Lecture Notes in Business Information Processing

This book contains the proceedings of two long-standing workshops: The 10th International Workshop on Business Process Modeling, Development and Support, BPMDS 2009, and the 14th International Conference on Exploring Modeling Methods for Systems Analysis and Design, EMMSAD 2009, held in connection with CAiSE 2009 in Amsterdam, The Netherlands, in June 2009.

The 17 papers accepted for BPMDS 2009 were carefully reviewed and selected from 32 submissions. The topics addressed by the BPMDS workshop are business and goal-related drivers; model-driven process change; technological drivers and IT services; technological drivers and process mining; and compliance and awareness.

Following an extensive review process, 16 papers out of 36 submissions were accepted for EMMSAD 2009. These papers cover the following topics: use of ontologies; UML and MDA; ORM and rule-oriented modeling; goal-oriented modeling; alignment and understandability; enterprise modeling; and patterns and anti-patterns in enterprise modeling.

LNBIP reports state-of-the-art results in areas related to business information systems and industrial application software development – timely, at a high level, and in both printed and electronic form.

The type of material published includes

- Proceedings (published in time for the respective event)
- Postproceedings (consisting of thoroughly revised and/or extended final papers)
- Other edited monographs (such as, for example, project reports or invited volumes)

In parallel to the printed book, each new volume is published electronically in LNBIP Online.

Detailed information on LNBIP can be found at
http://www.springer.com

Proposals for publication should be sent to
LNBIP Editorial, Tiergartenstr. 17, 69121 Heidelberg, Germany
E-mail: lnbip@springer.com

ISSN 1865-1348

ISBN 978-3-642-01861-9

9 783642 018619

Lecture Notes in Business Information Processing
LNBIP

Halpin et al. (Eds.)

LNBIP 29

Enterprise, Business-Process and Information Systems Modeling

BPMDS
EMMSAD

LNBIP 29

Terry Halpin John Krogstie
Selmin Nurcan Erik Proper
Rainer Schmidt Pnina Soffer
Roland Ukor (Eds.)

Enterprise, Business-Process and Information Systems Modeling

10th International Workshop, BPMDS 2009
and 14th International Conference, EMMSAD 2009
held at CAiSE 2009, Amsterdam, The Netherlands, June 2009
Proceedings

Springer

Lecture Notes
in Business Information Processing 29

Series Editors

Terry Halpin John Krogstie
Selmin Nurcan Erik Proper
Rainer Schmidt Pnina Soffer
Roland Ukor (Eds.)

Enterprise, Business-Process and Information Systems Modeling

10th International Workshop, BPMDS 2009
and 14th International Conference, EMMSAD 2009
held at CAiSE 2009
Amsterdam, The Netherlands, June 8-9, 2009
Proceedings

Springer

Volume Editors

Terry Halpin
LogicBlox, Atlanta, GA, USA
E-mail: terry.halpin@logicblox.com

John Krogstie
Norwegian University of Science and Technology, NTNU, Trondheim, Norway
E-mail: john.krogstie@idi.ntnu.no

Selmin Nurcan
University of Paris 1 Pantheon Sorbonne, Paris, France
E-mail: selmin.nurcan@univ-paris.fr

Erik Proper
Capgemini and Radboud University Nijmegen, Nijmegen, The Netherlands
E-mail: erikproper@gmail.com

Rainer Schmidt
University of Applied Sciences, Aalen, Germany
E-mail: rainer.schmidt@htw-aalen.de

Pnina Soffer
University of Haifa, Carmel Mountain, Haifa, Israel
E-mail: spnina@is.haifa.ac.il

Roland Ukor
University of Manchester, Manchester, UK
E-mail: roland.ukor@cs.man.ac.uk

Library of Congress Control Number: Applied for

ACM Computing Classification (1998): J.1, D.2, H.4, H.3.5

ISSN 1865-1348
ISBN-10 3-642-01861-0 Springer Berlin Heidelberg New York
ISBN-13 978-3-642-01861-9 Springer Berlin Heidelberg New York

springer.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12681877 06/3180 5 4 3 2 1 0

Preface

This book contain the proceedings of two long-running workshops held in connection to the CAiSE conferences relating to the areas of enterprise, business-process, and information systems modeling

- The 10^{th} International Workshop on Business Process Modeling, Development and Support (BPMDS 2009)
- The 14^{th} International Conference on Exploring Modeling Methods for Systems Analysis and Design (EMMSAD 2009)

BPMDS 2009

BPMDS 2009 was the tenth in a series of workshops that have successfully served as a forum for raising and discussing new ideas in the area of business process development and support.

The topics addressed by the BPMDS workshops are focused on IT support for business processes. This is one of the keystones of information systems theory. We strongly believe that any major conference in the area of information systems needs to address such topics independently of the current fashion. The continued interest in these topics on behalf of the IS community is reflected by the success of the last BPMDS workshops and the recent emergence of new conferences devoted to the theme.

During the previous BPMDS workshops, various issues were discussed that could be related to different but isolated phases in the life cycle of a business process. In the previous edition we arrived to a focus on the interactions between several phases of the business process life cycle.

In BPMDS 2009 the focus was on the drivers that motivate and initiate business process design and evolution. We distinguished three groups of drivers, which can exist separately or in any combination in real-life situations. These include (a) business-related drivers, where processes are changed to meet business objectives and goals, (b) technological drivers, where change is motivated or enabled by the availability, the performance or the perceived quality of IT solutions, and (c) drivers that stem from compliance requirements, facing standards and interoperability challenges.

The workshop discussions mainly dealt with the following related questions:

- What are the drivers or factors that initiate/demand change in business processes?
- How to cope with/introduce changes required by different drivers
- How to discover that it is time for a change
- How to discover that change has already happened (uncontrollable changes), and there is a need to explicitly change process definitions/operational instructions

The 17 papers accepted for BPMDS 2009 were selected from among 32 papers submitted from 14 countries (Australia, Brazil, France, Germany, Israel, Italy, Japan, Latvia, The Netherlands, South Africa, Spain, Switzerland, Tunisia, United Kingdom). They cover a wide spectrum of issues related to the drivers of business process change and how these affect the change process and are reflected in it. They are organized under the following section headings:

- Business and goal-related drivers
- Model-driven process change
- Technological drivers and IT services
- Technological drivers and process mining
- Compliance and awareness

We wish to thank all the people who submitted papers to the workshop for having shared their work with us, as well as the members of the BPMDS 2009 Program Committee and the workshop organizers of CAiSE 2009 for their help with the organization of the workshop. The conference was supported by IFIP WG 8.1

March 2009

Selmin Nurcan
Rainer Schmidt
Pnina Soffer
Roland Ukor

EMMSAD 2009

The field of information systems analysis and design includes numerous information modeling methods and notations (e.g., ER, ORM, UML, DFDs, BPMN), that are typically evolving. Even with some attempts to standardize (e.g., UML for object-oriented design), new modeling methods are constantly being introduced, many of which differ only marginally from existing approaches. These ongoing changes significantly impact the way information systems are being analyzed and designed in practice. EMMSAD focuses on exploring, evaluating, and enhancing current information modeling methods and methodologies. Although the need for such studies is well recognized, there is a paucity of such research in the literature.

The objective of EMMSAD 2009 was to provide a forum for researchers and practitioners interested in modeling methods in systems analysis and design to meet and exchange research ideas and results. It also provided the participants with an opportunity to present their research papers and experience reports and to take part in open discussions.

EMMSAD 2009 was the 14th in a very successful series of events, previously held in Heraklion, Barcelona, Pisa, Heidelberg, Stockholm, Interlaken, Toronto, Velden, Riga, Porto, Luxembourg, Trondheim, and Montpellier. This

year we had 36 papers submitted from 18 countries (Argentina, Austria, Brazil, Canada, China, France, Germany, Israel, Italy, Latvia, Luxembourg, The Netherlands, Norway, South Africa, Spain, Sweden, Switzerland, United Kingdom). After an extensive review process by a distinguished international Program Committee, with each paper receiving at least three reviews, we accepted the 16 papers that appear in these proceedings. Congratulations to the successful authors!

Apart from the contribution of the authors, the quality of EMMSAD 2009 depends in no small way on the generous contribution of time and effort by the Program Committee and the additional reviewers. Their work is greatly appreciated. We also express our sincere thanks to the CAiSE Organizing Committee, especially the CAiSE Workshop and Tutorial chairs Paul Johannesson (KTH, Stockholm, Sweden) and Eric Dubois (CRP Henri Tudor, Luxembourg).

Continuing with our very successful collaboration with IFIP WG 8.1 (http://home.dei.polimi.it/pernici/ifip81/) that started in 1997, this year's event was again a joint activity of CAiSE and WG 8.1. The European INTEROP Network of Excellence (http://www.interop-vlab.eu/) has also sponsored this workshop since 2005, as has AIS-SIGSAND (http://nfp.cba.utulsa.edu/bajaja/SIGSAND/).

For more information on EMMSAD, see our website www.emmsad.org

March 2009

John Krogstie
Erik Proper
Terry Halpin

Organization

Manfred Reichert	University of Ulm, Germany
Michael Rosemann	Queensland University of Technology, Brisbane, Australia
Rainer Schmidt	University of Applied Sciences, Aalen, Germany
Pnina Soffer	University of Haifa, Israel
Markus Strohmaier	University of Toronto, Canada
Lars Taxén	Linköping University, Sweden
Roland Ukor	University of Manchester, UK
Barbara Weber	University of Insbruk, Austria
Jelena Zdravkovic	Royal University of Technology, Stockholm, Sweden

BPMDS 2009 Additional Reviewers

Martin Henkel
Joy Garfield

EMMSAD Steering Committee

Keng Siau	University of Nebraska - Lincoln, USA
Terry Halpin	LogicBlox, USA
John Krogstie	NTNU, Norway

EMMSAD 2009 Organizing Committee

John Krogstie	NTNU, Norway
Terry Halpin	LogicBlox, USA
Erik Proper	Radboud University Nijmegen and Capgemini, The Netherlands

EMMSAD 2009 Program Committee

Wil van der Aalst	Eindhoven University of Technology, The Netherlands
Antonia Albani	Delft University of Technology, The Netherlands
Annie Becker	Florida Institute of Technology, USA
Egon Berghout	University of Groningen, The Netherlands
Giuseppe Berio	University of Turin, Italy
Nacer Boudjlida	Loria, France
Sjaak Brinkkemper	Utrecht University, The Netherlands
Andy Carver	Neumont University, USA
Olga De Troyer	Vrije Universiteit Brussel, Belgium
Mathias Ekstad	KTH, Sweden
John Erickson	University of Nebraska-Omaha, USA

Peter Fettke	Institute for Information Systems (IWi) at the DFKI, Germany
Ulrich Frank	University of Duisberg-Essen, Germany
Andrew Gemino	Simon Fraser University, Canada
Göran Goldkuhl	Linköping University, Sweden
Frank Harmsen	Capgemini and University of Maastricht, The Netherlands
Reimigijus Gustas	Karlstad University, Sweden
Wolfgang Hesse	Philipps - University Marburg, Germany
Stijn Hoppenbrouwers	Radboud University Nijmegen, The Netherlands
Jon Iden	Norges Handelshøyskole, Bergen, Norway
Paul Johannesson	Royal University of Technology, Stockholm, Sweden
Peri Loucopoulos	Loughborough University, UK
Graham McLeod	University of Cape Town, South Africa
Jan Mendling	Humboldt University of Berlin, Germany
Tony Morgan	Neumont University, USA
Michele Missikoff	LEKS, IASI, Italy
Andreas L. Opdahl	University of Bergen, Norway
Hervé Panetto	University Henri Poincaré Nancy I, France
Barbara Pernici	Politecnico di Milano, Italy
Anne Persson	University of Skövde, Sweden
Michaël Petit	University of Namur, Belgium
Jolita Ralyté	University of Geneva, Switzerland
Sudha Ram	University of Arizona, USA
Jan Recker	Queensland University of Technology, Brisbane, Australia
Colette Rolland	University of Paris 1, France
Michael Rosemann	Queensland University of Technology, Brisbane, Australia
Matti Rossi	Helsinki School of Economics, Finland
Kurt Sandkuhl	Jönköping University, Sweden
Peretz Shoval	Ben-Gurion University of the Negev, Israel
Il Yeol Song	Drexel University, USA
Janis Stirna	KTH, Sweden
Johan Versendaal	University of Utrecht, The Netherlands
Carson Woo	University of British Columbia, USA
Martin Zelm	CIMOSA, Germany
Pär Ågerfalk	Uppsala University, Sweden

EMMSAD 2009 Additional Reviewers

Namyoun Choi
Jens Gulden
Martin Henkel
Heiko Kattenstroth
Ki Jung Lee
Dominique Mery
Chun Ouyang
Pascal Ravesteyn
Ornsiri Thonggoom

Table of Contents

BPMDS 2009

Business and Goal Related Drivers

Model-Driven Process Change

Technological Drivers and IT Services

Patterns and Anti-patterns in Enterprise Modeling

Towards a BPM Success Model: An Analysis in South African Financial Services Organisations

Gavin Thompson, Lisa F. Seymour, and Brian O'Donovan

Information Systems Department, University of Cape Town, South Africa
Lisa.Seymour@uct.ac.za

Abstract. The improvement of business processes has recently emerged as one of the top business priorities for IT, and Business Process Management (BPM) is currently being seen as the best way to deliver process improvements. This research explores the enablers of BPM success, expanding on the Rosemann, de Bruin and Power theoretical BPM success model [1]. Qualitative research was conducted in four South African Financial Services Organisations with developing BPM capability. The research identified multiple success enablers categorised around Strategy, Culture, People / Resources, Governance, Methods and IT. Correlation between these factors was proposed and BPM, process and business success defined. Poor understanding of BPM within the participating organisations was found as well as insufficient supporting IT resources. It was found that the benefits of BPM investment had not yet been realised, which, increased the threat of funding being withdrawn.

Keywords: Business Process Improvement, BPM, Innovation Driver / Enabler / factors / process, IT Business Alignment / Value.

1 Introduction

For many organisations, success is based on how well they can model and optimise their processes in order to better manage the external value that the processes provide [2]. In a number of industries, organisations need to be able to create or modify business processes quickly to launch new product in a timely manner [3]. In the financial services industry, an increase in business competition and the amount of legislation being imposed by regulatory bodies has made it more difficult for companies to meet customer's service demands. This has resulted in process optimisation becoming a key strategic focus [4] and BPM (Business Process Management) being adopted.

BPM is the most recent stage in the advancement of process-oriented management theory with the overall goal of improving operational performance and increasing an organisation's agility in responding to dynamic market forces [5]. Although BPM is sometimes viewed as an IT focused extension of business process automation [6], we use the Melenovsky [7] definition of BPM as a management approach supported by technology components. By de-coupling the process from the underlying business application, BPM technology enables the business to design, deploy, change and optimise its business processes. As BPM is a fairly new discipline, there is limited research into the factors that contribute positively to BPM success. However recent

T. Halpin et al. (Eds.): BPMDS 2009 and EMMSAD 2009, LNBIP 29, pp. 1–13, 2009.

Analysis and Validation of Control-Flow Complexity Measures with BPMN Process Models

Elvira Rolón[1], Jorge Cardoso[2], Félix García[3], Francisco Ruiz[3], and Mario Piattini[3]

[1] Autonomous University of Tamaulipas – FIANS
University Center Tampico-Madero S/N, Tamaulipas, México
erolon@uat.edu.mx
[2] SAP Research, Germany and University of Coimbra, Portugal
jorge.cardoso@sap.com
[3] Alarcos Research Group, University of Castilla La Mancha
Paseo de la Universidad No. 4, 13071 Ciudad Real, Spain
{felix.garcia, francisco.ruizg, mario.piattini}@uclm.es

Abstract. Evaluating the complexity of business processes during the early stages of their development, primarily during the process modelling phase, provides organizations and stakeholders with process models which are easier to understand and easier to maintain. This presents advantages when carrying out evolution tasks in process models – key activities, given the current competitive market. In this work, we present the use and validation of the CFC metric to evaluate the complexity of business processes modelled with BPMN. The complexity of processes is evaluated from a control-flow perspective. An empirical evaluation has been carried out in order to demonstrate that the CFC metric can be useful when applied to BPMN models, providing information about their ease of maintenance.

Keywords: Business process models, BPMN, measurement, validation.

1 Introduction

Business process modeling is the first step towards the achievement of organizational goals, because its importance resides not only in the description of the process, but in that it also usually represents a preparatory phase for activities such as business process improvement, business process reengineering, technology transfer and process standardization [1].

But in all these activities the business process models are managed by different stakeholders (business process analysts, domain experts, technical analysts, software developers, among others). Therefore, one of their main purposes is support communication between stakeholders, and to fulfil this purpose business process models should be easy to understand and easy to maintain. High complexity in a process has several undesirable drawbacks: it may result in bad understandability, errors, defects, and exceptions, thus leading to the need for more time to develop, test and maintain the processes. Therefore, the first step towards reducing the complexity of processes is to first recognise its existence, and, then, measure it.

T. Halpin et al. (Eds.): BPMDS 2009 and EMMSAD 2009, LNBIP 29, pp. 58–70, 2009.

In this context, Cardoso [2] has defined process complexity as *the degree to which processes are difficult to analyze, understand or explain.* Along with this definition Control-Flow Complexity (CFC) metric for analyzing the degree of complexity of business processes has been presented. The metric is independent of the language used to model business processes. On the other hand, another stream of research [3, 4, 5] has concentrated efforts to develop a set of measures for the evaluation of models developed with BPMN (Business Process Modeling Notation) [6] which have been empirically validated. They are based on the measurement of the structural properties of process models. As a result of this empirical validation, several measures were correlated with the usability and maintainability of processes. However, we believe that since the measures proposed in both research streams are based on the analysis of the complexity of business processes models, it is important to analyze the influence of the CFC metric on the complexity of BPMN models from a control-flow perspective.

This paper therefore presents the analysis and empirical validation of the influence of the CFC metric on the usability and maintainability of BPMN process models. This is done by using the data obtained from two families of experiments which had previously been carried out to validate measures of the structural complexity of BPMN models [5].

This paper is organized as follows. Section 2 provides an overview of the related work in this area of research and Section 3 introduces the CFC metric, presenting an example of computation in a business process modelled with BPMN. Section 4 provides an overview of the two families of experiments carried out to empirically validate measures for BPMN process models. Section 5 presents the analysis of results in the validation of the CFC metric, using the data obtained from the experiments with the BPMN models. Finally, conclusions are outlined in Section 6.

2 Related Work

The complexity and other characteristics and aspects of business processes models (BPMs) such as size, density, cohesion, and coupling have been analyzed and measured by researchers who agree that, as with software processes, business processes should minimize their complexity in order to provide adequate support to the various stakeholders. The vast majority of the measures proposed for analyzing the complexity of BPMs have their origin in, or are adaptations of, measures previously defined for the evaluation of software. For instance, in [7, 8, 9, 10], this topic is analyzed and software complexity metrics (or other characteristics of software) are analyzed and compared with corresponding metrics for BPMs.

However, it is important to highlight the different perspectives from which the complexity of a business process has been evaluated. For instance, Gruhn and Laue [11] have adopted complexity measures based on cognitive weights, assuming that this is a good manner in which to measure the difficulty of understanding the BPM elements. In [12], Mendling investigates how the complexity of models influences errors observed in a wide range of existing BPMs by developing a set of metrics to measure the probability of error and testing 28 business process metrics as error predictors on a set of over 2000 process models from different samples [13, 14]. In [7], some ideas from McCabe's cyclomatic complexity are used and the CFC metric is defined, which can be used to analyze the complexity of business processes from a work-flow perspective (see Section 3).

Nonetheless, while a number of metrics have been proposed the work published about empirical validation of the measures is almost inexistent. In a recent study, the use of BPMN elements in practice and their implications were analyzed [15]. In this context, we use the CFC metric defined by Cardoso [2] to evaluate the control-flow complexity of several BPMs developed with BPMN standard notation [6]. The work presented in [2, 3] coincide in the study of the metrics defined for evaluating software processes complexity and their extension and adaptation to business processes. In addition, both share the idea that when information regarding process model complexity is obtained, the model is easier to understand and modify in order to perform maintenance tasks, and process quality improvement is more likely to occur.

3 Control-Flow Complexity Measure

An important aspect to consider in the quest to achieve an effective process management is the complexity analysis of processes. This is the aim of the CFC metric, whose definition is based on the hypothesis that the complexity of a process can be derived from its control-flow behaviour and it is affected by constructs such as splits and joins. As a result, the formula developed captures the complexity of XOR-split, OR-split and AND-split constructs as follows:

XOR-split Control-flow Complexity. Determined by the number of mental states that are introduced with this type of split. The function $CFC_{XOR\text{-}split}(a)$, where *a* is an activity, computes the control-flow complexity of the XOR-split *a*. For XOR-splits, the control-flow complexity is simply the fan-out of the split.

$$CFC_{XOR\text{-}split}(a) = \text{fan-out}(a) \quad (1)$$

OR-split Control-flow Complexity. Determined by the number of mental states that are introduced with the split. For OR-splits, the control-flow complexity is 2^(n-1), where n is the fan-out of the split.

$$CFC_{OR\text{-}split}(a) = 2^{\text{fan-out}(a)-1} \quad (2)$$

AND-split Control-flow Complexity. For an AND-split, the complexity is simply 1. The process designer needs only to consider and analyze one state that may arise from the execution of an AND-split construct, since it is assumed that all the outgoing transitions are selected and followed.

$$CFC_{AND\text{-}split}(a) = 1 \quad (3)$$

Mathematically, the Control-Flow Complexity metric is additive. This is done by simply adding the CFC of all the split constructs and is calculated as follows:

$$CFC = \sum CFC_{XOR\text{-}split}(a) + \sum CFC_{OR\text{-}split}(a) + \sum CFC_{AND\text{-}split}(a) \quad (4)$$

The greater the value of the CFC, the greater the overall structural complexity of a process will be. CFC analysis seeks to evaluate complexity without the direct execution of processes.

3.1 Example of CFC Calculation

Figure 1 shows a business process for an online ticket purchase modelled with BPMN. This process states that a customer has to choose between different outgoing paths once the process is initiated. Basically, it consists of selecting the type of tickets that is being sought on the Web, and for each option there are diverse outgoing paths. The process finishes when the purchase is carried out satisfactorily or when the customer cancels the purchase process. As example, the results of the Control-Flow Complexity calculation carried out in the process of Figure 1 are shown in Table 1.

The calculation of the overall CFC value basically consists of adding the individual CFC of each split. The value obtained gives an indication of the complexity of the ticket purchase process. With this example, it has been possible to verify that CFC metrics can be used to measure the complexity of BPMN models, thus fulfilling their objective of analyzing the control-flow complexity of business processes.

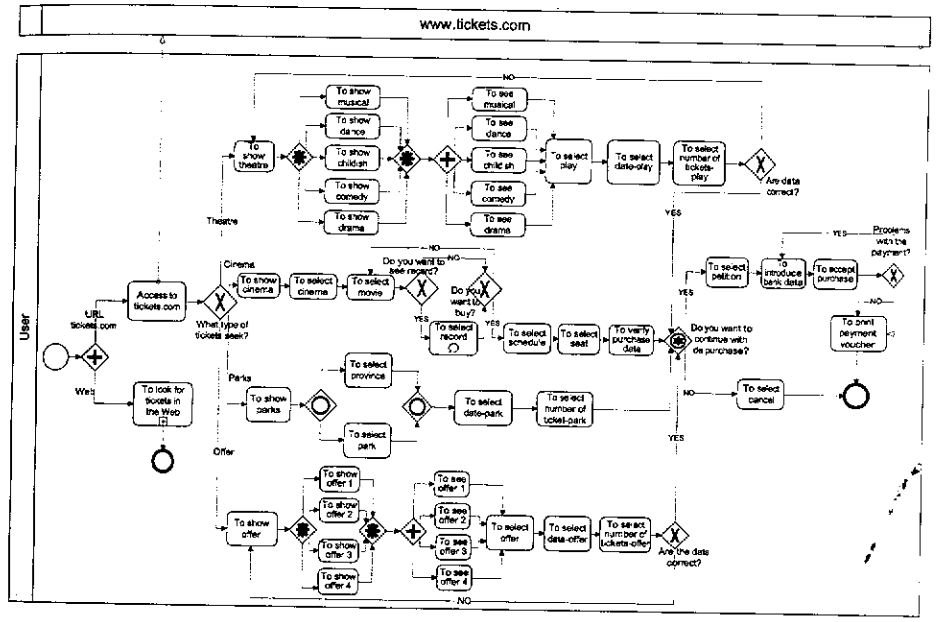

Fig. 1. Online ticket purchase process

Table 1. Values of CFC metrics for the process from Fig. 1

CFC Metric	Value	CFC Metric	Value
$CFC_{XOR\text{-}split}$(tickets type?)	4	$CFC_{XOR\text{-}split}$(shows offer type?)	4
$CFC_{XOR\text{-}split}$(to select theatre type)	5	$CFC_{XOR\text{-}split}$(offer data correct?)	2
$CFC_{XOR\text{-}split}$(theatre data correct?)	2	$CFC_{OR\text{-}split}$(to select province/park)	2^2-1
$CFC_{XOR\text{-}split}$(wants to see record?)	2	$CFC_{AND\text{-}split}$(to access the web)	1
$CFC_{XOR\text{-}split}$(wants to buy?)	2	$CFC_{AND\text{-}split}$(to select theatre type)	1
$CFC_{XOR\text{-}split}$(wants continue the purchase?)	2	$CFC_{AND\text{-}split}$(to select offer type)	1
$CFC_{XOR\text{-}split}$(payment' problems?)	2	**CFC (Online ticket purchase)**	**31**

4 Measures for BPMN Models

Our work consists of analyzing and empirically validating the CFC metric on the basis of previous work carried out to evaluate models developed with BPMN. Hence, in this section a summary of our previous works is included in order to place the results presented in this paper in context.

With the aim to evaluate the complexity of business processes by starting from the model which is a conceptual representation, we have previously defined a set of measures grouped into two categories: Base Measures and Derived Measures. Table 2 shows an example of some derived measures (the complete list of measures can be found in [3]).

Table 2. Derived measures for BPMN models

Measure	Definition	Formula
TNE	Total Number of Events of the Model	TNE = NTSE + NTIE + TNEE
TNG	Total Number of Gateways of the Model	TNG=NEDDB+NEDEB+NID+NCD+NPF
TNDO	Total Number of Data Objects	TNDO = NDOIn + NDOOut
CLA	Connectivity Level between Activities	$CLA = \frac{TNT}{NSF}$
PDOPOut	Proportion of Data Object as Outgoing Product and the total of Data Objects	$PDOPOut = \frac{NDOOut}{TNDO}$
PDOTOut	Proportion of Data Object as Outgoing Product of Activities of the Model	$PDOTOut = \frac{NDOOut}{TNT}$

The following subsections present the research context and an overview of the two families of experiments which were conducted to empirically validate the relationship between the proposed measures and the usability and maintainability of BPMN models.

4.1 Research Context

The objective of carrying out families of experiments to empirically validate the measures presented in [3] was to discover which of the measures defined could provide useful and objective information about the external quality of business process models. They focused mainly on two characteristics of the ISO 9126 external quality: usability (understandability) and maintainability (modifiability). The results obtained in the empirical validation of the first family are presented in [5].

Initially, the measures were theoretically validated according to the Briand *et al.* theoretical framework [16]. As a result, it was possible to group them in relation to the different properties of structural complexity (size, coupling and complexity) they evaluate (Fig. 2). So, the next step consisted of carrying out the empirical validation.

A set of experiments was planned and designed for the empirical validation of the measures defined. The GQM template (Goal Question Metric) [17], was used to define the research objectives as *analyse* measures of the structural complexity of BPMs *with the purpose of* evaluating them *as regards* their capability of being used as indicators of the understandability and modifiability of BPMs, *in the context of* PhD students, research assistants and others.

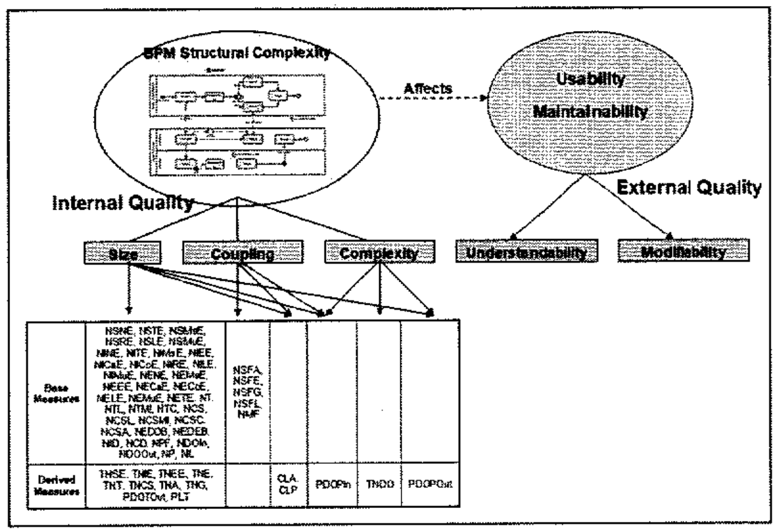

Fig. 2. Relationship between structural complexity and quality attributes

The hypothesis proposed with regard to the research objective was to ascertain whether there is a significant correlation between the measurements of structural complexity and the understandability and modifiability subcharacteristics. The independent variables were the measures defined for BPMN models and the dependent variables were those relating to the understandability and modifiability of BPMs. Later these were measured according to answer times, number of correct answers, subjective evaluation and the efficiency in the accomplishment of the tasks.

4.2 First Family

The first family of experiments was composed of five experiments. The experimental design used was the same for all five experiments. Thus, they were carried out in similar circumstances and in the same context, that is to say, by applying the same research objective, hypotheses and variables. In the experimental design a within-subjects design was carried out in which all the subjects had to do all the tests.

Material composed of ten randomly ordered BPMN models was given to each subject. These BPMN models had different structural characteristics; that is to say, different degrees of complexity; they included two questionnaires formulated for each process model. The first one was related to the understandability, and the second to the modifiability. A subjective question about the complexity of the model was also included. A more detailed description of the material can be found in [4].

The subjects (Table 3) were chosen since all of them had sufficient knowledge of modelling to carry out the experimental tasks. To leverage their knowledge about process modelling, a training lesson was carried out before the experiments run. This session consisted of an introduction to business processes and training about the BPMN standard notation.

Table 3. Groups of participants in the first family of experiments

Exp	Group	Nº Sub.	Profiles
1	UCLM (Spain)	27	PhD students, research assistants and lecturers in Computer Engineering.
2	UAT (Mexico)	31	Master's students in Information Systems.
3	University of Sannio (Italy)	37	Master's Students in: • Software Technology • Software Management and Technology • Computer Science Technology for Organizational Management and Knowledge.
4	HGCR (Spain)	6	Health professionals.
5	UCLM (Spain)	8	PhD students

4.3 Second Family

The second family of experiments included the development of five experiments. In the experiments of the second family, understandability and modifiability aspects were also analyzed, the difference being that in this instance separate experiments were designed to analyze each aspect. From the five experiments included in the second family, the first three were carried out to analyze the understandability of the models, and the modifiability was evaluated in the last two experiments.

The experimental material used to analyze the understandability consisted of fifteen BPMN models with different structural characteristics and degrees of complexity. For each model, a questionnaire with three questions related to the understandability of the process model was elaborated. In order to analyze the modifiability, the experimental material consisted of twelve BPMN models and a questionnaire with two modification requirements for each model. Moreover, in all cases the subjects answered a subjective question regarding the complexity of the process model [5].

As with the first family of experiments, the participant subjects in the second family (Table 4) received a training session about BPMN.

Having described the families of experiments the following step in this paper is to present the descriptive and statistical analysis that was carried out to validate the CFC metric. This was done by taking the data obtained concerning the dependent variables to determine the feasibility of using the CFC metric to measure the structural complexity of business process models developed with BPMN.

Table 4. Groups of participants in the second family of experiments

Exp.	Group	Nº Sub.	Profiles
1 (U)	UCLM (Spain)	22	PhD students and students in Computer Engineering.
2 (U)	UCLM (Spain)	40	Students of 4th year in Computer Engineering.
3 (U)	UCLM (Spain)	9	PhD students and students in Computer Engineering.
4 (M)	University of Bari (Italy)	29	Students in Computer Engineering
5 (M)	UAT – (Mexico)	15	Master's students in Information Systems.

As both the CFC metric and the measures proposed in [3] evaluate the structural complexity of BPMs, the same experimental design, hypothesis and variables in the two families of experiments can be stated. Consequently, the data obtained in the two former empirical studies can be used to analyze whether a correlation between the CFC metric and the maintainability of the BPMN models exists. The results of the CFC validation are shown in the next section.

5 Analysis and Validation of the CFC Metric

The CFC metric, presented in section 3, has been previously validated, by analyzing its values in different process models represented with the METEOR workflow management system and with regard to the subjective evaluation of such models by process designers [18]. As a result, the authors concluded that the CFC metric is highly correlated with the complexity of processes and, therefore can be used by business process analysts and designers to analyze the complexity of processes and to develop simpler processes when possible.

In this paper our aim is to corroborate whether the CFC metric can be used to analyze the complexity of business processes developed with a standard notation such as BPMN. Our goal is also to provide some insight, based on objective data, into the metric's influence on the ease of understanding and modifying BPMN models. With this objective in mind the stated research hypotheses are:

- Null hypothesis, H_{0u}: There is no significant correlation between the CFC metric and understandability.
- Alternative hypothesis, H_{1u}: There is a significant correlation between the CFC metric and understandability.
- Null hypothesis, H_{0m}: There is no significant correlation between the CFC metric and modifiability.
- Alternative hypothesis, H_{1m}: There is a significant correlation between the CFC metric and modifiability.

5.1 Descriptive Analysis

The first step, in order to carry out the descriptive analysis, was to obtain the values of the CFC metric of the models used in all the experiments (Table 5). The values of the CFC metrics reflect the degree of complexity of control-flows between process models. For example, process models 7 and 10 of the first family have the highest values of CFC. It is therefore possible to state that they have a greater structural complexity than process model number 1. In the second family, the highest CFC values were obtained with the first five models, as these models contained more gateways.

In both families of experiments, the dependent variables were measured based on: 1) the times that the subjects needed to carry out the required tasks, 2) the percentage of correct answers, 3) the subjective evaluation with regard to the complexity of the models, and 4) the efficiency of the answers (calculated as the ratio between the number of correct answers and the time needed to respond).

Table 5. Values of the CFC metric in experimental material

Process Model	1st Family			2nd Family	
	Exp. 1, 2 and 5	Exp. 3	Exp. 4	Exp. 1, 2 and 3	Exp. 4 and 5
1	2	2	2	25	25
2	2	2	2	25	25
3	6	6	6	33	33
4	8	8	8	31	--
5	7	7	7	2	2
6	6	6	6	7	--
7	11	11	8	9	9
8	2	2	3	5	5
9	2	2	8	8	8
10	14	15	15	0	0
11				2	--
12				4	4
13				8	8
14				4	4
15				0	0

Table 6 shows a summary of the results obtained from the experiments carried out, with regard to the time (in minutes) that the subjects needed to respond to the tasks related to understandability and modifiability.

By analyzing the time taken by the subjects to carry out the required tasks, it is possible to identify the process models in which more time was needed. For instance, for the understandability tasks in the first family, the subjects took more time to analyse process models 5, 7 and 10, whilst they took more time to carry out the modifications requested with process models 3, 4 and 7. On the other hand, the time taken by the subjects in the second family of experiments to carry out the tasks relating to the model's understandability is greater for process models 1, 2, 3, 4 and 13. For the modifiability tasks, the models 1, 2 and 13 had more spend time.

The results in both families reflect, in the first instance, the relationship of the understandability times - degree of model complexity, when comparing tables 5 and 6, since process models 7 and 10 in the first family and process models 1 to 4 in the second family coincide as being those of greater complexity. The descriptive analysis relating to correct answers, subjective evaluation and efficiency was carried out in a similar manner. Once the descriptive analysis of the data had been completed, the statistical correlation analysis was carried out and it is presented in the next section.

Table 6. Values of answer times (first family)

Process Model	First Family										Process Model	Second Family				
	Understandability Times					Modifiability Times						Underst. - Times			Mod. - Times	
	Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5	Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5		Exp. 1	Exp.2	Exp. 3	Exp. 4	Exp. 5
1	121	181	230	178	132	327	323	325	316	247	1	135	137	178	308	137
2	166	159	218	134	148	401	454	450	305	581	2	137	124	137	331	124
3	185	182	228	174	189	291	384	418	348	773	3	238	245	331	253	245
4	149	175	214	164	362	306	2546	1509	420	272	4	135	137	205	~	~
5	280	248	295	337	293	375	438	384	519	407	5	52	53	63	181	53
6	279	220	270	142	205	345	409	383	196	540	6	120	122	163	~	~
7	221	230	307	145	284	416	473	419	453	405	7	102	114	142	242	114
8	211	193	225	143	218	305	392	416	284	379	8	101	96	108	180	96
9	187	240	225	101	241	392	362	343	306	527	9	92	97	159	294	97
10	238	247	277	243	187	319	454	461	319	364	10	56	53	57	171	53
											11	123	126	178	~	~
											12	94	97	122	144	97
											13	174	161	262	312	161
											14	111	112	192	184	112
											15	49	53	116	162	53

5.2 Correlation Analysis

The first step of the correlation analysis was to ascertain whether the distribution of the data was Normal. Therefore the Kolmogorov-Smirnov test was applied. Since the data distribution was not Normal, we decided to use a non-parametrical statistical test. We have used Spearman correlation coefficient with a level of significance of α = 0.05, which indicates the probability of rejecting the null hypothesis when it is certain (type I error). That is to say, a confidence level of 95% exists. The Spearman correlation coefficient was used to separately correlate each of the measures with the dependant variables as regards each of the aspects evaluated in the descriptive analysis (answer times, correct answers, subjective evaluation and efficiency). The following subsections show the results obtained in the two families of experiments.

5.2.1 Results of the First Family

Table 7 shows the results of the correlation of the CFC metrics with regard to the measures of the dependent variables. With regard to understandability, only the $CFC_{AND\text{-}split}$ metric was validated in the fourth experiment in correlation with the answer times and subjective evaluation. In this case, we can conclude that the number of AND-split construct affects the understandability of the model, which is reflected in the answer time.

With regard to modifiability, the correlation analysis shows that the $CFC_{XOR\text{-}split}$ and CFC metrics were validated in experiments 2 and 3 in relation to the answer times, subjective evaluation and efficiency. On the other hand, only the $CFC_{AND\text{-}split}$ was validated in the third experiment in relation to the subjective evaluation.

From the results of the correlations analysis obtained in the first family of experiments, we can observe that the relationship of CFC metrics to process complexity is greater with regard to the modifiability aspect in particular. Specifically, these results show that the XOR-split construct affects above all the modifiability of the model. In addition, the validation of the CFC metric (which adds all the split constructors) gives us an indication that the structural complexity of a process, from the point of view of control flows, affects modifiability.

5.2.2 Results of the Second Family

In the second family of experiments, understandability and modifiability aspects were also evaluated, but in separate experiments designed to analyse each aspect. By following the same procedure as the one carried out in the first experimental family, once we had obtained the summary of data for each of the dependent variables measures (answer times, correct answers, subjective evaluation and efficiency) we carried out the analysis of correlations.

Table 7. Correlations of the CFC metrics and understandability (first family)

Measure	Understandability		Modifiability				
	Times	Sub. Eval.	Times	Subj. Eval.		Efficiency	
	Exp-4	Exp-4	Exp-2	Exp-2	Exp-3	Exp-2	Exp-3
CFC (XOR)			X	X	X	X	X
CFC (OR)							
CFC (AND)	X	X			X		
CFC			X	X	X	X	

Table 8. Correlations of the CFC metrics and Understandability (second family)

Measure	UNDERSTANDABILITY												MODIFIABILITY							
	Times			C. Answer			Sub. Eval.			Efficiency			Times		C. Answer		Sub. Eval.		Efficiency	
	E-1	E-2	E-3	E-1	E-2	E-3	E-1	E-2	E-3	E-1	E-2	E-3	E-4	E-5	E-4	E-5	E-4	E-5	E-4	E-5
CFC (XOR)	X	X	X		X		X	X	X	X	X	X	X	X			X	X	X	
CFC (OR)	X	X	X		X		X	X	X	X	X	X	X	X			X		X	
CFC (AND)	X	X	X		X		X	X	X	X	X	X	X				X		X	
CFC	X	X	X		X		X	X	X	X	X	X	X	X			X	X	X	

Table 8 shows that the CFC metrics were, on the whole, validated in relation to the variables analyzed. With regard to the understandability the correlations with the answer time the CFC metrics were validated in all experiments. This same correlation exists with regard to the variables of subjective evaluation and efficiency. The correct answers were only validated in the second experiment.

The correlation analysis results with regard to the modifiability also indicate that all the CFC metrics are highly correlated with the modifiability of the process models. The influence of the control-flow complexity on the modifiability of the BPMN models is reflected essentially both in the answer time in the required tasks and in the subjective evaluation and efficiency in the accomplishment of the tasks.

There are significant differences between the results obtained from the correlation analysis in the experiments of the second family as compared to the first one. These differences were also observed when the validation of measures for BPMN models was carried out. One reason for this was that the experimental material used in the second family of experiments was an improved version of that used in the first one (which, according to the feedback obtained, did not have much variability in its structural complexity). The accomplishment of the second family was therefore based on two main characteristics: a) the selection of a subset of structural complexity measures which included only the most significant measures (29 from the 60 initially defined) according to empirical results and an analysis of principal components, and b) an increase in the variability of the structural complexity of the models. We can thus consider the results obtained in the second family to be more conclusive.

The results obtained indicate that XOR-split, OR-split, and AND-split constructors affect the understandability and modifiability of the model. Therefore, based on the results and as regards the hypothesis proposed, it is possible to reject the null hypotheses and to conclude that there is a significant correlation between the CFC metric and the understandability and modifiability of BPMN models.

Finally, as a result of this empirical study, we consider that the CFC metric is a suitable complement in measuring the structural complexity of business processes models with BPMN alongside the measures proposed in [3]. With the use and validation of the CFC metrics it is possible to obtain additional information with regard to the structural complexity of BPMs, in this case from a control-flow perspective. This allows designers building process models (given more than one possible and equivalent modelling alternative) to determine which of those models is more usable and maintainable.

6 Conclusions

In this work we have presented the evaluation and empirical validation of the CFC metric for measuring BPMN business process complexity from the point of view of control-flows. The empirical validation relied on the results obtained from two families of experiments which included the carrying out of a total of ten experiments. Initially, these experiments were carried out with the aim of evaluating the structural complexity of BPMs, as a means to obtain useful information concerning their understandability and modifiability.

The CFC is a design-time measure. It can be used to evaluate the difficulty of producing a BPMN process design before implementation. When control-flow complexity analysis becomes part of the process development cycle, it has a considerable influence on the design phase, leading to further optimized processes. It is a well-known fact in software engineering that it is cost-effective to fix a defect earlier in the design lifecycle rather than later. To enable this to be done we introduce the first steps with which to carry out process complexity analysis.

As a result of applying the CFC metric, we were able to obtain additional information regarding the structural complexity of business processes. It was also possible to validate the CFC metric and to establish that it is highly correlated with the control-flow complexity of a business process and, therefore with its understandability and modifiability. These results, along with the results on the validation of BPMN measures previously obtained, provide valuable information when carrying out improvements or maintenance tasks in process models. A better understanding of the process facilitates its later modelling and evolution.

We believe that the evaluation and measurement of business process complexity in early phases of development (such as design and modeling phases) can help to identify problems in a process model and, therefore, assist designers to create or choose process models that are easy to understand for all stakeholders. Understandable models also facilitate maintenance tasks, thus reducing implicit costs. Models that are easy to understand and maintain can provide support to development tasks, such as process reengineering, the redesign of business processes on a large-scale and refactoring.

Acknowledgments. This work has been partially financed by the INGENIO Project (Junta de Comunidades de Castilla-La Mancha, Consejería de Educación y Ciencia, PAC 08-0154-9262) and ESFINGE Project (Ministerio de Educación y Ciencia, Dirección General de Investigación/Fondos Europeos de Desarrollo Regional (FEDER), TIN2006-15175-C05-05.

References

1. Succi, G., Predonzani, P., Vernazza, T.: Business Process Modeling with Objects, Costs and Human Resources. In: Bustard, D., Kawalek, P., Norris, M. (eds.) Systems Modeling for Business Process Improvement, pp. 47–60. Artech House (2000)
2. Cardoso, J.: How to Measure the Control-flow Complexity of Web Processes and Workflows. In: WfMC (ed.) Workflow Handbook, pp. 199–212. Lighthouse Point, FL (2005)

3. Rolón, E., Ruiz, F., Garcia, F., Piattini, M.: Applying Software Metrics to evaluate Business Process Models. CLEI-Electronic Journal 9(1) (paper 5) (2006)
4. Rolón, E., Garcia, F., Ruiz, F., Piattini, M.: An Exploratory Experiment to Validate Measures for Business Process Models. In: First IEEE International Conference on Research Challenges in Information Science (RCIS 2007). IEEE, Ouarzazate (2007)
5. Rolón, E., Garcia, F., Ruiz, F., Piattini, M., et al.: Evaluation of BPMN Models Quality: a Family of Experiments. In: 3rd International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE 2008), Funchal, Madeira (2008)
6. OMG, Business Process Modeling Notation (BPMN) Specification (2006)
7. Cardoso, J., Mendling, J., Neumann, G., Reijers, H.A.: A Discourse on Complexity of Process Models. In: Eder, J., Dustdar, S. (eds.) BPM Workshops 2006. LNCS, vol. 4103, pp. 117–128. Springer, Heidelberg (2006)
8. Ghani, A.A.A., Wei, K.T., Muketha, G.M., Wen, W.P.: Complexity Metrics for Measuring the Understandability and Maintainability of Business Process Models using Goal-Question-Metric (GQM). International Journal of Computer Science and Network Security (IJCSNS) 8(5), 219–225 (2008)
9. Gruhn, V., Laue, R.: Complexity Metrics for Business Process Models. In: 9th Int. Conference on Business Information Systems (BIS 2006), Klagenfurt, Austria (2006)
10. Cardoso, J.: Control-flow Complexity Measurement of Processes and Weyuker's Properties. In: 6th International Enformatika Conference. Transactions on Enformatika, Systems Sciences and Engineering, Budapest, Hungary (2005)
11. Gruhn, V., Laue, R.: Adopting the Cognitive Complexity Measure for Business Process Models. In: 5th IEEE International Conference on Cognitive Informatics (ICCI 2006), Beijing, China (2006)
12. Mendling, J., Neumann, G.: Error Metrics for Business Process Models. In: 19th International Conference on Advanced Information Systems Engineering (CAISE 2007), Trondheim, Norway (2007)
13. Mendling, J.: Detection and Prediction of Errors in EPC Business Process Models. Vienna University of Economics and Business Administration, Vienna, Austria (2007)
14. Mendling, J., Neumann, G., van der Aalst, W.M.P.: Understanding the Occurrence of Errors in Process Models based on Metrics. In: Meersman, R., Tari, Z. (eds.) OTM 2007, Part I. LNCS, vol. 4803, pp. 113–130. Springer, Heidelberg (2007)
15. Muehlen, M.z., Recker, J.: How Much Language is Enough? Theoretical and Practivcal Use of the Business Process Modeling Notation. In: Bellahsène, Z., Léonard, M. (eds.) CAiSE 2008. LNCS, vol. 5074, pp. 465–479. Springer, Heidelberg (2008)
16. Briand, L., Morasca, S., Basili, V.: Property-Based Software Engineering Measurement. IEEE Transactions on Software Engineering 22(1), 68–86 (1996)
17. Basili, V., Rombach, H.: The TAME Project: Towards Improvement-Oriented Software Environments. IEEE Transactions on Software Engineering 14(6), 728–738 (1988)
18. Cardoso, J.: Process control-flow complexity metric: An empirical validation. In: IEEE International Conference on Service Computing (SCC 2006), Chicago, USA (2006)

Vertical Alignment of Process Models – How Can We Get There?

Matthias Weidlich[1], Alistair Barros[2], Jan Mendling[3], and Mathias Weske[1]

[1] Hasso Plattner Institute, Potsdam, Germany
{matthias.weidlich, weske}@hpi.uni-potsdam.de
[2] SAP Research, CEC Brisbane, Australia
alistair.barros@sap.com
[3] Humboldt-Universität zu Berlin, Germany
jan.mendling@wiwi.hu-berlin.de

Abstract. There is a wide variety of drivers for business process modelling initiatives, reaching from business evolution and process optimisation over compliance checking and process certification to process enactment. That, in turn, results in models that differ in content due to serving different purposes. In particular, processes are modelled on different abstraction levels and assume different perspectives. Vertical alignment of process models aims at handling these deviations. While the advantages of such an alignment for inter-model analysis and change propagation are out of question, a number of challenges has still to be addressed. In this paper, we discuss three main challenges for vertical alignment in detail. Against this background, the potential application of techniques from the field of process integration is critically assessed. Based thereon, we identify specific research questions that guide the design of a framework for model alignment.

Keywords: process model alignment, business-IT gap, model consistency, model correspondences.

1 Introduction

The broad field of application of Business Process Management (BPM), from process analysis to process enactment, results in a variety of requirements for BPM methods and techniques. In particular, there is a huge difference in the appropriate level of abstraction of processes, as well as the assumed perspective. Both, abstraction level and perspective, depend on the purpose of the model and the involved stakeholders.

Evidently, real-world scenarios require multiple process models, each of them created for a specific objective. Such a model has to be *appropriate* in the sense that it incorporates a reasonable level of detail, focus on certain properties, and neglects unrelated aspects. As diverging modelling purposes cannot be organized in a strict top-down fashion, it is unrealistic that the corresponding models can always be derived through hierarchical refinement. Consequently, and most likely,

T. Halpin et al. (Eds.): BPMDS 2009 and EMMSAD 2009, LNBIP 29, pp. 71–84, 2009.