

BMSD 2011

First International Symposium on
Business Modeling and Software Design



Proceedings

Sofia, Bulgaria • 27 - 28 July, 2011

Organized by:



In Cooperation with:



TECHNICAL UNIVERSITY OF SOFIA



Technical Co-Sponsorship:



QlikView

BMSD 2011

Proceedings of the
First International Symposium on
Business Modeling and Software Design

Sofia, Bulgaria

July 27 – 28, 2011

Organized by
**IICREST - Interdisciplinary Institute for Collaboration and Research on
Enterprise Systems and Technology**

In Cooperation with
CTIT - Center for Telematics and Information Technology
**INSTICC - Institute for Systems and Technologies of Information, Control
and Communication**

Technical University of Sofia

Technical Co-Sponsorship by
Sofia Municipality
QlikTech Netherlands B.V.

Copyright © 2011 SciTePress - Science and Technology Publications
All rights reserved

Edited by Boris Shishkov

Graphics Production by Bozhana Yankova

Printed in Portugal

ISBN: 978-989-8425-68-3

Depósito Legal: 330150/11

<http://www.is-bmsd.org>

secretariat@iicrest.eu

BRIEF CONTENTS

KEYNOTE SPEAKERS	IV
CHAIR AND PROGRAM COMMITTEE.....	V
BEST PAPERS SELECTIONS	VIII
FOREWORD.....	IX
CONTENTS.....	XI

KEYNOTE SPEAKERS

Mehmet Aksit

University of Twente

The Netherlands

Dimitar Christozov

American University in Bulgaria - Blagoevgrad

Bulgaria

Bart Nieuwenhuis

University of Twente

The Netherlands

Hermann Maurer

Graz University of Technology

Austria

CHAIR AND PROGRAM COMMITTEE

CHAIR

Boris Shishkov, IICREST, Bulgaria

PROGRAM COMMITTEE

Mehmet Aksit, University of Twente, The Netherlands

Antonia Albani, University of St. Gallen, Switzerland

Ognyan Andreev, Technical University of Sofia, Bulgaria

Paulo Anita, Delft University of Technology, The Netherlands

Rumen Arnaudov, Technical University of Sofia, Bulgaria

Colin Atkinson, University of Mannheim, Germany

Csaba Boer, TBA, The Netherlands

Boyan Bontchev, Sofia University St. Kliment Ohridski, Bulgaria

Frances Brazier, Delft University of Technology, The Netherlands

Barrett Bryant, University of Alabama at Birmingham, USA

Cinzia Cappiello, Politecnico di Milano, Italy

Kuo-Ming Chao, Coventry University, UK

Ruzanna Chitchyan, Lancaster University, UK

Samuel Chong, Capgemini, UK

Dimitar Christozov, American University in Bulgaria - Blagoevgrad, Bulgaria

Selim Ciraci, University of Twente, The Netherlands

José Cordeiro, Polytechnic Institute of Setúbal, Portugal

Dumitru Dan Burdescu, University of Craiova, Romania

Joop De Jong, Delft University of Technology, The Netherlands

Jan L. G. Dietz, Delft University of Technology, The Netherlands

Lyubka Doukovska, Bulgarian Academy of Sciences, Bulgaria

Joaquim Filipe, Polytechnic Institute of Setúbal, Portugal

Boris Fritscher, University of Lausanne, Switzerland

J. Paul Gibson, T&MSP - Telecom & Management SudParis, France

PROGRAM COMMITTEE (CONT.)

Eduardo Goncalves Da Silva, University of Twente,
The Netherlands

Rafael Gonzalez, Javeriana University, Colombia

Clever Ricardo Guareis De Farias, University of São
Paulo, Brazil

Markus Helfert, Dublin City University, Ireland

Philip Huysmans, University of Antwerp, Belgium

Ilian Ilkov, IBM, The Netherlands

Ivan Ivanov, SUNY Empire State College, USA

Dimitris Karagiannis, University of Vienna, Austria

Marite Kirikova, Riga Technical University, Latvia

Samuel Kounev, Karlsruhe Institute of Technology,
Germany

Kecheng Liu, University of Reading, UK

Leszek Maciaszek, Macquarie University, Australia /
University of Economics, Poland

Jelena Marincic, University of Twente, The
Netherlands

Michele Missikoff, Institute for Systems Analysis and
Computer Science, Italy

Dimitris Mitrakos, Aristotle University of
Thessaloniki, Greece

Preslav Nakov, National University of Singapore,
Singapore

Ricardo Neisse, University of Kaiserslautern,
Germany

Bart Nieuwenhuis, University of Twente, The
Netherlands

Selmin Nurcan, University of Paris 1 Pantheon-
Sorbonne, France

Olga Ormandjieva, Concordia University, Canada

Robert Parhonyi, Inter Access, The Netherlands

Marcin Paprzycki, Polish Academy of Sciences,
Poland

Oscar Pastor, Universidad Politécnica de Valencia,
Spain

Erik Proper, Public Research Centre - Henri Tudor,
Luxembourg

Jolita Ralyte, University of Geneva, Switzerland

Gil Regev, EPFL / Itecor, Switzerland

Ella Roubtsova, Open University, The Netherlands

Irina Rychkova, University of Paris 1 Pantheon-
Sorbonne, France

Shazia Sadiq, University of Queensland, Australia

Brahmananda Sapkota, University of Twente, The
Netherlands

Tony Shan, Keane Inc., USA

Kamran Sheikh, IBM, The Netherlands

Valery Sokolov, Yaroslavl State University, Russia

Richard Starmans, Utrecht University, The
Netherlands

PROGRAM COMMITTEE (CONT.)

Cosmin Stoica Spahiu, University of Craiova,
Romania

Coen Suurmond, RBK Group, The Netherlands

Bedir Tekinerdogan, Bilkent University, Turkey

Linda Terlouw, ICRIS B.V., The Netherlands

Yasar Tonta, Hacettepe University, Turkey

Roumiana Tsankova, Technical University of Sofia,
Bulgaria

Marten Van Sinderen, University of Twente, The
Netherlands

Mladen Velez, Technical University of Sofia,
Bulgaria

Kris Ven, University of Antwerp, Belgium

Maria Virvou, University of Piraeus, Greece

Martijn Warnier, Delft University of Technology,
The Netherlands

Shin-Jer Yang, Soochow University, Taiwan

Benjamin Yen, University of Hong Kong, China

Fani Zlatarova, Elizabethtown College, USA

BEST PAPERS SELECTIONS

The authors of around ten selected papers presented at BMSD 2011 will be invited by Springer-Verlag to submit extended and revised versions of their papers for publication in a Springer LNBIP Series book, and the authors of around five selected papers presented at BMSD 2011 will be invited to submit extended and revised versions of their papers for publication in a special issue of the international journal Enterprise Modelling and Information Systems Architectures (EMISA).

FOREWORD

This book contains the proceedings of the First International Symposium on Business Modeling and Software Design (BMSD 2011) which was organized and sponsored by the Interdisciplinary Institute for Collaboration and Research on Enterprise Systems and Technology (IICREST), in cooperation with the Center for Telematics and Information Technology (CTIT), the Institute for Systems and Technologies of Information, Control and Communication (INSTICC), and Technical University of Sofia; the symposium was technically co-sponsored by Sofia Municipality and QlikTech Netherlands B.V.

BMSD 2011 provides a platform to researchers and practitioners interested in business modeling and its relation to software design. The scientific areas of interest to BMSD 2011 are: (i) Business Models and Requirements; (ii) Business Models and Services; (iii) Business Models and Software; (iv) Information Systems Architectures. Each year, a special theme is chosen, for making presentations and discussions more focused. The theme of BMSD 2011 is: **Business Models and Advanced Software Systems**.

Building adequate business models is of huge importance not only for understanding and re-engineering an organization but also for automating (part of) its processes by means of software systems. Not grasping correctly and exhaustively a business system would inevitably lead to consequent software failures. BMSD 2011 has addressed these challenges, by considering a large number of research topics: from more abstract ones, such as essential business models to more technical ones, such as software specification, from more business-oriented ones, such as business process management and coordination, and requirements specification to IT architectures -related topics. We expect the technical program that approaches these topics to appeal to a global audience of engineers, scientists, and business people interested in the BMSD areas. We believe the current proceedings highlight challenging technical problems and present innovative solutions relevant to the mentioned topics.

BMSD 2011 received 58 paper submissions from which 22 papers were selected for publication in the current proceedings. From these, 10 papers were selected for a 30-minutes oral presentation (Full Papers) and 12 papers (including several Invited Papers) were selected for a 20-minutes oral presentation (Short Papers and Special Session Papers). Hence, the full-paper acceptance ratio of 17% shows a high level of quality, which we intend to maintain and reinforce in the following editions of this symposium. All presented papers will soon be available at the SciTePress digital library. Furthermore, the authors of around ten selected papers presented at BMSD 2011 will be invited by Springer-Verlag to submit extended and revised versions of their papers for publication in a Springer LNBIP Series book, and the authors of around five selected papers presented at BMSD 2011 will be invited to submit extended and revised versions of their papers for publication

in a special issue of the international journal Enterprise Modelling and Information Systems Architectures (EMISA).

The high quality of the BMSD 2011 programme is enhanced by four keynote lectures, delivered by distinguished guests who are renowned experts in their fields, including (alphabetically): Mehmet Aksit (University of Twente, The Netherlands), Dimitar Christozov (American University in Bulgaria - Blagoevgrad, Bulgaria), Hermann Maurer (Graz University of Technology, Austria), and Bart Nieuwenhuis (University of Twente, The Netherlands). In addition, the keynote lecturers and other BMSD'11 participants will take part in informal discussions focused on community building and project acquisition. These high points in the symposium program would definitely contribute to positioning BMSD as a high quality event driven by a stable and motivated community.

Building an interesting and successful program for the symposium required the dedicated effort of many people. Firstly, we must thank the authors, whose research and development efforts are recorded here. Secondly, we thank the members of the program committee for their diligence and expert reviewing. Further, we appreciate the willingness of SciTePress to publish the current proceedings, expressing also special gratitude to Vitor Pedrosa for the brilliant work and support in preparing the proceedings. We also wish to include here a word of appreciation for the excellent organization provided by the IICREST team, who have smoothly and efficiently prepared the most appropriate environment for a productive meeting and scientific networking. Last but not least, we thank the keynote lecturers for their invaluable contribution and for taking the time to synthesize and deliver their talks.

We wish you all an inspiring symposium and an unforgettable stay in the beautiful city of Sofia. We hope to meet you again next year in Geneva, for the Second International Symposium on Business Modeling and Software Design (BMSD 2012), details of which will be made available at <http://www.is-bmsd.org>.

Boris Shishkov

IICREST, Bulgaria

CONTENTS

KEYNOTE SPEAKERS

SOFTWARE TECHNOLOGY MATURITY MODELS <i>Mehmet Aksit</i>	3
BUSINESS MODELING AS FOUNDATION IN DEVELOPING DATA MINING TOOLS - Case of Automobile Warranty Data <i>Dimitar Christozov</i>	5
BUSINESS MODELLING FOR SOFTWARE BASED SERVICES <i>Bart Nieuwenhuis</i>	7
WE HAVE SEEN NOTHING YET <i>Hermann Maurer</i>	9

FULL PAPERS

ASPECT-ORIENTATION IN MODELLING: LESSONS LEARNED <i>Ella Roubitsova</i>	13
PERFORMANCE ENGINEERING OF BUSINESS INFORMATION SYSTEMS - Filling the Gap between High-level Business Services and Low-level Performance Models <i>Sammel Kounev</i>	25
LEVERAGING THE PUBLISH-FIND-USE PARADIGM OF SOA - Supporting Enterprise Collaboration Across Organisational Boundaries <i>Brahmananda Sapkota and Marten van Sinderen</i>	34
REGULATION, THE INVISIBLE PART OF THE GOAL ORIENTED REQUIREMENTS ENGINEERING ICEBERG <i>Gil Regev and Alain Wegmann</i>	42
DESIGNING FOR INNOVATION BY APPLYING ORGANIZATIONAL MODULARITY <i>Philip Huysmans</i>	51
ASSESSING THE BUSINESS RISK OF TECHNOLOGY OBSOLESCENCE THROUGH ENTERPRISE MODELLING <i>Cameron Spence and Vaughan Michell</i>	61
ICT USE AS A DETERMINANT FACTOR IN STUDENTS' SATISFACTION AND SKILLS DEVELOPMENT - A New Model for the New Higher Education Context <i>Ángel F. Villarejo-Ramos, Manuel J. Sánchez-Franco and Fernando Criado-García-Legaz</i>	71
USING AHM APPROACH TO POSITION CDDP <i>Mladen Velez, Ognjan Andreev and Tanya Panayotova</i>	80
DEVELOPING THE BUSINESS MODELLING METHOD <i>L. O. Meertens, M. E. Jacob and L. J. M. Nieuwenhuis</i>	88
THE ACCCOUNTING CONCEPT AS KEY CONCEPT FOR BUSINESS MODELLING <i>Coen Suurmond</i>	96

SHORT PAPERS

A FOCUSED APPROACH TO BUSINESS CAPABILITY <i>Vaughan Michell</i>	105
BUSINESS MODELLING FOR GENERATION OF KNOWLEDGE FROM EXPLICIT DATA - Considering Administrative Management Processes <i>Anna Rozeva, Martin Ivanov and Roumiana Tsankova</i>	114
INCORPORATING COLLABORATION IN BUSINESS PROCESSES - How to Merge Business Processes and Collaboration Activities in an Efficient and Agile Way <i>Stoian Boev, Elina Surova, Kaloyan Nikolov and Viktor Zhivkov</i>	122
MMALL - Multilingual Mobile-assisted Language Learning <i>Maria Virvou, Eftimios Alepis and Christos Troussas</i>	129
TOWARDS THE AUTOMATIC ARCHITECTURE DRIVEN SOFTWARE MODERNIZATION <i>Dimitar Birov and Yanka Todorova</i>	136

SPECIAL SESSION ON EMBEDDED SYSTEMS AND BUSINESS MODELS

ANALYSIS OF METHODS AND TECHNICAL TOOLS FOR TRAFFIC CONTROL SYSTEMS <i>Rosen Vitanov, Plamen Balzhiiev and Roumen Arnaudov</i>	145
---	-----

SPECIAL SESSION ON HUMAN-CENTERED INFORMATION SYSTEMS MODELING

WHY DOES SOCIAL CONTEXT MATTER? - Integrating Innovative Technologies with Best Practice Models for Public and Behavioral Health Promotion <i>Luba Botcheva and Shobana Ragbupathy</i>	151
--	-----

SPECIAL SESSION ON INTELLIGENT ALGORITHMS FOR REAL-TIME INFORMATION SYSTEMS SUPPORT

APPLICATION OF MATHEMATICAL TRANSFORM IN DETECTION ALGORITHMS <i>Lyubka A. Doukowska</i>	161
RESOURCE MANAGEMENT FOR ONE CLASS OF WORKFLOW PROCESSES - A Stochastic Petri Net Based Approach <i>Oleg Gorbunov and Valery Sokolov</i>	168

SPECIAL SESSION ON BUSINESS MODELS AND INFORMATION SYSTEMS FOR SUSTAINABLE DEVELOPMENT

BUSINESS MODELS AND INFORMATION SYSTEMS FOR SUSTAINABLE DEVELOPMENT <i>Marten van Sinderen</i>	177
---	-----

INDUSTRIAL SESSION

BUILDING A BUSINESS PROCESS MAP USING BPM TOOLS AND BPMN NOTATION <i>Francesc Xavier Berjano Treserras</i>	185
---	-----

SEPARATING AND QUANTIFYING VALUE AND WASTE TO IMPROVE OPERATIONAL PERFORMANCE IN SOFTWARE DEVELOPMENT <i>Darius Khodavandi</i>	189
--	-----

AUTHOR INDEX	193
--------------	-----

**KEYNOTE
SPEAKERS**

SOFTWARE TECHNOLOGY MATURITY MODELS

Mehmet Aksit

University of Twente, Enschede, The Netherlands

M.Aksit@ewi.utwente.nl

Abstract: From enterprise systems to embedded systems, software is the key enabling force of today's businesses. Although this fact is recognized by the current business and technology managers, the complexities that come along with software, and how to deal with these, are hardly understood. This is mostly because software is "invisible" and the professional skills that are required to deal with complex software are generally unknown to the management. For this reason, software system development is largely considered as a bunch of coding activity plus some nasty process management. The businesses that are willing to apply the maturity models such as CMMI mainly focus on the processes without being conscious about the depth of the required solution techniques. The increased emphasis on architectures is mostly limited to considering the enabling technologies in system realization. On the other hand, the language of the researchers in computer science are unintelligible to the technical managers. We believe that the so called 'software crisis' is partly created due to the above listed problems. To address these challenges, we will first identify the so-called key quality and technology domains. Then we will introduce the concept of technology maturity models. A technology maturity model identifies the advancements an organization may master in due time within a key technology domain through adoption of increasingly more advanced and beneficial state-of-the-art methods, techniques and tools. Finally we will conclude the talk by emphasizing the advantages of adopting the software technology maturity models for creating successful businesses.

BRIEF BIOGRAPHY

Mehmet Aksit holds an M.Sc. degree from the Eindhoven University of Technology and a Ph.D. degree from the University of Twente. Currently, he is working as a full professor at the Department of Computer Science, University of Twente and affiliated with the institute Centre for Telematics and Information Technology. He has served many conferences and symposia. For example, he was the program (co-)chair of ECOOP'97, SACT'00, HQSAD'00, NoD'02 and AOSD2003. He was the tutorial chair of the ECOOP'92 conference and the organizing chair of the AOSD'02 conference. He has been also serving as a program committee member of many international conferences and as a reviewer of several journals. He is the co-founder and has been the co-editor in chief of Transactions on Aspect-Oriented Software Development (published by Springer-Verlag) until March 2007. Currently, he is at the editorial board of this journal. He has organized special journal issues as a co-guest editor on topics such as "Computational Intelligence in Software Engineering", "Auto-adaptable Systems", "Model Driven Architecture". In addition, since 1988, he has been serving as a reviewer of various

European projects. He has given numerous invited presentations and keynote talks. Examples in 2008 are keynote talks in Software Composition conference in Budapest, Aspect-Oriented Modeling workshop in Brussels, Informatics conference in Cesme, Software Quality and Tools Conference in Istanbul, Sysem Integration Conference in Brasilia. He is the co-founder of Aspect-oriented association, where he has served as the steering committee member until March 2008. He is the steering committee member of AITO, which organizes the ECOOP conference series. He is the steering committee member of the Turkish Software Architecture Group, which organizes National conferences on this topic. Since 1990, he has given more than 110 international and in-company courses and conference tutorials mainly in the Netherlands, but also in Canada, Denmark, France, Germany, Hungary, Ireland, Italy, Portugal, Spain, Sweden, Switzerland, Turkey and in the United States. For more than 10 years long, he has received (one of) the highest evaluations for the courses given for the post-academic organization (PAO-Informatica). He has organized special training programs for a number of multi-national companies, where he trained hundreds of software designers and

architects. As a visiting scientist, in 1989 he was at the IBM T. J. Watson Research Laboratory, New York, in 1993 at the University of Tokyo, and in 1994 at the New Jersey Institute of Technology. He has been involved in the design and implementation of many software systems. When he was working for Océ Nederland from 1981 – 1982 and 1983 - 1987, first he worked on image processing and coding techniques to be used in digital copiers. Later he worked on office system software. After moving to the University of Twente in 1987, he has been involved in many practical projects and designed various large-scale software architectures, which some of them are currently being utilized in products. Some of the research tools developed by the chair are now being used in some industrial applications. He has served as a consultant for large organizations such as in 2006 the Dutch Ministry of Traffic where he has evaluated large-scale applications of software systems managing traffic-flow. Also, in 2007 he has served the Dutch Tax office by giving consultancy and training.

BUSINESS MODELING AS FOUNDATION IN DEVELOPING DATA MINING TOOLS

Case of Automobile Warranty Data

Dimitar Christozov

American University - Blagoevgrad, Blagoevgrad, Bulgaria

Dgc@aubg.bg

Abstract: Using business models is essential for successful contemporary management. A model represents understanding about the real life domain and reflects its essential traits. Another characteristic of today's management is availability of huge amount of accessible and searchable data accumulated over executing business activities. Exploring those data to increase domain understanding and to build models, which further can be used to support decision making is the essential task of the emerging area of business intelligence and especially data mining. The process of modeling based on business data analysis to build decision support tools is discussed via sharing the experience acquired in a data mining project. The project aimed at discovering factors influencing the warranty cost in automobile industry. The warranty data analysis has to serve for solving two different tasks: (i) design of warranty policy. Warranty is an important marketing tool, used to share the risk of failures between all customers. This share is included into the product's price. Also, warranty is an important advertising and promotional tool – warranty coverage encourages the customers to purchase the product. In its both purposes, the warranty policy requires careful analysis related to the cost, which influence pricing and overall marketing policy. (ii) increase the reliability of cars. Warranty data contains information about the most common problems leading to failures. Warranty analysis helps to identify priorities and directions for improving the products. Or how to improve the cars and to reduce the warranty cost. An iterative research process of developing and exploring models to facilitate data analysis is presented. The process includes the following phase: data collection; data research; modeling; defining analytical procedures to expose the discovered patterns; development software tools to support use of the analytical procedures. Interpretation of discovered patterns provides the necessary arguments in design of software tools for regular ongoing business analysis to support decision making. The applied methodology represents a good practice in specifying, designing and implementing components of a data mining application.

BRIEF BIOGRAPHY

Dimitar Christozov is a Professor of Computer Science at the American University in Bulgaria, Blagoevgrad 2700, Bulgaria since 1993 and at the University of Library Studies and Information Technologies since 2002. He has more than 30 years of research and education experience in areas as computer science, applied statistics, information systems. His recent interests are in the field of business intelligence and data mining. He graduated Mathematics from Sofia University "St. Kliment Ohridski" in 1979. He completed his PhD thesis "Computer Aided Evaluation of Machine Reliability" in 1986. and DSc thesis "Quantitative measures of the quality of informing" in 2009. In ICTT "Informa" (1986-1993) Dr. Christozov was

involved in establishing the national information network for technology transfer and research in the areas of technologies assessment, integral quality measures and information systems for quality management. In these areas he was recognized as one of the leading experts in Bulgaria. At the American University in Bulgaria, he was the leading person in curriculum development, launching and development of the majors of Computer Science (1993) and Information Systems (2008). At the University of Library Studies and Information Technologies he proposed and implemented the major of Information Brokerage. Professor Christozov has more than 80 publications as separate volume, journal papers and papers in refereed proceedings. He is a founding member of Informing Science Institute and chair of Bulgarian Informing

Science Society; and founding member of the Bulgarian Statistical Society and the Bulgarian Telework Association.

BUSINESS MODELLING FOR SOFTWARE BASED SERVICES

Bart Nieuwenhuis

University of Twente, Enschede, The Netherlands

l.j.m.nieuwenhuis@utwente.nl

Abstract: During the 1970s the business model concept was used for describing IT-related business processes. More recently, the business model concept is used for analysing market structures as well as strategic choices related to positioning of organisations within these market structures. Organisations commercialise new ideas and technologies through their business models. The business model design can be seen as a key decision for new firm entrepreneurs. The research field is still lacking a common and general accepted definition of a business model. Chesbrough and Rosenbloom define a business model as 'a blueprint for how a network of organisations cooperates in creating and capturing value from technological innovation'. Essentially, a business model can be seen as a definition of the manner by which an organisation delivers value to customers, entices them to pay for value and converts those payments to profit. Initially, attention has been paid to empirically defining business model typologies. In recent years, business model research started focusing on exploring business model components and developing descriptive models. Osterwalder and Pigneur use a decomposition consisting of nine components: value proposition, customer segments, client relationships, distribution channels and revenue flows on one hand and key activities, key resources, cost structure, partner network on the other hand. These models can also be used to develop business models for software-based products and services. Software can be part of a tangible product that is being paid for by customers. Due to developments such as Application Service Provisioning (ASP), Software as a Service (SaaS) and more recently Cloud Computing, software is more and more the essential building block of services sold to customers. Due to these developments, a business model design process heading for delivering new experiences to customers is guiding the software development process. The state in which the business modelling field finds itself can be characterized as the pre-scientific chaos (Kuhn): there are several competing schools of thought, and progress is limited because of a lack of cumulative progress. Because of this, there are no clear and unique semantics in the research related to business models. During the last years we have been researching business models and are investigating possibilities to apply well-known engineering principles for this application field. We present a business modelling approach as well as some software business modelling cases.

BRIEF BIOGRAPHY

Bart Nieuwenhuis is part-time professor at the School of Management and Governance at the University of Twente. He is member of the Research Group Information Systems and Change Management (ICMS), holding the chair in QoS of Telematics Systems. He is working as advisor and consultant for his own consultancy firm K4B Innovation. His research focuses on generic service provisioning platforms including Quality of Service mechanisms. Application domains comprise telemedicine as well as billing and payment services. His research interests include service innovation and business modelling. Bart Nieuwenhuis supervises PhD students and publishes scientific articles and conference papers on services provisioning

platforms and middleware technologies for Quality of Service and Context Awareness. Bart Nieuwenhuis is chairman of the innovation-driven research programme Generic Communication, part of R&D programmes funded by the Ministry of Economic Affairs. For K4B Innovation, Bart Nieuwenhuis works as an advisor to The Netherlands ICT Research and Innovation Authority. He is the managing director of Exser, the center of service innovation in The Netherlands, founded in 2008. In this center private companies, academic institutions and governmental organization co-operate in order to realise open innovation initiatives. The centre is sponsored by various large, innovative service companies and governmental organizations in The Netherlands. Before joining the ISCM group, Bart Nieuwenhuis

was part-time full professor at the Architecture and Services of Network Applications (ASNA) group within the Faculty of Electrical Engineering, Mathematics & Computer Science (EEMCS) of the University of Twente. He joined the ASNA group in Twente after a period of five years at the University of Groningen, where he was Tele-Informatics professor at the Computer Science Faculty. Before starting his own company, he worked more than 20 years for KPN Research, the R&D facility of KPN, the telephony and Internet market leader in The Netherlands. He served as manager of R&D departments and Head of Strategy of KPN Research. Bart Nieuwenhuis worked on behalf of KPN for the European Institute for Research and Strategic Studies in Telecommunications (EURESCOM) in Heidelberg and was leader of various international, cooperative projects of European public network operators. Bart Nieuwenhuis holds a PhD in Computer Science and a MSc (cum laude) and BSc in Electrical Engineering, all from the University of Twente.

WE HAVE SEEN NOTHING YET

Hermann Maurer

Graz University of Technology, Graz, Austria

hmaurer@iicm.tu-graz.ac.at

Abstract: I will first present some unusual arguments that will show that we are not in for small changes in the near future, but for massive restructuring of how we live, think and learn. I will then explain in what way the strong convergence of cell phones and PCs is likely to develop. I will show clips of prototypes of new devices that overcome the small screen-size of cell phones and their small keyboards. I will then explain how dramatically this will change society and education: we will have with us a permanent powerful assistant. This brings both great benefits and great dangers. I will also address the issue whether large amounts of information help us or rather brainwash us and influence our decision making in a negative way. I will show at least one example that will surprise the audience. I will then argue that no matter how much one is using modern information technology oneself, one does not react the same way when compared to "digital natives", i.e. people (like our students in university) that have grown up in such new environments. I will report on two surprising recent experiments that seem to prove this point conclusively. Finally, I will show that the statement we always hear that the value of technology is ambivalent (like "a hammer can be used to drive in a nail or to kill a person by hitting the head") is not even close to the truth. Rather, some technologies are inherently good, others inherently bad, and many in between. I will present a number of examples to verify this provocative idea and show how computers and networks are classified if one looks at them that way.

BRIEF BIOGRAPHY

Dr. Hermann Maurer is Professor Emeritus at Graz University of Technology. He started his career at the University of Calgary as Assistant and Associate Professor, was appointed full professor at Karlsruhe just before he turned 30, and has been now Professor and Dean in Computer Science at Graz University of Technology since 1978, with some interruptions, like guest-professorships of more than a year at Denver University, University of Auckland, and shorter visits to Edith Cowan University in Perth, SMU in Dallas, Waterloo, Brasilia and others. Chair of the Informatics Section of Academia Europaea, "The Academy of Europe" since April 2009, and receiver of many national and international distinctions, Professor Maurer is author of over 650 papers and 20 books, founder of a number of companies, supervised some 60 Ph.D. and over 400 M.Sc. students and was leader of numerous multimillion Euro projects. More about him than you ever want to read under <http://www.iicm.tugraz.at/maurer>.

FULL PAPERS

ASPECT-ORIENTATION IN MODELLING: LESSONS LEARNED

Ella Roubtsova

Open University of the Netherlands, the Netherlands

Munich University of Applied Science, Germany

ella.roubtsova@ou.nl

ella.roubtsova@hm.edu

Keywords: Separation of concerns, Behaviour models, Semantics, Composition techniques.

Abstract: Model-driven software engineering community faces the problems related to the growing complexity of system models and their rapid evolution. These problems are similar to the problems of programming. Driven by the ideas of Aspect-Oriented Programming many modelling techniques were revised in attempts to find their ways to deal with model complexity and evolution. This paper presents an analysis of existing semantic groups of modelling approaches, their goals, pros and cons, advantages and restrictions. It is aimed to show the deep reason of the fact that particular composition semantics combined with the elegant idea of aspect weaving leads to heavy verification procedures after any change in the AOM model. This analysis reveals the elements of modelling semantics that simplify the verification procedures in AOM.

1 INTRODUCTION

Model-driven software engineering community faces the problems related to the growing complexity of system models and their rapid evolution. The analysts (Chwif et al., 2000) name many reasons for model complexity starting from the “show off” factor, the “possibility” to execute complex model on powerful computers and ending with the “lack of understanding of the real system” and “unclear modelling objectives”. Although it is universally acknowledged that the simplest model is the most beautiful one (Robinson, 1994), there are inescapable reasons for model complexity. The modelled system covers several interfering and non-localizable problem frames (Jackson, 2005) and the validity and completeness of the model may depend on the coverage of relevant problem frames.

The non-localizable problem frames in models are often called aspects following the terminology of Aspect-Oriented Programming (AOP). The modelling approaches that have the goal to separate aspects in models and then compose them into the complete model or produce the behaviour of the complete model, form the research area called Aspect-Oriented Modelling (AOM).

This paper presents a survey of existing semantic groups of aspect-oriented approaches to identify and

explain their goals, abilities and restrictions. There are many AOM surveys, for example (Reina et al., 2004), (Chitchyan et al., 2005), (Op de beeck et al., 2006), (Reddy et al., 2006) and (Schauerhuber et al., 2007). These surveys mostly analyze the UML-based approaches. The deviations of semantics is not the point of attention. The criteria of comparison are often of practical nature such as the position of the approach within the full life cycle, evaluation of granularity and scalability. The aim of this paper is different. It is aimed to show the deep reason of the fact that particular composition semantics combined with the elegant idea of aspect weaving leads to heavy verification procedures after any change in the AOM model. This analysis reveals the elements of modelling semantics that simplify the verification procedures in AOM.

The reminder of the paper is the following. Section 2 briefly presents the achievements of AOP. Section 3 observes the semantic groups of Aspect-Oriented Modelling approaches on the basis of the used composition techniques and explains the applicability and restrictions, PROS and CONS of those groups of AOM approaches. Section 4 draws conclusions and identifies the semantic elements that provide the necessary flexibility and may result in a breakthrough in the line of heavy design approaches.

2 ACHIEVEMENTS OF AOP

It was recognized by (Cardone and Lin, 2004) that the difficulties with separation of concerns in AOP were caused by the composition support of the mainstream object-oriented programming languages that “restricted to single inheritance and subtype polymorphism”. For example, there are three possible ways to organize two features *A* and *B*. into classes: (1) put them in the same class, (2) make *A* a subclass of *B*, (3) make *B* a subclass of *A*. The first two choices force *B* to be included whenever *A* is included. By forcing to choose a single fixed hierarchy orthogonal composition of features is restricted. “As the number of features grows, this problem becomes more severe, because the number of possible feature combinations grows rapidly, but the number of feature combinations that can be practically supported does not”.

In order to solve this problem, Aspect-Oriented Programming (AOP) invented a modular unit called *aspect* to implement a crosscutting concern. This unit contains an *advice* in form of a code presenting a concern and *pointcut designators* being the instructions on where, when and how to invoke the advice. The well defined places in the structure of a program where an advice should be attached were named *join points*. The join point model, using method calls as join points, were generally accepted (Filman et al., 2004) and implemented as extensions of programming languages. Such extensions gave a new task for compilers to produce the code with hard woven aspects. Another branch of AOP developed middleware able to fulfil run-time aspect composition without producing the code of the complete program.

The base programs used for weaving aspects are designed to be oblivious (Filman and Friedman, 2004) to aspects. The advantage of this feature is that “Concerns are separated not only in the structure of the program but also in heads of their creators”. The disadvantage is that the systems “melded from separate minds” can only function if the properties of aspects are preserved in the composed system. This preservation has not been achieved in AOP, although this is an active research area (Kiczales and Mezini, 2005). Current AOP techniques allow producing spectative, regulative and invasive categories of aspect (Katz, 2006). Spectative aspect can change the values of variables local to the aspect, but can’t change the value of any variable of other aspects and the flow of method calls. Regulative aspects affect the flow of control by restricting operations, delaying, or preventing the continuation of a computation. Invasive aspects change the values of variables in the underlying program. In case of invasive aspects no

guaranty can be given about preserving of dynamic properties of the underlying program.

3 AOM AND MODELLING SEMANTICS

The complexity problems caused by the non-localizable overlapping concerns was identified in AOP and AOM was trying to map the findings of AOP on models and understand the borders of application of the AOP solutions. The AOM started with modelling and analysis of AOP programs. The next step was to evaluate the ability of existing modelling techniques to capture aspects and compose them. This research activity resulted in two streams of approaches. The dominant stream localizes concerns at the modelling level but do not simplify model analysis. There are approaches that also localize reasoning on concern models about the behaviour of the whole model.

3.1 Revision of Sequence Diagrams

The sequence diagrams were the first group of modelling notations revised for the AOM purposes.

A set of sequence diagrams with conventional semantics is aimed to present only a part of possible sequences of system behaviour. Sequence diagrams are mostly used to illustrate behaviour of programs and therefore use operation calls and returns as elements of behaviour. The composition techniques used in sequence diagrams are restricted to sequential composition, alternatives, cycles and insertion of sequences.

For the AOM needs the conventional sequence diagrams were extended with Join Point Designation Diagrams (JPDD) (Stein et al., 2005). An advice behaviour and the join points of this advice are specified with different sequence diagrams. JPDDs are modelling means to graphically represent join point queries on sequence diagrams. Figure 1 borrowed from (Stein and Hanenberg, 2006) shows the sequence diagrams presenting a join point and an advice-sequence of an aspect “Access control policy” that prescribes that the owner management data are used to ensure that particular units (i.e. tasks, contacts, or appointments) are modified or viewed by their proper owner. The advice presented with the diagram *afterAdvice_DenyAccess* is executed when one of the join points specified with wildcards *move|setProgress|setPriority* is called. The information about the current user is obtained *getCurrentUser()* and compared with the information about the owner of the item. If the current

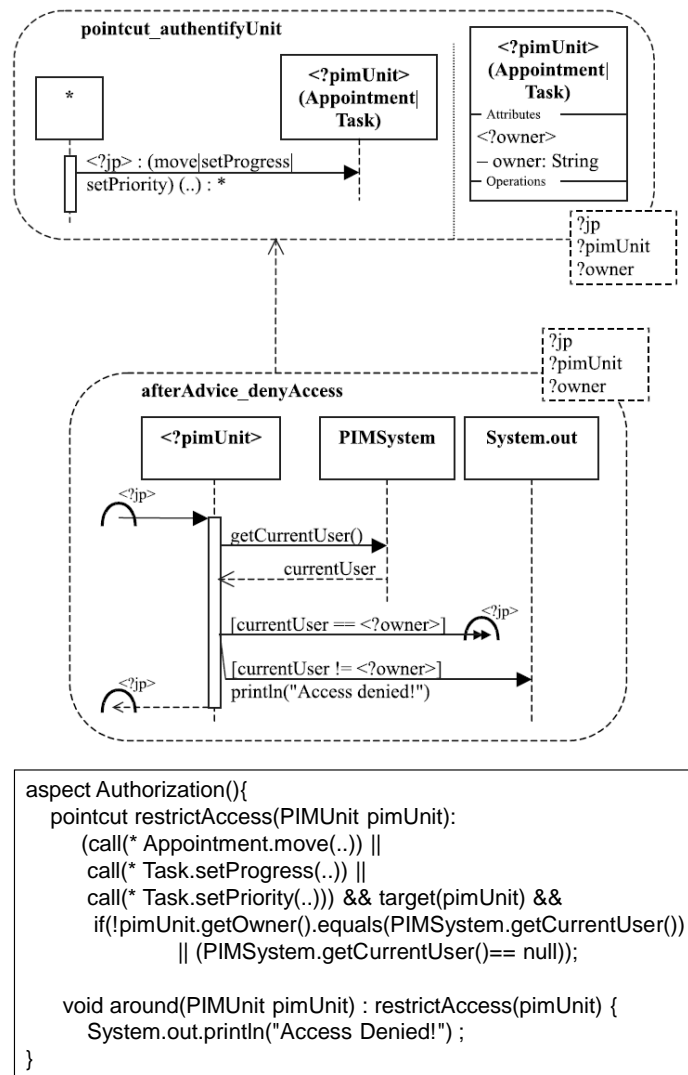


Figure 1: JPDD model.

user is not the owner `currentUser! – <?owner >` then the access is denied.

The sequence diagrams with JPDDs imitate aspect-oriented programs. The composition of sequences presenting the base behaviour, advice and join points is implemented as a compiler task. As the sequence diagrams are close to the notation of aspect-oriented languages, the code is generated from them. As sequence diagrams present partial behaviour of the system, then result of their composition is a partial code.

- PROS: The AOM version of sequence diagrams with JPDDs supports more compact presentation of aspects than the conventional sequence diagrams. Models of join points become reusable artefacts. Ability

to present aspects with sequence diagrams eases evolution of sequence diagrams.

- CONS: Understanding of models and reasoning on models presented with AOM sequence diagrams with JPDDs become more difficult than understanding and reasoning with plain sequences diagrams. It is mostly because the weaving of aspects should be done in human heads. However, with tool support, the aspects are woven into base sequences and presented to the designers for understanding and linear reasoning.

Although sequence diagrams are found to be very intuitive means for behaviour modelling, their main issues still remain. Namely, even in combination with class diagrams sequence diagrams are not able to present complete behaviour of complex business systems having infinite set of traces. Moreover, sequence

diagrams with JPDDs use the same composition techniques as the AOP programs and cannot prevent specification of invasive aspects, i.e. they cannot prevent changes of attribute values of one object or aspect by another aspect.

3.2 Reuse and Revision of Workflows

Activity and workflow based approaches are aimed to specify complete system behaviour that can be analyzed and verified against required properties. The workflows are often used in AOM approaches as integration means to combine specified aspects. There are also AOM approaches that define fragments of workflows and compose workflow fragments into the complete workflow.

Workflows for Integration. The Theme (Clarke and Baniassad, 2005) approach is a good example of the first type of workflow use. At the level of requirements a designer identifies actions as verbs in textual requirements. Each action potentially becomes a *theme* depicted as a rhomb (Figure 2). The actions have relations via concepts depicted as parallelograms with rounded corners. Actions found in different themes become potential aspects. An action view of each theme is a graph that contains actions and entities. The size of this view grows with the number of entities and actions in the model. The scalability is achieved by the separation of major and minor actions. “Major actions become themes, while minor actions are slotted to become methods within a theme” (Baniassad and Clarke, 2004). This means that the major actions in the Theme approach are not elementary, they are activities. The elements of behaviour are method calls and returns.

The action view is analyzed to produce the Theme/UML specification of actions-aspects as combination of class and sequence diagrams. The Design level Theme/UML supports modelling of aspects in UML. Theme/Doc views are mapped onto the Theme/UML model allowing traceability of requirements. If a Theme/action is reused in the model, the join points are specified in the Theme/UML view. Figure 2 shows that Themes *register* and *logged* identified in the workflow are specified as combinations of class and sequence diagrams called $\langle theme \rangle Register$ and $\langle theme \rangle Logger$. In this case the methods of themes *register*, *unregister* and *give* should be specified as join points of the $\langle theme \rangle Logger$.

The action view “drives composition semantics for design in Theme/UML”. The composition techniques are concatenation and hierarchy of *Major* and *Minor* actions. From the Theme/UML and

Theme/Doc specifications the corresponding computation tree can be generated and may be compared with the workflow built at the level of requirements. For example, such approach as GrACE (Graph-based Adaptation, Configuration and Evolution) (Ciraci et al., 2009) uses different graph transformation techniques in order to compose behaviours presented by class, sequence diagrams and the action view in the corresponding computation tree of the designed system and apply model checking and verification algorithms to analyse the properties of the system.

- PROS: In addition to conventional activity diagrams Workflow diagrams in Theme/UML separate aspects and minor actions to increase the abstraction level and decrease the size of workflows.

- CONS: The used composition techniques, namely concatenation and hierarchy of activities, make it impossible to add new aspect models at the requirements level without verification of the complete model. The evolution of the model is handled by re-generating the views for any new set of requirements. Reasoning on models does not become easier as the minor actions are specified by sequence diagrams, but the composition of them as a workflow. Reasoning demands relatively heavy procedures of building the computation tree, reachability analysis and verification of dynamic properties.

Composed workflows. Another direction in the use of workflow semantics for aspect-oriented modelling is specifying fragments of a workflow and then composing them into the complete workflow. The composition techniques used in workflows are restricted to concatenation and hierarchical inclusion of fragments.

The ADORE approach (Mosser et al., 2010) (Activity moDel supOrting oRchestration Evolution) re-examines the Business Process Execution Language (BPEL) to enable separation of concerns.

An orchestration of services is defined in ADORE as a partially ordered set of activities. The types of activities that can be defined in ADORE include service invocation (denoted by *invoke-inv*), variable assignment (*assign-a*), fault reporting (*throw-t*), message receiving (*receive-rcv*), response sending (*reply-rp*), and the null activity (*nop*), which is used for synchronization purpose (Figure 3).

Each process (fragment) corresponds to a specific concern and uses a partial point of view on its target. A fragment contains special activities, called predecessors *P*, successors *S* and a hook (assimilated as a proceed in AspectJ) that represents a part of a trace where the fragment is connected into an existing or-

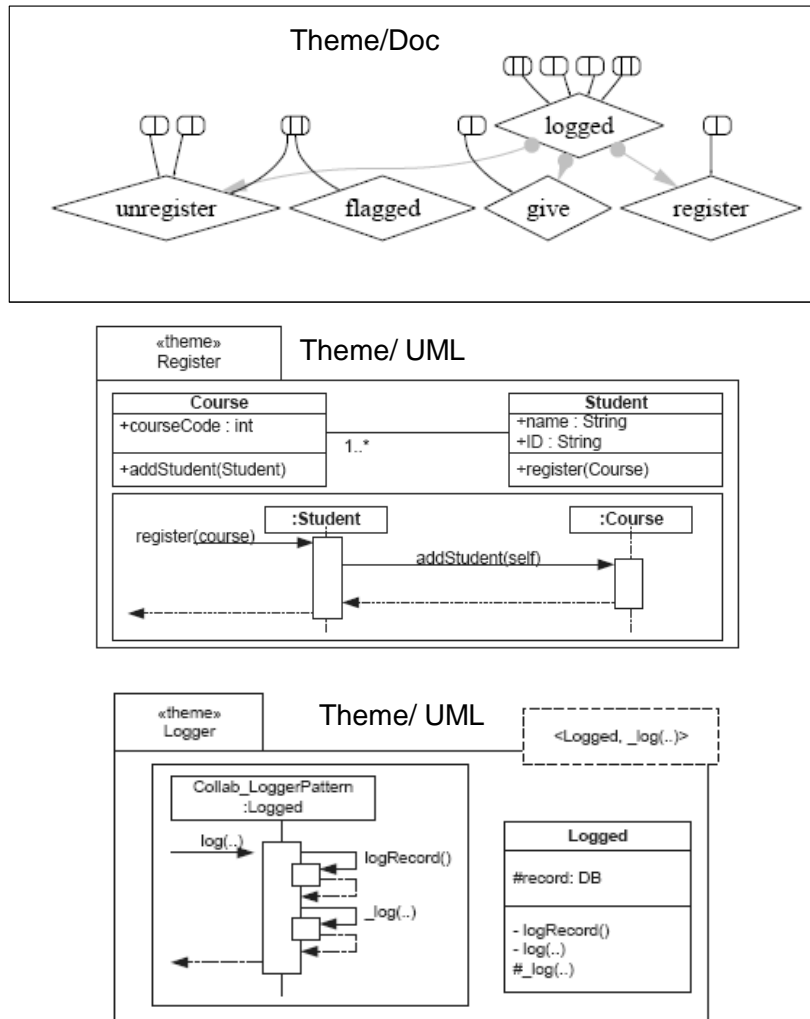


Figure 2: Theme model.

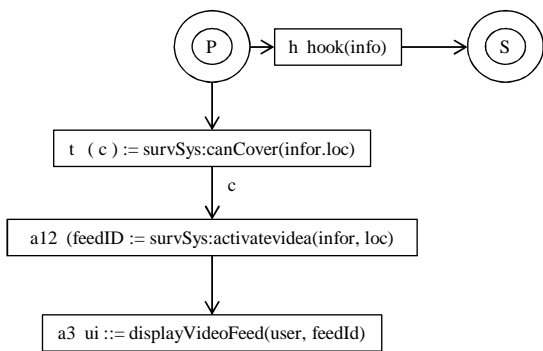


Figure 3: ADORE model. Fragment requestVideo<user>.

chestration. The hook predecessors (*P*) are the immediate predecessors of the first activity in the target block, and the hook successors (*S*) are the immediate

successors of the last activity in the block.

Fig. 3 describes the following behaviour. After the execution of the hook predecessors (*P*), perform the activity block referred to by (*h*) and then continue with the hook successors (*S*). The vertical branch of the fragment represents the additional behaviour. In parallel to the behaviour described by the hook (*h*) the system first determines if the surveillance system can cover the crisis area (*t*). If this functionality is available for this location, the process requests a video feed (single ADORE activity *a12*) and then broadcasts it to the Coordinator interface (*a3*).

The binding instructions are specified in a separate file. For example, the fragment from Fig. 3 can be bound on the block of activities

a3 : cms : buildchecklist(i)

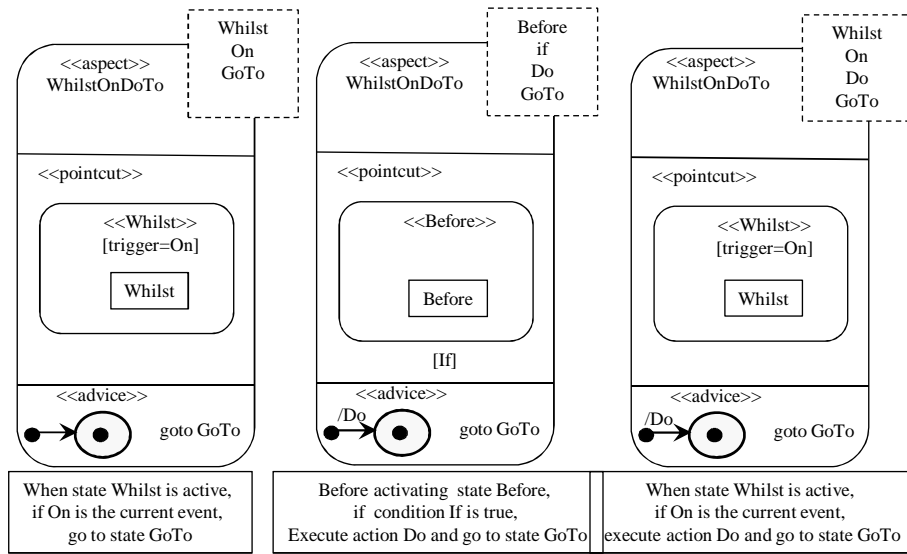


Figure 4: HiLA patterns.

and

$$a4 : (ci)ui := promptchecklist(coord, id, cc1)$$

of the base orchestration. The predecessor of $a3$ is assigned to P , the successor of $a4$ is assigned to S and the *hook* is assigned to $\{a3, a4\}$. The complete orchestration is generated from fragments according to the binding instructions.

As with such an approach there is no guarantee that any particular local properties of aspects are propagated to the complete orchestration, the side effects of separating of concerns and composition in ADORE are formulated as rules. The violating of rules not always signals about a mistake, this may indicate a "bad-smell", like, for example, the non-determinism caused by two conditions evaluated to false at the same time. The "bad-smells" are analyzed by the designer of the orchestration.

- PROS: ADORE extends BPEL with means for modelling of aspects. Models become closer to requirements.
- CONS: The composition techniques in workflows cause the need of re-generating and "bad-smells" re-analyzing of the complete orchestration after any change or adding any new fragment. Reasoning on models is not localized and demands reachability analysis and theorem proving.

3.3 Revision of State Machines

A UML Behaviour State Machine (BSM) (OMG, 2003) usually presents behaviour of one classifier.

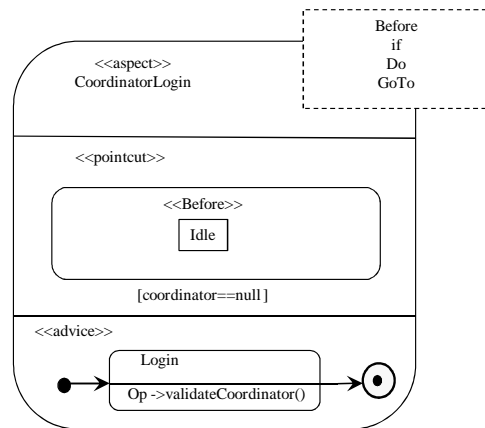


Figure 5: HiLA model of the Login Concern.

The execution semantics of BSM (McNeile and Roubtsova, 2009) is complex, involves a queue or a stack of active events and rules of their handling or keeping in the queue until the state is appropriate for their handling.

There are several approaches revising Behaviour State Machines (BSM) trying to extend behaviour specified by BSM for one classifier. The approach by (Mahoney et al., 2004) exploits the *AND-composition* of several independent (orthogonal) statecharts defined by D.Harel (Harel and Gery, 1997). "The key feature of orthogonal statecharts is that events from every composed statechart are broadcast to all others. Therefore an event can cause transitions in two or more orthogonal statecharts simultaneously" (Mahoney et al., 2004).

The ideas proposed by Mahoney et al. were further developed in the approach called High-Level Aspects (HiLA) (Hözl et al., 2010). HiLA modifies the semantics of BSM allowing classifiers to apply additional or alternative behaviour. Aspects extend the behaviour specified for classifiers.

The basic static structure usually contains one or more classes. Each base state machine is attached to one of these classes and specifies its behaviour. HiLA offers two kinds of aspects to modify such a model. Static aspects directly specify a model transformation of the basic static structure of the model. Dynamic (high-level) aspects only apply to state machines and specify additional or alternative behaviour to be executed at certain "appropriate" points in time in the base machines execution.

HiLA introduces patterns for specification of dynamic aspects (Figure 4). A pattern *Whilst* always has an annotation *Trigger = e*. "Conceptually it selects the compound transition from *State** to *State*** with trigger *e*, but if this transition does not exist, it is created." This means that an aspect is added while an action is in the stack of active actions.

A pattern labelled with *Before* selects each transition *T* entering any state contained in *State**, but only in active state configurations where after taking *T* all states in *State***, will become active" (Hözl et al., 2010). An aspect may extend the behaviour of a class or introduce a new class to the specification.

Figure 5 shows the aspect validating if the *Coordinator* is logged in. The aspect is taken from the model of a crisis management system (Holzl et al., 2010). Before the *Idle* state of system becomes active, the system checks whether the coordinator is not logged in (*coordinator == null*). If this is the case, the *Login* sub-state machine is triggered and operation *validateCoordinator()* becomes possible. The aspect is regulative, it changes the flow of control in the base state machine.

- PROS: HiLA patterns support specification of aspects in BSM.

- CONS: Evolution of models and reasoning on them do not become easier. Active objects represented with classifiers with aspects and new aspect classes execute asynchronously resulting in non-deterministic system behaviour. HiLA uses formal methods of model validation. The application of aspects to BSM results in another UML state machine which is analyzed using the model checking component of Hugo/RTmodel checking tools. Hugo/RT translates the state machine and the assertions into the input language of a back-end model checker SPIN. SPIN then verifies the given properties presented in Linear Temporal Logic.

3.4 Revision of Design by Contract

The search of composition techniques suitable for AOM led to the new look at the set-theory composition. Such approaches as Package Merge (Ammour and Desfray, 2006) and Visual Contract Language (VCL) (Amálio and Kelsen, 2010) explore the declarative way of aspect specification based on composition of sets.

VCL revises the Design By Contract (DbC) principle. An element of behaviour is an operation. According to the DbC (Meyer, 1991) the pre-conditions and post-conditions have to be defined for any specified behaviour.

- If the precondition is satisfied then the result of the corresponding behaviour is defined. It satisfies the post-condition.

- If a precondition is violated, the effect of the corresponding behaviour becomes undefined.

This means that only the non-interactive part of behaviour that takes place, when the preconditions are not violated, can be composed and analyzed. VCL inherits this property of DbC.

A VCL model is organized around packages, which are reusable units encapsulating structure and behaviour. Packages represent either a traditional module or an aspect. VCL's package composition mechanisms allow larger package to be built from smaller ones. A package is denoted by a cloud. In Figure 6 package *Authentication Ops* extends package *Authentication*. A package encapsulates state and behaviour in form of attributes and operations. Package *Authentication Ops* contains operations *Login* and *Logout*.

Ordinary packages define global state. Abstract packages do not define global state, they act as containers for state structures and their local behaviour to be used in other packages.

VCL behavioural diagrams (BDs) identify the operations of a package. There are two types of operations: update and observe (or query). Update operations perform changes of state in the system; they involve a pair of states: before-state (described by pre-condition) and an after-state (described by post-condition). They are defined in VCL contract diagrams. Figure 6 shows examples of VCL contract diagrams for *LogIn* and *LogOut* operations that belong to the package *Authentication Ops*.

Aspects are composed using join extensions, which is illustrated in Figure 7. In a join extension, there is a contract that describes the joining behaviour of an aspect (a join contract) that is composed with a group of operations placed in a join-box. All operations of package *CrisisWithJI* are conjoined with

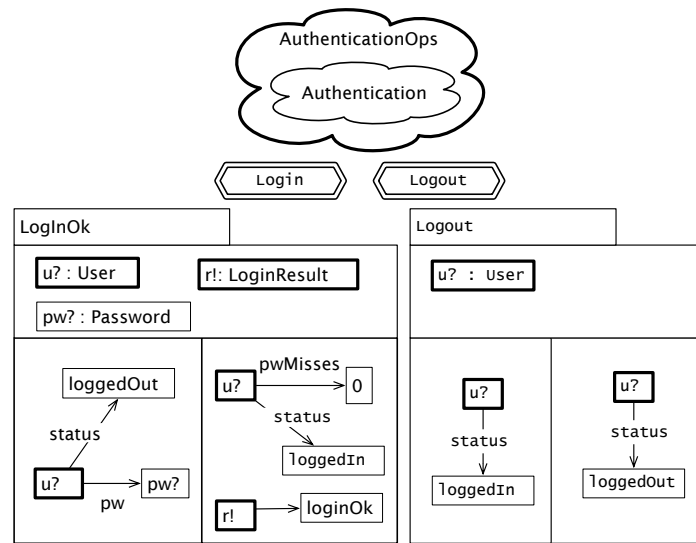


Figure 6: Packages in VCL.

join contracts *LoggingOp*, *SessionMgmtOp* and *AuthorisationOp*. Join contract *AuthorisationOp* specifies the extra behaviour of the *Authorisation* concern by adding an extra pre-condition to all operations of package *CrisisWithJI*; this specifies that the users executing operations of package *CrisisWithJI* must be logged-in and have the required permissions to execute that task.

VCL is designed with a formal Z semantics and so it has the potential for verification and global reasoning using theorem proving.

- PROS: The VCL extends the DbC approach by visual means to capture aspects. Modelling of packages is simple. It demands only knowledge of the set theory and predicate logic.

- CONS: The use of the DbC does not allow localizing of reasoning on aspects about behaviour of the whole model. Verification of the complete model is needed after every modification in the number and the content of aspects.

3.5 Revision of Mixins

Mixins are applied in the Protocol Modelling approach defined by McNeile and Simons (McNeile and Simons, 2003). A protocol model of a system is a composition of protocol machines. Protocol machines are partial descriptions of class behaviours. The composition operator for protocol machines is a variant of the parallel composition operator defined by Hoare (Hoare, 1985) in his process algebra Communication of Sequential Processes (CSP). This operator was extended by McNeile and Simons (McNeile

and Simons, 2003) for machines and events with data.

A protocol machine has its own alphabet of recognized events and a local state. An event type is specified as a data structure (Figure 8). An event instance contains values of the attributes.

The local state of a machine is presented as a set of own attributes and attributes derived from the state of other machines. A machine can read the state of other machines but cannot alter it. The derived states should not be topologically connected. They are already spectative aspects inside of protocol machines.

Moreover, a machine has a set of transitions being triples (*state, event, state*). Events are presented to the model by the environment. Being in a suitable state, a protocol machine accepts the presented event, otherwise it refuses the event. A protocol model handles one event at a time and reaches a well defined quiescent state between each event. An element of behaviour is an accepted event. The behaviour of a protocol model is a set of sequences of accepted events.

Protocol machines use the CSP parallel composition algorithm to form more complex protocol machines. This is the description of the CSP parallel composition algorithm:

- If all machines of the protocol model having an event in their alphabet accept the event, the protocol model accepts it.

- If at least one of protocol machines having this event in its alphabet refuses the event, the composition of machines refuses it.

So, the semantics of refusal which is absent in other modelling semantics (BSM, workflows, sequence diagrams and contracts) is the key to synchronization and CSP parallel composition.

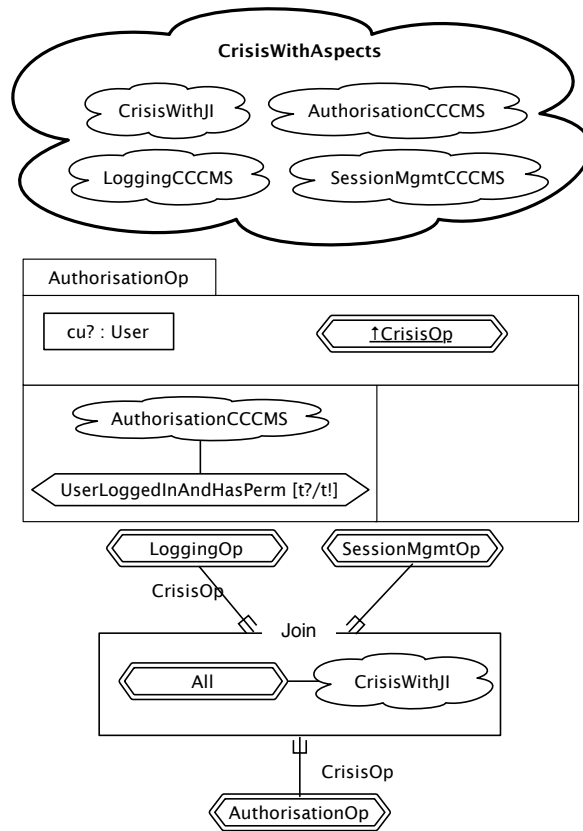


Figure 7: Aspect Composition in VCL.

It is recognized in (McNeile and Roubtsova, 2008) that protocol machines are natural abstractions to specify aspects. For example, a protocol model of a bank *Account* and a *Customer* is shown in Figure 8. Behaviours of both objects include and equally compose behaviours of aspects *Freezing* and *Freeze Control*. These aspects model the possibility to freeze the behaviour of an account or a customer. The INCLUDE-relation shown as a half-dashed triangle gives to Protocol Models the expressiveness of multiple inheritance. Behaviours of aspects are instantiated with instantiation of objects, each of which has its object identifier.

All behaviours are equally composed on the basis of the CSP parallel composition as shown in Figure 9. A new registered *Customer* and a new open *Account* are *not-frozen*. A bank security service may freeze both an *Account* and(or) a *Customer* submitting events *Freeze.Customer* and(or) *Freeze.Account*. Then such an *Account* or *Customer* transits to state *Frozen*. A *Frozen* behaviour cannot accept any event. Its state is derived from the instances of the behaviour *Freezing* included into the objects *Account* and *Customer* correspondingly.

A join point in protocol model is a set of events or

states that can be seen identical in the context a particular protocol machine.

For example, the join point *Generic Operate* matches events *Deposit*, *Withdraw*, *Transfer*, *Leave*, and *Close* and any of these events can be accepted only in state *Freeze Not Active* which is derived from state *not Frozen*.

The proof presented in (McNeile and Roubtsova, 2008) shows that the CSP parallel composition of Protocol Machines guarantees preservation of ordering of traces of aspects in the whole specification. This property is called *local reasoning* or *observational consistency* (Ebert and Engels., 1994). It simplifies reasoning on models as the traces can become longer or shorter but the order of events will be never changed. For example, trace *Open*, *Withdraw*, *Freeze Account* is a trace of the composition of *Freezing* and *Account*, but the sub-trace *Open*, *Withdraw* is a valid trace of *Account*. Small and deterministic protocol machines are verified or tested by direct execution. Then the syntactic checks of specifications of join point are sufficient for verification of the whole model.

- PROS: Protocol machines with generic events and

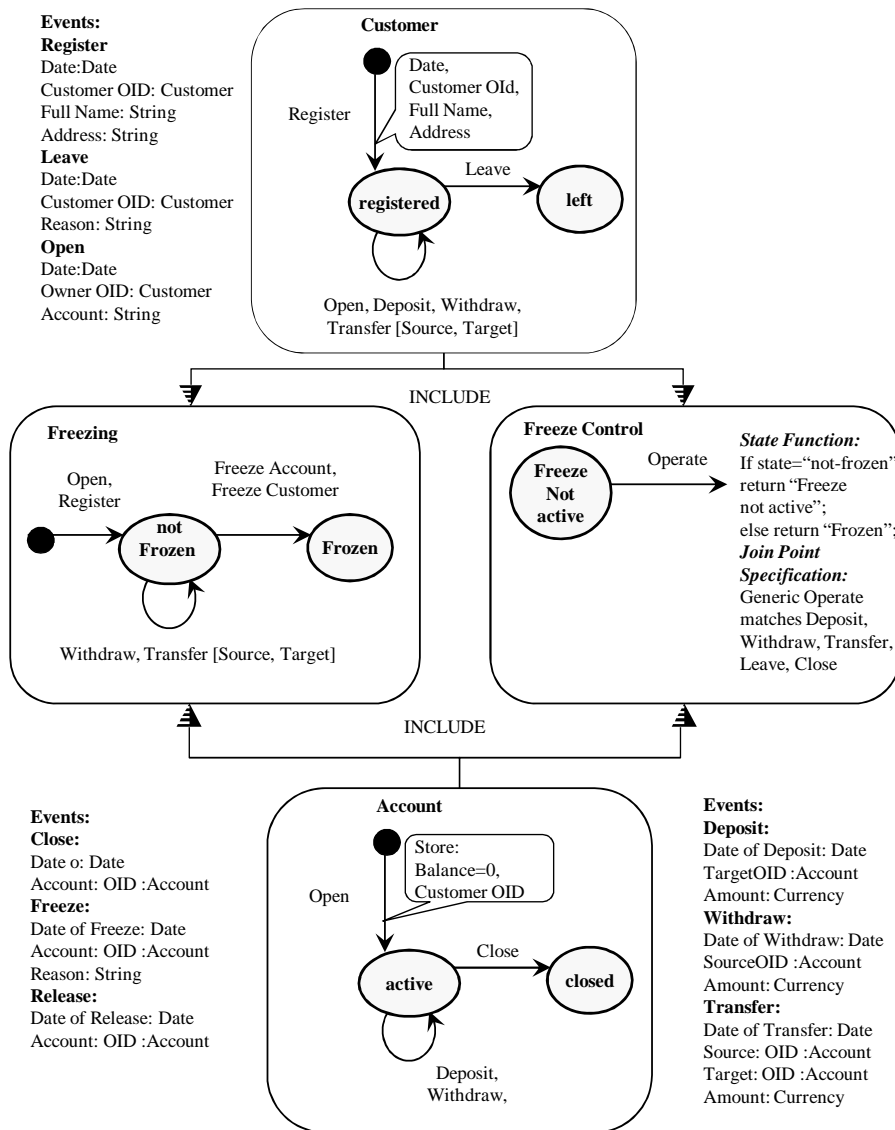


Figure 8: Protocol Model.

states and the synchronous composition are natural means for modelling of aspects. The CSP parallel composition provides flexible support for model evolution. Protocol machines are not hierarchical, they are equally composed. The composition operator guarantees preservation of behaviour and properties of aspects in the behaviour of the whole system.

- CONS: The understanding of a protocol model is not very intuitive. It takes some time to begin freely apply the CSP parallel composition. However, the tool Modelscope¹ supports execution and understanding of protocol models.

¹<http://www.metamaxim.com/pages/download.htm>

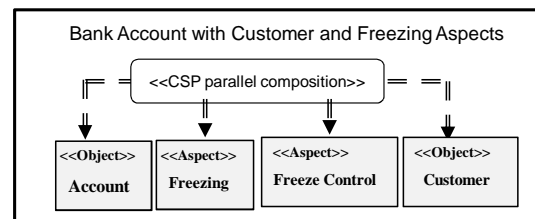


Figure 9: CSP composition of Protocol Machines.

4 CONCLUSIONS

Models should have higher abstraction level than programs and may more freely experiment with compo-

sition techniques unavailable in the main stream modelling languages. However, the composition techniques of many modelling techniques nowadays does not deviate from composition forms available in dominant programming languages. These languages are designed to cover asynchronous communication and have restricted means to deal with shared data. That is why it is not a surprise that the extension of models with the aspect building constructions causes the same problems as in aspect-oriented programs. Namely, such AOM models do not prevent specification of invasive aspects, do not guarantee local reasoning and have to be globally verified after every modification. Using AOM approaches with the same composition techniques as in programs we just move the complexity from models to analysis techniques and tools and rely heavily on results of this analysis. As the models are subject of frequent changes, after any new change the analysis should be repeated. This dependence on analysis and verification after any step of modelling slows down the modelling process, restricts the number of demonstrations of the model to the client and makes the modelling process inflexible.

However, the understanding of the nature of aspects can simplify the analysis. Aspect-abstractions exist at the same level as objects and they are not distributed from object. Therefore they need synchronization with objects and access to shared data. The lessons learned from the analysis of aspect-oriented modelling approaches trying to use asynchronous composition for aspects shows that approaches confuse aspects and other abstraction sorts. Each type of abstraction should be used on its own place. Aspects and objects present problem frames, share data and communicate synchronously. The distributed components or services built from aspects and objects communicate asynchronously (Roubtsova and McNeile., 2009). Such separation will simplify the AOM analysis.

The mixin-based group of AOM approaches offers to AOM a synchronous and agile way of modelling.

- The semantic restriction when one machine cannot alter the state of other machines makes it impossible to specify invasive aspects.
- Using sets of events and states as join points is more abstract and technology independent than using operation calls and returns.
- Derived states provide unlimited weaving abstractions.
- Determinism of machines is assured by the concept of quiescent states.
- The CSP parallel composition technique guarantees the property of local reasoning on aspects

about the behaviour of the whole system. The local properties of mixins survive their composition and are preserved in the whole behaviour of the system.

As a consequence of local reasoning, analysis and execution of mixin-based models does not require generation of the complete model or a computation tree after adding/removing/changing of aspects.

The semantics of event refusal splits the global computation tree into sets of traces of woven protocol machines. The traces of the whole model are generated on the fly by a tool or a middleware implementing CSP parallel composition algorithm.

- All these features reduce the analysis of mixin-based models to verification or testing of separate machines and syntactic checks of specification of join points.

As the proposed semantics is notation independent and it can be applied in many approaches, the next step is to apply it and investigate the implementation of systems in mixin-based and object-oriented languages. Most probably that offering of a CSP parallel composition middleware service will result in a new group of aspect-oriented languages.

ACKNOWLEDGEMENTS

The author thanks all participants of the workshops ABMB'05, ABMB'06, AOM'08, BM-MDA'09, AOM'10 in Bellairs, and BM-FA'10 for fruitful discussions of different modelling approaches.

REFERENCES

- Amálio, N. and Kelsen, P. (2010). VCL, a Visual Language for Modelling Software Systems Formally. In *Diagrams*, pages 282–284.
- Ammour, S. and Desfray, P. (2006). A Concern-based Technique for Architecture Modelling Using the UML Package Merge. *M.Aksit and E.Roubtsova eds. Proceedings of ABMB2005, ENTCS*, 163(1):7–18.
- Baniassad, E. and Clarke, S. (2004). Theme: An Approach for Aspect-Oriented Analysis and Design. *ICSE 2004. Proceedings. 26th International Conference on Software Engineering. IEEE*, pages 158–167.
- Cardone, R. and Lin, C. (2004). Using Mixin Technology to Improve Modularity. *Aspect-Oriented Software Development. R.Filman, T.Elrad, S.Clarke, M.Akşit eds. Addison-Wesley*, pages 219–241.
- Chitchyan, R., Rashid, A., Sawyer, P., Garcia, A., Alarcon, M., Bakker, J., Tekinerdogan, B., Clarke, S., and

- Jackson, A. (2005). Survey of Aspect-Oriented Analysis and Design Approaches. *Technical Report D11 AOSD-Europe-ULANC-9, AOSD-Europe*.
- Chwif, L., Barretto, M. R. P., and Paul, R. (2000). On simulation model complexity. In *WSC '00: Proceedings of the 32nd conference on Winter simulation*. Society for Computer Simulation International.
- Ciraci, S., Havinga, W. K., Aksit, M., Bockisch, C. M., and van den Broek, P. M. (2009). A Graph-Based Aspect Interference Detection Approach for UML-Based Aspect-Oriented Models. Technical Report TR-CTIT-09-39, Enschede.
- Clarke, S. and Baniassad, E. (2005). *Aspect-Oriented Analysis and Design: The Theme Approach*. Addison Wesley.
- Ebert, J. and Engels., G. (1994). Observable or invocable behaviour-you have to choose. *Technical report. Universitt Koblenz, Koblenz, Germany*.
- Filman, R., Elrad, T., Clarke, S., and Akşit, M. (2004). *Aspect-Oriented Software Development*. Addison-Wesley.
- Filman, R. and Friedman, D. (2004). Aspect-Orineted Programming is Quantification and Obliviosness. *Aspect-Oriented Software Development*. Addison-Wesley, pages 21–35.
- Harel, D. and Gery, E. (1997). Executable Object Modelling with Statecharts. *IEEE Computer*, 30(7), pages 31–42.
- Hoare, C. (1985). *Communicating Sequential Processes*. Prentice-Hall International.
- Hözl, M. M., Knapp, A., and Zhang, G. (2010). Modeling the Car Crash Crisis Management System Using HiLA. *T. Aspect-Oriented Software Development*, 7:234–271.
- Holz11, M., Knapp, A., and Zhang, G. (2010). Modeling the car crash crisis management system using HiLA. *S. Katz et al. (Eds.): Transactions on AOSD VII, LNCS 6210*, pages 234–271.
- Jackson, M. (2005). *Problem frames and software engineering*, volume 47. Butterworth-Heinemann.
- Katz, S. (2006). Aspect Categories and Classes of Temporal Properties. *Transactions on Aspect-Oriented Software Development, LNCS 3880*, pages 106–134.
- Kiczales, G. and Mezini, M. (2005). Aspect-Oriented Programming and Modular Reasoning. *Proc. of the International Conference on Software Engineering*, pages 49–58.
- Mahoney, M., Bader, A., Elrad, T., and Aldawud, O. (2004). Using Aspects to Abstract and Modularize Statecharts. *In the 5th Aspect-Oriented Modeling Workshop In Conjunction with UML 2004*.
- McNeile, A. and Roubtsova, E. (2008). CSP parallel composition of aspect models. *AOM'08: Proceedings of the 2008 AOSD Workshop on Aspect-Oriented Modeling*, pages 13–18.
- McNeile, A. and Roubtsova, E. (2009). Composition semantics for executable and evolvable behavioral modeling in MDA. In *BM-MDA '09: Proceedings of the 1st Workshop on Behaviour Modelling in Model-Driven Architecture*, pages 1–8, New York, NY, USA. ACM.
- McNeile, A. and Simons, N. (2003). State Machines as Mixins. *Journal of Object Technology*, 2(6):85–101.
- Meyer, B. (1991). Design by contract. *Advances in Object-Oriented Software Engineering*, eds. D. Mandrioli and B. Meyer, pages 1–50.
- Mosser, S., Blay-Fornarino, M., and France, R. (2010). Workflow Design Using Fragment Composition - Crisis Management System Design through ADORE. *T. Aspect-Oriented Software Development*, 7:200–233.
- OMG (2003). *Unified Modeling Language: Superstructure version 2.1.1 formal/2007-02-03*.
- Op de Beeck, S., Truyen, E., Boucke, N., Sanen, F., By-nens, M., and Joosen, W. (2006). A study of aspect-oriented design approaches. *Technical Report CW435, Department of Computer Science, Katholieke Universiteit Leuven*.
- Reddy, R., Ghosh, S., France, R., and Straw, B. (2006). Directives for composing aspect-oriented design class models. *Transactions on AOSD, LNCS 3880*, pages 75–105.
- Reina, A. M., Torres, J., and Toro, M. (2004). Separating Concerns by Means of UML-profiles and Metamodels in PIMs. In *Proc. of the 5th Aspect-Oriented Modeling Workshop (UML04), Lisbon, Portugal*.
- Robinson, S. (1994). *Successful simulation: A practical approach to simulation projects*. McGraw-Hill, Maidenhead, UK.
- Roubtsova, E. and McNeile., A. (2009). Abstractions, composition and reasoning. *Proceedings of the 13th Intl Workshop on Aspect-Oriented Modeling, Charlottesville, Virginia, USA, ACM DL*, pages 19–24.
- Schauerhuber, A., Schwinger, W., Kapsammer, E., Retschitzegger, W., and Wimmer, M. (2007). A survey on aspect-oriented modeling approaches. Technical report, E188 - Institut für Softwaretechnik und Interaktive Systeme; Technische Universität Wien.
- Stein, D. and Hanenberg, S. (2006). Why Aspect-Oriented Software Development And Model-Driven Development Are Not The Same - A Position Paper. *M.Aksit and E.Roubtsova eds. Proceedings of ABMB2005, ENTCS*, 163(1):71–82.
- Stein, D., Hanenberg, S., and Unland, R. (2005). Visualizing Join Point Selections Using Interaction-Based vs. State-Based Notations Exemplified With Help of Business Rules. *EMISA-2005*, pages 94–107.

PERFORMANCE ENGINEERING OF BUSINESS INFORMATION SYSTEMS

Filling the Gap between High-level Business Services and Low-level Performance Models

Samuel Kounev

Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany

kounev@kit.edu

Keywords: Business information systems, Performance, Scalability, Predictive modeling, Simulation.

Abstract: With the increasing adoption of virtualization and the transition towards Cloud Computing platforms, modern business information systems are becoming increasingly complex and dynamic. This raises the challenge of guaranteeing system performance and scalability while at the same time ensuring efficient resource usage. In this paper, we present a historical perspective on the evolution of model-based performance engineering techniques for business information systems focusing on the major developments over the past several decades that have shaped the field. We survey the state-of-the-art on performance modeling and management approaches discussing the ongoing efforts in the community to increasingly bridge the gap between high-level business services and low level performance models. Finally, we wrap up with an outlook on the emergence of self-aware systems engineering as a new research area at the intersection of several computer science disciplines.

1 INTRODUCTION

Modern business information systems are expected to satisfy increasingly stringent performance and scalability requirements. Most generally, the *performance* of a system refers to the degree to which the system meets its objectives for timeliness and the efficiency with which it achieves this (Smith and Williams, 2002; Kounev, 2008). Timeliness is normally measured in terms of meeting certain response time and/or throughput requirements, *response time* referring to the time required to respond to a user request (e.g., a Web service call or a database transaction), and *throughput* referring to the number of requests or jobs processed per unit of time. Scalability, on the other hand, is understood as the ability of the system to continue to meet its objectives for response time and throughput as the demand for the services it provides increases and resources (typically hardware) are added. Numerous studies, e.g., in the areas of e-business, manufacturing, telecommunications, health care and transportation, have shown that a failure to meet performance requirements can lead to serious financial losses, loss of customers and reputation, and in some cases even to loss of human lives. To avoid

the pitfalls of inadequate Quality-of-Service (QoS), it is important to analyze the expected performance and scalability characteristics of systems during all phases of their life cycle. The methods used to do this are part of the discipline called *Performance Engineering*. Performance Engineering helps to estimate the level of performance a system can achieve and provides recommendations to realize the optimal performance level. The latter is typically done by means of performance models (e.g., analytical queueing models or simulation models) that are used to predict the performance of the system under the expected workload.

However, as systems grow in size and complexity, estimating their performance becomes a more and more challenging task. Modern business information systems based on the Service-Oriented Architecture (SOA) paradigm are often composed of multiple independent *services* each implementing a specific business activity. Services are accessed according to specified workflows representing business processes. Each service is implemented using a set of software components distributed over physical tiers as depicted in Figure 1. Three tiers exist: presentation tier, business logic tier, and data tier. The pre-

sentation tier includes Web servers hosting Web components that implement the presentation logic of the application. The business logic tier normally includes a cluster of application servers hosting business logic components that implement the business logic of the application. Middleware platforms such as Java EE, Microsoft .NET, or Apache Tomcat are often used in this tier to simplify application development by leveraging some common services typically used in enterprise applications. Finally, the data tier includes database servers and legacy systems that provide data management services.

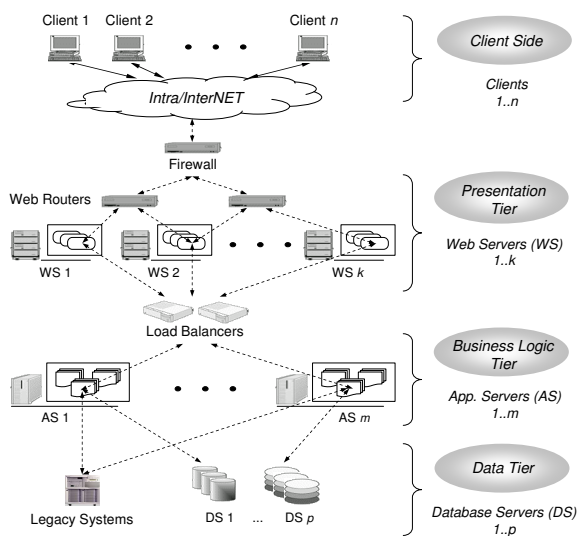


Figure 1: Modern business information system.

The inherent complexity of such architectures makes it difficult to manage their end-to-end performance and scalability. To avoid performance problems, it is essential that systems are subjected to rigorous performance evaluation during the various stages of their lifecycle. At every stage, performance evaluation is conducted with a specific set of goals and constraints. The goals can be classified in the following categories, some of which partially overlap:

Platform selection: Determine which hardware and software platforms would provide the best scalability and cost/performance ratio?

Platform validation: Validate a selected combination of platforms to ensure that taken together they provide adequate performance and scalability.

Evaluation of design alternatives: Evaluate the relative performance, scalability and costs of alternative system designs and architectures.

Performance prediction: Predict the performance of the system for a given workload and configuration scenario.

Performance tuning: Analyze the effect of various deployment settings and tuning parameters on the system performance and find their optimal values.

Performance optimization: Find the components with the largest effect on performance and study the performance gains from optimizing them.

Scalability and bottleneck analysis: Study the performance of the system as the load increases and more hardware is added. Find which system components are most utilized and investigate if they are potential bottlenecks.

Sizing and capacity planning: Determine how much hardware resources are required to guarantee certain performance levels.

Run-time performance and power management: Determine how to vary resource allocations during operation in order to ensure that performance requirements are continuously satisfied while optimizing power consumption in the face of frequent variations in service workloads.

Two broad approaches are used in Performance Engineering for performance evaluation of software systems: performance measurement and performance modeling. In the first approach, load testing tools and benchmarks are used to generate artificial workloads on the system and to measure its performance. In the second approach, performance models are built and then used to analyze the performance and scalability characteristics of the system.

In this paper, we focus on performance modeling since it is normally much cheaper than load testing and has the advantage that it can also be applied in the early stages of system development before the system is available for testing. We present a historical perspective on the evolution of performance modeling techniques for business information systems over the past several decades, focusing on the major developments that have shaped the field, such as the increasing integration of software-related aspects into performance models, the increasing parametrization of models to foster model reuse, the increasing use of automated model-to-model transformations to bridge the gap between models at different levels of abstraction, and finally the increasing use of models at run-time for online performance management.

The paper starts with an overview of classical performance modeling approaches which is followed by an overview of approaches to integrate performance modeling and prediction techniques into the software engineering process. Next, automated model-to-model transformations from architecture-level performance models to classical stochastic performance models are surveyed. Finally, the use of models at

run-time for online performance management is discussed and the paper is wrapped up with some concluding remarks.

2 CLASSICAL PERFORMANCE MODELING

The performance modeling approach to software performance evaluation is based on using mathematical or simulation models to predict the system performance under load. A *performance model* is an abstract representation of the system that relates the workload parameters with the system configuration and captures the main factors that determine the system performance (Menascé et al., 1994). A number of different methods and techniques have been proposed in the literature for modeling software systems and predicting their performance under load. Most of them, however, are based on the same general methodology that proceeds through the steps depicted in Figure 2 (Menascé et al., 2004; Smith and Williams, 2002; Kounev, 2006). First, the goals and objectives of the modeling study are specified. After this, the system is described in detail in terms of its hardware and software architecture. Next, the workload of the system is characterized and a workload model is built. The workload model is used as a basis for building a performance model. Before the model can be used for performance prediction, it has to be validated. This is done by comparing performance metrics predicted by the model with measurements on the real system obtained in a small testing environment. If the predicted values do not match the measured values within an acceptable level of accuracy, then the model must be refined and/or calibrated. Finally, the validated performance model is used to predict the system performance for the deployment configurations and workload scenarios of interest. The model predictions are analyzed and used to address the goals set in the beginning of the modeling study.

Performance models have been employed for performance prediction of software systems since the early seventies. In 1971, Buzen proposed modeling systems using queueing network models and developed solution techniques for several important classes of models. Since then many advances have been made in improving the model expressiveness and developing efficient model analysis techniques as well as accurate approximation techniques. A number of modeling techniques utilizing a range of different performance models have been proposed including standard queueing networks, extended and layered queueing networks, stochastic Petri nets, queueing Petri nets,

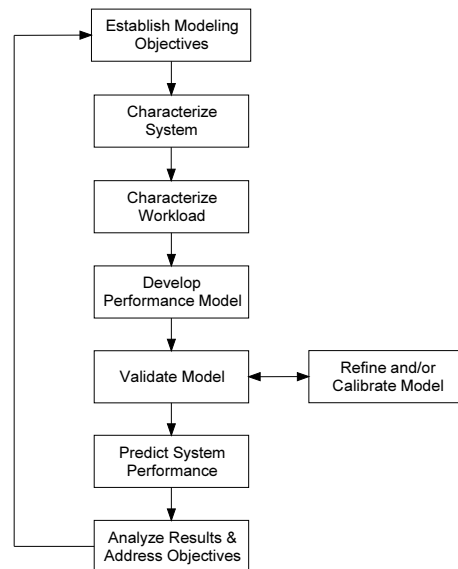


Figure 2: Performance modeling process.

stochastic process algebras, Markov chains, statistical regression models and simulation models. Performance models can be grouped into two main categories: *simulation models* and *analytical models*. One of the greatest challenges in building a good model is to find the right level of abstraction and granularity. A general rule of thumb is: Make the model as simple as possible, but not simpler! Including too much detail might render the model intractable, on the other hand, making it too simple might render it unrepresentative.

2.1 Simulation Models

Simulation models are software programs that mimic the behavior of a system as requests arrive and get processed at the various system resources. Such models are normally stochastic because they have one or more random variables as input (e.g., the times between successive arrivals of requests). The structure of a simulation program is based on the states of the simulated system and events that cause the system state to change. When implemented, simulation programs count events and record the duration of time spent in different states. Based on these data, performance metrics of interest (e.g., the average time a request takes to complete or the average system throughput) can be estimated at the end of the simulation run. Estimates are provided in the form of confidence intervals. A confidence interval is a range with a given probability that the estimated performance metric lies within this range. The main advantage of simulation models is that they are very general and can be made as accurate as desired. However, this accuracy comes at the cost of the time taken to develop

and run the models. Usually, many long runs are required to obtain estimates of needed performance measures with reasonable confidence levels.

Several approaches to developing a simulation model exist. The most time-consuming approach is to use a general purpose programming language such as C++ or Java, possibly augmented by simulation libraries (e.g., CSIM or SimPack, OMNeT++, DESMO-J). Another approach is to use a specialized simulation language such as GPSS/H, Simgen II.5, or MODSIM III. Finally, some simulation packages support graphical languages for defining simulation models (e.g., Arena, Extend, SES/workbench, QPME). A comprehensive treatment of simulation techniques can be found in (Law and Kelton, 2000; Banks et al., 2001).

2.2 Analytical Models

Analytical models are based on mathematical laws and computational algorithms used to derive performance metrics from model parameters. Analytical models are usually less expensive to build and more efficient to analyze compared to simulation models. However, because they are defined at a higher level of abstraction, they are normally less detailed and accurate. Moreover, for models to be mathematically tractable, usually many simplifying assumptions need to be made impairing the model representativeness. Queueing networks and generalized stochastic Petri nets are perhaps the two most popular types of models used in practice.

Queueing networks provide a very powerful mechanism for modeling hardware contention (contention for CPU time, disk access, and other hardware resources). A number of efficient analysis methods have been developed for a class of queueing networks called product-form queueing networks allowing models of realistic size and complexity to be analyzed with a minimum overhead (Bolch et al., 2006). The downside of queueing networks is that they do not provide direct means to model software contention aspects accurately (contention for processes, threads, database connections, and other software resources), as well as blocking, simultaneous resource possession, asynchronous processing, and synchronization aspects. Even though extensions of queueing networks, such as extended queueing networks (MacNair, 1985) and layered queueing networks (also called stochastic rendezvous networks) (Woodside et al., 1995), provide some support for modeling software contention and synchronization aspects, they are often restrictive and inaccurate.

In contrast to queueing networks, generalized

stochastic Petri net models can easily express software contention, simultaneous resource possession, asynchronous processing, and synchronization aspects. Their major disadvantage, however, is that they do not provide any means for direct representation of scheduling strategies. The attempts to eliminate this disadvantage have led to the emergence of queueing Petri nets (Bause, 1993), which combine the modeling power and expressiveness of queueing networks and stochastic Petri nets. Queueing Petri nets enable the integration of hardware and software aspects of system behavior in the same model (Kounev and Buchmann, 2003). A major hurdle to the practical use of queueing Petri nets, however, is that their analysis suffers from the state space explosion problem limiting the size of the models that can be solved. Currently, the only way to circumvent this problem is by using simulation for model analysis (Kounev and Buchmann, 2006).

Details of the various types of analytical models are beyond the scope of this article. The following books can be used as reference for additional information (Bolch et al., 2006; Trivedi, 2002; Bause and Kritzinger, 2002). The Proceedings of the ACM SIGMETRICS Conferences and the Performance Evaluation Journal report recent research results in performance modeling and evaluation. Further relevant information can be found in the Proceedings of the ACM/SPEC International Conference on Performance Engineering (ICPE), the Proceedings of the International Conference on Quantitative Evaluation of Systems (QEST), the Proceedings of the Annual Meeting of the IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), and the Proceedings of the International Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS).

3 SOFTWARE PERFORMANCE ENGINEERING

A major hurdle to the adoption of classical performance modeling approaches in industry is the fact that performance models are expensive to build and require extensive experience and expertise in stochastic modeling which software engineers typically do not possess. To address this issue, over the last fifteen years, a number of approaches have been proposed for integrating performance modeling and prediction techniques into the software engineering process. Efforts were initiated with Smith's seminal work pioneered under the name of *Software Perfor-*

formance Engineering (SPE) (Smith, 1990). Since then a number of languages (i.e., meta-models) for describing performance-relevant aspects of software architectures and execution environments have been developed by the SPE community, the most prominent being the UML SPT profile (UML Profile for Schedulability, Performance and Time) and its successor the UML MARTE profile (UML Profile for Modeling and Analysis of Real-time and Embedded Systems). The latter are extensions of UML (Unified Modeling Language) as the de facto standard modeling language for software architectures. Other proposed architecture-level performance meta-models include SPE-MM (Smith et al., 2005), CSM (Petriu and Woodside, 2007a) and KLAPER (Grassi et al., 2007a). The common goal of these efforts is to enable the automated transformation of architecture-level performance models into analytical or simulation-based performance models that can be solved using classical analysis techniques (see Section 4).

In recent years, with the increasing adoption of Component-Based Software Engineering (CBSE), the SPE community has focused on adapting and extending conventional SPE techniques to support component-based systems. A number of architecture-level performance meta-models for component-based systems have been proposed as surveyed in (Koziolek, 2009). Such meta-models provide means to describe the performance-relevant aspects of software components (e.g., internal control flow and resource demands) while explicitly capturing the influences of their execution context. The idea is that once component models are built they can be reused in multiple applications and execution contexts. The performance of a component-based system can be predicted by means of compositional analysis techniques based on the performance models of its components. Over the last five years, research efforts have been targeted at increasing the level of parametrization of component models to capture additional aspects of their execution context.

An example of a mature modeling language for component-based systems is given by the Palladio Component Model (PCM) (Becker et al., 2009b). In PCM, the component execution context is parameterized to explicitly capture the influence of the component's connections to other components, its allocated hardware and software resources, and its usage profile including service input parameters. Model artifacts are divided among the developer roles involved in the CBSE process, i.e., component developers, system architects, system deployers and domain experts.

4 MODEL-TO-MODEL TRANSFORMATIONS

To bridge the gap between architecture-level performance models and classical stochastic performance models, over the past decade the SPE community has focused on building automated model-to-model transformations which make it possible to exploit existing model solution techniques from the performance evaluation community (Marco and Mirandola, 2006). In the following, we provide an overview of the most common transformations available in the literature.

(Marco and Inverardi, 2004) transform UML models annotated with SPT stereotypes into a multichain queueing network. UML- ψ , the UML Performance Simulator (Marzolla and Balsamo, 2004), transforms a UML instance annotated with the SPT profile to a simulation model. The results from the analysis of the simulation model are reported back to the annotated UML instance (Marco and Mirandola, 2006). Another approach uses the stochastic process algebra PEPA as analysis model (Tribastone and Gilmore, 2008). In this case, only UML activity diagrams are considered, which are annotated with a subset of the MARTE profile. A software tool implementing this method is also available. (Bertolino and Mirandola, 2004) integrate their approach into the Argo-UML modeling tool, using the RT-UML performance annotation profile. An execution graph and a queueing network serve as the target analysis formalisms.

Other approaches use UML, but do not use standardized performance profile annotations: (Gu and Petriu, 2002) use XSLT, the eXtensible Stylesheet Language Transformations, to execute a graph pattern based transformation from a UML instance to LQNs. Instead of annotating the UML model, it has to be modeled in a way so that the transformation can identify the correct patterns in the model. (Bernardi et al., 2002) consider only UML statecharts and sequence diagrams. A transformation written in Java turns the model into GSPN sub-models that are then combined into a final GSPN. (Gomaa and Menascé, 2001) use UML with custom XML performance annotation. The performance model is not described in detail, but appears to be based on queueing networks. (Wu and Woodside, 2004) use UML component models together with a custom XML component performance specification language. LQN solvers are used for the analysis.

Further approaches exist that are not based on UML: (Bondarev et al., 2004; Bondarev et al., 2005) build on the ROBOCOP component model and use proprietary simulation framework for model analysis.

(Eskenazi et al., 2004) propose a custom control flow graph model notation and custom simulation framework. (Hissam et al., 2002) employ the COMTEK component technology, coupled with a proprietary analysis framework. (Sitaraman et al., 2001) specify component composition and performance characteristics using a variant of the big-O notation. The runtime analysis is not discussed in detail.

Several model-to-model transformations have been developed for the Palladio Component Model (PCM). Two solvers are based on a transformation to Layered Queueing Networks (LQNs) (Koziolek and Reussner, 2008) and a transformation to Stochastic Regular Expressions (Koziolek, 2008), respectively. Stochastic Regular Expressions can be solved analytically with very low overhead, however, they only support single user scenarios. (Henss, 2010) proposes a PCM transformation to OMNeT++, focusing on a realistic network infrastructure closer to the OSI reference network model. The PCM-Bench tool comes with the SimuCom simulator (Becker et al., 2009a) which is based on a model-to-text transformation used to generate Java code that builds on DESMO-J, a general-purpose simulation framework. The code is then compiled on-the-fly and executed. SimuCom is tailored to support all of the PCM features directly and covers the whole PCM meta-model.

Finally, a number of intermediate languages (or kernel languages) for specifying software performance information have been proposed in the literature. The aim of such efforts is to reduce the overhead for building transformations, i.e., only $M + N$ instead of $M \cdot N$ transformations have to be developed for M source and N target meta-models (Marco and Mirandola, 2006). Some examples of intermediate languages include SPE-MM (Smith et al., 2005), KLAPER (Kernel LAnguage for PErformance and Reliability analysis) (Grassi et al., 2007b) and CSM (Core Scenario Model) (Petriu and Woodside, 2007b).

5 RUN-TIME PERFORMANCE MANAGEMENT

With the increasing adoption of virtualization and the transition towards Cloud Computing platforms, modern business information systems are becoming increasingly complex and dynamic. The increased complexity is caused by the introduction of virtual resources and the resulting gap between logical and physical resource allocations. The increased dynamism is caused by the complex interactions between the applications and services sharing the physical in-

frastructure. In a virtualized service-oriented environment changes are common during operation, e.g., new services and applications can be deployed on-the-fly, service workflows and business processes can be modified dynamically, hardware resources can be added and removed from the system, virtual machines (VMs) can be migrated between servers, resources allocated to VMs can be modified to reflect changes in service workloads and usage profiles. To ensure adequate performance and efficient resource utilization in such environments, capacity planning needs to be done on a regular basis during operation. This calls for *online* performance prediction mechanisms.

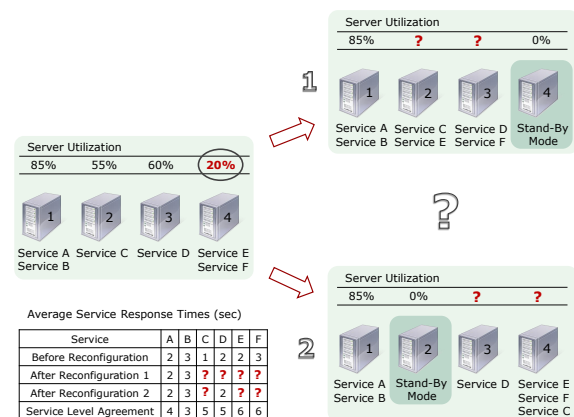


Figure 3: Online performance prediction scenario.

An example of a scenario where online performance prediction is needed is depicted in Figure 3. A service-oriented system made of four servers hosting six different services is shown including information on the average service response times, the response time service level agreements (SLAs) and the server utilization. Now assume that due to a change in the demand for services E and F, the average utilization of the fourth server has dropped down to 20% over an extended period of time. To improve the system's efficiency, it is considered to switch one of the servers to stand-by mode after migrating its services to other servers. Two possible ways to reconfigure the system are shown. To ensure that reconfiguring the system would not break the SLAs, the system needs a mechanism to predict the effect of the reconfiguration on the service response times.

Given the variety of changes that occur in modern service-oriented environments, online performance prediction techniques must support variations at all levels of the system including variations in service workloads and usage profiles, variations in the system architecture, as well as variations in the deployment and execution environment (virtualization, middleware, etc). To predict the impact of such variations,

architecture-level performance models are needed at run-time that explicitly capture the influences of the system architecture, its configuration, and its workload and usage profiles.

While, as discussed in the previous two sections, many architecture-level performance prediction techniques exist in the literature, most of them suffer from two significant drawbacks which render them impractical for use at run-time: i) performance models provide limited support for reusability and customization, ii) performance models are static and maintaining them manually during operation is prohibitively expensive (Woodside et al., 2007).

While techniques for component-based performance engineering have contributed a lot to facilitate model reusability, there is still much work to be done on further parameterizing component models before they can be used for online performance prediction. In particular, current techniques do not provide means to model the layers of the component execution environment explicitly. The performance influences of the individual layers, the dependencies among them, as well as the resource allocations at each layer should be captured as part of the models. This is necessary in order to be able to predict at run-time how a change in the execution environment (e.g., modifying resource allocations at the virtualization layer) would affect the overall system performance.

As to the second issue indicated above, the heart of the problem is in the fact that architecture-level performance models are normally designed for offline use and as such they are decoupled from the system components they represent. Thus, models do not capture dynamic aspects of the environment and therefore they need to be updated manually after every change in the system configuration or workload. Given the frequency of such changes, the amount of effort involved in maintaining performance models is prohibitive and therefore in practice such models are rarely used after deployment. Even though some techniques have been proposed to automatically construct stochastic performance models at run-time, e.g., (Menascé et al., 2007; Mos, 2004), such techniques abstract the system at a very high level without taking into account its architecture and configuration.

To address the challenges described above, current research efforts are focusing on developing *on-line* architecture-level performance models designed specifically for use at run-time. Such models aim at capturing all information, both static and dynamic, relevant to predicting the system's performance on-the-fly. They are intended to be integrated into the system components and to be maintained and updated automatically by the underlying execution platform

(virtualization and middleware) reflecting the evolving system environment.

Online performance models will make it possible to answer performance-related queries that arise during operation such as: What would be the effect on the performance of running applications if a new application is deployed in the virtualized infrastructure or an existing application is migrated from one physical server to another? How much resources need to be allocated to a newly deployed application to ensure that SLAs are satisfied? How should the system configuration be adapted to avoid performance issues or inefficient resource usage arising from changing customer workloads?

The ability to answer queries such as the above provides the basis for implementing techniques for *self-aware* performance and resource management (Kounev et al., 2010). Such techniques will be triggered automatically during operation in response to observed or forecast changes in application workloads. The goal will be to *proactively* adapt the system to such changes in order to avoid anticipated QoS problems or inefficient resource usage. The adaptation will be performed in an autonomic fashion by considering a set of possible system reconfiguration scenarios (e.g. changing VM placement and/or resource allocations) and exploiting the online performance models to predict the effect of such reconfigurations before making a decision.

Self-aware systems engineering is currently emerging as a new research area at the intersection of several computer science disciplines including software architecture, computer systems modeling, autonomic computing, distributed systems, and more recently, Cloud Computing and Green IT (Descartes Research Group, 2011; Kounev, 2011). It raises a number of big challenges that represent emerging hot topics in the systems engineering community and will be subject of long-term fundamental research in the years to come. The resolution of these challenges promises to revolutionize the field of systems engineering by enabling guaranteed QoS, lower operating costs and improved energy efficiency.

6 CONCLUDING REMARKS

We presented a historical perspective on the evolution of model-based performance engineering techniques for business information systems, focusing on the major developments over the past four decades that have shaped the field, such as the increasing integration of software-related aspects into performance models, the increasing parametrization of models to

foster model reuse, the increasing use of automated model-to-model transformations to bridge the gap between models at different levels of abstraction, and finally the increasing use of models at run-time for online performance management. We surveyed the state-of-the-art on performance modeling and management approaches discussing the ongoing efforts in the community to increasingly bridge the gap between high-level business services and low level performance models. Finally, we concluded with an outlook on the emergence of self-aware systems engineering as a new research area at the intersection of several computer science disciplines.

REFERENCES

- Banks, J., Carson, J. S., Nelson, B. L., and Nicol, D. M. (2001). *Discrete-Event System Simulation*. Prentice Hall, Upper Saddle River, NJ 07458, third edition.
- Bause, F. (1993). Queuing Petri Nets - A formalism for the combined qualitative and quantitative analysis of systems. In *Proceedings of the 5th International Workshop on Petri Nets and Performance Models, Toulouse, France, October 19-22*.
- Bause, F. and Kritzinger, F. (2002). *Stochastic Petri Nets - An Introduction to the Theory*. Vieweg Verlag, second edition.
- Becker, S., Koziolok, H., and Reussner, R. (2009a). The palladio component model for model-driven performance prediction. *J. Syst. Softw.*, 82(1):3–22.
- Becker, S., Koziolok, H., and Reussner, R. (2009b). The Palladio component model for model-driven performance prediction. *Journal of Syst. and Softw.*, 82:3–22.
- Bernardi, S., Donatelli, S., and Merseguer, J. (2002). From UML sequence diagrams and statecharts to analysable petri net models. In *Proc. on WOSP '02*, pages 35–45.
- Bertolino, A. and Mirandola, R. (2004). CB-SPE tool: Putting component-based performance engineering into practice. In *Component-Based Software Engineering*, volume 3054 of *LNCS*, pages 233–248.
- Bolch, G., Greiner, S., Meer, H. D., and Trivedi, K. S. (2006). *Queueing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications*. John Wiley & Sons, Inc., 2nd edition.
- Bondarev, E., de With, P., Chaudron, M., and Muskens, J. (2005). Modelling of input-parameter dependency for performance predictions of component-based embedded systems. In *Proc. on EUROMICRO '05*, pages 36–43.
- Bondarev, E., Muskens, J., de With, P., Chaudron, M., and Lukkien, J. (2004). Predicting real-time properties of component assemblies: a scenario-simulation approach. pages 40–47.
- Descartes Research Group (2011). <http://www.descartes-research.net>.
- Eskenazi, E., Fioukov, A., and Hammer, D. (2004). Performance prediction for component compositions. In *Component-Based Software Engineering*, volume 3054 of *Lecture Notes in Computer Science*, pages 280–293.
- Gomaa, H. and Menascé, D. (2001). Performance engineering of component-based distributed software systems. In *Performance Engineering*, volume 2047 of *LNCS*, pages 40–55.
- Grassi, V., Mirandola, R., and Sabetta, A. (2007a). Filling the gap between design and performance/reliability models of component-based systems: A model-driven approach. *Journal of Systems and Software*, 80(4):528–558.
- Grassi, V., Mirandola, R., and Sabetta, A. (2007b). Filling the gap between design and performance/reliability models of component-based systems: A model-driven approach. *J. Syst. Softw.*, 80(4):528–558.
- Gu, G. P. and Petriu, D. C. (2002). XSLT transformation from UML models to LQN performance models. In *Proc. on WOSP '02*, pages 227–234.
- Henss, J. (2010). Performance prediction for highly distributed systems. In *Proc. on WCOP '10*, volume 2010-14, pages 39–46. Karlsruhe Institute of Technology.
- Hissam, S., Moreno, G., Stafford, J., and Wallnau, K. (2002). Packaging predictable assembly. In *Component Deployment*, volume 2370 of *Lecture Notes in Computer Science*, pages 108–124.
- Kounev, S. (2006). Performance Modeling and Evaluation of Distributed Component-Based Systems using Queuing Petri Nets. *IEEE Transactions on Software Engineering*, 32(7):486–502.
- Kounev, S. (2008). *Wiley Encyclopedia of Computer Science and Engineering*, chapter Software Performance Evaluation. Wiley-Interscience, John Wiley & Sons Inc, ISBN-10: 0471383937, ISBN-13: 978-0471383932.
- Kounev, S. (2011). Self-Aware Software and Systems Engineering: A Vision and Research Roadmap. In *Proceedings of Software Engineering 2011 (SE2011), Nachwuchswissenschaftler-Symposium*.
- Kounev, S., Brosig, F., Huber, N., and Reussner, R. (2010). Towards self-aware performance and resource management in modern service-oriented systems. In *Proc. of the 7th IEEE Intl. Conf. on Services Computing (SCC 2010)*. IEEE Computer Society.
- Kounev, S. and Buchmann, A. (2003). Performance Modelling of Distributed E-Business Applications using Queuing Petri Nets. In *Proceedings of the 2003 IEEE International Symposium on Performance Analysis of Systems and Software - ISPASS2003, Austin, Texas, USA, March 20-22*.
- Kounev, S. and Buchmann, A. (2006). SimQPN - a tool and methodology for analyzing queuing Petri net models by means of simulation. *Performance Evaluation*, 63(4-5):364–394.

- Koziolok, H. (2008). *Parameter dependencies for reusable performance specifications of software components*. PhD thesis, University of Karlsruhe (TH).
- Koziolok, H. (2009). Performance evaluation of component-based software systems: A survey. *Performance Evaluation*.
- Koziolok, H. and Reussner, R. (2008). A model transformation from the palladio component model to layered queueing networks. In *Proc. on SIPEW '08*, pages 58–78.
- Law, A. and Kelton, D. W. (2000). *Simulation Modeling and Analysis*. Mc Graw Hill Companies, Inc., third edition.
- MacNair, E. A. (1985). An introduction to the Research Queueing Package. In *WSC '85: Proceedings of the 17th conference on Winter simulation*, pages 257–262, New York, NY, USA. ACM Press.
- Marco, A. D. and Inverardi, P. (2004). Compositional generation of software architecture performance QN models. *Software Architecture, Working IEEE/IFIP Conf. on*, 0:37.
- Marco, A. D. and Mirandola, R. (2006). Model transformation in software performance engineering. In *QoSA*.
- Marzolla, M. and Balsamo, S. (2004). UML-PSI: The UML performance simulator. *Quantitative Eval. of Syst.*, 0:340–341.
- Menascé, D., Ruan, H., and Goma, H. (2007). QoS management in service-oriented architectures. *Performance Evaluation*, 64(7-8):646–663.
- Menascé, D. A., Almeida, V., and Dowdy, L. W. (1994). *Capacity Planning and Performance Modeling - From Mainframes to Client-Server Systems*. Prentice Hall, Englewood Cliffs, NG.
- Menascé, D. A., Almeida, V., and Dowdy, L. W. (2004). *Performance by Design*. Prentice Hall.
- Mos, A. (2004). *A Framework for Adaptive Monitoring and Performance Management of Component-Based Enterprise Applications*. PhD thesis, Dublin City University, Ireland.
- Petriu, D. and Woodside, M. (2007a). An intermediate metamodel with scenarios and resources for generating performance models from UML designs. *Software and Systems Modeling (SoSyM)*, 6(2):163–184.
- Petriu, D. and Woodside, M. (2007b). An intermediate metamodel with scenarios and resources for generating performance models from UML designs. *Softw. and Syst. Modeling*, 6(2):163–184.
- Sitaraman, M., Kulczycki, G., Krone, J., Ogden, W. F., and Reddy, A. L. N. (2001). Performance specification of software components. *SIGSOFT Softw. Eng. Notes*, 26(3):3–10.
- Smith, C. U. (1990). *Performance Engineering of Software Systems*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.
- Smith, C. U., Llad, C. M., Cortellessa, V., Di Marco, A., and Williams, L. G. (2005). From UML models to software performance results: an SPE process based on XML interchange formats. In *WOSP '05: Proceedings of the 5th International Workshop on Software and Performance*, pages 87–98, New York, NY, USA. ACM Press.
- Smith, C. U. and Williams, L. G. (2002). *Performance Solutions - A Practical Guide to Creating Responsive, Scalable Software*. Addison-Wesley.
- Tribastone, M. and Gilmore, S. (2008). Automatic extraction of PEPA performance models from UML activity diagrams annotated with the MARTE profile. In *Proc. on WOSP '08*.
- Trivedi, K. S. (2002). *Probability and Statistics with Reliability, Queueing and Computer Science Applications*. John Wiley & Sons, Inc., second edition.
- Woodside, M., Franks, G., and Petriu, D. (2007). The future of software performance engineering. In *Future of Software Engineering (FOSE'07)*, pages 171–187, Los Alamitos, CA, USA. IEEE Computer Society.
- Woodside, M., Neilson, J., Petriu, D., and Majumdar, S. (1995). The Stochastic Rendezvous Network Model for Performance of Synchronous Client-Server-Like Distributed Software. *IEEE Transactions on Computers*, 44(1):20–34.
- Wu, X. and Woodside, M. (2004). Performance modeling from software components. *SIGSOFT Softw. Eng. Notes*, 29(1):290–301.

LEVERAGING THE PUBLISH-FIND-USE PARADIGM OF SOA

Supporting Enterprise Collaboration Across Organisational Boundaries

Brahmananda Sapkota and Marten van Sinderen

*Faculty of Electrical Engineering, Mathematics and Computer Science
University of Twente, Drienerlolaan 5, Enschede, The Netherlands
b.sapkota@utwente.nl, m.j.vansinderen@utwente.nl*

Keywords: Service Oriented Architecture, Enterprise Application Integration, Cross-organisational Collaboration, Business Requirements, Business Processes, Business-IT alignment.

Abstract: Service Oriented Architecture (SOA) has been widely recognized as an approach for flexible integration of enterprise applications across organisational boundaries based on service abstractions and Internet standards. Enterprise collaboration and application integration is driven by business requirements, which in turn are translated to business models and expressed as business processes. For example, business processes can be used to represent the coordination of several published services as well as the implementation of a composite value-added service. In this way, enterprise functions are aggregated using multi-stage business processes fulfilling the specific requirements of an enterprise. In order for enterprises to stay competitive in their respective businesses, such solutions must evolve in a timely and appropriate way in response to changes in market demands and opportunities that inevitably occur. Therefore, a mechanism is required for business-IT alignment during the complete lifecycle of SOA-based enterprise collaboration and application integration. In this paper, we discuss issues related to cross-organisational collaboration and how service-oriented principles and architectures can be applied to address these issues.

1 INTRODUCTION

Service Oriented Architecture (Erl, 2005) has been widely recognized as an approach for flexible integration of enterprise applications across organizational boundaries based on service abstractions and Internet standards. On-demand enterprise collaboration and application integration is driven by business requirements, which in turn are translated to business models and expressed as business processes. For example, business processes can be used to represent the coordination of several published services as well as the implementation of a composite value-added service. In this way, enterprise functions are aggregated using multi-stage business processes fulfilling the specific requirements of an enterprise (van Sinderen and Almeida, 2011). In order for enterprises to stay competitive in their respective businesses, such solutions must evolve in a timely and appropriate way in response to changes in market demands and opportunities that inevitably occur. Service Oriented Architecture (SOA) supports adaptation of such an evolution through the concept of service composition. The

composition of services is made possible because of *service discovery* which allows to find and select the suitable services. The service discovery thus plays a central role in enterprise collaboration and application integration.

One of the fundamental requirements for the service discovery are service descriptions and service registries. In general, service descriptions are used for specifying: 1) what functionalities are offered by the service to its users (i.e., the interface definition); 2) how the service is provided (i.e., the service binding); and 3) where the service can be accessed (i.e., the service endpoint information). The service registries are used to announce the offered services and thus play crucial role for a successful on-demand enterprise collaboration and application integration. In a typical scenario, service providers publish their service descriptions to a publicly accessible service registry. The service users search over these registries and find the information needed to use the required services. This approach of publishing, finding and using the services forms the so called SOA triangular operational model which clearly separates the

role of service provider, service user and service registry (van Sinderen, 2009). SOA, thus, enables flexible on-demand enterprise collaboration and application integration. Despite these sound principles of SOA, a number of practical complexities still exists, which are preventing enterprises to fully exploit the potential benefits of SOA.

An on-demand collaboration can only be achieved if 1) the participants required for the collaboration can be found and 2) the participants can communicate with each other. In the SOA based approach, the former requirement is supported by publishing the descriptions of the offered services and latter requirement is supported by defining the messages being sent. This can be realised relatively easily by using pre-meditated message structures and function libraries, if the enterprises collaborate in closed environment. If the autonomous enterprises are to collaborate, pre-meditated message structures or function libraries cannot be used. Therefore, semantics of the information provided through the service descriptions should be well defined and the service repository should contain valid service descriptions. While the initiatives around Semantic Web services have defined formalisms such as WSMO (Roman et al., 2006), OWL-S (Martin, 2004) and SAWSDL (Farrell and Lausen, 2007), to semantically define service descriptions, they cannot ensure that the published descriptions correctly reflect the offered services at the time of their discovery. Therefore, a mechanism is required for business-IT alignment during the complete lifecycle of SOA-based enterprise collaboration and application integration.

In this paper, we propose a *Summary of Services per Provider* (SSP) description based approach for service discovery and to ensure that the published descriptions are valid. An SSP description provides a means for describing the collection of services offered by a single service provider. It specifies what type of services are offered by a particular service provider. In the proposed approach, we use Web Service Definition Language (Chinnici et al., 2007) as the language for describing services and allow service providers to store them in their local repository. We provide a mechanism to generate SSP description based on the Web Service Definition Language (WSDL) documents stored at the local service registry of the service providers. These SSP descriptions are then published to the public service registry. Though the service provider still has to generate and publish these descriptions, it can be automated. Through this separation, we aim at reducing the registry updating burden at the side of the service providers. The proposed approach, therefore, follows the SOA triangular oper-

ational model except that SSP descriptions are published to the service registry instead of publishing the service descriptions.

The rest of the paper is structured as follows: Some of the highly relevant existing works are discussed in Section 2. The service discovery challenges are discussed in Section 3. The proposed solution is presented in Section 4 and its use to support cross-organisational collaboration is presented in Section 5. Finally, the work presented in this paper is concluded in Section 6 by highlighting possible future directions.

2 RELATED WORKS

The problem of guaranteeing correctness of the published service descriptions and reducing the effort required for providing such guarantees is starting to attract attention from the research communities. The work presented in (Küster and Köning-Ries, 2007) follows an approach similar to the one presented in this paper. The focus of this work is to ensure that the published service description indeed represents the concrete service. In order to support this, an estimation step followed by a single execution step is proposed. In the estimation step, additional information than that is available in the service description itself is gathered whereas in the execution step, the actual invocation of service is performed.

A preference-based selection of highly configurable web services is presented in (Lamparter et al., 2007). It focuses on defining algorithms required for finding optimal configurations while selecting the services. Unlike the work presented in this paper, their work neither considers minimisation of the extra effort required to update the service registry when service descriptions are changed nor maximising the correctness of the published service descriptions. To address the problem due to changes in service descriptions a RSS-based mechanism to announce such changes is proposed in (Treiber and Dustdar, 2007). The RSS-based approach is service provider dependent because the changing information should come from them.

Treating services from the economical point of view, (Cardoso et al., 2009) defines universal service description language. The proposed language is defined to describe both the IT and non-IT services. In the proposed language, provisions for defining dynamic information is poorly defined. The approach presented in (Truong et al., 2010) defines mechanism for identifying and reducing irrelevant information in service composition and execution. This approach

is target at increasing efficiency and correctness of the composition and execution. This approach works only after the services are discovered and does not eliminate the possibility of discovering services with incorrect information.

In contrast to many other traditional approaches, (Speiser and Harth, 2011) proposes a LinkedData based approach for integrating data providing services. Their approach is suitable for sharing data which might change over time. In comparison to the work presented in this approach, their approach requires the service providers to describe the offered services using LinkedData principles and does not support sharing of already existing service descriptions which are described using WSDL. We propose mechanisms to allow usage of WSDL while still dealing with changing and state dependent data.

A crawl based approach for collecting, annotating and classifying public Web services has been proposed in (AbuJarour et al., 2010) attempting to increase the role of service registry in service-oriented architecture by providing correct information. In their approach, publicly available web service descriptions are crawled, annotation information is gathered, Web services are annotated and classified based on this information. Through such classification, authors aim at supporting better discovery of services. In this direction, the work presented in (Obrst et al., 2010) aims at enabling rich discovery of Web services by projecting weak semantics from structural specifications. These approaches, however, cannot guarantee that the information required for service discovery is gathered.

Combination of document classification and ontology alignment schemes is proposed in (CRASSO et al., 2010) to semantically enrich Web services. Though this scheme helps in efficiently discovering required services, it does not tackle the problem of outdated service descriptions. The work presented in (da Silva et al., 2011) specifies mechanisms for runtime discovery, selection and composition of semantic services. The proposed approach supports semantic descriptions of the services but lacks support for dealing with outdated service descriptions.

3 SERVICE DISCOVERY CHALLENGES

In an open environment, it is difficult to support on-demand collaboration if the published service descriptions are either outdated or provide ambiguous or incorrect information. This difficulty escalates when service descriptions contain limited information, ei-

ther because the service providers are unwilling to share all the information or because the information is state dependent (Treiber and Dustdar, 2007; Küster and Köning-Ries, 2007), i.e., the information may change as the service is invoked. This will result in a poor discovery results.

The correctness problem arises due to the fact that service registries are passive. If the functionalities of the offered services are changed, service providers should take the initiative to update the corresponding descriptions published in the service registry. In practice, the published descriptions are rarely updated. Instead of publishing the service descriptions to the service registries, they are published on the Web (Michlmayr et al., 2007). Such a practice, violates the original SOA model (i.e., the triangular operational model) and consequently undermines the role of service registries in SOA (AbuJarour et al., 2010). This shift in practice is due to the lack of efficient and elegant support to update the service registry whenever a service description is updated. The latter essentially requires an additional effort on part of the service providers. It becomes more problematic because the existing service registries emerge and disappear (Sabou and Pan, 2007) and cannot be fully relied upon for service discovery. The results of the investigation of Web services on the Web published in (Al-Masri and Mahmoud, 2008) reveals that only around 63% of the discovered Web services are in fact active. This lack of reliable service repositories makes on-demand collaboration between autonomous enterprises difficult, if not impossible.

Besides these technical difficulties, there are other reasons why public service registries have not been successful so far. One of the reasons is that a considerable amount of the published service descriptions are unusable (Treiber and Dustdar, 2007). Some of the information (e.g., the information that depends on the change in state) which might be important for discovery purposes cannot be included in the service descriptions in a simple way. In addition, some providers may not be willing to disclose information related to non-functional properties and the quality of service parameters because of the fear of bargain or competition from other providers (Küster and Köning-Ries, 2007). This contributes to the retrieval of imprecise service descriptions leading to false positives and consequently reducing the usability and the reliability of service registries.

The above mentioned problems could be resolved if service providers are allowed to store their service description locally and provide them with a tool that extracts summary information from their repository, builds an SSP description and publishes it to the ser-

vice registry. The benefits of this approach are: 1) service providers do not have to publish all the information. It will also reduce the extra effort required for maintaining and updating the service registry. Service providers do not need to update service registry every time the service description is updated, updates are done locally. 2) service users can use these SSP descriptions to find the potential providers, and finally obtain the up-to-date information. Figure 1 shows the overall architecture of this approach.

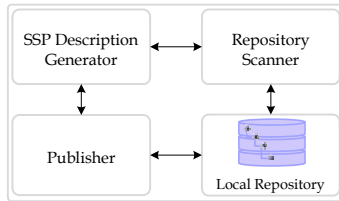


Figure 1: Overall architecture.

4 SERVICE DISCOVERY SOLUTIONS

We describe a mechanism to extract type information from the WSDL files stored in the service provider’s local repository and a mechanism to publish these information as SSP description to the service repository.

4.1 Generating Type Information

One of the purposes of using type information, in the proposed approach, is to provide indication of what type of services are offered by the service providers. This kind of information serves the purpose of *guiding* service requests towards the potential service providers. The type information based approaches are expected to help in narrowing down the search space and allowing applications to deal with the ever increasing number of Web services. We extract these information from service descriptions encoded in WSDL, which is a commonly used service description language. A WSDL document is structured into four elements describing *Service*, *Bindings*, *Interface* and *Types* definitions. A *Service* definition specification specifies a collection of endpoints (i.e., URLs). The *Bindings* definition typically specifies what data formats and communication protocols to use when invoking the service. The operations, message exchange patterns and mechanisms for fault handling are specified in the *Interface* definition. The data types used in messages and faults are specified in the *Types* definition. Figure 2 shows some of these

parts pictorially where as the Listing 1 shows the type and service definition parts of a WSDL document.

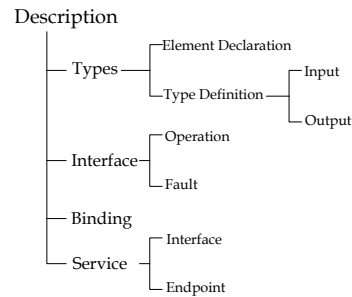


Figure 2: Pictorial representation of WSDL structure.

The data types specified in the *Types* definition essentially model the domain knowledge and thus are useful for extracting the necessary information for generating SSP description from a given WSDL document.

```

<wSDL:description
  targetNamespace="http://org.example.com/services/AvailabilityService/"
  xmlns="http://org.example.com/services/AvailabilityService/"
  xmlns:wSDL="http://www.w3.org/ns/wSDL"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema">

  <wSDL:types>
    <xsd:schema targetNamespace="http://org.example.com/resources/AvailabilityService">
      <xsd:element name="ServiceRequest">
        <xsd:complexType>
          <xsd:sequence>
            <xsd:element name="product" type="xsd:string"/>
            <xsd:element name="date" type="xsd:string"/>
            <xsd:element name="quantity" type="xsd:float"/>
          </xsd:sequence>
        </xsd:complexType>
      </xsd:element>
      <xsd:element name="ServiceResponse" type="response"/>
      <xsd:element name="ServiceResponse" restriction="xsd:boolean"/>
      <xsd:simpleType name="response">
        <xsd:restriction base="xsd:boolean"/>
      </xsd:simpleType>
    </xsd:schema>
  </wSDL:types>
  <wSDL:interface name="AvailabilityServiceInterface">
    <wSDL:operation name="RequestOperation" pattern="http://www.w3.org/ns/wSDL/in-out">
      <wSDL:input element="ServiceRequest"/>
      <wSDL:output element="ServiceResponse"/>
    </wSDL:operation>
  </wSDL:interface>
</wSDL:description>
  
```

Listing 1: Fragment of a WSDL document.

Given a WSDL document, we extract the Types definitions and represent them as RDF (Klyne and Carroll, 2004) triples. Representing those definitions in RDF has several advantages. RDF provides an abstract model for describing resources with properties, scoping them to a particular application domain through RDF Schema (Hayes, 2004) and defining relationships between these resources. In addition, standard RDF query language SPARQL (Prud’hommeaux and Seaborne, 2007) can be used to provide flexible means to allow users to express their requests. The RDF triples are generated based on the XML

Schema to RDF Schema mappings approaches presented in (Thuy et al., 2008) and (CRASSO et al., 2010). The resulting RDF triples are shown in Listing 2.

```
<rdfs:Class rdf:ID="http://org.example.com/resources/AvailabilityService#
ServiceRequest"/>
<rdfs:Class rdf:ID="http://org.example.com/resources/AvailabilityService#
ServiceResponse"/>
<rdf:Property rdf:about="http://org.example.com/resources/AvailabilityService#
response">
  <rdfs:domain rdf:resource="http://org.example.com/resources/
AvailabilityService#ServiceRequest"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#boolean"/>
</rdf:Property>
<rdf:Property rdf:ID="http://org.example.com/resources/AvailabilityService#
product">
  <rdfs:domain rdf:resource="http://org.example.com/resources/
AvailabilityService#ServiceRequest"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</rdf:Property>
<rdf:Property rdf:ID="http://org.example.com/resources/AvailabilityService#date">
  <rdfs:domain rdf:resource="http://org.example.com/resources/
AvailabilityService#ServiceRequest"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</rdf:Property>
<rdf:Property rdf:ID="http://org.example.com/resources/AvailabilityService#
quantity">
  <rdfs:domain rdf:resource="http://org.example.com/resources/
AvailabilityService#ServiceRequest"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#int"/>
</rdf:Property>
```

Listing 2: RDF representation of extracted information.

In Listing 2, each triples are encoded in the form $\langle s, p, o \rangle$ where s , p , and o are called the *subject*, *predicate* and the *object* respectively. In a triple, a predicate is called the property of the triple and denotes the relationship between the subject and the object that it connects. The subject always appears at the left whereas the object appears at the right side of the predicate in the triple.

4.2 SSP Description

An SSP description models the service provider as a collection of services and enumerates all the offered services. In particular, the SSP description specifies the type of services that are offered by a particular service provider. We use the information extracted from the locally stored WSDL files to describe domain specific concepts and to avoid ambiguities between services from different application domains. Using the SSP description, functional properties of the offered services are described collectively. These descriptions provide information sufficient enough to filter out the completely irrelevant service providers.

We define SSP description as $\langle n, e, Q \rangle$, where n is the URL of the service provider, e is the SPARQL endpoint of the local repository and Q is the collection of $\langle p(s_i, o_i), f \rangle$ pairs, where s_i and o_i represents the type of the subject and object connected by the predicate p whereas f represents the total number of occurrences of the predicate p together with

the s_i and o_i . The frequency of occurrences of subject and object type combination of each predicate is measured to indicate the number of services from a particular domain. We included the subject and object types in the SSP description because they represent the domain and range of a predicate and hence is useful to unambiguously select the required information. The structure of the SSP description is defined as RDF graph as shown in Listing 3.

```
<?xml version="1.0"?>
<rdf:RDF xmlns:ex="http://www.example.org/ontology#"
  xmlns:ont="http://www.example.org/ns/ms-ontology#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <ont:Service rdf:about="http://www.example.org/provider/xyz">
    <ont:endPoint rdf:resource="http://www.example.org/services/abc"/>
    <ex:name>
      <rdf:Description>
        <ex:subjectType rdf:resource="http://www.example.org/ontology#Hotel"/>
        <ex:objectType rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
        <ex:count>435</ex:count>
      </rdf:Description>
    </ex:name>
    <ex:product>
      <rdf:Description>
        <ex:subjectType rdf:resource="http://www.example.org/ontology#
Automobile"/>
        <ex:objectType rdf:resource="http://www.example.org/ontology#Wheel"/>
        <ex:count>1481</ex:count>
      </rdf:Description>
    </ex:product>
  </ont:Service>
</rdf:RDF>
```

Listing 3: An example SSP description.

The number of instances of objects and subjects in a repository is typically far larger than the number of distinct predicates. Inclusion of the objects in the SSP description is therefore likely to increase its size as close to as the size of the repository. In order to avoid such problems, the SSP description includes only the subject and object types and not their instances. This allows to select the relevant service provider based on this type information. For example, if there are a large number of printing service providers and only few of them are providing book printing services it is much more efficient to send a book printing service requests only to those that provide the book printing services. This type of SSP description will be helpful in selecting the service providers in situations where fine grained information is needed for answering the service requests.

4.3 Finding Potential Service Providers

In order to find the required service we follow two step discovery mechanism. In the first step, the discovery request is sent to the service registry from which a list of potential service providers are identified. In the second step, direct communication with the potential service providers is established to find the most appropriate service provider. The goal of the first step is to reduce the search space whereas the

second step is intended to select the service provider based on up-to-date information.

The major design goals of the proposed approach is simplicity and extensibility. The SSP description is essentially the collection of summary of the WSDL files stored in the local repository of the providers which simplifies the process of service description updates. The service providers need to update only the local repository. If the new service providers arise, they can simply publish the SSP description to the service registry.

We assume that the service providers employ the mechanisms discussed above for extracting the information needed for generating SSP descriptions. First, the service requester initiates the lookup over SSP descriptions and obtains a list of endpoints of the potential partners. These endpoints are then queried to obtain further information for selecting the potential service providers.

4.4 Maintaining Freshness

The SSP descriptions are published to the service registry and hence can still pose the same problems as with publishing the service descriptions if the domain information is changed. In order to ensure that the SSP descriptions are still up-to-date, the service repository requires a mechanism to reverse look up the service providers local repository and synchronize the information that is being provided through the SSP descriptions.

5 ENTERPRISE COLLABORATION

Let us now return to our original goal of facilitating enterprise collaboration using SOA. We previously concluded that the SOA architectural triangle with its 'publish-find-use' paradigm is in principle very convenient to enterprises to utilize distributed capabilities that may be under the control of different ownership domains. The convenience stems from the loose coupling of services - supporting flexible composition - and the external-oriented representation of services - allowing interactions between users and providers irrespective of their internal implementation. However, we also concluded that the 'publish-find' part of the triangle has practical limitations, which so far has prevented the successful uptake of public-registry/open-discovery based enterprise collaboration. The main limitations are: (a) it is hard to find and compose services based on current service descriptions, since the descriptions lack unambiguous and precise semantics;

(b) it is expensive and laborious to maintain service descriptions, as the corresponding services continuously evolve and therefore the descriptions require frequent and manually managed updates; (c) trust is a hindrance for publishing service descriptions and using services discovered with public registries.

We propose to leverage the publish-find-use paradigm by using public descriptions that are automatically generated and semantically enhanced, as described in Section 4. In the following, we first show which interactions are necessary for enterprise collaboration using our approach, and subsequently discuss the potential benefits of our approach.

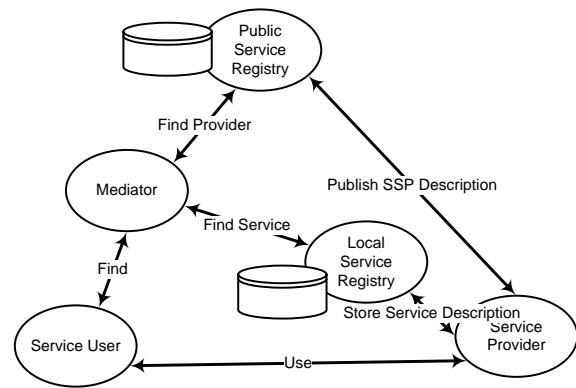


Figure 3: Proposed service discovery approach.

Figure 3 illustrates the basic interactions:

1. A business organisation acting as a service provider can publish relevant information on all its services using a single keyword-based description (SSP) at a service broker that maintains a public registry. If services offered by the business organization evolve, the SSP is re-generated or incrementally updated, depending on the nature of the change. The updated SSP can be pushed to the service broker, in similar to the original publication, where it is used to replace the old SSP. Alternatively, the service broker periodically asks the service provider for updates.
2. A business organisation looking for partners that can offer certain services can contact the service broker and find SSPs based on keyword matching. Although keyword matching can already be reasonable efficient (Obrst et al., 2010), we can further improve the recall and precision of service discovery by exploiting the RDF-based semantics of SSPs.
3. If an SSP fulfils the search criteria, the service provider is contacted via the endpoint that is part

of the SSP. Then the local repository of the service provider is used to find available services. Both step (2) and step (3) may be performed via an intermediary, making the two-step service discovery transparent to the requesting business organization. For example, the service broker may take the intermediary role. Possibly, the requested services are not available as single services or are not available from a single service provider. In that case, the services have to be offered as a bundle or must be composed. Again, an intermediary may automate or support this process (da Silva et al., 2011).

4. Once the (composed) services are found, the requesting business organization can start using them, effectively entering collaborations with one or more partner business organizations that are involved in the offering of these services.

Comparing this with the enterprise collaboration through public UDDI registries, we observe the following benefits:

- The SSP is based on extracting keywords from WSDL type definitions, and represents these keywords and their relationships with RDF. In this way, the semantic properties of keywords can be captured (CRASSO et al., 2010; Thuy et al., 2008). This addresses the limitation (a) mentioned above.
- Since the extraction is automatic, the burden for service providers to update descriptions is dramatically lowered. Moreover, if the service broker is able to poll for updates, the problem of 'disappearing' business organizations and 'ghost' services can be tackled. If a business organization no longer supports its previously published services, e.g. because it no longer exists, a poll for updates by the service broker gets no reaction and the service broker can decide to remove the SSP from its registry. This addresses the limitation (b) mentioned above.
- The two-step service discovery approach has the advantage that it first determines the services providers that offer potentially relevant services, and then limits the search for services to those of the selected service providers. Although we still have to confirm this with experiments, we believe that this approach has a better scalability than one-level semantic search. Furthermore, by favoring services from the same or a few providers, it is more likely that these services are defined and implemented in a consistent way, making search and composition easier and more efficient (Forestiero et al., 2010).

- In order to address limitation (c) mentioned above, the local registry of a service provider may be enhanced in two ways. First, the service provider may monitor who wants to access the local registry, and expose information on its services depending on some trust classification scheme (e.g., based on previous collaborations). Secondly, the service provider may provide additional information through the local registry, which facilitates the (non-) selection of services. For example, non-functional properties based on resource availability or historical data may be published, including information on trust, security or privacy aspects.
- The implications for existing standards, most notably UDDI, is minimal. Most of the interactions described above can be supported with UDDI as is.

6 CONCLUSIONS

Our research demonstrates how SSP descriptions allows us to reduce maintainability cost at the service providers side while still providing the relevant information required for service discovery. This type of approach has the ability to guarantee better service results due to the separation of abridged service descriptions and the actual detailed descriptions. Current approaches either do not provide adequate support for publishing accurate information of the offered services or the offered solutions are too restrictive in terms of cost and time required for maintaining the published descriptions. This is mainly because the service descriptions are valid only at the time they are created and subject to frequent change depending on changes in market trends.

ACKNOWLEDGEMENTS

This material is based upon works jointly supported by the IOP GenCom U-Care project (<http://ucare.ewi.utwente.nl>) sponsored by the Dutch Ministry of Economic Affairs under contract IGC0816 and by the DySCoTec project sponsored by the Centre for Telematics and Information Technology (CTIT), University of Twente, The Netherlands.

REFERENCES

- AbuJarour, M., Naumann, F., and Craculeac, M. (2010). Collecting, Annotating, and Classifying Public Web Services. In *Proc. of International Conference on On the Move to Meaningful Internet Systems*, pages 256–272.
- Al-Masri, E. and Mahmoud, Q. H. (2008). Investigating Web Services on the World Wide Web. In *Proc. of the World Wide Web Conference*, pages 759–804.
- Cardoso, J., Winkler, M., and Voigt, K. (2009). A Service Description Language for the Internet of Services. In *Proc. of the International Symposium on Services Science*.
- Chinnici, R., Moreau, J.-J., Ryman, A., and Weerawarana, S. (2007). *Web Services Description Language (WSDL) Version 2.0*.
- CRASSO, M., ZUNINO, A., and CAMPO, M. (2010). Combining Document Classification and Ontology Alignment for Semantically Enriching Web Services. *New Generation Computing*, 28:371–403.
- da Silva, E. G., Pires, L. F., and van Sinderen, M. (2011). Towards runtime discovery, selection and composition of semantic services. *Computer Communications*, 34(2):159–168.
- Erl, T. (2005). *Service-Oriented Architecture Concepts, Technology, and Design*. Prentice Hall Professional Technical Reference.
- Farrell, J. and Lausen, H. (2007). *Semantic Annotations for WSDL and XML Schema*.
- Forestiero, A., Mastroianni, C., Papuzzo, G., and Spetzano, G. (2010). A Proximity-Based Self-Organizing Framework for Service Composition and Discovery. In *Proc. of the 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing*, pages 428–437.
- Hayes, P., editor (2004). *RDF Semantics*. W3C Recommendation.
- Klyne, G. and Carroll, J. J., editors (2004). *Resource Description Framework: Concepts and Abstract Syntax*. W3C Recommendation.
- Küster, U. and Köning-Ries, B. (2007). Supporting Dynamics in Service Descriptions - The Key to Automatic Service Usage. In *Proc. of the 5th International Conference on Service-Oriented Computing*, pages 220–232.
- Lamparter, S., Ankolekar, A., and Grimm, S. (2007). Preference-based Selection of Highly Configurable Web Services. In *Proc. of the 16th International Conference on World Wide Web*, pages 1013–1022.
- Martin, D., editor (2004). *OWL-S: Semantic Markup for Web Services*. W3C Member Submission.
- Michlmayr, A., Rosenberg, F., Platzer, C., Treiber, M., and Dustdar, S. (2007). Towards Recovering the Broken SOA Triangle: A Software Engineering Perspective. In *Proc. of the 2nd International Workshop on Service Oriented Software Engineering*, pages 22–28.
- Obrst, L., McCandless, D., and Bankston, M. (2010). Enabling Rich Discovery of Web Services by Projecting Weak Semantics from Structural Specifications. In *Proc. of Semantic Technology for Intelligence, Defense, and Security*.
- Prud'hommeaux, E. and Seaborne, A., editors (2007). *SPARQL Query Language for RDF*. W3C Candidate Recommendation.
- Roman, D., Lausen, H., and Keller, U., editors (2006). *Web Service Modeling Ontology (WSMO)*. WSMO Working Group.
- Sabou, M. and Pan, J. (2007). Towards semantically enhanced Web service repositories. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5(2):142–150.
- Speiser, S. and Harth, A. (2011). Integrating Linked Data and Services with LIDS. In *Proc. of the 8th Extended Semantic Web Conference*.
- Thuy, P. T. T., Lee, Y.-K., Lee, S., and Jeong, B.-S. (2008). Exploiting XML Schema for Interpreting XML Documents as RDF. In *Proc. of the 2008 IEEE International Conference on Services Computing*, pages 555–558.
- Treiber, M. and Dustdar, S. (2007). Active Web Service Registries. *IEEE Internet Computing*, 11(5):66–71.
- Truong, H.-L., Comerio, M., Maurino, A., Dustdar, S., Paoli, F. D., and Panziera, L. (2010). On Identifying and Reducing Irrelevant Information in Service Composition and Execution. In *Proc. of the International Conference on Web Information Systems Engineering*, pages 52–66.
- van Sinderen, M. (2009). From Service-Oriented Architecture to Service-Oriented Enterprise. In *Proc. of the Third International Workshop on Enterprise Systems*, pages 3–16.
- van Sinderen, M. and Almeida, J. P. A. (2011). Empowering Enterprises through Next-Generation Enterprise Computing. *Enterprise Information Systems*, 5(1):1–8.

REGULATION, THE INVISIBLE PART OF THE GOAL ORIENTED REQUIREMENTS ENGINEERING ICEBERG

Gil Regev^{1,2} and Alain Wegmann¹

¹*Ecole Polytechnique Fédérale de Lausanne (EPFL), School of Computer and Communication Sciences
CH-1015 Lausanne, Switzerland*

²*Itecor, Av. Paul Cérésole 24, cp 568, CH-1800 Vevey 1, Switzerland
{gil.regev, alain.wegmann}@epfl.ch, g.regev@itecor.com*

Keywords: Goals, Requirements, Regulation, Survival, Systems, Norms, Beliefs.

Abstract: Goal-Oriented Requirements Engineering (GORE) is considered to be one of the main achievements that the requirements of the Requirements Engineering field has produced since its inception. Several GORE methods were designed in the last twenty years in both research and industry. Curiously, GORE methods seem to have emerged out of nowhere in the early 1990s, the concept of Goal appearing as a natural element in explaining human and organizational behaviour. We have found no theoretical or philosophical work that explicitly link GORE to an underlying organizational model. In this paper, we show that most GORE methods are implicitly based on the goal-seeking, decision making organizational model. We argue that there are other organizational models that may better explain human behaviour, albeit at the expense of more complex models. We present one such alternative model that explains individual and organizational survival through continuous regulation. We give our point of view of the changes needed in GORE methods to support this alternative view through the use of maintenance goals and beliefs.

1 INTRODUCTION

In the 25 years or so since the advent of requirements engineering (RE) research as an academic discipline, its flagship methods have been the Goal-Oriented Requirements Engineering (GORE). The GORE methods have been lauded as one of the main achievement of the RE community, especially in its first 15 years of existence (van Lamsweerde, 2001), (Mylopoulos, Kolp and Castro, 2006). GORE is still a very active field of research with dedicated workshops and conference tracks (e.g. the i* International workshop series, International Workshop on Requirements, Intentions and Goals in Conceptual Modeling).

In RE practice, GORE methods developed in research have had less influence but they are matched by goal-oriented methods that were created by RE practitioners, e.g. Goal-Oriented Use Cases (Cockburn, 2001), Essential Use Cases (Constantine, 1995). In this paper we designate all these methods as GORE, whether they have emerged from research or from practice.

The emergence of GORE methods has coincided with a less software centric view of requirements.

RE has evolved out of software specification methods by capturing more and more of the environment of the envisioned software system (Nuseibeh and Easterbrook, 2000), i.e. the composite system (van Lamsweerde, 2001). The GORE movement has been based on the understanding that goals justify and explain requirements that are assigned to agents in the composite system (software system and its environment), and that they help detecting and resolving conflicts among different stakeholder viewpoints (Dardenne, van Lamsweerde and Fickas, 1993). In GORE methods and subsequently in RE, it has been assumed that the behaviour of the stakeholders in the environment of the software system is predominantly goal-oriented, see for example (Loucopoulos and Kavakli, 1995), (Nuseibeh and Easterbrook, 2000).

There has been little debate concerning the epistemological roots of GORE methods (for an exception, see our own work (Regev and Wegmann, 2005)). Very few RE researchers have challenged this assumption. The few papers that have discussed problems with GORE methods have sought to supplement them with more artefacts, such as

subject matters (Zave and Jackson, 1997) and scenarios (Rolland, Souveyet and Ben Achour, 1998).

For Nuseibeh and Easterbrook (Nuseibeh and Easterbrook, 2000) there are several types of RE depending on the end product that is envisioned. They give the following examples: RE for information systems, RE for embedded control systems and RE for generic services e.g. networking and operating systems. In this classification our discussion applies mostly to RE for information systems. In this type of RE, the composite system is the organization that the envisioned computer-based system serves. Nuseibeh and Easterbrook further state that the context of most RE and software activities is in this field of information systems development. Hence our discussion is applicable to many RE projects. We hope that readers interested in other types of RE will benefit from this discussion as well.

We take Gause and Weinberg's view that RE is about discovering what is desired (Gause and Weinberg, 1989). In this paper, we show that what people desire has more to do with the way they regulate their affairs than with the goals they seem to pursue. We base our proposal mostly on Vickers concept of Appreciative System (Vickers, 1968), (Vickers, 1987), (Regev, Hayard and Wegmann, 2011) and show that goals are only the visible part of the way individuals and organizations regulate their relationships in order to survive. We also show that GORE methods have most of the necessary constructs to model this behaviour, e.g. maintenance goals and beliefs. We advocate a more systematic use of and widespread use of these concepts for better understanding regulation.

We begin by reviewing the GORE research (Section 2). We then introduce the regulation organizational model and show that it can be used as the underlying mechanism of which goal-oriented behaviour is but the visible part (Section 3). We identify some of the shortcomings of GORE methods and propose remedies based on the regulation-oriented view (Section 4). We review some related work in Section 5.

2 AN OVERVIEW OF GORE METHODS

GORE methods use the concept of Goal as the main construct for defining requirements. The first papers linking goals and requirements date to the beginning of the RE discipline, e.g. (Dubois, 1989), (Robinson,

1989), (Dardenne, Fickas and van Lamsweerde, 1991). The link with the concept of goal seems to have emerged naturally from research into software specification e.g. (Robinson, 1989) that relied on Artificial Intelligence (AI) artefacts, e.g. (Nilsson, 1971). Hence, most GORE methods use a vocabulary inherited from Artificial Intelligence, e.g. goals, agents, roles, objectives, constraints, and obstacles.

Many GORE methods have been defined over the years. The most prominent are: KAOS, GBRAM (Anton, 1996), i* (Yu, 1997), GRL (ITU-T, 2008), TROPOS (Mylopoulos, Kolp and Castro, 2001), ESPRIT CREWS (Rolland et al., 1998), and Goal-oriented Use Case (Cockburn, 2001).

Numerous goal types have been defined in these methods. Following is a very partial list of goal types: Achievement, maintenance, softgoal, feedback, satisfaction, etc.

Very quickly goals have become a central concept in RE. Zave and Jackson, for example, offer the following definition, "Requirements Engineering is about the satisfaction of goals" (Zave and Jackson, 1997). The call for papers of the Requirements Engineering conference series also strongly links requirements and goals, e.g. Requirements Engineering Conference 2004 (RE04):

Requirements Engineering (RE) is the branch of systems engineering concerned with the goals, desired properties and constraints of complex systems, ranging from embedded software systems and software-based products to large enterprise and socio-technical systems that involve software systems, organisations and people.

The focus on goals is understandable, not only because of the AI roots of GORE methods, but also because it relates with the goal-seeking organizational model prevalent in the neighbouring discipline of Information System (Checkland and Holwell, 1998). Many GORE methods take for granted this goal-seeking model. Modelling the organization, through Enterprise modelling, for example, is assumed to include goals, e.g. (Loucopoulos and Kavakli, 1995): "Enterprise modelling is about describing, in some formal way, a social system with its agents, work roles, goals, responsibilities and the like." (Nuseibeh and Easterbrook, 2000): "Enterprise modelling and analysis deals with understanding an organisation's structure; the business rules that affect its operation; the goals, tasks and responsibilities of its constituent members; and the data that it needs, generates and manipulates."

It is thus assumed that high-level enterprise goals can be gradually refined into requirements that can be assigned to the envisioned system (Dardenne et al., 1993), (Nuseibeh and Easterbrook, 2000). This refinement is most often done with the help of and/or goal trees inherited from (Nilsson, 1971).

Several RE researchers have reported problems with this assumption, e.g. (Anton, 1996), (Zave and Jackson, 1997), (Rolland et al., 1998), stating that goal discovery and goal refinement are not straightforward tasks, that enterprise goals are not necessarily a good starting point for goal refinement and that goal abstraction may lead to unrealistic or unwanted alternatives. The proposed remedies were to bound goal abstraction and refinement with the subject matter of the organization (Zave and Jackson, 1997), to use interview transcripts and organizational documents for goal discovery (Anton, 1996) and to use scenarios and goals reasoning together so that they inform one another (Rolland et al., 1998).

There have been surprisingly few attempts to link GORE to other organizational viewpoints, the only exception known to us is our own previous work (Regev and Wegmann, 2005). This is surprising because in their suggested roadmap for RE, Nuseibeh and Easterbrook (2000) suggest that, "RE is a multi-disciplinary, human-centred process." They further state that (Nuseibeh and Easterbrook, 2000),

RE needs to be sensitive to how people perceive and understand the world around them, how they interact, and how the sociology of the workplace affects their actions. RE draws on the cognitive and social sciences to provide both theoretical grounding and practical techniques for eliciting and modelling requirements

Nuseibeh and Easterbrook (2000) list the following disciplines as part of RE: Cognitive psychology, anthropology, sociology and linguistics. They proceed to prescribe that (Nuseibeh and Easterbrook, 2000):

There is an important philosophical element in RE. RE is concerned with interpreting and understanding stakeholder terminology, concepts, viewpoints and goals. Hence, RE must concern itself with an understanding of beliefs of stakeholders (epistemology), the question of what is observable in the world (phenomenology), and the question of what can be agreed on as objectively true (ontology). Such issues become important

whenever one wishes to talk about validating requirements, especially where stakeholders may have divergent goals and incompatible belief systems.

While prescribing the need to incorporate different epistemological, phenomenological and ontological systems, Nuseibeh and Easterbrook perpetuate the already prevalent organizational model in RE, namely goal achievement by repeating the main assumption of GORE methods about goal achievement.

By focusing on goal achievement, GORE methods have set aside all these aspects of epistemology and phenomenology and have applied a unique ontology supposed to be universal. In the recent years, KAOS and more so i* have become the main GORE methods with dozens of research papers devoted to them. Most of these are papers that use these methods in specific areas or that propose extensions to them. i* has been incorporated into the URN ITU-T standard (ITU-T, 2008) with seemingly hardly a questioning of its assumptions. Recently, a well received study of the i* graphical notation was published (Moody, Heymans and Matulevičius, 2009), noting that major improvement is needed in order to make i* user friendly. The study however remained at the graphical level and did not investigate the epistemological or ontological aspects of i*.

Sutcliffe and Maiden (1993) proposed a notable kind of goal in a paper that seems to have received little attention by GORE researchers. They proposed 6 classes of goals. One of the classes is called "feedback goals." They describe these goals as maintaining a desired state with a related tolerance range, spawning corrective actions when the state is considered to be outside the tolerance range (Sutcliffe and Maiden, 1993). Curiously, this class of goals has not been picked up by subsequent GORE research. We have ourselves added feedback into GORE research about 10 years later (Regev and Wegmann, 2004), (Regev and Wegmann, 2005) without noticing the significance of Sutcliffe and Maiden's feedback goal class at the time.

In the following section we propose an epistemological view for GORE methods that may help to alleviate some of their shortcomings. This view is an extension of our previous work (Regev and Wegmann, 2005). To clarify our discussion, we use Zave and Jackson's example of the development of a turnstile for a zoo (Zave and Jackson, 1997) as a running example. We reproduce their problem description here verbatim for the clarity of the discussion (Zave and Jackson 1997):

Requirements engineering is about the satisfaction of goals [...]. But goals by themselves do not make a good starting point for requirements engineering. To see why, consider a project to develop a computer-controlled turnstile guarding the entrance to a zoo [...].

If the engineers are told that the goal of the system is to deny entrance to people who have not paid the admission charge, they may decide that the goal has been stated too narrowly. Is not the real goal to ensure the profitability of the zoo? Should they consider other ways of improving profits, such as cutting costs? What if there is more money to be made by closing the zoo and selling the land? And what is the goal of profit? If the goal of profit is the happiness of the zoo owner, would religion or devotion to family be more effective? Obviously there is something wrong here. Almost every goal is a subgoal with some higher purpose. Both engineering and religion are concerned with goal satisfaction; what distinguishes them is their subject matter.

The engineers should be told, in addition to the goal, that the subject matter is the zoo entrance. This information should take the form of designations of phenomena observable at the zoo entrance, such as visitors, coins, and the action of entering the zoo. These designations circumscribe the area in which alternative goal satisfaction strategies can be considered, at the same time that they provide the basis for formal representation of requirements.

3 SURVIVAL AND REGULATION AS THE SOURCE OF GOALS

As we have seen, GORE methods seek to define the highest-level goals that are adequate for defining requirements for an envisioned system. Despite their important advancement, this is one aspect that the mainstream GORE methods (e.g. i*, KAOS) have not defined yet. So-called high-level goals are often described as strategic goals in GORE papers (see High-level goals entry in the previous section). It is therefore important to understand what is a strategic goal. If we define the strategic level as being geared toward the survival of the organization we have to define what survival means: We have shown

elsewhere, e.g. (Regev and Wegmann, 2005), (Regev and Wegmann, 2009), that survival can be understood in terms of the maintenance of relatively stable states, often called norms, that an observer can use to identify an organization. The zoo, for example, maintains a large set of norms for its animals, employees, visitors, owners, animal rights activists etc. For the animals, it maintains a stable place in which to live, get fed and cared for, and maybe reproduce. For the visitors the zoo maintains a stable place where they can view animals while being certain that they are well cared for. The location of the zoo, its name, its layout, the animals it houses are all aspects that change very little over time. It is quite probable that the zoo you visited as a child still exists and you can visit it with your children or grandchildren. It is also quite probable that most of the animals and employees who were at the zoo when you were a child are no longer there. But the zoo is still there. Hence the zoo maintains its identity despite changes in all of its components (e.g. employees, suppliers, management, animals, visitors).

Maintaining norms is a powerful motivator for action. The zoo's management might want to maintain its image as a modern facility and therefore maintain the fit with the state of the art in visitor management and feel that its entrance system is not up to date and needs to be replaced. It may believe that manual entrance control is not efficient enough and that by cutting costs on entrance control, it could use the money thus saved to better maintain its animals' health.

To maintain its norms relatively stable doesn't mean that the norms don't change at all. Organizations that survive over the long run make changes to their norms but in a very controlled way. This means that changes must be controlled for the organization to maintain its identity for its stakeholders. Consider what may happen if the zoo changed its location or name twice a year or more often. Would it be still recognized by its visitor? Would its employees continue to work there? What if it changed its mission from a zoo to a theme park, to a museum etc.? But of course the zoo also changes its norms, it may add new kinds of animals, playgrounds, activities, partnerships with other likeminded organizations. By saying that its norms should not change too much, we are actually placing bounds on how much change is acceptable. Remember the tolerance range defined by Sutcliffe and Maiden for feedback goals (1993).

What enables an organization to maintain and change its norms are the relationships it has with

individuals and organizations, both inside and outside the organization. This is a consequence of the open system model of organizations (Regev and Wegmann, 2005), (Regev and Wegmann, 2009). These relationships must themselves be maintained within very specific bounds (norms) for the organization to be able to leverage them for maintaining other norms. Thus, the zoo needs a continuous flow of visitors to maintain its funding and it places very strict bounds on what the visitors can and cannot do. For starters, visitors must in most cases pay their admission. They may pay a bit less if they are part of specific classes of people (seniors, children group members). Some visitors may not pay at all if the zoo has a real reason to want them to visit the zoo (e.g. VIPs, teachers accompanying a school class, very young children). Hence, there is a tolerance range for the admission price but it is not a wildly changing aspect. There are other bounds on what visitors are allowed to do in the zoo so that their presence will be beneficial for the zoo and not detrimental, e.g. it is probably forbidden to feed the animals outside of very controlled food dispensed by the zoo to specific animals, it is forbidden to mistreat animals etc. Visitors who do not conform to these bounds may be denied further admission to the zoo, despite the zoo's willingness to have as many visitors as possible. Applying different price schemes and denying entrance are likely to be translated into goals for the turnstile. Understanding the norms and where they come from helps analysts to understand the goals expressed by the stakeholders and discover non-expressed goals as well.

The study of the way norms are maintained stable is called regulation (or control). A feedback regulator (or controller), e.g. a thermostat, an automatic pilot, maintains a given state stable, e.g. temperature, course, by sensing the current state comparing it with the given state, and applying some action if the difference is above the tolerance level. Vickers (Vickers, 1968), (Vickers, 1987) proposed the concept of the appreciative system, by extending this model of a feedback regulation to human and organizational regulation. Vickers's appreciative system has three components (Vickers, 1968), (Vickers, 1987), (Regev, Wegmann and Hayard, 2011): Reality judgments, Value judgments and Action judgments. With reality judgments, some aspect of reality is singled out to be the study of attention. In value judgments, this reality judgment is matched to a category within which it is then compared to the norm (what ought to be). In Action judgment, some action might be taken to bring the

reality judgment closer to the norm. The three judgments function as a complete system so that change to one of them requires change to the others. Hence, the way we view and judge the world affect our actions and our actions affect our view and judgments (Vickers, 1968), (Vickers, 1987).

As we have seen before, the value judgment is done with some tolerance around the norm. In a simple automaton the information to be sensed, the norm and tolerances to be compared with and the kind of actions to be taken are all given by its designer. In an appreciative system, they are all subject to continuous change; nothing is set once and for all by a designer. Hence, an appreciative system creates its own dynamics, which an automaton does not.

Whereas the appreciative system creates its own judgments, it is nevertheless a rather stable construct, i.e. it creates its own norms. Vickers calls these norms readiness. Hence an organization has a readiness (a tendency) to see things in certain ways, to value them in certain ways and to act in certain ways. All these are rather stable in time. It is often the role of the analyst to try and shake these readinesses.

Linking the appreciative system and GORE aspects, we can see that maintenance goals can be an approximation of norms, beliefs can be an approximation of reality and value judgments, and achievement goals can be an approximation of action judgments.

From this point of view, GORE methods have concentrated on the third stage, action judgments (achievement goals), and neglected the other two, norms (maintenance goals), and reality and value judgments (beliefs). This is easily understandable if we consider that actions judgments are much more visible than reality and value judgments.

Because the three judgments of the appreciative system are interlocked, goals cannot be changed without changing the beliefs that justify them.

4 IMPROVING GORE METHODS

Based on the regulation view we proposed in the previous section, we identified a number of shortcomings common to most GORE methods. We explain them and provide pointers for addressing them.

Maintenance is Higher-level than Achievement. Although many types of goals have been defined in GORE methods, the most popular goal type has been

and remains the achievement goals; a goal that is to be achieved once and for all. The next most popular goal type is the softgoal. Both KAOS and GBRAM have introduced the concept of maintenance goal, a goal that “is satisfied as long as its target condition remains true” (Anton and Potts, 1998). Maintenance goals have not received much attention and remains largely unused. This is particularly unfortunate because maintenance goals have been identified as “high-level goals with which achievement goals should comply” (Anton and Potts, 1998).

As we have shown, the concept of maintenance goal is very useful to model norms. Consider some of the norms maintained by the zoo. A zoo just as any other organization is an organism that seeks some permanency and which takes actions to maintain this permanency. Thus, the zoo needs a continuous in-flow and out-flow of visitors, animals, feed, medicine, employees etc. Without these flows the zoo may not survive. The zoo is likely to take many actions to restore any one of these flows to the norm it depends on if the flow comes to be below a given tolerance level. These actions can be modelled as maintenance, softgoals, feedback and achievement goals.

If maintenance goals are augmented with tolerance levels (imported from feedback goals) they can be used to model norms.

The Concept of High-level Goal. While both why and how questions are encouraged by GORE methods, the how is much more prevalent in GORE publications. Most often, a so-called high-level goal is postulated to be strategic for the organization under analysis and is refined into subgoals.

For example, van Lamsweerde gives the following examples for “high-level, strategic concern”: “serve more passengers” for a train transportation system” and “provide ubiquitous cash service for an ATM network system.” (van Lamsweerde, 2001). It is not clear why these should be considered as high-level, strategic goals and how they can be satisfied. If the train transportation system serves a few more passengers, is this goal achieved? What will the system do next once this goal is achieved? What if this goal is never achieved? What would the ATM network system do once the ubiquitous cash service is provided? What are the criteria of achievement for a ubiquitous cash service? Looking closely at these goals, it becomes clear that they cannot be achieved once and for all but should be seen as maintenance goals, which are never fully satisfied. From the regulation viewpoint they therefore represent norms that may be considered essential for the survival of the system.

Identifying the norms that are considered essential for survival helps solve the unacceptable goal alternatives identified by Zave and Jackson. Identifying these strategic norms places the appropriate bounds on what is acceptable and not acceptable. For the zoo, these norms are the result of the relationships it maintains, between animals, visitors, employees, veterinarians, investors, buildings, donors, etc. Hence, rather than limiting the field of investigation to the objects and actions observable at the zoo entrance, the turnstile can be seen as an essential artefact in this regulation of relationships, even those that are not immediately observable at the zoo entrance, which will expand the field of investigation without crossing into subjects such as selling the zoo or religious devotion.

For example, asking what norms does the turnstile maintain may lead to answers such as that it regulates the in-flow of people, filtering between people who have the right to enter the zoo and those that don't. Knowing more about the norms maintained by the rest of the zoo, we can infer more about current and possible future norms. It is clear that the turnstile cannot maintain the expected norm if the zoo is not protected from rogue entrance elsewhere. The zoo must maintain this protection (probably in the form of a fence). Knowing that people can enter and leave the zoo only through the turnstile, we can see that the turnstile can maintain the number of people presently inside the zoo, which can help maintain security norms (preventing overcrowding, helping to insure complete evacuation in case of emergency and at closing time (another norm). Knowing about new norms of potential visitors, the turnstile can be made to accept cash, credit card and smartphone payments. It can allow quick admission of groups (a norm).

A more widespread use of maintenance goals as the highest-level goals that model an organization survival is therefore necessary.

Goal Refinement and Goal Abstraction. All GORE methods place goals in a hierarchy. Goal refinement is used to identify lower-level goals by asking how a given goal is achieved. Goal abstraction is used to identify higher-level goals by asking why a given goal needs to be achieved.

GORE methods have very good constructs for formalizing more or less predefined goals but goal discovery is still a challenge. Hence, despite arguments to the contrary (e.g. (van Lamsweerde, 2001)), GORE publications often describe tools for goal refinement but very little goal abstraction. The

problem may come from the difficulty in identifying the sought-out high-level goals.

Both goal refinement and goal abstraction create a rather simplistic model of human and organizational behaviour.

Consider goal abstraction, in Zave and Jackson's (1997) example of the zoo turnstile, the engineers ask themselves why questions and come up with higher level goals and alternatives. This is not usually the way requirements are identified. The engineers are supposed to ask the zoo employees or owners why they need a turnstile. In our experience, asking stakeholders why they need some solution that they identified is as likely to lead to answers such as, "because I want this other thing" as to answers such as "because I believe this." In the zoo example, asking the owners why they want to install a computer-controlled turnstile may yield answers such as:

- We want to have a more efficient admission system.

Or

- We believe that a computer-controlled turnstile will be more efficient than our manually controlled admission system.

Or

- The zoo in the neighbouring town recently installed a computer-controlled turnstile and are very happy with it.

These answers can be translated into higher-level goals (i.e. improve efficiency, match the competition) but this would mask the issue that this is just the owners' beliefs about their zoo. If these issues are captured as goals they are more likely to go unchallenged than if they are treated as beliefs.

Goal refinement can equally be improved by the use of beliefs. Remember that goals and beliefs are co-dependent as a result of the interlocking of the three judgments that form the appreciative system.

Consider asking the zoo owners how they would like to make their zoo more efficient. If they believe that the best way is to install a computer-controlled turnstile, they will ask for one. If they believe that cost cutting will yield better results, they will ask for cost cutting measures. If they think that both are needed, they will ask for both. Capturing the goal refinement as an and/or tree will result in the loss of these beliefs and more importantly in the opportunity to challenge them. For a more elaborate explanation of beliefs and their use in GORE methods, see (Regev and Wegmann, 2005). More research can be done on modelling reality and value judgments as beliefs.

A more widespread use of maintenance goals in conjunction with achievement goals and beliefs can help in modelling regulation and may help to better understand requirements. A beginning of a solution is proposed in our previous work (Regev and Wegmann, 2004).

5 RELATED WORK

Several conceptual studies of GORE methods have been published over the years, e.g. (Kavakli, 2002), (Kavakli and Loucopoulos, 2005). These studies assume a viewpoint from within the RE research paradigm. They do not ground their research in an external body of knowledge, which limits their explanatory power of goals. Moody et al. (2010) have studied the graphical language of *i** and its fit with users' understanding. We have proposed an explanation of goals based on General Systems Thinking and Vickers's work in (Regev and Wegmann, 2005).

Our work is similar in nature to Checkland and Holwell's conceptual cleansing (1998) of the field of information systems. Checkland (Checkland and Scholes, 1990) has worked extensively to popularize Vickers's work with Soft System Methodology (SSM). Ours is a very short description of Vickers's appreciative system. More elaborate descriptions are available in Vickers's writings (Vickers, 1968), (Vickers, 1987), and in (Checkland and Scholes, 1990), (Checkland, 2005), (Regev et al., 2011).

6 CONCLUSIONS

RE is about understanding peoples' desires and maybe designing some automated system to help them to obtain or maintain them. RE must therefore make the balance between what is desired and what is feasible. To understand what is desired, it is above all necessary to understand how people behave. GORE methods have made major contributions to the practice of RE but have modelled human and organizational behaviour in too simplified terms, mostly as goals to be achieved. In this paper, we have shown that goals can be seen as the visible part of regulation. Regulation models the way organizations (be they people or organizations) attempt to survive in a changing environment.

Regulation results in the establishment of norms, stable states that define the identity and therefore the survival of an organization. A long lasting

organization manages its internal and external relationships in a way that controls the changes to these norms but still allows them to change when needed. Looking at regulation rather than goals shifts the attention to the way people manage stability and change by managing relationships.

To understand people's desires it is important to analyse this more fundamental aspect of their behaviour, which is regulation rather than goals. RE has already defined most of the useful concepts needed for the study of regulation, e.g. maintenance goals, and beliefs. What is needed is a paradigm change from goals to regulation. This paradigm change will hopefully result in the more widespread use of these goal-oriented concepts.

A useful stream of research with which to connect, is General Systems Thinking, which has most of the necessary constructs to understand regulation in all kinds of systems, e.g. (Weinberg, 1975), (Weinberg and Weinberg, 1988), (Ashby, 1956). Organizational Semiotics can also be useful because it is the study of norms in organizations (Chong and Liu, 2002), (Shishkov, Xie, Liu, Dietz, 2002).

REFERENCES

- Anton, A. I., 1996. Goal-based requirements analysis. In *ICRE'96 Second International Conference on Requirements Engineering*, IEEE.
- Anton, A. I., Potts, C., 1998. The use of goals to surface requirements for evolving systems. In *ICSE 98 International Conference on Software Engineering*, IEEE.
- Ashby, W. R., 1956. *An Introduction to Cybernetics*, Chapman Hall, London.
- Checkland, P., Scholes, J., 1990. *Soft System Methodology in action*, Wiley, Chichester, UK.
- Checkland P, Holwell, S., 1998. *Information, systems and information systems - making sense of the field*. Wiley, Chichester, UK.
- Chong, S., Liu, K., 2002. A Semiotic Approach to Improve the Design Quality of Agent-Based Information Systems. In Liu, K., Clarke, R.J., Anderson, P.B., and Stamper, R.K. *Coordination and Communication Using Signs: Studies in Organizational Semiotics*, Kluwer.
- Cockburn, A., 2001. *Writing Effective Use Cases*, Addison-Wesley, Reading, MA.
- Constantine, L., 1995. Essential modeling: Use cases for user interfaces. In *ACM Interactions*, 2(2), 34-46.
- Dardenne, A., Fickas, S., van Lamsweerde, A., 1991. Goal-directed concept acquisition in requirements elicitation. In *IWSSD '91, Sixth International Workshop on Software Specification and Design*, ACM.
- Dardenne, A., van Lamsweerde A., Fickas, S., 1993. Goal Directed Requirements Acquisition. *Science of Computer Programming*, 20(1-2), 3-50.
- Dubois, E., 1989. A Logic of Action for Supporting Goal-oriented Elaborations of Requirements. In *IWSSD '89, Fifth International Workshop on Software Specification and Design*, ACM.
- Gause, D.C., Weinberg, G. M., 1989. *Exploring Requirements: Quality BEFORE Design*, Dorset House, N.Y.
- ITU-T, Telecommunication Standardization Sector of ITU, 2008. *User requirements notation (URN) – Language definition (Z.151)*.
- Kavakli, E., 2002. Goal-Oriented Requirements Engineering: A Unifying Framework. *Requirements Engineering*, 6(4), 237-251.
- Kavakli, E., Loucopoulos, P., 2005. Goal Modeling in Requirements Engineering: Analysis and Critique of Current Methods. *Information Modeling Methods and Methodologies*, 102-124.
- Loucopoulos, P., Kavakli, E., 1995. Enterprise Modelling and the Teleological Approach to Requirements Engineering. *Intelligent and Cooperative Information Systems*, 4(1), 45-79.
- Moody, D., Heymans, P., Matulevičius, R., 2010. Visual syntax does matter: improving the cognitive effectiveness of the i* visual notation. *Requirements Engineering*, 15(2), 141-175.
- Mylopoulos, J., Kolp, M., Castro, J., 2001. UML for Agent-Oriented Software Development: The Tropos Proposal. In *UML 2001, Fourth International Conference on the Unified Modeling Language*, Springer.
- Mylopoulos J., 2006. Goal-Oriented Requirements Engineering: Part II. Keynote Talk, *RE'06, 14th IEEE Requirements Engineering Conference*, Minneapolis, MS, September 2006. http://www.ifi.uzh.ch/req/events/RE06/ConferenceProgram/RE06_slides_Mylopoulos.pdf, accessed May 2011.
- Nilsson, N. J., 1971. *Problem Solving Methods in Artificial Intelligence*, McGraw-Hill.
- Nuseibeh, B., Easterbrook, S., 2000. Requirements engineering: a roadmap. In *ICSE '00, International Conference on The Future of Software Engineering*, ACM.
- RE04, 2004. *Requirements Engineering International Conference 2004*, <http://www.re04.org/>, accessed May 2011.
- Regev, G., Wegmann, A., 2004. Defining early it system requirements with regulation principles: the lightswitch approach. In *RE'04, 12th IEEE International Requirements Engineering Conference*, IEEE.
- Regev, G., Wegmann, A., 2005. Where do Goals Come From: the Underlying Principles of Goal-Oriented Requirements Engineering. In *RE'05 13th IEEE International Requirements Engineering Conference*, IEEE.

- Regev, G., Hayard, O., Gause, D. C., Wegmann, A. 2009. Toward a Service Management Quality Model. In *REFSQ'09, 15th International Working Conference on Requirements Engineering: Foundation for Software Quality* Springer.
- Regev, G., Hayard, O., Wegmann, A., 2011. Service Systems and Value Modeling from an Appreciative System Perspective. In *IESS1.1, Second International Conference on Exploring Services Sciences*, Springer.
- Robinson, W. N., 1989. Integrating multiple specifications using domain goals. In *IWSSD '89, Fifth International Workshop on Software Specification and Design*, ACM.
- Rolland, C., Souveyet, C., Ben Achour, C., 1998. Guiding goal modeling using scenarios. *IEEE Trans. Software Eng.*, 24, 1055–1071.
- Sutcliffe, A. G., Maiden, N. A. M., 1993. Bridging the Requirements Gap: Policies, Goals and Domains. In *IWSSD '93 7th International Workshop on Software Specification and Design*, IEEE.
- Shishkov, B., Xie, Z., Liu, K., Dietz, J. L. G., 2002. Using Norm Analysis to Derive Use Cases from Business Processes, In *OS 2002, 5th Workshop On Organizational Semiotics*.
- van Lamswerde, A., 2001. Goal-Oriented Requirements Engineering: A Guided Tour, in *RE'01, 5th IEEE International Symposium on Requirements Engineering*, IEEE.
- Vickers, Sir G., 1968. *Value Systems and Social Process*, Tavistock, London.
- Vickers, Sir G., 1987. *Policymaking, Communication, and Social Learning*, eds, Adams, G. B., Forester, J., Catron, B. L., Transaction Books. New Brunswick NJ.
- Weinberg, G. M., 1975. *An Introduction to General Systems Thinking*, Wiley, New York.
- Weinberg, G. M., Weinberg, D., 1988. *General Principles of Systems Design*, Dorset House.
- Yu, E. S. K., 1997. Towards modelling and reasoning support for early-phase requirements engineering. In *RE'97, Third IEEE International Symposium on Requirements Engineering*, IEEE.
- Zave, P. and Jackson, M., 1997. Four Dark Corners of Requirements Engineering, *ACM Transactions on Software Engineering and Methodology*, 6(1), 1-30.

DESIGNING FOR INNOVATION BY APPLYING ORGANIZATIONAL MODULARITY

Philip Huysmans

*Department of Management Information Systems, University of Antwerp, Prinsstraat 13, Antwerp, Belgium
philip.huysmans@ua.ac.be*

Keywords: Enterprise Architectures, Business Innovation and Software Evolution, Business-IT Alignment and Traceability.

Abstract: In volatile and customer-driven markets, the ability to innovate is a key success factor. Innovations have to be implemented at a steady pace to ensure business sustainability. However, the innovation process is only poorly understood. Therefore, many organizations and governments have difficulties stimulating and managing innovation. Several authors have proposed organizational modularity as a theoretical basis to understand and manage innovation. Their main argument is that a modular structure enables parallel evolution of different organizational modules. Consequently, innovations can be implemented without being limited by implementation aspects of other organizational modules. Modularity has been applied by various authors on different levels of the organization, such as products, processes, departments, and supporting IT systems. Moreover, enterprise architecture frameworks allow the modeling of different viewpoints of an organization. However, few organizations can be considered to have completely decoupled organizational layers. Consequently, dependencies between organizational modules on different enterprise architecture layers can heavily impact the ability to introduce innovations. In this paper, we demonstrate how modular dependencies impact enterprise architecture projects in two case studies. We then present a case study to illustrate how a modular dependency approach can be used to complement existing modeling approaches.

1 INTRODUCTION

Contemporary organizations are faced with rapidly changing environments. As a result, innovations need to be introduced in the organization in order to remain competitive in these markets. The introduction of innovations often impacts many different aspects of the organization. For example, Barjis and Wamba describe the impact on organizations caused by the introduction of RFID technology (Barjis and Wamba, 2010). Besides the adaptation of the business processes through, for example, BPR, it can be expected that other organizational artifacts need to be adapted as well. Data management systems need to be able to handle “enormous” data volumes (Barjis and Wamba, 2010), hardware infrastructure and software applications need to be purchased to handle accurate tracking, and customer acceptance needs to be handled by privacy commissions and marketing departments.

Enterprise architecture projects are frequently initiated to guide the change process needed to imple-

ment such innovations (Scheckerman, 2005). Enterprise architecture frameworks help to identify the different elements to which changes need to be applied. In such frameworks, models are created from different perspectives or viewpoints. Complexity from other viewpoints is abstracted away to be able to focus on the concerns of a specific viewpoint. While this approach aids understandability of different aspects of the organization, it can lead to unexpected results when artifacts from different viewpoints impact each other. This can be observed in the case of RFID, where changes to processes can be hindered by the data structures on which they operate. When innovative processes are designed using a BPR approach, but data systems are unable to support the required data types, implementation of the new processes will be problematic. Possibly, the redesigned processes will then need to be altered in order to be supportable. Consequently, adapting processes can be impacted by the concrete data implementation, which is not completely visible in the models which are created in a

process viewpoint. While aspects of this implementation are not visible in these models, they do impact the way these models can be used, and therefore need to be considered.

In this paper, we argue that *modular dependencies* can be used to explicitly identify such impacts. This approach can be applied complementary to existing enterprise architecture modeling. We introduce relevant literature for this approach concerning modularity and enterprise architecture frameworks in Section 2. We then present three case studies to illustrate the application of modular dependencies in organizations which are faced with innovations which require changes in several viewpoints. In Section 3, we present two case studies to demonstrate the importance of modular dependencies between layers in the enterprise architecture. In the first case study, we describe how dependencies between enterprise architecture layers limit the ability to react to market changes in a public broadcasting company. In the second case, we elaborate on a previously published case study to show the importance of eliminating such dependencies in order to design a structure which allows the implementation of innovations. In Section 4, we demonstrate in more detail how an approach using modular dependencies can be applied during an enterprise architecture project. Finally, we present our conclusions on these cases in Section 5.

2 RESEARCH BACKGROUND

In this section, we introduce a necessary background of modularity and enterprise architecture research literature.

2.1 Modularity

Organizational modularity recently receives much attention in both research and practice. Campagnolo and Camuffo provide a literature overview of 125 management studies which use the modularity concept (Campagnolo and Camuffo, 2010). They define modularity as “an attribute of a complex system that advocates designing structures based on minimizing interdependence between modules and maximizing interdependence within them” (Campagnolo and Camuffo, 2010). They argue that organizational artifacts such as products, production systems, and organizational structures can be regarded as modular structures. A characteristic of a perfect modular structure is that it allows parallel evolution of different modules (Baldwin and Clark, 2000). However, dependencies between the modules of such a structure limit

the autonomy of the individual modules. Baldwin and Clark state that “[b]ecause of these dependencies, there will be consequences and ramifications of any choice” made during the design of the artifact (Baldwin and Clark, 2000). A design choice for a given parameter can limit or affect the possible design choices concerning other parameters. In traditional modularity approaches (e.g., product modularity), dependencies between and within modules are visualized by Design Structure Matrices (DSM). In a DSM, an artifact is described by a set of design parameters. The matrix is then filled by checking for each parameter by which other parameters it is affected and which parameters are affected by it. The result is a map of dependencies that represent the detailed structure of the artifact. An example design structure matrix is shown in Figure 1. Dependencies are represented by an “x”. The intersection of identical design options is marked with a “.”. Consider the design dependency which is represented by the “x” in the intersection of the column of design option A2 and the row of design option A1. This signifies that design option A2 influences design option A1: the design decision for design option A1 will be dependent on the decision taken for design option A2. This dependency does not break the modular structure of the artifact, since design options A1 and A2 both belong to the same module. Now consider the dependency of design option B1 on design option A2. Since these design options belong to different modules, it can be concluded that these modules are directly dependent on each other. Therefore, this dependency does violate the modular structure. Indirect or chained dependencies can occur as well. While design option B2 does not seem to affect any design options of module A, it does affect design option B1. As we discussed before, design option B1 does affect design option A1. Therefore, a so-called chained dependency exists between design options B2 and A2.

		Module A		Module B	
		Option A1	Option A2	Option B1	Option B2
Module A	Option A1	.	x		
	Option A2		.	x	
Module B	Option B1			.	x
	Option B2				.

Figure 1: An example Design Structure Matrix.

Originally, DSMs were used to model modular products. In later publications, a DSM was also applied to model organizational departments (Baldwin

and Clark, 2003). This indicates that similar tools can be used for analyzing modularity on various levels. In this paper, we will focus on modularity on the organizational and software levels. Organizational modularity focuses on other artifacts than product modularity within the same organization. Research on this level has been performed by, for example, Galunic and Eisenhardt (Galunic and Eisenhardt, 2001). Galunic and Eisenhardt consider organizational divisions as modular organizational building blocks. These divisions, which have independent decision power, cost structures and profit responsibility, are a combination of capabilities and charters. Charters represent the task, market and customer a division is concerned with. Charters need to be able to change as markets evolve. By dynamically attributing these charters to organizational divisions, a flexible organization is created which can adapt to changing market conditions. This kind of modularity therefore enables the flexibility required on a business level. On the software level, modularity is used to achieve a flexible structure as well. For example, Parnas argued that a modular decomposition in software systems should be made to isolate the impact of changes (Parnas, 1972). When the impact of a change remains within a module, changes can be applied to individual modules without requiring changes in the rest of the system. While modularity is applied to both the organizational and IT level by various researchers, most research project focus on a single level. However, interactions between a modular approach on the organizational and the software level remain an important issue. Given the high dependence on IT systems, it is important that a change on the organizational level (e.g., changing charters) can be handled by the supporting IT systems as well. Otherwise, the inability to change the IT systems will restrict the ability to change organizational modules. Indeed, modularity needs to be considered as a *relative* attribute of complex systems (such as organizations), meaning that within a single artifact, different levels of modularity can exist (Simon, 1962).

2.2 Enterprise Architecture

Enterprise Architecture (EA) frameworks provide insight to the structure of organizational goals, divisions, and supporting IT systems. By specifying separated viewpoints on organizational artifacts, an overview is provided of specialized models created by different stakeholders (The Open Group, 2009). Most enterprise architecture frameworks use a top-down perspective. They start by defining business goals and a high-level artifacts to realize these goals

(e.g., an organizational structure). Based on these artifacts, lower-level artifacts are defined which offer services to support the business level. The following architectural levels are usually identified (Schöenherr, 2008):

- Business layer
- Functional layer
- Information layer
- Infrastructure layer

For example, TOGAF suggests to use an iterative process consisting of eight phases to develop an enterprise architecture. Based on the business goals which are defined in the first phase, different architectures need to be developed in the following order: business architecture (second phase), information systems architecture (third phase), and technology architecture (fourth phase). These architecture correspond to the defined layers. The business architecture is defined on the business layer. The information systems architecture consist of both the functional and information layer. The technology architecture is addressed on the infrastructure layer. This approach assumes that supporting services can be created straightforwardly based on business requirements. However, when implementation-focused approaches are described (e.g., Service-Oriented Architecture), the existing infrastructure is often an important restriction on the services which can be provided. Consequently, the focus of such approaches is often on subjects such as legacy integration. Approaches which explicitly integrate an implementation approach with business goals (e.g., SOMA) therefore use a middle-out approach (Arsanjani et al., 2008). Such an approach takes into account the possibilities of the supporting systems, and attempts to find the best solution for the business requirements. Indeed, the use of additional methods illustrates that a top-down approach in enterprise architectures needs complementary approaches to deal with complex environments. In a literature review on enterprise architectures, Lucke et al. identify several issues which motivate this need (Lucke et al., 2010). First, *complexity* is referred to as an underestimated issue. Not only the complexity of the models themselves, but also the dependencies between the different layers remain problematic. Second, *rapidly changing conditions* imply that a top-down specification of an enterprise-wide architecture can become out-dated before it is even implemented. Third, a top-down specification of the architectural layers results in issues regarding *scoping of architectural descriptions*. Rather than being straightforward, the identification of organizational and technical services to support business requirements is often considered to be

problematic. Therefore, a complementary approach is required to represent the impact of lower-level layers on higher-level layers.

2.3 Modularity in Enterprise Architectures

Several authors link the explicit decomposition of viewpoints in enterprise architecture frameworks to the ability to independently change artifacts. This indicates how enterprise architecture frameworks can be linked to the modularization of organizations. For example, business modularity is considered to be the highest level of enterprise architecture maturity (Ross and Beath, 2006). On this maturity level, the role of IT in an enterprise architecture is to “provide seamless linkages between business process modules” (Ross and Beath, 2006). Such business process modules allow “strategic experiments that respond to changing market conditions” (Ross and Beath, 2006). Based on a practitioners survey, it seems that many business users indeed expect an enterprise architecture to enable their ability to change in response to market conditions (Schekkerman, 2005). A modularity perspective can aid to specifically focus on the issues specified by Lucke et al. (Lucke et al., 2010), which have been mentioned above: identifying modular dependencies reveals how complexity is introduced when modules are added; the parallel evolution of modules limits the impact of rapidly changing conditions; the explicit specification of modules aids scoping architectural descriptions.

Consequently, independence between enterprise architecture layers should be achieved by defining business components and standardized interfaces based on the artifacts modeled in the different layers. This shows that an explicit focus on the coupling of artifacts from different enterprise architecture layers is required to gain insight in the kind of changes which can be supported. In the case studies we performed, dealing with this kind of coupling adequately often seemed to be an important success factor. An approach which explicitly shows these impacts can therefore help to improve insight in the change process.

3 CASE STUDY OBSERVATIONS

In this section, we illustrate the occurrence and impact of modular dependencies between artifacts from different levels in an enterprise architecture. The presented case studies are part of a larger group of case studies which aim to apply insights from modularity

to enterprise architectures. These case studies have been performed adhering to the case study methodology (Yin, 2003). Given the ambiguity of the definition of enterprise architecture in both practice and academic literature, and the large scope of enterprise architecture frameworks, it is difficult to clearly distinguish between the research subject itself and the organizational environment. Therefore, we selected an exploratory case study approach, since it is well suited for research goals where the boundaries between phenomenon and context are not clearly defined (Yin, 2003). In the various cases, we have used the key informant method to identify informants who were highly knowledgeable about and involved in the enterprise architecture projects. The primary mode of data collection consisted of face-to-face interviews. In preparation for these interviews, various documents (e.g., documentation and presentation materials, documents on the organizational structure) have been consulted to gain an initial understanding of the organization and the project itself. A case study protocol was crafted, including an initial set of questions. These questions concerned various topics (e.g., architecture definition, expected benefits and barriers), in order to obtain a thorough description of the project. Follow-up questions took place via e-mail. During the interview, additional sources of evidence were collected, such as articles and internal documentation. The interview was digitally recorded and transcribed for future reference.

3.1 Public Broadcasting Company

In a first case study, we observe a public broadcasting company (PBC) which is faced with changing customer demands. We focus on the division which is responsible for broadcasting news journals. Traditionally, this organization offers radio and television transmissions, which follow a clearly defined schedule. The radio and television business units import news items from different sources, such as feeds from external agencies, or items made by reporters. Items can be imported using physical tapes, digital files or through satellite transfers. The items are then edited and transmitted in the form of a news journal. However, since the introduction of the internet, customers demand personalized and real-time access to transmissions. Therefore, the content of news journals needs to be approached differently. The PBC decided to create a dedicated business unit to create an online channel, next to the existing radio and television units. This new business unit could reuse content from both the radio and television units, and create dedicated online news items as well. Adding the online channel

was considered to be a necessary strategic move in order to serve a changing market. For the PBC, the radio, television and online business units are therefore situated on the *business layer* of their enterprise architecture.

Adding this additional business unit posed serious problems due to the supporting structure of the PBC on the lower enterprise architecture levels. For example, accessing existing radio and television items proved to be complex, since they were stored in specialized applications. For every import channel (tape, digital or satellite), different applications needed to be used. This is a direct result of the principles used to develop the application portfolio, which was located on the *functional layer* in the enterprise architecture. As a general principle, the organization always selects best-in-class software solutions for specialized media editing. This principle ensured the most efficient editing process. However, this results in an application portfolio which is not well integrated. Employees with specialized competences are required to operate these software packages. Therefore, in order to reuse audio and video fragments for the online channel, employees with these competences needed to be made available for the online business unit. This resulted in a duplication of skill sets, which was not efficient. The inclusion of employees with the competence to use the specialized software packages was not foreseen when the new business unit was defined. However, the coupling between software packages and employee competence on the information layer necessitates this inclusion. Put differently, the required competences defined on the business layer need to be adapted to account for a dependency on the functional level. It should be noted that coupling on the functional layer is solved by adapting the artifacts which are defined on the business layer. Consequently, the ability to take strategic decisions to serve an emerging market are impacted by decisions made on the functional layer. While the principle to select best-in-class applications may be justified for this sector and the performance of the organization, the lack of integration which it causes and the restrictions it places on business flexibility need to be understood as well. When considered as a modular structure, this dependency can be represented using a DSM. The DSM is presented in Figure 2. It shows how a design parameter from the applications, which are part of the functional layer, affect the organization chart design parameter of the business unit, which is part of the business layer.

Since the PBC was unable to motivate the costs of specialized employees for the new channel, a structural solution was proposed. The dependency on the

		Business unit		Application	
		Organization chart	Charter	Competence	Platform
Business unit	Organization chart	.		x	
	Charter		.		
Application	Competence			.	
	Platform				.

Figure 2: The DSM for PBC.

required application capabilities forces the the organization to deal with concerns from the functional layer on the business layer, i.e., the different handling of tape, digital and satellite items. This dependency was removed by developing an abstraction layer on top of the functional layer, which provided only the functionality needed on the business layer. This abstraction layer was defined based on concepts known in the business layer: the basic entity on this abstraction layer was a news item. A news items entity abstracts from the method which was used to import the source material. Therefore, it served as a kind of interface to the specialized software applications. For every application, audio and video fragments can be extracted, together with the required meta-data, in a uniform way. The specific functionality of the editing programs is not available through this interface. In order to use the editing functionalities, specialized competences remain necessary. However, the need for employees with these competences is now not imposed on the business layer, only on the functional layer.

This introductory case shows that decisions taken on the functional layer of enterprise architectures can make changes to the business layer more complex. By adequately encapsulating the complexity from the functional layer, dependencies between the different layers can be removed, and necessary services can still be offered. As a result, decisions on the business layer only need to deal with the inherent complexity imposed by the business environment, instead of dealing with complexity originating from the supporting layers as well. In this case, the number of modules is low and the dependency which limits the business layer is clear. However, this case provides a clear illustration of the issue we address in this paper.

3.2 Gas Flow Manager Company

In a second case study, we observed an enterprise architecture project at a gas flow manager company. The company offers gas transport services to its cus-

tomers based on a grid consisting of entry points, nodes, and pipelines connecting the nodes. The functional and information layer of the enterprise architecture had to be rebuilt after the company was separated from a gas trading company. The legislation concerning liberalization of the energy market demanded this separation, as the company had to offer its gas transport services to other gas trading companies as well. Prior to the architectural redesign, IT was generally considered to be a bottleneck during the implementation of changes by business users. The new architecture needed to be able to respond better and more quickly to changing business requirements and had to be understandable for business users, so they could more realistically estimate the impact of the changes they requested. Therefore, it was decided to clearly align the functional architecture with a high-level enterprise architecture model on the business level. This model needed to be constructed in such a way that the *stable* operation and stakeholders of the organization are represented. If the subsequent changes required changes in the high-level model, it would not be useful as a basis for the functional architecture. The issue of rapidly changing conditions has indeed been addressed as an important enterprise architecture issue by Lucke et al. (Lucke et al., 2010). Put differently, business changes need to be attributed to the implementation of the model elements, not on the nature of the elements itself. The model which was designed is shown in Figure 3. The modeled activities represent generic descriptions of the business-level construction of the organization (e.g., *Define commercial services*). In order to achieve a well-aligned business and functional architecture, a separate application has been developed to support the scope of exactly one activity in this model.

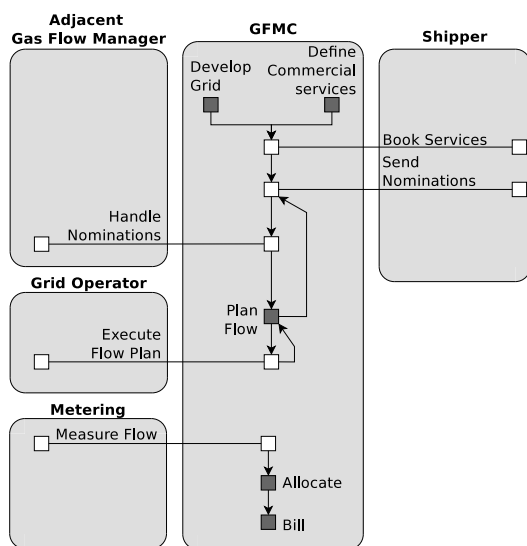


Figure 3: Stable GFMC model.

In a previous publication, we focused on a repeatable and reproducible method to design such models (Huysmans et al., 2010), and the benefits which can be achieved with such an approach. That description was limited to the top layers of the enterprise architecture, i.e., the business and functional layers. In this paper, we focus on the preconditions for applying this approach, which are situated on the functional and information layers. For this approach to work, the specification the functional layer may not be restricted by the information layer. Modular dependencies between artifacts from the information layer and artifacts from the functional layer can be used as an indication for such restrictions. Before the organizational split, applications had separate data sources, which impacted the specification of application scope. Put differently, dependencies existed between the functional and information layers. Because the semantic differences between entities in these data sources, it was not feasible to design applications which relied on data from different sources without creating complexity within the application. In the enterprise architecture maturity model, this indicates an architecture which is not correctly modularized (Ross and Beath, 2006). For example, the entity “customer” existed in different data sources. In some sources, this entity referred to end consumers of the gas which is transported. In other sources, it contained the customers of the gas flow manager company (i.e., the gas traders). As a result, the scope of applications was defined based on the scope in the data sources. During the enterprise architecture project, a glossary was developed iteratively with business users from different areas to ensure consistent terminology across the organization. The glossary defines the entities and their relationships to each other. Based on this glossary, a database scheme was developed and imposed on the different data sources. Consequently, the dependency of the application scope on the data source scope was resolved. The coupling between the functional and information layer was thus removed. As a result, the modules on the functional layer could be designed to be well-aligned with the business layer, without being restricted by aspects from the information layer.

Moreover, the solution applied in this case follows the solution presented in modularity theory. In order to resolve modular dependencies, modularity theory proposes to specify an *architectural rule*. Such a rule limits the design freedom for the implementation aspects of modules by prescribing a certain design choice. In this case, the glossary limits the allowed interpretation of, for example, the “customer” entity. As a result, no conversions need to be performed to ensure a correct use of data entities. Other

solutions are possible to resolve dependencies as well. Currently, we focus on the ability of modular dependencies to identify coupling between modules of different architectural layers as a cause for restrictions on higher-level layers. In future research, we elaborate on different methods to deal with this kind of coupling. However, the current case illustrates how a modularity approach can be complementary to enterprise architecture frameworks. While the different applications and data sources can be represented in an enterprise architecture, no indications of the coupling between the different viewpoints can be represented. When we consider the applications and data sources as modules, we can use a Design Structure Matrix (DSM) to represent aspects of the module implementation which affect each other.

4 APPLICATION OF MODULAR DEPENDENCIES

In the previous cases, we illustrated the relevance of modular dependencies in enterprise architecture projects. We now apply the insights gained from the analysis described above to a case study in a governmental organization. The mission of the organization is to introduce and implement e-government solutions. To achieve this goal, it undertakes projects in the field of back-office reengineering, and tries to leverage these improvements by supporting projects with governmental partners. In this paper, we focus on a project that improves the way data from various sources is used in governmental processes. We will refer to this project as the Data Usage in Governmental Processes (DUGP) project. Similar to the description in Section 3.2, the structure of the data sources limits the design of governmental processes. Because of the political situation, different data sources are controlled by different governmental entities, which belong to governments on different levels (e.g., federal, regional, local level). As a result, different implementations exist for a large amount of design parameters. For example, the data delivery design parameter may be implemented by an online web interface, an FTP transfer, or through web services. Consider the partial Design Structure Matrix represented in Figure 4, which has been developed to describe the situation before the DUGP project. In this DSM, multiple design parameters are considered simultaneously. Compare this to the coupling in the GFMC enterprise architecture, where we focused on a single design parameter (i.e., data semantics).

The “x”-es with the grey background represent (1) the dependency of the data retrieval design parameter

		Process			DS			
		Throughput	Data Retrieval	Data Syntax	Capacity	Data Delivery	Data Syntax	Data Dictionary
Process	Process Throughput	.	x		x			
	Data Retrieval	.	x		x	x		
	Data Syntax		.	x			x	x
	Data Dictionary		x	.			x	x
DS	Capacity				.	x		
	Data Delivery				x	.	x	
	Data Syntax					x	.	x
	Data Dictionary						x	.

Figure 4: DSM before the DUGP project.

of the processes on the data delivery design parameter of the data sources, and (2) the dependency of the data syntax used in the processes on the data syntax used in the data sources. Consider the impact of these two design parameters in the following example. A process to request construction premiums requires personal data of the citizen requesting the premium (from data source A) as well as geographical data of the construction site (from data source B). Since the databases which contain the personal and geographical data are not integrated or standardized, various conversions between the implemented data parameters of these data sources may be necessary. Suppose that the information required from data source A needs to be obtained by invoking a single web service call, providing the address from the end user using four data fields (street name, street number, bus number, city name). In order to query geographical information in data source B, a request file containing two data fields (street name and street number, and ZIP code) has to be transferred using the FTP protocol. Since the construction premium process depends on both data sources, it needs to be able to communicate using two different versions of address data syntax, and two different versions of data retrieval technology. If the construction premium process needs to be changed, and an additional data source is required, the process owner needs to be aware of the data syntax and data retrieval method offered by the new data source. Moreover, when changes are made to these implementation aspects of data sources, additional versions need to be supported by the processes. The resulting complexity of these conversions is a barrier for the use of these data sources. Moreover, this example shows that different design parameters are intertwined in a certain implementation. As a result, it is hard to resolve these dependencies individually. The goal of the DUGP project is therefore

to eliminate these dependencies at once in order to reduce the complexity of using data sources in governmental processes. However, the solution which is suggested by modularity theory, i.e., the definition of architectural rules, was not feasible because of the governmental structure. The different data sources are controlled by different governmental entities, who can decide independently on the implementation of design parameters. Declaring an architectural rule for a design parameter therefore requires an agreement between all governmental units responsible for a data source. However, since most units rely heavily on legacy systems to provide data services, changes to the implementation of design parameters are not easily realized. Therefore, it is difficult to reach such an agreement if an organization which can impose rules to these governmental units is not in place.

Consequently, the e-government organization developed a platform to consolidate data sources and provide uniform data access. Similar to the PBC case discussed in Section 3.1, an abstraction layer was developed to offer the required functionality without exposing the complexity of the layer offering these services. This abstraction layer provides services from the information layer to governmental processes, which are considered to be on the functional layer. The platform is based on two existing data sources from the federal government. Data from these data sources will be augmented with data available in data sources from other governments (e.g., geographical data, which is offered by regional governments). The first data source which is used focuses on data concerning organizations. In this data source, data such as registration number, official addresses and legal statute of enterprises can be obtained. We will refer to this data source as the Data Source for Organizations (DSO). The Federal Public Service Economy is responsible for this data source. The second data source offers data concerning individuals. It refers to data such as employment and social status of citizens. We will refer to this data source as the Data Source for Individuals (DSI). This data source is governed by a separate organization created by the federal government. The platform will maintain this distinction, and offer its data services grouped in an Enhanced Data Source for Organizations (EDSO) and an Enhanced Data Source for Individuals (EDSI).

Since the EDSO relies on the DSO for its data, and the EDSI relies on the DSI, they need to consider the implementation of these data sources. Many implementations of design parameters are quite different. For example, the DSI has webservices available to query its data. As a result, these webservices can be used to develop webservices in the EDSI. These

webservices are not directly offered in their original form. Instead, a facade pattern is used. This enables the creation of a uniform web service syntax throughout the platform. Otherwise, a dependency on the data syntax design parameter would be introduced. In contrast, the DSO has no webservices available. It is a mainframe which operates using batch requests. Therefore, a copy is made from the original DSO every night. This copy is then augmented with data from other governmental authorities, and used as a central database on which the services from the EDSO are provided. In order to simplify data access, the new platform provides three data delivery methods which will be available for all data sources: data repositories, an online application and webservices. Customized data repositories are large data files, which are copied to the process owner. After this initial data provision, automatic updates are transferred when data changes. These repositories are offered to enable process owners to incorporate the data from the new platform in their processes, without having to implement a webservices-based data access. Since many organizations are accustomed to using their own data sources in their processes, a customized data repository can be implemented without many changes in the processes. However, the unauthorized data sources which have been collected by the organizations themselves will then be replaced with authentic data. The online application allows for manual consultation of the data with a much smaller granularity: instead of a single large data file, only the result of a single query is returned. The same result can be obtained automatically through the use of webservices. Webservices offer the same data granularity, but can be implemented to automate processes.

From a modularity perspective, the platform can be considered as an additional module to eliminate dependencies between data and process modules. The DSM for the DUGP project is shown in Figure 5. When comparing DSM of the platform in Figure 5 with the DSM in Figure 4, we can conclude that some of these dependencies are indeed eliminated. Consider for example the data syntax and data delivery. We included an empty grey background to mark the previous existence of these dependencies. The syntax of webservices offered by EDSI is decoupled from the naming conventions of the DSI by using a facade pattern. As a result, naming conventions can be kept internally consistent with custom-built webservices for EDSO. The data syntax can be considered as an architectural rule which is maintained by the e-government organization. By adhering to this data syntax, process owners no longer need conversions between different data syntax versions. Another example is the data de-

		Process				Platform			DS				
		P. Throughput	Data Retrieval	Data Syntax	Data Dictionary	Capacity	Data Delivery	Data Syntax	Data Dictionary	Capacity	Data Delivery	Data Syntax	Data Dictionary
Process	Process Throughput	.	x			x							
	Data Retrieval	x	.	x		x							
	Data Syntax	x		.	x				x				
	Data Dictionary				.					x			
Platform	Capacity					.	x			x	x		
	Data Delivery					x	.	x		x			
	Data Syntax						x	.	x			x	
	Data Dictionary				x		x	.			x	x	
DS	Capacity									.	x	x	
	Data Delivery									x	.	x	
	Data Syntax									x	.	x	
	Data Dictionary									x	.		

Figure 5: DSM of the DUGP project.

livery design. In data sources from the platform, data can be provided through customized data repositories, an online application or webservices. Here, a single design option has not been selected, but the consistent offering of all data delivery techniques allows process owners to implement a single design for data delivery. Again, process owners no longer depend on the specific data delivery technique of the individual data sources. Consequently, it seems that the platform aids to decouple the process owners from the design decision of the data sources.

However, as stated by Baldwin and Clark, eliminating all dependencies in a modular structure is not a trivial task (Baldwin and Clark, 2000). Consider the design option *process throughput* in the case of an automated process. Our respondents indicated that the number of processes which can be supported is limited by, amongst others, the capacity of the data delivery implementation. In Figure 5, this is visualized by the “x” where the column of the capacity of the platform intersects with the row of process throughput. A possible data delivery implementation in the platform are webservices. As described above, webservices from the platform can be either custom-built by the e-government organization (e.g., webservices for EDSO), or can be part of a facade-pattern, calling underlying webservices (e.g., webservices for EDSI). The custom-built webservices operate on a local database. Consequently, their capacity is limited by the servers of the e-government organization itself. However, webservices which are part of the facade pattern are dependent on the capacity of the underlying services of DSI. Based on the implementation, issues with webservice capacity need to be discussed with the e-government organization

or with the organization responsible for the original data source. The difference between the implementation of webservices in the platform is based on the available data delivery techniques from the original data sources. In Figure 5, this is visualized by the “x” where the column of the data delivery of a data source intersects with row of the capacity of the platform. It therefore seems that an unexpected dependency can be identified: when the platform is used, the process throughput design decision is impacted by the data delivery technique of the original data source. When the data delivery of DSO is changed (e.g., webservices become available) and used by the platform, process performance may be impacted. This is an example of a chained design dependency which propagates through the design structure matrix. Such dependencies are difficult to trace and to account for in change projects.

5 CONCLUSIONS

In this paper, we explored a concrete application of modularity on the organizational level. We have shown that by modeling modular dependencies, interactions between layers in enterprise architecture models can be represented. Such interactions are not explicitly focused on in enterprise architecture frameworks. Therefore, this approach is complementary to existing frameworks. These frameworks usually focus on the top-down specification of different viewpoints. However, these viewpoints can not be considered to be independent from each other in complex organizations. As a result, one needs to be aware of the

impacts and restrictions imposed by lower-level layers during a top-down specification of enterprise architecture models. Therefore, research on this subject should be based on observations in real-life case studies instead of on theoretical examples. We focused on the restrictions of modeling artifacts on higher-level layers based on dependencies on the implementation of design parameters of lower-level artifacts. First, we demonstrated how this effect occurs in the PBC case study. Second, we discussed how the elimination of such dependencies can be a prerequisite in successful enterprise architecture projects. We illustrated this prerequisite in the context of the GFMC case study, which was published earlier. Moreover, we explored the applicability of dealing with modular dependencies as suggested by modularity literature. This solution implies the definition of architectural rules to limit the implementation possibilities of design parameters. Consequently, artifacts which are dependent on these design parameters can assume that a fixed implementation will always be supported. Third, we applied the insights from the observations in these case studies more concretely to the DUGP project. We showed that a DSM can be used to represent relevant issues for the enterprise architecture project as modular dependencies. This perspective allows to objectivate the issues which are resolved by the project. Moreover, a structured analysis can lead to the discovery of remaining issues after the project. Remaining issues can be unresolved dependencies, or newly introduced dependencies. This was illustrated by identifying the *capacity* design parameter as a chained dependency.

Future research needs to be conducted to gain insight on how modular dependencies on this level can be dealt with. In some cases, the definition of architectural rules seems appropriate. However, instead of choosing a single implementation option, it has already been observed that a consistent offering of multiple implementations can be required as well. Moreover, in some cases, imposing architectural rules does not seem feasible, because of the organizational structure. In the DUGP project, it was impossible to impose architectural rules to different organizational units. Therefore, an additional module was added. However, the introduction of new dependencies shows that a structured approach is required to adequately resolve modular dependencies.

REFERENCES

- Arsanjani, A., Ghosh, S., Allam, A., Abdollah, T., Gariapathy, S., and Holley, K. (2008). Soma: a method for developing service-oriented solutions. *IBM Syst. J.*, 47:377–396.
- Baldwin, C. Y. and Clark, K. B. (2000). *Design Rules, Volume 1: The Power of Modularity*, volume 1 of *MIT Press Books*. The MIT Press.
- Baldwin, C. Y. and Clark, K. B. (2003). The value, costs and organizational consequences of modularity. Working Paper.
- Barjis, J. and Wamba, S. F. (2010). Organizational and business impacts of rfid technology. *Business Process Management Journal*, 16(6):897–903.
- Campagnolo, D. and Camuffo, A. (2010). The concept of modularity in management studies: A literature review. *International Journal of Management Reviews*, 12(3):259–283.
- Galunic, D. C. and Eisenhardt, K. M. (2001). Architectural innovation and modular corporate forms. *The Academy of Management Journal*, 44(6):1229–1249.
- Huysmans, P., Ven, K., and Verelst, J. (2010). Designing for innovation: using enterprise ontology theory to improve business-it alignment. *Proceedings of the 1st International Conference on IT-enabled Innovation in Enterprise*, pages 177–186.
- Lucke, C., Krell, S., and Lechner, U. (2010). Critical issues in enterprise architecting - a literature review. In *AMCIS 2010 Proceedings*.
- Parnas, D. L. (1972). On the criteria to be used in decomposing systems into modules. *Communications of the ACM*, 15(12):1053–1058.
- Ross, J. and Beath, C. M. (2006). Sustainable it outsourcing success: Let enterprise architecture be your guide. *MIS Quarterly Executive*, 5(4):181–192.
- Schekkerman, J. (2005). Trends in enterprise architecture: How are organizations progressing? Technical report, Institute For Enterprise Architecture Developments.
- Schönherr, M. (2008). Towards a common terminology in the discipline of enterprise architecture. In *ICSOC Workshops'08*, pages 400–413.
- Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*, 106(6):467–482.
- The Open Group (2009). The open group architecture framework (togaf) version 9. <http://www.opengroup.org/togaf/>.
- Yin, R. K. (2003). *Case Study Research: Design and Methods*. Sage Publications, Newbury Park, California, 3rd edition.

ASSESSING THE BUSINESS RISK OF TECHNOLOGY OBSOLESCENCE THROUGH ENTERPRISE MODELLING

Cameron Spence

*Capgemini, 1 Forge End, Woking, Surrey, U.K.
cameron.spence@capgemini.com*

Vaughan Michell

*Informatics Research Centre, Henley Business School, University of Reading, U.K.
v.a.michell@reading.ac.uk*

Keywords: Enterprise architecture, Obsolescence, Risk, Modelling.

Abstract: The problem of technology obsolescence in information intensive businesses (software and hardware no longer being supported and replaced by improved and different solutions) and a cost constrained market can severely increase costs and operational, and ultimately reputation risk. Although many businesses recognise technological obsolescence, the pervasive nature of technology often means they have little information to identify the risk and location of pending obsolescence and little money to apply to the solution. This paper presents a low cost structured method to identify obsolete software and the risk of their obsolescence where the structure of a business and its supporting IT resources can be captured, modelled, analysed and the risk to the business of technology obsolescence identified to enable remedial action using qualified obsolescence information. The technique is based on a structured modelling approach using enterprise architecture models and a heatmap algorithm to highlight high risk obsolescent elements. The method has been tested and applied in practice in two consulting studies carried out by Capgemini involving three UK police forces. However the generic technique could be applied to any industry based on plans to improve it using ontology framework methods. This paper contains details of enterprise architecture meta-models and related modelling.

1 INTRODUCTION

As the pace of technology introduction quickens and IS becomes more pervasive the rate of change of technology and the forced obsolescence has also increased (Bulow, 1986). This very pervasiveness increases the connectedness and reliance on specific technology which can quickly become obsolete (Whelan, 2000). This increases both the cost of maintaining existing and or replacing the technology (Solomon et al., 2000). Doing nothing is not possible due to risk of loss of service provision and hence the management of obsolescent technology is becoming critical. The current economic focus on austerity has increased the need for better obsolescence awareness and management as businesses seek to consolidate and reduce their costs whilst maintaining their technology competitiveness.

Much of the existing literature regarding obsolescence has focused on its definition and

relationship to specific business contexts. (Lemer, 1996) has focused on evaluation of the performance and eventual failure of obsolescent infrastructure facilities (e.g. public works, sewers, pavements etc) and the need to establish reliable design service life metrics. (Solomon et al, 2000) explores the obsolescence of electronic integrated circuit components such as DRAMs, via an adapted stages of growth model that includes obsolescence. Whelan (Whelan, 2000) discusses the impact of obsolescence on stock management and the value and effective mathematical productivity of stock ranging from computers to industrial machinery and its relationship to computer usage. Feldman (Feldman and Sandborn, 2007) looked into the problems of obsolescence with respect to the parts procurement lifecycle to improve algorithms for parts forecasting and management. (Aversano et al., 2004) examined strategies for evolving existing software and included obsolescence as a factor along

with quality, economic and data value in their metrics. Similarly (Boehm, 1999) includes technical obsolescence as a factor in a similar paper on software productivity and reuse. Little work has been conducted to develop a practical methodologies approach to finding obsolete software and related components.

Obsolescence results in an inability to meet performance criteria (Lemer, 1996), for example when the requirement has moved on, or the technology has been superseded. Our concern is with the former. We propose an approach to identifying and managing obsolescence and risk using a simplified enterprise architecture and heat map approach that can be scaled to different size companies. The approach has also been successfully trialled to identify and manage IS infrastructure obsolescence in a specific business case of a police service where low cost obsolescence risk assessment was required to support change decisions.

There are three main issues to address when considering the impact of technology obsolescence on a business: we need to define obsolescence and its impact factors; we need a way of identifying the other aspects of technology and services impacted by that obsolescence; and finally, we need a way of identifying the greatest risk to the business from those obsolete technologies and services. These three issues are addressed sequentially in the next section on approach.

2 APPROACH

2.1 Obsolescence and Impact

2.1.1 What is Obsolescence?

As Sandborn suggests definitions of software obsolescence for commercial off the shelf (COTS) products vary, as there is often a big gap between the end of the product sale date and the date of the end of the support for the product (Sandborn, 2007). The withdrawal of support however, may not lead to immediate loss or degradation, but acts as a trigger for the business to make the necessary decisions to preserve the capability provided by the technology, subject to its criticality. For our purposes we will define obsolescence as the loss or impending loss of support for technology that reduces its ability to continue to function in the organisation (Feldman and Sandborn, 2007). Whilst the date of end of production of the technology and the date of end of support is reasonable to identify, the way in which

the technology will be affected by degraded or no further support is more difficult to determine and will depend on the capabilities and resources of the organisation as well as the technology being considered (Lemer, 1996).

2.1.2 The Impact of Obsolescence

Many organisations are running with a technology estate that is substantially obsolete. This causes problems of cost and risk. Increased costs are introduced by (a) the requirement for special support arrangements; (b) the extra development required to provide applications that can cope with obsolete technology (a notable case in point is the continued use of Internet Explorer 6 and Windows 2000 in the UK public sector); and (c) difficulty/high cost of obtaining support for very out-of-date technologies (both hardware and software).

Many factors may contribute to a high cost of ownership of an IT estate; it is often instructive to keep recursively asking the typical root cause analysis question “why” to determine root causes (Ginn et al., 2004). For example, high costs associated with running applications in data centres could be traced back to costs for the actual infrastructure (servers, switches etc.) and costs for the datacentres themselves (cooling, power, rent etc.) For the servers, then costs could be associated with both software and hardware – and the costs for both of these are likely to go up over time as the products on which they are based start to become obsolete and thus the subject of special (custom) support arrangements. Eventually the products involved will become unsupportable, which transforms an issue with high cost into an operational risk to the business. A root cause analysis in graphical format made this very apparent and understandable to the client (fig. 1).

The increased costs and the lack of knowledge results in increased operational risk of service failure, especially where obsolete technology is a key part of the core service delivery capability of the business. This operational risk if not attended to can then result in reputational risk where business service performance is badly impacted by IS/IT failure. With no action there is a risk of degradation of business service capability through the inability to process and disseminate quality (correct, complete and timely) data and information through the enterprise. However, it is not enough simply to identify the technology components that are subject to obsolescence. We also need to understand how the technology supports the business and specifically

the processes and services delivered by it. Critically we also need to quantify the importance of the relationship between in order to understand the impact of the obsolescence.

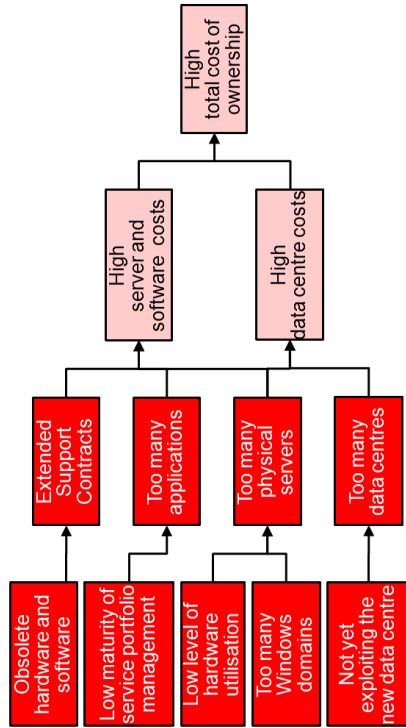


Figure 1: TCO Root Cause Analysis.

A near obsolete IS component will be dependent on a cascade or chain of factors that determine the level of risk and impact of the obsolescence as seen in the four step chain in Figure 2.

The first factor is the proximity of the technology component to obsolescence. This will depend on both the vendor support and maintenance to allow the product to function after obsolescence. The second factor is obsolescence criticality: even if the obsolete technology is supported after obsolescence the service provided by it might be degraded, as it may not be able to provide the IS service performance of newer or competitor products. Alternatively an upgrade path may allow the degradation to be reduced. The third factor is impact of the IS component on the business service. For example is the performance and capability of the IS component critical to the business service, or does it add very little value? Finally the fourth factor is the criticality of the business service in terms of delivering value to the customer. Ideally all these factors should be taken into account on the basis of a failure mode and effect analysis (Ginn et al, 2004)

but this can increase cost and complexity. Some analysis frameworks already provide useful metrics for the criticality of the business and IS service relationships (Liu et al., 2011) but these are better suited to comprehensive architecture analysis. For our purposes we propose a more simplistic risk framework based on a simplified metamodel explored in the later sections.

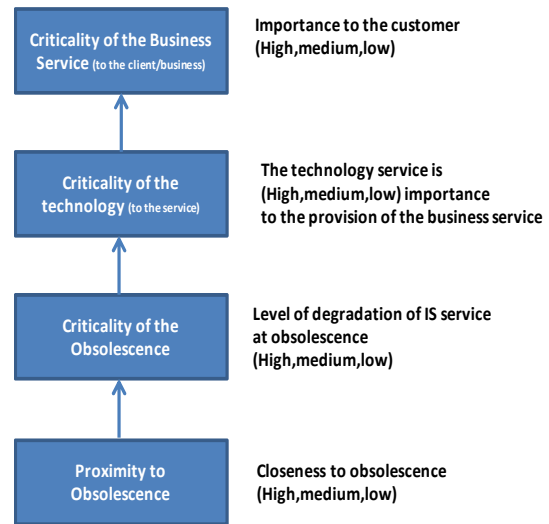


Figure 2: Obsolescence Impact Factor Chain.

2.2 Services and Technology Impacted

2.2.1 Identifying the Services and Technology Impacted

Information-intensive industries have very large amounts of IS/IT with attendant software and technology hardware and component risks. If the maturity of the architectural and operational processes have not kept track with the size of their IT estate, then there may be a lack of understanding of exactly what technology the organisation has, who is using it, and what its vulnerabilities are.

2.2.2 Understanding the IT Estate

The term “IT Estate” is generally used to refer to the complete portfolio of technology used within a business – including applications and infrastructure. If the knowledge of such an IT Estate is not encapsulated and maintained in some kind of ‘living repository’, then the lack of corporate knowledge can be exacerbated over time by changes to personnel within the organisation, so that key knowledge as to what exists and for what purpose, is lost as people’s roles change. The federated nature

of some large organisations can make this problem worse, because from the outset, no one part of the organisation ever has a complete picture of the business and technology architecture. The best that can be achieved in these circumstances is that individual parts of the organisation try to document solution architectures that capture their piece of the picture. A lack of knowledge of an Enterprise Architecture can make it hard to plan for the future, because without knowing the ‘as-is’ state, it is difficult to know what needs to change in order to achieve a ‘to-be’ state. Thus, it is possible to view the ‘status quo’ as a safer option, putting off the required upgrade and modernisation projects. Without knowledge of what applications are being used, by whom, and their underlying technologies, it is not possible to get a view as to the risks being posed to the business due to this obsolescence.

2.2.3 Case Studies

The aims of the first project, from the client’s perspective, was (a) to understand the potential cost savings associated with rationalising the application platforms, and (b) understand the degree of risk associated with technology, and provide a roadmap for addressing it. Thus, the idea of heat-mapping the business risk is highly relevant because it provides part of the ‘business case’ for making the relevant upgrade / replacement projects to address the risks thus identified.

The aim of the second project, a partial merger of two UK police forces, was to seek cost savings by identifying duplicate applications. However, the heatmaps were also relevant here, both in pointing

out risks to the businesses, as well as assisting in the choice of applications to retain.

2.2.4 Enterprise Architecture Meta Model

It is necessary to gain an understanding of the business technology architecture estate in enough detail so that, obsolete technology types can be traced through to the applications relying upon them, and thence through to the business services and functions that in turn rely upon those applications. Many enterprise architecture models have been developed to make sense of business and technology components found across a variety of businesses (Lagerstrom et al., 2009) and aligning business services with IT capability (Strnadl, 2005). Our specific issue requires a focused model that is easy for business and technical users to understand, but also shows dependencies between components. For this reason we have adapted and extended part of the TOGAF (OpenGroup, 2009) content meta-model to build a set of artefacts and inter-relationships that are particularly relevant to our area of interest.

Although standards exist for modelling Enterprise Architectures, it is frequently necessary to adapt those standards for different client situations. Reasons for this include: (a) the client prefers a particular standard (e.g. TOGAF or Zachman (Noran, 2003) or one of their own frameworks and (b) there are requirements for a particular engagement that demand a change to the standard model. In the first example discussed in this paper, the client was particularly interested in rationalising the application platforms rather than the applications themselves. Therefore, there was a heavy focus on infrastructure. The model used is illustrated below:

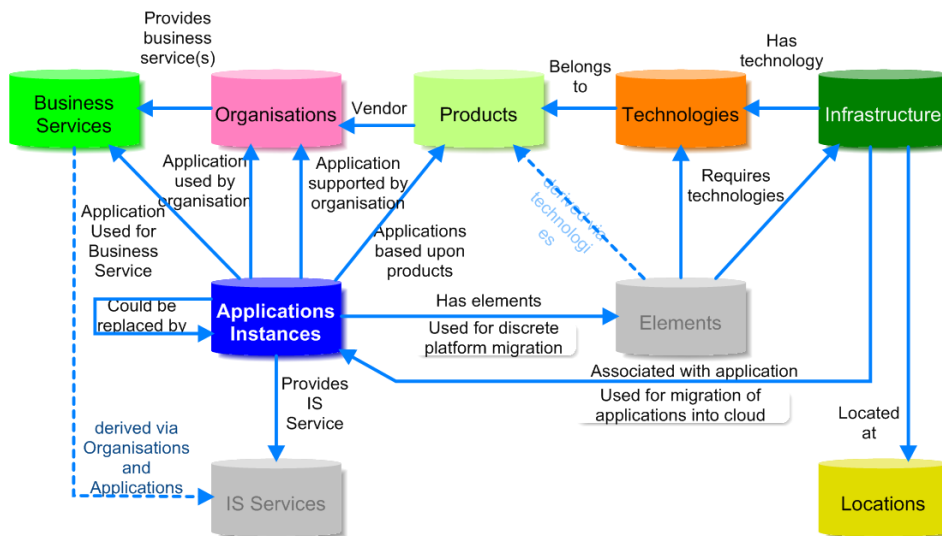


Figure 3: Sample Meta Model for Analysing Cost and Business Risk due to Obsolescence.

Figure 3 represents the actual model that was used to capture, analyse and report on various aspects of the IT portfolio for the first police force. With this particular client, two aspects of the model (shown in grey) were not used at the time of writing. For the other study involving two police forces, heavy use was made of the IS Services component, critical for application de-duplication.

Parts of this model had previously been used with another client to carry out an analysis of their application portfolio with a view to rationalising that portfolio (removing duplication). This gave rise to the two elements focusing on Applications (or **Application Instances**) and **IS Services** (both taken from TOGAF). By definition, any two applications that are labelled as offering the same IS Services are duplicates; and one of the aims of that study was to aim for a ‘minimum set’ of applications that gave the full range of required IS Services (functionality). Part of the business case for rationalising applications is of course the cost of running those applications, and part of the cost of an application comes from the servers hosting that application. Servers are of course a particular kind of infrastructure, which is why the model includes **infrastructure**. This infrastructure resides in physical **locations**, which are important to know for a number of reasons, especially when part of the rationalisation design includes closing one or more data centres. This has no bearing on the question of obsolete technologies, but was critical to the analysis and creation of the rationalisation design.

The servers in the IT estate for any client will have installed on it a number of software **products**, for example applications, databases, middleware, operating systems, monitoring and so on. In addition, the servers themselves are of course products from a hardware manufacturer. Thus, the products need to include a list of all software and hardware in the IT estate. In practice, the terminology used to describe a particular product may be very different to the terminology used by auto-discovery software, which sometimes goes to the extent of looking at versions of libraries installed on the servers (for example, Dynamic Link Libraries, or DLL files, on Windows platforms). Examples of this are (see table 1).

When looking up products on manufacturers’ websites, the term in the left column needs to be used. However, when auto-discovery tools are run, the terms in the right column are those that are generated.

Table 1: Sample Product-Technology Mappings.

Product	Technology
Windows Server 2003 Enterprise x64 (SP2)	Microsoft(R) Windows(R) Server 2003 Enterprise x64 Edition Version 5.2.3790 Build 3790 SP2
RDBMS 10g Release 2	Oracle Database Server 10.2.0.4.0
Solaris 10	SunOS 5.10

With this particular client, there were several thousand servers that needed to be matched up to products. Using this intermediate mapping meant that this could be done largely automatically. Once it had been calculated to which product a particular ‘technology’ corresponded, then that mapping was automatically applied to all instances of that technology.

Another reason for the use of this intermediate layer was the fact that in many cases, there were multiple pieces of software that corresponded to the same product, which were ‘discovered’ separately. For example, two separate pieces of software were discovered (“Oracle Net Services (TNS) Listener 9.0.1” and “Oracle Database Server 9.0.1.0.0”) that both corresponded to the same product (“RDBMS 9i Release 1”). This intermediate mapping of **technologies** provides the ability to cope with multiple synonyms and multiple pieces of technology that belong to a single product. The intention of the **elements** artefact was to allow the modelling of major components of the application (e.g. web tiers, database tiers, business logic tiers, storage allocations) so that they could then be rationalised. This is not relevant to the obsolescence discussion.

The final pieces of the model are both drawn from TOGAF. The **organisation** information allows us to represent the structure of an organisation, which is where the users of the applications reside. This is also useful for representing external organisations, for example the vendors of the products, or those involved in some way in supporting the applications.

The **business services** are the services provided by the business to its client. Clearly defining business services is necessary in order to be able to make a correlation between the services provided by the business and the IT that supports those services. Also the criticality of the service to the end client, will in turn affect the importance of the IT service and hence the impact and risk associated with the obsolescence of the technologies and components.

2.2.5 Populating the Model

The model was implemented using a particular modelling tool, MooD®. The functionality provided by this tool was critical to our ability to import, analyse and report on the data that was captured.

Many different information sources were used to populate this model, including but not limited to Active Directory, Tideway (auto-discovery software), a Configuration Management DataBase (CMDB) product and various spreadsheets populated manually.

The applications were populated using a combination of reports from the CMDB and spreadsheets provided by the client. The infrastructure was populated using a set of spreadsheets from various sources. The technologies were populated using spreadsheets from the auto-discovery software. The products were populated partially manually, interpreting the technology lists in the light of the team’s knowledge of the marketplace, and partially using data from the CMDB.

The organisation was populated using information on the client’s public web site.

The business services were populated from a generic UK police business service architecture published by the National Police Improvement Agency, called the “Policing Activities Glossary” (NPIA, 2011). This provided a hierarchical representation of the services provided by UK police forces, which we represented graphically in the modelling tool:

2.2.6 Tracing Obsolescence to Applications

Obsolescence as applied to technology (such as hardware and software products), using the definition previously offered, means “loss or impending loss of support for *hardware or software products* that reduces their ability to continue to function in the organisation”. These products are produced by various organisations (vendors / manufacturers), who often specify a date beyond which support for their products will either cease, or become more restricted (and perhaps substantially more expensive). Some manufacturers specify “End of Support” (EOS) and “End of Extended Support” (EOES) dates for their products.



Figure 4: Obsolescence from Products to Infrastructure.

Starting from the products, therefore, and knowing that several products may relate (via the Technologies intermediate layer) to a piece of Infrastructure (a server), then it is possible to say, for each piece of infrastructure, what is the earliest EOES date for any product that relates to that piece of infrastructure. For example, if that infrastructure was based upon a server model that had an EOES date of February 2013, but ran Windows 2000 that had an EOES date of 13th July 2010, then we can say that the earliest EOES date of these is the latter – so give the infrastructure as a whole, an EOES date of 13th July 2010. In other words, there is something about this piece of infrastructure that will be difficult and/or costly to support beyond that date.

The next step is to roll this up into the application layer.

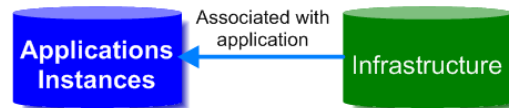


Figure 5: Obsolescence from Infrastructure to Application Instances.

Following the same approach, we can look at all the servers that are associated with a particular application (or instance of an application), and pick the earliest EOES date for each of these servers. In other words, we are saying that for a particular application, there is a particular piece of technology (software or hardware product) somewhere in the supporting infrastructure that will make application difficult and/or costly to support beyond that date.

2.3 Identifying Business Risk

An approach to identifying business risk from IT in general is discussed in (Halliday et al., 1996). This paper suggests a four-quadrant model for categorising risk, including notably the “avoid/prevent risks” which are viewed as being most critical because of the impact to the business should they be triggered, as well as the probability of them occurring. In this particular paper, we are focusing on the risk to the business from technology obsolescence that would fall within this particular quadrant. In other words, these are risks that a business can and should manage down to an acceptable level, and perhaps would if those business stakeholders were actually aware of their existence in the first place.

The use of ‘heatmaps’ is appropriate to demonstrate business risk, if backed up by proper evidence.

2.3.1 Algorithm for Generating a Risk Heatmap

The heat map, a colour-coded display of the intensity of a result has been used in various forms (Wilkinson and Friendly, 2009) to provide immediate visual understanding of multi-objective optimisation processes. It has been widely used in consultancy and problem solving as a means of highlighting critical obsolescence information in an easy to understand form (Miyake et al., 2004) (Detre et al., 2006).

As discussed in the risk section, whilst the Obsolescence Impact Factor Chain helps to cover a range of appropriate factors, this can quickly become complex and costly and hence we adopt a simplified approach that assumes full support is provided by the IT service and focuses only on the criticality of the service and whether the technology is deemed obsolescent as defined by the client.

Continuing the algorithm started in the IT domain, it is possible to look at all the applications used to support a particular business service, at each of their EOES dates, which in turn are rolled up from infrastructure and products. We ask the question “are there any applications required by this service where the EOES date has already passed?”. Where this is true, this obviously indicates that there is a degree of risk associated with the continued operation of that business service. The mapping of applications to business services can either be done via the organisation structure (i.e. for each organisational unit, determine which applications they use; and also which business services they provide) – used for our second study with the two police forces; or in the case of first study for a single police force, they were able to provide a direct correlation between applications and business services, as shown below.

```

FOR each business service
IF ∃ supporting applications of criticality 1 and EOES date in the past
THEN
    SET risk to MAJOR RISK
ELSE
    IF ∃ supporting applications of criticality > 1 and EOES date in the past
    THEN
        SET risk to SOME RISK
    ELSE
        IF we cannot associate any infrastructure with this service
        THEN
            SET risk to UNKNOWN RISK
        ELSE
            SET risk to NOT AT RISK
        ENDIF
    ENDIF
ENDIF
ENDIF
    
```

Figure 7: Algorithm for Calculating Business Risk in a Heatmap.

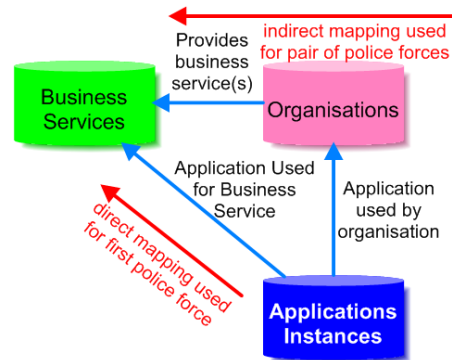


Figure 6: Obsolescence from Applications to Business Services.

However the mapping is done, the following algorithm was then used, which relies upon knowing the EOES dates for all the support applications, along with an indication of their criticality to the business (1 being the most important) (Fig.7).

Like all of the algorithms in the model, this calculation was automated for all of the business services using the tool’s ability to calculate and store intermediate results, including where necessary the ability to ‘call out’ to Excel. For example, the above algorithm looks like we can see in figure 8.

2.3.2 Risk Heatmap for a Police Force Requirement

By applying the above algorithm to each business service (from the PAG) in turn, it was possible to assign a risk value to each business service. The modelling tool was able to assign a colour to the business services dependent upon the risk value, and so the resulting heatmap looked like the figure 9.

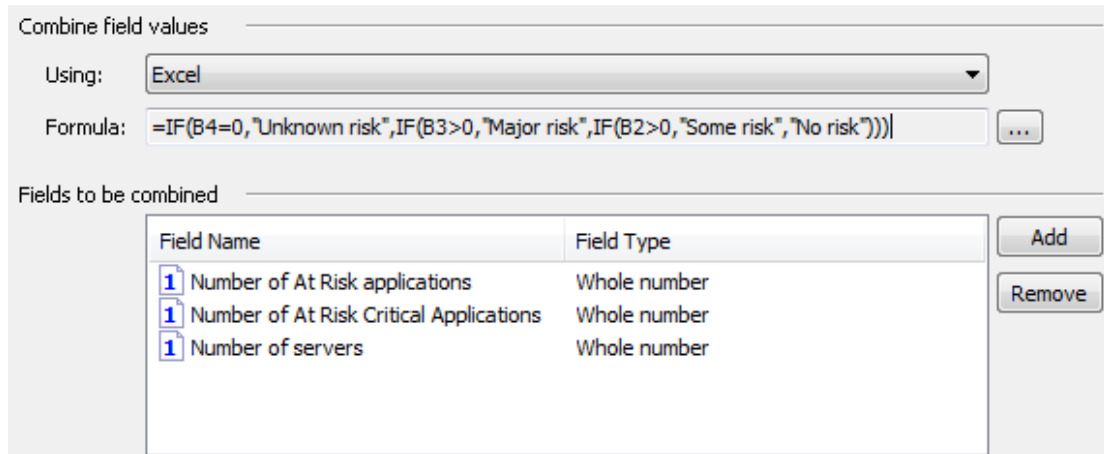


Figure 8: Automated calculation of business risk using MoodD®.

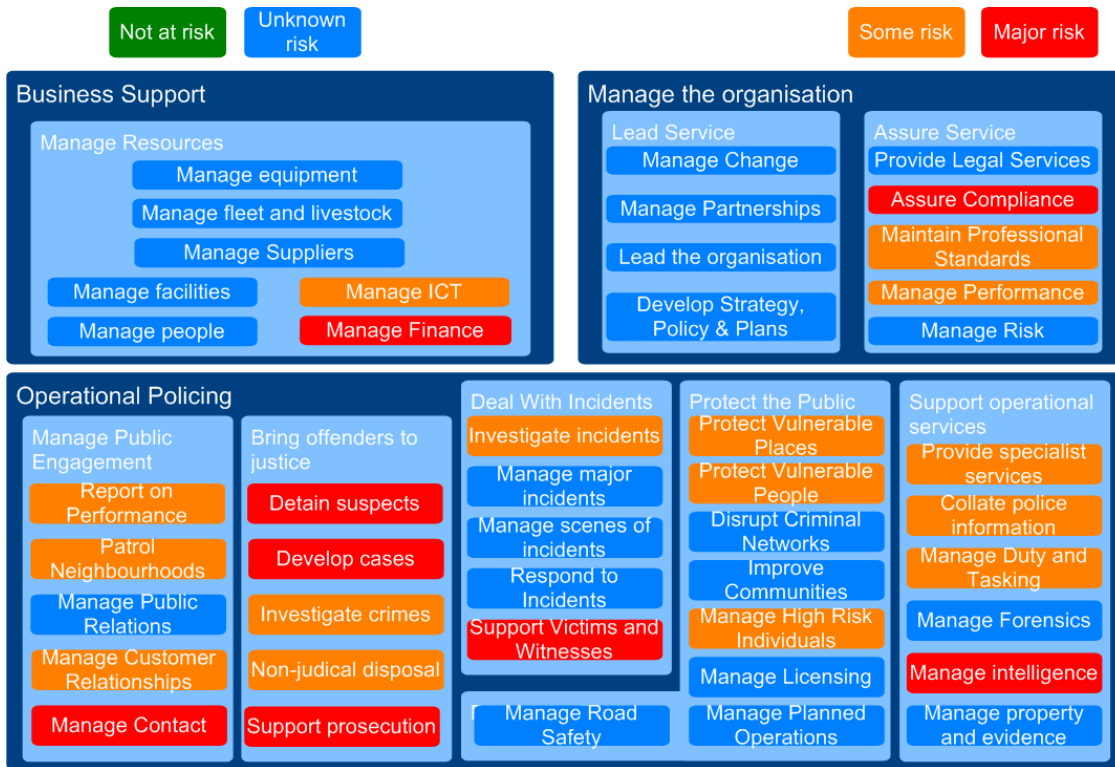


Figure 9: Police Business Services Risk Heatmap.

The value of this model lies in its ability to convey a technical concept (out-of-date products) and the link through the Enterprise Architecture model to business stakeholders in terms that are easy to understand and completely non-technical. Thus, this is a useful tool in demonstrating the risk component of the business case for making the necessary changes to the IT portfolio, to remove this risk.

3 LESSONS LEARNT

3.1 Building the Model

In order to carry out this kind of analysis, two things are needed: firstly, the raw information, with consistent terminology and all the relationships between the various kinds of information; and

secondly, some kind of toolset to enable the import, analysis and reporting of that information. For organisations with a relatively mature EA function, the first of these should not present a huge challenge, however in many cases even getting the client to produce a single definitive list of applications is difficult. It is also difficult to see how this could be done without some kind of modelling and requirements capture tool. However, one approach we are exploring is the development of an ontology chart based on organisational semiotics principles. The MEASUR (Stamper, 1994) model approach offers a structured method to interview and model the ontology of an organisation and has been used in related work applying the techniques to enterprise architecture and consulting modelling frameworks (Liu et al, 2011). Excel can handle two- or perhaps three-dimensional data with the help of pivot data; complex meta-models such as Figure 3 are probably beyond the ability of such tools to handle. A more robust technology is required, ideally layered over some kind of relational database to ensure the integrity of the data. Some candidate tools can be seen in (Short and Wilson, 2011).

3.2 What We Got Out of It

The team involved in the project felt that by the end of the study, a good deal of evidence had been collected that gave a very strong business case to continue work with this client, to address the cost and risk issues identified so clearly by the work so far. In the subsequent study; a partial merger of two UK Police Forces, the approach from the first study was readily re-usable, using the easy to understand heat-map, even though the meta-model was significantly different for the merger scenario. The resulting heat-map for the pair of police forces was all red, due to a critical corporate application, used across the whole of the organisations that uses Oracle 8. This ‘all-red’ picture gave a very powerful and well-received message to the client about the urgency of the situation. The meta-model used for the second case study, built using lessons learnt from the first, omitted the infrastructure and technology catalogues, relating applications directly to underlying products. It also differed in that the linkage from applications to business services went via the organisation structure. This linkage was much simpler in the first study, going directly from applications to business services. Nevertheless, the heat-map was still able to be calculated in a similar fashion.

3.3 What the Client Got Out of It

The main deliverable being sought by the client, regarding obsolescence, was a view as to the motivation (business case) for making change. The use of the business heatmap, along with the TCO root cause analysis (shown above) and other financial information outside the scope of this paper, provided a clear business case at low cost. The approach and model are capable of being extended to accommodate increased technology and risk complexity if required.

4 CONCLUSIONS

Having used the approach successfully in two separate cases with very different meta-models, albeit both in the policing sector, we have concluded that the approach is readily re-usable.

The library of product obsolescence data captured during the first engagement was very useful in terms of shortening the research required during the second engagement to produce the obsolescence heat-map.

5 FUTURE WORK

In retrospect, the terminology used for some of the artefacts in the meta-model need further work. In particular, the word ‘technologies’ is perhaps misleading in the way it is being used in this kind of analysis. Further research is required of existing architecture frameworks and methods to identify more commonly used and industry accepted terms that enable new clients to quickly come up to speed. As mentioned we are considering ontological analysis for specific industry terminologies to identify the relevant terms (Guarino, 1998). Also we intend to review and expand the obsolescence impact factor chain and investigate how this could be embedded in developing architecture based consulting analysis frameworks, potentially using the Capgemini Integrated Architecture Framework or the BTS analysis framework (Liu et al., 2011) which has embedded structures for identifying the relationships between business services and IS services and their criticality.

TRADEMARKS

MooD is a registered trademark of Salamander Enterprises Ltd.

Microsoft and Windows are registered trademarks of Microsoft Corporation in the United States and/or other countries.

Oracle and Solaris are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

REFERENCES

- Aversano, L., Esposito, R., Mallardo, T. & Tortorella, M. 2004. Supporting decisions on the adoption of re-engineering technologies. *Eighth euromicro working conference on software maintenance and reengineering (csmr'04)*. Tampere, finland.
- Boehm, B. 1999. Managing software productivity and reuse. *Computer*, 32, 111-113.
- Bulow, J. 1986. An economic theory of planned obsolescence. *The quarterly journal of economics*, 101, 729-749.
- Detre, J., Briggeman, B., Boehlje, M. & Gray, A. W. 2006. Scorecarding and heat mapping: tools and concepts for assessing strategic uncertainty. *International food and agribusiness management review*, volume 9.
- Feldman, K. & Sandborn, P. 2007. Integrating technology obsolescence considerations into product design planning. *Proceedings of the asme 2007 international design engineering technical conferences & computers and information in engineering conference, idetc/cie 2007*.
- Ginn, D., Streibel, B. & Varner, e. 2004. *The design for six sigma memory jogger : tools and methods for robust processes and products*, salem, nh : goal/qpc, cop. 2004.
- Guarino, N. Formal ontology and information systems. Proceedings of fois'98, 1998 trento, italy. Ios press, amsterdam, pp. 3-15.
- Halliday, S., badenhorst, k. & solms, r. V. 1996. A business approach to effective information technology risk analysis and management. *Information management & computer security*, 4, pp. 19 - 31.
- Lagerstrom, R., Franke, U., Johnson, P. & Ullberg, J. 2009. A method for creating enterprise architecture metamodels - applied to systems modifiability analysis. *International journal of computer science and applications*, vol. 6, no. 5, pp. 98 - 120.
- Lemer, A. C. 1996. Infrastructure obsolescence and design service life. *Journal of infrastructure systems*.
- Liu, K., Sun, L., Jambari, D., Michell, V. & Chong, S. 2011. A design of business-technology alignment consulting framework. *Caise conference 2011*.
- Miyake, M., Mune, Y. & Himeno, K. 2004. Strategic intellectual property portfolio management-technology appraisal by using the technology heat map. *Nri papers*. Nomura research institute.
- Noran, O. 2003. An analysis of the zachman framework for enterprise architecture from the geram perspective. *Annual reviews in control*, 27, 163-183.
- Npia. 2011. *Police activities glossary* [online]. Available: <http://pra.npia.police.uk/> [accessed 7th april 2011 2011].
- Opengroup 2009. The open group architectural framework (togaf 9).
- Sandborn, P. 2007. Software obsolescence - complicating the part and technology obsolescence management problem. *Ieee trans on components and packaging technologies*, 30, 886-888.
- Short, J. & Wilson, C. 2011. Gartner assessment of enterprise architecture tool capabilities. Gartner.
- Solomon, R., Sandborn, P. A. & pecht, m. G. 2000. Electronic part life cycle concepts and obsolescence forecasting. *Components and packaging technologies, ieee transactions on*, 23, 707-717.
- Stamper, R. 1994. Social norms in requirements analysis: an outline of measur. *Requirements engineering*. Academic press professional, inc.
- Strnadl, G. F. Aligning business and it: the process-driven architecture model. Computer as a tool, 2005. Eurocon 2005.the international conference on, 21-24 nov. 2005 2005. 1048-1051.
- Whelan, K. 2000. Computers, obsolescence, and productivity. *Ssrn eLibrary*.
- Wilkinson, L. & Friendly, M. 2009. The history of the cluster heat map. *The American statistician*, 63, 179-184.

ICT USE AS A DETERMINANT FACTOR IN STUDENTS' SATISFACTION AND SKILLS DEVELOPMENT

A New Model for the New Higher Education Context

Ángel F. Villarejo-Ramos, Manuel J. Sánchez-Franco

Department of Business Management and Marketing, University of Seville, Av. Ramón y Cajal, s/n, Seville, Spain
curro@us.es, majesus@us.es

Fernando Criado-García-Legaz

Department of Business Management and Marketing, University of Seville, Av. Ramón y Cajal, s/n, Seville, Spain
fcriado@us.es

Keywords: ICT, EHEA, Students' satisfaction, Skills development, Active learning methodologies.

Abstract: The important changes in learning methodologies in the context of the European Higher Education Area (EHEA) improve new educational methods. Therefore, the use of information and communication technologies (ICT) and their application in the new higher education model are considered as the background to student satisfaction and their development of generic and specific skills for professional results after graduating. This paper, firstly, provides an analysis of the expectations of a sample of second course undergraduate students in the business and management area (University of Seville, Spain). Secondly, it analyzes a causal model related to the use of participative teaching methodologies, perceived institutional support, and the ICT use in the teaching-learning process, as determinants of students' satisfaction and skills development in the context of the EHEA. The results of the descriptive analysis showed that students consider the use of specific materials and the use of e-learning platform to be important. The contrast with the estimation of the structural equations model and the results were satisfactory. Finally, the paper proposes some indications for higher education institutions and educators to further the right adaptation to the new context of the European Higher Education Area; for example, institutional promotion of learning technologies (e-learning instruments), training programs in ICT and teaching innovation.

1 INTRODUCTION

The context of university teaching and management is the result of numerous changes in past decades. The quantitative and qualitative increase in the university offer in the 80s meant a growth in the number of students and professors, as well as the incorporation of new study plans adapted to the technological, cultural and social changes. The new university management model in the 90s was the paradigm of the quality improvement of the teaching service. Concepts of business management were adapted, such as continuous improvement, process assessment and quality standards. The 21st Century has begun with the process of convergence towards the European Higher Education Area (EHEA). This context means that important methodological changes, the involvement of the university

institutions and the development and use of information and communication technologies (ICT) have become key elements to achieve success in the teaching-learning process. As Hernández, González and Guerra (2006) remark, this change in the EHEA is, mainly, a cultural change, a new vision of the professor-student relationship that implies a conception of learning in higher education in line with the changing demands of the labor environment and a new model of university management.

Currently, having surpassed the first phase of adaptation towards European convergence, the broad offer of Degree studies in Spanish universities has us facing the need to reflect about how these changes will be met with and adopted in the university context during the next decade.

This convergence framework does not mean homogenizing curriculums or procedures but rather

the abilities and the skills that are characteristics of each degree (Fernández, Suárez and Villarejo, 2008). This is why each university has tried to develop plans and strategies that are coherent with their socio-educational and professional realities (De Miguel, 2005a). The key to the success of implanting Degree studies can therefore be in detecting those needs and in adapting the new study plans to the achievement of the specific skills and the accomplishing of the students' expectations in the socio-professional placement that they will experience in the 21st. century. A high satisfaction level as a student and as a graduate will then be attained. Yet, as Correa and Paredes (2009) remarked, Spanish universities do not always incentivize the teaching staff taking part in processes of transforming their methods, which is why the need for a change in the teaching training model has been detected. This implies real adjustments towards more participatory methodologies.

On the other hand, learning with these characteristics requires methods that favour reflection about what is done, how it is done and what results are achieved. The students will in this way be able to use this as a strategy to improve their own performance, thus developing the most complex competence of all: that of learning how to learn with a critical sense (Fernández, 2006).

As is logical, insecurity and uncertainty have appeared in the renovation process in the university area. This has been particularly so when they are linked to a change that has the scope of the adaptation to the new pedagogical methods and a new philosophy of university teaching (Monereo y Pozo, 2003; Alvarez y Pérez, 2004). All this is immersed in a framework of the students' delocalization and globalization (Mavondo, Tsarenko and Gabbott, 2004) and the multiculturalism of the educational communities at which it is aimed (Castells 2000a; 2000b). Also, constant ICT development is opening routes to disseminate information and generate new learning modalities, but it is also opening digital gaps of varying importance between the students, the professors and the university administration staff (Área, 2000; Cabero, 2000; Zabalza, 2002).

The main aim of this current work is to analyze the students' perceptions about their university experience in the new Degree study plans and the incorporation of new didactical methods, the involvement of the university institution and ICT use as a tool in the teaching-learning process. All of this is from the students' perspective: accomplishing

their expectations with the satisfaction level and the skills development needed in their future professional work. This general aim can be specified in the following particular objectives: (1) to detect the methodology that the students consider should be the aim of preferential attention by their class professors and how knowledge and the assessment process should be acquired; (2) to explore the degree of ICT use by the teaching staff; and (3) to analyze the influence of the institutional support, the teaching methodology and the ICT use on the satisfaction and learning of the university student, both of them within the framework of the change of the teaching-learning process design model with respect to the traditional method of higher education (Fernández and Cabreiro, 2003).

2 RESEARCH METHOD

To attain these aims we have carried out an empirical study whose target population is the current second year students of different social sciences Degrees that are taught in the faculties of the University of Seville. The information gathering was carried out during the beginning of the 2010-11 course via an own questionnaire in which previous research was taken into account (De Miguel, 2005a; 2005b; Fernández *et al.*, 2008).

The final sample is made up of 402 valid surveys. The information from the sample has been analyzed via the statistical software *PASW 18*.

2.1 Characteristics of the Survey

In the questionnaire we included an initial block of variables that allows us to identify the most relevant characteristics of those who make up the sample. Specifically, we know the gender, age, marital status, type of High School and the average marks achieved, the degree of command in computer studies and their main software applications, the availability of a computer in the home and Internet access and the average time of daily computer use, both for study and leisure. The sample was extracted from undergraduate students of EHEA degrees. They are of an average age of 19.9 years, 54.7 % are female, 66.6 % studied in state schools and had Good or above in High School grades. The great majority of them (95.3%) have Internet access at home, and they mainly use the *Microsoft Office* pack. 68.4% of the students spend more than an hour per day on Internet-leisure, and almost half of the

sample spends more than an hour per day on Internet-study.

2.2 Basic Descriptive Analysis

The variables were analyzed via reflective indicators measured on the Likert 1-5 scale (1 - *strongly disagree*; 5 - *strongly agree*).

We can see that the main perceived institutional support is the promoting of virtual communication through web pages and virtual teaching platforms (68.2%). To a lesser extent (40.2%), the students consider that this support simplifies administrative tasks. However, a great number of them reckon that institutional support does not sufficiently respond to their study needs, training, exchange programs and internships; nor does it cover their communication and information needs either within or outside the institution.

We analyze the main concerns of professors in the organizing of classes. Students have shown a great interest in the teachers' orienting towards comprehending the contents they transmit and encouraging the students' interest in their subjects. On the other hand, students give less importance to the fact that the professors foster their taking part in classes and make them responsible for the learning process.

Regarding the resources they consider the teaching staff should use, the students prefer specifically-prepared printed materials to other types of support materials, both the classic ones and those connected with new technologies. 24% of the students do not consider the use of the books of publishing houses or the teachers' recommendation of them to be relevant.

With regards to the professors' means for knowledge transmission, more than 70% have shown an interest in the development of cases and simulation activities and class notes. They show a strong disinterest in traditional oral presentations, but also in methods linked to bibliographical investigation and self-learning. Nor do they give much importance to the organization of debates and work preparation, be it individual or in a group. Finally, the majority of the students (75.2%) consider that when marking the diverse materials, the professor should take into account the quality of the work carried out. What's more, it seems important to them for the level of knowledge attained in connection with the aims set out to be valued. The quality of the answers in oral tests (60.4%) and taking part in class and set activities (62.2%) has a lesser importance. The least valued

assessment systems are classroom tutorials and argued self-assessment: less than half of the sample does not consider them of much importance.

The assessment of the students about ICT show that the effects most valued by the students are associated with the better quality of access to the didactic materials and the disposition of better information access channels. The students give less importance to the possibility of the teaching staff orienting and following up the work carried out thanks to ICT, and the development of search and information selection skills. The least valued effects are associated with the possibility of generating contact and debate networks from virtual communities and the chance of increasing the student's interest and motivation. Lastly, we analyze the difficulties that students perceive in the teaching staff's ICT use. The factor that is considered to be most relevant, though for less than half of the sample, is the low training level perceived in the teaching staff when they use new technologies. On the other hand, a majority consider that the technological resources are available and adapted for their use in teaching.

3 THEORETICAL BACKGROUND AND RESEARCH HYPOTHESES

The interest that the teaching staff shows in fostering the students' communication and participation and promoting their responsibility in learning are key aspects in the skills that, once the years of study have finished, these students will have obtained (Fernández *et al.*, 2008). To achieve the students' participation via interaction with the teaching staff and with their companions favors their satisfaction level (Fredericksen, Shea and Pickett, 2000). A systematic teaching methodology that is intentionally organized to favor participation even it does not directly promote learning does favor the probability of this taking place (De Miguel, 2005b; 2009). What's more, when the development of skills linked to higher teaching aims such as the development of critical thought and autonomous learning is sought, methods centered on the students are more appropriate and efficient (Fernández, 2006).

As Fernández (2006) remarks, skills training even brings about the students' contact with the social and professional contexts in which the future graduate will perform. It also fosters the capacity to

learn with others, encouraging team work to exchange ideas, opinions, points of view, etc. The use of participatory teaching methods therefore means the possibility of developing the abilities and skills necessary for the university leaver in their professional career (Villarejo, Fernández, Suárez, Sánchez and Álvarez, 2010).

As a consequence of what has been presented, we can put forward the following hypotheses:

- *Hypothesis 1a*: Participative learning methodology proposed by the professor in the classroom has a direct and positive impact on students' satisfaction.

- *Hypothesis 1b*: Participative learning methodology proposed by the professor in the classroom has a direct and positive impact on the skills the students acquire through their academic training.

As well as the aspects referring to the use of the most appropriate methodologies, we must study the influence of the institutional support received during the period as a university student. The students' satisfaction is at times determined by multiple factors connected with the institution, such as the professors' level of preparation, the teaching styles proposed and even the support for research (Appleton-Knapp and Krentler, 2006).

Currently, universities need to implant *e-learning* systems, virtual campuses and *blended learning* to develop teaching practices in which research communities take part (Bonk, 2003; Anderson, 2004; Correa and Paredes, 2009). Moreover, the university institutions will be the ones in charge of fostering the working out of training offers. These are aimed at the teaching staff in order for them to acquire the skills that are necessary for them to attend to the new needs of Degree students and facilitate the resources needed to carry out these changes effectively (Álvarez and Romero, 2007).

In previous studies (Fernández *et al.*, 2008; Villarejo *et al.*, 2010) it was shown that: (1) to favor virtual communication via on-line teaching platforms; (2) to facilitate access to carrying out internships promoted by the University; (3) to encourage mobility and exchange programs with other universities; and (4) to simplify administrative tasks are matters that can influence the satisfaction levels attained by the students as well as the skills that they finally acquire. We propose the following hypotheses as a consequence of this:

- *Hypothesis 2a*: The institutional perceived support that students receive during their university stay has a positive influence on their greater satisfaction.

- *Hypothesis 2b*: The institutional perceived support

that students receive during their university stay has a positive influence on the skills that they acquire.

Additionally, ICT use in university teaching is a basic element of differentiation, compared to the traditional system that bears witness to the convergence of the teaching activity and the technological advances present in society (Área, 2000; Cabero *et al.*, 2003). Likewise, as in the previous descriptive analysis, students' familiarity with ICT is high and their expectations about training in the handling of them once their university studies are finished are also high. Technological skills and abilities are evident even in studies among pre-University students, in which a direct relationship is observed between ICT use and the development of specific skills in the subjects taught (Oliver and Corn, 2008). Therefore, the active learning methodologies using ICT have a positive influence on technological skills development, such as the use of Web 2.0 tools by students (Guerra, González and García, 2010).

However, and as Correa and Paredes (2009) pointed out, incorporating ICT into university life has fostered an important change in resources and infrastructures. Above all, it has modified the management and academic organization model, though this change has been less in teaching innovation. So, for students to satisfy their training expectations, university professors must use different technological resources to those that they dispose of to give their classes. This will require the teaching staff to control and handle ICT and renovate certain methodological aspects that this new form of teaching entails (Añel, 2008). Nevertheless, the professors are not in all cases prepared or accustomed to the use of these technologies available in their centers (Cabero *et al.*, 2003), especially when some studies show a difficulty in the teaching staff's ICT acceptance and use (Mahdizadeh, Biemans and Mulder, 2008).

Our proposal thus sets out the following hypotheses:

- *Hypothesis 3a*: The degree of the professors' use of ICT has a positive influence on the students' degree of satisfaction.

- *Hypothesis 3b*: The degree of the professors' use of ICT has a positive influence on the students' acquiring of skills.

Students' satisfaction increases to the extent that teaching methods based on inter-disciplinarity are favored [*inter-professional* education (*IPE*); Barr, Ivan, Reeves, Hammick and Freeth, 2005]. This entails an improvement in learning and putting it

into practice in the future professional work (Curran, Sharpe, Forristall and Flynn, 2008). What's more, and as Mavondo *et al.* (2004) pointed out, ICT use as a source of students' satisfaction provides them with the possibility of getting to know and trying out hardware and software that can encourage the development of skills needed in their professional work. Finally, let us consider that skills development can be influenced by the satisfaction level attained by the student:

- *Hypothesis 4*: The student's satisfaction experience has a positive influence on the acquiring of skills.

In the conceptual model (Figure 1), we set out relationships between the variables analyzed by our work. The aim of this is to favor recommendations for professors that may bring about a better perception of the teaching quality and the achieving of students' skills and abilities. Its contrast is carried out through a system of structural equations, using the statistical pack *SmartPLS 2.0*. The hypotheses set forth are of an exploratory nature. This is because their contrast gives rise to a more accurate knowledge of the relationships between the behavior of the professors, the institutional support and the students' satisfaction and learning that will orientate future research in an area that has not, until now, been the aim of many studies.

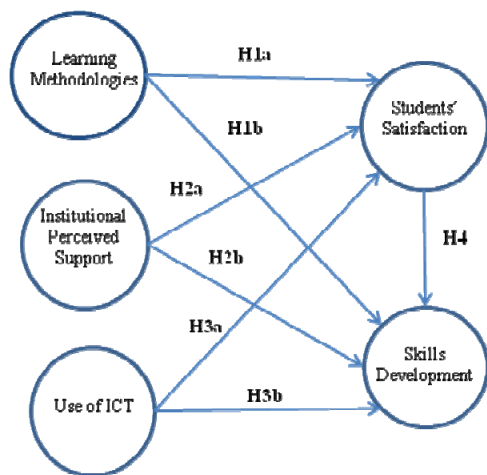


Figure 1: Causal model

4 DEVELOPMENT OF THE MODEL

As a previous step to their use in the causal model, we have assessed the reliability and validity of the scales that make up the measurement model. Initially, all the items in the questionnaire for the

five constructs of the causal model were considered. Nevertheless, an exploratory analysis showed the need to eliminate some of them to optimize the making up of the scales. When assessing the suitability of eliminating some of the initial items, we have taken into account the value of the Cronbach α coefficient, noticing, when its value was less than 0.7, its improvement in the case of suppressing some of the items.

Regarding the constructs ICT use and skills development we have opted for keeping items with values between 0.6 and 0.7 in their individual reliability, with the aim of not losing information. We checked that keeping them did not have a negative repercussion on the reliability of the construct measured. Throughout this process we used the statistical packs *PASW 18* and *SmartPLS 2.0* (Ringe, Wende and Will, 2008).

4.1 Reliability and Validity of the Measurement Scales

We based the individual reliability of the scales (Table 1) on the correlation coefficients of the items with the total of the measurement scale and the compounded reliability index. This must reach a value over 0.7. For each of the measurement models, the correlations of each indicator with the total of the scale satisfy the levels required, above 0.6 (Bagozzi and Yee, 1988). Advancing in the analysis of the psychometric properties of the scales, we went on to study their validity. The validity of content is accepted: the scales were designed from attributes contained in measurement instruments validated in previous studies.

Table 1: Convergent Validity.

CONSTRUCTS	AVE	Composite Reliability	Cronbach's α
IPS	.5577	.8345	.7372
SS	.5864	.8501	.7659
ICTU	.5608	.8355	.7359
PLM	.5323	.8501	.7799
SD	.6432	.8780	.8153

To check the convergent validity, we saw that the reliability of the constructs and their average variance extracted are over the recommended values of 0.7 and 0.5, respectively (Carmines and Zeller, 1979; Fornell and Larcker, 1981). It was also seen (see Table 2) that in all the constructs the reliability index measured via Cronbach's alpha is high, surpassing the 0.7 required in exploratory research (Cronbach, 1970; Ninnally, 1978).

Table 2: Individual Reliability.

PARTICIPATIVE LEARNING METHODOLOGIES (PLM)		.8501
Try to understand the contents transmitted	.7637	
Encourage students' interest in the subject(s)		
Foster students' communication and participation	.8539	
Stimulate students' responsibility for their learning	.8093	
PERCEIVED INSTITUTIONAL SUPPORT (PIS)	.7781	
Gives an answer to my needs concerning studies, internships, training, exchanges and so on	.7399	.8345
Covers my communication needs – information within and outside the institution	.7592	
Simplifies administrative tasks (there are support documents, data bases, etc.)		
Favors virtual communication (Web CT, domain us.es, etc.)	.7479	
ICT USE (ICTU)	.7399	
Projection systems (Rear projection of slides, video-projector, etc)		.8355
Computer, canon or multimedia systems	.7813	
Internet and telematic systems		
Virtual platforms (Web CT, etc.)	.8211	
STUDENTS' SATISFACTION (SS)		
The relationships that I establish	.6740	
The interest shown in my speeches, in the classroom, in the net and so on	.7103	.8501
The quality of the work and the tests that are set	.7734	
Passing the subjects based on the set criteria	.7656	
SKILLS DEVELOPMENT (SD)	.7801	
A better quality of access to the didactic materials	.7436	
An increase of our interest and motivation		
The development of search and information selection skills	.7620	.8780
The possibility of solving doubts about the program and the subject in the most efficient way	.7670	
A good enough command of ICT to apply them in my professional field	.7501	
	.7040	
	.6590	

To set up the discriminant validity, the AVE value must be above the variance shared between the construct and the rest of the constructs represented. For suitable discriminant validity, and to simplify the comparison, each element of the main diagonal

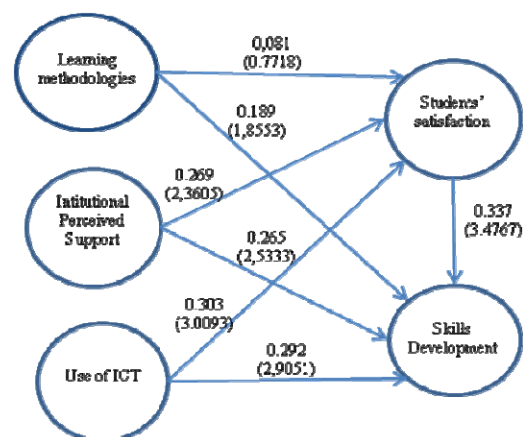
(the square root of the AVE) must be above the remaining elements of their corresponding row and column (correlations between constructs) (Barclay, Higgins and Thompson, 1995). In the model set out, the constructs satisfy the condition imposed, which leads us to accept the discriminant validity, as we can see in Table 3. Once the suitability of the measurement model scales has been checked, the next phase of the empirical study is the estimation of the causal model to contrast the hypotheses set out.

Table 3: Discriminant Validity.

	PIS	SS	ICTU	PLM	SD
PIS	.7467				
SS	.4179	.7657			
ICTU	.4051	.4342	.7488		
PLM	.3258	.2518	.2751	.7295	
SD	.4446	.5349	.4512	.3542	.8019

4.2 Contrast of the Model Set Out

Through the development of a system of structural equations, we have studied the relationships between the latent variables, the estimation of the parameters and their level of meaning (see Figure 2). We can accept five hypotheses set out in the model: *H2a*, *H2b*, *H3a*, *H3b* and *H4*. However, the relationships between participative learning methodologies and students' satisfaction and skills development, have not been able to be accepted due to the lack of significance of the relationship with the confidence level required.



Note: t-value in brackets. Based on $t_{(499)}$. Two-tailed test; $t_{(0.05; 499)} = 1.9635$; $t_{(0.1; 499)} = 1.6479$

Figure 2: Results.

From the contrast of the model, we highlight that the participatory methodology does not affect the students' satisfaction. This is justified because the

students perceive that the methods directly affect their learning and skills development. The students consider that the teaching activity serves to generate and develop skills and abilities.

Institutional support directly affects the students' satisfaction (*H2a*). This reflects the need for institutional support for the adaptation to the new EHEA as a priority of university institutions: directly or indirectly, it influences the development of the students' competences and skills through satisfaction (*H2b*).

ICT use significantly affects the students' satisfaction level (*H3a*), which corroborates the importance that Degree students give to new ICT in carrying out their work as students. ICT use in teaching directly and significantly affects the skills the student acquires (*H3b*). Fulfilling this last hypothesis serves to emphasize that students' learning in the context of university teaching is linked to the personal characteristics of the teaching staff. That is, their personal disposition and interest shown in ICT use, as well as the level of use of resources linked to ICT in teaching.

Finally, the students' satisfaction positively and significantly influences the development of their competences as university students (*H4*). This is justified by the fact that the students find, in the satisfaction of their needs, a possibility for the development of their skills needed for later professional work.

To sum up, the results of our research allow us to accept the positive influence of perceived institutional support on the students' satisfaction level and the development of their professional skills in the study area. Likewise, we can establish the positive influence that ICT use in university teaching has on the satisfaction level attained by the students in their study period and on the development of their skills for their future professional work. However, the positive influence of the use of participatory teaching methodologies on students' satisfaction or on the development of their skills is not accepted, though the latter could be accepted as significant for a lower confidence level. Finally, the satisfaction attained by the student favors the development of their skills.

5 CONCLUSIONS AND SOME PROPOSALS FOR IMPROVEMENT

As a result of the exploratory study carried out, there

is a series of contributions related to the students' perceptions that we consider relevant:

- The importance of perceived institutional support for the development of their university activities. They have shown a special interest in the development of techniques that favor virtual communication through e-learning platforms (Bonk, 2003; Anderson, 2004). This is the result of a growing interest of the universities in involving themselves in the adaptation process towards European convergence and the favoring of a more participatory teaching.

- The students give great importance to the concern shown by the teaching staff in the development of the training activity. They consider that it is very important to encourage interest in the subjects and to favor the communication and participation of the students. This is why they prefer specific materials for the subject and more practical classes via the use of cases and simulations.

- On the other hand, the professors do not show a special interest in the development of new teaching methodologies that substitute more traditional methods. Indeed, the students have shown little interest in self-assessment and assisting tutorials as assessment mechanisms. This leads us to think that the teaching staff will need to not only adapt their teaching methods to the new demands of the EHEA, but also take part in the design and use of these new methods to get more out of their work.

- The students' valuing of the lecturers' ICT use is especially favorable when it has a positive influence on the possibility of accessing to information or problem solving. However, they do not perceive that much importance is given to their motivation, the generating of networks of flexible contacts or the setting up of virtual learning communities. These aspects mean an important lack for students who find themselves, beginning a university career in which they are starting to build their own systems of relationships. We add to this the importance that students give to the low training level for ICT use that they perceive among the teaching staff (Cabero *et al.*, 2003).

- The students are not very satisfied with the interest that the teaching staff shows in their participation in the classroom and online as well as the quality of work and tests that the professors present in the development of the teaching-learning process.

In a previous paper (Fernández *et al.*, 2008), we analyzed the perceptions of students from two Spanish universities about the educational methodology used, as well as their opinion about the

possible aspects of ¿es un “the”? improvement in the ICT use. Establishing comparisons between the expectations of the present and previous studies related to the management and business area, the results have shown that the EHEA students give less importance to ICT use in learning methodologies and they think that professors have a good level of access to ICT and consider that they are trained in their use.

From the analysis of the causal model, it emerges that institutional support is important to achieve the adaptation of these new systems and the involvement of the university institution and the teaching staff in the setting up of new methods adapted to the new technological environment (Ávarez and Romero, 2007). This perceived institutional support positively affects the students' satisfaction level and the possibility of developing the specific and generic skills necessary for their professional work after graduation.

ICT use in teaching favorably affects the Degree students' level of satisfaction and the development of their professional skills. We must, however, pay special attention to this point as the restructuring of the educational and training systems may not be enough. Especially so when the ICT development programs in university training are only centered on technological aspects (Gromaz, Rodríguez and García, 2008), thus abandoning the true methodological character that the teaching-learning process should signify. In this sense, an efficient ICT use by the professors depends on, amongst other factors, the talents and skills that they develop (Martín-Sobrado, Ceinos and Fernández, 2010).

Due to all this, with regards to the university institutions and from the perspective of the teaching staff, it seems necessary to develop: 1) institutional policies to promote and deploy technological infrastructures and ICT access ; 2) spaces for personal and cooperative work, bearing in mind the communication tools available via ICT—mails/ emails/ notice boards and messages, forums and chat groups, videoconferences, and so on; 3) systems of inter and intra-institutional information dissemination ; 4) training models paying attention to the indicators diagnosed, taking into account the most deficient projects, areas and variables ; and 5) stimulating change processes oriented at reducing the digital gap in the teaching sectors that are least adapted to ICT use, as well as processes associated with the use of technologies to improve teaching in all its dimensions.

ACKNOWLEDGEMENTS

This paper has been supported by the *Consejería de Economía, Innovación y Ciencia (Junta de Andalucía-Spain)* through the Excellence Research Project (P09-SEJ-4568 and SEJ-5801).

REFERENCES

- Álvarez, E. and Pérez, R. (2004). Dimensión estructural y funcional de la comunidad educativa en una perspectiva de cambio: sistemas de relaciones y cultura de cambio. In Murillo, P., López, J. and Sánchez, M. (Coord.), *Cambiar con la sociedad, cambiar la sociedad* (pp. 344-358). Secretariado de Publicaciones Universidad de Sevilla: Sevilla
- Álvarez, V. and Romero, S. (2007). Competence-based education and training for guidance professionals. *Educación XX1*, 10, 15-37.
- Anderson, T. (2004). Teaching in an Online Learning Context. In Anderson, T. and Elloumi, F. (Ed.), *Theory and Practice of Online Learning* (pp. 273-294). Athabasca University: Athabasca.
- Añel, M. E. (2008). Formación on-line en la universidad. *Pixel-Bit. Revista de Medios y Educación*, 33, 155-163.
- Appleton-Knapp, S.L. and Krentler, K.A. (2006). Measuring Student Expectations and Their Effects on Satisfaction: The Importance of Managing Student Expectations. *Journal of Marketing Education*, 28, 254-264.
- Área, M. (2000). ¿Qué aporta Internet al cambio pedagógico en la Educación Superior? In Pérez, R. (Ed.), *III Congreso Internacional de Comunicación, Tecnología y Educación: Redes, Multimedia y Diseños virtuales* (pp.128-135). Servicio de Publicaciones Universidad de Oviedo: Oviedo.
- Bagozzi, R.P. and Yi, Y. (1988). On the Evaluation of Structural Equation Models. *Journal of the Academy of Marketing Science*, 16(1), 74-94.
- Barclay, D., Higgins, C. and Thompson, R. (1995). The partial least Squares (PLS) approach to causal modelling: personal computer adoption and use as an illustration. *Technology Studies, Special Issue on Research Methodology*, 2 (2), 285-309.
- Barr, H., Ivan, K., Reeves, S., Hammick, M. and Freeth, D. (2005). *Effective Inter-professional Education: Argument, Assumption & Evidence*. UK: Blackwell Publishing.
- Bonk, C. (2003). I should have known this was coming: computer-mediated discussions in teacher education. *Journal of Research on Technology in Education*, 36(2), 95-102.
- Cabero, J. (dir.), Castaño, C., Cebreiro, B., Gisbert, M., Martínez, F., Morales, J.A., Prendes, M.P., Romero, R. and Salinas, J. (2003). Las nuevas tecnologías en la actividad universitaria. *Pixel-Bit. Revista de Medios y Educación*, 33, 81-100.

- Cabero, J. (2000). La formación virtual: Principios, bases y preocupaciones. In Pérez, R. (Ed.). III Congreso Internacional de Comunicación, Tecnología y Educación. Redes, multimedia y diseños virtuales (pp.83-102). Servicio de Publicaciones Universidad de Oviedo: Oviedo.
- Carmine, E.G. and Zeller, R.A. (1979). Reliability and Validity Assessment. Thousand Oaks, CA: Sage Publications.
- Castells, M. (2000a). La era de la información. La sociedad red, Vol. 1. Madrid: Alianza.
- Castells, M. (2000b). La era de la información. El poder de la identidad, Vol. 2. Madrid: Alianza.
- Correa, J.M. and Paredes, J. (2009). Technological change, e-learning platform uses and education transformation in the Spanish Universities. *Revista de Psicodidáctica*, 14(2), 261-278.
- Cronbach, L. J. (1970). *Essentials of psychological testing*. New York: Harper & Row.
- Curran, V. R., Sharpe, D., Forristall, J. and Flynn, K. (2008). Student satisfaction and perceptions of small group process in case-based inter-professional learning. *Medical Teacher*, 30, 431-433.
- De Miguel, M. (2005^a). Adaptación de los planes de estudio al proceso de convergencia europea, Oviedo: Servicio de Publicaciones Universidad de Oviedo.
- De Miguel, M., (Coord.) (2005b). Modalidades de Enseñanza centradas en el desarrollo de Competencias: orientaciones para promover el cambio metodológico en el Espacio Europeo de Educación Superior. Proyecto EA2005-0118.
- De Miguel, M. (Coord.) (2009). Metodologías de enseñanza y aprendizaje para el desarrollo de competencias: orientaciones para el profesorado universitario ante el Espacio Europeo de Educación Superior. Madrid: Alianza.
- Fernández, A. (2006). Metodologías activas para la formación de competencias. *Educatio Siglo XXI*, 24, 35-56.
- Fernández, B., Suárez, L. and Villarejo, A.F. (2008). Determinantes del aprendizaje del alumno ante el reto de la convergencia europea. In Pindado, J. and Payne, G. (Coord.), *Building Bridges in a Global Economy*. ESIC: Madrid.
- Fernández, M.C. and Cebreiro, B. (2003). Evaluación de la enseñanza con TIC. *Pixel-Bit. Revista de Medios y Educación*, 21, 65-72.
- Fornell, C. and Larcker, D.F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 18(1), 39-50.
- Fredericksen, E., Shea P. and Pickett, A. (2000). Factors influencing student and faculty satisfaction in the SUNY learning network. New York: State University of New York.
- Gromaz, M., Rodríguez, M. J. and García, J. (2008). Experiencias exitosas de e-learning en la universidad gallega. *Pixel-Bit. Revista de Medios y Educación*, 33, 173-182.
- Guerra, S., González, N. and García, R. (2010). Study on the use of ICTs and teaching tools by University instructors. *Comunicar*, 35, 141-148.
- Hernández, A., González, N. and Guerra, S. (2006). Design of a portfolio in the High Education by competences. *Revista de Psicodidáctica*, 11 (2), 227-240.
- Mahdizadeh, H., Biemans, H. and Mulder, M. (2008). Determining factors of the use of e-learning environments by university teachers. *Computers and Education*, 51(1), 142-154.
- Martín-Sobrado, L., Ceinos, M.C. and Fernández, E. (2010). Planning and development of an ICT-skills map in guidance. *Comunicar*, 35, 167-174.
- Mavondo, F.T., Tsarenko, Y. and Gabbott, M. (2004). International and Local Student Satisfaction: Resources and Capabilities Perspective. *Journal of Marketing for Higher Education*, 14(1), 41-60.
- Monereo, C. and Pozo, J. I. (Ed.) (2003). *La universidad ante la nueva cultura educativa. Enseñar y aprender en la autonomía*. Barcelona: Institut de Ciències de l'Educació de la UAB-Síntesis.
- Nunnally, J.C. (1978). *Psychometric Theory*. New York: McGraw-Hill.
- Oliver, K. M. and Corn, J. O. (2008). Student-reported differences in technology use and skills after the implementation of one-to-one computing. *Educational Media International*, 45(3), 215-229.
- Ringle, C.M., Wende, S. and Will, A. (2008). *SmartPLS 2.0 (Beta)*. University of Hamburg: Hamburg, Germany. Retrieved March, 2010 from <http://www.smartpls.de>.
- Villarejo, A.F., Fernández, B., Suárez, L., Sánchez, M.J. and Álvarez, E. (2010). Participative methodologies and the use of ICT like determinant factors in the students' satisfaction and the development of competencies before the challenge of the European Higher Education Area. In *GUIDE International Workshop 2010, Association Global Universities in Distance Education*: Roma.
- Zabalza, M. A. (2002). *La enseñanza universitaria. El escenario y sus protagonistas*. Madrid: Narcea.

USING AHM APPROACH TO POSITION CODP

Mladen Velev, Ognyan Andreev

*Faculty of Management, Technical University of Sofia, 8, Kl. Ochriski Blv., Sofia, Bulgaria
{mvelev, oandre}@tu-sofia.bg*

Tanya Panayotova

*Department of Economics and Management, Technical University of Varna, 1, Studentska Str., Varna, Bulgaria
tagea@abv.bg*

Keywords: Customer order decoupling point, Analytical hierarchy process, Mass customization, Co-creation, Co-development.

Abstract: The concept of Customer Order Decoupling Point/CODP is a popular approach to increasing the diversity of end items, while taking advantage of standardization/unification due to increased repetitiveness of operations devoted to components and/or subassemblies manufacturing. CODP marks the place (the operation, the phase of the process etc.) where customer intervention occurs, in order to define, according to his/her wishes, the final mode of the end item (product or service). An underlying issue here is to make an economically motivated decision about the exact CODP position (1) among different end items of the company's product mix, and (2) inside a particular product/service line. Inside the operations process, before CODP, forecasts are usually used (Make-To-Stock), and after it – Make-To-Order. Consequently, the opportunities to achieve economies of scale are different before and after CODP. Therefore, the opportunities for optimizing the total operating costs are different as well. In the present paper, an approach is suggested for applying the Analytical Hierarchy Process/AHP in solving such a problem. Some examples of criteria are also presented to give reasons for the “pros” and “cons” during the decision making process about CODP position.

1 INTRODUCTION

CODP stands for “Customer Order Decoupling Point”. Often, for the same meaning, many authors use different terms and abbreviations, like “Customization Point/CP” (Ramachandran et al., 2002) “Delay of Product Differentiation” (Gupta & Benjaafar, 2004), “Point of Postponement/ PP” (Feitzinger & Lee, 1997) etc.

CODP is a widely used tool in the process of applying Mass Customization and Co-Creation. It is a popular approach to increase the diversity of end items, while using the advantages of the standardization/modularization due to an increased repetitiveness of operations devoted to manufacturing of components and/or subassemblies. CODP defines the stage in the manufacturing value chain, where a particular product is linked to a specific customer order. In fact, it marks the place (the operation, the process phase etc.) where the customer's intervention occurs, in order to define the final mode and appearance of the end item (no matter product or a

service), according to his/her wishes and preferences.

In general, the idea of CODP is presented in the Figure 1 (Andreev, 2009). On the top of the figure, a simplified view is used to depict the sequence of operations and supplier-client relationships. It is represented by the subsequent steps of the whole supply chain – from the suppliers of raw materials downstream to the end client – the customer. According to the position of CODP, the customer is “allowed to penetrate” through the operational process, by the act of his/her order, using different options to choose at the CODP itself. Thus, the customer could define one or more particular sub-assemblies/components of the end item to be used in the final assembly, or the components of any particular subassembly, or a given combination of both, as well as to define certain component parts, and so on – upstream to the beginning of the process.

In fact, each of the end items built this way is a different customized product/service assembled according to the choice/preferences of the particular

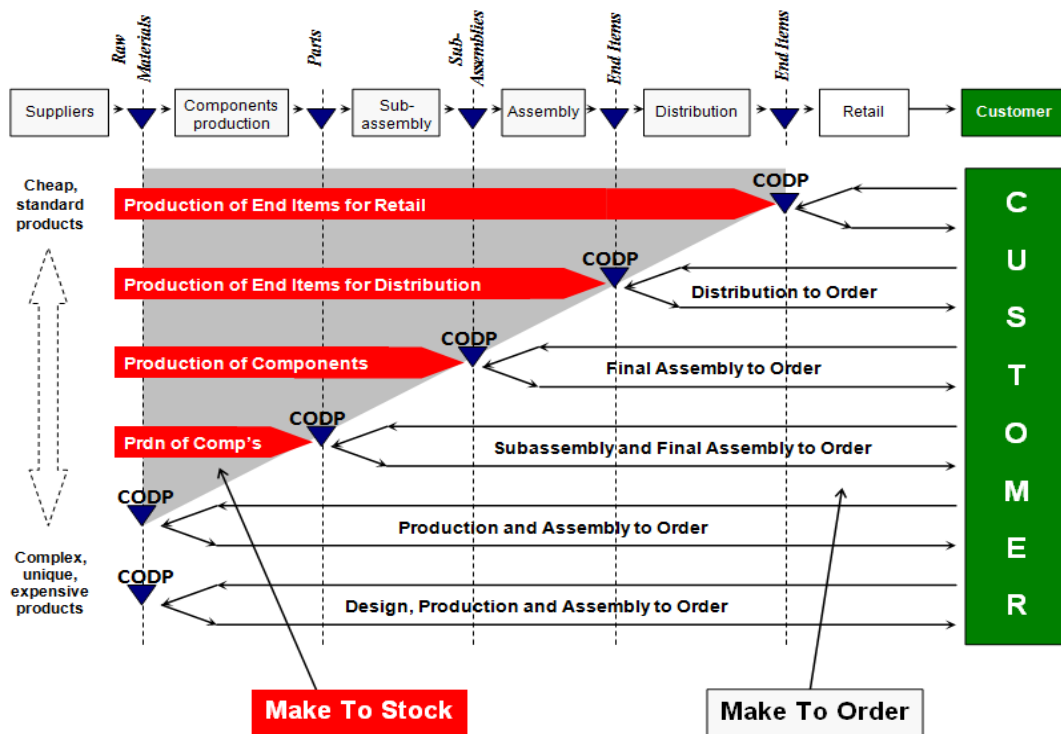


Figure 1: Variety of possibilities to position Customer Order Decoupling Point (Andreev, 2009).

customer. Moreover, the customer could be involved not only in choosing component options, but in performing some of the operations as well, likewise the case of home assembled furniture etc.

2 SPECIFICS OF CODP ISSUES

It can easily be seen from Figure 1 that moving CODP rightwards leads to a decrease in the end product diversity, and vice-versa. Also the opportunities for increasing product diversity by moving CODP leftwards are enhanced by, so to say, a kind of geometrical progression, due to the nature of the product breakdown structure, which “explodes” in quantity of generic items shifting down to the lower levels (Figure 2).

Acting this way, the company could combine the advantages of the Economies of Scope (a relatively rich diversity in the product mix, defined by the customer after CODP) with the ones of the Economies of Scale (aggregating and consolidating in batches as big as possible before CODP) with an aim to achieve a higher degree of customized variety of end products/services (Pine, 1992).

One can also see that both “boundary cases” are represented respectively: the upper one – 100% Sale-to-Forecast, and the one in the bottom – 100%

Engineer-to-Order. In between, hybrids are possible to be performed that complement mutually each other, so that CODP sets their share (Figure 1):

- Distribution/Shipment to Order
- Packaging/Labeling to Order
- Assembly to Order
- Make to Order
- Purchase to Order

The problem of deciding on the right position of CODP for different product families, as well as for the products inside the product families arises periodically in the companies not only in connection with the continuously changing environment – both outside the company (new market conditions) and inside (new technologies, materials, operations etc.), but also because of the pursuit of continuously improving their competitiveness.

However, from an economic point of view, the companies aim to such a CODP position, which is not only going to help them take a competitive market advantage, but to provide also maximum benefits/profits.

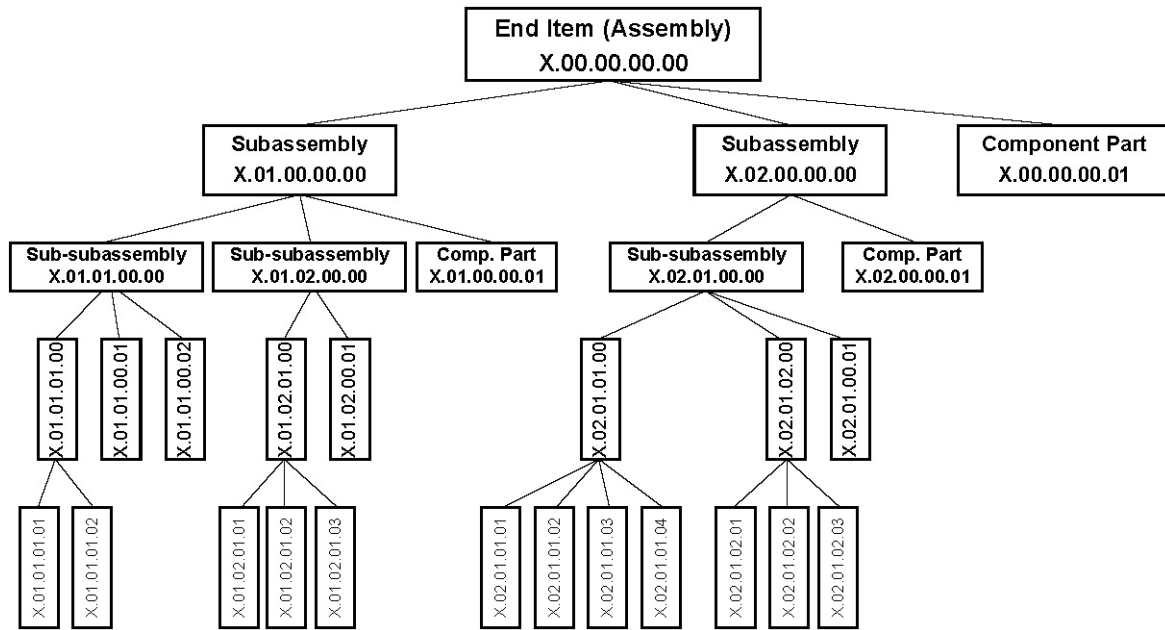


Figure 2: An example of a Product Breakdown Structure.

In determining the position of CODP, enterprises take into account the effect of a range of internal and external factors, which perform the role of selection criteria. These criteria build up a system and each of them can have its sub-criteria too. The simultaneous use of the whole range of criteria and their sub-criteria ensures a greater precision in the final estimates, but makes it quite difficult to determine them.

In their operations, enterprises are confronted with two major problems in determining CODP position: the first one is related to the correct establishment of the assessment criteria (influence factors) and their sub-criteria, and the second one – to the simultaneous taking into account of their effect.

Although several papers have discussed the conception of CODP, little work has been done on its influence factors and positioning (Olhager, 2001). Most often, it is noted that in the capacity of CODP it is possible to point out almost every component in the product breakdown structure (Figure 2), provided the enterprise is capable of suggesting alternative options for that component. Here, a modular design and structure of the end item is presumed so that it is possible to use any of the options in the further realization of the process at random (Velev & Tsvetanova, 2010). That also means that the capacity, which has to meet such a demand, should be extremely flexible and, as is most often the case, it will be to a great extent ineffectively utilized.

Our analysis of the publications points to the conclusion that it is necessary to suggest (1) a way of building up a complete system of criteria for decision making about CODP position and (2) a method/set of tools, by means of which the decision will be based on a relatively large number of such criteria.

3 DECISION MAKING CRITERIA FOR CODP POSITIONING

In order to solve the first problem we suggest a system of criteria for assessment and selection. Their particular meanings, which take into account the impact of the internal and external factors, depend on the specifics of the activities of the enterprises and the conditions of the environment, in which they are carried out. The system includes a hierarchical decomposing of the criteria into sub-criteria, etc.:

3.1 Degree of Conformity with the Specifics of the Products/Services

3.1.1 Degree of Compliance with Customer Requirements

A higher degree of compliance with the characteristics of the products, provided with the customer requirements, presupposes a greater shift of CODP

leftwards. The less standard and the better the product suits the requirements of the customers, the more profitable it is for the company. With standard products the need to meet these requirements is decreased. The degree of necessary interaction with the customers is different and it depends on whether the enterprise provides a tangible product or a service. In case that services are provided the degree of interaction with the customers is generally greater, which causes a shift of CODP leftwards.

3.1.2 Strength of Technological Ties with Customers

The greater technological commitment to the customers requires a better co-ordination with them – for example, manufacturing of components to be assembled by the customer. On the opposite, when end products are manufactured, the technological ties are low and CODP shifts rightwards.

3.1.3 Mode of Contacts with Customers

When industrial products are being produced and delivered (i.e. machinery), the contacts with the buyers are much closer and direct. Most often, products are made to order by agreed specifications. No intermediaries are used in delivery, which facilitates the shift of CODP leftwards. However, that is not the usual case with consumer products.

3.1.4 Degree of Technical Complexity of the Product/Service

The technical complexity of the product/service requires greater involvement of the customer in developing, manufacturing and delivery of the product. It is necessary to have better interaction with the producer and that leads to a shift of CODP leftwards.

3.1.5 Degree of Diversity of the Product Variants

The greater degree of diversity of the product variants and the product quality characteristics require a greater degree of conformity and closer interaction with the customer requirements. Thus, opportunities for the realization of Co-Creation and Co-Development arise and CODP shifts leftwards.

3.2 Degree of Compliance with Market Conditions

3.2.1 Degree of Market Turbulence

The fast-changing customer requirements and the intensifying market fragmentation, which makes it necessary to satisfy the specific requirements of small groups of customers, have provoked immediate responses from the manufacturers. Modern business strategies have been employed, at the basis of which a shift of CODP towards the left is laid.

3.2.2 Degree, to which Competitors Make Use of the Practice to Involve Clients in the Process of Manufacturing and Delivering Products/Services

On the one hand, the practices of the competitors reflect the collective experience of the companies, which are active on the respective market, and the company must take them into consideration. On the other hand, the introduction of new business strategies by the competitors, related to the use of Co-Creation and Co-Development and a shift of CODP leftwards, forces the enterprise to react accordingly. Otherwise, it will lose its competitive edge.

3.2.3 Degree of Intensity of Market Competition

The strong competitive pressure forces enterprises to look for new ways of increasing their competitiveness. By shifting CODP leftwards, a greater degree of product customization is achieved, as well as a closer interaction with the customers and therefore – an improvement of competitive position. Usually, there is reduction in the manufacturing and marketing costs, and achievement of better business results.

3.2.4 Degree, to which Customers are Looking for Opportunities to participate in the process of manufacturing and delivering the product/service

3.2.5 Technological Opportunities for the Customers to participate in the process of manufacturing and delivering the product/service

3.3 Degree of Conformity with Enterprise Goals and Strategies

3.3.1 Expected Increase of Sales due to the CODP position in consideration

3.3.2 Expected Reduction of Costs due to the CODP position in consideration

3.3.3 Expected Increase of Profit due to the CODP position in consideration

3.3.4 Expected Return on Additional Investment needed to position CODP at the location under consideration

3.3.5 Degree of Conformity with the Company Strategies

The outcomes from a particular CODP position must fit into the company strategic choices and goals.

3.4 Degree of Conformity with the Production Capacity of Enterprise

3.4.1 Technical / Technological Capabilities of Enterprise, needed to provide the performance required

3.4.2 Innovative Capabilities of the Enterprise to meet customer requirements

3.4.3 Degree of Product Modularity

Greater modularity of the products presupposes greater variability of end items and higher degree of customization. It presupposes also easier realization of the Co-Creation and Co-Development practices.

3.4.4 Degree of Processes Flexibility

Flexibility of product development processes is yet another condition for greater variability and individualization. It presupposes a CODP shift leftwards.

3.4.5 Information and Communication Opportunities

ICT availability and status is vital in applying the practice of customer participation in manufacturing and delivering the products.

3.4.6 Degree of Integration with Intermediaries

The higher degree of integration of the enterprise with the intermediaries in the supply chain, such as suppliers or distributors, is a prerequisite for a shift of CODP rightwards.

3.4.7 Reputation of the Enterprise as a Loyal Partner

The positive reputation of the enterprise as a loyal

partner, and one that maintains good quality is a prerequisite for increased customer commitment, leading to possibilities to shift CODP leftwards.

The above mentioned criteria can be changed and complemented depending on the particular conditions in which the particular enterprise operates.

4 METHODOLOGY FOR ASSESSING THE IMPACT OF DIFERENT CRITERIA

The second major problem related to determining the CODP position is connected with the simultaneous consideration of the whole system of criteria and sub-criteria, being used. In order to solve it, it is necessary to assess how strong the impact of each of the criteria and sub-criteria is, to evaluate their importance for obtaining desired final results, and to use an appropriate method for calculating the results according to various alternative positions of CODP. In addition, the method of assessment should avert or keep to a minimum the possibilities of errors and subjectivity in taking the final decision. The selection of CODP position should include a sequence of judgments and decisions, which have a hierarchical structure, as it is indicated in Figure 3.

In order to facilitate the procedure of choosing a position and to avoid or at least reduce subjectivity in taking the final decisions for the CODP location, it is recommended that the right assessment methods and software are used. An appropriate method for that purpose here is the Analytical Hierarchy Process/AHP (Saaty, 1980). AHP is a useful tool for choosing an option to be used out of a whole range of criteria, especially where there are sub-criteria to the criteria, sub-sub-criteria etc. Using that method, a choice is made based on the relative importance of the criteria and sub-criteria for achieving the aims set by the company strategy, and the capabilities of the operations to fulfil them.

The present paper aims at demonstrating, by using a simple example, the applicability of the AHP method in determining the position of CODP.

4.1 Determining the Relative Importance of the Selection Criteria

The criteria have different importance in maximizing the profits of the enterprise and for responding to the priorities of the customer. Thus, they carry different relative weights in making the decision about the choice of an option for CODP position.

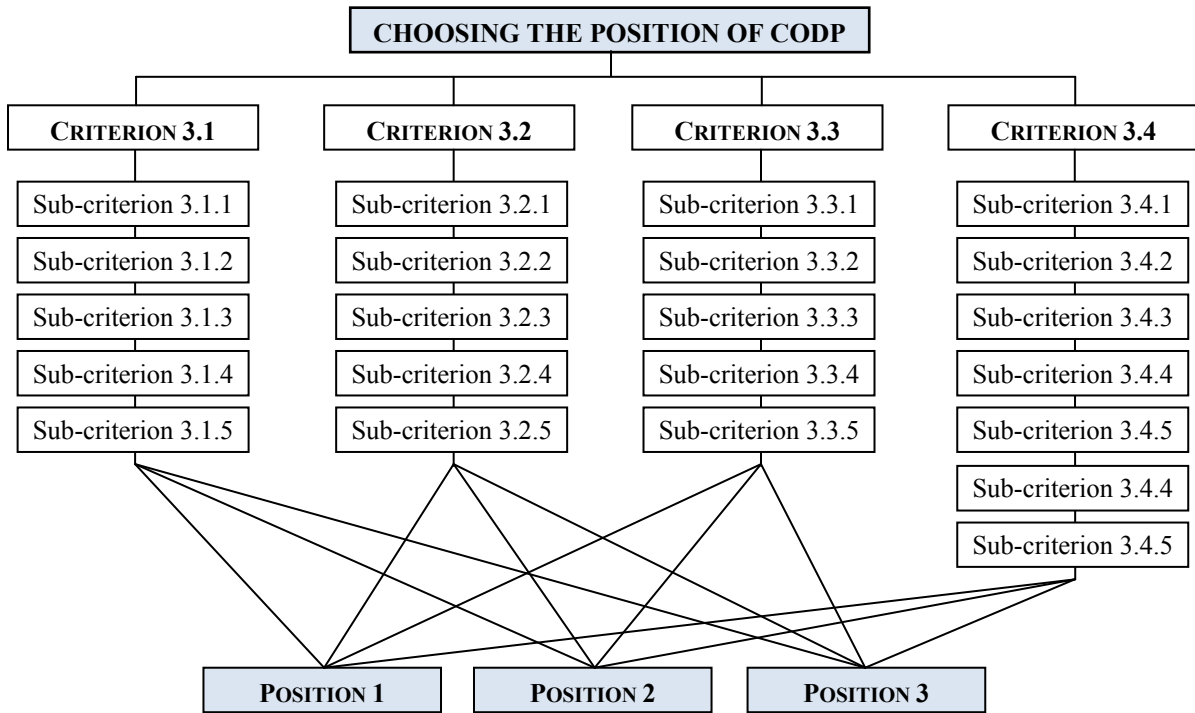


Figure 3: Decision hierarchy in choosing CODP position.

According to the AHP procedure, determining the relative weight of the selection criteria starts with their comparative assessment in pairs. This is done by a group of experts who use the assessment scale shown in Table 1 (Stevenson & Ozgur, 2007).

Table 1: Scale for determining the importance (priority)

Degree of Relative Importance	Score
Equal Importance	1
Between Equal and Moderate Great Importance	2
Moderate Great Importance	3
Between Moderate Great and Great Importance	4
Great Importance	5
Between Great and Very Great Importance	6
Very Great Importance	7
Between Very Great and Extremely Great Importance	8
Extremely Great Importance	9

The comparison stands for the relative importance of a certain criterion in relation to another one, in order to achieve that specific aim or in relation to another criterion at a higher level.

The levels of relative importance determined by the experts should be checked for inconsistency. If there is an inconsistency out of permitted boundaries, then the estimates should be reviewed. This is done by calculating the following ratio (Saaty, 1980)

$$CR = \frac{CI}{RI} \tag{1}$$

- CR – Consistency Ratio,
- CI – Consistency Index,
- RI – Random Index – defined according to the number (n) of objects compared in the following table (Panayotova, 2004):

n	2	3	4	5	6	7	8	9	10
RI	0	0,58	0,90	1,12	1,24	1,32	1,41	1,45	1,49

$$CI = \frac{\lambda - n}{n - 1} \tag{2}$$

- λ – Largest Eigen Value

The comparisons are considered to be consistent, when $CI < 0,10$. (Handfield et al., 2002).

In order to demonstrate the applicability of AHP in CODP positioning we suggest the following example:

A choice must be made among three alternatives for CODP: **POSITION 1** (situated in the left part of Figure 1), **POSITION 3** (situated in the right part), and **POSITION 2** – situated in the middle.

As a result of the expert group discussions on criteria 3.1 to 3.4, the following matrix for comparing their importance pair wise is suggested:

Table 2: Comparing criteria pair wise with respect to the objective.

Criterion	3.1	3.2	3.3	3.4
3.1	1	1,000	0,333	2,000
3.2	1,000	1	0,500	2,000
3.3	3,000	2,000	1	4,000
3.4	0,500	0,500	0,250	1

After normalizing and calculating the first normalized principal Eigen vector, we come to the following distribution of priorities of the criteria:

Table 3: Calculating criteria priorities with respect to the objective.

Criterion	2.1	2.2	2.3	2.4	Priority
2.1	0,182	0,222	0,160	0,222	0,197
2.2	0,182	0,222	0,240	0,222	0,217
2.3	0,545	0,444	0,480	0,444	0,479
2.4	0,091	0,111	0,120	0,111	0,108

Calculations show that Consistency Ratio is within the limit: $CR = 0.0076 < 0.1$

What follows is to determine the relative importance among the pairs of sub-criteria for each criterion. A view on the present example is shown in Tables 4 to 7:

Table 4: Comparing sub-criteria 2.1.1 – 2.1.5 pair wise with respect to the criterion 2.1. Inconsistency: 0.071.

Sub-Criteria	2.1.1	2.1.2	2.1.3	2.1.4	2.1.5	Priority
2.1.1	1	3,000	4,000	3,000	4,000	0,433
2.1.2	0,333	1	2,000	2,000	3,000	0,208
2.1.3	0,250	0,500	1	2,000	3,000	0,156
2.1.4	0,333	0,500	0,500	1	4,000	0,140
2.1.5	0,250	0,333	0,333	0,250	1	0,063

Table 5: Comparing sub-criteria 2.2.1 – 2.2.5 pair wise with respect to the criterion 2.2. Inconsistency: 0.049.

Sub-Criteria	2.2.1	2.2.2	2.2.3	2.2.4	2.2.5	Priority
2.2.1	1	2,000	4,000	3,000	5,000	0,410
2.2.2	0,500	1	3,000	2,000	4,000	0,258
2.2.3	0,250	0,333	1	2,000	4,000	0,158
2.2.4	0,333	0,500	0,500	1	2,000	0,114
2.2.5	0,200	0,250	0,250	0,500	1	0,060

Table 6: Comparing sub-criteria 2.3.1 – 2.3.5 pair wise with respect to the criterion 2.3. Inconsistency: 0.042.

Sub-Criteria	2.3.1	2.3.2	2.3.3	2.3.4	2.3.5	Priority
2.3.1	1	2,000	2,000	2,000	3,000	0,331
2.3.2	0,500	1	2,000	2,000	4,000	0,263
2.3.3	0,500	0,500	1	2,000	3,000	0,190
2.3.4	0,500	0,500	0,500	1	3,000	0,146
2.3.5	0,333	0,250	0,333	0,333	1	0,070

Table 7: Comparing sub-criteria 2.4.1 – 2.4.7 pair wise with respect to the criterion 2.4. Inconsistency: 0.086.

Sub-Cr.	2.4.1	2.4.2	2.4.3	2.4.4	2.4.5	2.4.6	2.4.7	Priority
2.4.1	1	3,00	2,00	2,00	4,00	5,00	4,00	0,335
2.4.2	0,33	1	2,00	2,00	3,00	4,00	4,00	0,210
2.4.3	0,50	0,50	1	2,00	4,00	4,00	4,00	0,185
2.4.4	0,50	0,50	0,50	1	3,00	4,00	4,00	0,130
2.4.5	0,25	0,33	0,25	0,33	1	3,00	4,00	0,057
2.4.6	0,20	0,25	0,25	0,25	0,33	1	2,00	0,040
2.4.7	0,25	0,25	0,25	0,25	0,25	0,50	1	0,043

4.2 Determining the Relative Importance of the Alternatives for CODP Position

After defining specific weights for the particular sub-criteria, a pair wise assessment of the three alternatives for CODP position is made, according to each sub-criterion. An example for the sub-criterion 2.1.1 is shown on the Table 8:

Table 8: Comparing three alternatives for CODP position pair wise with respect to the sub-criterion 2.1.1.

Alt.	Position 1	Position 2	Position 3	Priority
Position 1	1	2,000	4,000	0,571
Position 2	0,500	1	2,000	0,286
Position 3	0,250	0,500	1	0,143

In that particular case, the above calculation is repeated 22 times with respect to each sub-criterion, followed by weighting these estimates to the corresponding weights of the main criteria. The summary results are presented in Figure 4.

According to the results, the first alternative for the position of CODP (the one on the left-hand side of Figure 1) is most preferable and best fits the criteria and considerations formulated in section 3!

5 CONCLUSIONS

In this paper, an approach to CODP positioning was presented, using the Analytical Hierarchy Process/AHP. We summarized different approaches and existing methods to solving the problem. An in-depth analysis of some lacks in the literature in this direction enabled us to propose a different standpoint for the way of considering and solving the problem. The approach suggested here requires the decision to be made into two stages: (1) to build up a complete system of decision making criteria about CODP position and (2) to use the AHP in its capacity of a method/set of tools, by means of which the decision

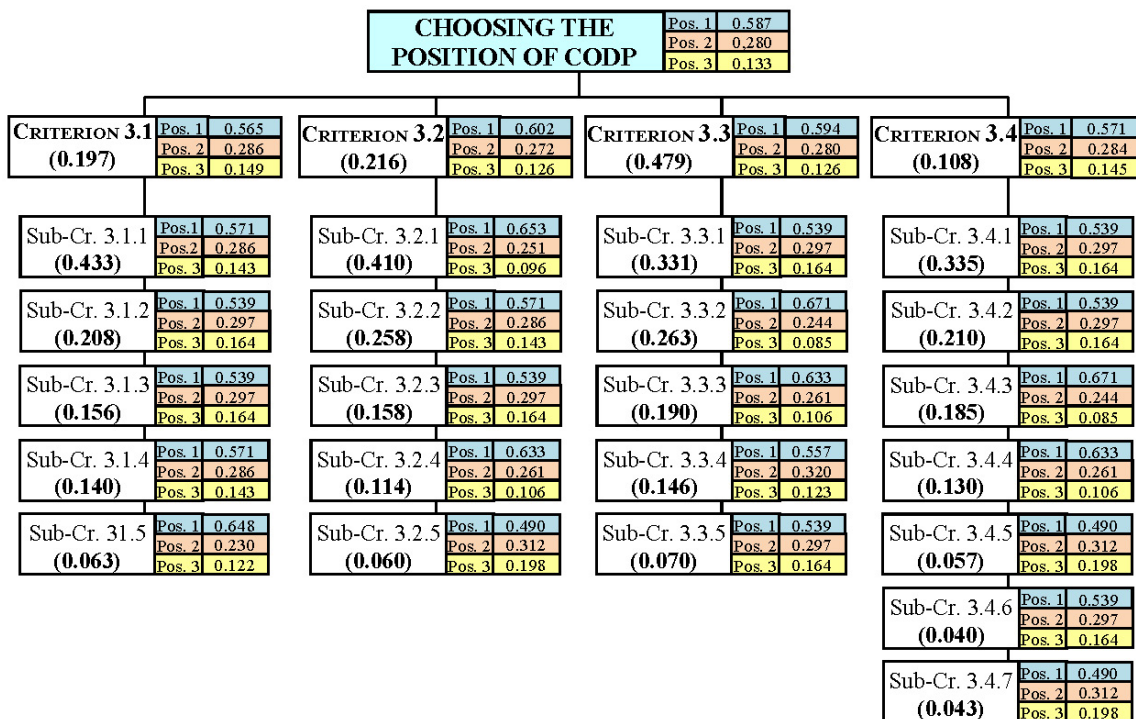


Figure 4: Comparing three positions with respect to the whole set of criteria and sub-criteria.

will be based on a large number of such criteria and sub-criteria.

In order to illustrate the approach, a hypothetical example was elaborated, based on the assumption of having three alternatives for the position of CODP.

First, a system of criteria and sub-criteria was build up, which reflects the particular circumstances and conditions influencing the hypothetical object under consideration, i.e. the specific branch of the company, the specific product/service, the specific competitive conditions at the market, financial considerations etc. Secondly, by using the AHP, these three alternative positions have been assessed with respect to the whole set of criteria and sub-criteria, and the one that fits best the criteria, was chosen.

Furthermore, it is necessary to make a vaster study in order to develop a method, which can consider the multi-positioning opportunities for CODP in the same product breakdown structure, as well as probabilistic behaviour of the units that manufacture/supply corresponding components.

REFERENCES

Andreev O., 2009. *Modern Production and Operations Systems – A Concept for Achieving Lean Mass Customization*, Softtrade. Sofia.

Feitzinger E., H. Lee, 1997. Mass Customization at Hewlett Packard: the Power of Postponement. *Harvard Business Review*, 75, 116-121.

Gupta D., S. Benjaafar, 2004. Make-to-Order, Make-to-Stock, or Delay Product Differentiation? A Common Framework for Modeling and Analysis. *IIE Transactions*, 36, 529-546.

Handfield, R., Walton, S., Sroufe, R. & Melnyk, S., 2002. Applying Environmental Criteria to Supplier Assessment: A Study in the Application of the AHP. *European Journal of Operational Research*, 141/1, 70-87.

Olhager J., 2001. Strategic Positioning of the Order Penetration Point. Linköping Institute of Technology. http://www.renard2kanak.eu/documents/Supply_Chain_Managment_-_Readings_-_Strategic_Positioning_-_A200.pdf

Panayotova T., 2004. *Organizational Connections Modeling in Complex Reengineering Projects*. Dissertation thesis. Sofia.

Pine II B. J., 1992. *Mass Customization: The New Frontier in Business Competition*. Harvard Business School.

Ramachandran K., L. Whitman & A. B. Ramachandran, 2002. *Criteria for determining the push – pull boundary*. Wichita State University, Kansas, USA.

Saaty, 1980. *The Analytic Hierarchy Process*, McGraw Hill.

Stevenson W., C. Ozgur, 2007. *Introduction to Management Science with Spreadsheets*. McGraw Hill.

Velev M., A. Tsvetanova, 2010. *Mass Customization – Strategy for Competitiveness Improving*, Softtrade. Sofia.

DEVELOPING THE BUSINESS MODELLING METHOD

L. O. Meertens, M. E. Iacob and L. J. M. Nieuwenhuis
University of Twente, PO Box 217, Enschede, The Netherlands
{l.o.meertens, m.e.iacob, l.j.m.nieuwenhuis}@utwente.nl

Keywords: Business modelling, Modelling method, Business models.

Abstract: Currently, business modelling is an art, instead of a science, as no scientific method for business modelling exists. This, and the lack of using business models altogether, causes many projects to end after the pilot stage, unable to fulfil their apparent promise. We propose a structured method to create “as-is” business models in a repeatable manner. The method consists of the following steps: identify the involved roles, recognize relations among roles, specify the main activities, and quantify using realistic estimates of the model. The resulting business model is suitable for analysis of the current situation. This is the basis for further predictions, for example business cases, scenarios, and alternative innovations. We offer two extra steps to develop these innovations and analyse alternatives. Using them may enable successful projects to be implemented, instead of ending on a shelf after the pilot stage.

1 INTRODUCTION: BUSINESS MODELLING BACKGROUND

A business model is critical for any company, and especially for any e-business. Its importance has been recognized over the past few years by several authors that have created different business model frameworks aimed at identifying the main ingredients of a business model (for example, Osterwalder (2004); for an overview, see Pateli & Giaglis (2004), and Vermolen (2010)). However, the state in which this field finds itself is one of “prescientific chaos” (Kuhn 1970): there are several competing schools of thought, and progress is limited because of a lack of cumulative progress. Because of this, there are no clear and unique semantics in the research related to business models. The very concept of “business model” is associated with many different definitions (Vermolen 2010). The elements of such a business model differ significantly from one approach to another. Furthermore, to the best of our knowledge, there are no methodological approaches in the literature for the design and specification of business models (Vermolen 2010). This lack of cohesion in the field clearly diminishes the added value of business models for companies and makes business modelling an art, rather than a science. This state of affairs motivated us to propose such a method, which enables the development of business models

in a structured and repeatable manner. Thus the contribution of this paper is three-fold:

- A business model development method;
- A definition of the concept of business model and the identification of its core elements, captured by the deliverables of the method steps;
- An illustration of the method by means of a case study from the healthcare domain.

The paper is structured as follows. Section 2 focuses on the discussion of the main concepts addressed in the paper, and positions our approach with respect to the existing design science and method engineering literature. In Section 3, we describe the steps of our business model development method. In Section 4, we demonstrate the method by means of a case study concerning the development of a qualitative business model of the elderly care in The Netherlands. Finally, we conclude our paper and give pointers to future work in Section 5.

2 THEORETICAL BACKGROUND

A simple analysis of the two words “business model” already gives an idea of what a business model is about. On the one hand, there is “business”: the way a company does business or creates value. On the other hand, there is “model”: a

conceptualization of something – in this case, of how a company does business.

We extend this common and simplistic interpretation of a business model as “the way a company earns money”, into a broader and more general definition of the concept: “a simplified representation that accounts for the known and inferred properties of the business or industry as a whole, which may be used to study its characteristics further, for example, to support calculations, predictions, and business transformation.”

The last part of the definition above, namely the indication of the possible uses of a business model is of particular importance in the context of this paper. The method we propose not only facilitates the development of such a design artefact - a business model - but also takes a business engineering perspective. Thus, its application will result in two (or more) business models: one that reflects the “as-is” situation of the business and one or more alternative “to-be” business models that represents possible modifications of the business as result of, for example, adoption of innovative technologies or more efficient business processes.

To the best of our knowledge, such a method does not exist yet for what we define as business models (Vermolen 2010). In the remainder of this section, we position our work in the contexts of design science and method engineering, to which it is related.

2.1 Design Science

A business modelling method can be seen as a design-science artefact. It is the process of creating a product: the business model. We use the seven guidelines of Hevner et al. (2004) to frame how we use the methodology engineering approach from Kumar & Welke (1992) to create our method.

The first guideline advises to design as an artefact. Design-science research must produce a viable artefact in the form of a construct, a model, a method, or an instantiation. As said, we produce a method.

The second guideline tackles relevance. The objective of design-science research is to develop technology-based solutions to important and relevant business problems. Viable business models lie at the heart of business problems. However, our solution is not yet technology-based. Partial automation of the method is left for future research.

The utility, quality, and efficacy of a design artefact must be rigorously demonstrated via well-executed evaluation methods. We demonstrate the

business modelling method using a case study. We leave more rigorous evaluation for further research.

Research contribution is the topic of the fourth guideline. Effective design-science research must provide clear and verifiable contributions in the areas of the design artefact, design foundations, and/or design methodologies. We provide a new artefact to use and study for the academic world. The methodology may be extended, improved, and specialized.

Guideline five expresses the scientific rigour: Design-science research relies upon the application of rigorous methods in both the construction and evaluation of the design artefact. We aim to be rigorous through using the methodology engineering approach. Existing, proven methods are used as foundation and methods where applicable. Evaluation was handled in the third guideline.

The sixth guideline positions design as a search process. The search for an effective artefact requires utilizing available means to reach desired ends while satisfying laws in the problem environment. Whenever possible, we use available methods for each of the steps. Following the methodology engineering approach helps us to satisfy the laws for creating a new methodology.

The final guideline instructs us to communicate our research. Design-science research must be presented effectively both to technology-oriented as well as management-oriented audiences. This article is one of the outlets where we present our research.

2.2 Methodology Engineering

Methodologies serve as a guarantor to achieve a specific outcome. In our case, this outcome is a consistent and better-informed business model. We aim to understand (and improve) how business models are created. With this understanding, one can explain the way business models help solve problems. We provide a baseline methodology only, with a limited amount of concepts. Later, we can extend, improve and tailor the methodology to specific situations or specific business model frameworks.

The business modelling method has both aspects from the methodology engineering viewpoint: representational and procedural (Kumar & Welke 1992). The representational aspect explains what artefacts a business modeller looks at. The artefacts are the input and deliverables of steps in the method. The procedural aspect shows how these are created and used. This includes the activities in each step, tools or techniques, and the sequence of steps.

3 DEFINING THE BUSINESS MODELLING METHOD

We define five individual step of business modelling, which the rest of this section elaborates. To describe each step, we use the following elements:

- inputs of the step,
- activities to perform during the step,
- possible techniques to use during the step's activities, and
- deliverables resulting from the step.

Each step in the proposed method requires specific methods, techniques or tools that are suitable for realizing the deliverables. We will mention examples of those. However, others may also be useful and applicable, and it is not our aim to be exhaustive in this respect. Table 1 shows an overview of our method.

3.1 Create As-Is Model

As mentioned in the previous section, our business model development method takes business engineering perspective. Thus, the first four steps of our method focus on creating a business model that reflects the current state of the business. Therefore, steps one through four create an as-is model.

3.1.1 Step 1: Identify Roles

Identifying the relevant parties (which we refer to as roles) involved in a business model should be done as systematically as possible. The aim is completeness in this case. The business modeller must carry out a stakeholder analysis, to identify all roles. The input to this step includes for example, documentation, domain literature, interviews, experience and previous research. The output is a list of roles.

For an example of stakeholder analysis method, we refer to Pouloudi & Whitley (1997). They provide an interpretive research method for stakeholder analysis aimed at inter-organizational systems, such as most systems where business modelling is useful. The method consists of the following steps:

1. Identify obvious groups of stakeholders.
2. Contact representatives from these groups.
3. (In-depth) interview them.
4. Revise stakeholder map.
5. Repeat steps two to four, until...

Pouloudi and Whitley do not list the fifth step, but mention that stakeholder analysis is a cumulative and iterative approach. This may cause the number of stakeholders to grow exponentially, and the question remains when to stop. Lack of resources may be the reason to stop the iterative process at some point. Closure would be good, but seems hard to achieve when the model is more complex. Probably, the modeller has to make an arbitrary decision. Nevertheless, one should choose stop criteria (a quantifiable measure of the stakeholder's relevance for the respective business model and a threshold for the measure) before starting the process (Pouloudi 1998).

“Revising the stakeholder map” (step four) could use extra explanation, which can be found in the description of the case Pouloudi and Whitley use to explain the method. The stakeholders gathered from interviews can be complemented with information found in the literature. The business modeller then refines the list of stakeholders by focussing, aggregating, and categorizing.

3.1.2 Step 2: Recognize Relations

The second step of our method aims to discover the relations among roles. The nature of these relations may vary substantially, but it always involves some interaction between the two roles and may assume some exchange of value of some kind. Much of the work and results from the previous step can be reused for this as input. In theory, all roles could have relations with all other roles. However, in practice, most roles only have relations with a limited number of other roles. Usually, these relations are captured in a stakeholder map, which often follows a hub-and-spoke pattern, as the focus is on one of the roles. This pattern may be inherent to the approach used, for example if the scope is defined as a maximum distance from the focal role.

To specify all relations, we suggest the use of a role-relation matrix with all roles on both axes as technique. Of this matrix, the cells point out all possible relations among the roles. Each of the cells could hold one or more relations between two roles. Assuming that roles have a limited number of relations, the role-relation matrix will be partially empty. However, one can question for each empty cell whether a relation is missing or not.

Cells above and below the diagonal can represent the directional character of relations. Usually, relations have a providing and consuming part. The providing part goes in the upper half of the matrix, and the consuming part in the bottom half. This

Table 1: Business Modelling Method.

Step	Inputs	Techniques or Tools	Deliverables
Identify Roles	Documentation, domain literature, interviews, experience, previous research	Stakeholder analysis (Pouloudi & Whitley 1997)	Role list
Recognize Relations	Role list, Stakeholder map, value exchanges	e3-value (Gordijn 2002)	Role-relation matrix
Specify Activities	Role-relation matrix, Role list, business process specifications	BPM methods, languages and tools	List of activities
Quantify Model	Process specifications, accounting systems and annual reports	Activity based costing	Total cost of the business "as-is"
Design Alternatives	As-is business model, Ideas for innovations and changes	Business modelling method (steps 1 to 4), Brainstorming	One or more alternative business models
Analyse Alternatives	Alternative business models	Sensitivity analysis, technology assessment, interpolation, best/worst case scenarios	Business case for each alternative

especially helps with constructions that are more complex, such as loops including more than two roles.

The output of this step is a set of relations.

3.1.3 Step 3: Specify Activities

For a first qualitative specification of the business model, the next step is to determine the main activities. Relations alone are not sufficient: the qualitative model consists of these main business activities (business processes) too. These activities originate from the relations identified in the previous step. Each of the relations in the role-relation matrix consists of at least one interaction between two roles, requiring activities by both roles. Besides work and results from the previous steps, existing process descriptions can be valuable input. Techniques from business process management may be used.

The output from these first three steps is a first qualitative business model, including roles, relations and activities. It reveals what must happen for the business to function properly.

3.1.4 Step 4: Quantify Model

Quantifying the business model helps us to see what is happening in more detail and compare innovations to the current situation. To turn the qualitative model into a quantitative model, numbers are needed. The numbers are cost and volume of activities (how often they occur). Together, these numbers form a complete view of the costs captured by the business model.

Several sources for costs and volumes are possible, such as accessing accounting systems or

(annual) reports. The resulting quantitative business model shows the as-is situation.

3.2 Develop To-Be Model

The as-is model, created in previous steps, is suitable for analysis of the current state only. However, from the as-is model, it is possible to derive alternatives. Such alternatives can be created to assess how reorganisations, innovations or other changes influence the business. These are the to-be models.

3.2.1 Step 5: Design Alternatives

From here on, we aim to capture a future state of the business in a business model. To make predictions, the model may need further instantiations. Each instantiation is an alternative development that may happen (to-be). Using techniques such as brainstorming and generating scenarios, business modellers create alternative, qualitative, future business models. These alternatives are used to make predictions. Usually, such alternatives are built around a (technical) innovation. This may include allocating specific roles to various stakeholders. A base alternative, which only continues an existing trend without interventions, may help comparing the innovations. Next to the business model, ideas for innovations serve as input. The resulting alternative business models show future (to-be) possibilities.

3.2.2 Step 6: Analyse Alternatives

The final step for a business modeller is to analyse the alternative business models. Besides the qualitative business models, several sources of input

are possible to quantify the alternatives. Applicable techniques include sensitivity analysis, technology assessment, interpolation and using best/worst case scenarios. Each alternative can be tested against several scenarios, in which factors change that are not controllable. We can use the models to predict the impact. This step and the previous one can be repeated several times to achieve the best results. The final output is a business case (including expected loss or profit) for each alternative.

4 THE U*CARE CASE: DEMONSTRATING THE BUSINESS MODELLING METHOD

U*Care is a project aimed at developing an integrated (software) service platform for elderly care (U*Care Project n.d.). Due to the aging population and subsequently increasing costs, elderly care - and healthcare in general - is one of the areas where governments fund research. However, many projects never get further than pilot testing. Even if the pilot is successful, the report often ends on a shelf. By applying the business modelling method, we plan to avoid this and put the U*Care platform into practice. Specifically, we aim to show how the technological innovations built in the U*Care platform will influence the business model for elderly care.

The U*Care case is in progress currently, and therefore, it is suitable for demonstrating the first three steps of the business modelling method. These three steps result in a qualitative business model of the as-is situation of the elderly care. At this time, we are collecting quantitative data. Therefore, steps four to six of the method are not yet possible and we will not address them here. However, they will be subject of future work.

4.1 Identify Roles

The first step of the stakeholder analysis, led to the identification of several groups of obvious stakeholders. The groups include all the project partners, as their participation in the project indicates their stake. Another group includes the main users of the platform: the clients and employees of the elderly care centre.

After identifying the obvious stakeholders, we contacted and interviewed representatives from all the project partners and several people in the care

centre. These interviews did not explicitly focus on stakeholder analysis, but served as a general step in requirements engineering. Table 2 displays a partial list of identified stakeholders after steps two and three of Pouloudi and Whitley's method for stakeholder analysis have been performed.

Table 2: Partial list of stakeholders after step three of Pouloudi and Whitley's method for stakeholder analysis.

Clients	Care (& wellness) providers
Volunteer aid	Hospitals
Nurses	Elderly care centres
Doctors	Psychiatric healthcare
Administrative employees	Homecare
General practitioners	Technology providers
Federal government	User organizations
Local government	Insurance companies

The fourth step includes a search for stakeholders in the literature. Besides identifying the extra stakeholders, the literature mentioned the important issue that some actors in the list are individual players, while other actors are organizations or other forms of aggregations (groups). Consequently, overlap can occur in the list of actors.

The final action of the first iteration is not a trivial one. Refining the stakeholder list requires interpretation from the researcher. Different stakeholder theories (for example, E. J. Emanuel & L. L. Emanuel (1996), J. Robertson & S. Robertson (2000), and Wolper (2004)) act as tools to minimize subjectivity.

The long list of identified stakeholders is not practical to continue with and has much overlap. Therefore, we grouped the stakeholders into a limited set of roles, shown in Table 3. This set of high-level roles is an interpretive choice. The small set helps to keep the rest of case clear instead of overcrowded. The larger set is kept in mind for the to-be situation to find potential "snail darters": stakeholders with only a small chance of upsetting a plan for the worse, but with huge results if they do (Mason & Mitroff 1981). The small set of stakeholders was subject to prioritization based on Mitchell et al. (1997). While the prioritization is subjective, it shows that all roles in the list are important.

4.2 Recognize Relations

The current situation consists of five categories of interacting roles. Table 3 shows them on both axes. The cells show relations between the roles. While

Table 3: Role-relation matrix.

Consumer Provider	Care consumers	Care providers	Technology providers	Government	Insurers
Care consumers	X	Pay for care		Pay for AWBZ Pay for WMO	Pay for insurance
Care providers	Provide ZVW care Provide WMO care Provide AWBZ care	X	Pay for (use of) technology or service	Provide care to citizens	Provide care to insured
Technology providers		Provide technology or service	X		
Government	Provide AWBZ insurance Provide WMO insurance	Pay for WMO care to citizens		X	Pay for AWBZ care to citizens
Insurers	Provide insurance Refund AWBZ and ZVW care			Ensure AWBZ care for citizens	X

the care provider has relations with all the other roles, it is not a clear hub-and-spoke pattern. Several of the other roles have relations outside the care provider.

The relations show that a main goal of the business is to provide care to the care consumer. The insurers and government handle much of the payment. Other (regulating) roles of the government remain out of scope, as the case is complex enough as it is.

The insurers handle most of the payments. The patient has to pay the care provider after receiving care. The patient can then declare the costs to the insurance company, which refunds the patient. The patient pays a premium to the insurance company. According to the Dutch Healthcare Insurance Act (*Zorgverzekeringswet, ZVW*), every citizen has to have basic care insurance (*ZVW*). For “uninsurable care” (including most home healthcare, similar to USA Medicare), the Dutch government set up a social insurance fund, termed General Exceptional Medical Expenses Act (*Algemene Wet Bijzondere Ziektekosten, AWBZ*). All employees and their employers contribute towards this fund. The AWBZ is similar to the regular insurance companies, except for collecting the premium. The premium is paid through taxation by the government, which outsources most of the further actions to insurers. A similar system is set up for wellness homecare, such as cleaning. This is the Social Support Act (*Wet Maatschappelijke Ondersteuning, WMO*). In contrast to the AWBZ, the government takes care of all actions itself, through its municipalities.

Several issues exist, which we do not handle in detail here. For example, it is inherent to insurance

that not all people who pay premium are also (currently) care consumers.

4.3 Determine Behaviour

Most of the relations between the roles in Table 3 start with verbs. This signals that they are (part of) behaviour. Any relation not beginning with a verb is a candidate for rephrasing or being split into smaller parts.

Besides the relations, we focussed on AWBZ to identify the main activities of the care providers. “Providing care” has four top-level functions: personal care, nursing, guidance/assistance, and accommodation. Each of these functions consists of many detailed activities, of which Table 4 provides an example.

We obtained these activities from documents made available by the government for reimbursement purposes (Ministry of Health, Welfare and Sport 2008). As it also provides an indication of volume (times a day) and a first indication of costs (minutes spent), it is a first step towards quantifying the model.

4.4 Further Steps

The other three steps are not possible for now, as detailed quantitative data is not available yet. However, when we have developed a quantitative business model, the final steps of the method help to generate business cases easily for various services that could be integrated into the U*Care platform. This should proof their viability and lead to implementation, instead of ending on the shelf.

Table 4: Example of personal care activities.

Activity	Actions	Time in minutes	Frequency per day
Washing	Whole body	10	1x
	Parts of body	20	1x
Dressing	(Un)dress completely	15	2x
	Undress partially	10	1x
	Dress partially	10	1x
	Put on compression stockings	10	1x
	Take off compression stockings	7	1x
Getting in and out of bed	Help getting out of bed	10	1x
	Help getting into bed	10	1x
	Help with afternoon rest (for example, get onto the couch)	10	1x
	Help with afternoon rest (for example, get off the couch)	10	1x
Eating and drinking	Help with eating cold meals (excluding drinking)	10	2x
	Help with eating warm meal (excluding drinking)	15	1x
	Help with drinking	10	6x

5 CONCLUSIONS: A FUTURE FOR BUSINESS MODELLING

Three contributions are made in this paper. Primarily, we created a business model development method. Secondly, we defined the concept of business model and identified its core elements, captured by the deliverables of the method steps. Finally, we demonstrated the method by means of a case study from the healthcare domain.

The business modelling method provides a way to create business models. Innovators can apply the steps to create business cases for their ideas systematically. This helps them to show the viability and get things implemented.

We provide a new design-science artefact to use and study for the academic world. As business modelling has several goals, conducting only the first few steps may be enough. For example, if your goal is to achieve insight in the current state only, the last two steps are not useful.

The business modelling method may be extended. Situational method engineering seems suitable for this (Henderson-Sellers & Ralyté 2010). For example, for information system development, it is interesting to research if steps towards enterprise architecture can be made from business models. This can be seen as a higher-level form of, or preceding step to, the BMM proposed by Montilva and Barrios (2004). On the other side, a step could be added before identifying roles. Other domains require different improvements.

In addition, the steps in the method can be further specified. The steps can be detailed further. One of the ways to do this is to tailor the techniques at each of the steps of this method. In the future, new tools and techniques may help provide partial automation.

ACKNOWLEDGEMENTS

This work is part of the IOP GenCom U-CARE project, which the Dutch Ministry of Economic Affairs sponsors under contract IGC0816.

REFERENCES

- Emanuel, E. J. & Emanuel, L. L., 1996. What is Accountability in Health Care? *Annals of Internal Medicine*, 124, pp.229-239.
- Gordijn, J., 2002. *Value-based Requirements Engineering: Exploring Innovative e-Commerce Ideas*. PhD Thesis. Vrije Universiteit Amsterdam.
- Henderson-Sellers, B. & Ralyté, J., 2010. Situational Method Engineering: State-of-the-Art Review. *Journal of Universal Computer Science*, 16(3), pp.424-478.
- Hevner, A. R. et al., 2004. Design Science in Information Systems Research. *MIS Quarterly*, 28(1), pp.75-105.
- Kuhn, T. S., 1970. *The Structure of Scientific Revolutions*, University of Chicago Press.
- Kumar, K. & Welke, R. J., 1992. Methodology Engineering: a proposal for situation-specific methodology construction. In *Challenges and*

- Strategies for Research in Systems Development*. John Wiley series in information systems. Chichester: Wiley, p. 257–269.
- Mason, R. O. & Mitroff, I. I., 1981. *Challenging strategic planning assumptions : theory, cases, and techniques*, Wiley.
- Ministry of Health, Welfare and Sport, 2008. Beleidsregels indicatiestelling AWBZ 2009.
- Mitchell, R. K., Agle, B. R. & Wood, D. J., 1997. Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. *Academy of management review*, pp.853-886.
- Montilva, J. C. & Barrios, J. A., 2004. BMM: A Business Modeling Method for Information Systems Development. *CLEI Electronic Journal*, 7(2).
- Osterwalder, A., 2004. *The Business Model Ontology - a proposition in a design science approach*. PhD Thesis. Universite de Lausanne.
- Pateli, A. G. & Giaglis, G. M., 2004. A research framework for analysing eBusiness models. *European Journal of Information Systems*, 13(4), pp.302-314.
- Pouloudi, A., 1998. *Stakeholder Analysis for Interorganisational Information Systems in Healthcare*. PhD Thesis. London School of Economics and Political Science.
- Pouloudi, A. & Whitley, E. A., 1997. Stakeholder identification in inter-organizational systems: gaining insights for drug use management systems. *European Journal of Information Systems*, 6, p.1–14.
- Robertson, J. & Robertson, S., 2000. *Volere: Requirements specification template*, Technical Report Edition 6.1, Atlantic Systems Guild.
- U*Care Project, U*Care website. Available at: <http://www.utwente.nl/ewi/ucare/> [Accessed May 2, 2011].
- Vermolen, R., 2010. Reflecting on IS Business Model Research: Current Gaps and Future Directions. In *Proceedings of the 13th Twente Student Conference on IT*. Twente Student Conference on IT. Enschede, Netherlands: University of Twente.
- Wolper, L. F., 2004. *Health care administration: planning, implementing, and managing organized delivery systems*, Jones & Bartlett Publishers.

THE ACCOUNTING CONCEPT AS KEY CONCEPT FOR BUSINESS MODELLING

Coen Suurmond

*RBK Group, Keulenstraat 18, 7418 ET Deventer, The Netherlands
csuurmond@rbk.nl*

Keywords: Business modelling, Software architecture, Accounting.

Abstract: Many problems in IT systems can be traced back to two misconceptions: that all information comes from these systems and that the information from these systems represents reality "as is". IT systems are often closed and inflexible. Modelling accounting subsystems, as closely as possible to operational practice, can prevent these problems. Current software architecture provide technical support for this approach. In modelling the formalisation and standardisation of the use of language is an important issue.

1 INTRODUCTION

The development, implementation and operational use of information systems within enterprises face certain persistent problems. These problems arise from misconceptions about the nature and role of information systems within an enterprise. I refer here firstly to the misconception that the role of the computer-based information systems is to cover all of the information supply within an enterprise. And I refer here secondly to the misconception that it is in the nature of an information system to represent reality "as is".

These misconceptions shape the analysis and the design of IT systems mostly as background assumptions, and are undiscussed. The nature and the role of information outside of the IT systems are often underexposed in IT projects, and such information is considered peripheral to these systems. Once an IT system is in use, the representations in the system are often considered as leading. E.g., a sales order is no longer seen as an agreement between a buyer and a seller; a sales order is defined by its representation in the system.

The result of these misconceptions is that IT systems often do not function properly in practice. Unforeseen limitations are imposed on business processes and on the commercial and logistic opportunities of the enterprise, as a consequence of the use of IT systems. Limitations that might cost money internally and that trade externally.

This can produce various reactions that aim to absorb these unforeseen negative consequences. First, the system can be expanded, to remove, to evade or to circumvent the limitations. Second, practice might create its own way of working and its own auxiliary systems for certain processes, alongside the IT systems that were meant to do this. Third, practice might learn to live with the limitations.

At the basis of any solution for the problems at issue must be the awareness that it is the task of computer-based information systems to facilitate business processes. These systems thus have an instrumental nature and a serving role. They cannot be the Archimedean point from which business processes are controlled. Neither should these systems determine how the business must be done.

To meet these demands IT systems should be open in three meanings of the word: (1) they have to be able to collaborate with other systems outside of their own functional areas, (2) they have to be able to collaborate with other systems within their own functional areas, when these systems have a lower-level execution task or a higher-level controlling task, and (3) they have to be able to deal with events within their functional area that (temporarily) circumvent the system.

If we want to achieve this openness, then I think it is useful, if not unavoidable, to exploit the concept of "accounting" as a core concept in the design of information systems. The accounting concept analyses the accounting processes separated from other processes, its operations are based on clearly

defined services and it implies clearly demarcated organisational responsibilities. Although the concept originates in the financial domain, it is very well applicable outside of the financial field; indeed, the concept should hold for all information that is used communally in an organisation.

2 ORGANISATIONAL ASPECTS OF ACCOUNTING PROCESSES

2.1 Division of Labour, Coordination, Responsibility

Mintzberg defines the structure of an organisation as “the sum total of the ways in which divides its labour into distinct tasks and then achieves coordination among them” (Mintzberg, 1979). He distinguishes five different kinds of coordination mechanisms, namely: (1) mutual adjustment, (2) direct supervision, (3) standardisation of work, (4) standardisation of outputs, and (5) standardisation of skills. From an organisational point of view an accounting department can be considered both as a specialisation of labour and as a coordination mechanism.

Accounting departments are often specialised in the processing of either financial data, employee data or production data. The justification of such departments in an organisation is in the required competencies of the employees, as well as in the sensitivity of data. The departments have to meet a variety of requirements of a variety of stakeholders. They have to meet both the internal requirements of the operational and management processes, and they have to meet the requirements of external stakeholders. The head of the department is responsible for the quality of the information supplied.

Accounting processes might also perform coordinating roles in an organisation. When the meaning and use of accounting data are clearly defined, and when the accounting processes are specified as well, this will in effect be a coordination mechanism by standardisation of output and by standardisation of skills. The required skills are related to the ways and means of collecting and verifying the data in the operational processes, and (equally important) to the interpretation of the information supplied.

2.2 Organisational Meaning

Directing information flows through an accounting

department or an accounting system also implies formalising the language used. The language used within an organisation is a mixture of everyday speech, jargon, and forms of more or less formalised language. Different departments can have different interpretations of the same concept. If an order is delivered to a customer in a truck with trailer, there is one shipment. If two trucks arrive together for the same delivery, does this involve one or two deliveries, and one or two shipments? If the customer demands that an order is delivered as a whole, then the answer is clear for the commercial department: in both cases one shipment is made. For the freight documents it is clear as well: in the first case one shipment is made with the accompanying freight documents and in the second case two shipments are made, each with their own documents. For the receiving DC of the customer it is also clear: in both cases two deliveries are made that must be docked and unloaded separately.

In practice people within an organisation use different kinds of sign systems simultaneously. The everyday use of language is fairly free and unconfined, even within the context of an organisation. For commercial and financial transactions the language used is more formalised and is ultimately grounded in written law and case law. The use of automated systems is another form of formalisation of sign use. Part of it is formatting (type and size of the fields), part of it is predefined categorisation (tick the correct box) and part of it is capturing some part of the organisational reality in IT artefacts.

To implement an accounting system implies the advance creation of conventions governing the terminology, relations and meaning. In this sense it is a formalisation of the use of language. It can also be considered as a coordination mechanism in the organisation: standardisation of meaning.

3 THE ACCOUNTING CONCEPT

3.1 Definitions

Starreveld arrives in his original work in the early 60s about information processes in organisations at this definition of accounting: “The systematic collection, recording, processing and supplying of information for purposes of the managing and functioning of a household and for purposes of the accountability thereof” (Starreveld, 1963). The American Accounting Association defines accounting as follows: “the process of identifying,

measuring and communicating economic information to permit informed judgements and decisions by users of the information". The definition of the AICPA in 1961: "Accounting is the art of recording, classifying and summarizing in a significant manner and in terms of money, transactions and events which are, in part at least, of a financial character, and interpreting the result thereof" (Glautier et.al., 2001). The value of these definitions is that they sketch a clear and normative view of the role of accounting in the business processes.

From the FASB, part 2: "The purpose of this Statement is to examine the characteristics that make accounting information useful. ...All financial reporting is concerned in varying degrees with decision making (though decision makers also use information obtained from other sources). ...The usefulness of information must be evaluated in relation to the purposes to be served, and the objectives of financial reporting are focused on the use of accounting information in decision making ... Even objectives that are oriented more towards stewardship are concerned with decisions. Stewardship deals with the efficiency, effectiveness, and integrity of the steward. To say that stewardship reporting is an aspect of accounting's decision making role is simply to say that its purpose is to guide actions that may need to be taken in relation to the steward or in relation to the activity that is being monitored."

3.2 Analysis and Expansion of the Definitions

It is clear that the term accounting traditionally has a strong connection to the financial management and to internal and external financial reporting. At the same time it becomes clear from the definition of Starreveld that this financial aspect is not an essential element of the definitions given. The essential elements of the definitions are: (1) the systematic nature of the collecting, processing and making available of data, (2) the separation of the collecting, processing and making available of data on the one side and the interpretation of the data on the other side, (3) the usage for operational decisions, and (4) the usage for internal and external reporting, analysis and accountability.

Regarding the systematic nature I would like to draw attention to an element that is only found explicitly in the definition of the AICPA: classifying and summarising. Boisot has made an analysis of the nature of information and he defines three aspects of

information (Boisot, 1998). One aspect concerns the extent to which the information has been codified, a second aspect concerns the degree of abstraction, and the third aspect concerns the degree of diffusion. In an accounting system information will have to be codified, the users request information at different levels of abstraction and an accounting system is a mechanism for diffusion both within and outside of the organisation.

4 OPENNESS OF INFORMATION SYSTEMS

4.1 Two Forms of Information Supply

A car navigation system is an example of an open system that gives its user the right information while leaving him the freedom to take his own decisions. Such a system continuously indicates the route to be followed, taking into account the actual position, possible routes and traffic intensity on the different routes.

This information system is open because it allows the driver to take his own decisions. He can divert from the route whenever he opts to do so. The navigation systems always works from the actual situation; what happened before and what motivates the choices of the driver are completely irrelevant.

To describe a route by step-by-step directions is quite a different story. See for example the directions given by the British AA for the trip from Corrie on Arran to Bridgend on Islay:

<u>Direction</u>	<u>Miles</u>
Start out on unnamed road	0.00
Turn left onto the A481	0.07
Turn right	8.78
Continue by vehicle ferry (Claonaig – Lochranza)	8.81
Turn right	13.47
Turn left onto the B8001	13.57
Turn right onto the A83 (signposted Glasgow)	18.63
Turn left (signposted Islay ferry)	19.00
Continue by vehicle ferry (Port Askaig – Kennacraig)	19.25
Turn left onto the A846	48.82
Arrive on the A846	56.66

Section time 6:16, Total time 6:16

Such a description loses a lot of its value when the driver is forced off the prescribed route. Apart

from that, these specific directions do not take into account that the driver should take either the ferry to Port Askaig (as described), or the ferry to Port Ellen (not described), depending on the time. And checking the miles travelled against the odometer of the car won't work either, because of the two ferries involved.

4.2 Two Forms of Information Supply

Starreveld writes in 1962 about information supply for making judgements *ex ante* (for decisions in the execution of processes) and *ex post* (for the accountability of processes) (Starreveld, 1963). He emphasises that on lower levels of the hierarchy there is mostly a need for all kinds of "grab"-information that is necessary to guarantee a correct and efficient execution of the tasks by good preparation.

Especially in production organisations there is a tendency to increasingly emphasise a cycle of planning and control where plans are made higher-up in the organisation and executed lower-down and where the results are reported back. This tendency is reflected in ERP systems with their modules for planning and control. The shop floor gets production orders to be executed, and can only report back in relation to those production orders. Registration of unplanned activities is difficult or even impossible, even if some problem on the shop floor made those unplanned activities necessary.

Compare this tendency with this quote from Robert Anthony: "Several authors state that the aim of control is to assure that the results of operations conform as closely as possible to plans. We emphasize that such a concept of control is basically inconsistent with the concept used in this study. To the extent that middle management can make decisions that are better than those implied in the plans, top management wishes it to do so. And the middle management can in fact make better decisions under certain circumstances; to deny this possibility is implicitly to assume that top management is either clairvoyant, or omniscient, or both, which is not so." (Anthony, 1965)

Current ERP systems could be compared to the directions given by the AA. The individual steps are determined in detail in advance of the execution of a process and employees in the execution of the process get information pushed about the steps to be taken. This way of working is vulnerable in case of deviations, the responsibilities are higher-up and those that perform the tasks are regarded as being just cogs in the machine. The approach described by

Starreveld and Anthony assumes a situation in which employees at every hierarchical level of the organisation have tasks and responsibilities and these are not frustrated by centralised and bureaucratic information systems. They can get the information they need whenever they ask for it, and they can make their own decisions within their domain. As a model this is more comparable to driver assisted by a navigation system.

5 SOFTWARE DEVELOPMENT

5.1 Software Engineering

Like organisations, software engineering has its methods for managing complexity. The multi-tier model, the client/server model and the service oriented architecture model are three examples of the principle of 'separation of concerns'. Earlier forms are the concept of structured programming, employing units (Pascal) or modules (Modulo) and later on object-oriented programming.

The mentioned mechanisms in software engineering are each directly concerned with the structure of the software as such. They do not or only tangentially concern themselves with how the software is used. In the last few decennia the discussion is increasingly about architecture. At first it was about the architecture of software, then about the architecture of information systems and finally about the architecture of the enterprise.

Within the software engineering as such Taylor states "By architecture we mean the set of principle design decisions made about a system; it is a characterization of the essence and essentials of the application" (Taylor et al., 2010). The architecture of this artefact consists of a number of more-or-less independent parts and the connections between them (static structure). Further there is a specific way in which the communication between the parts happens (dynamic structure). Both for the static and for the dynamic structure the architect makes use of a repertoire of standard patterns. This manner of working has first been charted by the architect Christopher Alexander and later on has spread widely within software engineering.

However, that similar mechanisms for managing complexity have been developed both in organisations and in software engineering does not mean that the mechanisms of both worlds should be considered equal. In the case of software we are dealing with a strictly formal and determined

system, whereas in the case of an organisation we are dealing with a social system.

5.2 RM-ODP (Reference Model for Open Distributed Processing)

The RM-ODP describes a model for the collaboration of interconnected autonomous, heterogeneous information processing systems. "The objective of ODP standardization is the development of standards that allow the benefits of distributing information processing services to be realized in an environment of heterogeneous IT resources and multiple organizational domains. These standards address constraints on system specification and the provision of a system infrastructure that accommodate difficulties inherent in the design and programming of distributed systems." (RM-ODP, 1998).

In the formulated aims of the RM-ODP it can be seen that accounting systems as described above fit beneath the heading "organisational" in such an IT structure. At the same time we must realise that a technical infrastructure cannot say anything about the organisational responsibilities and that the IT component often must be complemented with a human component for a complete information supply. It is ill conceivable that the financial department and the employee department would be fully automated and unmanned departments.

5.3 Software Architecture & Accounting Subsystems

The modern developments in software engineering, coupled with architectural ideas such as those expressed in RM-ODP are an solid basis for representing accounting subsystems in software. Such a separated accounting subsystem is accessed exclusively through well-defined interfaces, which are clear from a software perspective and the separate terms of which can be mapped clearly to definitions in the business processes. Such separated accounting subsystems can meet the demands of openness in structure and in management mentioned earlier. Ideally, the accounting system records just what actually is the case, regardless of any plans or intentions. When someone takes stock from the warehouse, this should be registered. The IT system should not forbid registration, because of some rule or constraint in the accounting system. It happened and it is relevant to the representation of the stock in the IT system, so it must be recorded in the accounting system.

In the same vein, employees and systems should be able to retrieve the information they need for their tasks (within the boundaries of authorisation) from the accounting subsystems involved and make their relevant facts available for the accounting subsystems. These processes should be based on a pull model for information retrieval and a push model for the information produced, all according to the organisational tasks and responsibilities.

6 KEY ISSUES IN THE DESIGN OF ACCOUNTING PROCESSES

6.1 Organisational Issues

Each accounting process has to be clearly grounded in the organisation. It should be clear who is responsible for the data. Accounting processes should be located as closely as possible to the operational processes involved, to ensure short communication lines. Those responsible for the accounting processes should have a clear understanding of both the operational processes and the accounting processes so that they are able to solve any problems occurring in the collecting, processing or interpreting of data. They should actively monitor the usage of the accounting processes, and indicate when they should be adjusted because of changing practices. This final point specifically concerns tracking change of meaning, either abrupt or slowly evolving. Consider for example the concept "customer" when an organisation first encounters the difference between the entity that places an order, the entity to which they should be delivered and the entity that pays for the delivered goods.

6.2 Modelling Issues

The domain of each accounting subsystem is clearly defined, together with the meaning of the main concepts. The rules for determining the identity of the individual accounting entities are explicit. The way in which both systems and human users refer to the separate accounting entities is defined and tested for practical applicability. Categories and range of values of the attributes of the entities are defined in advance.

The domain is determined by the question of what is it concerned with and of what is information requested. Essentially this is the same question as the one asked about objects in OO-thinking. An object is an identifiable unit with its own identity.

However, objects are no "ready-mades". They need to be carefully defined. See the problem above of what constitutes one shipment.

Another issue is how involved parties can identify the accounting entities. Systems use unique references, which must connect uniquely to physical or conventional reality. This can happen by physical identification such as barcodes or chips. It can also happen by conventional identification such as GTIN numbers for products or GLN numbers for addresses and locations, managed by the international GS1 organisation.

Human users have to know what they are dealing with as well. Either the references used are fit for both machine and human recognition, or different references are used for machines and for human users.

6.3 Technical Issues

The accounting processes can be supported by one or multiple IT subsystems with the responsibility for their proper working lying with those responsible for the accounting processes involved (and not with a central IT department somewhere far away in the organisation). Besides IT systems the accounting processes can be supported by locally developed solutions, in Excel for example, and by a systematic storage of forms, receipts, and lists. What medium is used to store data is less important, that the data are collected and made available according to the agreed procedures and in a correct, timely and complete manner is essential.

The interfaces of the IT subsystems supporting the accounting processes are explicit and complete. There is no tacit meaning and there are no hidden side effects. From the defined interfaces and from the defined technical implementation of the IT subsystem its behaviour can be fully explained. The technical interfaces of the subsystem can be directly mapped to the organisational aspects of the accounting processes. The subsystems can deal with the key characteristics as mentioned in RM-ODP 10746-1, paragraph 6.1: remoteness, concurrency, lack of global state, partial failures, asynchrony, heterogeneity, autonomy, evolution and mobility.

7 CONCLUSIONS

In the introduction two misconceptions were identified as a source of many problems in the use of IT systems, namely the idea that an information system would coincide with the IT systems used and

that the information in IT systems would represent the reality "as is". The need open systems to facilitate business processes instead of systems designed from closed and encompassing models was discussed.

It was then analysed that the accounting concept as defined by Starreveld should provide a good basis for open systems. Well-designed accounting subsystems that provide their services to the organisation and to applications independently of one another should make the systems more manageable. A condition for this is that the independent accounting subsystems are well embedded into the organisation, as closely as possible to the operational processes. Modern software architectures like RM-ODP should provide a good technical basis for this.

The design of an accounting subsystem for a specified area starts with abstracting and modelling it. This should be accompanied by a thorough analysis of the process logic in the area, to arrive at an adequate and practical choice of the entities and references. This should also be accompanied by a certain formalisation and standardisation of the use of language.

When the people involved are familiar with the ins and outs of the chosen model they will be able to work well with the model in the interaction with the accounting subsystem while retaining the freedom of interpretation of data from the subsystem because they know what is not represented and because they are able to combine the data with data from other sources autonomously.

In conclusion: adequate accounting systems function semi real time, provide crucial services to the business processes, and are driven by employees that have a thorough practical knowledge of these processes.

REFERENCES

- Mintzberg, H., 1979. *The Structuring of Organizations*, Prentice Hall Inc. Englewood Cliffs, N.J.
- Starreveld, R. W., 1963. *Leer van de administratieve organisatie*, N. Samsom NV. Alphen aan den Rijn, 2nd edition.
- Glautier, M. W. E., Underdown, B., 2001. *Accounting, Theory and Practice*, FT Prentice Hall. Harlow, 7th edition.
- Boisot, M., 1998. *Knowledge Assets*, Oxford University Press. Oxford.
- Anthony, R., 1965. *Planning and Control Systems*, Graduate School of Business Administration Harvard University. Boston.

- Taylor, R. N., Medvidovic, N., Dashofy, E.M., 2010. *Software Architecture*, John Wiley and Sons, Hoboken.
- RM-ODP, ISO/IEC 10746-1, 1998, *Information technology – Open Distributed Processing- Reference Model: Overview*. ISO/IEC, Genève

SHORT PAPERS

A FOCUSED APPROACH TO BUSINESS CAPABILITY

Vaughan Michell

*Informatics Research Centre, Henley Business School, University of Reading, Whiteknights, Reading, U.K.
v.a.michell@reading.ac.uk*

Keywords: Business capability, Core competence, Resource based theory, Enterprise architecture.

Abstract: The term ‘business capability’ is widely used and generally understood, but definitions vary greatly and are often insufficiently detailed to avoid confusion. Business processes are often mistakenly seen directly as a capability without any specific detail that would differentiate such a capability or process in a similar competitive firm or business. The increasing use of enterprise architecture approaches in consulting practice to analyse and make critical business change decisions such as business service divestment or outsourcing has resulted in the need to develop a more specific focused definition in order to differentiate between capabilities and their enabling resources. This paper seeks to create a focused and specific business capability definition that reduces confusion and enables clarity in defining capabilities within an enterprise. The approach reviews and categorises existing definitions, identifying four key elements of capability that are then analysed using resource based theory and operations theory to produce an integrated definition. The paper proposes an operational definition of resource capability relating to driving and passive resources. A structured tabulation is proposed that enables specific capabilities to be defined in terms of a delivery process, tangible and intangible resources used or consumed and the specific value added by the capability.

1 INTRODUCTION

The concept of capability is increasingly important due to the focus on enterprise architecture and the way a business serves its customers. Despite the wide use of the term, there are large differences in definitions (Curtis et al, 1995) based on their origins in business strategy, operations management and IS and service oriented architecture. Capability is often used to describe the generic potential ability of a business or parts of a business. There is an intuitive view of what capability may mean, but little work has been done to qualify and compare specific capabilities despite their use in analytical architecture frameworks. Students and practitioners often find it difficult to explicitly pin down and differentiate the key capabilities and services provided by a business. This can lead to problems in identifying different capabilities needed for strategic alignment when using enterprise and domain architecture tools (Liu et al, 2011). An explicit and measurable definition is necessary to make informed business decisions for people, systems and process change. For example when deciding whether to divest or outsource IT services that support specific business service capabilities (Liu et al, 2011), or the

allocation of the correct resources such as in medical process pathways and patient safety (Ball et al, 2003). This paper analyses the variety of existing definitions of capability to identify key themes that characterise the term. Working from first principles and using these themes and two examples it proposes more specific definitions to support architectural frameworks. The work is in the process of being tested with an industrial collaborator (for potential inclusion in EA analysis) and is also part of a current research project into the impact of capabilities and patient safety with a major hospital.

2 CAPABILITY DEFINITIONS

2.1 Business Strategy

There has been a focus on organisational aspects of capability as part of the research into business strategy to support the idea of core competences advanced by Porter (Porter and Millar, 1991) and Prahalad and Hamel (Prahalad and Hamel, 1990). This is based on the resource based view of the firm that sustainable competitive advantage is a result of the specific grouping and use of resources (Penrose,

Ref C:	Definition of Capability	Key Factors	Key Component	Source
1	<i>what a business function does and what its externally visible behaviour is</i>	what a function does and its behaviour	Capability is related to work functions and behaviour	Homann, 2006
2	<i>its fundamental purpose in terms of the outcomes of the activity</i>	purpose, outcome, activity	Capability relates to activity outcomes	Merrifield et al, 2008
3	<i>a firm's capacity to deploy Resources, usually in combination, using organizational processes, to effect a desired end.</i>	capacity to deploy resources using organisational processes	Capability is a capacity to use resources	Makadok, 2001
4	<i>operations strategy involves exploiting capabilities of operations resources'</i>	exploiting resources capabilities	Capabilities are functions of resources	Slack et al, 2004
5	<i>an organization's ability to assemble. Integrate, and deploy valued resources</i>	ability to assemble and integrate resources	Capability as an ability to do coordinate resources for action	Bharadwaj, 2000
6	<i>process capability describes the range of expected results that can be achieved by following a software process</i>	results of a process	Capability relates to the result of a process	Paul et al, 1993
7	<i>the ability that an organisation, person or systems possesses</i>	ability of a person or system	Capability depends on resource	Josey et al, 2009
8	<i>abilities within a firm which can be linked together as business processes in order to enable a specific purpose or outcome'</i>	abilities, process, outcome	Capability as ability related to the specific outcome of a business process	Beimborn et al, 2005
9	<i>capability is the capacity of a team of resources to perform some task or activity'</i>	capacity of resources to perform tasks	capability relates to groups of resources	Grant, 1991
10	<i>capabilities are 'formed through the coordination and integration of activities and processes</i>	coordination of activities and processes	capabilities depend on process integration/coordination	Hafeez et al, 2002

Figure 1: Content Comparison of Capability Definitions.

1959); (Wernerfelt, 1984). For example Grant's theory of organisational capability focused on how integration of specialist knowledge aids capability as a process input (Grant, 1991).

The resource based theory of competitive advantage suggests 'capability is the capacity of a team of resources to perform some task or activity'. Bharadwaj sees capabilities similarly referring to the assembly, integration and deployment of valued resources providing competitive advantage (Bharadwaj, A. S. 2000). A more explicit definition is provided by Makadok who describes a business capability in terms of a 'capacity to deploy resources in combination with processes. (Makadok R. 2001.). These definitions critically imply that the use of resources in a particular way is fundamental to the concept of capability. Other authors refer to capabilities of a firm and their importance in adding value and value creation. For example Moller et al (Moller and Torronen, 2003) develop a capability profile for suppliers, but don't specifically define what capability is. Gallouj et al (Gallouj and Weinstein, 1997) in their paper on innovation in services see capability in terms of competences bought to bear or mobilised by a service provider. They represent competences as a vector set and are careful to focus on teams or individual competences and exclude organisational competences. Nelson and Winter cited in Gallouj and Weinstein 1997, see competences as skills and intangible technical characteristics referred to codified routines and includes knowledge, knowhow and technical

characteristics which refer to methods and selection tests. This reminds us that competences include the intangible, but it misses the fact that capability and what we see as a subset; human and related competences are a function of resources, all resources and their organisation and application. Hafeez et al (Hafeez et al 2002) point out the relationship between core competence and capability and the fact that capability is formed by the integration of resources. They further suggest capabilities result from integration of activities, highlighting an important and intuitive linkage between resources and actions in a process.

2.2 Operations and Informatics

A second area to refer to capability is operations management theory, where capability and service are associated with of operational performance. For example Slack et al suggest 'operations strategy involves exploiting capabilities of operations resources' (Slack et al, 2004). This view further advocates capabilities as a function of a firm's resources. Capability is also used in informatics web and service orientated architecture. TOGAF makes an attempt to define capability as 'the ability that an organisation, person or systems possesses'. This reminds us that business structures possess capability (Josey et al, 2009). Merrifield (Merrifield et al, 2008) sees capability relating to the purpose of business activity and its outcomes. Beimborn (Beimborn et al, 2005) regards capabilities relating

to repeatable actions citing Wade & Hulland where capabilities relate to assets capable of producing products for the market. They also suggest that capabilities can be linked to together as business processes to deliver a specific outcome. This usefully relates capabilities to the outcomes of actions via processes. However, it adds to confusion by implying processes can simply be labelled as capabilities such as ‘pay employees’ and ‘ship products’, without understanding that capabilities perhaps relate more to the character of the resources delivering capability via the processes and not just the sequence of action or recipe for the series of activities. Homann (Homann, 2006) also relates capabilities to process and suggests capability is a model of the behaviour of a business function. However, whilst intuitively reasonable, both Homann and Merrifield’s approach (Merrifield et al, 2008) avoid mentioning or invoking the contributing resources and the often complex way they enable the capability. We can summarise the definitions of capability by comparing the semantic content in Figure 1.

3 CONTENT COMPARISON

3.1 Operations and Informatics

By comparing the content of these definitions from the summary table we can see that there are a number of key themes we can explore to better define capability. Firstly capabilities relate to work processes and the outcome of the work process (C1) (C2) (C6) (C8) (C9). We can also identify that capabilities relate to the capacity, promise or potential to do something (C3) (C5) and also to a specific outcome, value or result of an activity (C2) (C6) (C8). Finally capability is a characteristic of an organisation and specifically how it uses and applies its resources. However, an organisation is an organised collection of resources which themselves also provide the basis for value adding capability (C4) (C5) (C7) (C8) (C10). The four themes can be summarised as:

- Capabilities relate to work activities
- Resources possess capability
- Capabilities are the potential for action
- Capabilities relate to outcome/value.

We will now analyse these themes using two relatively easy to understand practical examples.

3.2 Capabilities and Activities

The traditional input-process-output model of business work activities sees business work activities as transformations of input resources via a transformation process or work activity into outputs (Slack et al, 2004). At the lowest level any work activity involves some kind of resource transformation. For example, reading a book transforms the reader’s input knowledge from one level to another. Assembling an electronic device transforms the components into a new structure as part of a saleable product. Thus any work process (activity acting on inputs) transforms the value of the inputs and their state from one value to another.

The transformation process involves the transformation of input resources in terms of a change of state (eg tangible; such as melting or forming materials or intangible in the form of a change in the level of knowledge, or 1s and 0s in a computer program). These input resources may be utilised to transform the inputs eg a person and a tool. Alternatively some of the input resources may be consumed as part of the conversion process, for example the solder used in soldering an ipod electrical component as part of the business of producing ipods, or the ingredients used in the act of making a pie. We need to differentiate between resources that are utilised to make the transformation process work vs. resources that are consumed in the process of transformation. The transformation process transforms the business inputs into outputs that, via a sequence of activities or process, are valued by the customer (Strnadl, 2006). We can say that a business is capable of producing a) products and services or at a lower level b) functions and components and service features that support the value added in the products and services delivered to the end customer. Each transformation activity adds value that leads to a product of service that can be sold to the customer as the output of a business. A process is a sequence for action or a recipe for the right combination of work activities and procedures to realise the potential to produce saleable products and services. Using the Davenport definition (Davenport and Short, 1990) we define a process as: a structured and defined sequence of activities that transforms a set of inputs (of specific value) to produce a specific output (of a different value). At this stage it is also useful to remind ourselves that a process may be tangible (Pt), i.e. we can see the activities being performed in a production line, or in an intangible process (Pi), where the process, e.g. a lecture, may transform knowledge in student’s

heads, without any obvious transformation being visible, i.e. a service process (Slack et al, 2004).

3.3 Resources and Capability

A resource has a capability for interaction with the environment to create value. One of the key characteristics of capabilities is that they are a property of an object or thing, i.e. a property of business resources as suggested by some of the definitions. We talk about people having capabilities, but we are also aware that objects have capabilities. We may consider a person in a particular role has a skill or capability to do something. For example an assembly worker has a capability to assemble an object based on their knowledge of the object, their training and the procedures and the skill to operate a machine or tool to assemble the components that form the object. However, it is not only the person that has the capability. If we consider the assembly of an ipod may involve soldering electronic components in place by means of a soldering iron. The soldering iron has a capability to melt specific materials, in this case solder, to attach the component to a circuit board. Unlike the person, the soldering iron needs the person to enable or execute its capability.

ICT technology also possesses capability. For example we can replace the person soldering the component with a dedicated programmable machine or robot. The difference here is that the soldering robot is capable of soldering with very limited human intervention; perhaps only turning on the machine and ensuring inputs and outputs. Also the components of intelligent products themselves, such as an ipod, may possess intelligent capability, for example the chipsets used in the phone are themselves capable of controlling devices. To efficiently manufacture an ipod, or most objects we may use intelligent agents (Russell and Norvig, 1995) that act on the environment in order to perform a transformation. These may be people or intelligent machines (robots, software etc) and are resources that configure and transform the consumable resources used within the process of production.

The capabilities these agents possess could be considered intelligent capabilities that involve the powerful ability to adjust their work actions to suit variations in the work conditions. Other resources, e.g. materials, and subassemblies are effectively passive and possess no intelligent capability. However, some are more or less capable than others. For example a component such as an on-off switch

is not intelligent, but has what we could call a potential or capability to turn the ipod on or off. This suggests that we should consider a hierarchy of resource capabilities to reflect the way that the resource possesses ability to impact the environment and the type of capability such as intelligent/passive and whether they are part of the process (utilised to make the transformation process happen) or are consumable resources to it.

Resource Capability		
	Intelligent	Passive
Used	intelligent process agents (eg production line workers), intelligent machines (robots, fixed production lines, ICT systems)	tools (eg soldering iron), codified information (procedures, drawings etc)
Consumed	intelligent components (eg computer chips, memory - ICT components)	materials, subassemblies, mechanical components

Figure 2: Resource Capability & Use.

3.4 Resource Capability Types

We can separate capabilities into three types; capabilities of people, capabilities of ‘intelligent’ machines and capabilities of inanimate objects. However, there is a fundamental difference intelligent resources (people, intelligent machines) have an intrinsic capability to act on the environment themselves to transform and add value, whereas the passive resources only add value when they are used by an intelligent agent. Very often these intelligent resources actively drive the process and coordinate the value added transformation. We can call these driving resources of the process. Driving resources use other resources, such as tools and machines and passive resources such as materials to achieve the transformation:

- Agent resources have a potential for interaction with the environment to create value
- Driving resources (people agents, intelligent agents) orchestrate and trigger the transformation process
- Passive resources require agents to realise their capability.

A person (human agent) or an intelligent agent e.g. a machine such as a robot or application is an agent resource of the business that will trigger and drive the process. For example a web bot may search, collect and group search terms (information resource) and return them as part of its search capability. An autonomous production line may act

as a driving agent to utilise tool resources (e.g. soldering tip) to solder and manipulate ipod components (passive resources). We should also think of capability as not just a property of individual resources, but of organised and structured groups of resources assembled in a process or a functional business group and ultimately as part of a whole business as per definitions C3, C5, C7.

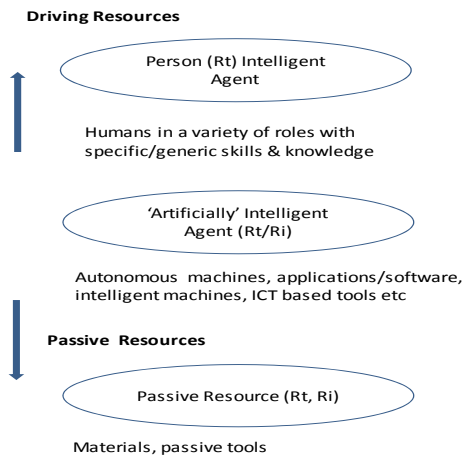


Figure 3: Driving & Passive Resources.

We can then relate capabilities to the basic resources of the firm. This approach allows us to identify core capabilities that ultimately relate to core competences of the firm as a whole and key capabilities relating to a specific functional group, as per the capability pyramid Figure 4.

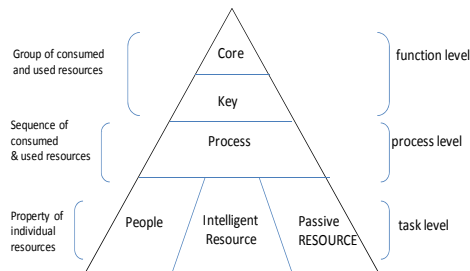


Figure 4: Business Resource Capabilities.

3.5 Tangible and Intangible Resources

Resources can also be tangible or intangible (Slack et al, 2004). For example the implicit knowledge in an employee’s head, e.g. a doctor is intangible, but critical in performing useful work activities, such as diagnosing an illness and providing capability based on the resource based theory of the firm Beimbom (Beimbom et al, 2005). For example one assembly worker may be capable of assembling ipods faster than others because of their better knowledge and

experience or by using the standard tools in superior ways. Similarly intellectual property such as a patent provides a unique license to use specific knowledge and prevent other businesses from using it in the same way.

<p>Tangible Resources People Materials Machines (passive/ICT enabled) Artefacts Tools Codified information and Knowledge Documents/forms/drawings Properties Can be physically touched Observably transformed Observable capability of physical action/use</p> <p>Intangible Resources Software Knowledge Encoded information/data Intellectual property Culture/values Skills Behaviour Properties Cannot be physically touched Not observably transformed Non observable capability for action/use</p>
--

Figure 5: Tangible/Intangible Resources.

The properties of tangible and intangible resources can be summarised in Figure 5.

3.6 Capabilities as Potential

The confusion of capabilities with processes is reasonable if we think of a capability as the potential for action is to create a transformation to produce a tangible or intangible output. We can say: A capability can only be realised by an action on an environment.

The capability needs to be realised by a work activity or sequence of actions/process that delivers the capability. In other words a capability requires a process to define its potential for action. We can therefore name capabilities using a verb for example a production line has a capability to manufacture, a person in the role/with the skills of a design engineer has the capability to design. A pizza delivery agent, with the right resources (address and price information, pizza, moped etc.) has the capability to deliver. Capabilities are realised by interaction with other resources and are inextricably linked to process.

3.7 Capabilities and Value

Another important factor is that capabilities must provide the ability to deliver an outcome or a change

Activity	Take order	Make dough	Form pie base & ingredients	Cook Pie	Serve pie	Transact bill
Value	V0	V1	V2	V3	V4	V4
Driving Resource	Waiter	Pastry Chef	Master Chef	Cook	Waiter	Front of house
Capability	To manage customer needs, take and enter order details	To make pastry	To assemble pie, To manage catering resources	To cook pie	To serve order and provide bill	To transact bill
Process	greet customer, take order, pass order to chef	place ingredients in dough machine, set machine controls, start and stop	collect, cut and shape pastry case, select and add ingredients	place raw pie in oven, set oven controls, remove cooked pie	serve pie, calculate bill	transact bill, receive payment
Output	pie order for table n	mixed dough	uncooked pie	cooked pie	served pie, billed order	paid bill
Utilised resource	ordering system	pastry mixer	pastry cutter, knife	oven	ordering system	Electronic till
Capability	store order, calculate price, display order	make pastry	shape ingredients	cook pie	display and print bill	display bill, calculate change
	link to electronic till					
Consumable resource	paper	flour, water	raw pie ingredients	electricity	pie	electricity, till roll

Figure 6: Capability Example.

adding value to the business/customer. Each internal activity or sub-process creates added value, which like the parts of a product are assembled and sum to a final value add. This value add must be sufficient for a customer to be willing to pay the price for it. We can illustrate the interaction of resource capabilities to add value by using an intuitive example of a restaurant that specialises in producing hand made pies. It has this capability as a result of specific capabilities of the restaurant staff and the resources they employ in the restaurant processes. Each capability transforms a consumable resource or utilises a resource to add value that the customer will pay for in the food and service they receive. Let us suppose the master chef has unique skills that enable him to design and skilfully make a unique range of pies that draws customers to the restaurant. The restaurant has the capability to make unique pies to order due to the capability of the master chef who organises other agent resources such as the pastry chef and the cook to work with him to make the pastry and cook the pies. This integrated capability of the restaurant may also provide it with a competitive advantage. However the capability also requires the use of utilised resources such as pastry mixers, ovens etc. and passive resources such as kitchen utensils, each with their own capability that contributes to the overall capability. In addition intelligent information processing resources such as the electronic ordering system and till are part of the capability. We also must not forget the consumable resources transformed by the driving resources as part of the process delivering the capability. For example pastry is created from flour eggs and water and shaped. Meat, vegetables and seasonings are cut/mixed and placed in the pie. Each activity has

used a capability of a driving resource to transform the consumed resources, or organise another resource to add value. But, we have only considered part of the mechanism that delivers the capability. The pies have to be served to the diners and the diners have to be managed and billed. This requires the coordination of other driving resources such as a waiter and front of house roles that deliver the pies and the bill to the customer and complete the business transaction by ensuring the customer's payment that pays for the value adding capabilities.

3.8 An Integrated Definition

From the previous sections we can see that a capability:

- Is possessed by a resource or resource groups of resources (tangible R_t and intangible R_i)
- Is the potential for action via a process P
- Produces a value for a customer V (internal/external)

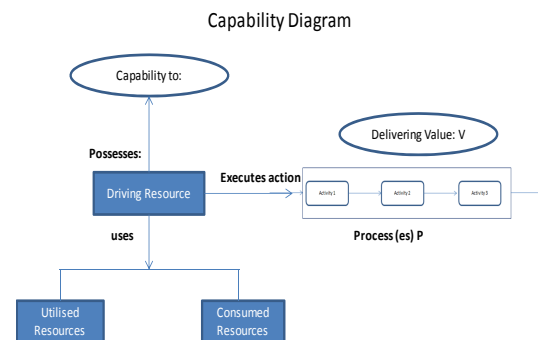


Figure 7: Capability Components.

Realising the capability requires interaction with the environment and the use of other resources: A capability is the potential of a business resource (or groups of resources) to produce customer value by acting on their environment via a process (P) using other tangible (Rt) and intangible Resources (Ri).

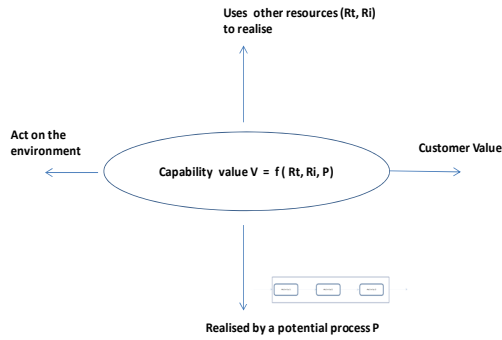


Figure 8: Capability Characteristics.

This definition allows us to account for the fact that individual resources possess active and passive capability as well as the organisation of groups of resources as in an organised business function and ultimately the complete business. Business or organisational capability refers to the collective capability or a group of resources with potential to deliver a specific business value output to an external customer. This capability depends on the way the grouping for action is made eg in a process structure or team structure etc.

4 DISCUSSION

4.1 Capability vs. Process

Using the pie example we can define a formal structure to document the capability. Firstly the capability must be labelled using a verb/noun combination to semantically describe the potential it provides. In this case we might choose a capability at the key level (not just a single process) for example ‘to provide customised pies to order’ note the operational action elements: ‘serve’ and the deliverable: ‘pie’ and the form in which it is provided ‘to order’. The semantic choices in the statement if chosen wisely will provide much specific evidence about the capability. For example the choice of ‘to serve’ vs. ‘to provide’, implies that the pies are not merely made available as in an environment, but served in controlled conditions such as a restaurant.

Capability f(P, Rt, Ri) to: <i>action/deliverable/conditions</i>		
Driving Resource: <i>process owner</i>		
Value: <i>related to customer need and business objectives and identifiable products/services</i>		
Process (P)	Tangible Resources (Rt)	Intangible Resources (Ri)
<i>description from process architecture</i>	<i>driving, passive, active</i>	
Capability f(P, Rt, Ri) to: <i>serve customised pies to order</i>		
Driving Resource: <i>Maitre de</i>		
Value: <i>hand crafted country pies</i>		
Process (P)	Tangible Resources (Rt)	Intangible Resources (Ri)
<i>Pie Making</i>	<i>master chef, pastry chef, cook</i>	<i>knowledge of pies, cooking skills</i>
<i>Restaurant service</i>	<i>waiter</i>	<i>service skills</i>
<i>Billing</i>	<i>front of house, order system, till</i>	<i>knowledge of order value, simple mathematics</i>

Figure 9: Capability Template/Example.

To enable this capability we need to include the driving resource responsible for the related process, in this case the Master Chef or Maitre de. The driving resource will need to consume and utilise other resources (tangible and intangible) within a process or series of processes to deliver the capability. The process names can be sourced from the business process architecture. The level of description of the tangible and intangible resources will depend on the capability level described (see Figure 4).

The other critical element is the value or benefit of the capability to enable use in evaluative decisions in enterprise architecture (Liu et al, 2011) and process engineering/reengineering decisions. For example, when deciding whether to develop a process or system or to divest it or whether to outsource etc.

This value is also critical for strategic decisions in understanding core competences as we discuss later. The value can be simply a statement of the benefit, in this example we may suggest a unique benefit, such as ‘make to order pies’ or ‘hand crafted country pies’. However, we need to ensure the value relates to customer need and expectations and also to the business objectives. Ideally the value should be linked to identifiable products and services. This link will enable a meaningful inventory of capability to be established that can help a business understand how value is developed by combinations of different resources and processes. Numerical values can also be included using activity based costing analysis or a similar approach to define the monetary value of this capability for the business.

4.2 Internal and External Capability

We need to consider two types of capability to sup-

port the idea that the value produced by the capability may be internal to the business or external and to the customer as well as the business. External customer focused capabilities occur where the potential output is of core importance to customer benefit. This relates to Merrifield's and Porter's definition relating capability to fundamental business success and customer value. In contrast an internal business capability is where the potential output is delivered within the business ie to help run the business and plays no part in adding customer specific value. This relates to Beimborn's (Beimborn et al, 2005) view about capability being 'what a business function does'.

4.3 Generic vs. Specific Capabilities

Driving resources using different combinations of activities and resources will produce variations in the realisation of the capability. For this reason we need to consider generic capabilities eg to manufacture something vs. the specific capabilities to manufacture a specific item such as an ipod, or a computer chip that is a component of the ipod in a dedicated and focused, reasonably invariant processes. We should also consider that some capabilities depend more on the linkage and coordination of resources than on a specific resources. Many capabilities reside in a purposely integrated group of people, process and technology and the way they are ingeniously organised and architected by the business. Hence capabilities may vary and actually disappear over time. This alludes to the organisational capability mentioned by some authors, i.e. the ability to coordinate disparate resource capabilities to produce a greater capability for the business is a key capability in itself.

4.4 Capabilities and Competences

Competences are abilities of a company to add value (customer benefit) to their products and services. (Prahalad and Hamel, 1990) As Gadrey (Gadrey, 2000) suggests 'competences are a network of capabilities' and not a single process. Competences are usually referred to in very generic terms such as the ability to do x or manufacture y, which makes use of the intelligent and wide ranging grouping of the capability of the driving resources and the intelligent machine resources as well as passive resources. As these authors mention core competences are seen as the cross functional integration and coordination of capabilities ie core capabilities support core competences. Core

competences are delivered by core capabilities. Defining capabilities as core or non-core in terms of being critical for the customer value add, embedded and unique (Prahalad and Hamel, 1990) enables us to understand at the resource level how the core competences of a business are developed and supported. The more the competence is interconnected, ie integrated, into a business structure the better you can retain control of it and the more sustainable it is. By definition then the more a resource capability or a group of core resource capabilities are integrated the greater the core competence. By definition you will not therefore be able to outsource it (ie find source another business that has the capability and can do it more cheaply/effectively) unless you give the core capability away to another business.

5 SUMMARY

We have analysed a range of definitions of capability from economics, business strategy, operations management and informatics and employed practical examples to develop a definition of a capability to include specific and potentially measurable elements; the potential of a business resource (or groups of resources) to produce customer value by acting on their environment via a process (P) using other tangible (Rt) and intangible Resources (Ri). We have shown that some resources may be intelligent or passive agents in a business process and that agents or driving resources use other active or passive resources to execute transformation activities that add business value.

The elements of the proposed concept can be included in an architecture model and have the following characteristics:

- Capability verb/noun definition forces specificity and clarity
- The output value of the capability is defined and could be enumerated and linked to economic information (finance architecture)
- Capability is linked to a named process (from enterprise architecture)
- Capability is linked to both driving and other resources that are critical to enable and ensure the quality of the output.

This approach enables a more structured definition of capabilities of an enterprise and its resources and a format to compare them at a variety of levels. It reduces confusion over meaning and enables different capabilities to be objectively assessed for

use in strategic and operational business decisions such as outsourcing and divestment selection and resourcing and creation of new processes such as in new product development or reengineering.

Further field research with industrial partners is planned to further test and develop the structure and analyse the semantic formulation of the capability description and its meanings. Opportunities to test the approach in other area are being taken up. A doctor or clinician needs to provide or source the right capability at the right time to ensure patient safety. Additional research in the medical patient safety arena (Rosenorn-Lanng et al, 2011) aims to investigate the link between this form of capability structure and resourcing to improve patient safety.

REFERENCES

- Curtis, B., Fefley, W. E., Miller, S., 1995. Overview of the people capability maturity model. *CMUSEI 95 MM01*.
- Porter, M E., Millar, V. E., 1991. How Information Gives You Competitive Advantage. *HBR*
- Prahalad, C. K., Hamel G., 1990. The Core Competence of the Corporation. *HBR*. May-June pp2-15.
- Grant, R. M., 1991. The Resource-Based Theory of Competitive Advantage- Implications for Strategy Formulation. *California Management Review*
- Bharadwaj, A. S., 2000. A Resource Based Perspective on Inf. technology Capability and Firm Performance: An empirical investigation. *MIS Quart.*, Vol 24 Issue 1
- Makadok, R. 2001. Towards a synthesis of the resource based dynamic capability views of rent creation. *Strategic Management Journal*.22: 387–401.
- Moller, K. E., Torronen P., 2003. Business suppliers value creation potential: A capability-based analysis *Industrial Marketing Management* 32) 109– 118
- Ball, M. J., Garets, D. E., Handler T. J., 2003. Leveraging IT to Improve Patient Safety. *IMIA Yearbook*
- Gallouj, F., Weinstein O., 1997. Innovation in services. *Research Policy*, Vol 26, Issues 4-5
- Hafeez, K., Zhang, Y., Malak, N., 2002. Determining key capabilities of a firm using analytic hierarchy process. *Int. J. Production Economics* 76 39–51CS9
- Slack, N., Chambers, S., Johnston, R., 2004. *Operations Management*. Prentice Hall 4th edition.
- Paul, M.C., Curtis, B., Chrissis, M.B., Weber C.V., 1993 Capability Maturity Model for Software, Vers. 1.1. *Tech. report CMU/SEI-93-TR-024 ESC-TR-93-177*.
- Josey, A., et al 2009. *Togaf Version 9 - A Pocket Guide*. The Open Group. Van Haren Publ. 2nd ed.
- Penrose, E. T. 1959. *The theory of the growth of the firm* Oxford University Press, Oxford.
- Wernerfelt, B. 1984. A resource-based view of the firm, *Strategic Management Journal* (5). pp 171-180.
- Beimborn, D., Martin S. F., Homann U. 2005. Capability Oriented Modelling of the Firm. *IPSI Conference*.
- Homan, U., 2006. A Business oriented foundation for service orientation. Microsoft Corporation. <<http://msdn.microsoft.com/enus/library/aa479368.aspx>>
- Merrifield, R., Calhoun J., Stevens D., 2008. The next revolution in productivity. *Harvard Business Review*.
- Russell, S., Norvig P., 1995. *Artificial Intelligence - A Modern Approach*. Prentice-Hall, Inc.
- Davenport, T., Short, J., 1990. The New Industrial Engineering: Inf. Technology and Business Process Design. *Sloan Management review* 31 - (4) Summer.
- Strnadl, C., 2006 Aligning Business and IT - The process driven arch. model. *Inf. Systems Management*.
- Gadrey, J., 2000. The Characterisation of goods and services: An alternative approach. *Review of Income and Wealth*. Series 46, Number 3.
- Rosenorn-Lanng, D. J., Michell, V., Fielden, J., 2011. Yes, yes, yes! *BMJ Letters*. 14 May.
- Liu, K., Sun, L., Jambari, D., Michell, V., Chong S., 2011. A Design of Business-Technology Alignment Consulting Framework. *CAiSE'11*.

BUSINESS MODELLING FOR GENERATION OF KNOWLEDGE FROM EXPLICIT DATA

Considering Administrative Management Processes

Anna Rozeva

*Department of Business Management, University of Forestry, 10 Kliment Ohridski blv., Sofia, Bulgaria
arozeva@hotmail.com*

Martin Ivanov

*New Bulgarian University, 21 Montevideo str., Sofia, Bulgaria
mivanov@nbu.bg*

Roumiana Tsankova

*Department of Management, Technical University, 8 Kliment Ohridski blv., Sofia, Bulgaria
rts@tu-sofia.bg*

Keywords: Knowledge generation, Business model, Text mining, Data mining, Administration management, Categorization, Clustering.

Abstract: The aim of the paper is to present a framework for designing business model for knowledge generation from explicit data on “good” administrative management practices. Knowledge discovery demands the availability and access to high volumes of data. There is such data collected in databases and files in different formats. Knowledge extraction is performed by statistical and machine learning mining methods of text mining and data mining. The proposed framework for business model design consists of structure and knowledge models. The two models refer to the text transformation and the knowledge extraction phases respectively. Structure model implements text mining methods for converting the text documents into structured objects. These objects form a data mining structure that is the source for the knowledge discovery models. They are oriented to descriptive and predictive modelling tasks which concern document clustering and categorization. The business model framework is trained on source text documents for “good” practices in public administration and business management which are classified according to preliminary established topics. The results obtained by the descriptive and predictive knowledge models are presented.

1 INTRODUCTION

Administration management processes generate continuously growing amount of data in a variety of digital formats. Part of it is structured in databases but there is a significant volume that is unstructured, i.e. plain text documents, images, sound and videos. This data is referred to as explicit. Structured data is processed for extracting business valuable information by means of standard database queries. At the same time the information contained in unstructured documents cannot be exposed by traditional data processing techniques. The explicit data contained especially in text documents represent a natural pool from which information of

high quality can be obtained. The higher quality is provided by the information that isn't explicitly stated but is “hidden”, “buried” or “derivable” from the data contained in the documents concerning administration management digital collections.

The information that is derivable and uncovered from digital data stores is considered knowledge. When extracted by implementing the proper information technology it turns into valuable asset for administration and business management that raises its quality and competitiveness. The generated knowledge is already in a structured form. It can be stored, processed, managed, combined and involved in further knowledge generation processes. The main challenge therefore is to turn the available unstructured digital data that is in a variety of

formats into numerical form that can be processed by standardized database techniques. The process to achieve this is text mining. The knowledge extraction by itself is performed by “mining” technique which is referred to as data mining. The aim of the paper is to design an approach for mining unstructured digital data documents on “good” practices for administration management resulting in extracted knowledge. The remainder of the paper is organized as follows: The second section is a review on mining approaches and results obtained. The third section presents a business model for knowledge generation from a data pool. The fourth section describes case study on documents concerning “good” practices in administration management and application of the model in selected platform for knowledge extraction. The last section concludes with discussion on the results obtained.

2 MOTIVATION FOR KNOWLEDGE GENERATION

Text data sourced from the World Wide Web, governmental and municipal digital repositories, blogs, e-mails, news, papers, articles, etc represent a data pool for analytical processing. Analysis of text resources performed by text mining technique is discussed in (Stavrianou and Anrditsos, 2007). It's considered to involve the following steps: parsing, pattern recognition, syntactic and semantic analysis, clustering, tokenization, application of statistics and machine learning algorithms. The analysis results are evaluated for emerging the previously unknown knowledge. It's used further on for database population or reconciliation.

Filtering and ranking of search results from bibliographic databases is considered in (Faulstich and Stadler, 2003). Text classifiers for filtering as automated text categorization and calculation of probabilities as document scores are implemented. The establishment of the training data sets from articles as PDF documents, data preparation for conversion into text and further conversion into vector representation are considered. The application for document classification with a variety of classifiers from WEKA (Witten and Frank, 2011) is presented.

A text mining system for knowledge discovery is presented in (Uramoto and Matsuzava, 2004). It represents an extension of a commercial mining system for a biomedical document base. It treats the information extraction and entity/relationship

mining by means of domain dictionary. Public ontological knowledge is used for constructing categories from extracted entities and relationships among them. Relationships are of type noun and a verb and two nouns and a verb. The information extraction process involves finding words by using a term dictionary. The obtained words are embedded in the text document as annotations in XML. The annotated text is passed to a syntactic parser. It outputs segments of phrases labelled with their syntactic roles, i.e. noun phrase or verb group. Further on categories are assigned to the terms in the segments and phrases. Mining functions are provided for discovering underlying information.

In order to analyze plain text documents it's necessary to pre-process them for obtaining data structure that is more applicable for mining with available statistical and machine learning methods. Traditional data mining (Nisbet and Elder, 2009) involves algorithms that are applied to structured data. Consequently in order to implement them to plain text documents text pre-processing is to be performed. It involves syntactic and semantic analyses of the text. For the purposes of both analyses the approach is to identify the words in the documents. And as a result the document is represented as a set (bag) of these words. Other approaches consider the importance of the words in a document which is measured by numerical value. A document is therefore represented as a vector in a multidimensional space with measures for the words contained therein. The resultant representation is structured and suitable for processing with data mining techniques. Basic text mining methods are presented in (Hotho and Nurnberger, 2005). They refer to either assigning keywords to documents based on a given keyword set (classification or categorization) or automatically finding groups of similar documents (clustering).

Another aspect of the knowledge generation process performed by text mining techniques refers to knowledge sharing. It requires linking of knowledge to ontologies as main repositories of formally represented knowledge. As shown in (Spasic and Ananiadou, 2005), ontologies provide the framework for the semantic representation of text documents. They serve for mapping terms extracted from documents to domain-specific concepts. By implementing them text can be mined for interpretable information and not only to discovering simple correlations based on co-occurrences of targeted classes of terms. Architecture for discovering conceptual structures and generation of ontologies is presented in

(Maedche and Staab, 2000). Text processing performs syntactic analysis of natural language documents. It involves word extraction (tokenization), lexical analysis and grammatical parser.

The phases of the text mining process after (Castellano and Mastronardi, 2007) are shown in Figure 1.

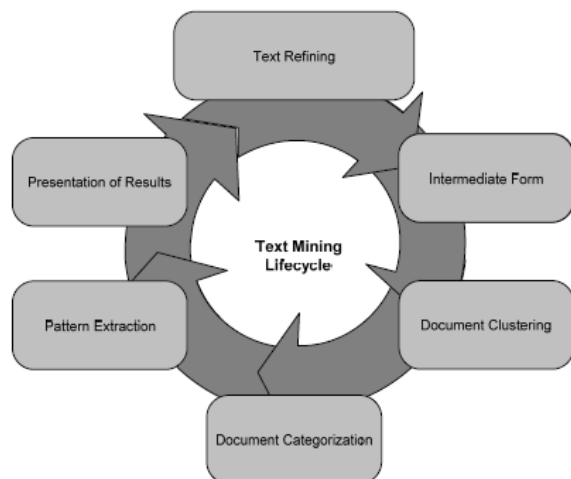


Figure 1: Text mining lifecycle.

The text mining phases are document clustering, document categorization and pattern extraction.

The review of research on knowledge generation from text documents motivates our work on establishing a business model for analysis resulting in extracted knowledge.

3 BUSINESS MODEL FOR KNOWLEDGE GENERATION

Architecture for knowledge generation system from digital document collection is presented in (Tsankova and Rozeva, 2011). The concept is the knowledge generation model to be comprised of text mining model and data mining model. The motivation is to separate the phase concerning the conversion of text documents into representation that is suitable for processing through data mining techniques and the real knowledge generation phase. Current work is further enhancement of the knowledge generation system design which concerns the mining model design. The proposed general model architecture is shown in Figure 2. The structuring phase produces text database from the documents by imposing a column/row structure. The least number of columns is two, one for the

document identifier and the other for the document text. The text being placed in a column becomes available to text mining tools. The text mining tool deals with text encoding and generates the structure model. It represents the mining structure as a source for the models that generate the knowledge.

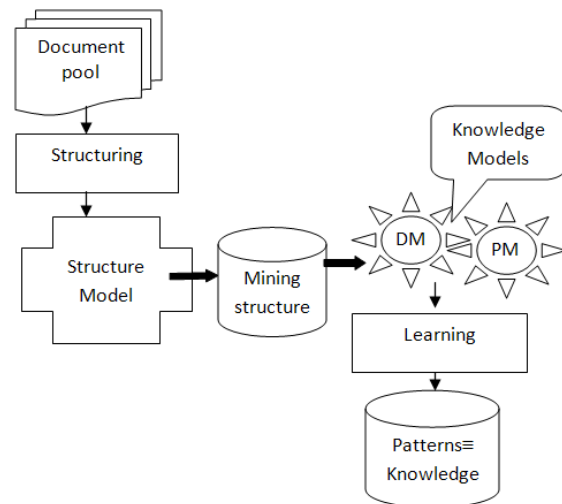


Figure 2: Knowledge generation model.

Each of the components is obtained as a result of a process that comprises several steps.

3.1 Mining Structure Design

The document pool is assumed as a set of plain text documents. The initial document format and the conversion to plain text aren't considered in the knowledge generation framework. Text structuring can be performed by ETL tool and imported in a database table.

3.1.1 Document Term Representation

The structure model generation deals with identification of words in the text. A way for designing it consists of the following steps:

- Tokenization – obtaining stream of words from documents and combining the different words into dictionary;
- Dictionary reduction – filtering, stemming;
- Term selection – further reduction of the dictionary by removing words based on calculating their entropy or another method.

Standard method of filtering the dictionary for reducing its size is stop word removal. Stop words bear little or no information content and also occur either extremely often or very seldom. The often

occurrence doesn't provide for distinguishing between documents. The seldom occurrence is of no particular statistical relevance.

Stemming methods produce basic word forms and hence groups of words with equal or similar meaning.

Term selection removes the words that are not suitable for separation of documents when searched by keywords.

The approach that has been selected for designing document structure model is document vector space. A document is represented as a vector with elements being terms, i.e. words or phrases. The vectors' size is determined by the number of words in the document collection. The selected document encoding is binary. The term vector element is set to 1 if the term is encountered in the document and 0 otherwise. In this representation terms are considered equally important. In order to take into account the term importance for describing a document, weighting scheme is applied. Several weighting schemes are shown in (Hotho and Nummerger, 2005). The weighting scheme selected for the structure model represents the product of the term frequency $tf(d,t)$ and the inverse document frequency $idf(t)=\log(N/n_t)$. The inverse document frequency takes into account the size of the document collection N and the number of documents that contain the term n_t . The term weighting scheme that is chosen implements length normalization factor as well. It eliminates the influence of document length on the chance for retrieving it. The term weight is computed after equation (1):

$$w(d,t) = \frac{tf(d,t) \log\left(\frac{N}{n_t}\right)}{\sqrt{\sum_{j=1}^m tf(d,t_j)^2 \log\left(\frac{N}{n_{t_j}}\right)^2}} \quad (1)$$

By implementing the weighting scheme a document in the vector space is represented by the term weights as (2):

$$w(d) = (w(d,t_1), \dots, w(d,t_m)) \quad (2)$$

The term weight representation allows for defining similarity between documents by means of the inner product of their vectors (3):

$$S(d_1, d_2) = \sum_{k=1}^m w(d_1, t_k) \cdot w(d_2, t_k) \quad (3)$$

The distance between the two vectors is calculated by Euclidean distance (4):

$$dist(d_1, d_2) = \sqrt{\sum_{k=1}^m |w(d_1, t_k) - w(d_2, t_k)|^2} \quad (4)$$

3.1.2 Linguistic Document Processing

Linguistic processing enhances document term representation by attaching labels for describing them as parts of speech, grouping them into sentences and eliminating disambiguities. The structure model's design involves the following linguistic processing steps:

- Assigning tags for part of speech to each term – noun, verb, adjective;
- Text chunking – grouping adjacent words into sentences;
- Determination of word sense – the semantics of words in the specific context is determined;
- Grammatical parsing – discovering the relation of each word to the others as well as its function in the sentence, i.e. (subject, object).

Linguistic processing is implemented by means of ontologies as shown in (Amardeilh and Laublet, 2005). The structure model from Figure 2 supports part of speech tagging. So far as the determination of word sense is considered the vector representation provides for automatic disambiguation by evaluation of term co-occurrence.

The process flow diagram for designing the document structure model is shown in Figure 3.

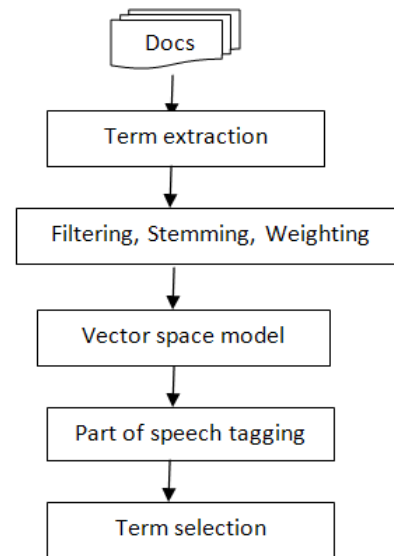


Figure 3: Document structure model design.

3.2 Knowledge Models Design

The document structure model represents structured database objects. Database objects comprise the dictionary of the document corpus and the document vectors. The structure model provides the source

data for designing data mining models for knowledge generation. Data mining methods and algorithms for building complex and powerful knowledge generation models are presented in (Larose, 2006) and in (Cios and Pedrycz, 2000).

The knowledge model consists of the mining structure and the mining algorithm. The algorithm to implement on the mining structure depends on the mining task. Hand and Mannila (2001) define the basic mining tasks as descriptive modelling and predictive modelling. The knowledge models in the knowledge generation framework from Figure 2 that perform these tasks are denoted as DM and PM.

The aim of descriptive modelling is to find models for the data. They are implemented in the setting of unsupervised learning. Typical methods of descriptive modelling are data segmentation and clustering. The reasoning behind the cluster analysis is that there are natural clusters in the data and the mining task consists in uncovering and labelling them.

Predictive modelling represents supervised learning. There is a target variable which has to be presented as function of other variables. The nature of the target variable determines the type of the model. Classification model applies to discrete variable and regression model – to continuous variable. Approaches for classification models are:

- Discriminative;
- Regression;
- Class-conditional.

The discriminative approach performs direct mapping of input variables to one of k possible target categories. The input space is partitioned into different regions which have a unique class label assigned. Examples of this approach are the neural networks and support vector machines.

The regression approach determines the posterior class distribution for each case and chooses the class for which the maximum probability is reached. Decision trees (CART, C5.0, CHAID) classify for both the discriminative approach and the regression approach, because typically the posterior class probabilities at each leaf are calculated as well as the predicted class.

The class-conditional approach starts with explicit specification of class-conditional distributions. After estimating the marginal distribution, Bayes rule is used to derive the conditional distribution.

The proposed knowledge generation model involves clustering descriptive model and categorization predictive model. The models' design

will be presented further on with the architecture and the implemented algorithm.

3.2.1 Descriptive Model (DM) Architecture

The descriptive model is designed as clustering model. It will produce groups with documents that are more similar than those in the other groups. Clustering is performed after the hypothesis that relevant documents tend to be more closely related to one another than to non-relevant documents. The clustering algorithm that is implemented is the k -means algorithm. It performs distance-based flat clustering as follows:

```

Input: D{d1, d2, ..., dn }; k-the cluster
number;
Select k document vectors as the
initial centroids of k clusters
Repeat
Select one vector d in remaining
documents
Compute similarities between d and k
centroids
Put d in the closest cluster and
recompute the centroid
Until the centroids don't change
Output: k clusters of documents

```

3.2.2 Predictive Model (PM) Architecture

The predictive model is designed as classification model since term weights in the document vectors are discrete. Classification task consists in assigning one or more predefined categories (topic themes) to a document.

Mining structure in the architecture that is provided by the structure model is document vector space. The new document for classification is turned into vector representation by passing it through the structure model. The vector space model together with the predefined topics is provided to the classifier the output of it being the predicted category for the input document.

The predictive model architecture is implemented with the Bayes classifier. The classifier's algorithm is the following:

```

Input: document d
Predefined topics T={t1, ..., tn}
Compute probability of d for t_j ∈ T

```

$$\Pr(t_j | d) = \frac{\Pr(d | t_j) \Pr(t_j)}{\Pr(d)} = \frac{\Pr(t_j) \prod_{j=1}^{|d|} \Pr(w_j | t_j)}{\sum_{t_k \in T} (\prod_{j=1}^{|d|} \Pr(w_j | t_k)) \Pr(t_k)}$$

Output: category c assigned to d with probability Pr

$$Pr(d|t) = \max_{t \in T} (Pr(d|\vec{t}))$$

The classification architecture is shown in Figure 4.

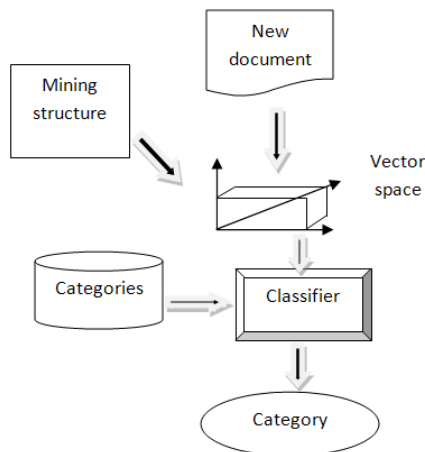


Figure 4: Categorization architecture.

4 KNOWLEDGE GENERATION MODEL IMPLEMENTATION

The knowledge generation model is implemented on document corpus of “good” practices for administration management. Documents on “good” practices are uploaded on web site and stored in a digital text database. There are topics defined for preliminary categorization of the submitted documents. The topics concerning administration management are the following:

- National administration and subsidiaries;
- Municipal administrations;
- E-solutions for effective administrative actions;
- Public- private partnership.

The topic national administration involves public policies, human resource management, new vision and performance evaluation. The municipal topic covers decentralization, policies and services. E-solutions address conceptual framework, e-government and e-community. And public-private topic involves national and municipal practices and solutions. The digital text database stores conference and workshop presentations and papers.

Since the uploaded documents are categorized after predefined topics they are suitable to be used in the predictive knowledge generation model for performing classification task. The available

documents will be implemented for training the model. It will be used further on for classifying newly uploaded documents on good practices. Thus the good practices digital store will hold automatically classified documents according to the stated topics.

The conference and workshop materials in the digital store provide for the descriptive knowledge generation model. The model will establish groups with similar documents. Thus the document corpus will be structured and the groups defined can be implemented further on in other mining tasks.

The test document pool for designing and training the knowledge models consists of 45 documents. The text represents paper titles and abstracts. The knowledge generation models are implemented with the WEKA open source data mining software (Witten and Frank, 2011) and (Hall and Frank, 2009). The text documents are structured by converting to the .arff format. The structured file for the descriptive model task has 2 attributes, i.e. filename and content. For the predictive model task third attribute is added for the predefined document topic (class).

4.1 Document Structure Model

The document corpus structure model is obtained as follows:

- Term extraction – Alphabetic tokenizer;
- Filtering – stoplist words as stoplist.txt file;
- Stemming – IteratedLovinStemmer;
- Term weighting – TFIDF transformation;
- Part of speech tagging – none;

The resulting vector space structure model is shown in Figure 5.

No.	about Numeric	accept Numeric	achief Numeric	act Numeric	administer Numeric	adv Numeric	aim Numeric
1	0.0	0.0	0.0	0.0	0.562093...	0.0	0.0
2	1.677...	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	1.115...	0.0	0.0	1.289...
4	0.0	0.0	0.0	1.115...	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	0.0	1.396...	0.0
6	0.0	0.0	0.0	0.0	0.0	1.396...	0.0

Figure 5: Document structure model.

The model contains the document vectors by rows, terms by columns and term weights by cells.

The structure model represents the mining structure for building the knowledge generation models.

4.2 Descriptive Model Implementation

The descriptive model will elaborate clusters with similar documents. There aren't predefined topics. The mining structure obtained is to be appended with selected term set that is descriptive of the document corpus content. Selection of descriptive attributes is performed by means of latent semantic analysis. It results in attribute set where attributes are obtained as a liner combination of the initially extracted terms. The selected latent attributes for clustering ranked by weight and filtered by AttributeSelection filter produce 3 clusters that are shown in Figure 6.

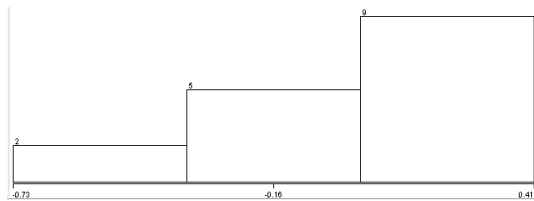


Figure 6: Clusters of selected terms.

The descriptive model is built with simple k-means clustering algorithm. The resultant clusters are shown in Figure 7.

Cluster centroids:

Attribute	Full Data (16)	Cluster# (12)
0.002addit+0.001addr+0	administer+0.002adopt+0.001aid...	0.197 0.1362
0 addit+0.001addr+0	administer+0.001adopt+0 aid...	-0.0251 -0.0467
0.004addit+0.001addr+0	administer+0.001adopt+0 aid...	0.0336 0.0636
-0.004addit+0.001addr+0	administer+0.001adopt+0.001aid...	0.0002 0.0619

Clustered Instances

Cluster	Count	Percentage
0	12	75%
1	1	6%
2	3	19%

Figure 7: Cluster content.

Clusters' visualisation is shown in Figure 8.

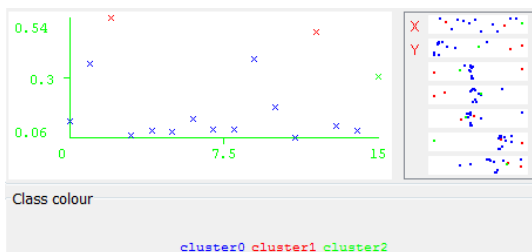


Figure 8: Cluster visualisation.

4.3 Predictive Model Implementation

Term selection has to be performed for designing the predictive classification model. The result is a set of

descriptive attributes. Term selection is performed with the following settings:

- Attribute evaluator - CfsSubsetEval;
- Search method - GreedyStepwise.

The resultant term set for implemented in the classification model is shown in Figure 9.

```
Selected attributes: 5,71,83,129,130,170,308,348,361,369,725 : 11
administer
datab
e
govern
governm
it
system
accord
both
client
way
```

Figure 9: Descriptive terms selection.

The predefined topics are encoded with A, B and C for national and municipal administration and e-solutions respectively. The classification model is built with Naïve Bayes classifier. The result is shown in Figure 10:

Correctly Classified Instances	43	95.5556 %
Incorrectly Classified Instances	2	4.4444 %
Kappa statistic	0.9322	
Mean absolute error	0.0501	
Root mean squared error	0.1399	
Relative absolute error	11.3351 %	
Root relative squared error	29.6954 %	
Total Number of Instances	45	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
1	1	0	1	1	1	1	A
0.857	0	1	0.857	0.923	0.995		B
1	0.074	0.9	1	0.947	0.996		C
Weighted Avg.	0.956	0.03	0.96	0.956	0.955	0.997	

=== Confusion Matrix ===

a	b	c	←- classified as
13	0	0	a = A
0	12	2	b = B
0	0	18	c = C

Figure 10: Predictive classification model.

The model precision is presented by correctly and incorrectly classified instances and as confusion matrix. The tree like view of the classification model is shown in Figure 11.

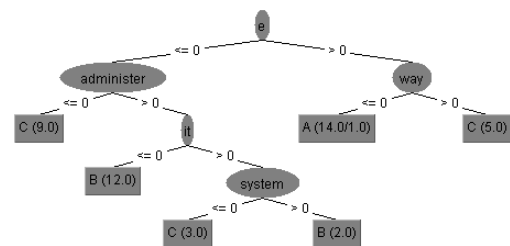


Figure 11: Classification model tree view.

5 CONCLUSIONS

Knowledge based management of administration and of business provides for enhanced competitive advantage and quality of services and business process flow. Practically knowledge doesn't exist in explicit form. It's hidden within data pools of different format and volume. The task of uncovering it poses significant challenge to information technologies and business modelling. Current paper contributes to the design and architecture of knowledge generation system with issues concerning data modelling and implementation for knowledge discovery. General model for mining knowledge from digital text stores is presented. The model framework involves structure model and knowledge models. The steps for designing them implement text and data mining techniques. Basic architecture and algorithms for performing descriptive (clustering) and predictive (classification) modelling tasks are presented. The knowledge generation model is trained on test document corpus for "good" practices for administration management. The models are established in WEKA and produce knowledge results for the clustering and categorisation tasks. Future work is intended in extracting associations between terms and implementation of ontologies.

ACKNOWLEDGEMENTS

The paper presents results of the project "Research and Education Centre for e-Governance" funded by the Ministry of Education in Bulgaria.

REFERENCES

- Amardeilh, F., Laublet, P., 2005. Document annotation and ontology population from linguistic extractions, In *Proceedings of the 3rd international conference on knowledge capture*, ACM New York, NY, USA.
- Castellano, M., Mastronardi, G., 2007. A web text mining flexible architecture, In: *Proceedings of World Academy of Science, Engineering and Technology, IV Int. Conf. on Computer, Electrical, and Systems Science, and Engineering Cesse Edited by:PWASET Vol. 26*, 78-85.
- Cios, K., Pedrycz, W., 2000. *Data mining methods for knowledge discovery*, Kluwer Academic Publishers, Netherlands, 3rd edition.
- Faulstich, L., Stadler, P., et al, 2003. litsift: Automated text categorization in bibliographic search. In *Data mining and text mining for bioinformatics Workshop at the ECML/PKDD Dubrovnik-Cavtat, Croatia*, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.10.8737&rep=rep1&type=pdf> , Retrieved May, 16th 2011.
- Hall, M., Frank, E., 2009. The WEKA data mining software: an update, *ACM SIGKDD Explorations newsletter, Vol.11, Issue 1*, New York, NY, USA
- Hand, D., Mannila, H., 2001. *Principles of data mining*, MIT Press.
- Hotho, A., Nurnberger, A., 2005. A brief survey of text mining, *LVD Forum, Band 20*, pp.19-62.
- Larose, D., 2006. *Data mining methods and models*, Wiley-IEEE Press.
- Maedche, A., Staab, S., 2000. Mining ontologies from text, In *R. Dieng and O. Corby (eds.): EKAW 2000, LNAI 1937*, Springer Verlag Berlin Heidelberg.
- Nisbet, R., Elder, J., 2009. *Handbook of Statistical Analysis and Data Mining Applications*, AP Elsevier Inc.
- Spasic, A., Ananiadou, A., 2005. Text mining and ontologies in biomedicine: Making sense of raw text, *Briefings in bioinformatics, Vol.6, No3*, Henry Steward Publications.
- Stavrianou, A., Andritsos, P., Nicoloyannis, N., 2007. Overview and semantic issues of text mining, *SIGMOD Record September 2007 (Vol.36, No3)*
- Tsankova, R., Rozeva, A., 2011. Generation of knowledge from "good practices" as open government procedure, In *CeDEM11 Proceedings of the International conference for e-democracy and open government, Peter Parychek, Manuel Kripp (eds)*, Danube University Krems, Austria, 209-219
- Uramoto, N., Matsuzava, H., 2004. A text mining system for knowledge discovery from biomedical documents, *IBM Systems Journal, vol.43, No3*.
- Witten, I., Frank, E., 2011. *Data Mining: Practical machine learning tools and technique*, Morgan Kaufmann Publishers. Burlington, 3rd edition.

INCORPORATING COLLABORATION IN BUSINESS PROCESSES

How to Merge Business Processes and Collaboration Activities in an Efficient and Agile Way

Stoian Boev, Elina Surova, Kaloyan Nikolov and Viktor Zhivkov

*SimPro – Business Processes and Simulation Competence Centre, Faculty of Mathematics and Informatics
Sofia University and Interconsult Bulgaria, 7, Indzhe Voivoda Street, Sofia, Bulgaria
icb@icb.bg*

Keywords: Business process, Business process model, Collaboration, Process execution, Automation, Task context, Enterprise portal

Abstract: Nowadays, in order to remain competitive, organizations strive to be efficient and agile. They need well defined and optimized business processes, without sacrificing the agility of collaboration. This can be achieved by applying a new approach for combining and tuning these two into a single concept. To implement the new approach characteristics of business processes such as task, goal, participating roles, resources, etc. and the characteristics of collaboration like flexibility, informal communication, collaborative tools, shared knowledge, the two should be merged together in a mixture called “task context”. Intelligent collaboration business portal solutions that utilize the task context approach provide a unified environment where users can collaborate while following the end-to-end process. Such solutions provide agility while maintaining adequate levels of process control thus improving efficiency and effectiveness.

1 INTRODUCTION

The organizations’ need for agility and efficiency improvement can be fulfilled through the use of intelligent collaboration portal solutions, which incorporate both the business processes of the organization and the collaboration. These new portal solutions introduce a novel approach for business process automation. It allows retaining process agility and covers the ad-hoc activities possible within the organization. At the same time it allows efficient performance of the daily tasks by providing all needed information and tools at the right time.

2 WHY IS A NEW APPROACH NEEDED

An agile and efficient work organization is a must if companies strive to quickly adapt to an ever changing external environment. Nowadays the successful management and execution of the tasks within the organization could be compared to jazz

improvisation which is a symbol of high creativity (Scheer, 2007, p.2).

To be adaptive to change and at the same time retain and improve its efficiency, an organization needs an innovative approach for combining and tuning its primary components:

- Business processes - the actions which the organization carries out;
- Resources - people (the human component which performs tasks), technological environment, facilities, knowledge, etc.;
- Collaboration and communication - working together to achieve a goal.

2.1 Current State of the Organization

If an organization is examined today, the following conclusions can be made:

- Importance of business processes is realised: “In the last 30 years companies have become more business process aware” (Huberts and Petten, 2007, p.3);
- Business processes are described but not used:

There are organizations in which the business processes are in fact described but not used in the actual carrying out of activities;

- Business processes have been automated without achieving the required effect:

The automation of the processes inside a single IT system cannot be a sufficient solution. (Huberts and Petten, 2007, p.3). Moreover the introduction of new processes in this type of system can result in the need for substantial additional investments for software development (Scheer, 2004, p.8).

Another option is to integrate multiple information systems to automate business processes. This approach handles the communication between IT systems and synchronization of the data, but it does not support the communication and collaboration at human level (Huberts and Petten, 2007, p.3).

- Collaboration could be inefficient. In every organization there are many communication and collaboration channels. Collaboration tools are powerful software products, but their uncoordinated usage leads to a loss of efficiency and can lead to the loss of valuable information.

The state of the organizations indicates that a new approach for their setup is needed. That approach should ensure the execution of processes in a collaborative environment without limiting users.

2.2 Company Characteristics

In order to extract this approach, the organization's components, that are subject to automation have been analyzed - namely processes and collaboration.

3 BUSINESS PROCESSES

The business processes are not only something that organization does, they are the business of the organization (Brabander and Davis, 2007, p.2).

“Business Processes are logically ordered sets of activities that produce a result of value to the customer” (Indulska, Kittel, and Muehlen, 2008, p.3);

“The process defines what activities are performed by who, when they are performed, and how they are performed” (Crow, 2002);

Business processes are a set of tasks in a given order. Each task has a trigger, executor, resources required, input, output and an expected result.

The main business process characteristics are:

- Predefined goal – each process has a goal that should be achieved;

- Set of activities – process contains a number of tasks. They can be individual – performed by single user or collaborative – a group of people working together to achieve common goal;

- Defined participants with their roles - the participant can have different relation to the tasks (e.g. execute, must be informed, etc);

- Required resources for execution - to execute a process certain resources are needed (data, documents, knowledge, applications, time, etc)

- Formal communication paths.

Processes can be classified as very rigid (production processes), more flexible (sales) and ad-hoc (not predefined process that can occur during the daily work) (Brabander and Davis, 2007, p.2).

The processes are usually used as means for analysis, optimization and control. During the daily execution of processes the executors see and carry out certain tasks without necessarily focusing on the process as a whole. (Figure 1: Process and tasks)

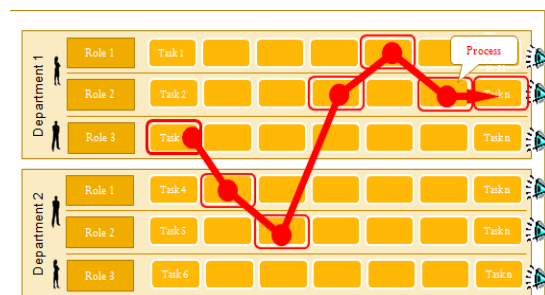


Figure 1: Process and tasks.

The main issue when automating processes is that described process models have some constrains:

- Too rigid - the defined process models represent a predefined set of activities in a given order. It is not possible to describe all variants, because they are even not known at that moment. The result is that flexible and ad-hoc processes are unsupported. Thus the direct automation and execution of the process models leads to rigidity, omission of some possible scenarios and cannot be used in a changing environment.

- Missing collaboration and real communication – the defined process models represent the formal communication. The possible additional actions and collaborations that could occur and are part of the process are not taken into consideration (i.e. while carrying out a task the executor may need help from a more qualified employee).

The question is how an organization can be flexible and at the same time controlled. For achieving this

balance in the organization it has to be positioned on the edge of chaos.

According to Scheer (2007):

Modern organizational theory pursues approaches that take into account the dynamics within companies and in markets, and that overcome inflexible organizational principles.

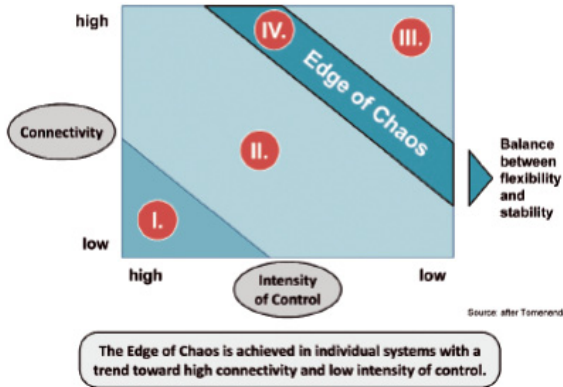


Figure 2: Balance Flexibility and Stability (Scheer, 2007).

Figure 2 shows an illustration of how connectivity – which can also be interpreted as communication and interaction – and the intensity of control with an organization determine the possibility for flexible, creative behavior (c.f. Tomenendal, 2002, and Scholz, 2000). If the organization has many rules, then all work processes are set. If, at the same time, there is little communication between the participants, then the organization stagnates. It is unable to react quickly to unexpected situations (lower-left section of Figure 2 Balance Flexibility and Stability). If, however, no rules exist, chaos rules (top-right section of Figure 2 Balance Flexibility and Stability). The shaded area represents a corridor of balance between minimal constraint and maximum communication. Area II describes a more stable organization that has not yet stagnated, but does not demonstrate spontaneity and flexibility as in the shaded area.

Having the business processes characteristics and the business models constraints on one hand and the theory of the edge of the chaos on the other the role of business processes in their automation and execution should be defined. Before that the hierarchical structure of process models and their tendency to change should be considered.

The structure of process models can be depicted as a pyramid having high-level abstract processes on top (see Figure 3 Current Structure of Business Process Models). Going down processes get more granular and volatile down to a point that requires

many details and loop-backs. The bottom zone is where things get done.

As the base provides the added value of the business process it is essential to ensure that it serves the purpose flawlessly and strictly. This usually means that business analysts try to define very rigid control flow of the execution. Taking processes a step further and trying to automate their execution it becomes obvious that the base is not stable. There are changes that occur frequently, but these changes are minor viewed from higher levels.

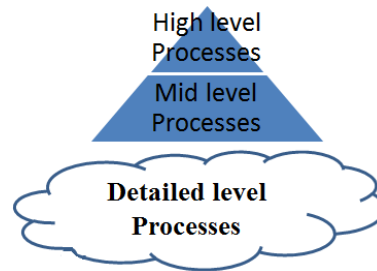


Figure 3: Current Structure of Business Process Models.

The root cause of the problem is often business processes models try to capture the collaboration activities together with the control flow of the process. This over-detailization induces instability. To achieve stability and manageability it is imperative to find a way to recognize individual tasks and collaborative tasks and enable users to convert individual tasks into collaborative if need arises. Each such task will be defined as a small perceivable brick in the foundation and will have strongly defined boundaries both in time and resources. The illustration will be changed from Figure 3 Current Structure of Business Process Models to Figure 4 Structure of Business Process Models with defined collaboration task.

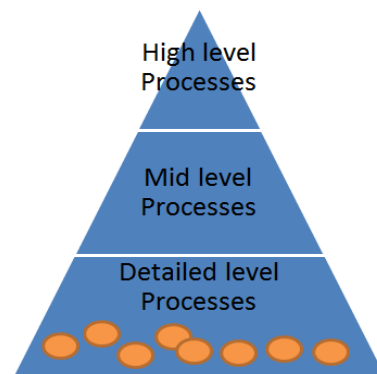


Figure 4: Structure of Business Process Models with defined collaboration tasks.

4 COLLABORATION

Collaboration is a process through which people who see different aspects of a problem can constructively explore their differences and search for solutions that go beyond their own limited vision of what is possible. Nowadays the collaboration is one of the most success factors for the organizations.

4.1 Collaboration Types

Examining the collaboration within the organization three types can be distinguished (see Figure 5 Collaboration types).

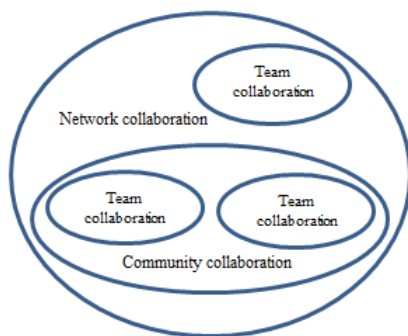


Figure 5: Collaboration types.

According to Callahan, Schenk, White (2008):

- Team collaboration – there are clear goals to achieve within explicit time-line. In such collaboration group all members are known and they have roles. Leadership increases the productivity of the team collaboration.
- Community collaboration – goals are more often focused on learning rather than on tasks. Members of a team contact another team in their community to ask a question or take advice.
- Network collaboration – it covers the collaboration beyond the team and the community. There are no explicit roles and time-lines.

The main collaboration characteristics are:

- Flexibility – collaboration is used for almost each problem area in the real life. The collaboration is in the human nature and a various ways of collaboration have been developed so far. There are various collaboration tools for different situations.
- Understanding – it is important to ensure that all collaborators know what the goals are and all of them can participate actively. For fruitful collaboration a common language should be set.
- No clear roles – how strict are the roles depend on the collaboration type but in general in most cases

all collaboration members are treated equally. The benefit is that the roles delusion encourages creativity and the set back is that focus is easily lost. That is why leadership is key for establishing good collaboration.

- No clear list of required participants – participants should be invited based on the collaboration goals and the knowledge of the problem domain that each one can contribute.
- Informal communication – the collaboration should be organized in such way that all members feel comfortable and can openly express their opinion. At the same time the result of the collaboration should be recorded and accessible for each contributor at any given time. A variety of technical tools could be used to support such collaboration.

4.2 Different Way to Collaborate

There are different ways how to collaborate, but most commonly they fall in two groups:

- Information sharing and decision making;
- Contribution to common knowledge or resources.

In addition collaboration channels can be various (ordered by their “synchrony” – from synchronous to asynchronous): Phone, Meeting, Instant messaging, E-Mail, Forum, Wiki, File sharing.

People usually get accustomed to their preferred channel of collaboration and rarely think if it is the most appropriate one for the task at hand. This leads to delays in collaboration and frustration. In order to prevent such inconveniences there should be some kind of aid that will present the right tool at the right moment to the user. Most of the time this is done using rigid business processes that confine the collaboration activity to a predefined sequence of tasks that should be done to produce the desired output. Such decision makes matters even worse:

- Business process analysts have a hard time defining the process and all loopbacks typical to a collaborative activity (e.g. brainstorming cannot be described as a linear predictable sequence of activities);
- People who have understanding of how to accomplish certain tasks get frustrated because they have to follow new strange rules and they cannot recognize easily the benefit of the “standardization”;
- It is hard to follow procedures that describe collaboration activities. Especially if business analysts have “captured” all the details and subtleties in the process.

Business processes models are not appropriate to describe collaborative activities. Rather than describing what to do it will be better just to define certain boundaries and tools that will help them steer in the right direction at maximum efficiency.

5 THE NEW APPROACH

Collaboration activities tend to lose efficiency as they start to go off-topic or loop around themselves. Thus a better boundary or constraints around the problem domain is needed. Business processes can lend several ideas how to set certain boundaries on the collaboration. Taking the best from the business processes and collaboration, a sensible mixture can be compiled with the following characteristics:

- Goal – a clear shared goal transmitted to all participants;
- List of participants – a way to include the necessary involved people and people that have knowledge to share or decision making power;
- Allowed time period for the activity – work tends to consume whole time dedicated to it according to Parkinson's Law. For activities such as collaboration that is often the case. It is better to restrict the time available than to have open-ended collaboration sessions;
- Resource availability – putting some constraints on the resources at hand can make many decisions easier and faster and will help focus on the real goals;
- Necessary information – only the relevant data and information to the current task is shown;
- Required knowledge – this includes the persistent knowledge base of the organization (like procedures and policies) and implicit knowledge from the employee experience. The second one requires that the experienced employee is invited as a participant in the task.
- Task execution tools – most of the time the task at hand will require a tool to be used in order to achieve the goal. Providing the right tool to people will eliminate the time and effort lost in seeking the tool and will minimize the risk of making a wrong choice.
- Collaboration tool/channel – standardization of tools and channels for each kind of collaboration activities within the organization brings the benefit of reduced number of tool and IT systems. This way organizations can reduce cost and clutter;
- Activity history – keep track of what and when happened prior to the current tasks so if there are

any notable changes with regards to the landscape everyone is alert of them;

- Dependencies – collaboration activities rarely exist by themselves and affect no one and nothing. All participants should be aware of the dependencies between tasks in the process.



Figure 6: Task context.

Based on the set of characteristics a new concept – task context, is introduced (see Figure 6 Task context). Collaboration activity can be wrapped into single task and attached to the proper task context that includes the fore-mentioned properties. The gap between business processes and collaboration activities can be bridged with the following rules how to integrate both:

- Individual tasks inherently receive task context from the business process. They are treated in the same manner as they are. If someone needs to collaborative support in such task the context easily allows facilitation of such collaboration.
 - Collaboration activities are modelled as a single task inside the business process without any collaboration implementation details. The task is given the necessary sufficient context.
- Let's take as an example the process Project Management. At a certain moment during the process the task Perform risk assessment should be done. The task should have the following context:
- Goal: assess relevant project risk;
 - Participants: project manager (Nikolay Nikolov), risk officer (Ivan Ivanov);
 - Deadline: March 25th;
 - Necessary information: project scope, plan, client information;
 - Required knowledge: risk assessment form and methodology;
 - Execution tool: document management;
 - Collaboration tool: chat/phone;
 - Activity history: previous meeting minutes;

- Dependencies: task should be performed after project plan and before kick-off meeting.

Note that there are no execution details, i.e. exact steps how risk assessment should be done. Note also that only information relevant to the project at hand and to the specific task is present.

The only obstacle to implement this approach is the lack of single orchestration tool that will drive the business process to follow the model and will provide task context to process and collaboration activities. It will be a benefit if the tool is collaboration enabled so it can host collaboration channels.

6 ENTERPRISE PORTALS

Current collaboration tools focus primarily on content management, specifically group editing and reviewing of documents but lack integration with the rest of the processes. On the other hand business applications generally follow a rigid process and collaborative tasks are done “outside” them. The idea for the context and how it benefits from the process and the collaborative approaches has been introduced previously.

The context materialization is done by Enterprise Portals or also known as Business Portals. The concept behind portals is a Web site that serves a single gateway to company information and knowledge base for employees and possibly for customers, partners, and the general public. Most of the existing portals are missing one vital for process and collaboration execution element – orchestration.

This can be solved by implementing Intelligent Collaboration Business Portal (in short ICB Portal) with the following main features:

- Access/search - allows a user to get all the information needed in the desired context. For example, a loan officer does not need marketing information to approve a loan. The portal shows the loan officer only the information needed;
- Categorization - the portal categorizes all information so that it is delivered to the user within the context needed;
- Collaboration - single web portal spanning to intranet or even internet allows individuals to collaborate regardless of geographical location;
- Business process execution – does the hard job to orchestrate the whole process;
- Personalization - The information provided to individuals using a portal is personalized to that person's role, preferences, and habits;

- Expertise and profiling - Expertise and profiling is essential for the collaboration element of a portal. Individuals within an enterprise are profiled according to their experience and competencies so that project members can be chosen according to their qualifications;

- Application integration - This allows individuals to deliver, access, and share information regardless of applications used, provided IT systems are SOA based and can provide parts of their functionality as a service;

- Security - This provides information to users based on security clearance. The user authenticates and accesses only information that has authorization for based on company rules and policies.

An ICB Portal allows following structured business processes, providing data to and from various orchestrated applications, and enabling collaboration whenever it is needed within the relevant task context.

7 CONCLUSIONS

If organizations embrace this new approach and implement the task context they will be able to preserve the controllable process execution in a collaborative environment enabling employees to be productive and use company knowledge. This single workspace will provide the executors knowledge, information and tools needed at the right time for the task execution. The concept of the task context can be expanded further with the introduction of the problem domain of project management. In this way organizations can get maximized view on the activities inside them and manage them better.

ACKNOWLEDGEMENTS

This research was partially supported by the Bulgarian National Science Fund under Grant DO 02-75/2008.

REFERENCES

- Brabanber, Eric, Davis, Rob, 2007. *ARIS Design Platform Getting Started with BPM*, Springer
- Callahan, Shawn, Schenk, Shawn, White, Nancy, 2008. *Building a collaborative workplace*
- Crow, Kennet, 2002. *Collaboration* <http://www.npd-solutions.com/collaboration.html>

- Huberts, Arjen, Petten, Alexander van, 2007. *Trends in Business Process Management*
- Indulska, Marta and Kittel, Kai and Muehlen, Michael, 2008. Towards Integrated Modeling of Business Processes and Business Rules
- Kirchmer, Dr Marthias, Jost, Dr Wolfram, Scheer, Prof. Dr. Dr. h.c. mult. August-Wilhelm, 2004. *Business Process Automation*.
- Scheer, Prof. Dr. Dr. h.c. mult. August-Wilhelm, 2007. *ARIS Expert Paper. Jazz improvisation and Management*

MMALL

Multilingual Mobile-assisted Language Learning

Maria Virvou, Efthimios Alepis and Christos Troussas

Department of Informatics, University of Piraeus, 80 Karaoli & Dimitriou St., 18534 Piraeus, Greece
{*mvirvou, talepis, ctrouss*}@unipi.gr

Keywords: Mobile-assisted language learning, Mobile learning, Multiple language learning, User modelling, Error diagnosis, Adaptive learning.

Abstract: Learning multiple foreign languages has become a necessity in modern life since globalization is a phenomenon responsible for joining different cultures from all over the world. Computer-assisted language learning tools exist over the last decades and their assistance has been considered as quite beneficial in the area of education. However, the incorporation of mobile facilities in these tools offers the quite important facility of time and place independence for the users that are going to use them. Towards addressing the problem of providing mobile-assisted language learning, in this paper we present a sophisticated educational system called m-MALL. The m-MALL system has the additional advantage of providing multilingual support, a facility that is not yet investigated in the related scientific literature.

1 INTRODUCTION

Mobile-Assisted Language Learning (MALL) represents a very recent research field in the domain of language learning where the educational process is assisted or enhanced through the use of handheld mobile devices. As a result, MALL is a subset of both Mobile Learning (m-learning) and Computer-assisted language learning (CALL). MALL has evolved to support students' language learning with the increased use of mobile technologies such as mobile phones (cellphones), PDAs and devices known as smartphones. Through MALL, students have the ability to access language learning materials, test their knowledge, as well as to communicate with their teachers and peers at any time and at any place.

In our century, we have witnessed major improvements in the areas of transportation and telecommunications, permitting globalization by which regional economies, societies, and cultures have become integrated through a global network of people. As a result, personal, professional, social, and economic considerations all point to the advantages of learning foreign languages (Kurata, 2010). Considering the scientific area of Intelligent Tutoring Systems (ITSs), there is an increasing interest in the use of computer-assisted foreign language instruction. Especially, when these systems

offer the possibility of multiple-language learning at the same time, the students may further benefit from this educational process (Virvou et al, 2000). The need for tutoring systems that may provide user interface friendliness and also individualized support to errors via a student model are even greater when students are taught more than one foreign languages simultaneously. Student modeling may include modeling of students' skills and declarative knowledge and can perform individualized error diagnosis of the student. However, in the recent scientific effort, it is not depicted the implementation of mobile-assisted language learning systems, which can support multilingual content in their domain of knowledge.

In view of the above, in this paper we propose a multilingual mobile-assisted language learning system which is the application of MALL in a multiple language learning environment. The prototype system combines an attractive multimedia interface and adaptivity to individual student needs in mobile learning. The communication between the system and its potential users as students is accomplished through the use of web services.

The paper is organized as follows. Firstly, we present the related work, concerning CALL systems, MALL systems and mobile learning in section 2. In section 3, we discuss our system's architecture. Following, in section 4, we present a description of

our system, namely a general overview accompanied by screenshots of the system. Finally, in section 5, we come up with a discussion about the usability of the resulting system and we present our next plans.

2 RELATED WORK

Teaching languages through computer-assisted approaches is a quite significant field in language learning. A small number of researchers in the subject area have been further attracted by mobile-assisted language learning (MALL) over the last decade but this interest is rapidly growing. The majority of scientists, who are interested in these fields, have been attracted by computer-assisted language learning (CALL) or even mobile learning (m-learning). In this section, we try to imprint the speckle of the scientific progress in assisted language learning.

2.1 Computer-assisted Language Learning

AutoTutor is a CALL system, developed by Graesser et al (2005), which simulates a human tutor by promoting the conversation and provides feedback to the learner, pumps him/her for more information, gives hints, fills missing information with assertions, identifies and corrects bad answers, answers learner's questions and summarizes answers. Another CALL system is rEcho, which is developed by Zhou et al (2007), can give relevance feedbacks through anatomy animation and is based on deliberate data trained recognition to give error trend relevant feedbacks. SignMT was implemented by Ditcharoen et al (2010) to translate sentences/phrases from different sources in four steps, which are word transformation, word constraint, word addition and word ordering. Another computer-based program on second language acquisition is Diglot Reader, which was developed by Christensen et al (2007) and is used in a way that students may read a native language text with second language vocabulary and grammatical structures increasingly embedded within the text. TAGARELA is an individualized instruction program, implemented by Amaral et al (2007), which analyzes student input for different activities and provides individual feedback. Finally, VIRGE, developed by Katsionis and Virvou (2008), works as an adventure virtual reality game but it has educational content as well and supports personalized learning based on a student modeling

component.

2.2 Mobile Learning

Mobile learning is a general aspect of assisted language learning and focuses on learning across contexts and learning with mobile devices. Jeng et al (2010) conducted an investigation of add-on impact of mobile applications in learning strategies. They surveyed recent researches including context awareness, pedagogical strategy-enhanced learning scenarios, as well as collaborative and socially networked mobile learning. Through their review study, essential characteristics of mobile learning were identified and discussed. Frohberg and Schwabe (2009) note in their critical analysis that mobile learning can best provide support for learning in context. There, learners are asked to apply knowledge and not just consume it. Furthermore, mobile learning should provide instruments to provoke deep reflection, communication and cooperation. Kuo and Huang (2009) propose an authoring tool named Mobile E-learning Authoring Tool (MEAT) to produce adaptable learning contents and test items. In addition, the visualized course organization tool has also been provided to teachers to organize their teaching courses. Moreover, Motiwalla (2005) proposes a project which explores the extension of e-learning into wireless/handheld (W/H) computing devices with the help of a mobile learning (m-learning) framework. This framework provides the requirements to develop m-learning applications that can be used to complement classroom or distance learning. A prototype application was developed to link W/H devices to three course websites. Gu et al (2011) make an effort to provide learning in informal settings through mobile. In order to learn how to develop usable learning content for lifelong learners on the move, a set of design principles from both pedagogical and usability concerns was identified. Next, a pilot system, based on the design principles, was developed to implement two prototype lessons. Finally, Mobile Author is an authoring tool, which was implemented by Virvou and Alepis (2005) and allows instructors to create and administer data-bases, concerning characteristics of students, of the domain to be taught and of tests and homework, through any computer or mobile phone.

2.3 Mobile-assisted Language Learning

Mobile-assisted language learning has evolved to

support students' language learning with the increased use of mobile technologies such as mobile phones and other devices. With MALL, students are able to access language learning materials and communicate with their teachers and peers. Chang and Hsu (2011) introduce a computer-assisted language-learning system for use on PDAs, which integrates an instant translation mode, an instant translation annotation mode and an instant multi-users shared translation annotation function in order to support a synchronously intensive reading course in the normal classroom. Zervas and Sampson (2010) propose an IEEE Learning Object Metadata (LOM) Application Profile that can be used for tagging educational resources suitable for Language Learning and supported by mobile and wireless devices. Fotouhi-Ghazvini (2009) et al also occupied with mobile-assisted language learning. Their paper concludes that using m-learning within the informal framework of learning provides a ubiquitous tool that can powerfully help adult learners and students in Iran during their continuous lifelong learning. Uther et al (2005) present a mobile adaptive computer-assisted language learning (MAC) software aimed to help Japanese-English speakers in perceptually distinguishing the non-native /r/ vs. /l/ English phonemic contrast with a view to improving their own English pronunciation. Sandberg et al (2011) conducted a research concerning the way that mobile learning may affect learning performance. Three groups participated in their study on the added value of mobile technology for learning English as a second language for primary school students.

However, after a thorough investigation in the related scientific literature, we came up with the result that teaching multiple languages through an integrated tutoring system via mobile is an approach that was not investigated before. For this reason, we implemented a prototype system, which incorporates user modeling and error diagnosis components, while teaching multiple languages through mobile devices. Our system offers the possibilities of e-learning and distance education as it operates in mobile devices, but also it offers interactivity and adaptivity to individual students needs. The resulting system includes all the standard attributes of an ITS and complies with its architecture, which consists of the domain knowledge, the student modeler, the advice generator and the user interface (Virvou and Alepis, 2005). A novelty of the system lies in the multilingual component and in the error diagnosis process, which is carried out through the m-MALL. The system provides estimation of the learner's proficiency in the domain as well as his/her

proneness to commit errors. The facility of individualized error diagnosis is particularly important for students, who can benefit from advice tailored to their problems.

3 GENERAL ARCHITECTURE OF THE SYSTEM

In this section, we describe the overall functionality and features of m-MALL. The architecture of m-MALL consists of the main educational application, a student modelling mechanism, a web service, a database and finally an educational application installed locally in each user's mobile device. The web service is responsible for transferring the available information from the main system's database to the mobile application. The database is divided in two logical partitions. One part of the database is used to store educational data and another part is used to store data related to user modelling, namely student profile information and error diagnosis data. Accordingly, the database is used to store user models and user personal profiles for each individual user that uses and interacts with the system, as well as stereotypic information about user profiles. Each user's initial profile is updated while s/he uses the educational system. The system's general 4-tier architecture is illustrated in figure 1.

Correspondingly, the student modelling mechanism consists of two sub-mechanisms. One mechanism that is responsible for each student's model and another mechanism that reasons about Multilanguage errors (figure 2). Each student's profile takes initial values by combining the student's personal information (such as age, gender and educational level) with pre-stored stereotypic information. Consequently, the student model is adapted according to the students' performance while using the educational application. The Multilanguage error diagnosis mechanism tries to find possible reasons about student errors in a Multilanguage domain of knowledge. As a next step, these errors are categorized in terms on five pre-defined categories of errors.

Error diagnosis in the student's domain of knowledge is accomplished by recognizing errors and trying to associate them with one of the following five categories of errors:

1. Article and pronoun mistakes
For example the user may have used "a" instead of "an" or "he" instead of "we".
2. Spelling mistakes

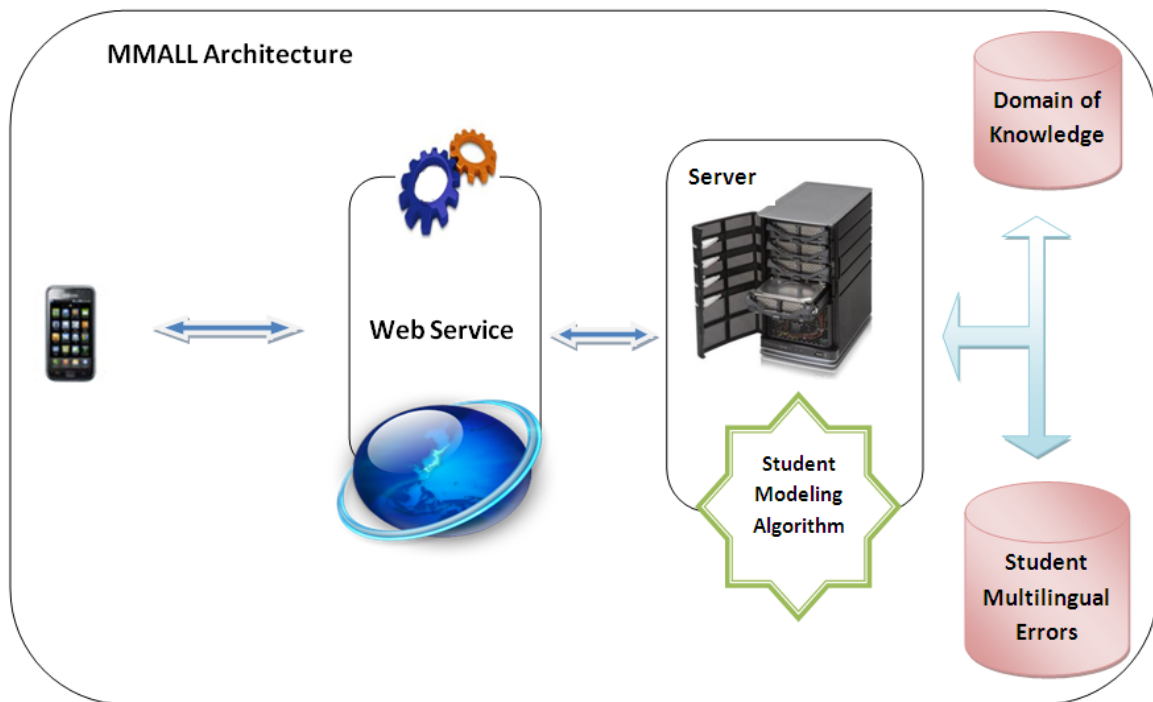


Figure 1: Architecture of m-MALL.

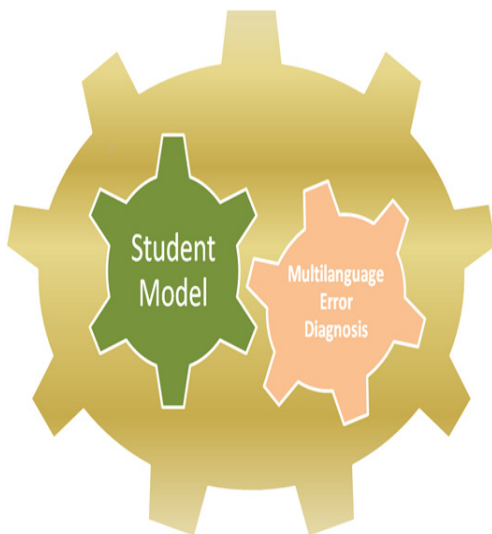


Figure 2: The Student modelling mechanism.

A spelling mistake is a result of letter redundancy, letter missing or interchange of two neighbouring letters.

3. Verb mistakes

Verb mistakes occur when the user has typed another person than the correct one, for example s/he may have typed "I has" instead of "I have".

4. Unanswered questions

The user may have no idea about what s/he should write and leave the question unanswered. That means that s/he has lack in theory.

5. Language Confusion

The resulting system is a multilingual learning system, which means that a student may learn two or more languages at the same time. However, there is the possibility of student's getting confused, concerning the proper use of an article or verb.

In order to successfully recognize one or more of the fore mentioned categories of errors, m-MALL incorporates two algorithmic approaches, as illustrated in figure 3. The first algorithm tries to find string similarities by matching a student's given "exact" wrong answer with the systems correct stored answer. If string matching occurs in a high percentage the system decides whether the mistake lies between categories 1-4. Correspondingly, using the second algorithm, the system also tries to find meaning similarities between the given and the correct answer by translating these two answers to the system's available supported languages. As an example, the student may have used "I am" instead of "Je suis", which is the French equivalent.

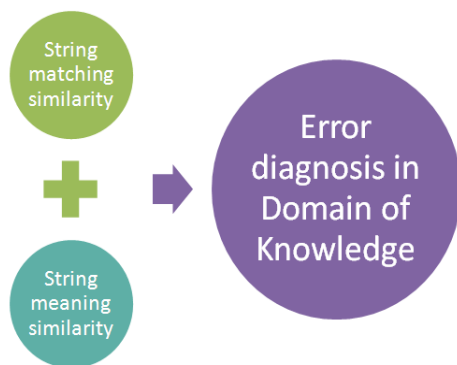


Figure 3: Error diagnosis mechanism.

A matter of great importance is the existence of a long term user model for each student. The system includes also a form, which keeps information about the student's progress in the three languages, the total grade in each one of the three languages and all the results of the tests. Students may benefit from viewing their own student models. For this reason, this form can be presented to students so that they stay aware of their advance of knowledge. M-MALL programs may promote noticing (focus on form) that will result in the improvement of students' existing grammatical knowledge. This can be accomplished by evaluating the students' performance through several tests.

The proposed architecture of the m-MALL prototype system gives great flexibility both to the students who are using the educational language learning application and to their instructors or supervisors, since the remote database can be easily updated and enriched with new knowledge domain data and user specific information. Furthermore, this is also the case of the student modelling mechanism since the algorithms may be independently modified or changed in order to provide more sophisticated feedback for the students

4 OVERVIEW OF THE SYSTEM

M-MALL has been developed to operate on the Android mobile operating system, while as for future work the authors are planning to provide implementations for other existing mobile phone platforms as well. Correspondingly, the system is programmed using JAVA as a programming language. This specific programming language is also compatible with the system's Object Oriented structure.

As we can see in figure 4, the m-MALL system is packaged and installed as an application in an

Android operating smartphone. Each user can run or stop the educational application by using to start or to stop m- MALL through his/her mobile device. However, this mobile application relies on data stored in m-MALL's server and transferred to each interacting user through a web service.

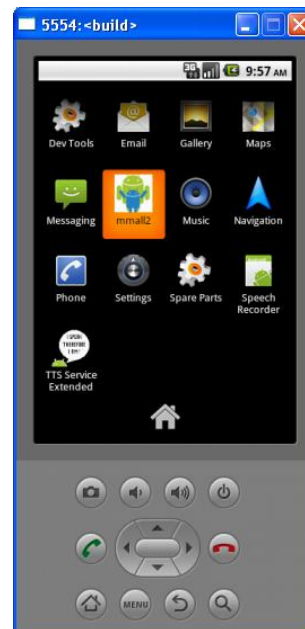


Figure 4: m-AWARE as an application in a smartphone.

Figure 5 illustrates a snapshot of the operating educational application where each student's personal profile may be updated. Personal student information is used for the student model to initialize using student stereotypic information which is stored in the system's database. According to each student's personal profile the system chooses which parts of the theory are appropriate for the student's learning level as well as the difficulty of each test.

Figures 6 and 7 illustrate snapshots of the operating educational application, where a student is completing a "fill in the gaps" exercise and taking the system's feedback. More specifically, in figure 6 we may see a student who has to fill in the gaps with the right word. The questions appear randomly and are adapted to each student model. It is quite important to note that the student must have acquired the knowledge offered by all the lessons as prerequisites.

Figure 7 illustrates a categorization to a student's specific errors. The student can be evaluated and check where s/he is wrong and what type of mistake s/he has made. The different colours indicate different type of errors:

- The red colour in the field means error in articles or pronouns.
- The green colour means a verb mistake.
- The yellow colour means a spelling mistake.
- The blue colour means confusion with the French language, while the purple means confusion with the English language.
- Finally, the grey colour indicates an unanswered question.

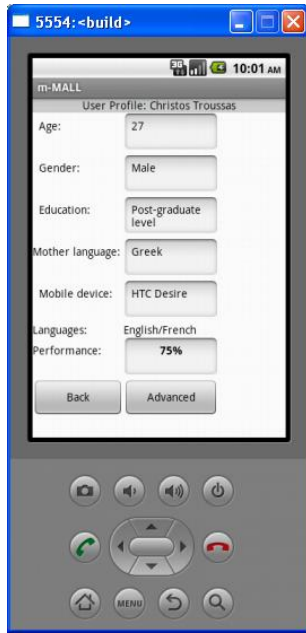


Figure 5: Snapshot of a user editing his/her profile.

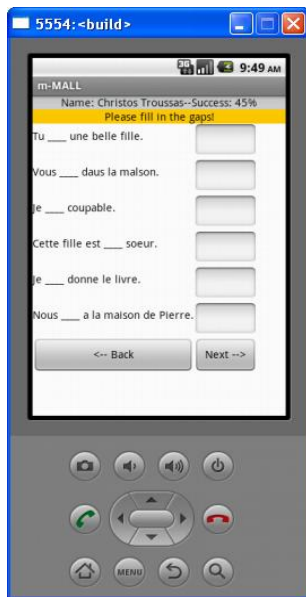


Figure 6: A student is taking a test.

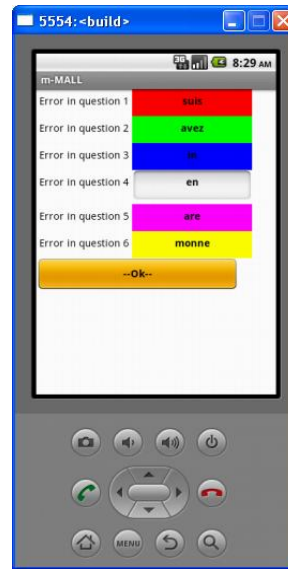


Figure 7: A student is viewing his/her errors.

At the same time, the system shows the grades of the students, along with the exact number of the errors in each category. The overall interaction through a mobile device is as much user friendly as possible in order to achieve high interactivity with regards to the limited functionalities and technical specifications of mobile devices in comparison with personal computers. To this end, part of the most “demanding” processing of data is carried through the system’s main server and then transferred to the mobile device using web services.

5 CONCLUSIONS AND FUTURE WORK

M-MALL is a multilingual mobile educational application which combines attractiveness and user-friendliness that usual desktop applications provide with the well-known advantages of mobile learning. It is not only a post-desktop model of human-computer interaction in which students can “naturally” interact with the system in order to get used to electronically supported computer-based learning, but it promotes the m-mall in a platform where the student interacts with his/her mobile phone. In particular, the system incorporates the student modeling component for each user and performs error diagnosis. Moreover, the system keeps each student’s error history in one language that is already taught and then provides advice in the tests of the other languages. In order to perform error diagnosis, the system bears a detailed

categorization of common student's mistakes. The error diagnosis process of the m-MALL system is especially focused on errors due to confusion of the other languages of the system, if the student learns more than one language at the same time. Furthermore, apart from the friendliness of the user interface, the system is oriented to offer adaptivity and dynamic individualization to each user that interacts with the educational application. All the available data that are used both for the domain of knowledge and for the student model are stored and processed in a server and transferred to each student's mobile device through the use of web services.

It is in our future plans to evaluate m-MALL in order to examine the degree of its usefulness as an educational tool for teachers, as well as the degree of usefulness and user-friendliness for the people who are going to use the educational system. We are also planning to extend the functionalities of our system by incorporating an authoring component into the m-MALL system in order to help teachers author their Multilanguage learning lessons. This incorporation will give the facility to teachers with limited computer skills to author important modules of the educational system in a short time and with less effort.

REFERENCES

- Amaral, L., Meurers, D., 2007. Conceptualizing student models for ICALL. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Volume 4511 LNCS, Pages 340-344.
- Chang, C.-K., Hsu, C.-K., 2011. A mobile-assisted synchronously collaborative translation-annotation system for english as a foreign language reading comprehension. In *Computer-Assisted Language Learning* Volume 24, Issue 2, Pages 155-180.
- Christensen, E., Merrill, P., Yanchar, S., 2007. Second language vocabulary acquisition using a diglot reader or a computer-based drill and practice program. In *Computer-Assisted Language Learning*, Volume 20, Issue 1, Pages 67-77.
- Ditcharoen, N., Naruedomkul, K., Cercone, N., 2010. SignMT: An alternative language learning tool. In *Computers and Education*, Volume 55, Issue 1, Pages 118-130.
- Fotouhi-Ghazvini, F., Earnshaw, R.A., Haji-Esmaili, L., 2009. Mobile-assisted language learning in a developing country context. In *2009 International Conference on CyberWorlds, CW '09*, Article number 5279514, Pages 391-397.
- Frohberg, D., Göth, C., Schwabe, G., 2009. Mobile Learning projects - a critical analysis of the state of the art: Original article. In *Journal of Computer-Assisted Learning*, Volume 25, Issue 4, Pages 307-331.
- Graesser, A.C., Chipman, P., Haynes, B.C., Olney, A., 2005. Auto tutor: An intelligent tutoring system with mixed-initiative dialogue. In *IEEE Transactions on Education*, Volume 48, Issue 4, Pages 612-618.
- Gu, X., Gu, F., Laffey, J.M., 2011. Designing a mobile system for lifelong learning on the move. In *Journal of Computer-Assisted Learning*, Volume 27, Issue 3, Pages 204-215.
- Jeng, Y.-L., Wu, T.-T., Huang, Y.-M., Tan, Q., Yang, S.J.H., 2010. The add-on impact of mobile applications in learning strategies: A review study. In *Educational Technology and Society*, Volume 13, Issue 3, Pages 3-11.
- Katsionis, G., Virvou, M., 2008. Personalised e-learning through an educational virtual reality game using Web services. In *Multimedia Tools and Applications*, Volume 39, Issue 1, Pages 47-71.
- Kuo, Y.-H., Huang, Y.-M., 2009. MEAT: An authoring tool for generating adaptable learning resources. In *Educational Technology and Society*, Volume 12, Issue 2, Pages 51-68.
- Kurata N., 2010. Opportunities for Foreign Language Learning and Use within a Learner's Informal Social Networks. In *Mind, Culture, and Activity*, Volume 17, Issue 4, pages 382 – 396.
- Motiwalla, L.F., 2007. Mobile learning: A framework and evaluation. In *Computers and Education* Volume 49, Issue 3, Pages 581-596.
- Sandberg, J., Maris, M., De Geus, K., 2011. Mobile English learning: An evidence-based study with fifth graders. In *Computers and Education*, Volume 57, Issue 1, Pages 1334-1347.
- Uther, M., Singh, P., Uther, J., 2005. Mobile Adaptive CALL (MAC): An adaptive s/w for computer-assisted language learning. In *Proceedings - International Conference on Pervasive Services, ICPS '05*, Volume 2005, Article number 1506556, Pages 413-416.
- Virvou, M., Maras, D., Tsiriga V., 2000. Student modelling in an intelligent tutoring system for the passive voice of English language. In *Educational Technology and Society*, pp. 139-150.
- Virvou, M., Alepis, E. (2005). Mobile educational features in authoring tools for personalised tutoring. *Computers and Education* (pp. 53-68).
- Zervas, P., Sampson, D.G., 2010. Enhancing educational metadata with mobile-assisted language learning information. In *Proceedings - 10th IEEE International Conference on Advanced Learning Technologies, ICALT 2010*, Article number 5571305, Pages 269-273.
- Zhou, W., Zheng, J., Lu, Q., Chiu, T., You, X., Ye, W., 2007. A computer-assisted language learning system based on error trends grouping. In *IEEE NLP-KE 2007 - Proceedings of International Conference on Natural Language Processing and Knowledge Engineering*, Article number 4368040, Pages 256-261.

TOWARDS THE AUTOMATIC ARCHITECTURE DRIVEN SOFTWARE MODERNIZATION

Dimitar Birov and Yanka Todorova

Faculty of Mathematics and Informatics, Sofia University, James Bouchier blvd. 5, Sofia, Bulgaria

Dimitar.Birov@fmi.uni-sofia.bg, Yanka@Todorff.co.uk

Keywords: Software modernization, Model driven software development (MDD), Model driven architecture (MDA), Model transformation, Domain specific language (DSL), Business process modelling.

Abstract: In this paper, we describe a strategy for implementing source code analysis, model extraction, edition and analysis and code generation tools that can be applied to a software modernization of existing legacy software intensive system. As well we present an integrated approach focused on model driven architecture for software modernizations. Started with an extraction models from source code and other available software artefacts, transforming these models in order to obtain modern structure of software, and then generate a code from these models. Bussiness modeling tools and models naturally fits in the proposed constuction. A tools implementing domain specific languages are integrated into framework in vertical pipeline toolsuite.

1 INTRODUCTION

The constant evolution of software technology leads to continuous modernization of IT systems and software. Software modernization is a main driver of software evolution. Current enterprise IT systems are very complex, large and dispersed, which makes tasks of modernization non-attractive from business point of view. Moreover modernization of complex, software intensive systems is very expensive task. The reasons for software modernizations can be very different: from technology driven (the obsolescence of a technology) through the increasing users business needs (continuously changing user requirements) to market and business reasons (integration of enterprise IT systems in merging companies).

Full redesign and redevelopment of legacy system is not possible in most cases due to lost domain knowledge and technical skills for modernization. Model-driven software development (Stahl, Voelter, Czarnecki, 2006) offers an opportunity for increasing the automation in software modernizations. The full automation of this process could not be performed because of internal and external software quality attributes have to be established like maintainability, testability, reliability, security, etc.

This paper revisiting the possibilities to semi-automate the processes of IT systems modernization. We based on extraction models from source code and other available software artefacts, transforming these models in order to obtain modern structure of software, and generate a code from these models. Adding bussiness modeling tools and models produced of them and merging these models with others is one of the contribution of this paper. Different tools supporting described process of software modernization exists, but they lack of integration in one and the same software development framework or ingerated development environment. Second contribution of the paper is a common framework or toolsuite with vertical integration of tools. As well the toolsuite reflects 4 model levels of abstraction.

The paper is organized as follows: Sections 2 gives an overview of domain specific languages. Domain specific languages are used in different stages of software development. They automate some software engineers and design activities. Section 3 overview some bridges between technologivcal spaces. During software development software engineers use knowledge, tools and experience from different areas - technology spaces. Section 4 presents proposed approach based on the vertical pipeline scenario for possible software

modernization focused on architecture driven approach of software development.

2 DOMAIN SPECIFIC LANGUAGES

A domain specific language (DSL) (Fowler, 2010) is a (programming, specification, modelling) language dedicated to a particular domain or problem. The advantage of a domain-specific language is that it provides appropriate built-in abstractions and notations according the problem domain. DSLs are used in a broad range of application domains—widely known example of DSL is MS Excel (Excel, 2010). Some software configuration or script languages can be viewed as domain specific. Software engineering and software architecture are very interesting problem domains concerning all process, stages and expertise of software development. In other words each DSL is specialised for a set of problems that share enough characteristics that it is worthwhile to study them as a whole. In this paper we focus on using DSL for software modernization. The problems of design and creating tools (compiler, interpreters) are out of scope of the paper and a lots of publications could be found.

Software modernization consists of following stages – first, extraction models from software artefacts (source code, databases, etc.), second extracted models are transformed to models, which are appropriate for modelling software components and software architecture, and third stage is automated code generation from obtained models. These three stages are supported with specific DSLs. The architecture driven software modernization is a specific problem domain with a knowledge, experience and technologies behind. This makes DSLs very useful instrument in the domain of modernization of software intensive systems.

DSL allows the different domain experts to be involved in the software development process. Domain experts can be software users who shares domain knowledge with the developers. It is not common software engineers to be an expert in the domain and additional resources have to be planned for their education. On the other hand, domain experts are often needed not only during the requirements specification, but also in the any stage of the development. These stages include design, modelling, verification and testing phases. If properly designed, DSLs provide a chance to

involve these domain experts in the design of complex models like software architecture or IT architecture. DSLs extend the range and collaboration of people being able to contribute to the software modernization of the product. If the software developing team use the specific languages that each party is a familiar, this will decrease time for production and will increase a quality of software product and will shorten the time for production.

Software (architecture) models can be specified using DSLs. DSLs provide enough abstraction that they can serve as model specification language during the design phase of a software development. Because many technical details are already built in the semantics of the DSL, the specification written in a DSL can often be used to automatically generate code that forms the implementation. Thus, DSLs often bridge the gap that exists between the phases of the software engineering process, especially between the design and the implementation phase. As we see later software designers can use DSLs for model transformation which allows in pipeline manner to transform one model to another, more abstract or more concrete model depending on purpose. Pipeline transformation allows not losing knowledge and expertise during transformation and same time increase an abstract level of models. As well each transformation step can be checked syntactically and semantically.

The DSL is good for documentation purposes in order to ease the communication between developers and customers due to semantics included in DSL. If the semantics of the DSL is formally specified using some mathematical notation, then the DSL can be used as specification language also, because an unambiguous description of the semantics exists.

2.1 Business Process Modelling

Business processes (Boev, Surova, Nikolov, Zhivkov, 2011) mirror business activities in the company. BPMN (Shapiro, White, Palmer, 2011) is a notation for graphical presentation of business processes and model business activities according to domain experts and business analysers. The knowledge encoded in diagrams will be saved during model transformation on later stages with purposes of optimization of business process. We could not generate executable code directly from these diagrams but after appropriate transformation we can reach models suitable for software assets (source code, database schema, configuration script, etc.) generation. After modification of the model

modified source code can be obtained semi automatically, which is syntax correct with respect of some general purpose language (GPL).

2.2 DSL Categories

Two types DSL are known – textual and graphical domain specific languages. Textual DSLs can be easily embedded in other general purpose languages (GPL). The editors for embedded languages are widely used and users are familiar with it. Graphical DSLs are more intuitive for domain experts and can be embedded in some graphical language like UML or BPMN language. Textual DSLs can be less understandable for domain experts, while for graphical DSLs can be difficult to develop tools with appropriate quality.

Two other categories of DSLs are internal (embedded) and external DSLs. Internal DSL is implemented inside of general purpose host language, and their characteristics vary depending on the features of base language. Sometimes embedded DSL is implemented as a library or framework. External DSL gives a maximum syntax and expressions freedom and requires a good language development support. DSLs discussed in the rest of the paper are external. This means that they have notation used by domain experts.

3 BRIDGING TECHNOLOGY SPACES

Model-driven approaches move focus of software modernization from last generation programming languages code to models expressed in some modelling language - UML for example. Models can graphically depict system's structure and behaviour at a certain point of abstraction. We can refer to a source code as a textual representation of a model of design concepts. In this paper this understanding is important, because we are not focused on the source code analysis and transformation techniques and approaches in details but we are going to treat source code as input for a model extraction and as an output after code generation from a model of design concepts.

3.1 Technological Spaces

The term technological spaces (TS) was initially proposed by Kurtev et al. (Kurtev, Bezivin, Aksit, 2002) to name “a working context with a set of

associated concepts, body of knowledge, tools, required skills, and possibilities" It is often associated to a given user community with shared know-how, educational support, and common literature. It is also special network of exchange expertise and ongoing research and a repository for abstract and concrete resources.

Five technological spaces are commonly recognized: Programming languages concrete and abstract syntax, Ontology engineering, XML-based languages, Data Base Management Systems (DBMS), Model-Driven Architecture (MDA). Each technology space is defined according to a couple of basic concepts: Program/Grammar for the Syntax TS, Document/Schema for the XML TS, Model/Meta-Model for the MDA TS, Ontology/Top-Level Ontology for the Ontology engineering TS and Data/Schema for the DBMS TS.

In this paper we outline the bridges between (abstract) Syntax TS (known as *grammarware*), XML TS (known as *documentware*) and Model TS (known as *modelware*) and integration of bodies of knowledge developed by different research software development communities. The grammarware technological space is concerned with grammars, grammar-based description languages, and associated tools. The modelware technological space is concerned with metamodels, model-based description languages, and associated tools, documentware is concerned with XML, XSLT and associated tools. The bridges between these three TS establish foundation of integration between transformation tools discussed in the paper. Integration is based on a physical, a logical, and a pragmatical bridge between grammarware language and modelling framework.

3.2 Bridge Grammarware to Modelware

Most of modernization scenarios (Ulrich, Newcomb, 2010) involve dealing with source code conforming to the grammar of a programming language. Some of additional software assets (like configuration files, script files, resource description files) which are formed using a formal language can be analyzed or manipulated by tools. These tools can automate knowledge extraction from software assets in some degree. This way the tools are significantly support bridging and understanding knowledge between the grammarware and modelware TS. Extracting and creating models from source code is a first step for architecture/model driven software modernization.

Two main groups of bridging approaches are known - approaches focused on grammars or syntax oriented approaches and approaches focused on models or model oriented approaches. Grammar-based approaches consider generation of metamodels from grammars, while (meta)model based approaches generate grammars from metamodel. Software modernization process starts with collection of some information from source code and other available software assets (user interfaces, databases, design documentation, configuration files, etc.) that's why grammar based approach prevails. Model-based approach suits significantly for code-generation phase in automated software modernization process, when from models a new code was generated which confirms a (previously specified) language grammar. Both process transforms text to model (T2M) and then model to text (M2T) in order to implement software modernization process.

The xText (Behrens, Clay, Efftinge, 2010) and the works of Wimmer et al. (Wimmer, Kramler, 2005) and (Kunert, 2006) are examples of grammar-based approaches. Operational semantics of modelled languages in modelware TS can be described formally as well (Sadilek, Wachsmuth, 2009).

This approaches lack of quality of generated models. M2M transformations are not easy and lots of manual work is needed to be performed in respect to obtain clean models suitable for code generation. These transformation languages do not provide construct to make transformation process easy. In order to obtain knowledge about the software system parsing of additional software artifacts has to be done. MoDisco (MoDisco, 2011) extract knowledge from different software artifacts during the model discovery phase - obtaining a model that represents a view on the legacy system (or at least parts of it) from its source code, raw data, available documentation, etc.. Next MoDisco phase consists of models analysis, particular model transformations are performed until the final (desired) software artifacts are obtained. MoDisco is an Eclipse open source project and is based on Eclipse Modelling Framework and integrates OMG/ADM standards (KDM, SMM).

Gra2Mol (Izquierdo, Cuadrado, Molina, 2008) is domain-specific model transformation language specially intended to deal with source code described by a grammar. Gra2MoL is a rule-based transformation language whose rules have a similar nature to that of other model transformation languages.

Each transformation definition consists of a set of transformation rules which specify relationships between grammar elements and metamodel elements. Gra2Mol raises significantly levels of abstraction of model extracted from source code. For example it is easy to extract knowledge KDM model from Java code.

3.3 Bridge Modelware to Grammarware

Widely used techniques in software development is (partially) code generation from models. The possibility to automate code generation process and obtain code straight from models adds to the flexibility, maintainability, and portability of application. The different tools exist depending on level of automation of the code generation process. In the next part of this section we will outline the WebDSL – DSL created for purposes of web applications. WebDSL (Hemel, Kats, Visser, 2008), (Hemel, Kats, Groenewegen, Visser, 2010), (Hemel, Groenewegen, Kat, Visser, 2011) allows to reduce amount of code developers need to write by introducing abstractions same time entire application is typechecked for errors.

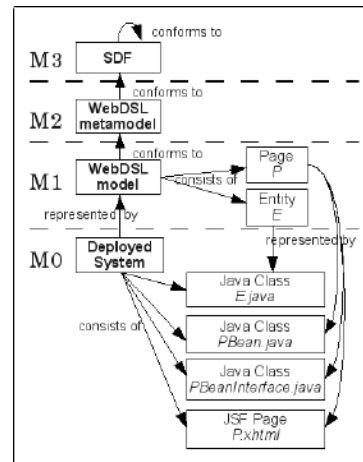


Figure 1: Organization of models of the WebDSL generator.

The architecture of WebDSL generator comprises the approach of code generation by model transformation, and follows the four-level model organization of Bezivin (Bezivin, 2005). Fig. 1. shows the model hierarchy – at top level (M3 meta-meta-model) is the grammar of the Syntax Definition Formalism SDF (Visser, 1997), (Hearing, Klint, Rekers, 1990). SDF is intended for the high-level description of grammars for wide spectrum

computer-based formal languages: general purpose programming languages, domain-specific languages, data formats and others. Any SDF definition describes the syntax of the language and the following step is to generate a working parser from this definition.

The grammar of WebDSL is defined in SDF at the M2 level meta-model and describes the valid sentences of the language. From the grammar, the parser automatically can be generated. Generated parser transforms the textual representation of a model to an abstract syntax tree (AST). All subsequent transformations are applied to the AST corresponding to the textual representation of the model.

The WebDSL generator transforms high-level models into Java code and XML files. The bridge between three TS – model TS, syntax TS (Java) and document TS (XML) is hardly coded in functionality of the WebDSL. On three of them we could apply transformation in syntax TS. We could transform from one language to another use code-to-code (C2C) transformation. In document TS we could apply different schema for XML transformation. Bridging between syntax TS and document TS is a subject of study during the last decade and it is well understood and established.

The transformations (of WebDSL model) are expressed in Stratego (Bravenboer, M., Kalleberg, K. T., Vermaas, R., Visser, E.. 2008),(Visser, 2004) transformation language. Stratego/XT is a high-level term rewriting system which implements the paradigm of rewrite rules with programmable rewriting strategies and integrates M2M, model-to-code (M2C), and code-to-code (C2C) transformations. A strategy is essentially function that controls the order of application of more basic transformations. As well Stratego provides programmable strategies for building complex transformations that control the application of rules. In Stratego, the application of rewrite rules is under the control of programmable strategies, such that transformations can be explicitly staged.

Using strategies, the WebDSL generator is divided into different transformation stages. Actually the generator is organized as a pipeline of model-to-model transformations. Each stage consists of a set of rewrite rules that rewrite extensions of the WebDSL core language to more primitive language constructs. This technique of compilation by normalization has advantage to reduce the semantic gap between input and output model, this way avoiding the complexity associated by directly generating code from the input mode.

Model level (M1, Fig 1) of WebDSL models web applications, which consisting of entity and page definitions. At this level not all models that conform to the WebDSL syntax are valid. That is why semantic analysis needs to be performed. A separate type checking stage of the generator performs checks. If static semantic constraints are violated an error reported. The semantic information gathered at this stage is also used to provide context information for other transformations.

Level M0 presents the actual web applications consisting of programming language constructs (like Java classes) and web (XHTML) pages. These software assets represent the models at the M1 level. M0 models can be implemented in different languages like PHP, Python, JScript etc. This is very useful and important for next modernization and software evolution. For example product lines and mobile application can be developed with a small amount of efforts. Moreover that M0 systems can consists of high-level application frameworks, in case of Java these are Java Persistence API (JPA), JavaServer Faces (JSF), etc. In some other cases more elegant and flexible approach for implementation is to insert middle level of intermediate embedded DSLs into the general purpose implementation language.

Extensions of the WebDSL language, such as the access control and workflow abstractions are realized as plug-ins to the base language, extending the generator with new normalization rules.

4 PIPELINED TOOLSUITE

In this section we present our proposal for software modernization based on previously described two bridges. First of it based on a source code (P), available artefacts, and documentation Artefacts(P) = P + DataBasesSchema (used in P) + UserInterface (HTML, Asp, Jscript, PHP web pages, dialog boxes, XML files, etc.) + DataBase(about design of P). Part of the artefacts confirms to some set of grammars GrSet(Art(P)). From this set we can generate models of representation MRep(Art(P))

$$\text{FMestr: GrSet(Art(P))} \rightarrow \text{MRep(Art(P))} \quad (1)$$

Model Extraction Functions (FMestr) is a set of functions which extract models from grammars for each artefact of software application and source code.

MRep(Art(P)) is a set of models. Each model confirms to model from a set of metamodels MMRep(Art(P)). These metamodels represent a

basic knowledge about the source code and other artifacts of the existing system. Part of this metamodels could be KDM of OMG group, which is basic knowledge repository. Another metamodels could be Abstract and/or Concrete Syntax Trees of the source P. Another example could be entity relationships schema for databases. This part is very similar to this one implemented in Gra2Mol.

If we look back on models hierarchy at Fig.1., M0 level of model abstraction is a source code P and Art(P), M1 level is presented by MRep(Art(P)) and M2 level is represented by MMRep(Art(P)). M3 could be SDF for description of metamodel construction.

Additionally to this some business process specification can be obtained – if it is available we can use it as an artefact of software. So this way BPMN notation is a part of Art(P)). If it not exist based on the users interviews we could create it using visual modelling tools. Visual languages for BP Modelling are part of Syntax TS supporting BPM. As well visual languages are DSL languages with appropriate tools supporting modelling. As a DSL language they have a syntax (graphical symbols) and semantic (incorporated into the models). Many (visual) languages for modelling web applications have been developed. WebML (Brambilla, Comai, Matera, 2007) support generation of (web) application from BP specifications (Brambilla, 2006). Transformation from existing BPMN to WebML can be automated using DSL. As it is shown on Fig 2. through manual activity we can create Choreography model. Then through model transformation we can obtain refinement set of models.

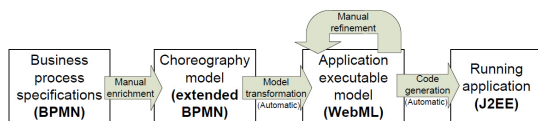


Figure 2: Automatic transformation of BPMN to WebML.

This way application executable model became a part of MRep(Art(P)) and respectively we can obtain extended part of MMRep(Art(P)). Merging models obtained both ways will enrich knowledge obtained from statically structured models with a knowledge of processes or dynamical models. Call Flow Graph (CFG) which is obtained from source code extraction as a part of MRep(Art(P)) carry some dynamical information but using BPMN models we could obtain almost complete information about dynamics of the process of execution.

The process of model transformation can be automated also. Using toolsuite like XTEAM-2 (XTEAM), (Edwards, Brun, Medvidovic, 2010) (eXtensible Tool-chain for Evaluation of Architectural Models) we can automate process of evaluation/creation of DSLs for manipulation of models. This type of tools performs model checking and model transformation. These tools operates on metamodel level M3 and transformed models to models through generated DSL.

$$\text{MMTools: MMRep(Art(P))} \rightarrow \text{MMRep(Art(Q))} \quad (2)$$

Art(Q) is a new set of models of artifacts of a new modernized system Q. So next step from our proposal goes down from refined and appropriate metamodels to obtain concrete models (one of them is software architecture model) which will represent concrete artefacts of the new system Q. We got idea for this from process of code generation, described for WebDSL. We need particular DSLs for obtaining models for each one of the metamodel in MMRep(Art(Q)). Automatic creation of DSLs could be done with XTEAM.

Code generation step is very trivial and well-studied – from models of artefacts of Q we can create the concrete exemplars for concrete platform. This approach makes process of software modernization very flexible.

5 CONCLUSIONS

In this paper we present an integrated approach focused on model driven architecture for software modernizations. Started with an extraction models from source code and other available software artefacts, transforming these models in order to obtain modern structure of software, and generate a code from these models. Bussiness modeling tools and models naturally fits in the proposed constuction. A recent (versions of) tools from different technological spaces are integrated into framework in vertical pipeline toolsuite.

Some details of integration can be a subjects for future research. The prototype of this framework in Eclipse integrated development environment is under development.

ACKNOWLEDGEMENTS

This research was fully supported by the Bulgarian National Science Fund under Grant DO 02-102/23.04.2009.

REFERENCES

- Behrens X., Clay M., Efftinge, C., et al. *Xtext User Guide*, http://www.eclipse.org/Xtext/documentation/1_0_1/xt_ext.pdf, 2010
- Bezivin, J. *On the unification power of models*. Software and System Modeling, 4(2):171–188, 2005.
- Boev, S., Surova, E., Nikolov, K., Zhivkov, V., *Incorporating collaboration In Business Processes.*, First International Symposium on Business Modeling and Software Design, BMSD 2011
- Brambilla, M., *Generation of WebML Web Application Models from Business Process Specifications*, Demo at 6th International Conference on Web Engineering (ICWE2006), July 2006, Palo Alto, CA, USA
- Brambilla, P. F. M., Comai S., Matera, M. Designing web applications with WebML and WebRatio. In G. Rossi et al., editors, *Web Engineering: Modelling and Implementing Web Applications*, Human-Computer Interaction Series. Springer, October 2007.
- Bravenboer, M., Kalleberg, K. T., Vermaas, R., Visser, E.. *Stratego/XT 0.17. A language and toolset for program transformation*. Science of Computer Programming, 72(1-2):52-70, 2008.
- Edwards, G., Brun, Y., Medvidovic, N., *Automated Analysis and Code Generation for Domain-Specific Models*, Technical Report USC-CSSE-2010-517, Center for Software and Systems Engineering, University of Southern California, August 2010.
- Excel, <http://office.microsoft.com/en-us/excel/>
- Fowler M., *Domain-Specific Languages*, book, Addison-Wesley, 2010, ISBN: 0321712943, 9780321712943
- Hearing J., Klint P., Rekers, J. *Incremental generation of parsers*. 1344--1350. IEEE Transactions on Software Engineering. 16. 12. 1990.
- Hemel Z., Kats L. C. L., Visser E. *Code Generation by Model Transformation. A Case Study in Transformation Modularity*. ICMT 2008, LNCS 5063, pp 183—198. Springer, June 2008. (An updated, extended version was published in 2009 in SoSyM.)
- Hemel Z., Kats L. C. L., Groenewegen, D., Visser, E. *Code Generation by Model Transformation. A Case Study in Transformation Modularity*. Software and Systems Modeling, Vol. 9, Iss. 3, pp 375—402, Springer, 2010.
- Hemel, Z. , Groenewegen D. M., Kats L. C. L., Visser, E.. *Static Consistency Checking of Web Applications with WebDSL*. Journal of Symbolic Computation, Volume 46, Issue 2, Elsevier, 2011.
- Izquierdo, J. L. C., Cuadrado, J. S., Molina, J. G.. *Gra2MoL: A domain specific transformation language for bridging grammarware to modelware in software modernization*. In Workshop on Model-Driven Software Evolution, 2008
- Kunert, A., *Semi-automatic generation of metamodels and models from grammars and programs*, Fifth Intl. Workshop on Graph Transformation and Visual Modeling Techniques, E. N. in Theoretical Computer Science, Ed., 2006.
- Kurtev I., Bezivin, J., and Aksit M.. *Technological spaces: An initial appraisal*. In CoopIS, DOA'2002 Federated Conferences, Industrial track, 2002
- MoDisco, <http://www.eclipse.org/MoDisco/>
- Sadilek, D., Wachsmuth, G., *Using Grammarware Languages To Define Operational Semantics of Modelled Languages*, TOOLS '09: 47th International Conference Objects, Models, Components, Patterns, Springer, 2009.
- Shapiro, R., White, St., Palmer, N., et al, BPMN 2.0 Handbook, book, 2011, ISBN: 9780981987033
- Stahl, T., Voelter, M., Czarnecki K., *Model-Driven Software Development: Technology, Engineering, Management*, book, John Wiley & Sons, 2006 ISBN:0470025700
- Visser, E. *Syntax Definition for Language Prototyping*. PhD thesis, University of Amsterdam, September 1997.
- Visser, E. *Program transformation with Stratego/XT: Rules, strategies, tools, and systems in StrategoXT-0.9*. In C. Lengauer et al., editors, *Domain-Specific Program Generation*, LNCS 3016, pp 216–238. Springer, June 2004.
- Wimmer, M., Kramler, G.: *Bridging Grammarware and Modelware*. 4th Workshop in Software Model Engineering (WiSME'05), October 3rd, 2005.
- Ulrich W, Newcomb Ph., *Information Systems Transformation: Architecture-Driven Modernization Case Studies*, Morgan Kaufmann OMG Press, 2010, ISBN: 0123749131
- XTEAM <http://softarch.usc.edu/~gedwards/xteam.html>

**SPECIAL SESSION ON
EMBEDDED SYSTEMS AND
BUSINESS MODELS**

ANALYSIS OF METHODS AND TECHNICAL TOOLS FOR TRAFFIC CONTROL SYSTEMS

Rosen Vitanov, Plamen Balzhiev and Roumen Arnaudov

Faculty of Telecommunication, Technical University, 8 Kl. Ohridski blvd., Sofia 1000, Bulgaria
rosenvitanov@mail.bg, baljiev@abv.bg, ra@tu-sofia.bg

Keywords: Road traffic and control, Counting vehicles, Measurement.

Abstract: In this article we analyse different methods for remote non-contact counting of moving objects at different speeds: video, mechanical, optical, ultrasonic. The principles of their implementation, advantages and disadvantages are described and the reliability of information processing methods is evaluated. A model of road traffic measurement and analysis system is presented. It is specially emphasized on application of different sensors in real-time data acquisition, management and analysis of traffic intensity.

1 INTRODUCTION

Nowadays, everyone is faced with road traffic problems due to the ever growing urban population, the formation of mega-cities, the inability of the road infrastructure to take the ever increasing road traffic stress and inefficiency of the existing traffic management systems. To solve those problems it is necessary to collect, analyse and systematize traffic information - when and where traffic is densest, what vehicles cross the road - cars, vans, trucks, buses. Obtaining the data requires the use of different methods and tools for identification and counting of the moving vehicles.

Once the information is obtained it can be organized and used to build intelligent systems for traffic management and statistical databases. It will be useful in various business areas such as logistics, navigation, courier services in search of the most efficient way of transportation to reduce fuel costs, saving time and hence reduce air pollution. Therefore, information needs to be frequently updated and to be with the highest precision. Collected information can be used for traffic signalization systems, opening or closing highway roadbeds. It can also be stored in databases and analysed at later stages in case a decision have to be taken if a road must be created, when a maintenance should be carried out or how road infrastructure must be changed.

This article aims to make comparisons between various methods of traffic detection, highlighting their strengths, weaknesses, areas of application,

efficiency, the most common errors in measurement and finally make brief conclusions.

2 ROAD TRAFFIC DATA COLLECTION SYSTEM

As defined by The National Electrical Manufacturers Association [NEMA] it is "a system for indicating the presence or passage of vehicles." An information system responsible for collecting data about road traffic in real time from one or many junctions, highways, roads or streets requires sensors to be deployed to count the passing cars. To obtain more detailed information multiple sensors have to be placed on a sufficiently large number of places. Therefore, the designed systems should be cost-effective, with great accuracy, not to be easily broken, to require minimal (or not any at all) changes on the existing road infrastructure, without interference with the natural flow of vehicles, to be easily moved from one place to another, to be easily accessible.

All mentioned requirement demand sophisticated, intelligent systems that can manage to process and analyse all the acquired data from the sensor network and detect certain events on the road. On Figure 1 presents intelligent traffic monitoring and analysis system. It includes several traffic detection units which count and transmit data to the data acquisition and analysis unit for processing. All

data and result could be monitored and further processed.

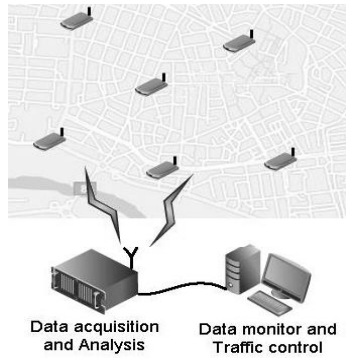


Figure 1: Road traffic data collection and analysis system.

3 ROAD TRAFFIC DATA COLLECTION METHODS AND SENSORS

Nowadays there is a great variety of efficient vehicle detection methods. “In order to assess the present and future traffic demands, for the development of need-based infrastructure accurate information and continuous monitoring of traffic is necessary” (Ministry of Works and Transport of , Gaborone, Botswana, 2004, Chapter 1.2)[MWTGB]. Traffic count technologies can be in general divided by the place, where sensors are situated - in-the-roadway (intrusive) and over-the-roadway (non-intrusive).

3.1 Intrusive Methods

Intrusive methods rely on direct interaction between the vehicle and the sensors along the road. The implemented sensors are deployed can be divided in three groups - embedded in the road pavement, embedded in the sub-grade of the road, taped or attached to the surface of the road. Typical representatives of the in-the-road methods include inductive-loop detectors, which are saw-cut into the pavement; magnetometers, which may be placed underneath a paved roadway or bridge structure; and tape switches, which are mounted on the roadway surface.

3.1.1 Pneumatic Sensors

Pneumatic sensors rely on a direct hit of a vehicle to detect it. In most cases a “rubber tubes are placed across the road lanes to detect vehicles from pressure changes that are produced when a vehicles’

tires passes over the tube. The pulse created is recorded and processed by a counter located on the side of the road. The main drawback of this technology is that it has limited lane coverage and its efficiency is subject to weather, temperature and traffic conditions. This system may also not be efficient in measuring low speed flows (Leduc, 2008).

Another drawback is the great wear factor of the tubes - the more vehicles cross the tubes the bigger chance of micro punctures to appear. Constant changes of seasons will make the rubber degrade more rapidly than usual. Another drawback is the tubes must be filled with temperature independent gas, like Nitrogen, so the pressure inside them will not rely on any environment changes. Also if two or more cars hit the tube at the same time, sensor will miscount. On the other hand they are very easily deployed on the road. They need no other preparations (like digging the road or mounting on poles). Suitable for short-term counts of roads with low traffic (suburb roads, etc.).

3.1.2 Optical Fiber Sensing

Optic fibers placed in plastic and/or rubber tubes could be implemented as detection sensors. When a vehicle passes over the fiber a fluctuation of the light stream appears. A photo-receiver detects the changes in the optical signal and converts them in electrical impulses. Due to tenderness of the fiber core it can be easily damaged by heavier or high-speed moving vehicles. Therefore this method needs a thicker housing for better protection of the optic fiber. Some of the drawbacks due to weather conditions in pneumatic sensors are missing – like weather and temperature dependencies. For insurance, the fiber optic cable can be put into pre-cut line in the road, and then fixed with flexible gum. Figure 2 presents a measured signal of external interference on the optical fiber. As seen there is a noticeable change in measured signal which could be utilized in traffic measurement and analysis.

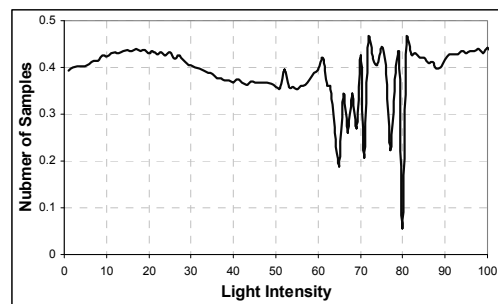


Figure 2: Fiber optic signal and movement detection.

3.1.3 Piezoelectric Sensors

Piezoelectric sensors described by Vehicle Detector Clearinghouse (2000)[VDC], are sensors that use piezoelectric effect to detect passing vehicles. Drawbacks of this method are that a certain part of the pavement must be cut. In winters if a pile of snow is on the road the sensor may miscount. Sensor can brake if it is exposed to extreme stresses. They are independent to weather and climate changes. Also a miscount may occur when the road's surface extrapolates due to extreme temperatures.

3.1.4 Inductive loop Sensors

Inductive-loops sensors (Figure 3) are wires that are placed under the road surface in square-like shape. They are connected to a detection unit which collects traffic data. Then an electrical signal is fed to the wires and electromagnetic field is created. When a car passes over it the induction of the loop changes and a vehicle is detected ([VDC, 2000],[MWTGB, 2004]). However this method is expensive due to the fact that the inductive-loop must be placed beneath the road surface.

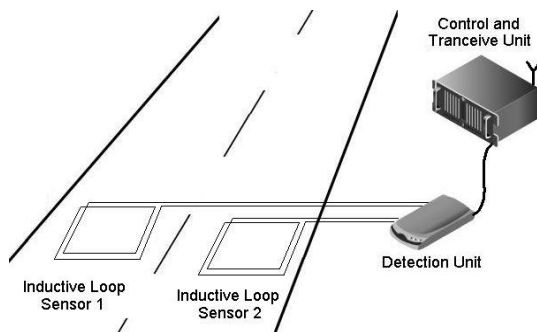


Figure 3: Implementation of inductive loop sensors.

Non intrusive methods are methods where traffic counting sensors do not interact directly with the passing vehicles. They need less preparation time, efforts and

3.2 Non-intrusive Methods

Non intrusive methods are methods where traffic counting sensors do not interact directly with the passing vehicles. They need less preparation time, efforts and investments. Most of them are located above the road lanes, others like video image processing cameras, can be placed next to the road, which leads to minimum or to none deployment time, money and work. „Examples of over-roadway sensors are video image processors that utilize

cameras mounted on tall poles adjacent to the roadway or traffic signal mast arms over the roadway; microwave radar, ultrasonic, and passive infrared sensors mounted in a similar manner; and laser radar sensors mounted on structures that span the lanes to be monitored“ [USD,2006].

3.2.1 Manual Counting

Manual counting is quite common method for counting traffic. Not too expensive, does not need any preparations. Data is recorded on a paper or in a computer by one or two specially trained professionals (Klein, Mils, Gibson 2006). However it's not suitable for long-term data collection or highways with intense traffic. Some counting error may occur due to “human error” factor.

3.2.2 Ultrasonic Sensors

Ultrasonic sensors can be also divided in two groups active and passive. Passive are not very accurate on high-intense traffic areas. That's why they are not widely-accepted. On the other hand active sensors count at greater precision. They emit their own sound wave which reflects form vehicle's surface and is received by a sensor.

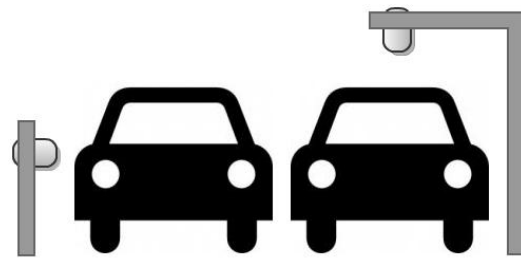


Figure 4: Sideway and over-road ultrasonic sensor implementation and detection.

Using some signal-processing algorithms speed and direction could be measured and vehicle classification could be estimated. They are weather-independent, cannot be easily tricked and can be used on high-traffic roads and highways. Also they can determine with increased accuracy vehicle types even detect trailers [VDC, 2006].

3.2.3 Video Image processing

Video image processing is a method where instead of sensors a video camera is used to detect traffic and vehicles. Images from the camera are sent by wire (or wirelessly) to a remote station where images are stored and processed. There are many

algorithms for detecting vehicles on the road – like signal processing, artificial neuron networks, bitmap processing etc. After a vehicle is detected another algorithm is started to detect its speed, direction, and even dimensions.

This method can provide us with different kind of information. As this is still a developing methodology there are some miscounts and system can be easily fooled. It depends on the weather conditions for example in foggy or too snowy conditions nothing will be detected. Front protecting glass of camera has to be cleaned frequently otherwise the camera will not be able to detect anything [VDC, 2007],[MWTGB, 2006], (Klein, Mils, Gibson 2006).

4 CONCLUSIONS

Intrusive methods for vehicle counting purposes need preparation steps, such as cutting big parts of the road surface, deploying sensors, reconstruction of the cut slots after sensors were deployed. This with no doubts leads to more expenses, inconvenience to drivers since the road has to be closed for a certain period of time (depending on the size of the project). Sensors can be easily broken due to constant interaction with the passing vehicles or due to cavities in the pavement, which will uncover the sensors. They are suitable for small roads, where traffic is not quite intensive and closing the road will not lead to major inconvenience for drivers. The accuracy and the reliability in some of them depend on weather conditions and sometimes they may give wrong information.

On the other hand non-intrusive methods are more easily-installable to the existing infrastructure. Their positioning is easier than intrusive sensors, because there is no need of any road modifications. Some of them are vulnerable to weather conditions, but measure with greater accuracy. Further more there is no need traffic to be stopped during their installation. If a short-term measurement is needed without road interventions, pneumatic method or optical sensors are most appropriate.

REFERENCES

Leduc, G., 2008. *Road Traffic Data: Collection Methods and Applications*. JRC 47967, European Commission, Joint Research Centre, Institute for Prospective Technological Studies.

2004. *Traffic Data Collection and Analysis*, Ministry of Works and Transport, Roads Department, Gaborone, Botswana.
- Klein L. A., Mills M. K., Gibson, D. R. P., 2006. *Traffic Detector Handbook: Third Edition*, Vol.1- FHWA-IP-90-002, US Department of Transportation publishing company.
- Gibson, D., Milton K. Mills, Klein L. A., 2007. *A New Look at Sensors*, Vehicle Detector Clearinghouse publishing company.
- Hoshide, G. *Evaluation of Recreation Traffic Counters*, United States Forest Service.
- Luz Elena Y. Mimbela. *A Summary of Vehicle Detection and Surveillance Technologies*.
- Luz Elena Y. Mimbela, 2000. *A Summary of Vehicle Detection and Surveillance Technologies use in Intelligent Transportation Systems book*, Vehicle Detector Clearinghouse.
2006. *Freeway Management and Operations Handbook book*, US Department of Transportation.
- Hoshide, G. *Evaluation of Recreation Traffic Counters*, United states forest service.
2004. *Traffic Data Collection and Analysis*, Ministry of Works and Transport, Roads Department, Gaborone, Botswana.

**SPECIAL SESSION ON
HUMAN-CENTERED INFORMATION
SYSTEMS MODELING**

WHY DOES SOCIAL CONTEXT MATTER?

Integrating Innovative Technologies with Best Practice Models for Public and Behavioral Health Promotion

Luba Botcheva and Shobana Raghupathy
Sociometrics Corporation, Los Altos, California, U.S.A.
{lbotcheva, shobana}@socio.com

Keywords: Technology, Innovation, Acceptance, Integration, Adoption, Behaviour, Change, Health, Social context.

Abstract: In this paper we will present a framework that is intended to guide synthesis of different theoretical perspectives for the purpose of developing strategies for integrating IT use in diverse social settings. First, we will briefly review existing theoretical models grounded in behavioural science; and present our company's approach for development of products using technology innovation that take in account the individual, organizational and contextual community characteristics. Secondly, we will illustrate this approach with three case study examples in the fields of public/behavioral health and education. Finally, we will conclude with theoretical and practical considerations that can be used by IT developers to maximize adoption and implementation of innovative technologies.

1 INTRODUCTION

It is universally acknowledged that information and communication technologies (ICTs or IT) hold huge potential for enhancing the effectiveness of services in different sectors. Their ability to reach new populations, improve communications, and transform service delivery can increase the effectiveness and social impact of private, public and non-profit initiatives. Yet, the positive effects of innovative IT will only be fully realized if, and when, they are widely spread and used. It is a well-accepted fact that the existence of technology does not guarantee its utilization. Attempts to promote the adoption and diffusion of innovative IT have often failed due to a lack of understanding of the factors that affect acceptance and use of technology by individuals and organizations. A notable example is the recent failure of the One Laptop Per Child (OLPC) project, which aimed to distribute millions of \$100 laptops to disadvantaged school children but failed to anticipate the social and institutional problems that could arise in trying to diffuse technology in the developing country context (Kraemer, Dedrick, & Sharma, 2009).

As the OLPC case demonstrates, investigating social context is vital to understanding the acceptance and use of technologies. Researchers have often addressed the issue of why individuals

and organizations who would benefit from technological systems do not use them but traditionally most of the research has focused on technological factors and has rarely been applicable to different sectors and social contexts (Al-Gahtani, 2008). With the rapid utilization of IT in different spheres of life and across geographical and economic dimensions, best practice models have shifted focus to the potential adopter and the organization or community into which the technology will be integrated. An adopter based, instrumentalist approach incorporating both macro- and micro-level perspectives now appears to be the most widely used to promote the adoption and diffusion of innovative ITs.

However, a gap exists between these best practice models and IT adoption strategies. In particular, the non-profit and public sectors as a whole lag behind the private sector in the adoption of technologies. There has been little scholarly research into the IT adoption in the non-profit and public sectors. This paper discusses ways that technology acceptance models can be utilized to develop multi-level approaches for facilitating IT adoption in diverse social settings (educational, health and community-based organizations) with special emphasis of the contextual characteristics that determine the success of this process.

2 BEST PRACTICE IN TECHNOLOGY ADOPTION

2.1 Existing Theoretical Models

A rich body of literature has emerged that employs behavioural science theories to model factors affecting the acceptance and use of technology at both the individual level and the organizational level. Prominent examples include the Technology acceptance model (Davis, Bagozzi, & Warshaw, 1989), Theory of planned behaviour (Ajzen, 1985), Unified theory of acceptance and use of technology (Venkatesh, Morris, Gordon B. Davis, & Davis, 2003), Diffusion of innovation (Rogers, 1995), and the Technology, organization, and environment framework (Tornatzky and Fleischer 1990). Some of these models focus on technological and individual factors influencing acceptance but as models have become more sophisticated and better validated, there has been an increasing acknowledgement of the centrality of environmental and contextual constructs. Key constructs in these models that relate to contextual factors include compatibility (with existing technology, work practices, beliefs and values), social influence, professional environment, organizational structure. In addition, many of the models discuss individual factors that are inevitably related to broader environmental and social contexts, such as attitudes, beliefs and subjective norms. A

comparison of theories of change and contextual constructs defined in models of technology acceptance is presented in the Appendix.

Drawing from these theories we present a revised multi-contextual model, adapted from the extended technology acceptance model (Dadayan & Ferro, 2005) to incorporate broad social influences such as culture, community context, and ideologies that are critical for technology adoption in real life contexts (see Figure 1).

- Individual context refers to the characteristics of individual end-users and their attitudes to technology;
- Technological context refers to the characteristics of the technology such as functionality and user-friendliness; and
- Implementation context refers to the user’s environment, including organizational factors (climate, support, readiness); broad social influences (community, culture, ideologies); and technology compatibility.

In the next sections drawing on the extensive experience of our company in developing innovative products for public and behavioural health promotion, we will outline the methodology that we use to examine these different characteristics for developing technology products that are adequately integrated within real-life contexts.

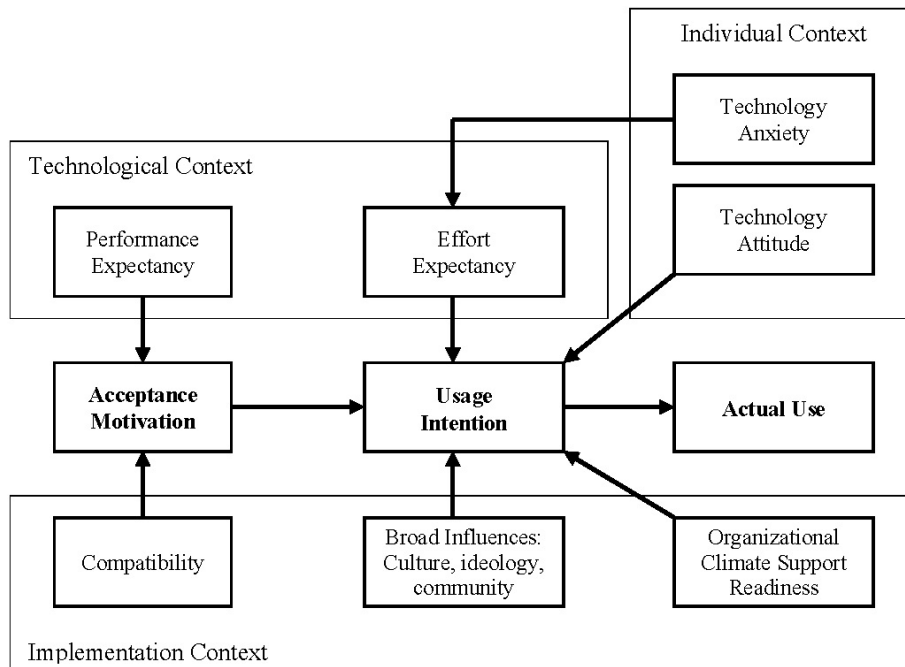


Figure 1: Revised Multi-contextual Model of Technology Acceptance.

2.2 Our Approach: The Intersection Between Research, Technology and User Needs

Established in 1983 and selected as an exemplary small business in 2007 by the National Academy of Sciences, Sociometrics' primary mission is to develop and disseminate behavioural and social science research-based resources for a variety of audiences in order to: 1) promote healthy behaviours; and 2) prevent or reduce behaviours that put an individual's health and well-being at risk.

During the last ten years, Sociometrics' staff has developed engaging IT public and behavioural health promotion products and has also accrued significant experience developing and disseminating research-based materials tailored for diverse target audiences through large scale websites, digitalized effective program materials, data libraries, and evaluation and training e-tools (see full description at www.socio.com).

All products are developed based on a thorough examination of user needs and preferences with a special attention to the contexts in which they will be implemented to assure their acceptance and relevance.

3 CASE STUDIES

3.1 e-Learning Products for Early Intervention Professionals

In this case example we present the process of deve-

loping interactive program tutorials tailored to the different learning styles of early intervention professionals working in diverse settings (child care centres, hospitals, and community-based centres). The project was part of a Sociometrics initiative funded by NIH aimed at assembling in one place—for public dissemination, distribution, and replication—treatment programs in the area of early childhood intervention. One of the important goals of the project was to design technology assisted professional tutorial materials to assist early childhood professionals to implement programs with fidelity in their professional settings. Following the principles of Dabbagh and Bannan-Ritland (2005) for on-line learning design, we first identified users' characteristics and learning styles and then explored their learning and professional context. Using interviews with potential users, observations of their professional context and reviews of relevant literature we were able to outline contextual characteristics that determined the design and technological modality of the professional tutorials (see Table 1).

3.1.1 Individual Context

The analysis of the individual context showed that technology based products are not used routinely by potential users, which determined relatively high technology anxiety; attitudes to technology vary among professionals with medical and administrative personnel being more positive. In response to this context we decided to create learning tools that are easy to use by people with no previous experience with technology.

Table 1: Context analysis for developing early childhood intervention professional tools.

Context	Early Childhood Intervention Contextual Characteristics	Solutions
Individual Context Technology anxiety Technology attitudes	Technology anxiety relatively high Attitudes vary among professionals but in most cases technology based products are not used routinely	Create learning tools that are easy to use by people with no previous experience with technology
Technological Context Performance expectancy Effort Expectancy	Performance expectancy is high; users will not use these tools if they do not believe that they will improve their direct work Based on the high work load and relative low level of technology skills, the effort expectancy is for easy use	Constructivist approach— learning by doing, e.g., cognitive apprenticeships, situated learning, problem-based learning, efficient learning (minimizes tangential activity) Practical (product of instruction is useful for everyday activity)
Implementation context Compatibility with existing technical systems Organizational support Professional culture	Technical system limited to basic software Technology support often is scarce Resistance to change; conservative organizational climates	System that is compatible with basic software Materials that will not require a lot of additional support Adaptable to diverse learners' styles and contexts

3.1.2 Technological Context

The technological context analysis showed that performance expectancy is high; users will not use these tools if they do not believe that they will improve their direct work; based on the high work load and relative low level of technology skills the effort expectancy is for easy use. To match these expectations we selected a constructivist approach for the pedagogical design of the products that is efficient and practical.

3.1.3 Implementation Context

We analysed several aspects of the implementation context: 1) compatibility with existing technological systems: most of the technical systems used by our users were limited to basic software, which meant that the professional tools should utilize basic software that is widely available; 2) Professional culture; most of the professionals worked in the non-profit sector, either at health or educational environments that have been shown to be relatively conservative and resistant to change (Botcheva et al.2003): thus the use of our tools should not require a lot of changes in the routine practices; they should be tailored to different professional environments and specific learner needs allowing customisation; 3) Organisational support of implementation: technology support is scarce in most of the organisations; thus the tools should be easy to maintain with minimum technical support.

This multi-level analysis led us to the decision to create e-learning materials using Adobe interactive PDFs that will include hands on examples and resources tailored to the different learning styles and experience of the individual learner. Interactive PDFs fit seamlessly into the complex pattern of diverse learners' needs, constraint and resources. On the one hand, they are: completely stable; typically get past firewalls; require no special software/system other than Adobe Reader; printable, and easy to maintain. On the other hand, they are: highly dynamic; allow audio, video, automations; indexing/bookmarking and easy linking. These characteristics helped us to create interactive and engaging e-learning tools that fit the context of and user characteristics of early childhood professionals.

3.2 Using the Individual, Technological and Implementation Context to Design e-Tools for Data Collection

Public schools in the United States are required to

annually collect and report data on drug use and other high-risk behaviours from elementary, middle and high school children. All schools receiving federal and state funding are expected to collect baseline data for establishing incidence or prevalence of data on truancy rates, drug and violence related suspensions and expulsions, drug incidence, and prevalence rates, and for demonstrating simple percentage changes in outcomes for end of the year performance reports.

It is often difficult, however, for schools to engage in periodic data collection efforts in the light of budget constraints and time constraints (Mantell, Vittis, & Auerbach, 1997; Sedivy, 2000). Teachers are expected to take on responsibilities other than teaching even at a time when there are increasing pressures on them to raise students' academic achievement levels. Thus, collection and monitoring of data on substance use or other health concerns are perceived as consuming valuable time (Hallfors, Khatapoush, Kadushin, Watson, & Saxe, 2000). In response to this need, Sociometrics designed a web-based survey development and analysis tool that would allow for swift, efficient and most importantly, cost-effective data collection, analysis and reporting. The online system would allow students to login to a pre-programmed survey with measures on drug use patterns, truancy and other high risk behaviours; answer the survey questions; then logout once he or she is finished. The survey data would automatically be deposited in a secure web server and can be accessed by the teacher for analysis. Such an e-tool is cost effective as it automates the survey creation and administration process, and relieves the teacher of burdensome tasks such as printing surveys, distributing them and then entering and processing the data.

In designing such a tool, we first started with a needs assessment that took account of the larger operational context: specifically, we investigated the wide range of constraints, limitations and facilitating factors at the individual (teacher), technological (school infrastructure), and implementation context (school). We first identified the primary consumers of the product. These included not just school teachers and principals who were responsible for the data collection and reporting, but also district and state level supervisors at the State Department for Education who were responsible for school funding allocations and monitored school progress. Next, we conducted numerous focus groups and interviews with the target audience. The qualitative studies yielded useful insights into current data collection efforts in schools, and offered valuable design,

content and dissemination guidelines for our e-tool. Some of these insights are outlined below.

3.2.1 Individual Context

Teachers began their data collection efforts by selecting measures, creating paper and pencil surveys, and administering the surveys in the classroom during recess, after school hours or whenever possible, during a health education class. The data were coded and entered by hand and then reported to the district. The schools had to provide districts with end of the year “performance reports” which reported simple changes in drug and violence incidents over the school year as part of their assessment. Numerous problems with the process were reported, such as insufficient funding and time, as well as lack of technical assistance. Those in the poorly funded districts mentioned high student movement and attrition, and problems in tracking students. Teachers complained about the time and effort involved in tracking such data and preparing reports. They also complained about not receiving any form of technical assistance from their districts in terms of selecting appropriate measures. Many teachers lacked the capacity to conduct basic statistical analysis such as means, and percentage changes in key behavioural indicators. It therefore became clear to us that any online data collection tool would need to pre-program measures and indicators that were popular, reliable and validated. We conducted a poll of the most popularly used measures and indicators for capturing school performance and developed pre-programmed surveys incorporating such indicators. A basic statistical tool was developed that would allow teachers to derive summary statistics such as means and frequencies (e.g. no. of student arrests on drug related charges, percentage of boys and girls referred to treatment services etc) without having download data or use any external software such as SPSS. (Teachers were also given the option to download the data if they desired). Finally, it became clear that because of time and capacity limitations, an interface needed to be created that would allow students to take the survey from multiple locations, and over different points in time. This realization led not only in interface design changes, but also the format and structure in how data were to be stored.

3.2.2 Technological Context

One major concern that emerged was whether the introduction of a new, online data collection system would introduce a steep “learning curve” for

teachers. Another concern was related to the investments that the school would be required to undertake in order to adopt it. To emphasize the utility and user-friendliness of the e-tool, we decided to present our product concept in images and terms that were already familiar to the groups. For example, we used an existing online survey system—SurveyMonkey—to illustrate our e-tool and the precise manner in which it would differ and be specially tailored for their school needs. Some of the school teachers in the focus group were already using online teacher evaluations: we encouraged them to speak to the group about their experiences (both positive and negative), and highlighted how our product would attempt to overcome the limitations and replicate the successes of their experience. By the end of our needs assessment, teachers and administrators were by and large, receptive to the idea of online data collection indicating that our strategies for reducing “technology anxiety” and establishing “performance expectancy” were successful.

3.2.3 Implementation Context

During the design and product development stage of any school-based e-tool, it is absolutely essential to ensure its compatibility with the school’s technological infrastructure. A technology “screeener” was mailed out to the focus group participants and interviewee’s participants in order to assess the basic “minimum” technological capacity that the e-tool would have to be compatible with. Questions included: number of computers in the school, Internet access, and bandwidth etc. While almost all schools in our focus group had Internet access, it became clear that lack of access to sufficient computers (along with time constraints) was yet another feature that necessitated group log-ins from multiple locations (such as libraries, computer labs and even homes). Privacy and confidentiality of student data subsequently emerged as a concern. As a first step, we designed a single question interface with autoprogession; the screen automatically gets refreshed once a question was answered thereby minimizing the amount of time a response was present on screen. The interface also included separate login IDs for students and administrators. Students could use their IDs to login from any location that was convenient to them while the administrator had sole control over the data collected. Administrators using the online statistical tool (described earlier) would be able to do so without accessing individual student data. If the

administrator did choose to download the data, the downloaded data was made available *without subject identifiers* in order to maintain confidentiality.

Besides confidentiality, another concern was related to product pricing and affordability. Participants identified “frontline” funding and decision making agencies and offices at the state, county and district level that could spearhead the use of online data collection mechanisms in their districts. We learned that pursuing business opportunities with state and district agencies (rather than individual schools) would allow costs to be incurred by these agencies and would facilitate large scale adoption of the technology at the ground level.

4 IMPLICATIONS FOR DEVELOPERS

There are several theoretical and practical implications for developers that stem from this analysis.

First, the review of existing theoretical models of technology acceptance highlight the importance of developing multi-dimensional approaches that take in account different social contexts to fully understand the processes of technology integration in real life contexts. Interdisciplinary teams incorporating the knowledge and skills of technology developers, social and behavioural scientists will be best suited to solve this problem.

Second, the analysis of technology acceptance in the non-profit sector highlights the critical importance of broad social context, such as culture, ideology, and community climate. While in industry, the transition from research and development to the field primarily focuses on the end user, in the non-profit sector, there is a range of intermediary factors (agencies, policies) that influence if and how the product reaches the end user. Thus, conventional theories regarding technology diffusion and adoption need to be modified with regard to the non-profit and public sectors.

Third future research and development effort should focus on development of practical tools and screeners that will facilitate the translation of contextual characteristics into technical requirements for development of products that can be easily adopted and integrated in real life contexts.

ACKNOWLEDGEMENTS

We want to thank Katie Boswell and Eileen Moyles for valuable input to this paper. The research developments reported in this paper are funded by National Institute on Deafness and Other Communication Disorders and National Institute on Drug Abuse.

REFERENCES

- Ajzen, I. (1985). From Intentions to Action: A Theory of Planned Behavior. *Action-Control: From Cognition to Behavior* (pp. 11-39). Springer.
- Dabbagh, N., & Bannan-Ritland, B. (2005). *Online learning: concepts, strategies, and application*. Pearson/Merrill/Prentice Hall.
- Dadayan, L., & Ferro, E. (2005). When Technology Meets the Mind: A Comparative Study of the Technology Acceptance Model. In M. A. Wimmer, R. Traunmüller, Å. Grönlund, & K. V. Andersen (Eds.), *Electronic Government* (Vol. 3591, pp. 137-144). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, 35(8), 982-1003.
- Hallfors, D., Khatapoush, S., Kadushin, C., Watson, K., & Saxe, L. (2000). A Comparison of Paper vs. Computer-Assisted Self Interview for School, Alcohol, Tobacco, and Other Drug Surveys. *Evaluation and Program Planning*, 23(2), 149-55.
- Kraemer, K. L., Dedrick, J., & Sharma, P. (2009). One laptop per child: vision vs. reality. *Communications of the ACM*, 52, 66-73. doi:10.1145/1516046.1516063
- Mantell, J. E., Vittis, A. T. D., & Auerbach, M. I. (1997). *Evaluating HIV prevention interventions*. New York: Plenum Press.
- Rogers, E. M. (1995). *Diffusion of innovations*. Simon and Schuster.
- Sedivy, V. (2000). *Is Your Program Ready to Evaluate Its Effectiveness? A Guide to Program Assessment*. Los Altos, CA: Sociometrics Corporation.
- Venkatesh, V., Morris, M. G., Gordon B. Davis, & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425-478.

APPENDIX

Table 2: Comparison of theories of change and key contextual constructs in models of technology acceptance.

Model and authors	Theory of change	Key constructs related to context
Technology, organization, and environment (TOE) framework (Tornatzky and Fleischer 1990)	At the organizational level, three aspects influence the process by which an enterprise adopts and implements a technological innovation: technological context, organizational context, and environmental context.	Environmental context is the arena in which a firm conducts its business—its industry, market structure, competitive pressures, technology support infrastructure and government regulation.
Theory of planned behavior (TPB) (Ajzen 1985, Ajzen 1991, Bajaj and Nidumolu 1998)	At the individual level, behavior is influenced solely by behavioral intention and behavioral intention in turn is influenced by attitudes toward behavior, by subjective norms and by perceived behavioral control.	Behavioral intention is influenced by attitudes, subjective norms and perceived control but these are not explicitly linked to broader environmental context.
Diffusion of innovation (DOI) (Rogers 1995)	At both individual and organizational level, innovations are communicated through certain channels over time and within a particular social system. Diffusion through an organization is related to individual (leader) characteristics, internal characteristics of organizational structure, and external characteristics of the organization.	The external characteristics element of the model refers to system openness.
Technology acceptance model (TAM) (Davis 1986, Davis 1989, Davis et al. 1989)	At the individual level, “perceived usefulness” (outcome expectation) and “perceived ease of use” (self-efficacy) influence decisions about how and when individuals will use a new technology, with intention to use serving as a mediator of actual use.	External variables may be antecedents or moderators of perceived usefulness and perceived ease of use. However, there is an assumption that when someone forms an intention to act, that they will be free to act without limitation.
Multi-contextual technology acceptance framework (Hu et al. 1999, Chau and Hu 2002)	At the individual level, technology acceptance behavior is influenced by factors pertaining to the individual context, the technological context, and the implementation context.	“Implementation context” refers to the user’s professional environment.
Unified theory of acceptance and use of technology (UTAUT) (Venkatesh et al. 2003)	At the individual level, technology use is directly determined by performance expectancy, effort expectancy, social influence, and facilitating conditions. The impact of these factors is moderated by gender, age, experience, and voluntariness of use.	Social influence refers to the degree to which an individual perceives that others believe he or she should use a particular technology. Facilitating conditions refer to the degree to which an individual believes that an organizational and technical infrastructure exists to support the use of a particular technology.
Extended technology acceptance model (Dadayan and Ferro, 2005)	At the individual level, technology acceptance is influenced by not only technological factors but also by the individual context and the implementation context.	The implementation context includes three determinants — compatibility, social influence, and organizational facilitation.
Not-for-profit internet technology adoption model (O’Hanlon and Chang 2007)	At the organizational level, technology adoption is influenced by technical capacity, compatibility (with the organization’s work practices, beliefs and values), support (of staff and donors), and organizational characteristics.	Organisational practices, beliefs and values are critical for the adoption process.

**SPECIAL SESSION ON
INTELLIGENT ALGORITHMS FOR
REAL-TIME INFORMATION
SYSTEMS SUPPORT**

APPLICATION OF MATHEMATICAL TRANSFORM IN DETECTION ALGORITHMS

Lyubka A. Doukovska

*Institute of Information and Communication Technologies, Bulgarian Academy of Sciences
Acad. G. Bonchev str., bl. 2, 1113, Sofia, Bulgaria
doukovska@iit.bas.bg*

Keywords: Hough transform, Radar detector, Target detection, Parameter estimation, Probability of detection, Constant false alarm rate.

Abstract: Recent trends in the design of highly efficient and fully automated systems for processing radar data in terms of a priori uncertainty about the targets and disturbances are causing the researchers to use the latest achievements in the design of real time computing architectures for optimum realization and high performance. The development of new algorithms that can be used to retrieve information about targets, applying a mathematical transformation on the received signals yielding estimates of the parameters of moving targets with extremely high precision in a dynamically changing radar environment is a new and very promising direction in modern information and communication technologies. This article discusses such an approach applying the Hough transform to determine the coordinates of the targets. The approach uses a finite set of preselected patterns of the target movement. The Hough transform, translates the set of measurements received in the space of patterns. Association to one or another specific pattern is done estimating the information about the coordinates extracted from the received signals. Thus the moving target parameters in the surveillance zone are uniquely determined by the parameters of the pattern.

1 INTRODUCTION

In the recent years development of modern highly effective algorithms with optimal statistic characteristics for real time radiolocation data processing is becoming a very actual scientific task. Nowadays algorithms that extract information about target's behavior through mathematical transformation of the signals reflected from a target, find ever-widening practical application. Applying signal transformation allows for higher accuracy of the estimated moving target parameters in dynamic radiolocation environment. That is why development of new robust and reliable algorithms for simultaneous trajectory and target detection applying the Hough transform is a perspective field of research, so the present paper considers this problem. The performance of original Hough detector structures maintaining constant false trajectory detection probability in intensive randomly arriving impulse interference environment is studied. Estimated are the efficiency and quality of the obtained algorithms for data streams with different distribution laws of occurrence of impulse

interference. A comparative analysis of the presented Hough detector structures is made. The practical effect of the obtained results lies in the possibility of development of radar signal processing algorithms for automated systems of air traffic control service.

On 18.Dec.1962 the American Patent Service issued a patent 3,069,654 "Method and means for recognizing complex patterns" on the name of Paul V. C. Hough (Hough, 1962). The Hough transform is a mathematical conversion, in which the task for finding specific features of the processed image consisting of points defined in the feature space is transformed to a task for finding groups of points in the parameter space. The Hough transform for straight lines detection is a sub case of the Radon transform which for the Euclidean two dimensional space and arbitrary generalized function $F(x, y)$ is as follows:

$$f(\theta, \rho) = R\{F\} = \iint_D F(x, y) \delta(\rho - x \cos \theta - y \sin \theta) dx dy \quad (1)$$

where $\delta(\cdot)$ is the Dirac delta function defining the integral on direction of a straight line defined by the

equation: $\rho = x \cos \theta - y \sin \theta$.

Initially Hough transform proposed in (Hough, 1962) presents the straight lines from a two-dimensional image in the features space (FS) with incidence m and segment c from $y = mx + c$, where x and y are the image coordinates and m and c are the coordinates in parameter space (PS).

The parameter space is sampled to a set of subspaces (accumulators). Each point of the input image is projected onto a straight line with coordinates (m, c) . Accumulators through which this line passes increase their content by one. Each of these accumulators corresponds to an area in the features space and the presence of a peak in the accumulator corresponds to a straight line or a segment of the image. Lines existing in the features space are detected according to the value accumulated into an accumulator in the parameter space. Hough transform proves to be a major tool in the analysis and algorithms for pattern recognition.

The concept of using the Hough transform for target detection improvement is introduced in (Carlson, Evans and Wilson, 1994). Regardless of the particular application of the Hough transform, different authors point three of its main properties that make it applicable to moving targets detection:

- Applicable for raster images;
- Applicable for fuzzy images processing;
- Effective when there is a lack of necessary information (measurements, observations).

2 HOUGH TRANSFORM FOR MOVING TARGET DETECTION

Let us consider the operation of surveillance radar which measures the distance, targets azimuth, elevation and Doppler velocity as a function of time. The sampling time is specified. Trajectory detection by means of Hough transformation can be made either having rotating antenna or phased array (Carlson, Evans and Wilson, 1994). In a single resolution cell in azimuth the traditional radars emit several pulses on a specified carrier frequency. The surveillance area is being consequently scanned with the radar antenna pointing in different directions. This procedure is repeated on successive periods of time equal to the sampling time. In each "azimuth-distance" resolution cell the station processes non-coherent accumulation of the emitted pulses. The target is considered detected if the pulse amplitude

in the (azimuth-distance) resolution cell exceeds a preset threshold. This approach has some difficulties to detect fast moving targets, because these objects move quickly from one to another resolution cell during one sampling period.

If received by the radar echo signals are arranged as discrete multidimensional array, i.e. discrete information card (with 5 dimensions - distance, azimuth, elevation, Doppler velocity and time), the target will appear as a curve, which intensity depends on the power level of the echo signals. If this curve can be monitored, it contains all the accumulated information about the target and complete history of its trajectory. Object with constant radial velocity appears as a straight line. The projection of this 5-dimensional information about distance and time is a convenient way to display the curve, while no interest in the other three levels for this target. The result is a so-called "range - time" (r-t) space.

Figure 1 shows (r-t) space of a target with a constant radial velocity, with given direction of the antenna beam to a specified resolution cell of the Doppler velocity. The slope of this line is determined by the radial velocity of the target. The trajectory of a stationary object will appear as a vertical line. All moving objects will have a certain angle, reaching zero for the fastest objects.

Time axis starts from zero to maximum. It is convenient to present the information about the past to a decent level, because too old information is not useful. The current information contains the disturbance, which is an internal white noise of the receiver. It has Rayleigh amplitude distribution and is summed in each cell of (r-t) space. The problem is to find a straight line on the background of the noise.

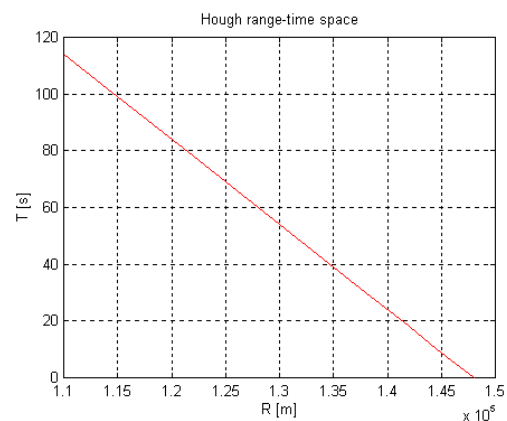


Figure 1: "Range-time" space.

The Hough transform is a method detecting curve elements, often used to detect location lines on

the noise background. Other forms of trajectories can be also detected, but so far only straight lines have been investigated. Figure 2 shows several data points that form a straight line in the “range – time” space. In polar coordinates, a straight line can be accurately defined by two parameters:

1. θ - the angle between the perpendicular from the coordinate system origin in (r-t) space to the straight line and the abscissa axis;
2. ρ - length of the perpendicular, i.e. the distance between the coordinates origin in (r-t) space and the line.

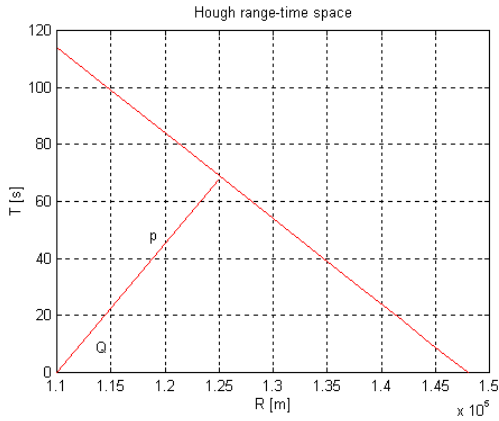


Figure 2: Relation between “range-time” (r-t) space and Hough transform.

The Hough transform translates the points from (r-t) information space to (ρ - θ) or Hough parameter space using the following expression:

$$\rho = r \cos \theta + t \sin \theta \quad (2)$$

where r and t are coordinates in the (r-t) space.

The Q grid in Hough space is formed by consequent change of θ angle from 0° to 180° and calculating the corresponding ρ . Sometimes another form of the Hough transform is used:

$$\rho = \sqrt{r^2 + t^2} \sin \left(\theta + \arctg \frac{r}{t} \right) \quad (3)$$

The transform results in a sinusoid with phase and amplitude defined by the (r-t) value of the information point. The maximal ρ value is equal to the length of the diagonal in the (r-t) space. The transformation according (2) is shown on Figure 3. Each point in (ρ - θ) space corresponds to a separate straight line in (r-t) space defined by the values of ρ and θ . Each sinusoid presents a set of possible straight lines through the point. If there are points forming a straight line in the (r-t) space this

corresponds to an intersection point of set of sinusoids in the Hough space. The (r-t) space is sampled to cells which number is equal to the number of the distance resolution elements and the sample numbers. The primary threshold is used for signal detection in each (r-t) cell. When the signal value in a specified (r-t) cell exceeds the primary threshold, its power gets added to the (ρ - θ) cell being intersected by the corresponding sinusoid in the Hough space. Thus the value of an accumulator cell in the intersection of several sinusoids will become higher.

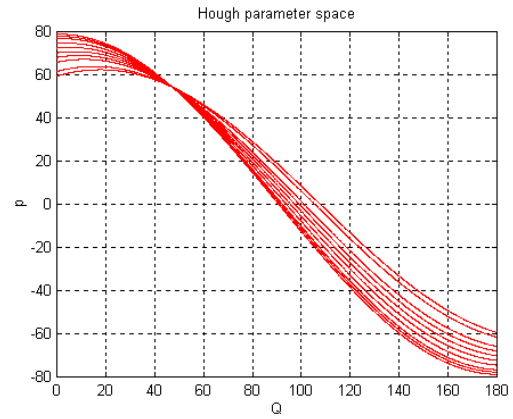


Figure 3: Hough parameter space.

The secondary threshold applied to each cell in the parameter space may declare straight line trajectory detection. This is the accumulated for several scans moving target echo signal. The ρ and θ parameters of the tracked straight line trajectory in Hough space might be transformed back to the (r-t) space indicating the current object position. Transition from (r-t) to parameter space is being made by means of a simple matrix manipulation. Matrix D contains I number of elements where the signal value exceeding the primary threshold.

$$D = \begin{bmatrix} r_1 & r_2 & \dots & r_I \\ t_1 & t_2 & \dots & t_I \end{bmatrix} \quad (4)$$

Transformation matrix H consists of sinusoids and cosinusoids from (2) defined as:

$$H = \begin{bmatrix} \sin \theta_1 & \cos \theta_1 \\ \sin \theta_2 & \cos \theta_2 \\ \dots & \dots \\ \sin \theta_{N_s} & \cos \theta_{N_s} \end{bmatrix} \quad (5)$$

where θ are discrete values of Q from 0° to 180° , obtained during the sizing of the parameter space.

The product of H and D is a matrix R of size $(N_s \times I)$, which contains the corresponding ρ values.

The indices of the ρ elements in matrix R are the indices of the points in (r-t) space where the primary threshold has been exceeded.

$$R = HD = \begin{bmatrix} \rho_1, \theta_1 & \dots & \rho_I, \theta_1 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \rho_1, \theta_{N_s} & \dots & \rho_I, \theta_{N_s} \end{bmatrix} \quad (6)$$

Each column of the R matrix contains the ρ values for one sinusoid on the parameter space. It is plain to see the more points in (r-t) space exceed the primary threshold the bigger D and R will be. The result is increasing number of calculations. The size of matrix H depends only on the parameter space sampling.

The advantage of the Hough transform application is the simultaneous target and its trajectory detection. The target is considered detected when its straight line trajectory is localized in the Hough space, i.e. ρ and θ parameters. When applying the Hough transform additional non-coherent integration of the signal obtained in several consequent scans is done. This signal integration for fast moving targets increases the detection probability compared to the conventional radars. In the recent years the Hough transform finds wide application in moving targets detection (Carlson, Evans and Wilson, 1994). This is a new and perspective direction of the Hough transform application and the results presented in (Kabakchiev, Doukovska, Garvanov, 2005, Doukovska, 2007, Doukovska 2007, Doukovska, 2008, Doukovska 2008, Doukovska, 2010) are dedicated to this problem.

3 TARGET DETECTION SIGNAL MODEL

The radar receives signal, noise and randomly arriving impulse interference. In the present paper a Swerling II target signal model is used. This model is a package of echo signals with Rayleigh distribution, reflected from a fast moving target. The noise is a stationary and internal for the receiver. The noise has normal distribution law which corresponds to a Rayleigh distribution of the envelope. Distribution function of the envelope of the signal and the noise and the corresponding

density are $F_{s1}(x)$ and $f_{s1}(x)$. If there is there is a possibility for randomly arriving impulse interference (RAII) - e_0 , (Poisson stream) the radar resolution cells are filled with signal, noise and RAII (Akimov, P., F. Evstratov, S. Zaharov, 1989). The function and density distribution of the envelope in this case are $F_{s2}(x)$ and $f_{s2}(x)$.

The overall distribution function of the above described case is obtained taking into account the probability of absence of RAII is $(1 - e_0)$ and for presence of RAII - (e_0) respectively. Now the distribution function of the envelope is (Akimov, P., F. Evstratov, S. Zaharov, 1989):

$$F_{sP}(x) = (1 - e_0)F_{s1}(x) + e_0F_{s2}(x) \quad (7)$$

The distribution density function is:

$$f_{sP}(x) = (1 - e_0)f_{s1}(x) + e_0f_{s2}(x) \quad (8)$$

Here it is assumed that the probability of RAII appearing in a resolution cell is an infinitely small quantity compared to the probability of single impulse occurrence. This is a typical feature for a Poisson stream.

When the duration of the impulse disturbance is not negligible compared to the average period of recurrence (high probability for RAII), a binomial model of the stream distribution is used. In this case the distribution function of the envelope is (Akimov, P., F. Evstratov, S. Zaharov, 1989):

$$F_{sB}(x) = (1 - e)^2 F_{s1}(x) + 2e(1 - e)F_{s2}(x) + e^2 F_{s3}(x) \quad (9)$$

For the distribution density of the envelope we have:

$$f_{sB}(x) = (1 - e)^2 f_{s1}(x) + 2e(1 - e)f_{s2}(x) + e^2 f_{s3}(x) \quad (10)$$

where $F_{s3}(x)$ and $f_{s3}(x)$ - are function and distribution density of the signal, the noise and two pulse interferences.

In the presented paper it is assumed that the distribution of the signal plus noise and the mixture of signal, noise and RAII after the quadratic detector have an exponential density (Akimov, P., F. Evstratov, S. Zaharov, 1989):

$$f_{s1}(x) = \frac{1}{\lambda_0(1+s)} \exp\left(\frac{-x}{\lambda_0(1+s)}\right) \quad (11)$$

$$f_{s2}(x) = \frac{1}{\lambda_0(1+s+r_j)} \exp\left(\frac{-x}{\lambda_0(1+s+r_j)}\right) \quad (12)$$

where s is the average value of the signal to noise ratio. In this case the probability density function for

Poisson distribution model of the RAI has the following expression (Bird J., 1982) – see (8):

$$f_{sP}(x) = \frac{(1-e_0)}{\lambda_0(1+s)} \exp\left(\frac{-x}{\lambda_0(1+s)}\right) + \frac{e_0}{\lambda_0(1+s+r_j)} \exp\left(\frac{-x}{\lambda_0(1+s+r_j)}\right) \quad (13)$$

For high probability of RAI, when the model is binomial the noise density distribution function is used as well as two pulse interferences:

$$f_{s3}(x) = \frac{1}{\lambda_0(1+s+2r_j)} \exp\left(\frac{-x}{\lambda_0(1+s+2r_j)}\right) \quad (14)$$

In this case the probability density function for binomial distribution of the RAI – see (10) is:

$$f_{sB}(x) = \frac{(1-e)^2}{\lambda_0(1+s)} \exp\left(\frac{-x}{\lambda_0(1+s)}\right) + \frac{2e(1-e)}{\lambda_0(1+s+r_j)} \exp\left(\frac{-x}{\lambda_0(1+s+r_j)}\right) + \frac{e^2}{\lambda_0(1+s+2r_j)} \exp\left(\frac{-x}{\lambda_0(1+s+2r_j)}\right) \quad (15)$$

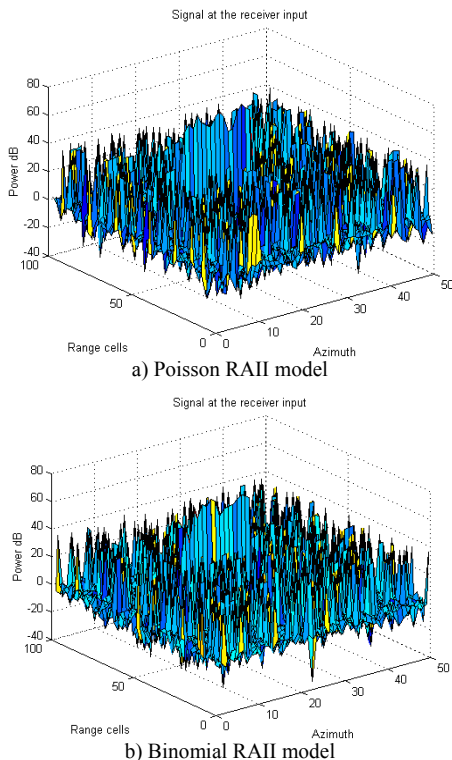


Figure 4: Radar resolution cell filled with signal ($s=70\text{dB}$), noise ($\lambda_0=1$) and impulse interference ($\text{INR}=30\text{dB}$, $e_0=0.1$).

Figure 4 shows the two streams – a) Poisson and b) Binomial. The cells that do not contain useful signal are filled with receiver internal noise and impulse interference. The cells containing signal are filled according (13) and (15). The results are obtained for: average receiver noise level $\lambda_0=1$, signal to noise ratio - $s=70\text{dB}$, impulse interference to noise ratio - $\text{INR}=30\text{dB}$ and probability for RAI 0.1 for both distributions.

4 EXPERIMENTAL RESULTS

Recently a lot of robust moving target detection algorithms for processing signals from noisy environments are developed. As a result a bank of Hough detectors making use of one and two dimensional signal processors was created (Kabakchiev, Doukovska, Garvanov, 2005, Doukovska, 2007, Doukovska 2007, Doukovska, 2008, Doukovska 2008, Doukovska, 2010). All these structures have been analytically studied and by means statistical analysis has been compared to each other as well as to results obtained by other authors (Carlson B., E. Evans, S. Wilson, 1994). On Figure 5 is presented the overall structure scheme of an adaptive to the environmental conditions detector. It consists of two main modules – signal processor and Hough detector. Maintaining constant false alarm rates at the detector's output depends on the chosen scalar factor (T_a) of the CFAR signal processor. The system input signal reflected from the target is filtered with a simple sinusoidal signal (complex signal compression), then it enters the quadratic detector where the signal matrix of the receiver is generated. This signal matrix is fed to the CFAR processor. As a result at the output a binary signal matrix is generated containing zeros and ones presenting absence or presence of a signal in a given radar resolution cell. The binary matrix is visualized on the plot extractor. Results are stored in the so called target coordinates record determined by the i -th radar observation. For several consecutive scans an interscan gathering of the plots of the target is done. Then the (r-t) space is formed and using the already processed data the trajectory is being determined. The Hough transform is applied over the points from the (r-t) space in order to transfer them to the Hough space. As a result there is a bunch of sinusoids which intersection point accumulates the energy reflected from the target. Comparing the value accumulated in this point (sum of zeros and ones) to a preliminary chosen threshold is the way to detect a target if the radar range resolu-

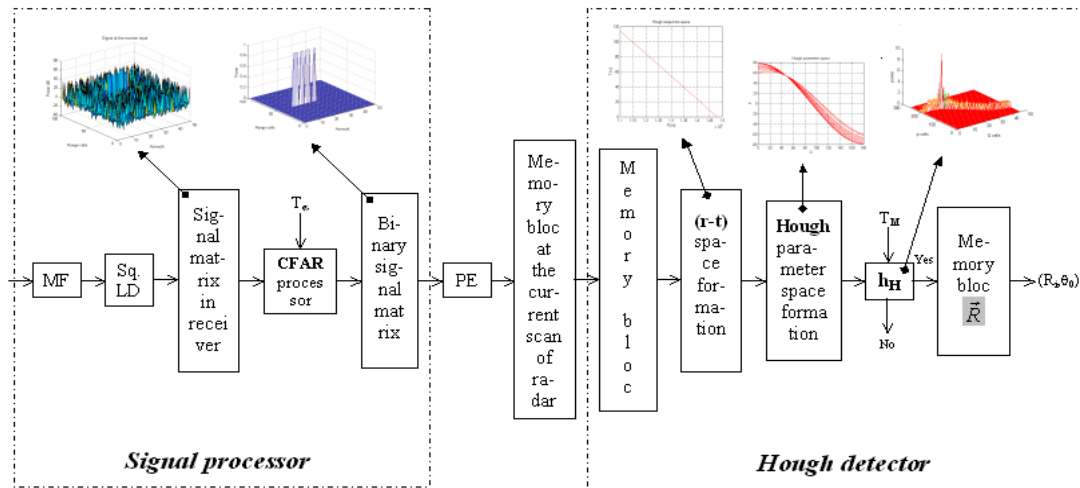


Figure 5: Generalized Hough detector structure.

tion cell. If the result is positive a reverse Hough transform is applied in order to determine the distance to the target for a given azimuth.

Presented paper considers the results obtained from the analysis of different Hough detectors with one and two dimensional signal processors maintaining constant false alarm rates. To make the results applicable they were compared in equal conditions using equal criteria. The efficiency of the Hough transform application was estimated by the profit, gained during the detection process, expressed by the signal to noise ratio as per the criterion presented in (Rohling H., 1983).

Choosing the appropriate threshold constants assures good detection results even for low values of the SNR (Doukovska, 2010). Table 1 presents the obtained threshold constants in equal experimental conditions for the different detection structures and different values of the binary rule in the Hough parameter space.

Table 1: Threshold constants for different Hough detectors.

Hough detectors	$T_M=2/20$	$T_M=T_{Mopt}/20$
CA Hough CFAR	672	1.186
EXC Hough CFAR	21880	3.225
Hough CFAR BI	0.000494	0.0000858
EXC Hough CFAR BI	1.1285	0.3161
API Hough CFAR	7.5	1.535

For comparison are shown the achieved results for the detection probability of different Hough detector structures, calculated for non optimal and optimal values of binary rule in Hough parameter space - $T_M = T_{Mopt}/20$, for following environment parameter values - average power of the receiver noise $\lambda_0=1$, average interference-to-noise ratio (INR)

$r_j=30\text{dB}$, probability for the appearance of impulse interference with average length $e_0=0.1$, $N=16$, $L=16$ and for probability of false alarm $P_{FA}=10^{-4}$.

The results presented in this paper are obtained after statistical analysis of the Hough detectors detection probability in intensive noise environment with very high probability for randomly arriving impulse interference. Different Hough detector structures with one and two dimensional CFAR signal processors are studied.

All analytical conclusions necessary to convey the experiments are considered in details in (Kabakchiev, Doukovska, Garvanov, 2005, Doukovska, 2007, Doukovska 2007, Doukovska, 2008, Doukovska 2008, Doukovska, 2010).

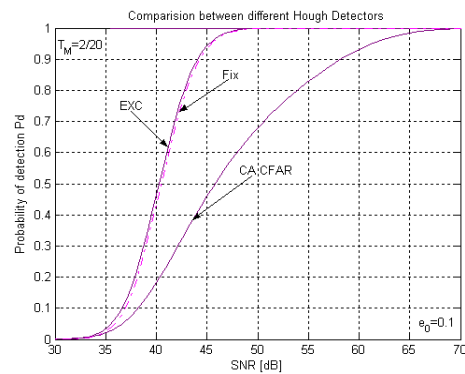


Figure 6: Probability characteristics of a Hough detector with one dimensional signal processors - cell averaging CFAR (CA CFAR), excision CFAR processor (EXC CFAR) and with fixed threshold, for RAII probability - e_0 .

It was shown that application of a binary CFAR processor significantly increases the detection quality (about 30dB) compared to the fixed

threshold algorithm (Doukovska, 2007). Analyzed is a Hough detector with a more efficient structure of the two dimensional CFAR processor with excision censoring procedure in the reference window (EXC CFAR BI). The hypothesis that censoring techniques increase the detection efficiency with about 5dB was confirmed (Doukovska, 2008).

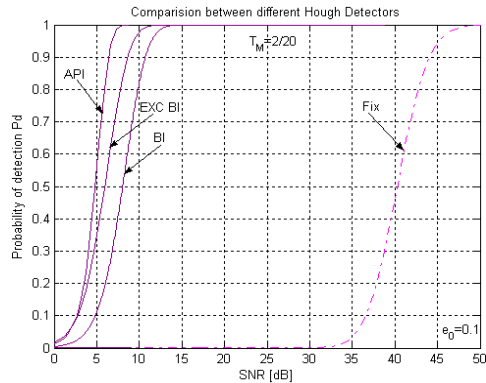


Figure 7: Probability characteristics of a Hough detector with two dimensional signal processors - adaptive CFAR processor (API CFAR), excision binary CFAR processor (EXC CFAR BI), binary CFAR processor (BI CFAR) and with fixed threshold, for RAII probability - e_0 .

The most effective for noisy environment with high probability for randomly arriving impulse interference is the Hough detector with adaptive non coherent CFAR signal processor (API CFAR). This structure is by 37dB more effective than the one with fixed threshold Hough detector (Doukovska, 2007).

5 CONCLUSIONS

In conventional signal detection approach the process of target detection is separate from its trajectory detection. Unlike this wide spread technique Hough transform application allows for simultaneous target and trajectory detection. To detect a trajectory data from several consecutive radar scans is processed.

The presented paper considers the results obtained by the proposed adaptive threshold determination procedure and analysis of different Hough detector structures in intensive RAII environment. The need of an adequate threshold analysis procedure allowing better detection results for low values of the SNR is considered.

The obtained results are applicable for wide range of tasks like synthesis of radiolocation detectors, communication systems, medicine and

other systems making use of infrared, ultrasonic and other sensor types.

ACKNOWLEDGEMENTS

The investigations in this work are within the frame of Project "Formation of Highly Qualified Young Researchers in Information Technologies for Optimization, Pattern Recognition and Decision Support Systems", Contract with the Ministry of Education, Youth and Science: BG051PO001-3.3.04/40/28.08.2009.

REFERENCES

- Hough P. V. C., Method and means for recognizing complex patterns, US Patent - 3,069,654, 18.XII.1962.
- Carlson B., E. Evans, S. Wilson, Search Radar Detection and Track with the Hough Transform, Parts I, II, III, *IEEE Trans.*, Vol. AES-30, pp. 102-124, 1994.
- Akimov P., F. Evstratov, S. Zaharov, Radio Signal Detection, *Radio and Communication*, Moscow, pp. 195-203, 1989 (in Russian).
- Bird J., Calculating detection probabilities for systems employing non coherent integration, *IEEE Trans.*, vol. AES-18, № 4, July, pp. 401-409, 1982.
- Kabakchiev Chr., L. Doukovska, I. Garvanov, Hough Radar Detectors in Conditions of Intensive Pulse Jamming, *Sensors & Transducers Magazine (S&T e-Digest)*, Special Issue (ISSN 1726-5479), "Multisensor Data and Information Processing", №8, pp. 381-389, 2005.
- Doukovska L., Hough Detector with Binary Integration Signal Processor, *Comptes rendus de l'Academie bulgare des Sciences*, vol. 60, №5, pp. 525-533, 2007.
- Doukovska L., Moving Target Hough Detector in Randomly Arriving Impulse Interference, *Cybernetics and Information Technologies*, vol. 7, №3, pp. 55-72, 2007.
- Doukovska L., Hough Target Detectors with Small Values of SNR, NATO Advanced Study Institute "Unexploded Ordnance Detection and Mitigation", Poster presentation, Il Ciocco, Italy, 2008, <http://www.prometheus-inc.com/asi/uxo2008/>
- Doukovska L., Hough Moving Target Detector with Fixed Threshold, *Comptes rendus de l'Academie bulgare des Sciences*, vol. 61, №11, ISSN: 0861-1459, pp. 1459-1466, 2008.
- Doukovska L., Adaptive Hough Detector Threshold Analysis in Presence of Randomly Arriving Impulse Interference, *Proc. of the International Radar Symposium – IRS'10*, Vilnius, Lithuania, pp. 142-147, 2010.
- Rohling H., Radar CFAR thresholding in clutter and multiple target situation, *IEEE Trans.*, vol. AES-19, 4, pp. 608-621, 1983.

RESOURCE MANAGEMENT FOR ONE CLASS OF WORKFLOW PROCESSES

A Stochastic Petri Net Based Approach

Oleg Gorbunov and Valery Sokolov

Department of Computer Science, Yaroslavl State University, Yaroslavl, Russia
{falcon, sokolov}@uniyar.ac.ru

Keywords: Workflow processes, Petri nets, Stochastic Petri nets, Stochastic Resource-Constrained Petri nets, Management of resources.

Abstract: In the paper the new approach for modeling workflow processes have been proposed. This approach is based on the special class of stochastic Petri nets and allows to model resources that is needed for tasks execution. Such models are well suited both for qualitative and quantitative analysis of workflow processes. In the paper the function reflecting the cost of waiting the task execution due to the lack of resources have been introduced. The problem of minimizing of this function have been stated. The decision approach for this problem have been introduced. This approach manages resources by means of priorities. In general, optimal priorities may be found during simulation.

1 INTRODUCTION

One of the most popular and relatively recent methodology of enterprise management is the management of workflow processes. This methodology deals with the models of workflow processes. Many languages and formalisms are proposed for modeling them. One of such formalisms is Petri nets which have been proven to be a successful formalism for this reason (van der Aalst, 1998; van der Aalst and van Hee, 2002).

Many special classes of Petri nets were proposed for modeling workflow processes. In general, workflow processes are modeled by *WF*-nets (van der Aalst, 1998), i.e. Petri nets with one initial and one final places and every place or transition being on a path from the initial place to the final one. The execution of a case is represented as a firing sequence that starts from the initial marking consisting of a single token on the initial place. The token on the final place with no tokens left on the other places indicates the proper termination of the case execution. A model is called sound iff every reachable marking can terminate properly. Such models reflect the partial ordering of activities in the process and abstract from resources, e.g. machines or personnel, that actually execute tasks and any quantitative measures of its ex-

ecution.

In (K.M. van Hee, 2005) a notion of *RCWF*-nets and a respective soundness property was introduced. Such models represent *WF*-nets that take resources into account.

The concept of time was intentionally avoided in the classical Petri net as timing constraints may prevent certain transitions from firing. Many different ways of incorporating time in Petri nets have been proposed. Some timed Petri net models use deterministic delays (Ramchandani, 1973; Sifakis, 1977). The others use interval timing (Merlin, 1974; van der Aalst, 1993) or stochastic delays (G. Florin, 1980; M.A. Marsan, 1984; M.A. Marsan, 1985).

In real systems execution of tasks in workflow processes depends on the various conditions, such as availability of free resources and all needed information and so on. Hence for simulation and quantitative analysis of workflow systems these external conditions are essential.

In (Reijers, 2003) a model that takes into account both resources and timing is proposed. In that book different heuristic rules for allocation of additional resources that minimize mean throughput time of the process are also discussed.

2 STOCHASTIC PETRI NETS WITH PRIORITIES

We propose a class of stochastic Petri nets with priorities SPN^e . These nets combine properties of $GSPN$ -nets proposed in (M.A. Marsan, 1985) and Interval Timed Petri nets (van der Aalst, 1993). SPN^e -nets are based on Petri nets with priorities, so begin with these formalism.

2.1 Petri Nets with Priorities

Definition 2.1 (Petri nets with priorities) *Petri net with priorities (PN_{pr} -net) is a tuple (P, T, R, Pr) , where (P, T, R) – Petri net; $Pr \in T \rightarrow \mathbb{N} \cup \{0\}$ – priority function, that assign for each transition t natural number $Pr(t)$, priority of the transition.*

The definition of marking is the same as for ordinary Petri nets but firing rule differs. The transition that is active in Petri net P, T, R is potentially active in Petri net P, T, R, Pr with priorities.

Denote $M_{pr}()$ a function $M_{pr}() \in 2^T \rightarrow 2^T$, that for any set of transitions from T returns the subset of transitions with maximal priority:

$$\forall J \subseteq T \quad M_{pr}(J) = \{t \in J \mid \exists t' \in J : Pr(t') > Pr(t)\}.$$

Potentially active transition t_j of the PN_{pr} -net $N = (P, T, R, Pr)$ is active in marking m , if there is no another potentially active transition $t_i \in T$: $Pr(t_i) > Pr(t_j)$. So, the existence of priorities restricts the number of active transitions in comparison with the same Petri net without priorities. Denote $At(m)$ the set of active transitions of the PN_{pr} -net N in marking m . From the definition of $At(m)$ follows that two transitions t_i and t_j , $t_i \neq t_j$ are active in marking m only if they have the same priority: $Pr(t_i) = Pr(t_j)$. Active transition may fire. Firing rule is the same as in the ordinary Petri net.

It is well known that the expressive power of Petri nets with priorities is greater than of the ordinary Petri nets. So, in general, if we model workflow processes by means of WF -nets with priorities, soundness property would be undecidable.

Let us consider free-choice Petri nets with priorities. Remind that by constructing WF -nets with task refinement approach using basic structures of choice, sequential and parallel execution, a free-choice WF -net will be obtained (van der Aalst, 2000). Relation SC (structural conflict relation on the set of T) for such nets is reflexive, transitive and symmetric. So, we may conclude that it is an equivalence relation and the set T may be divided into the disjoint subsets $SC_1, SC_2 \dots SC_k$: $SC_1 \cup SC_2 \cup \dots \cup SC_k = T$.

Obviously, for free-choice PN_{pr} -net N all transitions in any subset SC_i potentially active or not potentially active at the same time. It is easy to prove that for such nets if there exist at least two transitions $t_i, t_j \in T$: $t_i \neq t_j, t_i SC t_j, Pr(t_i) \neq Pr(t_j)$ then there exist dead transitions that will never be active.

Obviously, for free-choice Petri net $N = (P, T, R, Pr)$, with priorities such that $\forall t_i, t_j \in T$: $t_i SC t_j \Rightarrow Pr(t_i) = Pr(t_j)$, for any marking m $At(m)$ is empty or consists of subsets $SC_{a1}, SC_{a2}, \dots, SC_{ap}$ with the same priority.

If we constrain the structure of the net by the free-choice property and require certain rules on priorities assignment, the soundness property will be decidable.

The following theorem may be proved (Gorbunov, 2006).

Theorem 2.1 *If a free-choice WF -net $N = (P, T, R)$ is sound, the WF -net N' with priorities: $N' = (P', T', R', Pr')$: $P' = P, T' = T, R' = R, \forall t'_i, t'_j \in T'$: $t'_i SC' t'_j \Rightarrow Pr(t'_i) = Pr(t'_j)$ with the same initial marking $m'_i = m_i$ is sound.*

2.2 Stochastic Petri Nets

Tokens have time stamps that denotes time when token will be available for transition execution. While executing, transitions assign time stamps to the produced tokens.

Definition 2.2 (Stochastic Petri net with priorities) *A stochastic Petri net with priorities $N \in SPN^e$ is a tuple (P, T, R, W, Pr) :*

- (P, T, R, Pr) – free-choice PN_{pr} net;
- $W \in T \rightarrow \mathbb{R}^+$.

There exist two types of transitions: timed and immediate. A transition t is a timed transition iff $Pr(t) = 0$ and is an immediate transition otherwise.

Firing of immediate transitions takes no time. Delays of timed transitions are defined by the negative exponential probability function. For timed transitions W defines the rate of executions, i.e. the parameter of negative exponential probability function of the delays: $\forall t \in T, Pr(t) = 0$: $P\{\text{delay of execution of } t \leq x\} = 1 - e^{-W(t)x}$.

For immediate transitions W is used for resolving conflicts between transitions.

For SPN^e -net $N = (P, T, R, W, Pr)$ timed state space $S \subseteq P \rightarrow (\mathbb{R} \rightarrow \mathbb{N})$ is defined. For timed state $s \in S$ for any $p \in P, s(p)$ is multiset on \mathbb{R} . Timed state defines for any place p the number of tokens and their time stamps.

If in a timed state s we abstract from token time stamps, we obtain marking s_m as in ordinary Petri

nets. Function $U_t()$ makes an appropriate transformation:

Definition 2.3 (Function $U_t()$)

$$\forall p \in P : U_t(s, p) = |s(p)|.$$

So, $s_m = U_t(s)$.

For SPN^e -nets some initial timed state $m_0 \in S$ is fixed.

Function $first()$ for timed state $s \in S$ and position $p \in P$ returns the minimal time stamp of tokens in p :

$$first(s(p)) = \min(\{k \in \mathbb{R} | s(p)(k) > 0\}).$$

Definition 2.4 (Function $ttime(s, t)$)

$$\forall s \in S, t \in T \quad ttime(s, t) = \begin{cases} \max_{p \in \bullet t} \{first(s, p)\}, \\ \text{if } t \in E_t(U_t(s)); \\ \text{not defined otherwise.} \end{cases}$$

Function $ftime()$ for timed state $s \in S$ returns first moment of time, when some transition in $E_t(U_t(s))$ can fire:

$$ftime(s) = \min_{t \in E_t(U_t(s))} ttime(s, t).$$

Here E_t is function that for any marking of Petri net returns the set of active transitions.

Definition 2.5 (Function $fire()$) Function $fire()$ returns the set of transitions, that can fire in timed state $s \in S$:

$$fire(s) = M_{pr}(\{t \in E_t(U_t(s)) | ttime(s, t) = ftime(s)\}).$$

If $t \in fire(s)$, it can fire at the timed state s .

If $fire(s)$ consists of some transitions, they have the same priority value. It can be shown, that the set of $fire(s)$ is empty or consists of sets $SC_{a1}, SC_{a2}, \dots, SC_{ap}$ with the same priority value. Suppose $fire(s) = SC_{a1}, SC_{a2}, \dots, SC_{ap}$.

The probability of firing the transition $t_b \in fire(s)$ is defined by its relational weight among other transitions from $fire(s)$:

$$P\{t_b \text{ will fire in } s\} = \begin{cases} \frac{W(t_b)}{\sum_{u \in fire(s)} (W(u))}, \\ \text{if } t_b \in fire(s) \\ 0, \text{ otherwise.} \end{cases} \quad (1)$$

When a transition t is firing, tokens with the smallest time stamps are removed from its input places and tokens with time stamps equal to the moment of firing increased by the firing delay d are added to its output places. Firing delay d is sampled from the probability function associated with the delay of transition. The new timed state obtained from timed state $s \in S$ by

firing the transition $t \in T$ with delay $d \in \mathbb{R}$, is defined by the function g :

$$g(s, t, d)(p) = \begin{cases} s(p), \text{ if } p \notin \bullet t \text{ and } p \notin t^\bullet \\ s(p) - [first(s(p))], \\ \quad \text{if } p \in \bullet t \text{ and } p \notin t^\bullet \\ s(p) + [ftime(s) + d], \\ \quad \text{if } p \notin \bullet t \text{ and } p \in t^\bullet \\ s(p) - [first(s(p))] + \\ \quad [ftime(s) + d], \\ \quad \text{if } p \in \bullet t \text{ and } p \in t^\bullet. \end{cases}$$

2.3 Stochastic Workflow Nets

(SWF^e -nets)

A SWF^e -net N is a tuple (P, T, R, W, Pr) :

- (P, T, R) is a WF -net;
- (P, T, R, W, Pr) is a SPN^e -net.

For the SWF^e -net $N = (P, T, R, W, Pr)$ the initial timed state m_0 is defined as follows:

$$\forall p \in P, m_0(p) = \begin{cases} [0], \text{ if } p = i \\ 0, \text{ otherwise.} \end{cases}$$

At the initial timed state m_0 the net contains one token in the place i with the time stamp equal to 0, other places don't contain tokens.

2.4 Stochastic Resource-Constrained Workflow Nets ($SRCWF^e$ -nets)

In this section a stochastic extension for $RCWF$ -nets (K.M. van Hee, 2005) is proposed. Stochastic $RCWF$ -net ($SRCWF^e$ -net) N is a tuple $(P_p \cup P_r, T, R_p \cup R_r, W, Pr)$:

- $(P_p \cup P_r, T, R_p, W, Pr)$ – SWF^e -net;
- $(P_p \cup P_r, T, R_p \cup R_r)$ – $RCWF$ -net with places i and f as source and sink places.

Denote $P = P_r \cup P_p$.

Definition 2.6 (Initial timed state) $SRCWF^e = (P_p \cup P_r, T, R_p \cup R_r, W, Pr)$ with the timed state space S , the initial timed state $m_0 \in S$ is defined as follows:

$$\text{for } p \in P, m_0(p) = \begin{cases} [0] \text{ if } p = i, \\ l[0] \text{ if } p \in P_r, \\ \quad \text{where } l \in \mathbb{N}, l > 0 \\ 0 \text{ otherwise.} \end{cases}$$

In the initial timed state m_0 there is one token in place i with time stamp 0 and some tokens in places from P_r . Places from P_r (resource places) contain multisets of tokens with time stamp 0. Every place from P_r denotes a resource class. The quantity of tokens in the resource position denotes the quantity of resources of that class.

There are two possible classes of methods for quantitative analysis of $SRCWF^e$ -nets: simulation and analytical methods. Note, that analytical methods are applicable only for restricted subclasses of $SRCWF^e$ -nets with additional constraints on structure, initial timed state and so on. In general, at present, analytical methods are inapplicable for quantitative analysis (Gorbunov, 2005; Gorbunov, 2006).

3 RESOURCE MANAGEMENT FOR ONE CLASS OF $SRCWF^E$ -NETS

3.1 $SRCWF_s^{e1}$ -nets

Denote the set of transitions of the $SRCWF^e$ -net N which have input places from P_r as T_r : $T_r = \{t \in T | \exists p \in P_r : (p, t) \in R_r\}$. Suppose, that all places from P_r have some names p_{r1}, \dots, p_{rs} . Denote subsets of $t \in T_r$, such that $(p_{ri}, t) \in R_r$, as T_{ri} . Suppose that all transitions from T_{ri} are denoted as $t_{(ri)(1)}, t_{(ri)(2)}, \dots, t_{(ri)(n_i)}$, where $|T_{ri}| = n_i$.

We will use a special class of $SRCWF^e$ -nets (denote it $SRCWF_s^{e1}$ -nets) with some restrictions and modifications.

The net $N = (P_p \cup P_r \cup p_g, T \cup \{t_i, t_f\}, R_p \cup R_r \cup \{(t_i, i), (t_i, p_g), (p_g, t_i), (f, t_f)\}, W, Pr)$ is a $SRCWF_s^{e1}$ -net, iff $(P_p \cup P_r, T, R_p \cup R_r, W, Pr) - SRCWF^e$ -net with some restrictions:

1. the net (P_p, T, R_p) is a state machine non-cyclic net;
2. in the net $(P_p, T, R_p) \exists t_l, t_m : t_l \in T_r, t_m \in T, t_l \neq t_m, t_l SC t_m$;
3. in the net $(P_p, T, R_p, Pr) : \forall t_i, t_j \in T : t_i SC t_j \Rightarrow Pr(t_i) = Pr(t_j)$;
4. for any $t \in T_r : \exists! p_{ri} \in P_r : (p_{ri}, t), (t, p_{ri}) \in R_r$;
5. for any $t \in T_r : (Pr(t) = 0)$;

$Pr(t_i) = 0, Pr(t_f) = 1$.

In other words, the transition t_i is a timed transition and t_f is an immediate transition.

The initial marking $m_0 \in S$ is defined as follows:

$$\forall p \in P, m_0(p) = \begin{cases} 1[0], & \text{if } p = i; \\ 1[0], & \text{if } p = p_g; \\ l[0], & \text{if } p \in P_r, \\ & \text{where } l \in \mathbb{N}, l > 0; \\ 0 & \text{otherwise} \end{cases}$$

Note, that the transition t_i will generate the poison stream of tokens with the rate $W(t_i)$. After each

firing of t_i place p_g will contain one token with time stamp increased by the sampled delay of t_i .

The special transition t_f consumes tokens from the place f . The value of $W(t_f)$ is of no importance and, for certainty, let be 1.

Due to the constraints, if any resource position is connected by the output (input) arc with some transition, such position must be connected with it by the input (output) arc. At the same time this transition cannot be connected with any other resource positions. In other words, a resource becomes free after fulfilling the task (firing the transition). Obviously, places from P_r are bounded, moreover, at any reachable marking the number of tokens in any place from P_r is the same as in the initial marking.

All resource places are connected with timed transitions only. That is, if some resource is needed for some task, this task must consume time. At the same time, in the model there may be timed transitions which are not connected with resource places. Such transitions model some time delays which don't depend on resources and are defined by external factors.

Definition 3.1 (Function $ttime_p(s, t)$)

$$\forall s \in S, t \in T \quad ttime_p(s, t) = \begin{cases} \max_{p \in {}^*t \setminus p_p} \{first(s, p)\}, \\ \text{if } t \in E_t(U_t(s)); \\ \text{don't defined} \\ \text{otherwise} . \end{cases}$$

The function $ttime_p(s, t)$ for timed state s and potentially active transition $t \in T, t \in E_t(U_t(s))$ result in the moment of time when the transition t could fire, if we abstract from places in P_r .

Definition 3.2 (Waiting time) For $SRCWF_s^{e1}$ -net $(P_p \cup P_r \cup p_g, T \cup \{t_i, t_f\}, R_p \cup R_r \cup \{(t_i, i), (t_i, p_g), (p_g, t_i), (f, t_f)\}, W, Pr)$, that induces the stochastic process $\pi = \{(X_n, Y_n) | n = 0, 1, 2, \dots\}$, where X_n is a timed state after n firings, Y_n is a transition that will fire at the state X_n , define the stochastic variable $WT_t(j)$, waiting time of transition $t \in T$ due to the lack of resources:

$$WT_t(j) = \begin{cases} ttime(X_j, Y_j) - ttime_p(X_j, Y_j), \\ \text{if } t = Y_j, \\ \text{is not defined otherwise.} \end{cases}$$

$WT_t(j)$ equals the waiting time of firing of t due to the lack of resources, when t fires in the state X_j and is not defined otherwise. Denote by $E(WT_t)$ the mathematical expectation of $WT_t(j)$. Of course, there must be some restrictions on functions W to obtain finite values of $E(WT_t)$.

3.2 Soundness of the Nets Underlying $SRCWF_s^{e1}$ -nets

Consider $SRCWF_s^{e1}$ -net $N = (P_p \cup P_r \cup p_g, T \cup \{t_i, t_f\}, R_p \cup R_r \cup \{(t_i, i), (t_i, p_g), (p_g, t_i), (f, t_f)\}, W, Pr)$.

WF -net (P_p, T, R_p) is a state machine net. Hence, WF -net (P_p, T, R_p) is sound.

Due to the structural constraints of $SRCWF_s^{e1}$ -nets and theorem (2.1), Petri net (P_p, T, R_p, Pr) with priorities is sound.

It is obvious that $RCWF$ -net with priorities $P_p \cup P_r, T, R_p \cup R_r, Pr$ is sound due to the structural constraints.

3.3 Problem Statement

Let N be a $SRCWF_s^{e1}$ -net. Denote $PF \in T_r \rightarrow \mathbb{R}$ a function that assigns for each timed transition from T_r some penalty for waiting per unit of time due to the lack of resources. $PF(t)$ may reflect the cost of waiting or some measure of client dissatisfaction.

The problem is to minimize the function F :

$$F = \sum_{t \in T_r} E(WT_t)PF(t) \rightarrow \min. \quad (2)$$

Another important characteristic of workflow processes is throughput time (Reijers, 2003). Note that it is possible to vary function F without changing throughput time of the process.

3.4 Decision Approach

Let us introduce some transformation rule with $SRCWF_s^{e1}$ -net N . Denote γ the maximum value of the function Pr in N : $\gamma = M_{Pr}(T)$. For every time transition $t_{(ri)(j)} \in T_{ri}$ add (in the set T) new immediate transition $t'_{(ri)(j)}$: $W(t'_{(ri)(j)}) = W(t_{(ri)(j)})$, $PF(t'_{(ri)(j)}) = PF(t_{(ri)(j)})$, $Pr(t'_{(ri)(j)}) \in \{1 + \gamma, 2 + \gamma, \dots, n_i + \gamma\}$ and the new place $p'_{(ri)(j)}$ (in the set P_p), $\bullet t'_{(ri)(j)} = \bullet t_{(ri)(j)}$, $t'_{(ri)(j)} = \{p'_{(ri)(j)}\}$, $\bullet t_{(ri)(j)} = \{p'_{(ri)(j)}\}$.

Denote the modified $SRCWF_s^{e1}$ -net N as N_f . Note, that N_f is not a $SRCWF_s^{e1}$ -net. The set T_r of the net N_f consists of immediate transitions. This transformation rule preserves the soundness of $RCWF$ -net with priorities that underlies N_f .

Now we obtain the possibility to change the value of function F by changing the priorities of transitions from sets T_{ri} .

It may be shown that to obtain the same value of F in the net N_f , the priorities of transitions within each set T_{ri} must be the same (for example, $1 + \gamma$).

In general, simulation may be used to obtain some optimal result. In brute force approach, $|T_{r1}|^{|T_{r1}|} * |T_{r2}|^{|T_{r2}|} * \dots$ simulations may be carried out to obtain some optimal result. Moreover, if some transitions from some T_{ri} have the same priority, the value of function W may be changed to obtain the optimal result.

3.5 Example

Let us introduce an example that illustrates the approaches discussed above. In Figure 1 some $SRCWF_s^{e1}$ -net N is illustrated.

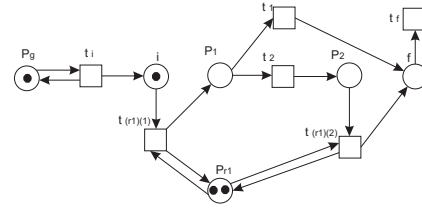


Figure 1: $SRCWF_s^{e1}$ -net N .

The characteristics of the net N are specified in Table 1.

Table 1: Characteristics of N .

Transition	W	Pr	PF
t_i	10	0	
t_1	0.2	1	
t_2	0.8	1	
$t_{(r1)(1)}$	20	0	50
$t_{(r1)(2)}$	30	0	100
t_f	1	1	

By applying transformation rules from 3.4 the net N_f illustrated in Figure 2 is produced. Priorities of transitions within set T_{r1} are selected arbitrarily.

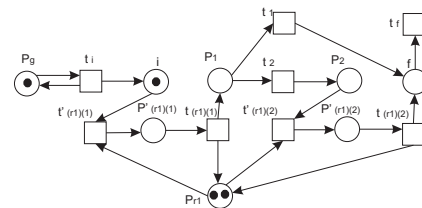


Figure 2: The net N_f obtained from N .

The characteristics of the net N_f are specified in Table 2.

Table 2: Characteristics of N_f .

Transition	W	Pr	PF
t_i	10	0	
t_1	0.2	1	
t_2	0.8	1	
$t_{(r1)(1)}$	20	0	
$t'_{(r1)(1)}$	20	2	50
$t_{(r1)(2)}$	30	0	
$t'_{(r1)(2)}$	30	3	100
t_f	1	1	

4 CONCLUSIONS

This paper opens many ways for further work. One way is to develop heuristic rules for assigning priorities without simulation for some classes of nets and functions F (may be not linear). The other way is to weaken the constraints of $SRCWF_s^{e1}$ -net such as the structural constraint of the state machine. Some heuristic rules may also be developed to obtain values of W for some T_{ri} of the net N_f in the case of deriving transitions with the same priorities.

ACKNOWLEDGEMENTS

Authors appreciate prof. E.A. Timofeev for the useful advices and discussion in the paper related topics.

REFERENCES

- G. Florin, S. N. (1980). Evaluation based upon stochastic petri nets of the maximum throughput of a full duplex protocol. In C. Girault, W. R., editor, *Applications and Theory of Petri Nets*, pages 280–288. Springer-Verlag.
- Gorbunov, O. (2005). Optimisation of workflow processes execution. *Modeling and Analysis of Informational Systems*, 12(1):45–51.
- Gorbunov, O. (2006). *Modeling and Optimisation of Workflow Processes*. Phd thesis, Yaroslavl State University, Yaroslavl.
- K.M. van Hee, A. Serebrenik, N. S. M. V. (2005). Soundness of resource-constrained workflow nets. In Giardo, G. and Darondeau, P., editors, *Applications and Theory of Petri Nets 2005*, volume 3536, pages 250–267. Lecture Notes in Computer Science, Springer-Verlag.
- M.A. Marsan, G. B. e. a. (1985). On petri nets with stochastic timing. In *International Workshop on Timed Petri Nets*, pages 80–87. IEEE Computer Society Press.

- M.A. Marsan, G. Balbo, G. C. (1984). A class of generalised stochastic petri nets for the performance evaluation of multiprocessor systems. *ACM Transactions on Computer Systems*, 2(2):93–122.
- Merlin, P. (1974). *A Study of the Recoverability of Computer Systems*. Phd thesis, University of California, Irvine.
- Ramchandani, C. (1973). *Performance Evaluation of Asynchronous Concurrent Systems by Timed Petri Nets*. Phd thesis, Massachusetts Institute of Technology, Cambridge.
- Reijers, H. (2003). *Design and Control of Workflow Processes*. Springer-Verlag.
- Sifakis, J. (1977). Use of petri nets for performance evaluation. In H. Beilner, E. G., editor, *Third International Symposium IFIP. Measuring, modeling and evaluating computer systems*, pages 75–93, Amsterdam. Elsevier Science Publishers.
- van der Aalst, W. (1993). Interval timed coloured petri nets and their analysis. In Marsan, M., editor, *Application and Theory of Petri Nets*, volume 691, pages 453–472. Lecture Notes in Computer Science, Springer-Verlag.
- van der Aalst, W. (1998). The application of petri nets to workflow management. *The Journal of Circuits, Systems and Computers*, 8(1):21–66.
- van der Aalst, W. (2000). Workflow verification: Finding control-flow errors using petri net-based techniques. *Business Process Management. Lecture Notes in Computer Science*, 1806:161–183.
- van der Aalst, W. and van Hee, K. (2002). *Workflow Management: Models, Methods and Systems*. MIT Press, London.

PETRI NET BASICS

Definition 4.1 (Petri net) *Petri net N is a tuple (P, T, R) , where*

- P – finite set of places;
- T – finite set of transitions, $(P \cap T = \emptyset)$;
- R – flow relation, $R \subseteq (T \times P) \cup (P \times T)$.

We use $\bullet t$ to denote the set of input places of a transition t : $p \in \bullet t$ iff $p \in R(p, t)$. t^\bullet have the similar meaning: it is the set of output places of a transition t : $p \in t^\bullet$ iff $p \in R(t, p)$.

Definition 4.2 (Petri net marking) *The marking (state) m of Petri net N is a mapping $m: P \rightarrow \mathbb{N}$. A marking is represented by the vector $(M(p_1) \dots M(p_n))$, where p_1, \dots, p_n is an arbitrary fixed enumeration of P .*

Definition 4.3 (Firing rule) *A marking m of a Petri net (P, T, R) enables a transition $t \in T$ if it marks every place in $\bullet t$. If t is enabled at m , then it can fire, and*

its firing leads to the successor marking m' (written $m \xrightarrow{t} m'$) which is defined for every place $p \in P$:

$$m' = \begin{cases} m(p) & \text{if } p \notin t^\bullet \text{ and } p \notin \bullet t, \text{ or } p \in t^\bullet \text{ and } p \in \bullet t \\ m(p) + 1 & \text{if } p \in t^\bullet \text{ and } p \notin \bullet t \\ m(p) - 1 & \text{if } p \notin t^\bullet \text{ and } p \in \bullet t \end{cases}$$

Definition 4.4 (Free choice Petri net) Petri net N is called a free choice Petri net, if for any transitions $t_1, t_2 \in T$: if $\bullet t_1 \cap \bullet t_2 \neq \emptyset$, then $t_1^\bullet = t_2^\bullet$.

Definition 4.5 (Structural conflict of transitions) A structural conflict of transitions is a relation SC on the set of transitions T : $\forall t_i, t_j \in T : t_i SC t_j$ iff $\bullet t_i \cap \bullet t_j \neq \emptyset$.

**SPECIAL SESSION ON
BUSINESS MODELS AND
INFORMATION SYSTEMS FOR
SUSTAINABLE DEVELOPMENT**

BUSINESS MODELS AND INFORMATION SYSTEMS FOR SUSTAINABLE DEVELOPMENT

Marten van Sinderen

*University of Twente, Drienerlolaan 5, Enschede, The Netherlands
m.j.vansinderen@utwente.nl*

Keywords: Business models, Information systems, Goal modelling, Value modelling, Model-driven development, Sustainable development.

Abstract: Businesses are expected to explore market opportunities in the area of sustainable development, thus contributing to finding solutions aiming at sustainable quality of life. This will require adaptation and innovation of business models and information systems, with challenges of particular interest to the business modeling and software design community. This paper briefly discusses two relevant topics in this respect, namely (i) goal and value modeling, and (ii) model-driven development. We mention existing work that can be taken as a starting point for addressing sustainability issues, and we make some observations that may be taken into account when extending existing work.

1 INTRODUCTION

The IT-driven globalization has changed business and its underlying economic concepts (Friedman 2007). Also, people start to realize that consumption, financial profits and material ownership are not always positively correlated to well-being. A new paradigm for business and economics is needed, one which emphasizes sustainable development, where ecological, social and psychological dimensions of economic activity are recognized (Tideman 2005). This new paradigm will give quality of life a more central role. Although measurable indicators for sustainable development and quality of life are hard to define, business organizations will play an important role in shaping and testing it in practice. While business may have significantly contributed to the problem, companies to date are definitely considering sustainability issues, be it for reasons of governmental regulation, public image or genuine concern.

Thus, business organizations will explore market opportunities and thus exert activities that are potential solutions to the challenges of sustainable development, including reducing poverty, enhancing livelihoods, protecting ecosystems and tackling climate change (Wilson 2009). They have to invest in and rely on business models and information systems to be able to rationally and effectively address these challenges.

Business models specify how companies achieve their goals and realize a value proposition to their customers. Sustainable development at this level implies that not only economical aspects are taken into consideration when designing business processes, but social and environmental aspects as well. At the technical level, information systems (partially) automate the business processes. The design of information systems for sustainability should be optimized not only for delivering economic value, but also for providing social and environmental benefits. Figure 1 shows the role of business models and information systems in achieving a sustainable and balanced trade-off between economic, social and environmental aspects, both for individual business organizations as well as for business networks in which multiple business organizations collaborate.

Examples of areas where business models and information systems can be applied for sustainable development include smart energy grids (Ipakchi 2009, Warmer 2009), energy-efficient buildings and spaces for public use (Sinderen 2010, Teherian 2010), energy-positive neighborhoods (Lopez 2011), and low carbon mobility and freight transport (Banister 2007, Steenwijk 2011).

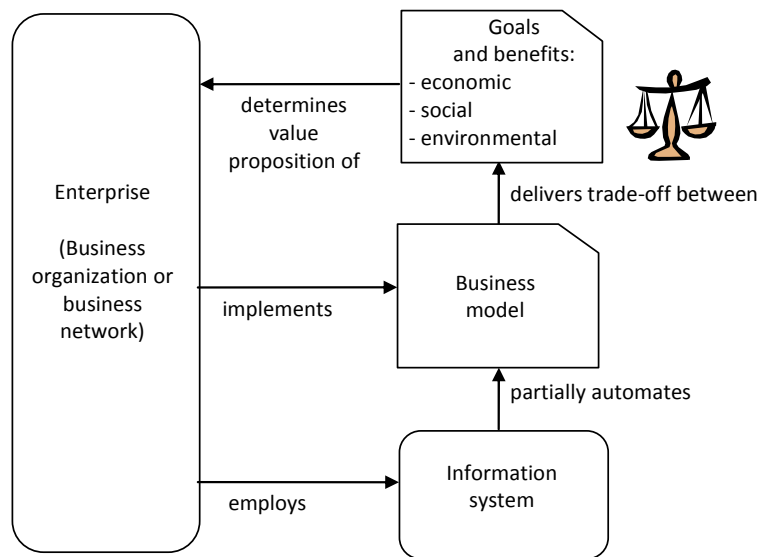


Figure 1: Role of business models and information systems in sustainability development.

We identify two important challenges related to sustainable development, which are of particular interest to the business modeling and software design community:

1. Goal and value modeling;
2. Model-driven development.

In the remaining of this paper we briefly explore these two topics. We mention existing work that can be taken as a starting point for addressing sustainability issues, and we make some observations that may be taken into account when extending existing work.

2 GOAL AND VALUE MODELING

Business models should be able to represent and enable reasoning about the goals and benefits of sustainable development. Trade-offs between the economic, social and environmental aspects of sustainable development goals need to be addressed (Wilson 2009). Goals of an enterprise lead to a value proposition to the enterprise's customers, which is expected to induce value exchanges between the enterprise and its customers. Value in this context can be anything worth in usefulness or importance to the stakeholders, ranging from money, goods, services (economic benefits), and equality, justice, liberty (social benefits), to a balanced ecosystem and stable climate (environmental benefits).

If the enterprise is a business network in which business partners collaborate, then the business partners have mutual value propositions and value

exchanges between each other, in addition to value propositions to and value exchanges with customers. Assuming that all partners are essential for the collaboration, the success of the collaboration will depend on the achievement of a net value (profit or other benefit) for each of the partners. Also, in general, partners want a fair distribution of the overall net value of the collaboration, i.e., no partner should have disproportionate benefits considering his efforts spent and risks taken.

Enterprises are complex systems, hence creating models of such systems in order to understand and analyze their properties normally requires a structured, or multi-viewpoint, approach. Sustainability is one of these properties, which, as we have seen, can be related to multifaceted goals and values. Several techniques exist for goal and value modeling. A well-known approach to problem-oriented goal modeling is goal oriented requirements engineering (GORE), with techniques such as i^* (Yu 1997) and KAOS (Dardenne 1993). A more recent development is ARMOR (Quartel 2009), which has the important benefit of being aligned with a standard enterprise modeling language, namely ArchiMate (TOG 2009). The most popular technique for value modeling is *e3value* (Gordijn 2002). The combined use of goal and value models has been explored in several works, including (Gordijn 2006, Mantovaneli 2009). However, none of the approaches and techniques explicitly addresses sustainability, and how economic, social and environmental aspects can be traded and balanced to allow sustainable development.

Therefore, an important challenge is to extend business modeling, possibly based on current goal and value modeling techniques, but specifically targeted to sustainability. Such extensions may take account of the following observations:

- Trust is an important lubricant for business collaboration. Without a sufficient level of trust among potential partners, collaboration will not take place. Moreover, discovery of misplaced trust in a partner will lead to exclusion of that partner, and possibly to disintegration of the collaboration if the partner cannot be replaced. Initial work on trust and value modeling is reported in (Fatemi 2011a). We expect that a business organization can procreate trust from its potential business partners if it shows to respect or embrace similar values, particularly with respect to social and environmental benefits.
- Local communities maybe easier to involve in building partnerships than remote communities. Availability of background knowledge and trust recommendations plays a role in this, as well as similarity in culture and mutual understanding of values. Moreover, elimination of travel and transport may contribute to reduction of environmental impact.
- Personalization of business offerings to customers is an important factor of reinforcing customer relationships (Marca 2011). Customers are more interested if they feel that their needs and beliefs are taken seriously and if they can influence the value objects in the value exchanges with business organizations. In general, business organizations that are context-aware, i.e., aware of their customers' context but also of their operational context, may be able to exploit this knowledge to save energy and limit environmental impact.

3 MODEL-DRIVEN DEVELOPMENT

As mentioned earlier, at the technical level, information systems (partially) automate the processes and activities as described or constrained by business models. An information system will only be effective if it correctly implements the business models. Furthermore, maintenance of the information system will only be viable if properties expressed by the business models can be traced to capabilities of information system components. Changes to business models, e.g., triggered by market demands, can then be addressed at the

technical level by adapting or replacing the components that support the affected business properties.

Model-driven development, based on the model-driven architecture initiative of OMG (OMG 2003), is a widely accepted approach to manage system and software complexity and to achieve and maintain alignment between business level models and technology level models (Sinderen 2009). In particular, model-driven development organizes and relates the modeling space using a classification of models that offer different abstractions (Frankel 2003). These abstractions cover enterprise design and business network design together with the development of supporting software solutions. A key technical feature of model-driven development is the use of model transformations to (semi-) automatically derive 'target' models from 'source' models, where target models are typically of a lower level of abstraction than source models.

The challenge that is posed related to model-driven development is then: can we propose model transformations that yield models of information systems to support sustainability, given one or more source business models. Given the complexity of enterprise systems, we cannot expect that a single transformation can be used to bridge between the business level and the information system level. Several approaches exist that consider transformations starting from either goal or value models. For example, (Mantovaneli 2009) explores the relations between value (e3value) models and goal (ARMOR) models; (Bandara 2004) presents an approach that applies goal decomposition to achieve policy refinement, using UML as high-level notation for both goals and policies; and (Fatemi 2011b) describes a transformation from a value (e3value) model to a coordination process (BPMN) model using the Groove transformation tool. The latter transformation result could be further exploited when targeting service-oriented information systems. For example, (Khadka 2011) proposes a transformation from a (CDL) model of an abstract coordination process between services to a concrete (BPEL) orchestration of services using the Atlas Transformation Language.

Many more transformation approaches exist, addressing a variety of modeling languages and employing various transformation techniques and tools. It is worthwhile to investigate whether existing approaches can be extended such that sustainability requirements can be preserved and sustainability solutions can be optimized. The

following observations may be taken into account when developing model-driven sustainability:

- Reference models can be useful to guide the transformation process. Transformations from a higher to a lower abstraction level allow many (in principle, an infinite number of) possible target models given a single source model. This is because there are many ways to add lower level details that do not contradict the higher level properties represented by the source model. It may be difficult to know and formalize in advance the stakeholders' considerations that would allow automated selection of the 'best' solution. Hence, such situations would require frequent interaction with and input from a designer acting on behalf of the stakeholders. A well-founded and proven architectural model to which any target model must comply can limit the number of possible choices considerably, thus reducing designer interaction. For instance, (Balmus 2011) proposes a trade promotion management reference architecture that can be used to guide transformations of marketing business models.

- Monitoring and evaluation of sustainability parameters should be anticipated in business models, and should be implemented as far as possible in the supporting information systems. Alencar Silva (2011) presents an initial framework for configuring monitoring strategies from business requirements. A more elaborate framework is proposed in (Wetzstein 2011), which allows business analysts to determine how KPIs for business process performance depend on lower level process metrics and quality of service (QoS) characteristics of the information system infrastructure.

- Additional concepts may be required in languages, both at the business and technical level, to express properties relevant for sustainability development. These concepts should have a careful definition of semantics, in order to be useful in automated transformation approaches. This is especially true at the business level where sustainability sometimes deals with subjective aspects of the enterprise. Moreover, enterprises generally have several stakeholders, involve collaboration between several business organizations, and require the integration of several architectural domains. The lack of a common interpretation of modeling concepts would inhibit model-driven development. The use of ontology-based semantics for concepts has been proposed as a viable approach to address this issue (Azevedo 2011). This work makes use of the Unified Foundational Ontology (UFO) as semantic foundation to evaluate, re-design and integrate

models of conceptual modeling languages. Importantly for our purpose, UFO covers social and intentional phenomena, which need to be modeled when addressing sustainability development.

4 CONCLUSIONS

Traditional business models and information systems must be adapted and innovated in order to fit the new paradigm of sustainable development. Particular challenges for the business modeling and software design community are in the areas of goal and value modeling and model-driven development. We discussed the relationship between multifaceted (i.e., economic, social, environmental) goals and value propositions of an enterprise as a single business organization and as a business network. Trust, local communities and personalization present important boundary conditions for successful business models, and thus have to be explicitly considered in goal and value modeling. Model-driven development can be used to achieve and maintain alignment between the business level models and technology level enterprise models. Particularly, automated model transformations are useful tools to cope with the dynamicity of business constellations and markets. However, current approaches in this area do not explicitly address sustainability. Extensions should consider the use of reference models (especially those that address sustainability), the need for monitoring and evaluation of sustainability parameters, and the definition of additional concepts with ontology-based semantics (using a foundational ontology).

REFERENCES

- Friedman, T. (2007). *The world is flat - The globalized world in the twenty-first century*. Penguin Books.
- Tideman, S. (2005). *Mind over matter: towards a new paradigm for business and economics*. In: *Proceedings of the 2nd International Conference on Gross National Happiness – Rethinking Development – Local Pathways to Global Wellbeing*.
- Wilson, E., J. MacGregor, D. MacQueen, S. Vermeulen, B. Vorley, L. Zarsky (2009). *Business models for sustainable development – Innovating for environment and society*. International Institute for Environment and Development (IIED), Briefing, May 2009.
- Ipakchi, A., F. Albuyeh (2009). *Grid of the future*. IEEE Power and Energy Magazine, Vol. 7, No. 2, pp. 52-62.
- Warmer, C., K. Kok, et al. (2009). *Web services for integration of smart houses in the smart grid*. In:

- Proceedings The Road to an Interoperable Grid, Grid-Interop Forum, pp. 207-211.
- Sinderen, M. van, M. Warnier, B. Shishkov (2010). Service-oriented architecture for household energy management. In: Proceedings of the 4th International Workshop on Enterprise Systems and Technology (I-WEST 2010), SciTePress, pp. 52-69.
- Taherian, S., M. Pias, G. Coulouris, J. Crowcroft (2010). Profiling energy use in households and office spaces. In: Proceedings of the 1st International Conference on Energy-Efficient Computing and Networking (e-Energy 2010), ACM, pp. 21-30.
- López, G., P. Moura, J. Ignacio Moreno, A. de Almeida (2011). ENERSip: M2M-based platform to enable energy efficiency within energy-positive neighbourhoods. In: Proceedings of the IEEE INFOCOM Workshop on Machine-to-Machine Communications and Networking (M2MCN 2011), IEEE, pp. 217-222.
- Banister, D. (2007). The sustainable mobility paradigm. *Transport Policy*, Vol. 15, No. 2, pp. 73-80.
- Steenwijk, M. (2011). A reference architecture for fuel-based carbon management information systems in the logistics industry. MSc thesis, University of Twente.
- Yu, E. (1997). Towards modelling and reasoning support for early-phase requirements engineering. In: Proceedings of the 3rd IEEE International Symposium on Requirements Engineering (RE 1997), IEEE, pp. 226-235.
- Dardenne, A., A. van Lamsweerde, S. Fickas (1993). Goal-directed requirements acquisition. *Science of Computer Programming*. Vol. 20, pp. 3-50.
- Quartel, D.A.C., W. Engelsman, H. Jonkers, M. van Sinderen, (2009) A goal-oriented requirements modelling language for enterprise architecture. In: Proceedings of the 13th IEEE International EDOC Enterprise Computing Conference (EDOC 2009), IEEE Computer Society, pp. 3-13.
- The Open Group (2009). ArchiMate 1.0 Specification. <http://www.opengroup.org/archimate>.
- Gordijn, J., H. Akkermans (2002). Value based requirements engineering: exploring innovative e-commerce ideas. *Requirements Engineering Journal*, Vol. 8, pp. 114-134.
- Gordijn, J., M. Petit, R.J. Wieringa, R.J. (2006). Understanding business strategies of networked value constellations using goal- and value modeling. In: Proceedings of the 14th IEEE International Requirements Engineering Conference (RE 2006), IEEE computer Society, pp. 129-138.
- Mantovaneli Pessoa, R., M. van Sinderen, D.A.C. Quartel (2009). Towards requirements elicitation in service-oriented business networks using value and goal modelling. In: Proceedings of the 4th International Conference on Software and Data Technologies (ICSOF 2009), INSTICC, pp. 392-399.
- Fatemi, H., M. van Sinderen, R. Wieringa (2011a). Trust and value webs. In: Proceedings of the 15th IEEE International EDOC Enterprise Computing Conference (EDOC 2011), IEEE Computer Society, 10 pp. (to appear).
- Marca, D., R. Bulander, C. Kruslin, B. Shishkov, M. van Sinderen (2011). E-business challenges and directions. In: Revised Selected Papers of the 7th International Joint Conference on e-Business and Telecommunications (ICETE 2010), Springer, 33 pp. (to appear).
- Open Management Group (2003). MDA Guide Version 1.0.1. Document No. omg/2003-06-01, 62 pp. <http://www.omg.org/mda>.
- Sinderen, M. van, M. Spies (2009). Towards model-driven service-oriented enterprise computing. *Enterprise Information Systems*, Vol. 3, No. 3, pp. 211-217.
- Frankel, D.S. (2003). Model driven architecture: applying MDA to enterprise computing. JohnWiley and Sons.
- Bandara, A.K., E.C. Lupu, J. Moffett, A. Russo (2004). A goal-based approach to policy refinement. In: Proceedings of the 5th Workshop on Policies for Distributed Systems and Networks (Policy 2004), IEEE Computer Society, 11 pp.
- Fatemi, H., M. van Sinderen, R. Wieringa (2011b). E3value to BPMN model transformation. In: Proceedings of the 12th IFIP Working Conference on Virtual Enterprises (PRO-VE 2011), Springer, 8 pp. (to appear).
- Khadka, R., B. Sapkota, L. Ferreira Pires, M. van Sinderen, S. Jansen (2011). Model-driven development of service compositions for enterprise interoperability. In: Proceedings of the 3rd International IFIP Working Conference on Enterprise Interoperability (IWEI 2011), LNCS 76, Springer, pp. 177-190.
- Bulmus, A.B. (2011). Reference system architecture for trade promotion management leveraging business intelligence technologies and decision support systems. MSc thesis, University of Twente.
- Alencar Silva, P. de, H. Weigand (2011). On the move to business-driven alignment of service monitoring requirements. In: Proceedings of the 3rd International IFIP Working Conference on Enterprise Interoperability (IWEI 2011). LNBIP 76, Springer, pp. 103-117.
- Wetzstein, B., P. Leitner, F. Rosenberg, S. Dustdar, F. Leymann (2011). Identifying influential factors of business process performance using dependency analysis. *Enterprise Information Systems*, Vol. 5, No. 1, pp. 79-98.
- Azevedo, C.L.B., J.P. A. Almeida, M. van Sinderen, D. Quartel, G. Guizzardi (2011). An ontology-based semantics for the motivation extension to ArchiMate. In: Proceedings of the 15th IEEE International EDOC Enterprise Computing Conference (EDOC 2011), IEEE Computer Society, 10 pp. (to appear).

INDUSTRIAL SESSION

BUILDING A BUSINESS PROCESS MAP USING BPM TOOLS AND BPMN NOTATION

Francesc Xavier Berjano Tresserras

*Mediterranean Consulting, Marinel.lo bosch 115-B, 08222 Terrassa, Barcelona, Spain
francesc@mediterranean-consulting.com*

Keywords: BPMN notation, BPM, BPM open source tools, BPMI.

Abstract: This paper aims at discussing business process management tools and the BPMN notation, identifying particular strengths of these tools and notation, supported by a case example, namely a project case on business process re-engineering. We will explicitly consider the phases that we normally follow in realizing such a project (project organization, key people, tasks, resources, deliverables and a project management method). This will support our concluding about strengths of the above-mentioned tools and notation.

1 INTRODUCTION

This paper aims at discussing business process management tools and the BPMN notation, identifying particular strengths of these tools and notation, supported by a case example, namely a project case on business process re-engineering. We will explicitly consider the phases that we normally follow in realizing such a project (project organization, key people, tasks, resources, deliverables and a project management method). This will support our concluding about strengths of the above-mentioned tools and notation.

In general, companies need Business Process Management (BPM) model(s), as a way of assuring good organization. The process mapping related to this is therefore essentially important. It appears like a key activity in managing and organizing all (business goal –driven) activities within a company.

In considering BPM strategies in our work, we are based on the PDCA interactive four-step management process (PDCA stands for ‘plan–do–check–act’); the PDCA process is typically used in considering quality issues. It is also known as the Deming circle (Walton & Edwards, 1988).

As suggested by Fig. 1, a problem definition should come first, followed by analysis & objective setting, measures determination, consideration of alternative solutions, and identification of the best solution. All this concerns the planning. Then solutions need to be applied and evaluated, and finally improvement actions and/or standardization

would need to be applied.

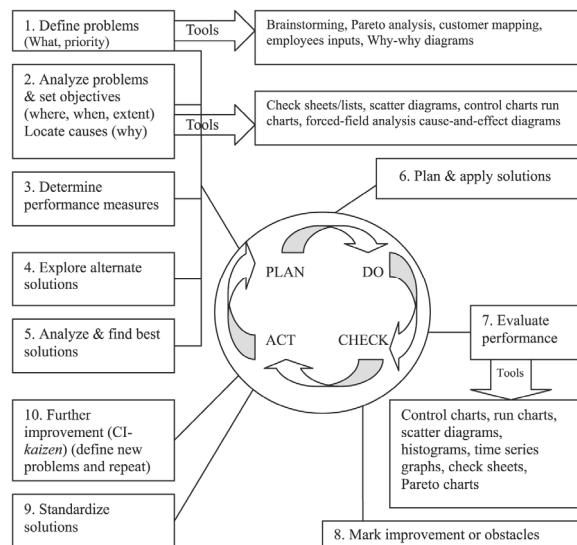


Figure 1: A typical PDCA model.

This is known in theory but applying adequate management and establishing appropriate control in practice remains a challenge.

As already mentioned, we aim in the current paper at considering BPM tools + BPMN notation, and identifying particular strengths of theirs with regard to the above challenge.

The remaining of this paper is organized as follows: Section 2 will briefly introduce and discuss the BPMN notation; further on, Section 3 will consider a business process re-engineering process;

Section 4 and Section 5 then discuss these in the light of a case study – introducing the case is done in Section 4 and tool consideration is done in Section 5; Section 6 contains the conclusions.

2 MODELING STRATEGY USING BPMN NOTATION

Being a business process modeling standard, BPMN was developed by Business Process Management Initiative (BPMI), and is currently maintained by the Object Management Group (BPMN Home, 2011).

With respect to this, it is interesting to consider BPMN from a methodological perspective - origin and which are the key elements, which it possesses. It is interesting not only discussing this but also establishing where to obtain more information and which are the principal institutions linked with BPMN.

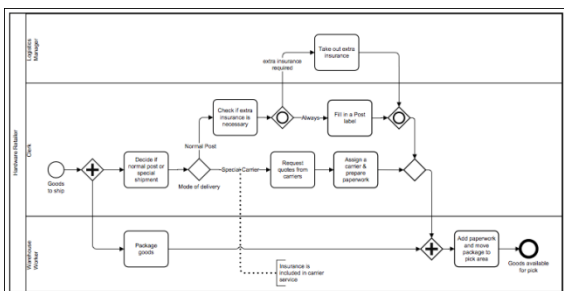


Figure 2: A shipment process of a hardware retailer (a sample).

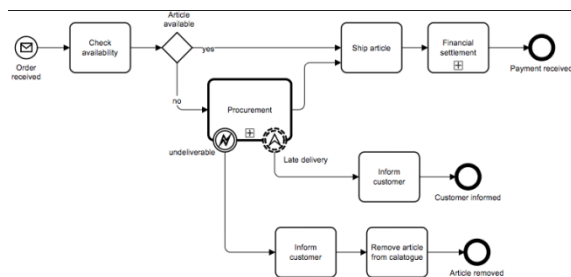


Figure 3: Order fulfilment and procurement (a sample).

Figure 2 and Figure 3 illustrate some important features of BPMN and for more information interested readers are referred to (BPMN Home, 2011).

For the sake of brevity, we will not discuss BPMN in more detail in the current section.

3 PROCESS RE-ENGINEERING PHASES

As according to Dietz (1994), the re-engineering of business processes needs sound modeling and proper notations, such that the re-designed (part of the) business processes are guaranteed to fit in the broader context.

As process improvement is concerned, the typical process re-engineering phases are to be taken in consideration, being reflected in addressing the following points: project organization, key people, tasks, resources, deliverables and project management method.

For the sake of brevity, we will not go into further detail here, presenting only (for partial exemplification) the "Buldeza's Project" that has been realized in the northern Bulgarian city of Pleven.

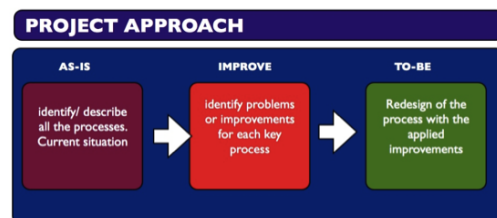


Figure 4: The approach followed in the “Buldeza’s Project.

As seen from Fig. 4, the ‘to be’ situation is reached by properly projecting improvements, having an ‘as is’ situation as starting point.

The way we are modeling and the notations we are using are thus of great importance for appropriately introducing our intention in a real-world business system.

4 DESCRIPTION OF THE “BULDEZA’S” PROJECT

As already mentioned in the current paper, we will discuss a case study – the “Buldeza’s” project.

4.1 Why a BPR Project

The company under consideration is undergoing reorganization in its plant in Bulgaria (Pleven), focused on the improvement of management processes efficiency. The need for reorganization has been motivated not only by the company’s

growth but also by a forecast concerning the support of this increase in next years.

4.2 The Project

The project has been structured in the improvement of three principal areas: production, confection and store of raw material. Also initially, has been setup the control of presence of the people, using an integrated system with the headquarters in Spain.

In production area has been proceeded the improvement of the manufacturing processes, as well as a reorganization of the human resources. There was established a performance control system in plant. After realizing this step, one proceeded to analyze the flow of confection and to establish mechanisms to increase the performance as well as the traceability of the garment inside the area. Later been setup the logistic control integrated to BULDEZA's central warehouse.

The system is based on a method of locations and areas management, , where, since the receipt of the raw material, a total traceability exists: the spinning is identified in the warehouse and is located in a specific position of the store sorted by: quality of the spinning, color, and dyed.

4.3 Results

After finishing this project, the company has in production area one performance control system in plant diary, in different perspectives: machines, articles, people or section.

The Warehouse of raw materials integrated with the production process, then when an order of manufacture is emitted, the spinning is visualized and the system tell us the number of kilos available in the Warehouse and his exact position (X,Y).

For this reason, the time of search of the spinning has diminished because now this information is obtained automatically.

Definitively, nowadays the set of processes of manufacture – confection are more efficient and have allowed to optimize the resources and at the same time to increase the productivity.

5 TOOLS USED

Both for the sake of managing the project and introducing improvements, it is necessary to apply modeling tools and graphical notations in directions discussed already in the paper; Fig. 5 and Fig. 6 are just illustrating this.

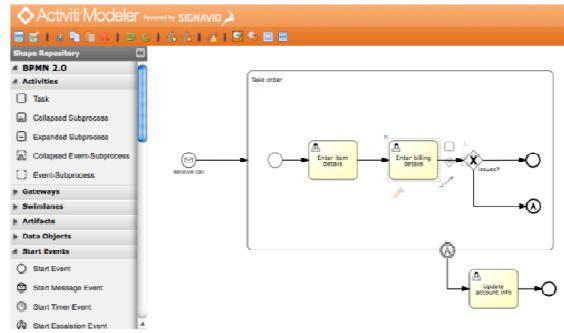


Figure 5: Tools used (Take Order).

As it is seen from these figures, the tools have potential for supporting the management with regard to possible (re-)design of business processes, simulating them, and visualizing their execution.

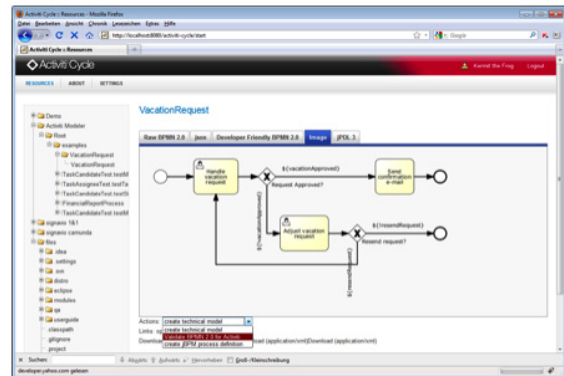


Figure 6: Tools used (Vacation Request).

This potential has been demonstrated in the case study considered.

6 CONCLUSIONS

Business Process Modeling Notation (BPMN) is a graphical representation for specifying business processes in a business process model. BPMN was developed by Business Process Management Initiative (BPMI), and is currently maintained by the Object Management Group since the two organizations merged in 2005. On the other hand, there are business process modeling tools available on the market, based on BPMN, that could be helpful in realizing process generation using an easy and understandable process map. These tools can be grouped by functions such as: modeling, documentation, simulation and execution (in essence, BPM tools take graphical process descriptions as input). A process is composed of

activities that are connected with transitions. Processes represent an execution flow. The graphical diagram of a process is used as the basis for the communication between non-technical users and developers. This all makes BPM tools and the BPMN notation useful at the management level of companies, as a way for mastering complexity and leading businesses properly.

REFERENCES

- BPMN Home (2011), <http://www.bpmn.org>.
- Dietz, J. L. G. (1994): Modelling Business Processes for the Purpose of Redesign. In: *Business Process Re-Engineering*.
- Walton, M., Edwards, W. (1988). The Deming Management Method. Berkley Publishing Group, New York (USA).

SEPARATING AND QUANTIFYING VALUE AND WASTE TO IMPROVE OPERATIONAL PERFORMANCE IN SOFTWARE DEVELOPMENT

Darius Khodawandi

*Wilhelm-Schickard-Institute, University of Tuebingen, Sand 13, D-72076, Tuebingen, Germany
darius.khodawandi@mac.com*

Keywords: Software development productivity, Productivity quantification, Lean software development, Agile software development, Operational performance.

Abstract: Software companies choose to implement Agile software development practices to increase the frequency of their release cycles, stabilize quality, reduce cost - or in short: improve operational performance by increasing productivity. Yet quantifying, what productivity improvement the implementation of Agile software development resulted or will result in is still a big challenge. This paper introduces a new concept to quantify productivity based on the theory of Lean Management - which is closely related to Agile software development - and the specific concept to separate value and waste in processes and products. The author claims that consideration of three dimensions is necessary in the context of software development: (1) the value adding share of product features, (2) the share of the product which contributes to the overall feature set in a value adding way, and (3) the value adding share of time employees can spend when creating software. The ideas proposed in this paper are subject to ongoing trials in an industry environment, which the author is directly involved in. The overall contribution of this paper is a) a new concept to quantify productivity in software development and b) results from initial application of the concept in a large enterprise environment to quantify one of the three dimensions of the overall quantification concept.

1 INTRODUCTION

The market for On-Demand software solutions is growing, and companies are adjusting their business models accordingly. When looking at the changes required more closely, they are all opportunities which - from the authors' point of view - will lead to an industry-wide improvement of customer-orientation and productivity. By offering On-Demand software solutions, installing learning cycles that quickly turn insights from the field into redesigned solutions becomes more easily possible than when solutions are deployed in a traditional fashion. At the same time, productivity will increase because the process by which software is developed has to change dramatically. Based on the previous statement, the process must be capable of handling frequent learnings (i.e. new requirements or changes) and turn them into useable software in a reasonable timeframe. These capabilities are inherent to the concepts of Agile and Lean Software Development, since they promise shorter lead time

from idea to customer use. One reason to be able to reduce lead time by implementing the concepts of Agile and Lean Software Development are productivity gains, which are achieved by e.g. assigning all relevant roles to one team, dedicating team members to only one team and one team to one task at a time, and by co-locating team members to make communication more efficient.

Even though it seems obvious that the concepts of Agile and Lean Software Development - of which only an extract was mentioned above - will lead to improved productivity, there are good reasons to try to quantify the effects of implementation. Besides the convincing nature of figures which is relevant from a pure organizational change view when implementing Agile or Lean Software Development, being able to quantify productivity improvement long-term is regarded as an essential capability for learning software companies. Especially when regarding the concept of Lean Software Development, it is an inherent objective to reduce waste within a company, i.e. any tasks or product

attributes which the customer would not be willing to pay for.

Therefore the objective of this research is to determine a method to quantify productivity of software development from an enterprise perspective, which helps to quantitatively incentivize and prove that working in an Agile or Lean Software Development Model leads to overall improvements. This includes determining what should be measured, and how to measure it in a way to be meaningful and economically implementable. This research is not intended to determine a figure of what improvement to expect from implementing Agile or Lean Software Development, nor how to create a measurement method which is suitable for industry-wide benchmarking.

Section 2 will describe relevant background knowledge regarding Agile and Lean Software Development, and existing quantification approaches for productivity in the software industry. Section 3 introduces and describes the quantification approach in a sequence of *what* to measure, then *how* to measure, and finally draws conclusions. Section 4 contains next steps.

2 BACKGROUND

The relevant existing research for the scope of this paper can be divided into the fields of Agile and Lean Software Development, as well as Productivity Quantification. A more generic background is seen in the work of Drucker (1999) relating to knowledge workers.

2.1 Agile and Lean Software Development

The evolution of process models and concepts for software development reach back to the traditional waterfall and V-Models, followed by Extreme Programming, Spiral or RUP. More recently, approaches have emerged that claim to be more reactive to the environment and the nature of software development - rapid new insights leading to frequent changes, its non deterministic nature (maybe even as a typical attribute to development in general), and with respect to software development specifically, the seemingly indefinite solution space to a given problem space.

These recent approaches go by the name of Agile or Lean Software Development. Both approaches have large commonalities with respect to their basic motivation (e.g. accelerate delivery, build quality in,

customer orientation) - and even though there are some differences (e.g. Agile does not explicitly ask for a separation of value and waste, nor does it define how continuous improvement should work), the author will refer to the two in combination for the purpose of this paper, as it makes no difference in this context. For further reading, Poppendieck (2008) is a frequently cited source for Agile and Lean Software Development

2.2 Productivity Quantification

Quantification of productivity is not a new challenge within the software industry (Scacchi, 1994). A commonly accepted approach is to put the input in relation to the output when calculating productivity. Input can be regarded as the resources utilized and output as what the result is worth. The most difficult aspect is the quantification of output, as it should not be a measure dependent on technical complexity or size (e.g. lines of code), but dependent on the "size" of the customer problem which it solves. The most commonly used and cited approaches are Function Points and User Stories and the corresponding Story Points. These approaches are also utilized to estimate effort based on expected complexity of code.

The existing approaches do not yet fully fit the intentions of Agile and Lean Software Development in respect to their contribution to productivity improvement:

1. Instead of coding efficiency or speed, the question is how efficiently the knowledge worker e.g. a developer can use his/her available time in total (i.e. 100%) - this means what occupies the developer, e.g. rework or creative work
2. An aspect not considered at all in alternative productivity quantification approaches is the question, how much of the assets created by knowledge workers are actually important from the customer point of view.

In this paper, a concept is introduced which attempts to address the above mentioned weaknesses of existing productivity quantification approaches in the industry. Parts of the concept have been applied in practice and are subject to further research.

3 QUANTIFICATION APPROACH

The first question this paper attempts to answer is *what* to measure. Compared to existing research, the author proposes to go down a different path than the

popular approaches. Instead of a white-box view, the author proposes a more black-box oriented selection of criteria to measure. The reasoning for this is to avoid trying to replicate the complexity of a knowledge worker organization in its cause and effect relationships within a productivity figure.

3.1 Overall Concept

Considering the identified weaknesses of the existing approaches and also the strategies which Boehm (2007) identified to improve software productivity (get the best people, make development steps more efficient, eliminate development steps, eliminate rework, build simpler products, and reuse components), the following 3 dimensions should be considered to determine the value-adding share (from the customer point of view) and thus the productivity of software development: (1) feature value add, (2) product value add, and (3) process value add. The order chosen is based on the perceivable sequence of the dimensions from the customers perspective, has no further relevance.

The result of this quantification approach is an overall percentage, resulting from the multiplication of the individual dimensions' value adding share. The overall value adding share is an indicator for the operational performance of software development, as it considers the share of resources on the input side spent value adding, as well as a differentiated view on the output side by distinguishing the customer relevance and the product architecture.

A benefit of the chosen approach is seen in the fact that the unit is percent. Maybe one of the most trivial issues is that a ratio is hard to relate to from an employees perspective (what does a productivity improvement from 2.3 to 3.1 really mean?). The percent figure gives an easy to understand value. The next section describes the concept of separating value and waste focusing on dimension 3 (process value add), as this is the dimension for which trial results are already available.

3.2 Separating Value and Waste

The remaining paper focuses on *how* to measure productivity. The main question of dimension 1 - process value add is: how much of an employees time can be spent on tasks, which are of value to the customer? In order to measure this dimension, the distinction of value adding compared to non value adding tasks needs to be clarified, as well as the level of measurement.

The most reliable distinction criteria for value

add from non value add is - according to Lean Management theory: value adding activities are those activities, which the (end) customer is willing to pay for and thus should be maximized. Non value adding tasks are the remaining activities which should either be minimized (necessary non value adding - e.g. testing) or eliminated (obvious non value adding - e.g. rework). According to this definition, testing as a task is non value adding - a definition which is unusual at first sight but makes sense when one considers the business rationale not to maximize testing, but to try to achieve better quality from the customer point of view with less testing.

Regarding the dimensions (1) feature value add and (2) product value add, research is still required to determine adequate measurement concepts. Dimension (1) can be based on direct customer feedback, in a form where product features are listed and customers are able to rank or classify features (or feature groups) according to their perceived value. Dimension (2) can be based on the concept of target-costing or the quality function deployment method, by which the value of features from the customers point of view is determined and then their cost of creation is put into relation.

3.3 Results to Date

So far, three trials have been executed in an attempt to quantify the process value add axis. Figure 1 shows the task types and their categorization into the classes value add, necessary non value add, and obvious non value add.

Task categories e.g. Code Development	Categorization		
	Value Add	Necessary Non-Value Add	Obvious Non-Value Add
Creating reports	X		
Design compliance checking		X	
Fixing		X	
Handover and review		X	
Information Exchange		X	
Office Activities		X	
Project planning and setup		X	
Reviewing content		X	
Searching			X
Testing		X	
Travelling		X	
Waiting			X
Meetings		X	

Figure 1: Task categories applied in trials.

The first method applied to quantify the value adding share of the process is based on individuals submitting data on a daily basis, about which time was spent for which predefined task category in a granularity of 15 minute intervals. The application

of this method was on-top to regular tasks of the day and for 5 days in a row to eliminate day-specific patterns (e.g. dedicated meeting or travel days). The submitted data was normalized in a way that the length of an individual's work day did not have an impact on the results, i.e. only the calculated percent-figures per individual were considered during aggregation. The second method applied is the integration of the above mentioned task categories into a time tracking system. This way, the generation of data was expected not to be on-top but part of the routine to track working hours. Also, the method allowed continuous generation of data. The third method applied was a simple daily indication, which requested from users their opinion on the value add share of the day. This daily submission was handled by a small tool, which collected the input in a database.

The trial application of the quantification concept follows the definition of measurement as "a quantitatively expressed reduction of uncertainty based on one or more observations" (Hubbard, 2010, p. 23), as it was rather pragmatic. The next section will elaborate on the learnings from the trials.

3.4 Evaluation and Conclusions

Overall, the following modifications were identified during the initial trials to be considered moving forward. (1) The duration of sampling should be aligned to the duration of development sprints. If sprints take e.g. 4 weeks, the duration to collect data should be at minimum 1 or a multiple of the sprint length. This is due to the fact, that otherwise the reported tasks would be biased due to the nature that e.g. at a beginning of a sprint there is more meeting time to discuss the expected outcome. The hypothesis moving forward is that data collection for one sprint is sufficient for a measurement, and that longer data collection would not lead to the equivalent of more informational value. (2) Measuring at the level of tasks may have disguised relevant productivity losses, which occur on the level of activities (i.e. inside tasks). Tadhani (1984) also suggests that limiting measurement to one level is insufficient. It is proposed to distinguish: the level of tasks and the (lower) level of activities. The necessity of this distinction is made clear by the following example: when considering the amount of time spent on testing, taking a closer look reveals that even if an employee spends 4 hours on testing, the actual net productive time is typically less due to necessary test infrastructure setup time, finding and

loading of test cases, initializing the test - all activity level steps that need to occur before the actual test can be conducted. Moving forward, the hypothesis is that at least another 25% of productivity is lost on the level of activities, and that especially system performance is a major contributor for this. (3) Regarding the general measurement method, a comparison of the three trials conducted shows that an improved version of method one seem to be most promising. Method two (work hour tracking) seems to deliver better quality data (based on a bigger sample size and a continuous generation), but has turned out to increase daily complexity of data entry because the number and granularity of task categories is higher than before. Adding the level of activities will even further worsen this situation. Method three seemed to be a simple approach for a large sample size, but the informational value is insufficient (only the value adding share is determined, with no data on the share of time spent on non value adding activities).

4 NEXT STEPS


The measurement method to determine the process value add will be reevaluated in another trial, incorporating the insights described above, first in a university environment, then in a large enterprise environment. In parallel, research to determine adequate measurement concepts for the feature and product axes will continue.

REFERENCES

- Boehm, B. W. (2007), *Improving Software Productivity*. In Selby R. W., *Software Engineering: Barry W. Boehm's Lifetime Contributions to Software Development, Management, and Research*, John Wiley & Sons.
- Tadhani, A. J. (1984). *Factors Affecting Programmer Productivity During Application Development*. In *IBM Systems Journal*, Vol. 23(1), 19-35.
- Scacchi, W. (1994). *Understanding Software Productivity*. In *Advances in Software Engineering and Knowledge Engineering*, D. Hurley (Ed.), Vol. 4, 37-70.
- Hubbard, D. (2010). *How to Measure Anything: Finding the Value in Intangibles in Business* (2nd ed.). John Wiley & Sons.
- Poppendieck, M. and T. (2008). *Lean software development: An agile toolkit*. Addison-Wesley.
- Drucker, P. F. (1999). *Management Challenges of the 21st Century*. New York: Harper Business.

AUTHOR INDEX

Aksit, M.	3	Panatoyotova, T.	80
Alepis, E.	129	Raghupathy, S.	151
Andreev, O.	80	Regev, G.	42
Arnaudov, R.	145	Roubtsova, E.	13
Balzhiev, P.	145	Rozeva, A.	114
Birov, D.	136	Sánchez-Franco, M.	71
Boev, S.	122	Sapkota, B.	34
Borcheva, L.	151	Sokolov, V.	168
Christozov, D.	5	Spence, C.	61
Criado-García-Legaz, F.	71	Surova, E.	122
Doukovska, L.	161	Suurmond, C.	96
Gorbunov, O.	168	Todorova, Y.	136
Huysmans, P.	51	Tresserras, F.	185
Iacob, M.	88	Troussas, C.	129
Ivanov, M.	114	Tsankova, R.	114
Khodawandi, D.	189	van Sinderen, M.	34, 177
Kounev, S.	25	Velev, M.	80
Maurer, H.	9	Villarejo-Ramos, Á.	71
Meertens, L.	88	Virvou, M.	129
Michell, V.	61, 105	Vitanov, R.	145
Nieuwenhuis, B.	7	Wegmann, A.	42
Nieuwenhuis, L.	88	Zhivkov, V.	122
Nikolov, K.	122		



Proceedings of BMSD 2011
First International Symposium on Business Modeling and Software Design
ISBN: 978-989-8425-68-3
<http://www.is-bmsd.org>

