzivanja in zaznavna občutljivost na tarčne in moteče dražljaje). Izkazalo se je, da se zdrave osebe, z bolj izraženimi dimenzijami BAS hitreje odzivajo na prijetne tarče in jih manj motijo neprijetni in prijetni distraktorji. Poleg tega so bolj zaznavno občutljive za prijetne in nevtralne tarče. Dimenzija BIS se ni pomembno povezovala z nobenim aspektom kognitivne kontrole inhibicije. Pri osebah z BMR se je pokazala drugačna slika. Dimenzija BIS je korelirala s počasnejšim odzivanjem v situacijah s prijetnim distraktorjem, dimenzije BAS-a pa se v glavnem niso povezovala s kognitivno inhibicijo, z izjemo BAS-skupno, ki se je povezoval z večjo zaznavno občutljivostjo za neprijetne distraktorje. Rezultate lahko intepretiramo v luči nekaterih nevrobioloških študij ki navajajo, da BAS odslkava predvsem regulativni mehanizem, ki je povezan z aktivnostjo prefrontalnega področja možganske skorje, BIS pa se povezuje zlasti z mehanizmom zaznavanja konfliktnih situacij in z aktivnostjo akumbensa (Amodio in sod., 2008). Za BMR je značilna zmanjšana aktivnost v področjih prefrontalne skorje in povečana aktivnost v področjih limbičnega sistema kar rezultira v okrnjenem delovanju možganskih področij, odgovornih za sprejemanje odločitev, povezanih z emocionalnimi dražljaji in kontrolo pozornosti med procesiranjem le teh (Keener in Phillips, 2007). Slednje bi se lahko odražalo tudi v izsledkih naše raziskave, če predpostavljamo, da se pri zdravih osebah zaradi neokrnjenega delovanja prefrontalnih področij odslkava povezanost BAS dimenzij z ustrenejšo kognitivno inhibicijo, medtem ko se pri osebah z BMR tovrstna povezava izgubi s prevlado aktivnosti v limbičnih področjih in posledično bolj občutljive BIS dimenzije.

6.0 ZAKLJUČEK

Predstavljena raziskava se uvršča med redke vedenjske študije, ki obravnavajo povezavo med BIS/BAS tendencami in kognitivno kontrolo inhibicije, saj se večina raziskav na tem področju pretežno usmerja na nevrofiziološke korelate kognitivne inhibicije in večinoma brez vključevanja vpliva emocionalnih dražljajev in kliničnih vzorcev. Raziskava potrjuje uveljavljen model hiperobčutljivih BAS in BIS dimenzij pri osebah z bipolarno motnjo ter podkrepljuje dokaze o splošni deficitarni sposobnosti kognitivne inhibicije pri osebah z bipolarno motnjo.

* Študija je bila opravljena v sklopu raziskovalnega projekta L3-9698-2485-06, ki sta ga sofinancirala ARRS in Astrazeneca.

7.0 LITERATURA

- [1] Alloy LB, Abramson LY, Walshaw P, Cogswell A, Grandin L, Hughes M idr. (2008). Behavioral Approach System and Behavioral Inhibition System sensitivities and bipolar spectrum disorders: prospective prediction of bipolar mood disorders. *Bipolar Disorders*, 10; 310-322.
- [2] Amodio DM, Master SL, Yee CM, Taylor SE (2008). Neurocognitive components of the behavioral inhibition and activation systems: Implications for theories of self-regulation. *Psychophysiology*, 45; 11 - 18.
- [3] Blasi G, Goldberg TE, Weickert T, Das S, Kohn P, Zolnick B idr. (2006). Brain regions underlying response inhibition and interference monitoring and suppression. *European Journal of Neuroscience*, 23; 1658-1664.
- [4] Carver CS., White TL. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology*, 67;319 - 333.
- [5] Campbell-Sills L, Liverant GI, Brown TA. (2004). Psychometric evaluation of the behavioral inhibition/behavioral activation scales in a large sample of outpatients with anxiety and mood disorders. *Psychol Assess*, 16; 244-54.
- [6] Depue RA, Iacono WG. (1989). Neurobehavioral aspects of affective disorders. *Ann Rev Psychol*, 40; 457-492.
- [7] Franken IH, Muris P, Georgieva I. (2006). Gray's model of personality and addiction. *Addict Behav*,31(3); 399-403.
- [8] Gray JA. (1990). Brain systems that mediate both emotion and cognition. *Cognition and Emotion*, 4; 269-88.
- [9] Gray JA. Neural systems, emotion, and personality. In: Madden J IV, ed. *Neurobiology of learning, emotion, and affect*. New York: Hillsdale Press, 1991: 273-306.
- [10] Gray JR, Braver TS. (2002). Personality predicts working-memory-related activation in the caudal anterior cingulate cortex. *Cogn Affect Behav Neurosci*, 2(1); 64-75.
- [11] Hewig J, Hagemann D, Seifert J, Naumann E, Bartussek, D. (2005). The relationship of cortical activity and personality in a reinforced Go/NoGo paradigm. *Journal of Individual Differences*, 26(2); 86-99.
- [12] Johnson SL, Turner RJ, Iwata N. (2003). BIS / BAS levels and psychiatric disorder: an epidemiological study. *Journal of Psychopathology and Behavior Assessment*, 25; 25-36.
- [13] Keener MT, Phillips ML. (2007). Neuroimaging in bipolar disorder: a critical review of current findings. *Curr Psychiatry Rep*, 9(6); 512-20.
- [14] Murphy FC, Sahakian BJ, Rubinsztein JS, Michael A, Rogers RD, Robbins TW, Paykel AS. (1999). Emotional bias and inhibitory control processes in mania and depression. *Psychological Medicine*, 29; 1307-1321.
- [15] Putman P, Hermans E, van Honk J. (2004). Emotional stroop performance for masked angry faces: it's BAS, not BIS. *Emotion*, 4; 305-11.
- [16] Stanislaw H, Todorov N. (1999). Calculation of signal detection theory measures. *Behavior Research, Methods, Instruments & Computers*, 31(1); 137-149.
- [17] Swann AC, Lijffijt M, Lane SD, Steinberg JL, Moeller FG. (2009). Severity of bipolar disorder is associated with impairment of response inhibition. *Journal of Affective Disorders*, 116; 30-36.
- [18] Wessa M, Houenou J, Paillere- Martinot ML, Berthoz S, Leboyer M, Martinot JL. (2007). Fronto-striatal overactivation in euthymic bipolar patients during an emotional Go/NoGo task. *American Journal of Psychiatry*, 164; 638-646.

ON CONSCIOUSNESS; INSIGHTS FROM INTUITION, REFLECTION AND DIALOGUE

Eva Žerovnik

Department of Biochemistry, Molecular and Structural Biology

Jozef Stefan Institute

Jamova 39, 1000 Ljubljana, Slovenia

Tel: +386 1 4773753; fax: +386 1 4773984

e-mail: eva.zerovnik@ijs.si

ABSTRACT

Philosophers and mystics (men and women) have tried to unravel mysteries of human consciousness for centuries, among them was for example Ibn Rushd and his followers, who believed it was a collective phenomenon. Being totally aware of the limits of my own knowledge, I do not see why modern men and women would not try to contribute some. Therefore, I am going to write an essay on how to define consciousness intuitively. What all might comprise this phenomenon, what is it after all: a feeling, a state, a process? As well intuitively, I will try to define experimental design to probe some of these guesses. I will later on mention some newest results on the quantum nature of light and quantum dots - with the processes taking place in and around our brain - in mind. Some parts are written in the form of exchange of opinions – as it is easier to discuss philosophical issues in a dialogue. Thinking on the aim (purpose) of consciousness, I will finally ask if it worth worrying for the future of the Earth and current civilization, which neither one individual neither all simulations can predict. Is the time critical to use reason and intuition or is it time to become more spiritual and rely on prayers? If I cite my grandfather here: “Help yourself (do what you think is best) and God will help you.”

Instead of an Introduction:

We two like to talk about hard issues, raise hard questions, don't we? (I would say »isn't it?« And you would not understand that I say so, even though, I know that this is not 100% correct; I still continue saying so to be funny, to joke...and expect you would guess my internal dialogue). So, to come on the track: Why not pose hard questions and

why not bring our thoughts to the limit? That does not cost much yet might be revealing or disappointing, depending on the point of view. If he/she wants to arrive at the »truth« or fact, he'll be disappointed. If, however, one wants just to develop possible theories, options, discover ideas, he might be rewarded. Hard to know where our »**clear, recognizing thoughts of the world and ourselves in it**« (defining consciousness) come from? Where do numerous similarities and subjective ways of thinking come from? Is there a well, a source to dive into or, is this a result of human evolution, only, all hidden in the individual's brain? Could that not be that evolution would lead to a being(s), who would be capable to plug in into a higher, all-surrounding, or innermost source of the knowledge, wisdom, experiences – when developed that highly?

Can one define Consciousness? Is it :

1. »clear, recognizing thoughts of the world and ourselves in it«
2. »to be aware of ones place in this world, of others, morals etc«
3. or, »Consciousness arises when higher reality (God, Tao, universal mind) gets in touch (resonates) with our emotions, thoughts and intuition. Consciousness itself feels alike staring (not much thoughts), like clarity, like recognition, like enlightenment.

Is there anything like animal consciousness?

Would it not be possible that when my dog looks smart and even wise that he is in a transient state, as if he meditated...? Starting by an example from animal world: when your dog looks at you with adoration, or observes a cat promenading itself in front of him, you are nearly sure he thinks: »oh, how happy I am that my master likes me (first case) and, oh,

how good-looking is this cat – I know I should chase her yet am a bit old and lazy (second case). He sometimes looks bored, the other time slightly sad (if you have been away travelling for a week and he did not believe he would see you again), he also can play and is amused when you cry, shout or jump – to make him jump. He is very responsive and is of course aware he is this and that particular dog, with his specific name. There are even smarter animals than dogs..

All these instantaneous emotions and thoughts can arise from experience and one would imagine that all processes allowing them are brain-based. But is that for sure ? Does that deep look of your dog not imply a »soul«, i.e., something beyond? Can we ask ourselves, where does consciousness: »to be aware of ones place in this world, of others, morals etc« – if one defines it that simple, come from? Some scientists prescribe all the higher processes of thought and awareness to the brain itself, which in a manner of electrical and chemical discharges, communicates between different parts: electrically – via electrical synapses and chemically – via synapses. But could it be that consciousness is a collective phenomenon, like reading patterns or catching radio waves (which need not be in the brain?), which become tuned and recognized (resembling quantum qualia ? [1])

What's new in physics and do we understand that?

I was surprised that so many new things are going on in material physics and nano-technology. I collected some papers and will just cite their titles here: *Experimental demonstration of quantum memory for light* [2], *Quantum Memory for Squeezed Light* [3] Would need a more practical explanation what does it mean: the “quantum memory of light”, mind. Then I read: *A novel molecular solution for ultraviolet light detection in Caenorhabditis elegans* [4]. The paper says that receptors have been found in C. Elegans, which transduce UV light into loco-motion. Does something like that exist in our own brains? Has a little worm remained from evolution? Very mysterious to general public are also the so called “quantum dots”. The newest papers on nano-technologically engineered quantum dots tell: *Optically Mapping the Electronic Structure of Coupled Quantum Dots* [5] and *Next-generation quantum dots* [6]. They even report: *The effect of quantum dots on synaptic transmission and plasticity in the hippocampal dentate gyrus area of anesthetized rats* [7]. I have an idea here – yet would withhold it for a moment.

From personal insights to experimental design

As we have eyes to record the outside forms and light, colours, so we may have some organ (in the middle of the brain – behind our eyes) which records the intuitive deeper lands: state of mind, peace, internal light, colours, warmth.

In calming one down by either shutting one's eyes (maybe shutting them strongly or even covering them by cooling hands) sitting and listening to silence one can sometimes see lightening, after which he may see colours – first all red, then, all orange, yellow, green and pale blue, finishing with purple. If he is paying attention, the electric discharges become synchronized and overwhelming extending one's own head. Very symmetrical and standing seems the field. Sometimes, one sees small lands of gardens and creatures (in rainbow colours), alike mandallas arising from the central spot, usually orange to purple light. One could argue that this is happening on the other side of our retina. OK, but what if that is coming from the land beyond this un-explored inner organ, which is harmonized to some vibrations from the Field surrounding all?

See some mandalas at:

<http://www.mandalanetdesign.com/website/gallery/art.html>

Scientists may do here some research about the notion of the (overwhelming) Field and energies of the thought – which – if existent - must be even faster than light...your suggestion). How to probe this? I would imagine this organ – in the middle of the brain would be of specific interest to explore. Might that be the hypothalamus or the pineal gland? Or, is there a »hole« into 4th dimension surrounded by receptors? Did anyone collect higher animals (mice for example) pineal glands and explore if there are special light receptors? Did anyone attach such separated glands to pulsating field and see if they emit light...?

Would that be possible to record something like functional magnetic resonance or fluorescence detected emitted light in the brains of a »meditating« mouse? Could colours, which would come from within – somewhere, not necessarily from the light emitting proteins, be detected? Have the so called quantum dots, which also exist naturally, made from protein-metal complexes, such as electron transport chain or photosynthesis system, any real function in the brains?

Figure 1.

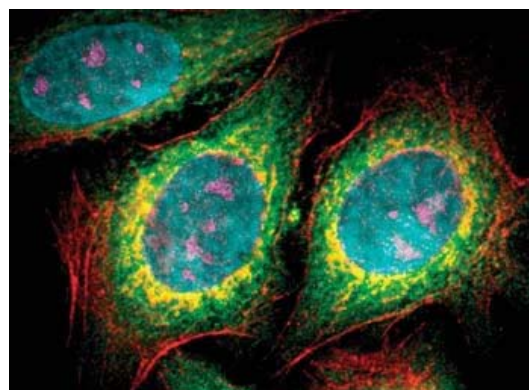


Fig.1 Quantum dots easily enter the cell. Image credit: *Quantum Dot Corp, CA*

What would count as »proof« of non-material, transpersonal basis of consciousness?

»And what?« you would likely comment. That would prove nothing. Just another physiological correlate to our psychological worlds. I would argue that physiology till now did not take into account energies (apart from EEG) but rather chemicals and flows. Such subtle energies are claimed to be recorded by some, who however, are not recognised as true scientists. How could physicists help here? Do they have new theories of particles and energies and ways to measure those, if anything emerges from a living (meditating) brain?...

What makes individual unique?

Genes, environment – which is substances one encounters through food, water, air - and social contacts. All 3 make a permanent trace in the systems biology: how many copies of the gene get expressed, how the proteins get post-translationally modified, how the synapses get potentiated or depressed... Thus, physical reality is for each Individual unique.

Is consciousness a collective phenomenon or is it also limited to one individual?

And, how can we explain some common underlying Unity, which is the basis of nearly all Religions or peak human experiences, such as art, music, extreme sports or ecstasies of any sort? These might indeed stem from (and return into) a well of collective consciousness. Similarly, more goodness and harmony in one's life, more goodness in town and in the whole world – which is experimentally proven of transcendental meditation. However, that does not imply a collective consciousness. Can be that harmonized individual consciousness acts synergistically with others.

Another argument for a common pool: some intelligent solutions often appear when the conditions are "ripe". They appear in many individuals at once, even if they do not communicate with each other. This would argue of some "evolution of common consciousness". Did you here of "indigo" children? People of the new, Aquarius era? Nothing has been proven, however, even some serious psychologists observe different traits of behaviour in nowadays youth: more resolution to follow one's own intellect and convictions. More confidence of natural "wisdom" – so to say.

Can one research Consciousness at all ?

If it is not material (neither matter or energy), if it exists just in the world of ideas, maybe one can use theoretical ways of thinking, mathematics, extrapolations by induction and phenomenological descriptions. Maybe what is known

as research one's own mind – is the way. First person research? [8].

Maybe one can extrapolate Consciousness to God and Love. Love and Consciousness might have in common that they both transcend one Individual. If you love One person you love all people (more so). That's why love (not war) means foundation for peaceful development ... Or, is that reasoning too simple?

Be conscious – what does that imply in practical terms?

Does that resemble being reasonable?

In the times of possible global crisis on the levels of financial, economic, ethical values and climate changes, what does consciousness bear ? We could safely assume, our dog is not worried, neither is some child who listens to the dark stories of the fate of this world as to fairy tales with rather bad outcome. But do grown-ups, intelligent and conscious people behave appropriately, rationally, to the best of their knowledge and capacity to solve the problem? Russell in his book [9] counted that we (as the civilization) have a few seconds left. There are two ways of dealing with that: either to do all what we think is wise and necessary to turn the destruction or, sit back, enjoy as long as we can, maybe rely on gods to come and save us? Even if it looks desperate, I believe humans still have a kind of sense of humour and can enjoy immediate on-going life. There already have been plays written, such as "Memory of Light" [10], which replay scenarios for the last judgement.

But to have "clear consciousness", could we not do more, globally, do we develop new ways of thinking, different scientific disciplines working hand in hand (interdisciplinary); do we learn from old civilizations regarding physical and mental health?

Are there initiatives of such Educational and Scientific institutions on-going?

For example at UNESCO there exists a site:
<http://www.eolss.net/Eolss-About.aspx>

There is something similar, oriented towards tourism more than general knowledge - supported by EU:
http://www.futourism.eu/top_menu/Home/tabid/1597/Default.aspx

Slovenia may be happy about the environment, which has been rather preserved: for how much longer?
http://www.slovenia.si/whats_new/200802/2008020614372609

T.C.Daffern, has founded the Global Green University (GGU) in UK and this has distinguished members elsewhere in the world:

<http://www.educationaid.net>
<http://slgandhi.blogspot.com/>

Having browsed the internet I could trace some more Green Universities:

<http://www.greenuniversity.net/>

<http://www.greenuniversity.net/Partners.htm>

<http://www.green.colostate.edu/>

<http://www.greencampus.harvard.edu/>

http://science.dal.ca/Files/ESS_proposal_Background_documents_.pdf

Will programs of the existing “Green universities” and special programs of normal high educational bodies, make a needed difference and lead to cut-through. Especially needed are new energy sources but also to understand needs of humans (research into consciousness, morals, values) to turn into less consumption and more happy society [11].

Does being Conscious mean being Ethical?

In Slovene language (as well as in English) to have clear consciousness and to act ethically mean more or less the same (Zavest – čista vest). To use reason we said would seem to help in solving the contemporary crisis. To use Ethical awareness – also. Should we raise that latter first? Should our educational systems involve more care to educate students in humanities (also Ethical answers) not strictly natural sciences? And if so, how should that be done?

Just to mention, that arts (and science) may be highly connected to Ethical awareness. I watched a TV show, an interview of our well known musician, a very talented flute player, I.G. Not only music but life itself is the field she succeeds to master and is searching to learn and apply high ethics to both.

Acknowledgement: Thanks for many inspiring discussions go to dr. Thomas C. Daffern.

References

- [1] Stuart Kauffman S. 2009. Physics and Five Problems in the Philosophy of Mind:
http://arxiv.org/PS_cache/arxiv/pdf/0907/0907.2494v1.pdf
- [2] Julsgaard B, Sherson J, Cirac I, Fiurášek J & Polzik ES (2004). Experimental demonstration of quantum memory for light *Nature* 432, 482-486 | doi:10.1038/nature03064
- [3] Appel J, Figueroa E, Korystov D, Lobino M, and Lvovsky AI (2008). Quantum Memory for Squeezed Light. *Phys. Rev. Lett.* 100, 093602.
- [4] Edwards SL, Charlie NK, Milfort MC, Brown BS, Gravlin CN, et al. (2008) A novel molecular solution for ultraviolet light detection in *Caenorhabditis elegans*. *PLoS Biol* 6(8): e198. doi:10.1371/journal.pbio.0060198
- [5] Optically Mapping the Electronic Structure of Coupled Quantum Dots. M. Scheibner et al. *Nature Physics* 4, 291 (2008).
- [6] Smith AM, Nie S (2009). Next-generation quantum dots *Nat Biotech*, Vol. 27, 732-733.
- [7] Mingliang Tang, Zhifeng Li, Liang Chen, Tairan Xing, Yong Hu, Bo Yang, Di-Yun Ruan, Fei Sun, Ming Wang (2009). The effect of quantum dots on synaptic transmission and plasticity in the hippocampal dentate gyrus area of anesthetized rats *Biomaterials* 30, 4948-4955
- [8] Kordeš U, How to research experience? pp. 191- 208, in “Philosophical Insights about Modern Science, Zerovnik, Markic, Ule (eds.), Scientific Revolutions Series, Uversky (Ser. Ed). Nova Sci publ. 2009, **ISBN:** 978-1-60741-373-8
- [9] Russel P. (2007). THE GLOBAL BRAIN The Awakening Earth for a New Millennium. **ISBN:** 978-086315-616-8
- [10] Daffern C.T. “Memory of Light” , 2008, : <http://www.lulu.com/content/paperback-book/the-memory-of-light/2211834>
- [11] Detela A, Ways to new sustainable energy sources, pp. 147 – 171, in “Philosophical Insights about Modern Science, Zerovnik, Markic, Ule (eds.), Scientific Revolutions Series, Uversky (Ser. Ed). Nova Sci publ. 2009, **ISBN:** 978-1-60741-373-8

K VPRAŠANJU O NACIONALNI SLIKI SVETA

Irina Shchukina

Kafedra ruskega jezika i stilistiki, Filologičeskij fakultet

Permskij gosudarstvennyj universitet

Bukireva, 15, 614990 Perm, Ruska federacija

Tel: +7 342 2 396 256

Fax: +(7) (342) 237 16 36

e-mail: yatik@mail.ru

IZVLEČEK

Narejen je poskus jasnega in razločnega razumevanja strukture (verbalne) slike sveta, ki je nujno potrebna jezikoslovcem aktivnim na področju kognitivne lingvistike. V članku je predlagana ena od možnosti predstavitve strukture slike sveta v kognitivni lingvistiki. V ruski lingvistiki obstoječi pristop k obravnavi structure slike sveta ko jo vsak avtor vidi po svoje in ne formulira eksplicitno svojega pogleda, ne dopušča možnosti analizepredlaganih zaključkov in trditev.

V zadnjih desetletjih se je zelo povečalo zanimanje za kognitivni vidik govorne dejavnosti, za formiranje vzorcev dejavnosti zavesti, za probleme odraza v jeziku slike sveta (SS) in postaja vse bolj popularen antropološki pristop k jezikoslovnim problemom. Raziskave, ki so bile izvedene v okviru kognitivne lingvistike imajo vse polno sodb o nacionalni sliki sveta (NSS), premislekov, ki so dejansko potrebni za analizo vseh teh procesov, ki jih avtorji analizirajo. Vendar pa se v procesu realizacije, kot bi se lahko zdelo, že davno proučenega pojava s strai filozofov, se znajdo v "košu" NSS stvari, ki sicer nimajo neposredne zveze s njo, ali morda bolje rečeno, "vse kar pride pod roko". Menimo, da je smiselno, da se še enkrat skoncentriramo na strukturo NSS, predvsem z namenom, da bi se izognili nenatančnosti in preveč ohlapni uporabi termina.

Idejo posebnega jezikovnega videnja sveta je sformuliral že v začetku 19. stol. W. von Humbolt: "Skozi mnogoterost jezikov...se odkriva bogastvo sveta in mnogoterosti tega, kar mi spoznavamo v svetu; in človekova bit postane (...) širša, ker nam jeziki v jasnih in dejanskih črtah dajejo različne načine mišljenja in zaznavanja" (Humbolt: 349).

V 20. stoletju raziskave SS nadaljujejo E. Sepir, B. Worf, L. Weisgeber in drugi. Prblem je blizu von Foersterju, L.N. Gumilevu, ki je vključil človeka v etnogezo, L. Wittgensteinu, ki je zamenjal tradicionalni problem odnosa biti in zavesti s problemom jezika in stvarnosti/realnosti. V okviru lingvistike se opisuje model sveta z naborom osnovnih semantičnih opozicij, ki imajo univerzalni značaj

za vsa ljudstva sveta (C. Levi – Straouss, Vjač. Vs. Ivanov, V.N. Toporov, A. Vežbickaya, V.V. Krasnyh in drugi).

Ne bomo dajali porobno karakteristiko v filozofiji sprejetim pojmom "znanstvena slika sveta", "naivna slika sveta" ali "vsakdanja slika sveta" in se bomo strinjali z mnenjem N.I. Beresneve in V.D. Beresneva, da "znanstvena" in "naivna" slika sveta nista izolirani, saj "Pomen znanosti je v življenju sodobnih družb ogromen. V zadnjih letih prodira znanost v vse sfere življenja in zato znanstvena slika sveta postaja z razvojem družbe vse bolj sprejemljiva za večino članov družbe in posledica tega je, da se tudi naivna slika sveta "poznanstvuje", menja se njena struktura." (Beresneva: 84).

Kot predstavnik živalskega sveta Zemlje, ima človek skupne z vsemi živalmi značilnosti. To so nadaljevanje vrste, samohranitveni gon in težnja k uživanju.

Konstruktivna pogojne structure slike sveta zahteva komparacijo in samo na ta način je možno razkriti posebnosti tega ali onega nivoja. Prvotni (osnovni) nivo ima splošnočloveške značilnosti in je torej potrebno odkriti razliko med človekovim sprejemanjem sveta in sprejemanjem sveta predstavnikov živalskega sveta. Človek je pokoril ogenj, zaveda se menjave dneva in noči, letnih časov, človek ima za razliko od živali sredstvo prenosa in ohranjanja informacije – jezik, ne glede nato ali je to najnovejša računalniška tehnologija ali pa ustno ljudsko izročilo avstralskih domorodcev. Človek za razliko od živali dela zavestno. To pomeni, da na prvem, baznem nivoju lahko zaznamujemo razlikovalne posebnosti SS, ki so značilne za vso človeško družbo, ne glede na stopnjo razvitosti tistega sociuma, kjer prebiva nosilec elementov te SS: torej zavesten odnos k naravi, zavesten odnos k delu, prisotnost sredstev ohranjanja in prenosa informacij.

Naslednji bazni nivo je tisti nivo, ki zakoliči razliko sprejemanja sveta človeka v odvisnosti od kraja prebivanja. In pri tem ni pomembno ali živi človek v mestu ali na vasi. Kot so pokazale zadnje raziskave, se prebivci velikih mest poslužujejo istih prijemov, kot so jih izdelali njihovi predniki, še predno so se pojavili megapolisi. Naravno, da bo človek, ki živi v gorah, človek, ki živi v neskončnih stepah, človek, ki živi v gozdnih predelih in človek, ki živi okoli morja, različno gledal na svet in ga na različne

načine sprejemal. Ta nivo pogojno poimenujmo "geografski", kjer se ne upošteva samo relief, ampak tudi klimatski pasovi in podnebni pogoji v katerih živi človek. "Naj bo še tako razvta tehnika, vse potrebno za ohranjanje življenja ljudje dobijo v naravi. To pomeni, da so v trofični verigi vrhnji, zaključni člen biocenoze tistega področja, ki ga naseljujejo. Kakor hitro pa je tako, predstavljajo element strukturno-sistemske celote, ki vključuje v sebi poleg ljudi še domesticated (to je domače živali in kulturne rastline); pokrajine, tako te, ki jih je človek spremenil, kakor tudi nedotaknjene pokrajine; bogastva pod zemljo; vzajemne odnose s sosedi, ki so lahko prijateljski ali pa sovražni; takšno ali drugačno dinamiko socialnega razvitja; prav tako pa tudi to ali drugo kombinacijo jezikov (od enega do mnogih) in elemente materialne in duhovne kulture." (Gumilev: 30) To pomeni da igra zelo pomembno vlogo v človekovem videnju sveta kraj njegovega odraščanja, ko se formirajo njegovi osnovni, tradicionalni pogledi na svet, poznejša sprememba kraja prebivanja lahko naredi SS širšo in bogatejšo, vendar jo, praviloma, ne spremeni.

Delo Maxa Weberja "Kapitalizem in protestantska etika" je iniciirala začetek civilizacijskega pristopa k proučevanju družbe, njenega razvoja, njene zgodovine. Razvijajoč se dogo časa izolirano druga od druge je vsaka civilizacija spoznavala in osvojila svet po svoje. O kitajski iznajdbi smodnika in papirja so danes poučeni vsi šolarji, vendar pa je, na primer, daleč od tega, da bi izrek v geometriji, da je vsota ploščin kvadratov nad katetama enaka ploščini kvadrata nad hipotenuzo, ki ima v evropski (šolski) zavesti stabilno asociacijo z imenom Pitagora, da bi imela v zavesti Kitajca isto asociativno zvezo, zdi se, kot da so Kitajski matematiki pršli do tega izreka s "svojimi salami". Predstavniki raznih civilizacij prav tako na poseben način gledajo na svet, čeprav včasih pogledi ljudstev, ki pripadajo različnim civilizacijam, so si v nekaterih parametrih blizu, naprimer, razredne zavesti v sliki sveta Rusov, Kitajcev, Kubancev in Kolumbijcev mnogem sovpadajo, čeprav to ni le davek ideologiji, takšne slike sveta, kot je že bilo rečeno, imajo dostikrat le začasen značaj. Tako torej, od tega h kateri civilizaciji se prišteva ljudstvo, v katerem se formira človekov pogled na svet, zavisi tudi eden od vidikov njegove slike sveta. Poimenujmo ta nivo "civilizacijski".

Prav gotovo je ena od najbolj pomembnih komponent SS religiozni svetovni nazor. Rzlike niso samo v hramih (cerkve, mošeje, sinagoge, pagode itd.) ampak je vsak korak v vsakdanjem življenju človeka pogojen s tistimi pravili, ki mu jih predpisuje njegova religija. Tako je za pogane Hante in Manse (Zahodna Sibirija) rojevajoča ženska – umazana žival. In celo dandanes, ko jo ne izženejo v hladni "čum" (rusko ime za mobilno prebivališče severnih ljudstev) v času rojevanja, ampak jo odpeljejo v lokalno porodnišico, je odnos do nje odgovarjajoč tradiciji. Celotno dandanes ortodoksni Judje na severu Permskega rajona vzamejo svoje šolarke po devetem razredu iz šole, da bi jih omožili. Ne moremo si predstavljati prestižno likovno galerijo v Evropi ali

Ameriki, brez slik bibličnih oseb, tu so svetniki, in Noe s svojimi sinovi, in Judita z glavo obglavljenega Holoferna, a za muslimane je naslikati človeka nekaj prepovedanega. Najbolj grobe ruske ali srbske kletvice vključujejo preklinjajne Boga, a za muslimane je že karikatura Preroka zadosten razlog za smrtno obsodbo. Kristjani, muslimani in Judje so nastopili proti kloniranju človeka, odnos budistov pa je do kloniranja umirjen, saj vendar reinkarnacija lahko predstavlja za dušo mnogo strašnejšo usodo v prihodnosti kot je kloniranje. Torej, religija je ena od zelo važnih komponent formiranja človekove slike sveta. Poimenujmo ta nivo "religiozni svetovni nazor".

V zadnjem desetletju so raziskave s področja lingvistike spolov postale skoraj glavna tema ruskega jezikoslovja. Proučevanje spolov kot sociokulturnega konstrukta, ki določa vlogo in pomen moškega in ženske v družbi, neizogibno pripelje k problemu stereotipnega zaznavanja teh vlog in njim pripisanih kvalit. Obstoječi spolni stereotipi začenjajo delovati na fantke in dklice že v samem zgodnjem otroštvu. Praviloma ti stereotipi podpirajo formiranje adekvatnih spolnih reakcij pri tradicionalni "moški" in "ženski" vzgoji. Spolni stereotipi "nastopajo kot kognitivne sheme, ki upravljajao s procesom predelave informacij, ki jih sprejemamo in to na tak način, da jo začenjamo sprejemati, pomniti in interpretirati v skladu z našimi predstavami o spolu. Stereotip je pomembna komponenta človekove zavesti. S pomočjo stereotipov se hitro in z gotovostjo orientiramo v vsakdanjem življenju, v mnogoterosti faktov in dogodkov, pojavov in procesov stvarnosti" (T.N. Rezvuškina: 56) To pomeni, da družba zavestno sodeluje v podpiranju stereotipov, da bi omejevala in kanalizirala obnašanje svojih članov. Ta nivo tako pogojene slike sveta poimenujmo "spolni nivo".

Pri sociolingvističnih raziskavah se skupine, ki prdstavljajo vzorec, praviloma formirajo v obliki opozicijskih strat: izobrazba – visoka/srednja, poklicni profil - filolog /nefilolog ali humanistična izobrazba /ne-humanistična. "V bistvu je to (...) proučevanje socialne structure družbe vendar k standardnim parametrom proučevanja socialne stukture (socialni status, izobrazba, dohodek, preživljanje prostega časa, politčne preference itd.) je dodan še parameter jezikovne diferenciacije: ljudje s končano sednjo šolo običajno govorijo tako-le, z visolo izobrazbo pa tako-le in tisti, ki imajo humaistično izobrazbo pa tako-le" (Mečkovskaya: 55) . Poimenujmo nivo, ki diferencira pogled na svet glede na "izobrazbo" kar "izobrazbeni".

Vrh pogojne structure SS je individualno videnje. Ne smemo pozabiti, da so pri kateremkoli individualnem pristopu predhodni nivoji že opravili svoje delovanje in so formirali bazično obliko sveta v zavesti individuuma. Praviloma, po eni strani, prav predhodni nivoji v celoti formirajo nacionalno sliko sveta. Po drugi strani pa so bazični nivoji osnova za oblikovnje individualne slike sveta (ISS). Ko govorimo o ISS, moramo upoštevati fizične lastnosti (ali je otrok rasel kot močan ali šibak, bolan ali zdrav, celo - nositi očala ali jih ne nositi ima tu svoj pomen), njegov temperament, v kakšni družini je živel

(popolna ali nepopolna, ali je imel stare starše ali jih ni imel), kakšne nadarjenosti je prinesel na svet in v kolikšni meri jih je mogel realizirati. Ta nivo vključuje ves spekter odnosa k svetu in predstavlja osnovno razliko med ljudmi, med posamezniki. Poimenujmo ta nivo "individualna slika sveta".

Shematsko lahko predstavimo strukturo ISS s tole shemo:



nadaljevanje vrste// zakoni preživetja// stremljenje k uživanju

Shematski prikaz ISS

Literatura

- Бердяев, Н. Евразийцы//Николай Трубецкой, М., 2007.
- Береснева Н.И., Береснев В.Д. Репрезентация философских знаний в картине мира студентов// Вестник Пермского университета. Философия. Вып. 5 (31). 2009.
- Булгаков, С.Н. Война и русское самосознание. Форум. Международный журнал. ? 5. М., 1993. Вебер М. Капитализм и протестантская этика// Макс Вебер Избранные труды. М., 1996.
- Вежицкая А. Язык. Культура. Познание. Ю М., 1996.
- Воробьев В.В. Лингвокультурология. М., РУДН, 2008.
- Гумбольдт ф.В., Язык и философия культуры., М., 1985.
- Гумилев Л., Этногенез и биосфера Земли, М., 2002.
- Гуревич А.Я. Категории средневековой культуры. М., 1972.
- Мечковская Н.Б. Социальная лингвистика. М., 2000.
- Петренко В.Ф. Основы психосемантики. Смоленск, 1997.
- Ратбиль Т.Б. Современные проблемы филологии. Когнитивные исследования. Язык как деятельность. Н.Новгород, 2006.
- Руднев В.П. Словарь культуры XX века. М., 1997.
- Сепир. Э. Избранные труды по языкознанию и культурологии. М., 1993.
- Трубецкой Н., К проблеме русского самопознания// Николай Трубецкой, М., 2007.

Кнап Ж.Р. Щукина И.Н. Болонский процесс и русская национальная картина мира// Слово. Словарь. Словесность. (К 225-летию основания Российской Академии). Материалы Всероссийской научной конференции Санкт-Петербург, РГПУ им. А.Герцена, 12-14 ноября 2008 глда. СПб, 2009.

Щукина И.Н. "И придут ко мне все трудящиеся..." (к проблеме интерпретации переводных религиозных текстов)//Концептосфера-дискурс-картина мира. Международный сборник научных трудов по лингвокультурологии. Самара, 2006.

Резвушкина Т. Использование метода семантического дифференциала при изучении гендерных стереотипов//Гендерные исследования в Центральной Ази. Алматы, 2002.

RAZMERJE DO VODENJA IN IZBRANE SOCIALNO VREDNOTNE ORIENTACIJE PRI RAZLIČNIH KULTURAH V SLOVENIJI

Dr. SLADANA MIHAJLOVIĆ
Partizanska 10 A
SI – 1000 LJUBLJANA, SLOVENIA
FILOZOFSKA FAKULTETA LJUBLJANA
Te.: +386 41 455 000
e-mail: sladjana.suzana@siol.net

POVZETEK

Socialne in socialno vrednotne orientacije v zvezi z vodenjem, samoznavo in življenjskim stilom so v vsakem kulturološkem okolju ene najpomembnejših socialnih orientacij.

V pričujočem prispevku me je zanimalo, ali se omenjene socialne reprezentacije pomembno ločijo med skupinami slovenskih in neslovenskih udeležencev oz. kultur v raziskavi. Naše udeležence je predstavljalo šest različnih skupin študentov (ekonomisti, strojniki, študentje managementa, pravniki, sociologi in psihologi), ki sem jih zajela v raziskavi. Vsega skupaj 269 študentov, od tega 246 Slovencev in 23 Neslovencev.

V vprašalniku, ki sem ga konstruirala skupaj s sodelavci, sem uporabila Rokeachovo lestvico vrednot, Hofstedejevo lestvico vodenja, Lernerjev vprašalnik o pravičnosti sveta ter lestvice v zvezi z evalvacijami vodenja in vodstvenih izkušenj.

Rezultati so pokazali, da se slovenski in neslovenski udeleženci ne ločijo pomembno v nobeni sumacijski odvisni spremenljivki, da pa se ločijo pomembno v preferencah vrednot Svet lepote in estetike, Enakost vseh ljudi, Družinska varnost, Vljudnost in Intelktualnost.

Ključne besede: vrednotne orientacije, vodenje, vodstveni oz. managerski in življenjski stil, zaznava pravičnosti, nacionalna pripadnost oz. kultura narodov, Slovenci – Neslovinci

ABSTRACT

Social and social value orientations related to management, self – perception and life style by chosen groups of Slovene and non – Slovene participants

Social and social value orientations related to management, self-awareness and life style are among the most important social orientations in any cultural environment.

In the research, I was interested if the mentioned social interpretations differ considerably among groups of Slovenian and non-Slovenian participants or cultures of the survey. Respondents have been consisted of six different

groups of students (economy students, mechanical engineering students, management students, law students, sociology students and psychology students), which have been included in the survey. A total of 269 students have been questioned, of which 246 have been Slovenians and 23 non-Slovenians.

In the questionnaire I compiled together with my colleagues, I applied the Rokeach Value Scale, Hofstede's Management Scale, Lerner's questionnaire on the justness of the world and scales related to evaluations of management and management experience.

Results have shown that responses from Slovenian and non-Slovenian respondents do not differ considerably in any summation dependable variable, while they differ considerably in the value preferences World of beauty and aesthetics, Equality of all people, Family security, Politeness and Intelctuality.

Key words: value orientations, management, management/managerial life style, justice awareness, nationality or culture of the nations, Slovenians – non-Slovenians

1 UVOD

Aktualni prispevek ima predvsem namen predstaviti enega od zanimivih aspektov empiričnega pristopa k raziskovanju izbranega problema. Obsežna in izčrpna elaboracija teoretičnega ozadja, ki jo je avtorica sicer že izpeljala, pa je v pričujočem članku skraćena na vsebinski in stilistični prikaz, ki naj bi poudaril tradicijo in pomembnost tem tudi za tiste, ki se z omenjeno problematiko sicer raziskovalno ne ukvarjajo, je pa le ta očitno ves čas prisoten vidik »psihologije in sociologije njihovega vsakdanjega življenja«.

Vodenje je večplastni pojem, zato se tudi definicije številčne ter se med seboj razlikujejo glede na aspekt, ki ga izpostavljajo. Sadler [1] je na podlagi analize večih definicij, določil glavne komponente vodenja. Tako naj bi bilo vodenje aktivnost ali proces. Ta proces obsega pojave kot so vpliv, zgledno vedenje¹ ali prepričevanje. Vključuje

¹ "Exemplary behaviour".

tudi udeležence, in sicer voditelje in privrženca. Kot zadnjo komponento avtor navaja različne izide procesa – najbolj očitni so doseženi cilji, pa tudi predanost posameznikov tem ciljem, povečanje skupinske kohezivnosti in okrepitev ali sprememba organizacijske kulture. Kulturo sistema Hofstede [2] definira kot psihološko imetje organizacije in kot taka predstavlja kolektivno programiranje mišljenj in delovanja a) družbene kulture (predvsem vrednote, socializacija) in organizacijske kulture (predvsem praktično delovanje). Vrednote pa so bile tudi predmet njegovega preučevanja in prav tako del moje raziskave, nanašajoč se na izbiro idealne službe. V zadnjih desetletjih so postale vrednote imeniten predmet medkulturnega preučevanja in so izraz samega jedra človeštva in človeških kultur. Hofstedejeva medkulturna raziskava, je pokazala, da so določilne dimenzije značilnosti različnih kultur močno povezane z vrednotnimi sistemi in dimenzijami. Vrednote lahko definiramo kot življenjska vodila in kot takšna pomenijo zgostitev vsega tistega, kar neka kultura ceni, kar jo normativno usmerja in motivira. Vrednote so nekakšen povzetek notranjih norm in ciljev neke družbe in kulture. Poleg tega pa so vrednote v primerjavi z drugimi psihosocialnimi prvinami neke kulture relativno stabilne in dolgotrajne. Kulturo opredeljujejo trajnostno in stabilno, kar pomeni, da so lahko njen najboljši identifikacijski znak. Če se pomembno spremenijo vrednotne usmeritve, se pomembno spremeni kultura in obratno.

Vrednotni sistem pa je trajnejša organizacija prepričan, ki se nanašajo na prednostne načine vedenja ali na eksistenčna stanja na kontinuumu pomembnosti. Trajnost vrednotnega sistema ni absolutna. Je toliko stabilnejši, kolikor zagotavlja kontinuiteto edinstvenih dimenzij osebnosti in različnih socializacijskih učinkov. Hkrati pa je v tolikšni meri nestabilen, da omogoča preurejanje vrednot glede na njihovo pomembnost kot izraz sprememb v kulturi, družbi ali na podlagi posameznikovih izkušenj [3]. Krech, Crutchfield in Ballachey [4], definirajo vrednote kot posebno vrsto prepričanja, ki se nanaša na to, kateri objekti so dobri in kateri so nezaželeni in slabi. Hkrati s tem poudarjajo tudi družbeno in kulturno pogojenost vrednot.

Kane [5], ki v svojem opisu izpostavlja moralno dimenzijo vodenja, poudarja neprisilno, vzajemno naravo odnosa med voditelji in privrženca oz. volivci (če uporabimo manj obremenjen oz. bolj korekten pojem)².

Bass [7] je vodenje predstavil kot fokus oz. žarišče skupinskih procesov, stvar osebnosti, sprožanje voljnosti oz. sporazumevanja, vajo vplivanja, posebno obnašanje, obliko prepričevanja, povezavo moči, instrument za doseganje ciljev, učinek interakcije, razločevalno vlogo,

začetek strukture in kot še veliko kombinacij omenjenih definicij.

Vodenje lahko vključuje uporabo moči, ne more se zreducirati zgolj na uporabo moči, saj se odločilno opira na prepričevanje. Čeprav politični voditelji uradno zavzemajo vodilni položaj, ne morejo dati zapovedi, saj volivci niso podrejeni.

Nasprotno je Weber [8] v preteklosti omenil, da se v politiki kot pogosto sredstvo uporablja prisila. V številnih primerih je doseganje "dobrih" ciljev vezano na to, da je treba vzeti v zakup moralno dvomljiva ali vsaj nevarna sredstva in možnosti ali tudi verjetnost neprijetnih stranskih posledic. Kane [5] dalje meni, da so voditelji neozogibno tudi simboli, saj vrhovni voditelj skupnosti ali naroda simbolizira kolektivno identiteto skupine in kontinuiteto. Vodenje naj bi se potemtakem na splošno razlikovalo od managementa³; delno na račun svoje simbolične vloge in delno na podlagi dejstva, da so voditelji manj tesno povezani z organizacijo kot managerji – in seveda, nekateri voditelji sploh niso vezani na nobeno organizacijo.

O skupinah, ki jim pripadamo, si oblikujemo različne predstave, misli, spoznanja, čustva in občutja. Celoto vseh teh prvin, ki jih doživljamo v odnosu do neke skupine, imenujemo skupinska zavest. S skupinsko zavestjo se povezuje občutje skupinske identitete. To zajema celoto predstav o skupini, s katero se posameznik identificira in po kateri se omenjena skupina razlikuje od drugih skupin istega razreda. Analogno kot o nacionalni zavesti lahko govorimo tudi o nacionalni identiteti.

Tradicija ter skupna in povezana usoda omogočajo, da se pojavijo podobne kognitivne in emocionalne vsebine nacionalne identitete pri mnogih pripadnikih narodnostne skupnosti oz. kulture. Tako lahko govorimo o pravi narodnostni samopodobi.

2 OPEDELITEV PROBLEMA

Problem lahko v splošnem izrazim z vprašanjem, ali med skupinama oz. kulturama slovenskih in ne – slovenskih udeležencev v aktualni raziskavi obstajajo pomembne razlike ($p \leq 0.05$) v izbranih odvisnih spremenljivkah, ki predstavljajo izbrane socialno – vrednotne orientacije oz. socialne reprezentacije v zvezi s problematiko vodenja, družine in samega sebe.

3 HIPOTEZA

³ Avtor uporablja izraz "leadership", ki ga prevajamo kot vodenje in "management", ki bi ga lahko prevedli kot poslovno vodenje. Možina [9] pri razločevanju obeh oblik vodenja trdi, da je glavni poudarek pri procesu vodenja kot managementa v doseganju organizacijskih ciljev, glavni poudarek pri procesu vodenja kot leadershipa pa je v vplivanju na ljudi, da bi dosegli cilje, pri čemer ni nujno, da gre za cilje organizacije. Vodenje (leadership) naj bi bil eden vitalnih vidikov procesa vodenja (management), ki vključuje tudi funkcije kot so planiranje, organiziranje, razdeljevanje in pogajanje (Hersey in Blanchard, 1982; po Možini [9]).

² Konotacije besede – privrženec – (follower) preveč napeljujejo na pasivnost in odvisnost, kar ne more veljati za vse, ki so v dvosmerni izmenjavi z voditeljem. Zato je morda bolj primerna uporaba besede – »volivec« (constituent), čeprav v določenih kontekstih lahko deluje okorno – Gardner [6].

H1:

Ne pričakujem pomembnih razlik med skupinama v preferencah a/ terminalnih in b/ instrumentalnih vrednot (pomembnost razlik sem preverjala z Mann-Whitneyevim neparametričnim testom neodvisnih spremenljivk).

H2:

Alternativna hipoteza znotraj od vsakih podskupin "Slovenci – Neslovenci". Pričakujem visoke in pozitivne korelacije med odvisnimi spremenljivkami: a/ samoevalvacija, b/ vodstveni stil, c/ življenjski stil, d/ klima v družini, e/ zaznava pravičnosti.

4 METODA

A) Udeleženske in udeleženci raziskave

V raziskavi je sodelovalo 269 študentov od tega 94 študentov in 173 študentk s povprečno starostjo 22 let in povprečnim odklonom 1.55 let (bilo je $n = 246$ Slovencev (od tega 87 moških in 159 žensk, povprečna starost = 22.29 let (SD = 1.89)) in 23 Neslovencev (od tega 7 moških in 16 žensk, povprečna starost = 21.52 let (SD = 1.12)).

V raziskavo so bili vključeni študentje tretjih letnikov Ekonomske fakultete v Ljubljani, Strojne fakultete v Ljubljani, Pravne fakultete v Ljubljani, Filozofske fakultete v Ljubljani in Fakultete za management v šolskem letu 2006 / 2007. Raziskava je bila izvedena marca 2007.

B) Pripomočki oz. instrumenti

Za zbiranje podatkov je bil uporabljen večinoma originalen konstruiran vprašalnik, ki je med drugim zajel Rokeachovo lestvico vrednot, vprašalnik o pravičnosti sveta (priredba Dalbert in Montana), Hofstede-jev vprašalnik vrednot, različne lestvice stališč do vodenja.

Vsi instrumenti, ki so bili prevedeni iz angleščine, so bili za uporabo v slovenščini prilagojeni z metodo "naprej – nazaj". Avtorja sumacijskih lestvic (do vodenja, zaznav zadovoljenosti lastnih psihosocialnih potreb, socialne klime v lastni družini, življenjskega stila sta V. S. Rus in M. Rus – Makovec).

A/ Pomembnost in zaznava uresničenosti osemnajstih terminalnih vrednot po Rokeach-u

Spremenljivke od A 68 do A 85 pomenijo pomembnost oziroma range terminalnih vrednot (od 1 do 18), spremenljivke od A 86 do A 103 predstavljajo zaznavo uresničenosti terminalnih vrednot (od 1, ki pomeni "prav nič ni uresničena" do 5, ki pomeni "v celoti je uresničena");

B/ Pomembnost in zaznava uresničenosti osemnajstih instrumentalnih vrednot po Rokeach-u

Spremenljivke od A 104 do A 121 pomenijo pomembnost oziroma range instrumentalnih vrednot (od 1 do 18), spremenljivke od A 122 do A 139 predstavljajo zaznavo uresničenosti instrumentalnih vrednot (od 1, ki pomeni "prav nič ni uresničena" do 5, ki pomeni "v celoti je uresničena");

5 REZULTATI

Tabela 1: Rezultati Mann-Whitneyevega testa razlik v rangih terminalnih vrednot [3] glede na kategorijo "Slovenci – Neslovenci"

Variable	Skupina	N	Srednji rang	Mann-Whitney	
				U	P
V068	1	246	132.77	2281.50	0.12
Vznemirlj. Življenje	2	23	158.80		
V069	1	246	135.56	2691.00	0.69
Udobno življenje	2	23	129.00		
V070	1	246	134.36	2671.00	0.66
Izpolnjen smisel ž.	2	23	141.87		
V071	1	246	135.36	2739.50	0.80
Mir na svetu	2	23	131.11		
V072	1	246	132.42	2193.50	0.07
Svet lepote in estet.	2	23	162.63		
V073	1	246	138.62	1938.00	0.01
Enakost vseh ljudi	2	23	96.26		
V074	1	246	138.07	2073.00	0.03
Družinska varnost	2	23	102.13		
V075	1	246	132.99	2334.50	0.16
Svoboda	2	23	156.50		
V076	1	246	132.97	2330.00	0.16
Sreča	2	23	156.70		
V077	1	246	135.23	2773.00	0.86
Notranja harmonija	2	23	132.57		
V078	1	246	132.43	2196.00	0.08
Zrela ljubezen	2	23	162.52		
V079	1	246	133.77	2526.00	0.39
Nacional. Suveren.	2	23	148.17		
V080	1	246	135.08	2808.50	0.96
Užitki	2	23	134.11		
V081	1	246	137.66	2175.00	0.05
Vera	2	23	106.57		
V082	1	246	134.80	2779.50	0.89
Samospošovanje	2	23	137.15		
V083	1	246	137.53	2206.50	0.08
Resnično prijatelj.	2	23	107.93		
V084	1	246	134.94	2815.00	0.97
Socialno priznanje	2	23	135.61		
V085	1	246	133.84	2544.00	0.42
Modrost	2	23	147.39		

Opombe: Skupina: 1 = Slovenci; 2 = Neslovenci; variable so označene z zaporedno številko v vprašalniku

V zvezi s preverjanjem H1 smo pomembne razlike med skupinama Slovencev in Neslovencev smo potrdili samo v primeru vrednot: svet lepote in estetike, enakost vseh ljudi, družinska varnost in vera (tabela 1).

Tabela 2: Rezultati Mann-Whitneyevega testa razlik v rangih instrumentalnih vrednot [3] glede na kategorijo "Slovenci – Neslovenci"

Variable	Skupina	N	Srednji rang	Mann-Whitney U	P
V104	1	246	135.19	2781.50	0.89
Prizadevnost	2	23	132.93		
V105	1	246	135.32	2750.50	0.83
Širokosrčnost	2	23	131.59		
V106	1	246	133.49	2457.00	0.27
Sposobnost	2	23	151.17		
V107	1	246	136.24	2525.00	0.39
Vedrost	2	23	121.78		
V108	1	246	136.16	2542.50	0.42
Urejenost	2	23	122.54		
V109	1	246	135.49	2708.00	0.73
Hrabrost	2	23	129.74		
V110	1	246	134.87	2797.50	0.93
Odpuščanje	2	23	136.37		
V111	1	246	135.62	2675.50	0.66
Uslužnost	2	23	128.33		
V112	1	246	135.79	2634.50	0.58
Iskrenost	2	23	126.54		
V113	1	246	134.10	2607.50	0.53
Iznajdljivost	2	23	144.63		
V114	1	246	134.30	2658.00	0.63
Neodvisnost	2	23	142.43		
V115	1	246	131.28	1914.00	0.01
Intelektualnost	2	23	174.78		
V116	1	246	133.79	2530.50	0.40
Ljubeznjiv	2	23	147.98		
V117	1	246	134.04	2593.00	0.51
Logičnost	2	23	145.26		
V118	1	246	135.97	2589.50	0.49
Ubogljivost	2	23	124.59		
V119	1	246	135.16	2790.50	0.91
Odgovornost	2	23	133.33		
V120	1	246	137.85	2127.50	0.05
Vljudnost	2	23	104.50		
V121	1	246	135.84	2622.50	0.56
Obvladovanje	2	23	126.02		

Opombe: Skupina: 1 = Slovenci; 2 = Neslovenci; variable so označene z zaporedno številko v vprašalniku

Pri verifikaciji H1 za preference instrumentalnih vrednot smo ničelno hipotezo zavrnili pri vrednotah intelektualnost in vljudnost (tabela 2).

Tabela 3: Korelacije med obravnavanimi variablami pri skupini "Slovenci" (n = 246)

Skupine	KD	Vod	VodMan	ZS	Prav
KD					
Vod	0.31**				
VodMan	-0.04	0.08			
ZS	0.26**	0.65**	0.15*		
Prav	0.07	0.22**	0.11	0.19**	

Opombe: 1) KD = zaznana klima v družini; Vod = evaluacija lastnega (stila) vodenja; VodMan = zaznava vodstvenega stila v slovenskem gospodarstvu; ZS = evaluacija lastnega življenjskega stila; Prav = zaznava »pravičnosti v svetu« – vsi sumarni skori

*, $p < 0.05$; **, $p < 0.01$

Pri preverjanju H2 smo odkrili pomembne ($p < 0.01$) visoke pozitivne korelacije v skupini Slovencev v primeru odvisnih spremenljivk, in sicer med evaluacijo vodenja in življenjskim stilom (tabela 3).

Tabela 4: Korelacije med obravnavanimi variablami pri skupini "Neslovenci" (n = 23)

Skupine	KD	Vod	VodMan	ZS	Prav
KD					
Vod	0.59**				
VodMan	0.05	-0.19			
ZS	0.46*	0.75**	-0.22		
Prav	-0.03	0.03	0.29	0.23	

Opombe: 1) KD = zaznana klima v družini; Vod = evaluacija lastnega (stila) vodenja; VodMan = zaznava vodstvenega stila v slovenskem gospodarstvu; ZS = evaluacija lastnega življenjskega stila; Prav = zaznava »pravičnosti v svetu« – vsi sumarni skori.

*, $p < 0.05$; **, $p < 0.01$

V zvezi s preverjanjem H2 smo odkrili pomembne ($p < 0.01$) visoke pozitivne korelacije v skupini Neslovencev v odvisnih spremenljivkah med evaluacijo vodenja in življenjskim stilom ter prav tako med evaluacijo vodenja in klimo v družini (tabela 4).

6 DISKUSIJA

V primeru nobene od posamičnih odvisnih spremenljivk (zaznana klima v družini, evaluacija lastnega (stila) vodenja, evaluacija/ zaznava vodstvenega stila v slovenskem gospodarstvu, evaluacija lastnega življenjskega stila ter zaznava »pravičnosti v svetu«), izraženih kot socialno – vrednotne orientacije oz. kot socialne reprezentacije in merjene kot sumacijske variable, nisem odkrila pomembnih razlik med skupinama udeležencev oz. kultur »Slovenci« vs. »Neslovenci«. Gre torej za izrazite podobnosti v nekaterih temeljnih socialnih reprezentacijah in self konceptih oz. samoevaluacijah med skupinama slovenskih in neslovenskih udeležencev v raziskavi. Iz omenjene podobnosti lahko z relativno visoko stopnjo verjetnosti sklepamo, da je bila vsaka od obeh skupin izpostavljena podobnim socializacijskim oz. akulturacijskim učinkom s strani slovenskega okolja.

Omenjene socialne/ akulturacijske učinke lahko opazimo tudi na nivoju podobnosti v preferenčnih zaznavah terminalnih oz. instrumentalnih vrednot Rokeacheve lestvice vrednot. Pomembne razlike v preferencah terminalnih vrednot sem odkrila samo v primeru treh vrednot (Svet lepote in estetike, Enakost vseh ljudi, Družinska varnost), v zvezi z instrumentalnimi vrednotami pa v primeru dveh vrednot (Intelektualnost in Vljudnost). Hkrati menim, da ravno odkrite pomembne razlike med skupinami v rangih vrednot Enakost vseh ljudi in Družinska varnost (tabela 1) pa vendarle kaže tudi na nekatere manjše razlike v socializacijskih učinkih med skupinama slovenskih in neslovenskih udeležencev. Zanimivo je, da je skupina neslovenskih udeležencev v raziskavi pomembno bolj pozitivno preferira vrednoto Intelektualnost, kot pa jo preferira skupina slovenskih udeležencev, kar morda kaže na to, da neslovenski udeleženci izobrazbo oz. intelektualnost zaznavajo kot pomemben faktor, ki je povezan s procesi socialne promocije.

V zvezi s korelacijskimi razmerji med obravnavanimi socialnimi reprezentacijami v skupini slovenskih udeležencev izstopata dve pomembni in pozitivni korelaciji (tabela 3), in sicer med evaluacijo vodenja in življenjskim stilom oz. med zaznano klimo v lastni družini ter evaluacijo lastnega stila vodenja. V zvezi s korelacijskimi razmerji med obravnavanimi socialnimi reprezentacijami v skupini neslovenskih udeležencev pa izstopajo tri pomembne in pozitivne korelacije (tabela 4), in sicer med zaznano klimo v lastni družini na eni strani ter evaluacijo lastnega stila vodenja oz. lastnega življenjskega stila na drugi strani.

Dobljene rezultate lahko strnemo v ugotovitve, da sta slovenska in neslovenska skupina udeležencev oz. kultur v raziskavi zelo podobni v vrednotnih preferencah.

Literatura

- [1] **Sadler, P.** Leadership. London: Kogan Page Limited. 1997.
- [2] **Hofstede.** Culture's consequences: International differences in work-related values. Beverly Hills CA: Sage Publications. 1980.
- [3] **Rokeach, M.** The nature of human values. New York: The Free Press. 1973.
- [4] **Musek, J.** Osebnost. Ljubljana: Dopisna delavska univerza Univerzum. 1982.
- [5] **Kane, J.** The politics of moral Capital Cambridge: Cambridge University Press. 2001.
- [6] **Gardner, H.** Leading Minds (London: Harper Collins). 1990;1996.
- [7] **Bass, B. M.** Handbook of leadership (theory, research, and managerial applications), Third edition. The Free Press - Collier Macmillan Publishers. London. 1990.
- [8] **Weber, M.** Politika kot poklic. v F. Adam (Ur.), Politika kot poklic. Ljubljana: Krt, 21 – 66. 1992.
- [9] **Možina, S.** Osnove vodenja. Ljubljana: Ekonomska fakulteta. 1994.

Zbornik 12. mednarodne multikonference
INFORMACIJSKA DRUŽBA – IS 2009

Proceedings of the 12th International Multiconference
INFORMATION SOCIETY – IS 2009

Robotika

Robotics

Uredili / Edited by

Jadran Lenarčič, Leon Žlajpah, Andrej Gams

<http://is.ijs.si>

16. oktober 2009 / October 16th, 2009
Ljubljana, Slovenia

PREDGOVOR

Letos prvič pod okriljem multikonference »Informacijska družba« organiziramo tudi konferenco Robotika. V Sloveniji ima raziskovalno področje robotike dolgo tradicijo in uveljavilo se je več raziskovalnih skupin, ki ne samo da sledijo trendom sodobne robotike, temveč jih so-ustvarjajo.

Robotika se je v zadnjih desetletjih utrdila v številnih industrijskih procesih kot nepogrešljiv del moderne, ekonomične in človeku prijazne tehnologije. Brez robotskih manipulatorjev si ne znamo več predstavljati varjenja avtomobilskih ohišij, vstavljanja obdelovancev v stiskalnice, razpršilnega barvanja. Ne presenečajo nas kirurški roboti ali servisni mobilni roboti, ki bodo kmalu čistili naša stanovanja. Vse večje število kvalitetnih mednarodnih robotskih strokovnih revij pa dokazuje, da postaja robotika priznana tudi kot samostojna znanost. V prvem sklopu prispevka pokrivata področje načrtovanja novih robotskih mehanizmov. Tudi samo načrtovanje mehanizmov lahko razdelimo na več področij. Predstavljen je prispevek o načrtovanju mobilne robotske platforme ter prispevek o uporabi elastomerov za robotske aktuatorje.

Sodobna robotika izrazito presega okvire industrijskih robotskih manipulatorjev. Označili bi jo lahko kot »inteligentno gibanje mehanizmov«. Sem sodijo tako mobilni roboti, hodeči roboti, roboti v medicini, servisnih robotov, ki nam bodo v pomoč v vsakdanjem življenju in pa raznovrstne industrijske naprave, ki so kakorkoli povezana z gibanjem: na primer dvigala, podajalniki. Eden izmed glavnih ciljev moderne robotike je razvoj metod in znanj, ki bodo omogočila prenos robotskih tehnologij iz tovarniških v domača okolja. V našem domačem okolju lahko roboti postanejo izvrstni pomočniki na več področjih, od pomoči ostarelim do pomočnikov pri zahtevnih in vsakodnevnih opravilih. V ta namen moramo usposobiti robote za avtonomno delovanje v nepričakovanih situacijah. Temu sledi drugi sklop prispevkov, ki združuje dela na področju humanoidne robotike, ki je v svetu raziskovalne robotike velik hit. Prispevki pokrivajo teme od samega načrtovanja programskih struktur za izvajanje humanoidnih nalog, pa do delovanja v spreminjajočem okolju in v interakciji z različnimi objekti.

Zadnji čas so zlasti popularni tudi rehabilitacijski roboti, ki so »pametni« funkcionalni rehabilitacijski pripomoček. Rehabilitacijski robot razbremeni fizioterapevte težkega fizičnega dela ter obenem poskrbi za učinkovito terapijo poškodovanca. V zadnjem sklopu je prispevek, ki opisuje uporabo robotike za rehabilitacijo pacientov po možganski kapi.

Na robotski konferenci vzpodbujamo interdisciplinarno debato med znanstveniki, ki se ukvarjajo s katerikoli od področij robotike. Prispevki predstavljajo inovativne raziskave sodelavcev dveh raziskovalnih skupin na katerih temelji prihodnost raziskovalnega dela na področju robotike pri nas.

Jadran Lenarčič
Leon Žlajpah
Andrej Gams

EKSPERIMENTALNA MOBILNA ROBOTSKA PLATFORMA

Peter Čepon^a, Roman Kamnik^a, Jernej Kuželički^b, Tadej Bajd^a, Marko Munih^a

^aUniverza v Ljubljani, Fakulteta za elektrotehniko, Tržaška 25, 1000 Ljubljana

^bIskra Avtoelektrika d.d., Polje 15, 5290 Šempeter pri Gorici

peter.cepon@robo.fe.uni-lj.si

Experimental Mobile Robotic Platform

Abstract

In this paper an experimental mobile robotic platform is presented. The system incorporates the mobile platform construction, the drive unit with drive and steering wheel, main controller, the drive controller and the software development environment. The system enables that all software development and control actions are accomplished on a remote supervisory computer. The main controller built on PC-compatible hardware running with xPC Target real time operating system is used for execution of the developed programs. The main controller communicates with the drive units via CAN communication. The software development is based on Mathworks Matlab tools Simulink, Stateflow and xPC Target.

1. Uvod

Mobilni roboti so naprave, ki so zmožne gibanja v danem okolju. Njihovo gibanje je gnano s pomočjo nog, koles ali drugih elementov, ki omogočajo gibanje po zemlji, vodi ali zraku. Zaradi hitrega razvoja na področjih lokalizacije, zaznavanja, kontrole in ostalih področij postaja mobilna robotika vedno bolj prisotna in razširjena na različnih področjih uporabe.

Od vseh izvedb je pri mobilnih robotih gibanje največkrat izvedeno z uporabo koles. Kinematika gibanja mobilnih kolesnih robotov je osnovana na osnovi trenja v dotiku koles s tlemi. V literaturi je mogoče zaslediti različne načine izgradnje kinematičnega modela [1], [2], [3].

Pri razvoju mobilnih robotov je pomembna uporaba programskih razvojnih okolij, ki omogočajo hiter razvoj algoritmov vodenja ter preprosto vključevanje novih senzorjev in ostalih komponent [4], [5].

V današnjem času postajajo mobilni roboti vse bolj komercialno dosegljivi in so na voljo vedno širši množici uporabnikov. Med podjetja, ki tržijo mobilne robote spadata podjetji MobileRobots (Pioneer 3, Patrol-Bot, Seekur) in K-Team (Hemisson, KoalaII).

V pričujočem delu je predstavljen razvoj eksperimentalne mobilne robotske platforme, ki je zasnovana na osnovi pogonskega sistema namenjenega pogonu električnih viličarjev. V prvem poglavju je predstavljen in opisan pogonski sistem ter mehanska zasnova mobilne robotske platforme. V drugem poglavju je razvit kinematični model mobilne platforme. V tretjem poglavju je predstavljen

razvojni sistem mobilne platforme, ki je osnovan na operacijskem sistemu xPC Target. V zadnjem poglavju je razvit preprost algoritem gibanja, ovrednoten z uporabo referenčnega kinematičnega merilnega sistema.

2. Mehanska konfiguracija

Mobilna robotska platforma je bila zgrajena na osnovi pogonskega sistema z enim pogonskim kolesom. Tovrsten pogonski sistem se uporablja za pogon industrijskih električnih viličarjev in je proizvod podjetja Iskra Avtoelektrika, d.d. iz Šempetra pri Gorici. Pogonski sistem sestavljata pogonski modul, ki je namenjen premikanju ter usmerjevalni modul, ki je namenjen usmerjanju pogonskega kolesa mobilne platforme. Pogonski modul vsebuje trifazni asinhronski motor (Iskra AML7103, 15 V, 2.2 kW), ki je krmiljen z ločenim AES1136 AC krmilnikom. Usmerjevalni modul vsebuje v svojem ohišju krmilnik, brezkrtačni DC motor (0.15 kW, max. navor 1.1 Nm pri 1600 obr/min) ter reduktor. Krmilnika modulov komunicirata z nadrejeno enoto preko vodila CAN [6] z uporabo CANopen komunikacijskega protokola [7]. Za zagotavljanje varnostni je na pogonskem modulu nameščena magnetna zavora.

Kinematična konfiguracija mobilne platforme je sestavljena iz aluminijastih profilov in vsebuje dve pasivni kolesi ter eno pogonsko kolo. Na konstrukcijo sta nameščeni dve akumulatorski bateriji, AC krmilnik, kontaktor in robotski krmilnik. Konstrukcija mobilne robotske platforme z mehanskimi elementi je prikazana na sliki 1.

3. Kinematični model

Kinematični model mobilne robotske platforme opisuje pozicijo in orientacijo glede na globalni koordinatni sistem. Pri izpeljavi kinematičnega modela je potrebno upoštevati kinematične omejitve koles, ki jih posamezno kolo vsiljuje mobilni robotski platformi med gibanjem. Lego mobilne platforme opišemo z vektorjem dveh translacijskih koordinat ter enim orientacijskim kotom:

$$\xi_o = \begin{bmatrix} x \\ y \\ \theta \end{bmatrix} \quad (1)$$

Ob predpostavki, da se mobilni robot giblje le v ravnini, lahko s pomočjo rotacije okoli navpične osi z ,

transformiramo vektor hitrosti $\dot{\xi}_o = [\dot{x} \quad \dot{y} \quad \dot{\theta}]$ v lokalni koordinatni sistem:

$$\dot{\xi}_r = \mathbf{R}(\theta) \cdot \dot{\xi}_o \quad (2)$$

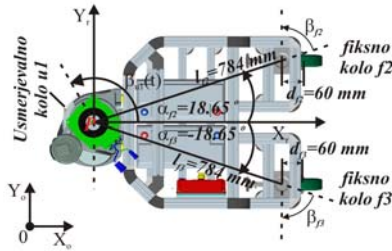
Vsako kolo robotu vsiljuje omejitve gibanja. Te lahko pri fiksnem standardnem kolesu izrazimo s petimi konstantnimi parametri, ki opisujejo položaj vpetja kolesa (α , β , l , r , d) in spremenljivko kota zasuka kolesa $\varphi(t)$ kot prikazuje slika 2.

Za posamezno fiksno standardno kolo zapišemo omejitve kotaljenja:

$$[\sin(\alpha + \beta) \quad -\cos(\alpha + \beta) \quad (-l) \cdot \sin \beta] \cdot \mathbf{R}(\theta) \cdot \dot{\xi}_o = r \cdot \dot{\varphi}$$



Slika 1: Mehanska konstrukcija mobilne platforme



Slika 2: Prikaz kinematičnih parametrov za trikolesnik s fiksnimi kolesi

Gibanje mobilnega robota omejujejo le standardna kolesa, zato z matrično enačbo združimo in zapišemo le omejitve kotaljenja za vsa standardna kolesa pritrjena na mobilni robot:

$$\mathbf{J}_1(\beta_{u1}) \cdot \mathbf{R}(\theta) \cdot \dot{\xi}_o - \mathbf{J}_2 \cdot \dot{\varphi}(t) = 0 \quad (3)$$

Na enak način lahko izrazimo omejitve drsenja za fiksno standardno kolo:

$$[\cos(\alpha + \beta) \quad \sin(\alpha + \beta) \quad d + l \cdot \sin \beta] \cdot \mathbf{R}(\theta) \cdot \dot{\xi}_o = 0$$

Tudi v tem primeru z matrično enačbo združimo in zapišemo omejitve drsenja standardnih koles:

$$\mathbf{C}_1(\beta_u) \cdot \mathbf{R}(\theta) \cdot \dot{\xi}_o = 0 \quad (4)$$

Parametri potrebni za določitev kinematičnega modela mobilne robotske platforme so zbrani v Tabeli 1.

kolo	α [°]	β [°]	l [mm]	d [mm]	r [mm]
u1	0	$\beta_{u1}(t)$	0	0	115
f2	18,65	71,35	784	60	100
f3	-18,65	108,65	784	60	100

Tabela 1: Parametri mobilne platforme z dvema fiksnima in enim usmerjevalnim kolesom

S pomočjo parametrov iz Tabele 1 zapišemo matriko kinematičnih omejitev kotaljenja in drsenja:

$$\mathbf{J}_1(\beta_{u1}) = \begin{bmatrix} \sin \beta_{u1}(t) & -\cos \beta_{u1}(t) & 0 \\ 1 & 0 & -251 \\ 1 & 0 & 251 \end{bmatrix}$$

$$\mathbf{J}_2 = \begin{bmatrix} r_{u1} & 0 & 0 \\ 0 & r_{f2} & 0 \\ 0 & 0 & r_{f3} \end{bmatrix} = \begin{bmatrix} 115 & 0 & 0 \\ 0 & 110 & 0 \\ 0 & 0 & 110 \end{bmatrix}$$

$$\mathbf{C}_1(\beta_{u1}) = \begin{bmatrix} \cos \beta_{u1}(t) & \sin \beta_{u1}(t) & 0 \\ 0 & 1 & 803 \end{bmatrix}$$

Vse kinematične omejitve kotaljenja in drsenja združimo ter jih zapišemo v kompaktni obliki:

$$\begin{bmatrix} [\sin \beta_{u1}(t) & -\cos \beta_{u1}(t) & 0] \cdot \mathbf{R}(\theta) & 0 & 0 & 115 \\ [\cos \beta_{u1}(t) & \sin \beta_{u1}(t) & 0] \cdot \mathbf{R}(\theta) & 0 & 0 & 0 \\ [0 & 1 & 803] \cdot \mathbf{R}(\theta) & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \dot{\xi}_o \\ \dot{\beta}_u \\ \dot{\varphi} \end{bmatrix} = 0$$

S pomočjo kinematičnih parametrov lahko določimo tudi stopnje mobilnosti δ_m , vodljivosti δ_v in manevrabilnosti δ_M mobilne platforme:

$$\delta_m = \dim N[\mathbf{C}_1(\beta_{u1})] = 3 - \text{rang}[\mathbf{C}_1(\beta_{u1})] = 1, \\ \delta_v = \text{rang}[\mathbf{C}_{1u}(\beta_{u1})] = 1, \delta_M = \delta_m + \delta_v = 2$$

4. Robotski krmilnik

Pri razvoju robotskega krmilnika je bilo izbrano okolje Matlab-Simulink in orodje xPC Target, ki deluje v realnem času in je osnovano na nadzorno/ciljnem principu delovanja (slika 3). Takšen razvojni sistem vedno tvori cilj in razvojni računalnik, ki ves čas delovanja komunicirata preko brezžične TCP/IP povezave. Za ciljni računalnik, ki je pritrjen na konstrukcijo mobilne platforme, je bil izbran Advantech-ov vgrajeni računalnik PCM3380. Njegova naloga je izmenjava podatkov z razvojnim računalnikom, zajem podatkov senzorjev ter komuniciranje z usmerjevalnim in pogonskim modulom preko vodila CAN. Vse operacije se vršijo v realnem času z vzorčno frekvenco 1 kHz. Razvojni računalnik (prenosni računalnik) je namenjen nadzoru in razvoju algoritmov vodenja, kar se opravlja v programskem okolju Matlab-Simulink.



Slika 3: Konceptualni prikaz delovanja nadzorno/ciljnega sistema

Razvoj programske opreme poteka v grafičnem okolju s povezovanjem funkcionalnih blokov iz Simulink knjižnic ali t.i. S-funkcij razvitih v programskem jeziku C.

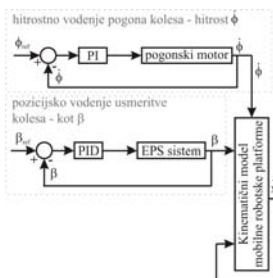
5. Eksperimentalna evalvacija delovanja mobilne robotske platforme

Prikazan je eksperiment na realni mobilni robotski platformi, ki jo prikazuje slika 4. Za namen evalvacije gibanja mobilne robotske platforme po vnaprej določeni trajektoriji ter preverjanja njegovega kinematičnega modela, je bila načrtana preprosta shema vodenja. Za preverjanje kinematičnega modela mobilne platforme je bil uporabljen optični merilni sistem Optotrak.

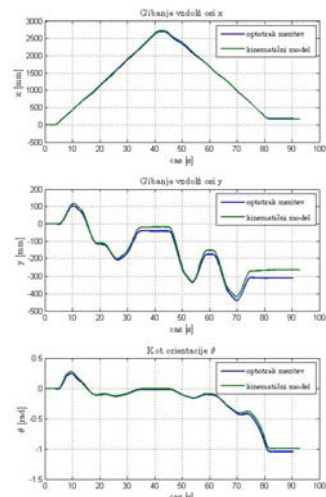


Slika 4: Prikaz mobilne robotske platforme

Za pogonski modul je bilo v eksperimentu uporabljeno hitrostno vodenje asinhronskega motorja s PI regulatorjem ter pozicijsko vodenje s PID regulatorjem za usmerjevalni modul. Informaciji o hitrosti vrtenja $\dot{\phi}$ in kotu usmeritve pogonskega kolesa β sta uporabljeni v kinematičnem modelu za izračun položaja mobilne platforme v prostoru. Blokovna shema vodenja mobilne robotske platforme je prikazana na sliki 5.



Slika 5: Blokovna shema vodenja mobilne robotske platforme



Slika 6: Primerjava izmerjene in izračunane lege mobilne robotske platforme med gibanjem

Rezultati gibanja mobilne robotske platforme po vnaprej načrtani trajektoriji so prikazani v grafični obliki na sliki 6. Na treh grafih so primerjane zunanje koordinate lege mobilne robotske platforme x , y in θ izmerjene z merilnim sistemom Optotrak ter vrednosti izračunane s kinematičnim modelom robota.

Iz rezultatov na sliki 6 je razvidno dobro ujemanje rezultatov kinematičnega modela in realnega sistema. Rezultati kažejo, da navkljub relativno zanesljivemu določanju lege, prihaja do odstopanj od dejanske lege izmerjene s kinematičnim merilnim sistemom Optotrak. Predvidevamo, da je vzrok odstopanj spodrsavanje pogonskega kolesa, ki ga notranji senzorji ne zaznajo.

6. Zaključek

V pričujočem delu je prikazan razvoj in izgradnja mobilne robotske platforme, ki je osnovana na pogonskem sistemu namenjenem pogonu in usmerjanju industrijskih vilicarjev. Za mobilno robotsko platformo je bila načrtana mehanska konstrukcija in izpeljan kinematični model, ki je bil uspešno implementiran in preizkušen na realnem primeru. V namen vodenja je razvit robotski krmilnik, ki deluje v realnem času. S pomočjo merilnega sistema Optotrak je opravljena tudi evalvacija gibanja mobilne robotske platforme v prostoru. Iz dobljenih rezultatov je razvidno, da je kinematični model pravilen in da krmilnik deluje robustno.

7. Zahvala

Projekt je bil sofinanciran v okviru programskega financiranja ARRS Slovenija »[Analiza in sinteza gibanja pri človeku in stroju](#)« P2-0228 in Ciljnega raziskovalnega programa MORS in ARRS, Slovenija "Znanje za varnost in mir 2004-2010": "Mobilni robotski sistem za izvidniške, raziskovalne in reševalne namene" M2-0116.

8. Literatura

- [1] G. Campion., G. Bastin, B. D'Andrea-Novel, Structural properties and classification of kinematic and dynamic models of wheeled mobile robots, *IEEE Transaction on Robotics and Automation*, 12(1), 47–62, 1996.
- [2] D. Kim, W. H. Kwon, H. S. Park, Geometric kinematics and applications of a mobile robot, *International Journal of Control, Automation, and Systems*, 1(3), 376–384, 2003.
- [3] P. Muir, C. P. Neuman, Kinematic modeling of wheeled mobile robots, *Technical report, Robotics Institute, Pittsburgh, PA*, 1986.
- [4] G. Dudek, M. Jenkin, *Computational Principles of Mobile Robotics*, Cambridge University Press, Cambridge, 2000.
- [5] R. Siegwart, I. R. Nourbakhsh, *Introduction to Autonomous Mobile Robots*, MIT Press, Cambridge, 2004.
- [6] M. Farsi, M. B. M. Barbosa, *CANopen implementation: applications to industrial networks*, Research Studies Press Ltd., 2000.
- [7] H. Boterenbrood, Canopen high-level protocol for can bus, NIKHEF Amsterdam, 2000.

DESIGN OF A CUSTOM ELECTRONICS DRIVER FOR DIELECTRICS ELASTOMER ACTUATORS

Mitja Babič¹, Rocco Vertechy², Giovanni Berselli³, Jadran Lenarčič¹

¹Department of Automation, Biocybernetics, and Robotics, Jožef Stefan Institute, Ljubljana, Slovenia

²PERCRO Laboratory, Scuola Superiore Sant'Anna, Pisa, Italy

³Department of Mechanical and Aeronautical Engineering, University of Bologna, Bologna, Italy

e-mail: mitja.babic@ijs.si

ABSTRACT

Actuators based on Dielectric Elastomers are a promising technology in robotic and mechatronic applications. The overall actuator performances are influenced by the non-linear and highly coupled electrical and mechanical behavior of the employed materials. To date, the practical electro-mechanical response and controllability of actuators based on Dielectric Elastomers are limited by the inadequacy of the employed driving circuits, which are based on voltage-regulated converters. In this paper a novel activation strategy is proposed for Dielectric Elastomer actuators, which is based on a custom electronic driver derived from the flyback transformer topology. The flyback converter design is based on the specifications given by the Dielectric Elastomer actuator mechanical design and the desired actuator response.

1 INTRODUCTION

Dielectric Elastomers (DE) are incompressible visco-elastic dielectrics which can experience deviatoric (isochoric) finite deformations in response to applied large electric fields while, at the same time, alter the applied electric fields in response to the deformations undergone [1–3]. Thanks to the strong electro-mechanical coupling, DE intrinsically offer great potentialities for conceiving novel solid-state mechatronic devices, in particular actuators, which are more integrated, lightweight, economic, silent, resilient and disposable than equivalent devices based on traditional technologies [4–5]. Up to now, practical application of DE-based actuators is hampered by poor material reliability, non-linear and non-repeatable electro-mechanical behavior, non-trivial control and inappropriateness of the electrical driving circuit which often degrade actuator response and energy efficiency [6–12].

From the electrical standpoint, DE actuators are capacitive loads. To date, the driving circuits employed for the activation of DE actuators are based on switch-mode voltage-regulated low-ripple High Voltage (HV) Direct Current (DC) bench-top power supplies (with moderate power and high costs) [11] or portable DC/DC converters (with low power and moderate costs) [12], which are both

not suited for capacitor charging applications since, by nature, they consume at least double the energy that is being stored in the capacitor (thus reducing DE actuator efficiency) and provide limited peak currents that slow-down the capacitor charging process (thus limiting DE actuator dynamic response). Instead, DE actuators should be better activated via purposely designed pulsed power supplies or current-regulated sources (e.g. constant current supplies), which deliver larger currents, are cheaper and smaller, and require smaller internal filter capacitors (that usually lead to significant energy losses in high repetition rate capacitor charging applications) as compared to voltage-regulated supplies with identical power ratings. Moreover, developed DE actuator controllers feature a nested loop architecture consisting of an external position loop closed around the DE actuator output link and of an internal voltage loop implemented within the DE actuator voltage-regulated low-power supply/converter, usually having scarce response time that drastically degrades DE actuator dynamics and controllability.

In this paper a flyback converter as an electronics driver for controlling DE actuators is presented. The flyback converter design is based on the specifications given by the DE actuator mechanical design and the desired actuator response.

2 BACKGROUND ON DIELECTRIC ELASTOMER ACTUATORS

For actuation usage, DE are typically shaped in thin films which are coated on both sides with compliant electrodes to form an Electro-Active Film (EAF) [2,3]. The placement of a differential electric potential (voltage) between the electrodes of the EAF generates an electric field within the DE that may induce an isochoric film area expansion (see Fig. 1). Such an electrically induced deformation can be used to produce useful mechanical work.

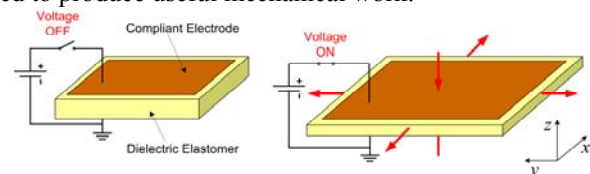


Fig. 1: Electrically induced isochoric area expansion of an Electro-Active Film (EAF).

An example of planar DE linear actuator is depicted in Fig. 2, where the independent variable x indicates the DE actuator length while the parameter b indicates the DE actuator width. The actuator features two identical rigid beams (one fixed and one movable) made of an acetal resin (Delrin) and a DE made by a visco-elastic acrylic film (VHBTM-4905 by 3MTM, characterized by non-linear and non-repeatable dynamic mechanical response) coated with compliant electrodes made of a silver conductive grease (CW-7100 by CircuitWorks, having good electrical conductivity). Because of the mechanical constraints, the activation of the DE makes it expand uniformly in the x direction only and shrink in the thickness direction. Such a uniform pure-shear deformation [13] is used to produce useful mechanical work by translating (in the x direction) the movable beam along with any external force that is applied to it.

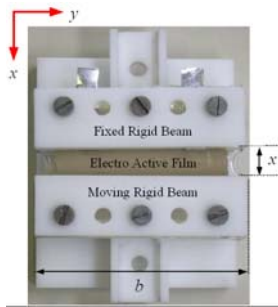


Fig. 2: Planar pure-shear type linear Dielectric Elastomer (DE) actuator.

In terms of potential performances, note that a DE actuator of the type depicted in Fig. 2, having dimensions $b = 100\text{mm}$ and $\underline{x} = 15\text{mm}$ (pre-stretched state configuration), weight lighter than 50gr and cost cheaper than 5€, can produce an actuator stroke, $x - \underline{x}$, larger than 13mm and an output force larger than 5N.

From the electrical point of view, the DE actuator is equivalent to a serial variable RC load depicted in Fig. 3 [4,5].

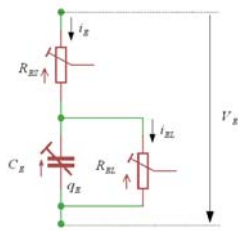


Fig. 3: DE actuator electrical model.

In Fig. 3, V_E is the voltage applied between the DE electrodes, i_E is the current feeding the DE actuator, C_E is the DE variable capacitance, i_{EL} is the leakage current flowing through the DE capacitance, represented as a variable resistance R_{EL} (dependent on the actuator configuration and the stored charge in C_E), R_{ES} is the DE electrode variable resistance (dependent on the actuator configuration) and q_E is the electric charge stored on the DE electrode surfaces. The DE variable capacitance is depending on actuator length x as

$$C_E(x) = \varepsilon \varepsilon_0 b x^2 / \underline{z} x = \underline{C}_E x^2 / \underline{x}^2, \quad (1)$$

where $\varepsilon_0 = 8.85 \times 10^{-12} \text{F/m}$ and $\varepsilon = 4.5$ being the dielectric constant of the vacuum and the relative dielectric permittivity of the acrylic DE respectively and \underline{x} and \underline{z} are the pre-stretched state configuration (\underline{C}_E being the minimum DE capacitance at \underline{x}).

3 FLYBACK CONVERTER DESIGN BASED ON DE ACTUATOR SPECIFICATIONS

The flyback converter schematic is depicted in Fig. 4. The operation principle is as follows [14].

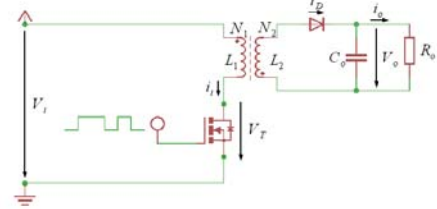


Fig. 4: Flyback converter schematic.

During transistor on time a constant voltage V_i is applied on the primary inductance L_1 , so current i_1 it ramps up linearly at a rate of

$$\frac{di_1}{dt} = \frac{V_i - V_{Rds(on)}}{L_1}, \quad (2)$$

where $V_{Rds(on)} = R_{ds(on)} \cdot i_1$ is the voltage drop across the power transistor (MOSFET). At the end of the on time, the primary current has ramped up to

$$I_{1P} = \frac{(V_i - V_{Rds(on)}) \cdot T_{ON}}{L_1}. \quad (3)$$

This current represents a stored energy of

$$W_i = \frac{1}{2} L_1 I_{1P}^2. \quad (4)$$

During discharge of this stored energy the output voltage will rise to whatever level is needed to cause current to flow, thus dissipating the stored energy. Because the current in an inductor cannot change instantaneously, at turn off, the primary current transfers to the secondary with amplitude

$$i_2 = i_1 \frac{N_1}{N_2}. \quad (5)$$

Secondary current decrease linearly at rate

$$\frac{di_2}{dt} = \frac{V_o + V_D}{L_2}, \quad (6)$$

where L_2 is the secondary inductance and V_D is the voltage drop across the output diode. If the secondary current has decrease to zero before the start of the next on time, all the energy stored in the primary has been delivered to the load and the circuit is said to be operating in the discontinuous mode.

The proposed DE actuator electrical driver is depicted in Fig. 5, which is obtained with the replacement of the filter capacitor C_o in Fig. 4 with the DE actuator electrical model in Fig. 3. The driver design parameters are defined by the DE actuator electrical constants, which are the initial pre-stretched capacitance \underline{C}_E , capacitance change ΔC_E during

actuator movement, electrode resistance R_{ES} and resistance R_{EL} representing the leakage current flowing through the actuator capacitance C_E .

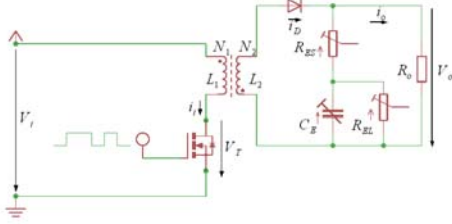


Fig. 5. Proposed electrical driver schematic for DE actuators.

Note that in this application the pulse control (PWM) will be a position feedback control loop, which will charge the actuator capacitance C_E to a given voltage level V_o in order to reach the desiderated position x_r . The capacitor will be charged and additional pulses will be applied to supply load resistance R_o currents $i_o = V_o/R_o$ and DE actuator losses. The DE actuator capacitor C_E can not be discharge throughout the flyback transformer, so an output load R_o is needed. Usually $R_{ES} \ll R_o$ and $R_{EL} \gg R_o$, so the electrical time constant can be estimated as

$$\tau_{el} = R_o C_E . \quad (7)$$

For fast actuator response the electrical time constant must be much smaller then the actuator maximum position rise time (τ_{rp}), defined from 10% to 90% of steady-state position value ($\tau_{el} \ll \tau_{rp}$).

The energy required to charge the capacitor and to supply load resistance currents is

$$W_E = \frac{C_E \cdot V_o^2}{2} + \frac{V_o^2}{R_o} \cdot \frac{1}{f} . \quad (8)$$

During calculations the DE actuator losses are ignored supposing $R_{EL} \gg R_o$. The needed number of charging pulses based on the switching frequency $f = 1/T$ and desiderated stated time t_s is

$$N_{puls} = t_s \cdot f . \quad (9)$$

The energy required per one charging pulse is

$$W_{EPuls} = \frac{W_E}{N_{puls}} . \quad (10)$$

Including an estimation of the circuit efficiency η the input energy that must be supplied is

$$W_i = \frac{W_{EPuls}}{\eta} . \quad (11)$$

Based on the selected input voltage V_i , the maximum on time $T_{omax} = T \cdot 45\%$ (discontinuous mode) and the calculated W_i , the primary peak current I_{1P} can be estimated from equations (3) and (4) as

$$I_{1P} = \frac{2W_i}{(V_i - V_{Rds(on)})T_{omax}} . \quad (12)$$

Determined the primary peak current the primary inductance is

$$L_1 = \frac{(V_i - V_{Rds(on)}) \cdot T_{ON}}{I_{1P}} . \quad (13)$$

From here the approach of designing the flyback converter is the same as described in [14].

In switching power converters the output capacitor is chosen on the basis of the specified output ripple. At maximum output current, during the transistor on time of T_{on} , the filter capacitor C_o , carries the maximum output current I_{omax} . It slopes in voltage an amount of

$$\Delta V_o = \frac{I_{omax} \cdot T_{on}}{C_o} . \quad (14)$$

The electrical time constant τ_{el} (equation (7)) and the maximum voltage ripple ΔV_o (equation (14)), which influence the position of the DE actuator, define the selection of the output load resistance R_o .

4 TEST BENCH

These experimental tests have been performed using a custom-designed test bench which is shown in Fig. 6.

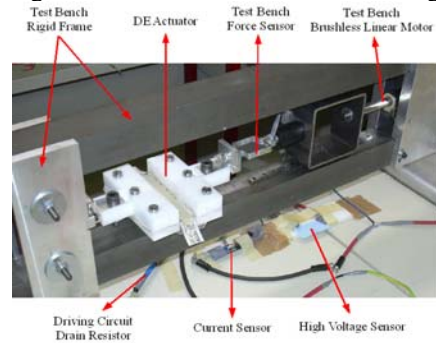


Fig. 6. Test-bench employed for testing the performances of DE actuator driving circuits and controllers.

The test-bench features a custom-made force sensor and a brushless linear motor which is commanded by a closed-loop force control so as to be able to generate on the tested DE actuator any desired profile of the external force F_{ext} . Note that this closed-loop force control has high bandwidth and has been properly designed and tuned in order not to influence DE actuator dynamics [15]. The test-bench interfaces with a standard PC (an AMD Athlon 64 2.2 GHz processor having 1 GB of RAM) equipped with a National Instruments PCI 6259 high-speed data acquisition board which is used for acquiring the data from the test-bench sensors as well as for delivering the control signals. The closed-loop controllers are developed through the Matlab/Simulink/RealTime Workshop environment and run together on the real-time operating system RTAI-Linux (OpenSUSE 10.3, Linux kernel 2.6.22.18 and RTAI 3.6.1). The adopted driving circuit components are: an HV cathode ray tube (CRT) monitor transformer with $L_1 = 1.5\text{mH}$; a PWM having switching frequency $f = 1\text{kHz}$; a load resistor $R_o = 100\text{M}\Omega$; and an input power supply with $V_i = 2.5\text{V}$. To prevent DE actuator failure this driving circuit has been dimensioned so that its maximum output voltage is lower than 5kV. For the experiments, the external load regulated by the test-bench has been set to $F_{ext} = 3\text{N}$.

4.1 Closed-loop position response

In order to show the benefits of the proposed electrical driver a self-tuning adaptive PID controller [16] is used for controlling the PWM module and flyback converter respectively.

Fig. 7 reports the system response to a sequence of forward and backward steps within 0±3mm range. From the actuator stroke plots we can see that the tracking error goes to zero. From the controller's duty cycle commands and EAF voltage plots we can note that the controllers, with an adequate control signals, overcomes the problem of non-linear and non-repeatable dynamic mechanical response of DE actuators.

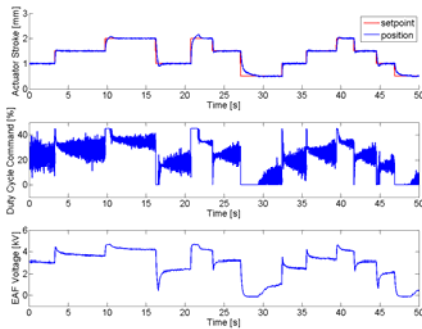


Fig. 7. Experimental closed-loop position response.

Fig. 8 reports the system trajectory tracking response. Good trajectory tracking is achieved with the control of the flyback converter PWM signal, and selecting an appropriate controller switching frequency, which results in fast DE voltage and position control.

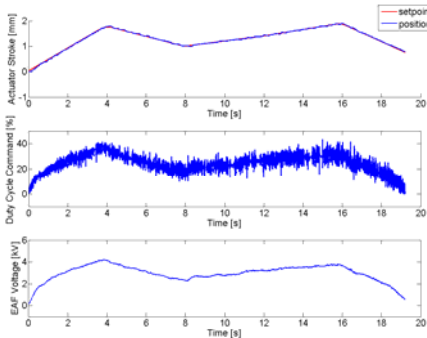


Fig. 8. Experimental trajectory tracking.

5 CONCLUSIONS

The performance of DE actuators depends greatly on actuator geometry and actuator mechanical design. Besides the mechanical design also an adequate electronic driver and controller significantly contribute to the performance and response of DE based actuators. The drawbacks of usually used electronics device for driving DE actuators is that they are not designed and not suitable for capacitors (DE actuators) charging.

The novel electronic driver, which is derived from the flyback transformer topology, is more performing, has a

simpler architecture, is composed by a smaller number of components, and should be more compact and less expensive than traditional driving circuits, which are based on high-voltage voltage-regulated DC/DC converters or DC power supplies.

REFERENCES

- [1] R. A. Toupin, The elastic dielectrics, *Archive for Rational Mechanics and Analysis* 5 (1956) 849–915.
- [2] R. E. Pelrine, R. D. Kornbluh, J. P. Joseph, Electrostriction of polymer dielectrics with compliant electrodes as a means of actuation, *Sensors and Actuators A: Physical* 64 (1) (1998) 77–85.
- [3] G. Kofod, P. Sommer-Larsen, R. Kornbluh, R. Pelrine, Actuation response of polyacrylate dielectric elastomers, *Journal of Intelligent Material Systems and Structures* 14 (12) (2003) 787–793.
- [4] Y. Bar-Cohen, *Electroactive Polymer (EAP) Actuators as Artificial Muscles: Reality, Potential and Challenges*, Vol. PM136 of 2, SPIE Press, 2004.
- [5] L. K. Kim, S. Tadokoro, *Electroactive Polymers for Robotic Applications: Artificial Muscles and Sensors*, Eds. Springer, 2007.
- [6] H. R. Choi, K. M. Jung, S. Ryew, J.-D. Nam, J. Jaewook, J. C. Koo, K. Tanie, Biomimetic soft actuator: design, modeling, control, and applications, *IEEE/ASME transactions on mechatronics* 10 (5) (2005) 581–593.
- [7] S. Q. Xie, P. F. Ramson, D. D. Graaf, E. P. Calius, I. A. Anderson, An adaptive control system for dielectric elastomers, *IEEE International Conference on Industrial Technology* 14 (17) (2005) 335–340.
- [8] R. Vertechy, V. P. Castelli, K. J. Waldron, On the driving circuits of electro-sensitive elastomers, in: *Proceedings of Actuator 2006*, 2006, pp. 1–4.
- [9] A. Wingert, M. D. Lichter, S. Dubowsky, and M. Hafez. Hyper-redundant robot manipulators actuated by optimized binary-dielectric polymers. *SPIE*, 4695 (2002) pages 415–423.
- [10] Pelrine R. E., Kornbluh R. D. and Joseph J. P. 1998 Electrostriction of polymer dielectrics with compliant electrodes as a means of actuation *Sensors and Actuators A-Phys.* 64 77-85.
- [11] P. Lochmatter, Development of a shell-like electroactive polymer (EAP) actuator, Ph.D. thesis, Swiss Federal Institute Of Technology Zurich (2007).
- [12] A. Wingert, M. D. Lichter, S. Dubowsky, On the design of large degree-of-freedom digital mechatronic devices based on bistable dielectric elastomer actuators, *IEEE/ASME Transactions on Mechatronics* 11 (4) (2006) 448–456.
- [13] R. W. Ogden, Large deformation isotropic elasticity: on the correlation of theory and experiment for incompressible rubber-like solids, *Proc. Roy. Soc. London A-326* (1972) 565-584.
- [14] A. I. Pressman, *Switching Power Supply Design*, 2 edition, McGraw-Hill Professional, 1998.
- [15] G. Palli, C. Melchiorri, Non-model based friction and load compensation in linear electric drives, *Motion Control 2007*, 51th National congress on motion control, Milan, Italy (2007) 10–11.
- [16] F. Lin, R. Brandt, G. Saikalis, Self-tuning of PID controllers by adaptive interaction, *American Control Conference* 5 (2000) 3676–3681.

ROBOTSKE PROGRAMSKE STRUKTURE V PROGRAMIRANJU HUMANOIDNIH ROBOTOV

Andrej Kos

Odsek za avtomatiko, biokibernetiko in robotiko

Institut "Jožef Stefan"

Jamova cesta 39, 1000 Ljubljana, Slovenija

e-mail: andrej.kos@ijs.si

ABSTRACT

This paper presents a description of various robotic system architectures and their usage in writing programs for robotic experiments. The robotic software applications are complex, as they need to monitor various sensors, analyse the signals, react to the signals and drive the actuators during execution of various tasks. Therefore, appropriate programming architecture and programming tools are needed. Various existing approaches to this problem are discussed and our approach to writing robotic system architecture for humanoid robot is presented.

1 UVOD

Robotski programi so kompleksni. Programi morajo v realnem času zajeti signale množice senzorjev, jih analizirati glede na nalogo, ki jo robot trenutno opravlja, določiti ustrezne odzive na signale in krmiliti aktuatorje. Poleg opravljanja primarne robotske naloge, mora robot sproti zaznavati in reagirati na vrsto neobičajnih situacij. Vse to se mora izvajati paralelno in asinhrono, kar je vzrok kompleksnosti sistema. Programska struktura robotskih programov pripomore k obvladovanju kompleksnosti programa. Programska struktura določa, kako je robotski program razdeljen na podsisteme in kakšne so povezave med njimi, ter na kakšen način je vse to implementirano z izbranim programskim jezikom. Za različne tipe robotskih aplikacij obstajajo različne strukture. Optimalne strukture, ki bi ustrezala vsem zahtevam, ni. V članku bomo omenili različne izvedbe robotskih programskih struktur. Predstavili bomo tudi, kako smo uporabili del tega znanja, pri izvedbi programske strukture humanoidnega robota, kot nadgradnjo obstoječega osnovnega vmesnika.

2 POSEBNE ZAHTEVE ROBOTSKIH PROGRAMOV

Dobro zasnovana programska struktura omogoča lažje načrtovanje, pisanje in testiranje aplikacij, ter pripomore k boljšemu razumevanju celotnega programa in posameznih podsistemov. Dobro definirana robotska programska struktura je koristna, ker nas pri načrtovanju in pisanju programa usmerja[1]. Struktura mora čim bolj omejiti stil strukture in način implementacije, ter čim manj omejevati uporabnika programske strukture glede funkcionalnosti

robota in enostavnosti implementacije raznovrstnih robotskih programov. Robotske programske strukture imajo posebne zahteve, ker moramo upoštevati značilnosti robotskih sistemov. Robot mora delovati asinhrono, v realnem času in v neznanem dinamičnem okolju[2]. Odzivi robota na signale senzorjev morajo obsegati zelo različne časovne intervale. Regulacijske zanke lahko delujejo z mikrosekundnim taktom, kompleksnejše naloge pa lahko delujejo z večminutnim taktom.

2.1 Modularnost

Robotska programska struktura je običajno zgrajena tako, da je sistem razdeljen na več podsistemov, ki si izmenjujejo informacije med seboj [3]. Sistem je ponavadi razdeljen hierarhično, tako da so podsistemi zgrajeni iz še manjših podsistemov. Tako se kompleksnost zmanjša s povečanjem abstrakcije sistema. Hierarhija je lahko urejena po različnih kriterijih. Podsistemi so lahko porazdeljeni časovno. To pomeni, da višji sloji hierarhije delujejo pri višjih frekvencah, ker rabijo veliko časa, da generirajo kompleksen odziv robota, nižji sloji delujejo hitreje, najnižji sloji, ki vsebujejo regulacijske zanke, ki so neposredno povezane z robotsko strojno opremo, pa delujejo z najvišjimi frekvencami, ki jih dovoljuje strojna oprema. Druge hierarhije so porazdeljene po nalogah. Tako so višje-nivojske naloge sestavljene iz nižje-nivojskih nalog. Visoko-nivojska naloga se tako izvrši z izvršitvijo njenih podnalog [4]. Različne aplikacije imajo različne zahteve po porazdelitvi problema.

2.2 Programska orodja

Dobro definirana pravila pripomorejo k bolj učinkovitemu načrtovanju programskega sistema. Še bolje je, če lahko uporabimo že obstoječe programske orodje, ki omogoča pisanje programa robotske aplikacije v izbrani programski strukturi. Ta orodja so lahko programske knjižnice funkcij, specializirani programski jeziki ali grafična programska orodja. Orodja zagotavljajo upoštevanje pravil strukture robotskega programa, obenem pa pred programerjem skrijejo kompleksnost implementacije teh pravil. Programi med drugim omogočajo komunikacijo med podsistemi, sočasno izvajanje več podsistemov in drugo.

2.3 Vrste struktur robotskih programov

Prve robotske programske strukture (ang. Robotic Systems Architectures) so bile narejene kot zaporedje treh funkcionalnih elementov: zajemanja podatkov s senzorjev, načrtovanje odziva in odziv. Podsystem za zajemanje podatkov s senzorjev je zgradil model naloge, podsystem za načrtovanje odziva pa je model uporabil pri izračunu odziva. Trije podsystemi so delovali zaporedno. Ker je podsystem za načrtovanje odziva bil počasen, je vnesel veliko zakasnitev v celoten sistem. Med tem ostala dva podsystema nista delovala.

Boljše strukture omogočajo, da so najnovejši podatki s senzorjev prisotni pri načrtovanju odziva. Višje-nivojski podsystemi odziva lahko prekinejo delovanje nižje-nivojskih in tako lahko robot pravočasno reagira na nepričakovano motnjo. Ena izmed teh struktur je T3 [5] robotska programska struktura. Ima tri nivoje kontrolnih procesov: načrtovanje, tvorjenje zaporedja in vodenje v realnem času.

2.4 Paralelizem

Robotske aplikacije potrebujejo možnost sočasnega izvajanja nalog, da lahko robot dela več stvari hkrati [6]. To je lahko realizirano z večnitnimi programi ali na druge načine. Pri tem mora biti poskrbljeno za dostop do skupnih programskih struktur in strojne opreme.

Primer paralelizma lahko pokažemo pri robotskem operacijskem sistemu URBI. Tu paralelizem ni implementiran večnitno, ampak znotraj ene niti poskrbi, da se ustrezen del vsake naloge izvrši znotraj ene periode. Kompleksnost kode, ki zagotavlja paralelizem, je pred uporabnikom skrit v semantiki jezika. Z visoko stopnjo abstrakcije uporabniku ponuja zelo enostaven in razumljiv zapis programa.

```
whenever(face.visible)
{
  headPan.val += camera.xfov * face.x
  &
  headTilt.val += camera.yfov * face.y
}
```

Zgornji zapis spreminja vrednosti spremenljivk headPan in headTilt hkrati – v istem taktu. Sočasnost izvajanja obeh vrstic programske kode je označena z znakom &.

3 ROBOTSKI OPERACIJSKI SISTEMI

Najnovejši tip robotske programske strukture so tako imenovani robotski operacijski sistemi. Vsak tip robota je ustvarjen za določene tipe nalog, kar pogojuje tako arhitekturo in način programiranja. Mnogi raziskovalci skušajo določiti skupne lastnosti robotov in kaj bi lahko standardizirali, da bi dobili univerzalni operacijski sistem, ki ga lahko vsak uporablja.

Standardizacija bi omogočila raziskovalcem, da se bolj osredotočajo na nove eksperimente, in ne toliko na implementacijo eksperimentov na posameznem robotu, ki ga uporabljajo. Prenosljivost programske kode med roboti bi lahko tudi okrepilo sodelovanje med raziskovalci. Brez skupnega operacijskega sistema koda praktično ni prenosljiva. To je eden izmed glavnih ovir pri ponavljanju, preverjanju in uporabi objavljenih eksperimentov z različnimi roboti. Raziskovalci pogosto objavijo izboljšavo kakšnega algoritma, na primer za navigacijo, ne objavijo pa programske kode, ki bi jo lahko drugi neodvisno preverili. Roboti so zelo kompleksni stroji, zato morajo robotiki veliko časa porabiti za reševanje in implementacijo najbolj običajnih problemov, ki so jih mnogi že večkrat rešili, da dobi robot tisto funkcionalnost, ki je potrebna za izvedbo njihovega eksperimenta.

Sestavljanje različnih algoritmov v robotski program je zamuden proces. Z nakupom robota dobimo programske funkcije za vodenje robota, za vse ostalo pa moramo poskrbeti sami.

Robotski operacijski sistemi omogočajo abstrakcijo strojne opreme, dostop do nizko-nivojskih robotskih strojnih ukazov, implementacijo funkcij, ki so pogosto v rabi in komunikacijo med podsystemi ter dobro definirano programsko sintakso.

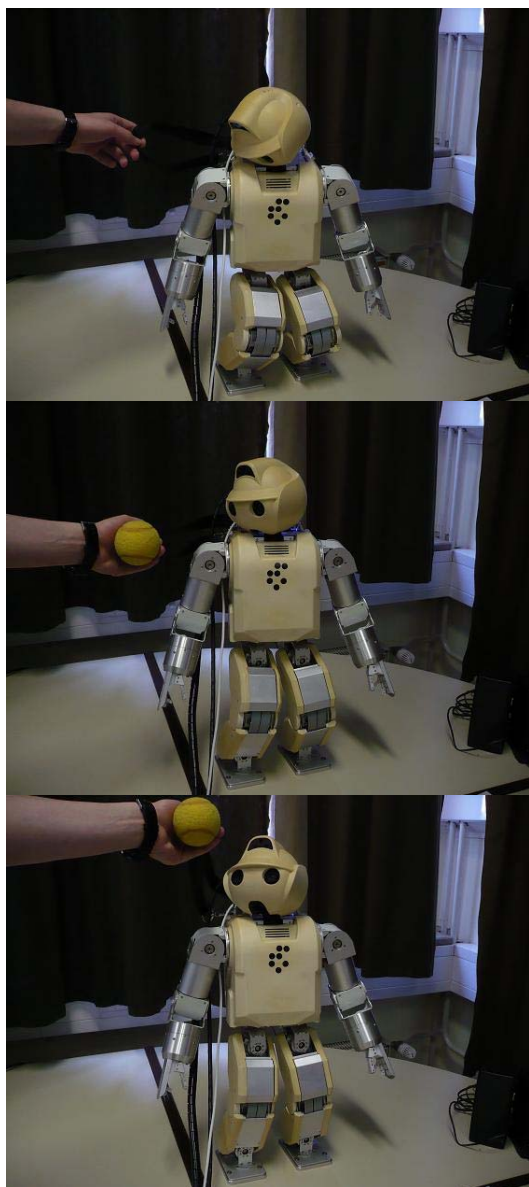
Napisanih je bilo že več različnih sistemov [7], s ciljem ustvariti standard za bodoče robotske aplikacije, kot naprimer Robot Operating System (ROS), Microsoft Robotics Developer Studio, Player Project, CLARAty, Marie, URBI, RUBIOS, Robotics suite, Webots, Carmen Toolkit [8], Orca in drugi. Tako je naprimer ROS (Robot Operating System) [9] odprtokodno programsko orodje za širok spekter robotov. Razvijajo in uporabljajo ga na Univerzi Stanford v Kaliforniji, na Tehnološkem inštitutu v Massachusettsu in na Tehnični univerzi v Minhnu. ROS, na primer, vsebuje programske ukaze za premik robotske roke, brez vnaprej podane informacije o zgradbi robota. Šele ko se ROS začne izvajati na robotovem računalniku, zahteva podatke o dolžini segmentov in obliki sklepov robota. To informacijo potem posreduje višje-nivojskim algoritmom, na primer tistemu za krmiljenje rok.

4 ŠTUDIJA EKSPERIMENTALNEGA DELA

Na osnovi predstavljenih pristopov smo zgradili programsko strukturo in uporabniški vmesnik za vodenje humanoidnega robota Hoap-3, ki naj bi poenostavil delo z robotom.

Hoap-3 je miniaturni humanoidni robot podjetja Fujitsu. Priložena programska oprema robota poseduje le osnovne funkcije, ki omogočajo komunikacijo z robotom. Mogoče je le brati vrednosti senzorjev in zapisovati vrednosti rotacijskih sklepov v regulatorje motorjev. Osnovni program deluje v realno-časovnem Linux okolju in omogoča periodično komunikacijo z robotom s taktom ene milisekunde. Poleg tega robot sprejme tudi nekatere nastavitvene ukaze, kot na primer vklop in izklop napajanja

posameznih motorjev, nastavitve maksimalne dovoljene rotacijske hitrosti sklepov in podobno.



Slika 1: Robot z vidom sledi predmetu.

Robot Hoap-3 je zasnovan tako, da ima uporabnik dostop do krmiljenja motorjev in odčitavanja vseh senzorjev. Motorje robota je moč krmiliti pozicijsko, hitrostno ali močnostno (z navori, preko krmiljenja električnega toka). Robot vsebuje vrsto senzorjev za detekcijo okolice. Na razpolago imamo merilnik pospeškov, žiroskop, infra-rdeči merilnik razdalje, senzorje sile v podplatih, senzor sile oprijema, mikrofoni in dve kameri. V robotu so nameščeni tudi zvočnik in dodatni digitalni vhodno-izhodni priključki, ter LED diode, nameščene okrog kamer.

Za potrebe vodenja je bilo potrebno razviti novo programsko orodje. Izdelali smo robotsko programsko

strukturo, ki omogoča abstrakcijo nizko-nivojskih ukazov, sočasno izvajanje različnih nalog ter komunikacijo z drugimi aplikacijami prek UDP / IP komunikacije, krmiljenje robota s tujega računalnika prek UDP / IP povezave ter vsebuje funkcije za izračun kinematike. Ta struktura ima funkcijo uporabniškega vmesnika, ker poenostavlja delo z robotom. Vmesnik je bil narejen z namenom, da bi uporabniki prek njega programirali krmiljenje robota.

S pomočjo naše izdelane robotske programske strukture smo implementirali več robotskih eksperimentov, kjer se je izkazala potreba po nadgradnji, oziroma izboljšavi sistema. Eden izmed teh eksperimentov je naloga sledenja predmetu z robotskim vidom humanoidnega robota, ki jo bomo na kratko opisali.

4.1 Naloga sledenja predmetu z robotskim vidom humanoidnega robota

Realizirali smo robotsko nalogo, ki robotu omogoča spremljanje opazovanega predmeta v prostoru (slika 1). Ker je vidno polje stereo para kamer robota majhno, je koristno, če lahko robot spremlja predmet z obračanjem glave. Aktivno vidno polje robota smo še povečali z dodatno mobilnostjo telesa robota. Gibanje telesa robota smo omejili na nekaj parametrov, ki smo jih lahko krmilili. Izbrali smo parametre, s katerimi smo nadzorovali obračanje trupa okrog navpične osi, pokrčenje nog in izvedbo predklona.

Za izvedbo robotske naloge smo uporabili algoritem, opisan v članku [5]. Ta algoritem so avtorji članka uporabili za krmiljenje sklepov oči, vratu, telesa in rok pri nalogi seganja roke v vidno polje robota. Opisali bomo bistvo algoritma in kako smo algoritem mi uporabili pri izvedbi naše naloge.

4.2 Lastnosti robotske naloge

Napisali smo algoritem vodenja z naslednjimi značilnostmi. Robot opazuje predmet s parom kamer, ki je nameščen v glavi robota. Robot s kamerami spremlja opazovani predmet in pri tem uporablja premik glave, nog in telesa. Uporablja vse aktuatorje hkrati, pri tem pa bolj premakne tiste aktuatorje, katerih vrednosti so blizu definirane poze mirovanja. Robot se vedno namesti v takšno lego, da so aktuatorji čim bližje vrednostim aktuatorjev v poziciji mirovanja-vzravnani pozici.

Poleg sklepov, ki premikajo glavo levo-desno in gor-dol ter sklepa v bokih, ki omogoča predklon, krmilimo tudi skrčenost nog in s tem spreminjamo višino glave. Krmilimo tudi zasuk telesa okrog navpične osi, kar dosežemo z ustreznimi usklajenimi premiki sklepov v nogah.

V regulatorju sklepov nog robota ne obravnavamo posamično, tako kot sklepe glave in bokov, ampak obravnavamo skupino sklepov kot en sam aktuator. Tako krmilimo počep robota s parametrom pokrčenosti nog. Za parameter počepa, s katerim počep krmilimo, smo si izbrali razdaljo med prvim in zadnjim sklepom noge. Obrat robota okrog navpične osi prav tako krmilimo, kot da bi bil en sam

aktuator, ker celotno kinematično strukturo nog krmilimo le z enim parametrom. Vrednosti sklepov nog robota, ki delujejo pri počepu in pri rotaciji telesa okrog navpične osi, nazadnje izračunamo prek izračuna inverzne kinematike. Iz navedenega sledi, da imamo neskončen nabor možnosti, s katerimi aktuatorji vplivamo na izvedbo zadane naloge. Problem rešujemo s postopki vodenja redundantnih robotskih sistemov [5].

5 ZAKLJUČEK

Dobra izbira strukture robotskega programa zelo pripomore k enostavnosti implementacije eksperimentalnih algoritmov na robotu. Naš sistem vsebuje različne elemente strukture robotskega programa in olajša delo z robotom, ni pa tako dovršen in vsestranski kot robotski operacijski sistem.

Ugotovili smo, da naj bo čimveč programske kode neodvisne od tipa robota, ki ga uporabljamo. Posamezni podsistemi morajo biti čim bolj samostojni in opravljati le eno jasno določeno nalogo.

Področje humanoidne robotike se hitro razvija. Končno postajajo humanoidni roboti dostopnejši večjemu številu raziskovalnih ustanov, ki ne gradijo lastnih robotov. Zato menimo, da bodo v prihodnje kvalitetne robotske programske strukture oziroma robotski operacijski sistemi in programska orodja za delo s humanoidnimi roboti zelo dobrodošli. Njihova kvaliteta utegne občutno vplivati na hitrost razvoja humanoidne robotike.

Načrtovanje robotske programske strukture je bolj umetnost kot znanost [5]. Cilj je narediti programiranje lažje in fleksibilnejše. Načrtovalci sistemov se odločajo glede na izkušnje in poznavanje programskih jezikov, ter glede na vrsto robotske naloge in robota, ki ga uporabljajo. Pri načrtovanju in izbiri primerne sistema, moramo vedeti, kako bomo sistem uporabljali, kakšne naloge bo robot opravljal, kako hitro se bodo naloge izvajale, kako bodo podatki v programu zapisani in kako bodo podatki predstavljeni uporabniku, kaj morajo uporabniki vedeti o robotu in ali bo programska struktura uporabna za več različnih nalog in več različnih robotov.

Robotske programske strukture omogočajo sočasno izvajanje več nalog in obnašanja. Omogočajo uporabo senzorjev in aktuatorjev, interpretacijo signalov in odločanje s pomočjo metod umetne inteligence. Omogočajo konceptualno osnovo, znotraj katere lahko pišemo svoj program. Razvito je bilo že mnogo različnih pristopov, ki jih lahko uporabljamo v različnih situacijah. Izbira prave strukture ni enostavna. Ta članek je predstavil nekaj kriterijev, na katere moramo biti pozorni pri izbiri. Izbira vsaj delno vpliva na vse nadaljnje delo z robotom.

Robot se mora sproti odzivati na zunanje dogodke, notranje dogodke in na minevanje časa. Zunanji dogodki so spremembe parametrov naloge in nepredvidljive motnje. Notranji dogodki so spremembe notranjih stanj robota, kot na primer, če roki robota med seboj trčita, če sklep robota doseže omejitev, če robot pade ali če tovor robota preseže maksimalno dovoljeno obremenitev robota.

Roboti so zelo kompleksni nestandardizirani stroji, zato morajo robotiki veliko časa porabiti za reševanje in implementacijo najbolj običajnih problemov, ki so jih mnogi že večkrat rešili, da dobi robot tisto funkcionalnost, ki je potrebna za izvedbo njihovega eksperimenta. S tem člankom skušamo vzbuditi zanimanje za uporabo robotskih operacijskih sistemov.

Literatura

- [1] E. Gat, Architecture, language, and non-compositional constraints, Aerospace Conference, Proceedings IEEE, 2003.
- [2] Reid G. Simmons, Structured Control for Autonomous Robots, IEEE Transactions on robotics and automation, vol. 10, no. 1, February 1994.
- [3] R. Simmons and D. Apfelbaum, A task description language for robot control, International Conference on Intelligent Robots and Systems, 1998. Proceedings, 1998 IEEE/RSJ, Vol. 3 (1998) , p. 1931-1937 vol.3.
- [4] Maja J Matarić, Integration of Representation Into Goal-Driven Behavior-Based Robots, in IEEE Transactions on Robotics and Automation, 8(3), Jun 1992, 304-312.
- [5] Bruno Siciliano, Oussama Khatib, Springer Handbook of Robotics, Springer, 2008.
- [6] M. Ingham, R. Rago, and B. C. Williams, A reactive model-based programming language for robotic space explorers, presented at the Int. Symp. Artif. Intell., Robotics, Automation Space (ISAIRAS-01), Montreal, QB, Canada, 2001.
- [7] B. Solvang, G. Sziebig, P. Korondi, Multilevel control of flexible manufacturing systems, Conference on Human System Interactions, Krakow, 2008.
- [8] M. Montemerlo, N. Roy, and S. Thrun, Perspectives on standardization in mobile robot programming: The Carnegie Mellon Navigation (CARMEN) Toolkit, in Proc. of the IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS), Las Vegas, Nevada, Oct. 2003, pp. 2436-2441.
- [9] Morgan Quigley, Brian Gerkey, Ken Conley, Josh Faust, Tully Foote, Jeremy Leibs, Eric Berger, Rob Wheeler, and Andrew Ng, ROS: an open-source Robot Operating System, International Conference on Robotics and Automation, 2009.

REGULACIJA TEŽIŠČA ROBOTA PRI VIZUALNO - MOTORIČNEM VODENJU

Blaž Hajdinjak, Jan Babič

Department of Automation, Biocybernetics and Robotics

Jozef Stefan Institute

Ljubljana Slovenia

e-mail: blaz.hajdinjak@ijs.si, jan.babic@ijs.si

ABSTRACT

Regulation of the center of mass in visuo-motor control.

We know various strategies toward teaching and controlling humanoid robots. Some refer to direct joint or tip control and others use more intuitive approach such as mimicking human motion in a certain task. This kind of robot control, where a robot is considered a tool, controlled by a human demonstrator, is called visuo-motion control. In this paper, we present an improved approach to overcome a problem of balance in visuo-motor control of humanoid robots.

Keywords: Center of mass, robot control, regulation, adaptation

Ključni pojmi: Težišče, vodenje robota, regulacija, adaptacija na okolje

1 UVOD

Učenje trajektorij robotskih gibov pri izvajanju različnih aplikacij je v osnovi precej kompleksna naloga, k kateri je mogoče pristopiti z zelo različnimi strategijami. Najbolj običajen pristop je vodenje robotskega vrha z inverzno Jakobijevo matriko, pri kateri dodatne omejitvene parametre priredimo kriterijem danega problema. Omenjeni pristop je do določene mere mogoče uporabljati tudi pri vodenju udov humanoidnih robotov, čeprav je tak način vodenja pri, denimo, vodenju spodnjih ekstremitet pri robotski hoji, nekoliko problematičen. Pri hoji namreč ni pomembno le zagotavljanje pravilne pozicije in orientacije vrhov robotskih manipulatorjev – v našem primeru stopal – ampak je zelo pomembno uravnoteženo in usklajeno delovanje vseh sklepov spodnjih ekstremitet, da o pomenu usklajenega sodelovanja ostalega telesa sploh ne govorimo. Zaradi tega je robotska hoja pogosto zelo toga, počasna in, za razliko od človeka, statično stabilna. Taka hoja na človeško hojo spominja zgolj po svoji funkcionalnosti. Če, torej, želimo doseči hojo, ki bi bila bolj podobna človeški, se je smiselno posluževati drugačnih strategij pristopa k učenju trajektorij robotske hoje. Ista ugotovitev velja tudi za ostale funkcionalne gibe humanoidnih robotov. Ena strategij, ki se že zaradi referenc našega končnega cilja – doseči gibanje, ki bi bilo kar najbolj podobno človeškemu –

ponuja kar sama je učenje robotskega gibanja s posnemanjem človeškega gibanja (ti. Visuo-motor control) [1 - 5]. Ta, na prvi pogled enostavna strategija pa ima nekaj ovir, ki jih je potrebno pri postavljanju sistema za tako učenje upoštevati. Eden teh problemov je neusklajenost položaja težišča določenega humanoidnega robota in človeka, ki robota uči trajektorij gibanja in v tem prispevku se bomo osredotočili prav na ta problem.

Človek in humanoidni robot kljub skladni sistemski zgradbi telesa, v osnovi nimata enakega položaja težišča. Razporeditev teže v robotu, je, namreč, precej drugačna od razporeditve teže pri človeku. Pri robotu je razporeditev mase odvisna od notranje arhitekture: Od razporeditve aktuatorjev, položaja akumulatorja, položaja procesorja in nenazadnje tudi od zunanje arhitekture, ki se odraža na razmerju velikosti trupa in posameznih sklepov. Le redki humanoidni roboti se po razporeditvi teže primerjajo z človekom. In to je tudi glavni razlog, da bi robot, ki bi neposredno posnemal človekovo hitro izgubil ravnotežje, oziroma, bi njegovo gibanje izgledalo precej nenaravno v smislu reguliranja lastne stabilnosti. Človeško gibanje je narava optimizirala tako, da je na vsaki točki dosežen kompromis med dosego cilja in ohranjanjem stabilnosti in prav zato se nam zdi človeško gibanje zelo naravno in logično. Tako izražena naravnost gibanja je brez dvoma cilj tudi v humanoidni robotiki.



Slika 1: Premična ploščad (platforma): s katero pri našem postopku reguliramo težišče človeka.

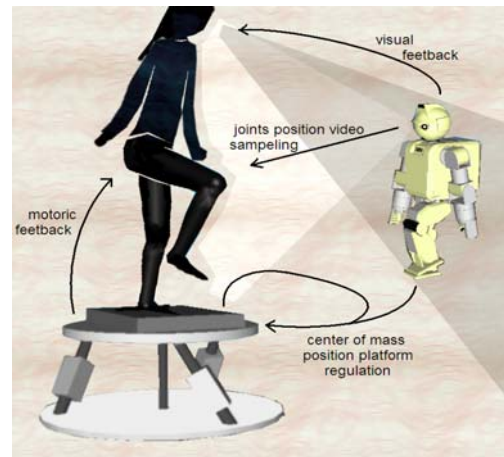
2 REGULACIJA NEUJEMANJA TEŽIŠČA ČLOVEKA IN ROBOTA

V nadaljevanju je opisan pristop krmiljenja robota z vizualno motoričnim vodenjem, pri čemer bomo najbolj osredotočeni k odpravljanju problema neujemanja položaja težišča robota in človeka pri vizualno-motoričnem vodenju robota.

2.1 vizualno motorično vodenje

Gibanje človeka zajema sistem kamer, ki informacijo o poziciji in orientaciji posameznih sklepov posreduje humanoidnemu robotu. Robot na podlagi teh informacij sledi gibanju človeka. Ker se poziciji težišč robota in človeka med gibanjem načeloma ne ujemata, je potrebno to neujemanje na nek način odpraviti. Prva strategija, ki se pri tem ponuja je samo-regulacija robota, pri kateri z merjenjem težišča robota le to z dodatnim gibanjem ohranjamo čim bolj v centru njegove stabilnosti, ali pa vsaj znotraj meja njegovega področja stabilnosti [1, 2]. Taka regulacija je, gledano iz stališča stabilnosti, povsem legitimna, po drugi strani pa biomehansko precej neargumentirana. Pri določenih gibih bi namreč upravičeno pričakovali, da bo robot, zaradi regulacije svojega težišča, pri uporabi takega sistema regulacije deloval precej nenaravno. Zato je v tem primeru bolje uporabiti nekoliko drugačen pristop. Namesto da bi z robotovo stabilnostjo regulirali robota, z njo »reguliramo« osebo, ki vodi robota. V našem primeru to storimo tako, da osebo postavimo na premično platformo (ravno podlago, ki jo je mogoče sukati v vseh treh smereh – sliki 1 in 2) in s signali, ki nosijo informacijo o položaju robotovega težišča, reguliramo njen nagib okrog posameznih osi. Če bi torej položaj težišča robota nakazal, da se bo robot prekucnil nazaj, bi se platforma premaknila tako, da bi prekucnila nazaj osebo, ki sto ji na njej, ta pa bi se s svojo naravno regulacijo stabilnosti na ta premik platforme odzvala tako, da bi z zasukom nekoliko naprej ta padec preprečila (slika 3). Ker robot posnema gibe te osebe, bi se naprej zasukal tudi sam, ter tako preprečil svoj padec. Očitno je, da je tak način regulacije precej bolj človeku-podoben, saj se robot na svojo nestabilnost odzove z regulacijo, ki posnema človekov odziv.

To regulacijo je mogoče še dodatno izboljšati, če pri opisanem procesu vodenja poznamo tudi trenutno pozicijo človekovega težišča. Na tak način namreč lahko vodimo robotovo težišče pri danem gibu tako, da je le-to v vsakem trenutku usklajeno s pozicijo človekovega težišča. Na ta način je doseženo gibanje, iz biomehanskega stališča človeku še bolj podobno. Človek, namreč, svojega težišča ne regulira vedno tako, da bi ga ne glede na nalogo ohranjal točno v centru svojega področja stabilnosti, pač pa vedno naredi določen kompromis med ohranjanjem stabilnosti in med ciljem zastavljene naloge.



Slika 2: Vizualno-Motorična kontrola robota. Humanoidni robot se pri tej metodi kontrole smatra kot orodje, s katerim upravlja človek. Robot posnema človekovo gibanje, sistem pa človeku daje povratno informacijo o robotovi stabilnosti.

2.2 merjenje trenutnega položaja težišča človeka

Merjenje trenutne pozicije težišča bo zagotovljeno tako, da bo na premično ploščad pritrjen še senzor sile (ti.: forceplate – Kistler 9281C), ki meri sile vzdolž in navore okrog vseh treh osi kartezičnega koordinatnega sistema. Pri tem nastane problem, saj je taka plošča običajno umerjena na nepremična tla, premična ploščad pa je, kot že ime pove – premična. Zato je potrebno dobljene sile ustrezno transformirati glede na zasuk ploščadi, tako da bodo izmerjene vrednosti položaja težišča vedno vezane na ravnino, ki je pravokotna na smer gravitacije. Če je, denimo, informacija o teži merjenega objekta pri senzorju na ravnih tleh skrita zgolj v sili, ki kaže pravokotno na ravnino merilne plošče (sili teže), postane pri zasukani plošči ta informacija kombinacija vsote projekcij sil v vseh treh kartezičnih smereh na premico nosilko smeri gravitacije. Tako transformacijo dosežemo s premultiplikacijo izmerjenih sil in navorov z inverzno rotacijsko matriko premične ploščadi:

$$F_{rot} = A \cdot F \quad (1)$$

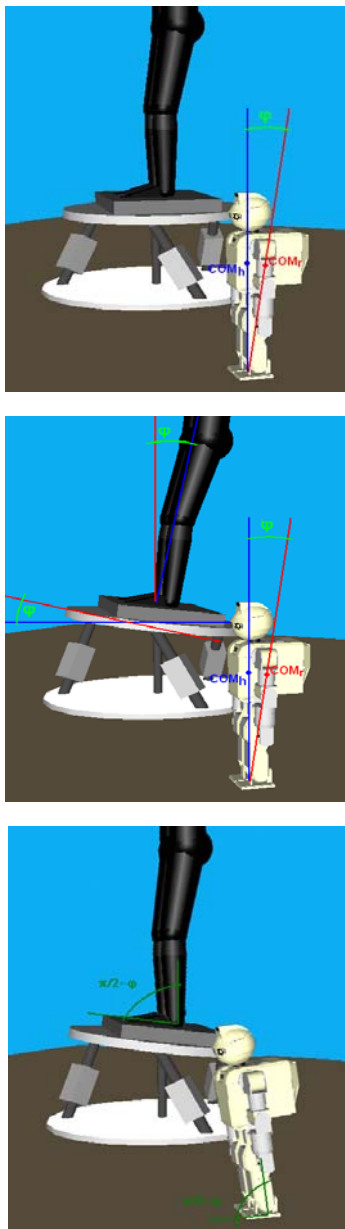
in

$$M_{rot} = A \cdot M \quad (2)$$

Na ta način zagotovimo, da je položaj težišča vezan na ravnino, ki je pravokotna na smer gravitacije.

Ko na ta način dobimo položaja obeh težišč – tako težišča robota, kot težišče človeka, ki robot vodi – lahko z njima vodimo nagib ploščadi. Kriterij tega vodenja je ohranjati enak položaj projekcije težišča na mnogokotnik področja stabilnosti glede na oddaljenost od centra tega

mnogokotnika, pri čemer bomo upoštevali tudi razmerje dimenzij človeka in robota. Tak način regulacije pravzaprav zagotavlja ujemanje kotov med premico, ki gre pravokotno skozi središče področja stabilnosti in med premico, ki gre skozi točko središča poligona in položaja težišča človeka oziroma robota (slika 3). Parameter razmerja dimenzij bo najverjetneje potrebno uglasti intuitivno.



Slika 3: Primer regulacije težišča – stoječ položaj. Ko človek stoji vzravnano, robot sledi njegovi vzravnani drži. Težišče robota je pri tem zaradi telesne arhitekture pomaknjeno nekoliko nazaj, medtem ko je težišče človeka centrirano (zgornja slika). Razlika v položajih težišč nagne premično ploščad nekoliko nazaj (srednja slika). Pri tem se človek avtomatsko upogne nekoliko naprej, da bi svoje težišče nekoliko stabiliziral. Temu premiku sledi robot, ki na ta način stabilizira svoje težišče. (spodnja slika). Ploščad se postavi v lego ravnovesnega težišča, v tej legi pa sta tako človek kot robot optimalno stabilizirana.

2.3 opis regulacijskega sistema premične ploščadi

V tej regulacijski zanki smo trenutno najbolj osredotočeni na optimalni način regulacije premične ploščadi. Naša teza je, da se bo sistem obnašal optimalno v primeru, ko bo ploščad regulirana tako, da bo odpravljala napako med relativnima položajema težišč robota in človeka glede na svoji področji stabilnosti. Če bo torej težišče robota, pri izvajanju določene naloge, na nekoliko različni relativni legi, kot težišče človeka, bo na vhodu regulatorja ploščadi neničelna vrednost in ploščad se bo nanjo odzvala tako, da bo z adekvatno spremembo nagiba človeka prisilila v to, da bo začutil enako nestabilnost kot robot. Ker se bo človek na to nestabilnost optimalno (če rečemo, da je optimalni odziv človeka podobna regulacija) odzval, mu bo robot sledil in na ta način optimalno odpravil svojo nestabilnost. Ker zajema projekcija področja stabilnosti dvodimenzionalni lik, bo ploščad odpravljala napako težišča po obeh dimenzijah plenarne ravnine.

Težišče človeka bomo po že opisanem postopku dobili s pomočjo ploščde za 3D merjenje sile, težišče robota (v našem primeru gre za robota Fujitsu HOAP III) pa bomo morali dobiti iz položaja težišč posameznih sklepov po enačbi:

$$\vec{r}_{tr} = \frac{\sum_{i=1}^n \vec{r}_i \cdot m_i}{\sum_{i=1}^n m_i} \quad (3)$$

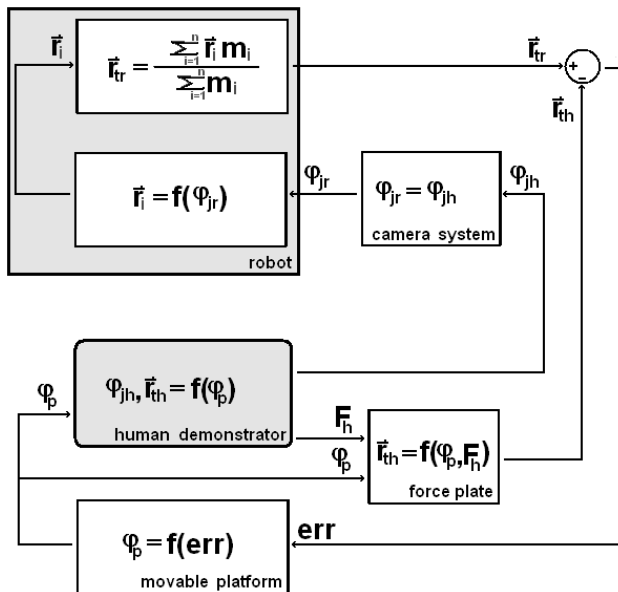
Pri čemer je n število sklepov, r z indeksom i je razdalja do težišča i -tega sklepa, m z indeksom i pa masa i -tega sklepa. Ker na ta način dobimo težišče robota zgolj relativno na samega sebe, je potrebno z uporabo žiroskopskega senzorja orientacije meriti tudi njegov trenutni nagib. Zgolj iz položaja težišča, dobljenega po enačbi 3, namreč ne moremo zaznati padanja, saj z robotom začne padati tudi njegova relativna projekcija področja stabilnost. Robot HOAP III sicer omogoča zajem položaja težišča tudi iz senzorjev, ki jih ima v stopalih, vendar je zajem iz teh senzorjev za naše potrebe najverjetneje nekoliko premalo natančen.

Z Razliko položajev težišč, oziroma napako (err) nato krmilimo premično ploščad. Velja torej:

$$\varphi_{ploščadi} = f(err) \quad (4)$$

Pri čemer bomo funkcijo ki preslika napako med položajema težišč v nagib oziroma premikanje (v tem primeru bomo uporabili $\Delta\varphi$) ploščadi določili intuitivno, v skladu z učinkovitostjo sledenja in z zahtevnostjo regulacijskega okolja za človeka, ki robota vodi. Tudi ta problem je namreč pri opisanem načinu vodenja robota potrebno vzeti v obzir.

Regulacijska shema celotne opisane regulacije je prikazana na sliki 4.



Slika 4: Regulacijski sistem vodenja: Humanoidni robot preko kamere sledi gibom človeka. Pri tem primerjamo položaja težišča robota in človeka. Robotovo težišče se izračuna iz položaja njegovih udov, človekovo težišče pa merimo s 3D senzorjem sile (force plate). Napaka med položajema težišč nato krmili premično ploščad ang. (movable platform), ki na ta način adaptira človekovo težišče.

3 ADAPTACIJA ČLOVEKA NA OKOLJE ZA VODENJE ROBOTA

Postaviti človeka v regulacijski sistem je pristop, ki zahteva nekaj pozornosti v smislu zmožnosti človekove adaptacije na tako okolje. Iz stališča človeka tak sistem vodenja robota ni bistveno različen od sistema vodenja kateregakoli drugega kompleksnega orodja, pri čemer nagib ploščadi predstavlja dodatno povratno informacijo o stanju tega orodja. Različne študije na ljudeh in na primatih [6] kažejo, da smo sposobni tudi kompleksnejša orodja hitro adaptirati tako, da se z orodjem poistovetimo do te mere, da ga smatramo kot podaljšek svojih udov. Izkušen upravljalec bagra tako med svojim delom ne razmišlja več o premikanju ročic ki vodijo lopato ampak neposredno o vodenju lopate, ročice pa ob tem premika podzavestno. Tako smemo odkrito pričakovati, da bo tudi človek, ki bo v našem sistemu vodil robota, po nekaj adaptacije postal osredotočen zgolj na vodenje robota, regulacijo težišča pa bo ob tem izvajal podzavestno. Jasno je, da bo za doseg tega stanja potrebna določena prilagoditvena doba, ki pa bo tem krajša, čim bolj bo naš regulacijski sistem uspešno izveden. V tem smislu bo prilagoditvena doba tudi eden od kriterijev optimizacije tega sistema.

4 UGOTOVITVE IN ZAKLJUČKI

V članku je opisan pristop k reševanju problema neujemanja težišča, eksperiment pa je trenutno še v fazi priprav, zato je še nemogoče podati kvantitativni rezultat. Začetni eksperimenti bodo vezani na vodenje robota pri opravljenih, ki zahtevajo uporabo telesa in rok in pri katerih so stopala fiksirana na mestu. Na ta način bomo dobili fiksirano področje stabilnosti, kar bo olajšalo računanje položaja težišča. Šele ko bo sistem uglašen bomo poskušali njegovo funkcionalnost nadgraditi tudi z vodenjem pri nalogah, ki zahtevajo dvigovanje stopal, denimo, pri stopicanju. Pričakujemo, da bodo eksperimenti pokazali, na naš način regulacije zaznavno izboljša gibanje robota v smislu stabilnosti in človeku-podobnosti, poleg tega pa adaptivno okolje za človeka, ki vodi robota. Pri taki regulaciji se namreč pričakuje precej bolj predvidljivo in precej manj agresivno gibanje premične ploščadi, ki regulira težišče človeka.

Reference

- [1] E. Oztop, J. Hale, J. Babic and M. Kawato, "Robots as complex tools for humans to control: Human visuo-motor learning for robot skill synthesis," International Conference on Intelligent Robots and Systems (IROS08 Workshop: Grasp and task learning by imitation), Nice, France, 2008
- [2] E. Oztop, J. Babic, J. Hale, G. Cheng, M. Kawato, "From Biologically Realistic Imitation to Robot Teaching Via Human Motor Learning," Neural Information Processing: 14th International Conference, ICONIP 2007, Kitakyushu, Japan, p. 214-221, Springer-Verlag, 2008
- [3] S. Obayashi, T. Suhara, K. Kawabe, T. Okauchi, J. Maeda, Y. Akine, H. Onoe and A. Iriki, "Functional Brain Mapping of Monkey Tool Use," Neuroimage, 14(4): p. 853-861, 2001
- [4] C.G. Atkeson and S. Schaal "Robot Learning From Demonstration", Machine learning: Proc. 14th Int. Conf. (ICML '97): p. 12-20, Nashville, Morgan Kaufmann, 1997
- [5] Y. Kuniyoshi, Y. Yozu, M. Inaba and H. Inoue, "From visuo-Motor Self Learning to Early Imitation - A Neural Architecture for Humanoid Learning," International Conference on Robotics & Automation, Taipei Taiwan: IEEE, 2003
- [6] S. Schaal, A. Ijspeert and A. Billard, "Computational Approaches to Motor Learning by Imitation," Philosophical Transactions of the Royal Society B: Biological Sciences 358 (1431), p. 537-547, London, 2003

RITMIČNO VODENJE NIHALA Z UPORABO NELINEARNEGA DINAMIČNEGA SISTEMA

Tadej Petrič, Andrej Gams, Leon Žlajpah
Odsek za avtomatiko, biokibernetiko in robotiko
Institute "Jožef Stefan"
Jamova 39, 1000 Ljubljana, Slovenia
Tel: +386 1 4773362;
e-mail: tadej.petrič@ijs.si

ABSTRACT

In the paper we address the problem of controlling an oscillation motion with a robot. As the object of task we have selected an object on a rope (mathematical pendulum). The first task is achieving the stable swinging. The second task is achieving the active dumping by moving the top end of the rope left and right. The third task is controlling the angle of pendulum. Although the human can quickly learn the requirement movement for each task, it is quite complex to perform it on robot. First of all, not all the states of the device are directly measurable, and secondly, synchronization with the object is required.

In the paper we describe the use of a two-layered imitation system with which we can perform all three tasks. The first layer of the system is used to adapt to the frequency of the input and is based on adaptive frequency oscillators. The second layer is used to output the desirable trajectory and is based on a set of differential equations. The ability of the imitation system to adapt to the frequency of the input signal and the ability to change the phase angle between the motion of the robot and motion of the object allowed us to control the swing.

1 UVOD

Ljudje stalno izvajamo različne periodične naloge, kot npr. hoja in različne naloge, ki jih opravljamo z rokami [6]. Skupno tem nalogam je, da se jih ljudje hitro naučijo s posnemanjem drugih ali s pomočjo enostavnih navodil [6]. Po drugi strani, pa je izvajanje takšnih nalog z roboti zahteven problem, ki zahteva izvajanje naprednih algoritmov in uporabo kompleksnih senzorskih sistemov [12]. Izvajanje periodičnih oziroma ritmičnih nalog (ang. rhythmic tasks) z robotom lahko izvedemo na različne načine [3,4]. Eden od teh je, da s pomočjo ustreznega vodenja ponovimo v naprej določen vzorec giba [10], ki je lahko posnetek demonstracije ali npr. rezultat matematične izpeljave [2]. Drug način je, da se trajektorija določa v odvisnosti od trenutnega stanja naprave [4]. Pri uporabi te metode mora imitacijski sistem poleg generiranja trajektorije, ujeti tudi pravilno frekvenco in fazo.

Robotsko vodenje različnih nihala je že bilo predmet raziskav [1]. Kjer je prikazano, da lahko togo

nihalo vodimo z uporabo Matsuoka oscilatorja in merjenjem sile na vrhu robota [5]. V tem prispevku smo se osredotočili na vodenje matematičnega nihala z merjenjem kota med nihalom in njegovo ravnovesno lego. Rezultati simulacij so pokazali, da predlagani sistem vodenja modulira frekvenco giba tako, da ujame pravilno fazo med stanjem nihala in robotom, ter s tem sinhronizira gibanje robota. Sistem tako omogoča izvedbo vseh treh zadanih nalog: stabilno nihanje, aktivno dušenje s premikanjem vrha robota levo in desno, ter regulacijo kota nihanja.

Za vodenje nihala je predlagana uporaba dvo-nivojskega imitacijskega sistema, ki temelji na nelinearnem dinamičnem sistemu [3,4]. Naloga prvega nivoja je prilagajanje oz. sinhronizacija frekvence s pomočjo nelinearnih frekvenčnih oscilatorjev. Naloga drugega nivoja je moduliranje želene trajektorije. Sestavljajo ga diferencialne enačbe.

V prispevku je opisana uporaba sistema za posnemanje periodičnih nalog. V drugem poglavju je opisan sistem za posnemanje gibanja, ter oba njegova podnivoja. V tretjem poglavju je predstavljen eksperiment ter rezultati vodenja. Zaključki so podani v četrtem poglavju.

2 IMITACIJSKI SISTEM

To poglavje opisuje imitacijski sistem, ki je bil podrobno opisan v [3]. Strukturo imitacijskega sistema sestavljata dva nivoja. Prvo plast sestavlja kanonični dinamični sistem (set diferencialnih enačb). Naloga tega sistema je izločiti osnovno frekvenco vhodnega signala. V tem prispevku je vhodni signal enak kotu nihala. V primeru, da uporabimo imitacijski sistem za učenje trajektorije kot v [3], je vhodni signal enak želeni trajektoriji. Drugi nivo sistema je izhodni dinamični sistem, katerega izhod je zelena trajektorija robota. Diferencialne enačbe sistema so drugega reda, kar zagotavlja gladkost izhodnih trajektorij.

Delovanje imitacijskega sistema je sledeče. Kanonični dinamični sistem izloči osnovno frekvenco Ω , in fazo ϕ . Ti veličini sta povezani z izhodnim dinamičnim sistemom. Izhodni dinamični sistem generira zeleno trajektorijo s setom diferencialnih enačb in vektorjem uteži $w = [w_1, \dots, w_n]$. Vektor uteži je pomnožen z preddefiniranimi jedrnimi funkcijami ψ , tako da je izhod enak

želeni trajektoriji. Trajektorija se sinhronizira s frekvenco Ω preko sidranja jedrnih funkcij na fazo ϕ . Naslednje poglavje podrobno opisuje obe posamezni plasti.

2.1 Kanonični dinamični sistem

Kanonični dinamični sistem ima dve pomembni nalogi. Kot prvo mora izločiti osnovno frekvenco Ω iz vhodnega signala in kot drugo, mora zagotoviti stabilno ciklično obnašanje sistema, tako da izloči še fazo ϕ .

Osnova kanoničnega dinamičnega sistema je nabor frekvenčno spremenljivih faznih oscilatorjev [7], ki so združeni v povratno zanko [8], tako da je izločena frekvenca natančno določena. Povratnozanka struktura M adaptivnih frekvenčnih faznih oscilatorjev je podana z

$$\dot{\phi}_i = \omega_i - Ke(t)\sin(\phi_i), \quad (1)$$

$$\dot{\omega}_i = -Ke(t)\sin(\phi_i), \quad (2)$$

$$e(t) = y_{in}(t) - \hat{y}(t), \quad (3)$$

$$\hat{y}(t) = \sum_{i=1}^M \alpha_i \cos(\phi_i), \quad (4)$$

$$\dot{\alpha}_i = \eta \cos(\phi_i)e(t), \quad (5)$$

kjer je K konstanta medsebojne povezanosti, ϕ_i , $i = 1, \dots, M$, faza i -tega oscilatorja, $e(t)$ vhod v oscilatorje, y_{in} je vhodni signal, M je število oscilatorjev, α_i je amplituda i -tega oscilatorja, in η je konstanta učenja.

Ker so vsi oscilatorji v negativni povratni zanki z enakim vhodnim signalom se njihova razlika približuje ničli, ko se vsota posameznih faznih oscilatorjev približuje vhodnemu signalu. Ta struktura na nek način izvede delno Fourierjevo analizo. Lahko se adaptira na množico različnih frekvenčnih komponent vhodnega signala. Število frekvenčnih komponent je odvisno od števila uporabljenih frekvenčnih oscilatorjev. Določevanje oscilatorja, ki nosi osnovno frekvenco je izredno pomembno, saj predstavlja osnovo za pravilen odziv drugega nivoja (izhodnega dinamičnega sistema).

Kanonični dinamični sistem se lahko uporabi tudi kot samostojen imitacijski sistem [8]. Z dodanim izhodnim dinamičnim sistemom zagotovimo večjo robustnost sistema na različne perturbacije. Prav tako izhodni dinamični sistem omogoča gladko moduliranje izhodnega signala. To dvoje je pomembno, kadar imamo opravka z hitrimi dinamičnimi nalogami.

2.2 Izhodni dinamični sistem

Izhodni dinamični sistem se uporablja za moduliranje in učenje zelene trajektorije. Vhoda v sistem sta frekvenca Ω in faza ϕ . V tem prispevku je tretji vhod v sistem konstanten vektor uteži w , ker je ta nivo uporabljen samo za moduliranje izhodnega signala. Vektor w skupaj z drugim redom diferencialnih enačb in jedrnimi funkcijami določa obliko trajektorije. Enačbe veljajo za eno prostostno

stopnjo. Za več prostostnih stopenj se algoritem poveže paralelno. Obnašanje dinamičnega sistema je podano z

$$\dot{z} = \Omega \left(\alpha_z (\beta_z (g - y) - z) + \frac{\sum_{i=1}^N \Psi_i w_i r}{\sum_{i=1}^N \Psi_i} \right), \quad (6)$$

$$\dot{y} = \Omega z. \quad (7)$$

Kjer je Ω frekvenca dobljena iz kanoničnega dinamičnega sistema, $\alpha_z = 8$ in $\beta_z = 2$ sta pozitivni konstanti, N je število Gaussovih jedrnih funkcij. Amplituda trajektorije je podana s parametru r . Nelinearna odvisnost določa želeno obliko trajektorije z množenjem jedrnih funkcij in pred-definiranega vektorja uteži. Jedrne funkcije so podane z

$$\Psi_i = \exp(h(\cos(\phi - c_i) - 1)), \quad (8)$$

in imajo Gaussovo obliko. Parametra jedrnih funkcij sta h , ki določa njihovo širino in c_i , ki določa njihovo porazdeljenost preko ene periode. Parameter c_i je enakomerno porazdeljen med 0 in 2π v N korakih.

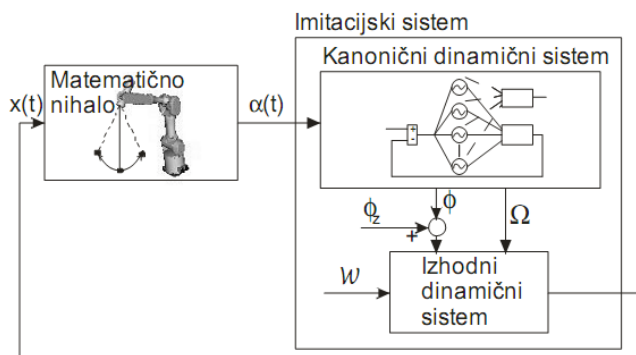
Predlagani dvo-nivojski imitacijski sistem se lahko uporablja tako v načinu moduliranja pred-definirane trajektorije, kot tudi v načinu učenja nove trajektorije. V tem prispevku je sistem uporabljen samo za moduliranje trajektorije. V [3,4] je prikazan tudi način učenja, kar dosežemo z inkrementalnim spreminjanjem vektorja uteži w v vsakem časovnem koraku z inkrementalno lokalno utežno regresijo (*ang.* incremental locally weighted regression) [9].

3 VODENJE NIHALA

Pri ritmičnem vodenju nihala z robotom ločimo tri naloge: stabilno nihanje, aktivno dušenje s premikanjem vrha robota levo in desno, ter regulacijo kota nihanja. Vse naloge so izvedene s pomočjo imitacijskega sistema. Uporabili smo simulacijo nihala na vrvi, ki se lahko prosto giblje v vseh smereh v okviru omejitev. Omejitev gibanja ter frekvenco nihanja določa dolžina vrvi nihala.

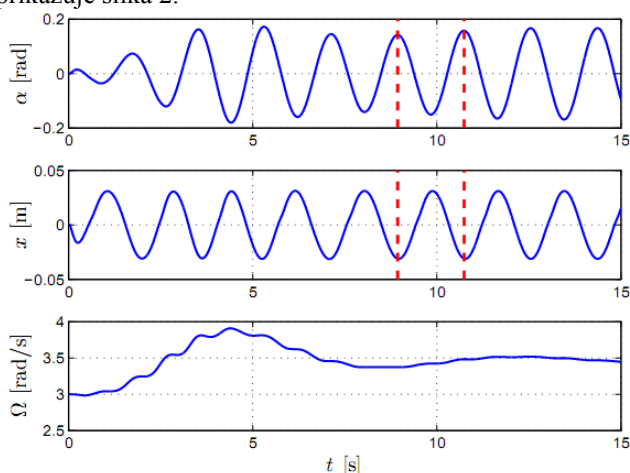
3.1 Stabilno nihanje in aktivno dušenje

Prva naloga je doseči stabilno nihanje s konstantno amplitudo. Za vodenje je uporabljen imitacijski sistem, ki temelji na adaptivnimi oscilatorji. V osnovi si je večina shem vodenja periodičnih nalog podobna. Izhod iz imitacijskega sistema vezemo na vhod sklepa mehanske strukture. Ta signal je lahko trajektorija vrha robota, navor, kot zasuka, ali kaj drugega. Vhod v imitacijski sistem je merjeni signal, ki je zopet lahko kot, navor ali kaj drugega. Pomembno je le, da je vhodni signal periodičen in odvisen od stanja naprave. S tem signalom je sistem sklopljen. Za vodenje matamatičnega nihala z imitacijskim sistemom smo uporabili shemo prikazano na Sliki 1.



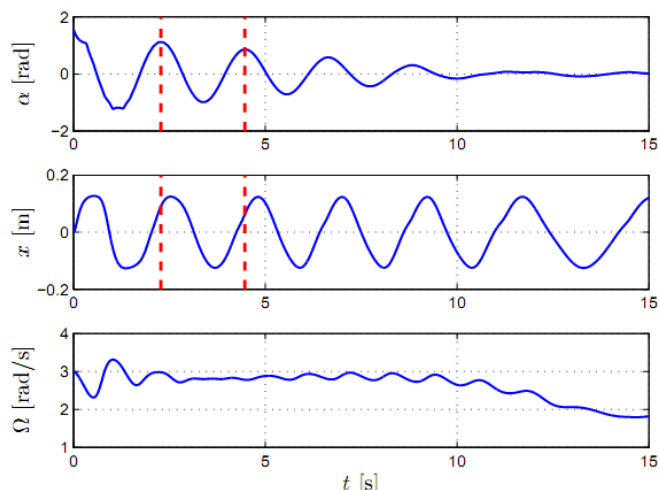
Slika 1: Blokovna shema vodenja matematičnega nihala

Uspešno izvedbo naloge ob pravilni shemi vodenja določajo parametri oscilatorja in oblika izhodnega signala. Vektor uteži w , ki predstavlja obliko izhodnega signala, je izbran tako, da je izhod sinusne oblike z konstantno amplitudo. Izvedba poizkusa stabilnega nihanja poteka tako, da je začetno stanje nihala v mirovni legi. Parameter, ki določa ali bo sistem zanihal ali ne, je konstanta faznega zamika ϕ_z . S skrbno izbranim parametrom ϕ_z lahko nihalo zanihamo ali zadušimo. Z izbranim faznim zamikom ($\phi_z = \pi$), dosežemo nihanje z konstantno amplitudo, kar prikazuje slika 2.



Slika 2: Rezultati simulacije stabilnega nihanja. Odmik nihala od začetne lege (zgoraj), gibanje vrha robota v smeri x (na sredini) in frekvenca oscilatorja Ω (spodaj). Parametri: $l = 0.7m$, $K = 8$, začetni kot $\alpha = 0$, $\phi_z = \pi$

Iz rezultatov je razvidno, da se nihanje hitro vzpostavi. Frekvenca oscilatorja se sinhronizira z nihalom, to pa začne nihati s konstantno amplitudo. Ohranjanje faznega zamika je ponazorjeno z črtkanimi navpičnicami na zgornjem in srednjem grafu. Druga naloga je ustavljanje nihala oziroma aktivno dušenje s premikanjem zgornjega konca vrvi (vrha robota) levo in desno. Za izvedbo naloge je uporabljena enaka shema vodenja s spremenjenim faznim zamikom ϕ_z . Rezultate vodenja s konstantno amplitudo vrha robota in faznim zamikom ($\phi_z = -\pi/4$) podaja slika 3.

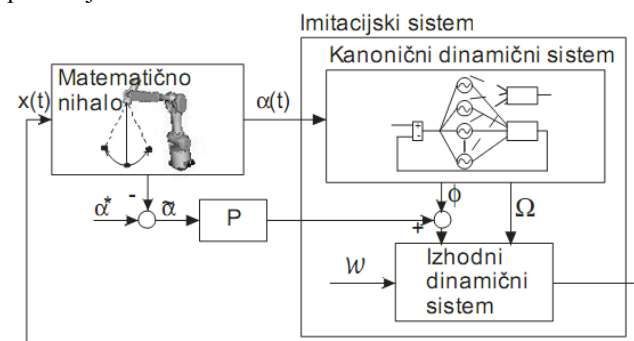


Slika 3: Rezultati simulacije aktivnega dušenja nihanja. Odmik nihala od začetne lege (zgoraj), gibanje vrha robota v smeri x (na sredini) in frekvenca oscilatorja Ω (spodaj). Parametri: $l = 0.7m$, $K = 8$, začetni kot $\alpha = 0$, $\phi_z = -\pi/4$

Samo s spremembo faznega zamika ϕ_z , ob enakih parametrih sistema, smo dosegli aktivno dušenje nihala. Nihalo se ustavi že po 4 nihajih oziroma slabih 10 s. Rezultat se lahko še izboljša z uporabo spremenljive velikosti amplitude giba, ter natančnejšo nastavitvijo parametra ϕ_z . Oba preizkusa vodenja kažeta na možnost krmiljenja odmika nihala s spremembo faznega kota ϕ_z .

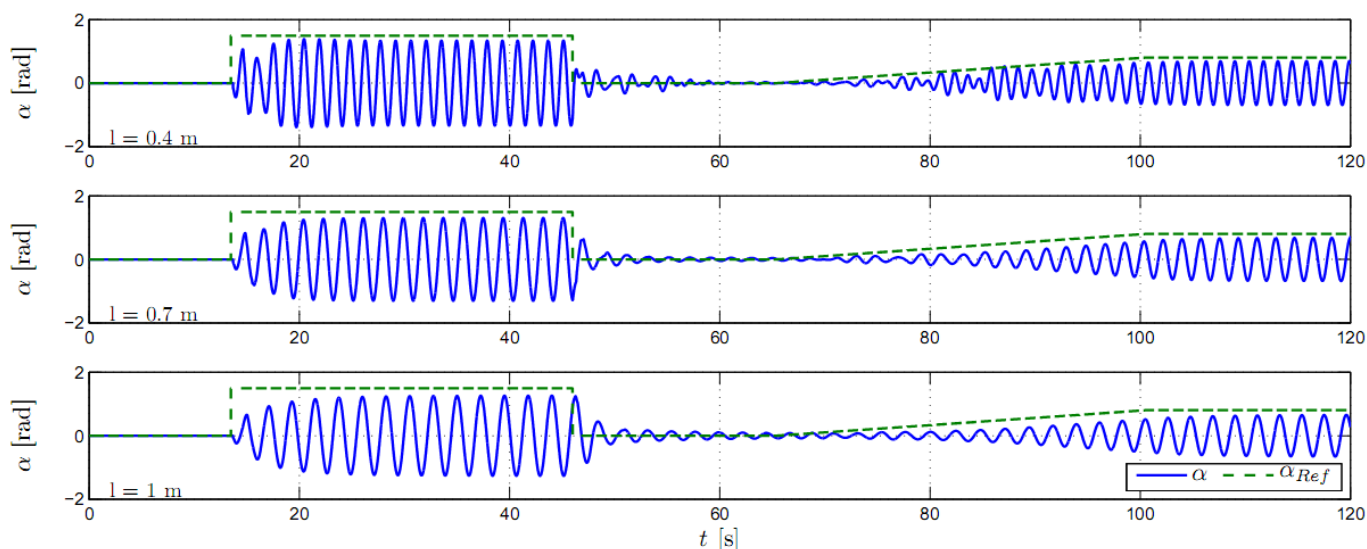
3.3 Regulirano vodenje nihala

Tretja naloga je regulacija kota nihanja. Krmilna veličina s katero reguliramo kot nihala, je fazni zamik ϕ_z , kot to prikazuje slika 4.



Slika 4: Blokovna shema regulacije matematičnega nihala

Rezultate vodenja podaja slika 5, kjer je na zgornjem grafu prikazan rezultat vodenja nihala z dolžino 0.4 m, na srednjem grafu je rezultat vodenja nihala z dolžino 0.7 m in na spodnjem grafu je rezultat vodenja nihala z dolžino 1 m. Razvidno je, da je sistem, poleg tega da sledi zelenemu kotu, tudi izredno robusten. Saj se ob nespremenjenih parametrih sistema in spremembi dolžine nihala, uspešno sinhronizira in sledi zelenemu kotu odmika.



Slika 5: Rezultati simulacije regulacije kota nihanja identičnega sistema z identičnimi parametri in različno dolgimi nihali. Dolžina nihala $l = 0.4\text{m}$ (zgoraj), dolžina nihala $l = 0.7\text{m}$ (na sredini) in dolžina nihala $l = 1\text{m}$ (spodaj).

4 ZAKLJUČEK

V prispevku smo pokazali način vodenja ritmičnih robotskih nalog z uporabo imitacijskega sistema. Enostavnejše naloge se vse vodi na enak način (s sklopitvijo procesa in oscilatorja). Težji del naloge predstavlja nastavitve parametrov oscilatorja in definiranje želene trajektorije robota. Še pred tem pa tudi izbira merjene in krmilne veličine. Za vodenje nihala na vrvici nastavitve parametrov oscilatorja in definiranje trajektorije vrha robota ni zahtevna. Merjeno veličino nam določaja stojna oprema. Pomembno je le, da je merjena veličina ena od periodično spremenljivih veličin odvisnih od procesa.

Z predefinirano obliko giba in merjenjem kota odmika nihala od ravnovesne lege, je sistem pravilno razbral frekvenco in fazo vodenja. Z spreminjanjem parametra faznega zamika ϕ_z , pa sistem omogoča, da lahko dosežemo različna stanja: mirovanje, stabilno nihanje ter dušenje. Kar omogoča, da se fazni zamik ϕ_z lahko uporabi kot krmilno veličino. Izkaže se, da je takšno vodenje nihala izredno robustno. Ob nespremenjenimi parametri sistema in spremembi dolžine nihala, ki rezultira v počasnejšem oziroma hitrejšem nihanju, se sistem uspešno sinhronizira in sledi referenčnemu kotu odmika.

Literatura

[1] Gams A., L. Lahajnar, J. Lenarčič, L. Žlajpah. 2007. Kombinirano ritmično in aperiodično vodenje nihala z uporabo nevronskega oscilatorja. Proceedings ERK 2007, Portorož (Slovenia), September 24-26, pp. B:156-159

[2] Gams, A., L. Žlajpah, J. Lenarčič. 2007. Imitating human acceleration of a gyroscopic device, Robotica, Vol. 25, pp. 501-509

[3] Gams, A., A.J. Ijspeert, S. Schaal, J. Lenarčič. 2009. On-line learning and modulation of periodic

movements with nonlinear dynamical systems, Autonomous Robots, Vol. 27(1), pp. 3-23

[4] Gams A., T. Petrič, L. Žlajpah. 2009. Controlling yo-yo and gyroscopic device with nonlinear dynamic systems, Proceedings of the 17th International Workshop on Robotics in Alpe-Adria-Danube Region, RAAD, Brasov (Romania), May 25-28

[5] Omrčen D., B. Nemec. 2002. Measuring knee movement using an industrial robot-gravity compensation for the automatic gripper, Strojniški vestnik, Vol. 48(2), pp. 87-98

[6] Petrič, T., A. Gams, L. Žlajpah. 2009. Modelling and control strategy of robotic Powerball, Proceedings of the 18th International Workshop on Robotics in Alpe-Adria-Danube Region, RAAD, Brasov (Romania), May 25-27

[7] Righetti L., J. Buchli, A. Ijspeert. 2006. Dynamic hebbian learning in adaptive frequency oscillators, Physica, Vol. D 216(2), pp. 269-281

[8] Righetti L., A. Ijspeert. 2006. Programmable central pattern generators: an application to biped locomotion control, Proceedings of the 2006 IEEE International Conference on Robotics and Automation, pp. 1585-1590

[9] Shall S., C. G. Atkeson. 1998. Constructive Incremental Learning From Only Local Information, Neural Computation, Vol. 10(8), pp. 2047-2084

[10] Schaal, S. 1999. Is Imitation Learning The Route to Humanoid Robots?, Trends in cognitive sciences, Vol3(6), pp. 233-248

[11] Žlajpah L. 2003. Modeling and control strategy of robotic yo-yo, Proceedings of the 12th International Workshop on Robotics in Alpe-Adria-Danube Region, RAAD, Cassino (Italy), May 7-10

[12] Žlajpah L. 2005. Robotic yo-yo: modeling and control strategies, Robotica, Vol 24(2), pp. 211-220

POSPLOŠEVANJE PERIODIČNIH GIBANJ ZAPISANIH Z NELINEARNIMI DINAMIČNIMI SISTEMI

Andrej Gams, Aleš Ude

Odsek za avtomatiko, biokibernetiko in robotiko

Institut "Jožef Stefan"

Jamova cesta 39, 1000 Ljubljana, Slovenia

e-mail: andrej.gams@ijs.si

ABSTRACT

Acquiring new sensorimotor knowledge by imitation is a promising paradigm for robot learning. For effective use, however, general action learning should not be limited to direct replication of movements obtained during training, but must enable also the generation of actions in new situations. This paper describes a methodology that enables the generalization of the accumulated example movements and synthesizes new periodic movements with statistical methods. Nonlinear dynamic systems are employed as an underlying motor representation. The proposed approach enables the generation of a wide range of policies without the need of an expert to appropriately modify the underlying representation for task specific features. A demonstration on a CB-I humanoid robot is presented.

1 UVOD

Učenje robotskega gibanja za reševanje izbrane naloge ne glede na trenutno stanje okolja je zahtevno, kajti potencialni proctor, ki ga mora robot pri tem preiskati je potencialno ogromen [1]. Njegova velikost je odvisna od števila robotovih stopenj prostosti in predmetov, ki so vključeni v samo nalogo.

Reševanje takih problemov zahteva vodenje procesa isknja, pri tem pa je eden najbolj uspešnih pristopov učenje z demonstracijo [2]. Z neposrednim posnemanjem so raziskovalci rešili že več nalog kompleksnega premikanja, recimo ples [3], oziroma vodenje žiroskopske naprave [4].

Neposredno posnemanje gibanja je neprimerno za naloge, ki zahtevajo manipulacijo z objekti, kajti take naloge morajo biti prilagojene trenutnemu stanju okolja, npr. lokaciji objekta, ki ga hočemo premikati. Dokaj malo verjetno je, da bi lahko prav vse primerne gibe posneli vnaprej.

Omeniti je treba tudi, da lahko začetno znanje pridobimo tudi na druge načine, ki ne vključujejo imitacije oz. kinestetične demonstracije. Obstajajo tudi možnosti avtonomnega raziskovanja [5].

Glavni doprinos tega prispevka je metodologija

posploševanje dostopnega znanja, ne glede na to, kako je bilo to znanje pridobljeno. Metode, ki modificirajo eno samo naučeno trajektorijo, npr. z učenjem z okrepitvijo [6], sicer lahko dosežejo podobne rezultate, vendar je pri tem treba upoštevati dolgotrajen proces učenja. To pomeni, da metode ne moremo uporabiti v sprotnem (on-line) načinu.

Eden od načinov zapisa trajektorij je z uporabo nelinearnih dinamičnih sistemov [7], poimenovanih Dynamic Movement Primitives oz. DMP. Osnovani so na diferencialnih enačbah drugega reda, s katerimi zapišejo karakteristike želenega giba. Podane so bile metode za aperiodične in periodične gibe, v tem prispevku pa smo se osredotočili samo na periodično gibanje. Ena glavnih prednosti tega pristopa je neodvisnost izvajanja giba od časa, kar omogoča ustavljanje in ponavljanje gibov ter možnost odzivanja na motnje.

Glavni anmen tega prispevka je metodologija posploševanja naučenih periodičnih trajektorij za nove situacije. Vsaki naučeni trajektoriji se asociira parametre, ki opisujejo njene značilnosti, tipično cilj, in ki služijo kot povpraševanje v knjižnico naučenih gibov. Metoda je andgradnja metode za posploševanje aperiodičnih gibov [8].

Osnova zapisu gibov so nelinearni dinamični sistemi oz. DMPji. DMP-je lahko uporabimo za modificiranje naučenih gibov tudi drugače [9], vendar to zahteva eksperta, da zapis priredi. Tak pristop je za avtonomno robotsko delovanje neprimeren. Pristop opisan v tem prispevku omogoča posploševanje naučenih periodičnih gibov specifikam naloge, ki jih podamo na povsem naraven in intuitiven način. Ker je končni rezultat tudi DMP, lahko uporabimo tudi vse alanitične spremembe, ki jih tak zapis omogoča.

V nadaljevanju je v drugem poglavju na kratko predstavljen pristop z DMPji, v tretjem poglavju metoda posploševanja gibanja, v četrtem pa rezultati naloge posploševanja bobnanja na humanoidnem robotu CB-i.

2 Posnemanje periodičnega gibanja z DMPji

V tem poglavju je na kratko predstavljen pristop posnemanja periodičnega gibanja z nelinearnimi dinamični sistemi oz. DMPji. Ijspeert in drugi [7] so predlagali zapis gibanja s sistemom nelinearnih enačb, ki imajo dobro definirano dinamiko atraktorja.

Posamezni stopnji prostoti gibanja v notranjih koordinatah ali pa v prostoru nalog naslednji sistemo diferencialnih enačb poda osnovo za določitev specifig gibanja.

$$\tau \dot{z} = \alpha_z (\beta_z (g - y) - z) \quad (1)$$

$$\tau \dot{y} = z \quad (2)$$

Pri tem modajo biti parametri α_z, β_z in $\tau > 0$ primerno določeni recimo $\alpha_z = 4\beta_z$ (v tem primeru je sistem celo kritično dušen) in sistem ima samo en atraktor pri $y = g, z = 0$.

Enačbam (1) – (2), ki opisujeta zelo omejen nabor gibanja, dodamo nelinearni člen, s katerim preoblikujemo rezultirajočo trajektorijo. Za periodično gibanje dodamo

$$f(\phi) = \frac{\sum_{i=1}^N w_i \Psi_i(\phi)}{\sum_{i=1}^N \Psi_i(\phi)}, \quad \Psi_i(\phi) = \exp(h_i (\cos(\phi - c_i) - 1)) \quad (3)$$

kjer je r parameter izhodne amplitude in $h_i > 0$. Z vpeljavo fazne spremenljivke ϕ se znebimo direktne odvisnosti od časa, ϕ pa podaja fazni oscilator

$$\dot{\phi} = \Omega. \quad (4)$$

Pri tem je Ω frekvenca oscilatorja ter $\tau = 1/\Omega$. S tem se enačbi (1) – (2) spremenita v

$$\tau \dot{z} = \alpha_z (\beta_z (g - y) - z) + f(\phi) \quad (5)$$

$$\tau \dot{y} = z, \quad (6)$$

s katerima pa lahko zapišemo poljubno periodično trajektorijo. Enačbi (5) in (6) opisujeta en DMP.

3 Posploševanje trajektorij z DMPji

V tem poglavju je opisan algoritem posploševanja periodičnih gibanj iz knjižnice naučenih gibov. Najprej je predstavljena metoda učenja ene same trajektorije, kasneje pa posploševanje iz večih trajektorij.

3.1 Učenje ene trajektorije.

Poljubno gladko trajektorijo lahko ponovimo s spreminjanje parametrov w_i iz enačbe (3). Pri tem začnemo iz posnetih demonstracijskih pozicij, hitrosti in pospeškov $\{y_d(t_j), \dot{y}_d(t_j), \ddot{y}_d(t_j)\}$, kjer je t_j čas vzorčenja.

Podatke dobimo s kinestetičnim vodenjem robota ali pa človeško demonstracijo. Enačbi (5) – (6) zapišemo v eno samo enačbo drugega reda

$$\tau^2 \ddot{y} + \alpha_z \tau \dot{y} - \alpha_z \beta_z (g - y) = f, \quad (7)$$

kjer je f definiran z enačbo (3). Z zapisom

$$F(t_j) = \tau^2 \ddot{y}_d(t_j) + \alpha_z \tau \dot{y}_d(t_j) - \alpha_z \beta_z (g - y_d(t_j)),$$

$$f = \begin{bmatrix} F(t_1) \\ \dots \\ F(t_T) \end{bmatrix}, \quad w = \begin{bmatrix} w_1 \\ \dots \\ w_N \end{bmatrix}, \quad (8)$$

dobimo naslednji set linearnih enačb:

$$Xw = f, \quad (9)$$

ki ga moramo rešiti, da lahko ocenimo DMP, ki opisuje želeno gibanje. X podaja

$$X = r \begin{bmatrix} \frac{\Psi_1(\phi_1)}{\sum_{i=1}^N \Psi_i(\phi_1)} & \dots & \frac{\Psi_N(\phi_1)}{\sum_{i=1}^N \Psi_i(\phi_1)} \\ \dots & \dots & \dots \\ \frac{\Psi_1(\phi_T)}{\sum_{i=1}^N \Psi_i(\phi_T)} & \dots & \frac{\Psi_N(\phi_T)}{\sum_{i=1}^N \Psi_i(\phi_T)} \end{bmatrix}, \quad (10)$$

x_i in ϕ_i pa dobimo z integracijo enačbe (4). Frekvenco gibanja moramo poznati v vsakem trenutku. Za to smo uporabili frekvenčno prilagodljive oscilatorje [10], [11]. Namesto (4) uporabimo

$$\dot{\phi}_i = \Omega_i - ke(t) \sin(\phi), \quad (11)$$

$$\dot{\Omega}_i = -Ke(t) \sin(\phi_i), \quad (12)$$

$$\dot{\alpha}_i = \eta \cos(\phi_i) e(t), \quad (13)$$

kjer je $e(t) = y_d(t) - \hat{y}(t)$, $\hat{y}(t) = \sum_{i=1}^L \alpha_i \cos(\phi_i)$. Če je $e(t) = 0$, sistem (11) – (13) postane enak (4). Z opisanim sistemom lahko določimo frekvenco vhodnega signala [10]. S to frekvenco lahko ocenimo parametre oblike w_i z inkrementalno metodo najmanjših kvadratov

$$P_j = P_{j-1} - \frac{P_{j-1} x_j x_j^T P_{j-1}}{1 + x_j^T P_{j-1} x_j} \quad (14)$$

$$w_j = w_{j-1} + (f_j - x_j^T w_{j-1}) P_j x_j \quad (15)$$

kjer je $P_0 = \sigma^2 I$, $w_0 = 0$, $f_j = F(t_j)$, x_j je M dimenzonalen stolpični vektor s primerno vrstico matrike X in končne uteži so $w = w_T$. Frekvenco

ocenjujemo sproti z enačbami (11) – (13). Trenutno vrednost uporabimo za izračun $F(t_j)$.

V izračunu uteži vektorja w v našem primeru upoštevamo vse vrednosti na celotnem linearnem sistemu (9) in ne samo lokalne utežene regresije. Na ta način lahko uporabimo manjše število jedrnih funkcij.

3.2 Posploševanje iz več trajektorij

Sočasno ocenjevanje frekvence Ω in parametrov uteži w robotu omogoča sprotno ponavljanje oz. reproduciranje demonstracijske trajektorije. Na ta način lahko demonstrator opazuje obnašanje robota ter neha z demonstracijo, ko je robotska trajektorija primerna. Podobna ideja je opisana v [12], kjer je prenos znanja na robota vključen v sistem z operaterjem v zanki.

Za posploševanje trajektorij shranimo ocenjene vrednosti frekvenc in trajektorije zadnjih nekaj period učenja. V našem primeru smo jih uporabili 5. Pri tem je kritičnega pomena sprotno ocenjevanje frekvence, saj posploševanje zahteva isto število period demonstracijskega signala.

Proces učenja zagotovi podatke trajektorijah (pozicije, hitrosti in pospeške) ter frekncah. Vsako situacijo opisuje neka točka povpraševanja, v našem primeru je bila to razlika v višini med dvema bobnoma.

Za posploševanje smo uporabili lokalno uteženo regresijo (LWR). Za dano točko povpraševanja q se optimalni parametri izračunajo iz dostopnih podatkov z oceno funkcije

$$C(q) = \sum_k L(\Xi(q_k, w), f_k) K(d(q, q_k)). \quad (16)$$

Lokalne modele opisuje

$$L(\Xi(q_k, w), f_k) = \|X_k w - f_k\|^2 \quad (17)$$

in glede na w moramo minimizirati funkcijo

$$\sum_{k=1}^M \|X_k w - f_k\|^2 K(d(q, q_k)), \quad (18)$$

kjer je K jedrna funkcija in d metrika v prostoru točke povpraševanja q . Izbrali smo naslednje utežitveno jedro

$$K(d) = \begin{cases} (1 - |d|^3)^3 & \text{če } |d| < 1 \\ 0 & \text{sicer} \end{cases}, \quad (19)$$

ki je končno ter ima zvezen prvi in drugi odvod. Ker podatke pridobimo inkrementalno, tako tudi posplošujemo, kar pomeni, da moramo pred posploševanjem poskrebeti, da so vsi podatki v primernem vrstnem redu. To vključuje posebej vodenje faz za posamezno demonstracijo.

Posplošiti moramo tudi posamezne frekvence, slednje pa lahko obravnavamo tudi kot funkcije točke povpraševanja.

4 Evalvacija metode posploševanja

Za evalvacijo metode posploševanja smo izvedli poskus na humanoidnem robotu CB-i, ki je bil razvit v sodelovanju med japonskim institutom ATR ter podjetjem Sarcos.

Humanoidni robot CB-i je človeške velikosti in ima 39 stopenj prostosti. Za naš eksperiment smo uporabili vseh 7 stopenj v desni roki ter 3 stopnje v trupu.

Za eksperiment smo posneli 10 trajektorij bobnanja z eno roko na dva bobna oz. boben in činelo. Pri vsaki demonstracijski trajektoriji je bila činela malce višje. Cilj eksperimenta je bil bobnanje na bobna različnih višin, pri čemer je točka povpraševanja v knjižnico naučenih gibov razlika v višini med njima. Razliko smo za opisani eksperiment določili z merjenjem, prav tako pa bi lahko uporabili robotski vid [13].

Slika 1 prikazuje rezultate posploševanja trajektorije za naljučno izbrano višino. Na sliki je prikazanih 5 trajektorij, ki prikazujejo rezultat posploševanja (v modri barvi) ter trajektorije, ki so bile uporabljene za generacijo te trajektorije. Z rdečo so prikazane trajektorije nad točko povpraševanja in z zeleno trajektorije pod točko povpraševanja. Opazna je podobnost med trajektorijami. Če bi bile demonstracijske trajektorije, ki smo jih generirali z roko, bolj enake, bi jim bila tudi posplošena trajektorija bolj podobna.

Slika 2 prikazuje demonstracijske trajektorije za učenje bobnanja na bobna različnih višin. Modre trajektorije prikazujejo demonstracijski signal, rdeče pa naučene trajektorije.

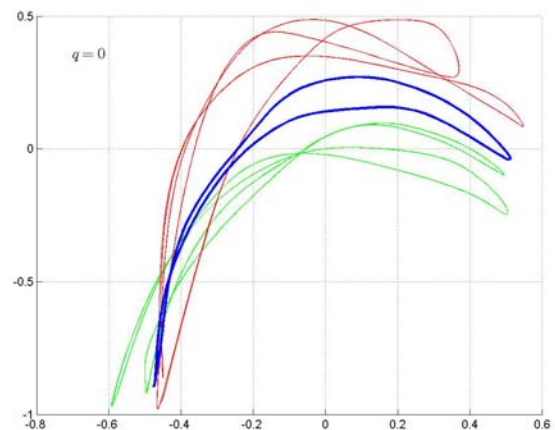


Figure 1: Rezultati posploševanja periodične trajektorije. Slika prikazuje rezultate v modri barvi, trajektorijo nad točko povpraševanja v rdeči barvi ter trajektorije pod točko povpraševanja v zeleni barvi.

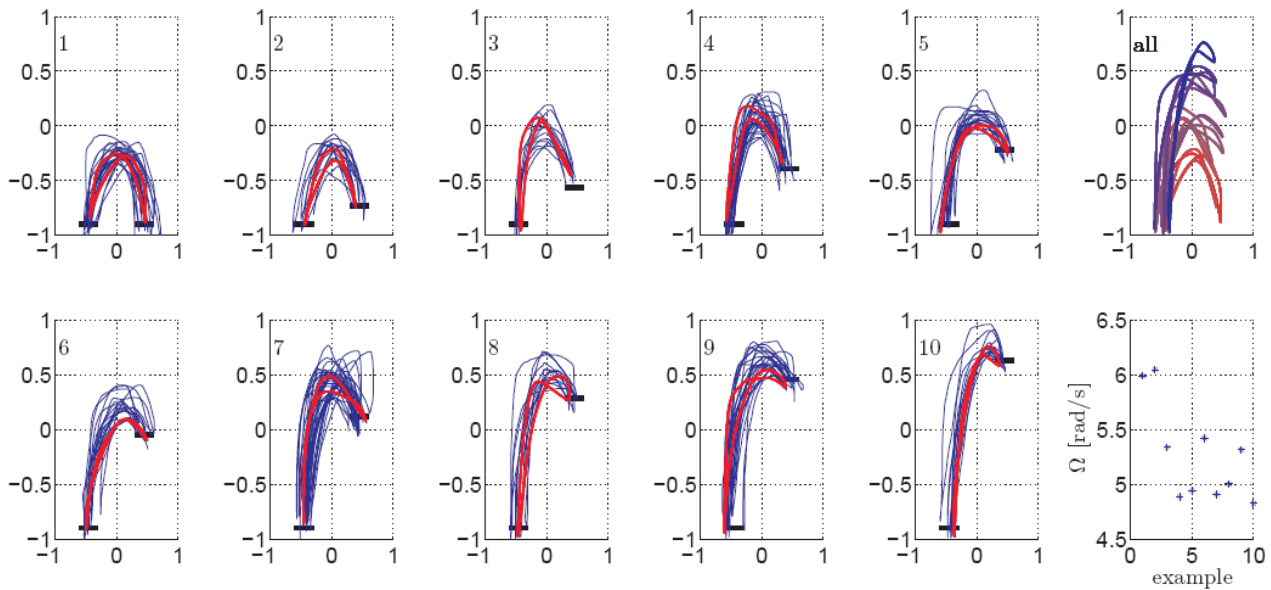


Figure 2: Demonstracijske trajektorije bobnanja na dva bobna različnih višin. Modre trajektorije prikazujejo demonstracijski signal, rdeče pa naučeni gib. Desno zgoraj so prikazane vse trajektorije na enem samem grafu. Desno spodaj so prikazane frekvence demonstracijskih gibov.

5 Povzetek.

V prispevku smo predstavili metodo posploševanja periodičnih trajektorij iz knjižnice naučenih gibov. Osnova zapisu so nelinearni dinamični sistemi oz. DMP-ji. Že sami DMP-ji omogočajo moduliranje naučene trajektorije ter so robustni na motnje. Sama modulacija pa ne zagotovi ohranjanje specifik posameznega giba, kajti oblika izvedenega giba prevzame obliko trajektorije kritično dušenega sistema drugega reda. Poleg tega ne omogočajo enostavnega spreminjanja trajektorije.

Z opisano metodo lahko iz knjižnice naučenih gibov izločimo primerne ter s statistično metodo posplošimo gibe med sabo, da dobimo gib, ki ohranja karakteristike naučenega giba. Metoda deluje povsem avtonomno in ne potrebuje intervencije eksperta. Metodo smo tudi uspešno preizkusili na humanoidnem robotu.

References

- [1] S. Schaal. Is imitation learning the route to humanoid robots? *Cognitive Sciences*. vol. 3. pp 233-422. 1999.
- [2] R. Dillmann. Teaching and learning of robot tasks via observation of human performance. *Robotics and Autonomous Systems*. vol. 47. no. 3. pp. 109-116. 2004.
- [3] A. Ude. C.G. Atkeson. M. Riley. Programming full-body movements for humanoid robots by observation *Robotics and Autonomous Systems*. vol. 47. no. 3. pp. 93-108. 2004.
- [4] A. Gams. L. Žlajpah. J. Lenarčič. Imitating human acceleration of a gyroscopic device. *Robotica*, vol. 25. no. 4, pp. 501-509. 2007.
- [5] M. Lungarella. G. Metta. R. Pfeifer. G. Sandini. *Developmental robotics: a survey*. Connection Science. vol. 15. no. 4. pp. 151-190. 2003.
- [6] J. Peters. S. Schaal. Reinforcement learning of motor skills with policy gradients. *Neural Networks*. vol. 21. pp. 682-697. 2008.
- [7] A. J. Ijspeert. J. Nakanishi. S. Schaal. Movement imitation with nonlinear dynamical systems in humanoid robots. *Proc. IEEE Int. Conf. Robotics and Automation*. pp. 1398-1404. Washington, D.C. 2002.
- [8] A. Gams. A. Ude. Generalization of Example Movements with Dynamic Systems. *Proc. IEEE Int. Conf. On Humanoid Robotics*. To appear. 2009.
- [9] P. Pastor. H. Hoffmann. T. Asfour. S. Schaal. Learning and generalization of motor skills by learning from demonstration. *Proc. IEEE Int. Conf. Robotics and Automation*. pp. 763-769. Kobe. 2009.
- [10] A. Gams. A. J. Ijspeert. S. Schaal. J. Lenarčič. On-line learning and modulation of periodic movements with nonlinear dynamical systems. *Autonomous Robots*. vol. 27. no. 1. pp. 3-23. 2009.
- [11] L. Righetti. J. Buchli. A. J. Ijspeert. Dynamic Hebbian learning in adaptive frequency oscillators. *Physica D*. vol. 216. pp. 269-281. 2006.
- [12] E. Oztop, J. Babič. J.G. hale. G. Cheng. M. Kawato. From biologically realistic imitation to robot teaching via human motor learning. *Neural Information Processing: 14th International Conference. Ser. Lecture Notes in Computer Science*. pp. 214-222. Berlin. 2008.
- [13] A. Ude. E. Oztop. Active 3-D vision on a humanoid head. In *Proc. 14th Int. Conf. Advanced Robotics*, Munich, Germany, June 2009.

ROBOTSKA REHABILITACIJA Z NAVIDEZNO RESNIČNOSTJO IN PSIHOFIZIOLOŠKIMI MERITVAMI

Domen Novak, Jaka Ziherl, Andrej Olenšek, Janez Podobnik, Matjaž Mihelj, Marko Munih

Laboratorij za robotiko in biomedicinsko tehniko

Fakulteta za elektrotehniko

Tržaška 25, 1000 Ljubljana, Slovenia

Tel: +386 1 4768196; fax: +386 1 4768239

e-mail: domen.novak@robo.fe.uni-lj.si

POVZETEK

V prispevku je predstavljen sistem za rehabilitacijo roke po možganski kapi. Sistem temelji na haptičnem robotu HapticMaster, ki je povezan z navideznim okoljem, v katerem lahko uporabnik vadi gibanje roke in prijemanje predmetov. Senzorji na robotu merijo gibe, sile in navore, poleg tega pa psihofiziološki senzorji merijo elektrokardiogram, prevodnost kože, dihanje in temperaturo kože uporabnika. S pomočjo psihofizioloških meritev lahko ocenimo psihološko stanje uporabnika. Sistem je bil preizkušen na devetnajstih subakutnih pacientih po možganski kapi.

ABSTRACT

This paper presents a system for post-stroke upper extremity rehabilitation. The system is based on the HapticMaster haptic robot, which is connected to a virtual environment that allows the user to practice arm movement and grasping. The robot's sensors measure movement, forces and torques. They are complemented by psychophysiological sensors that measure the user's electrocardiogram, skin conductance, respiration and skin temperature. Using these four measurements, it is possible to assess the user's psychological state. The system was tested with nineteen subacute stroke patients.

1 UVOD

V rehabilitaciji oseb z motnjami gibanja (zaradi kapi, poškodb hrbtenice itd.) se vedno bolj uveljavlja uporaba robotskih naprav, opremljenih s senzorji za zaznavanje pozicij in sil [1, 2]. Te naprave omogočajo bolj intenzivno terapijo, zato skrajšajo čas, potreben za terapijo, poleg tega pa nudijo tudi objektivno oceno motoričnih sposobnosti in napredka pacientov [3, 4]. Rehabilitacijo lahko dodatno izboljšamo s kombinacijo robotov in navidezne resničnosti. Navidezna okolja izboljšajo motivacijo pacientov, saj omogočajo vadbo v zanimivih, raznolikih okoljih. Izkušnje, pridobljene v teh okoljih, lahko pacienti s pridom uporabljajo v vsakdanjih opravilih [5].

V prispevku bomo predstavili sistem za robotsko rehabilitacijo, podprt z navidezno resničnostjo in psihofiziološkimi meritvami. Psihofiziologija je znanost, ki proučuje vpliv človekovega duševnega stanja na telo. Vključuje meritve funkcij vegetativnega (bitje srca, dihanje ...) in osrednjega živčnega sistema (elektroencefalografija, fMRI). Merjencu ni potrebno aktivno sodelovati pri meritvah, zato so psihofiziološke meritve neinvazivne in objektivne. Pogosto se uporabljajo za ocenjevanje stanja oseb npr. pri zdravljenju duševnih motenj s pomočjo navidezne resničnosti [6], v rehabilitaciji motenj gibanja pa do sedaj še niso bile uporabljene. S psihofiziološkimi meritvami smo želeli pridobiti dodatne informacije o počutju pacientov med vadbo z robotom.

2 ARHITEKTURA SISTEMA

Sistem za rehabilitacijo je sestavljen iz treh glavnih komponent: navideznega okolja, haptičnega vmesnika in psihofiziološkega merilnega sistema.

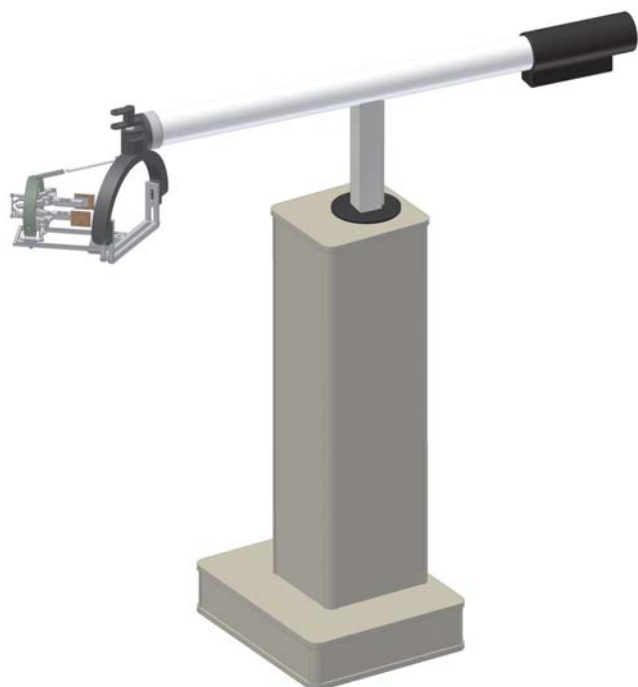
2.1 Haptični vmesnik

Haptični vmesnik je osnovan na haptičnem robotu HapticMaster [7], ki ima tri prostostne stopnje. Prvi sklep omogoča translacijo v navpični smeri, drugi rotacijo okrog navpične osi, tretji pa translacijo v vodoravni ravnini. Na vrhu robota se nahaja tudi senzor za merjenje sile, ki vsebuje tri merilne celice z uporovnimi lističi za merjenje sile v treh oseh kartezičnega koordinatnega sistema senzorja sile. Robot HapticMaster je prikazan na sliki 1.

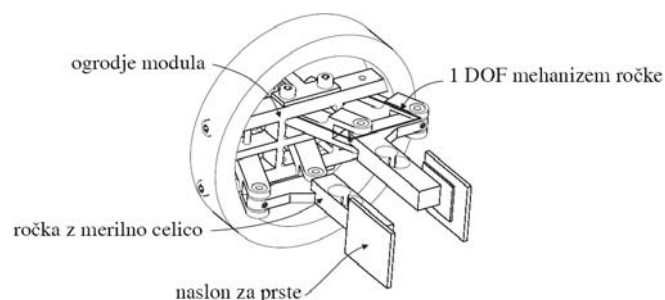
Na senzor za merjenje sile je pritrjen modul za prijemanje, razvit v našem laboratoriju [8]. To je pasivna naprava, ki omogoča pasivno haptičnost za prijemanje navideznih predmetov v navideznih okoljih. Ima dve pasivni prostostni stopnji. Na vsako od njiju je pritrjen senzor za merjenje sile. Prvi senzor meri silo, ki jo uporabnik izvaja s palcem, drugi pa skupno silo, ki jo uporabnik izvaja s kazalcem in sredincem. Občutek pasivne haptičnosti je dosežen prek dveh elastičnikov, ki sta dodani med premične dele modula in njegov okvir. Tako uporabnik dobi občutek, da je predmet resnično prijel. Modul za prijemanje je prikazan na sliki 2.

Poleg robota HapticMaster z modulom za prijemanje ima haptični vmesnik še sistem za kompenzacijo teže roke. Na

roko uporabnika sta pritrjeni dve manšeti (ena nad in ena pod komolcem), ki sta preko močnih vrvi povezani z dvema motorjema na stropu. Navor, ki ga izvajata motorja, lahko ročno nastavljamo. Tako na uporabnikovo roko stalno deluje sila, ki kompenzira silo teže. Sistem se je izkazal za zelo uporabnega pri pacientih po kapi, ki roko sicer lahko premikajo levo in desno, ne morejo pa je dvigniti.



Slika 1: Robot HapticMaster z modulom za prijemanje.



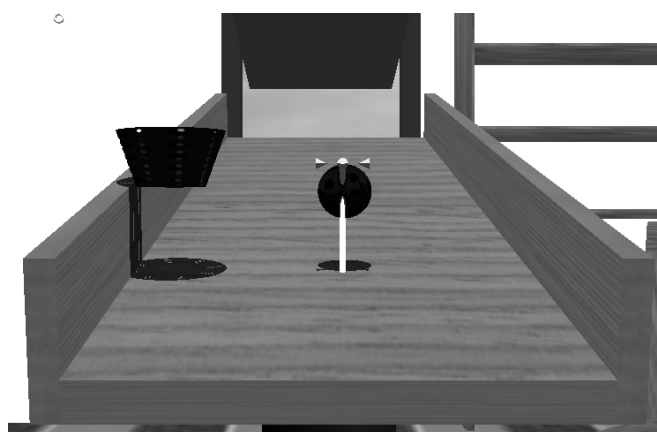
Slika 2: Modul za prijemanje. Na vsaki od ročk se nahaja senzor sile. Vsaka od ročk je pritrjena na ogrodje modula prek mehanizma z eno prostostno stopnjo. Na drugi strani modula so ročke vezane na vzmeti.

2.2 Navidezno okolje in naloga

V sistemu navidezne resničnosti lahko ustvarimo veliko različnih okolj in nalog, vendar smo se za prve klinične preizkuse sistema odločili, da vsem pacientom ponudimo enako nalogo. Naloga je bila načrtovana z namenom omogočati hkratno vadbo gibanja roke in prijemanja, saj pri večini opravil v vsakdanjem življenju roko uporabljamo za različne gibe ter za prijemanje predmetov [9].

Navidezno okolje je prikazano na sliki 3. V sredini navideznega okolja je miza, nagnjena navzdol proti gledalcu. Vrh robota HapticMaster je predstavljen s tremi rumenimi oznakami v vodoravni črti. Razdalja med temi tremi oznakami je obratno sorazmerna sili, ki jo uporabnik izvaja na modul za prijemanje. Na vrhu mize se pojavi žoga, ki se začne kotaliti navzdol. Naloga uporabnika je, da s pomočjo haptičnega vmesnika prime žogo, preden le-ta doseže spodnji rob mize. Ko je žoga prijeta, se nad mizo pojavi koš. Uporabnik mora žogo držati in jo postaviti v koš. Ko je žoga v košu ali pa pade z mize, se na vrhu mize pojavi nova žoga in naloga se nadaljuje. Preko haptičnega vmesnika lahko uporabnik čuti vse predmete v navideznem okolju.

Naloga omogoča različne nivoje haptične podpore. Če uporabnik ni zmožen premikati roke levo in desno, se lahko robot samodejno premakne do žoge. Če uporabnik ni zmožen prijemanja, lahko robot žogo tudi prime. Končno lahko robot žogo tudi postavi v koš. Pot, po kateri se robot premika, je izračunana s pomočjo B-zlepkov in aproksimira naravno gibanje človeške roke [10]. Vsako od teh treh možnosti je mogoče vključiti neodvisno od drugih dveh. Uporabnik samostojno opravlja gibe, ki jih zmore, robot pa pomaga pri razgibavanju mišic, ki jih uporabnik ne more premikati.



Slika 3: Navidezno okolje: miza, žoga (sredina), koš (levo).

2.3 Psihofiziološke meritve

Med nalogo lahko s pomočjo ojačevalnika g.USBamp (proizvajalec g.tec Medical Engineering GmbH, Avstrija) in pripadajočih senzorjev merimo štiri različne fiziološke signale: elektrokardiogram, prevodnost kože, frekvenco dihanja in temperaturo kože. Standardne elektrode za elektrokardiogram so nameščene na trup. Senzor za prevodnost kože je realiziran v obliki dveh elektrod, nameščenih na kazalec in sredinec neprizadete kože. Na elektrodi je priključen konstanten vir napetosti (0,1 V), senzor pa meri prevodnost posredno preko toka. Senzor za temperaturo kože je termistor, nameščen na konico mezinca. Tudi senzor dihanja je realiziran v obliki termistorja, ki je nameščen pod nos merjenca. Termistor meri razliko med

temperature vdihanega ter izdihanega zraka in lahko zaradi svoje oblike meri tako dihanje skozi nos kot skozi usta. Iz elektrokardiograma izračunamo frekvenco bitja srca in različne ocene variabilnosti bitja srca [11]. Prevodnost kože se zviša, kadar je merjenec aktiven tako v fizičnem kot v miselnem smislu [12]. Prav tako se s fizično in miselno aktivnostjo zviša frekvenca dihanja [13], temperatura kože pa se zniža s povečano miselno aktivnostjo [14].

3 KLINIČNI PREIZKUS SISTEMA

3.1 Pacienti in postopek

Sistem smo preizkusili z devetnajstimi subakutnimi pacienti po možganski kapi (13 moških, 6 žensk, starost 49,3 +/- 13,5 let), ki so se zdravili na Inštitutu za rehabilitacijo Republike Slovenije. Devet pacientov je imelo levostransko hemiparezo, devet pacientov desnostransko hemiparezo, en pacient pa tetraparezo z izrazitejšo levostransko hemiparezo. Vsi pacienti so opravljali preizkus po istem postopku. Ob prihodu je vodja študije vsakemu pacientu predstavil robota in študijo. Po podpisu soglasja o sodelovanju v raziskavi se je pacient usedel pred robota, terapevt pa je nanj namestil manšeti za kompenzacijo sile teže ter psihofiziološke merilne senzorje. Pacient je nato opravil kratko vajo naloge, opisane v poglavju 2.2. Haptična podpora je bila določena individualno med vajo glede na uspešnost pacientov. Dvanajst pacientov ni potrebovalo haptične podpore, eden je potreboval pomoč pri postavljanju žoge v koš, eden pri prijemanju žoge in postavljanju žoge v koš. Pri petih pacientih so bili vključeni vsi trije nivoji haptične podpore.

Po kratki vaji je pacient tri minute počival. Ta interval je bil namenjen merjenju psihofizioloških signalov v mirovanju. Po počitku je pacient tri minute opravljal nekoordinirane gibe. Zaslona je bil ugasnjen, robot pa je omogočal prosto premikanje roke. Pacientu je bilo naročeno, naj roko premika levo in desno s poljubno hitrostjo in občasno stisne prste. Po opravljanju nekoordiniranih gibov je pacient znova tri minute počival, nato pa tri minute opravljal nalogo, opisano v poglavju 2.2.

Statistična značilnost razlik je bila izračunana bodisi s t-testom bodisi s t-testom parov, odvisno od tipa primerjave (ponovljene meritve na istih osebah ali primerjava med skupinami oseb).

3.2 Rezultati

Meritve sil so uspešno pokazale razliko med pacienti, ki so potrebovali haptično podporo, ter pacienti, ki niso potrebovali haptične podpore. Pri pacientih, ki niso bili sposobni samostojnega dviganja roke, je bila izmerjena sila v navpični smeri (pozitiven predznak pomeni navzgor) med postavljanjem žoge v koš enaka -1,8 +/- 1,5 N. Povprečna sila, ki so jo pacienti izvajali na senzor v navpični smeri, je bila torej usmerjena navzdol. Pri pacientih, ki so roko lahko samostojno dvigali, je bila izmerjena sila v navpični smeri med postavljanjem žoge v koš enaka 6,1 +/- 0,9 N. Razlika

med skupinama je bila statistično značilna ($p < 0,001$). Z meritvami sil tako lahko ugotovimo, ali gib pretežno opravlja pacient ali robot.

Psihofiziološke meritve so pokazale veliko fiziološko razliko med mirovanjem in opravljanjem nalog. Prevodnost kože, ki sicer s časom v povprečju pada, se je med opravljanjem nekoordiniranih gibov zvišala za 10,0 +/- 21,8 μ S, med opravljanjem naloge z žogo pa za 20,0 +/- 21,8 μ S. Temperatura kože se je med opravljanjem nekoordiniranih gibov zvišala za 1,0 +/- 1,4 K, med opravljanjem naloge z žogo pa se je znižala za 1,1 +/- 1,7 K. Razlike v prevodnosti kože in temperaturi kože med mirovanjem in nalogo je bila statistično značilna tako za nekoordinirane gibe kot za nalogo z žogo ($p < 0,001$ za vse 4 primerjave). Naloga z nekoordiniranimi gibi (ki ne zahtevajo miselnega navora) in naloga z žogo (ki zahteva precejšnjo koordinacijo) sta se statistično značilno razlikovali tako v spremembi prevodnosti kože ($p = 0,02$) kot v spremembi temperature kože ($p < 0,001$). Prevodnost in temperatura kože tako lahko ločita ne samo med mirovanjem in opravljanjem nalog, marveč tudi med opravljanjem miselno nezahtevnih in miselno zahtevnejših gibov. Temperatura kože je še posebej uporabna za ločevanje nezahtevnih in zahtevnih gibov. Pri teh meritvah je pomembno predvsem, da nam nudijo pogled v subjektivno počutje pacienta. Tako dopolnjujejo informacije, ki jih dobimo s pomočjo senzorjev sil in navorov.

Spremembe v bitju srca so odražale predvsem fizični napor. Razlika v frekvenci bitja srca med nekoordiniranimi gibi in nalogo z žogo ni bila statistično značilna. V obeh primerih se je frekvenca bitja srca v povprečju povečala za 4 udarce/minuto. Tudi razlika v variabilnosti bitja srca med nekoordiniranimi gibi in nalogo z žogo ni bila statistično značilna. Frekvenca dihanja je bila med opravljanjem obeh nalog višja kot v mirovanju, med nalogo z žogo pa je bila višja kot pri opravljanju nekoordiniranih gibov ($p = 0,046$). Žal pa so bile razlike relativno majhne (nekoordinirani gibi: 0,8 +/- 2,3 vdiha/minuto več kot v mirovanju; naloga z žogo: 1,9 +/- 3,0 vdiha/minuto več kot v mirovanju). Pri nekaterih pacientih se je frekvenca dihanja med opravljanjem nalog celo znižala kljub opaznemu fizičnemu naporu.

3.3 Odziv pacientov

Odziv pacientov na navidezno okolje in haptičnega robota je bil večinoma pozitiven. Nekateri pacienti so se robota najprej bali, vendar pa so se ga večinoma hitro navadili. Pri tem je zelo pomembno, da se robot giblje zvezno, brez kakršnihkoli sunkovitih gibov, ki bi pacienta lahko prestrašili. Pacienti so bili sicer zelo zadovoljni s sistemom za kompenzacijo teže roke in s haptično podporo gibanja. Nekateri pa so zaradi posledic kapi imeli težave z razumevanjem naloge in z gibanjem v tridimenzionalnem navideznem prostoru. Pri teh pacientih je ob začetku okrevanja smiselno uporabiti bolj preproste naloge v manj

dimenzijah in postopoma preiti na kompleksnejše naloge. Senzorja za prevodnost in temperaturo kože za pacienta nista bila moteča. Tudi meritve elektrokardiograma niso bile moteče, vendar pa je bilo potrebno več časa, da smo jih namestili na pacienta. Kot glavni problem se je izkazal senzor za dihanje, ki je bil zaradi svoje postavitve (pod nosom) zelo moteč. Pacienti so se pogosto pritoževali, da jih senzor žgečka ali srbi. Zaradi teh težav so senzor pogosto premikali s prsti, kar je povzročalo motnje. Do motenj je prihajalo tudi zaradi govora pacientov. Pri nadaljnjih meritvah bi lahko za merjenje dihanja uporabili piezoelektrični senzor dihanja, nameščen okoli trupa, vendar pa je namestitev takega senzorja veliko bolj zamudna.

Na haptičnem področju bi bila najbolj bistvena izboljšava razvoj adaptivnega sistema, ki bi se lahko sam odločil, koliko pomoči pri gibanju pacient potrebuje. To sedaj izbira operater ročno, možno pa bi bilo potreben nivo podpore računati na podlagi izmerjenih sil in navorov. V poznejši stopnji razvoja bi lahko rehabilitacijski sistem nalogo spreminjal tudi na podlagi psihofizioloških odzivov. Iz njih bi lahko ugotovili, ali se pacient dolgočasi ali pa je preobremenjen, na podlagi tega pa spreminjali težavnost naloge ali kompleksnost navideznega okolja.

Še vedno pa je pri rehabilitaciji nujna prisotnost terapevta. Nekateri pacienti so namreč sami popolnoma sposobni določenih gibov, a jih zaradi pomanjkanja motivacije ali pa zaradi nerazumevanja naloge ne opravljajo dovolj hitro. Tu lahko človeški pristop najboljše pomaga.

4 ZAKLJUČEK

Uspešno smo razvili sistem za rehabilitacijo roke, ki omogoča vajo tako gibov roke kot prijemanja. Sistem smo nadgradili z navideznim okoljem, ki pacientu omogoča vadbo v različnih zanimivih okoljih in tako poživi terapijo. Dodatno smo z meritvijo psihofizioloških odzivov omogočili vpogled v čustveno in miselno stanje pacienta, ki je pri rehabilitaciji zelo pomembno. Pričakujemo, da se bo razvoj sistema nadaljeval, hkrati pa bo ta metoda rehabilitacije zaradi svojih prednosti uspešno vstopila v klinično prakso.

Zahvala

Raziskavo sta podprla Javna agencija za raziskovalno dejavnost Republike Slovenije ter Sedmi okvirni program Evropske skupnosti (projekt MIMICS, grant 215756). Zahvaljujemo se tudi vsem prostovoljcem, ki so s svojim sodelovanjem v kliničnih preizkusih omogočili uspešen razvoj sistema.

Viri

[1] R. C. V. Loureiro, C. F. Collin, W. S. Harwin. Robot aided therapy: Challenges ahead for upper limb stroke rehabilitation. V: *5th International Conference on*

Disability, Virtual Reality and Associated Technologies, Reading, UK, The University of Reading, 2004: 3-39.

- [2] T. Nef, M. Mihelj, G. Colombo, R. Riener. ARMIN – A robot for patient-cooperative arm therapy. *Medical and Biological Engineering and Computing* 45: 887-900, 2007.
- [3] W. S. Harwin, J. L. Patton, V. R. Edgerton. Challenges and opportunities for robot-mediated neurorehabilitation. *Proceedings of the IEEE*, 94: 1717-1726, 2006.
- [4] L. E. Kahn, M. L. Zygmant, W. Z. Rymer, D. J. Reinkensmeyer. Robot-assisted reaching exercise promotes arm movement recovery in chronic hemiparetic stroke: a randomized controlled pilot study. *Journal of Neuroengineering and Rehabilitation*, 3, 2006.
- [5] M. K. Holden. Virtual environments for motor rehabilitation: Review. *CyberPsychology & Behavior*, 8: 187-211, 2005.
- [6] F. H. Wilhelm, M. C. Pfaltz, J. J. Gross, I. B. Mauss, S. I. Kim, B. K. Wiederhold. Mechanisms of virtual reality exposure therapy: The role of the behavioral activation and behavioral inhibition systems. *Applied Psychophysiology and Biofeedback*, 30: 271-284, 2005.
- [7] R. Q. van der Linde, P. Lammertse. HapticMaster – a generic force controlled robot for human interaction. *Industrial Robot: An International Journal*, 30:515-524, 2003.
- [8] J. Podobnik. Seganje in prijemanje v haptičnih navideznih okoljih. *Doktorska disertacija*. Fakulteta za elektrotehniko, Univerza v Ljubljani, 2009.
- [9] S. L. Fritz, K. E. Light, T. S. Patterson, A. L. Behrman, S. L. Davis. Active finger extension predicts outcomes after constraint-induced movement therapy for individuals with hemiparesis after stroke. *Stroke*, 36:1172-1177, 2005.
- [10] J. Zihelr, J. Podobnik, M. Sikič, M. Munih. Pick to place trajectories in human arm training movement. *Technology and Health Care*, in press.
- [11] Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal*, 17, 354-381, 1996.
- [12] C. Collet, P. Averty, A. Dittmar. Autonomic nervous system and subjective ratings of strain in air traffic control. *Applied Ergonomics*, 40, 23-32, 2009.
- [13] J. A. Veltman, A. W. Gaillard. Physiological workload reactions to increasing levels of task difficulty. *Ergonomics*, 41, 656-669, 1998.
- [14] M. Ohsuga, F. Shimono, H. Genno. Assessment of phasic work stress using autonomic indices. *International Journal of Psychophysiology*, 40, 211-220, 2001.

Zbornik 12. mednarodne multikonference
INFORMACIJSKA DRUŽBA – IS 2009

Proceedings of the 12th International Multiconference
INFORMATION SOCIETY – IS 2009

Kognitonika

Cognitonics

Uredila / Edited by

Vladimir A. Fomichov, Olga S. Fomichova

<http://is.ijs.si>

12.–13. oktober 2009 / October 12th–13th, 2009
Ljubljana, Slovenia

Preface/Predgovor

First International Workshop on Cognitonics (Cognit 2009)

The swift technological changes in the modern information society cause not only positive but also negative shifts in the system of humanitarian values. The first shift is the shift from the eternal values to commercialized values. The second shift is a consequence of the underestimation of the value of national cultures in contrast to the overestimation of the on-going processes of globalization (by the young generation, in particular). The third shift is dealing with the existing gap between the rational, pragmatic, self-oriented development and spiritually-coloured emotional development of the personality. The expanding gap destroys the integrity of the human nature.

Cognitonics is a new scientific discipline aiming (a) at explicating the distortions in the perception of the world caused by the information society and globalization and (b) at coping with these distortions in different fields by means of elaborating systemic solutions for compensating the negative implications for the personality and society of the stormy development of information technologies and globalization processes, in particular, for creating cognitive-cultural preconditions of the harmonic development of the personality in the information society and for ensuring the successive development of national cultures and national languages. The birth of Cognitonics has been stimulated by the ideas of Philosophy, Cognitive Linguistics, Artificial Intelligence theory, Applied Linguistics, Art theory, Cognitive Psychology, and Cognitive Biology.

Cognitonics not only has numerous connections with Philosophy, Cognitive Linguistics, Applied Linguistics, Art theory, Cognitive Psychology, Education theory but also formulates a new, large-scale goal for the software industry and Web science: to develop a new generation of culture-oriented computer programs and online courses (in the collaboration with educators, linguists, art historians, psychologists) - the programs and online courses intended for supporting and developing creativity, cognitive-emotional sphere, the appreciation of the roots of the national cultures, the awareness of the integrity of the cultural space in the information society, and for developing symbolic information processing and linguistic skills, associative and reasoning abilities of the learners.

Editors and Program Chairs / Urednika

- Vladimir Fomichov
- Olga Fomichova

COGNITONICS AS AN ANSWER TO THE CHALLENGE OF TIME

Olga S. Fomichova (1), Vladimir A. Fomichov (2)

(1) State Educational Centre “Dialogue of Sciences”,
Universitetsky prospect 5, 119296 Moscow, Russia

(2) Faculty of Business Informatics

State University – Higher School of Economics

Kirpichnaya str. 33, 105679 Moscow, Russia

Tel: +7-495-771-3238; fax: +7-495-771-3238

e-mail: vfomichov@hse.ru and vdrfom@aha.ru

ABSTRACT

This paper aims at explaining the principles and the model of the educational process underpinned by the ideas of Cognitonics. It is a new scientific discipline aimed at explicating the distortions in the perception of the world caused by the information society and globalization and at coping with these distortions in different fields, in particular, conducing to overcoming a gap between spiritual (emotional) and intellectual (professional) development of the person in the new reality of swift changes, technological challenges, and globalization.

1 INTRODUCTION

In the modern information society, the following tendencies concerning the development of the personality are observed: the shift from the eternal values to commercialized values, an increasing gap between the intellectual (professional) and spiritual (emotional) development of the personality, and the underestimation of the national cultures in the age of globalization. To overcome these negative shifts, it was proposed to develop a new scientific discipline called Cognitonics [1, 2]. It aims (a) at explicating the distortions in the perception of the world caused by the information society and globalization and (b) at coping with these distortions in different fields by means of elaborating systemic solutions for compensating the negative implications for the personality and society of the stormy development of information technologies and globalization processes, in particular, for creating cognitive-cultural preconditions of the harmonic development of the personality in information society and for securing the successive development of national cultures [1, 2, 4].

The birth of Cognitonics has been stimulated by the ideas of Philosophy, Cognitive Linguistics, Artificial Intelligence theory, Applied Linguistics, Art theory, Cognitive Psychology, and Cognitive Biology. The goals of Cognitonics are to be achieved, first of all, through education. The main subject of this paper is a working model of educational process underpinned by the ideas of

Cognitonics. This working model is a cross-disciplinary program covering 12 years of continuing studies, the starting age of the students is 5 - 6 years. The positive experience of implementing this program accumulated by the authors covers 19 years, the total number of taught students exceeds six hundred fifty, the age varies from 6 to 19 years.

2 BASIC PRINCIPLES UNDERPINNING A WORKING MODEL OF COGNITONICS

The main social function of education is the reproduction of the culture. We regard cognition as an active transfiguration, as an introduction of beauty into the world by means of creativity. Cognition is an active transfiguration in comparison with the passive reflection. Our approach to solving educational problems corresponds to the ideas of constructivist theory, according to which knowledge is not just simply transmitted from the teacher to the pupil but rather is being constructed in the mind of the pupil in the process of active learning. Under the framework of our new approach, underpinned by Cognitonics as a new scientific discipline, creative work aims at transfiguration of the world around by introducing beauty into it. Such kind of creativity introduces humans into the cultural space, the space of eternal values and eternal questions.

The so-called world of grown-ups, revealed in the world of art, music, poetry, literature, becomes very tempting for children. The process of knowledge accumulation is speeding up day by day. Children are taught to make their own discoveries while decoding the pictures and establishing the ties between painting, poetry, literature. The starting age of the children initiated into the cultural space and into the world of art and thoughts in particular is very important. Taking into account that the cultural space of the mankind is like an ocean and own cultures – like rivers flowing into it to enrich it, children should be taught how to swim first in their “local rivers” and then they will be bestowed by the intellectual and spiritual pleasure to breathe in the air of the ocean, and some of them won’t resist the temptation to sail or to swim in that ocean.

The aim of cultural reproduction and enrichment as well as the aim of intellectual and spiritual development of the child

should be put together and incorporated in the educational process. The peculiarity of educational process is defined by putting together, intertwining in one course different humanitarian branches: literature, poetry, art, music, philosophy, language, and classic dances. This cross-disciplinary program across traditional educational disciplines is developed to achieve the goal. Learning at a high cognitive level requires the understanding of facts and concepts in their context, applying them to new situations, analyzing them, evaluating them, and synthesizing new concepts from them.

The efficiency of such idea is explained by the natural ability of the human brain to perceive the world in its various manifestations. The program also takes into account the peculiarity of the modern society – a self-oriented, strongly pragmatic generation.

2.1 Basic Principle 1: Establishing Numerous Correspondences Between the Topics Under Discussion Throughout the Course

Let's begin to consider the main principles of the educational process model underpinned by Cognitonics. First of all, show (as an example) the way the basic notion of the Russian culture – Beauty – is introduced into the conceptual picture of the world of the students. This will help us to get an initial impression about the principle of establishing numerous correspondences between the topics under discussion throughout the course.

1 st year	2nd year	3rd year	4th year
Beautiful princess: the phenomenon of Cinderella's sisters	The beauty of the seasons: subtle beauty	aesthetics and ethics	Beauty as a philosophical notion
5th year	6th year		7th year
The difference between cold glamorous beauty and subtle beauty	The idea of beauty in the female portraits (Russian painting of the 19th -20th centuries)		The beauty of the parks and estates
8th year	9th year	10th year	
The idea of beauty in Russian culture revealed in painting at the beginning of the 20 th century. "Blue Rose" ("the pictures like prayers")	The way the human consciousness resists the destruction of Beauty as a philosophical notion (the beginning of the 20th century)	Linden avenues (dark linden alleys of Russian estates) as an embodiment of Russian culture, based on the notion of beauty (the end of the 19th – the beginning of the 20th century)	

2.2 Basic Principle 2: A Particular Notion as the Point of Intersection of the Ideas

Example 1. The notion "calling" is the main notion connecting a human being with the society. It is the strongest social link, providing the sound relationships between the

human beings, forming the sound society. This notion is the basic idea for entering the society, revealing one's self, and being a success.

Calling is considered as the most important work a person can do in which he/she would be most difficult to replace. This notion is discussed together with such notions as assignment and occupation. As an example of the point of intersection between the eternity and every-day life, the idea of uncertainty is discussed in connection with Matthew the Levite, who was a tax collector in Capernaum (Bulgakov's "Master and Margarita"). A tax collector was one of the most certain occupations in Rome. When God called Matthew into service, he didn't present him with an employment contract. We cannot know the future exhaustively, an element of faith exists in every decision. On the other hand, the life experience proves that without faith it will be hard to be creative, innovative, and curious about new technologies and computers.

The notion "etiquette" is introduced in the context of social agreement and social ties. Students are taught that the violation of social agreement causes serious private problems for the person as a result of misunderstanding. It leads to expanding the space of uncertainty for the person.

Example 2. Predictive behavior caused by etiquette minimizes uncertainty and makes social ties much stronger and much more comfortable. According to [3], social agreements are aimed at making the behavior predictable. In "Sleeping Beauty", for instance, the King behaved against etiquette. He sent the invitations to the fairies but he didn't receive the confirmations. As a result, he didn't know whether the invitations were received and accepted. So the violation of the social agreement caused the provocative behavior on the part of the King. As a result, the King faced unpredictable behavior of the 13th fairy.

Social agreements are based on communication. To attain better communication within the community, to make it relevant, children should be taught to listen to the partner of communication and try to hear with all their hearts.

Children should be aware of the problem of understanding, because understanding is based on the conceptual picture of the world. Fyodor Tyutchev, a well-known Russian poet, philosopher, and a diplomat of the 19th century, wrote: "A thought when spoken becomes a lie." Being decoded by students, this idea will conduce to minimizing the examples of misunderstanding between people: between grown-ups and children, between people belonging to different cultures.

2.3 Basic Principle 3: Purposeful Socialization (How to Avoid the False Impression of Impinging upon an Individual Freedom)

To overcome the difficulties and misunderstanding that cause rebellious behavior and deep conflicts, it is necessary to make children understand from the very beginning that they and adults *are eating one and the same apple but from different sides*.

Example. One of the notions forming that so called *apple* is “offence”. We have to consider offence as a result of a situation when someone’s behavior doesn’t meet the expectations of the other person. The idea of putting together in one bundle these main principles (our educational model is based also on several

other principles formulated in our previous publications) is to make the child be reflected “in all the mirrors” round him/her in one moment in order to have a multi-facet reflection. It is an attempt to implement a cross-disciplinary approach to viewing oneself as a multi-facet personality.

3 THE METAPHOR OF AN EGG: A YOUNG STUDENT AT THE AGE OF 12 - 13 IS INSIDE AN “EGG”

Let’s consider the structure of three metaphoric concentric layers forming the cognitive nourishing shell of an egg (at the moment when the egg hatches)

Denotations: OutLr = Outer Layer, MidLr = Middle Layer, InnLr = Inner Layer

SECTOR 1	→	SECTOR 2
OutLr: Basic notions MidLr: Good – Bad, Beauty InnLr: Life experience		OutLr: Reality MidLr: Uncertainty, swift changes InnLr: The phenomenon of Cinderella’s sisters; “Spring” (Boris Pasternak)
SECTOR 3	→	SECTOR 4
OutLr: Anchorage = Sweet Home (Dulce domum) MidLr: Certain, clear to children values; moral values InnLr: Revealing the picture of the world of the student		OutLr: Awareness of another look MidLr: (1) the basis for misunderstanding; (2) mind-expanding experience; (3) thought-provoking experience InnLr: thaw → a dapple-gray stallion; thunderstorm → blinding stills; definition of offence
SECTOR 5	→	SECTOR 6
OutLr: Thought-Producing Self MidLr: The student generates His/her own vision, idea, metaphor InnLr: “Whipped cream on the surface of the lake” (the image of white lilies)		OutLr: Become noticeable MidLr: Language (discussion): “I agree, disagree, I would like to add” InnLr: Awareness of his/her thoughts as a precondition of the discussion
	→	SECTOR 7
		OutLr: Reward for the work of soul and mind + new look MidLr: Fresh experience InnLr: awareness, appreciation of the society

Comments on the Metaphor. A young student (the age 6 – 12 years) can be depicted as an imaginary egg. Until he/she is not a teenager, until the “egg” hatches and a teenager comes out of it and starts considering himself/herself as a member of the society, we propose the way of handling the “egg”, and this new way is underpinned by the ideas of Cognitonics.

IF AN EGG THEN:

1. Never press, never impinge upon. Teach to take the responsibility for the deeds and views.
2. Put it into the warmth of concern and protection.
3. Always watch and be attentive to the surrounding lest the egg should be destroyed before due time.
4. Don’t be too attentive. Remember: once in due time the egg will hatch!

4 CONCLUSION

Our working model of educational problem deals with the existing gap between the intellectual and spiritual development of the personality. The notion “human being”

can’t be regarded without considering such notions as “body”, “soul”, “spirit”. The expanding gap destroys the integrity of the human nature. Spiritual values are rooted in the developed emotional sphere of the personality. The emotional sphere is closely connected with the sensitivity. Modern education is focused on the development of cognitive mechanisms being necessary for improving the

information processing abilities of the child. Informational technologies don't require the spiritual development of the person. But it is clear that every intellectual discovery should be followed by equal to it spiritual discovery, lest the present time and challenging prospects should split off the eternity. In this context:

1. The paper reveals the main principles of the working model of educational process underpinned by Cognitonics. This model has been tested for 19 years, more than 650 students have taken part in this work. The age of students varies from 5 till 19.
2. The model is strictly oriented at the peculiarities of the different ages of the students.
3. It takes into account the peculiarities of information society, globalization, and the existing shift towards the commercialized values.
4. The model is aimed, in particular, at balancing the work of soul and the work of mind and at bridging the gap between the rational and ideal entities in the students' conceptual picture of the world.
5. The basic ideas of the working model of educational process underpinned by Cognitonics are published in the monograph [4], in over 50 papers in English (numerous references can be found in [1-2, 4]), and in 3 papers in French.

References

- [1] V.A. Fomichov, O.S. Fomichova. Cognitonics as a New Science and Its Significance for Informatics and Information Society; Special Issue on Developing Creativity and Broad Mental Outlook in the Information Society (Guest Editor Vladimir Fomichov), *Informatica. An Intern. Journal of Computing and Informatics (Slovenia)*, 2006, Vol. 30, No. 4, pp. 387-398.
- [2] O.S. Fomichova, V.A. Fomichov. Cognitonics as a New Science and Its Social Significance In the Age of Computers and Globalization. *IIAS-Transactions on Systems Research and Cybernetics. Vol. VII, No. 2. Intern. Journal of The International Institute for Advanced Studies in Systems Research and Cybernetics*. Published by The International Institute for Advanced Studies in Systems Research and Cybernetics (IIAS), Tecumseh, Ontario, Canada, 2007, pp. 13-21.
- [3] J. R. Searle. *The Construction of Social Reality*. The Penguin Press. 1995.
- [4] O.S. Fomichova. *Humanitarian Education – an Answer to the Challenge of Time*. Moscow State University Publishing House. 2009. In Russian.

AN INFORMATION SYSTEM IN SCHOOL FOR A RISK MANAGEMENT OF THE INTERNET: PREVENTING CYBERBULLYING WITHOUT PROHIBITIONS

Hirohiko Yasuda

Shimonoseki Technical High School
Tomitou-cyo, 4-cyome, 1-1, Shimonoseki City,
postcode 759-6613, Japan
Tel: +81-83-258-0065; fax: +81-83-258-0685
e-mail: yasuda.hirohiko@ysn21.jp
konippon@pluto.dti.ne.jp

ABSTRACT

The aim of this study is preventing cyber-bullying with the risk management system. This paper examines the importance of the guardians role in preventing cyber bullying and considers how the schools can support guardians and students. School organizes Information System to support students and guardians. School supports them, without the means of class, to gain extensive knowledge on an information society from the view point of Web Science [1].

1 INTRODUCTION

There is no solid solution or countermeasure to stop and prevent cyber-bullying. There are two reasons why school can not stop cyber-bullying. The first reason is that technology is changing rapidly and knowledge does not last long. The second reason is that High School students do not like any kind of moral education. We have been practicing Information Education in interdisciplinary approach since 2002 [2]. We have used two methods. One is using leaflets which mainly deal with case studies of computer science and social science. The other is organizing Information System of school to support students and guardians.

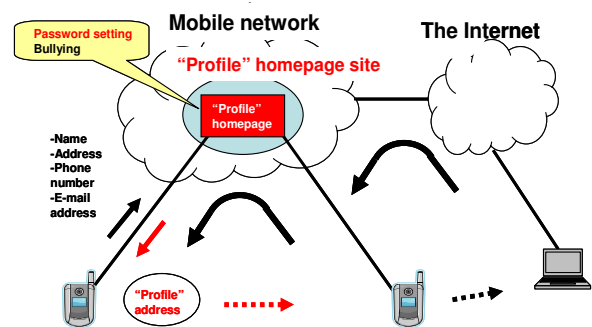
In 2006, Tim Berners-Lee and Wendy Hall, Nigel Shadbolt, Daniel J. Weitzner have advocated Web Science and founded Web Science Research Initiative (WSRI). Web Science focuses on understanding, designing, and applications that make up the World Wide Web (WSRI) [3]. The method of this paper and Web Science method have similar aims. This paper shows that Web Science method is useful not only for High School education but also for guardian's education.

2 BACKGROUND

On January 31, 2009, Ministry of Education, Culture, Sports, Science and Technology (MEXT) notified all Prefectural Boards of Education in Japan that mobile phones are now

banned in principle in Elementary and Junior High Schools, and no use of mobile phones allowed in High Schools [4]. High school students most often use a Profile home page service, which is a free home page service for an individual person on the mobile phone internet (Figure 1).

Fig. 1: Profile homepage



Bullying using "Profile" homepage password

Profile home pages are a breeding ground for cyber-bullying now. They often use a social networking service (SNS) too. "MOBAGE Town" is a free game site on the mobile phone internet. 50 percent of Junior High and High School students use it (April 2008). It has the same function as a dating site. Most of Junior High and High School bullying are related to profile home pages and SNS. They provide the students with opportunities for contact with harmful information and adults with inappropriate intentions.

3. PROBLEMS OF MOBILE PHONES

3.1 The crux of the problems with mobile phones

The most remarkable characteristics of mobile phones are their mobility and high performance. Children can use mobile phones privately and are able to accomplish many extraordinary things without adults knowing.

The biggest reason that cyber-bullying and illegal acts are continuing is the misunderstanding that the net is anonymous. Students do not understand what strong

weapons mobile phones and the Internet are. They can not realize how many people read their messages and pictures. Once the message has been sent on the mobile phone internet and the Internet, it is hard to delete them because of the constitution, “secret of communication”, “freedom of expression”, “freedom of speech” and the provider law. Even though the sender is detected and the original information or data defaming others is deleted, a defaming message, or illegal personal information or pictures, which seriously invade people’s human rights, have been already copied repeatedly through the bulletin board system and peer-to-peer file exchanging software. Actually it is impossible to delete them all.

3.2 The limits of school guidance

It is essential for guardians to instruct their children in use of mobile phones and the Internet to prevent cyber-bullying. Schools do not even have the authority to ask providers and administrators of the sites to delete the illegal defamations on the net, because of the laws and regulations (Figure 2).

Fig. 2: School and mobile phones

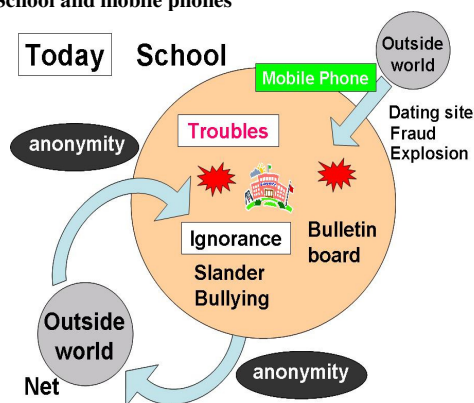
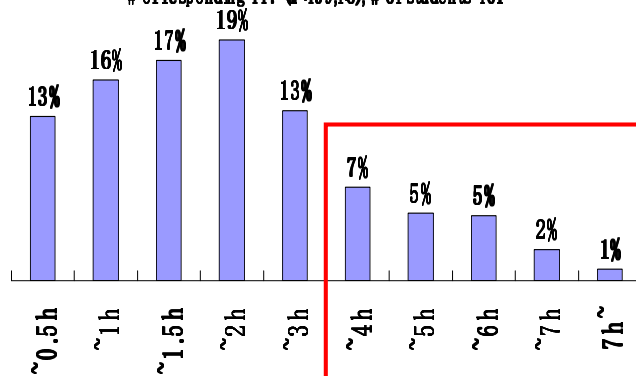


Fig. 3 :

Students: How many hours do you use your mobile phone?

Questionnaire: Shimonoseki Technical High School, June 2008
of responding 447 (n=439, f=8), # of students 461



3.3 The biggest reason guardians can not instruct

Guardians have the lack of knowledge on technology and Information Society. It is the biggest reason that they can not instruct on the safe use of mobile phones. Guardians do not understand the increased importance of mobile phones in a student’s life. They can hardly keep up with teenagers’ use of mobile phone. 20 percent of students use mobile phones for more than 3 hours a day (Figure 3).

Guardians do not understand how students use the mobile phones and the Internet. Only 47 percent of guardians know the actual use by students [5].

A fixed sum system of paying for mobile phones makes it difficult for guardians to know the exact use by students. They are confused about the student’s way of thinking about mobile phones and the anxiety over the threat of troubles caused by it.

Guardians should have responsibility of minors’ use of mobile phones and the Internet. School have to make their position clear, inside and out-side of school, that guardians should make themselves responsible for the use of mobile phones and the Internet by students.

4. PURPOSE OF THE STUDY

School help students to learn information morality through case studies and understand how our information society will be changing.

Students learn how information technology influences people, their daily life, the market and the law. They learn to recognize the information society as a system composed of many social factors. Then they will be able to adapt themselves to deal with unknown situations more effectively in the future.

School help students to understand the intention of the information sender on the Internet by asking them why these cases or events have happened. We can prevent students from getting involved in troubles as a victim or a perpetrator, resulting from a lack of knowledge of the Internet and information technology.

School help students and parents to obtain the ability to adapt themselves to the information society, which is rapidly changing. School help students to obtain the ability and aptitude to make judgments using their own initiative without being confused by transient incidents or social trends.

4.1 Merits of a Web Science approach in school: Moving from a ban to understanding information society

Students neglect school instructions to keep safe on the Net. They do not like any kind of moral education and never listen to serious instructions of principles. A merit of a Web Science approach is that it does not depend on compulsion of morality or prohibition. Learning the events or cases, students understand why they happened, or think how technology related to them. Then they come to recognize the changes of society as a process of social development,

they gain an insight into information society, and adapt themselves to the changes of a society. An insight into a technological society is useful to decide their courses in the future. It helps students to choose an occupation that they will not loose in 10 or 20 years hence.

5. TARGET AND DEVELOPMENT OF A WEB SCIENCE METHOD

The followings are the aims of this method.

- (1) To put an emphasis on timely instruction.
- (2) To reduce the frequency of moral education.
- (3) To keep a way to always provide information to guardians.
- (4) To utilize and maximize the instructions of guardians.
- (5) Guidance to be short.
- (6) To require no prior knowledge on Web Science.
- (7) To be used repeatedly.
- (8) To be a Web Science approach.

6. METHOD

6.1 Information System

School provides Information System for complete prevention and understanding of information society (Figure 4). Information and warning of the Net are given to students and guardians with leaflets at every opportunity, morning homeroom class, classes, long homeroom class, school meetings, PTA committee, PTA general meetings. Emergency staff prevents repeated suffering and secondary victims. Emergency staff solves cyber crime at the early stage, keeping victim's privacy.

Every homeroom teacher hands out leaflets to students during their morning homeroom sessions, and comments on a story just for one minute. Sometimes, the teacher will caution the students not to bully with mobile phones or the Internet. After students read each leaflet, they hand it over to their parents. We help parents to understand new technologies and what students are doing at school or in their daily lives. Every year, we edit all of the distributed leaflets for that year and write an Annual Textbook of Information-related Education.

6.3 Guidance given to new freshmen and guardians prior to starting school

The Annual Textbook will be given to all new freshmen and their parents of next year when they attend a briefing. They are promoted to understand the danger of the Net with the Textbook and to talk about that before entering High School. Freshmen have less knowledge on mobile phones and the Internet. They are a high-risk group, who can easily become involved in trouble. This intensive instruction targets the high-risk group.

6.4 An immediate and intensive instruction

An immediate and intensive program of guidance is given to freshmen in their first three months. Freshmen are taught a program about characteristics and dangers of the Net again and again. The program prevents freshmen becoming either a victim or perpetrator.

6.5 Contents of leaflet

At least one quiz is made to present a problem, and let students think why the event happened and what the point of the problem is. We have issued 220 leaflets science 2002. Contents cover the following;

- (1) Events in school
- (2) Risk information
- (3) About cyber-bullying
- (4) Request of guardians' instruction
- (5) Explanation on information technology
- (6) Impact given to industry by information technology
- (7) Information-oriented society
- (8) The future built by technology

7. RESULTS AND CONCLUSION

The methods using leaflets are very effective. School has practiced these methods for three years, and could get rid of all trouble caused by mobile phones and the Internet in school. Leaflets are so flexible that the school can use it for any purpose, in combination with various methods. This practice proved that these combinations have been very effective. We can make a leaflet quickly, take it anywhere and teach immediately. The merits of flexibility, which class lessons do not have, have extended the target from students to guardians, and increased the chances of timely intervention.

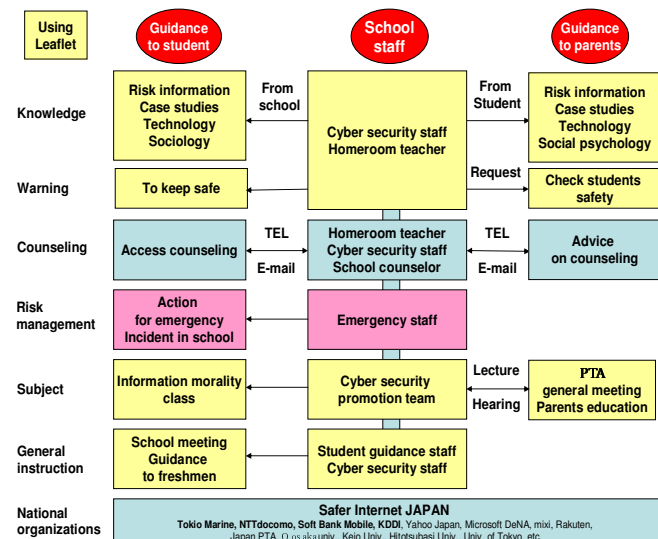


Fig. 4 Information System

6.2 Guidance with leaflets in morning home room class

We have published and produced a regular series of leaflets, "The SEIKO Good Net News for Family" [6].

All results show that there is no difference between the basic concept of prevention against cyber-bullying and that of traditional bullying. The different point is that significant knowledge of the power of technology is necessary to prevent cyber-bullying.

7.1 Effect of timely publishing

As soon as a serious event or a dangerous case takes place, school stopped students having to face the same risks as these cases (Pic. 1). Repeated warning to the freshmen year and reduced exposure of personal information on the Net, which is often sent with mobile phone cameras from school.



Pic. 1: A lecture to students



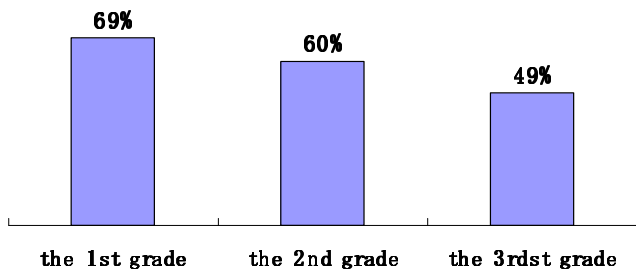
Pic. 2: A lecture to parents

With a leaflet school give information to guardians at anytime (Pic. 2). As soon as a profile homepage caused a suicide, school informed both guardians and students details of an event, and gave instruction and guidance to check the students' use of mobile phones is safe (Figure 5).

Fig. 5

Rate of guardian who guideedd children according to warning of school

Questionnaire:
Shimonoseki Technical High School, Dec. 2008
of responding 387, # of guardians 434



7.2 Campaign for communication with guardians

73 percent of guardians think that school guidance with leaflets is useful (Figure 6, Figure 7).

Fig. 6:

Guardians: Are leaflets usefu? Questionnaire: Shimonoseki Technical High School, Dec. 2008 # of responding 401, # of guardians 434

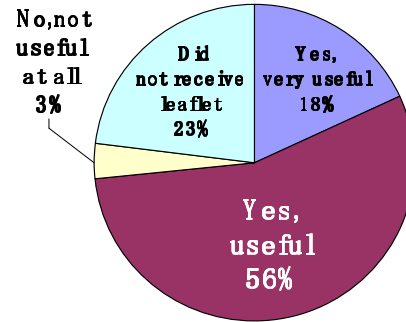
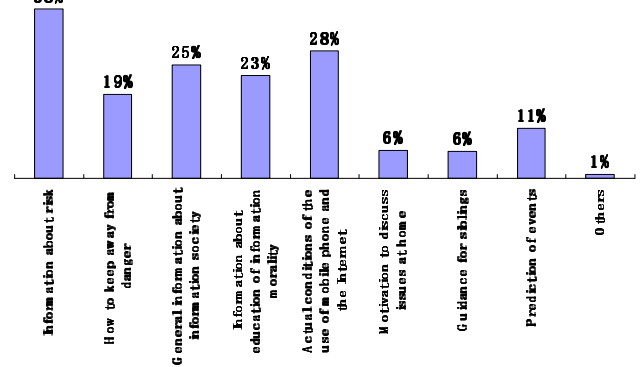


Fig. 7:

Guardians: What content of leaflets is usefu? Questionnaire: Shimonoseki Technical High School, Dec. 2008 Multiple answers: # of responding 414, # of guardians 434



As a result, leaflets give an opportunity to talk about the use of mobile phones and the Internet at home. In response to the schools request for safety checks, more guardians discussed safety and use of mobile phones with students.

7.3 Monitoring the effect of our guidance

According to the survey, 41 percent of students don't keep to the school regulations about using their mobile phone in school [7]. Also, only 45 percent of parents agreed to students using mobile phones at school [8]. Therefore, mobile phones remain banned at school.

8. ISSUES AND SOLUTIONS

Actually, there is the possibility that cyber-bullying goes underground, continues secretly and is becoming more serious. School keeps sufficient communication with guardians. School has established a system that guardians inform the school if they suspect any symptoms of cyber-bullying at home. 25 percent of students do not read the leaflet [9]. We take a countermeasure that the homeroom

teacher urges students to read leaflet and give short comment.

A leaflet can not substitute a class. It is important that essential instruction should be done in a class and a whole school meeting too, especially a whole school meeting has been very effective. A whole school meeting can make students understand the matter is a serious problem.

25 percent of guardians do not receive the leaflet at all. Students do not pass them on to parents [10]. Students don't want to pass the warning information to ban the use of mobile phones and the Internet. We take a countermeasure that school send a leaflet to guardians through their e-mail.

References

- [1] J. Hendler, N. Shadbolt, W. Hall, T. Berners-Lee, D. Weitzner, *Web science: An interdisciplinary approach to understanding the World Wide Web*. New York: Communications of the ACM, Volume 51, Issue 7 (July 2008)
- [2] H. Yasuda, *SEKIKO good-net NEWS*. Shimonoseki Technical High School, Shimonoseki, Japan, 2002 (http://sweb.nctd.go.jp/g_support/others/k_ys_s01.pdf)
- [3] Web Science Research Initiative (<http://webscience.org/>)
- [4] MEXT, *Survey on school bans on mobile phones*. Tokyo, 2008 (http://www.mext.go.jp/b_menu/houdou/21/01/1234723.htm) (http://www.mext.go.jp/b_menu/houdou/21/01/_icsFiles/afieldfile/2009/02/02/1234723_1_1.pdf)
- [5] Shimonoseki Technical High School, *Survey on guardians 'opinions about use of mobile phones*, Shimonoseki, Japan, July 2007, Questionnaire: # of responding 360, # of guardians 479
- [6] Hirohiko Yasuda, *SEKIKO good-net NEWS for family*. Shimonoseki Technical High School, Shimonoseki, Japan, 2007 (<http://www.jasrac.or.jp/seminar/5th.html>) (http://www.jasrac.or.jp/seminar/pdf/5th_pdf010.pdf)
- [7] Shimonoseki Technical High School, *Survey on students use of mobile phones*, Shimonoseki, Japan, June 2007, Questionnaire: # of responding 300, # of students (1st, 2nd grade) 309
- [8] Shimonoseki Technical High School, *Survey on guardians 'opinions about use of mobile phones*, Shimonoseki, Japan, July 2007, Questionnaire: # of responding 378, # of guardians 479
- [9] Shimonoseki Technical High School, *Survey on Information Morality of students*, Shimonoseki, Japan, Dec. 2005, Questionnaire: # of responding 524, # of students 560
- [10] Shimonoseki Technical High School, *Survey on guardians 'opinions about use of mobile phones*, Shimonoseki, Japan, July 2007, Questionnaire: # of responding 374, # of guardians 479.

A NEW MODEL FOR ONLINE READING COMPREHENSION RESEARCH

S. Ottaviano, A. Chifari, L. Seta, G. Chiazzese, G. Merlo and M. Allegra

Institute for Educational Technologies

Italian National Research Council

Via Ugo La Malfa, 153, Palermo - Italy

e-mail: [simona.ottaviano; antonella.chifari; luciano.seta; giuseppe.chiazzese; gianluca.merlo; mario.allegra]@itd.cnr.it

ABSTRACT

In this paper some important aspects related to on line reading comprehension are discussed; in particular, we describe a new model to orient researchers in the different dimensions and interactions involved in the development of new hypermedia educational support tools for enhancing online reading comprehension. The model named Interactive REading Comprehension (IREC) is represented by a triangle whose vertices symbolize three different dimensions: online reading comprehension instruction, web tool features and users. The main aim of the proposed model is to stress the relationships between these various aspects and their impact on the two principal processes characterizing any study activity on the web: surfing and comprehension.

1 INTRODUCTION

The last three decades have seen significant cultural and economic changes, many of them influenced by the introduction of new Internet technologies, which have had an effect on the way in which people approach reading, process contents and interact with different media. In the knowledge society, reading means dealing with complex scenarios like the Web and extending reading skills beyond the confines of a book. Blogs, wikis, virtual libraries and, in general, hypermedia information are online reading resources which students can use for learning. Online learning activities on the Internet and hypermedia are characterized by the action of two cognitive processes applied by students: the *reading comprehension process* and the *surfing process*. In this domain, cognitive strategies for orienting one's self in hypertext reading compete for cognitive capacity that might otherwise be devoted to the comprehension of text information [1]. *Reading comprehension* is a complex active process of constructing meaning that is *interactive*, *strategic* and *adaptable*. It is interactive because it involves not just the reader but also interaction with the text in which reading takes place [2]; it is strategic since readers have goals that guide their reading and they use different strategies and skills as they construct

meaning [3, 4]; it is adaptable, because readers change the strategies they use as they read different kinds of text or as they read for different purposes [5]. At the same time we define the *surfing process* as a complex, active process of constructing paths and finding directions. It is *interactive*, because it involves hypertext links through which browsing takes place; it is *strategic* because surfers have information to find that orients their choice of links and they use different strategic skills while they are following a path; it is *adaptable* since surfers change their strategies according to the design characteristics of different hypermedia interfaces. The combination of these cognitive strategies employed by students during hypermedia learning and online learning in general changes the architecture of acts of reading [6] and consequently learning. We prefer to call this scenario the *online reading comprehension* domain. Therefore, skilled online readers should be able to balance both the demands for comprehension and for thoughtful navigation to provide direction in hypermedia [6, 7].

Recent studies [8, 9] indicate clear differences in the quality of strategies used by readers in traditional and online reading contexts and suggest introducing new literacy into the reading comprehension instruction curriculum.

The need to promote reading literacy is a central issue in lifelong learning. As a consequence, psychologists, computer scientists, researchers and educators are studying online reading processes and the evolution of reading comprehension and surfing skills in web-based learning environments and how hypermedia educational tools can contribute to the development of online reading comprehension. The present paper follows this direction of studies and aims to provide a model to orient researchers in the different dimensions and interactions involved in the development of new hypermedia educational support tools for enhancing on line reading comprehension. The following section presents the model and its characteristics.

2 THE MODEL

The model of Interactive REading Comprehension (IREC) can be represented by a triangle whose vertices symbolize three different dimensions: *Online reading comprehension instruction*, *web tool features* and *users*. The first vertex

represents the main theoretical aspects of reading comprehension instruction, the second one is focused on the characteristics of web-based tools and their relationship with the reading process, while the last vertex describes the learners' characteristics in terms of *prior strategic knowledge* and *prior contents knowledge*.

The model is based on the most recent theories according to which skilled readers are able to balance both the demands for comprehending and for orienting themselves in hypertexts [9]. This concept must be borne in mind while providing instruction, designing web based tools and evaluating user characteristics (Figure 1).

This paper focuses on the *web tool features* vertex and how the features can be suitably developed according to the interaction with the characteristics of the other two vertices.

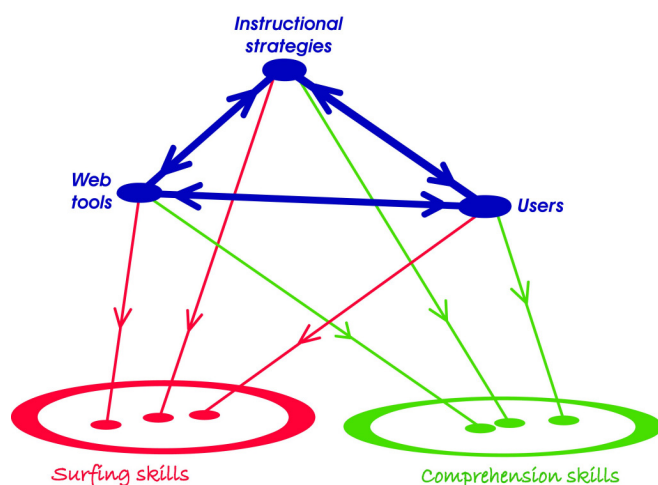


Figure 1: The graphical representation of the model of Interactive Reading Comprehension.

2.1 Vertex: Online reading comprehension instruction

The ongoing technological evolution requires the acquisition of new online reading comprehension skills and the consequent redefinition of many aspects of traditional comprehension instruction [10]. In fact, according to Coiro e Dobler [11] online reading comprehension is defined not only by the purpose, task and context but also by a process of self-directed text construction that occurs as readers navigate their own paths through an infinite informational space to construct their own versions of the online texts they read.

While it is still true that instruction is effective if it is based on modeling processes and is well organized, explicit, intensive and long lasting, and if students are made aware of text organization and are motivated to read widely [12], new instructional principles are based on certain key elements which make offline and online reading significantly different. The first of these elements is related to the hypertextual structure of digital resources in which «readers construct individual texts through hyperlinks and the unique textual paths each chooses to follow on the Internet» [13]. The second element concerns the approach

to reading which requires additional reading comprehension strategies [13] in online reading contexts: generating online research questions, locating relevant and reliable information, critically evaluating information, synthesizing ideas across multiple texts and media, communicating and exchanging online reading comprehension strategies. Online comprehension strategy instruction can be organized into a three-part framework, as in a traditional context, with specific activities used before, during and after reading to enable the student to become competent in recognizing how, when and where to flexibly apply and modify strategies.

Independently of the models such as peer tutoring, collaborative learning, reciprocal teaching, etc. that can be applied in the definition of the instructional framework of an online reading comprehension program, and combined in different ways depending on the characteristics of the other two vertices, an online training program must promote the students' awareness of their study strategies and how to control them metacognitively.

2.2 Vertex: web tool features

The *web tool features* vertex aims to describe some basic principles that should be taken into account when designing new web-based environments.

According to Leu et al. [13], «online reading comprehension is not isomorphic with offline reading comprehension». In fact, the online comprehension activity requires the management of:

- 1) *Surfing processes* that include all the processes that a user carries out while managing his/her online activities. For instance, the skills required to use search engines and browsers and to identify the key elements of a hypertext (a link, an image, a menu entry, a form, a filter, etc.).
- 2) *Comprehension processes* that refer to the cognitive and metacognitive processes involved in knowledge construction based on hypermedia resources.

Afflerbach and Cho [6] underline the importance of balancing these two kinds of processes to promote both meaningful content comprehension and thoughtful navigation.

Starting from these considerations, some basic principles to take into account during the design phase of a new web-based comprehension tool are:

- 1) setting out a clear purpose for the tool;
- 2) identifying a well-defined target and taking into account the users' characteristics (see the *Users* vertex of the IREC model);
- 3) identifying a comprehension instructional model and evaluating how it could be applied in a web-based environment (see the vertex *Online reading comprehension instruction* of the IREC model);
- 4) balancing surfing and comprehension features according to the established goal (Table 1 provides examples of features in each domain);
- 5) including motivational features/activities to promote greater user participation [14].

Surfing process features	Comprehension process features
Aids for monitoring surfing behaviour (e.g. navigation maps)	Aids for monitoring comprehension (e.g. answering questions, prompts)
Aids for using a search engine (e.g. aids for choosing the keywords, identifying the best query results, evaluating web credibility)	Aids for organizing contents graphically (e.g. conceptual maps, flow charts)
Aids for managing web page storage (history, bookmarks/social bookmarks)	Aids for identifying the main concepts of web pages and summarizing information (e.g. highlighters, notes)

Table 1: *Examples of possible features to support both comprehension and surfing processes.*

2.3 Vertex Users

This vertex guides the development of features of hypermedia and web-based tools in the construction of users' profiles. The design of an effective support tool to help students in their learning process should take into consideration the users' characteristics in order to provide effective aids according to the students' surfing and reading comprehension proficiency levels. We believe that an important element for the construction of effective tools is the assessment of user characteristics in terms of prior strategic knowledge and prior contents knowledge (applied to the online reading comprehension domain). The first type of knowledge is related to the introduction of tools to assess students' strategic metacognitive background with reference to their knowledge of surfing and reading comprehension strategies and their ability to self-regulate surfing and reading comprehension processes by applying the strategies during learning activities.

The second type of knowledge regards students' prior knowledge of the topics they are going to read. This type of knowledge has a direct effect on the sensitivity of learners to the organizing characteristics of hypermedia tools [7]. Learners with little prior knowledge of a topic are more sensitive to the hypermedia system characteristics while learners with a sound knowledge are less sensitive. We define the online reading comprehension proficiency level as a qualitative or quantitative indicator of this type of knowledge. We argue that the construction and introduction of new assessment tools to assess the online reading comprehension proficiency level of students can have a positive effect on the definition of student profiles and on the development and introduction into the system of specific aids, such as training activities to enhance this type of knowledge, designed and provided ad hoc to the students.

3 THREE VERTICES AND TWO PROCESSES

The main intention of the IREC model is not to highlight the most important aspects for the success of an educational activity in improving online reading comprehension, but rather to stress the relationships between these various aspects and their impact on the two principal processes characterizing any study activity on the web: surfing and comprehension.

While it is sufficiently clear that the choice of an instructional model must be the result of a careful evaluation of the characteristics of both the users and the technological tools used, it is less evident that also an assessment of users cannot be made on a positivistic base either: users should be assessed according to the instructional model applied and the web tools used. In the same way, the technological design must consider the educational context and, in particular, didactic methodologies and users' characteristics.

But the model also wants to emphasize that these characteristics are not static, they evolve over time and so some specific tools are needed to keep pace with this evolution. Generally, this aspect is discussed in relation to the rapid changes in information and communication technologies: the development of new techno-based systems appears to be a major stress factor in the educational environment, with teachers and students running a non-stop race to acquire the latest novelties. Little attention is paid to the parallel evolution of pedagogical methodologies and users' behavioral habits. One of the reasons why the diffusion of some technology based educational tools has also been possible is because of the increasing familiarity of learners and teachers with new modalities of interaction as well as the development of pedagogical approaches based on, for example, simulation and visual knowledge management.

Moreover, an evaluation of the various model dimensions needs to recognize that they have different effects on surfing and on comprehension. For example, any operation to improve features of the surfing process might have a negative effect on comprehension, and vice versa. So, web tools designed to facilitate the storage of web pages, a typical surfing feature, might limit the students' ability to identify the main concepts of a text, an important comprehension feature (see table 1); likewise, tools designed to organize the contents graphically, a comprehension feature in table 1, might hamper orientation on the web, a surfing aid for monitoring surfing behavior.

All these effects need to be carefully evaluated and the IREC model may prove to be a useful guide for the identification of the most relevant variables involved in a web-based didactic experience.

On the basis of this model it will also be possible to develop a meta-language to describe the online reading comprehension process, to define relevant dimensions, and to guide the design of web based tools for improving

navigational and comprehension skills of students engaged in online activities.

4 CONCLUSION

The rapidity of technological change and the more and more frequent use of Internet for educational purposes have increased the learning demands for comprehension and for thoughtful navigation.

Our paper addresses some issues related to the development of new educational hypermedia for enhancing online reading comprehension, starting from the analysis of the relationships among the characteristics of instructional strategies, the web tools features and the users. This is still an open research subject aiming at the development of new educational tools that should help students to become good and effective readers during online surfing.

References

- [1] Eveland , W., & Dunwoody , S. (2000). Examining information processing on the World Wide Web using think aloud protocols. *MEDIA PSYCHOLOGY*, 2(3), 219-244.
- [2] Heilman, A., Blair, T. R. & Rupley, W. R. (1998). Principles and practice of teaching reading. Upper saddle River, NJ: Merril/Prentice-Hall.
- [3] Baker, L., & Brown, A. L. Metacognitive Skills and Reading. (p. 74). Technical, Cambridge, Mass: Bolt, Beranek and Newman, Inc. retrived May 29, 2009, from http://www.eric.ed.gov/ERICWebPortal/custom/portlets/recordDetails/detailmini.jsp?_nfpb=true&_ERICExtSearch_SearchValue_0=ED195932&ERICExtSearch_SearchType_0=no&accno=ED195932
- [4] Wasik, B., & Turner, J. C. (1991). The development of strategic readers. In *Handbook of Reading Research* (Vol. 2, pp. 609-640). New York: Longman.
- [5] Dole, F., Duffy, G., Roehler, L., & Pearson, P. (1991). Moving from the old to the new: Research on reading comprehension instruction. *Review of Educational Research*, 61(2), 239-364. doi: DOI: 10.3102/00346543061002239.
- [6] Afflerbach, P., & Cho, B. (2009). Responsive Comprehension Strategies in New Traditional Forms of Reading. In *Handbook of Research on Reading Comprehension* (pp. 69-90).
- [7] Shapiro, A. M.. (2008). Hypermedia design as learner scaffolding. *Educational Technology Research and Development*, 25(1), 29-44.
- [8] Coiro, J.& Dobler, E. (2004). Investigating how less-skilled and skilled readers use cognitive reading strategies while reading on ihe Internet. *Proceedings from 54th Annual Meeting of the National Reading Conference*, San Antonio, Texas.
- [9] Leu, D., Zawilinski, L., Castek, J., Banerjee, M., Housand, B., & Liu, Y. (2008). What is new about the new literacies of online reading comprehension? In *Secondary school literacy: What research reveals for classroom practices* (pp. 37-68). L.S. Rush, A.J. Eakle, & A. Berger (Eds) Urbana, IL:NCTE/NCRL.
- [10] Coiro, J. (2003). Reading comprehension on the Internet: Expanding our understanding to encompass new literacies. *Reading Teacher*, 56, 458-464.
- [11] Coiro, J. & Dobler, E. (2007). Exploring the comprehension strategies used by sixth-grade skilled readers as they search for locate information on the Internet. *Reading Research Quarterly*, 42, 214-257.
- [12] Texas Education Agency (2002). *Comprehension Instruction*. Retrieved June, 15, 2009, from http://ritter.tea.state.tx.us/reading/products/TRI-Comprehension_Instr.pdf
- [13] Leu, D. J., Coiro, J., Castek, J., Hartman D., K., Henry, L., A. & Reinking, D. (2008). Research on Instruction and Assessment in the New Literacies of Online Reading Comprehension. In Block, C. C. & Parris, S. R. (Eds), *Comprehension Instruction: Research-Based Best Practices* (321-341). New York: The Guildford Press.
- [14] Giacoma, G. & Casali, D. (2008). *Design motivazionale. Usabilità Sociale e Group Centered Design*. Retrieved June, 15, 2009, from <http://ibridazioni.com/2008/11/01/design-motivazionale-usabilita-sociale-e-group-centered-design/>

LANGUAGES AND LANGUAGE: THE GREEK CASE

Maria Bontila⁽¹⁾, Vassilios Dagdilelis⁽²⁾

⁽¹⁾ History of Education, Aristotile University of Thessaloniki,

Tel & Fax +30 2310 923 247

e-mail: bontila@edlit.auth.gr

⁽²⁾ Department of Educational and Social Policy, University of Macedonia

Egnatia Street 156, 540 06, Thessaloniki, Greece

Tel: +30 2310 891 336 Fax: +30 2310 891 388

e-mail: dagdil@uom.gr

ABSTRACT

The phenomenon of globalization constitutes a stable point of reference in the analysis of many social sciences. Globalization presupposes a homogenization of the market and its exchanges, among other things. Language is called upon to serve this new reality and is required to adapt to unrestricted competition, obeying indirectly the rules of the free market. At the same time, communication technologies greatly facilitate this process. In our paper, we propose a study of the greek language (a rather “weak” language) and its relation to globalization and to “strong” languages. We will try to generalize our findings.

1. INTRODUCTION

Globalization not only constitutes a typical point of reference in the analysis of many social sciences but seems to displace the term “post-modernism” (as observed by many researchers). Globalization, as an analytical concept, in some cases appears to have a productive character whereas in others it appears to have a rather restricting one, as often happens with related concepts.

By “globalization”, as a rule, we consider two sets of phenomena, which are similar but not identical. On the one hand, we consider that globalization is a expansion process of social phenomena at a global level. A basic characteristic of *this* globalization is homogenization: phenomena on a global scale are inclined to acquire a homogenous form and a common basis, either through processes of convergence, or in most cases, through processes of an almost *natural selection*, where the survival of the strongest prevails. Globalization, for example, imposes a uniform way of how credit cards operate, as well as homogenous protocols of communication between computing systems, in order to enable the existence of the Internet. Noteworthy is the fact that in one case the uniformity of credit cards most likely derives from a gradual, mandatory convergence of all the different types and forms of credit cards (a process of natural selection), whereas in the second case there is a specific organization which is concerned with welfare. As it is natural, the predominance of the “strongest model” gives rise to concerns for the fate of the “weak models”.

Concurrently, at another level, by “globalization” we mean both the worldwide transportation of people and goods as well as the accomplishment of international communications, which are constantly becoming faster, more efficient and cheaper.

It is within this double framework of globalization, where the pressing issue of language arises. In other words, the fact that there is concern at an international level for the future of those languages with a weaker position in contrast to the allegedly “strong” languages, of which American -English is first and foremost.

In this paper the particular character of the relationship between “strong” and “weak” languages in Greece is presented. In the first section of the paper we refer in brief to the terms under which this domination of American-English appears. Then we examine the particularity of the Greek language, and finally we present the conclusions of our analysis.

2. THE DOMINATION OF THE “STRONG” LANGUAGES

As referred in the introduction, globalization presupposes a homogenization of the market and its exchanges, among other things. The uniformity, which tends to prevail in production, the market and the consumption of goods, also demands mass means of communication. Language is called upon to serve this new reality and is required to adapt to unrestricted competition, obeying indirectly the rules of the free market. At the same time, communication technologies greatly facilitate this process.

The outcome of this new situation is the worldwide trend of certain so-called “strong” languages to prevail, comprising the *linguae mundi* with English as the leader becoming the new *lingua franca*. Nowadays, English has precedence over all the other languages due largely to the economic superiority of the countries, which have it as their official language. In addition, it forms the instrument of technological inventions, research, and science, as well as being the dominant language in trade transactions.

Furthermore, these “strong” languages representing the powerful nations of the world, and dominating the “weak”

ones, mirror intensely the existing situation of the free market economy. Therefore, a series of questions arise: What is the future of the less “strong” languages? Will they be able to maintain their independence? Will they continue to play a role in their own societies or will they be gradually marginalized and as a consequence dwindle? Can this happen at a time when language is identified with the nation and constitutes a vehicle for national values?

All countries and communities are aware of the value of language and its symbolic role in the creation of national identity. For this reason they have developed a protective rationale for their language, which may often reach the point of presenting a nationalistic character. In addition, they have generated a scheme of the dangers concerning the threat to national identity and the loss of national ideals, since language is also considered responsible for the acquisition of ones specific identity and the preservation of national unity. Even the most radical and open-minded react against the cultivation and adoption of a supra-national language, which will restrict their own. Many countries have developed a language policy related to education and in particular institutionalized education. They display, at least in words, a language protectionism, which however, is not always successful or effective.

Another issue is related to the way that the “strong” languages dominate the “weak” ones. In the past, within a spirit of cultural expansionism, the powerful countries would exercise a specific language policy, setting up foreign language schools, (Colleges or Schools for instance) where their language could be taught, such as the British Council or various American Colleges in Greece, also offering scholarships to students attending their institutions. Now within the framework of the global market they have changed to strategies of indirect domination by opening up other channels of language and cultural input. In this way a network of interdependent interests has developed concerning products, ideas, arguments and ideologies. The trend is to create a global culture in conjunction with the production of a global market with global customers.

3. THE NATURE OF THE “DANGER”

It is perhaps desirable to define in more precise terms the nature and the extent of the “danger” which weak languages face. We must mention that on the issue of language dominance, we do not mean language borrowing, because this has always existed among spoken, living languages for reasons well-known and much discussed. We are of course pertaining to the intrusion of a language to such an extent as to isolate or restrict the use of the local language. We refer to language conflicts, which will most likely result in the disturbance of the global linguistic ecological system.

Economic and cultural superiority is what alters the language landscape, reinforcing the “strong” languages and downgrading the “weak” ones. Could these fears, however, be exaggerated and the influence of a “strong” language, for example, be limited to the use of some words or phrases by a

certain section of people with particular characteristics? Linguistics accepts that *units of loan* derive as much from the productive sector, like science and technology, as from sectors to do with people’s daily life, such as music, fashion, sport, entertainment, leisure time [2, p. 280]. These units of loan, however, do not always fill a lexical gap, as at various times it has been stated, but most times they co-exist with the local words. Usually, they meet at the crossroads of youth’s culture and technology. The degree and the variety of use depend on the identity of each speaker; the wish for accurate expression, when there is no available word in his/her language or when the existing word is not adequate to express the desired meaning, are some of the reasons for using a foreign language. At other times this borrowing occurs as an exhibition of knowledge of that language, since using a foreign language has a symbolic value, contains an evaluation identifiable code, which informs others that the speaker is of a certain culture and probably of a social status; in other words, he/she is classed in the elite educational minority, which exudes glamour and prestige.

However, the point should not escape us that along with the “weak” languages, the “strong” are also under threat. Languages are of course vehicles of culture, but what type of culture does English convey in both the way and the extent to which it is being spoken in the new countries where it is received? The standard English used most times serves only superficial needs of communication and does not progress to a deeper level. Within this context the alleged strong languages are affected and thus are also in “danger”.

In any case, globalization needs the cooperation of technology (a rather regulatory framework), as well as symbolic power, symbolic authority.

In earlier times, when language was taught systematically rather than arrive in countries of reception as technical or technological terminology or by way of song verses, the cinema or advertisements, it was essential to cultivate it. And so, Shakespeare’s language, for example, would be revived again in different countries attaining, if not new glory, then at least admiration and/or at times awe. Nowadays, the learning and adoption of a foreign language operates mainly within the context of a *linguistic fast food style*, in other words, for purposes of superficial and fast communication. All corporations, for instance, require that their employees have a good knowledge of English, whose use however, is limited and serves an understanding and communication at a primary level. The senior executives, who form company policy, are usually the ones who have a better command of the foreign language. Furthermore, it could be said that since language is not socially innocent, it is perhaps more desirable for some if it were not spoken well to an adequate degree by all, because in this way those who lack sufficient knowledge of the language yield more readily to their superiors and in addition to this, they are not in any position to exercise criticism. Language becomes the means for a monologue and in this way protects the social and economic prerogatives of the powerful.

4. THE DIFFERENCE BETWEEN “WEAK” AND MINORITY LANGUAGES

It should perhaps be emphasized that minority languages are oddly enough less in danger than the more prevalent “weak” ones. In addition, the level of danger is different for each language. The closed, non-developing minority languages, are not so much “in danger” of globalization as they are of the effects of the formal national language of the country in which they simultaneously exist. The political and economic submission of a community whose identification is based on a language, causes the minority language to yield to the strong formal language. The formal, dominant speech is adopted and a linguistic shift is observed from the minority to the official linguistic norm. The two linguist codes may not come into direct conflict, however, one yields to the other gradually, and is limited to specific functions only within the boundaries of the minority group, obeying the demands of the linguistic market, which in turn, obeys and is adapted to the demands of the economy [4, p. 174]. So it seems that the open/weak languages are more under the threat of globalization, since the new invasion tends to be mainly cultural.

5. PARTICULARITIES OF THE GREEK LINGUISTIC ENVIRONMENT

After the introduction, which attempted to present concisely the general problem of “strong” and “weak” languages, we come to the example of Greece. Greek presents a series of particularities as far as language is concerned.

In our opinion, there are three main points, which characterize the particular Greek linguistic reality.

The first is the *diglossia*, i.e. the dual form of the language, which prevailed from 1832, the year the modern, Greek nation was formed until 1974, the year in which *Demotic* Greek was instigated as the official national language. During this time, there existed two languages concurrently *Katharevousa* (Purist Greek) the written, cultivated language and *Demotic*, the popular one. The so-called Purist Greek language (*Katharevousa*) was both syntactically and grammatically very close to Ancient Greek and had been enforced by the then governments, who driven by reasons of nationalism, in this way, projected the continuation of the race and Greece’s greatness based on its glorious historical past. At the same time, while the official language of the nation and of education was Purist Greek (*Katharevousa*), the populace, even those who were educated, spoke *Demotic* in their daily lives. Purist Greek (*Katharevousa*) was so different to *Demotic*, concerning the structure and form, that fifty years later, the young people of this generation do not comprehend it, believing it to be Ancient Greek. As a consequence we have reached the point today where literature written in Purist Greek (*Katharevousa*) (60-70 years ago) is translated into *Demotic* in order to be understood by the new generation (a translation from modern

Greek to modern Greek !). One result of this linguistic dichotomy is the partial inability to render contemporary or technical terms into modern Greek – since the 1.5 century-old language *savante* (*Katharevousa*) has now been marginalized.

The second particularity of Greek language is the following: the language is very closely connected to its past and the history of the nation. At the present time, although *Demotic* is the official national language and in the form of the monotonic system (as opposed to the polytonic) there exist many people who write using the older polytonic system and in so doing advocate the language’s historical formation. In spite of this, in the Greek mentality, foreign languages are regarded as an important professional qualification and are an essential aspect of education.

The third particularity of Greek is that it is in a disadvantageous position in regards to the other Latin based languages, due to the different alphabet it uses. Of course, there are 14 characters in the upper case and one in the lower case, which are common between the Latin and Greek alphabets, but often in a different order (“Z” is the 7th letter in Greek but the 26th in Latin). On the one hand this fact creates certain problems, such as the classification of words – as opposed to languages with a Latin alphabet. On the other hand, added problems are created for electronic communication. Greeks have devised a new form of Greek writing, that of *Greeklish* – greek words with latin characters -, which they use only when communicating on a computer. These elements compose a particular linguistic landscape in Greece.

After the induction of Greece into the European Community, the Greek language suffered a big cultural and linguistic shock in that it was believed it was threatened by the “stronger” European languages. At some point, in fact, the danger was even more manifest, when, in a spirit of economizing, the European Union announced that the Community’s decisions will not be translated into all the languages of the member states. The danger was overcome because the particular decision was revoked after strong protests and discussions in the European Parliament. However, now a new threat for language is lurking within the context of globalization. The invasion of English is intense and aggressive and is apparent in all areas of language functioning. Foreign shop signs, names of radio stations and television channels, titles of newspapers and magazines, songs and cinema, as well as common phrases intermingled with Greek, which are encountered on a daily basis in Greek society. Many have sounded the alarm, developing rhetoric and conjuring up the glorious past of the Greek language, its diachronic nature and the era when it was itself a *lingua franca*.

Notwithstanding, two observations need to be stated:

Greek, like all other languages, does not constitute one single core. For example, in the official dictionaries of the language no definitions of foreign words exist (even though naturally there are thousands of borrowed words, which have acquired a Greek structure and form). However, the

dictionaries express only one partial reality of Greece. Greek language, is composed, in part, of idiomatic expressions, such as the slang of young people, occupational jargon, the language of trade of technology. In this way, the invasion of the English language and of strong languages generally, does not contribute to the same degree in all cases. *RU ready 2 go?* (Are you ready to go?) is a very common SMS message [1, p. 35], however, naturally expressions of this type are practically non-existent in the academic field - of philology for instance. At the same time, however, in technical branches such as Informatics, the jargon is of a hybrid form. Thus, the phenomenon of the “invasion” of strong languages appears to be more complex than the single act of the total substitution of one language for another.

The second observation is related to the likelihood of “weak” languages actually benefiting from the new state of globalization (where globalization is here perceived in its second form: the globalization of communications). The development of digital communication allows today the creation of a special forum where all languages can participate. In other words, the opportunity is given to create at least one “space” where the “small” languages can also be used and developed by people who speak them, even though they may be dispersed in different parts of the world. Therefore, the “small”, “weak” languages, whether local or not, benefit from the Internet, within the framework of global communications. Due to the fact that there is a very large number of Greek migrants and Greek speakers around the world and also because there is an apparent relationship between modern and ancient Greek culture, the Greek nation is trying to develop its digital presence on the Internet.

6. CONCLUSIONS

The direct danger of the disappearance of weak languages, Greek included, does not appear to be so great and it is evident that there exists the possibility for regulations of control, as well as the prospect of support for the national languages. Once the linguistic and cultural areas, which are most under pressure from the “strong” languages, have been pinpointed, and the actual ways that the foreign languages penetrate the native ones have been detected, effective measures can thus be taken. Languages are preserved when they function and serve, express ideas and meanings, and that is why we must turn in that direction in order for every language to perform all the functions within the society where it was developed. The expression ‘*ou on parle francais on achete francais*’ certainly determines the relationship between language and the market.

Despite the invasion of English, however, we do not believe that the Greek language is under threat. The use of English is restricted at a first level more so to make an impression or from acquired speed, when used by young people from the songs they listen to or the video games they play. Scientists and professionals use English as well as other “strong” languages, including scientific and technological terminology, wherever necessary or where there is no

alternative expression in Greek language. This usually happens at the initial introduction of a product or concept into the country and with the passing of time the term or expression either gets translated or becomes Greek-a-fied, which is due to the fact that Greek has a large capacity for assimilation and is able to incorporate many foreign words into its grammatical system. Even when the foreign words are not assimilated, they are functionally but not however, phonologically or grammatically incorporated in their daily use.

Despite all the warnings from the lips of responsible cultured people, that in future the language will dwindle, it is difficult to establish or support such a threat. English serves practical communication needs, and it is with difficulty that it will ever become the cultural language of Greece. The Greek education system places great emphasis on the subjects of language, both ancient and modern, the teaching of which has been associated with the uninterrupted continuance of the country’s civilization and its past history. The language, perhaps, is more in danger of internal forces, such as the linguistic anarchy in its use rather than the invasion by foreign languages – in fact, the same causes seem to also threaten the “strong languages. The Greek language is endangered by users, who do not respect its structure and form or who are indifferent to the lack of respect shown by third parties. It is likewise more in danger from the Internet, the modern telematics and telecommunications, the visual information of image and the technical languages of Microsoft. A simple techno-English language is used, which imposes its own terms on global communication. The conflict, therefore, is not to be found between universal English and other native languages, but rather between the new techno-English language [4] and the national languages whose use of the former weakens the latter (American-English included).

The issue of language is therefore, clearly one of politics because it is through language, considered as a social institution, that hierarchies are propagated, identities are constructed and power is controlled. Besides, bear in mind that language in itself is a form of power.

References

- [1] M. Bontila. SMS generation: Grammar and Linguistic Choices of an Almost-Online Communication. *Proceedings 13th Conference of the Greek Applied Linguistic Association*. Thessaloniki, 2003.
- [2] C. G. Lazos. National languages and European coin. In A.F. Christidis (Editor), “*Strong*” and “*Weak*” languages in the European Union, Centre for the Greek language, Thessaloniki, 1997, vol.1, p.273-285.
- [3] L. Tsitsipis. Discourse, language shift, and mutual implications for linguistic anthropology and political philosophy. In A.F. Christidis (Editor), “*Strong*” and “*Weak*” languages in the European Union, Centre for the Greek language, Thessaloniki, 1997, vol.1, p.173-179.
- [4] K. Tsoukalas. The idiomatic language in a new world. *Greek Journal “To Vima”*, 6 August 2000.

AESTHETICS AND LOGIC, THE TWO MAIN BRANCHES OF ONE SINGLE TREE

Nicole Szendy

Ass. Résonances 75015 PARIS

e-mail: nicole_szendy@yahoo.com

ABSTRACT

The principal subject of this paper is the analysis of an analogy from the history of architecture in France. This analysis aims at helping to find a direction of establishing an optimal balance between the work of two hemispheres of the personality in the information society.

1 INTRODUCTION

Almost forty years ago, Dr. George Gaal invented a new concept which he called "Logical Aesthetics". He started from the fact that our brain is made of two completely separated hemispheres that join together at the back of our head in the spinal cord. As we have known for a long time that one hemisphere is the centre of logical reasoning and mathematical order, and the other one the centre of emotions and feelings that gives us the sense of beauty, Dr. Gaal developed the theory that there could not be harmony without taking into account these two aspects of our being.

Thanks in particular to Dr. Jill Bolte Taylor [1], a neuro-scientist working at the department of psychiatry at Harvard, we know much better how our two hemispheres are working. In fact, on the 10th of December 1996, Dr. Jill Bolte Taylor had herself a cerebral stroke and lost all her abilities to speak, walk, move or do anything. She had a strong haemorrhage on her left hemisphere, and after an operation, little by little, she managed to get back all her faculties. It took her about eight years to completely recover. But, as she is a neuro-scientist, now she can tell and explain with precision the process of a cerebral stroke and how our two hemispheres are working. She tells us that these two hemispheres have entirely different functions :

One, the right hemisphere works like a parallel processor and connects us with the present time, it is to say now and here. With this hemisphere we feel that we are energy beings and that we are not separated from all the energy that surrounds us. In fact, we can feel no limits between us and the world around, it is as if we were the soul of the world musical symphony, we feel linked, and we see, we hear, we taste all the beauty of the Universe.

The other one, the left hemisphere works like a serial processor. It is the past and the future, it classifies and organizes all the possibilities in connection with the outer world. It tells us what we are. Thanks to this hemisphere we

feel that we are isolated, solid and separated from the rest of the world. This hemisphere is a calculating intelligence.

By considering these two hemispheres, we can see that there are two plans of reality, that we have two cognitive minds, and that it is quite impossible to have a normal life or even to simply live if one of the two hemispheres is not working properly. The same way, we cannot walk with only one leg !

The principal subject of this paper is the analysis of an analogy from the history of architecture in France. This analysis aims at helping to find a direction of establishing an optimal balance between the work of two hemispheres of the personality in the information society. It seems that the search for the ways of establishing an optimal balance of the kind is just a reformulation of one of the principal goal of Cognitronics described in [2].

2 THE LOSS OF OUR NATURAL BALANCE

What is happening today ? On the one hand, many artists work with their right hemisphere in order to show and make us feel strong emotions, but unfortunately, it also happens that they wander away, and attracted by the showing off of instinctive urges, they may fall into a sort of chaotic sexual drift and lose their sense of beauty.

On the other hand, many intellectual tend to develop their left hemisphere and its logical deductions and inductions. This trend is supported by all the electronic devices which through their data processing help them expand their knowledge in quite an extraordinary way. So, with them, everyday we learn something new, we experience something new. In a way these people feel no limits to the extension of their knowledge and power. But here again another drift may appear : this calculating intelligence that they develop, suddenly neglects the realities of life, everything becomes virtual and then people do no longer see the consequences of such a frenetic development. Without being aware of it, we are erecting a new Babel tower where everyone with his performing language adds a new store to the building ignoring more or less the others and, all of a sudden we feel that the whole building is about to collapse !

But, why ? Mainly because our earth planet has become a global village, and, despite that, everyone, every group, whatever it is, is working for himself, for his own benefit and glory, and in his own direction. Therefore, our world appears at the same time very efficient and chaotic, extremely rich

and terribly poor. As a consequence, men and goods are moving in all directions. In this fierce competition for immediate profit, we make the most of everything, we make the most of our earth planet and forget that we are but passengers on the earth. In such a situation, whether we are rich or poor, anxiety prevails, we have lost our natural balance.

3 AN ANALOGY FROM THE HISTORY OF ARCHITECTURE IN FRANCE

To imagine what can be done to re-establish confidence and hope, and also to avoid conflicts, it might be interesting and useful to study how societies reacted when there were such important crisis and changes in their country. Quite a number of examples could be taken in different places, at different times (of the silk road for instance), but let us speak now of the French history.

In the twelfth century France experienced a very severe crisis. At that time, like today there were very important moves of populations, from the country to the towns, from their home-villages to Middle East crusades. An important part of the French population was on the road and a number of people, having no marks, went astray.

How did the clerical society manage to re-establish order and harmony ? The bishop and the secular clergy, who were the masters of the growing cities started to build large temples that they called cathedrals, that is to say, "the place where the bishop is seated". Everyone in the city and in the surrounding country participated in the building of these temples either with their money or with their work and more than one hundred cathedrals were built in France. Once the cathedral had been consecrated, it became the living centre of the city. It was a place for business, a shelter for the sick and the travellers, even for the thieves ; it could also be changed into a theatre, but it was above all an intellectual and spiritual centre.

How did this new architecture appear ?

First, there was in Chartres a school which had been created in the 11th century where students learned Plato and Socrates, and where mathematics were considered as the very link between God and the world. Geometry becomes Theology, but this geometry is also music and determines the cathedral measures according to musical ratios.

Second, during the crusades, ten knights of the Temple went to Jerusalem to study the ruins of Solomon temple and try and find the Ark of the Covenant. They did not only study those ruins but also met learned Arab people and studied their architecture which was much lighter than ours and which was lit by beautiful glass windows whose colours and particularly the blue were extremely pure.

The conjunction of these two trends led to the concept of a new art, called : "Art Goth" or Gothic Art. Argot in French means a language which can only be understood by those who have been initiated.

To understand all this, let us go to Paris and visit Notre Dame cathedral.

Let's imagine that we are now looking at the cathedral from the banks of the Seine river, the temple appears as an upside-down vessel whose flying buttresses would be the ribs of the boat. It is navigating in the heaven-sea and yet it is clinched to the earth of the city island which is itself in the middle of the river. The beautiful flying buttresses give the possibility of opening large windows in the walls and let light enter inside the temple.

Now we move to the main entrance on the West side. There are three gates and three different levels, above the gates in the middle, there is a circle with a rose window framed inside a square.

If we take into account that a cathedral is a reduced picture of the Universe, we will find out how important geometrical figures are in Gothic cathedrals [3]. There is : (Figure A/ A-B)

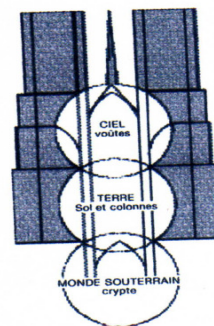


Figure A. Vertical view

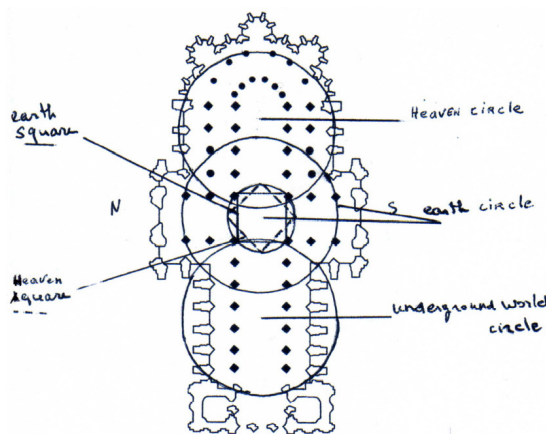


Figure B. Horizontal view

- the circle, symbolizing eternity and movement with the unity in its centre.

It is to be found almost everywhere in the cathedral. In the façade of course, but not only, if we take a vertical view of the temple and if we include the crypt, we will also have three circles (Figure A) representing :

- The underground world circle
- The earth circle
- and the heaven circle

Similar circles are to be found if we take a horizontal view of the cathedral (Figure A/ A - B)

- The square, symbolizing stability, realization and solidification.

Like the circle we have it in many places, in the façade particularly as a frame in which is inserted a rose window, but also in the very centre of the cathedral at the transept crossing. There, we have two very important squares, one is called Heaven square (Figure A/ B, dotted lines) which determines the orientation of the cathedral, and the other, earth square which determines the width of the nave.

- The triangle, especially the equilateral triangle whose angles make the conciliation of opposites become possible, represents God. With this triangle the ratio is fixed between the width of the nave and the one of the transept, as well as all the geometrical module of the cathedral.

In connection with them, we will see how important numbers are.

The Book tells : “God created everything with numbers, weights and measures”

3 - represents :

- the three levels of the universe : Heaven Earth and Underground world
- the three primary colours : Blue Red and Yellow out of which all the other colours can be obtained
- and also : Man :
- the mind represented by the number
- the soul represented by geometry
- the body represented by architecture

and God : the Trinity

- the Spirit
- the Father
- the Son

and the three aspects of Creation :

- Power
- Intelligence
- Love

4 – represented by the square symbolizes :

- the four elements : earth water fire air
- the four directions : North South East West
- the four Seasons of the year

4X3 =12 it is to say the Perfection of the Universe

5 – is at the very basis of the golden number and represents Harmony.

With it, we obtain the most beautiful proportions which are directly connected with the dimensions of our body, with our five senses, five fingers (four fingers plus one thumb),

four limbs plus one head, and with also five musical intervals ...

The cathedral represents the incarnation of God in Christ where every man who is also a Christ travels in the vessel of light to the meeting of God. (Figure C).

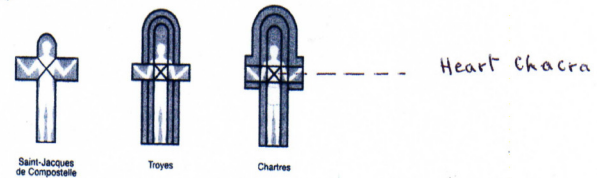


Figure C.

The link between God and man is made possible thanks to Mary, Our Lady, Notre Dame, who is at the same time Christ’s Mother and our Mother.

We can perfectly see that the shape of the cathedral is the one of Christ on the Cross, but also simply represents the image of Man :

The head being the choir

The heart being the transept crossing

The hands and feet being the doors

But let us enter the Cathedral and go through the Door of Saint Ann, which tells us Mary’s story-life and the birth of Jesus, that is to say our own birth.

Once in the cathedral, we first have to get used to the light, then instinctively, we walk to the middle of the nave and look at the vault. The nave and its vault sustained by powerful columns look like a tree-avenue whose branches by joining together in the centre make us think of a celestial vault. But, we go on, we are attracted by the light coming from the choir-windows. Only at the transept crossing we stop, because there, all of a sudden, lights are pouring upon us in the most magnificent way. They come from the dark-blue-rose-window on the North, referring to the Old Book, from the yellow-red narrow windows of the choir which usually refer to Christ’s death and also from the purple-rose-window of the South referring to the New Book ; after, almost instinctively, we turn round, and at the other end of the nave, we can see the West rose-window where Christ in Glory appears to us. The stained-glass windows are perfect jewels, they do not let the rays of the sun go through the glass but keep all this light inside them, that is why colours are so very bright, and what is more, that is why these colours perpetually change according to the weather outside and to the time of the day .

Always the same and always different. Whether we have a faith or not, we just marvel at so much beauty !

But imagine in the Middle Age, when people truly believed in God and could see all this, and, at the same time, could hear the choir singing (the music being attuned to the proportions of the temple) and could also smell the incense floating in the air, just imagine what they could feel ! It was

not their intellectual mind which was pleased but their heart which was bursting with emotion (Figure A/D).

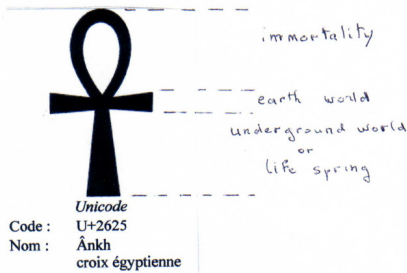


Figure D. Ankh – the Egyptian life-cross.

And, if we compare the image of Man and the one of the cathedral, we see that this place of marvel (the transept crossing) precisely matches with the place of our heart or more exactly with our heart-chakra. In an antique Egyptian Hymn to Atum, the Sun-God, we can read :

“What the Heart thinks, the Mouth tells”.

Yet, our journey is not over, we shall now walk around the choir and look at the different chapels, there, we shall see windows representing Christ’s death, symbol of our own death. Then, we shall come back to the transept and back to the entrance. This time, we shall leave through the door of Mary. Outside, on the lower part of the tympanum, we see Mary in “dormition”, it is to say : asleep after her death, while on the upper part she is Queen in Heaven sitting by her Son Jesus Christ. Between the two doors, the door of Saint Ann and the door of Mary, there is the Doom’s Day. All this makes it clear that this journey in the cathedral is in fact our life journey and symbolizes our walk and everybody’s walk towards Heaven.

It can remind us of another civilization - the Egyptian civilization and the Ankh (Figure D). The design of the Egyptian life-cross is very similar to the plan of the cathedral, here again, we find the three different levels as well as the intersection point which matches exactly with the heart-chakra.

4 CONCLUSION

So, we can see that French people in the Middle Age, like Egyptians, a long time before, knew how to attune the intelligence of the mind with the intelligence of the heart. Doing so, they gave a meaning to our life on earth which for them was a journey towards Light. Other civilizations using different means, but always taking into account the two cognitive parts of our mind, also reached a similar harmony and achieved masterpieces. Their search was so deep and so beautiful that even today we are touched by so much beauty and harmony, and we partake emotion and feel joy at visiting those places, since that is our inner nature, we are all longing for love and harmony, and longing for knowledge.

Therefore, it is reasonable to believe that today we should work in this direction in order to give hope and beauty for the generations to come.

References

- [1] J. Bolte Taylor. Video of the lecture “My stroke of insight”. Postatheisme.
- [2] V. A. Fomichov, O. S. Fomichova. Cognitonics as a New Science and Its Significance for Informatics and Information Society. *Special Issue on Developing Creativity and Broad Mental Outlook in the Information Society (Guest Editor Vladimir Fomichov), Informatica. An Intern. Journal of Computing and Informatics (Slovenia)*, 2006, Vol. 30, No. 4, pp. 387-398.
- [3] F. Schwarz. *Symbolique des cathedrals*. Ed. Du huitième jour.

AN EXTENDED CONCEPT OF MULTI-MEDIA AND ITS ROLE IN CREATIVITY IN BASIC EDUCATION

G. Tsayang, D.M.Totev

Faculty of Education, University of Botswana

P/Bag 00702, Gaborone, Botswana;

Tel 00267 355 2390; Fax 00267 3185096

tsayangt@mopipi.ub.bw ; totevdm@mopipi.ub.bw

ABSTRACT

Creativity is very pertinent to both individual and societal growth and therefore it must be enhanced in the education systems. The paper argues that in order to achieve universal basic and quality education we must consider using even the indigenous multi-media (culture based) to reach even the most modern technologically disadvantaged learners in order to enhance their creativity. Deliberate efforts must be taken by governments to infuse the culture based ways of enhancing creativity into their school syllabi and curriculum in general.

1 INTRODUCTION

Creativity plays a central role in education and especially in Primary education, because of its fundamental nature and great importance for life. Guilford [2] is remembered for his significant contribution to the idea that intelligence is not a unitary concept. He introduced the theory of Divergent Thinking and on its basis proved that standard intelligence tests do not reflect the divergent thinking of individuals and are better oriented to convergent thinkers. His Structure of Intellect Theory is considered to be a general theory of human intelligence with many applications in real life practices.

Torrance [10] with his measurement of the creative potential is another pillar of modern research on creativity and human intelligence. Torrance elevated the understanding of creativity at a new level, differing from the concepts of Wallach, Kogan, Guilford and other pioneers of the creativity research, e.g. the Threshold Hypothesis.

In this research Torrance introduced a method of quantifying creativity and invented the Torrance Tests of Creativity Thinking. Consequently, he proved that the IQ test is not the only one to measure human intelligence.

Osborn together with Parnes [7] developed the modern version of the Creative Problem Solving (CPS) technique, widely used in schools and corporate organizations. The CPS process is the conceptual framework of Future Problem Solvers International organization (FPSPI).

Taylor and Williams [9] worked on establishing a relationship between creativity and instruction. Taylor devoted most of his research to the development of Multiple Creative Talent Teaching Approach. In

accordance with this concept students would be able to learn a great deal about themselves. In such a way students are led to the optimization of their productivity and motivation. Taylor chaired 19 summer workshops for teachers with the main focus on creativity and talent development in class.

The commonly accepted definition of creativity has three parts [1]: 1. Creativity as a complex process, subject to studies by Behavioral Psychology, Social Psychology, Psychometrics, Cognitive Science, Artificial Intelligence, Philosophy, Economics, Business and Management. 2. Creativity as an interpretation of past knowledge and experience in a new way and 3. Enrichment of the existing knowledge base.

George Keller, <http://dusanstojakovic.com/dusan%201.swf> (2008), expressed this definition thus "... creativity consists largely of re-arranging what we know in order to find out what we do not know". Van Gundy [12] considers the creative process as a five-step process as follows: 1. Fact-finding 2. Problem-finding 3. Idea-finding 4. Solution-finding 5. Acceptance-finding.

Certain cognitive characteristics contribute to one's creative behavior (<http://tip.psychology.org/create.html>). These are noted as *Fluency, Flexibility, Visualisation, Imagination, Expressiveness and Openness*.

Such skills could be learned or they may be situational. Fluency is related to the number of meaningful responses to a challenge, while flexibility reflects the diversity of the responses provided. The following features are involved in the psychometric study of creativity too [3]:

Encouraging confidence and a willingness to take risks; Focusing on mastery and self-competition; Promoting supportable beliefs about creativity; Providing opportunities for choice and discovery; Developing self-management (meta-cognitive skills); Teaching techniques and strategies for facilitating creative performance and Providing balance.

It is argued therefore that conventional educational systems do not provide an environment conducive enough for the development of creative thinking. This is the main reason why special attention should be paid to pre-schools and primary schools in order to rectify such a situation.

The pupils, teachers and administration at that level need to be a focus for creativity efforts. Such an approach calls for a wider interpretation of the 'Multimedia' concept, an

essential tool to enhance creativity. The advantages of the conventional application of multimedia techniques are well proven by many researchers. Recently, a study on the implementation of a powerful tool for interactive training [11] could illustrate the importance of the conventional multimedia to creativity.

The developed Interactive Tool for Intercultural Education Software is a powerful tool for interactive training applying all modern trends of multimedia in education which in turn can trigger a lot of creativity in the young ones.

Some other authors consider creativity from a philosophical point of view and try to define the major social and political factors that contribute to its stimulation. Hiwaki [4] asserts that culture plays a significant role in enhancing creativity and emphasizes that Sound Creativity results from Sound Culture. Hiwaki's contention regarding creativity is that creativity can be explored to the maximum if it embraces all features of culture values: societal (own), societal (foreign), material, traditional, collective, spiritual, individual and progressive values. So the paper is based on the well balanced cultural environment as an appropriate tool for promoting creativity in Basic Education in Botswana.

2 CULTURE AS BASIS FOR SOUND CREATIVITY IN BASIC EDUCATION

While Hiwaki's view on creativity provides the heuristic and philosophical framework of the paper, constructivist theory forms the basis for application.

Constructivist theory of knowing implies humans construct knowledge and learn from their experience. Individuals have live experiences. Learners acquire skills better through doing things or through activities (<http://en.wikipedia.org/wiki/constructivism>).

An individual's view the world manifested through their culture will to a larger extent influence the way they learn. One's background (socio-economic, educational, cultural, traditional, etc.) plays an important role in the process of learning. Constructivists believe learning is a lifespan activity resulting from the learner's interaction with their environment which could be through the game they play, for example. "It is thus important to take into account the background and culture of the learner throughout the learning process, as this background also helps shape the knowledge and truth that the learner creates, discovers and attains in the learning process" (<http://en.wikipedia.org/wiki/constructivism>).

Young boys and girls play traditional games, they learn these from an informal point of view. There are a number of traditional games which are mind boggling or challenging the young ones to solve problems. One's observation of these has led to a belief that they reflect commitment and desire to excel and outperform. According to [8], the motivation to solve the traditional games puzzles is more powerful than situations where the motivation is brought into the learners through an external source or a teacher. The argument here according to social constructivists is that self-motivation is much more

powerful than being motivated by others. The desire to achieve in traditional games is more an intrinsic factor than an extrinsic one.

It is also noted that in the Botswana Educational system a large number of schools are found in the remote and not so developed villages and/or settlements. They are generally characterized by poor or lack of utilities such as electricity which is pertinent to the modern technological world. The development of modern technology is viewed as contributing a lot to creativity of learners, in particular, the computer world.

Despite the divergent cultures in the country, Botswana has as one of the aims of education, the creation of a creative and innovative nation. As indicated in Vision 2016, the education system will empower citizens to become the best producers of goods and services. The country's education system aims at producing citizens who have entrepreneurial aptitudes and who will therefore be able to create employment.

It is the contention of the authors that providing an enabling environment as early as at primary education level is the basic foundation for creativity. Adopting the Hiwaki theoretical framework of a balanced culture composed of the modern and traditional attributes, it is argued that basic education can be very meaningful and therefore an important preparatory basis for a sustained future of Botswana, economically, socially and politically. It is argued further that since all normal children have the potential which they are born with, it is important for such potential to be triggered to be live. It is the contention of the authors of the paper that sound creativity in learners is based on sound culture.

Research has found out that children who are exposed to multi-media perform better than those who are not [11, 13]. A comparison is made between multi-media in the conventional sense of using modern technology in the likes of a computer in particular and non use of the modern technology. This paper argues that, the conventional conception of multi-media is limiting because it excludes the traditional indigenous way in which people communicated then and now. The early indigenous people survived, produced things and lived on because they had their own technology. They were able to be creative because of their art, skill, and craft. They had their own techniques or methods of performing specialized tasks. They have their own ways of teaching their own young ones how to critically look at the world through games, riddles poems and songs. All these reflected their ability to be creative. They did not need a computer to be creative. In other words, ICT is not a pre-requisite to creativity but an enhancement of it. A computer is, of course, an important tool to enhance that natural creativity but not a replacement of culture as a source of creativity. Therefore, indigenously, there has always been traditional multi-media which the authors feel is not fully explored to benefit those who may not be directly exposed to modern multi-media. If this is not explored, we cannot think or dream of universal education for all.

The traditional or the cultural way of communicating in its multi-faceted ways is hardly regarded as part of multi-media let alone taken on board in modern learning environments. It is further argued that, creativity does not come with the advent of modern technology, it has always been there with the existence of humankind hence they were able to survive various conditions. It is further argued that learning and creativity are enhanced by experiential circumstances. People learn better what they live, which is culture and therefore, if the traditional lives which we are sure, offered creativity were to be taken on board in the modern learning environments, they would comfortably complement modern technology and enhance creativity. In particular, the authors feel, where modern technology is not available it would be wise to utilize traditional indigenous ways of knowing and ways of being creative as a result. This implies taking on board what Prof. Hiwaki refers to as using Sound Culture to enhance Sound Creativity [5].

If the words multi- and media are interrogated further, and Webster Dictionary meaning is any thing to go by then the word itself is not fully interpreted if its meaning is mainly confined to the use of modern technology. This further interrogation can be seen to mean the following:

Multi as in 'many' or 'more than one'.

Media as in a 'means of communication'.

According to [6, p. 835], Multi-Media means, "relating to or using two or more media, especially a combination apprehended by different senses as sight and hearing". Looking at the literal meaning of the word itself, it can be argued that indeed every culture or every society had its own way of communicating which is multi-faceted and comprehended by the different senses as sight and hearing, hence every society can be said to have its own multi-media even before the advent of modern multi-media.

The narrow conception of multi-media is indeed limiting with regards to the training of and understanding by the teachers who must stimulate children creativity. The concept of multi-media in the modern or conventional context, it is argued, appears to be perceived as alien by teachers as it is associated with modern technology.

It must be noted that indeed multi-media is not a new concept to Botswana and the Batswana and the indigenous world as a way of communication and encouraging creativity amongst the young ones. One thing that is disturbing is that the indigenous people have not and do not respect what the authors of this article refer to as indigenous multi-media, their culture which is a basis for creativity. The colonial mentality has eroded the pride they should have about their own culture which is a basis for Sound Creativity. These indigenous multimedia sources are indeed very rich, in abundance and accessible to everyone to encourage the young ones to be creative. It is the most relevant and would be better understood by most rural children who are not exposed to modern multi-media. The authors see the traditional games and other indicators of the Botswana culture as indigenous multi-media which can enhance children's creativity if utilized by the teachers

in the same way that they regard modern technology. For instance, there are a number of Botswana traditional games whose role was not only to provide entertainment but also to develop creativity and critical thinking in the young minds, e.g. Mmele, Morabaraba, Diketo, etc.

Mmele is a Setswana traditional game which encourages creativity, decision making, critical thinking and sharpens geometric skills in young learners.

The game is played by two people. The aim of the game is for either of the parties to win. The game is a triangle subdivided into 4 parts as shown in the diagram (Figure 1).

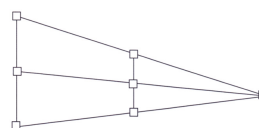


Figure 1: *The Diagram of Mmele Game.*

The players each use 3 pebbles called (cows/cattle) of different shapes or colours to distinguish between the two players. Players alternate in placing their cows one at a time such that at the end whoever wins has his cows all in one straight line without the opponents on the way. Therefore, each player must strive to ensure they block their opponent's cows from being in a straight line while they strive to have their own in that straight line. After all the cows have been placed on the triangle, they must be moved around. Each player makes one move at a time and so on until one of them is able to align theirs in one line. Then they win.

This exercise is a challenge to the mind because either player must strategize and be creative enough to defend their play, block the opponent and win their opponent. It sharpens the critical thinking skills, the decision making and therefore problems solving skills. The players have to be really creative in their strategies. This game represents old technology which not also sharpens the mind but also teaches mathematical concepts such as geometry.

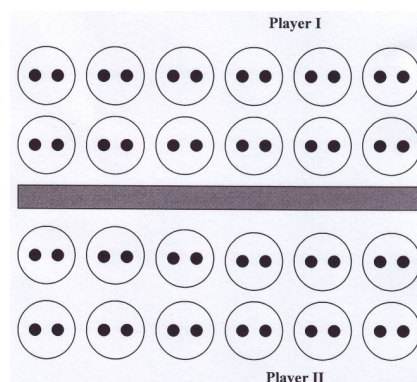


Figure 2: *The Diagram of Morabaraba Game.*

The game has traditionally been used by young head boys as they pass their times looking after either, goats, sheep or cattle. The technology is still used and what needs to be done is to relate and infuse it to modern pedagogy since most children no longer spent their time looking after

livestock and do not get to play it much. In terms of resources, it is the cheapest as it uses ordinary boxes and the pupils can even play it on the ground. Classrooms can be decorated by drawing/painting such triangles on the floors and then letting pupils play such games as part of classroom activities intended to sharpen learners' creativity and decision making as well as providing entertainment at an early age.

Morabaraba (Figure 2) is yet another Setswana traditional game which tests the player's creativity, decision making and critical thinking.

The game is played by two people (Player I and Player II). Each player has twelve holes having two pebbles into each one. The rules of the game are as follows:

- 1 player at a time.
- Player picks all pebbles from any hole of his/her choice.
- Moving in an anticlockwise direction he/she drops 1 pebble in each hole at a time.
- If pebbles picked end in a hole that already has some, player is required to pick those pebbles from that particular hole and if that hole is directly opposite to the opponent's hole, the player will first pick the pebbles from the two opposite holes and continue with the movement, but if the pebbles end in an empty hole, the opponent is now given a chance.
- Whoever gathers all the opponent's pebbles first is declared as a winner. The idea is for each player to strategize their movement such that they collect as much from the opponent as possible. This requires a lot of creativity, critical thinking and it trains young minds into being able to solve tricky problems. Such training was used before the introduction of formal schooling and even now, as some informal mathematics, decision making creativity skills development, hence preparing youth for their adulthood decision making roles.

It is argued that indeed those not exposed to multi-media are more likely not to perform like those exposed. The point of contention is that the concept of multi-media should be enhanced and treated in a broader way to accommodate the indigenous cultural values and create a conducive and balanced environment for sound creativity. Where there are no computers and modern technology, there is the other component of the enhanced multi-media which is sound culture which if explored can enhance creativity. It is emphasized therefore that if vision 2016 and achievement of universal basic education for all is to be achieved, and access to educational resources is to be achieved, there is need to explore these other forms of multi-media, the indigenous traditional multi-media and utilize it to enhance creativity by those children in areas far from the modern technology.

Recommendations

- Teachers must be instructed to appreciate that which is traditional and therefore context based and indeed readily available to everyone as a learning resource base. This could be accomplished through including topics on such in the official school curriculum.

- UNESCO and other Educational Organisations could be asked for assistance to help infuse traditional cultural attributes into school curriculum to enhance creativity.
- The concept of multi-media should be explored further to include any other form of knowledge or experience that could be beneficial to the creativity of pupils through research.
- The education systems must make a deliberate effort to infuse indigenous multi-media in teacher training curriculums.

References

- [1] Chakalisa, P., Mapolelo, Totev, D. M, Totev, E. D. Modern Methods for Stimulating Creativity in Education. *Special Issue on Developing Creativity and Broad Mental Outlook in the Information Society (Guest Editor: Vladimir Fomichov). Informatica (Slovenia), 2006, Vol. 30, No. 4, p. 421-425.*
- [2] Guilford, J. P. *The Nature of Human Intelligence.* 1967.
- [3] Guilford, J. P. Creativity. *American Psychologist*, 1950, 5, 444-454.
- [4] Hiwaki, K. Developing Creativity: A Sound Culture for Sound Creativity. *A paper presented at the Intern. Conf. ISSEI 2006, University of Malta.* 2006.
- [5] Hiwaki, K., Tong, J. Credibility Trap: Japan Today and China Tomorrow. *Human Systems Management*, 2006, Vol. 25, No 1, pp.31-50
- [6] Markwardt A. H., Cassidy, F. G, McMillan J. G. *Webster Comprehensive Dictionary Encyclopedia Edition, Volume One.* Chicago, Ferguson Publ. Company, 2000.
- [7] Parnes, S. J. *Creative Behavior Guidebook.* New York: Scribners. 1967.
- [8] Prawat, R. S., Floden, R. E. Philosophical Perspectives on Constructivist Views of Learning. *The Educational Psychologist*, 1994, Vol. 29, No, pp. 37-48.
- [9] Taylor, C., Williams, F. *Instructional Media and Creativity.* New York: Wiley. 1966.
- [10] Torrance, E. *The Search for Satori and Creativity.* Buffalo, NY: Bearly Ltd. 1979.
- [11] Turrini, G., Baroni, P., Paccosi, A. AddizionarioPLUS: a Creative Approach to Linguistic and Intercultural Education. *Special Issue on Developing Creativity and Broad Mental Outlook in the Information Society (Guest Editor: Vladimir Fomichov). Informatica (Slovenia), 2006, Vol. 30, No. 4, p. 407-412.*
- [12] Van Gundy, A.B. *Creative Problem Solving.* New York: Quorum. 1987.
- [13] Zanetti, M. A., Turrini, G., Miazza, D. "ADDIZIONARIO": a New Tool for Learning between Metacognition and Creativity. *Special Issue on Developing Creativity and Broad Mental Outlook in the Information Society (Guest Editor: Vladimir Fomichov). Informatica (Slovenia), 2006, Vol. 30, No. 4, p. 399-405.*

SCHOOL READINESS: AN ITALIAN TOOL WITH A MULTIFACTORIAL APPROACH FOR ACADEMIC SUCCESS

D. Miazza, M.A. Zanetti

Department of Psychology, Botta Square

University of Pavia

27100 Pavia, Italy

Tel.: (+39) 0382986273

e-mail: daniela.miazza@unipv.it

ABSTRACT

Many theoretical and experimental approaches have underlined the role of the development of cognitive and self-regulation abilities, motivation, attention, inhibition control in the concept of *academic readiness*. The *academic readiness* concept has amplified and developed the previous one known as *school-readiness*.

The presented work includes two batteries of tests (one for 4 year-old children, one for 5 year-old children) with their standardization procedures conceived to verify the presence of the necessary skills which make children able to face without difficulty the passage from kindergarten to primary school.

The elaborated batteries investigated the following areas: linguistic, phonological, logical-mathematical and numerical ability, the development of motor ability, the use of symbols and children's social relationships.

1 INTRODUCTION

School readiness is the result of a complex and articulated perspective, a multifactor construct in which the important cognition-emotion relationships, as well as levels of family, peer, school and community influence converge.

As for what concerns the concept of academic readiness, recently many theoretical and experimental approaches have underlined the role of the development of cognitive, self-regulation abilities but also of variables related to motivation towards learning, together with attentional and inhibition control processes [1-3, 6, 10]. The *academic readiness* concept has in fact replaced and partly expanded the better-known one of *school-readiness*.

Recently, many authors also have focused on the role of logical-mathematical and numerical abilities which seem to be the best predictors of cognitive- academic skills, followed by linguistic development and attention abilities [6].

Within this complex construct, we concentrated our attention mainly on the cognitive aspects of learning.

Borkowski's et al. work [4, 5] has highly influenced studies exploring the relationship between one's learning and deficits in strategies and in attributions; Dweck's work [7-9] on implicit intelligence is also strongly in tune with the school-readiness approach we present below. In fact, all these contributions, which considerably influence international literature, as well as applicative research, study small children's development of beliefs which serve as a basis for organizing their world and giving meaning to their experiences. These beliefs organize into "meaning systems" which are able to create different psychological worlds, thus inducing those who have built them to think, feel and act differently in similar situations. Until a few years ago, we thought small children were far more protected from negative effects of failure, for example, compared to older ones.

This may have represented an important protective factor for children dealing with difficult learning tasks they face from the beginning of elementary school, like learning to read or write; it wouldn't have caused negative self-evaluations or impotence reactions in case of failure, given that, not knowing what intelligence is, a failure could not have caused them to doubt about their intelligence and thus experience vulnerability or impotence reactions. They are rushed into a world of roles and expectations, learning what being a good or a bad boy means. This is how failure or criticism becomes emblematic for their being bad or not worthy, and, as many older children do with intelligence, they believe being bad is a stable trait. We believe that a different attention to the concept of school readiness could be an important protective factor.

Recent research on learning, within Blair's theoretical perspective [1-3], refers in particular to the growing interest in the necessary relationships that unfold between thought ability, motivation and self-regulation, confirming that an elevated level of motivation and self-regulation, mediated by parents, teachers or peers, is clearly linked to success and engagement in school.

One last aspect seems noteworthy: in life as in school there are boring things we simply must do. When a task is tiring or uninteresting to learn, it is necessary to learn to apply oneself. This is where self-regulation - the series of organizational and self-monitoring abilities that are particularly useful when tasks are long, complex or unpleasant – comes into the picture. Children can be taught these abilities, and those who acquire them can work more efficiently.

If learning, because of implicit intelligence theory, is dedicated to performance versus competence, the process may be seriously compromised in its motivational components when exposed to frustration, impairing all subsequent learning.

This aspect seems particularly relevant, in that beyond the abilities required by school readiness, an adequate level of development and control of emotional and motivational functions is just as necessary in order to support the child in this new learning adventure, which should not be solely understood in terms of cognitive performance, but also as active construction of the self.

2 GOALS AND CONSTRUCTION CRITERIA

The aim of our tests is the assessment of kindergarten children's aptitude to enter primary school, offering an objective instrument for the evaluation of some skills that, even if they must not be considered, strictly speaking, school "prerequisites", are nevertheless necessary to attend primary school without difficulty.

Our instruments – administered to 4 and 5 year-old children – were not conceived as achievement tests, but they aimed at investigating the ways of developing of basic skills, in order to act preventively when the child's mind has great plasticity and is utterly open to environmental influence.

Besides, competences that consolidate during kindergarten years represent the essential prerequisites to face primary school learning tasks; the early assessment of impairments in a child makes it possible for his-her primary school teachers to structure from the very beginning personalized academic curricula aimed at empowering the better-developed areas and making them able to act vicariously, countervailing the deficits of the weaker areas, thus letting the child enter a process of co-construction of his-her possible learning pathways.

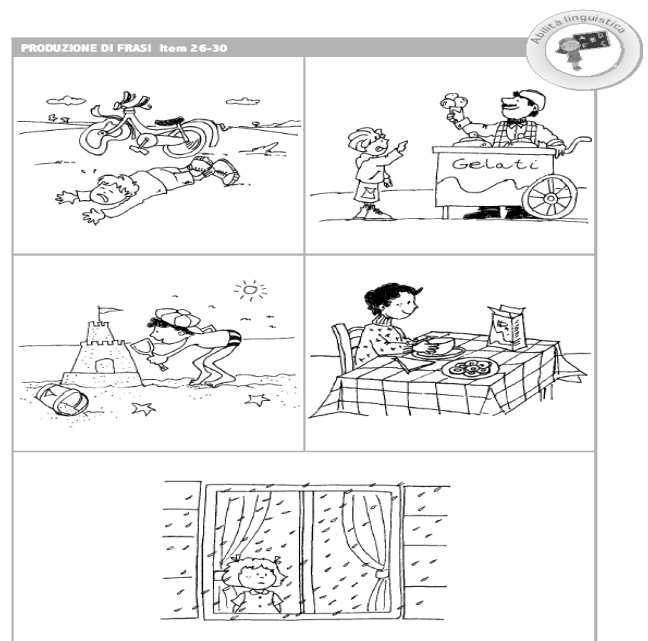
The present work includes two batteries of tests (one for 4 year-old children, one for 5 year-old children) with their standardization procedures conceived to verify the presence of the necessary skills which make children able to face without difficulty the passage from kindergarten to primary school.

Construction criteria were similar for the two versions.



© 2002, M.A. Zanetti e D. Miazza, SR 4-5, Trento, Erickson

Figure 1: Naming example.



© 2002, M.A. Zanetti e D. Miazza, SR 4-5, Trento, Erickson

Figure 2: Comprehension of morphosyntactic structures examples.

The **linguistic subtest** investigates the main aspects of comprehension and linguistic competence following the acquisition phases. The early assessment of a delayed development, at both levels of comprehension and production of linguistic structures, can make it possible for teachers to implement focused interventions for the improvement of the mentioned structures. For the construction of this subtest we considered the ways of evolving of linguistic ability and of oral and pictorial task comprehension.

The **phonological subtest** was conceived for the assessment of phonemic identification and discrimination in similarly-sounding stimuli; we identified phonologically similar words differing only for the initial consonant of for an initial or final syllable, in order to assess the level of metaphonological awareness, an important prerequisite for the acquisition of reading skills. We also included in the subtest some statements containing articulatory difficulties for the assessment of dyslalia or pronunciation impairments concerning single consonants or group of consonants. Children are also asked to recall a series of growing-difficult statements in order to evaluate the accuracy of their phonological reproduction.

The **symbolization subtest** (only for 5 year-old children) explores the child's familiarity and expertise with alphabetical and numerical symbols and, compared to the other subtests, is more influenced by family context, socio-cultural level and educational choices made by adults; still it is possible to consider the familiarity with symbols a good predictor of academic achievement. This subtest gives us clues on the child's familiarity with graphic signs (both letters and digits) and their phonological decoding.

The **logical-mathematical and numerical subtest** assesses skills that represent the prerequisite for the achievement of logical-mathematical and numerical knowledge in subsequent years; we chose items referring to basic clusters like quantification: considering that the child's familiarity with such fundamental operations underlies the acquisition of logical-mathematical and numerical competences, it is important to assess this prerequisite in order to promote an adequate development.

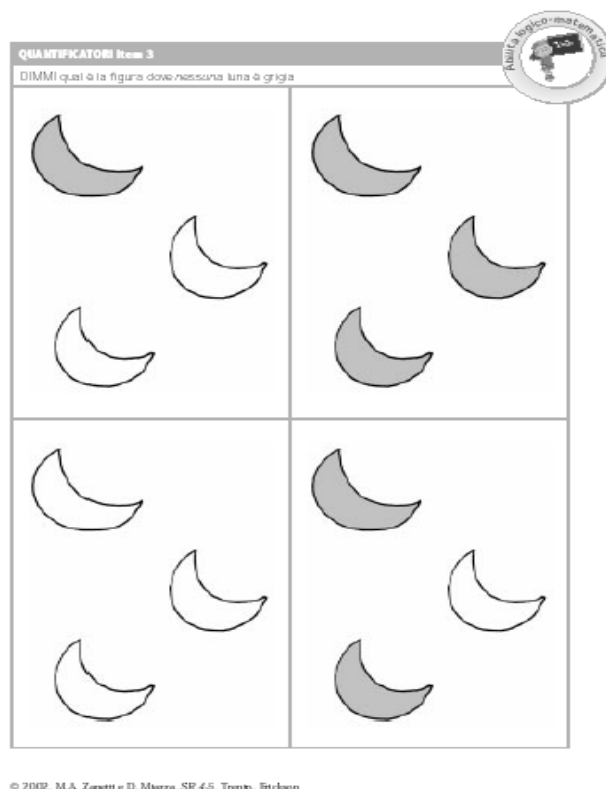


Figure 3: *Quantificators example.*

The **motor development subtest** assesses general motricity, the acquisition of body scheme, fine motor skills, visuo-spatial and oculo-manual co-ordination.

For the evaluation of **social development** (only for 5 year-old children) we examined the child's social behaviour using an observational scale which gave us clues about the child's ability to form relationships with both adults and peers and about the presence of aggressive behaviour, thus obtaining an appraisal of the child's positive and negative socialization that makes it possible to investigate more deeply socially unacceptable conducts.

The areas investigated by means of our batteries concern, in detail:

- linguistic ability (naming, production of words, comprehension of morphosyntactic structures, production of statements, text comprehension, text production);
- phonological ability (discrimination of phonemes and syllables, double consonants, phonological similarities, reproduction of articulatory difficulties);
- logical-mathematical and numerical ability (quantificators, biunivocal correspondence,

classification, serialization, spatial localization, counting skills, geometry);

- development of motor ability (general and oculo-manual co-ordination, praxia, lateralization, body scheme);
- use of symbols (reading and recognition of symbols and graphemes, word writing, recognition, reading and writing of numbers);
- social relationships (autonomy, social closeness to adults and other children, leadership behaviours, aggressive behaviours).

Our batteries were standardized on a sample of 300 children attending their last year of kindergarten in Lombardy.

All the tests were administered by teachers during curricular activities; their global evaluation is the final result of an articulated program that, beginning when children are 4 years old, allows to follow every child until his-her entering primary school with a portfolio of abilities that could help the teachers to implement better-focused didactic activities, thus making learning more suitable to the needs of each pupil and maximizing the children's potential during daily teaching work, for better academic outcomes.

TABELLA 2.2
Tabelle di conversione punteggi grezzi/centili - SR 4-5 (5 anni)

Centili	Abilità linguistica	Abilità fonologica	Abilità logico-matem.	Simbolizzazione	Abilità psicomot.	Abilità sociale	Centili
99	50	20	27	30	20	115	99
95	50	20	27	30	20	108	95
90	50	20	27	30	20	103	90
85	50	20	26	30	20	100	85
80	49	20	26	30	20	97	80
75	49	20	26	30	20	94	75
70	48	20	26	30	19	92	70
65	48	20	25	30	19	91	65
60	48	20	25	30	19	89	60
55	47	20	25	30	19	88	55
50	47	20	25	29	19	88	50
45	47	20	25	29	18	86	45
40	46	19	24	28,8	18	85	40
35	45,2	19	23	28	18	83	35
30	45	19	23	27	17	81	30
25	44	19	23	27	17	78	25
20	43	19	22	26	15,4	76	20
15	41,8	18	21	24	14,8	73	15
10	40	16,2	16,8	20	13	66	10
5	34,6	15	13,2	14	10,6	57	5

Figure 4: Score table.

References

- [1] Blair C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of child functioning at school entry. *American Psychologist*, No. 57, p. 111-127.
- [2] Blair C., Vernon-Feagans L. (2006). Measurement of School Readiness. *Early Education and Development*, Vol. 17, No. 1, p. 1-5.

- [3] Blair, C., Knipe, K., Cummings, C., Baker, D., Eslinger, P., Gamson, D., & Thorne, S. (2007). A developmental neuroscience approach to the study of school readiness. In R. Pianta, M. Cox, & K. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 149-174). Baltimore MD: Brookes.
- [4] Borkowski J.G., Estrada T.M., Millstead M. e Hale C.A. (1989). General problem-solving skills: relations between metacognition and strategic processes. *Learning Disabilities Quarterly*, No. 12, p. 57-70.
- [5] Borkowski J.G., Carr M., Rellinger L. e Pressley M. (1990). Self-regulated cognition: interdependence of metacognition, attributions and self-estimate. In Jones B. e Idol L. (a cura di), *Dimensions of thinking and cognitive instruction*, Hillsdale, Erlbaum, Vol. 1, p. 53-92.
- [6] Duncan, G.J, Dowsett, C.J, Claessens, A., Magnuson, K., Huston, A.C., Klebanov, P., Pagani, L.S., Feinstein L., Engel M., Brooks-Gunn J., Sexton H., Duckworth K., Japel, C. (2007). School readiness and later achievement. Peer Reviewed Journal. *Developmental Psychology*. Vol. 43, No. 6, 2007, p.1428-1446.
- [7] Dweck C.S. (1975). The role of expectations and attributions in the alleviation of learned helplessness, *Journal of Personality and Social Psychology*, No.31, p.674-685.
- [8] Dweck, C.S. (1986). Motivational processes affecting learning. *American Psychologist*, No. 41, p. 1040-1048.
- [9] Dweck C.S. (2000). *Teorie del sé. Intelligenza, motivazione, personalità e sviluppo*. Erickson, Trento
- [10] McClelland M., Cameron C. E., McDonald Connor C., Farris C. L., Jewkes A. M., Morrison F. J. (2007). Links Between Behavioral Regulation and Preschoolers' Literacy, Vocabulary, and Math Skills. *Developmental Psychology*, Vol. 43, No. 4, 947-959.

EXPANDING MENTAL OUTLOOK BY USING CONCEPT MAPS

Dumitru Dan Burdescu¹, Marian Cristian Mihăescu¹, Bogdan Logofatu², Costem Marian Ionașcu³

¹Department of Software Engineering, University of Craiova

Bvd. Decebal, Nr. 107, 200440, Craiova, Dolj, Romania

Tel/fax: +40 251 438198;

e-mail: {burdescu, mihaescu}@software.ucv.ro

²CREDIS Department, University of Bucharest

Bulevardul Mihail Kogalniceanu 36-46, sector 5, Buharest, Romania

Tel: +4 021 315 80 95; Fax: +4 021 315 80 96;

e-mail: logofatu@credis.ro

³Analysis, Statistics and Mathematics Department, University of Craiova

Str. A.I. Cuza 13, 200585, Craiova

Tel/Fax: +4 0251 411 317

e-mail: icostelm@yahoo.com

ABSTRACT

Expanding mental outlook of learners is one of the important open problems in e-Learning. Online should have the property of expanding and enhancing the mental outlook of learners in general and also, in particular, concerning the studied discipline. This paper presents an approach to this issue. The tool used for expanding the mental outlook is represented by concept maps. Concept maps are used for representing relationships among concepts that define a certain area (e.g. discipline in e-Learning). The concept map is built in such a way that it represents a concrete and broad representation of the domain. As an example in this paper, a concept map is presented built for *Data Structures and Algorithms* course and more exactly for *Binary Search Trees* chapter.

1 INTRODUCTION

Concept maps are a result of Novak and Gowin's [3] research into human learning and knowledge construction. Novak [1] proposed that the primary elements of knowledge are concepts, and relationships between concepts are propositions. Novak [2] defined concepts as "perceived regularities in events or objects, or records of events or objects, designated by a label". Propositions consist of two or more concept labels connected by a linking relationship that forms a semantic unit. Concept maps are a graphical two-dimensional display of concepts (usually represented within boxes or circles), connected by directed arcs encoding brief relationships (linking phrases) between pairs of concepts forming propositions.

This paper uses concept maps for presenting the very higher general structure of a studied discipline. The concept map is used by learners as a top level reference material that may be consulted. This structured high level overview of the discipline is aimed to expand the mental outlook of the learners in general by exercising this ability on a particular discipline.

2 CONCEPT MAPS

Concept mapping may be used as a tool for understanding, collaborating, validating, and integrating curriculum content that is designed to develop specific competencies. Concept mapping, a tool originally developed to facilitate student learning by organizing key and supporting concepts into visual frameworks, can also facilitate communication among faculty and administrators about curricular structures, complex cognitive frameworks, and competency-based learning outcomes. To validate the relationships among the competencies articulated by specialized accrediting agencies, certification boards, and professional associations, faculty may find the concept mapping tool beneficial in illustrating relationships among, approaches to, and compliance with competencies [4].

Concept maps are also effective in identifying both valid and invalid ideas held by students, and this will be discussed further in another section. They can be as effective as more time-consuming clinical interviews for identifying the relevant knowledge a learner possesses before or after instruction [7].

Recent decades have seen an increasing awareness that the adoption of refined procedures of evaluation contributes to

the enhancement of the teaching/learning process. In the past, the teacher's evaluation of the pupil was expressed in the form of a final mark given on the basis of a scale of values determined both by the culture of the institution and by the subjective opinion of the examiner. This practice was rationalized by the idea that the principal function of school was selection - i.e. only the most fully equipped (outstanding) pupils were worthy of continuing their studies and going on to occupy the most important positions in society.

Ausubel [1] made the very important distinction between rote learning and meaningful learning. Meaningful learning requires three conditions: 1. The material to be learned must be conceptually clear and presented with language and examples relatable to the learner's prior knowledge. Concept maps can be helpful to meet this condition, both by identifying large general concepts held by the learner prior to instruction of more specific concepts, and by assisting in the sequencing of learning tasks through progressively more explicit knowledge that can be anchored into developing conceptual frameworks; 2. The learner must possess relevant prior knowledge. This condition can be met after age 3 for virtually any domain of subject matter, but it is necessary to be careful and explicit in building concept frameworks if one hopes to present detailed specific knowledge in any field in subsequent lessons. We see, therefore, that conditions (1) and (2) are interrelated and both are important; 3. The learner must choose to learn meaningfully. The one condition over which the teacher or mentor has only indirect control is the motivation of students to choose to learn by attempting to incorporate new meanings into their prior knowledge, rather than simply memorizing concept definitions or propositional statements or computational procedures. The indirect control over this choice is primarily in instructional strategies used and the evaluation strategies used. Instructional strategies that emphasize relating new knowledge to the learner's existing knowledge foster meaningful learning. Evaluation strategies that encourage learners to relate ideas they possess with new ideas also encourage meaningful learning. Typical objective tests seldom require more than rote learning [9].

According to this approach, the responsibility for failure at school was to be attributed exclusively to the innate (and, therefore, unalterable) intellectual capacities of the pupil. The learning/ teaching process was, then, looked upon in a simplistic, linear way: the teacher transmits (and is the repository of) knowledge, while the learner is required to comply with the teacher and store the ideas being imparted [5]. Usage of concept maps may be very useful for students when starting to learn about a subject. The concept map may bring valuable general overlook of the subject for the whole period of study. It may be advisable that at the very first meeting of students with the subject to include a concept map of the subject.

3 DATA STRUCTURES AND ALGORITHMS / BINARY SEARCH TREE CONCEPT MAP

The experimental concept map was used on Tesys e-Learning platform [6]. On this platform there was set an Algorithms and Data Structures discipline. The tests were performed for five chapters: Simply/Double Linked Lists, Binary Search Trees, Height Balanced Trees, B Trees and Graphs.

The concept map for Binary Search Trees is presented in figure 1. It contains 16 concepts, 11 linking phrases and 16 propositions.

The concepts are presented in table 1.

ID	Concept	ID	Concept
C1	BST	C9	Right child
C2	Dynamic Structure	C10	No child
C3	Node(s)	C11	Root
C4	Traversed	C12	Leaf
C5	Key	C13	Preorder
C6	Parent	C14	Inorder
C7	Child	C15	Postorder
C8	Left child	C16	Ascending order

Table 1: List of concepts.

The list of propositions with two concepts and one linking phrase is presented in table 2. The list of propositions with three concepts and two linking phrases is presented in table 3.

Once the concept map has been built, the general graph of the each chapter may be created. In this graph, each proposition will become an edge that links the first concept and the last concept. The domain knowledge expert will assign a weight for each edge. While the students answer the questions, the number of correct and wrong answers will determine the knowledge weight of that edge.

There is one proposition with five concepts and four linking phrases: "BST" may be "Traversed" in "Preorder" determines "Key" in "Ascending Order". The concepts are bolded and put between quotation marks, while linking phrases are italic and underlined.

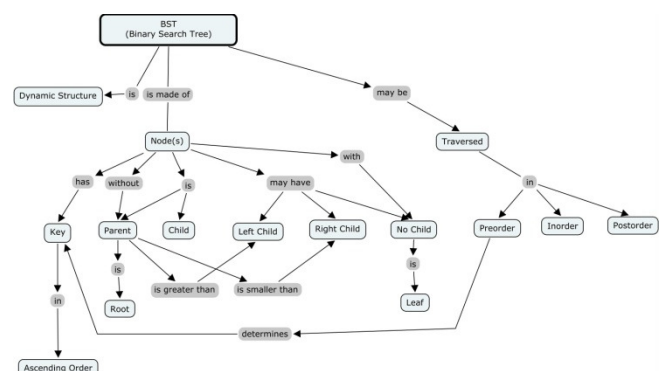


Figure 1: Binary Search Tree Concept Map.

Id	Concept	Linking phrase	Concept
P1	BST	is	Dynamic Structure
P2	BST	is made of	Node(s)
P3	Node	has	key
P4	Node	is	Parent
P5	Node	is	Child
P6	Parent	is greater than	Left child
P7	Parent	is smaller than	Right child
P8	Node	may have	Left child
P9	Node	may have	Right child
P10	Node	may have	No child

Table 2: List of propositions with two concepts and one linking phrase.

Id	C	LP	C	LP	C
P11	Node	without	parent	is	root
P12	BST	may be	traversed	in	preorder
P13	BST	may be	traversed	in	inorder
P14	BST	may be	traversed	in	postorder

Table 3: List of propositions with three concepts and two linking phrases.

Once the Concept Map has been set up, the professor has to set up the quiz questions for the chapter. A certain number of quiz questions will correspond to each edge in the map. There is no specification regarding the number of quiz questions, but a minimum (e.g. five) number is still required. The activity performed by a learner is monitored and stored. The Calculus engine will reconstruct an Annotated Concept Map which will present to the learner the current status of his/her knowledge level at Concept level. In this way, the learner will have an exact outlook of his/her knowledge level regarding that chapter. This information may be used by the learners and course managers to obtain important information regarding the reached knowledge level.

Defining a metric for computing, the accumulated knowledge is accomplished. If:

- W- is the weight of the edge
- CA – is the number of correct answers
- WA – is the number of wrong answers
- N – the number of questions

Then:

KW – is the knowledge weight of the edge

$$KW = \frac{CA - WA}{N} \frac{1}{W} * 100$$

Under these circumstances, the knowledge weight may also be negative. At any time, the overall knowledge level of the learner may be estimated as the ratio between overall knowledge weight and overall weight.

Proposition	Weight	Number of questions
P1	10	8
P2	4	7
P3	7	6
P4	3	5
P5	2	7

Table 4: Sample setup of BST chapter.

Proposition (weight)	No. of questions	CA	WA	KW (%)
P1 (10)	8	3	2	1.25
P2 (4)	7	4	2	7.14
P3 (7)	6	1	3	-4.76
P4 (3)	5	3	1	13.3

Table 5: Sample values for learner's associated graph.

Chapter	Weight
Simply/Double Linked Lists	15
Binary Search Trees	15
Height Balanced Trees	25
B Trees	25
Graphs	20

Table 6: Sample weights assigned to chapters.

Table 4 presents a sample of the setup of the Binary Search Trees chapter. Table 5 presents a sample of the values of the Learner's Associated Graph corresponding to BST chapter.

The values from table five are marked in an Annotated Concept Map that is finally presented to the learner. The Annotated Concept Map is the final outcome of the Decision Support System and is supposed to guide the learner regarding the necessary future efforts.

Table 6 presents the weights of chapters as they were assigned by the domain expert.

6 CONCLUSIONS AND FUTURE WORKS

Tesys e-Learning platform has been designed such that on-line testing activities may be performed as they were set up by course managers.

A Concept Map has been created for a Binary Search Trees chapter as well as for each chapter of Algorithms and Data Structures course. The Concept maps have been the starting point in creating the sets of quiz questions. Each quiz question refers to a certain proposition from the concept map.

After the setup has been put in place, the learners started using the platform. At request, from the general concept map there was derived the learner's associated concept map. This concept map provides important information regarding the level of knowledge reached by learner.

The business logic computes the knowledge of the student regarding the chapter as a knowledge weight and regarding the discipline as a percentage of covered concepts. This weight is computed as a function of proposition's weight, number of questions assigned to that proposition, the

number of correct answered questions and number of wrong answered questions.

This whole mechanism represents the functionality of a decision support system that runs along the Tesys e-Learning platform.

Whenever needed, the learner may study his/her own associated concept map at discipline level or at chapter level. This functionality has as main benefit expanding the mental outlook in general and for studied discipline in particular.

As for future works, a computational framework may be created which estimates the proficiency of learners after using the concept map as structuring facility. This kind of analysis may give important information that regards the improvement of used concept map.

References

- [1] Novak, J. D. (1977). *A Theory of Education*. Ithaca, NY: Cornell University Press.
- [2] Novak, J. D. (1998). *Learning, Creating, and Using Knowledge: Concept Maps as Facilitative Tools in Schools and Corporations*. Mahwah, NJ: Lawrence Erlbaum Associates.
- [3] Novak, J. D., and Gowin, D. B. (1984). *Learning How to Learn*. New York: Cambridge University Press.
- [4] McDaniel, E., Roth, B., and Miller, M. Concept Mapping as a Tool for Curriculum Design. In *Issues in Informing Science and Information Technology*.
- [5] Vecchia, L., Pedroni, M. (2007). Concept Maps as a Learning Assessment Tool. In *Issues in Informing Science and Information Technology*, Vol. 4.
- [6] Burdescu, D.D., Mihăescu, M.C. (2006). Tesys: e-Learning Application Built on a Web Platform. In *Proceedings of International Joint Conference on e-Business and Telecommunications*, pp. 315-318, Setubal, Portugal. INSTICC Press.
- [7] Edwards, J., and Fraser, K. (1983). Concept maps as reflections of conceptual understanding. *Research in Science Education*, 13, 19-26.
- [8] Ausubel, D. P., Novak, J. D., and Hanesian, H. (1978). *Educational Psychology: A Cognitive View (2nd ed.)*. New York: Holt, Rinehart and Winston.
- [9] Holden, C. (1992). Study flunks science and math tests. *Science Education*, 26, 541.

ACADEMIC-SCHOOL READINESS: AN ITALIAN TRAINING

M.A. Zanetti, D. Miazza

Department of Psychology, Botta Square

University of Pavia

27100 Pavia, Italy

Tel.: (+39) 0382986275

e-mail: zanetti@unipv.it

ABSTRACT

The paper describes a study aimed at promoting and maximizing children's academic readiness and achievement. In order to support an adequate development of scholastic performances, a specific training to be carried out in classrooms, following the battery School- Readiness 4-5, was created. 81 children at the age of 4 and 139 children at the age of 5 who had scored under the average in the battery S.R. were selected and divided into two experimental groups (43 and 64 children), to which an ad-hoc training was proposed, and into two control groups (38 and 75 children). The analysis of the data seems to confirm the efficacy of the training in promoting academic-school readiness.

1 INTRODUCTION

Before discussing academic-school readiness and achievement, it is necessary to find an adequate and sharable definition, which is not a marginal matter given that the most significant factors and variables remarkably change if we take into consideration different perspectives: in literature a wide range of definitions, some very dissimilar from the other ones, can be found. From Gesell's [6] maturational perspective, readiness is the result of an organic inner process that allows to achieve a given maturational level thanks to the activation of variables of an endogenous kind. Empirists, on the other hand, assume that all these processes are externally driven and, focusing their attention on the tangible characteristics of behaviour, prefer to concentrate on exogenous variables.

Socio-constructivist perspective [9] conceives academic-school readiness as the result of a social co-construction in which the context is not merely "the background", a container in which the development of skills takes place, but the common and constant referring that makes it possible. The last classic perspective, the interactionist one [11, 12], postulates school-readiness as the product of the interaction

experience-genetic patrimony – context, in other words the interaction between individuals and their contexts.

It's a widely shared opinion nowadays that when examining a child's school readiness both the complexity of the concept and its multidimensionality must be considered in their relationships to the context that contributes to co-construct its meaning. Recently some authors have tried to investigate more specifically some of the most significant involved variables: if Blair [1] focused on the relationship between emotional development and social skills, La Paro, Pianta [8] and Fantuzzo, McDermot and McWayne [7] chose to concentrate on cognitive-academic skills. The theme of the complexity and multidimensionality of the construct was explored by McClelland et al. [10], Duncan, Japel et al. [4] and once again Blair et al. [3], who investigated the role of self-regulatory abilities, motivation, attention, attitudes towards learning, inhibition control and executive functions in academic-school readiness and achievement. In a sort of ranking of the main predictors of academic achievement, pre-mathematics appears to be the most powerful, followed by linguistic development and the ability to sustain attention [4]. In a constructivist perspective, the growth process must be regarded more and more in developmental terms: at present it is defined as an open system, if one takes into account that development is not understandable through a stiff causalism and determinism paradigm, but as the outcome of constant self-regulations, capable to re-orientate its trajectories. Presently, the attention focused on the context aims in particular at concentrating on the conditions that can determine the quality of an intervention, mostly when its effectiveness is related to the possibility to exert some influence on processes that are still to be entirely defined and/or completed. The role played by the environment modifies the schemes of innate responses, which in turn, transforming, produce changes that have repercussions on the environment.

These considerations disclose all their importance whenever we study basic skills with a preventive intent, for the implementation of early intervention programs designed to support their adequate development. Taking into account individual variables is not enough; it appears more suitable

to consider those conditions that could significantly widen the distances that separate children before the intervention begins, with potential consequences on their future achievement too. We know [5] that experiences of lack of mastery or generalized difficulties can mark an individual at more levels: school (academic failure), emotions (excess of anxiety), social relationships (aggression, withdrawal).

2 GOALS AND CONSTRUCTION CRITERIA

In order to sustain an adequate development of scholastic performances a specific training to be carried out in classrooms was created, in respect of the study that motivated the construction of the battery School-Readiness 4-5 [13]. The goal of our training is to improve or regain specific prerequisite abilities; these acquired competences must not be the result of curricular activities nor of a time effect, but outcomes of the specific intervention we implemented. We expected the effectiveness of the training to be stronger in the logical-mathematical area, as literature indicates, also as a consequence of the most involved variables.

2.1. Measures and methodological characteristics

At the end of the previous school year we administered the batteries S.R. 4-5 to all the children aged 4 and 5 in the sample; on the basis of the results we built a general mapping of each child's needs referred both to basic skills and specific indicated subscales.

As areas to be improved, we chose the linguistic and logical-mathematical ones, given that literature indicated them as the fundamental prerequisites, especially logical-mathematical skills [4]; our indications as for what concerns the implementation of the training referred to: organizational aspects, the concept of learning in a constructivist perspective, the role of teachers, the zone of proximal development, scaffolding, monitoring, metacognition and motivation.

2.3. Subjects

81 children aged 4 who had scored under the average in the battery SR were selected and subsequently divided into an experimental group (43), to which we offered an ad-hoc training, and into a control group (38 children).

At the same time we found 139 children attending their last year of kindergarten with low scores in the last tests of the battery SR: also this sample was divided into an experimental group (64), sustained during the period of the first year in the primary school and into a control group (75 children).

2.4. Training

Once the "target" children individuated, the training focused its attention on the non total competence of the progressed competences, using specific charts graduated in difficulties and in competences required, proposed individually or to small groups. Otherwise the control group continued the normal activity.

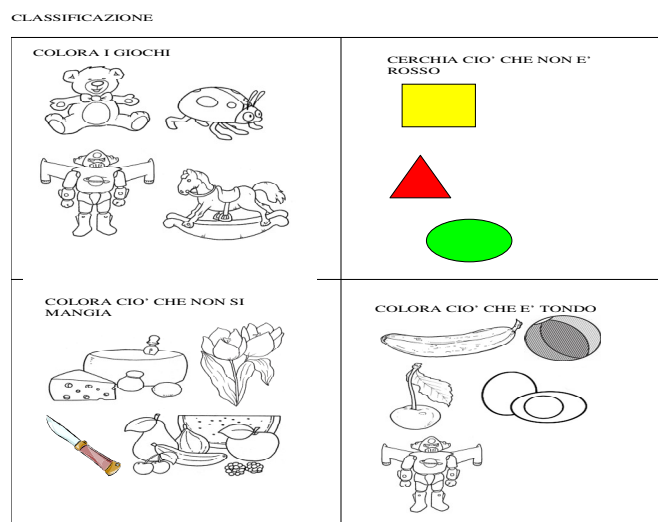


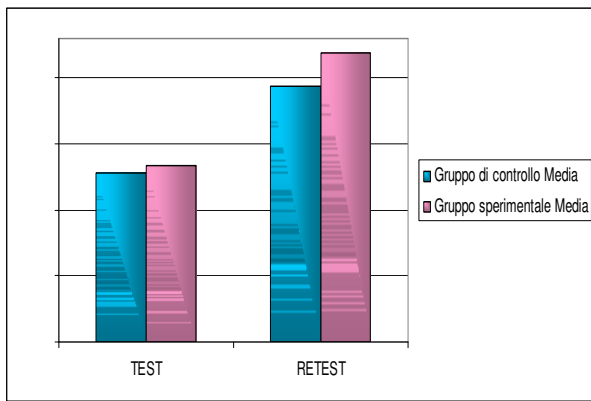
Figure1: An example for linguistic area

3 RESULTS

During the last stage, all the children were retested. The analysis of the data seems to confirm the efficacy of the training .

The comparison between the groups, non significant in the first stage of the research for what concerns the investigated areas , showed relevant differences not just in the comparison pre/post-test for the experimental group, but also as an improved competence in all the areas with reference to the final performances of the control group.

For instance, in the logical-mathematical subscale the initial score was similar for the control group of children of 5 years old ($M=3,3$ $DS=1,28$) and the experimental one ($M=3,10$ $DS=1,30$), but after the training the performance of the experimental group increased significantly ($M=4,54$ $DS=0,65$, $p= 0.001$), when compared to the pre-test and to the controls that, in this case, obtained lower scores ($M=2,89$ $DS=1,36$).



$F= 18.556$; $gdl= 1$; $p<=.001$

Figure 2: Logical mathematical area

4 CONCLUSION

The reduction of the initial disadvantage cannot be explained just with maturation or as a consequence of scholastic activities, but rather as the result of the opportunities promoted by the training activities, which became important stimulations capable of sustaining the development.

With this study we tried to demonstrate that it is possible to integrate the assessment phase with a specific intervention in order to maximize children's academic success in their learning environment.

References

- [1] C. Blair. School readiness: Integrating cognition and emotion in a neurobiological conceptualization of child functioning at school entry. *American Psychologist*, 20, No. 57, p. 111-127. 2002
- [2] C. Blair, L. Vernon-Feagans. Measurement of School Readiness. *Early Education and Development*, Vol. 17, No. 1, p. 1-5. 2006.
- [3] C. Blair, K. Knipe, C. Cummings, D. Baker, P. Eslinger, D. Gamson, S. Thorne. A developmental neuroscience approach to the study of school readiness. In R. Pianta, M. Cox, and K. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 149-174). Baltimore MD: Brookes. 2007
- [4] G. J. Duncan, C. J. Dowsett, A. Claessens, K. Magnuson, A. C. Huston, P. Klebanov, L. S. Pagani, L. Feinstein, M. Engel, J. Brooks-Gunn, H. Sexton, K. Duckworth, C. Japel. School readiness and later achievement. Peer Reviewed Journal *Developmental Psychology*. Vol. 43, No. 6, p.1428-1446. 2007.
- [5] C.S Dweck. *Teorie del sé. Intelligenza, motivazione, personalità e sviluppo*. Erickson, Trento. 2000.

[6] A. Gesell. *Studies in Child Development*. Westport, CT, Greenwood Press. 1948.

[7] J. W. Fantuzzo., P. A. McDermott, C. M McWayne. Preschool Competency in Context: An Investigation of the Unique Contribution of Child Competencies to Early Academic Success. *Developmental Psychology*, Vol. 40, No. 4, p. 633-645. 2004.

[8] K. M. La Paro, R. C.. Pianta. Predicting Children's Competence in the Early School Years: A Meta-Analytic Review. *Review of Educational Research*, Vol. 70, No. 4, p. 443-484. 2000.

[9] J. M.. Love, J. A. Aber, J. Brooks-Gunn. Strategies for assessing community progress toward achieving the first national educational goal. Princeton, NJ: Mathematica Policy Research, Inc. 1994.

[10] M. McClelland, C.E. Cameron, McDonald Connor C., C. L. Farris, A. M. Jewkes, F.J. Morrison. Links Between Behavioral Regulation and Preschoolers' Literacy, Vocabulary, and Math Skills. *Developmental Psychology*, Vol. 43, No. 4, 947-959. 2007.

[11] S. L. Kagan. Readiness 2000: Rethinking rhetoric and responsibility. *Phi Delta Kappan*, No. 72, p. 272-279. 1990.

[12] S. J. Meisels. Performance in context: Assessing children's achievement at the outset of school. In A. Sameroff and M. Haith (Eds.), *The five to seven year shift* (pp. 407-434). Chicago, IL: University of Chicago Press. 1996.

[13] M.A. Zanetti, D. Miazza. *SR 4-5 School Readness. Prove per l'individuazione delle abilità di base nel passaggio dalla scuola materna alla scuola elementare*, Trento, Edizioni Erikson. 2002.

COGNITONICS: A SOPHISTICATED LOOK AT SOCIALIZATION VIA VOGUE

Olga S. Fomichova (1), Anna V. Molyukova (2)

(1) State Educational Centre “Dialogue of Sciences”,
Universitetsky prospect 5, 119296 Moscow, Russia
e-mail: vfomichov@hse.ru and vdrfom@aha.ru
(2) 3rd year student, Faculty of Applied Art,
Kossygin Moscow State Textile University, Moscow, Russia

ABSTRACT

The paper aims at studying, under the framework of cognitronics, the ways the clothes influence the self-representation of the individual and his/her world perception. It is an attempt to represent the vogue as a tool of active transfiguration of the conceptual picture of the world of the young, of improving their sense of beauty and making the picture of the world more harmonized. Vogue is regarded as an effective tool for transforming the process of upbringing in the consciousness of the young from perceiving it as an impinging upon an individual's freedom to purposeful socialization.

1 INTRODUCTION

The general approach of Cognitronics as a new scientific discipline aimed at keeping the integrity of the human nature in the face of the ground-breaking modern shifts in the system of humanitarian values of the humans is based on a correlation of the achievements of various sciences dealing with the human's nature with the challenges of the information society and, in this respect, with the society's ultimate purpose [2 – 6]. The young are the most sensitive to the changed, pragmatic, commercialized profile of the society.

Evidently, fashion plays an important role in the modern life. Fashion can be treated as a message to the minds of people. That is why Vogue should answer the challenge of time, and it might be regarded as a tool for alleviating the existing distortion in the emotional and intelligent development of the teenagers, when eternal values are substituted by the commercialized ones.

An example of this distortion is as follows: the basic notions “like – dislike” are strongly underpinned by other, much more up-to-date notions “profitable – non profitable”. It is similar to the situation described by Oscar Wilde in his tale “Happy Prince”. Admiring the beauty of the statue of the Happy Prince, the Town Councillor remarked that it is as beautiful as a weather-cock, only not quite so useful, fearing

lest people should think him unpractical, which he really was not.

In the emotional sphere, everything is on the surface now. Deep emotions caused by magnanimity, compassion, anguish, remorse, uncertainty are not allowed any longer. The dress is provocatively open. One shouldn't expect to derive pleasure from solving the riddle: no mystery any more. Marketing logic is revealed in the personal relationships: a very short dress with a plunging neckline is very primitive in comparison with that designed by Coco Chanel, the Little Black.

The subject of this paper is studying, under the framework of cognitronics, (a) the ways the clothes influence the self-representation of the individual and his/her world perception, (b) the ways of forming the vogue as an effective tool for transforming the process of upbringing in the consciousness of the young from perceiving it as an impinging upon an individual's freedom to purposeful socialization.

2 THE BLACK DRESS OF COCO CHANEL. THE GIFT AND FUTURE GENERATIONS

Spiritual and intellectual maturity of a human being is revealed in the understanding of the necessity to keep the balance between the intellectual breakthrough and equal to it spiritual interpretation of what one has discovered. This balance is important lest the inner harmony, the essence of things should be destroyed by the transformation, transfiguration of the existing reality.

In this connection let us consider, following [5], the black dress of Coco Chanel, which is an example of the breakthrough that influenced the society a lot. Coco Chanel kept the enigma of the Woman, that is the inner essence of the Woman but, on the other hand, made the Woman much more free and independent. She transformed the nature of the Woman but kept the core, the essence of that nature – enigma – untouched. Coco Chanel created a new aesthetics, a new look without destroying the enigmatic essence of the Woman's nature.

A person who is aware of his/her intellectual and spiritual maturation regards his/her intuition and the moments of

inspiration as a treasure he/she is bestowed with. That is why the response of a human being is his/her hard work, his/her spiritual ascending in order to descend via objects of art, literature, scientific articles. This process provides an opportunity for the culture in general to ascend via the masterpieces in all domains of culture. In this way people can imbibe the Spirit of culture and be brought up as bearers of the culture.

According to the theory of Vyacheslav Ivanov (1866 – 1949), a world-known Russian philosopher and poet, an individual as a human being is always ascending and as an artist (in general) is always descending. Spiritual ascending of an individual, his/her spiritual development can't create objects of art, literature, scientific articles.

Descending is the process of searching for the forms of adapting cognition, spiritual discoveries to the world of humans. The culture in general is the process of descending, but it is rooted in the process of ascending of the individual. In all spheres of culture there are the points of intersection of the creative ascending and descending as the opposite directions of the first process.

But if those individuals who are bestowed with the gift that conduces to intellectual breakthrough are deeply influenced by the peculiarities of our time and, as a result, most of them become profit oriented and consequently they become dependable on that current profit, then they deprive their contemporaries and the following generations of spiritual revelation and intellectual breakthrough, because they can't enrich the culture by their discoveries in any domain caused by the revelation and, as a consequence, they can't contribute to the spiritual maturation of the current generation. That is why they become responsible for the gap between the previous generations and the following ones.

Art is a manifestation of emotions. Inspiration is an inclination of the soul to the active perception and transfiguration of the impressions. Intuition is a strong feeling concerning something without conscious reasoning. Emotion speaks a language that all may understand in case the emotional, spiritual sphere of the human being is developed. Strictly profit oriented people with "squashed" souls and feelings resist the reproduction of the culture and lead astray the people. It is revealed in the shift to commercialized values and in the fact of spiritual underdevelopment. The responsibility for the development and for the experience of evoking strong emotions in the hearts while looking at the masterpieces in all domains of culture should be taken by the mature people, especially by the gifted people.

The peculiarities of the time condition the changes in the consciousness of the young generation, in particular, when they are not aware of the gift as the necessity and possibility to use it not only for one's own sake but for the sake of other people. They don't regard this duty and this opportunity as their highest assignment. Their selfishness and materialistic approach to everything prevent them from understanding their social duty. In fact, their own personality is damaged, and the intellectual and spiritual potential of the personality

is artificially limited and mislead. It is one more example that explains the consequence of the lack of balance in our contemporary world between individual and collective, material and spiritual.

3 THE INHERENT SENSE OF BEAUTY AND THE DISTORTIONS OF ANY KIND

The child is born with an inherent sense of beauty. This unconscious sense of beauty conditions the aesthetic feeling which leads to the further development of the ethic feeling. If a child is asked to choose between a ripe apple and a rotten one, a child will choose a ripe one without hesitation. Any balanced situation ensures and encourages a human being and makes him/her feel safe, sound, and confident. On every level of perception one can find examples of balanced entities: good and evil, material and spiritual, intellectual and emotional.

The semantic trace from the experienced emotions evoked by strong impressions can't be defined as a passive reflection of a situation or a person. It is an active transfiguration according to the inherent in the consciousness of the human being ability to see a harmony or a lack of harmony, to put it in another words, to see whether the situation is balanced or not. It can be called the Phenomenon of Cinderella's Sisters [5, 6]: the girls were ordinary ones, well dressed (because Cinderella did it herself), but the lack of inner beauty contradicting their beautiful dresses and coiffures makes the readers think that the sisters were ugly.

If a person is shown a distorted square or any other figure as an image of some ongoing process then a person realizes that something is going wrong. This understanding is caused by the sense of beauty, sense of harmony being inherent in the consciousness of the human being.

On the other hand, a distorted square or lack of harmony manifested in any other way triggers the curiosity, attracts attention, because something is out of the way, something that doesn't meet the expectations of the inner vision, underpinned by the strong semantic trace from the experienced emotions caused by strong impressions from the observed or contemplated beauty/harmony revealed in various situations.

It leads to the intention of the person to correct, to harmonize the situation, to bring it to the perfect 50-50 balances. A person with an underdeveloped sense of beauty assumes the lack of harmony to some extent, doesn't see the ugliness of Cinderella's sisters and is much more tolerant to any one-sided shifts and distortions of any kind.

On the one hand, the distortion in this case doesn't evoke the curiosity which is the main precondition of cognition; on the other hand, a response to the distortion on the part of a person with an underdeveloped sense of beauty will be much weaker, less active, and it may lead to the underestimation of the on-going process or even to mistakes in one's private life.

4 THE CHALLENGE OF TIME AND AN ANSWER OF THE VOGUE

Vulgar market excludes the human nature, that is, the integrity of the soul, spirit, and body. People plunge at once at private relationships. They don't anticipate, don't contemplate, don't feel the whole cocktail of emotions, don't discover the taste little by little lest they should be fed up with the relationships. They are led by rigid pragmatism without even undertaking some efforts, as Antoine De Saint-Exupery wrote, "to prepare my heart to greet you".

"Men have no more time to understand anything. They buy ready-made things in the shops. But since there are no shops where you can buy friends, men no longer have any friends" [1, p. 78]. People don't lavish on their friends or beloved. It leads to the catastrophe that may be defined as the emasculation of feelings.

Vogue for the young should show the difference between primitivism and simplicity, between a trivial look and a sophisticated one. Clothes for the young should develop their taste and metaphoric way of perceiving the world, as in the poem by Boris Pasternak. It is simple but far from being primitive [7, p. 171]:

"But the quiver of intrusion / Through those curtains folds will run./ Measuring silence with your footsteps / Like the future, in you'll come./

You'll appear there in the doorway/Wearing something white and plain, / Something in the very stuff from / Which the snowflakes too are sewn."

Vogue can enhance the comprehension of the preciousness of the fleeting moment, making it brighter, stressing the exclusiveness of every minute, and providing the particular clothes. Clothes help to create special atmosphere of any social event or private meeting. This idea underpins the idea of dress code.

On the other hand, teenagers are sure that the idea of dress code is an example of impinging upon his/her individual freedom. In general, the upbringing is regarded by children as a kind of "taming of the shrew".

5 VOGUE AS A TOOL FOR TRANSFORMING CHILDREN'S ATTITUDE TOWARDS UPBRINGING

The purpose of fashion is to cause happiness. Beauty can be modern. Teenagers want to be appreciated, noticeable, but, on the other hand, they aspire to wear something what is in vogue. Children like everything extraordinary. They know Goethe as a poet but know nothing about his theory of colour. Taking into account the theory of colour elaborated by Goethe (1810), the teenagers should be explained the links between colours, colour and inner freedom, colour and "feeling of sight" aspiring to harmony. It is interesting to observe that, via the "feeling of sight", the colour influences the mood. Goethe came to the conclusion that colour adjoins the domain of morality, because colour harmonizes the state of soul.

J.M.W. Turner, an English painter (1775-1851), was deeply impressed by that theory. It influenced his painting a lot. He rejected the lines, his main concern became colour and light. Turner's works of art together with the Goethe's theory should help to develop children's taste, emotional sphere, and to establish the ties between the perception of the world and revealing oneself.

As for the cut of clothes, it should be simple and stylish, because the young are beautiful and like simple things. What really helps the teenagers to reveal taste and individuality, improve the sense of beauty and be changeable according to the mood is an accessory, either hand-made or bought.

6 CONCLUSION

1. It is a misapprehension to think that humans can do without the sense of harmony/ beauty. In fact, the first question people ask themselves is whether they like something or not, and their further behavior depends on the answer. That is why the sense of beauty is regarded as a basic element of the educational process in the frame of Cognitronics.
2. Vogue reveals the atmosphere in the society like a portrait of the time. But any portrait is a double reflection: reflection of the time and the reflection of the perception of time in the painter's conceptual picture of the world. In our case, a teenager is a painter: he/she paints his/her self-portrait, he/she is not sure in lines, in profile but colour, light, and details will work wonders.
3. The style is everything to do with casual chic. Chic reveals the developed taste of the personality and the degree of spiritual and intellectual maturity.
4. A young girl's love to cloth is a strong motive for the designers to show the difference between just putting on something and being born to wear beautiful things, being an embodiment of beauty. The girls should be brought up with the idea that they are born to introduce beauty into the world.
5. The beliefs are formed in one's childhood and during school years. There is no contradiction between being delicate and light and being intelligent and strong. *The ballerinas seem to be delicate and light as feathers, but they are strong as steel, pitting themselves against the body's limitations. The illusion of effortlessness that comes about through endless practice and the impression of frailty conveyed by women with the strength of professional athletes – that is of the modern women that should be explained via vogue for girls.* It is to prevent the girls from wearing masculine clothes and acquiring masculine ways (boyish ways).
6. The combination of delicacy and daring should be revealed in the vogue for the youth (with the help of the colour and accessories, in particular).
7. Gait is very important, it speaks volumes. Clothes and shoes influence it a lot.

8. Getting to know the rules of vogue and the power of casual chic should not become a peeling-the-onion experience for the teenagers. They should be reared with it, and it will improve their sense of self.
9. The vision of Cognitonics makes it possible to put together the investigations in various sciences and art: in poetry and mathematics, art (in particular, Whistler “Harmony in Grey and Pink”) and psychology (“to make peace with oneself”), cognitive science and economics, linguistics and computer science, philosophy and gardening in order to elaborate a new conceptual picture of the world underpinned by the eternal values, the well developed sense of beauty, which may harmonize the on-going processes and keep the integrity of the human nature.

(ed. by G.E. Lasker and K. Hiwaki), Vol. IV. Addresses and Papers from the 4th International Symposium on Personal and Spiritual Development in the World of Cultural Diversity, August 1, 2007, held at the 19th International Conference on Systems Research, Informatics and Cybernetics – InterSymp-2008, July 30 – August 4, 2007, Baden-Baden, Germany; The International Institute for Advanced Studies in Systems Research and Cybernetics, Tecumseh, Ontario, Canada, 2007 (pp. 27-31)

- [6] O.S. Fomichova. *Humanitarian Education – an Answer to the Challenge of Time*. Moscow State University Publishing House. 2009. In Russian.
- [7] Pasternak, B. (1990). *Poems*. Raduga Publ. House, Moscow (317 pp)

References

- [1] De Saint-Exupery, A. (1995). *The Little Prince*. Translated by Irene Testot-Ferry; Wordsworth Editions Limited, Hertfordshire (109 pp)
- [2] Fomichov, V.A. and O.S. Fomichova (2006). Cognitonics as a New Science and Its Significance for Informatics and Information Society. *Special Issue on Developing Creativity and Broad Mental Outlook in the Information Society (Guest Editor Vladimir Fomichov), Informatica. An International Journal of Computing and Informatics (Slovenia)*, Vol. 30, No. 4 (pp. 387-398)
- [3] Fomichova, O.S. and V.A. Fomichov (2007). Cognitonics as a New Science and Its Social Significance in the Age of Computers and Globalization. *G. E. Lasker and M. Malatesta (Eds.), Systems Research and Cognition*, Vol. 1. The International Institute for Advanced Studies in Systems Research and Cybernetics, Tecumseh, Ontario, Canada (pp. 67-75)
- [4] Fomichova, O.S., Fomichov, V.A. and L.D. Udalova (2004). The Ways of Cultural Interaction and a Personal Response to It; In *Personal and Spiritual Development in the World of Cultural Diversity (ed. by G.E. Lasker and K. Hiwaki), Vol. 1. Addresses and Papers from the 1st International Symposium on Personal and Spiritual Development in the World of Cultural Diversity, July 30, 2004, held at the 16th International Conference on Systems Research, Informatics and Cybernetics – InterSymp-2004, July 29 – August 5, 2004, Baden-Baden, Germany*; The International Institute for Advanced Studies in Systems Research and Cybernetics, University of Windsor, Ontario, Canada, 2004 (pp. 47-51)
- [5] Fomichova, O.S., Fomichov, V.A. and Udalova, L.D. (2007). Cognitonics: The Phenomenon of the Cinderella’s Sisters // In *Personal and Spiritual Development in the World of Cultural Diversity*

Zbornik 12. mednarodne multikonference
INFORMACIJSKA DRUŽBA – IS 2009

Proceedings of the 12th International Multiconference
INFORMATION SOCIETY – IS 2009

2. mini konferenca iz teoretičnega računalništva 2009
The 2nd Mini Conference on Theoretical Computer Science
2009

Uredil / Edited by
Andrej Brodnik

<http://is.ijs.si>

14.–15. oktober 2009 / October 14th–15th, 2009
Koper, Slovenia

PREDGOVOR

Drage bralke in bralci!

Pred vami so povzetki druge Minikonference iz teoretičnega računalništva. Minikonferenca je letos prvič del multikonference Informacijska družba, a bo v prihodnje brez dvoma njen reden član. Slednje pomeni velik doprinos tako za področje teoretičnega računalništva, ki se pojavlja v širši družbi, kot tudi celotni stroki računalništva in informatike, ki ga pokriva multikonferenca, saj ji daje tudi tisto temeljno podstat, ki jo včasih pogrešamo.

Minikonferenca se odvija na Univerzi na Primorskem že drugo leto. Njen namen je vzpostaviti okolje za srečevanje med najboljšimi podiplomskimi študenti in strokovnjaki s področja teoretičnega računalništva. Temu sledi tudi struktura minikonference, ki je dvodnevna. Prvi dan so na sporedu študentski članki in drugi dan znanstveni članki. Študentski članki so letos razdeljeni na 3 sekcije, v katerih je predstavljenih 9 prispevkov s področja podatkovnih struktur in 2 sekciji s sedmimi članki s področja teoretičnih osnov računalništva.

Letos bo na konferenci predstavljenih 8 znanstvenih člankov. Kar je posebno vzpodbudno je to, da članki posegajo na področje uporabnega teoretičnega računalništva. Članki, ki jih predstavljajo József Békési, David Paš in Atilla Tóth sodijo na področje razporejanja. Vsi njihovi prispevki so motivirani z realnimi problemi, za katere predlagajo model reševanja in tudi hevristične rešitve.

Problem, ki ga predstavlja Carlos Kavka, je s področja načrtovanja vezij, ki prav tako sodi med NP-polne probleme. V svojem prispevku Carlos predstavlja rešitev z uporabo orodja modeFRONTIER. Tudi problem polnjenja košev, ki ga opisuje Gábor Galambos v svojem vabljenem predavanju, je v osnovi NP-poln. Različica sprotne algoritma je še toliko bolj zanimiva ter dovoljuje zgolj približno rešitev. V članku se bomo sprehodili skozi nekaj spodnjih mej možnih približkov.

Miklós Krész se v svojem prispevku loteva problema učinkovite redukcije grafa ter pokaže, da je možna samo ob določenih pogojih. Posebnost sta članka Gregorja Smogavca in Janeza Žiberta. Oba se ukvarjata z obdelova velikih količin podatkov. Podatki Gregorja Smogavca so geografski oziroma geometrijski, medtem ko so pri Janezu Žibertu zvočni. Pri Gregorju Smogavcu imamo opravka z geometrijskim približkom terena, po katerem se mora gibati avtonomen robot. Dobra stran rešitve je v tem, da je zaradi približnostnega pristopa hitra in omogoča procesiranje v stvarnem času. Prispevek Janeza Žiberta sodi na področje razpoznave govorcev v zvočnem toku.

Naj vam še enkrat v imenu programskega in organizacijskega odbora izrazimo dobrodošlico ter povabimo, da nas obiščete tudi drugo leto!

Andrej Brodnik, David Paš

PREFACE

Dear readers,

here are the abstracts of the second Miniconference on theoretical Computer Science. The Miniconference is for the first time part of a multiconference Information Society, but we believe it will become its regular member. This means a twofold advantage: first for the Theoretical Computer Science itself for its appearance in a general society, and second for the whole Computer Science and a multiconference which got a firm theoretical basis that is sometimes missed.

The Miniconference is taking place at the University of Primorska for the second year. Its purpose is to establish an environment for meeting the best graduate students and professionals in the area of the theoretical Computer Science. Therefore the conference consists of student sessions (the first day) and regular sessions (the second day). Three student sessions with 9 papers cover the area of data structures and 2 sections with 7 papers cover the area theoretical basis of Computer Science.

This year we have at the conference 8 regular papers and we are especially happy to notice that all of them are from the area of applied theoretical Computer Science. Papers presented by József Békési, David Paš and Atilla Tóth address scheduling problems. They describe models and hevristical solutions for various real world problems.

Carlos Kavka in his paper addresses the problem of chip design. This is also NP-hard problem and Carlos in his contribution presents a solution obtained using tool modeFRONTIER. Further, bin packing problem addressed by Gábor Galambos in his invited lecture is in its basis also NP-hard. The variant he studies is an on-line version of the problem and he presents some lower bounds on the approximate solutions.

An efficient graph reduction method is a topic studied by Miklós Krész. He shows that the efficient graph reduction is possible only under some specific conditions. The papers presented by Gregor Smogavec and Janez Žibert are different from other papers as the emphasis in their papers is on the amount of data they deal with. The data processed by Gregor are geographic or geometric while Janez processes sound. Gregor Smogavec deals with a terrain approximation through which a robot has to move. His solution is slightly inaccurate though fast which permits processing in real-time. Janez in his paper studies the problem of speaker recognition in continuous audio streams and presents several methods and approaches to solve it.

At last but not least we want to express once more our welcome to you and invite you to join us the next year again!

Andrej Brodnik, Davi Paš

Programski odbor / Programme committee:

Andrej Brodnik (predsednik/chair), Miklós Krész, József Békési, Janez Žibert

Organizacijski odbor / Organizing committee:

David Paš (predsednik/chair), Katjuša Gorela, Andrej Kramar

Regular Session

Lower Bound on On-Line Bin-Packing Algorithms

Gábor Galambos¹, János Balogh¹, József Békési¹

¹*Department of Informatics' Applications, Faculty of Teachers Training, University of Szeged, Hungary*

keywords: bin packing, online algorithms, analysis of algorithms

Abstract

One of the most frequently studied combinatorial problems is one-dimensional bin packing: We are given a list $L = \{x_1, x_2, \dots, x_n\}$ of real numbers in $[0, 1)$, and an infinite list of unit capacity bins. Each number x_i has to be assigned to a unique bin such that the sum of the elements in each bin does not exceed 1. Our aim is to minimize the number of used bins. It is well-known that finding an optimal packing is NP-hard. Consequently, large number of papers have been published which look for polynomial time algorithms with an acceptable approximative behavior. Some of these algorithms are on-line algorithms: They put items into a bin as they appear without knowing anything about the subsequent elements (neither the size nor the number of elements).

In this talk first we will overview the history of the certain lower bounds for the on-line algorithms. Thereafter we will consider a special class of lists: we suppose that the elements are preordered in decreasing order, and we need to pack the elements by the on-line rule. We will show an improvement of a 26 years old result, where Csirik et al. proved an $8/7$ lower bound on these type of lists. Our new lower bound is $54/47$.

A Flexible Method for Driver Scheduling in Public Transportation

Attila Tóth¹

¹*Department of Informatics' Applications, Faculty of Teachers Training, University of Szeged, Hungary*

keywords: public transportation, driver scheduling

Abstract

Nowadays scheduling of public transportation is a very important and crucial question for transportation companies. However, the development of an automatic scheduling system is a very complex task, since the public transportation of even a middle size city induces a large scale NP hard scheduling problem. Usually two subproblems are distinguished, such as vehicle scheduling and driver scheduling, the objectives. Nevertheless, both vehicle and driver scheduling are special assignment problems, but for drivers, the number of the constraints and their complexity are significantly higher. This talk describes the problem statement of the driver scheduling and introduces a flexible system. This system divides the scheduling into five different sequent steps and provides the possibility to combine different methods to produce the solution.

Design Space Exploration for Embedded Parallel System-on-Chip Platforms using modeFRONTIER

C. Kavka¹, L. Onesti², P. Avasare³, G. Vanmeerbeeck³, M. Wouters³ and H. Posadas⁴

¹ESTECO SRL, Trieste, Italy, carlos.kavka@esteco.com

²ESTECO SRL, Trieste, Italy, luka.onesti@esteco.com

³IMEC vzw, Leuven, Belgium, {avasare, vanmeerb, woutersm}@imec.be

⁴University of Cantabria, Santander, Spain, posadash@teisa.unican.es

Abstract

A complete design space analysis of a System-on-Chip (SoC) architecture is practically prohibitive due to the large number of possible architectural configurations and the long time required to perform system simulations. The problem is usually more complex due to the existence of multiple and often conflicting objectives to optimize. In this type of problems, carefully chosen optimization algorithms which explore wisely the design space by considering all objectives at the same time need to be applied. These optimization algorithms do not produce a single solution, but a set of non-dominated solutions named Pareto front, which represent a good compromise between all conflicting objectives.

Many methods have been proposed in the literature to identify the Pareto front for SoC design problems. However, the use of powerful design exploration methods by itself does not guarantee that optimum solutions will be found in reasonable time with a reasonable use of resources. It is important for the SoC designer to have access to a design environment that provides the ability to express the design problem in clear terms, understanding the relevant characteristics of the problem, and allowing them to discover how these characteristics change with the problem specification and parameter values. In this way, it is possible to avoid unrealistic combinations of parameters, choose representative points and prune the design space as much as possible.

The modeFRONTIER design environment is one of the most widely used tools for multi-objective optimization in complex engineering domains. In the EU MULTICUBE project, modeFRONTIER is being retargeted to the domain of Embedded Parallel SoC design. The interaction between modeFRONTIER and the high level simulators is performed by using an open XML specification, which allows the integration with various simulators (or models) for SoC platforms and architectures. During the optimization process, modeFRONTIER provides values for the system configuration parameters and expects back from the simulator the corresponding system metrics. The work presents how an engineering specific workflow can be schematized into the modeFRONTIER's framework, how local or remote application can be transparently executed and how the complex output (measures) can be assessed using traditional and innovative Data Mining techniques.

Initial optimization experiments have been performed using two high level simulators (IMEC-HLsim and MULTICUBE-SCoPE) running an MPEG4 encoder multimedia application. Configuration parameters considered till now are the number of CPUs, the instruction cache size and the processor frequency. System metrics like power consumption, latency and execution time are used as optimization objectives. The Design of Experiments (DOE) module helps to define the initial set of designs for the exploration and the set of points for reliable Response Surface Models (RSM). The RSM module allows users to save simulation time by creating statistically validated mathematical models that approximate the behavior of the simulator. Post-processing tools allow users to obtain information from the SoC system and determine correlations, extracting useful design features. The optimization experiments performed till now show promising results, suggesting that modeFRONTIER can provide an easy to use and flexible design environment for SoC platforms.

Improved Analysis of an Algorithm for the Coupled Task Problem with UET Jobs

József Békési¹, Gábor Galambos¹, Marcus Oswald², Gerhard Reinelt²

¹*Department of Informatics' Applications, Faculty of Teachers Training, University of Szeged, Hungary*

²*Institute of Informatics, University of Heidelberg, Germany*

keywords: scheduling, coupled task problem, analysis of algorithms

Abstract

The coupled task problem is to schedule n jobs, each one consisting of two subtasks with exact delay times between them, on a single machine. In this talk we present a new lower bound for the problem variant with unit execution times and correct an analysis of Ageev and Baburin.

A Framework for a Flexible Vehicle Scheduling System

David Paš¹, József Békési², Miklós Krész², Andrej Brodnik¹

¹*Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Slovenia*

²*Department of Informatics' Applications, Faculty of Teachers Training, University of Szeged, Hungary*

keywords: transportation, multiple depot vehicle scheduling

Abstract

A flexible framework for vehicle scheduling is presented which was developed for public bus transportation systems. The framework builds upon the theoretical foundation that has been established in this area and extends it to incorporate additional real-world constraints. We have applied our approach to the urban bus system of two middle-sized cities, and computed substantially better schedules than the previous ones prepared by-hand.

Medial Axis Approximation of Simple Polygons

Gregor Smogavec¹, Borut Žalik¹

¹Faculty of Electrical Engineering and Computer Science, University of Maribor, Slovenia

keywords: medial axis, triangulation, Centers of maximal disks

Abstract

Medial axis is a set of centers of maximal inscribed discs [1]. It is used as an essential part in a variety of applications like character recognition, GIS, robotics, collision detection, image description and analysis. Since exact computation of the medial axis is difficult and time consuming, it limits the above applications in a sense of real-time execution [2]. In this paper we propose an algorithm, which finds an approximation of a medial axis. The algorithm works in two steps. Firstly, a constrained Delaunay triangulation is constructed. Polygon is after that further triangulated with Steiner points to meet the criteria of introduced heuristics. The centers of gravity of neighboring triangles are connected and an approximation of the medial axis of a polygon is obtained (see Figure below).

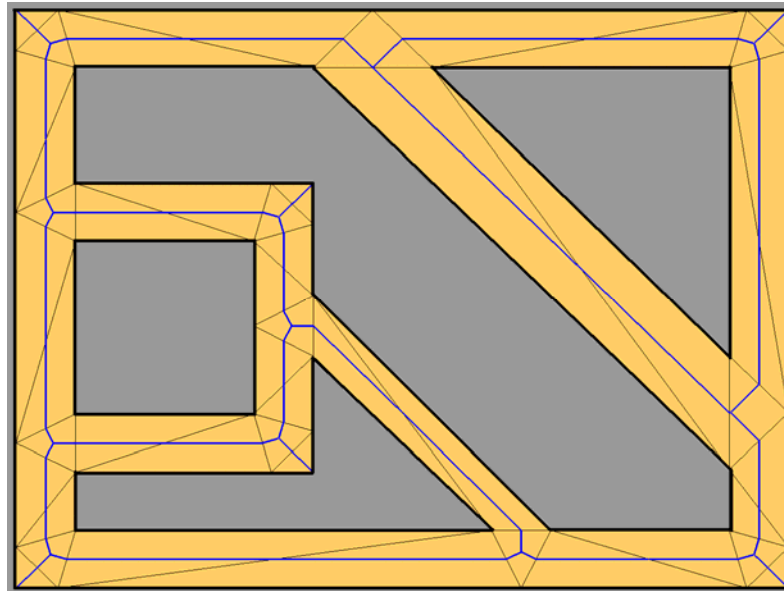


Figure: An approximate polygon's medial axis

Experiments show that approximation of the medial axis is obtained very rapidly, what gives the possibility to use the algorithm in real-time applications; for example in robotics.

References:

- [1] E. Remy, E. Thiel, Exact medial axis width Euclidean distance, *Image and Vision Computing*, vol. 23, no. 2, pp. 167-175, 2005
- [2] [Tim Culver](#), [John Keyser](#), [Dinesh Manocha](#), Accurate computation of the medial axis of a polyhedron, *Proceedings of Fifth Symposium on Solid Modeling and Applications (ACM Solid Modeling '99)*, pp. 179-190, 1999.

An Efficient Graph Reduction Method

Miklós Krész¹, Miklós Bartha²

¹*Department of Informatics' Applications, Faculty of Teachers Training, University of Szeged, Hungary*

²*St. John's, University of Newfoundland, Canada*

keywords: reduction, unique perfect matching, depth-first trees, implied redex

Abstract

A redex in a graph G is a triple $r = (u; c; v)$ of distinct vertices that determine a 2-star. Shrinking r means deleting the center c and merging u with v into one vertex. Reduction of G entails shrinking all of its redexes in a recursive way, and, at the same time, deleting all loops that are created during this process. It is shown that reduction can be implemented in linear time only, when implemented on depth-first trees in a smart way.

Speaker Tracking in Broadcast News: a Case Study

Janez Žibert¹

¹*Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Slovenia*

keywords: speaker diarization, speech detection, audio segmentation, speaker clustering, audio indexing, speaker recognition, speaker tracking

Abstract

A system for tracking speakers within audio data of broadcast news shows is presented and the impacts of the main components of the system to the overall speaker-tracking performance are evaluated. The process of speaker tracking in continuous audio streams involves several processing tasks and is therefore treated as a multistage process. The main building blocks of such system include the components for audio segmentation, speech detection, speaker clustering and speaker identification. The aim of the first three processes is to find homogeneous regions in continuous audio streams that belong to one speaker and to join each region of the same speaker together. The task of organizing the audio data in this way is known as a speaker diarization and plays an important role in various speech-processing applications. In our case the impact of speaker diarization was assessed in a speaker-based audio indexing system by performing a comparative study of how each of the component influenced the overall speaker-detection results. The valuation experiments were performed on broadcast-news audio data with a speaker-tracking system, which was capable of detecting 41 target speakers. We implemented several different approaches in each component of the system and compared their performances by inspecting the final speaker-tracking results. The evaluation results indicate the importance of the audio-segmentation and speech-detection components, while no significant improvement of the overall results was achieved by additionally including a speaker-clustering component to the speaker-tracking system.

Student Session

IPTOR

David Božič, University of Primorska, Gene network analysis

David Paš, University of Primorska, Networks of Evolutionary Processors

Milan Djordjević, University of Primorska, Complexity Theory and Randomized Algorithms: The Perspective of Grafted Evolutionary Algorithm

Peter Race, University of Primorska, DNA computation via Self-Assembly

Jakob Bartolj, University of Primorska, Quantum Computing

Gregor Ambrožič, University of Primorska, Communication complexity as a lower bound for learning in games

Aleš Kalin, University of Primorska, Parsing formal languages

IPPS

Jakob Bartolj, University of Primorska, Finger search trees with constant insertion time

David Božič, University of Primorska, Funnelheap: a cache-oblivious priority queue

Marko Grgurovič, University of Primorska, Dynamic LCA queries on trees

Marko Mikolavčič, University of Primorska, Optimal pointer algorithms for finding nearest common ancestors in dynamic trees

Gregor Tomazin, University of Primorska, Overcoming the memory bottleneck in suffix tree construction

David Paš, University of Primorska, Linear time construction of suffix arrays

Aleš Kalin, University of Primorska, Sorting and searching in faulty memories

Peter Race, University of Primorska, Cache-oblivious priority queue and graph algorithm applications

Tomasz Świderek, Maria Curie-Skłodowska University, Cache-oblivious data structures for orthogonal range searching

Indeks avtorjev / Author index

Allegra M.	440
Amberg Michael.....	298
Arčon Iztok.....	186
Arh Tanja	154
Avasare P.	476
Avbelj Viktor	362
Babič Jan	411
Babič Mitja.....	403
Bajd Tadej	399
Bajgot Mateja	26
Bajželj Jelka	147
Balogh János	475
Bartha Miklós.....	479
Bauer Andrej.....	240
Bazhenova Elena.....	360
Begg Graham	30
Békési József.....	475, 477
Bele Julija Lapuh.....	166
Bernik Mojca.....	151
Bernsteiner Reinhard.....	294
Berselli Giovanni.....	403
Bizant Sašo.....	148
Bohanec Marko	85, 351
Bontila Maria	444
Bosnić Zoran	7, 46
Bregar Andrej.....	11
Brimicombe Allan J.	311
Brodnik Andrej.....	477
Brožič Liliana.....	149
Bučar Urška.....	150
Budihna Alenka.....	179
Budimac Zoran.....	307
Burdescu Dumitru Dan.....	460
Burgar Bojan	151
Butka Peter	249
Čepon Peter	399
Cestnik Bojan	85
Chalupová Naděžda.....	15
Chiazzese G.....	440
Chifari A.....	440
Chisăliță-Crețu Camelia	269
Colja Nevenka	152
Corduneanu Maria	19, 22
Cosoi Carmen Maria	19, 22
Cosoi Catalin Alexandru	19, 22
Curkova Matea	172
Cvitkovič Dejan.....	153
Daffern Thomas C.....	365
Dagdilelis Vasssilios	444
Dali Lorand	203, 211
Debevec Aleksandra.....	176
Dinevski Dejan.....	154
Dolenc Barbara.....	371
Dovgan Erik	89, 97
Džeroski Sašo	30
Ferlež Jure	222

Florjančič Jože.....	151
Fomichov A. Vladimir.....	431
Fomichova S. Olga.....	431, 467
Fon Miroslava.....	155
Fortuna Blaž.....	203
Fošnarič Samo.....	154
Frece Aleš.....	273
Galambos Gábor.....	475, 477
Gamberger Dragan.....	240
Gams Andrej.....	415, 419
Gams Matjaž.....	89, 93, 116, 333
Georgiev Strahil.....	171
Gerlec Črt.....	265
Gerlič Ivan.....	156
Gider Franc.....	157
Godena Giovanni.....	253
Goljat Uroš.....	281
Gorenjak Mario.....	26
Grobelnik Marko.....	203, 218, 222
Hajdinjak Blaž.....	411
Heričko Marjan.....	261, 281, 294, 303
Hlade Milan.....	158
Horvat Boris.....	159, 170
Huber Jernej.....	315
Ijeh Anthony C.....	311
Imafidon Chris O.....	311
Ionascu Costem Marian.....	460
Ivanišin Marko.....	160
Ivanović Mirjana.....	307
Ivanovska Aneta.....	30
Jaakkola Hannu.....	285
Jankovič Zoran.....	307
Jarc Boštjan.....	166
Jarc Tomo.....	171
Kaluža Boštjan.....	97
Kamnik Roman.....	399
Kastelic Alenka.....	161
Kavka C.....	476
Klauser Julian.....	34
Knap Žiga.....	355
Kocjan Egon.....	38
Kogoj Mojca.....	162
Kononenko Igor.....	34, 38, 46
Kordeš Urban.....	329
Korošec Darko.....	163
Kos Andrej.....	407
Košič Kristjan.....	294
Kozoderc Danilo.....	164
Kreševič Irena.....	165
Krész Miklós.....	477, 479
Krivec Jana.....	101, 116
Kryštof Jan.....	257, 277
Kuhar Saša.....	261
Kukar Matjaž.....	50
Kuželički Jernej.....	399
Lavrač Nada.....	85, 211, 240
Lenarčič Jadran.....	403
Linna Petri.....	285
Logofatu Bogdan.....	460
Lokar Matija.....	159, 167, 170

Loskovska Suzana.....	62
Lotrič Komac Tatjana.....	168, 169
Lukman Tomaž.....	253
Lukšič Primož.....	159, 170
Luštrek Mitja.....	105
Mahnič Viljan.....	50, 171
Malnarič Sonja.....	175
Marija Kotyurova.....	360
Marinčič Domen.....	109
Markič Olga.....	325
Marko Ribič.....	184
Mastnak Tanja.....	179
Mastouri Azza.....	54
Mastouri Manal.....	54
Merlo G.....	440
Miazza Daniela.....	456, 464
Mihaescu Marian Cristian.....	460
Mihajlović Slađana.....	390
Mihelač Lorena.....	173
Mihelj Matjaž.....	423
Mircevska Violeta.....	112
Mladenec Dunja.....	207, 222
Mladenić Dunja.....	203, 214, 218
Mlakar Broder Jožica.....	174
Molyukova V. Anna.....	467
Motyčka Arnošt.....	15, 58, 257, 277
Možina Miran.....	337
Munih Marko.....	399, 423
Nančovska Šerbec Irena.....	192
Novak Domen.....	423
Novak Tatjana.....	379
Novalija Inna.....	214
Ogorevc Marko.....	42
Olenšek Andrej.....	423
Omerza Damijan.....	159
Onesti L.....	476
Orbanić Alen.....	159
Oswald Marcus.....	477
Ottaviano S.....	440
Ožek Matej.....	116
Pajntar Boštjan.....	218
Paš David.....	477
Pavlakovič Tomaž.....	175
Pejčić Biljana.....	26
Pekolj Rok.....	175
Petrič Tadej.....	415
Pevec Darko.....	46
Pfeifer Jože.....	290
Piltaver Rok.....	89, 120
Pirtošek Zvezdan.....	323
Plankar Matej.....	375
Podobnik Janez.....	423
Pogorelc Bogdan.....	124
Pograjc Debevec Marjanca.....	176
Pollak Senja.....	228
Popelka Ondrej.....	66
Posadas H.....	476
Potočar Zdenko.....	177
Poženel Marko.....	50
Preston David S.....	311

Prezelj Andreja	178
Procházka David	257, 277
Purg Peter	179
Pušnik Maja.....	303
Raggad Bel G.	54
Rakar Stanka	180
Rakovec Gorazd	181
Rakovec Žumer Irena.....	182
Raszková Magdalena.....	58
Ratej Mateja	357
Rebolj Vanda.....	183
Reinelt Gerhard	477
Reinhardt Michael	298
Rijavec Darja.....	185
Ristevski Blagoj	62
Rozman David.....	166
Rugelj Jože	192
Rusu Delia	203
Šajne Damjana.....	186
Šef Tomaž	128
Seta L.	440
Sgarciu Valentin.....	19
Shchukina Irina	387
Škarja Peter	187
Slokar Čevdek Magda	188
Sluban Borut.....	240
Šmitek Branislav	187
Smogavec Gregor.....	478
Soini Jari.....	285
Šorgo Andrej	189
Sovec Andrej.....	26
Špernjak Andreja.....	189
Šprah Lilijana	371, 379
Štajner Tadej	203
Stastny Jiri	66
Stencl Michael.....	66
Štiglic Gregor	26
Stiplovšek Amalija	190
Stojmenovič Jelena.....	191
Štrajhar Jožica	191
Strnad Mateja	192
Strniša Gašper	193
Štrucl Jože	193
Styskin Andrey	70
Šumak Boštjan	303
Sušnik Dušan.....	149
Szendy Nicole	448
Tancig Simona	343
Tavčar Aleš	132
Tisnikar Viljem	74
Todorovski Ljupčo	30
Tomac Stanojev Polona.....	194
Tomaško Mladen.....	232
Totev M. Dimitar.....	452
Tóth Attila	475
Trampuš Mitja	207
Tsayang Gaba.....	452
Tubin Pavel	78
Turčinek Jan.....	82
Ude Aleš.....	419

Ule Andrej	347
Urbančič Tanja	157, 186
Vanmeerbeeck G	476
Vavpetič Anže	236
Vehar Jerman Andreja	195
Vertechy Rocco	403
Vidmar Barbka	196
Vidulin Vedrana	136
Vlad Madalin	19
Vlad Madalin Stefan	22
Weng Jiangping	298
Wouters M.	476
Yasuda Hirohiko	435
Žagar Pernar Tina	168
Zakrajšek Srečo	197
Žalik Borut	478
Zanetti Maria Assunta	456, 464
Žepič Mojca	198
Žerovnik Eva	383
Žibert Janez	479
Ziherl Jaka	423
Živkovič Aleš	265
Žlajpah Leon	415
Žnidaršič Martin	85

prof. dr. Marko Bohanec
prof. dr. Matjaž Gams

Intelligentni sistemi
Intelligent Systems

prof. dr. Vladislav Rajkovič
prof. dr. Tanja Urbanžič
Mojca Bernik

Vzgoja in izobraževanje v informacijski družbi
Education in Information Society

dr. Dunja Mladenič
Marko Grobelnik

Izkopavanje znanja in podatkovna skladišča
Data Mining and Data Warehouses (SiKDD-2009)

prof. dr. Marjan Heričko

Sodelovanje, programska oprema in storitve v informacijski družbi
Collaboration, Software and Services in Information Society

dr. Urban Kordeš
prof. dr. Olga Markič
prof. dr. Matjaž Gams

Kognitivne znanosti
Cognitive Sciences

prof. dr. Jadran Lenarčič
dr. Leon Žlajpah
dr. Andrej Gams

Robotika
Robotics

dr. Olga S. Fomichova
dr. Vladimir A. Fomichov

Kognitonika
Cognitonics

dr. Andrej Brodnik

Druga mini-konferenca iz teoretičnega računalništva
The Second Mini Conference on Theoretical Computer Science