6th International Workshop on

Enterprise & Organizational Modeling and Simulation

EOMAS 2010

Hammamet, Tunisia, 7-8 June 2010

In conjunction with the CAISE'10
22nd International Conference on Advanced Information Systems Engineering

Workshop Proceedings

Editors

EOMAS 2010 Chair
Joseph Barjis
Delft University of Technology, The Netherlands

Murali Mohan Narasipuram
City University of Hong Kong, China

Ghaith Rabadi
Old Dominion University, USA

CAiSE'10 Workshop Chairs
Jolita Ralyté
University of Geneva, Switzerland

Pierluigi Plebani
Politecnico di Milano, Itally

CAiSE 2010 Workshop EOMAS

Proceedings

This volume contains the original articles presented at the 6th International Workshop on Enterprise & Organizational Modeling and Simulation – EOMAS'10. The workshop was held in conjunction with the 22nd International Conference on Advanced Information Systems Engineering, in Hammamet, Tunisia, June 7-8, 2010.

CEUR Workshop Proceedings, CEUR-WS.org, ISSN 1613-0073.

Preface

The 21st century enterprises are crucial components in delivering services to society and contributing to economic prosperity. Service is provided when an enterprise is conducting its business within its business environment. With growing complexity of modern business processes and evolving business environments, enterprises require profound engineering approaches with properties such as ability for reengineering, scalability, adaptability, and reimplementation.

Enterprises are purposefully designed and implemented systems to fulfill certain functions. As any system, enterprises are objects of continuous improvements, redesign and reimplementation. Usually, a redesigning activity is triggered by changes in the business environment, where the enterprise is functioning (delivering its service), or need for efficiency. The departure point for any design or redesign is understanding the enterprise business processes. Therefore, in the overall enterprise engineering activities, business process modeling plays a central role. However, an extended enterprise and organizational study involves both analysis and design activities, in which Modeling and Simulation play prominent roles.

The growing role of Modeling and Simulation attracts serious attention of researchers in the context of enterprises. Modeling and simulation are the tools and methods that are effective, efficient, economic, and widely used in enterprise engineering, organizational study, and business process management. Complementary insights of modeling and simulation in enterprise engineering constitute a whole cycle of study of these complex sociotechnical systems enterprises. In order to monitor and study business processes and interaction of actors in a realistic and interactive environment, animation and gaming are the other two rapidly growing fields associated with enterprise and organizational study, and business process management.

In order to explore these topics, address the underlying challenges, find and improve solutions, and demonstrate application of modeling and simulation in enterprise engineering, its organization and underlying business processes, peer refereed papers have been accepted for presentation at EOMAS 2010. These proceedings include only a subset of the fully reviewed papers while the other subset is published in a book format in the LNBIP series by Springer.

May 2010

Joseph Barjis
Workshop Chair
EOMAS 2010

Organization

The EOMAS workshop is annually organized as an international forum for researchers and practitioners in the field of Enterprise & Organization Modeling and Simulation. Organization of this workshop and peer review of the contributions made to this workshop are accomplished by an international team of researchers in the fields of Enterprise Modeling and Simulation.

Workshop Chair

Joseph Barjis, Delft University of Technology, Netherlands

Program Co-Chairs

Ghaith Rabadi, Old Dominion University, USA
Murali Mohan Narasipuram, City University of Hong Kong, China

Program Committee

Antonia Albani, Delft University of Technology, Netherlands
Anteneh Ayanso, Brock University, Canada
Joseph Barjis, Delft University of Technology, Netherlands
Ygal Bendavid, Polytechnic and Academia RFID, Canada
Peter Bollen, Maastricht University, Netherlands
Mahmoud Boufaida, Mentouri University of Constantine, Algeria
Tatiana Bouzdine-Chameeva, BEM - Bordeaux Management School, France
Manuel I. Capel-Tuñon, University of Granada, Spain
Jan Dietz, Delft University of Technology, Netherlands
Samuel Fosso, Wamba University of Wollongong, Australia
Jose Luis Garrido Bullejos, University of Granada, Spain
Ashish Gupta, Minnesota State University Moorhead, USA
Oleg Gusikhin, Ford Research and Advanced Engineering, USA
Johann Kinghorn, Stellenbosch University, South Africa
P. Radha Krishna, Infosys Technologies Ltd., India
Peggy Daniels Lee, Penn State Great Valley, USA
Oswaldo Lorenzo, Instituto de Empresa, Spain
Prabhat Mahanti, University of New Brunswick, Canada
Yuri Merkuryev, Riga Technical University, Latvia
Vojtech Merunka, Czech U of Life Sciences Prague, Czech Republic
Alta van der Merwe, University of South Africa, South Africa
Martin Molhanec, Czech Technical U in Prague, Czech Republic

Murali Mohan Narasipuram, City University of Hong Kong, China
Ghaith Rabadi, Old Dominion University, USA
Srini Ramaswamy, University of Arkansas at Little Rock, USA
Han Reichgelt, Southern Polytechnic State University, USA
Peter Rittgen, University College of Boras, Sweden
Mamadou Seck, Delft University of Technology, Netherlands
Natalia Sidorova, Eindhoven University, Netherlands
Michel Soares, Federal University of Uberlandia, Brazil
David Sundaram, The University of Auckland, New Zealand
Nick Szirbik, University of Groningen, Netherlands
Yutaka Takahashi, Senshu University, Japan
Andreas Tolk, Old Dominion University, USA
José Tribolet, Technical University of Lisbon, Portugal
Alexander Verbraeck, Delft University of Technology, Netherlands

Auxiliary Reviewers

Kawtar Benghazi Akhlaki, University of Granada, Spain
Michele Fumarola, Delft University of Technology, Netherlands
Yilin Huang, Delft University of Technology, Netherlands
Rick van Krevelen, Delft University of Technology, Netherlands
Max Erik Rohde, The University of Auckland, New Zealands
Christian Stahl, Eindhoven U of Technology, Netherlands

Sponsoring Institutions

- SIGMAS (Special Interest Group on Modeling And Simulation of the Association for Information Systems) - in collaboration
- SIGSIM (Special Interest Group on Simulation of the Association for Computing Machinery) – in Collaboration
- CAiSE 2010 (International Conference on Advanced Information Systems Engineering)
- TU Delft (Delft University of Technology, Department of Systems Engineering)

Table of Contents

Business Process Simulation Revisited

Wil van der Aalst

Eindhoven University of Technology
Department of Mathematics and Computer Science (HG 7.75)
PO Box 513
NL-5600 MB Eindhoven
The Netherlands
w.m.p.v.d.aalst@tue.nl

Presentation Summary

Although simulation is typically considered as relevant and highly applicable, in reality the use of simulation is limited. Many organizations have tried to use simulation to analyze their business processes at some stage. However, few are using simulation in a structured and effective manner. This may be caused by a lack of training and limitations of existing tools, but in this talk we will argue that there are also several additional and more fundamental problems. First of all, the focus is mainly on design while managers would also like to use simulation for operational decision making (solving the concrete problem at hand rather than some abstract future problem). Second, there is limited support for using existing artifacts such as historic data and workflow schemas. Third, the behavior of resources is modeled in a rather naive manner. This keynote will discuss these problems and show solutions that are based on using process mining results and improved modeling of resources. By incorporating better resource characterizations in contemporary tools, business process simulation can finally deliver on its outstanding promise.

Brief Biography

Prof.dr.ir. Wil van der Aalst is a full professor of Information Systems at the Technische Universiteit Eindhoven (TU/e) having a position in both the Department of Mathematics and Computer Science and the Department of Technology Management. Currently he is also an adjunct professor at Queensland University of Technology (QUT) working within the BPM group there. His research interests include workflow management, process mining, Petri nets, business process management, process modeling, and process analysis. Wil van der Aalst has published more than 115 journal papers, 15 books (as author or editor), 230 refereed conference/workshop publications, and 40 book chapters. Many of his papers are highly cited (he has an H-index of more than 70 according to Google Scholar, making him the Dutch computer scientist with the highest H-index) and his ideas have influenced researchers, software developers, and standardization committees working on process support. He has been a co-chair of many conferences including the Business Process Management conference, the International Conference on Cooperative Information Systems, the International conference on the Application and Theory of Petri Nets, and the IEEE International Conference on Services Computing. He is also editor/member of the editorial board of several journals, including the Distributed and Parallel Databases, the International Journal of Business Process Integration and Management, the International Journal on Enterprise Modelling and Information Systems Architectures, Computers in Industry, Business & Information Systems Engineering, IEEE Transactions on Services Computing, Lecture Notes in Business Information Processing, and Transactions on Petri Nets and Other Models of Concurrency. He is also a member of the Royal Holland Society of Sciences and Humanities (Koninklijke Hollandsche Maatschappij der Wetenschappen).

Automated Model Transformations Based on STRIPS Planning

Oldřich Nouza[1], Vojtěch Merunka[2], and Miroslav Virius[3]

[1] Czech Technical University in Prague, Faculty of Nuclear Sciences and Physical Engineering, Department of Software Engineering in Economy,
nouza@fjfi.cvut.cz
[2] Czech University of Life Sciences Prague, Faculty of Economics and Management, Department of Information Engineering,
merunka@pef.czu.cz
[3] Czech Technical University in Prague, Faculty of Nuclear Sciences and Physical Engineering, Department of Software Engineering in Economy,
virius@fjfi.cvut.cz

Abstract. This paper deals with application of STRIPS planning in automated model transformations. Object oriented model is viewed as a state space containing possible models as states and elementary transformations as state transitions. A source model is represented by an initial state, a target model by a goal state. Automation of model transformation consists in finding a plan to reach a goal state in this state space.

Key words: automated model transformations, modeling, object oriented approach, refactoring, SBAT, STRIPS planning

1 Introduction

Transformations play a key role in software engineering. Although there exist satisfiable solutions of automated transformations of models to text, the same cannot be said about transformations of models to models. This area is still in phase of research and exploring of possibilities [1].

According to the approaches of present implemented in several CASE tools, every model transformation requires application of the corresponding transformation rule with suitable parameters [3, 4, 6, 13] Unfortunately, as mentioned in [4], this approach has one big disadvantage, which is a low reusability. For every transformation that has no rule defined, it is necessary to either apply a composition of other rules, or define a new transformation rule. In this paper, we introduce transformation engine named SBAT (STRIPS Based Transformation Engine) that does not require such steps.

2 Model Transformations

There are several ways how to define model transformation. In this paper, we introduce the definition presented in [5] and cited by [8]:

J. Barjis, M.M. Narasipuram, G. Rabadi, J. Ralyté, and P. Plebani (Eds.):
CAiSE 2010 Workshop EOMAS'10, Hammamet, Tunisia, pp. 1-13, 2010.

> *A* transformation *is the automatic generation of a target model from a source model, according to a transformation definition. A* transformation definition *is a set of transformation rules that together describe how a model in the source language can be transformed into a model in the target language. A* transformation rule *is a description of how one or more constructs in the source language can be transformed into one or more constructs in the target language.*

2.1 Classification of Transformations

Publication [8] describes two criteria of classification of model transformations. The first one is a difference of abstraction level of source and target models:

Horizontal transformation – A transformation where source and target models have the same level of abstraction. A typical example is refactoring.

Vertical transformation – A transformation where source and target models have a different level of abstraction. A typical example is refinement.

The second classification criteria is a difference of modeling languages in which the source and targets models are expressed:

Endogenous transformation – A transformation where source and target models are expressed in the same language. Typical examples are refactoring and normalization.

Exogenous transformation – A transformation where source and target models are expressed in different languages. Typical examples are code generation and reverse engineering.

In this paper, we focus more closely on refactoring.

3 Refactoring

There exist several ways how to define refactoring. The definition presented in [7] says that refactoring is an improvement of software system without changing its behavior. In other words, for the same input, the refactored software must return the same output as the original software. Detail information on refactoring is available in [2].

3.1 Complex Refactorings and Primitive Refactorings

The idea of complex refactoring as composition of finite primitive (atomic) refactorings was first published in [10], where formal definition of C++ code refactoring was discussed, and later in [12], where it was demonstrated on refactoring of UML models. We have used the same idea to construct the SBAT transformation engine.

4 STRIPS Planning

Technical report [9] defines STRIPS as following:

> *STRIPS (STanford Research Institute Problem Solver) belongs to the class of problem solvers that search a space of "world models" to find one in which a given goal is achieved. For any world model, we assume there exists a set of applicable operators each of which transforms the world model to some other world model. The task of the problem solver is to find some composition of operators that transforms a given initial world model into one that satisfies some particular goal condition.*

The STRIPS language is based on the calculus of first-order predicate logic. Formally, the STRIPS problem can be expressed by the following definitions:

Definition 1. STRIPS problem *is an ordered triple* (I, O, G)*, where* I *is an initial state,* O *is a set of operators, and* G *is a goal state condition.*

Definition 2. Operator $o(\bar{x})$ *is defined as an ordered triple*(P, A, D)*, where*$P =$ $(\bar{x})$ *is an* application condition*,* $A = [A_1(\bar{x}), \ldots, A_l(\bar{x})]$ *is a set of formulas which become true after operator application,* $D = [D_1(\bar{x}), \ldots D_m(\bar{x})]$ *is a set of formulas which become false after operator application,* $\bar{x} = (x_1, \ldots x_n)$ *are free variables contained in formulas* $P, A_1, \ldots, A_l, D_1, \ldots D_m$*,* $n \in \mathbb{N}^+, l, m \in \mathbb{N}^0$*. Elements of* A *are called* add-effects*, elements of* D delete-effects*.*

Definition 3. *Let* $o(x_1, \ldots, x_n) = (P, A, D) \in O$ *be operator.* A transition function $o' : (x_1 \times x_2 \times \ldots \times x_n \times S) \to S$*, where* S *is a state set, is defined in the following way:*

$$o'(x_1, \ldots, x_n, s) \stackrel{def}{=} \frac{(s \cup A) - D \qquad \varnothing}{s \vDash P} \tag{1}$$

A goal is to find such list of operators applications, which cause transition for initial state to the state satisfying the goal state condition. More formally:

Definition 4. *State* s_m *is a* solution *of problem* (I, O, G)*, if there exists a list of operators applications* $o_1\left(\overline{h_1}\right), \ldots, o_m\left(\overline{h_m}\right)$*, where* $\overline{h_i}$ *are vectors of constants,* $m \in \mathbb{N}$*, and:*

1. $s_0 = I$
2. $(\forall i \in \hat{m})\, s_i = o'_{i-1}\left(\overline{h_{i-1}}, s_{i-1}\right)$
3. $s_m \vDash G$

If $I \vDash G$*, then the solution, which is* I *in this case, is called trivial solution.*

5 SBAT Transformation Engine

5.1 Requirements

Resulting from the facts on present state model transformation discussed in the introduction of this paper, we have decided to design a new transformation engine fulfilling the following requirements:

1. The engine will support model transformations such as refactoring.
2. The input of transformation process will be the source model and conditions of a target model.
3. The output of transformation process will be a target model, or information that the source model cannot be transformed to any model fulfilling the input conditions.
4. The source model and the target model will be consistent in light of behavior of modeled system.
5. The transformation process will be fully automatized, without need to specify transformation rules on input.
6. The engine will be universal and sufficiently reusable for wide scale of object models.

5.2 Formal Definition of Object Model

In this paper, formal definition of object model is based on simplified metamodel of UML 2.0 class diagrams, in detail described in [11]. Because of intended generality, we do not focus on implementation details, such as method parameters and bodies, access mode of class members, etc.

5.2.1 Model as State Space For our purposes we use the primitive refactoring composition mentioned earlier. For this reason, we define model as a state space, where state set represents model in all possible states and state transitions represent primitive refactorings.

Definition 5. *Let C be a class universe, A attribute universe, and F method universe. Let$Object, Client \in C$ be classes, where $\forall (x \in C - [Object]) (Object \prec c)$, which means that Object is parent of all other classes and Client represent a client class, whose object sends messages to objects of classes in model. Letϵ be "empty value".* Model state *s is an ordered 5-tuple $(C_s, in_s, super_s, type_s, send_s)$, where*

- *$C_s \subset C - [Object, Client]$ is a finite set of* classes of model in state *s,*
- *$in_s \subseteq ((A \cup F) \times C_s)$ is a binary relation named* "is in class" *defined as*

$$in_s \overset{def}{=} \{(x, y) \,|x \in (Attr\,(y) \cup Meth\,(y)) \wedge y \in C_s\}\,, \tag{2}$$

where $Attr\,(y)$ and $Meth\,(y)$ are sets of attributes and methods respectively of class y,

- *$super_s \subseteq ((C_s \cup [Object]) \times C_s)$ is a binary relation* "is superclass" *defined as*

$$super_s \stackrel{def}{=} \{(x,y)\,|x = super\,(y) \wedge x,y \in C_s\}\,, \tag{3}$$

where $super\,(y)$ means "a superclass of y",
- *$type_s \subseteq (A \times C_s) \cup (F \times (C_s))$ is a binary relation named* "is of type" *defined as*

$$type_s \stackrel{def}{=} \{(x,y)\,|x = type\,(y) \wedge y \in C_s\}\,, \tag{4}$$

where $type\,(y)$ means "a type of y",
- *$send_s \subseteq (F \times (C_s \cup [Client]) \times (A \cup F) \times C_s)$ is a 4-ary relation named* "sends message" *defined as*

$$\begin{aligned} send_s \stackrel{def}{=} \{(x,y,u,v)\,|&(x,y) \in in_s \\ &\wedge (\exists w)\,((u,w) \in in_s \wedge (u \prec w \vee u = w)) \\ &\wedge \langle x, \Lambda\rangle \in Meth\,(y)\}\,, \end{aligned} \tag{5}$$

where lambda-expression Λ contains message sending $o \lhd u$, where o is an instance of class w.

The following conditions must be fulfilled:

- *in the class hierarchy, each attribute appears at most once, so*

$$(\forall x \in A)\,(\forall y,z \in C)\,(in_s\,(x,y) \wedge in_s\,(x,z) \rightarrow \neg\,(y \prec z) \wedge \neg\,(z \prec y))\,, \tag{6}$$

- *each attribute is of some type, so*

$$(\forall x \in A)\,(\exists y \in C_s)\,(type_s\,(x,y))\,, \tag{7}$$

- *each attribute is of at most one type and each method has at most one return value type, so*

$$(\forall x \in (A \cup F))\,(\forall y,z \in C_s)\,(type_s\,(x,y) \wedge type_s\,(x,z) \rightarrow y = z)\,. \tag{8}$$

Definition 6. Model *is a state space (M,Φ), where M is a finite set of model states and $\Phi = \bigcup_{i=1}^{n} \left[\varphi_i : M \times E^{k_i} \rightarrow M\right], (\forall i \in \hat{n})\,(k_i \in \mathbb{N})$ is a set of transformation rules.*

5.3 Model Transformation Problem

Each model transformation requires answers to the following questions [8]:

1. What needs to be transformed?
2. What will be the result of the transformation?

To find answers, we have to formulate the transformation problem and set the principle of its solution. This can be done by several ways, we have decided to apply the STRIPS planning.

5.4 STRIPS Planning application

Let's assume any finite subset B of object universal, containing elements of all possible model states. To formulate STRIPS problem (I, O, G) for model transformation, we must describe the model states and transformation rules using first-order predicate logic calculus. For this purpose, we define predicates shown in table 1.

Table 1. Predicates for STRIPS problem formulation in SBAT engine

Predicate	Declaration	Definition
class c is in model	$inModel\,(c)$	$c \in C_s$
class c is not in model	$outOfModel\,(c)$	$c \in (B - C_s)$
attribute a is in class c	$attrInClass\,(a, c)$	$(a, c) \in in_s$
attribute a is not in class c	$attrOutOfClass\,(a, c)$	$\neg\,((a, c) \in in_s) \wedge a \in (B \cap A)$
method μ is in class c	$methInClass\,(\mu, c)$	$(\mu, c) \in in_s$
method μ is not in class c	$methOutOfClass\,(\mu, c)$	$\neg\,((\mu, c) \in in_s) \wedge \mu \in (B \cap F)$
attribute a is of type t	$hasType\,(a, t)$	$(a, t) \in type_s$
method μ has return value type t	$hasRetType\,(\mu, t)$	$(\mu, t) \in type_s$
class c is superclass of d	$superClass\,(c, d)$	$(c, d) \in super_s$
class c is a parent of d	$parent\,(c, d)$	$c \prec d$
method $\mu$of class c sends messageη to objects of class d	$sending\,(\mu, c, \eta, d)$	$(\mu, c, \eta, d) \in send_s$

5.4.1 Initial State Formulation Let $m_I = (C_I, in_I, super_I, type_I, send_I)$ be a model in initial state. We construct initial state I of STRIPS problem by the following steps:

1. Put $I := \varnothing$.
2. Add formulas about existence of classes in model:
   a) $(\forall c \in C_I)$ put $(I := I \cup [inModel\,(c)])$ and
   b) $(\forall c \in (B - C_I))$ put $(I := I \cup [outOfModel\,(c)])$.
3. Add formulas about class attributes:
   a) $(\forall\,(a, c) \in (in_I \cap (B \cap A, C_I)))$ put
   $(I := I \cup attrInClass\,(a, c))$ and

   b) $(\forall (a, c) \in ((B \cap A, B \cap C) - in_I))$ put
      $(I := I \cup attrOutOfClass\,(a, c))$.
4. Add formulas about class methods:
   a) $(\forall (\mu, c) \in (in_I \cap (B \cap F, C_I)))$ put
      $(I := I \cup methInClass\,(\mu, c))$ and
   b) $(\forall (\mu, c) \in ((B \cap F, B \cap C) - in_I))$ put
      $(I := I \cup methOutOfClass\,(\mu, c))$.
5. Add formulas about inheritance:
   $(\forall (c, d) \in super_I)$ put $(I := I \cup superClass\,(c, d))$.
6. Add formulas about attribute types and return value types of methods:
   a) $(\forall (a, c) \in (type_I \cap (B \cap A, C_I)))$ put
      $(I := I \cup hasType\,(a, c))$ and
   b) $(\forall (\mu, c) \in (type_I \cap (B \cap F, C_I)))$ put $(I := I \cup hasType\,(\mu, c))$.
7. Add formulas about message sending:
   $(\forall (a, c, b, d) \in send_I)$ put $(I := I \cup sending\,(a, b, c, d))$.

5.4.2 Formulation of Goal State Condition A goal state condition G is defined by the formula that is true in any goal state.

5.4.3 Formulation of Operator Set The set of operators O is defined identically for each particular problem, because it represents a set of primitive refactorings. The complete definition of the operator set is described in table 3 in appendix.

5.5 Example

5.5.1 Problem Let's suppose a simplified class model of file system (see fig. 1).

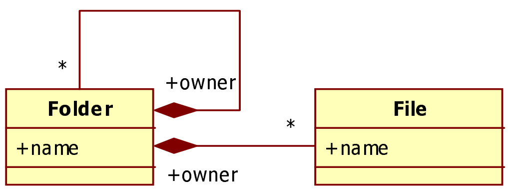


Fig. 1. File system class model

The goal is to transform this model into state which satisfies the composite design pattern.

5.5.2 Solution The initial model state $m_s = (C_s, in_s, super_s, type_s, send_s)$, where:

- $C_s = [Folder, File, String]$
- $in_s = [(name, Folder), (owner, Folder), (name, File), (owner, File)]$
- $super_s = [(Object, Folder), (Object, File), (Object, String)]$
- $type_s = [(name, String), (owner, Folder)]$
- $send_s = \{(Client, main, x, y) | (x, y) \in in_s\}$

The corresponding initial state of STRIPS problem is the following:

$$
\begin{aligned}
I = [&inModel(Folder), inModel(File), inModel(String), \\
&outOfModel(Element), \\
&attrInClass(name, Folder), attrInClass(owner, Folder), \\
&attrInClass(name, File), attrInClass(owner, File), \\
&attrOutOfClass(name, Elememt), attrOutOfClass(owner, Element), \\
&superClass(Object, File), superClass(Object, Folder), \\
&hasType(owner, Folder), hasType(name, String), \\
&sending(Client, main, Folder, name), \\
&sending(Client, main, Folder, owner), \\
&sending(Client, main, File, name), \\
&sending(Client, main, File, owner)]
\end{aligned} \tag{9}
$$

The goal state condition is the following:

$$
\begin{aligned}
G = &inModel(Element) \wedge attrInClass(name, Element) \wedge \\
&attrInClass(owner, Element) \wedge superClass(Element, File) \wedge \\
&superClass(Element, Fodler)
\end{aligned} \tag{10}
$$

The STRIPS planner reaches the goal state by application of satisfiable operators (see table 2).

Table 2. List of operators application to reach the goal state

Operator application	Description
$addClass\,(Element)$	Add class *Element* to the model.
$changeSup\,(File, Object, Element)$	Change superclass *Object* of class *File* to *Element*.
$changeSup\,(Folder, Object, Element)$	Change superclass *Object* of class *Fodler* to *Element*.
$attrUp\,(name, Element, File, Folder)$	Move attribute *name* to class *Element* from its subclasses *File* and *Folder*.
$attrUp\,(owner, Element, File, Folder)$	Move attribute *owner* to class *Element* from its subclasses *File* and *Folder*.

The goal state is as follows:

$$\begin{aligned} Goal = [&inModel\,(Folder)\,, inModel\,(File)\,,\\ &inModel\,(String)\,, inModel\,(Element)\,,\\ &attrInClass\,(name, Element)\,, attrInClass\,(owner, Element)\,,\\ &attrOutOfClass\,(name, Folder)\,,\\ &attrOutOfClass\,(owner, Folder)\,,\\ &attrOutOfClass\,(name, File)\,, attrOutOfClass\,(owner, File)\,,\\ &superClass\,(Element, File)\,, superClass\,(Element, Folder)\,,\\ &hasType\,(owner, Folder)\,, hasType\,(name, String)\,,\\ &sending\,(Client, main, Folder, name)\,,\\ &sending\,(Client, main, Folder, owner)\,,\\ &sending\,(Client, main, File, name)\,,\\ &sending\,(Client, main, File, owner)] \end{aligned} \tag{11}$$

A class model in UML notation corresponding to the goal state is shown in fig. 2.

6 Conclusion and Future Work

In this paper we have introduced SBAT transformation engine based on STRIPS planning system. This engine automates refactoring of the given source model to a target model fulfilling the input condition.

The main asset of SBAT engine for practice is a contribution to an improvement of automation of object model transformations, which consequently would implicate saved human resources for software projects. Then, these resources could be allocated for example on software debugging or testing tasks, rather than on model transformation ones. Another asset should be a theoretical background for research activities in the area of model transformations.

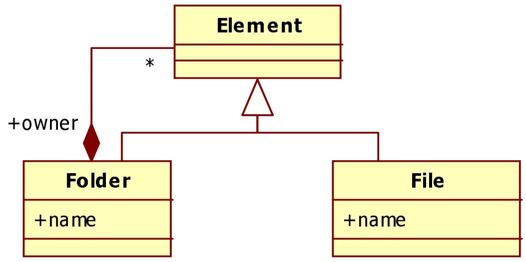


Fig. 2. File system as Composite design pattern

Several consecutive research topics appeared during the SBAT development, e.g. the model state space optimization by reduction of states count or implementation of SBAT in some CASE tool.

Acknowledgment

The authors would like to acknowledge the support of the research grant project SGS10/094 of the Czech Ministry of Education, Youth and Sports.

References

1. Czarnecki K. and Helsen S.: Feature-based survey of model transformation approaches. In: IBM Systems Journal 2006, vol. 45, no. 3, pp. 621-645. (2006)
2. Fowler M.: Refactoring. Addison-Wesley. ISBN 0-201-48567-2. (1999)
3. Gray J., Lin Y., and Zhang J.: Automating Change Evolution in Model-Driven Engineering. In Computer 2006, vol. 31, no. 2, pp. 51. (2006)
4. Jézequel J-M.: Model Transformation Techniques. Available online at `http://modelware.inria.fr/static_pages/slides/ModelTransfo.pdf`. (2005)
5. Kleppe A. G., Warmer J., and Bast W.: MDA Explained: The Model Driven Architecture: Practice and Promise. Boston (MA, USA): Addison-Wesley Longman Publishing Co., Inc., 170 p. ISBN:032119442X. (2003)
6. Lin Y amd Gray J: A model transformation approach to automatic model construction and evolution. In Proceedings of the 20th IEEE/ACM international Conference on Automated Software Engineering, pp. 448-451. (2005)
7. Markovic S.: Composition of UML Described Refactoring Rules. In OCL and Model Driven Engineering, UML 2004 Conference Workshop, pp. 45-59. (2004)
8. Mens T., Czarnecki K, and Gorp P. V.: A Taxonomy of Model Transformations. In Language Engineering for Model-Driven Software Development, ser. Dagstuhl Seminar Proceedings. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Dagstuhl (Germany). (2005)
9. Nilson N. J. and Fikes R. E. STRIPS: A new approach to the application of theroem proving to problem solving. Standford Research Institute, Menlo Park (California) 34 p. (1970).

10. Opdyke W. Refactoring Object-Oriented Frameworks. University of Illinois at Urbana-Champaign, Champaign, (IL, USA), PhD. thesis, 197 p. (1992)
11. Object Management Group (OMG): OMG Unified Modeling Language (OMG UML), Infrastructure: Version 2.2. 226 p. Available online at www.omg.org. (2009)
12. Sunyé G., Pollet D., Le Traon Y., Jézéquel J-M.: Refactoring UML Models. In Proceedings of the 4th International Conference on The Unified Modeling Language, Modeling Languages, Concepts, and Tools. Springer-Verlag, London (UK), pp. 134-148. ISBN 3-540-42667-1. (2001)
13. Zhang J., Lin Y., and Gray, J.: Generic and Domain-Specific Model Refactoring using a Model Transformation Engine. In Volume II of Research and Practice in Software Engineering, pp. 199-218. (2005)

Appendix: Primitive Refactorings as STRIPS Operators

Table 3: Primitive refactorings as STRIPS operators

Primitive refactoring	**STRIPS Operator**	
Add class c	Declaration	$addClass\,(c)$
	Condition	$outOfModel\,(c)$
	Add-effects	$inModel\,(c)$ $(\forall x \in B \cap C)\, attrOutOfClass\,(x, c)$ $(\forall x \in B \cap F)\, methOutOfClass\,(x, c)$
	Delete-effects	$outOfModel\,(c)$
Add attribute a of type t into class c	Declaration	$addAttr\,(a, t, c)$
	Condition	$(\forall x)\,(attrOutOfClass\,(x, d)$ $\vee \neg superClass\,(x, c) \vee \neg superClass\,(c, x)$ $\vee x \neq c) \wedge$ $(\forall x)\,(\neg hasType\,(a, x) \vee (x = t))$
	Add-effects	$attrInClass\,(a, c)$
	Delete-effects	$attrOutOfClass\,(a, c)$
Add method μ to class c	Declaration	$addMeth\,(\mu, c)$
	Condition	$methOutOfClass\,(\mu, c) \wedge (\forall x, y, z)$ $(\neg sending\,(x, y, \mu, z) \vee \neg parent\,(z, c)$ $\vee (z \neq c))$
	Add-effects	$methInClass\,(\mu, c)$
	Delete-effects	$methOutOfClass\,(\mu, c)$
Remove class c	Declaration	$removeClass\,(c)$
	Condition	$inModel\,(c) \wedge (\forall x)\,(\neg parent\,(c, x)) \wedge$ $(\forall x)\, \neg attrInClass\,(x, c) \wedge$ $\neg methInClass\,(x, c)$
	Add-effects	$outOfModel\,(c)$
	Delete-effects	$inModel\,(c)$

Primitive refactoring	**STRIPS Operator**	
Remove attribute a from class c	Declaration	$removeAttr\,(a,c)$
	Condition	$attrInClass\,(a,c) \wedge (\forall x,y,z)$ $(\neg sending\,(x,y,\mu,z) \vee \neg parent\,(c,z)$ $\vee\,(z \neq c))$
	Add-effects	$attrOutOfClass\,(a,c)$
	Delete-effects	$attrInClass\,(a,c)$
Remove method μ from class c	Declaration	$removeMeth\,(\mu,c)$
	Condition	$methInClass\,(\mu,c) \wedge (\forall x,y,z)$ $(\neg sending\,(x,y,\mu,z) \vee \neg parent\,(c,z)$ $\vee\,(z \neq c))$ $\wedge\,(\forall x,y)\,(\neg sending\,(x,y,\mu,c) \vee (y = c))$
	Add-effects	$methOutOfClass\,(\mu,c)$
	Delete-effects	$methInClass\,(\mu,c)$
Change type t of attribute a to u	Declaration	$changeAtrrType\,(a,t,u)$
	Condition	$hasType\,(a,t) \wedge super\,(u,t)$
	Add-effects	$hasType\,(a,u)$
	Delete-effects	$hasType\,(a,t)$
Change type t of values returned by method μ to u	Declaration	$changeMethType\,(\mu,t,u)$
	Condition	$hasRetType\,(\mu,t) \wedge super\,(u,t)$
	Add-effects	$hasRetType\,(\mu,u)$
	Delete-effects	$hasRetType\,(\mu,t)$
Change superclass b of class a to c	Declaration	$changeSup(a,b,c)$
	Condition	$inModel\,(c) \wedge$ $superClass\,(b,a) \wedge (\neg parent\,(a,c)) \wedge$ $\forall\,(x,u,v,w)\,((\neg superClass\,(x,c)) \wedge$ $(\neg superClass\,(x,b)) \vee$ $(\neg sending\,(v,w,u,x))$ $\wedge\,(\neg sending\,(u,x,v,w)))$
	Add-effects	$superClass\,(c,a)$
	Delete-effects	$superClass\,(b,a)$
Move attribute a from class c to its all subclasses $b_1,\ldots,b_n$	Declaration	$attrDown\,(a,c,(b_1,\ldots,b_n))$
	Condition	$attrInClass\,(a,c) \wedge (\forall x)$ $(x = b_1 \vee \ldots \vee x = b_n$ $\vee \neg superClass\,(c,x))$ $\wedge\,(\forall x,y)\,\neg sending\,(x,y,a,c)$
	Add-effects	$attrOutOfClass\,(a,c)$ $(\forall i \in n)\,attrInClass\,(a,b_i)$
	Delete-effects	$attrInClass\,(a,c)$ $(\forall i \in n)\,attrOutOfClass\,(a,b_i)$
Move attribute a to class c from all its subclasses $b_1,\ldots,b_n$	Declaration	$attrUp\,(a,c,(b_1,\ldots,b_n))$
	Condition	$(attrInClass\,(a,x) \vee \neg superClass\,(c,x))$ $\wedge\,(\forall x)\,((x \neq b_1 \vee \ldots \vee x \neq b_n)$ $\vee \neg superClass\,(c,x))$
	Add-effects	$attrInClass\,(a,c)$ $(\forall i \in n)\,attrOutOfClass\,(a,b_i)$
	Delete-effects	$attrOutOfClass\,(a,c)$ $(\forall i \in n)\,attrInClass\,(a,b_i)$

Primitive refactoring	STRIPS Operator	
Copy method μ from class c to its subclass b	Declaration	$methDown\,(\mu, c, b)$
	Condition	$methInClass\,(\mu, c) \wedge methOutOfClass\,(\mu, b) \wedge (superClass\,(c, b))$
	Add-effects	$methInClass\,(\mu, b)$
	Delete-effects	$methOutOfClass\,(\mu, b)$
Copy method μ to class c from its subclass b	Declaration	$methUp\,(\mu, c, b)$
	Condition	$methOutOfClass\,(\mu, c) \wedge methInClass\,(\mu, b) \wedge (superClass\,(c, b))$
	Add-effects	$methInClass\,(\mu, c)$
	Delete-effects	$methOutOfClass\,(\mu, c)$
Add message η sent to objects of class d from method μ of class c	Declaration	$addSend\,(\mu, c, \eta, d)$
	Condition	$(\neg sending\,(\mu, c, \eta, d) \wedge methInClass\,(\mu, c) \vee (\exists x)\,((parent\,(x, d) \vee (x = d)) \wedge (methInClass\,(\eta, x)) \vee (attrInClass\,(\eta, x)))) \wedge (\forall x, y)\,(\neg sending\,(x, y, \mu, c) \wedge (c \neq Client))$
	Add-effects	$sending\,(\mu, c, \eta, d)$
	Delete-effects	$\varnothing$
Remove message η sent to objects of class d from method μ of class c	Declaration	$removeSend\,(\mu, c, \eta, d)$
	Condition	$sending\,(\mu, c, \eta, d) \wedge (c \neq Client) \wedge (\forall x, y)\,(\neg sending\,(x, y, \mu, c))$
	Add-effects	$\varnothing$
	Delete-effects	$sending\,(\mu, c, \eta, d)$
Split message sending $d \lhd \eta$ to $(d \lhd l) \lhd \eta$, where l is of type e.	Declaration	$splitSend\,(\mu, c, \eta, d, l, e)$
	Condition	$sending\,(\mu, c, \eta, d) \wedge attrInClass\,(l, d) \wedge hasType\,(l, e)$
	Add-effects	$sending\,(\mu, c, l, d)$ $sending\,(\mu, c, \eta, e)$
	Delete-effects	$sending\,(\mu, c, \eta, d)$
Merge message sending $(d \lhd l) \lhd \eta$ into $d \lhd \eta$, where l is of type e.	Declaration	$mergeSend\,(\mu, c, \eta, d, l, e)$
	Condition	$sending\,(\mu, c, \eta, e) \wedge sending\,(\mu, c, l, d) \wedge attrInClass\,(l, d) \wedge hasType\,(l, e)$
	Add-effects	$sending\,(\mu, c, \eta, d)$
	Delete-effects	$sending\,(\mu, c, l, d)$ $sending\,(\mu, c, \eta, e)$

A Framework and Methodology for Enterprise Process Type Configurations

Peter Bollen

School of Business and Economics
Maastricht University, Tongersestraat 53,
6200MD, Maastricht, the Netherlands
p.bollen@maastrichtuniversity.nl

Abstract. In this paper we will present the results of research into the semantics of modeling constructs for the process-oriented perspective for the conceptual modeling of enterprise subject areas. We will distinguish 3 conceptual process types that will be the building blocks for any enterprise process base. The definition of these conceptual process types will be anchored in existing process- and decision making frameworks within the fields of management information systems and the administrative sciences

Keywords: Conceptual Modeling, Process Modeling, Process Configurations, Business Process Modeling

1 Introduction

In [1, 2] we have introduced a number of conceptual process configurations in organizations. In this paper we will apply the framework, thereby introducing a modelling methodology for the process-oriented perspective and we will relate our conceptual framework to earlier decision-making frameworks from the fields of (management) information systems and administrative sciences. Furthermore, we will present the meta-model for our process modelling language; the meta-process model (see figure 1).

In this paper we will derive the semantic bridges that we need for instantiating the process modeling constructs (as defined in [1, 2]), in an enterprise subject area under the restriction that they are 'compatible' with the models for the data-oriented perspective in the fact-based approach [3, 4]. In line with the IFIP-CRIS framework [5] we will assume that an application model for the data-oriented perspective is available (see figure 1). Subsequently, we can derive a model for the process-oriented perspective that will use the model in the 'data-oriented' perspective as a starting point, thereby constraining the possible 'process-oriented' models that can exist for the application area and respecting the borders of the application that are imposed by the *Universe of Discourse* (UoD) in the 'data-oriented' perspective.

The remainder of this paper is organized as follows: in section 2 the methodology for instantiating the process modeling constructs in a specific application area will be given, in section 3 some methodological backgrounds will be provided. In section 4 a

J. Barjis, M.M. Narasipuram, G. Rabadi, J. Ralyté, and P. Plebani (Eds.):
CAiSE 2010 Workshop EOMAS'10, Hammamet, Tunisia, pp. 14-29, 2010.

process modelling procedure will be given and in section 5 the meta process model will be given, finally, in section 6 conclusions will be given.

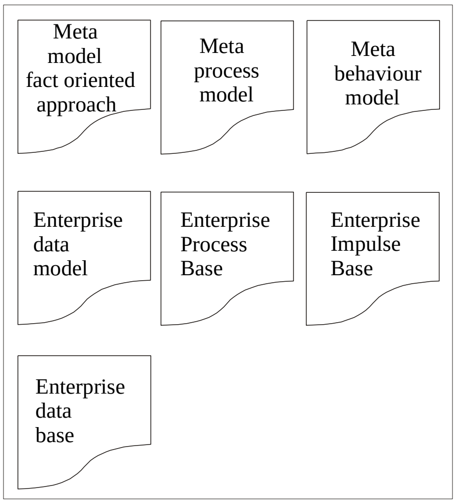


Fig. 1. Documents in the data-, process- and behaviour-oriented perspectives.

In the (information- and knowledge-) management literature different definitions of ‘process’ can be found. Davenport [6] defines a process as “…a structured, measured set of activities designed to produce a specified output for a particular customer or market.”. Hammer and Champy [7] define a process as : “..a collection of activities that takes one or more kinds of input and creates an output that is of value to the customer.” Nickols [8] addresses the issues of identification and analysis of business processes as follows: “..identification and analysis of business processes must be anchored to something concrete.”

1.1 The Enterprise Process Type Configurations in our Framework

In this article we will ‘anchor’ the concept of a conceptual (business) process to the ‘tangible’ result of such a business process in terms of ‘knowledge’. In line with Nickols our position is that processes are not discrete sets of related activities but rather selected portions of larger streams of activity. In figure 2 we have summarized the different conceptual process configurations that we have introduced in [1, 2]. We will provide the definitions of these conceptual process types here.

Definition 1. A *derivation process type* is a conceptual process type whose process instances create fact instances by applying the same derivation rule on instances of the same ingredient fact type(s) (from the enterprise data model).

Definition 2. A *mixed determination process type* is a conceptual process type in which the fact generator uses instances of the same ingredient fact types (that are contained in the application's data model) for all process instances.

Definition 3. A *strict-determination process type* is a conceptual process type in which the fact generator does not use a known derivation rule all the time and the fact generator does not use instances of the same ingredient fact types (that are contained in the application's data model) in all process instances.

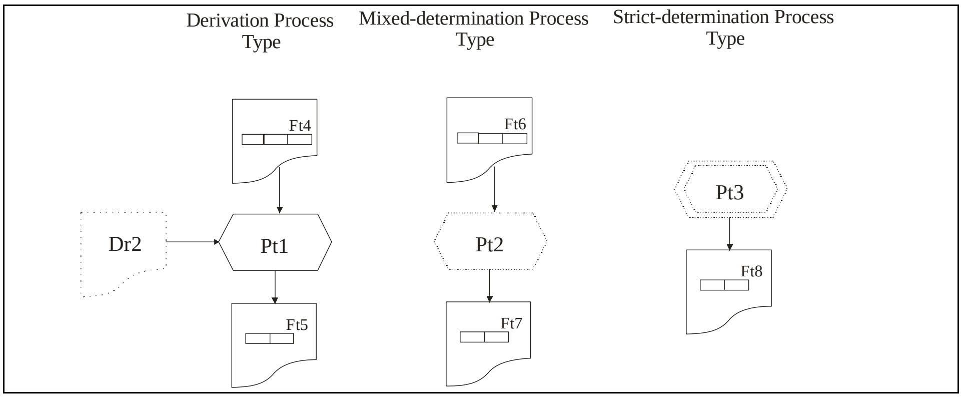


Fig. 2. Conceptual process configuration types (in[1]).

2 Typology of Process Types in Management Information Systems and Knowledge Management

2.1 The Gorry and Scott-Morton Framework

In the field of *management information systems,* Gorry and Scott-Morton [9] have introduced a typology for managerial decision making. In their framework, managerial decisions can considered to be *structured*, *semi-structured* or *unstructured*. This framework was based on the work of Simon who made a distinction into 'programmed' and 'non-programmed' decisions [10]: " The basis for these differences is that in the unstructured case the human decision maker must provide judgement and evaluation as well as insight into problem definition. In a very structured situation, much if not all of the decision-making process can be automated." [9]. Although Gorry and Scott-Morton, refer to managerial decision making in a broad sense, we will interpret their framework in the context of active users that create fact instances as 'materilization' of a decision making process.

Definition 4. A conceptual process type is *structured* if ingredient fact types are known to exist (within or outside the enterprise data model) AND if a derivation rule for the process type is known to exist (either accessible or not accessible by the active users within the SoI).

Definition 5. A conceptual process type is *semi-structured* if ingredient fact types are known to exist (within or outside the enterprise data model).

Definition 6. A conceptual process type is *unstructured* if ingredient fact types are not known to exist (within or outside the SoI) AND a derivation rule for the process type is not known to exist (either accessible or not accessible by the active users within the SoI).

Definitions 4, 5 and 6 are based upon the extent in which ingredient fact types and derivation rules are known to exist. This means that the fact instances of these ingredient fact types are recorded within the focal UoD or within at least one different UoD, and that derivation rules should be potentially accessible by active users within or outside the focal Sphere of Influence.

2.2 The (revisited) Polanyi Framework

Polanyi classifies knowledge into *tacit* knowledge and *explicit* knowledge: " Tacit knowledge is personal, context-specific, and therefore hard to formalize and communicate." 'Explicit' or 'codified' knowledge, on the other hand, refers to knowledge that is transmittable in formal, systematic language" [11]. In Den Hertog et al. [12] tacit knowledge is defined as: " (implicit) knowledge stored in the brains of human beings rather than material knowledge carriers." McBriar et al. [13] define explicit knowledge as: "knowledge that can be represented in words, drawings, plans, equations or numbers, which can easily be communicated between people".

Kim et al. [14] studied the existing distinction into 'tacit' and 'explicit' knowledge in the literature and concluded that a revised epistemology was necessary in order to make a distinction into the concept of 'tacit' knowledge as defined by Polanyi [11] (in which tacit knowledge cannot be expressed externally) and the concept of 'tacit' knowledge as defined by Nonaka [15] (in which tacit knowledge is defined as knowledge that is (currently) not expressed externally). They revised the existing epistemology by replacing the old concept of 'tacit' knowledge by the revised concepts of *tacit* knowledge and the new concept of *implicit* knowledge: "tacit knowledge is knowledge that cannot be expressed externally and implicit knowledge is knowledge that can be expressed externally when needed, but currently exists internally" [15] (p.3).

Definition 7. A conceptual process type is *explicit* if a derivation rule is known to the user groups within the SoI AND all ingredient fact types are contained in the enterprise data model.

Definition 8. A conceptual process is *implicit* if a derivation rule is known to exist outside the SoI but currently is not accessible to the active users of the user group(s) within the SoI OR[1] ingredient fact types are known to exist outside the enterprise data model but currently are not contained in the enterprise data model.

Definition 9. A conceptual process type is *tacit* if a derivation rule does not exist outside or within the SoI.

Definitions 7, 8 and 9 are based upon the extent in which ingredient fact types exist in the enterprise data model and the extent in which the derivation rules are known to user groups in the SoI.

The modalities in the different knowledge typologies, however, cannot be matched 1-on-1 because these three different 'knowledge typologies' are based on different domain paradigms. Gorry and Scott-Morton take managerial decision making as the foundation for their classification. Kim et al. take the extent in which knowledge can be made explicit as a starting point. Our process-typology that we have introduced in [1] has systems theory as its foundation.

3 The Modeling Methodology for the Process-Oriented Perspective

In order to be able to model the process-oriented features for fact types that are contained in the application's data model but that are created in conceptual process instances that are executed by active users *outside* the focal SoI we need to introduce a fourth conceptual process configuration to our framework from section 1.1: the *enter* process type.

Definition 10. An *enter* process type models the process-oriented characteristics for those fact instances of fact types that are contained in the enterprise data model but that are 'created' in conceptual processes by active users *outside* the SoI of the enterprise subject area.

Definition 10 implies that every 'potential' process type that can not be executed under the responsibility of one or more user groups within the SoI will be considered an enter process type when the enterprise process base is created.

If we consider all possible combinations of ingredient fact types, conceptual process types and resulting fact types in terms of whether the fact types are contained in the application's data model and/or whether the instances of the conceptual process types are performed under the responsibility of active users within the application's SoI, we yield 14 possible process types and sphere of influence/UoD combinations. If a *declarative document* or *user example* is within this border, its fact types can be considered to be part of the enterprise data model. If a *conceptual process type* lies within the rectangle it can be considered to be executed under the responsibility of the

[1] To be interpreted as an *inclusive* OR.

user groups within the SoI. If a *prescriptive document* lies within the rectangle, it means that the *derivation rule* on that prescriptive document is accessible to active users in the enterprise SoI that tells potential user groups *how* to execute a process instance.

The duality that exists regarding *declarative* versus *prescriptive* documents can be explained as follows. If a document is qualified as a prescriptive document it means that within that part of the enterprise subject area the document must be considered to specify a course of action. The same document can however, be considered as an instance of a declarative document in the process base of a different (part of the) enterprise subject area. In Anthony's hierarchy [16] three types of management control can be distinguished: *operational-*, *tactical* (or *management*)- and *strategic* control. Within the UoD and SoI of tactical control a derivation rule might be an outcome of a fact-generating activity. In this situation the derivation rule is an instance of a declarative document (for example a lot-sizing decision rule). Such an instance of a declarative document, however, can be used as a prescriptive document for operational control in another UoD and SoI when it is a derivation rule in a prescriptive document (see figure 3).

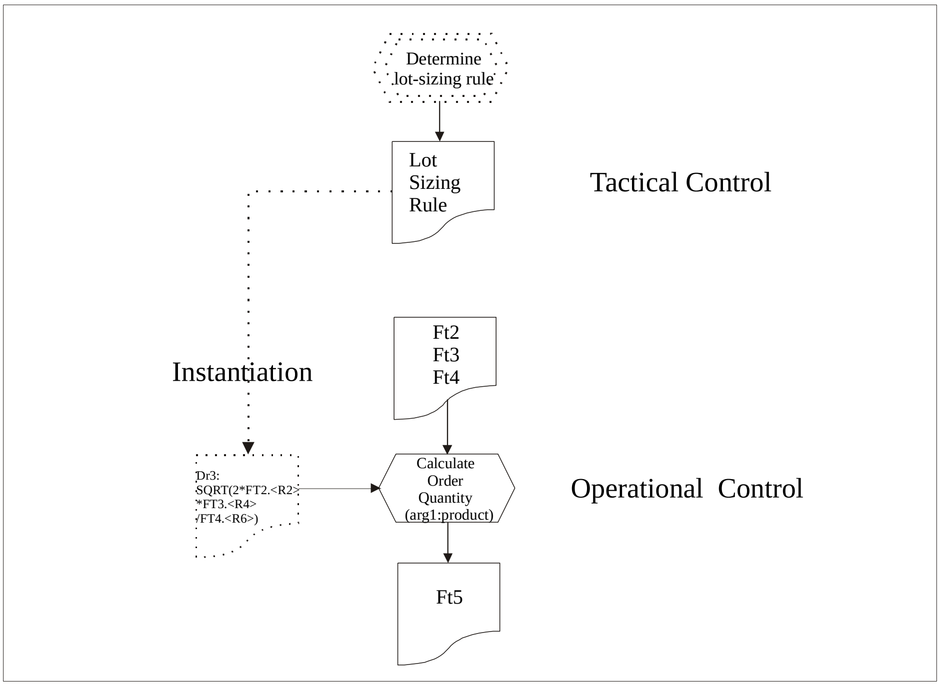


Fig. 3. Application data model for the payroll example.

As a running example for the remainder of this paper we will use the ABC payroll example.

Example: The ABC company

The example application is the payroll department of the branch X of the ABC company. The relationship between this department and the other parts of the organization are shown in figure 4.

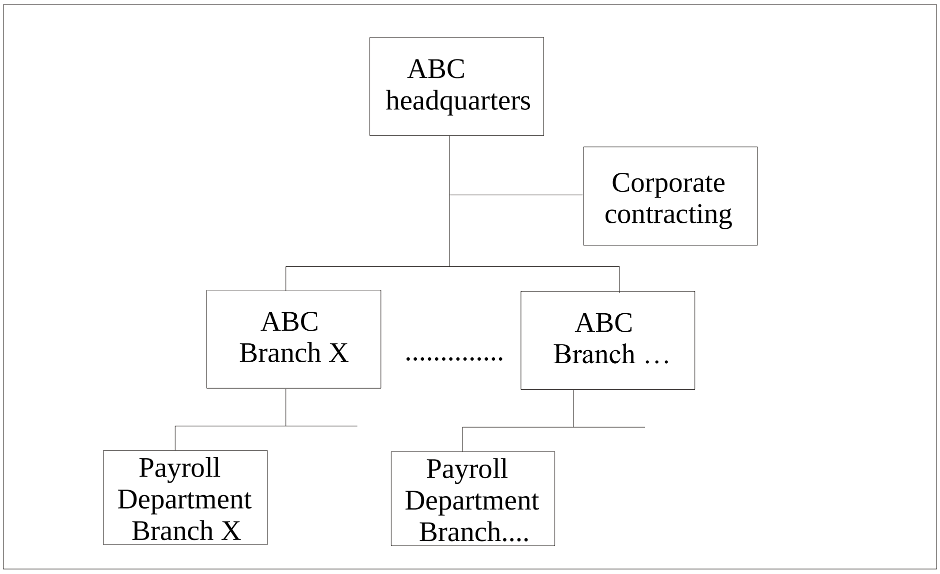


Fig. 4. Organization chart ABC company.

The users in the user groups of the payroll department of branch X, 'decide' how many hours an employee has worked in a given week by inspecting work-order documents and taking additional information into account, e.g. traveling time and information that was obtained in personal contact with the employees. For some employees no work-order documents exist, and therefore the determination of their work-hours is entirely based upon facts that are not contained in the current UoD of the ABC example. The active users in this department furthermore decide upon the gross salaries for the employees that are directly recruited. Although the criteria that determine the salary for each employee are known, the facts that are needed for applying these criteria are not available in the current UoD[2]. The net salary that will appear on the salary slip for the employees is calculated outside the payroll's enterprise area by a payroll service provider. The *gross-to-net* calculation rules are applied by this outside service-agency, and therefore are not accessible by the active users payroll department of the ABC company. Under some conditions it is possible

[2] This would normally suggest, that we will extend the UoD by those documents that contain this information. However, in order to be able to illustrate these 'border' concepts precisely, for now, we assume that we can take any UoD as a starting point and illustrate how we can derive an enterpise process base that belongs to such an arbitrary UoD.

that the working hours for contractors must be recorded although these contractors are not on the company's payroll. In addition it is possible that employees are on the payroll who are hired under the responsibility of a temping-agency. The users in the user groups of the payroll department of the branch X of the ABC company, are also responsible for knowing the highest (gross) salary for an employee at any time.

The fact-based model for the data perspective for this UoD and SoI is given in figure 5. For a brief explanation of the modelling concepts in fact-based modelling we refer to the appendix. The SoI consists of the users in the user group of the payroll department of branch X.

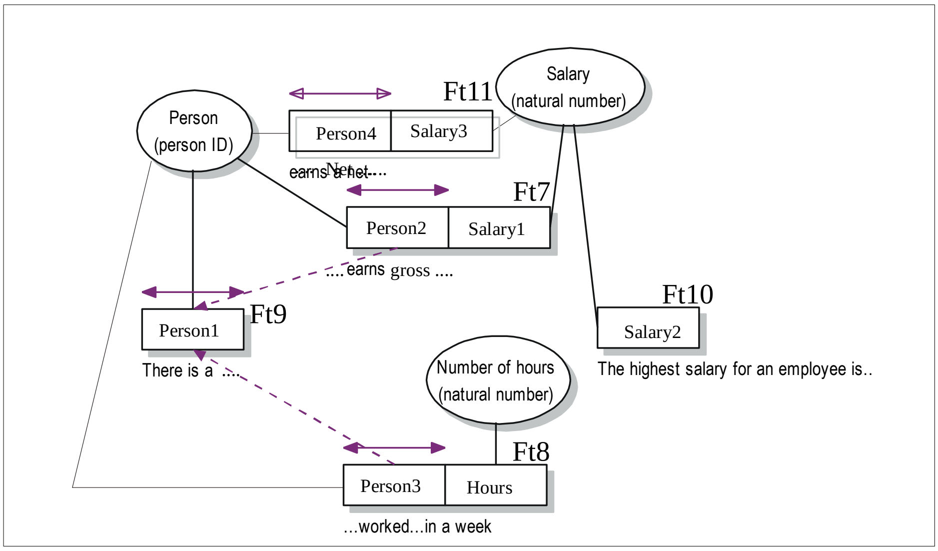


Fig. 5. Fact-based data model for the payroll example.

The content of the fact-based data model in figure 5 can be summarized as follows. There exists fact types that declare the existence of a person (Ft9), that declare that a person earns a gross salary (Ft7), that a person has worked a specific number of hours in a week (Ft8), that there is a highest (gross) salary for an employee (Ft10), and that a person earns a net salary (Ft11). An overview of the fact-based modeling constructs, that are used in figure 5 is given in [17].

We have now discussed all possible situations under which fact instances can be created. We will now synthesize the modalities under which a conceptual process type within an indefinite SoI and UoD can be transformed onto a specific conceptual process type that is defined within the borders of a known application UoD and SoI. The most important modality is the responsibility under which a process instance is executed. If this responsibility lies outside the application's SoI the process will always be modeled as an *enter* process type. We will now consider the modalities from the frameworks of Gorry and Scott-Morton [9] and Kim et al. [14] to characterize the proto-process type configurations and how they map onto the actual processs type configurations when the processes are performed under the responsibility of active users within the SoI.

From table 1 we see that when a conceptual process is performed under the responsibility of (a) user group(s), within the SoI and the ingredient fact types are not known within the enterprise data model or the process type is unstructured this will always lead to a *strict-determination* process type in the the typology of this paper. If a conceptual process type is not explicit but has (an) ingredient fact type(s) that are contained in the enterprise data model then a process type will always be modeled as a *mixed-determination* process configuration. Finally a structured and explicit process type will always be modeled as a derivation process configuration.

Table. 1. Additional modalities for process configurations performed under responsibility of user groups within enterprise SoI.

Gorry-Scott Moton	**Ingredient fact type(s)**	**Derivation rule**	**Kim**	**Conceptual process types in this paper**
structured			explicit	Derivation
structured	In enterprise Data model	Outside SoI	implicit	mixed-determination
structured	Not in entpr Data model	Within SoI	implicit	strict-determination
structured	Not in entpr Data model	Outside SoI	implicit	strict-determination
semi-structured	In enterprise Data model		tacit	mixed-determination
semi-structured	Not in entpr Data model		tacit	strict-determination
unstructured			tacit	strict-determination

We can conclude from the analysis in this section that the two frameworks that we have used from the literature (Gorry and Scott-Morton [9] and Kim et al. [14]) are not sufficient for determining the exact process configuration in case the UoD and SoI are finite. It turns out that the responsibility of the user who ‘performs’ the process instances of the conceptual process type in combination with the precise knowledge on the ‘status’ of the ingredient fact types, in terms of whether they are contained in the enterprise data model, determine the resulting process configuration as defined in this article. The main problem in terms of the applicability of the Gorry and Scott-Morton and the Kim et al. frameworks lies in the notion of ‘falsifying’ the claim that something does not exist. The framework that we have introduced in this paper, however, only needs an answer to the question whether ingredient facts can be found on example documents within the UoD of the enterprise subject area, or whether there exist active users within the given SoI, that have access to a given derivation rule.

4 A Procedure for Deriving the Process Base of an Enterprise Subject Area

The interaction between the UoD (what fact types are relevant for the enterprise subject area) and the SoI (what active users are contained in the enterprise subject area) if not properly managed can be a risk resulting in project delays and project cost overruns in the development life cycle of business information systems. This phenomenan is known as ‘scope-creep’ [18] and is characterized by a human tendency to widen the *SoI* over and over again, thereby extending the application’s UoD with new examples who in turn lead to an extension of the SoI and so on. To overcome these problems concepts like *Rapid Application Development (RAD)* [19] and *timeboxing* [20] emerged in the project management of IT development. These approaches have had a big impact on the project lead times and they enforce information and business analysts and user management to clearly demarcate the enterprise subject area (UoD and SoI) in the analysis stage of the project. The embodying of these demarcation requirements within the *process-oriented perspective* enforces business analysts to decide what informational activity belongs to the *environment* of the 'system' and what informational activity has to be considered part of the system that is subject of the analysis. It should be noted that the *enter* process types **never** have a process type argument, because instances of such a conceptual process type do not have to be instantiated within the *SoI* under consideration. In figure 6 the procedure for the determination of process type signature for given *UoD* and a known *SoI* is summarized.

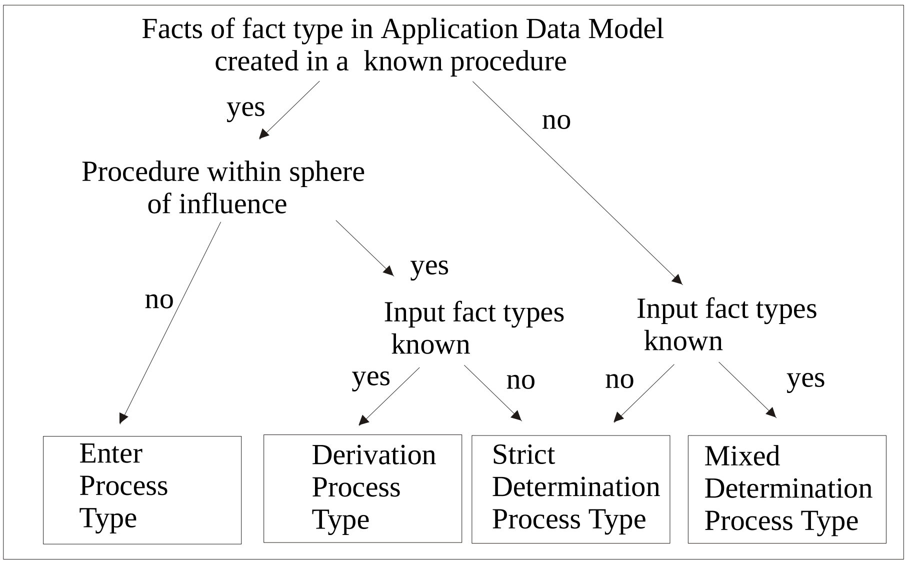


Fig. 6. Procedure for the determination of process type signature for given UoD and known SoI

We can now easily derive an application process base for a given UoD and SoI by applying the decision tree from figure 6. We note that different user groups might use different conceptual process types to create instances of a given fact type. After we have determined all relevant conceptual process types within the sphere of influence we in principle have atomic process type that subsequently can be grouped within a user group to form compound process types.

Definition 11. A process base for a given UoD, user group and SoI contains all *conceptual process types* and *enter process types* that exist within the *SoI* for the fact types in the information grammar of that UoD and user group.

In figure 7 we have given the complete process base for the salary example in a graphical format. We note that for each fact type from the models in the data-perspective at least one process configuration must be contained in the application's process base. To determine to what process type a process instance belongs, that creates an instance of a fact type (that can be created in 2 or more process types), we need an enterprise impulse base, that specifies under what conditions a *specific* process type will be instantiated to create an instance of such a fact type.

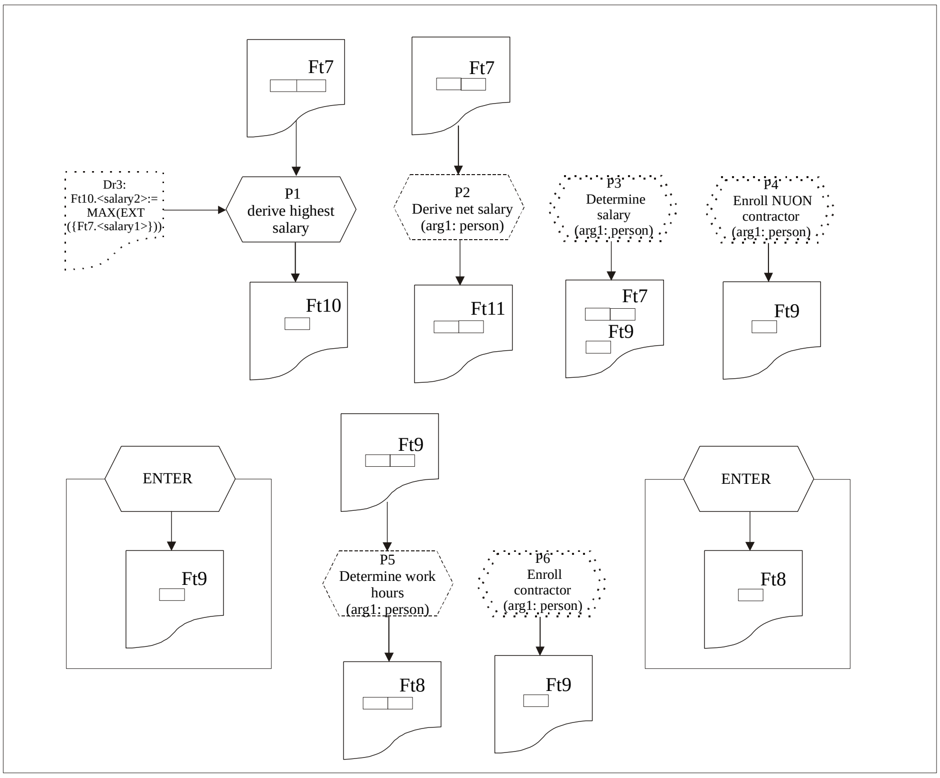


Fig. 7.'As-is' application process base for salary example.

5 The Meta-Process Model

In this section we will give the meta model for the process-oriented perspective. The meta model for the process-oriented perspective or *meta process model* (see figure 1) is a specific application model for the data perspective that is based upon the UoD of a process analyst. The meta process model determines the possible contents of any process base. So a process base of any application UoD and SoI must be an instance of the meta process model. In figure 8 we have given the meta process model expressed as a UML class diagram [21].

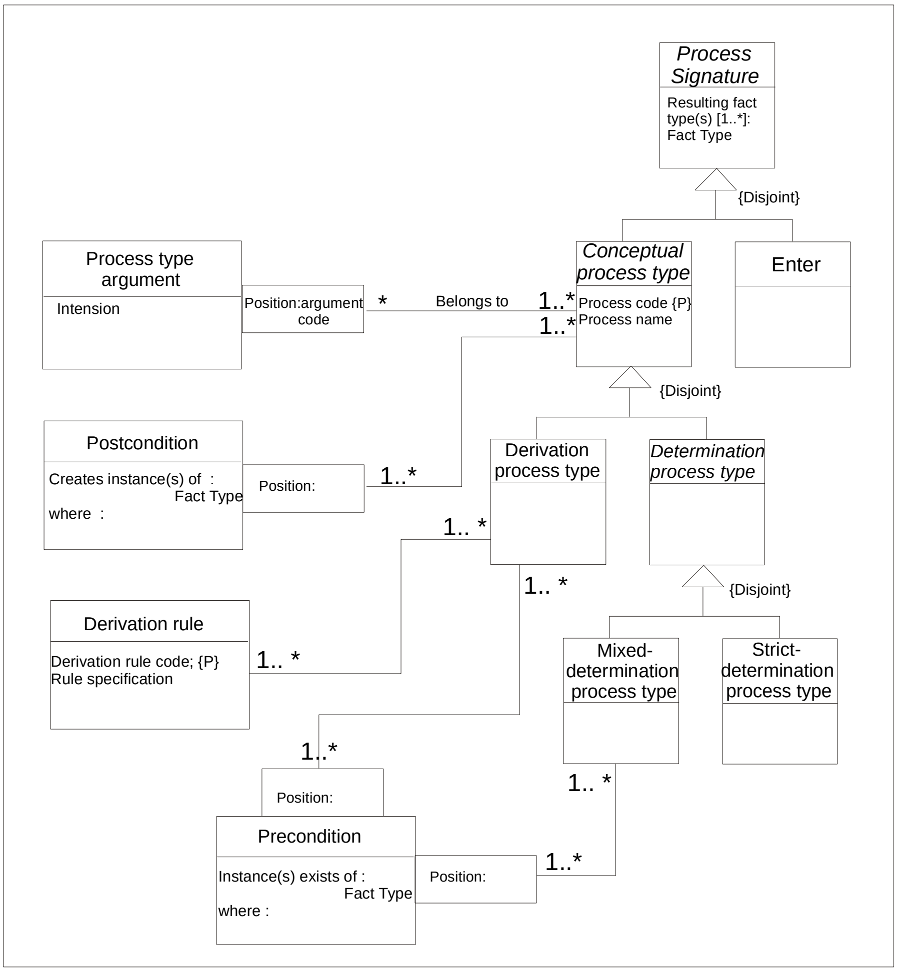


Fig. 8. Meta process model expressed as UML class diagram.

The root of the object-class hierarchy is the abstract object class *process signature*. The root class process signature has two subclasses namely the classes *conceptual process type* and the *enter* configuration. Every instance of the conceptual process type class has a post-condition. The object class conceptual process type, furthermore, has two subclasses: *derivation process type* and *determination process types.* The latter subclass is an abstract sub-class which has two leaf-classes; *strict-determination* process type and *mixed-determination* process type. The meta process model is linked to the meta model for the fact-based approach (for an example see [4]) via the (implicit) object class *fact type*.

6 Conclusion

The definition of three different conceptual process types in combination with the process border-concept of Sphere of Influence (SoI) has resulted in the existence of 3 conceptual process configurations (plus the *enter* configuration) for a given enterprise subject area with a known *UoD* and a known *SoI*. We have shown in this paper that the process configurations are not only determined by the level of 'structuredness' and 'tacitness' in a general sense, but in many instances they are determined by the borders in the data- and process perspectives, respectively. The ability to model conceptual knowledge processes that have a 'tacit' nature and the extent in which the 'codifiable' properties of these tacit knowledge processes can be modeled makes the constructs in the meta process model in this paper applicable in the field of Knowledge Management. The modeling constructs in the framework for the process base in this paper turn out to be applicable in every enterprise subject area, whereas the earlier frameworks of Gorry and Scott-Morton and Kim et al. are hard to apply in real-life situations because they do not have 'finite' border constructs for the enterprise subject areas at hand.

References

1. Bollen, P. *A Conceptual Modeling Language for the Adequate Design of Business Processes*. in *BPMDS '07* 2007. Trontheim, Norway.
2. Bollen, P., *Conceptual process configurations in enterprise knowledge management systems*, in *Applied computing 2006*. 2006, ACM: Dijon, France.
3. Halpin, T. and T. Morgan, *Information Modeling and Relational Databases; from conceptual analysis to logical design* 2nd ed. 2008, San-Francisco, California: Morgan-Kaufman.
4. Nijssen, G. and T. Halpin, *Conceptual schema and relational database design: A fact based approach*. 1989, Englewood Cliffs: Prentice-Hall.
5. Olle, T.W., et al., *Information Systems Methodologies- A Framework for Understanding*. 1988: North-Holland, Amsterdam.
6. Davenport, T., *Process innovation: reengineering work through information technology*. 1993, Cambridge: Harvard Business Press.
7. Hammer, M. and J. Champy, *Reengineering the corporation: a manifesto for business revolution*. 1993, Harper Collins: New York.

8. Nickols, F., *The difficult process of identifying processes.* Knowledge and process management, 1998. **5**: p. 14-19.
9. Gorry, G. and M. Scott Morton, *A framework for management information systems.* Sloan Management Review, 1971(fall).
10. Simon, H., *The new science of management decision making*. 1960: Harper & Brothers publishers.
11. Polanyi, M., *The tacit dimension*. 1966, London: Routledge & Kegan Paul ltd. .
12. Hertog, F.d. and E. Huizenga, *The Knowledge Enterprise: implementation of intelligent business strategies*. 2000: Imperial College Press.
13. McBriar, I., et al., *Risk, gap and strength: key concepts in knowledge management.* . Knowledge-Based Systems 2003. **16**: p. 29-36.
14. Kim, T.-G., S.-H. Yu, and J.-W. Lee, *Knowledge strategy planning: methodology and case.* . Expert Systems with Applications, 2003. **24**(3): p. 295 - 307.
15. Nonaka, I., *A dynamic theory of organizational knowledge creation.* . Organization Science, 1994. **5**(1): p. 14-37.
16. Anthony, R., *Planning and control systems: A framework for analysis*. 1965: Harvard University Press
17. Halpin, T., *Information Modeling and Relational Databases; from conceptual analysis to logical design*. 2001, San Francisco, California: Morgan Kaufmann.
18. Verner, J., S. Overmyer, and K. McCain, *In the 25 years since The Mythical Man-Month what have we learned about project management ?.* . Information and Software Technology 1999. **41**: p. 1021-1026.
19. Beynon-Davies, P. and M. Williams, *The diffusion of information systems development methods.* . The journal of strategic information systems, 2003. **12**: p. 25-46.
20. Jalote, P., et al., *Timeboxing: a process model for iterative software development.* Journal of Systems and Software, 2004. **70**: p. 117-227.
21. Booch, G., J. Rumbauch, and I. Jacobson, *Unified Modeling Language User Guide*. 2nd ed. The Addison-Wesley Object Technology Series). 2005: Addison-Wesley Professional. 496.

Appendix: Fact-Based Modeling concepts

Fact-Based Modeling (FBM) is a methodology for modeling information systems on the conceptual level. It is named after its main constituents: objects that play roles in relationships. The ‘role-based’ FBM notation makes it easy to define static constraints on the data structure and it enables the modeler to populate FBM schemas with example sentence instances for constraint validation purposes. In FBM (and other fact oriented approaches) the fact construct is used for encoding all semantic connections between entities. Figure 9 summarizes the symbols in the FBM modeling language that we will use in this paper.

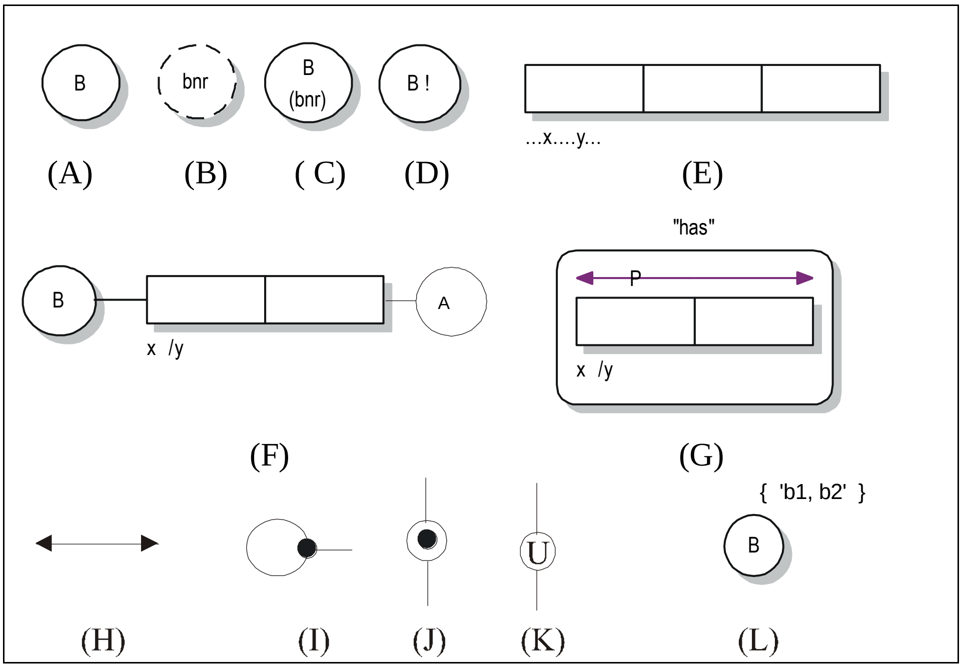


Fig. 9.: Main symbols in Fact-Based Modeling (FBM).

Atomic entities (figure 9A) or data values (figure 9B) are modeled in ORM as simple (hyphenated) circles. Instances of an entity type furthermore can exist independently (e.g. they are not enforced to participate in any relationship), which is shown by adding an exclamation point after the entity type's name (figure 9D). Simple reference schemes in ORM are abbreviated by putting the value type or label type in parenthesis beneath the name of the entity type (figure 9C). Semantic connections between entities are depicted as combinations of boxes (figure 9E) and are called facts or fact types in ORM. Each box represents a role and must be connected to either an entity type, a value type or a nested object type (see figure 9F). A fact type can consist of one or more roles. The number of roles in a fact type is called the fact type arity. The semantics of the fact type are put in the fact predicate (this is the text string …x…y… in figure 9E). A nested object type (see figure 9G) is a non-atomic entity type that is connected to a fact type that specifies what the constituting entity types and/or values types are for the nested object type.

Figures 9H through 9L illustrate the diagramming conventions for a number of static population constraint(s) (types) in ORM. A double-arrowed line (figure 9H) that covers one or more 'boxes' of a fact type is the symbol for an internal uniqueness constraint. The symbol in figure 9K stands for an external uniqueness constraint. A(n) uniqueness constraint restricts the number of identical instances of a role combination 'under' the uniqueness constraint to one. A mandatory role constraint (figure 9I) can be added to a role. It specifies that each possible instance of such an object type must play that designated role at all times. A disjunctive mandatory role constraint (figure 9J) is defined on two or more roles and specifies that each possible instance of the object type connected to these roles must at least play one of these roles at any time. A subset constraint in figure 9K is sometimes depicted as an arrow: ----> between roles or role-combinations. It enforces that the population of the ‘source’ role at all times must be a subset of the population of the ‘target’ role. An in-depth treatment of Fact-Based Modeling can be found in [17].

A Structural Verification of Web Services Composition Compatibility

Kamel Barkaoui , Maryam Eslamichalandar , Meryem Kaabachi

CEDRIC - CNAM Paris, France
kamel.barkaoui@cnam.fr

Abstract. A fundamental feature of service oriented computing is that simple services need to be composed for generating complex services. This work focuses on the analysis and verification of behavior models of web services composition. In particular, we have to check that neither deadlock nor livelock occurs in this composition. Usually, the verification of such integration, with or without mediators, is achieved by using techniques based on state space exploration of a given service formal model. In this paper, we present an approach based on structure theory of Petri nets allowing the recognition of necessary and/or sufficient conditions ensuring compatible composition and a better understanding of the incompatibility sources.

Key words: Web service composition, Compatibility, Structure theory of Petri Nets

1 Introduction

With the increasing use of the platform independent software architecture such as web-based applications, web services exist in distributed environments. Therefore a web service often depends on other web services which have been implemented by different vendors and their correct usage is governed by constraints specified on their interfaces. Whilst different languages such BPEL4WS [7] have been proposed for describing and executing workflow specifications for a web service composition invocation, we still have a critical need of methods and tools to solve many problems related to service interaction [1,5,8]. In this paper, we deal with the issue of verification of web services composition compatibility by using the structure theory of Petri nets formalism. A WS composition is called compatible if its underlying interaction service is such that each service can terminate properly. Our approach is mainly motivated by the fact that verification techniques particularly structural techniques and tools developed for Petri nets can be fully exploited in the context of web services described by BPEL4WS [3], or others. The main goal of this paper is to show how structure theory of Petri nets can provide some guidelines and solutions for ensuring the correctness of web services composition. This paper is organized as follows. Section 2 gives a brief summary of basics of Petri nets and of its related structure theory. Section 3 introduces the open nets as a formal model for web services and their

J. Barjis, M.M. Narasipuram, G. Rabadi, J. Ralyté, and P. Plebani (Eds.):
CAiSE 2010 Workshop EOMAS'10, Hammamet, Tunisia, pp. 30-41, 2010.

composition. In Section 4, using recent results of structure theory of Petri, we deal with the correctness of the WS composition in particular with behavioral compatibility and provide new way of looking at interaction services permitting us the identification of some interface patterns ensuring compatibility between two or more services. Section 5 concludes this paper.

2 Basics of Petri nets

In this section, after giving basic definitions and properties of Petri nets, we present some recent structure theory results.

2.1 Definitions and notations

A Petri Net (P/T) $N = (P, T, F, m_0)$ consists of :

- P a finite set of places and T a finite set of transitions with (P ∪ T) ≠ ∅ and (P ∩ T = ∅),
- F ⊆ (P × T) ∪ (T × P) is the flow relation
- m_0 is the initial marking where a marking m is a mapping m : P → ℕ.

Each node x ∈ P ∪ T of the net has a pre-set and a post-set defined respectively as follows : $^{\cdot}$x = {y ∈ P ∪ T / (y, x) ∈ F}
and x$^{\cdot}$ = {y ∈ P ∪ T / (x, y) ∈ F}.

The incidence matrix of the net is the matrix C indexed by P × T and defined by C(p, t) = W(t, p) - W(p, t) with W(u) = 1 if u ∈ F and W(u) = 0 otherwise.

An integer vector f ≠ 0, indexed by P (f ∈ Z^P) is a P-invariant iff it satisfies ^{t}f·C = 0.
An integer vector g , g ≠ 0, indexed by T (g ∈ N^T) is a T-invariant iff it satisfies C·g = 0.
we denote by $\|f\|$= {p ∈ P / f(p) ≠ 0} the support of f ;
$\|f\|^+$= {p ∈ P / f(p) > 0} and $\|f\|^-$= {p ∈ P / f(p)< 0}, if there exists a P-invariant f /$\|f\|^+$ = P then N is said to be conservative.
and by $\|m\|$ = {p ∈ P / m(p)> 0} the support of marking m .

A transition t is said to be enabled under m iff $^{\cdot}t \subseteq \|m\|$ (i.e. there is a token on every place of $^{\cdot}$t). A transition t enabled under a marking m can be fired, leading to a new marking m' such that : $\forall p \in P : m'(p) = m(p) + C(p,t)$. The set of reachable markings from a marking m in N is denoted by R(N, m).

We recall some behavioral properties of a Petri net N.

- A marking m^* is a home state if and only if $\forall m' \in R(N,m)$, m* $\in R(N,m')$.
- N is reversible iff m_0 is a home state.
- N is bounded iff $\forall p \in P : \exists k \in \mathbb{N}, \forall p \in R(N,m_0), m(p) \leq k$
  i.e. $R(N,m_0)$ is finite.
- N is structurally bounded iff N is bounded for any m_0.
- If N is conservative then N is structurally bounded.
- N is quasi-live iff ∀ t ∈ T ,∃ m ∈ R (N, m_0) for which t is enabled.

– N is deadlock-free (or weakly live) iff $\forall\ m \in$ R (N, m_0), $\exists$ t $\in$ T enabled in m.
– N is live iff $\forall$ t $\in$ T , $\forall$ m $\in$ R (N, m_0) $\exists m' \in$ R (N, m) for which t is enabled.
– N is structurally live iff $\exists m_0$ / (N, m_0) is live.
– A bounded and live Petri net is said to be well formed.

2.2 Basics of Structure Theory of P/T nets

Structure theory of Petri nets investigates the relationship between the behavior and the structure of the net. The use of structural methods for the analysis of systems present two major advantages with respect to other approaches :
the state explosion problem inherent to concurrent systems is avoided , otherwise limited , and this relationship usually leads to a deep understanding of the system.
A remarkable sub structure of Petri nets is that of Siphon.
Let N be a P/T system. A non empty set S$\subseteq$ P is called a siphon if and only if $^{\cdot}$S$\subseteq$ S$^{\cdot}$.
S is said to be minimal if and only if it contains no other siphon as a proper subset.
Due to its structure a siphon which is unmarked will never becomes marked.
In this case, transitions of S$^{\cdot}$ cannot be live so S need to be controlled.
S is said to be controlled if and only if S is marked at any reachable marking i.e. $\forall\ m \in$ R (N, m_0), $\exists$ p $\in$ S / m(p) > 0.
CS-property: N is said to be satisfying the controlled-siphon property if and only if all its minimal siphons are controlled.
We recall below two well-known basic relations between liveness and the CS-property [2]. The first states that the CS-property is a sufficient deadlock-freeness condition while the second states that the CS-property is a necessary liveness condition.

Proposition 1. *Let N be a P/T net. If N satisfies the CS-property then N is deadlockfree (weakly live).*

Proposition 2. *Let N be a P/T net. If N is live then N satisfies the CS-property.*

The following proposition recall the two structural (sufficient but not necessary) conditions permitting us to check if a given siphon is controlled or not.

Proposition 3. *Let S be a siphon of N satisfying one of the two following conditions, then S is controlled [2] :*

i) $\exists\ R \subseteq S$ *such that* $R^{\cdot} \subseteq {}^{\cdot}R$ *and* $R \cap \|m_0\| \neq \emptyset$
ii)$\exists$ *P-invariant* $f \in Z^P$ *such that* $S \subseteq \|f\|$, $\|f\|^+ \subseteq S$
and , $\sum_{p\in P}$ *[f(p). $m_0(p)$]>0*

K-Systems : P/T nets for which CS-property is not only necessary but also sufficient liveness condition, in other words , systems for which there is equivalence between liveness and CS-property are called K-systems [2].

Root-place : Let $t \in T$ be a transition of a P/T net N and $r \in {}^{\cdot}t$; r is called a root place for t if and only if $\forall\ p \in {}^{\cdot}t$, $r^{\cdot} \subseteq p^{\cdot}$.
Ordered transition : A transition $t \in T$ is said to be ordered if and only if $\forall$ p, $q \in {}^{\cdot}t$, $p^{\cdot} \subseteq q^{\cdot}$ or $q^{\cdot} \subseteq p^{\cdot}$, an ordered transition has at least one root place. A transition admitting a root place is not necessarily ordered.
We denote by:

- Root (t) the set of root places of t.
- T_0(N) the set of ordered transitions of N.
- T_R(N) the set of transitions of N admitting a root (i.e. $T_O(N) \subseteq T_R(N)$).
- Root (N) the set of root places of N.
- The Root Component of N is the net $R_C(N) = (P_C(N), T_C(N), F_C(N))$ defined as follows:
  - P_C = Root (N), $T_C = T_R(N)$.
  - F_C is the restriction of F such that:
  (p,t) $\in F_C$ iff p $\in$ Root(t) and (t, p) $\in F_C$ iff (t, p) $\in$ F.

Two main subclasses of K-systems namely ordered nets and root nets can be recognized structurally and effectively [2].

(1) N is called an Ordered net iff $T_O(N) = T$ (i.e. all its transitions are ordered).

(2) N is called a Root net iff $T_R(N) = T$, ($T_O(N) \neq T$) and its root component $R_C(N)$ is bounded and strongly connected.

Note that by definition these two subclasses are disjoint.

Theorem 1. *Let N be an Ordered net or a Root net. N is live if and only if it satisfies the CS-property [2].*

In particular for well known subclasses of ordered nets for which
Root (t) = ${}^{\cdot}t\ \forall\ t \in T$, (therefore $R_c(N) = N$) such Extended Free Choice (EFC) nets, the cs-property (by condition i) is a necessary and sufficient liveness condition . Moreover, if such nets are bounded then liveness property (i.e. here condition (i)) can be decided in polynomial time [10]. Also as control by trap (i) is preserved after increasing marking (contrary to control by invariant (ii)), liveness property is monotonic for K-systems satisfying CS-property by condition (i). Based on these structural theoretical results , we show in the rest of the paper , how compatibility analysis and verification of web services composition can be under taken efficiently .

3 Modeling and Specification of web Service Composition

For specification and modeling services, we focus on the concepts which are independent of a given implementation language [7,9] . First of all, a service has a definition describing its behavior and its interface . An instance of a given service

corresponds to an execution of the activities of this service .These activities are atomic units of work specified in the service definition. The interface of a service consists of a set of ports. A pair of ports can be connected using a channel, thus enabling the exchange of messages sent or received by services .
In this work , we abstract from non-functional properties, data and information semantics. Hence , web service can be viewed as a control structure describing its behavior according to an interface to communicate asynchronously with other services in order to reach a final state (i.e. a state representing a proper termination). As for modeling of business processes and workflows, P/T nets are well appropriate to model such control perspective of web services. A web service is modeled by a P/T net called open net [5] which is an extension of a workflow net [6] or service net [4] by adding, to the internal places, two specific disjoint sets of input and output places (called interface places) modeling the service interface. Services can be composed by connecting the interfaces.
More precisely, each input place (i.e. with empty pre-set) corresponds to an input port of the interface (used for receiving messages from a distinguished channel) whereas an output place (i.e. empty post-set) corresponds to an output port of the interface (used for sending messages via a distinguished channel).

Definition 1. *An open net $N = (P, T, F, I, O, m_0 , m_f)$ consists of :*
A Petri net $N^* = (P, T, F, m_0 , m_f)$ such that :

– $m_0 = s$ *(the initial marking of the service)* , $^{\bullet}s = \emptyset$
– $m_f = o$ *(the final marking of the service)* , $o^{\bullet} = \emptyset$

with an interface places $(I \cup O \subseteq P\ )$ *such that :*

– $\forall\ p \in I \cup O, m_f(p) = m_o(p) = 0$
– $\forall\ p \in I\ ,\ ^{\bullet}p = \emptyset$ *(input interfaces places)*
– $\forall\ p \in\ O\ ,\ p^{\bullet} = \emptyset$ *(output interface places)*

Our definition of open nets is not restrictive . Indeed any P/T net N* with an initial marking defined on more than one initial place , or admitting a set of final markings (with mutually exclusive supports) , can be transformed easily to an equivalent open net. Also our open nets are not elementary communicating in the sense that a transition can be connected to more than one interface place.

The basic web services infrastructure provides simple interactions between a client and a web service. However, the implementation of a web services business needs generally the invocation of other web services. Thus it is necessary to combine the functionality of several web services. The process of developing a composite service is called service composition. Composite services are recursively defined as an aggregation of elementary and composite services. The composition of two or more services generates a new service providing both the original behavior of initial services and a new collaborative behavior for carrying out a new composite task [7].
From a modeling point of view, a composite service can be described as a

recursive composition of open nets [8]. Communication between services takes place by exchanging messages via interface places. Thus, composing two open nets is modeled by merging their respective shared constituents which are the equally labeled input and output interface places. Such a fused interface place models a channel and a token on such a place corresponds to a pending message in the respective channel. As it is convenient to require that all communications are bilateral and directed, i.e. every interface place p $\in (I \cup O)$ has only one open net that sends into p and only one open net that receives from p.
Thereby, open nets involved in a composition are pairwise interface compatible i.e. only input interface places of the one open net overlap with output interface places of the other. This interface compatibility is a basic and first requirement for services composition .

Definition 2. *Let N1 and N2 be two open nets with pairwise disjoint constituents except for the interfaces . If* $I = (I_1 \cap I_2) = \emptyset$ *and* $(O_1 \cap O_2) = \emptyset$ *then* N_1 *and* N_2 *are interface compatible.*

Definition 3. *Let N1 and N2 two interface compatible open nets. Their composition* $N = N_1 \oplus N_2$ *is the open net defined as follows:*

– $P = P_1 \cup P_2$; $T = T_1 \cup T_2$; $F = F_1 \cup F_2$;
– $I = (I_1 \cup I_2) \setminus (O_1 \cup O_2)$; $O = (O_1 \cup O_2) \setminus (I_1 \cup I_2)$;
– $m_0 = m_{01} \oplus m_{02}; m_f = m_{f1} \oplus m_{f2}$

Open net composition is commutative and associative i.e. for interface compatible open nets N1, N2 and N3 : $N_1 \oplus N_2 = N_2 \oplus N_1$
and $(N_1 \oplus N_2) \oplus N_3 = N_1 \oplus (N_2 \oplus N_3)$.
An open net with an empty interface (I = $\emptyset$ and O ˙= $\emptyset$) is called a closed net.
By choreography , we refer to the coordination of messages between services involved in a composite service. Therefore a service choreography can be described as a closed net .
The next section is devoted to check the verification of behavioral properties of a closed obtained by composing open nets.

4 Structural Verification of Composition Compatibility

A composite web service modeled as a closed net is a service that consists of coordination of several conceptually autonomous but interface compatible services. Although it is not easy to specify how this coordination should behave, we focus here on these three behavioral requirements :

- *Weak-Compatibility* . A closed net N is said to be *weak-compatible* iff N is *deadlock-free.*
- *Compatibility* which excludes not only deadlocks but also livelocks. A closed net N is said to be compatible iff m_f is *home state*(final state is always reachable).
- *Strong-Compatibility.* A closed net N is said to be strong compatible iff N is compatible and quasi-live (proper termination and no dead activities).

Our contribution in this paper is to show how using recent results of structure theory of Petri nets (that can be interpreted as restrictions or operating guidelines on service interaction patterns), we can check or ensure structurally these behavioral properties.

Let us precise that a deadlock state m in a closed net N is a reachable state (m $\neq m_f$) under which no transition is enabled.
Obviously, compatibility implies weak compatibility.
Let N $= N_1 \oplus N_2 \oplus \ldots N_k$ be a closed net.
Let N_i be an open net , $N_i^* = (P_i, T_i, F_i, m_{0i}, m_{fi})$ is called the inner subnet of N_i . We denote by N_i^{**} the subnet obtained from N_i^* by connecting the initial place s_i to the terminal place o_i by an additional transition t_i^*.
Let N $= N_1 \oplus N_2 \oplus \ldots N_k$ we denote by θ(N) the net obtained by substituting in each N_i, N_i^* by N_i^{**}.

First of all, from the two well known propositions (1) and (2), we can deduce easily the two following propositions :

Proposition 4. *Let $N = N_1 \oplus N_2 \oplus \ldots N_k$ be a closed net. If θ(N) satisfies the cs-property then N is weak compatible.*

Proposition 5. *Let $N = N_1 \oplus N_2 \oplus \ldots N_k$. If N is strong compatible then θ(N) satisfies CS property.(we prove that θ(N) is live)*

Let us consider the closed net obtained by the two open nets of Fig.1 described in [5] .As the cs-property , is not satisfied : the siphon S = (food , money, P7 , P3) is empty at m_0 , N cannot be live neither deadlockfree.
Consequently N is not weak compatible.

Now, Consider the two interface compatible open nets of Fig.2 , the corresponding closed net N is such that θ(N) satisfies the cs-property therefore N is weak compatible.

However N is not compatible : indeed the final marking $m_f = p_4 + p_{14}$ cannot be reached from the accessible marking $m^* = p_4 + p_{14} + p_7$.

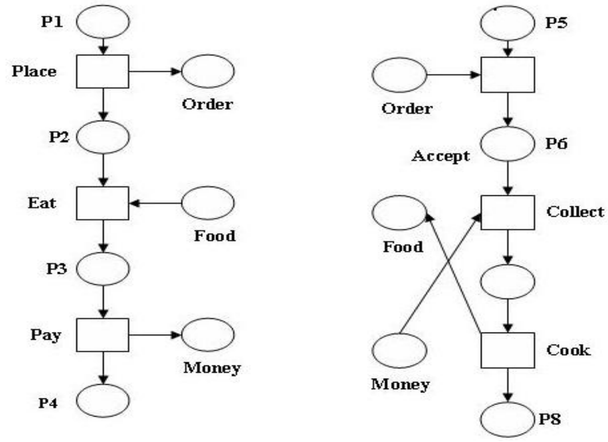


Fig. 1. A not weak compatible closed net

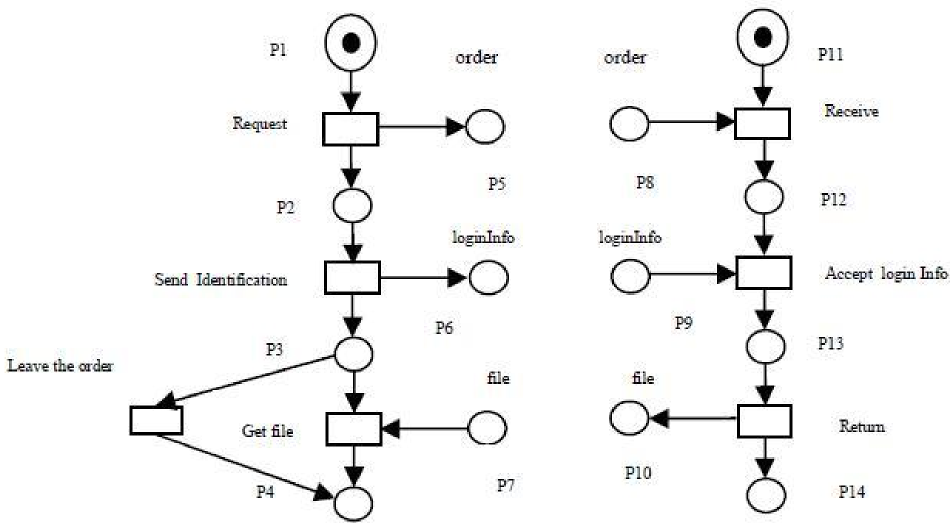


Fig. 2. A weak compatible (but not compatible) closed net

Theorem 2. *Let $N = N_1 \oplus N_2 \oplus \ldots N_k$ be a closed net. If N is strong compatible then all N_i^* are sound.*

Proof. Suppose there exists N_i^* not sound, i.e. N_i^{**} is not live or not bounded. Case (1): N_i^{**} is not live i.e. there is transition t $\in T_i$ not live in N_i^{**}. As (input) interface places only limit the behavior of the associated open net N_i^*, t remains not live in θ(N), thus N cannot be strong compatible.

Case (2): N_i^{**} live but not bounded, thus m_f cannot be a home state and N is not compatible.

According to previous results, strong compatibility of open nets requires not only interface compatibility of open nets but also soundness of their inner subnets. We define now two classes of open nets namely Ordered open nets and Root open nets for which soundness is equivalent to cs-property[6].

Definition 4. *Let N be an open net. N is called an Ordered open net if and only if N** is an ordered net.*

Definition 5. *Let N be an open net. N is called a Root open net if and only if N** is a Root net.*

From this two classes of open nets , we define a large subclass of closed nets called Root closed nets presenting realistic interfaces patterns and for which compatibility can be structurally decided. In this subclass we impose a restriction on the connection nature of interface places such that root internal places are preserved after composition i.e. an input interface place can be a root place but it cannot take the place of another internal one. A larger subclass of composite service can be obtained by applying the basic building process of Root closed nets in a recursive way, i.e. modules can be root closed nets or more complex nets defined in this way.

Definition 6. *A P/T system N = (P, T, F, m_0) is called a Root Closed net (or simply an RC net) if and only if P is the disjoint union $P_1, \ldots, P_n$ and B , T is the disjoint union $T_1, \ldots, T_n$ and the following holds:*

i) For every $i \in \{1, \ldots, n\}$, let $N_i = < N_i^{**}, I_i, O_i >$ be an open net such that :
- *$(I_i \cup O_i) \subseteq B$*
- *$N_i^{**} = (P_i, T_i, F_i, m_{0i}, m_{fi})$ where $F_i \subseteq (P_i * T_i) \cup (T_i * P_i)$ is an ordered or root open net satisfying CS-property.*

ii)For every $N_i^{**} i \in \{1, \ldots, n\}$: $\forall b \in B$, b preserves the sets of root places of N_i^{**} (i.e $\forall t \in T_i$, $Root(t)_{Ni^{**}} \subseteq Root(t)_{Ni}$

iii)There exists a subset $B' \subseteq B$ such that the sub net induced by the inner subnets $\subseteq N_i^{**} (i \in \{1, \ldots, n\})$ and B' (denoted by $\theta(N)_{B'}$) is conservative and strongly connected (if $B' = B$, $\theta(N)_{B'} = \theta(N)$)

Theorem 3. *Let N be a Root Closed net. The three following assertions are equivalent :*

- *N is deadlock free*
- *N satisfies CS property*
- *N is live*

Proof. Root Closed nets are , by construction , a subclass of Synchronized Dead Closed Systems (SDCS) [2] which are a K-systems. Therefore this equivalence holds.

Corollary 1. *Let N be a Root Closed net. If $\theta(N)_{B'}$ satisfies cs-property then N is weak compatible. This means that N is deadlock free but some interface places can be unbounded.*

Corollary 2. *Let N be a Root Closed net such that $B' = B$. If $\theta(N)$ satisfies cs property, then N is strong compatible.*

Proof. Since $B' =$ B , $\theta(N)$ is live and bounded. This means that N is deadlock free and the final marking is well a home state.

Let us consider now the root closed net N = $N_1 \oplus N_2$ of Fig.3 where N_1(on the right) is a sound root open net and N_2 is a sound ordered open net. As $\theta(N)$ satisfies the cs-property we can claim that N is strong compatible.

We consider now the closed net N obtained by composition of $N_1, N_2 and N_3$ of Fig.4 from [5]. N_1 ,N_2 and N_3 are sound ordered nets, moreover N satisfies cs-property.However as N is not a root closed net (the input interface place CMoney does not preserve the root place of transition t*) we cannot claim that N is strong compatible. In fact N is compatible but not strong compatible (t* is not live in N).

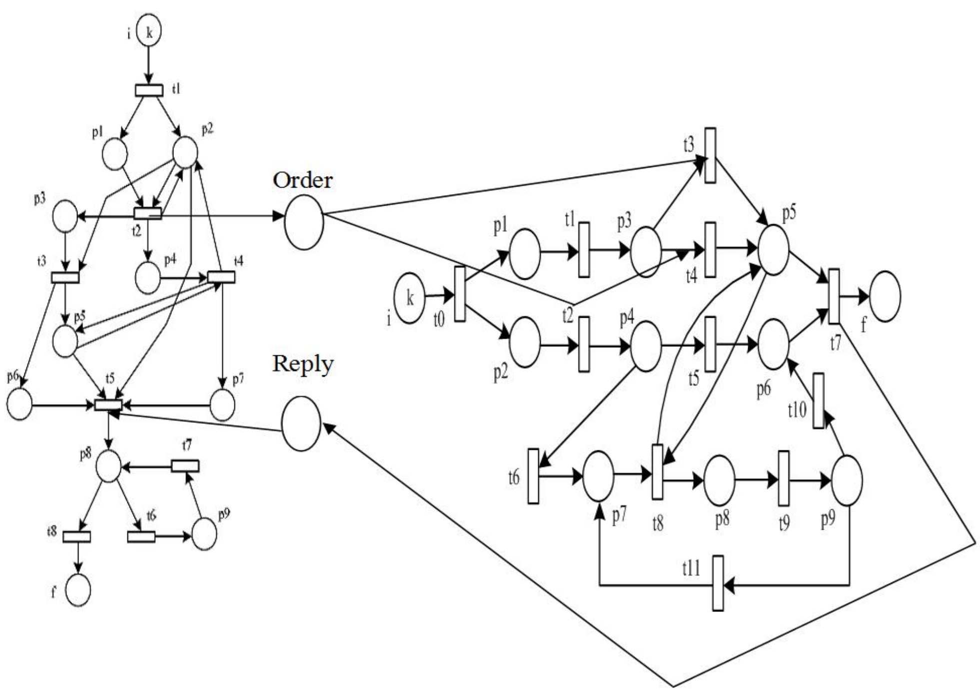


Fig. 3. A strong compatible root closed net

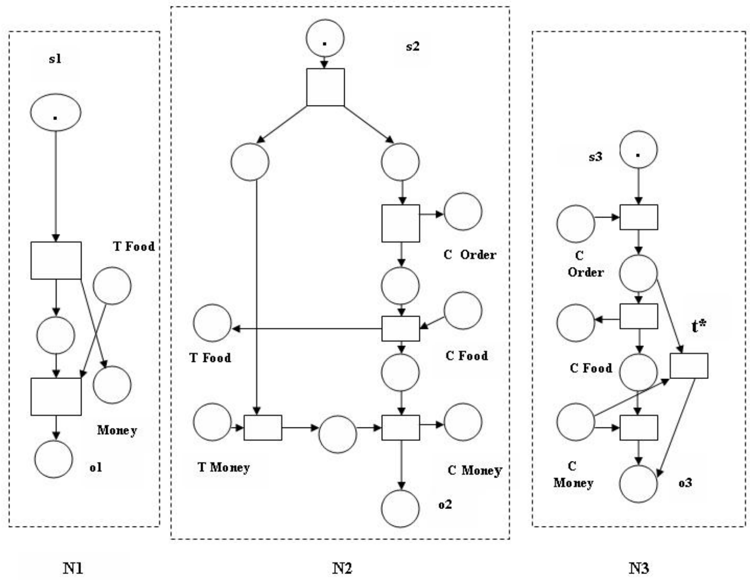


Fig. 4. A compatible (but not strong compatible)closed net

5 Conclusion

This paper presented a structural approach to verifying process interactions for coordinated web services composition. Using results of structure theory of Petri net , we have identified necessary and /or sufficient structural conditions on web services interfaces ensuring the composition compatibility.
The main contribution of this paper is to provide a structural technique to check if two or more web services are compatible and a better understanding of the incompatibility sources.
A direction for further work is to exploit these results to develop efficient solutions for the substitutability problem (i.e. the assurance that a given service can be replaced by another one as a better partner in a given composition).

References

1. Cheng Y., Wang Z.: Research on Reachability Verification of Web Service Composition. World Congress on Software Engineering. pp. 233237 (2009).
2. Barkaoui K., Couvreur J.M., Klai K.: On the equivalence between liveness and deadlock-freeness in Petri nets. Lecture Notes in Computer Science. 3536, 90 107 (2005).

3. Dun H., Xu H., Wang L.: Transformation of BPEL Processes to Petri Nets. Theoretical Aspects of Software Engineering, 2008. TASE'08. 2nd IFIP/IEEE International Symposium on. pp. 166173 (2008).
4. Hamadi R., Benatallah B.: A Petri net-based model for web service composition. Proceedings of the 14th Australasian database conference-Volume 17. p. 200 (2003).
5. Aalst, W.M., Mooij, A.J., Stahl, C., Wolf, K.: Service Interaction: Patterns, Formalization, and Analysis. Formal Methods for Web Services. p. 88 (2009).
6. Barkaoui, K.; Ben Ayed R.; Sbai, Z.: Workflow soundness verification based on structure theory of Petri nets. International Journal of Computng and Information sciences (IJCIS) Vol.5, n1, April 2007.
7. Alves, A., Arkin, A., Askary, S., Barreto, C., Bloch, B., Curbera, F., Ford, M., Goland, Y., Gu?zar, A., Kartha, N., Liu, C., Khalaf, R., Koenig, D., Marin, M., Mehta, V., Thatte, S., Rijn, D., Yendluri, P., Yiu, A.: Web Services Business Process Execution Language Version 2.0 (OASIS Standard). WS-BPEL TC OASIS (2007), http://docs.oasis-open.org/wsbpel/2.0/wsbpel-v2.0.html.
8. Lohmann, N., Massuthe, P., Stahl, C., Weinberg, D.: Analyzing interacting WS-BPEL processes using flexible model generation. Data and Knowledge Engineering 64(1), 38-54 (2008).
9. J.Zhang,J-Y.Chung,C. K.Chang,and S.Kim.WS-Net: A Petri-net based specification model for web services. Proc. ICWS,pp. 420-427.IEEE,2004.
10. K.Barkaoui et M.Minoux.A Polynomial Time Graph Algorithm to Decide Liveness of some basic classes of Bounded Petri Nets. In Application and theory of Petri Nets.LNCS 616.pp.62-75,Springer Berlin / Heidelberg.1992

Proposition of a Generic Metamodel for Interorganizational Business Processes

Khoutir Bouchbout[1], Jacky Akoka[2], Zaia Alimazighi[1]
1: Computer Science Department, USTHB University, Algiers, Algeria
2: équipe ISID, Laboratoire CEDRIC/CNAM-Paris, Paris, France
kbouchbout@gmail.com, alimazighi@wissal.dz, jacky.akoka@cnam.fr

Abstract. An Interorganizational Business Process (IOBP) is an organized group of related activities carried out by multiple organizations to accomplish a common business goal. A consequence of this is that business process modeling and design used inside an organization have to be enhanced and extended to cope with interorganizational business relationships. Modeling business processes that span multiple organizations involves new challenges, mainly the ability to cope with autonomy, privacy, heterogeneity, and the support for coordination trough mutual agreements. As a contribution to this area, this paper presents a metamodel that captures a wide range of IOBP elements.

Keywords: Interorganizational business process, business process modeling, interorganizational business process metamodel, MDA-based framework, B2B.

1 Introduction

Collaboration and coordination between companies are considered necessary in a business environment, where companies focus on their competitive advantage, perform only those functions for which they have expert skills and they complement their offering through partners and suppliers. Interorganizational business processes are the enabler of such business environments. The modeling of IOBP is a challenging task, due to the high degree of autonomy and heterogeneity of the cooperative organizations. The paper proposes an IOBP generic metamodel which depicts the nature of interaction between organizations through business processes under specific business requirements that emphasize the heterogeneity, privacy and autonomy of the participating organizations.

For this purpose, we conducted explorative research which is considered appropriate for gaining better insight and for analyzing particularities of interorganizational processes in comparison with internal processes. Having answered this question, metamodel elements should be derived that are considered necessary for metamodeling IOBP.

The remainder of the paper is structured as follows. In section 2, we present basic concepts of IOBP. Section 3 highlights the framework for IOBP design. In section 4, we discuss the related works. Section 5 describes the aspects of the proposed IOBP metamodel. Finally, the paper finishes giving a summary and an outlook to future work.

J. Barjis, M.M. Narasipuram, G. Rabadi, J. Ralyté, and P. Plebani (Eds.):
CAiSE 2010 Workshop EOMAS'10, Hammamet, Tunisia, pp. 42-56, 2010.

2 Interorganizational Business Processes Basic Concepts

A business process is a continuous series of organizational tasks, undertaken for the purpose of creating output [25]. Among the forms of information that people ordinarily want to extract from a process model are *what* is going to be done, *who* is going to do it, *when* and *where* will it be done, *how* and *why* will it be done, and *who* is dependent on its being done [7].

This section aims to present the basic definitions and concepts of IOBP modeling.

2.1 Intra-organizational versus Interorganizational Business Processes

While intra-organizational processes comprise activities executed inside one organization only, the activities comprised in an IOBP are executed by different organizations that are working together to reach a common objective. Hence, a number of particularities arise in comparison with intra-organizational business processes [12]. IOBP usually do not have a centralized control instance or process owner. Coordination between the different organizations requires an agreement on how to interact and exchange information. However, autonomy of the different parties has to be taken into account when designing IOBP.

In order to illustrate the IOBP concepts, we regularly refer to the following corporate procurement process example scenario depicted by the figure 1.

Application example: The Procurement application concerns two organizations (enterprises) – a buyer and a seller – which are collaborating and need to interlace their business processes. The Buyer sends an initial request for quote to the Seller. The Seller checks if the requested product is offered, i.e. listed in its product catalogue. If so, then the stock information is required in order to see if the product is kept in stock. If the product is out of stock, product information is needed to check if the product can be produced or not. In cases of either having the product in stock or having to produce the product, the Seller needs to calculate its price and to send back a quote to the Buyer. If the requested product is not offered by the Seller and cannot be produced, a rejection is sent back to the Buyer. In case of having received a quote for the requested product, the buyer checks if the price corresponds to the price limit set; if so, it sends a PO to the Seller. The Seller then verifies the credibility of the Buyer. If the credibility is ok, the Seller returns an order response to the Buyer.

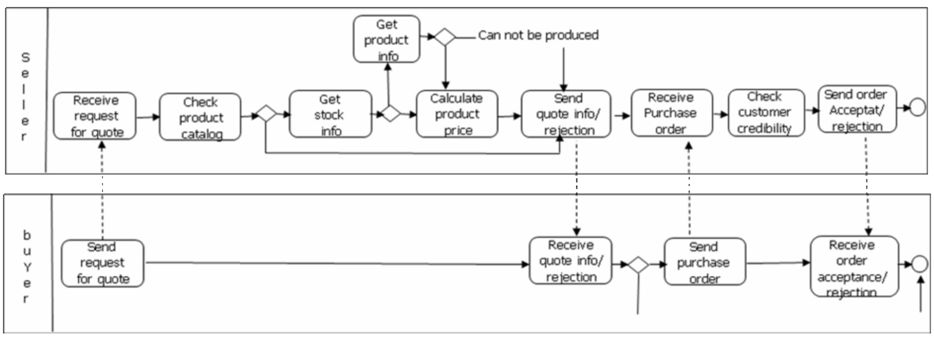


Fig.1. An example of an IOBP

In order to make IOBP work, each involved organization has to implement not only its internal processes (private processes), but also its external behavior (public processes). Hence, in IOBP modelling it is common to distinguish between internal and external activities of business processes. Adopting the approaches used by ([5],[16],[12],[19],[26],[33],[35],[36,[41]), we consider a process described either as a Private (internal or executable), Public (abstract or view), and Collaborative (cross-organizational or interorganizational) as illustrated in figure 1, figure 2, and figure 3.

• *Collaborative Business Processes* define the interactions (vertical dashed arrows in figure 1) between two or more business companies. These interactions take place between the defined public processes and are defined as a sequence of message and/or other material input/output exchange as depicted in figure 2. The collaborations between the involved parties are modeled as interaction patterns between their roles. It is shown by two or more public processes communicating with each other.

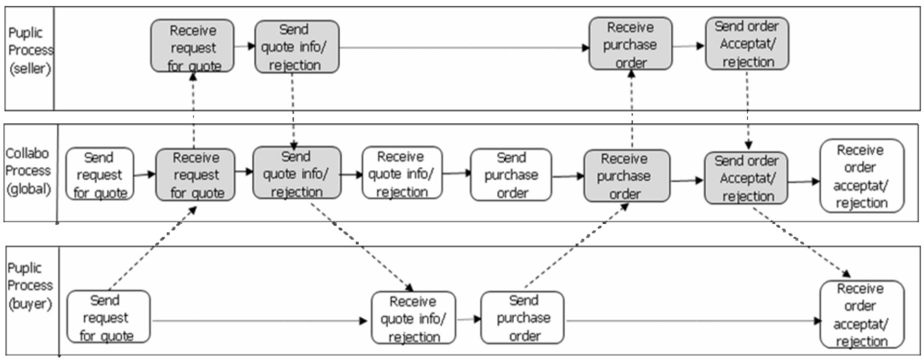


Fig.2. An example of an IOBP: Collaborative & Public processes

• *Public Process* abstracts information from one or more private processes and thus enables companies to hide critical information from unauthorized partners. It is an interface to the outside world which extracts only that kind of information which is necessary for interaction with one or more potential partners. A public process defines an external message exchange of an organization with its partners according to a message exchange protocol like PIPs (www.rosettanet.org). Thus a public process can be seen as general interaction description of one or more private processes from the perspective of one partner. Seller's public activities are represented in grey and Buyer's public activities are represented in white in figure 2.

• *Private Processes* are internal to an organization. They contain data not to be revealed by default (private activities are represented in grey and public activities are represented in white in figure 3). On private process level, organizations model their internal business processes according to a modeling approach that is most suitable for internal demands independently of the modeling methodologies used by the business partners [34]. For example a Seller wants to hide the "Check product catalog", "Get stock info", "Get product info" ,"Calculate product price", and " Check customer credibility" activities from Buyer.

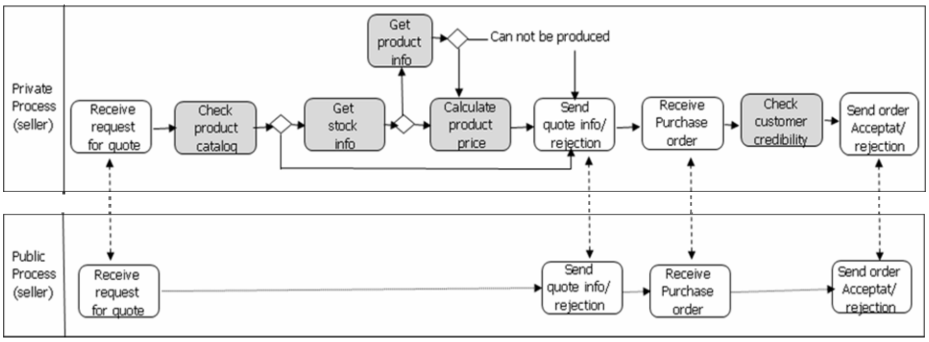


Fig.3. An example of an IOBP: Public vs. Private processes

To explain specifics of IOBP modeling, we will discuss the requirements and particularities of interorganizational business processes.

2.2 Particularities of Interorganizational Business Processes

The approaches investigated until now introduce a representation of the interorganizational business process, which uses either an existing modeling notation or its extensions. Specific artifacts are necessary for describing interorganizational business processes, among them external organizations, roles or partner types as well as messages, business documents and channels. With regard to the allocation of tasks to the actors in the interorganizational business process, partition concepts have become popular.

Important contributions to handling the specifics of IOBP come from research on workflow management, e.g. the Public-To-Private Approach ([1], [40]) and the View-based Process Model ([11], [26], [27], [33], [35], [36]). Hence and based on a literature review ([8],[19],[20],[22], [23],[41]), we deduce in the following the most important specific particularities for IOBP modeling.

1. *No central governance of the global process*. There is no entity that designs, implements, executes, and monitors the end-to-end process [24]. This requirement follows the assumption that central governance reduces the autonomy of the parties and may require visibility to details that are not necessarily visible in a purely distributed process. Then the organizations can hide their processes from other organizations. In interorganizational environments the internal business processes are one key competence of the organizations they want to preserve from the other organizations. In order to support these requirements, a flexible information hiding mechanism is required.

2. *Autonomy of business Partners*. Each partner has a full autonomy to design, implement, execute and monitor its internal processes, provided they comply with the partner's obligations toward the other partners [23]. The IOBP participants act autonomously and must coordinate themselves by means of interactions.

3. *Generation of executable processes*. The distributed execution of an IOBP starts with a common process model that all partners share and that is business oriented.

From this model every partner extracts those parts that he has to execute and augments them with arbitrary information he needs for execution [12]. Thus the used modelling language should be able to transfer the IOBP from business level into an IT-oriented workflow model on technical level like e.g. BPEL.

4. *Support of organizational units and roles*. Because different partners are involved in an IOBP, it is important to describe the organizational units with the communication and reporting relationships within the IOBP. Furthermore, the role defines the requirements profile of an organizational unit, particularly necessary for workflow applications. Thus, the modelling language should be able to describe the different organizational units and roles of the partners within the IOBP [41].

5. *Support of activity semantics.* For interoperability reasons, trading partners have to exchange electronic business documents. Since each partner has its own systems and culture, they could use different terms and metadata structures to represent their data, even when referring to the same domain of interest. Inefficiencies concerning the electronic exchange of data and information can be eliminated by the definition of central semantic and syntactic standards for exchange objects (for example business documents) as well as transfer methods (transmission medium, exchange protocols etc.) in order to achieve semantic interoperability [13].

2.3 Overview of the Current Business Process Modeling Languages

To specify IOBP, big efforts have been made during recent years and many languages have been proposed. The origins of process modeling languages are quite diverse (see Table 1). Today, there are a lot of conceptual business processes modelling languages available. This section discusses the evaluation of four well-known of them which are used generally in the intra-organizational case.

Table 1. Comparison of business processes modeling languages.

	Petri Net	EPC	UML 2 AD	BPMN 2
Issue Edition	C. A. Petri, 1962	Keller, Nüttgens & Scheer, 1992	OMG, 2004	BPMI, OMG, 2009
Perspective	IT Perspective	Business perspective	Object-Oriented perspective	IT/Business perspective
Source domain	Formal specification	Business engineering	Software engineering	Business engineering
Specification	Academic	Proprietary	Open	Open
Purpose	Analysis, verification	Description, Analysis	Description, enactment	Description, enactment

Although there is an abundance of business process modeling languages, only a few were applicable for IOBP modeling in practical cases. One major requirement of business/IT specialists in practice is that business process modeling languages should be widely used in industry and in commercial products. This is the case for EPC [32] and UML [39]. UML is of additional importance because there is a strong organization (OMG [39]) behind UML pushing it. This is also the case for BPMN [4]. Another reason of importance is the ability of formal analysis, optimization and verification, which is the case for high level Petri Nets [29]. Thus, in this section we analyze and compare the modeling languages High Level Petri Nets, EPC (ARIS), UML2 AD, and BPMN. An overview of the evaluation results can be found in Table 2. Our evaluation scale ranges from comprehensively fulfilled (depicted by +), partially fulfilled (+/-) to not fulfilled (-).

Note that we selected those four languages among many others like IDEF3, BPML, RAD, and DFD ([7],[8],[9],[24],[37]) because they either provide a set of interesting concepts and/or are supported by a prominent industrial consortium.

Despite their diversity, all the different modeling approaches have their pros and cons. However, the comparison presented in this section shows that no single language fulfills all requirements identified for specifying IOBP.

Table 2. Comparison of business processes modeling languages (cont.).

	Petri Net	EPC	UML 2 AD	BPMN 2
Support of private, public, IO Processes	+/-	+/-	+/-	+
Representation power	+/-	+/-	+/-	+
Support of analysis control	+	+/-	-	-
Interaction patterns protocols	+	+/-	+/-	+/-
Semantic annotations	-	+	+/-	+
Support of involved role	+/-	+	-	+/-
Tool support notation	+/-	+	+/-	+
Mapping to exec. language	+ (BPEL)	+/-(BPEL)	+(BPEL)	+(BPEL)

3 Framework for Interorganizational Business Process Design

To meet the IOBP particularities we have presented above, we propose a novel approach for building an IOBP based on a Model-Driven Architecture (MDA) as illustrated in figure 4. The framework is characterized by a set of transformations/mappings (horizontal and vertical) at and across different layers.

The MDA is a framework for software development driven by the Object Management Group (www.omg.org). The following three models are at the core of the MDA: (1) *Computation Independent Model (CIM):* This is the most abstract model within MDA independent of computational technology; (2) *Platform Independent Model (PIM):* This model is defined at a high level of abstraction; it is independent of any implementation technology. It describes a software system that supports some business; (3) *Platform Specific Model (PSM):* A PIM is transformed into a PSM for each specific technology platform. Processes at PIM level shall be described in such a way, that they can be transformed to process execution languages on PSM level.

We note that the term of *metamodel* is put in relation with the OMG's MOF (Meta Object Facilities) [28], which is an abstract framework, and a four-layered architecture for defining and managing metamodels, neutral of any technology ([3], *[16] [18])*. Metamodels are simply referred to as being just "models of models" [13]. While a *model* is an abstraction of phenomena in the real world, a *metamodel* is an abstraction of the model itself. A Metamodel comprises an explicit description (formalized specification) of constructs, rules and notation for building domain-specific models.

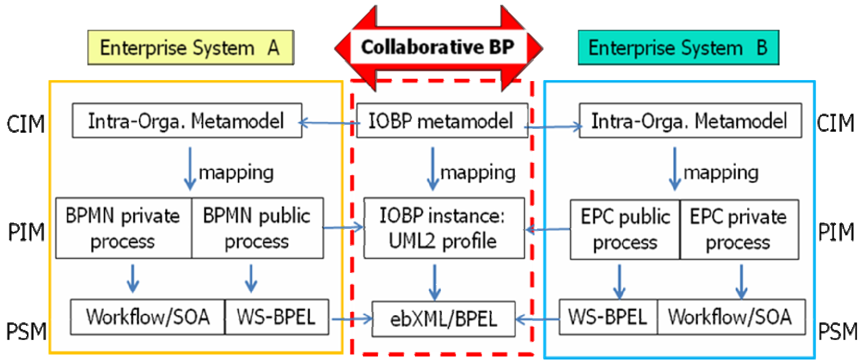


Fig.4. MDA-based framework for IOBP design

The vertical dimension distinguishes the different layers of abstraction applied in MDA and the horizontal dimension represents the collaborative modelling between two enterprises A and B. Business process models of enterprise A and B have to be shared at different level of abstraction in order to agree on and develop IOBP. The gaps between these abstraction levels are overcome by vertical transformations like presented in [19]. We assume that enterprise A and B use different business process modelling tools and languages at the PIM/PSM MDA layers. To develop IOBP both enterprises have to provide public parts of their models as basis for discussion for collaborative modeling. The vertical transformation in the downward direction corresponds to process automation approaches where conceptual models are transformed to executable processes. Both enterprises have to exchange at least parts of their models as a basis for collaborative modeling (UML 2 [39], ebXML [41]). Hence, models of enterprise A (BPMN [4]) and B (EPC [32]) are shared at PIM layer.

4 Description and Evaluation of the Related Works

Before we present our approach for IOBP metamodeling we will briefly refer to some related work in the following propositions done in the field of Workflows/Business Process metamodels.

List et al. [22] developed a generic metamodel composed of 5 contexts. They are inspired from the work of [7]. On a high level, this metamodel addresses the following views: Business Process Context Perspective, Behavioral Perspective, Functional Perspective, Informational Perspective, and Organizational Perspective. The functional perspective represents what process elements are being performed, and what flows of informational entities, are relevant to these process elements. The behavioral perspective basically describes the order in which the different activities are executed. The organizational perspective describes the organization structure and, in particular, the resources and in which way these are involved in the BP. The informational perspective describes the information that is involved in a BP, how it is represented, and how it is propagated among the different activities. The business process context perspective captures important business process context information like process goals and performance measures or process type. However, this metamodel is not well adapted to represent interorganizational relationships.

In the proposition of Morley [25], the metamodel is based on two assumptions. First, the underlying modelling approach is a top-down one. A Process is initially defined by the Purpose assigned to it. It can be described at several levels of granularity, but the last level is the only one to be detailed. This is particularly useful when drawing cartography of all the enterprise business processes. Then, the notion of Activity is a central concept, due to the influence of the standard process definitions dedicated to enterprise modelling. The starting point is to model the Activities and then to define the suitable organization with Roles and Actors. However, this metamodel is not well adapted to represent interorganizational relationships.

Kradolfer [17] develops a workflow metamodel that allows defining the functional/structural, informational, behavioral, and organizational aspects of workflows. The workflow metamodel is modular in the sense that the various elements (workflows, organizational entities, etc.) can be specified independently of each other, and in that no assumptions are made in which context the elements are going to be used. For instance, activity assignment, control and data flow are not specified with the activities themselves, but only when the activities are used within a workflow. The metamodel is activity-centric in that it allows to “think” in terms of activities/workflows and their results instead of states or transitions. However, this metamodel lacks the representation of some IOBP elements (private and public business processes).

Saidani and Nurcan ([30], [31]) provide a start points for the definition of a methodology allowing the design of adaptive and flexible BP metamodels according to the situation at hand. They have introduced the concepts of BPM-chunk and business method. They promote the fact that the final business process model has to be created from the set of proposed chunks in order to suit to a particular situation. Their approach aims to make easier the definition of flexible and customized meta-models. However, this metamodel do not consistently support interorganizational model requirements and concepts.

5 The Proposed IOBP Generic Metamodel

There is a lot of work done on the definition of intra-organizational business process metamodel ([2],[3],[5],[10],[12],[16],[17],[22],[25],[30],[34]), but it misses some research clearly addressing the case of the interorganizational business process metamodel. Relying in the approaches of these metamodels, we extend and adapt them into one combined high level generic metamodel addressing all the requirements of the IOBP seen before.

For this aim we structure our metamodel into four aspects according to the metamodel developed by Curtis et al. [7]: *functional, behavioral, organizational, and informational*. Besides the four business process aspects, there are further non-functional requirements a business process metamodel should fulfill: *enactability, ease of use, correctness criteria, evolution, and reuse.*

For readability reasons, we display the most important concepts in each of the four aspects separately in a UML class diagram.

- *Functional aspects: What has to be performed?*

The functional aspects of the metamodel are shown in figure 5. The concept of activity is one of the core concerns of every metamodel we studied. To enable an exchange of process data using IOBP, information might be hidden via "private process", "public process", and "collaborative process" process elements, which hide critical private process data.

Events are things that "happen" during the course of a business process. There exist three types of events: *interrupt*, *temporal*, and *trigger*. Examples of these events include change in delivery date, change in price, etc.

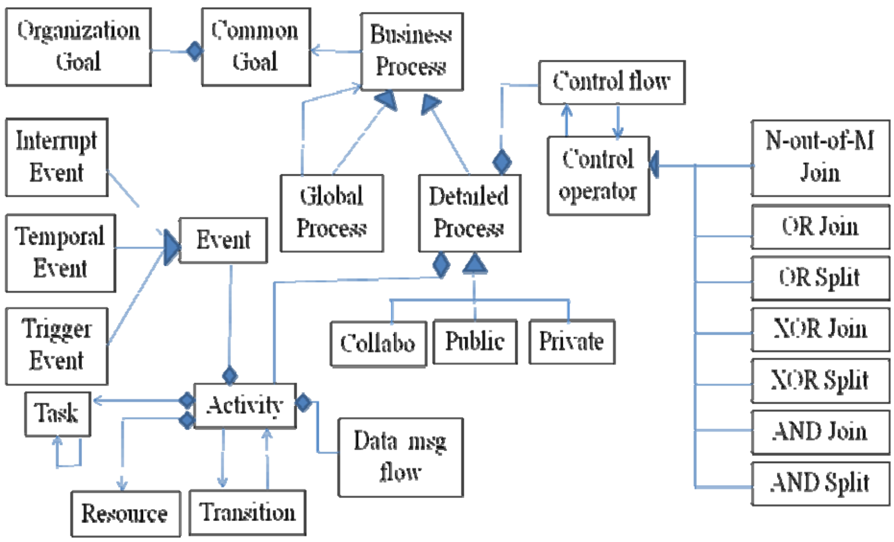


Fig. 5. Functional & Behavioral aspects metamodel

The figure 6 depicts the application of the metamodel functional aspects to the purchase order processing IOBP example seen before in the section 2.1.

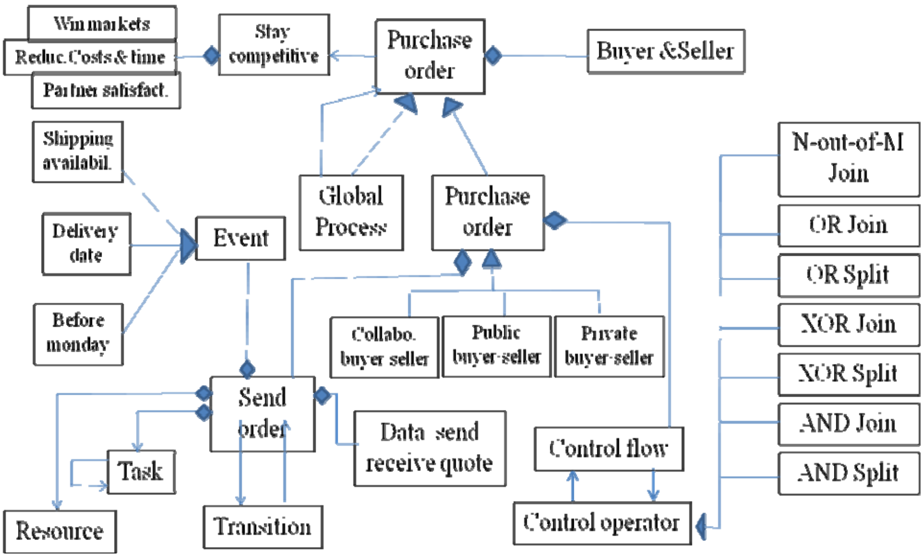


Fig. 6. Example of Functional & Behavioral aspects metamodel

- *Behavioral aspects: How is produced? (Control flow, data flow,, rules)*

The behavioral aspects of the metamodel are shown in figure 5 (analogue to [22]). Specification of control flow is essential in IOBP for the coordination of business process participants. Our metamodel supports the basic (sequence, branch) structures in order to be programmatically complete. Activities and events are connected by sequence flows indicating the order in which activities will be performed or events occur in a business process. Conditional expressions and various split and join restrictions are provided for advanced branching and synchronization patterns.

- *Organizational aspects: Who does it? (Stakeholder, role, and organizational unit)*

The organizational aspects of the metamodel are shown in figure 7. We can distinguish two major categories of "process *stakeholder*". In the first one the stakeholder is concrete. It may be a person, a computer program, a department, a position in the enterprise but it is an entity which exists apart from the process. In the second category the stakeholder is abstract. It defines a role which is played in the process and covers a set of properties (skills, capabilities, degree of responsibility …) which may be expected from the concrete stakeholder which will be assigned to this role. Hence, the modelling of IOBP requires an additional role model different to the internal role model. It should allow specifying the role of the organization as a whole in a public business process. A "*collaboration role*" defines the observable behavior that a party exhibits when collaborating with other parties in the public process.

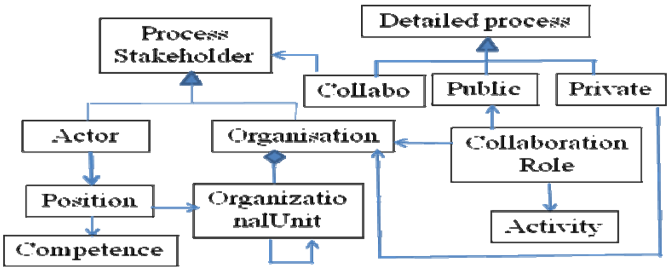


Fig. 7. Organizational aspects metamodel

The figure 8 depicts the application of the metamodel organizational aspects to the purchase order processing IOBP example seen before in the section 2.1.

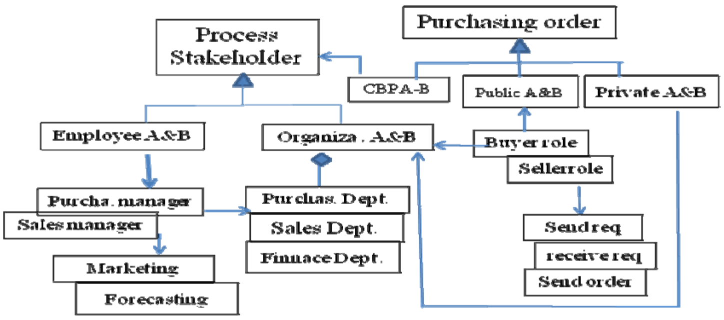


Fig. 8. Example of organizational aspects metamodel

- *Informational aspects: What is produced/exchanged? (physical resource, business document, service, software application, information object)*

The informational aspects of the metamodel are shown in figure 9. There may be lot of things behind the resource concept. On one hand resource artifacts are considered to be pieces of information. On the other hand they are concrete products like material, service or information. The resource may be of different nature according to the nature of the field covered by the metamodel. Some focuses on software process and others on manufacturing or service supplying processes.

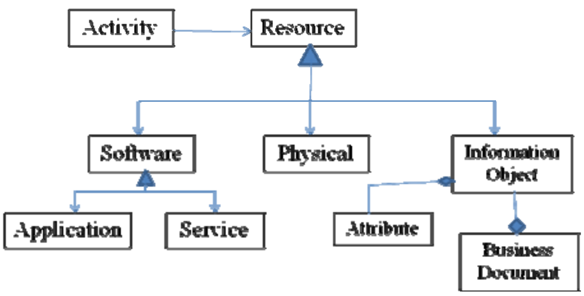


Fig. 9. Informational aspects metamodel

In other words, informational aspects represent elements describing information, material or other artefacts that are objects used by the process activities, e.g. Business documents, material that is to be sent, money that is to be received, etc. This is inspired by the workflow data patterns as well as by the input/output view of ARIS [22]. Specification of data flow must additionally consider autonomy and privacy of organizations.

The figure 10 depicts the application of the metamodel informational aspects to the purchase order processing IOBP example seen before in the section 2.1.

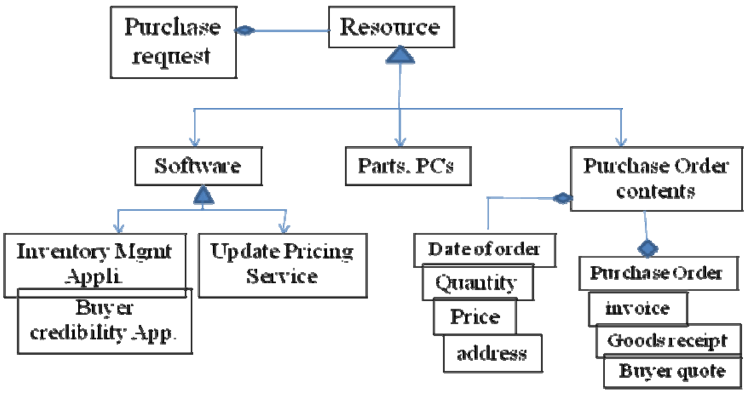


Fig. 10. Example of Informational aspects metamodel

6 Conclusion and Future Research

The increasing interest in process engineering and application integration has resulted in the appearance of various intensive works related to business process metamodeling both in academia and in the industry. The importance of IOBP has been widely recognized, leading to a variety of approaches and proposed solutions to their design and implementation. To describe and analyze existing approaches to model business processes we first described requirements distinct for interorganizational scenarios. For the representation of the IOBP elements the approaches of the intra-organizational business process modeling languages like EPC, BPMN, and UML 2 AD were adapted and extended because they do not address conveniently the particularities of the interorganizational business process. So, we developed an IOBP independent generic metamodel common to these languages which ensures the best suitability to model IOBP.

Modelling IOBP requires specific constructs and methodologies, and requires a high-level model and the corresponding executable one for exchanging and merging behaviors, resources and activities. Our current research activities focus on employing the MDA approach such that, based on a platform independent model of an IOBP, it is possible to automatically derive business process specifications expressed in the specification languages best suited for any of the different activities.

The developed generic IOBP metamodel provides the capability to represent and model business processes independent of notation or methodology, thus bringing these different approaches together into a cohesive capability.

As further work, we will validate the metamodel by instantiating it with a case study example in order to verify the completeness of the proposed concepts, then completed it with the necessary transformations to the involved business process models (EPC, UML2 AD, and BPMN) according to the MDA approach.

Acknowledgment

The work published in this paper was achieved within the ISID research team of CEDRIC/CNAM-Paris, France that the authors want to thank all members for their valuable help.

References

1. Aalst Van der W.M.P. and Weske M.: The P2P approach to Interorganizational Workflows. In K.R. Dittrich, A. Geppert, and M.C. Norrie, editors, Proceedings of CAiSE'01, LNCS Vol. 2068, 140--156. Springer-Verlag, Berlin, (2001).
2. Akoka J., Lang D.: Places de marché électroniques et reconfiguration des processus interorganisationnels, Actes du 7ème colloque de l’Association Information & Management, Hammamet, Tunisie, (2002).
3. Breton Erwan, Bézivin Jean: An Overview of Industrial Process Meta-Models, ICSSEA’2000, 14 December 2000, Paris, France, (2000).
4. BPMN Business Process Modeling Notation specification: OMG Final Adopted Specification, OMG, February 2006, URL: http://www.bpmn.org/OMGFinalAdoptedBPMNl-0Spec2006-02-01.pdf, (2006)
5. Chebbi Issam, Dustdar Schahram, Tata Samir: The view-based approach to dynamic inter-organizational workflow cooperation, Data & Knowledge Engineering, Vol. 56, 139--173 (2006).
6. Chiu D. K. W., Karlapalem K., Li Q., and Kafeza E.: Workflow View Based E-Contracts in a Cross-Organizational E-Services Environment, Distributed and Parallel Databases, vol. 12, pp. 193--216, (2002).
7. Curtis Bill, Kellner Marc I. and Over Jim: Process Modeling, Communications of ACM, Vol. 35(9), (1992).
8. Dorn J., Grün C., Werthner H., and Zapletal M.: A Survey of B2B Methodologies and Technologies: From Business Models towards Deployment Artifacts. Proceedings HICSS’07, (2007).
9. Freiheit, J.; Matheis, T., Ziemann, J.: Definition of static and dynamic models of collaborative workflow interoperability. Deliverable D4.1, R4eGov – Towards e-Administration in the large. Proceedings IST’2004, (2004).
10. Ghannouchi Sonia Ayachi, Jamel Menzli Leila: Proposition et expérimentation d’un métamodèle pour la réingénierie des processus métier, Ingénierie des Systèmes d’Informations, Vol. 13(3), 105--129(2008).
11. Grefen Paul, Ludwig Heiko, and Angelov Samuil: A three-level framework for process and data management of complex e-services, International Journal of Cooperative Information Systems, Vol. 12, No. 4, 487--531 (2003).

12. Greiner, U., Lippe, S., Kahl, T., Ziemann, J., Jäkel, F.W.: Designing and implementing cross-organizational business processes - description and application of a modelling frame-work. In Proceedings I-ESA'2006, (2006).
13. Höfferer, Peter: Achieving business process model interoperability using metamodels and ontologies, Proceedings ECIS'2007, (2007).
14. Huemer C., Liegl P., Schuster R., Werthner H., and Zapletal M.: Inter-organizational Systems: From Business Values over Business Processes to Deployment. Proceedings DEST'2008, (2008).
15. Koehler, J., Hauser, R., Kapoor, S., Wu, F.Y. & Kumaran, S.: A Model-Driven Transformation Method, Proceedings EDOC'2003, p. 186, (2003).
16. Korherr Birgit, List Beate: Extending the EPC and the BPMN with Business Process Goals and Performance Measures. 9th International Conference on Enterprise Information Systems (ICEIS07), Portugal, ACM Press, (2007).
17. Kradolfer Markus: A workflow metamodel supporting dynamic reuse-based model evolution, PhD thesis, University of Zurich, Switzerland, (2000).
18. Lazarte Ivanna M., Chiotti Omar and Villarreal Pablo D.: Transforming Collaborative Process Models into Interface Process Models by Applying an MDA Approach, AIS Transactions on Enterprise Systems, Vol. 2, 13--23(2009).
19. Legner Christine, Vogel Tobias, Löhe Jan, Mayerl Christian: Transforming Inter-Organizational Business Processes into Service-Oriented Architectures. Method and Application, Proceedings of KiVS 2007, Bern, Switzerland, (2008).
20. Legner Christine, Wende Kristin: The Challenges Of Inter-Organizational Business Process Design – A Research Agenda, University of Saint Gallen, Switzerland,n (2007).
21. Lin, F-R., Yang M-C., and Pai, Y-H.: A generic structure for business process modeling, Business Process Management Journal, Vol. 8 No. 1, 19--41, (2002).
22. List Beate, Korherr Birgit: An Evaluation of Conceptual Business Process Modelling Languages, Proceedings of the 2006 ACM Symposium on Applied Computing (SAC'06), Dijon, France, ACM Press, (2006).
23. Medjahed, B., Benatallah, B., Bouguettaya, A., Elmagarmid, A.: Business-to-business interactions issues. The VLDB Journal, Vol. 12, 59--85, (2003).
24. Mendling, J., Nüttgens, M. and Neumann, G.: A Comparison of XML Interchange Formats for Business Process Modelling. In F. Feltz, A. Oberweis and B., Proceedings of EMISA 2004, LNI 56, 129--140, (2004).
25. Morley Chantal : Un cadre unificateur pour la représentation des processus, Pre-ICIS. Cahiers de recherche, INT Management, Ivry, France, (2005).
26. Norta Alexis, Grefen Paul: Discovering Patterns for Inter-Organizational Business Collaboration; International Journal of Cooperative Information Systems; Vol. 16, No. 3/4, 507—544 (2007).
27. Nurcan Selmin, Etien Anne, Kaabi Rim, Zoukar Iyad, and Rolland Colette: A strategy driven business process modelling approach. Business Process Management Journal, 11(6), 628--649 (2005).
28. OMG: Meta Object Facility (MOF) Core Specification. Version 2.0 (2006).
29. Petri, C.A.: Kommunikation mit Automaten. Dissertation, Technische Universität Darmstadt. Institut für Instrumentelle Mathematik, (1962).
30. Saidani Oumaima, Nurcan Selmin: Towards Situational Business Process Meta-Modelling, Proceedings of CAISE'08, 93-96 (2008).
31. Saidani Oumaima, Nurcan Selmin: A Role-based approach for modeling flexible business processes, Proceedings of CAISE/BPMDS'06, (2006).
32. Scheer, A.-W.; *ARIS Business Process Modeling*. Springer Verlag, (1999).
33. Schulz K. A. and Orlowska M. E.: Architectual Issues for Cross-Organisational B2B Interactions, in 21st International Conference on Distributed Computing Systems Workshops (ICDCSW '01), pp. 79--87, (2001).

34. Seel Christian, Vanderhaeghen Dominik: Meta-Model based Extensions of the EPC for Inter-Organisational Process Modelling, Proceedings "EPK 2005", Hamburg, Germany, (2005).
35. Shan Z., Chiu D.K.W., Li Q.: Systematic Interaction Management in a Workflow View Based Business-to-business Process Engine, Proceedings HICSS'2005, (2005).
36. Shen, M., & Liu, D.R.: Coordinating interorganizational workflows based on process-views. Proceedings of the 12th International Conference on Database and Expert Systems Applications, 274-283, (2001).
37. Söderström E, Andersson B, Johannesson P, Perjons E, Wangler B.: Towards a Framework for Comparing Process Languages. Proc. of the 14th International Conference on Advanced Information Systems Engineering; pp. 600--611 (2002)
38. Strahonja Vjeran: The Evaluation Criteria of Workflow Metamodels, Proceedings of the ITI 2007 29th Int. Conf. on Information Technology Interfaces, June 25-28, 2007, Cavtat, Croatia, (2007).
39. Unified Modeling Language Specification, Object Management Group (OMG), version .2.0. http://www.omg.org.
40. Wombacher, A.; Aberer, K.: Requirements for Workflow Modeling in P2P-Workflows derived from Collaboration Establishment; in Proc. 1st Intl. Workshop on Business Process Integration and Management (BPIM 04), Zaragoza, Spain, IEEE Computer Society Press, 1036--1041 (2004).
41. Ziemann Jorg, Matheis Thomas, Freiheit Jorn: Modelling of Cross-Organizational Business Processes - Current Methods and Standards, Proceedings EMISA'2007, LNI, Vol.119, 87--100 (2007).

Reengineering the Learning Process in a Transport Company

Pieter de Vries[1], Heide Lukosch[1]

[1] Delft University of Technology, Jaffalaan 5,
2628 BX, Delft, Netherlands
{Pieter de Vries, Heide Lukosch, pieter.devries@tudelft.nl)

Abstract Swift changes in the regulations and the organization of the transport sector make innovation an absolute necessity. A company's ability to cope with these changes depends largely on the learning capacity of the organization. Therefore in the transport case described in this paper the company decided that reengineering of the learning strategy was needed to handle the needs of their mobile workforce also in the long term. The strategy is based on a multi dimensional design using sound educational concepts and recent insights in the changing knowledge landscape in combination with Learning 2.0 elements to handle the ever changing learning demands. The multi faceted solution is an integrated cross functional business portal with information, learning and performance improvement as the essential processes in the design. The portal is online and acceptance is clearly growing. Process monitoring and evaluation supply data on the usage and success of this integrated approach.

Keywords: Reengineering, Learning strategy, Transport, Learning 2.0, Business portal

1 Introduction

The transport world in general shows a mastery of logistic processes including the coordinated transport of goods, transport chains, freight handling, time windows, and planning arrangements. This practice is supported by well organized business processes in combination with dedicated transport software to ensure a minimum of errors and a maximum of efficiency. Rapid innovations in the industry and frequent changes in legal issues and governmental regulations endanger the conservation and improvement of that mastery when modernization becomes a burden. One of the actions taken by Van der Wal International Transport, the case discussed here, was to improve the learning capacity of the company to allow for adequate response to the demands for change and innovation. An important issue is the mobility of the workforce of which the majority are truck drivers, who are away from home and office most of their time. A main challenge is to involve this target group in the

J. Barjis, M.M. Narasipuram, G. Rabadi, J. Ralyté, and P. Plebani (Eds.):
CAiSE 2010 Workshop EOMAS'10, Hammamet, Tunisia, pp. 57-68, 2010.

information, learning and performance improvement processes that largely define the capacity for innovation in the organization.

Van der Wal international transport is a family owned business founded in Utrecht in the Netherlands in 1924 with presently about 450 employees. The company has a firm base in the Netherlands with branches in Belgium, Poland, Russia, Azerbaijan and Kazakhstan. The company has received several rewards as a fair and green enterprise that operates with respect for people and the environment.

The aspiration of the company is to better integrate information and learning into the business processes to enhance the learning capacity of the organization. This paper covers the analysis of the actual and the desired situation, the concept development based on educational theories and concepts, the implementation and the experiences so far with an innovative information and learning approach for these mobile workers in the transport sector using web 2.0 functionalities.

2 The Research Flow

The overall research plan was based on the Corporate Learning Strategy model, the CLS model [1] [2]. This model is marked by a holistic approach and subsequent phases for the development of a sustainable learning strategy (see figure 1). This allows the development of a clear picture of the actual situation at both the strategic, tactical and operational level of an organization.

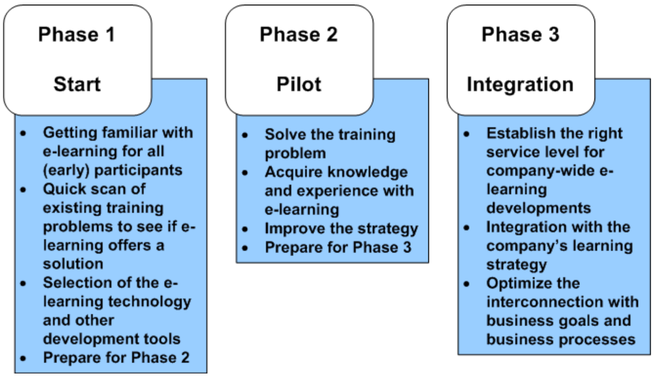


Fig. 1. The Corporate (e-)Learning Strategy Model [3]

The exploration using the CLS analysis framework is structured around a quick scan on five major issues: organization, process, and content, infrastructures and business model. Phase two of the model is the pilot phase, in which solutions are tested in practice and in phase three the focus is on the structural integration of these solutions in the actual business processes of the organization. This paper is about the experiences in the first two phases of the development and the transition process of the innovation into phase three.

2.1 Analysis of Learning Requirements

Although information management, training and learning had never been a priority, management of the transport organization was convinced that these issues had become essential and that there was a strong need for a strategic solution using good theory, good structured processes, a realistic timeline and a sound business plan to recoup the investment. An important goal of information and learning in the eyes of management was to familiarize the employees with the companies' vision, mission and values and have a tool to cope in a sustainable way with new development, change and innovation.

Dominant in the overall analysis is the mobility of the employees, which makes it almost impossible for them to attend courses, workshops or information sessions. Formal training and learning events that are very much time and place restricted have shown to be of little value for the mobile learner, as we define the type of learners in this case. For this learner, other ways must be found to reinforce the information and learning processes [3]. From the analysis it became also clear that learning in a transport company has a multi-dimensional character which can not be served adequately with formal and traditional learning models. The day to day reality comprises a broad spectrum of information exchange associated with learning and communication processes very much linked to workplace related needs for performance improvement.

On the basis of the analysis the following three dimensions of requirements were identified: (1) Information: An information system as a low threshold resource of news and other relevant resources; the need to better connect with the mobile colleagues to reduce isolation; the need to engage all employees in the company's values and vision viewpoint; an information resource as reference for daily tasks (drivers handbook). (2) Learning: A sustainable alternative learning concept to replace ad hoc and incidental learning practices; a concept which is attractive for the target group that arouses interest and fits into the daily working schedule; the 'learning' activity can be monitored for supervision, support and planning, adding transparency to the mostly informal learning process. (3) Performance: Plain job descriptions which allow for KPI's (Key Performance Indicator) reference to establish a clear picture of the performance requirements; the transparency of the job descriptions should allow employees to continuously be aware of the requirements they must comply. These requirements must be translated into clear operational instructions including learning activities when needed. In this way the employee receives guidance which may serve as a structuring element for learning.

2.2 Learning concepts to match requirements

The three dimensions of requirements clarified the goals for the new learning strategy and emphasized the fact that the solution would need to have a multidimensional character. The first step made was to match theoretical concepts with these requirements to be able to cope with the multitude of issues and establish a firm theoretical basis to work with.

Social constructivism. Training Models based on socio-constructivist principles seem to work well for informal, practical, ad-hoc-like, spontaneous learning [4]. Core of the constructivist concept is the premise that we all have our own 'mental model' which is developed in interaction with the world around us [5], [6]. The main principles are that each person is unique with regard to knowledge and experience and that people primarily learn by actively trying and learning always is related to a particular social context. Therefore the learner plays a central role in the development of learning activities with the sole aim to make learning more efficient. As a consequence motivation becomes an important stimulating factor and the individual approach to learning relates very much to the knowledge and experience acquired in the past. The relevance of this concept for the transport case is that the learning demand of the employer is an important trigger for learning to take place.

Connectivisme. The constructivist approach is a strong advocate for the inclusion of informal learning as part of the learning strategy. This approach is supplemented with the concept of Connectivisme that focuses on the changes taking place in society when it comes to knowledge and learning [7]. The knowledge landscape is changing by the array of new information media, like You Tube, Facebook, Wikis, which emerge in a rather quick pace. In this new situation learning is the ability to connect to different 'nodes' of knowledge, which are spread over a network of data, information and people and is called ‘connected knowledge' [8]. This observation is relevant in our transport case, because the lorry drivers are increasingly using mobile phones and the Internet in their day to day working environment, which affects their communication patterns, information acquisition, their learning and the development of their social network.

Levels of experience. Assuming that people are unique learners, as conceived in the social constructivist model, they also have a different experience level to be taken into account [9], [10], [11]. Three levels are distinguished: early development, competent and experienced, expert (see figure 2). A newcomer needs good formal and

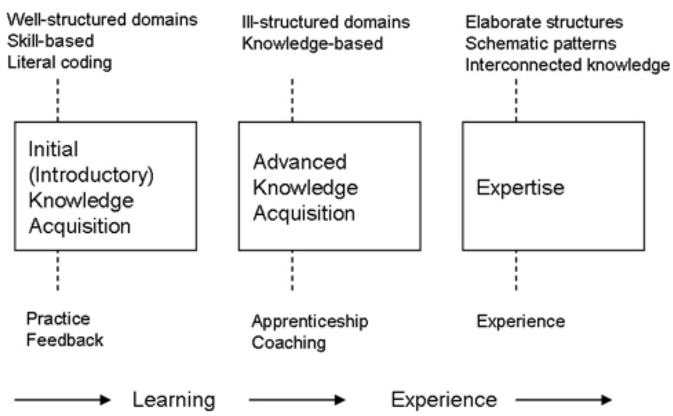


Fig. 2. Levels of Experience [10]

structured learning to acquire basic knowledge and skills in an appropriate fashion. More experienced employees with more knowledge and skills are better served with a largely informal learning situation that better fits the rather ad hoc learning needs, so common for workplace related learning. For people with expert knowledge and experiences, formal and structured learning can even become counter-productive, because it does not fit the very personal learning needs. So in the case of the transport organization this means that a well structured on-boarding program might work very well for newcomers, but for the experienced workers one need to rely much more on the self-initiative of the worker to support his or her learning process.

Formal and informal. The distinction between formal and informal is not based on a strict separation between different learning activities, but the notion that formal education is traditional, class and curriculum bound and informal learning is a predominantly social activity consisting of a mix of learning related actions. In reality we learn all the time, under different circumstances, but apparently more informal than formal. Cross [12] states that 80% of the knowledge we need to do our work well is obtained through informal processes, the remaining 20% through formal channels. Obviously this is not yet common ground for learning professionals, because about 80% of the training budget is spent on formal business training efforts. This can be qualified as an over-investment which shows that many professionals are unaware of the possibilities offered by more natural and informal learning to increase the learning capacity of an organization. Cross [12] describes this phenomenon as the 'Spending & Outcome' paradox. The notion of formal and informal learning very much relates to the concepts of social constructivism, connectivism and levels of experience and clarifies that in daily practice the company can profit much more from the support of informal learning than continue to invest in traditional learning practices which increasingly show to be ineffective.

Microtraining. Microtraining is a mechanism to support informal learning with the learning demand of the worker as the main point of reference. This mechanism is based on a development scheme to organize self learning or group learning, to develop learning materials, e-learning and other learning activities, while focusing on the applicability of what is learned and needed in the workplace [13]. It supports the development of short, customized learning sessions for different groups of employees with the focus on workplace related learning demands (see figure 3). It is in fact a framework for types of self promoted learning support to be used by the employee or manager with a high level of practical relevance.

A Microtraining arrangement comprises a time span of 15-20 minutes for each learning occasion, which can activate and maintain learning processes for a longer period if they are bundled up in series, being face-to-face, online or in an e-learning situation. Each session starts actively, followed by a demonstration or exercise, feedback or short discussion, and ends with directions for further development and a brief preview of the next sessions.

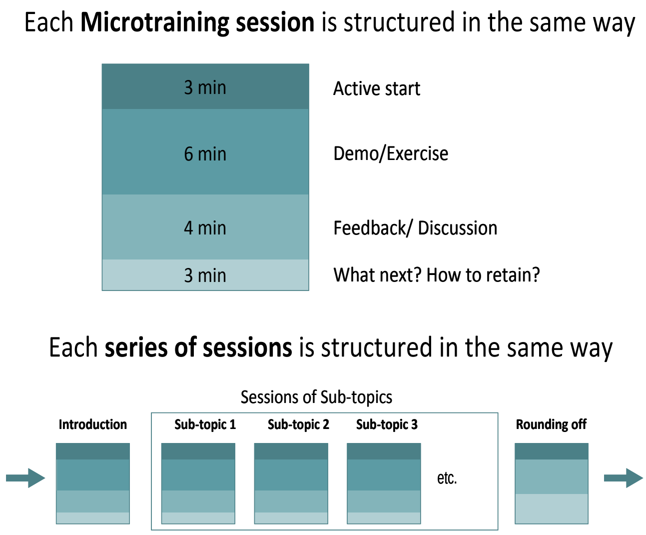


Fig. 3. The Microtraining Workflow [14]

'The Strength of weak Ties' [15] Granovetter [15] clarifies in his theory that weak ties (see figure 4) play an important role in the functioning of social networks and should be used to optimize the links between employees in an organization. Nowadays this theory again receives much attention in relation to the emerging use of social software. This discussion is relevant as the use of mobile devices, laptops and internet related applications like email and social software like You tube and Facebook or just a blog, enable the weak links (drivers) to be better informed and have more opportunities to communicate, which can strengthen the social network of the company.

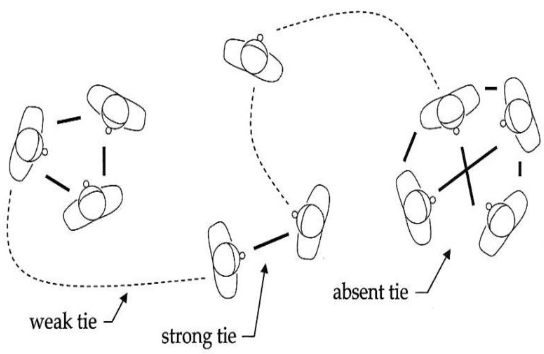


Fig. 4. A Social Network of Strong and Weak Ties [16]

Lorry drivers as mobile workers have a 'weak tie' with colleagues, unless they are friends. If so then they are part of a strong tie group that tends to become isolated if they do not interact with 'weak ties'. In the context of a strong tie group, learning is not well served, because of the internally directed focus of the group.

3. Learning 2.0 Ingredients

The analysis of the actual and desired situation in the company clarified the need of an integrated information and learning strategy that comprise three dimensions as there are information; learning and performance. As a first step we matched theoretical concepts and visions with these requirements to establish a theoretical basis, which was used to develop solutions. Evident in the solutions approach is that learning comprises all interactions with information and as a consequence deals with the capacity of the employees to manage information in a meaningful way. The following section illustrates the main actions taken to integrate the different functionalities into a complete cross functional process. For phase 1 and 2 of the project, five work packages were defined: 1. Learning policy, 2. Communications Requirements, 3. Online Newsletter, 4. Micro Training and 5. Online courses. Each of the items is discussed below.

Integration of learning policy in strategic business plans. In the initial phase consensus was achieved with management about the importance of learning and the need to make this an integral part of the company’s strategic business plan. The learning chapter should provide the link between the business plan and the curriculum. This learning plan should contain the vision, the approach and tools for the business plan as prepared for the holding and the different companies, the departments, the job positions and the related tasks. In the framework of this work package basic job profiles and annual job profiles were developed. The basic profile comprises the minimum knowledge, skills and experience levels. The annual profiles relate to instant changes in work practices and depending on the needs and the time period, these are added to the basic profile.

Communications Requirements. In the initial phase of the project the existing information resources were complemented through interviews and a survey focusing on the communications needs of employees and management. The aim of the study was to determine what the perception was of the employees and their wishes. The survey was distributed among all employees with a reaction level of 65%. In addition a selection of 13 employees were interviewed, taking into account age, group representation (drivers, planners, administration, management) or department, talkative or not, men and women, ICT minded or not. From the interviews and survey it is shown that the employees have a strong commitment towards the company. Generally there is a strong need for news about the business (state of affairs) with the emphasis on reliable and inclusive information. Especially the drivers are in need of good professional information. Their main sources of information are colleagues (85%). The monthly paper newsletter was considered an important good quality resource. Online information had not arrived yet as an alternative. To foster understanding of the situation a workshop was held with the management team on the outcome of the surveys and interviews to clarify the views on the role of communications in and for the business. Managers underlined that communication should be open, honest and inspiring, respectful, interactive and available for everyone, requiring good solutions for 2-way communication, communication rules, clear (communication) tasks, responsibilities, sufficient support and commitment throughout the organization.

Online newsletter as the main entry for the Business portal. Since the paper newsletter was highly appreciated, it was decided to extend this tool to an electronic online newsletter and pilot this product over a period of time. The aim of the pilot was to find out which functionalities were needed to establish such a 24/7 news service. An online version requires a different strategy concerning content selection, production process, format, design, navigation, storage, and offers other features, such as the comments, which can be used and read by everyone. In that context it was important to establish a workflow for the online newsletter and analyze the work processes involved to assure that this service could in the long term be managed by the company itself. In addition, the technical infrastructure had to be designed and adapted to the actual ICT infrastructure. The Drupal content management system was selected because this open source software was a better fit for the company's demands for tailor made solutions within the margins of available skills and budget.

Microtraining. The managers were trained in the use of Microtraining. In a series of four workshops we dealt with the topics: 1. The need for well trained and informed staff, 2. Microtraining methodology, 3. Microtraining development, 4. Online course production based on the Microtraining concept. In this way the managers became familiar with the methodology and gained experience in using Microtraining through a group assignment of developing the content for an online module based on existing learning demands.

Online courses. The opportunities for mobile workers to attend place and time fixed learning events like courses are extremely limited. E-learning helps to make learning less place and time dependent. In the pilot phase learning by means of an online course has been examined using a course on the Digital Tachograph. The performance risks for drivers and company are high, since the introduction of this device is related to lots of implementation problems. The existing analogue equipment is gradually replaced by a digital version. The Microtraining concept was the basis for the development of this course, which consists of a series of learning events that could be used by the drivers online at any time or place.

4. Key Experiences

The project is moving from phase two to three as presented by the CLS-model [3] and data are being collected at the time of writing. Nevertheless we want to present some key experiences in the project up till now. The integration of redesigned information and learning policy in the strategic business plans has moved forward and basic-profiles are now used to assess performance. The ongoing evaluation of the basic tasks and the required tasks show that employees are in continuous need for new information, new learning and new skills and in this framework personal development plans are being prepared. Experiences with the personal profiles so far, show that communication with and among the employees are of great importance to pass on the changes that are taking place. The electronic newsletter allows conveying the message in the right context and evoking public interaction.

The online newsletter (see figure 5) started with a minimum of functionalities to allow for tracking the impact of the intervention. At the same time the workflow of newsletter production was developed for the people formerly responsible for the paper version. At the moment an editor is responsible for the content and first line technical support. The editor, employees and managers all contribute to the news flow by sending in messages, links, reports, pictures, mobile phone video clips, etc. All this information is being tagged that allows for easy access to topic related information, which could also be learning related content like a course, the drivers handbook, legal documents and others.

Fig. 5. The Homepage of the Business portal

Since news attracts most users, the homepage of the online newsletter is qualified as the main entrance for the business portal. The portal shows a multi dimensional approach reflecting the integration strategy of information and learning in the business processes. Employees were asked about their appreciation of the new system. The research shows that the drivers greatly appreciate the new online newsletter. Nevertheless there is still a large group that has difficulties with online information acquisition. This group is yet facilitated by the paper version of the newsletter, but this will come to an end. An ongoing activity is the further integration of the online newsletter (news portal) into the working practices. Reports on current diesel prices, pump instructions and so on should encourage people to use the system. Other incentives are the option for each employee to develop an online profile. This concept should allow for better use of the connections with the weak tie colleagues and strengthen the company's network.

As mentioned before, the online course on the Digital Tachograph (see figure 6), as a first product, was based on the Microtraining concept. The course was developed in close collaboration with the end users. A preliminary classification of subjects was used by the fleet manager and drivers to determine their needs and express their wishes. The content then was collaboratively developed and tested by the users. The course is online and frequently being used. A first test and survey revealed that the users liked the online opportunity, had virtually no difficulties with the navigation in the course, considered the information as informative and useful, but were confronted with the 'lonely learner experience' as being very different from what they were used to.

Fig. 6. The Homepage of the Tachograph Course

The holistic and multi dimensional approach in the project is a complicating factor, because of the number of interdependencies. Although the CLS model helps to manage the complexity, the day to day business reality is a rather unpredictable factor that influences the pace and the planning in the project. Urgent needs receive much more attention than planned for and other interventions have to be delayed. This is a reality test for the project and is a crucial element in the integration with the existing business processes as part of the new learning strategy. The online newsletter or business portal, as some call it, continues to be the central point for access to the news, business resources, learning environment, personal profiles and workshops (communities) that start developing around topics such as the Digital Tachometer. Key words (tags) are the main links between these various components and lower the threshold for the user to acquire any information, including the access to learning related resources offered in the business context. The most important next step is the emerging integration with the daily working practices that should coincide with the emerging acceptance of this new approach.

Some interesting numbers come from the comparison of the analysis of the target group at the beginning of the project at the end of 2007 with the state of affairs at the end of 2009 after the business portal was successively introduced during 2008. In 2007 the employees were not very satisfied with the communication in the company, especially not with the way management conveyed their information. The business portal apparently has a positive effect on the perception of the employees concerning the communication. The employees express that they receive more information on the topics they consider relevant, like about their work, business results, customer satisfaction, management information and human resource information. Also they consider this information to be more relevant than before. In 2007 32% of the employees indicated that they received information from management on a regular basis. By the end of 2009 this had risen to 76% and they considered this information more relevant, 39% – 54%.

An important issue was when, how often and from where the employees were using the portal. This is relevant to judge if the portal serves as low threshold information and learning resource at all times. It’s 37% of the employees that uses the portal daily and 27% weekly. On a two week basis it is 84% that uses the portal. 56% use an office pc to log on, 46% a home pc and 19% use mobile computers and other devices. 46% of the employees use during working days, 41% uses it in the weekend.

5. Conclusions

The reengineered information and learning strategy as an integrated service has been rewarded by management and employees of the transport company as a business process for innovation. Change though takes time especially when change of behavior is part of the innovation. Some employees think the development does not go fast enough, which is an indication for the emerging momentum in the development which should be addressed and used to better manage the expectations. The news portal is up and running and is used more and more by the employees, which shows the increasing integration of the service in the daily work processes.

The role of managers in this development needs to be strengthened. It remains difficult for them to maintain an overview of the new methodologies and materials and to actively work on the integration with their tight agenda. Manager involvement should carefully be planned. The results so far increase the demand and therefore it is necessary to agree on the pace and scope, so expectations stay in line with the possibilities. The managers are generally satisfied with the more flexible training options and start using the Microtraining concept of short focused sessions also in their group meetings. The Microtraining method is perceived as clear and logical. For the overall acceptance of the innovation, this is a promising development also because regular use will foster other applications. The use of the theoretical concepts and visions as a development framework plays an important role in discussions on choices to be made regarding the feature development of the learning strategy. On virtually any level of development, these concepts supply guidance which is very helpful to transfer the main ideas.

Two items seem to be crucial. A holistic approach is not only useful to get a good overview, but provides a better insight in the business processes, the opportunities and barriers and the relevance of the redesign. This project is not just about the introduction of some information management and e-learning applications, but about the development of a sustainable learning strategy that builds on the qualifications of the worker population. The second item is involvement. The needs and wishes of management and employees are at the core of what this strategy is about. If it fits their needs, they will use it and start depending on it.

The numbers of the user evaluation show very positive results in regard to the objectives of this project. Reengineering the learning strategy has shown so far to be useful and rewarding as a cross-functional process.

6. References

1. De Vries, P.: An Analysis Framework Approach for Managing Corporate E-learning Development. Brussels: De Vries. Doctoral dissertation. (2005)
2. De Vries, P. & Leege. T.: Final Report WP 1: Baumaerkte und Lernen: eine Bedarfsanalyse. Reload project. DE/07/LLP-LdV/TOI/147058. Leonardo Project European Union, (2008)
3. De Vries, P. & Lukosch, H.: Supporting Informal Learning at the Workplace. In: IJAC, International Journal of Advanced Corporate Learning. Vol 2, No 3. August 2009. Pp. 39-44. www.i-jac.org (2009)
4. De Vries, P. & Brall, S.: Microtraining as a Support Mechanism for Informal Learning. In: Elearningpapers of Elearningeuropa, on: http://www.elearningpapers.eu/, (2008)
5. Vygotsky, L.S. Mind and society: The development of higher mental processes. Harvard University Press, Cambridge, MA (1978)
6. Wertsch, J.V. Vygotsky and the Formation of the Mind. Cambridge, MA (1997)
7. Siemens, G.: Connectivism: A Learning Theory for the Digital Age. International Journal of Instructional Technology and Distance Learning, Vol. 2, No. 1 (2005)
8. Siemens, G.: Knowing Knowledge. Lulu Publishers (www.lulu.com), (2006)
9. Jonassen, D., Mayes, T., & McAleese, R.: A manifesto for a constructivist approach to uses of technology in higher education. In T.M. Duffy, J. Lowyck, & D.H. Jonassen (Eds.), Designing Environments for Constructive Learning (pp. 231-247). Springer-Verlag, Heidelberg (1993)
10. Rosenberg, M.J. Beyond E-learning: Approaches and Technologies to enhance Organizational Knowledge, Learning and Performance. (2006).
11. Jonassen, D. H.: "Instructional Design Models for Well-Structured and Ill-Structured Problem-Solving Learning Outcomes". In: Educational Technology Research and Development 45 (1): 65-94 (1997)
12. Cross, J.: Informal Learning: Rediscovering the Natural Pathways that Inspire Innovation and Performance. Pfeiffer , San Francisco (2007)
13. De Vries, P.; Lukosch, H.: Supporting Informal Learning at the Workplace. In: IJAC, International Journal of Advanced Corporate Learning. Vol 2, No 3. Pp. 39-44. www.i-jac.org. (2009)
14. Overschie, M.: Microteaching Manual: Effective transfer of knowledge for Sustainable Technological Innovation. On http://www.microteaching.org (15.8.2008), (2007)
15. Granovetter, M. The Strength of Weak Ties: A Network Theory Revisited. In: Sociological Theory, Volume 1, p. 201-233 (1983)

Gross Product Simulation with pooling of Linear and Nonlinear Regression Models

Ahmad Flaih[1], Abbas Abdalmuhsen[1], Ebtisam Abdulah[1], Srini Ramaswamy[1]

[1] University of Arkansas at Little Rock
Little Rock, AR 72204, USA
{anflaih, akabdalmuh, ekabdulah, sxramaswamy}@ualr.edu

Abstract. This paper discusses the problem of decision support systems in the organization. The procedure (linear combination) developed with the aim to combine some predicted results obtained with simulation of linear and nonlinear regression models (experts), multiple regression model, nonparametric regression model, and semi parametric regression model. This adjustment procedure enforce some statistical characteristics like the expected value of the gross production rate based on Cobb-Douglas production function is unbiased for the actual value, and the total weights(importance) of all models(experts) is equal to one. We used modeling and simulation techniques to generate our data and to apply the procedure.

Keywords: Regression Models, Linear combination, Cobb-Douglas production function, Predict of the gross product.

1 Introduction

In economics, the relationship between the production (outputs) and the production elements (inputs) known as production function. Usually decision makers in the organization and companies used some statistical and economics methods to support the production policy. Cobb-Douglas (CD) production function is one of the most useful tools to support the production policy, [1] and [9] used the CD function and found that public investment has a large contribution to production, [7] argued that public-sector capital has no effect on production after controlling for location characteristic, [2] and [5] pursued the insignificance of public capital. All of the literature above assumes a Cobb-Douglas (CD) production function to estimate productivity of inputs. One property of the CD functional form is that elasticity of an input is the estimated slope coefficient when the output and inputs are measured in logs, [6] used nonparametric perspective, which allows elasticities to vary across location and time, [15] use the CD functions proposed by [9]. [2] and [6] considers the gross product(GP) as responding variable, public capital (PUC), private capital (PRC), the employment rate (EMP), and the unemployment rate (UEPR) as explanatory variables, to predict of the gross product by using linear and nonlinear regression models. [13] they have propsed linear pooling method to combine the probability forecasts and proved the propsed method gave them more accurate results.

In this paper we used the same variables and methodology to predict gross production, but we use linear combination method to combine the results of the three experts (multiple linear regression, non-linear regression, and semi-parametric) as

J. Barjis, M.M. Narasipuram, G. Rabadi, J. Ralyté, and P. Plebani (Eds.):
CAiSE 2010 Workshop EOMAS'10, Hammamet, Tunisia, pp. 69-76, 2010.

input variables, we choose these three types of models because we think there is linear relationship, non-linear relationship and mixture relationship between gross production and the predictor variables. It is based on our hypothesis that the linear combination method will give us more accurate predictive gross production since we believe this is a better way to take into account the results of multiple models accompanied by an appropriate weighting scheme. Furthermore, we prove that this method gives an unbiased estimator to the gross production. Our contribution is that we use linear combination method to pool three regression models for predicting the gross production.
Linear combination is linear pooling of some sets of ordered pairs (importance, gross production), we have explored the prediction associated with different regression model can be combined into one final model via weighted linear combination will give more accurate results. Mathematically linear combination of the sequence $y_1, y_2, \ldots, y_n$ each with mean μ is:

$Y = \sum_{i=1}^{n} a_i y_i$, where a_i is the weight of y_i ,and

$$\mathrm{E}(Y) = \sum_{i=1}^{n} a_i \mathrm{E}(y_i) = a_1\mu + a_2\mu + \ldots + a_n\mu \qquad \textbf{(1)}$$

The $\mathrm{E}(Y) = \mu$ if $\sum_{i=1}^{n} a_i = 1$

2 The Models

Regression analysis is a technique used in data analysis; we use regression technique to predict the value of the response (dependent) variable given any value of the predictor (independent) variable. A general regression model is [15]:

$$y_i = E(y_i | x_i) + e_i \qquad \textbf{(2)}$$

Where i=1, 2... n denoting an observation of a subject. y_i is the response variable and x_i is a $k \times 1$ vector of independent variables. E $(y_i | x_i)$ is the expectation of y_i conditional on x_i, and e_i is the error term. In this paper we will use the following types of regression model:

2.1 Parametric Regression Model:

In this model, it is assumed that y_i is linearly related with x_i, so we can say it is linear regression model:

$$E(y_i | x_i) = \alpha + x_i' \beta$$

Thus a linear regression model is written as:

$$y_i = \alpha + x_i'\beta + e_i \qquad \textbf{(3)}$$

Where α is the intercept and β is $k \times 1$ vector of parameters. Under Gauss-Markov assumptions, the estimators of α and β are Best Linear Unbiased Estimators (BLUE), and can be estimate by using Ordinary Least Squares method (OLS).

2.2 Nonparametric Regression Model

If we do not know the data generating process, it is very unlikely that a linear regression (parametric) model is exactly the appropriate model specification, so the estimators α and β are not BLUE (best linear unbiased estimator). Instead of making assumptions for the functional form of E ($y_i | x_i$), nonparametric regression methods do not require any presumptions about the underlying data generating process [10]. Let

$$y_i = m(x_i) + e_i \quad \textbf{(4)}$$

Where m (.) is some function of unspecified functional form. Some basic assumptions about m (.) are commonly made. Many methods have been devised to estimate the regression function m (.), but we will just consider a simple but effective estimate known as the Kernel regression estimate. Suppose we have a random sample (x_1, y_1), (x_2, y_2) …, (x_n,y_n). The Kernel regression estimate of m(x) =E (y | X=x) is given by:

$$\hat{m}\,(\mathrm{x}) = \frac{\sum_{i=1}^{n} k\,(\frac{x_i - x_j}{d})\,y_i}{\sum_{i=1}^{n} k\,(\frac{x_i - x_j}{d})} \quad \textbf{(5)}$$

Here, K is a nonnegative symmetric function that is not increasing as its argument gets away from zero, and d is a parameter called the smoothing parameter that is selected by the user to control the amount of smoothing. The estimator in equation (4) is called the Local- Constant Least- Squares, which can be interpreted as a weighted average of y_i, where:

$$\frac{k\,(\frac{x_i - x_j}{d})}{\sum_{i=1}^{n} k\,(\frac{x_i - x_j}{d})}$$ is the weight attached to y_i

It should be noted that the product Kernel function $\mathrm{K}\,(\frac{x_i - x_j}{d})$ is the product of the Kernel function of all components of x. That is:

$$\mathrm{K}\,(\frac{x_i - x_j}{d}) = \prod_{s=1}^{k} k\,(\frac{x_{is} - x_{js}}{d_S}) \quad \textbf{(6)}$$

Where x_{js} is the s^{th} component of x_j and d_s the s^{th} component of d. The Kernel function k (.) can take several forms. In this paper we used the Gaussian Kernel function is defined as:

$$k\,(\frac{x_{is} - x_{js}}{d_S}) = \frac{1}{\sqrt{2\pi}} \exp[-\frac{1}{2}(\frac{x_{is} - x_{js}}{d_S})^2] \quad \textbf{(7)}$$

The smoothing parameter d is generally the most important factor when performing nonparametric regression. The bandwidth is chosen to obtain a desirable trade-off between the bias and the variance of estimation, so we need a method that could balance the bias and variance of the resulting estimate.

We have used Leave-One-Out Cross-Validation to select the optimal bandwidth. This method is depending on the principle of selecting bandwidth that minimizes the sum squared error of the resulting estimates. We will try to find the minimum MSE, mathematically; we are trying to minimize the following function [15]:

$$\text{CV (d)} = \frac{1}{n}\sum_{i=1}^{n}[y_i - \hat{m}_j(x_j)]^2 \quad \textbf{(8)}$$

Where $\hat{m}_j(x_j)$ is the Leave-One-Out estimator of m (.) evaluated at x_j. SAS software uses this method to find the optimal bandwidth values.

2.3 Semi parametric regression model:

Nonparametric techniques are very flexible because they do not need any assumptions about functional form. However, there are cases in which the relationship between the dependent variable and some independent variables is known to be linear and the relation between the dependent variable and other independent variables remain undetermined. In this situation semi parametric models are developed to solve this problem. Semi parametric models have both parametric and nonparametric components [15].

2.3.1 Semi parametric Partially Linear Model

Consider the semi parametric partially linear model:

$$y_i = x_i{}'\beta + m(z_i) + e_i \quad \textbf{(9)}$$

Where β is a $k \times 1$ vector of parameters, z_i is a $q \times 1$ vector of independent variables and e_i is an additive error, is assumed to be uncorrelated with xi and z_i. β is the parametric part of the model and the unknown function m (.) is the nonparametric part. [8] [11] proposed an estimate of β and m (z_i) as follows. Using ordinary least squares, the estimator of β is:

$$\hat{\beta} = (\sum_{i=1}^{n}\tilde{x}_i\tilde{x}_i{}')^{-1}\sum_{i=1}^{n}\tilde{x}_i{}'\tilde{y}_i \quad \textbf{(10)}$$

Where:

$\tilde{x}_i = x_i - E(x_i | z_i)$ and $\tilde{y}_i = y_i - E(y_i | z_i)$

Once we obtain $\hat{\beta}$, the nonparametric part m (z_i) is easy to estimate. From equation (8), we get: m (z_i) = $y_i - \tilde{x}_i\beta - e_i$

Then we can get the estimator of m (z_i) as follow:

$$\hat{m}(z_i) = \frac{\sum_{i=1}^{n}(y_i - \tilde{x}_i\beta)k(\frac{z_i - z_j}{d})}{\sum_{i=1}^{n}k(\frac{z_i - z_j}{d})} \quad \textbf{(11)}$$

Where d can be estimate similar to the nonparametric model? SAS software uses Cross-Validation to find the optimal bandwidth.

3. Simulation Study

We will generate a random sample of size 30 observations for each explanatory variables: public capital (PUC), private capital (PRC), the employment rate (EMP), and the unemployment rate (UERP). The Cobb-Douglas (CD) production function is,

$$y_i = f\,(PUC_i\,, PRC_i\,, EMP_i\,, UERP_i\,) \qquad \textbf{(12)}$$

Where y_i =gross product, PUC_i =public capital,

PRC_i =private capital, EMP_i =employment rate (labor)

$UERP_i$ =unemployment rate.

We can rewrite the CD function as,

$$\mathrm{E}(y_i) = \beta_0 PUC_i^{B_1} PRC_i^{B_2} EMP_i^{B_3} UERP_I^{B_4}\,\varepsilon_i$$

The log-linear CD production function is

$$y_i = \beta_0 + \beta_1 PUC_i + \beta_2 PRC_i + \beta_3 EMP_i + \beta_4 UERP_i + \varepsilon_i \qquad \textbf{(13)}$$

Further,

1- If $\beta_1 + \beta_2 + \beta_3 + \beta_4$ =1, the product function has constant returns to scale.

2- If $\beta_1 + \beta_2 + \beta_3 + \beta_4$ <1, returns to scale are decreasing.

3- If $\beta_1 + \beta_2 + \beta_3 + \beta_4$ >1, returns to scale are increasing.

3.1.1 Parametric model Results

We apply multi-linear regression model, covariance model, and variance component model to our generated data to find the estimated values by using SAS procedure proc reg [12] and the output are listed in Tables 1, follow:

Table 1: Estimates of output Elasticity: parametric approaches

Parametric	β_0	β_1	β_2	β_3	β_4
OLS S.E	15.089 (5.4239)	-0.0157 (0.3465)	0.29952 (0.218)	-1.709 (1.107)	-0.6721 (0.2114)
covariance model	1.87 (4.23)	0.342 (0.034)	0.005 (0.1933)	-.0127 (0.039)	-.004 (0.001)
variance component model	0.018 (0.023)	0.266 (0.020)	0.755 (0.023)	0.347 (0.052)	-0.005 (0.001)

The results of Table 1 are based on the model in [12].

3.1.2 Non Parametric Model

We used this model to find the estimated values of the intercept and elasticity values by using generated data and SAS procedure (proc gam). The output from proc gam is listed in following table, Table 2:

Table 2: Estimates of output Elasticity: Nonparametric approach

Nonparametric	β_0	β_1	β_2	β_3	β_4
Mean	14.834	-0.211	0.2409	-1.242	-0.653
S.E	(6.342)	(0.37)	(0.391)	(1.401)	(0.27)

3.1.3 Semi- Parametric Model

We used this model to find the estimates values of Elasticity by using SAS procedure (proc gam) and we considered the independent variable (Public Capital) is linearly related to the gross production. The output from proc gam is listed as follows in Table 3.

Table 3: Estimates of output Elasticity: semiparametric approach

Semiparametric	β_0	β_1	β_2	β_3	β_4
Mean	19.892	-0.459	0.288	-2.010	-0.642
S.E	(7.369)	(0.470)	(0.296)	(1.50)	(0.28)

3.2 Linear Combination Method

We will use the estimated models (experts), multiple parametric regression, nonparametric regression, and Semiparametric regression to predict o f the log-gross production, for sample of size equal to 10 observations, and by using the values of the columns(Parametric y_1,Nonparametric y_2,Semiparametric y_3) in the following table, we can estimates the weight corresponding to each expert(model) with SPSS software by using neural networks procedure and estimate the log-gross production Y as follow:

$$Y = 0.194 y_1 + 0.595 y_2 + 0.255 y_3$$

Correlation is a measure of the association between two variables; it is a very important part of statistics. One of the most fundamental concepts in many applications is the concept of correlation. If two variables are correlated, this means that you can use information about one variable to predict the values of the other variable. The correlation matrixes between the predicted values in the table 4 as follows:

Table 4: Correlation between Predicted Values

i	y_1	y_2	y_3	Actual	Y
1	12.445	11.234	12.876	12.98	12.41
2	12.458	11.011	12.455	12.23	12.08
3	12.765	10.203	12.897	11.70	11.84
4	11.239	11.054	12.987	12.12	12.07
5	12.456	12.899	12.765	13.25	13.31
6	12.345	12.876	12.849	13.56	13.33
7	12.455	11.112	12.098	12.32	12.11
8	13.345	13.453	12.069	13.55	13.66
9	13.678	12.456	12.932	13.40	13.36
10	12.546	12.453	12.948	13.23	13.14

From Table 5 we can find that the correlation coefficient between the predicted values of the linear combination(Y) and the actual variable of gross production is (0.954) that means the linear combination method is more associated with the actual values than the expert1, expert2, and expert3.this strongest association supports the suggestion that the predicted value of the gross production which associated with the different experts that combined is better than taking the predictive values of each expert individually. The average difference is (0.261, 0.959, 0.146, 0.103) of the three regression models (parametric, non-parametric, and semi-parametric) respectively and linear combination procedure. The proposed method (linear combination) is more accurate than regression models because the difference (actual and Y) in table 4 is more less than the other models, so this means that the proposed method give us unbiased estimator.

Table 5: Correlation Matrix

	y_1	y_2	y_3	Actual	Y
y_1	1	0.407	-0.268	0.457	0.525
y_2	0.407	1	-0.149	0.935	0.981
y_3	-0.268	-0.149	1	-0.015	-0.047
Actual	0.457	0.935	-0.015	1	0.954
Y	0.525	0.981	-0.047	0.954	1

4. Conclusion

In this paper, table 5(correlation matrix) shown that the correlation coefficients between the linear combination method, parametric model, nonparametric model and semi-parametric and the actual values of the gross production are (0.954, 0.457, 0.935, and -0.015), because we believe the public capital (PUC) is linearly related to gross production, so we choose it as parametric variable and that is why the correlation between y_3 and actual value is -0.015. Linear combination method is more associated with the actual value than the others experts (regression model), i.e. linear combination method reflects the experts views to find the most fitted predicted value to the actual value. Finally as forecasters often wish to provide an accurate predicted value, so we will use the linear combination method to predict of the gross production.

References

1. Aschauer David A. : public investment and productivity growth in the Group of Seven. 89-13, Federal Reserve Bank of Chicago (1989).
2. Baltagi,B.and Pinnoi,N. : Public capital stock and state productivity growth:Further evidence from and error component model.Empirical Economics 20,351-359 (1995).
3. Clemen, R. T. and Winkler, R. L. :"Combining Probability Distributions from Experts in Risk Analysis", Risk Analysis 19,187-203 (1999).
4. Dong Xiang "Fitting Generalized Additive Models with the GAM Procedure", SAS Institute Inc., Cary, NC 27513 (1990).
5. Garcia-Mila,Teresa,Therese.J.McGuire,and Robert H. ,Porter: The effect of public capital in state-level production function reconsidered .Review Economics and Statistics forthcoming (1996).
6. Henderson,D.J. and Kumbhakar,S. :Public and private capital productivity puzzle:A nonparametric approach,southern Economic Journal 73(1),219-232 (2006).
7. Holtz-Eakin,D. :Public-sector capital and the productivity puzzle,Review of Economics and statistics 76,12-21 (1994).
8. Li, Q. and Racine, J.S. : Nonparametric econometrics: theory and practice, Princeton University Press (2007).
9. Munnell,A. :How does public infrastructure affect regional economic performance? New England Economic Review September,11-32 (1990).
10. Par Osterholm :"Incorporating Judgment in Far Chart", Finance and Economics Discussion Series. Screen Reader Version (2006).
11. Robinson, P.M. : Root-n-consistent semiparametric regression, Econometrica 56, 931-954 (1988).
12. Ronald P. Cody and Jeffrey K. Smith : "Applied Statistics and the SAS Programming Language" 5^{th}.ed (2006).
13. Roopesh R., and Tilmann G. : Combining Probability Factors", Technical Report no.543 (2008).
14. Schick, A. : On asymptotically efficient estimation in semiparametric models, Annals of Statistics 14, 1139-1151 (1986).
15. Xianghang S.:"Application of Nonparametric and Semi-parametric methods in Economics and Finance", PhD dissertation, University of Southern California (2009).

Actionable Meta Models to Support Inter Organizational Business Processes Modeling for e-Services

P. Radha Krishna[1] and Murali Mohan Narasipuram[2, *]

[1]SET Labs, Infosys Technologies Limited, Hyderabad, India.
[2]Department of Information Systems, City University of Hong Kong,
83 Tat Chee Avenue, Hong Kong.
radhakrishna_p@infosys.com, ismohan@cityu.edu.hk

Abstract. E-services usually span multiple organizations and evolve in time-domain due to the changes in the environment pertaining to each organization. Enacting composite e-services over evolving inter-organizational business processes is a challenging task. In this paper, we highlight modeling of business processes for e-services and present enactment of the e-services with the help of meta models in order to cope with the changing needs. We also describe high-level framework to enact e-services through actionable meta models.

Keywords: Meta models, Business process modeling, Run-time environment

1 Introduction

Modeling e-services is an important task in business processes specifications and execution. A single model may not be enough to design e-services when it spans multiple organizations. At the same time, it is hard to describe a generic model for any kind of e-services. This necessitates an abstract level of modeling, that is meta-modeling, which helps in instantiating context-specific model instances to suit the needs of e-services. It also provides required facilities to support functionality for adapting the model to new requirements.

In the domain of business process modeling, meta-models take a leadership role in defining how the data and corresponding application/business logic is handled right from the modeling to all the way to enactment of business processes. E-services involve inter-organizational business processes. Enactment of such business processes involves multiple parties and is modeled, using workflows by introducing e-services in business process models [1].

---

[*] currently visiting International Institute of Information Technology, Hyderabad (IIITH), INDIA

J. Barjis, M.M. Narasipuram, G. Rabadi, J. Ralyté, and P. Plebani (Eds.):
CAiSE 2010 Workshop EOMAS'10, Hammamet, Tunisia, pp. 77-83, 2010.

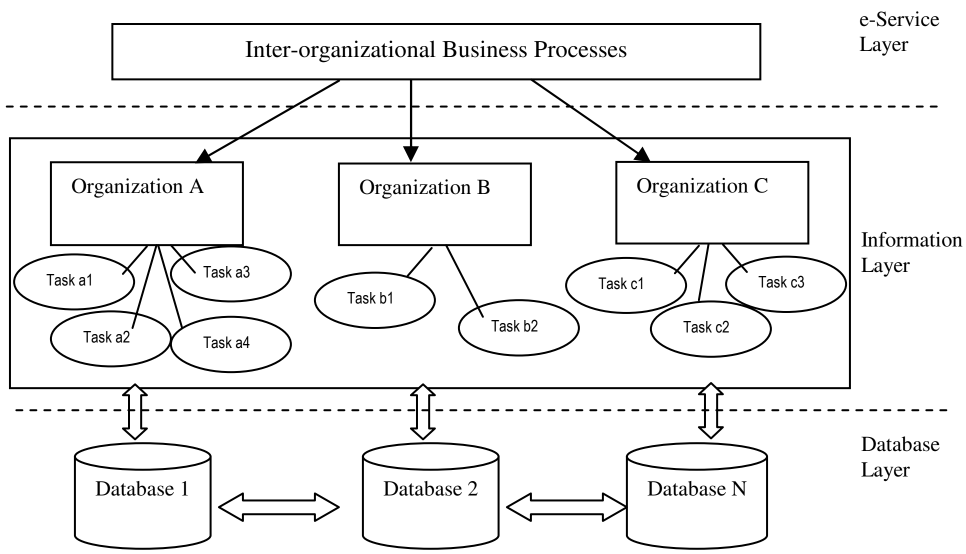


Fig. 1. An example scenario of inter-organizational business processes

Figure 1 shows a layered diagram to illustrate an example of e-services involving multiple organizations. The top layer is concerned to e-service layer, where an e-service spans multiple business organizations and carried out using inter-organizational business processes. The middle layer is information layer, which outlines information about the tasks of each organization and the e-service will involve some or all of these tasks. The bottom layer is a database layer where all the data pertaining to respective organizations are stored for further processing. Consider an example scenario where a particular e-service may need task a1, task a4, task b2 and task c1. So there is a need for collaborative business processes from different organizations. Web service applications integrate inter-organizational and heterogeneous services on the Web. Due to the changes in application/business logics at run-time, Web services are created and updated dynamically.

Several researchers have developed meta-model for different domains. Karagiannis and Höfferer [5] presented a survey on meta-models and developed taxonomy for classifying application scenarios of meta modeling according to domain, design, and integration. In [2], a formalism-independent meta-model is described to model a business process and presented the interactions among the different aspects in order to capture the dynamic behaviour of a business process model. The ER^{EC} meta-model proposed in [6] is useful to instantiate appropriate data model for electronic contracts. Chiu et al., [3] presented a meta-model approach to develop web services based e-Negotiation framework. Rosemann and Muehlen [7] presented a meta model approach to evaluate and compare different workflow management systems. They used extended entity relationship models to design the meta data models and provided a generic organizational reference meta model for enacting inter-organizational workflows. Design of all these meta models are specific to an application and assumes complete specifications of it is known in advance.

In this work, we consider services as workflows and present how meta-models are useful to facilitate the changes in specifications during run-time.

Actionable meta modeling enables upkeep of meta model and the corresponding mapping all the way to run-time caused by the changes in organizational environment at run time, or sometimes changes initiated by users due to external factors. The main feature of this framework is that it facilitates to systematically address the problem of proactively responding to changes driven from environment.

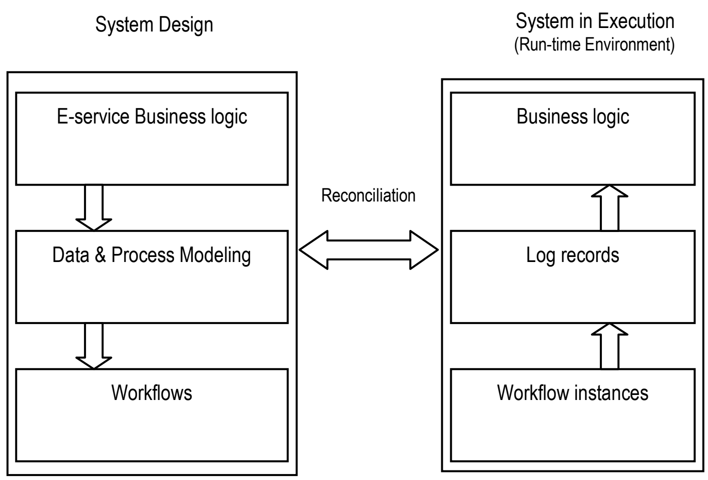


Fig. 2 Reconciliation of Meta Model with Run-Time Environment Changes

During e-service system design, the requirements are collected and the e-service elements such as activities and payments are extracted to model an e-service. Workflows will be identified to execute the e-service activities carried out by organizations. The log records keep track of the e-service execution. However, changes in organizational environment during e-service enactment need modification at conceptual level as well as at logical level. In order to propagate changes that occur during enactment to conceptual and logical level, there is a need to obtain business logic from log records and should be reconciled with the actual business logic of the e-service (Fig. 2).

In the next section, we present the framework for enactment of e-services for evolving inter-organizational business processes.

2 Architecture for Actionable Meta-modeling

E-service execution requires the modification of model schema definitions and in-turn changes in workflow instances, as a remedy. Moreover, additional constructs are needed in the system during work in progress. The present approach offers a practical solution from modeling and enactment of e-services, driven by meta-models. We maintain business policies that refer to adaptable instances of workflow schema while the service is being executed. Workflow patterns [4, 8] are useful to describe how the changes will be specified, implemented and perceived in e-services.

Figure 3 shows architecture of a system to model business processes and its instances. Web Service Server provides the services and transport of inter-organizational communications among business partners involved in the service. The *Run-time environment* (RTE) details such as workflows, rules, etc. are maintained in the database. *Dynamic Workflow Instance Generation (DWIG)* generates workflows on-the-fly and rules. It also allows the administrators to customize and edit them. Workflow definitions created or specified are executed by the *Meta-workflow Driven Workflow execution* engine. That is, the workflow engine enacts the workflows specified by the dynamic workflow instance Generation. The *Event handler* manages the events occurring during the execution of workflows. It handles events in a unified manner for both normal and exception parts of a business process workflow. The Event-Condition-Action (*ECA) Rule Manager* initiates appropriate ECA rules based on the input from Event Handler. It also keeps track of generated rules with their corresponding actions and allows users to define additional rules if necessary.

The workflow engine and the ECA rule manager works in a synchronized manner. Thus, the ECA rules control the workflow execution and the events that occur during the workflow execution result in appropriate actions. The changes in the design-time update the corresponding database. The *Knowledge Base* maintains the currency of information such as application policies, versions, new requirements etc. that governs execution of the workflow. It also stores the vocabulary of an e-service domain. Knowledge Base captures the updates that take place in the run-time as well as in the design-time. These updates, in turn, become input to the *Run-time Business Processes evaluator (RTBPE),* which generates candidate meta-model(s) and add/modify the meta-model repository. *Meta-Model repository* contains the meta-models that are specific to a service under consideration. Meta-models are added or modified based on the requirements collected from run-time environment changes as well as design-time changes.

Business Processes logic and events are captured by execution support of meta-models. *Control and Monitoring Unit* (CMU) generates/modifies the *workflow patterns* and also specific elements namely *Events, States, Processes and Rules (ESPR)* in order to keep track of changes and their progression with respect to business processes under consideration. CMU receives events and results from the run-time environment and service specific components and policies. The modeling of changes during service

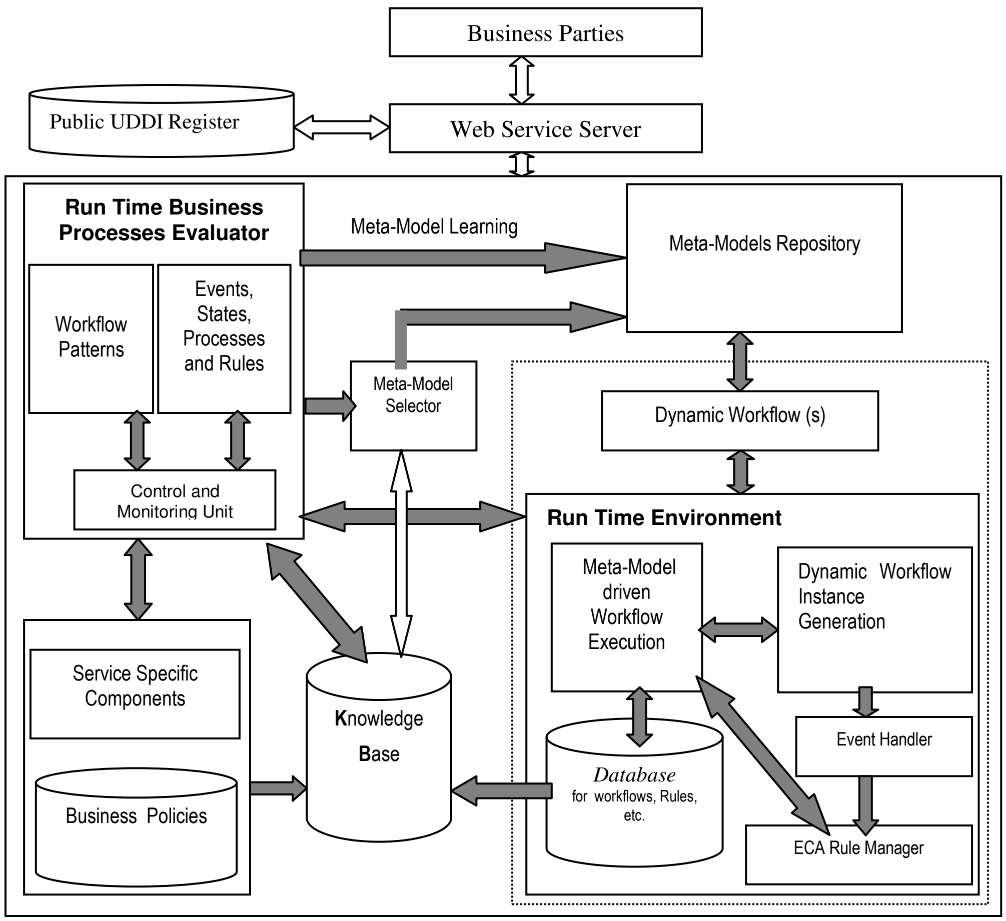


Fig. 3. Architecture for actionable meta-modeling to support evolving businesses processes

enactment can be seen as a different kind of meta-processes (tasks). Further, *Service Specific Components* are required for encapsulation and realization of the domain-specific logic for the service. The three components namely CMU, workflow patterns and ESPR will serve as RTBPE. Meta-model selector selects the appropriate meta-model(s) from the meta-model repository, which in turn drives the dynamic workflows to carry out business processes execution according to the changed context.

3 Discussion on the Proposed Meta-modeling Architecture

Proposed architecture facilitates learning from Business Activity Monitoring (BAM) to feed into the meta model updates. In the current Business Process Management domain, there is a lack of suitable architectures for BAM to support learning. The RTBPE proposed in this architecture provides the BAM metrics and their respective evaluations. These evaluations are the primary drivers for the learning required to update the meta models. The meta model learning is fed back into RTE through the dynamic work flow link with meta model repository. The benefits derivable from the proposed architecture depend on the e-service BAM design in terms of the RTBPE. Another advantage is that the dynamic workflow link between the meta model repository and RTE makes this architecture actionable.

Further work is needed to enhance the proposed architecture by explicitly specifying the service design issues such as granularity, complexity, flexibility, interoperability, autonomy, contract specifications, etc. in the RTBPE. Such enhancements will help the Business Process Management Systems (BPMS) software developers to bench mark their products and the BPMS users to assess and evaluate the BPMS products in a more comprehensive manner.

4 Conclusion

Modeling of e-services facilitates effective specification and execution of business processes. In the case of inter-organizational business processes, development of a generic model is a complex task. In this paper, we presented a meta modeling approach which helps to create context based model instances according to the changes that occur during run-time. Moreover, the meta-models learn from the services environment and update the models as necessary and hence making the meta-models actionable.

References

1. L. Aversano and G. Canfora, Introducing eservices in business process models, Proceedings of the 14th international conference on Software engineering and knowledge engineering, 2002.
2. B. Axenath, E. Kindler and V. Rubin, AMFIBIA: a meta-model for integrating business process modelling aspects, International Journal of Business Process Integration and Management, 2 (2), pp. 120 – 131, 2007.
3. D. K. W. Chiu, S. C. Cheung, P. C. K. Hung, S. Y. Y. Chiu and A. K. K. Chung, Developing e-Negotiation support with a meta-modeling approach in a web services environment, Decision Support Systems, 40 (1), pp. 51-69, 2005.

4. G. Greco, A. Guzzo, G. Manco, and D. Sacca`, Mining and Reasoning on Workflows, IEEE Transactions on Knowledge and Data Engineering, 17 (4), pp. 519-534, 2005.
5. D. Karagiannis and P. Höfferer, Metamodels in Action: An Overview. In J. Filipe, B. Shishkov, and M. Helfert, eds. ICSOFT 2006 - First International Conference on Software and Data Technologies, Insticc Press, Setúbal, 2006.
6. P. Radha Krishna, Kamalakar Karlapalem and D. K. W. Chiu, An EREC Framework for E-Contract Modeling, Enactment and Monitoring, Data and Knowledge Engineering, 51 (1), pp 31-58, 2004.
7. M. Rosemann, and M. zur Muehlen, Evaluation of Workflow Management Systems: A Meta-Model Approach, Australian J. of Infor. Systems, 6(1), p.103-116, 1998.
8. W.M.P. van der Aalst, A.H.M. ter Hofstede, B. Kiepuszewski, and A.P. Barros, Advanced Workflow Patterns, In Proc. of CoopIS 2000, pp. 18-29, 2000.

Discovering Organizational Perspective in Workflow using Agent Approach: An illustrative Case Study

Mahdi ABDELKAFI and Lotfi BOUZGUENDA

MIRACL Laboratory / ISIMS, Route de Tunis, km 10.
BP 242, 3021 Sakeit Ezzit, Sfax, Tunisia.
mahdi.abdelkafi@hotmail.fr,lotfi.Bouzguenda@isimsf.rnu.tn

Abstract. The existing Workflow Management Systems, WfMS (ACTIONTm, Xsoft InConcert, Team Ware) show some insufficient due essentially to two factors. First, the enterprises express a real need of WfMS that deal with flexible business processes. These later must adapt to the optimization of enterprises work methods and the evolution of their needs. Second, the enterprises also need of control mechanisms and modeling tools for their WfMS in order to design robust workflow models assuring an efficient and flexible processing of errors and exceptions of execution. In others words, the WfMS must deal with Workflow mining to support Business Processes Re-engineering (BPR). A BPR is a method to improve the effectiveness and efficiency of business processes. Existing propositions in the literature are rather dedicated to the process perspective mining and they neglect the discovering of organizational perspective although it is decisive for the enhancement of existing workflow or the proposing of a new Workflow. The paper addresses the workflow mining by considering the organizational perspective. By organizational perspective, we mean the organizational structures (Federation, Hierarchical, Market and so on) and interaction protocols (Contract net, Auction...) describing respectively the forms of work and the activities allocation process between actors. This paper defends the idea that organizational dimension in multi-agent system is an appropriate approach to discover this organizational perspective since it (i) provides inter-agents communication languages based on performatives which allow capturing the semantic of exchanges between actors, (ii) it introduces organizational structures to highlight the forms of work and (iii) it offers powerful interactions protocols which permit to identify the dynamic organizational structures. First, this paper proposes a Workflow log meta-model which extends the classical one. Then, it provides some algorithms to permit the discovering of the main organizational structures and interaction protocols. Third, it presents the Management process of Water distribution crisis case study to illustrate our approach. Finally, the paper describes the developed prototype, called DiscopFlow, in order to validate the proposed solution. Keywords: Interaction Protocols, Organizational Structures, Workflow log Meta-Model, Algorithms and DiscopFlow.

J. Barjis, M.M. Narasipuram, G. Rabadi, J. Ralyté, and P. Plebani (Eds.):
CAiSE 2010 Workshop EOMAS'10, Hammamet, Tunisia, pp. 84-98, 2010.

Key words: Interaction Protocols, Organizational Structures, Meta-Model, case study, FIPA-ACL Performatives.

1 Introduction

Workflow Mining Context. The Workflow is a key technology which automates the coordination of activities composing business process [1]. A Workflow Management System, WfMS is a software which permits to define, implement and execute one or several business processes. In order to support the interoperability between WfMS, the WorkFlow Management Coalition, WFMC has defined reference architecture [1]. The main component of this architecture is the Workflow Enactment Service (WES) that manages the execution of business processes and that interacts, on the one hand, with business processes (or workflows) definitions, execution and monitoring components and on the other hand, with external WES. The communication among the different components is supported by five interfaces of which Interface 5 is of most interest to us. This interface supports the connection between the WES and the monitoring tools. Unfortunately, a good standard for this interface has never been proposed. In this context, the Workflow Mining area has recently appeared and considered as a new research domain. More precisely, the purpose of Workflow mining is to analyze the Workflow log in order to discover the Workflow perspectives such as the Organizational Perspective (OP), the Informational Perspective (IP) and the Process Perspective (PP) [2], which help the monitor to improve or propose a new workflow. The OP has two objectives. First, it structures actors in classes playing each one a specific role (for instance subordinate and chief roles in simple hierarchical structure). A class is called an organizational structure defining an interaction space between actors. Of course, each actor belongs to its own organizational unit where it plays a predefined role. Second the organizational perspective describes the allocation activities between actors according to interaction protocols. In this paper we are interested to discover the organizational structures and interaction protocols and not the concepts of classical organizational model such as role, organizational unit, etc. The IP describes the structure of forms, documents and data which are consumed and produced by processes. The PP defines activities components, their coordination, information and actors involved in each activity. This perspective refers to both the organizational perspective, which defines and organizes the set of potential actors, and the informational perspective, which allows access to the objects to be processed. Most of the work deals with the discovery of process and/or informational perspective(s) while few researches were investigated around the discovery of the organizational perspective although it is decisive for the enhancement of existing workflow or the proposing of a new Workflow ([2], [3], [4], [5]).
A good application domain justifying the need to discover the organizational perspective is that of crisis Workflows as much as it is important to understand the social behavior of the actors during a crisis process. In others terms, we need to know the involved organizational structures in this crisis and the in-

teraction protocols used by the potential actors.In this paper, we have chosen the well-known "Management process of Water Distribution crisis" case study for three reasons. First, the water is a main need for the human and must be protected against any incident. Second, the frequency of water pollution is very high. Third, this case study is complete in so far as it illustrates clearly the interaction protocols used by the actors (Analyst, Controller, Investigator and so on) and the deployed organizational structures.
The problem being addressed in this paper is how to discover the organizational perspective from workflow log? We recall that organizational perspective refers to the organizational structures and interaction protocols. One possible way to deal with the discovering of organizational perspective in workflow is to use the agent approach and more precisely:

– the FIPA-ACL performatives which define the semantic of messages and notably the agents intentions,
– the interaction protocols such as contract net, auction and negotiation which constraint the conversations between actors during the activities allocation,
– the organizational structures like coalition, hierarchical, market and federation which define the behavior of actors ([6], [7]).

The goal of this paper is to propose a Workflow log Meta-model to help the discovery of organizational structures and interaction protocols.
Our solution is based on the following principles:

– The inclusion of the three complementary workflow perspectives. As a consequence we can support any discovering of perspective.The use of FIPA-ACL Performatives which are appropriates to define the organizational structures referencing interaction protocols,
– The simplicity by including the basic workflow concepts.

Organization of the paper. The reminder of this paper is organized as follow. Section 2 motives the use of Agent approach for discovering organizational structures and interaction protocols in Workflow. Section 3 describes the Workflow log Meta-model that we propose to deal with organizational perspective mining. Section 4 presents our case study useful for instantiating this model. First, it introduces the Petri Nets with Objects formalism and exposes its three models which are in interaction. Then, it gives the instantiated Meta-model. Finally, it discusses some organizational structures and interaction protocols which can be discovered according to this instantiation. Section 5 briefly discusses the related Works and concludes the paper.

2 Motivations for using Agent Approach

The objective of this section is to briefly motive the use of agent approach for organizational perspective discovering in workflow and more precisely:

– the FIPA-ACL communication language,

– the organizational structures,
– the interaction protocols,

FIPA-ACL (Foundations of Intelligent Physical Agents-Agent Communication Language, `http://www.fipa.org`) is one of the inter-agents communication languages proposed by the agent community. It is composed of a set of performatives and integrating of protocols. The FIPA-ACL message structure is shown in figure 1. For more information on FIPA-ACL performatives the interested reader can refer to [8].

```
(Performative
                    :sender A
                    :receiver B
                    :content XXX
                    :in-reply-to YYY
                    :language ZZZ
                    :reply-by CCC
                    :conversation_ID VWV)
```

Fig. 1. The FIPA-ACL message structure

These FIPA-ACL performatives play a key role in the activities allocation between actors. First, they clearly define the semantic of messages and namely the agents intentions (delegate, subcontract, negotiate,). Second, they highlight, during the interaction between agents, the organizational structures such as coalition, sub contracting, federation and so on. Third, they capture without ambiguity the interaction protocols like contract-net, auction, negotiation and heuristic. These interaction protocols could be used with benefits to allocate activities to actors in workflow. The table 1 describes the performatives used in this paper.

The organizational structures mentioned above can model the behaviour of the systems under consideration. I.e. they describe the macro-level dimension of the coordination among actors in terms of externally observable behaviour, independently of the internal features of each participating component.

3 The Workflow log Meta-model

3.1 Key requirements for a Workflow log Meta-model

The most important requirements for a Workflow log Meta model are the following:

Performative	Description
Delegate	The initiator delegates the receiver the Activity A of Case C
Inform	Concerns any information dispatch concerning Activity A of Case C
Execute	The initator informs that it performed the Activity A of Case C
Call for Proposal	Call for Offers: announces a negotiation for performing of the Activity A of case C
Propose	The initiator proposes to execute an Activity
Refuse	The initiator refuses the proposition of the receiver
Accept Proposal	After propose: to mean the acceptation
Reject Proposal	After propose: to mean the reject

Table 1. The performatives description

- It must allow the expression of the three workflow perspectives such as the informational perspective, the organizational perspective and the process perspective,
- It must be simple and comprehensive: it must define the core concepts of the three complementary Workflow perspectives,
- It supports the discovering of organizational structures and interaction protocols without ambiguity as mentioned previously.

Most of the work concerning Workflow perspectives mining ([2], [3], [4]) only focuses on the process perspective by providing a workflow log containing only concepts of the process perspective. Other also considers the informational perspective besides to the process perspective [5]. To the best of our knowledge, any work supports the last requirement and makes really the glue between the three workflow perspectives. Consequently, we have defined our own meta-model which fulfills the previous requirements. This meta-model is shown in the UML diagram of figure 2.

3.2 The proposed Workflow log Meta-model

In this UML Meta-model, a Process is composed of one (or several) Process Instance(s). Each Process Instance is composed of one (or several) Event(s). Each Event makes reference to the following elements:

- An Activity which is described through the Act_Name attribute,
- A Document which is described through the Doc_Name attribute,
- An Actor which is described through the Actor_ID and Actor_Name attributes, Its a member of organizational unit and plays one (or several) Role(s),
- A Role which is described through the Role_Name attribute,
- An Organizational Unit which is described through the Org_Unit_Name attribute,

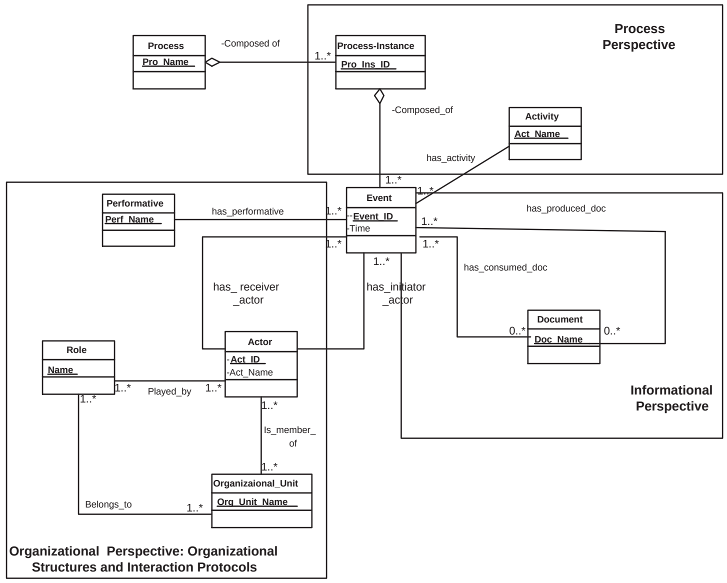


Fig. 2. The Workflow log Meta-model

– A Performative which is described through the Per_Name attribute.

Besides, the consumed documents and the produced documents are respectively represented by the Has_Consumed_Doc and Has_Produced_Doc relationships. The initiator Actor and the receiver actor are represented respectively by the Has_Initiator_Actor and Has_ Receiver_Actor relationships.

The process perspective is described with the Process Instance, Event and Activity classes (or concepts). According to the literature, these concepts are sufficient to discover the process structure. The informational perspective is described with the Process Instance, Event and Document concepts. We have chosen only to add the document concept because we think that this later will be very useful in crisis process. It is important to know on what such actor is based to produce such document. For instance, in our case study (Management Process of Water Distribution Crisis) the Malik actor playing the penal role can produce the document Juridical Report only on the basis of a document Investigation Report produced by the Walid actor playing the investigator role (see figure 4). The organizational perspective as mentioned in introduction is represented with Actor, Role, Organizational Unit and Performative concepts.

4 An illustrative case study

The Objective of this section is (i) to illustrate the workflow log Meta-model instantiation through the well-known Management process of Water Distribution crisis case study and (ii) to show the organizational structures and interaction protocols which can be discovered. First, it briefly introduces Petri Net with Objects (PNO) formalism as an appropriate language for modelling processes. Second, it models the case study through three communicating models (information, organization and process). Third, it presents the instantiated Meta-Model. Finally, it exposes some organizational structures and interaction protocols which can be discovered.

4.1 What are Petri Nets with Objects?

Petri Nets with Objects (PNO) [9] are a formalism combining coherently Petri nets (PN) technology and Object-Oriented (OO) approach. While PN are very suitable to express the dynamic behavior of a system, OO approach enables the modeling and the structuring of its active (actor) and passive (information) entities. In a conventional PN, tokens are atomic and undissociable, whereas they are objects in a PNO. As any PN, a PNO is made up of places, arcs and transitions, but in PNO, they are labeled with inscriptions referring to the handled objects. More precisely, a PNO features the following additive characteristics:

- Places are typed. The type of a place is a (list of) type of some object-oriented sequential languages. A token is a value matching the type of a place such as a (list of) constant (e.g. 2 or hello), an instance of an object class, or a reference towards such an instance. The value of a place is a set of tokens it contains. At any moment, the state of the net, or its marking is defined by the distribution of tokens onto places. A transition is connected to places by oriented arcs as it aims at changing the net state, i.e. the location and value of tokens.
- Arcs are labeled with parameters. Each arc is labeled with a (list of) variable of the same type, as the place the arc is connected to. The variables on the arcs surrounding a transition serve as formal parameters of that transition and define the flow of tokens from input to output places. Arcs from places to a transition determine the enabling condition of the transition: a transition may occur (or is enabled) if there exists a binding of its input variables with tokens lying in its input places. The occurrence of an enabled transition changes the marking of its surrounding places: tokens bound to input variables are removed from input places, and tokens are put into output places according to variables labeling output arcs.
- Each Transition is a complex structure made up of three components: a precondition, an action and emission rules. A transition may be guarded by a precondition, i.e. a side effect free Boolean expression involving input variables. In this case, the transition is enabled by a binding only if this binding evaluates the precondition to true. Preconditions allow for the fact that the enabling of a transition depends on the location of tokens and also on their value. Most transitions also include an action, which consists in a piece of code in which transitions variables may appear and object methods be invoked. This action is executed at each occurrence of the transition and it processes the values of tokens. Finally, a transition may include a set of emission rules i.e.

side-effect free Boolean expressions that determine the output arcs that are actually activated after the execution of the action.

4.2 The Informational Model

The informational model that we propose focuses only on the documents consumed or produced by the case study process (see figure 3). More precisely, it contains the following documents:

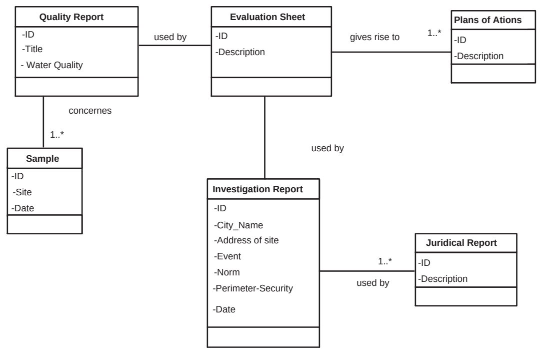


Fig. 3. The UML Classes diagram of the informational model

- An Investigation Report which can be used by one or several Juridical Report,
- An Evaluation Sheet which is based on Quality Report and gives rise to one or several Plans of Actions,
- A Quality Report concerns one or several Sample to be analysed.

4.3 The Organizational Model

Our organizational model is organized around the following components (see figure 4).
Thirteen roles:

- Investigator: Consists in elaborating an investigation report, informing the penal and delegating the triggering of alarm to the reporter,
- Penal: Consists in elaborating the juridical report,
- Reporter: consists in informing the controller of the incident arrival,
- Controller: consists in realizing the taking of sample from the pumping site in order to prepare the samples and delegate their analysis,
- Analyst: consists in analysing of samples, elaborating a quality report and informing the authority to trigger the crisis cell,
- Authority: consists in activating and inactivating the pumping site,
- Evaluator: consists in informing the scheduler of the risks evaluation result,

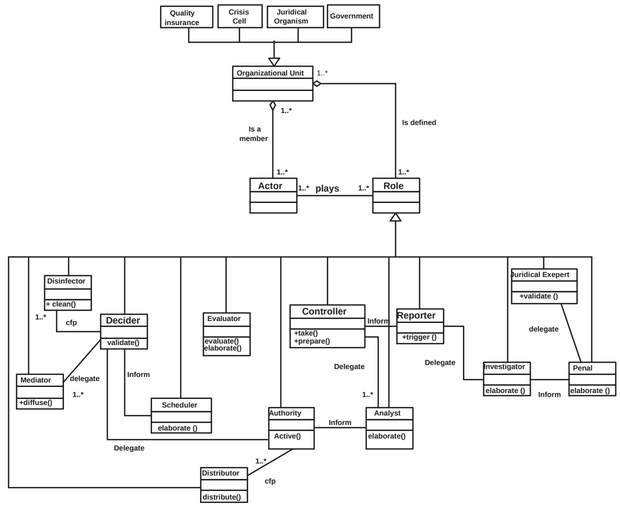


Fig. 4. The UML Classes diagram of the organizational model

- Scheduler: consists in informing the decider of the plans of actions,
- Decider: consists in validating one among proposed plans of actions,
- Distributor: consists in distributing of water bottles to the concerned users,
- Juridical expert: consists in validating the juridical report,
- Disinfector: Consists in cleaning the polluted site,
- Mediator: Consists in diffusing information to the concerned users.

Four organizational units:

- Quality insurance: charged to control the water quality,
- Crisis cell: charged to assure the return to the normal functioning of the Water distribution system,
- Juridical Organism: charged to take the juridical decisions,
- Government: charged to activate/inactivate the pumping site.

4.4 The Process model

The objective of the management process of Water Distribution crisis is to assure the return to the normal functioning of the Water distribution system. More precisely, this process is made of several coordinated activities and described using Petri Nets with Objects (see figure 5). Let us detail the activities (transitions) composing the process:

Investigation report establishment (A1) is realized by the investigator and consists in elaborating an investigation report containing all the information around the incident, the security perimeters and transport norms.

Alarm triggering (A2) is realized by the reporter and consists in triggering an alarm in order to inform the arrival of incident on water distribution system.
Juridical report establishment (A3) is realized by the penal and consists in establishing a juridical report describing the juridical decisions.
Juridical report Validation (A4) is realized by a juridical expert and consists in applying the juridical decisions taken by the penal.
Taking of samples (A5) is realized by a controller and consists in realizing Taking of samples on water distribution system, preparing of samples and sending them to the analysts.
Analysis of samples (A6) is realized by an analyst and consists in handling the samples and elaborating a quality report.
Risks Evaluation (A7) is realized by an evaluator and consists in elaborating an evaluation sheet containing the eventual risks.
Actions scheduling (A8) is realized by a scheduler and consists in proposing more actions plans to assure the return to the normal functioning of the water distribution system.
Actions plans validation (A9) is realized by the decider and consists in validation of one actions plan.
Disinfection of polluted site (A10) is realized by a disinfector and consists in realizing the cleaning work on the polluted site.
Information diffusion (A11) is realized by media (TV or Journal) and consists in informing the concerned users.
Inactivation of the site (A12) is realized by the authority and consists in interdiction the functioning of the water distribution system.
Water bottles distribution (A13) is realized by a distributor and consists in distributing water bottles to the concerned users.

4.5 The Instantiated Meta-Model

In order to illustrate the Workflow log meta-model instantiation, we propose to use the process model presented previously.
This instantiation is visualized in figure 6.

For instance, the grey part of this figure represents the contract net protocol employed between actors Malik, Walid and Mahdi (see figure C in table 2). According to this protocol, Malik launches by a call for dealing with given activity (Cfp). The other actors Walid and Mahdi propose their bid (Propose). Finally, Malik notifies each participant either by acceptation (Accept-propose) or by rejection (Reject-propose). This information is useful to understand how to allocate activities to actors.

4.6 The Organizational Structures and Interaction Protocols which can be discovered from our Workflow Log

According to the instantiated meta-model, we give here some organizational structures and interaction protocols that can be discovered. For each organiza-

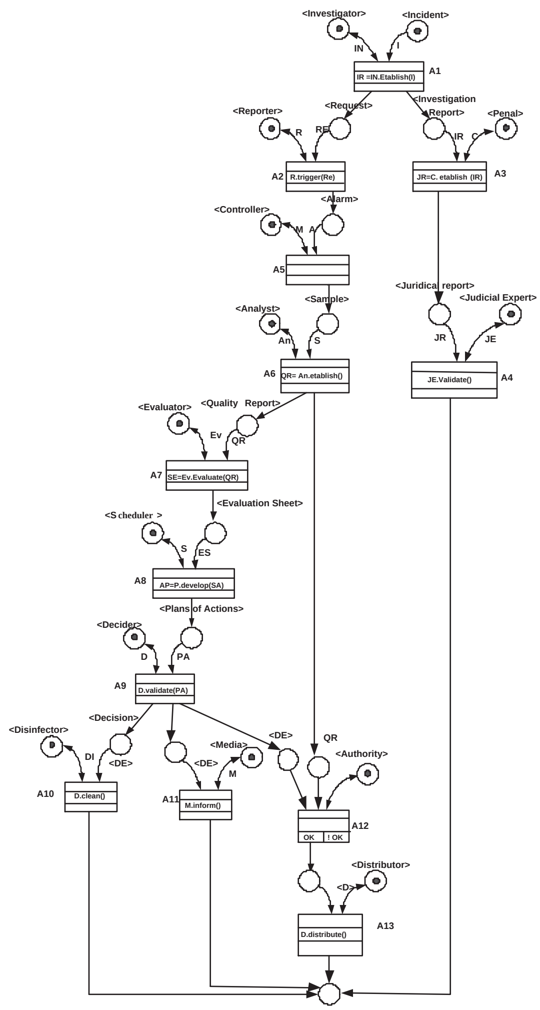


Fig. 5. The Process Model described using Petri Nets with Objects.

tional structure or interaction protocol we only give its definition and its associated graphic representation.

5 Discussion and Conclusion

The organizational perspective mining remains insufficiently addressed([10], [11]). Existing propositions in the literature are rather dedicated to the process perspective mining([3], [4], [5])and they neglect the important point that workflow is much more that process perspective. We believe that these works

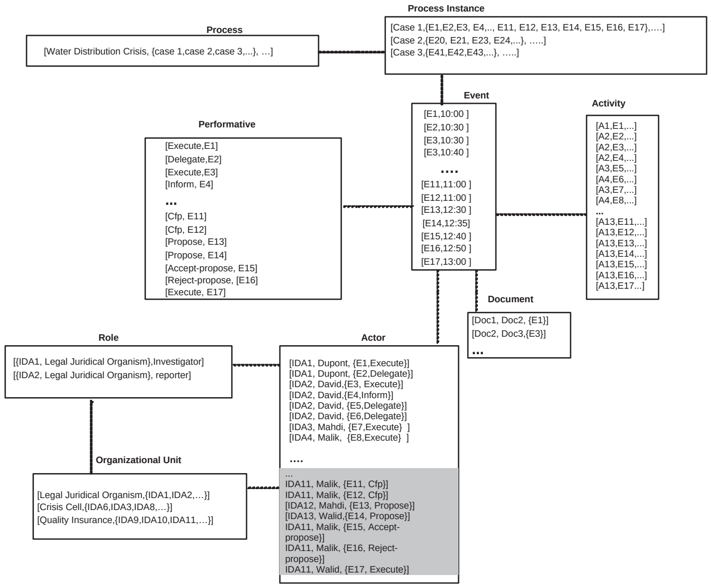


Fig. 6. The workflow log meta-model instantiation.

limit the scope and utility of workflow mining. For instance [3]only focuses on process mining by providing a Workflow log meta-model which defines the minimum concepts of process. [4]proposes a statistical technique to discover Workflow patterns from event log.[5]deals with the workflow mining in more workflow perspectives (process and informational perspectives) but without consider the organizational perspective.[10]defines three methods for mining of organizational structures from workflow logs such as default mining, mining based on the similarity of activities and mining based on the similarity of cases. Even if the proposed solution in [10]deals with the organizational perspective mining,it does not exploit the agent approach and as a consequence it does not support the interaction protocols and organizational structures so far discovered. [11]proposes an approach based on metrics expressing the relationships among actors deriving social networks. We believe our solution is currently unique in trying to take into account the three Workflow perspectives in unique model.
In this paper, we have presented an approach based on agent technology to deal with organizational perspective discovering.
To better illustrate our solution, we have chosen the well-known Management process of Water Distribution crisis case study.
The proposed meta-model is characterized by:

– its inclusion of three complementary workflow perspectives and as a consequence it supports any perspective discovering although in this paper we only focus on the organizational perspective,

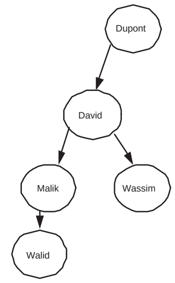
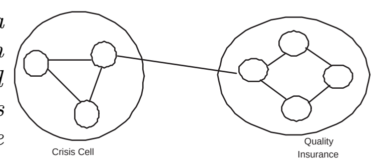
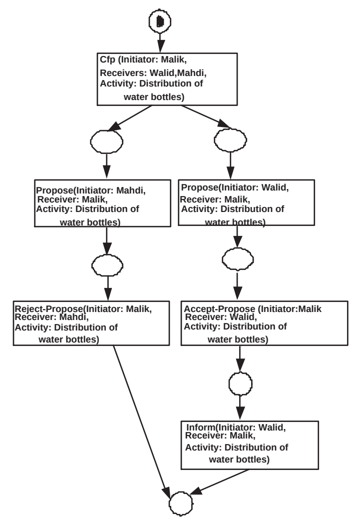
<table>
<tr><th>Organizational Structure or Interaction Protocol</th><th>Definition</th><th>Graphic representation</th></tr>
<tr><td>Simple Hierarchical</td><td>It's a flat hierarchy at single level. A single agent has the authority to catch decision while all other members are his subordinates.</td><td>A)

</td></tr>
<tr><td>Federation</td><td>A federation is a set of groups. Each one is represented by a delegate. This later plays the role of mediator between its group and the exterior.</td><td>B)

</td></tr>
<tr><td>Contract Net Protocol, CNP</td><td>In the CNP, one agent takes the manager role which desires to have one or more agents (the participants) to execute some activities. Each participant submits its bid. The manger selects the well offer and rejects the others bids.</td><td>C)

</td></tr>
</table>

Table 2. Some Organizational structures and interaction protocols that can be discovered

– its simplicity since it defines the core concepts of the three workflow perspectives,

– and finally, it exploits the Agent technology concepts (organization abstractions, FIPA-ACL performatives and interaction protocols) which ease the organizational structures and interaction protocols mining.

Actually, the DiscoopFlow (Discovering Organizational Structures and Interaction Protocols in WorkFlow) prototype development is in progress.
We have implemented some algorithms for discovering organizational structures like simple hierarchical and federation structures. The figure 7 gives an idea about the overview of our prototype.
As future work, we plan to implement the discovering algorithms concerning

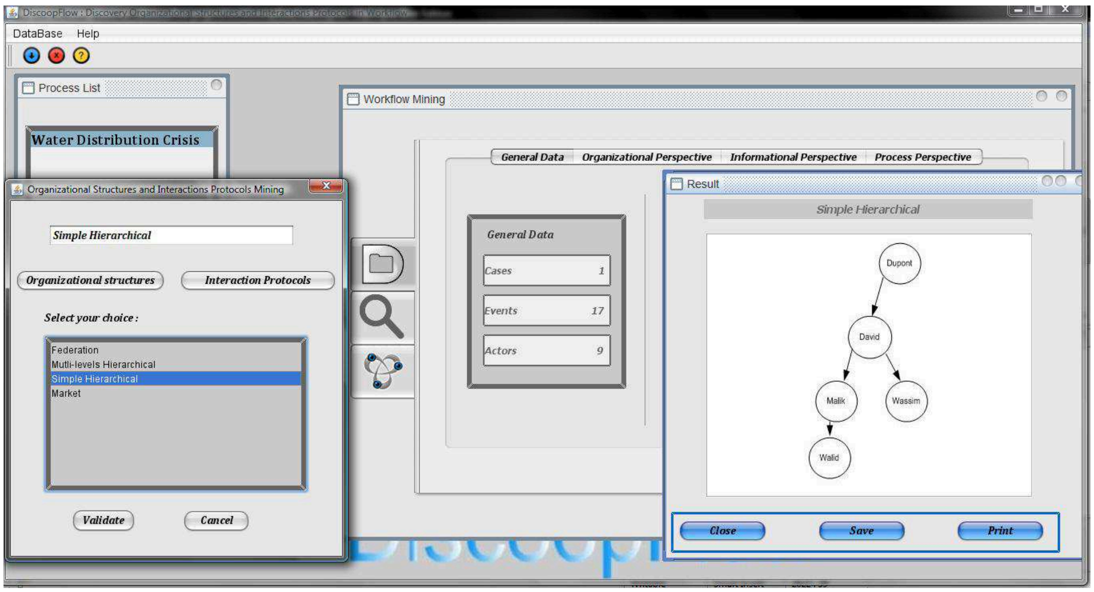


Fig. 7. Overview of DiscoopFlow prototype

interactions protocols and others organizational structures (market, multi-levels hierarchical, etc.).
We also aim to discover conjointly the three workflow perspectives in a coherent framework and extend our Workflow Meta model by adding the intentional perspective. We believe that this perspective justifies the choice of such execution scenario for a given crisis process.

Acknowledgment : The authors acknowledge all the participants involved in the DiscoopFlow project development. More precisely, they acknowledge Foued BEN MOALLEM, Achraf BOUDAWARA, Wassim CHTOUROU and Akram BEN AMOR.

References

1. Workflow Management Coalition.: The Workflow Reference Model. Technical Report WFMC-TC-1003, November (1994).
2. Van der Aalst, W.M.P., Van Dongen, BF., Herbst, J., Maruster, L., Schimm, G., Weijters, AJMM.: Workflow mining: A survey of issues and approaches. In Data & Knowledge Engineering, volume 47(2), pages 237 - 267, (2003).
3. Van Dongen, BF., Van der Aalst, WMP.: A Meta Model for Process Mining Data. In Advanced Information Systems Engineering, volume (161),(2005).
4. Gaaloul W., Godart, C.: Mining Workflow Recovery from Event Based Logs. In 3rd international conference on Business Process Management, BPM, Nancy,(2005).
5. Liu, P., Zhou, B.: Workflow Mining of More Perspectives of Workflow. In Software Engineering and Applications,in SciRes (www.SciRP.org/journal/jsea),Volume 1, Pages 83-87, (2008).
6. Agent Communication language Specifications. (http://www.fipa.org/repository/aclspecs.html).
7. Mathieu, P., Routier, J.C., Secq, Y.: Dynamic Organization of multi-agents systems. In proceeding of the first international joint conference on Autonomous Agents and Multi-Agent Systems, ACM Press, (2002).
8. Koning, JL., Pesty, S.: Modles de communication. Principes et architectures des systems multiagents. Chapitre 3, J. P. Briot, Y. Demazeau (eds), paris, Hrmes science publications, (2001).
9. Sibertin-Blanc, C.: High Level Petri Nets with Data Structure. 6th Int. Workshop on PetriNets and Applications Espoo (Finland),(1985) .
10. Gao, A., Yang, Y., Zeng, M., Le Zhang, J., Wang, YW.: Organizational Structure Mining Based on Workflow Logs. In International Conference on Business Intelligence and Financial Engineering, Beijing, China, July 24-July 26, (2009).
11. Van der Aalst, W.M.P and Song, M.: Towards comprehensive support for organizational mining. In Decision Support Systems, Volume 46, Issue 1, Pages 300-317, (2008).

Towards a Center for Modeling and Simulation: The Case for Jordan

Ghaith Rabadi, Ph.D.

Engineering Management and Systems Engineering
Old Dominion University
Norfolk, Virginia, USA
grabadi@odu.edu

Hazem Kaylani, Ph.D.

School of Technological Sciences
German-Jordanian University
Amman, Jordan
hazem.kaylani@gju.edu.jo

Abstract. Modeling and Simulation (M&S) has recently become an important area that is pursued by many researchers and practitioners due to the role it plays in understanding complex systems and problems. We have therefore witnessed the establishment of many M&S organizations in the last two decades especially in the more developed world. Less developed countries are starting to recognize the need for such capability especially that the problems they face are not less complex. In this paper, we present a preliminary study towards a business plan for establishing a scientific center for Modeling, Analysis, Simulation and Animation in Jordan (JoSAMA) and the value it can bring to the academic, industrial and governmental communities in Jordan and potentially in the Middle East. This effort was funded by the Fulbright Specialist Program and hosted by the German-Jordanian University, Amman, Jordan.

1 Introduction

It should be clear to everyone by now that information is among the most valuable and critical drivers for decision making and training. In fact, Information Technology (IT) in general has become almost an integral part of life in most of the world. Modeling and Simulation (M&S) is a field that has advanced significantly with the advancement in IT and has shown to be very useful in many applications especially the field of decision-making for complex systems. The Blue Ribbon Panel on Simulation-Based Engineering Science reported that "in the past decade, computer simulation has emerged as a powerful tool; one that promises to revolutionize the way engineering and science are conducted in the twenty-first century" [1].

Computer modeling can be defined as representing real existing (or future) systems or processes using computer software and hardware to develop better and deeper understanding of the behavior of these systems. *Simulation modeling* has become one of the most popular modeling paradigms as it can handle very complex and dynamic systems without making many crude assumptions [2]. Nonetheless, there are other modeling approaches including mathematical, analytical, graphical, and animated methods that may be deemed appropriate or necessary for certain applications at a certain level of fidelity. Such modeling approaches can also fall under the umbrella of *modeling* as they could be integrated into complex (simulation) models.

Experimenting with real systems could be costly, dangerous, and/or time consuming. Modeling systems can therefore become a much less costly alternative compared to the potential consequences of making suboptimal or wrong decisions directly on the real system. Computer simulation emerges as a powerful modeling tool as it mimics the behavior of the real system to enable the users experiment with the

J. Barjis, M.M. Narasipuram, G. Rabadi, J. Ralyté, and P. Plebani (Eds.):
CAiSE 2010 Workshop EOMAS'10, Hammamet, Tunisia, pp. 99-112, 2010.

simulated system before experimenting with the real one. The most common and important benefits of simulation include [3]:

- One of the best ways to study complex, dynamic, and stochastic systems
- One can easily run "what-if" analysis
- Does not disturb the real system
- Cheaper than experimenting with the real system
- Time is compressed
- No need for crude assumptions
- Animation gives us insight into what is happening
- Training using computer simulation, animation and Virtual Reality has proved to be very cost effective compared to training on real systems

The return on the investment of modeling a problem or a system may be quite significant given the potential consequences in real life. Over the last a few decades, we have witnessed numerous areas of civilian and military applications for M&S such as Manufacturing Systems and Processes, Logistics and Distribution, Emergency-Response System and Crisis Management, Homeland Security, Computer and Communication Networks, Transportation, Healthcare and Hospital facilities, Space Applications, and Queuing Systems (e.g., Banks, restaurants, parks, supermarkets, stadiums, etc).

The importance of the field is not in question here as it is clear that the area is growing rapidly especially in developed countries. What we argue in this paper is that in developing countries and regions such as the Middle East, the need for such a scientific engineering approach and the capabilities to support this approach both exist but they need to be organized into a viable and sustainable business model like the one proposed here in order to benefit the academic and industrial communities in these regions.

2 Vision and Objectives of the Study

The objective of this paper is to conduct a preliminary study towards a business plan for establishing a scientific center in **Jo**rdan for **M**odeling, **A**nalysis **S**i**m**ulation, **A**nimation, and Visualization (JoSAMA). In this paper, we propose an organizational structure for this center and the value it can bring to the academic, industrial and governmental communities in Jordan and the Middle East.

The proposed vision is for JoSAMA to be a center of excellence for Modeling, Simulation, Analysis, Animation and Visualization at both the national (Jordan) and regional (Middle East) levels. The center can potentially be an icon for (i) Modeling complex systems in different areas of applications (ii) Research and development in M&S by bridging the gap between academia and industry (iii) Training to support various industries, and (iv) Creating job opportunities.

3 Why Jordan?

M&S research and consulting communities in developed countries such as those in the U.S., Europe, and Australia have recognized the importance of this field in the last two to three decades and have done significant amount of development in terms of research, software, theory, and the necessary hardware to support M&S needs. This can be easily seen by inspecting the number of books, technical journals, software packages, conferences, consultants, and research centers around the world that arose in the past two decades. To focus on centers similar to some extent to the one proposed here, Appendix A includes a list of existing centers that conduct research, development and consulting in the area of Modeling, Simulation, Analysis and Visualization. In Jordan, and the Middle East, there does not exist a center or an organization that can capture the advancements and momentum in this field to generate products and tools that can benefit the academic and industrial community although the regional demand for such products exists. At this point, M&S knowledge, technology and research are being transferred and shared around the world, and Jordan has all the criteria to host a successful center in M&S for the following reasons:

1. Jordan has a strategic location in the Middle East to serve the region's needs in M&S. Since the products that will be produced by this center are software, reports, and information that can be electronically transmitted easily over the Internet to any location in the region and worldwide, the physical location may not be of a significant importance. However, being geographically close to clients is very important to maintain good client relations, face-to-face meetings when necessary, conducting demonstrations, and hosting and attending relevant events. Furthermore, neighboring countries that may need the technical expertise of the center (e.g., Saudi Arabia, United Arab Emirates, Qatar, and Iraq) add tremendous value to the business relations. Businesses in Jordan, especially in the IT sector, have been performing great work for neighboring countries like those mentioned earlier.

2. Human capital and talent are arguably the country's most important resource, as it does not enjoy as much of other natural resources. Jordan ranks at 91% literacy rate according to [4] and people in Jordan are highly educated in Engineering, Sciences, and other technical domains. There are currently 30 public and private universities in Jordan, a strong indication of people's awareness of the importance of higher education especially in technical fields given that the population of Jordan is estimated to be a little six million. Among these universities is the German Jordanian University (GJU), the university host of this effort, has strong ties to the industry and universities in Jordan, the Middle East, Europe and the U.S. Appendix B includes a list of universities that offer technical degrees and programs that can support the fields of Modeling, Simulation, Analysis, Animation, and Visualization. Almost all technical departments and programs in these universities teach in English, which makes their interaction and communication with the center and future clients and supporters much easier. Several universities in Appendix B are the result of partnership with leading universities in the U.S.A. and Europe.

3. Jordan is a modern and moderate country that enjoys excellent relations with states and countries in the U.S. and Europe, which can support advancing the M&S field in the Middle East through various funding programs. There are currently several international funding programs that Jordan can benefit from in establishing a M&S center and advancing the field in Jordan. Jordan also enjoys friendly business relations with countries in the Middle East and many of its expatriates work in these countries.

4. Jordan is also considered an attractive educational destination for students from the neighboring countries. Promoting a field like M&S and establishing a research center that will be affiliated with multiple universities will help the universities attract foreign students.

5. The top leadership in Jordan is keen on promoting the human capital and investing in higher education as it is the most precious resource the country has. For a relatively young field like M&S, the potential to attract interest, funding, and investors from the region will create more jobs in a specialized and technically advanced field.

4 Current State of Modeling and Simulation (M&S) in Jordan

The field of M&S is generally young worldwide including the Middle East and Jordan. However, since higher education in Jordan has been a focal point for Jordanians, the country now has many core pillars upon which fields like Modeling, Simulation, Analysis, Animation and Visualization can be founded in a much easier and quicker fashion than other countries in the region. There are currently 30 public and private universities in Jordan (see Appendix B for a complete list), the majority of which offer technical bachelor and masters programs such as Engineering, Computer Science, and Information Technology (IT). Most Engineering colleges that offer Industrial Engineering and Engineering Management teach simulation modeling and analysis, stochastic processes and statistical analysis for their senior students. Also some Computer Engineering and Computer Science Departments teach simulation-related courses such as Computer Modeling, Computer Graphics, Animation and Parallel Computing.

Jordanian universities, however, do not have academic programs in M&S; however, the technical courses they currently offer can be a very good starting point in establishing the discipline and supporting the center. Establishing a M&S center in Jordan will help generate momentum for this important field to help the country model and solve complex problems in various areas of applications for itself and the region.

On the industrial consulting front, there are many large and small companies that conduct various types of studies. However, it was not easy to survey which companies are involved in M&S projects. By reviewing the various technical academic programs in higher education institutions, it seems that the academic side can quickly adapt to support the field. On the industrial side, the main challenge is the lack of awareness of M&S importance and the value it can add to their businesses. After meeting with several key personnel at universities and industries, it became clear that the first step to successfully develop the field through this center is to start with a M&S awareness campaign for the industrial and academic communities. It is worthwhile noting that in the past two decades there has been some sporadic M&S activities in Jordan and the Middle East in the form of conferences such as:

- The First Middle East Workshop on Simulation and Modelling, 1999 at the University of Jordan, Amman, Jordan

- The International **M**iddle **E**astern **S**imulation Multiconference, 2008, at Philadelphia University, Amman, Jordan

- International **M**iddle **E**astern **S**imulation Multiconference, 2009, at Lebanese American University, Beirut, Lebanon

- The International **M**iddle **E**astern **S**imulation Multiconference, 2010, Alexandria, Egypt

- International Conference on Modeling, Simulation, and Applied Optimization, 2005, 2007 and 2009 in the United Arab Emirates

5 Short, Mid, and Long-term Objectives

The center must have clear objectives to achieve over its first 7 years and performance measures based on these objectives to track its success. In this section, we define these objectives keeping in mind that such objectives can be modified and refined by the Center Director and Advisory Board upon establishing the center and its bylaws. It is assumed that a Center Director with the appropriate academic background and industrial experience is selected early on to lead the effort and achieve the center's goals and objectives. It is also assumed that sufficient startup fund will be available for the center to have a reasonable level of resources to operate (e.g., space, equipment, personnel, etc). The amount of funds and its potential sources will have to be defined in a formal business plan that can be developed as a next step to this effort.

Short Term Objectives (1 – 2 years)

1. Equip the center with the necessary office and lab space, computer hardware and software necessary to achieve the center's short and mid term objectives.
2. Employ a full time Chief Scientist, a technical M&S simulation developer and an administrative assistant to perform the objectives listed above.
3. Conduct a M&S awareness campaign including developing a web site for the center. The campaign should be executed on two fronts: on the industrial front– the center will promote the field for potential clients, and at the academic front – the center will make the faculty and students

of various universities aware of the effort and encouragement to participate in supporting the center.
4. Run M&S short courses and workshops for various industrial sectors to spread awareness and generate revenue for the center.
5. Secure small and medium size funded project through various sources to demonstrate capabilities and generate more revenue.
6. Recruit at least 12 charter members and 12 corporate members
7. Assemble the Advisory Board members according to the structure proposed later in this paper.
8. Conduct industrial survey to collect information regarding to what extent the M&S tools are used by the industry and what potential areas of application are.

Mid Term Objectives (2 –3 years)

1. Secure major funding for R&D, Training, and Consulting divisions through marketing and competing for major agencies and companies. Acquire more hardware and software to support the projects and center as necessary.
2. Hire more technical staff as needed (Engineers, Simulation Analysts, Computer scientists, IT specialists, etc) to perform project work. The hiring should be consistent with the organizational structure discussed later.
3. Hold semi-annual M&S workshops for a small fee to strengthen the name of the center and establish industry confidence.
4. Participate in academic and industrial local and regional conferences and exhibits and present the work that has been developed at the center so far. Participate if possible in international conferences to keep the staff update with the latest technologies and research.
5. Conduct annual workshops to present successful projects and potential future applications.

Long Term Objectives (5 – 7 years)

1. The recognition of JoSAMA as a premier center of Excellence for Modeling, Simulation, Analysis, Animation and Visualization in Jordan and the Middle East.
2. Establishing a warehouse of M&S products that resulted from previous projects and which can be reusable in future projects.
3. Continuous self-sustaining growth by winning more local and regional consulting and research contracts from local and regional private industries, and governments.
4. Advancing academic research by partnering with different universities to recruit more skilled personnel and remain on the cutting edge of technology. This objective will be further supported by holding international conferences and workshops in coordination with local and international universities.
5. Develop new academic programs (B.S., M.S. and certifications) in partnership with various universities.
6. Stimulate technology-based economic development in Jordan and the region, which should lead to more job opportunities for Jordanians

6 A Proposed Organizational Structure of the Center

The proposed organizational structure in Figure 1 is based on the authors' research and experience with other similar centers in the U.S.A. The center's organizational units and functions are discussed next.

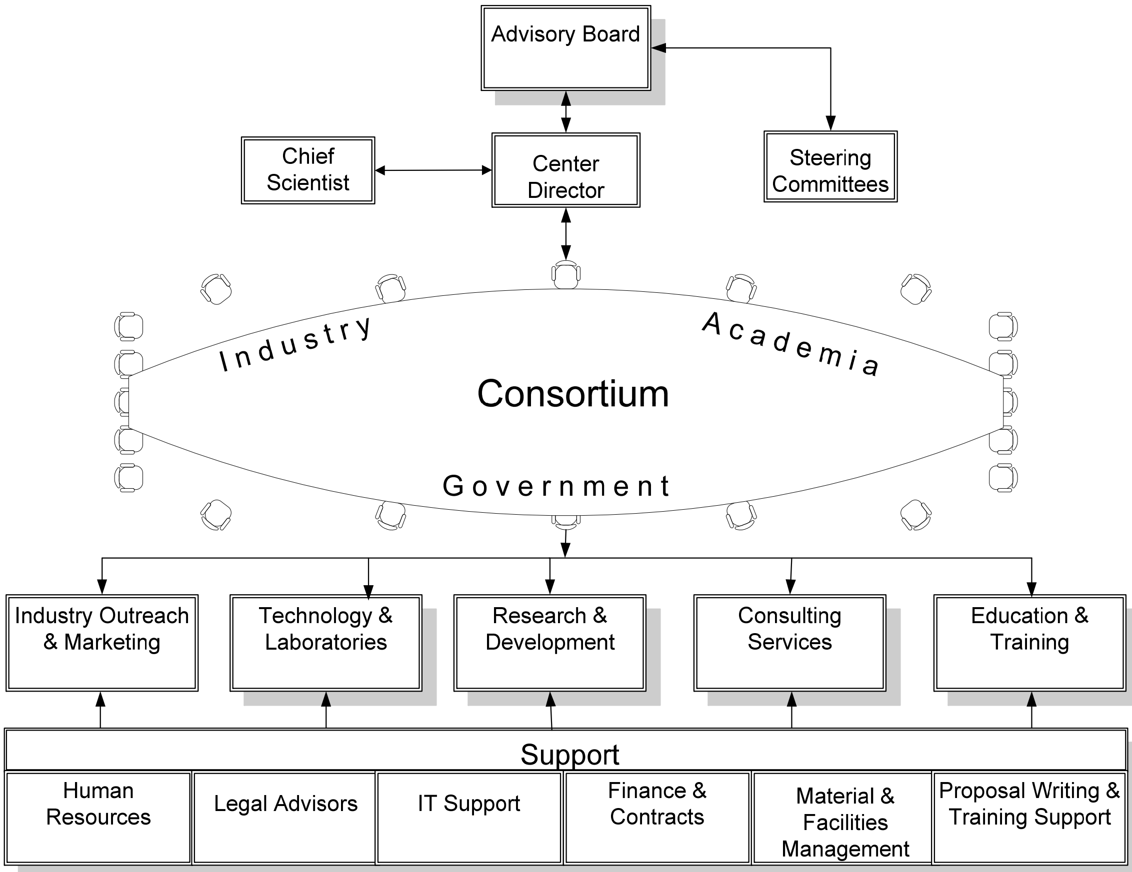


Fig. 1. Organizational Structure of the Proposed Center

<u>Advisory Board</u>

The main purpose of the Advisory Board is to advise the Center Director on various issues including actions and directions that the center should take to accomplish its objectives. The Advisory Board consists of 9 to 12 members, one of which is assigned by the Center Director to be the Board's Chairman. The Board shall be made up of members from large, medium, and small companies depending on the their revenue and number of employees. At the discretion of the Center Director, the Board may be augmented by 3 to 6 At-Large members who are invited to participate in Board meetings and activities for a certain period (e.g., one-year) and may serve for consecutive terms. A nomination and an election process by the board members and Center Director will be followed to assign new members and fill any vacancies. The initial Advisory Board can be assigned by the Center Director and a committee of stakeholders.

<u>Steering Committees</u>

JoSAMA can charter committees made up of individuals who are interested in certain aspects of modeling, simulation, analysis, animation and/or visualization. They should also have the necessary expertise to perform tasks assigned to them by the Director and Board, or they may undertake on their own. At least one of the members (preferably the committee chair) needs to be from the Advisory Board. The rest of the members can be from JoSAMA experts mixed with volunteers from outside the center. Committees serve important roles including (a) representing the views and opinions of their respective communities of interest to the Board and (b) using their influence to promote the implementation of resolutions. Using outside committee members is especially important as they can promote outreach and spread awareness about the field and the center.

<u>Academic Partners</u>

Technical universities will play a very important role in the proposed center as they will be the source of human knowledge and skill in the field. Very few universities around the world have academic programs that award full academic degrees in M&S. Old Dominion University is one of the few to offer B.S., M.S., Ph.D. and DEng (Doctor of Engineering) in M&S (http://eng.odu.edu/msgp/) as well as in various Engineering disciplines with M&S focus. While, the universities in Jordan currently do not offer such degrees, they collectively can offer courses that can become a strong foundation for M&S knowledge-base infrastructure for the center. If the center proved to be successful, it may in turn promote academic programs and degrees for the universities to sponsor in the future.

While the center has to have full time researchers and developers, the majority of the Ph.D. level personnel are going to be recruited from Jordanian universities as part time affiliates where they perform projects and supervise students and researcher at the center. This way, the university faculty can use the center's capacity (name and resources) to pursue projects and conduct more research in collaboration with the center. The incentive for the faculty members to do so will be to advance their research and interface with the industry to apply their knowledge, research and expertise to real-life problems in addition to receiving financial compensation for their work. The center's incentives, on the other hand, include having a sustainable source of knowledge and research, create more products based on the project work performed, and increasing its revenue for future growth and cover its cost. Moreover, the more projects are performed with university professors and researchers, the stronger the academic relationship will be.

In summary, there are mutual benefits for both the universities and the center to collaborate and succeed. This proposed approach to support the center with knowledgeable Ph.D. level professors and researchers has been implemented at almost similar centers including VMASC at ODU, which has been very successful at attracting grants and contracts for both the faculty members and the center. The academic members at VMASC are not limited to ODU where other universities and colleges participate and collaborate with VMASC. The same should be the case for JoSAMA where the partnership should not be limited to one university only.

The proposed center needs to act as a connecting point between academic institutions, government, and the private sector. Not only would all parties benefit from grants and contracts, but they would also enrich the research and knowledge in the field by having the academic institutions participate in offering degree programs and certificates in M&S to support the center's mission and increase their student credit hours. The list of universities that can teach courses and support the center's mission along with some of the relevant programs and courses are listed in Appendix B.

Member Consortium

The Consortium must include members from the Private Industry, Government and Academia. While academic institutions will play a significant role as partners to support the center, they can become consortium members like other members. Organizations from different sectors can become members for a fee that can be decided by the center's bylaws. There should be membership classes (e.g., Gold, Silver, Regular, academic and In-kind memberships) depending on the level of involvement and benefits that a member organization wishes to have. The exact roles and benefits of the members can be specified in center's bylaws; however, in general the members should have an influence on the work areas and applications and should also have access to some of the resources and expertise available at the center. The members will also provide access to the industry and government to help them compete for project work. Their feedback to the Advisory Board via the center director will be an important feedback loop that will help the center remain relevant.

Center Divisions

The center can consist from the following divisions, which are suggested based on the authors' research and experience with other similar centers:

1. Industry Outreach and Marketing
2. Technology and Laboratories
3. Research and Development

4. Consulting Services

5. Education and Training

The previous divisions with the appropriate personnel to run them should cover the project activities and objectives to be achieved by the center.

Support Functions

For the center to function effectively and progress towards achieving its objectives, it needs the following functional areas to support the different divisions and management team:

1. Human Resources

2. Legal Advisors

3. Information Technology Support

4. Finance and Contracts

5. Material and Facilities Management

6. Proposal Writing and Training Support

Conclusions and Future Work

In this paper, we presented a preliminary work on the need to establish a center for Modeling, Simulation, Analysis, Animation and Visualization in Jordan. Such a center can play an important academic and industrial role in the Middle East in a technical field that has witnessed significant momentum in the last two decades. The necessary human talent, geographic location, and demographic makeup, as well as its relations with developed countries, makes Jordan a very attractive for locating such a center. More work is still needed to make the business case however. The next step is to develop a business plan that can be presented to different stakeholder including government, local and international funding agencies, universities, and industries. The plan needs to show that the proposed center is viable and sustainable on the long run. Once finalized, the business plan in conjunction with proposals for funding should next be used to raise enough capital to start the center. The implementation plan can be executed over multiple phases that are consistent with the short, mid, and long-term objectives and timeline as discussed in the paper.

Acknowledgement

We are grateful for the Fulbright Specialist Program for funding the first author to conduct this study in Amman, Jordan. We are also thankful for the host institution, the German Jordanian University, for supporting both authors and providing the necessary data access and workplace during the duration of the project. Finally, we would like to thank Mike Mcginnis, John Sokolowski**,** Mike Robinson and the staff at Virginia Modeling, Analysis, and Simulation Center (VMASC) for providing valuable information on VMASC's bylaws, structure, history, and the lessons learned in the past 10 years [5], which gave the authors a lot of insight into this paper.

References

1. Oden, J.T., Belytschko, T., Fish, J., Hughes, T.J.R., Johnson, C., Keyes, D., Laub, A., Petzold, L., Srolovitz. D., and Yip, S.: Simulation-Based Engineering Science: Revolutionizing Engineering Science through Simulation. A Report of the National Science Foundation Blue Ribbon on Simulation-Based Science Engineering (2006).
2. Law, A.M., and Kelton, W.D.: Simulation Modeling and Analysis, 3rd ed., McGrow Hill, New York, NY (2000)
3. Kelton, W.D., Sadowski, R.P., and Swets, N.B.: Simulation with Arena, 5th edition. McGrow Hill, New York, NY, (2010)
4. United Nations Development Programme Report, Page 171 (2009)
5. Banks, C.M.: From Concept to Leader: The Virginia Modeling, Analysis, and Simulation Center: A Ten Year Narrative. 10-Year Anniversary Report, Old Dominion University, Norfolk, VA. (2008)

Appendix A: List of Selected Modeling, Simulation, Analysis and Visualization Centers

1. Virginia Modeling, Analysis and Simulation Center (VMASC)

   Old Dominion University, Norfolk, Virginia, U.S.A
   http://www.vmasc.odu.edu/

   Mission:

   - Engage in collaborative research and development in Modeling and Simulation (M&S) and in Visualization
   - Develop the M&S workforce through education
   - Provide M&S technical expertise to government and industry
   - Stimulate technology-related economic development in Hampton Roads and the Commonwealth of Virginia

2. Institute for Simulation & Training

   the University of Central Florida, Orlando, Florida, U.S.A
   http://www.ist.ucf.edu/

   Mission: Advance the state of the art and science of modeling and simulation by

   - performing basic and applied simulation research
   - supporting education in modeling and simulation and related fields
   - serving public and private simulation communities

3. Institute for Advanced Simulation (IAS)

   Jülich, Germany
   http://www.fz-juelich.de/portal/about_us/institutes_facilities/institutes/ias

   The Institute for Advanced Simulation unites Simulation Sciences and supercomputing under one roof. Thus, disciplinary, methodic and technological competences can be combined to manage the future challenges in the Simulation Sciences. It is considered one of the largest research centers in Europe.

4. Center for Advanced Studies, Research and Development (CRS4)

   Sardinia, Italy
   http://www.crs4.it/index.html

   CRS4 carries out applied research across a range of scientific and technological areas. Areas of expertise encompass large-scale computational problems requiring advanced simulation techniques supported by High Performance Computing, along with the timely integration of newly emerging Information and Communications Technology. It focuses on problems stemming from natural, social and industrial environments. The center co-operates with Industry, the Academic world and European Scientific Laboratories.

5. International Research Institute for Simulation, Motion and Navigation (SIMONA)

Delft University of Technology, Delft, Netherlands
http://www.simona.tudelft.nl/

The objectives of SIMONA include:

- Development of an international research institute in the field of simulation with an advanced flight-simulator and research facilities.
- Research in aspects of safe and economical transport.
- Combining and applying new methods and technologies for simulation and transport.
- Participate in national and international research projects.
- Accomplish collaboration between national and international universities, research institutes and industry.

6. McLeod Institute of Simulation Sciences
California State University, Chico, California, U.S.A.

http://www.ecst.csuchico.edu/~mcleod/overview.html

The original goal was to create a self-supporting, ongoing Institute supported in a variety of ways (funds, equipment, visiting personnel) by industry and state and federal contributions. Faculty and student efforts are supported by grants and contracts supporting research and training projects. The intent is to involve as many faculty and students as can be reasonably accommodated based on the space, equipment, and funds available. Activities of the Institute include: Applications of the High level Architecture (HLA), Intelligent software environments for simulation, Programming languages for simulation, Techniques for evaluating the performance of computer simulation systems, High-speed real-time simulation systems and the development of software for engineering education.

7. Modeling, Virtual Environments, and Simulation Institute (MOVES)
Naval Postgraduate School, Monterey, California, U.S.A.
http://www.movesinstitute.org/

MOVES mission is to enhance the operational effectiveness of the joint forces and their allies by providing superior training and analysis products, education, and exemplary research in the field of modeling and simulation.

8. Center for Modeling, Simulation, and Analysis (CMSA)
University of Alabama, Huntsville, Alabama, U.S.A
http://cmsa.uah.edu

CMSA's work centers on modeling and simulation and systems engineering. CMSA has broad capabilities in both of those disciplines, with special expertise in these areas: Physics-based modeling, Model validation methodologies, Spacecraft propulsion modeling, Simulation interoperability and composability, Discrete event simulation, Mathematical modeling and analysis, Finite element modeling and computational fluid dynamics, Modeling and simulation education, System-level modeling for design analysis and trade studies, Using modeling and simulation in the systems engineering process, Systems engineering methods, and Technical team performance.

9. The Center for Simulation and Modeling (SAM)

University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.
http://www.sam.pitt.edu/

SAM is dedicated to supporting and facilitating computational-based research across campus. Faculty across the University are using modeling and simulation to further their research. SAM serves as a catalyst for multidisciplinary collaborations among professors, sponsors modeling-focused seminars, teaches graduate-level modeling courses, and provides individual consultation in modeling to all researchers at the University.

Faculty and staff using simulation and modeling at the University come from a wide range of disciplines, including astronomy, biology, chemistry, economics, engineering, health, and medicine. Areas of research include: energy and sustainability, nanoscience and materials engineering, medicine and biology, and economics and the social sciences.

Appendix B: A List of Jordanian Universities that Can Support the Proposed Center

University Name	Relevant Departments	Courses (Course Number)
Ajloun National Private University	New and has the potential to add programs/courses that support M&S	None
Al al-Bayt University	Computer Information Systems	Modeling and Simulation for business (902352)
Al-Ahliyya Amman University	Computer Engineering	Modeling & Simulation (82589)
Al-Hussein Bin Talal University	Computer Engineering	
Al-Isra Private University	Computer Engineering	Simulation & Modeling (404542)
Al-Zaytoonah University of Jordan	Computer Science	Computer Simulation (0102441)
Amman Arab University	College of Computer Science and Informatics	Information is not available
Applied Science Private University	Industrial Engineering	Simulation (803482)
Arab Academy for Banking & Financial Sciences	No programs to support M&S	No courses to support M&S
Arab Open University	IT Department	Various IT Courses
Balqa Applied University	Master Degree in Computer Science	Simulation Methods (501001788)
	Communication & Electronics Engineering	Simulation & Modeling (404542)
Columbia University Global Center: Amman	New and has the potential to add programs/courses relevant to M&S	Currently None
DePaul University in Amman, Jordan	School of Computing	CSC 332 Simulation and Modeling and many other supporting courses
German-Jordanian University	Industrial and Management Systems	Simulation (IE 532)
Hashemite University	Computer Engineering	Modeling & Simulation (408360)
Irbid National University	No information is available about programs	No information is available about courses
Jadara University for Graduate Studies	Computer Science and Information Technology	Modeling and Simulation (501351)
Jerash Private University	College of Engineering	No information is available about programs/courses
Jordan Academy of Music	No relevant programs to M&S	No relevant courses to M&S
Jordan University of Science and Technology	Engineering: Industrial Engineering, Electrical Engineering and several others	Simulation (IE 512), Industrial Simulation (IE751), Modeling and Simulation of Nuclear Reactors (NE472), Modeling, Simulation and Analysis of Physical Systems (ME503)
	Computer and Information Technology: Computer Engineering, Computer Information Systems, Computer Science, Software Engineering	SIMULATION AND MODELING (CPE412), Modeling and Simulation of Business Cases (CIS385)
Mutah University	Industrial Systems Engineering, Computer Engineering, Engineering Management Program, Dept of Information Technology	No information is available about courses
New York Institute of Technology, Jordan	School of Engineering & Computer Sciences with several department s	Systems Simulation (IENG 425), other supporting courses

	including Engineering Management	
Philadelphia University	Computer Engineering	Modeling & Simulation (630573)
		Virtual Reality Systems (630582)
	Mechatronics	Modeling and Simulation (640645)
Princess Sumaya University for Technology	Computer Science	
	Computer graphics and animation	
	Computer Engineering	
	Communication Engineering	Simulation Tools (5244)
Red Sea Institute of Cinematic Arts	Many related programs and courses	Many related programs and courses
Tafila Technical University	Computer Engineering	Modeling and Simulation (0107552)
University of Jordan	Computer Science	Modelling and Simulation (1901352)
	Business Information Systems	Computer Simulation in Business (1903442)
	Industrial Engineering	Simulation, Advanced Simulation, Systems Simulation
University of Petra	Computer Science	
	Computer Information Systems	
	Programming Engineering	
	Computer Information Systems	Computer Simulation (0102441)
	Management Information Systems	Modeling & Simulation (303016337)
	Computer Engineering	Modeling & Simulation (302007547)
	Programming Engineering	Discrete Systems Simulation (301002481)
	Computer Engineering	Simulation and Modeling (CPE412)
		Simulation and Modeling LAB (CPE 412A)
	Computer Science	Simulation and Modeling (CS487)
Yarmouk University	Industrial Automotive Engineering	System modeling and identification (CE 623)
	Industrial Engineering	Simulation (403413)
Zarqa Private University	Computer Information Systems	Simulation by Computer (0306382)
	Programming Engineering	
	Computer Science	

Modeling Cross-Docking Operations using Discrete Event Simulation

Georges Arnaout[1], Elkin Rodriguez-Velasquez[1],
Ghaith Rabadi[1], Rami Musa[2],

1 Engineering Management and Systems Engineering Department,
Old Dominion University, Norfolk,VA 23529 USA
{ garna001, erodrig ,grabadi}@odu.edu
2 Agility Logistics, Atlanta, GA, 30339, USA{rmusa@vt.edu}

Abstract. Cross-docking is a distribution system where products transported from warehouses to destinations are unloaded temporarily at intermediate facilities, and loaded onto outbound trucks to be shipped to their final destination. The complexity of cross-docking operations has been considered a problem of great interest to researchers and practitioners in the areas of Scheduling, Optimization, Supply Chain Management, and Operations Research. In this paper, a cross-docking simulation model is proposed to study the cross-docking problem. This problem has been addressed in the literature previously, but was considered in a deterministic fashion and with several strong assumptions. The contribution of our study is to present a stochastic discrete-event simulation for the problem where several scenarios will be studied aiming at relaxing some of the assumptions set by previous work. Finally, the study reveals some of the most important parameters that must be investigated in a cross-docking problem, and analyzes their significance on the final results.

Key words: Scheduling, Cross-Docking, Discrete-Event Simulation, Supply Chain Management.

1 Introduction

As companies try to cut down shipping costs and provide a better service to their customers, the attention of supply chain managers turned to cross-docking. Cross-docking is described as the flow of material directly from receiving to shipping, where the goal is minimal handling and virtually no storage [1]. This new method of distribution management has gained a great reputation in helping companies cut their overall shipping costs and achieving better service levels.

As cross-dock operations were first pioneered in the US trucking industry in the 1930s, they became exceptionally popular after the great success of Wal-Mart utilizing cross-docking in the retail sector in the 1980's [2].

Cross-docking is the practice of routing trucks from origin to destination either directly or indirectly through cross-dock facilities (CDF). At the CDFs, trucks unload their shipments and become available for other requests, and the shipments are consolidated and loaded onto other outgoing trucks in order to reach their final destinations. The customer is known in advance before the product arrives to the CDF so that the storage is minimized or eventually eliminated [3].

In this study, we propose a discrete event simulation model of cross-docking operations where orders are placed randomly through three different warehouses. Trucks transporting different orders are routed from a warehouse to a destination directly or indirectly through CDFs. Orders to be shipped are subject to delivery time constraints. The service level, truck utilization, the number of CDFs, and the number of trucks scheduled are taken into account.

The paper is organized as follows: In section 2, a literature review about cross-docking simulation models is presented. In section 3, we present a brief description of our proposed simulation model with our assumptions. Section 4 describes the technical details of our simulation model. In section 5, we present our

J. Barjis, M.M. Narasipuram, G. Rabadi, J. Ralyté, and P. Plebani (Eds.):
CAiSE 2010 Workshop EOMAS'10, Hammamet, Tunisia, pp. 113-120, 2010.

different scenarios and the results accordingly. Finally, conclusions and further research are presented in section VI.

2 Literature Review

Starting from the earliest research that tackled cross-docking simulation problems, Rohrer [1] explained how simulation helps ensuring the success of cross-docking operations. The paper that targeted simulation practitioners introduces metrics that may help analyzing cross-docking problems efficiently. The problem of allocating the incoming and outgoing trucks to the doors of a CDF has been previously studied by simulation approaches. In 1999, Gue [4] proposed a rule where a rank of doors is assigned to each incoming truck based on the load's weight and the distances to each door. Regan et al [5] proposed a rule based on the transfer time of the order from incoming to outgoing door. This problem has been studied in the Optimization area as well by several researchers. Several scheduling problems have been indentified and different approaches have been proposed to solve it, see for example [5], [6], and [7].

Maglabeh et al. [8] introduced a simulation model of a CDF having several suppliers and destinations. The study focuses on the internal CDF operations (inbound and outbound trucks, consolidation, loading/unloading, etc) but does not address the aspects of the policies concerning the scheduling and the assignment decisions made by the dispatcher.

Cross-docking problems have been addressed by many researchers, but mostly in a deterministic fashion and with several strong assumptions. Deterministic cross-docking optimization problems were studied by [9] and [10]. This study proposes a discrete-event simulation model to attempt solving the cross-docking problem and evaluate different scenarios where some of the assumptions found in the literature are relaxed.

3 Model Description

In cross-docking, orders with specific load size are placed by customers from specific warehouses, where trucks pick the orders and unload them at the CDFs. At the warehouses, the orders are generated according to uniform schedules (depending on the scenario being modeled). Additionally, order attributes are assigned to each order upon generation. The attributes are: order id, order size, order due date, order destination, and order release time. The orders are then sorted and consolidated at the CDF to be ready for pick up by outgoing trucks to the destination of the customer who placed the order initially. At the CDF, the orders remaining slack time are updated (by subtracting current time t from order release time). The orders are finally shipped to the final destination according to the orders' attribute "order destination", and by priority depending on the updated slack time.

In our model, we are assuming three warehouses (origins), and a number of CDFs and a number of trucks that change according to the scenario being studied. Additionally, we are assuming three different standard truck sizes of 20, 40, and 53 units and truck types will be used according to the scenario chosen. A truck cannot be preempted during its trip to a certain CDF, origin, or destination, and becomes available right after offloading its carried load at a CDF or a destination.

The assumptions used in our model are the following:

1. Truck availability is taken into consideration. After offloading a load, the truck becomes available and free to be assigned to any other location.
2. All the products being transported are assumed to have a similar shape.
3. The origins and CDFs have infinite capacity (i.e. no queuing delays will occur at these locations) – as this assumption may seem unrealistic, the next version of our study will relax this assumption and focus on studying the overall behavior with queuing delays occurring at the CDFs.
4. The loading/unloading delays at the warehouse and CDFs are considered to be negligible.

5. The inventory at the warehouses is considered infinite. The demand is always met without any production delay.
6. All order sizes are generated according to a triangular distribution where orders vary from 5 units (minimum) to 53 units (maximum). In the case of an order size being equal to the requested truck size, the order is shipped directly to its final destination without passing through CDFs.
7. The three truck types have three different sizes.

4 Discrete Event Simulation Model

In our cross-docking simulation model, we use discrete event simulation in order to capture the effects of the chronological sequence of events occurring at specific times, and discretely changing states. The simulation model proposed was created using the simulation package Arena 10.0 that uses a graphical user interface (GUI) for SIMAN.

Our simulation model consists of three warehouses (or origins) and five customers (or destinations). Trucks carrying requested loads from origins to destinations pass through CDFs which can be 0, 2, or 4 CDFs depending on the scenario being studied. Trucks can move directly from origin to destination (and vice versa), from origin to CDFs (and vice versa), and from CDFs to destinations (and vice versa), as shown in Fig 1.

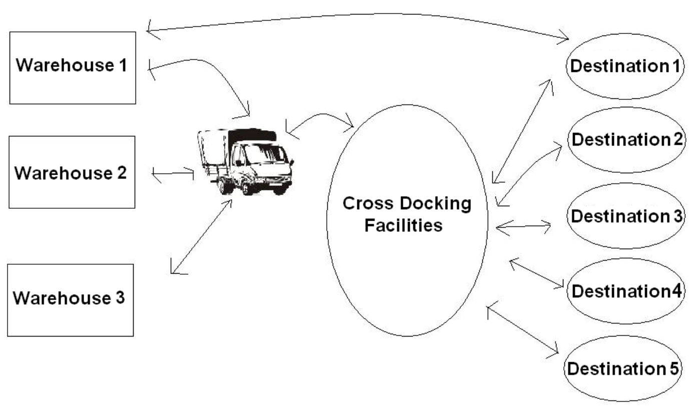


Fig. 1. The different routings of a truck.

When an order is placed by a customer, the order size, the supplier, and the due date are assigned as order's attributes. Following assumption number 5 in section 3, no delay occurs in generating the order and the load is ready to be picked up by a truck as soon as the order is placed. Orders are picked up by the closest available trucks according to the closest due date (orders with earlier due dates have higher priority). Once a load is picked up, if the load size is less than the truck size, it is transferred to the cross-dock representing the shortest route between origin and destination. Our assumption in choosing the cross-dock solely based on distance is that the origins, destinations, and CDFs are static in nature, thus, the distance between them remains the same throughout the entire model's lifecycle. If the load size is equal to the truck size, then the order is shipped directly from origin to destination without passing through CDFs. Note that the distances are generated randomly between each node in our system, following a general rule that the direct distance from origin to destination is larger than the total distance from origin to destination through an optimal cross-dock. The generated distances and the optimal cross-docks are discussed in more details in section 5.

When orders reach the CDFs, they are unloaded and batched with other orders according to the final destination. As soon as a truck is unloaded, it becomes available for other requests. Following our third and fourth assumptions in section 3 respectively, the CDFs have infinite capacities (i.e. no queuing delays will occur resulting from a CDF oversaturation) and loading/unloading delays are negligible.

When orders reach their final destination, they are disposed from the system, and necessary statistics are recorded. The orders at the CDFs are decided according to a priority rule using the following Slack function:

$$\text{Slack} = \text{DueDate} - \text{CurrentTime} \qquad (1)$$

where DueDate is the due date of the order placed by the customer, and CurrentTime is the current time of the simulation clock. Higher priority is given to orders with smaller slack. Note that the CDFs and origins compete for the same resources which are the available trucks.

Controlling the flow of operations is critical in cross-docking. Cross-docking demands that an order must be delayed (intentionally) at the CDF, in order to wait for other incoming orders that could be scheduled to be shipped to the same destination. In that case, the orders are batched together in the same truck (according to their final destination) and depart from the CDF to the customer. Therefore, we delay the truck at the CDF using the following intended delay formula:

$$\text{Intended delay} = \text{DueDate} - \text{CurrentTime} - \text{SafetyFactor} \qquad (2)$$

where SafetyFactor is the time amount that can be tolerated before a truck must depart the CDF in order to avoid missing the order's due date.

5 Experiments and Results

The experiments and different scenarios are conducted with the same distances and same average truck velocity of 55 miles per hour. All distance units are in miles and time units are in hours. In the first scenario, we assumed a constant rate of 4 orders per hour placed through each of the three warehouses. This excessively high order rate was chosen deliberately in order to test our system in the worst case scenario. We present in Table 1 the order size and the orders' due dates placed by each customer that are generated according to triangular distributions (i.e., TRI (minimum, most likely, maximum)). We used triangular distributions for the order sizes and due dates since there is only limited sample data concerning those parameters. The values of the distributions (minimum, most likely, maximum), were chosen based on a knowledge of the minimum and maximum possible values, or simply by "inspired guess" [1].

Table 1. Orders Generation Attributes

Destination (customer)	Order Size	Due Date
1	TRI(11,21,33)	TRI(25,30,50)
2	TRI(10,40,53)	TRI(25,30,50)
3	TRI(5,10,20)	TRI(25,30,50)
4	TRI(8,25,35)	TRI(25,30,50)
5	TRI(20,40,53)	TRI(25,30,50)

In Table 2, we present the optimal cross-docks selection according to randomly generated distances in the case of having four CDFs. The optimal cross-docks in the case of having two CDFs are presented in Table 3.

The running time for the simulation was set at 168 hours (7 days a week, 24 hours a day). After running the experiments for 100 replications, we obtained the results shown in Table 4. Two metrics are used in all of our experiments to investigate the model's performance of Truck Utilization and the order's Service Level. Note that orders received on Sundays are supposed to be produced and transported for the next week (i.e. their statistics were ignored in our study).

Table 2. Optimal cross-dock allocation in the case of 4 CDFs.

Origin	Destination	Optimal CDF	Total Distance
1	1	4	56
1	2	1	67
1	3	1	72
1	4	4	62
1	5	2	65
2	1	3	80
2	2	2	70
2	3	3	65
2	4	3	67
2	5	2	60
3	1	3	60
3	2	4	74
3	3	3	45
3	4	3	47
3	5	3	50

Table 3. Optimal cross-dock allocation in the case of 2 CDFs.

Origin	Destination	Optimal CDF	Total Distance
1	1	1	62
1	2	1	67
1	3	1	72
1	4	1	87
1	5	2	65
2	1	1	85
2	2	2	70
2	3	1	95
2	4	2	87
2	5	2	60
3	1	1	74
3	2	1	79
3	3	1	84
3	4	2	97
3	5	2	70

When having a small number of trucks (i.e. 2 trucks available), the truck utilization tends to be at its highest. For the case of having no CDFs, the average service level is low with two trucks being scheduled, and tends to increase when we assign more trucks with larger size. For instance, the service level increased from 57.52% (two size-20 trucks scheduled) to 100% (four size-40 or size-53 trucks scheduled). In the case of having two CDFs available, having chosen the safety factor variables according to the quality of the results generated, we basically notice an improvement in the system as the service level increases (82.52% if two size-20 trucks were scheduled and 100% if four trucks were scheduled of size-40 or size-53). If we average the service level results according to the number of CDFs, the best scenario would be to have 2 CDFs, having an average service level of 91.49%, compared to a service level of 91.11% (with 4 CDFs), and 81.52% (with 0 CDFs).

For all the cases in scenario 1 in Table 4, the truck utilization was extremely high resulting in more gas spending and higher level of carbon emissions (due to the excessive demand). However, the truck utilization was lower in cases having active CDFs especially with higher number of trucks. In the 2 active CDFs case, the truck utilization was 80.46% when scheduling 4 trucks. In the 4 active CDFs case, the truck utilization was 80.1% when scheduling 4 trucks. However, in the 0 active CDFs case, the truck utilization stayed high at 97.07% when having the same number of trucks scheduled.

Table 4. Scenario 1 results (constant rate of 4 orders/hour).

Number of CDFs	Truck Size	Number of Trucks	Safety Factor	Avg. Order Cycle	Truck Utilization	Service Level
0	20	2	n/a	36.9	100	57.53
0	40	2	n/a	30.78	100	66.52
0	53	2	n/a	30.78	100	66.52
0	20	3	n/a	21.87	99.89	79.77
0	40	3	n/a	18.43	99.69	84.44
0	53	3	n/a	18.43	99.69	84.44
0	20	4	n/a	12.17	99.23	94.42
0	40	4	n/a	4.89	96.07	100.00
0	53	4	n/a	4.89	96.07	100.00
2	20	2	12	21.66	100	82.52
2	40	2	12	19.45	100	83.92
2	53	2	30	25.45	99.91	74.32
2	20	3	30	14.73	99.7	91.55
2	40	3	30	9.4	98.91	99.07
2	53	3	30	12.79	98.2	96.17
2	20	4	15	12.5	98.9	95.90
2	40	4	12	4.97	89.11	100.00
2	53	4	12	6.33	80.46	100.00
4	20	2	30	23.43	99.85	81.06
4	40	2	30	16.36	99.77	89.98
4	53	2	15	25.34	98.79	78.99
4	20	3	15	17.1	98.63	90.08
4	40	3	15	11.1	98.66	94.92
4	53	3	30	15.12	98.36	90.54
4	20	4	15	11.91	97.02	95.20
4	40	4	15	7.99	83.99	99.22
4	53	4	15	10.66	80.18	100.00

In the second scenario in Table 5, the same experiments are conducted but according to a uniform schedule of 3 to 6 orders per day (no orders are placed on Sundays). Clearly, in this scenario, the demand is less than the previous scenario. The results are shown in Table 5.

The service level in this scenario is evidently higher than the previous scenario since the demand is significantly less. In the experiments where two trucks were available, the system having active CDFs outperformed the one without any active CDFs. In the experiments of having three or four trucks available, the results varied depending on the number of active CDFs and the safety factor parameter.

It is worthy to note that in both scenarios, when having 0 active CDFs, the results did not vary between assigning trucks of size-40 and size-53. However, it was noticed that having larger truck types assigned always yields better results.

6 Conclusions and Future Research

The proposed discrete event simulation model of cross-docking operations, where trucks must be routed from origins to destinations, is presented. The stochastic nature of the system allows us to analyze different scenarios and to reveal the importance of the model's parameters such as the safety factor, the number of trucks assigned, the types of the trucks (according to their size), and the number of active CDFs.

Two scenarios were conducted, one having a constant rate of 4 orders/ hour requested through each supplier, and a second scenario having a uniform schedule of 3 to 6 orders per day (excluding Sundays). The results of the experiments conclude that the use of cross-docking configurations produces higher service levels if the right safety factor was chosen. Additionally, truck utilizations tend to be lower when using CDFs, which according to our assumptions will result in lower level of trucks' carbon emissions. Finally, we conclude that larger trucks result in better service levels and overall better performance of the system.

Table 5. Scenario 2 results (schedule of 3 to 6 orders per day).

CDF	Truck Size	Number of Trucks	Safety Factor	Avg. Order Cycle	Truck Utilization	Service Level
0	20	2	n/a	28.11	99.83	66.99
0	40	2	n/a	24.15	99.68	77.65
0	53	2	n/a	24.15	99.68	77.65
0	20	3	n/a	15.39	98.34	89.88
0	40	3	n/a	10.19	95.86	97.15
0	53	3	n/a	10.19	95.86	97.15
0	20	4	n/a	7.35	88.63	99.10
0	40	4	n/a	5.02	79.59	99.99
0	53	4	n/a	5.02	79.59	99.99
2	20	2	15	18.04	99.78	87.07
2	40	2	15	13.38	99.69	92.12
2	53	2	15	17.35	98.93	88.65
2	20	3	15	13.06	98.07	94.01
2	40	3	15	8.23	88.66	98.69
2	53	3	15	9.87	88.42	99.57
2	20	4	17	7.03	87.57	99.66
2	40	4	17	5.37	58.77	100.00
2	53	4	10	8.16	56.68	100.00
4	20	2	30	18.29	99.4	86.12
4	40	2	15	12.69	97.65	93.45
4	53	2	15	20.55	97.61	88.36
4	20	3	15	15.76	95.21	91.60
4	40	3	15	9.9	87.45	97.95
4	53	3	15	12.36	79.69	99.08
4	20	4	15	10.96	84.21	96.64
4	40	4	15	8.70	64.51	99.32
4	53	4	15	9.86	62.28	100.00

Note that in both scenarios, the service level and truck utilization varied depending on the following variables of the model: number of CDFs, number of trucks available, value of the safety factor, and the type of trucks used. This implies that a more extensive study is needed in order to choose the optimal safety factor, the optimal number of trucks assigned, and the most favorable truck assignment of truck types (depending on their sizes) in the system.

As for future research, the authors plan to extend the proposed simulation model by making it more generic allowing the user to control the number of CDFs, the number of warehouses, and the number of customers at the initiation of the simulation. As a second step, more assumptions are to be relaxed making the model more realistic. Possible follow-on work for the model includes the possibility of adding different product types and different due dates.

References

1. Rohrer, M.: Simulation and Cross Docking. In: 27th conference on Winter Simulation Conference, pp. 846--849. IEEE Computer Society Washington, DC. (1995).
2. Apte, U.M., S. Viswanathan: Effective Cross Docking for Improving Distribution Efficiencies. International Journal of Logistics Research and Applications: A Leading Journal of Supply Chain Management, 3(3), 291—302 (2000)
3. Li, Y., Lim, A., Rodrigues, B.: Crossdocking - JIt scheduling with time windows. Journal of the Operational Research Society. 55, 1342--1351. (2004)
4. Gue, K.R.: The effects of trailer scheduling on the layout of freight tenninals. Transportation Science, 33, 419--428 (1999)

5. Wang, J.F., Regan, A.: Real-time trailer scheduling for crossdock operations. Transportation Journal. 47(2), 5--20 (2008)
6. Soltani, R., Sadjadi, S.J.: Scheduling trucks in cross-docking systems: A robust meta-heuristics approach. Transportation Research Part E. To be published.
7. Ley, S., Elfayoumy, S.: Crossdock scheduling using genetic algorithms. In: 7th IEEE Int. Symposium in Computational Intelligence in Robotics and Automation, pp. 416--420. Jacksonville, FL, (2007)
8. Magableh, G. M., Rossetti, M. D., Mason, S. Modeling and analysis of a generic cross-docking facility. In: 37th Conference on Winter Simulation, pp. 1613--1620. Orlando, FL (2005)
9. Derbes, D. Rabadi, G., Musa, R., Transportation problem for a cross-dock network with time constraints. In: Industrial Engineering Research Conference, pp. 1393--1398. Miami, FL, (2009)
10. Musa, R., Arnaout, J.-P., Jung, H.: Ant Colony Optimization algorithm to solve for the transportation problem of cross-docking Network. Computers & Industrial Engineering. To be published.
11. Weisstein, E.W., Triangular Distribution, MathWorld.

Feasibility Study Inputs based on Requirements Engineering

Robert Pergl

Department of Information Engineering,
Faculty of Economics and Management,
Czech University of Life Sciences,
Prague, Czech Republic
pergl@pef.czu.cz

Abstract. Theoretically, every software project can be successful if it has unlimited resources and does not care about the profit. Because this is not true in practice, the feasibility study is an important step in the software project initial phase. To achieve a valuable analysis, it is important to identify crucial aspects related to the feasibility. Most of the aspects come from the software product requirements.
An original method for identifying and documenting inputs to a feasibility study is presented in this paper. The method takes the requirements on the software product and provides a structured quantified framework for analysis of the requirements' impacts on the project infrastructure needs. The method formalism is based on the theoretical background of systems theory and modelling.

Key words: feasibility study, requirements engineering, software engineering, information systems development, managing software projects

1 Introduction

The *feasibility study* (or the *analysis of alternatives*[1]) is used to justify a project. It compares the various implementation alternatives based on their economic, technical and operational feasibility [2]. The steps of creating a feasibility study are [2]:

1. Determine *implementation alternatives.*
2. Assess the *economic feasibility* for each alternative. The basic question is "How well will the software product pay for itself?" This is decided by performing a cost/benefit analysis.
3. Assess the *technical feasibility* for each alternative. The basic question is "Is it possible to build the software system?". The set of feasible technologies is usually the intersection of the following main aspects:

[1] In many corners of the software and other system development universes, the term, "cost/benefit analysis" or CBA, has been supplanted by "analysis of alternative", itself a term encompassing not only economic feasibility (and, hence, also the practice of CBA) but technical and operational feasibility, as well.

J. Barjis, M.M. Narasipuram, G. Rabadi, J. Ralyté, and P. Plebani (Eds.):
CAiSE 2010 Workshop EOMAS'10, Hammamet, Tunisia, pp. 121-132, 2010.

  - Requirements on the technology.
  - Available licences and their costs.
  - Abilities of the developers and maintainers to master the technologies.
  - Maturity of the technology, its support.
  - Technologies to cooperate with / integrate with.
4. Assess the *operational feasibility* for each alternative. The main question is "Is it possible to maintain and support this application once it is in production?"
5. *Choose an alternative.*

We will let aside the operational feasibility which is not so tightly related to the requirements engineering compared to the remaining feasibilities (economic and technical). The requirements analysis implies many input parameters to the feasibility study and the feasibility study implies input parameters to the Define Infrastructure stage [2] (Figure 1).

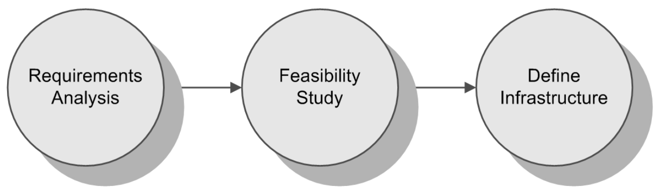


Fig. 1. Data flow from requirements analysis to the Define Infrastructure stage

In an ideal (but a rather rare) case, the project infrastructure fits all the project needs: the process is fine-tuned, all the necessary technologies are available and ready and all the team members have all the needed skills and knowledge. However, typically something is missing in this mosaic. For the project to be successful, a certain *adaptation* to the implied requirements must take place. The role of the method presented in this contribution is to help to identify and quantify such adaptation needs.

The starting point for the method is the demand that the transformation of the requirements to the feasibility study (further denoted as **RFST**, requirements-feasibility study transformation) should have the following attributes:

- *structured*, so that the logic of the transformation is easy to comprehended,
- *recordable*, so that the conclusions may be verified and/or used for increasing the evaluation accuracy in the future,
- *traceable*, so that the conclusions may be analysed and audited.

Here follows an original method based on the analogy between the software project and the systems theory and modelling that provides a quantified structured framework for this transformation.

2 The Method's Formal Background

The method is based on the analogy between systems theory and software project. A formal model of a general system consists generally of inputs, outputs, inner elements and relations [6]. Inputs are divided into endogenous ones, which are inputs crucial for the system model and exogenous, which are other inputs that must be taken into account. The analogy between the general system model and a software engineering project is shown in the Table 1[2].

[2] Not all the terms in the table are important for the goal of the contribution, however we find the analogy very inspiring and thus we wanted to present it in its fullness

Systems modelling term	**Software project analogy**	**Set symbol**
endogenic inputs (crucial inputs interesting for the modelling)	explicit software product requirements	I_S
exogenic inputs (other inputs)	external conditions both predictable and unpredictable (environment)	I_E
inputs (union of endogenic and exogenic inputs)	all external factors and impacts influencing the project	$I = I_S \cup I_E$
outputs	– software product and its parametres, – technology environment for running the product, – documentation and other artefacts, – trainings	O
inner elements	– team (project roles), – subcontractors, – tools (both development and supporting), – artefacts (code and documentation)	P
inner relations (relations between inner elements)	– process, – project management, – intra-team communication, – subcontractors communication.	R_I
relations from inside to outside	information to the customer	R_{IO}
relations from outside inside	information about the requirements changes	R_{OI}
relations from inside to outside and to inside	cooperation requests to customer from the team	R_{IOI}
relations from outside to inside and outside	team responds to immediate customer requests for cooperation	R_{OIO}
relations (union)	all relations	$R = R_{OIO} \cup R_{IOI} \cup R_{OI} \cup R_{IO} \cup R_I$

Table 1: Analogy between the general system model and a software engineering project

Inner elements may be based on the concrete purpose and goal of the model

– objects,
– classes.

Objects are chosen in the case when it is necessary to model the project system in a detailed level of single objects: documents, team roles, etc. In general methodological models, classes will be usually used. Each element then represents a whole class, not an individual object: e.g. "programmer" represents all the programmers; We do not care about structure and dynamics of objects inside the class. For the purpose of this paper, we will assume that all the inner elements are classes.

Now we can speak about *software project management* from the perspective of systems modelling. Input-output mappings constitute functional requirements. The inputs are given (the customer specifies the requirements). The outputs are implied by the inputs. This implication is methodologically not trivial at all, however we may assume that the outputs are created according to some best-practices and their characteristics are thus given. Inputs and outputs are thus out of our management attention here, while the other elements are the core of software process management:

– inner elements,
– relations between inner elements,
– relations from inside to outside.

Managing the software project thus means managing those elements. It may be also comfortable to work with groups of those elements:

Definition 1. ***Project factor*** *is*

– *An inner element.*
– *A relation between inner elements,*
– *A relations from inside to outside,*
– *Any sub graph of a graph consisting of a set of nodes P and a set of edges $R_I \cup R_{IO}$.*

The set of project factors will be denoted C.

If we suppose, that the goal of the project is to achieve the relation between the inputs and outputs, then the project management means fine-tuning the project factors. This happens based on the inputs and represents an *adaptation* of the project system according to inputs in such a way, that the project system achieves its goal in an optimal way, i.e. with minimum resources (time and costs).

Generally, there are two possibilities how to detect an adaptation need:

1. Adaptation *ex post*, that is based on past. The adaptation driver is discrepancy between outputs and inputs. This type of adaptation is used in the Adaptive Software Development (ASD) methodology [3].
2. Adaptation *ex ante*, that is considering the future. This type of adaptation is performed based on prediction about the needs of structure changes. This is the type of adaptation that is relevant to the RFST.

3 The Method

3.1 Definitions

The above formal model as a general software project model may be used for further research by introducing appropriate relations and factors in the system model. The method for supporting the RFST is based on introducing the relation representing demands on adaptation for the project system:

Definition 2. ***Demand*** *$dem(a, s)$ of the input a for the project factor s, where $a \in I$ and $s \in C$ is a mapping $I \times C$ onto an ordinal scale $\langle 0, 10\rangle$. 0 means, that the input a does not require the adaptation of the factor s. The higher the value, the higher the adaptation needs.*

As for the relations between the inner elements, substitutability has been chosen to be included in the method. Demands for adaptation of one factor may be mitigated by substitution of another factor. An example may be providing training to team members instead of hiring a new needed expert role (or vice versa). Substitutability is defined as:

Definition 3. ***Substitution*** *of project factors s_1 and s_2 $sub(s_1, s_2)$, where $s_1, s_2 \in C$, is a mapping $C \times C$ onto an ordinal scale $\langle 0, 10\rangle$. The substitution represents a possibility of substituting a demand for project factor s_1 by a substitute s_2. In case where substitution is not possible, the function has value 0. The higher value, the higher substitutability. The value 10 means perfect substitutes.*

Individual demands for factor adaptation are added and we get the total demand:

Definition 4. ***Difference*** *of the factor s_j, where $j = 1, \ldots, n$ is the value of the function $dif(s_j)$, that assigns a non-negative whole number to every project factor $s_j \in C$:*

$$dif(s_j) = \sum_{i=1}^{m} dem(a_i, s_j) \tag{1}$$

where a_i is input, n is the number of project factors and m is the number of inputs.

A factor may be substituted by substitutes, which is covered by the following definition:

Definition 5. ***Total substitution of the project factor*** *s_j, where $j = 1, \ldots, m$ is the value of the function $csub(s_j)$, that assigns a non-negative whole number to every project factor $s_j \in C$:*

$$csub(s_j) = \sum_{k=1, k \neq j}^{n} sub(s_j, s_k), \tag{2}$$

where n is the number of project factors.

The resulting demands for factor adaptation may be thus mitigated by inner substitution relations. We get the resulting difference of the factor:

Definition 6. ***Resulting difference of the project factor*** *b_j is the function*

$$vdif(b_j) = max(0, dif(b_j) - csub(b_j)) \tag{3}$$

The resulting difference represents overall clean demands for factor adaptation. Factors with highest values are the most crucial topics for the RFST, however also high differences mitigated by high substitutions will result in an action (ensuring the substitution works).

3.2 Inputs and Factors Selection

The next step in the method construction is to select appropriate inputs and project factors. The sets I and C are naturally very large. For practical applications it is necessary to specify a subset of the inputs $I_2 \subset I$ and a subset of the project factors $C_2 \subset C$. Ideal attributes of those sets should be:

- completeness,
- independence,
- minimalism.

In practice, it is very hard to achieve perfection in all those parameters and we make a balance between completeness and model comprehensibility and manageability, for the time complexity of processing all demands according to the definition is $\theta(|I_2| \times |C_2|)$.

During the method development, 32 inputs and 57 factors were included in the method. Software product quality characteristics and subcharacteristics according to ISO/IEC 9126-1 [5] and additional aspects were selected as the inputs. The factors were divided into the following categories[3]:

- human resources (the team),
- the management process,
- the artefacts,
- software and hardware tools,
- communication and collaboration.

3.3 The Evaluation

The process of evaluating the inputs for RFST using the method is as follows:

1. *Requirements gathering.* Requirements gathering by ordinary methods (interviews, questionnaires, etc.) [4].
2. *Structuring requirements.* Informal requirements are transformed to the method's inputs.

[3] The list and description of the inputs and the factors is out of scope of this paper. Please contact the author of the contribution if interested.

3. *Requirements analysis.* This step means identification and quantification of demand functions. For all the pairs of inputs and factors, we analyse whether the input implies some sort of adaptation of the project infrastructure. For factors that are not present in the project infrastructure yet, the adaptation means the adoption of this factor. The demand value then represents the complexity of the factor implementation.
4. *Difference function evaluation* according to the Equation 1.
5. *Substitution functions and total substitutions evaluation.* For the factors with high differences, the high adaptation demand may be mitigated by identifying some substitution relations. Substitutions for each factor are then summed according to the Equation 2.
6. *Resulting differences evaluation* according to the Equation 3.
7. *Results interpretation.* Non-zero resulting differences represents overall clean demands for factor adaptation and are thus a topic for the RFST. Generally, the higher the value, the higher the overall adaptation needs. Factors with highest values are the most crucial topics for the RFST, however also high differences mitigated by high substitutions need attention (ensuring the substitution takes place). For a further discussion about results interpretation see the next section.

3.4 Results Interpretation

The quantification of the demand and substitution values is performed based on expert estimations of the method's user. The correctness of the values and thus the correctness of the results depends on the ability of the user to estimate the adaptation needs and to assign ordinal numbers to them.

If the method's user sticks to the recommended ordinal scale $\langle 0, 10 \rangle$, the values of the resulting differences lie in the interval $\langle 0, 10m \rangle$, where m is the number of inputs. The upper limit represents the situation when all the inputs imply the maximum demand.

For each specific input set it is necessary to define certain results interpretation ranges having an adequate message. For example for the set of 32 inputs the resulting differences are in the range $\langle 0, 320 \rangle$ and we may decide to define the following ranges categories:

1. The range $\langle 0, 50 \rangle$ means "Neglectable adaptation need".
2. The range $\langle 51, 250 \rangle$ means "Necessary to adapt".
3. The range $\langle 251, 320 \rangle$ means "Too high adaptation needs".

When interpreting the values, it is necessary to keep in mind that the resulting difference is a scalar value, while the demands have in nature also various qualitative characteristics, as well. Those qualitative characters are not expressed in the calculation. Thus situations may occur, when two demands for adaptation may even neutralise each other. The resulting difference value represents just a sum of all the demands. Its high value represents the message "this factor needs an attention" and must be interpreted correctly according to the nature of demands that contribute to it.

4 A Practical Example

Let us illustrate the whole concept on a small example. Let us imagine that we need to develop an information system for a cattle farm. The information system should be used to administrate information about the cattle, the information about the veterinary inspections and the lactation data.

Let us suppose that we chose the quality characteristics according to ISO/IEC 9126-1 [5] as inputs I_2. The norm defines six characteristics:

- functionality,
- reliability,
- usability,
- efficiency,
- maintainability,
- portability.

As for the project factors C_2, let us suppose, we use the following factors in several categories:

- team characteristics:
  - qualification,
  - personal stability,
  - personal commitment,
- roles:
  - team leader,
  - analyst,
  - developer,
  - tester,
  - technical writer,
  - subject matter expert.
- process:
  - development process flexibility,
  - risk management,
  - quality assurance.

We gather requirements using suitable methods, structure them, map them to the inputs and evaluate the demands. For simplicity of the example, let us assume that we learnt that the information system must be highly *reliable* and this makes demands for our *team qualification*, on the *tester role* and also makes the *quality assurance* process a crucial one. The project is large and complex and the schedule is tight. It makes demands for the *team commitment.* Unfortunately, it looks like the team commitment is not high and needs increasing, fortunately, the team is at least *stable* and the personality of the *team leader* makes him a team authority. Users request the possibility of remote lactation data gathering. This will be solved by porting the solution to mobile devices. This *portability* of the solution will require more *team qualification* and will enhance demands for the *tester role* and overall *quality assurance.*

First, we fill in the demands table Table 2. Only rows and columns with at least one non-zero demand are shown.

		Inputs a			$dif(s)$
		reliability	functionality	portability	
Factors s	team qualification	6	2	5	13
	commitment	0	8	0	8
	tester role	3	0	8	11
	quality assurance	8	0	5	13

Table 2: Demands analysis $dem(a, s)$

Next, we quantify the substitution functions like:
sub(commitment, personal stability) = 2
sub(commitment, team leader role) = 3

By incorporating these substitutions we obtain a table with resulting differences (Table 3).

		Total difference $dif(s)$	**Total substitution** $csub(s)$	**Resulting difference** $vdif(s)$
Factors s	team qualification	13	0	13
	commitment	8	5	3
	tester role	11	0	11
	quality assurance	13	0	13

Table 3: Resulting differences

Now we can interpret the results. The analysis shows us, that the most crucial areas that will require adaptation is team qualification and quality assurance. There is a high demand for the tester role. Increased personal commitment will be required, but it may be partly mitigated by substitutes.

This output provides a structured input into the feasibility study elaboration. The inputs should be further analysed, the necessary actions formulated and evaluated from the three feasibility perspectives described in section 1.

5 Discussion and Conclusions

The feasibility study is a crucial input to decision about the go/not-go for a project. In the case of software projects, a quality feasibility study has a great importance, because the success rate of the software projects is far from 100% (Standish Group CHAOS reports). One of the input sources for the feasibility study are the resulting software product requirements.

The presented method provides a way how to quantify the impact of user requirements and other software project aspects on the project infrastructure. The demands for infrastructure changes resulting from the requirements represent required adaptation that should take place for the project to be successful. The presented method represents one possible approach to this issue and meets the stated criteria: to be structured, recordable and traceable. It is based on the analogy between the system modelling and a software project and formalizes the software project as a system with inputs, outputs and inner structure which is represented by project factors crucial in the project management.

The presented example should provide the reader with a feeling how the method works and how it can be used. In this simple example, the conclusions may be made just with common sense, but in real situations, typically tens of inputs and factors will be involved and the conclusions will not be that apparent.

The selection of appropriate inputs and project factors sets is an import step of the method. The balance between the accuracy and the manageable number of elements should be maintained.

The method does not automate the process, but helps to cope with the analysis in a structured, manageable way that simplifies discussion, reasoning and makes decision making process a recordable, traceable action with better reuse options. The method also inspires the user to think about possible relations between the requirements and the infrastructure, which serves as a kind of checklist for not forgetting some important issue.

The economic aspects of the analysis are not covered in the method, however they should be taken into account during the analysis: e.g. when deciding about the possible substitutions. Enhancement of the method in this area is one of the topics of the further development.

The method is now being further developed and tested in practice with the support of the IZMAN project (see Acknowledgement).

Acknowledgement

The research was supported by the Ministry of Education, Youth and Sports of the Czech Republic (Grant No. 2C06004 Information and knowledge management IZMAN).

References

1. Beck K., Andres C.: Extreme Programming Explained: Embrace Change (2nd Edition), Addison-Wesley Professional, (1981)
2. Ambler S.W.: Process Patterns: Building Large-Scale Systems Using Object Technology, Cambridge University Press, (1998)

3. Highsmith III J.A.: Adaptive Software Development: A Collaborative Approach to Managing Complex Systems, Dorset House Publishing Company, (1999)
4. Hull M.E.C., Jackson K., Dick J.: Requirements Engineering, Springer, (2004)
5. ISO/IEC IS 9126-1 Information Technology - Software product Quality Part 1: Quality model
6. Skyttner L.: General Systems Theory, World Scientific Publishing Company, (2001)
7. Standish Group International: CHAOS Report 2007, (2007)

Towards the Conceptual Normalisation

Martin Molhanec[1]

[1] Czech Technical University in Prague,
Faculty of Electrical Engineering, K13113
Technicka 2, Prague 6, Czech Republic
molhanec@fel.cvut.cz

Abstract. This article deal with some considerations of conceptual normalisation and very briefly how can be derived the relational and object normalisation from it. This contribution is only an introduction to this very interesting and serious issue, but infrequently discussed. Nor it is a comprehensive paper covering all particular problems and questions and offering precise mathematical proofs of authors theses presented herein. The idea of this contribution is to be a proper starting point for the discussion of this issue in the frame of international expert community engaging in conceptual, object and data modelling.

Keywords: data normalisation, conceptual, object and relational normal forms, conceptual modelling

1 Introduction

This article deals with an issue of conceptual normalisation. The author engaged in this issue some years ago but the results from his work have been published only at local Czech conferences and seminars [8, 9, 10, 11, 12, 13, 14 and 15] till now. So this article is first attempt to disseminate the author's opinions in wider scope of international expert community engaging in conceptual, object and data modelling. The author hopes that his contribution can be a starting point of wider discussion about these very interesting and serious problems.

2 Motivation and Problem Statement

The paradigm of relational normalisation (1st to 3rd normal form and other not such common as BNCF, etc.) is a very common practise in university courses in the field of database theory and design. No one have a doubt about usefulness and advisability of it. Regrettably, there are many opinions about nonsensicality of the normalisation principle in the field of object oriented paradigm by the reason that object oriented databases do not use the concept of primary and foreign keys. It is a false opinion according with my view.

J. Barjis, M.M. Narasipuram, G. Rabadi, J. Ralyté, and P. Plebani (Eds.):
CAiSE 2010 Workshop EOMAS'10, Hammamet, Tunisia, pp. 133-141, 2010.

It is clear that we do not need the primary and foreign keys for object oriented databases, because the normalization is particularly a process – how we can remove off redundancy from our data model and the definitions of normal forms for relation database with the aid of keys are purely the very proper formulation of this fact in the relational area. This approach is supported by many experts [1, 5, 16, 18 and 19]. But still, there exist some problems:

- There are not any definitively and generally accepted definitions of object normal forms.
- Most of the authors have a problem how to replace the concept of relational keys with another proper concept in the definition of object normalisation.
- It is not clear if there exist in the object area the same count of normal forms as well as in the relational area.

3 Approach

My approach rise from the following conjecture:

> ***Conjecture:*** *Both the relational and object models are specialisations of more common conceptual model; hence both the relational and object normalisations are specialisations of conceptual normalisation as well.*

I would like to mention herein one misguided argument with which I very frequently meet in the course of conversation about the necessity of object normalisation and differences between the object and relational approaches.

> ***False conjecture:*** *The object approach has not and do not need any primary and foreign keys, therefore it is better than the relational approach and its normalisation is thus nonsense.*

This argument along with the idea about normalisation as a something what relates with the primary and foreign keys implicates an idea, that normalisation in the area of object oriented paradigm is nonsense. However, as I mentioned before, the normalisation paradigm is not about the keys, that are in the relational area only proper devices how to define the normal forms, but the normalisation is essentially about a redundancy which rises from absence of redundancy in the real world.

In addition if we reformulate the definitions of normal forms by means of functional dependencies or by means of internal identifiers in the case of the object area we eventually get more or less the same conclusion.

In subsequent text we can show how it is possible to define the conceptual normalisation and only very briefly how to deduce the relational and object normalisation from it as well. But we have to overcome one basic problem – we have not any keys, the same way as in the object area, with the help of them we would very easily realised the definitions of normal forms in the conceptual area.

3.1 Conceptual Model

The first thing we must remember is that the conceptual model has nothing common with computer technologies, programming and databases as well. This is, as mentioned above, a model of concepts occurred in the real world around us. So the conceptual model is an outcome of the ontological observation of the real world and rises from the ontological model of the universe. Consequently I will use only an ordinary conceptual model without a need to deal with ontology too much. In this article I will use only the following concepts.

- *Object*
- *Class*
- *Inheritance*
- *Attribute*
- *Composition*
- *Relationship*

Apart from the fact that we can see herein proposed conceptual model identical with any general object model, we cannot anticipate about this conceptual model any characteristic related to particular real implementation in computer hardware or object oriented programme language.

> ***Conjecture:*** *There is not any unique internal identifier of the object in contrast to any usual object oriented systems.*

This is an elementary assertion. We cannot predicate that any natural world objects have any real internal identifiers. Any unique identification of object such as serial number is an artificial property created by human being just for intent of unique identification. But majority of natural world objects have not any such identification at all. Of course we do not take, for example, into consideration the unique genetic identification. Finally, simple to say, in the real world modelled by means of conceptual model do not exist any simple natural object identifiers.

> ***Lemma:*** *There are not any identical classes in light of its intension.*

We need a brief explanation now. The real world classes are created by different way in contrast to the object oriented programming classes. In usual object oriented programme languages we can created two classes with the same intension, i. e. they will be defined by the same set of attributes and methods and nevertheless they will be comprehend as two different classes by the computer system despite the fact that their behaviour and properties will be identical. In the real world the class as a matter of fact do not exist at all (regardless of the Plato's concept of ideas). The real world class is a concept arising from the human being capability to comprehend and perceive the fact that specific sets of objects are similar to each other. Therefore the conceptual model class is defined solely by its extension; the intension is not exactly obtainable. We can only assign different names to the class, but it is just the same class all the time. We have only different synonyms or other names for it.

3.2 The Base for Conceptual Normalisation

All subsequent considerations used by myself in this article arise from very simple and basic assumption.

Axiom: *The real world has not any redundancy.*

Let appreciate following facts. All objects in the real world exist only in one occurrence. There is only one "Martin Molhanec" (it's me), there is only one occurrence of EOMAS 2010 conference etc. In other words, in the real world are not any two identical instances of the same object at the same time and space. The informational systems, that we used, hold by means of included data a model of the real world, so it is clear that as the real world exist without any redundancy, thus the model of that world has not any redundancy at all.

Of course, it is not the case of warehouses that use redundancy for purpose of achievement certain specific features. Similarly, in database practise we break above mentioned principle to achieve an increasing throughput of database system.

3.3 Conceptual Object and its Features

Herein we introduce certain features of conceptual objects. Let commence with very natural lemma.

Definition: *Property is certain characteristic of object, which has its value.*

An example of such property can be possibly: colour or age. Generally, the property can be set as well. Actually, a relationship to other objects can be a property too. Thus, a property is certain abstraction of object characterisation, constituent of its intension and distinguishable or perceivable by a human being.

Lemma: *A set of object properties is unique in relation to him.*

Really, a car has only one colour property. Of course the colour of car body is another property than the colour of car chassis. These two colours are two different properties. A person has only one age property, denominate "age" etc. A proof can be done with respect to the fact, that property is a concept as well, thus it is abstract object and so the aforementioned lemma is true for them too. In other words, any real car has not two colours or any person has not two ages at the same time.

The following theorems proposed by us relate to object properties as well.

Theorem: *Object property is not dividable, in other words, object property is atomic. If we need to work with a part of it, the part becomes to be an object of one's own, often abstract, with its own properties.*

Theorem: *If we need to work with a group of object properties as if it was one concept in one's, the group becomes to be an object of one's own, often abstract, with its own properties.*

Theorem: *Prospective atomicity of any property depends on domain oriented point of view.*

Herein we explain above mentioned theorems with the aid of very simple example relates to the name of person. If we work, in the domain of our special-interest, always with the person name just like a single property including both first and last person name, we can comprehend this property as atomic and unique in the frame of the concept of person.

But, if we need to work, in the domain of our special-interest, with the first and last name separately, then the person name becomes an abstract concept by itself with two properties (the first name property and the last name property). On the contrary, if the concept of person has two properties, a first and last name, and we need to work with this pair of properties any time jointly, the pair becomes a single concept by itself with its own denomination, in others words, the pair is a new named concept.

Finally, we may not overlook the fact that in the conceptual world it is principally possible to think about any property always by a dual manner.

Lemma: *A property of any object can be perceived simultaneously as another object related to just aforementioned object.*

From this duality and above mentioned lemmas result finally an explication why it is possible to model the real world by many different ways and all that ways can be right.

3.4 Property, Object and Class

Herein we define a conceptual class as follows. (I strongly note that we deal with conceptual classes and not with programmer classes.)

Definition: *The conceptual class is a named set of properties belonged to each other similar object.*

We must be aware that in the real world the class is only an abstract concept denominating a set of abstract and real objects which are similar to each other. Thus the class named "car" is only a notation of the set of objects with such similar properties that we have a need to have a common notation for them. This notation is the name of the class. We notate, the conceptual class do not exist as a real world object at all.

Herein we submit another two definitions:

Definition: *An intension of the class is a set of all possible properties of this class.*

Definition: *An extension of the class is a set of all possible objects of this class.*

In contrast to the world of programmers where is possible very easily to define the intension of certain class and the extension of this class results from this definition. In the real world exist only real natural objects which constitute the extension of similar objects and the intension is constructed by us on the basis of investigation of their similarities.

3.5 Object, Property, Class and Relationship

Up to now we have not deal with relationships yet. In agreement with an ontological approach [3] we recognize in conceptual model proposed by us four different types of relationships.

- *Relationship between the class and its property.*
- *Relationship of inheritance between classes.*
- *Relationship which represents a composition of classes.*
- *General relationship between objects.*

The above mentioned relationships are generally used at different modelling techniques with exception of the first which is implicitly supposed. It is worth mentioning that although we named all above mentioned relationships with one common lexical label, namely "relationship", they are essentially different. Actually, all that relationships present unrelated ontological concepts and common denomination of them can be very misguided.

4 Results – Definitions of Conceptual Normal Forms (CNF)

Herein we present our definitions of conceptual normal forms understood by us as the rules of redundancy prohibition already introduce above and here altogether mentioned again.

0. CNF: The real world has not any redundancy.

1. CNF: A set of object properties is unique in relation to him.

2. CNF: Object property is not dividable, in other words, object property is atomic. If we need to work with a part of it, the part becomes to be an object of one's own, often abstract, with its own properties.

3. CNF: *If we need to work with a group of object properties as if it was one concept in one's, the group becomes to be an object of one's own, often abstract, with its own properties.*

The grounds for these conceptual normal forms have been introduced hereinbefore. I believe that the relational and object forms can be derived from these more common conceptual forms as can be very briefly shown in hereinafter text.

4.1 Relational Normal Forms (RNF)

The author of this article suggests that 1. RNF can be substantiating by our 0. to 2. CNF. The basis for this suggestion laid in the fact that 1. RNF deals with atomicity of data attributes, prohibition of multi-attributes and necessity of primary key existence. Evidently the issue of atomicity relates to herein proposed 2. CNF, the prohibition of multi-attributes results from the 1. CNF and issue of necessity of primary key existence relate to very principal 0. CNF.

According to author's humble opinion the 2. RNF is a specialisation of more common 3. RFN, but more detailed discussion of this issue is outside this article for now. That means that both the 2. and 3. RFN follow from the above suggested 3. CNF. The evidence can be done on the basis of consideration that transitive dependency between relational keys at the level of relation data paradigm is simply an implication of incorrect recognition of conceptual object at higher level of comprehensibility. It is worth to note that the concept of relational keys in relational database systems by self presents only the programmer implementation of the concept of functional dependency by implication of mutual relationships among conceptual objects. Thus, the incorrect recognition of object at conceptual level leads to transitive dependency between relational keys in the relational level.

4.2 Object Normal Forms (ONF)

At this time there is not any standard and commonly accepted concept of object normal forms, notwithstanding there are many scientists engaged in this issue. The brief summary of them is included in [7]. This work contains definitions of object normal forms based on an approach introduce originally by Ambler in [1] and further elaborated in aforementioned work by Merunka and Molhanec.

Despite the fact that the authors deal with the object oriented paradigm, the definitions of object normal forms are simply based on analogical relational forms, i. e., 1. ONF is based on 1. RNF and so on. The fundamental difference lays in the fact, that the definitions of all ONF are constructed without the use of the relational keys of course. Thus, the presented definition of ONF at aforementioned work is very similar to the definitions of CNF proposed by us herein and we can apply the same or very similar argumentation for their reasoning.

5 Conclusion

The author of this article suggests that relational as object normalisation arises from the same source i.e. from the conceptual normalisation. Surprisingly, this principal and serious subject matter is not widely discussed in expert community at all. In addition, the need of object normalisation that directly arises from the conceptual normalisation is often disputed.

The author believes that his contribution in this very interesting and serious subject matter can be a good starting point for the discussion of this issue in the frame of international expert community engaging in conceptual, object and data modelling.

References

1. Ambler Scott: Building Object Applications That Work, Your Step-By-Step Handbook for Developing Robust Systems Using Object Technology, Cambridge University Press/SIGS Books, 1997, ISBN 0521-64826-2
2. Edgar F. Codd, Wikipedia, online: [http://en.wikipedia.org/wiki/Edgar_F._Codd]
3. Heller, B., Herre, H. 2004. „Ontological Categories in GOL.“, Axiomathes 14(1):57-76 Kluwer Academic Publishers.
4. Christopher J. Date, Wikipedia, online: [http://en.wikipedia.org/wiki/Christopher_J._Date]
5. Khodorkovsky V. V.: On Normalization of Relations in Databases, Programming and Computer Software 28 (1), 41-52, January 2002, Nauka Interperiodica.
6. Merunka, Vojtěch, „Datové modelování“, ALFA Publishing, Praha 2006. ISBN 80-86851-54-0.
7. Molhanec, M., Merunka, V., Object normalization as the contribution to the area of formal methods of object-oriented database design, In: Advances in Computer and Information Sciences and Engineering. Berlin: Springer Science+Business Media , 2007, p. 100-104. ISBN 978-1-4020-8740-0.
8. Molhanec, M., Ontologické základy konceptuální normalizace, In: Objekty 2006. Ostrava: Technická universita Ostrava - Vysoká škola báňská, 2006, s. 81-90.
9. Molhanec, M., Konceptuální modelování, formální základy a ontologie, In: Tvorba softwaru 2006. Ostrava - Poruba: VŠB - Technická univerzita Ostrava, 2006, s. 121-127. ISBN 80-248-1082-4.
10. Molhanec, Martin. „Konceptuální modelování“, In Objekty 2005, Ostrava: VŠB, Technická Universita, 2005. ISBN 80-248-0595-2.
11. Molhanec, Martin. „Zásady konceptuálního totálně objektově orientovaného modelování“ In: Tvorba softwaru 2005. Ostrava: VŠB, 2005, s. 153-158. ISBN 80-86840-14-X.
12. Molhanec, Martin. „Několik poznámek k porozumění objektového paradigmatu“. Sborník konference Objekty 2004, Praha. ČZU PEF 2004, s. 189-197. ISBN 80-248-0672-X.
13. Molhanec, Martin. „Kritika některých výkladů objektově orientovaného paradigmatu“. Sborník konference Tvorba software 2004. Ostrava. Tanger, 2004, s. 163-172. ISBN 80-85988-96-8.
14. Molhanec, Martin. „Objektové metodologie – jejich užití a výklad“. Sborník konference Tvorba software 2003. Ostrava. Tanger, 2003, s. 111-115. ISBN 80-85988-83-6.
15. Molhanec, Martin. „UML – několik kritických poznámek“. Sborník konference Tvorba software 2002. Ostrava. Tanger, 2002, s. 150-159. ISBN 80-85988-74-7.
16. Nootenboom Henk Jan: Nuts - a online column about software design. http://www.sum-it.nl/en200239.html

17. Opdahl, A. L., Henderson-Sellers, B., Barbier, F., “Ontological analysis of whole–part relationships in OO-models”, Information and Software Technology, Volume 43, Issue 6, 1 May 2001, Pages 387-399
18. Tari Zahir, Stokes John, Spaccapietra Stefano: Object Normal Forms and Dependency Constraints for Object-Oriented Schemata, ACM Transactions on Database Systems 513-569, Vol 22 Issue 4, December 1997.
19. Yonghui Wu, Zhou Aoying: Research on Normalization Design for Complex Object Schemes, Info-Tech and Info-Net, vol 5. 101-106, Proceedings of ICII 2001, Beijing.

Real-time Web Services Orchestration and Choreography

Kawtar Benghazi, Manuel Noguera, Carlos Rodríguez-Domínguez, Ana Belén Pelegrina, and José Luis Garrido

University of Granada
Department of Languages and Computer Systems
GEDES Research Group
C/ Periodista Daniel Saucedo Aranda S/N. 18071 Granada, Spain
{benghazi,mnoguera,carlosrodriguez,abpelegrina,jgarrido}@ugr.es

Abstract. Real-Time issues are not usually considered when describing and composing web services. However, modern web services are usually involved in the software implementation of time-constrained business processes [1]. The satisfaction of time constraints is crucial in order to ensure the validity of systems where the response to a certain stimulation has to take place in a shortened period of time. Thus, the service composition problem becomes more complex, since time restrictions should be taken into account both in the choreography and orchestration processes in order to establish the temporal consistency of the web services. In this paper, we present a formal approach for real-time service orchestration and choreography. In this regard, we use UML-RT as a visual and user-friendly notation in order to model services and their interactions, Timed CSP as an underlying formal grounding to enable services verification and WS-BPEL as an execution language.

Key words: Timed-web services, web-services composition

1 Introduction

Web services are loosely-coupled modular software applications that interact with each other through web technologies. In order to enable inter and intra organizational integration, these web services should be composed to work together to carry out the integrated business process goals.

The choreography and orchestration are important emerging mechanism that deal with the problem of web services composition. Choreography is concerned with the web services interaction coordination, and orchestration is concerned with the creation of high level web services (called orchestrators) from existing ones.

Whenever web services are involved in the implementation of time-constrained business processes [1], their composition becomes very complex, since time restrictions should be taken into account both in the choreography and orchestration processes in order to establish the temporal consistency of the web services.

J. Barjis, M.M. Narasipuram, G. Rabadi, J. Ralyté, and P. Plebani (Eds.):
CAiSE 2010 Workshop EOMAS'10, Hammamet, Tunisia, pp. 142-153, 2010.

To this matter, it is important to introduce the notion of *timed web services*, which are services whose behavior and interactions with other services have to accomplish pre-established time restrictions.

Capturing time restrictions in orchestration and choreography requires the usage of languages with enough expressive power to represent complex relations about timely interaction between services. Moreover, it is very important to use formal methods with a well defined syntax and semantics in order to ensure the *compositionally* of timed web services and in order to verify the temporal properties that they should fulfill. Several authors (e.g., [2]) have advocated for the usage of process algebra for describing services and for reasoning on their properties. In this paper, we use Timed CSP [3] for the orchestration and choreography of timed web services. Timed CSP has several characteristics that make it suitable in order to describe both the behavior and the interactions of timed web services:

– It allows the description of temporal constraints.
– It has a well defined denotational and operational semantics [4]. On that basis, it can be checked whether processes satisfy properties of models by means of model checker tools (e.g., HORAE [5], FDR [6]) or *bisimulation*
– Compositionally: the denotation of a process is constructed from denotations of its parts.

Timed CSP and formal methods have the disadvantage of being tedious to use. In order to overcome this problem we combine this language with UML notations, specifically with timed sequence diagrams [7] and timed state diagrams [7]. Thus, both the choreography and the orchestration of timed web services are carried out in a visual manner and, then, are mapped to Timed CSP processes by applying a set of transformation rules.

In the field of web services, the Web Service-Business Process Execution Language (WS-BPEL) is a commonly adopted standard with a rich set of constructs to compose services. In this paper, we establish a set of mapping rules to systematically derive code from Timed CSP processes to WS-BPEL.

Several works in the literature have addressed the composition of web services [8]. In [9], sequence diagrams are used for the choreography of web services. Also, the transformation from sequence diagrams to WS-BPEL is given. In [10], UML-RT is used in order to describe timed web services. Other Works use formal methods for describing and reasoning on web services. For example, [11] [12] uses petri nets, [2] and [13] use CCS and [14] uses π-calculus. However, the time constraints and the temporal consistency between services was not addressed in any of these approaches.

This paper is organized as follows: Section 2 gives an overview of Timed CSP language. In Section 3, the timed web services and its constituents are defined. Both Section 4 and Section 5 describe proposed web services composition techniques. In Section 6, a mapping from Timed CSP to WS-BPEL is defined. Finally, brief conclusions drawn from this paper and some ongoing work is described in Section 7.

2 An Overview of Timed CSP

Timed CSP has been proposed as a specification language in order to model and to reason on concurrent, parallel systems that must fulfill explicit time constraints. Timed CSP in an extension over CSP [15] that introduces the capability of quantifying temporal aspects of sequencing and synchronization [16]. In this paper, Timed CSP terms are constructed according to the following grammar:

$$P ::= Stop \mid Skip \mid Wait\ t \mid a \rightarrow P \mid P; Q \mid \\ P \Box Q \mid \ P \overset{t}{\triangleright} Q \ \mid P \parallel Q$$

These terms has the following intuitive interpretations:

- P is a process.
- *Stop* is a deadlocked, stable process which will never engage in any external communication. It is the "broken program".
- *Skip* does nothing except to end successfully $\surd$. This process is equivalent to $\surd \rightarrow STOP$.
- $Wait\ t$ does nothing, but it is ready to end after a delay t.
- $a \rightarrow P$ is a process that is initially prepared to engage in an event a, and then it behaves as P.
- $P; Q$ corresponds to the sequential composition of P and Q. It represents a process that behaves like P until P chooses to terminates and becomes to behave like Q. The process Q is turned on at the exact moment that P terminates.
- $P \Box Q$ corresponds to a process that is willing to behave like either P or Q. The choice is made by the environment. The decision is taken on the first visible event (The process is non deterministic only if the first visible event of P and Q are equal).
- $P \overset{t}{\triangleright} Q$, is the process that initially gives P the priority of turning on. If no visible event from P has occurred in t units of time, the process behaves as Q, and P never turns on.
- $P \parallel Q$ represents a parallel composition of processes P and Q.

A variety of semantic models were defined for the Timed CSP language.

The Timed CSP *operational semantic model* is given in terms of two relations: an evolution relation, which describes the situation in which a process becomes another one by simply allowing time to pass, and a timed transition relation, which describes when a process becomes another one by performing an action at a particular time.

The *denotational model* is used in order to provide formal semantics to Timed CSP terms. It also allows to specify the required behavior of a process, that is, its desired properties.

The properties to be satisfied by a system or a process are defined in terms of timed failures (or timed traces). This definition characterizes some timed failures (timed traces) as acceptable and some others as non-acceptable. A process complies with its specification only if all its executions are acceptable, that is, none of its executions violates its specification.

Likewise, if we denote $(tr, \aleph)$ as a timed failure and $S(tr, \aleph)$ as a predicate in the timed failure, then it is said that P *satisfies* or *complies* with $S(tr, \aleph)$ if $S(tr, \aleph)$ holds for any timed failure (trace) of P.

$$P \ \mathbf{sat} \ S(tr, \aleph) \Leftrightarrow \forall (s, \aleph) \in tfailure(P) : S(tr, \aleph)$$

The specification of a process in terms of timed traces allows the description of both *safety* and *liveness* properties [17].

3 Real-Time Service Elements

Services are self-contained active entities. Real-time services are services that must behave under certain time restrictions. Thus, the operation of a real-time service can only be considered correct if it delivers a valid result in a predefined time range.

A real-time service is conceptually equal to a *capsule* in UML-RT. Thus, in this paper it is used the same visual representation of an UML-RT capsule to represent real-time services. Figure 1 shows the elements involved in a communication with services.

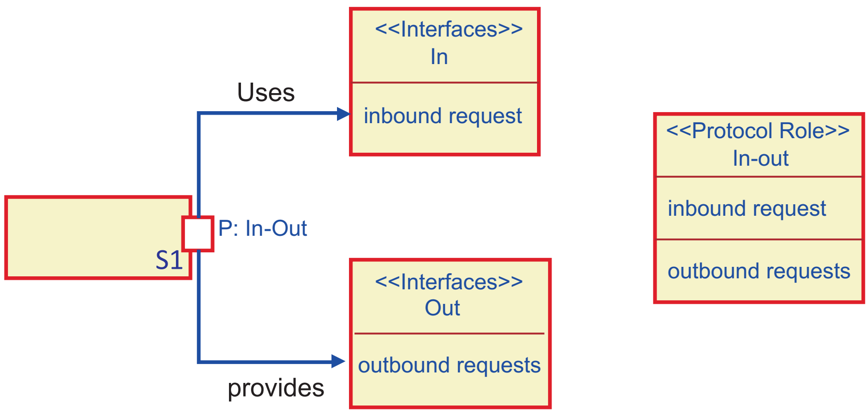


Fig. 1. Service, port, interfaces and protocol role.

A real-time service S_i communicates with other services through a set of interaction ports P_k and according to a pre-established communication protocol. It is important to mention that ports in our approach are active and therefore they have an associated behavior.

Each port can be associated to one or two interfaces. This interfaces represents access points that control the behavior of the whole service. We can distinguish between two types of interfaces:

– *Provided interfaces* that describe the set of input requests and represent the set of public operations provided by the service.
– *Required interfaces* that describe the set of output requests and represent the set of public operations required by the service.

The set of input and output requests are encapsulated in protocols and protocol roles.

Definition 1: Real-time services Real-time services are modular software applications that interact with each other through interfaces. This definition coincide of the definition of a process in Timed CSP conceptual framework. Thus, a real-time service can be represented by a process P in Timed CSP.

A real-time service has an internal and an external timed behavior:

– *Internal timed behavior.* It describes the behavior of the service that is associated with the accomplishment of its internal tasks. This behavior can be represented as a set of timed events that may be observed during the service execution. These are the timed traces $\sigma = \langle(\sigma_1, t_1), (\sigma_2, t_2) \ldots (\sigma_n, t_n)\rangle$ with $\forall i \in 1..n$, $\sigma_i \in \alpha(P)$, $\alpha(P)$ the set of all action of P and t_i the time occurrence of the event σ_i.
– *External timed behavior.* It describes the set of interactions between a service and its environment (i.e., a system composed by other services). Moreover, it provides a *black box* that only shows an abstraction over the functionality that is provided by a service, which will also hide its internal actions. Formally, the internal behavior σ_{int} of a timed-service can be deduced from its internal one by hiding all its internal actions: $\sigma_{int} = \sigma \setminus INT$, with INT a set of internal action of a service.

It is possible to distinguish two types of real-time services:

– *Primitive services* are services that are not able to be divided.
– *Composite services or orchestrators* are a set of services that are composed of existing services.

4 Real-time Service Choreography

The choreography is concerned with the interaction between services. This interaction is carried out through a set of messages exchanged between the interfaces or the ports of the services. Due to this message passing nature, undesirable situations such as deadlocking can be reached if the behavior of the services is inconsistent. In order to ensure behavior consistency between communicating services, the communications between their interfaces or ports must be coordinated.

Definition 2: Business Protocol The business protocol represents a valid communication sequence (conversation) and can be specified by a timed sequence diagram (as it is shown in figure 2) or its corresponding timed traces (by applying the rules described in [18]) and a set of temporal restrictions. Each trace has the form $\langle(e_1, t_1), (e_2, t_2) \ldots (e_n, t_n)\rangle$, where each t_i represents the instant of time when the corresponding event e_i occurs and, therefore, it establishes a partial order in the occurrence of events. The set of the temporal restrictions has the form $\{t_i\ Op\ t_j[+d]\}$, being Op a relational operator and d an optional amount of time.

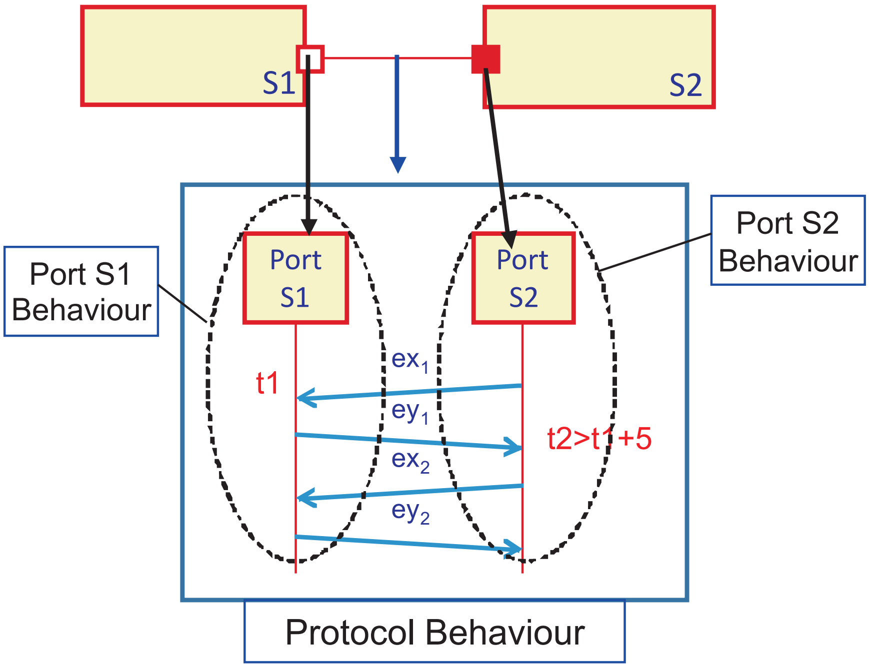


Fig. 2. Services choreography using Timed sequence diagram.

Definition 3: Deadlock free A communication between two services σ_1 and σ_2 is deadlock free if the parallel composition of the two services preserves the execution order of the messages and the temporal constraints specified in a conversation SD. Formally,

$$\sigma_1 || \sigma_2 \models_T SD$$

As an application example, let us consider a wire transfer order in an on-line banking system. Once the client has confirmed all data and has ordered to proceed through his/her web client, the transfer validation web service at the bank server generates a random key and sends it out to the client's mobile phone by invoking another web service of the client's mobile network operator. It also randomly selects one position in the matrix of secret codes that every client has and that each branch office provides to its clients. We call that value the code card matrix position (CCMP). Then, the transfer validation service requests the web client for the two codes: the one randomly generated and sent to the client's cell phone and the one from the client's code card. This process has to be completed in 15 seconds, from the time the proceed order is received until the time an SMS is delivered to the client's cell phone. Afterwards, the client has 150 seconds to provide the requested codes. Otherwise the transfer would be canceled. The interaction between the transfer validation and the web client services is described in figure 3.

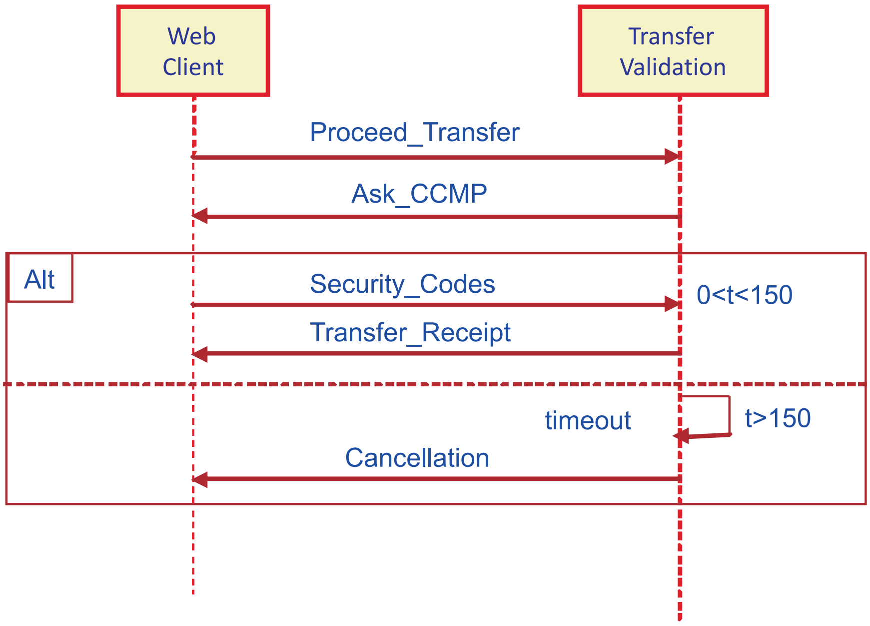


Fig. 3. Timed sequence diagram

5 Bottom-Up Service Orchestration

In our approach, the internal behavior of primitive services is modeled by transforming its timed sequence diagram into a corresponding timed state diagram, following a set of rules established in [18]. A specification of this behavior in Timed CSP syntaxis can be derived from a timed state diagram. Figure 4 shows the internal behavior of the web client and the transfer validation services that were introduced in previous section.

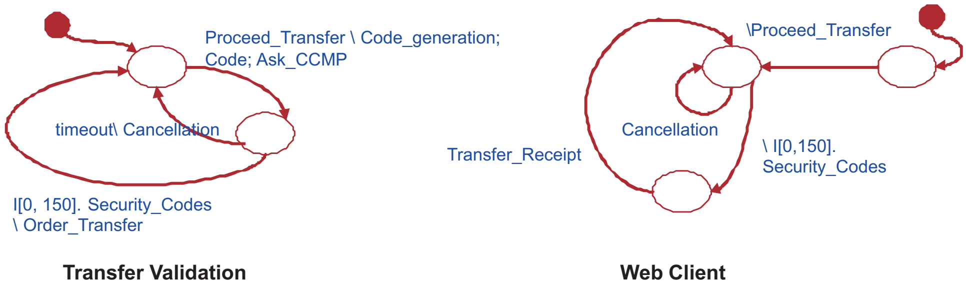


Fig. 4. Internal behavior of web client and transfer validation services

By deriving timed state diagrams in figure 4 to Timed CSP, it is obtained the formal specification of the transfer validation service (annotated as TV) and the formal specification of the client (annotated as WC) [1]:

$$\begin{aligned} TV = \quad & C1?Proceed\_Transfer \rightarrow Code\_Generation \rightarrow C2!Code \rightarrow C1!Ask\_CCMP) \\ & (C1?Security\_Codes \rightarrow (C1!Transfer\_Receipt \rightarrow SKIP)) \\ & \stackrel{150}{\triangleright} (C1!Cancellation \rightarrow SKIP) \\ WC = \quad & C1!Proceed\_Transfer \rightarrow (C1!Security\_Code \rightarrow C1?Transfer\_Receipt \rightarrow \\ & SKIP) \stackrel{150}{\triangleright} (C1?Cancellation \rightarrow SKIP) \end{aligned}$$

The specification of composite services or orchestrators is deduced from the formal specification of its components (i.e, existing services) following a bottom-up process, as it is shown in figure 5. These new services can be considered as black-boxes that interact with the environment through their ports.

Definition 4: Orchestrators The behavior of an orchestrator Σ is defined by the parallel composition of its lower level timed web services $\sigma_i, ..., \sigma_n$, by hiding all their internal actions. Formally,

$$\Sigma = \|_{i=1..n} \, \sigma_i \setminus INT$$

For example, the specification of the orchestrator that results from the composition of both services TV and WC services is like follows:

$$Orchestrator = (TV \parallel WC) \setminus \{Code\_Generation, Code\}$$

[1] In Timed CSP, the ? symbol represents "a waiting for requests" operation and the ! symbol represents a "send request" operation.

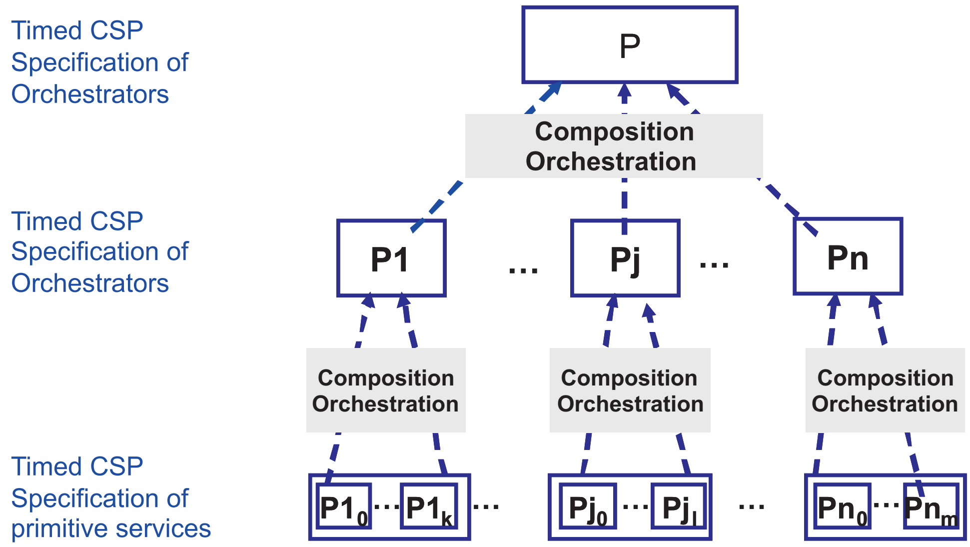


Fig. 5. Bottom-up service orchestration

6 Mapping from Timed CSP to WS-BPEL

In this section we establish the transformation rules between Timed CSP and WS-BPEL XML format. Table 1 summarizes these rules.

Events to basic activities The main elements of Timed CSP framework are events. *Events* represent atomic indivisible actions in the behavior of a process and, thus, correspond to messages in WS-BPEL. These events may be either reception or sending actions.

- The reception actions correspond to the basic activity $< receive > .. < /receive >$.
- The sending actions may correspond to $< reply > .. < /reply >$ or to $< invoke > .. < /invoke >$ basic activities, depending on whether the sending action is an output-response to a previous request or a call to another service with input parameters, respectively.

Processes to structured activities

- *SKIP* is a process that indicates a successful termination and can be mapped to the *terminate* activity.
- *STOP* is a deadlock stable process that does not engage in any external communication. This process models abnormal activity termination and, thus, it corresponds to *terminate* activity.

Timed CSP	WS-BPEL
SKIP	empty
STOP	terminate
$P_1; P_2; \ldots; P_n$	<sequence> Activity P_1 …. Activity P_n <\sequence>
$P_1 \parallel P_2 \parallel \ldots \parallel P_n$	<flow> Activity P_1 …. Activity P_n <\flow>
$P \overset{t}{\triangleright} Q$	<pick> <Onmessage…..\> <OnAlarm…\> <\pick>
Wait t;	Wait for= “0 ” until=“t”
a ? P	< sequence > < invoke, receive or reply.../ > activityP < /sequence >

Table 1. Mapping from Timed CSP to WS BPEL

– the sequence $P_1; P_2; ..P_n$ obviously, corresponds to

$$\begin{array}{c} < sequence > \\ P_1 \\ P_2 \\ .. \\ P_n \\ < /sequence > \end{array}$$

– $a \rightarrow P$ is the process that is initially prepared to engage in an event a and, then, it behaves like P. This process corresponds to:

$$\begin{array}{l} < sequence > \\ \qquad\qquad < invoke, receive \ \ or \ \ reply.../ > \\ \qquad\qquad\qquad\qquad activityP \\ < /sequence > \end{array}$$

– $P1||P2$ corresponds to a parallel execution of the structured activities $P_1, P_2.., P_n$, expressed as $< flow > P_1, P_2, P_n < /flow >$.
– Delay *wait t* corresponds to:

$$
\begin{array}{l} < wait > \\ \qquad < for > t < /for > \\ < /wait > \end{array}
$$

– A time-sensitive choice ($a \to P \overset{t}{\triangleright} Q$ is a process that initially gives the turning on priority to the process ($a \to P$). If the event a does not occur in t units of time, a timeout occurs and the process behaves as Q and P never turns itself on. This choice corresponds to a "pick" activity with an $< onMessage >$ activity and an $< onAlarm >$ one.

$$
\begin{array}{l} < pick > \\ \qquad < onMessage > .. < /onMessage > \\ \qquad\qquad ActivityP \\ \qquad < onAlarm >< for > n < /for >< /onAlarm > \\ \qquad\qquad activityQ \\ < pick > \end{array}
$$

7 Conclusions and Future Work

Despite the fact that time in which processes and services have to take place is capital in system design, there exists few proposals that address the capture of time constraints when composing services. In this paper we have presented a set of techniques for the orchestration and choreography of timed web services that combines UML-RT, Timed CSP and WS-BPEL. This approach enables the verification of some of the properties of real-time services while using a user friendly notation.

The techniques presented above will set the grounding for a model based approach intended to develop time-constrained business processes based on web services. This approach will be defined as a set of transformations between different models, as in a Model-Driven Engineering (MDE) methodology.

The intended development process can be summarized as follows:

1. **CIM level.** The business processes are designed using timed activity diagrams as presented in previous work [1].
2. **PIM level.** The specification of services and its composition is carried out by UML diagrams and its corresponding Timed CSP and WS-BPEL processes.
3. **PSM level.** An implementation of the models is specified in programming languages like RT-CORBA, Java-RT, etc..

References

1. Benghazi, K., Garrido, J.L., Noguera, M., Hurtado, M.V., Chung, L.: Extending and formalizing uml 2.0 activity diagrams for the specification of time-constrained business processes. In: RCIS. (2010)
2. Salaün, G., Bordeaux, L., Schaerf, M.: Describing and reasoning on web services using process algebra. In: ICWS '04: Proceedings of the IEEE International Conference on Web Services, Washington, DC, USA, IEEE Computer Society (2004) 43
3. Schneider, S.: Concurrent and Real–Time Systems – The CSP Approach. John Wiley & Sons, Ltd., Chichester, England (2000)
4. Schneider, S.: An operational semantics for timed csp. Inf. Comput. **116**(2) (1995) 193–213
5. Hao, P, J.S.D.J.S., Zhang, X.: Reasoning about timed csp models. In: 14th International Symposium on Formal Methods (FM'06). (2006)
6. Roscoe, A.W., Hoare, C.A.R., Bird, R.: The Theory and Practice of Concurrency. Prentice Hall PTR, Upper Saddle River, NJ, USA (1997)
7. Benghazi, K.: Medistam-RT: Metodología de Diseño y Análisis de Sistemas de Tiempo-Real. Phd Thesis, University of Granada, Spain (2009)
8. Rauf, I., Iqbal, M.Z.Z., Malik, Z.I.: Uml based modeling of web service composition - a survey. In: SERA '08: Proceedings of the 2008 Sixth International Conference on Software Engineering Research, Management and Applications, Washington, DC, USA, IEEE Computer Society (2008) 301–307
9. Bauer, B., Müller, J.P.: Mda applied: From sequence diagrams to web service choreography. In: ICWE. (2004) 132–136
10. Cambronero, M.E., Diaz, G., Pardo, J.J., Valero, V.: Using uml diagrams to model real-time web services. In: ICIW '07: Proceedings of the Second International Conference on Internet and Web Applications and Services, Washington, DC, USA, IEEE Computer Society (2007) 24
11. Yeung, W.: Csp-based verification for web service orchestration and choreography. Simulation **83**(1) (2007) 65–74
12. Martens, A.: Analyzing web service based business processes. In: FASE. (2005) 19–33
13. Cámara, J., Canal, C., Cubo, J., Vallecillo, A.: Formalizing wsbpel business processes using process algebra. Electron. Notes Theor. Comput. Sci. **154**(1) (2006) 159–173
14. Puhlmann, F.: Soundness verification of business processes specified in the pi-calculus. In: OTM Conferences (1). (2007) 6–23
15. Hoare, C.A.R., Hoare, C.A.R.: Communicating sequential processes. Communications of the ACM **21** (1985) 666–677
16. Dong, J.S., Zhang, X., Sun, J., Hao, P.: Reasoning about timed csp models (2004)
17. Dong, J.S., Hao, P., Sun, J., Zhang, X.: A Reasoning Method for Timed CSP Based on Constraint Solving. In: Formal Methods and Software Engineering. Springer Berlin / Heidelberg (November 2006) 342–359
18. Benghazi, K., Tuñón, M.I.C., Holgado, J.A., Mendoza, L.E.: Towards uml-rt behavioural consistency. In: ICEIS (3). (2007) 612–615
19. Miller, J., Mukerji, J.: MDA Guide Version 1.0.1. Technical report, OMG (2003)