Universität Augsburg University

Value-Based and Automated Process Design

Dissertation

der Wirtschaftswissenschaftlichen Fakultät

der Universität Augsburg

zur Erlangung des Grades eines

Doktors der Wirtschaftswissenschaften

(Dr. rer. pol.)

vorgelegt

von

Manuel Bolsinger

(Master of Science with honors)

Augsburg, Juli 2014

Erstgutachter:	Prof. Dr. Hans Ulrich Buhl
Zweitgutachter:	Prof. Dr. Yarema Okhrin
Vorsitzender der mündlichen Prüfung:	Prof. Dr. Marco C. Meier
Datum der mündlichen Prüfung:	25.11.2014

Table of Contents

Please note: References are provided at the end of each section and each research paper, respectively.

Index of Research Papers

This doctoral thesis contains the following research papers:

Research Paper 1:

Bolsinger M, Bewernik MA, Buhl HU (2011) Value-Based Process Improvement. In: Tuunainen VK, Rossi M, Nandhakumar J (eds) *Proceedings of the 19th European Conference on Information Systems* (ECIS), Helsinki, Finland, Paper 21

VHB-JOURQUAL 2.1: 7.37 points, category B

Research Paper 2:

Bolsinger M (2014) Bringing Value-Based Business Process Management to the Operational Process Level. Appears in: *Information Systems and e-Business Management*

VHB-JOURQUAL 2.1: 6.52 points, category C; Impact Factor 2012: 0.605

Research Paper 3:

Bolsinger M, Elsäßer A, Helm C, Röglinger M (2014) Process Improvement through Economically Driven Routing of Instances. Resubmitted (major revision) to: *Business Process Management Journal*

VHB-JOURQUAL 2.1: 6.58 points, category C

Research Paper 4:

Heinrich B, Bolsinger M, Bewernik MA (2009) Automated Planning of Process Models: The Construction of Exclusive Choices. In: Nunamaker JF, Currie WL, Chen H, Slaughter SA (eds) *Proceedings of the 30th International Conference on Information Systems* (ICIS), Phoenix, Arizona, USA, Paper 184

VHB-JOURQUAL 2.1: 8.48 points, category A

I Introduction

A number of studies have provided evidence that different economic indicators, such as the growth rate of the global gross domestic product, increase and decrease regularly in a wavelike movement, pointing to cyclical economic upswings and downswings (Kitchin 1923; Korotayev and Tsirel 2010; Kondratieff and Stolper 1935; Tinbergen 1981). These recurring changes in the economic environment force companies – being corporative actors within the economies – to continuously question their working activities (Trkman 2010). The "art and science of overseeing how work is performed in an organization to ensure consistent outcomes and to take advantage of improvement opportunities" (Dumas et al. 2013, p. 1) is called Business Process Management (BPM). BPM is about "managing entire chains of events, activities and decisions that ultimately add value to the organization and its customers" (Dumas et al. 2013, p. 1). This narrows the focus on BPM and puts it into a prominent role for companies facing an ever-changing environment. The importance of BPM is further substantiated by the fact that companies spend considerable amounts of money on BPM (Harmon and Wolf 2014) and that Chief Information Officers position BPM among their top concerns (Luftman et al. 2013). In addition, in Germany, there is a persistent high demand for BPM consulting. In 2013, consulting regarding "organization and processes" amounted to 43.6% (>10 billion Euro) of the consulting revenue and "process optimization and performance management" was ranked second among the consulting areas in terms of predicted growth for 2014 (Bundesverband Deutscher Unternehmensberater BDU e.V. 2014). Beyond that, the volume of BPM research has also increased significantly during the last decades (Sidorova and Isik 2010).

In BPM, the objects of investigation are called processes. Processes are "chains of events, activities and decisions" that must be actively managed. Such management of a process is described by means of the so-called BPM lifecycle (Hammer 2010; van der Aalst 2013). On a high level of abstraction, the BPM lifecycle consists of three phases: (1a/b) (re)design, (2) implement/configure, and (3) run and adjust (van der Aalst 2013), as illustrated in Figure I.1.

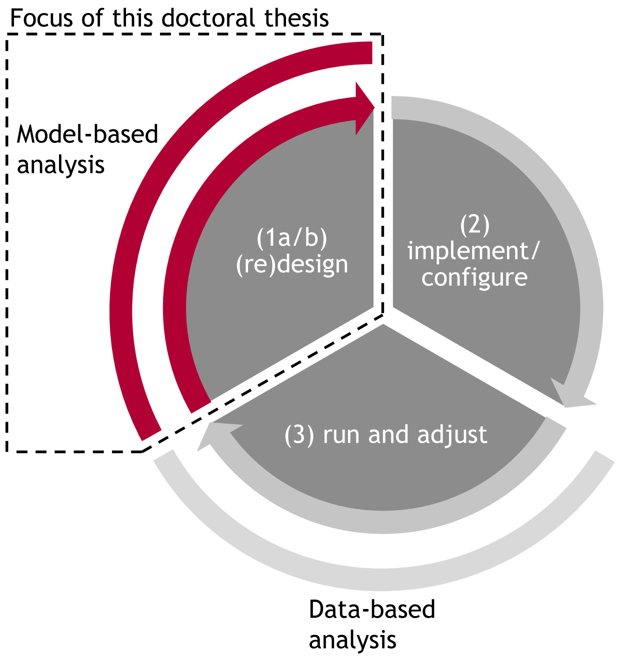


Figure I.1: BPM lifecycle on the basis of van der Aalst (2013, p. 5)

The first phase of the BPM lifecycle is twofold: (1a) design and (1b) redesign. In case a process does not exist then (1a) the process is initially developed. In this first phase, process models are commonly used to specify and visualize the process (van der Aalst 2013; Harmon and Wolf 2011). As soon as the process fits its intended purpose, (2) the process is realized in a real-world setting. Thereafter, the process is (3) executed and, if needed, modified within the boundaries set in the first phase. During the third phase, data is collected and experience is gained, which can lead to possibilities for process improvement. Such improvement can be induced by changes in the environment as for example regulatory changes, technological changes, a shift in demand, or the emergence of a competitor. This improvement can also be caused by internal reasons as for example the removal of identified bottlenecks or the need for more standardization and automation. In order to realize the possibilities for improvement, (1b) the process is redesigned followed again by the second and the third phase.

The possibility to initially design a process and the chance to redesign a process in order to improve a process shows the importance of this phase within the BPM lifecycle. Henceforth, it is referred to both (1a/b) as “design phase”. This importance is also reflected by vom Brocke et al. (2010) stating that during a process management project “the degree of influence on the business value typically aligns with the design phase of processes“ (p. 335). In addition, the improvement of processes is among the top BPM priorities for companies (Gartner 2013; Harmon and Wolf 2014; Palmer 2007; Thome et al. 2011). Within the design phase of the BPM lifecycle, there are two particular challenges that are the focus of this doctoral thesis and are detailed hereafter:

(i) Determine the business value of a process.
(ii) Create process models in an automated manner.

As for the first challenge: Designing a process should be goal-oriented (Kueng and Kawalek 1997). Besides functional goals that ensure consistent outcomes, there are in particular non-functional goals. In BPM such non-functional goals typically are: short cycle time, low execution costs, high quality, or high flexibility (González et al. 2010; Reijers and Liman Mansar 2005; Vergidis et al. 2008). However, the goal dimensions of time, cost, quality, and flexibility build the so-called “devil’s quadrangle”, which emphasizes that the improvement of one dimension may have a negative effect on another dimension (Reijers and Liman Mansar 2005). Furthermore, “the goal-oriented view of […] process engineering dictates that business goals are the driving force for structuring and evaluating […] processes” (Neiger and Churilov 2004, p. 150). Not only may these dimensions have negative effects on one another, but also these typical BPM goals per se are not the primary business goals of a company. The primary goal of a company is to sustainably increase its long-term value (Coenenberg and Salfeld 2007; Koller et al. 2010; Young and O'Byrne 2001). This overall goal is also explicitly stated, for example, in the 2013 annual reports of 19 out of 30 companies on the German stock index (DAX). Thus, all activities within a company need to contribute to this goal. This implies that all processes must be aligned to this company goal as well (Trkman 2010). Hence, when designing a process there should be a focus on the value that this process adds to a company, which reflects the business value of a process. Such process valuation is so far not performed adequately in BPM (vom Brocke et al. 2009; vom Brocke et al. 2010; vom Brocke et al. 2011), in particular with respect to the possible complex structure of processes. This can be achieved by considering the principles of value-based management during the design phase (Buhl et al. 2011), providing an overarching valuation framework that has the potential to incorporate the different dimensions typically used in BPM. Value-based management aims to increase the value of a company in an enduring manner (Ittner and Larcker 2001; Koller et al. 2010; Young and O'Byrne 2001). It has its origins in the shareholder value approach (Rappaport 1986) and was later developed further by Copeland et al. (1990) as well as Stewart and Stern (1991). In more detail, this means, when aligning a process to this primary company goal, during the design phase of a process, such aspects as cash flows, risk, and a sustainable, forward-looking increase in value need to be considered on a level of detail that accounts for the structure of a process.

As for the second challenge: Regarding a processes that potentially adds value to a company, this process adds value to a company as soon as it is running in a correct way. The shorter its design time, the more quickly value is created and the fewer errors such a process has, the more value is created. This means, the time to design a process and the freedom from errors are important in the design phase. When designing a process, usually manually created process models are used (van der Aalst 2013; Harmon and Wolf 2011). A process model represents a process graphically and specifies a process with the goal to "capture working procedures at a level of detail appropriate to fulfill its envisioned tasks" (Polyvyanyy et al. 2010, p. 150). However, the manual creation of process models can be time-consuming and error-prone (Hornung et al. 2007; vom Brocke et al. 2011), which in turn contributes to a longer design phase and to error-prone processes. These issues can partly be overcome by automating the creation of process models (Clever et al. 2013; Heinrich et al. 2012; Krause et al. 2013). Aside from systems that provide recommendations on how to proceed when modeling a process manually (Clever et al. 2013; Koschmider et al. 2011), there are approaches that target on creating whole process models automatically (Heinrich et al. 2008; Henneberger et al. 2008). Such approaches aim to reduce the time to create process models and aim to provide processes models that are free of both syntactical and semantic errors. However, the automated creation of process models for any kind of process is not yet fully possible. Notably, not all of the so-called workflow patterns (van der Aalst et al. 2003), which are reoccurring structures in process models, can be considered. Since it is only possible to create a model of any process when all workflow patterns can be considered, the possible application of such approaches is still limited.

In summary, in the design phase of the BPM lifecycle there are challenges regarding (i) the business value of a process and regarding (ii) the creation of process models in an automated manner. This doctoral thesis aims to contribute to overcome these challenges and to provide insights for research and practice. The following Section I.1 illustrates the objectives and structure of the doctoral thesis. In the subsequent Section I.2, the corresponding research papers are embedded in the research context and the fundamental research questions are highlighted.

I.1 Objectives and Structure of this Doctoral Thesis

The main objective of this doctoral thesis is to contribute to the field of BPM by focusing on a value-based and automated process design as prominent topics in research and practice. Table I.2 gives an overview of the pursued objectives and structure of the doctoral thesis.

I Introduction	
Objective I.1:	Outlining the objectives and the structure of the doctoral thesis
Objective I.2:	Embedding the included research papers into the context of the doctoral thesis and formulating the fundamental research questions
II Value-Based Process Design (Research Papers 1, 2, and 3)	
Objective II.1:	Making decisions at the process level that are in the best interest of a company as a whole considering the risk attitude of a company/person in charge
Objective II.2:	Considering the impact of a redesign on both the expected return of a company and the risk contribution
Objective II.3	Deciding between process alternatives by only having to account for the differences in the expected returns and the risk contributions of the process alternatives
Objective II.4	Providing a valuation calculus for determining the risk-adjusted expected net present value of a process
Objective II.5	Deriving concrete recommendations for process improvement that do not require extensive re-engineering projects and align with economic objectives
III Automated Process Design (Research Paper 4)	
Objective III.1:	Developing a formal language as foundation to construct the basic workflow pattern “exclusive choice” in an automated manner
Objective III.2:	Developing an algorithm to construct the basic workflow pattern “exclusive choice” in an automated manner
IV Conclusion and Outlook	
Objective IV.1:	Presenting the key findings of the doctoral thesis
Objective IV.2:	Identifying and highlighting areas for future research

Table I.2: Objectives and structure of the doctoral thesis

I.2 Research Context and Research Questions

The design phase of the BPM lifecycle needs a clear business perspective for processes to be in line with the business model of a company and, thus, to support the value creation goal of a company. Introducing the principles of value-based management into BPM has a great potential to better connect the business model of a company with its processes. Moreover, the design phase of the BPM lifecycle is typically performed in a rather manual way, although there is information technology (IT) that supports the design phase. However, this potential for IT support is not fully tapped yet. In particular for the modeling of processes, there is the potential of IT to enable a more automated process modeling. This would introduce more IT into the design phase of BPM.

In Figure I.3, the layers of the enterprise architecture show the close connection between the business model and processes as well as between processes and IT. The papers included in this doctoral thesis add, within the design phase of the BPM lifecycle, to the interface of these layers, as shown in Figure I.3.

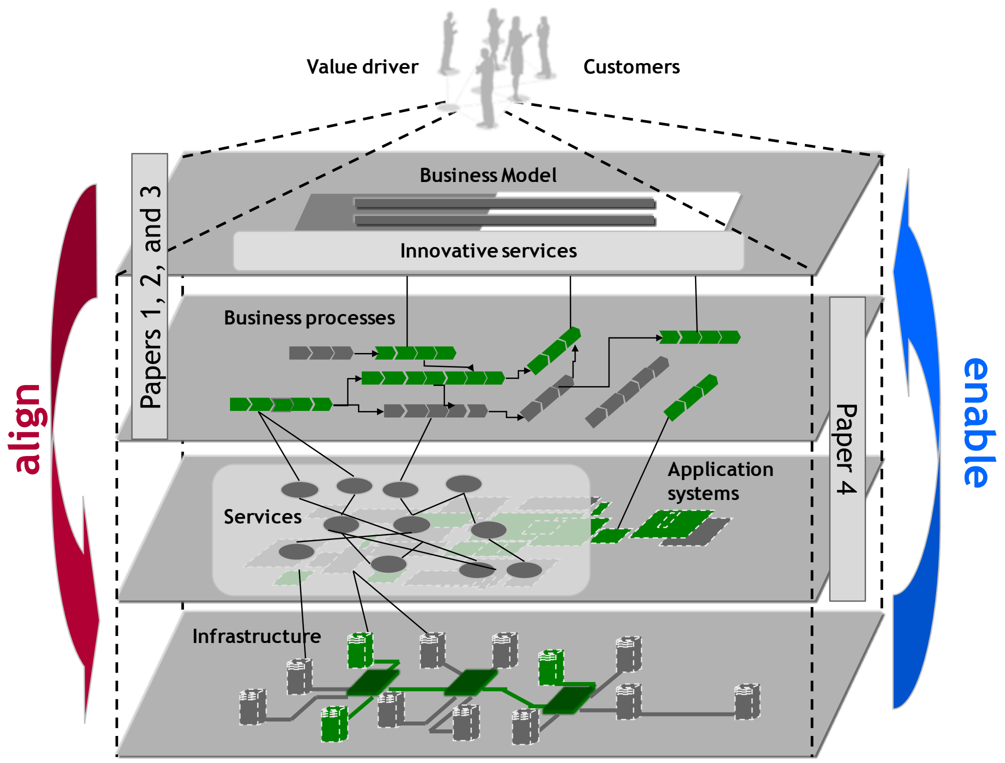


Figure I.3: Enterprise architecture on the basis of Buhl and Kaiser (2008, p. 47)

At the interface of the business model and processes, this doctoral thesis (i) extends the body of knowledge in the area of BPM by providing a valuation calculus for processes in order to design processes in line with the value creation goal of a company. At the interface of processes and IT, this doctoral thesis (ii) extends the body of knowledge regarding the automated construction of process models. This means, this doctoral thesis contains research regarding the business perspective as well as the IT perspective on processes, which are the integral views of Business and Information Systems Engineering.

In the following section, the research papers included in this doctoral thesis are embedded in the research context with respect to the above stated objectives and the respective research questions are motivated.

I.2.1 Section II: Value-Based Process Design

Research Paper 1*: "Value-Based Process Improvement"*

For companies, the increase of their value is their primary focus (Coenenberg and Salfeld 2007; Koller et al. 2010; Young and O'Byrne 2001) and the improvement of their processes is a top BPM priority (Gartner 2013; Harmon and Wolf 2014; Palmer 2007; Thome et al. 2011). Bringing together these two issues can be achieved by considering the principles of value-based management during process improvement projects (Buhl et al. 2011). In order to identify if process improvement adds value to a company, it needs to be possible to determine the amount of company value that a process adds to a company, and in particular, how much the process contributes additionally when the process is changed towards a potential process alternative. In case there is one process alternative that adds more value than any other alternative, then this process alternative should be implemented from a business perspective. This improves a process in a value-based manner, contributing to the primary goal of companies. This change in company value can only be determined if the company as a whole is considered. However, process changes are often decided considering only the single process in focus. Hence, decisions are often made due to the best knowledge of the process manager, leaving out the interest of the company/person in charge. In addition, when determining the value that is added to the company, the risk a process contributes to the overall risk of the company is omitted. All of these issues require BPM to connect the business model layer with the process layer considering value-based management. This first research paper aims at closing these gaps.

This research paper provides insights in what way a company can be regarded as a portfolio of processes and how this portfolio is related to the company value. It provides a theoretically well-founded approach to combine the overall return of these processes with the overall risk of this return. In a next step, it is shown how this value-based approach that considers a company as a whole is transferred to a single process, allowing process managers to decide in line with the value creation goal of their company. In addition, as this approach focuses on the value that a process alternative contributes additionally to a company and not on the absolute process value, it is less complex to determine the best alternative to improve a process. Doing that, the research paper addresses the following research questions, providing the foundation for a value-based process improvement and setting the stage for the following two research papers:

- How to make decisions regarding a single process that are in the best interest of a company as a whole considering the risk attitude of a company/person in charge?
- How to consider the impact of a redesign on both the expected return of a company and the risk contribution?
- Is there an efficient way to decide between process alternatives via the differences in the expected returns and the risk contributions of the process alternatives?

Research Paper 2*: "Bringing Value-Based Business Process Management to the Operational Process Level"*

In order to take on a business perspective on process design, the design phase needs to focus more on the value that a process adds to a company. For this reason, value-based BPM is discussed in Buhl et al. (2011). This is done by introducing the principles of value-based management to BPM on a rather general level. In a next step, research paper 1 focuses on value-based process improvement with a company as a whole in focus and connecting the business model layer with the process layer. Research paper 1 sets the stage for designing processes in a value-based manner that is well-founded in decision and investment theory (Buhl et al. 2011). It shows in more detail that the net present value should be used to valuate processes, which represents the value that a process adds to a company. However, when valuating a single process, a more detailed consideration of that process is required. In particular, the structure of the process needs to be considered when valuating the process (Rotaru et al. 2011; vom Brocke et al. 2010).

The second research paper focuses on the net present value of processes, and, in particular, its uncertainty. Not only does it show in detail how cash flows and multiple periods can be considered, but also how the structure of processes can be considered. Furthermore, it presents how the risk of a process can be measured via the variance of the uncertain net present value of this process. Altogether, this results in the mathematically sound determination of a risk-adjusted expected net present value of a process, for which a valuation calculus is given in this paper. It addresses the following research question:

- How to determine the risk-adjusted expected net present value of a process, in particular considering the structure of a process?

Research Paper 3*: "Process Improvement through Economically Driven Routing of Instances"*

In continuation of the previous research papers, which add to the rather conceptual body of knowledge of value-based BPM, providing insights into the valuation of processes from a general perspective, research paper 3 focuses on the use of these insights at an operational process level.

While process improvement plays an important role within companies, it mostly seems to be a black box, because it is often unclear how to improve a process. Often, processes are improved according to the feeling of a process manager, qualitative criteria, or plausibility considerations (Buhl et al. 2011; Neiger et al. 2006). Hardly any approaches provide guidance or suggest concrete ideas for improving a process (van der Aalst 2013). Existing mathematical approaches are very complex and primarily focus on very specific areas of application and, thus, are very limited regarding their scope (Vergidis et al. 2008). Furthermore, these approaches usually target on non-economic objectives, like cycle time (Buhl et al. 2011; vom Brocke et al. 2010), which is not necessarily in line with the goal of a company. Moreover, process improvement mostly results in projects that require large investments and are very risky (Devaraj and Kohli 2002). This is due to the often considerable structural process changes that go along with process improvement projects. Companies can face these problems by continuously improving their processes rather than conducting expensive re-engineering projects (Trkman 2010).

In order to address these challenges, research paper 3 aims to give clear guidance on how to improve a process from a business perspective. This is achieved by providing a concrete suggestion on how to change decision parameters within a process, which leads to a change in how a process is executed in the future, because for the same case different decisions might be taken. Such change in decision parameters aims to maximize the expected cash flow of a process, which in turn aims to increase the value of a process. This contributes to a continuous improvement of processes without engaging in large-scale process improvement projects. This is addressed in the following research question:

- How to derive concrete recommendations for process improvement that do not require extensive re-engineering projects and align with economic objectives?

I.2.2 Section III: Automated Process Design

Research Paper 4: *"Automated Planning of Process Models: The Construction of Exclusive Choices"*

In order to complement the previous three papers, which connect the business model with the processes and take on a business perspective, research paper 4 provides insights into how IT can enable a faster design phase and improve the quality of process models at the same time, taking on an IT perspective on BPM.

While improving a process in the design phase of the BPM lifecycle, process models are usually used to visualize and specify a process (van der Aalst 2013; Harmon and Wolf 2011). However, major drawbacks of process models are their time-consuming and error-prone creation (vom Brocke et al. 2011; Hornung et al. 2007). To overcome these drawbacks, approaches that aim to create whole process models in an automated manner can be used (Heinrich et al. 2008; Henneberger et al. 2008). Although these approaches have the potential to reduce the design time and to improve the quality of process models, it is not completely possible to create process models for every kind of process in an automated manner. In BPM, the so-called workflow patterns (van der Aalst et al. 2003) describe structures within process models that reoccur on a regular basis. Only if an automated approach to process modeling covers all workflow patterns, can it create process models for any process. However, the mentioned approaches do not consider all workflow patterns.

One of the most important workflow patterns, because it occurs in almost all processes and, thus, belongs to the so-called "basic workflow patterns", is the workflow pattern "exclusive choice". Due to its importance in process modeling, it is of particular interest to be able to create this workflow pattern automatically. Research paper 4 provides the formal foundation to capture the properties of an "exclusive choice" and then provides the algorithm to create an "exclusive choice". By addressing the following research questions, this research paper contributes to an IT-enabled BPM, potentially reducing the time of the design phase and increasing the quality of process models:

- What is needed as formal foundation to construct the basic workflow pattern "exclusive choice" in an automated manner?
- What is an algorithm to construct the basic workflow pattern "exclusive choice"?

I.2.3 Section IV: Summary and Future Research

After this introduction, which aims at outlining the objectives and the structure of the doctoral thesis as well as at motivating the research context and formulating the fundamental research questions, the respective research papers are presented in Sections II and III. Subsequently, Section IV presents the key findings and highlights areas for future research in the field of BPM.

I.3 References

Buhl HU, Kaiser M (2008) Herausforderungen und Gestaltungschancen aufgrund von MiFID und EU-Vermittlerrichtlinie in der Kundenberatung. Zeitschrift für Bankrecht und Bankwirtschaft 20(1):43-51

Buhl HU, Röglinger M, Stöckl S, Braunwarth K (2011) Value Orientation in Process Management - Research Gap and Contribution to Economically Well-founded Decisions in Process Management. Business & Information Systems Engineering 3(3):163-172

Bundesverband Deutscher Unternehmensberater BDU e.V. (2014) Facts & Figures zum Beratermarkt 2013/2014. http://www.mediaproverlag.de/pdf/personalintern/PERSONALintern-Links/2014_PI_16/BDU_Facts-Figures%202013.pdf. Accessed 2014-06-24

Clever N, Holler J, Shitkova M, Becker J (2013) Towards Auto-Suggested Process Modeling - Prototypical Development of an Auto-Suggest Component for Process Modeling Tools. In: Jung R, Reichert M (eds) Proceedings of the 5th International Workshop on Enterprise Modelling and Information Systems Architectures (EMISA), St. Gallen, Switzerland, pp 133-145

Coenenberg AG, Salfeld R (2007) Wertorientierte Unternehmensführung: Vom Strategieentwurf zur Implementierung. 2 ed. Schäffer-Poeschel, Stuttgart

Copeland TE, Koller T, Murrin J (1990) Valuation: measuring and managing the value of companies. Wiley, New York

Devaraj S, Kohli R (2002) The IT Payoff: Measuring the Business Value of Information Technology Investments. Prentice Hall, Upper Saddle River

Dumas M, La Rosa M, Mendling J, Reijers H (2013) Fundamentals of Business Process Management. Springer, Heidelberg

Gartner (2013) Hunting and Harvesting in a Digital World: The 2013 CIO Agenda.

González LS, Rubio FG, González FR, Velthuis MP (2010) Measurement in business processes: a systematic review. Business Process Management Journal 16(1):114-134. doi:10.1108/14637151011017976

Hammer M (2010) What is Business Process Management? In: vom Brocke J, Rosemann M (eds) Handbook on Business Process Management 1 - Introduction, Methods and Information Systems. Springer, Heidelberg, pp 3-16

Harmon P, Wolf C (2011) Business Process Modeling Survey. http://www.bptrends.com/members_surveys/deliver.cfm?report_id=1005&target=Process_Modeling_Survey-Dec_11_FINAL.pdf&return=surveys_landing.cfm. Accessed 2014-06-24

Harmon P, Wolf C (2014) The State of Business Process Management 2014. http://www.bptrends.com/bpt/wp-content/uploads/BPTrends-State-of-BPM-Survey-Report.pdf. Accessed 2014-06-24

Heinrich B, Bewernik M, Henneberger M, Krammer A, Lautenbacher F (2008) SEMPA - Ein Ansatz des Semantischen Prozessmanagements zur Planung von Prozessmodellen. Wirtschaftsinformatik 50(6):445-460

Heinrich B, Klier M, Zimmermann S (2012) Automated Planning of Process Models: Towards a Semantic-Based Approach. In: Smolnik S, Teuteberg F, Oliver T (eds) Semantic Technologies for Business and Information Systems Engineering: Concepts and Applications. IGI Global, pp 169-194

Henneberger M, Heinrich B, Bauer B, Lautenbacher F (2008) Semantic-Based Planning of Process Models. In: Bichler M, Hess T, Krcmar H, Lechner U, Matthes F, Picot A, Speitkamp B, Wolf P (eds) Tagungsband der Multikonferenz Wirtschaftsinformatik (MKWI) 2008, München, pp 1677-1689

Hornung T, Koschmider A, Oberweis A (2007) Rule-based Autocompletion of Business Process Models. In: CAiSE Forum 2007, Conference on Advanced Information Systems Engineering, Trondheim, Norway

Ittner CD, Larcker DF (2001) Assessing empirical research in managerial accounting: a value-based management perspective. Journal of Accounting and Economics 32(1-3):349-410

Kitchin J (1923) Cycles and trends in economic factors. The Review of Economic Statistics 5(1):10-16

Koller T, Goedhart M, Wessels D (2010) Valuation: Measuring and Managing the Value of Companies. 5 ed. John Wiley, New Jersey

Kondratieff ND, Stolper WF (1935) The Long Waves in Economic Life. The Review of Economics and Statistics 17(6):105-115

Korotayev AV, Tsirel SV (2010) A spectral analysis of world GDP dynamics: Kondratieff waves, Kuznets swings, Juglar and Kitchin cycles in global economic development, and the 2008–2009 economic crisis. Structure and Dynamics 4(1)

Koschmider A, Hornung T, Oberweis A (2011) Recommendation-based editor for business process modeling. Data & Knowledge Engineering 70(6):483-503. doi:10.1016/j.datak.2011.02.002

Krause F, Bewernik M, Fridgen G (2013) Valuation of manual and automated process redesign from a business perspective. Business Process Management Journal 19(1):95-110

Kueng P, Kawalek P (1997) Goal-based business process models: creation and evaluation. Business Process Management Journal 3(1):17-38

Luftman J, Zadeh HS, Derksen B, Santana M, Rigoni EH, Huang ZD (2013) Key information technology and management issues 2012–2013: an international study. Journal of Information Technology 28(4):354-366

Neiger D, Churilov L (2004) Goal-Oriented Business Process Engineering Revisited: a Unifying Perspective. In: Cordeiro J, Felipe J (eds) Proceedings of the 1st International

Workshop on Computer Supported Activity Coordination (CSAC 2004) in conjunction with the 6th International Conference on Enterprise Information Systems (ICEIS 2004), Porto, Portugal, pp 149-163

Neiger D, Churilov L, zur Muehlen M, Rosemann M (2006) Integrating Risks in Business Process Models with Value Focused Process Engineering. In: Ljungberg J, Andersson M (eds) Proceedings of the 14th European Conference on Information Systems, Gothenburg, Sweden, pp 1606-1615

Palmer N (2007) A Survey of Business Process Initiatives. http://www.bptrends.com/bpt/wp-content/surveys/FINAL%20PDF%201-23-07.pdf. Accessed 2014-07-21

Polyvyanyy A, Smirnov S, Weske M (2010) Business Process Model Abstraction. In: vom Brocke J, Rosemann M (eds) Handbook on Business Process Management 1 - Introduction, Methods and Information Systems. Springer, Heidelberg, pp 149-166

Rappaport A (1986) Creating Shareholder Value. The Free Press, New York

Reijers HA, Liman Mansar S (2005) Best practices in business process redesign: an overview and qualitative evaluation of successful redesign heuristics. Omega 33(4):283-306. doi:10.1016/j.omega.2004.04.012

Rotaru K, Wilkin C, Churilov L, Neiger D, Ceglowski A (2011) Formalizing process-based risk with Value-Focused Process Engineering. Information Systems and E-Business Management 9(4):447-474. doi:10.1007/s10257-009-0125-5

Sidorova A, Isik O (2010) Business process research: a cross-disciplinary review. Business Process Management Journal 16(4):566-597. doi:10.1108/14637151011065928

Stewart GB, Stern JM (1991) The quest for value: the EVA management guide. 27 ed. HarperBusiness, New York

Thome R, Müller T, Vogeler K (2011) Zukunftsthema Geschäftsprozessmanagement. www.pwc.de/de_de/de/prozessoptimierung/assets/pwc-gpm-studie.pdf. Accessed 2012-11/11

Tinbergen J (1981) Kondratiev cycles and so-called long waves: The early research. Futures 13(4):258-263

Trkman P (2010) The critical success factors of business process management. International Journal of Information Management 30(2):125-134. doi:10.1016/j.ijinfomgt.2009.07.003

van der Aalst WMP (2013) Business Process Management: A Comprehensive Survey. ISRN Software Eng vol. 2013(Article ID 507984). doi:10.1155/2013/507984

van der Aalst WMP, ter Hofstede AHM, Kiepuszewski B, Barros AP (2003) Workflow patterns. Distributed and Parallel Databases 14(3):5-51

Vergidis K, Tiwari A, Majeed B (2008) Business Process Analysis and Optimization: Beyond Reengineering. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 38(1):69-82. doi:10.1109/TSMCC.2007.905812

vom Brocke J, Becker J, Maria Braccini A, Butleris R, Hofreiter B, Kapočius K, De Marco M, Schmidt G, Seidel S, Simons A (2011) Current and Future Issues in BPM Research: A European Perspective from the ERCIS Meeting 2010. Communications of the Association for Information Systems 28(1):25

vom Brocke J, Recker J, Mendling J (2010) Value-oriented Process Modeling: Integrating Financial Perspectives into Business Process Re-design. Business Process Management Journal 16(2):333-356

vom Brocke J, Sonnenberg C, Simons A (2009) Wertorientiertes Prozessmanagement: State-of-the-Art und zukünftiger Forschungsbedarf. In: Hansen HR, Karagiannis D, Fill H (eds) Business Services: Konzepte, Technologien, Anwendungen. 9. Internationale Tagung Wirtschaftsinformatik, Wien, pp 253-262

Young SD, O'Byrne SF (2001) EVA and value-based management: a practical guide to implementation. McGraw-Hill, New York, NY

II Value-based Process Design

The primary focus of companies is on increasing their value. Thus, the primary focus in BPM should be on the value that a process adds to a company. This motivates the need for a value-based process design that considers the principles of value-based management during the design phase. Section II contributes to the alignment of processes to the value creation goal of a company by bringing the principles of value-based management to the operational process level.

The first research paper *"Value-Based Process Improvement"* (Section II.1) transfers the principles of value-based management at the business model layer to the operational process level in terms of improving a process. The process value is connected to the company value. In particular, the value contribution and risk contribution of a process to the overall company value and company risk are presented in detail. This sets the stage to conduct the design phase at the operational process level in line with the company goal.

The second research paper *"Bringing Value-Based Business Process Management to the Operational Process Level"* (Section II.2) details value-based process design at the operational process level by showing how the value of a process can be determined via the risk-adjusted expected net present value. This presents in detail how the structure of a process can be considered. The given valuation calculus provides the necessary theoretical foundation to use the risk-adjusted expected net present value as central figure for a value-based process design.

The third research paper *"Process Improvement through Economically Driven Routing of Instances"* (Section II.3) makes use of the insights of the previous papers to give concrete guidance on how to improve a process in a value-based manner. Furthermore, these suggestions do not cause large re-engineering projects, but are suggestions that can be implemented with little effort, following a continuous process improvement rather than a revolutionary process improvement paradigm.

II.1 Research Paper 1: "Value-based Process Improvement"

Authors:	Manuel Bolsinger[a], Marc-Andre Bewernik[b], Prof. Dr. Hans Ulrich Buhl[a] [a] FIM Research Center, Department of Information Systems Engineering & Financial Management (Prof. Dr. Hans Ulrich Buhl), University of Augsburg, Germany manuel.bolsinger@fim-rc.de hans-ulrich.buhl@fim-rc.de [b] SOFORT AG, Gauting, Germany m.bewernik@sofort.com
Published in:	Proceedings of the 19th European Conference on Information Systems (ECIS), Helsinki, Finland, Paper 21

Abstract:

For years, "improving business processes" has been and is the primary business priority of IT. In business process management (BPM), common criteria to evaluate the improvement of a process are time, costs, customer satisfaction and output quality. In contrast, the management of companies focuses on increasing the company's value, using a value-based management approach, which is hard to be linked to these criteria. A value-based process improvement can alleviate this drawback by incorporating value-based management into the area of BPM. In this paper we introduce, based on the design science paradigm, an approach that is suitable for the value-based improvement of processes. Demonstrating the feasibility and the advantage of our approach, we show its applicability within a real world scenario and evaluate it by comparing it to a competing work in the field of value-based process management.

II.1.1 Motivation, Aim, and Contributions

In May 2003, Nicholas G. Carr published his widely debated article "IT Doesn't Matter" in the Harvard Business Review. In response, Howard Smith and Peter Fingar published the book "IT Doesn't Matter—Business Processes Do", in which they critically analyze Nicolas Carr's article. They state that "*Business processes are the main intellectual property and competitive differentiator manifest in all business activity, and companies must treat them with a great degree of skill and care*." (Smith and Fingar, 2003) Thus, companies must manage their business processes in an *effective* and *efficient* manner. To do so, one particular important area of business process management (BPM) deals with the improvement of business processes. This is also validated by the recent worldwide survey "Leading in Times of Transition: The 2010 CIO Agenda" (Gartner, 2010), which interviewed nearly 1,600 CIOs. This yearly performed survey found that since 2004 "improving business processes" has been and is the primary business expectation of IT as well as the top business priority of the CIOs. However, what does "improving" refer to? Is it decreasing the costs of a process, decreasing the processing time, decreasing the risks of a process, increasing the quality of products or services that are the result of a process, all of these together or some other factor? An objective definition of "improvement" within the context of business emerges as the first step in achieving the goal of "improving business processes" from a business view.

Since the 1990s, managers have been striving to increase the value of their companies (Koller et al., 2005), using a value-based management (Coenenberg and Salfeld, 2007; Ittner and Larcker, 2001). Hence, in order to improve a process from a business view, this paper defines "improving business processes" as the change of an existing process (redesign), which increases a company's value. In order to effectively decide what change of a process will increase said value, decision makers should not only consider the resulting change of the (expected) return[1] of a process, but also the uncertainty of this return. That means, decision makers should also consider risk that is determined by the processes and influences the value of the company (*risk contribution of a process*).

Based on the design science paradigm, this paper aims to develop an effective and efficient value-based model (cp. figure 1) to support decisions on how to improve a process with the goal of increasing the value of a company (value-based process improvement). Such an approach should be crucial when it is necessary to redesign an existing process, for example due to new regulations, to decide how to effectively change the process. This is done by

[1] In this paper, *return* refers to the uncertain (stochastic) net present value (NPV) of all uncertain future cash flows.

comparing the different changes in the value of a company caused by different possible redesigns (process alternatives). The process alternative that has the greatest increase in the value of a company is the best process alternative to be used.

The following aspects help to achieve this aim and they are the key elements that this paper adds to the existing research in the area of value-based process management:

❶ *The possibility to make decisions at the process level that are in the best interest of a company as a whole with consideration of the risk attitude of a company/person in charge (decision maker):* There are different stakeholders to consider when redesigning a process, for example process analysts, organizational strategists, workflow designers and workflow managers (Lewis et al., 2007), all of which might have different objectives for a redesign. We will show how these stakeholders could select the best process alternative from a company's point of view at a process level considering the risk attitude of the decision maker.

❷ *An effective approach to consider the impact of a redesign on both the expected return of a company and the risk contribution:* A redesign of a process can have two effects on the value of a company, which can be considered to be a combined (risk-adjusted) figure of the expected return of a company and the risks that are contributed by all activities of a company to its value (*risk contribution of a company*) (Bamberg et al., 2006; Faisst and Buhl, 2005). Both the expected return and the risk contribution can change. For example a decrease in the expected return of a company as a result of a redesign may be acceptable if the risk contribution decreases even more, as this could result in an increase of the company's value. We will present how both quantities need to be considered and combined.

❸ *A model to efficiently decide between process alternatives by only having to account for the differences in the expected returns and the risk contributions of the process alternatives:* In this paper, a process alternative improves an existing process if it increases the value of a company. This already implies that it is not necessary to know the total amount of the value of the company before and after a redesign, but only the difference. This is more efficient, because it is easier, faster and cheaper to determine the difference than the total amount of the values.

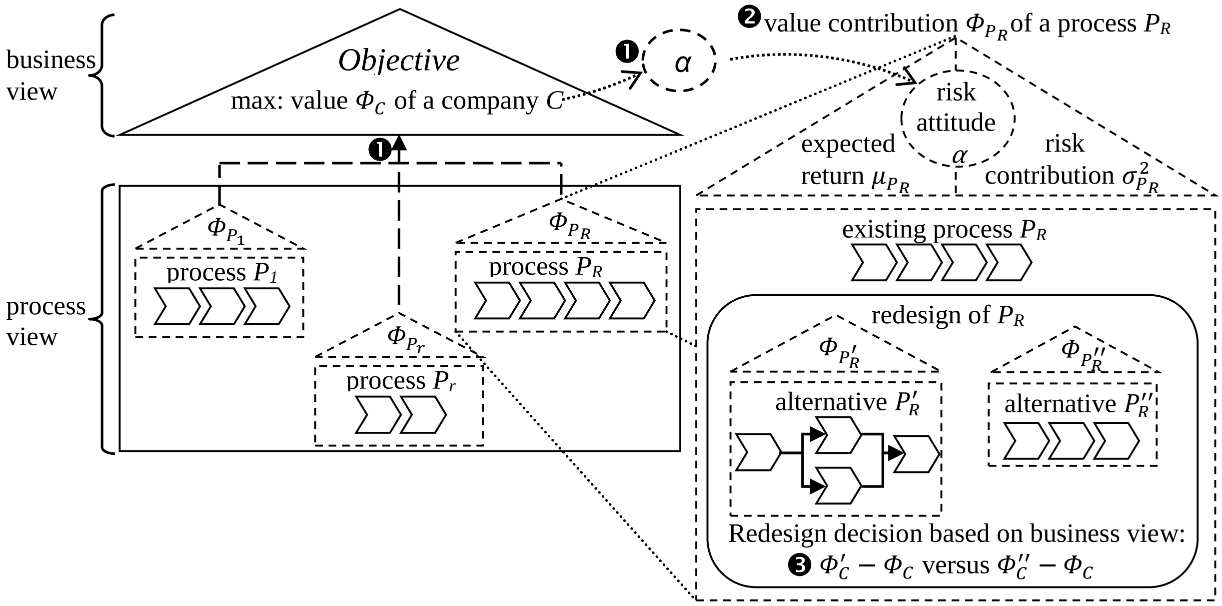


Figure 1. Value-Based Process Improvement

Considering the guidelines for conducting design science research by Hevner et al. (2004) and following the process for design science research in Peffers et al. (2008), we have organized the paper as follows: After having *identified the problem and motivated* its importance in this section, the design process continues in the next section. There, we identify the requirements for our approach which also *define its objectives*. These requirements in combination with a discussion of the related work show the research gap that our approach proposes to fill. Section three answers the key research question of how to perform a value-based process improvement. We *design* an *artifact* that can be used by a technology-oriented audience, and that should be used to *communicate* value-based process improvement to a more general managerial audience. In section four, we *demonstrate* the use of our model by illustrating its application within a real world scenario (*problem instance*). The penultimate section is dedicated to the *evaluation* of our model. Finally, the last section summarizes our considerations and provides an outlook on future steps.

II.1.2 Requirements and Related Work

We begin with the formulation of requirements, which the model to perform a value-based process improvement must meet, and that are used during the design process to guide the development of the model. At the same time, these requirements are the source for the subsequent analysis of the related work to identify the need for research. In addition, our proposed model is evaluated against those requirements, after the model has been presented and applied.

II.1.2.1 Requirements

The requirements result from the preceding remarks and stem from the area of value-based management. They are listed as follows:

(R1) *Multiple periods*: When comparing alternative processes it is not enough to consider only the current or a single period cash flow, but also future cash flows and multiple periods. A process might have a higher cash flow than another process when just looking at one period, which could lead to wrong decisions if a lower cash flow would result from a comparison over several periods.

(R2) *Objective function of a company*: A model for a value-based process improvement focuses on the increase of the value of a company. Therefore, an objective function, representing that value of a company as a combined figure of the expected return of a company and a risk contribution, is required, which takes the value that a process contributes to a company (*value contribution of a process*) into account.

(R3) *Decision at the process level in the interest of the company:* As mentioned, there are different stakeholders during a redesign, all with different attitudes towards risk. It must be possible for them to decide in the best interest of the company, considering the risk aversion of the decision maker even at the process level.

While (R1) and (R2) are obvious requirements, we take a closer look at (R3). For instance, if a manager that is risk neutral needs to redesign a process of his department, he would disregard a potential risk contribution that is caused by the redesign and focus only on the expected return. In contrast, the CEO with its averse attitude towards risk does consider a risk contribution that narrows the return. This means that also the department manager needs to know and apply the decision makers risk attitude for the redesign, in order to decide in the same way as the CEO (decision maker).

II.1.2.2 Related Work

The costs/cash outflows of a process are one of the major criteria regarding decisions in BPM. This is criticized by Kanevsky and Housel (1995). They show that it is important to consider cash inflows as well. In addition, they show how the cash inflow of a process can be allocated to its components and how the value added by the components can be expressed. However, they do not consider multiple periods (R1), nor do they consider the impacts of a redesign on the value of a company including the risk contribution, as they consider the return on investment (ROI) of the process (R2). They do not account for the risk attitude of the decision maker (R3). Still, the use of both, cash outflows and inflows, and their allocation to the process components, represents an important step towards a value-based process management. Gulledge et al. (1997) state that for “the cost evaluation of […] business processes within the value-based approach, Action-Based Costing […] can be used.“ In this related work, they show that besides costs the process revenues/cash inflows are equally important. It is noted how these cash inflows might be assigned to a process. However, the authors do not consider multiple periods (R1). There is no consideration of the impacts of a process on the value of a company (R2). It is not known if the risk attitude of the decision maker is being considered or if the decision at the process level is in the best interest of the company (R3). Another work in the area of value-based process management is Neiger et al. (2006). They base the decision of which process alternative to choose on the expected value of the cash flows in one period with no consideration of multiple periods (R1). They do not connect the value contribution of a process to the value of a company (R2). In order to consider the uncertainty of the return, they perform a sensitivity analysis, but this is not included in their utility function, as this was not the primary focus of this paper. Finally, they base their decision on the expected value, implying that the decision maker is risk neutral, and therefore achieving (R3) only to a certain extent. However, they introduce the utility of process alternatives as a basis for a decision, in the special case of a risk neutral decision maker. Besides similar works of the authors, a fourth one – integrating previous research results – is vom Brocke et al. (2010). In this paper, they consider the terminal value of the investment after multiple periods (R1) and the ROI to decide which process alternative to choose. However, they focus on one process and not on the effect of that process on the value of a company (R2). They consider the expected value of each cash flow, which is used to calculate the terminal value. Just as in the previous work, implicitly, they consider a risk neutral decision maker but do not allow for risk aversion (R3). Just as the other related works, this paper increases the general knowledge in the area of value-based process management as multiple periods are considered to compare process alternatives.

II.1.3 Model for Value-Based Process Improvement

The previous section showed that existing approaches do not completely fulfill the defined requirements, which provides a research gap that we strive to close in this section. Thus we develop a basic model to perform a value-based process improvement, which simplifies certain aspects. These simplifications allow us to present the idea and a model that can be used in practice. Thus, we strive in this approach to be practical for a managerial audience, rather than complex, in order to more easily communicate it, which is an important demand of design science (Hevner et al., 2004; Peffers et al., 2008). We start by stating necessary assumptions. Afterwards, we describe how the value of a company can be calculated, representing the objective function of a company. In a next step, we will present how the value of a company is connected with the process level. There, the best process alternative can be selected by considering differences in the expected return of a process and the variance of this return, which is the risk contribution.

II.1.3.1 Assumptions

If one process r of the R processes in a company is executed, a *process instance* PI_r is triggered. A process instance is the execution of certain activities from the beginning to the end of a process. As a result of a process instance, there is a cash flow CF_{PI_r} that is caused by different kinds of certain characteristics of this process instance (e.g. cash outflow for wages, cash outflow for materials, cash inflow for selling the product, etc.). In reality, this cash flow CF_{PI_r} is uncertain (stochastic) before the process instance is executed completely, since processes often include choices (e.g. exclusive choice), which means there are different possibilities how a process can be executed. In addition, even if there is only one possibility to execute a process, the activities are most likely not executed the same way every time (e.g. activities use different amounts of material in different process instances). Therefore, CF_{PI_r} is a random variable, and so is the cash flow $CF_{P_{r_j}}$ of a process P_r in a certain period j, with $CF_{P_{r_j}} = \sum_{i=1}^{n_{r_j}} CF_{PI_{r_i}}$, where n_{r_j} is the number of process instances of process P_r in period j. In order to include multiple periods in our model (R1), we consider the (net) present value NPV_{P_r} of process P_r. It is $NPV_{P_r} = -I_{P_r} + \sum_{j=0}^{J} \frac{CF_{P_{r_j}}}{(1+w)^j}$, where I_{P_r} is an initial investment that causes a certain cash outflow (e.g. to implement or redesign a process, to analyze a process domain, etc.), $J+1$ is the number of periods and w "is the rate of interest which properly reflects the investor's time value of money." (Hillier, 1963) This uncertain net present value of all uncertain cash flows of J future periods (fulfilling (R1)) builds the return NPV_{P_r} of a process P_r, $r=1,\dots,R$. The fact that $CF_{P_{r_j}}$ is a random variable makes the return NPV_{P_r} a random variable as well. First, we will make assumptions regarding the properties of this random variable. Before we assume that it follows a normal distribution, we describe briefly why it is plausible to make this assumption.

Normally, processes in a company are executed several times, which means there are several process instances in every period, resulting in several cash flows CF_{PI_r} per period. The sum $CF_{P_{r_j}}$ of cash flows in one period is again a random variable, which can be approximated by a random variable that is normally distributed, since the $CF_{PI_{r_i}}$ are identically distributed and we assume in the following that they are independent of each other (central limit theorem (Feller, 1968)). Hence, for each future period *j, j=0,…, J*, the sum of its cash flows can be represented by a normally distributed random variable. As a result, the net present value NPV_{P_r} follows a normal distribution as well (see Hillier (1963)). Accordingly, we formulate our first assumption.

(A1) *There are no kinds of dependences between the processes, i.e. between the* NPV_{P_r}*, as well as no dependences between process instances and between periods. Each return* NPV_{P_r} *is normally distributed.*

The assumption, that there are no dependences, is a simplification in this first approach that reduces the formalism significantly and eases the communication of this approach. This way, we can focus on one process and not on all processes in a company, just as in Davamanirajan et al. (2006). In addition, practical experience shows that it is difficult to measure these dependences, for example by using correlation coefficients, and it is very unlikely that the values of the correlations are known. Since NPV_{P_r} follows a normal distribution, it is fully described by its expected value and its variance, which are considered by our next assumption.

(A2) *The expected value* $E\left[NPV_{P_r}\right]$ *and the variance* $Var\left[NPV_{P_r}\right]$ *of* NPV_{P_r}*, r=1,…,R, of a process* P_r *are finite.*

We want to point out, that we do not assume to know the exact values of both the expected value and the variance of NPV_{P_r}, but that they are finite. So far, we have been at the process level. In the following, we will assume how the processes are connected with the value of a company.

The value of a company includes the net present values of all cash flows of a company, which means we must consider all of these cash flows. This could be done by separating cash flows that are caused by processes and cash flows that are caused by anything else, which would give us one random variable that represents all these other cash flows. However, for reasons of simplicity, we assume that all cash flows of a company are due to processes, i.e. the return of a company is the sum of the returns of the processes. This simplification can be assumed if a company is seen as a portfolio of processes that cause all cash flows of a company, as everything could be considered to be a process.

(A3) *The (risk-adjusted) value Φ_C of a company C is entirely caused by its processes P_r, r=1,…,R. The return of a company C is represented by the random variable NPV_C. It is the sum of the returns of the processes NPV_{P_r} of the company, i.e. $NPV_C = \sum_{r=1}^{R} NPV_{P_r}$.*

The fact that the return of a company is the sum of the returns of the processes, implies that NPV_C is the uncertain net present value of all uncertain future cash flows inside the company (R1) and that its expected value (expected return) and variance are finite. In addition, with assumption (A1) the return of a company NPV_C follows a normal distribution. With this assumption, we connected the return of a company with the return of its processes, which is essential to fulfill (R2).

We aim to decide between process alternatives based on the change of the random variable NPV_C. Therefore, it seems reasonable to use decision theory under uncertainty. In particular, we use the expected utility theory (Copeland et al., 2005). As the stakeholders should decide based on the best interest of the company/person in charge (*decision maker*) (R3), we make the following assumption regarding the decision maker, similar to Fridgen and Müller (2009) and Zimmermann et al. (2008).

(A4) *The decision maker has a constant risk aversion with respect to returns (Pratt, 1964) and maximizes the expected utility.*

As stated in Bamberg and Spremann (1981), a constant risk aversion is "flexible enough to cover a broad spectrum of risk averse patterns", which is why we can assume the risk aversion to be constant.

II.1.3.2 Value-based Selection of Processes

Since NPV_C follows a normal distribution, it is fully described by its expected value and its variance, which means we can make a decision based on the change of these two quantities. In order to fulfill requirements (R2) and (R3), we need an objective function that combines this expected return of a company and a risk contribution (the variance), as well as the risk attitude of the decision maker and is compatible with assumption (A4). The following function fulfills the requirements and is based on expected utility theory to decide between different process alternatives:

$$\Phi_C \coloneqq \Phi(\mu_C, \sigma_C) = \mu_C - \frac{\alpha}{2}\sigma_C^2, \tag{1}$$

where μ_C is the expected value of NPV_C, σ_C^2 is the variance and α is the risk attitude of a decision maker, the so called risk aversion constant (Freund, 1956). For a risk averse decision maker it is $\alpha > 0$ (Pratt, 1964). Although it is not an easy task to determine α, Bamberg and Spremann (1981) show how α could be determined. They show that in order to determine α, the decision maker is asked certain questions in order to elicit the risk attitude. In addition, the difficulties to choose the right questions to elicit the required information are presented.

This function was introduced by Freund (1956), and applied in more recent works such as Fridgen and Müller (2009), Longley-Cook (1998) and Zimmermann et al. (2008). According to Freund (1956), function (1) can be used, if the decision maker has a utility function of the form $u(x) = 1 - e^{-\alpha \cdot x}$ and if NPV_C is normally distributed. This is due to the fact that if function (1) is maximized, the expected utility $E[u(NPV_C)] = 1 - e^{-\alpha \cdot (\mu_C - 0.5 \cdot \alpha \cdot \sigma_C^2)}$ is maximized. As we assume the decision maker to have a constant risk aversion (A4), the decision maker has indeed an exponential utility function (Bamberg and Spremann, 1981). Such exponential utility function can be $u(x)$, which is also similar to empirically found utility functions by Swalm (1966). Furthermore, since NPV_C is normally distributed, Φ_C is the certainty equivalent $u^{-1}(E[u(NPV_C)])$ (Copeland et al., 2005) of the normal distributed return of a company with constant risk aversion and can therefore be seen as the (risk-adjusted) value of a company (Bamberg et al., 2006).

We will show that in order to know which process alternative increases Φ_C the most, we do not need to know the value of μ_C and σ_C^2. It is enough to know how much the expected value and the variance of the return of the process, which is to be redesigned, change through the redesign, making the model more efficient. For a formal way to show this, we will introduce some additional notations.

Let, without loss of generality, P_R be the process that a company C might want to redesign. Further, let P_R' be any process alternative of P_R and Φ_C' the new value of C, if the modifications in P_R' would be implemented, i.e if P_R' would be selected as the alternative to P_R. Then let be

- $NPV_C = \sum_{r=1}^{R} NPV_{P_r}$ the return of C without redesign, with $\mu_{P_R} := E\left[NPV_{P_R}\right]$ and $\sigma_{P_R}^2 \coloneqq Var\left[NPV_{P_R}\right]$,

- $NPV'_C := \sum_{r=1}^{R-1} NPV_{P_r} + NPV_{P'_R}$ the return of C if P_R would be redesigned to P'_R, with $\mu_{P'_R} := E\left[NPV_{P'_R}\right]$ and $\sigma^2_{P'_R} \coloneqq Var\left[NPV_{P'_R}\right]$,
- $\Phi_C := \Phi(\mu_C, \sigma_C) = \mu_C - \frac{\alpha}{2}\sigma_C^2 = E[NPV_C] - \frac{\alpha}{2}Var[NPV_C]$ the value of C without redesign,
- $\Phi'_C := \Phi(\mu'_C, \sigma'_C) = \mu'_C - \frac{\alpha}{2}{\sigma'_C}^2 = E[NPV'_C] - \frac{\alpha}{2}Var[NPV'_C]$ the value of C if P_R would be redesigned to P'_R,
- $\Delta\Phi'_C := \Phi'_C - \Phi_C$ the difference in the values of C if P_R would be redesigned to P'_R, and
- $\Delta\mu_{P'_R} \coloneqq \mu_{P'_R} - \mu_{P_R}$ and $\Delta\sigma^2_{P'_R} \coloneqq \sigma^2_{P'_R} - \sigma^2_{P_R}$ the differences between the existing process and any process alternative P'_R in terms of the expected return and the variance of the return, respectively.

With this, it can be formally shown that

$$\Delta\Phi'_C = \Delta\mu_{P'_R} - \frac{\alpha}{2}\Delta\sigma^2_{P'_R}. \tag{2}$$

If $\Delta\mu_{P'_R}$ and $\Delta\sigma^2_{P'_R}$ can be determined for all process alternatives of P_R, then we can select the best alternative in the interest of the company at the process level (R3) by calculating $\Delta\Phi'_C$. Therefore we extend the assumption (A2) to the following assumption (A2)'.

(A2)' Assumption (A2) holds. In addition, although the expected value and variance of NPV_{P_R} and $NPV_{P'_R}$ are not known, the differences $\Delta\mu_{P'_R}$ and $\Delta\sigma^2_{P'_R}$ can be determined. This is true for every process alternative P'_R.

Therefore, if we know how much the return of a process changes, in terms of its expected value and its variance, we can calculate how much the change of this process return NPV_{P_R} – ceteris paribus – changes the value of the company, connecting the process level with the value of the company. Thus, the stakeholders at the process level can decide at that level in the best interest of the company, using the same risk attitude as the decision maker and not their own attitude towards risk (R3). In the end, we select the process alternative with the highest $\Delta\Phi'_C$ to realize a value-based process improvement. Of course, if $\Delta\Phi'_C$ is negative for all process alternatives, there would be no redesign unless a redesign is necessary, for example due to new regulations. Equation (2) presents the heart of our approach (our artifact) for a value-based process management.

In addition, we can also select between newly designed processes P'_R, and not just redesigns, in the special case of the quantities $\mu_{P'_R}$ and $\sigma^2_{P'_R}$ being known, and setting $\mu_{P_R} = 0$ and $\sigma^2_{P_R} = 0$, using equation (2). Furthermore, with the function $\Phi_{P_r} := \Phi(\mu_{P_r}, \sigma_{P_r}) = \mu_{P_r} - \frac{\alpha}{2}\sigma^2_{P_r}$, we can obtain the stand-alone value contribution of a process P_r to the value of a company.

II.1.4 Application

In Neiger et al. (2006), a scenario is given and four alternatives (including the existing process) for a process are presented. We will use our approach to select an alternative. The scenario is given as:

"In June 2005, the payroll process of a large educational institution failed. More than 4,000 employees were not paid on schedule, but on the following day instead. This unanticipated delay resulted in bounced checks, rejected automatic bill payments and declined check card purchases by staff and faculty, who did not receive information about this delay in time. A hastily installed mediation procedure allowed employees to receive their compensation as a cash payout, which was then deducted from their following month's paycheck, depleting cash reserves of the university.

An investigation of the problem revealed that the cause for the delay was a data entry mistake made by a staff member who entered the wrong payroll date in one step of the payroll process. Two administrators signed off on the scheduled payroll run and did not notice the wrong date. The payroll run order was transmitted to the university's bank for processing and when the error was discovered it was too late to re-schedule the payroll run."

We do not know the whole payroll process P_r, which means we cannot determine Φ_{P_r}. However, as equation (2) states, this is not necessary, as it is enough to know the effects of the possible redesigns, which result in different $\Delta\Phi'_C$. The existing, and to be changed, sub-process of the payroll process, can be represented by a sequential process *SP*. The existing process *SP* has one activity "Enter Payroll run information", with a cash outflow of \$1,000 per process instance, and two separate activities "Approve Payroll run", with each having a cash outflow of \$500 per process instance. In case the process *SP* goes wrong, the rectification costs are \$250,000. In addition to the existing process *SP*, they give three process alternatives. All four alternatives have a probability that a problem occurs (*failure probability*), that could result in the cash outflow of \$250,000. The alternatives are:

- *Alternative 1*: One entry activity and one approval activity; failure probability: 1.5%.
- *Alternative 2*: Two separate entry activities, where two different persons enter the same data, and one approval activity; failure probability: 0.075%.
- *Alternative 3* (existing process): One entry activity and two separate approval activities, where two different administrators have to approve the data; failure probability: 0.45%.
- *Alternative 4*: Two separate entry activities and two separate approval activities; failure probability: 0.0225%.

We state some assumptions that are not explicitly made in Neiger et al. (2006), but are implicit to some extent, and need to be made to use our approach, before we present $\Delta\Phi_C'$ for each alternative. In Neiger et al. (2006), the alternatives are compared on the basis of the expected cash flow per process instance for each alternative. This is only valid, if the number n_{r_j} of process instances, which are executed in each period, is the same for each alternative, i.e. the redesign does not have any effect on the number of process instances. For each activity "Enter Payroll run information" is $E[CF_{enter}] = -1{,}000$ and $Var[CF_{enter}] = 0$ and for each activity "Approve Payroll run" is $E\left[CF_{approve}\right] = -500$ and $Var\left[CF_{approve}\right] = 0$ for every process instance. When modeling the four alternatives it is easy to add this information of the cash flows CF_{enter} and $CF_{approve}$ to the activities, just as the rectification costs to the process, and simulate the alternatives to determine a sample mean and sample variance of CF_{SP_j} for this sub-process. Since such a payroll process exists in many educational institutions and companies, the number of simulated process instances is set to $n_{r_j} = 25{,}000$. Only one period is considered in Neiger et al. (2006) and there are no cash outflows to change the existing process. Thus, for the sake of simplicity and to be comparable with Neiger et al. (2006), we set $J=0$ and $I_{SP}=0$. A comparison with $J>0$ would be easily possible with the NPV. However, in this scenario it would lead to the same result, as n_{r_j} is the same for each alternative. This means, CF_{SP_0} of the sub-process can be determined and it is $CF_{SP_0} = NPV_{SP}$ of the sub-process. In terms of the whole process, we assume that the whole payroll process P_r can be represented as a sequential process, with all activities having a variance of zero, just as it is with the activities "Enter Payroll run information" and "Approve Payroll run". The rectification costs represent all cash outflows if something goes wrong with P_r. With this, and since n_{r_j} is fairly high, it can be assumed that the return NPV_{P_r} of the process P_r as well as the return NPV_{SP} of the sub-process SP

follow a normal distribution (A1), where $E[NPV_{SP}]$ is set to the sample mean of CF_{SP_j} and $Var[NPV_{SP}]$ is set to the sample variance of CF_{SP_j} (law of large numbers) (A2). As nothing differently is stated in Neiger et al. (2006), we can assume that there are no kinds of dependences (A1). Further, we assume that (A3) and (A4) hold as well and the risk aversion of the decision maker is assumed to be α*=0.0001*.

With this, we can determine $\Delta\Phi_C'$ for each alternative via several simulation runs of *SP*. The model to determine $\Delta\Phi_C'$ can easily be implemented into a process modeling tool. The result is presented in table 1. It is P_R the existing payroll process (alternative 3), P_R' represents the alternatives.

Alternative	$\Delta\mu_{P_R'}$	$\Delta\sigma^2_{P_R'}$	$\Delta\Phi_C'$
1	$-53.1 \cdot 10^6$	$16.1 \cdot 10^{12}$	$-8.6 \cdot 10^8$
2	$10.9 \cdot 10^6$	$-5.8 \cdot 10^{12}$	$3.0 \cdot 10^8$
3	0.00	0.00	0.00
4	$1.7 \cdot 10^6$	$-6.7 \cdot 10^{12}$	$3.3 \cdot 10^8$

Table 1. Changes of the Value Contribution of the Payroll process with different redesigns

Of course, alternative 3 has $\Delta\Phi_C' = 0$, because it is the existing process. With the presented approach we select the alternative with the highest $\Delta\Phi_C'$ which results in the decision to use alternative 4.

II.1.5 Evaluation

First, we analyze to what extend we achieve our aim and how far we close the research gap, which is identified in section 2. Then we compare our approach with the model used in Neiger et al. (2006).

II.1.5.1 Closing the Research Gap

In subsection 2.1, we state three requirements for an approach to enable a value-based process management, providing the objectives for such an approach. We concluded, that there are works that already satisfy (R1) and (R3) to a certain extent. With our approach, we can consider multiple periods as we use net present values (R1). We present an appropriate objective function for a company with equation (1), fulfilling requirement (R2). With equation (2) we provide a way to make decisions at the process level in the best interest of the company (R3). Thus, we closed the research gap and developed a complete and effective artifact (Hevner et al., 2004).

II.1.5.2 Comparison with Competing Artifact

Hevner et al. (2004) stressed that an artifact must be evaluated with respect to the practical utility provided. We presented the practical utility in the application, as we used it in a real world scenario. We used our approach to select a process alternative and will now examine the result in comparison to Neiger et al. (2006). Their recommendation is to choose alternative 2. Our approach results in the same recommendation in the special case of a risk neutral decision maker ($\alpha=0$). However, if we consider the decision maker to be risk averse (with $\alpha=0.0001$), we recommend to use alternative 4, although $\Delta\mu_{P_R'}$ is higher for alternative 2 and α is close to zero. It is the same result when taking the sensitivity analysis in Neiger et al. (2006) into account. This is due to the lowest failure probability in alternative 4, lowering the variance of CF_{SP_0} of that alternative, and the risk attitude of the decision maker. This demonstrates how important it is to consider the risk aversion of a decision maker and the deviation of the return, which is possible with our approach. We can show this importance even more, if we do not set $Var[CF_{enter}] = 0$ and $Var[CF_{approve}] = 0$, as it is done in table 2.

$\sqrt{Var[CF_{approve}]}$		0	50	100	200	300	400
$\sqrt{Var[CF_{enter}]}$	0	Alt. 4	Alt. 2	Alt. 2	Alt. 2	Alt. 2	Alt. 2
	50	Alt. 4	Alt. 2	Alt. 2	Alt. 2	Alt. 2	Alt. 2
	100	Alt. 4	Alt. 2	Alt. 2	Alt. 2	Alt. 2	Alt. 2
	200	Alt. 3	Alt. 3	Alt. 3	Alt. 1	Alt. 1	Alt. 1
	300	Alt. 3	Alt. 3	Alt. 3	Alt. 1	Alt. 1	Alt. 1
	400	Alt. 3	Alt. 3	Alt. 3	Alt. 1	Alt. 1	Alt. 1

Table 2. Resulting Alternative with different Standard Deviations of the Activities' Cash Flows

In table 2 it can be seen that if the deviations of the activity cash flow are small, then the rectification costs are the major risks to be considered, which is why it is reasonable to use more activities to look for mistakes. However, as the deviations of the activity cash flow increase, it is better to use fewer activities to lower the risk of deviations and take a higher risk that an error occurs. This demonstrates how important it is to consider deviations from expected values and not only the expected values. It shows the advantage of our approach as compared to a model that decides solely on the basis of expected values. Neiger et al. (2006) do a sensitivity analysis to account for the risk that these returns may vary, but such variation is not part of the function to decide for a process alternative.

It can be noticed, that we talk about the deviation from the expected value as a risk. However, Neiger et al. (2006) consider the risk that data is entered wrongly and the mistake is not discovered. There seems to be a different understanding of risk, which makes it questionable as to whether the two approaches can be compared. According to Hansson (2005) there is not only one single meaning of risk. The author also states that “at present, by far the most common technical definition of risk” is “risk as statistical expectation value of unwanted events, which may or may not occur.” In this *first definition*, risk is seen as probability multiplied by the consequence of an unwanted event, which is an expected loss. These kinds of risks are part of the expected return. This means, they are included in both approaches. However, since we want to provide a value-based approach, we have to consider the meaning of risk from a finance perspective. In finance, risk “refers to the likelihood that we will receive a return on an investment that is different from the return we expect to make.” (Damodaran, 2002) This *second definition* sees risk as difference from an expected return and therefore considers good and bad unwanted events. With our approach, the expected loss is part of the expected return, and the variation of this loss/cash outflow is part of the risk contribution. Therefore, our approach can handle this kind of risk (second definition) as well, extending the model used in Neiger et al. (2006).

II.1.6 Conclusion

In this paper, we describe how to perform the improvement of processes in an *effective* and *efficient* manner. It is effective, since it directly targets the value of a company which is the main focus of managers, and it is efficient, since only $\Delta\Phi'_C$ is necessary, but not Φ'_C as a whole. Related to the guidelines for conducting design science research by Hevner et al. (2004), we can summarize as follows: Our *artifact* is an approach to support decisions on how to improve a process with the goal of increasing the value of a company. We regard this as an important step to improve processes from a business view during a process (re-)design. The model is formally noted and can thus be evaluated. This builds the basis to use the common evaluation criteria of process improvements like time, costs, customer satisfaction, output quality, etc. Those criteria need to be specified on process level and transformed into monetary values, so that their return and risk contribution can be determined. A detailed analysis of how to incorporate these criteria should be addressed in further research.

Our artifact is thought to contribute to process management, to design and adapt processes in the interest of a company and to be useful regarding decisions at the process level. Since such a statement cannot hold for every process, the question of when to apply a model to perform a value-based process improvement needs to be clarified. Such a clarification is required to specify the boundaries within which the model is expected to be applied. The amount of information that is needed could put a limitation to the processes. In order to get the information, there are initial costs to analyze the problem domain. If this information can later be reused during further (re-)design projects, then the costs to retrieve the information might be worthwhile. This might limit the approach to processes that are redesigned more often. However, since BPM is an ongoing task inside a company and the risks of processes can be quite considerable, we assume that it is worthwhile in many cases, to gather the information. Another limitation is the assumption of normal distribution. This assumption holds for instance, due to the central limit theorem, if there are no dependences and if processes are executed several times, which limits the approach for example to highly repetitive processes.

Further work is proposed on the question of how dependences can be considered, as it might have a big impact on the selection of the right process alternative. Other work is necessary if the number of process alternatives that need to be compared is very high. For efficiency reasons, this task should be automated. Thus, we would need the corresponding process models that are extended with financial values. With these process models, combined with the use of our or similar approaches to select process alternatives on the basis of financial values, the selection could be automated. This would also allow the valuation of complex processes, where the gathering of the required amount of information limits the applicability of our approach. For this future work, the designed model is a reliable basis for value-based process improvement, to support the CIO to meet the primary business priority of IT.

II.1.7 References

Bamberg, G., Dorfleitner, G., and Krapp, M. (2006). Unternehmensbewertung unter Unsicherheit: Zur entscheidungstheoretischen Fundierung der Risikoanalyse. Zeitschrift für Betriebswirtschaft, 76 (3), 287-307.

Bamberg, G., Spremann, K. (1981). Implications of Constant Risk Aversion. Mathematical Methods of Operations Research, 25 (7), 205-224.

Coenenberg, A.G., Salfeld, R. (2007). Wertorientierte Unternehmensführung: Vom Strategieentwurf zur Implementierung. 2nd Edition, Schäffer-Poeschel, Stuttgart.

Copeland, T.E., Weston, J.F., and Shastri, K. (2005). Financial Theory and Corporate Policy. 4th Edition, Pearson Education Inc., Boston, MA.

Damodaran, A. (2002). Investment Valuation: Tools and Techniques for Determining the Value of Any Asset. 2nd Edition, Wiley & Sons, Inc, New York, NY.

Davamanirajan, P., Kauffman, R.J., Kriebel, C.H., and Mukhopadhyay, T. (2006). Systems Design, Process Performance, and Economic Outcomes in International Banking. Journal of Management Information Systems, 23 (2), 65-90.

Faisst, U., Buhl, H.U. (2005). Integrated Enterprise Balancing mit integrierten Ertrags- und Risikodatenbanken. Wirtschaftsinformatik, 47 (6), 403-412.

Feller, W. (1968). An Introduction to Probability Theory and Its Applications. Vol. 1. 3rd Edition, John Wiley & Sons, Inc., New York, NY.

Freund, R.J. (1956). The Introduction of Risk into a Programming Model. Econometrica, 24 (3), 253-263.

Fridgen, G., Müller, H. (2009). Risk/Cost Valuation of Fixed Price IT Outsourcing in a Portfolio Context. In Proceedings of the 30th International Conference on Information Systems (ICIS), (Chen, H. and Slaughter, S.A. Eds.), Phoenix, Arizona.

Gartner. (2010). Leading in Times of Transition: The 2010 CIO Agenda. 2010, 06.04, http://drishtikone.com/files/2010CIOAgenda_ExecSummary.pdf.

Gulledge, T.R., Hirschmann, P. and Scheer, A. (1997). Value-Based Management of Inter-Organizational Business Processes. In Proceedings of the third international conference on Wirtschaftsinformatik '97 - Internationale Geschäftstätigkeit auf der Basis flexibler Organisations-strukturen und leistungsfähiger Informationssysteme (Krallmann, H. Ed.), p. 73-98, Berlin.

Hansson, S.O. (2005). Seven Myths of Risk. Risk Management: An International Journal, 7 (2), 7-17.

Hevner, A.R., March, S.T., Park, J., and Ram, S. (2004). Design Science in Information Systems Research. MIS Quarterly, 28 (1), 75-105.

Hillier, F.S. (1963). The Derivation of Probabilistic Information for the Evaluation of Risky Investments. Management Science, 9 (3), 443-457.

Ittner, C.D., Larcker, D.F. (2001). Assessing empirical research in managerial accounting: a value-based management perspective. Journal of Accounting and Economics, 32 (1-3), 349-410.

Kanevsky, V., Housel, T.J. (1995). Value-Based Business Process Reengineering: An Objective Approach to Value Added. In Business Process Change: Reengineering Concepts, Methods and Technologies (Grover, V. and Kettinger, W.J. Eds.), Idea Group Publishing, Hershey, PA.

Koller, T., Goedhart, M., and Wessels, D. (2005). Valuation: Measuring and Managing the Value of Companies. 4th Edition, John Wiley, New Jersey.

Lewis, M., Young, B., Mathiassen, L., Rai, A., and Welke, R. (2007). Business process innovation based on stakeholder perceptions. Information, Knowledge, Systems Management, 6 (1), 7-27.

Longley-Cook, A.G. (1998). Risk-Adjusted Economic Value Analysis. North American Actuarial Journal, 2 (1), 87-100.

Neiger, D., Churilov, L., zur Muehlen, M. and Rosemann, M. (2006). Integrating Risks in Business Process Models with Value Focused Process Engineering. In Proceedings of the 14th European Conference on Information Systems (Ljungberg J. and Andersson M. Eds.), p. 1606-1615.

Peffers, K., Tuunanen, T., Rothenberger, M.A., and Chatterjee, S. (2008). A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems, 24 (3), 45-77.

Pratt, J.W. (1964). Risk Aversion in the Small and in the Large. Econometrica, 32 (1), 122-136.

Smith, H., Fingar, P. (2003). IT Doesn't Matter--Business Processes Do: A Critical Analysis of Nicholas Carr's I.T. Article in the Harvard Business Review. Meghan-Kiffer Press, Tampa, FL.

Swalm, R.O. (1966). Utility Theory - Insights into Risk Taking. Harvard Business Review, 44 (6), 123-136.

vom Brocke, J., Recker, J.C., and Mendling, J. (2010). Value-oriented process modeling: integrating financial perspectives into business process re-design. Business Process Management Journal, 16 (2), 333-356.

Zimmermann, S., Katzmarzik, A. and Kundisch, D. (2008). IT Sourcing Portfolio Management for IT Service Providers - A Risk/Cost Perspective. In Proceedings of 29th International Conference on Information Systems (ICIS) (Boland, R., Limayem M. and Pentland B. Eds.), Paris.

II.2 Research Paper 2: "Bringing Value-Based Business Process Management to the Operational Process Level"

Authors:	Manuel Bolsinger[a] [a] FIM Research Center, Department of Information Systems Engineering & Financial Management (Prof. Dr. Hans Ulrich Buhl), University of Augsburg, Germany manuel.bolsinger@fim-rc.de

Appears in:	Information Systems and e-Business Management, 2014. The final publication is available at http://link.springer.com

Abstract:

For years, improving processes has been a prominent business priority for Chief Information Officers. As expressed by the popular saying, "If you can't measure it, you can't manage it," process measures are an important instrument for managing processes and corresponding change projects. Companies have been using a value-based management approach since the 1990s in a constant endeavor to increase their value. Value-based business process management introduces value-based management principles to business process management and uses a risk-adjusted expected net present value as the process measure. However, existing analyses of this issue operate at a high (i.e., corporate) level, hampering the use of value-based business process management at an operational process level in both research and practice. Therefore, this paper proposes a valuation calculus that brings value-based business process management to the operational process level by showing how the risk-adjusted expected net present value of a process can be determined. We demonstrate that the valuation calculus provides insights into the theoretical foundations of processes and helps improve the calculation capabilities of an existing process-modeling tool.

II.2.1 Introduction

Constant change in their economic, political, and social environments is forcing companies to strive for increased efficiency and more frequent innovation (Becker and Kahn 2005, p. 3), a situation in which the management and, in particular, the improvement of processes play a considerable role (González et al. 2010; Thome et al. 2011; van der Aalst 2013; vom Brocke et al. 2011a). One indicator of process improvement's prominent role is the fact that companies invest considerable amounts of money to develop their business process management (BPM) capabilities and realize improvement activities (Wolf and Harmon 2012). The volume of research on process improvement has also increased (Sidorova and Isik 2010, p. 572).

In their efforts to improve processes, researchers and practitioners alike must establish a basis on which it can be decided that an alternative (or "to-be") process is better than an existing (or "as-is") process. The instruments deemed appropriate for determining the extent to which a process alternative improves an existing process are called "process measures" (González et al. 2010; Tregear 2012; zur Muehlen and Shapiro 2010). When the value of a process measure of an alternative process is greater than that of an existing process, it might be reasonable to implement the alternative process and thus improve the existing process. However, there are many process measures, and, while the value of one measure may suggest a process improvement, the value of another may indicate the opposite. For example, the dimensions of time, cost, quality, and flexibility, often used to evaluate process improvement, comprise the so-called "devil's quadrangle" because, "in general, improving [a process] upon one dimension may have a weakening effect on another" (Reijers and Liman Mansar 2005, p. 294). Hence, process managers have to consider these complementary and competitive goal relations when determining whether an alternative process improves an existing process. In order to resolve potential conflicts among goals, process managers need integrated approaches that consolidate various goals into one overall goal, thus allowing them to make decisions based on that overall goal.

Value-based BPM introduces into BPM an overall goal in line with economic theory (Buhl et al. 2011). Value-based BPM applies value-based management principles to process decision-making and aims to increase company value from a long-term perspective (Ittner and Larcker 2001; Koller et al. 2010; Young and O'Byrne 2001), thus supporting process improvement from a monetary-centered view of BPM. Companies have been using value-based management since the 1990s in their constant endeavor to increase their value

(Coenenberg and Salfeld 2007, p. 3). Almost two thirds of the 30 companies on the German stock index (DAX), representing Germany's major companies, explicitly state in their 2013 annual reports that they follow a value-based management approach. Moreover, the 2013 CIO agenda (Gartner 2013) identified "harvest value from business process changes" as one of their three performance profiles. Hence, value-based BPM not only provides an approach for integrating different goals but also takes on a business perspective by facilitating the overall goal of increasing company value, wherein a process' value contribution is determined by its risk-adjusted expected net present value, or "rNPV" (Bolsinger et al. 2011; Buhl et al. 2011). A process alternative should be implemented as an improvement whenever its rNPV is higher than that of the existing process.

However, although research suggests the transferability of value-based management to BPM, current studies operate at a high (i.e., corporate) level and do not show how the rNPV is to be calculated in detail, particularly with reference to a process' control flow, which is important to connect the corporate level with the operational level (Rotaru et al. 2011; vom Brocke et al. 2010). Furthermore, in the practice of BPM, modeling tools (e.g., IBM WebSphere Business Modeler Advanced, Bonita Studio, TIBCO Business Studio, ibo Prometheus Klassik and Bizz Designer) cannot determine the rNPV and, thus, do not support value-based BPM. In order to substantiate value-based BPM from both theoretical and practical points of view, additional research capable of establishing the appropriate theoretical foundations is necessary (Vergidis et al. 2008).

This paper contributes to the literature *by providing a valuation calculus for determining the risk-adjusted expected net present value of a process*. After the valuation calculus is implemented, a process-modeling tool could calculate the rNPV for various process alternatives, from which a process manager could choose for a process improvement project. This functionality would provide a valuable asset for process managers (van Hee and Reijers 2000; Vergidis et al. 2008) and bring value-based management into the practice of BPM.

This paper, reflecting the design science research process presented in Peffers et al. (2008), is organized as follows. After *motivating the importance of the problem* in this section, Section 2 provides more background information about value-based BPM and positions it against other BPM approaches related to value-based BPM. Based on this theoretical background, we derive the requirements for the valuation calculus that *define its objectives* before discussing related work. In Section 3, we introduce a basic illustrative example to provide a better understanding of the issues raised in the subsequent sections. In Section 4, the valuation calculus (our *artifact*) is *designed* using a formal-deductive research approach (Meredith et al. 1989). In Section 5, we focus on the *evaluation* of the valuation calculus in an artificial setting (Sonnenberg and vom Brocke 2012; Venable et al. 2012). We then present a feature comparison, a comparison with a related artifact, and a *demonstration* of the feasibility of the artifact by solving an exemplary problem instance and by illustrating how the knowledge of the valuation calculus corrected the calculation logic of the process-modeling tool of the CubeFour company. Finally, the last section summarizes our results and provides an outlook for future study.

II.2.2 Theoretical Background

II.2.2.1 Value-based Business Process Management

The value-based BPM paradigm focuses on the value that a newly designed process or a change in an existing process contributes to a company (Buhl et al. 2011; vom Brocke et al. 2010). In doing so, value-based BPM introduces value-based management principles to BPM, thus motivating process-related decisions according to a well-established management approach. Before discussing value-based BPM in detail, we will first outline the principles of value-based management.

Value-based management aims to sustainably increase a company’s value from a long-term perspective (Ittner and Larcker 2001; Koller et al. 2010; Young and O'Byrne 2001). It extends the shareholder value approach that traces back to Rappaport (1986) and was further advanced by Copeland et al. (1990) and Stewart and Stern (1991). Taking a long-term perspective, value-based management complies with the stakeholder value approach (Danielson et al. 2008), which is important for a less decision-making oriented perspective on value-based BPM (vom Brocke et al. 2009). For value-based management to be fully realized, all activities on all company levels must be aligned with the goal of maximizing company value (Coenenberg and Salfeld 2007). The same holds true for company processes: each process has to contribute to the value of the company, and a process should be changed only if its value contribution can be increased.

Following a value-based management in BPM requires that process decisions be based on cash flows, that the time value of money be considered, and that the risks associated with the cash flows be taken into account (Buhl et al. 2011), all of which support process improvement from a monetary-centered view of BPM. The risks arise because cash flows are uncertain; thus, cash flows are modeled as random variables. These cash flows originate from every execution of a process, each of which is executed not only a few times but several times within a given planning horizon. This cash flow structure is brought together into one quantity through the net present value (NPV). The NPV of a process is thus uncertain, which is why it is also modeled as a random variable, and builds the foundation of a value-based BPM. As described in Bolsinger et al. (2011), the NPV of a process is expressed as follows:

$$NPV = -I + \sum_{t=0}^{T} \frac{\sum_{j=1}^{n_t} CF_{P_j}}{(1+i)^t}, \qquad (1)$$

where I denotes an initial process investment,

$T+1$ the number of periods that a process will be executed within a certain planning horizon,

n_t the number of times a process is executed within a period t,

CF_{P_j} the process cash flow of the *j*th execution of process P, and

i "the rate of interest which properly reflects the investor's time value of money" (Hillier 1963, p. 447).

The initial process investment can be, for example, the cash outflow needed to design an alternative process or change to one. This investment is different for each process alternative and can be set to zero for the existing process when comparing process alternatives to the existing one.

As mentioned, *NPV* is an uncertain quantity because CF_P is uncertain. Therefore, comparing the *NPV*s of different processes is difficult because no process (alternative) has a single value by which the best process (alternative) (i.e., that with the best NPV) may be determined. To comply with value-based management, value-based BPM uses the expected utility theory to determine a single value per process (alternative) by using the certainty equivalent Φ of *NPV* (Buhl et al. 2011; Copeland et al. 2005, p. 54). The certainty equivalent corresponds to the process' contribution to company value and is (as mentioned) the rNPV. The certainty equivalent is expressed as follows:

$$\Phi = E[NPV] - \frac{\alpha}{2} Var[NPV], \qquad (2)$$

where $E[NPV]$ denotes the expected value of *NPV*,

$Var[NPV]$ the variance of *NPV*, and

α the risk aversion constant, representing the risk attitude of the decision maker (Freund 1956).

The expected value is used as a process measure to capture the expected return of a process, while the variance is used to measure the risk of a process. The expected value is adjusted by the risk, depending on the risk attitude of the decision maker. The adjustment of the expected value results in the risk-adjusted value the decision maker assigns to the process. Bamberg and Spremann (1981) show how it is possible to elicit the needed information from decision makers to determine their utility function and translate it into a value of α. Decision makers must be asked certain questions, from which the utility function is then determined. More about preference elicitation for utility measurement can be found in works such as Abdellaoui et al. (2013), Andersen et al. (2008), Beer et al. (2013), Friedman and Savage (1948), Mosteller and Nogee (1951), and Swalm (1966). Another approach to determining α is the market price perspective, which uses the capital asset pricing model (CAPM). In this model, α/2 is the market price of risk, which can be determined through the CAPM's so-called "price equation" (Kruschwitz and Husmann 2010). Kasanen and Trigeorgis (1994) show how α can be calculated within the CAPM and it is estimated using actual market data (the authors' parameter *m* corresponds to our α).

The result is an integrated risk/return decision function based on a theoretically well-founded method, which is also used to make decisions in other domains (Datar et al. 2001; Fridgen and Müller 2009; Gibbons 2005; Longley-Cook 1998; Sen and Raghu 2013; Zimmermann et al. 2008). The certainty equivalent is used to decide if a process alternative improves an existing process (see Figure 1).

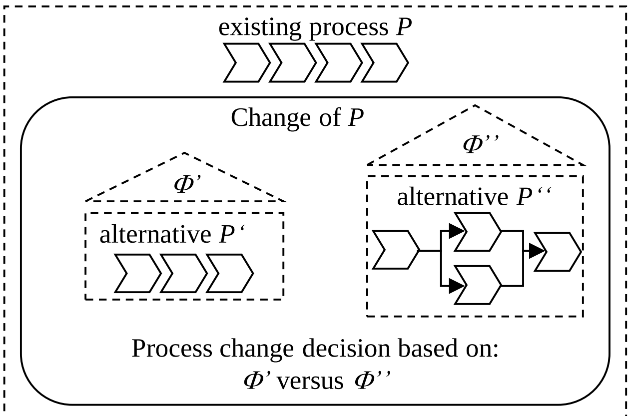


Figure 1. Process change decisions regarding process improvement

The merits and limitations of value-based BPM become clearer when positioned against related approaches such as goal-oriented BPM (Kueng and Kawalek 1997; Neiger and

Churilov 2004a), value-focused BPM (Neiger and Churilov 2004b; Rotaru et al. 2011), value-driven BPM (Franz et al. 2011), and value-oriented BPM (vom Brocke et al. 2010).

Goal-oriented BPM demands that processes fulfill certain goals, which must be clearly stated in order to clarify what the process must achieve or avoid (Kueng and Kawalek 1997); the goals can be either functional (e.g., “sell insurance”) or non-functional (e.g., low operational costs, short cycle time). Whatever goals are chosen, “the goal-oriented view of business process engineering dictates that business goals are the driving force for structuring and evaluating business processes” (Neiger and Churilov 2004a, p. 150). Thus, the goals provide the basis for evaluating how well a process is designed, but the process managers have to decide what those goals will be.

Value-focused BPM shows how value-based thinking (Keeney 1994) helps elicit essential goals from decision makers, facilitating goal-oriented BPM. In this context, values are “principles for evaluating the desirability of any possible alternative or consequence. They define all that you care about in a specific decision situation” (Keeney 1994, p. 33). Value-focused BPM shows how value-based thinking can substantiate the goals of a process and be incorporated into process modeling (Neiger and Churilov 2004b).

Value-driven BPM provides the values to which organizations aim when beginning a BPM initiative. These values consist of the core value “transparency” and the three value pairs “efficiency-quality,” “agility-compliance,” and “integration-networking” (Franz et al. 2011). These values are suggested as BPM goals, each pair consisting of “two values that tend to be oppositional” (Franz et al. 2011, p. 6) therefore presenting conflicting goals. Thus, possible goals of goal-oriented BPM have been provided, but how to measure them or consolidate them into one overall goal and resolve their conflicts is not stated.

Finally, value-based and value-oriented BPM both have the goal of determining processes’ and process changes’ long-term business value (Buhl et al. 2011; vom Brocke et al. 2010), substantiating the goals of goal-oriented BPM. Both approaches are also based on capital budgeting methods. While, as discussed in vom Brocke et al. (2010), value-oriented BPM uses the Visualization of Financial Implications (Grob 1993) to valuate a process, value-based BPM, as illustrated in Buhl et al. (2011), uses the certainty equivalent method (Copeland et al. 2005, p. 54). Both methods are based on cash flows and consider the time value of money. The Visualization of Financial Implications provides in-depth insights into the payment structure of a process and can be used in a detailed analysis of processes from a financial perspective. The certainty equivalent method brings decision theory, in the form of the expected utility theory (Bernoulli 1954), into capital budgeting and represents a kind of semi-subjective valuation (Kruschwitz and Löffler 2003). This valuation considers a decision maker's estimation of the utility of a financial value and allows the incorporation of the risk associated with that value as well as the risk attitude of the decision maker. Thus, while value-oriented BPM provides more detail about the payment structure, value-based BPM proposes an objective function that is "well-founded in terms of investment and decision theory" (Buhl et al. 2011, p. 170). Overall, both approaches are closely related and provide an important economic perspective to BPM, adding the well-founded, non-functional goals to goal-oriented BPM, as deemed necessary in Kueng and Kawalek (1997). As noted in vom Brocke et al. (2010), the value-oriented/value-based perspective has its limitations in that it does not necessarily consider other drivers for process improvement, such as compliance management. However, process improvement projects "in their essence present significant investments (Devaraj and Kohli 2001) to project sponsors who, ultimately, are interested in the return-on-investment from engaging in process re-design projects" (vom Brocke et al. 2010, p. 335). Hence, project sponsors are interested in the bottom line impact of their investment, thus focusing on the value-oriented/value-based perspective.

II.2.2.2 Requirements

We condense the remarks made so far regarding value-based BPM into the requirements below, which serve as our design objectives and the considerations we use to calculate the rNPV of a process; we also use the requirements when analyzing related studies in the next section:

(R1) *Control flow*: Value-based BPM relies on a process' rNPV as a process measure. To calculate the rNPV, the control flow of the process under consideration must be considered; this details how the corporate level is connected to the operational level because even a minor change in the control flow can result in a major change of the rNPV.

(R2) *Cash flows*: The rNPV is based on the cash flows at the operational level.

(R3) *Long-term perspective*: The rNPV does not consider only one period but can cope with a time horizon of several periods, incorporating a long-term perspective into value-based BPM and allowing the consideration of money's time value.

(R4) *Risk*: In value-based BPM, process risk is measured as the variance of its NPV, making it necessary to be able to calculate not only the NPV's expected value but also its variance.

II.2.2.3 Related Work

This paper contributes to the value-based BPM literature, as described in Section 2.1, by attempting to connect the corporate level with the operational level by substantiating process rNPV calculation. We now review the relevant research in the BPM field that brings a value-oriented/value-based perspective to BPM. We discuss how this work addresses the requirements for value-based BPM outlined in Section 2.2. The overview on value orientation in BPM by Buhl et al. (2011) contains relevant papers. We briefly discuss the three that best fulfill the requirements: vom Brocke et al. (2010), Linderman et al. (2005), and Bai et al. (2007). We also discuss Buhl et al. (2011) because it not only surveys the literature but also contributes to economically well-founded BPM decisions. In addition to the works included in the overview on value orientation in BPM, we add others published after the overview appeared in order to include more recent research. These works are Bolsinger et al. (2011), Sampath and Wirsing (2011), and Wynn et al. (2013).

The work that best fulfills the requirements is vom Brocke et al. (2010), previously discussed in Section 2.1. The authors choose among process alternatives in order to improve a process on the basis of the (expected) terminal value of the investment and/or the return on investment (ROI). The terminal value considers cash flows and takes a long-term perspective, fulfilling (R2) and (R3). Moreover, the determination of the terminal value considers the process' control flow. However, the example process includes only one exclusive choice and one simple merge (van der Aalst et al. 2003). How the terminal value could be calculated for more complex control flows is not explained. Hence, (R1) is only partly fulfilled. Although probabilities are included, thus considering risk to a certain extent, risk is not measured via the variance of the values, leaving (R4) unfulfilled. Overall, however, this work contributes significantly to the literature on value orientation in BPM.

Linderman et al. (2005) present a model for minimizing the expected costs of process maintenance. Although their approach considers costs and not cash flows, we regard (R2) as being partially fulfilled because this approach can be applied to cash flows as well. This work considers specific kinds of costs for a process as a whole, without considering the control flow; hence, (R1) is not fulfilled. As the authors do not determine the variance of the costs, risk is not considered, as is required in value-based BPM. Thus, (R4) is not fulfilled.

A long-term perspective is included to some extent because average long-term costs are used. However, the time value of money is not incorporated. Therefore, (R3) is met in only a limited way.

Bai et al. (2007) and its most recent version, Bai et al. (2013), present a framework for determining where within a process to include control mechanisms for mitigating risk exposure. The paper focuses on the costs of executing a process to determine the best location. As with the previous paper, (R2) is partially fulfilled because the approach could have focused on cash flows instead. They consider risk measures such as expected loss, Value-at-Risk, and Conditional Value-at-Risk to determine the "optimal control structure design model." However, the variance is not included, leaving (R4) unfulfilled. Nevertheless, the paper contributes to the consideration of risks within BPM. The risk measures are determined with the help of simulations. Thus, the control flow is considered, fulfilling (R1). A long-term perspective is not included (R3), however.

The work of Buhl et al. (2011) also contributes to the value-oriented/value-based perspective in BPM. The rNPV is introduced as a process measure within value-based BPM, meeting the requirements of (R2) and (R3). Although the work argues that the variance of a process' NPV should be considered, methods of calculation are not discussed; thus, (R4) is not fulfilled. Moreover, the paper remains on the corporate level rather than the operational process level, and control flow is thus not considered (R1).

Bolsinger et al. (2011) extend the work of Buhl et al. (2011) by providing detail about the rNPV, fulfilling (R2) and (R3). However, their paper also remains on the corporate level, without considering the operational process level, as required by (R1). Nor does the paper discuss how the variance can be determined (R4).

Sampath and Wirsing (2011) illustrate how the expected costs of a process can be determined using a process pattern based approach, which can also be applied to cash flows, partly fulfilling (R2). Since there is no consideration of costs in different periods, a long-term perspective is not included. This is also true for the calculation of the variance, which is not considered as well. Therefore, (R3) and (R4) are not fulfilled. Since the calculation of the costs is based on process patterns, the control flow of a process is considered. However, it is not stated, how to do so for a process that includes several different patterns. Nevertheless, (R1) is fulfilled to a considerable extent.

Wynn et al. (2013) incorporate the "cost perspective in the BPM Systems with the view to enable cost-aware process mining" (p. 87). This paper focuses on the reporting of costs, which could also be used for cash flows. As with previous papers, then, (R2) is partly fulfilled. The calculation of costs is confined to single process executions, without considering the long-term perspective, as required for (R3). A risk perspective is not incorporated; thus, (R4) is not fulfilled. The costs for all tasks within an execution are considered to determine the costs for a single process execution; process control flow is thus considered. However, the featured approach uses existing data about a process, which is possible only for existing processes and not for alternatives. Nevertheless, this approach fulfills (R1).

The contributions to the study of value-based BPM offered by the papers discussed above, all of which take a value-oriented/value-based perspective on BPM, are summarized in Table 1. Though the works all provide important contributions to value orientation in BPM, none fulfills every requirement. None of the works considers the operational process level, the long-term

Papers	(R1) *Control flow*	(R2) *Cash flow*	(R3) *Long-term perspective*	(R4) *Risk*
vom Brocke et al. (2010)	partly fulfilled	fulfilled	fulfilled	not fulfilled
Linderman et al. (2005)	not fulfilled	partly fulfilled	partly fulfilled	not fulfilled
Bai et al. (2007, 2013)	Fulfilled	partly fulfilled	not fulfilled	not fulfilled
Buhl et al. (2011)	not fulfilled	fulfilled	fulfilled	not fulfilled
Bolsinger et al. (2011)	not fulfilled	fulfilled	fulfilled	not fulfilled
Sampath and Wirsing (2011)	partly fulfilled	partly fulfilled	not fulfilled	not fulfilled
Wynn et al. (2013)	Fulfilled	partly fulfilled	not fulfilled	not fulfilled

Table 1. Summary of discussed papers with a value-oriented/value-based perspective

perspective, and risk together. Thus, none of the studies shows how to determine $E[NPV]$ and $Var[NPV]$ while considering processes' control flow, which is important to connect the corporate level with the operational process level (Rotaru et al. 2011; vom Brocke et al. 2010). Section 4 strives to close this gap by providing a valuation calculus for determining this expected value and process variance.

II.2.3 Illustrative Example

To provide a better understanding of the issues raised in the sections below, we briefly discuss an example of a process. We refer to this process whenever necessary to add an example in Section 4. In Section 5, we use the example process for evaluation purposes. Although the following valuation calculus is, of course, valid for more complex processes, we use this rather simple process, which nevertheless contains the five control flow patterns—XOR-split, XOR-join, AND-split, AND-join, and structured loop (van der Aalst et al. 2003)—for illustrative purposes.

Suppose there is an existing payroll process *PR* and a process alternative *PR'*, both of which are modified versions of real-world processes discussed in Neiger et al. (2006), as presented in Figure 2. The processes differ in their control flow, number of actions, and transition probabilities, which we briefly describe below. We use the term "action" for a fundamental component of a process, which "takes a set of inputs and converts them into a set of outputs" (Object Management Group 2011, p. 225), in line with the OMG Unified Modeling Language Superstructure (Object Management Group 2011).

The process *PR* has one action, "Enter Payroll run information" (a_1), with an expected cash outflow of $1,000 per execution. This action is followed by two parallel actions, "Approve Payroll run" (a_2, a_3), each of which has an expected cash outflow of $500 per execution. If

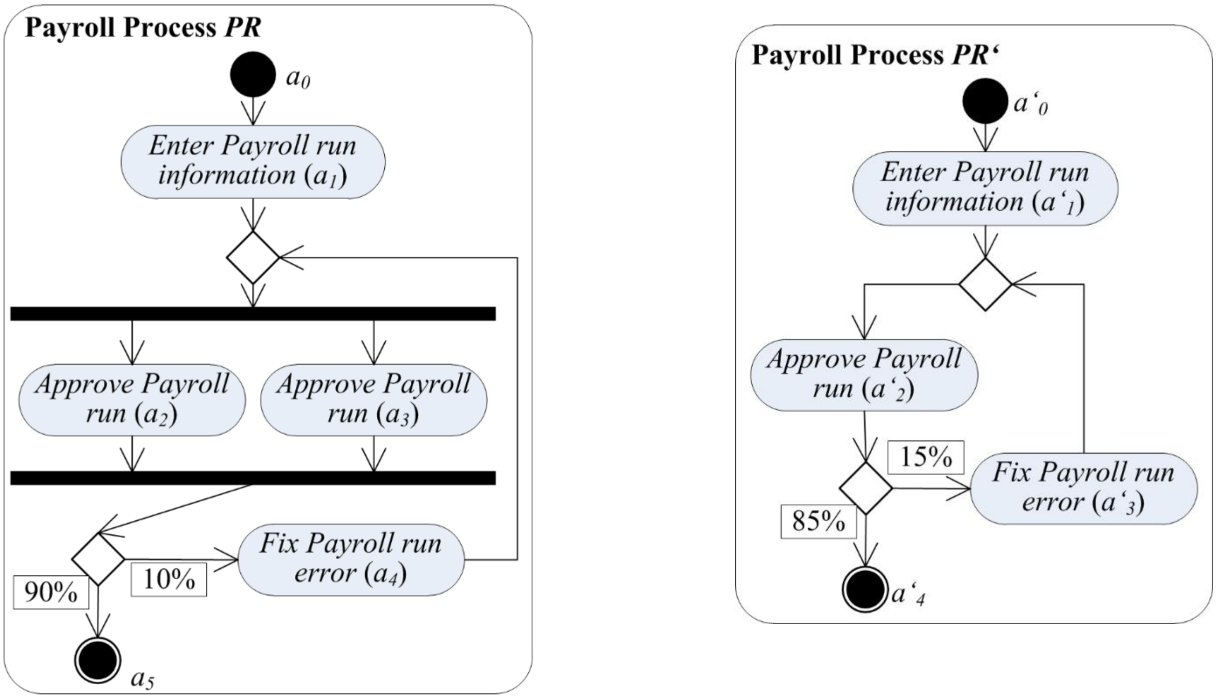


Figure 2. Existing payroll process *PR* and process alternative *PR'*

data are entered incorrectly during the execution of the first action without being discovered and corrected in either of the following two actions, the expected cash outflow to fix the error in the payroll run is \$5,000. This is done in the action “Fix Payroll run error” (a_4) and occurs with an estimated probability of 10%, which has to be approved again. Suppose that the process alternative PR’ has only one action, “Approve Payroll run” (a'_2). The action “Fix Payroll run error” (a'_3) will then occur with an estimated probability of 15%, due to the less thorough approval.

The process manager’s challenge is to determine if the existing process PR is better or worse than PR’ from a value-based BPM perspective. It is not easy just knowing the rNPV or the expected value, and particularly the variance of NPV. This is because the control flow structure of the processes needs to be considered. This structure can be very complex. Thus, the cash flows for the process’ actions need to be provided, and then the rNPV for the process as a whole can be calculated. If using a modeling tool that can calculate the rNPV, a process manager can determine if the existing process PR is better or worse than PR’ in terms of the rNPV and how much better or worse it is.

II.2.4 Valuation Calculus

To determine the rNPV, as shown in expression (2), the expected value of the uncertain net present value of a process $E[NPV]$ and its variance $Var[NPV]$ need to be calculated. This is the focus of this section, whereas other papers deal with the determination of the risk aversion constant as the third component of the rNPV (see Section 2.1). Before we show how $E[NPV]$ and $Var[NPV]$ are connected with the process cash flow in Section 4.2, we state the assumptions of our valuation calculus in Section 4.1. Finally, in Section 4.3 we go into more detail about the process cash flow, while considering the control flow of a process.

II.2.4.1 Assumptions

The execution of a process is an important part of the determination of the expected value and variance. A closer look at the “execution of a process” and a more precise definition are necessary. Every time a process is executed, a *process instance PI* is performed. The Workflow Management Coalition (WfMC) defines a process instance in Hollingsworth and WfMC (2003) as the “representation of a single enactment of a process…including its associated data. Each instance represents a separate thread of execution…of the process…which may be controlled independently and will have its own internal state and externally visible identity” (p. 269). In order to specify an “enactment of a process” more precisely, we consider the term *process*. According to Hollingsworth and WfMC (2003), a process represents a “co-ordinated (parallel and/or serial) set of [actions] that are connected in order to achieve a common goal” (p. 275). When a process is executed (enacted) the whole set of actions is not necessarily executed, but only a subset, because there can be points in the “process where, based on a decision or workflow control data, one of several branches is chosen” (van der Aalst et al. 2003, p. 11). However, although the actions are connected to achieve a common goal, the process might fail to achieve the process goal because of errors in the process execution. Thus, a rather informal definition, similar to that in Braunwarth et al. (2010), is proposed below in order to ease the communication of the approach, which is in line with design science research.

Definition 1 (*Process instance and process path*). A *process instance PI* is the execution of a certain (sub)set of actions of a *process* (coordinated set of actions). The execution of this set is intended to achieve a common goal, has its own internal state, and an externally visible identity. In case of error, the set is only partly executed, and the process reaches the end of the process. Both a set of actions that achieves the process goal as well as the partly executed set form a path through the process, from start to end, called *process path pp*.

A process path is not necessarily a sequence of actions. It can include actions that are executed in parallel or executed more than once. Due to the structured loop in both processes seen in Figure 2, an infinite number of process paths is possible, although there is a finite number of actions, as, for example, in the left process in Figure 2, with one process path consisting of the actions a_1, a_2, and a_3 (case: no fixing is needed), another path of the actions a_1, a_2, a_3, a_4, a_2, and a_3 (case: fixing is needed once), a third path with the actions a_1, a_2, a_3, a_4, a_2, a_3, a_4, a_2, and a_3 (case: fixing is needed twice), and so on. The number of different coordinated sets of actions is the number of process paths, which can be an infinite number, as is in the example in Figure 2. However, infinite numbers of process paths are uncommon. In reality, the probabilities at an exclusive choice would likely be very different every time a process instance reaches the same exclusive choice. In the example from Figure 2, with process *PR* it can be 90% and 10% the first time the exclusive choice is reached, 99% and 1% the second time, and 100% and 0% the third time. This eases the calculation because it results in a finite number of process paths while being closer to reality. This consideration about changing probabilities is possible with the expressions used below but is, to the best of our knowledge, not possible with any current process-modeling tool. A process instance executes exactly one possible process path. From this, we can make an assumption about how the considered processes are to be structured:

(A1) *A process P consists of a set A of actions* $a_d \in A$, *d = 1,…, D, one starting point* a_0, *one final point* a_{D+1}, *transitions between the actions, and routing constructs (van der Aalst et al. 2003). A process instance PI starts in* a_0 *and ends in* a_{D+1}. *The probability that a process instance follows a process path* pp_k *is denoted by* p_k, *called "path probability." Each path probability can be determined and is fixed. No logical error in the process can prevent a process instance from reaching* a_{D+1}. *The probability of an action's execution failure is known.*

Within a process (model), identical tasks may be done more than once. For example, in the left process in Figure 2, "Approve Payroll run" is done twice, but we label one of them a_2 and the other a_3. We consider everything modeled within a process as a different action, even if the same task is done, thus considering each to be a different action. This allows us to label all the tasks in a process differently in order to consider all of them separately in the valuation calculus. Action a_0 designates the (fictitious) point where the process starts, and a_{D+1} designates the (fictitious) point towards which a process instance proceeds and at which it always ends. The path probability p_k can be determined and is fixed (for more details on the determination of path probabilities, see appendix A). If no process action fails its execution, then every possible process instance starts in a_0 and ends in a_{D+1}. Hence, it is assumed that the process is correct and sound (van der Aalst et al. 2011). The execution of an action may fail with a known probability. Such failure of an action a_d can be modeled as an exclusive choice before a_d, with one choice going to a_{D+1}, which is taken with the probability that a_d fails, and a choice to continue the process, which is taken with the probability that a_d does not fail. Such explicit modeling of action failure would result in a new process path, to which a probability can be assigned. Thus, it is assumed that all known errors are modeled as described.

The cash flow of a process is caused by its actions. Thus, the cash flow of each action is important. Each action's cash flow is caused by different action characteristics (e.g., wages, material). These characteristics result in different cash flows (e.g., cash outflow for wages, cash outflow for material; cp. vom Brocke et al. [2010]). In reality, the cash flow of an action might be different with each process instance. Hence, the cash flow of an action a_d is uncertain and thus modeled as a random variable CF_{a_d}. In addition to the cash flows caused by actions, some cash flows are caused each time a process is executed, independent of the executed actions (e.g., cash outflows for overheads, cash inflows resulting from purchase transactions, cash outflows for process maintenance). These are cash flows of the characteristics of a whole process, called *process attributes*. These process attribute cash flows must be combined with the cash flows of actions to determine the cash flow of a process.

(A2) *The random variables CF_{a_d} represent the uncertain cash flows of the actions. The random variables CF_{pa_s}, s = 1,…, S, represent the uncertain cash flows of process attributes, which are cash flows that are relevant for a process as a whole for every process instance. The expected values $E[CF_{a_d}]$ and $E[CF_{pa_s}]$ as well as the variances $Var[CF_{a_d}]$ and $Var[CF_{pa_s}]$ are finite and known.*

The expected value and variance of the cash flows of actions and of process attributes must be determined. Direct cash flows can be easily assigned to an action or process attribute. In terms of indirect cash outflows, Action-Based Costing can be used, as stated in Gulledge et al. (1997). This is also possible when accounting is linked with process-aware information systems (vom Brocke et al. 2011b). For cash inflows, the price of a product or service can be used and assigned to the process. Another possible method of determining the expected values and variances is to identify and use the subjective probability distributions of the cash flows. Suggestions on how to determine these distributions and elicit the necessary data from individuals can be found in Hubbard (2007).

Every planning horizon period contains several process instances, resulting in many process cash flows CF_P. Concerning the process instances, we assume the following:

(A3) *There are no dependencies between process instances.*

The process instances of a process are independent of the process instances of other processes; there is a high degree of autonomy (Feiler and Humphrey 1993). This is in line with Davamanirajan et al. (2006) because we concentrate on one process only. Moreover, process instances are independent of the process instances of the same process, as assumed in Bolsinger et al. (2011). In fact, a more general version of the valuation calculus is able to deal with dependencies through correlation coefficients. However, in order to prevent the presentation becoming overly complex, we assume independent process instances here.

II.2.4.2 Corporate Level

While the managers at the corporate level are interested in the rNPV, this value is based on the cash flows at the operational process level. Thus, the following expressions show how $E[NPV]$ and $Var[NPV]$ are connected with the process cash flow. With expression (1), it follows as expressed below:

$$E[NPV] = -I + \sum_{t=0}^{T} \frac{\sum_{j=1}^{n_t} E\left[CF_{P_j}\right]}{(1+i)^t} = -I + \sum_{t=0}^{T} \frac{n_t \cdot E\left[CF_{P_j}\right]}{(1+i)^t}. \tag{3}$$

It is $\sum_{j=1}^{n_t} E\left[CF_{P_j}\right] = n_t \cdot E\left[CF_{P_j}\right]$, because CF_{P_j} are identically distributed (Bolsinger et al. 2011). In combination with (A3), the random variables CF_{P_j} are independent and identically distributed (iid).

Then, it follows for $Var[NPV]$ that

$$Var[NPV] \overset{(A3)}{\triangleq} \sum_{t=0}^{T} \frac{\sum_{j=1}^{n_t} Var\left[CF_{P_j}\right]}{(1+i)^{2t}} \overset{iid}{\triangleq} \sum_{t=0}^{T} \frac{n_t \cdot Var\left[CF_{P_j}\right]}{(1+i)^{2t}}. \tag{4}$$

Hence, the corporate level puts the focus on the expected value of the process cash flow $E[CF_P]$ and its variance $Var[CF_P]$. In the following section, we show how $E[CF_P]$ and $Var[CF_P]$ are calculated including a consideration of the operational process level.

II.2.4.3 Operational Process Level

When a process instance “reaches” a routing construct upon which the process can “continue” in different ways (e.g., after an exclusive choice), the process instance “continues” depending on which condition(s) hold (e.g., depending on process inputs, on the environmental state). Thus, a process consists of multiple process paths, each executed with a certain probability. Every process path describes a possibility of executing a process from start to finish, which is why each process instance may result in a different cash flow depending on the control flow. This demonstrates the importance of process paths in considerations of processes as a whole. Thus, the expected value and variance of the cash flow of a single process path are first determined before the expected value and variance of the process as a whole are calculated.

II.2.4.3.1. Process Path

A process path pp_k contains actions from the start to the end of a process (see definition 1). Each process path is assigned a natural number k to make it formally distinct. The actions of a process path pp_k plus a_0 and a_{D+1} form (in a first step) an action multiset AS_k, whose elements are out of $A \cup \{a_0, a_{D+1}\}$. It is important that it be a multiset, so that loops can be considered, as the same actions can occur several times. Each action a_d in AS_k that occurs more than once (in a second step) is given an index $n \in \mathbb{N}$ in the form $a_d^{(1)}, a_d^{(2)}, \ldots, a_d^{(n)}, \ldots$. The index indicates the number of the loop iteration to which the action is assigned in order to distinguish among the actions, each of which is from different iterations, with different probabilities of being executed. In the process seen on the left in Figure 2, there are the action sets

$AS_1 = \{a_0, a_1, a_2^{(1)}, a_3^{(1)}, a_5\}$,

$AS_2 = \{a_0, a_1, a_2^{(1)}, a_3^{(1)}, a_4^{(1)}, a_2^{(2)}, a_3^{(2)}, a_5\}$,

$AS_3 = \{a_0, a_1, a_2^{(1)}, a_3^{(1)}, a_4^{(1)}, a_2^{(2)}, a_3^{(2)}, a_4^{(2)}, a_2^{(3)}, a_3^{(3)}, a_5\}$, and so on.

The path probabilities are $p_1 = 0.9$, $p_2 = 0.1 \cdot 0.9 = 0.09$, $p_3 = 0.1^2 \cdot 0.9 = 0.009$ (for more details, see appendix A). Given that exactly one process path is taken if a process is executed and that they are mutually exclusive, the probabilities p_k sum up to 1. A process path has only sequential and parallel actions. Thus, the actions of a process path could be transformed into a sequential order without changing the result of the process path or the cash flow CF_{pp_k} of a process path pp_k. In addition to the cash flows of the actions, there are also the cash flows of process attributes CF_{pa_s}, which are considered with every execution of a process. Hence, it is

$$CF_{pp_k} = \sum_{a_d \in AS_k} CF_{a_d} + \sum_{s=1}^{S} CF_{pa_s}\,. \tag{5}$$

The expected value of CF_{pp_k} is

$$E\left[CF_{pp_k}\right] = \sum_{a_d \in AS_k} E\left[CF_{a_d}\right] + \sum_{s=1}^{S} E\left[CF_{pa_s}\right] \tag{6}$$

and the variance of CF_{pp_k} is

$$Var\left[CF_{pp_k}\right] = \sum_{a_d \in AS_k} Var\left[CF_{a_d}\right] + \sum_{s=1}^{S} Var\left[CF_{pa_s}\right]$$
$$+ \sum_{\substack{a_d, a_j \in AS_k \\ d \neq j}} \rho_{a_d,a_j} \cdot \sigma_{a_d} \cdot \sigma_{a_j} + 2 \sum_{a_d \in AS_k} \sum_{s=1}^{S} \rho_{a_d,pa_s} \cdot \sigma_{a_d} \cdot \sigma_{pa_s} \qquad (7)$$
$$+ 2 \sum_{s=1}^{S-1} \sum_{j=s+1}^{S} \rho_{pa_s,pa_j} \cdot \sigma_{pa_s} \cdot \sigma_{pa_j} \,,$$

where $Var\left[CF_{a_d}\right] = \sigma_{a_d}^2$ and $Var\left[CF_{pa_s}\right] = \sigma_{pa_s}^2$. The correlations ρ_{a_d,a_j}, ρ_{a_d,pa_s}, and ρ_{pa_s,pa_j} may reflect dependencies between the actions and process attributes. In Figure 2, it is possible that the lower the cash outflow of a_1, (because the payroll run information is entered very quickly), the higher the cash outflow of a_2 and a_3, (because they must make more corrections during the approval since the information was entered quickly and less carefully). Such dependencies could be reflected with correlations ρ_{a_1,a_2} and ρ_{a_1,a_3}. If there are no dependencies, all correlations ρ_{a_d,a_j}, ρ_{a_d,pa_s}, and ρ_{pa_s,pa_j} are zero, (7) simplifies to

$$Var\left[CF_{pp_k}\right] = \sum_{a_d \in AS_k} Var\left[CF_{a_d}\right] + \sum_{s=1}^{S} Var\left[CF_{pa_s}\right]. \qquad (8)$$

This determines the expected value and variance of the cash flow of one process path. The next step extends this to a process, where we need to consider all process paths at once in order to consider the control flow. Therefore, to determine the expected value and variance of a cash flow of a process, we must take into account the control flow of a process. A process may be not only a sequence of actions (as possible in a process path) but may also contain control flow patterns, like exclusive choice, simple merge, parallel split, synchronization, and loops (van der Aalst et al. 2003). Due to loops, each process can have an infinite number of process paths, which need to be considered using the valuation calculus below.

II.2.4.3.2. Process

To consider all process paths at once, a process is modeled as a probability space, which is a "triple $(\Omega, \mathcal{F}, \mathbf{P})$ of a sample space Ω, a [sigma]-algebra $\mathcal{F}$ of sets in it, and a probability measure **P** on $\mathcal{F}$" (Feller 1971, p. 116).[1] This is a stochastic model that provides the formalism necessary for determining the expected value and variance of process cash flows. The sample space Ω is the set of all possibilities that the object under consideration can take; it is thus the set of all possible process paths. A sigma-algebra $\mathcal{F}$ is a family of sets over Ω (a set of sets), and a set in $\mathcal{F}$ is called "event" (Feller 1971, p. 112). The probability measure **P** assigns a certain probability to each event (Feller 1971, p. 115), thus to each set of process paths. In definition 2, a process is modeled as a probability space:

Definition 2 (*Process-probability-space*). A process P is a probability space $(\Omega, \mathcal{F}, \mathbf{P}_P)$ consisting of:

- the sample space $\Omega = \{pp_k \mid k \in \mathbb{N}\}$, which is the set of all possible process paths of a process P,
- the sigma-algebra $\mathcal{F} = 2^{\Omega}$, which is the power set of Ω and therefore a set of subsets of Ω, which are the events of this probability space, and
- the probability measure

  $$\mathbf{P}_P\{PP\} = \sum_{pp_k \in PP} f_{PI}(k) = \sum_{pp_k \in PP} p_k \qquad \text{for all } PP \subseteq \Omega\,,$$

  with the probability mass function

  $$f_{PI}(k) = Prob(PI = k) = p_k\,,$$

  where the process instance *PI* is a random variable

  $$PI(\omega) = \begin{cases} 1 & \text{if } \omega = pp_1 \\ \dots & \dots \\ k & \text{if } \omega = pp_k \\ \dots & \dots \\ |\Omega| & \text{if } \omega = pp_{|\Omega|} \end{cases}\,,$$

  which takes on the value *k* for the *k*th process path with probability p_k.

In definition 2, a process is formally described as a probability space. In appendix B, it is formally shown that this process-probability-space is indeed a probability space. Definition 2 presents a process as a stochastic model and displays the formal differences and interplay among a process, a process instance, and a process path. As when modeling a process with UML activity diagrams, for example, a process model defines the process as a whole and does not change when a process is executed. The process paths are also fixed by the process

[1] The text is italicized in the source. The symbol $\mathfrak{A}$ for the sigma-algebra and the symbol $\mathfrak{S}$ for the text's sample space were replaced by the now more commonly used symbols $\mathcal{F}$ and Ω, respectively.

model, which are fixed in the process-probability-space as well. As in every process, the process instance is the random component. Before executing a process, it is unknown which process path will be executed by a process instance; it could be any of them. In the process-probability-space, this randomness is represented by the random variable PI, which takes a certain process path pp_k with a certain probability p_k. Thus, in definition 2, it is possible to see a process, a process instance, and a process path explicitly within one model. If a process contains loops, an infinite number of process paths are possible. This is accounted for in definition 2 via the possibly infinite sample space. According to definition 2, the expected value of the cash flow of a process path pp_k is more precisely

$$E\left[CF_{pp_k}\right] = E[CF_P \mid PI = k]\,. \tag{9}$$

Expression (9) shows that the expected value of the cash flow of the process path pp_k is equal to the expected value of the cash flow of a process P given that process path pp_k is executed.

Now the expected value $E[CF_P]$ and the variance $Var[CF_P]$ of the cash flow of a process P can be determined. We want to express the expected value and variance only with the information about the actions and the additional process attributes.

In order to determine $E[CF_P]$ and $Var[CF_P]$, let $Pr(a_d)$ be the probability that an action $a_d \in AS$, with $AS \coloneqq \bigcup_{k=1}^{|\Omega|} AS_k$, is executed when executing a process with

$$Pr(a_d) := \mathbf{P}_P\{PP_{a_d}\} = \sum_{pp_k \in PP_{a_d}} p_k \tag{10}$$

where PP_{a_d} is the set of process paths that contain the action a_d:

$$PP_{a_d} = \{pp_k \in \Omega \mid a_d \in AS_k\}\,. \tag{11}$$

It is $|\Omega|$ the number of process paths, which can be set to infinity for a process with loops. Expression (10), in combination with expression (11), shows that the probability that an action a_d is executed is the sum of the path probabilities p_k assigned to the process paths pp_k that contain action a_d. The expected value $E[CF_P]$ can be determined as follows, where (12) corresponds to the determination of expected costs in Linderman et al. (2005); for details, see appendix C:

$$E[CF_P] = \sum_{k=1}^{|\Omega|} E[CF_P \mid PI = k] \cdot Prob(PI = k) \tag{12}$$

$$= \sum_{k=1}^{|\Omega|} E\left[CF_{pp_k}\right] \cdot p_k \tag{13}$$

$$= \sum_{a_d \in AS} E\left[CF_{a_d}\right] \cdot Pr(a_d) + \sum_{s=1}^{S} E\left[CF_{pa_s}\right]. \tag{14}$$

The variance $Var[CF_P]$ can be similarly determined. Let $Pr\left(a_d, a_j\right)$ be the probability that both actions $a_d \in AS$ and $a_j \in AS$ are executed when executing a process with

$$Pr\left(a_d, a_j\right) := \mathbf{P}_P\left\{PP_{a_d,a_j}\right\} = \sum_{pp_k \in PP_{a_d,a_j}} p_k \tag{15}$$

where PP_{a_d,a_j} is the set of process paths that contain both actions a_d and a_j:

$$PP_{a_d,a_j} = \left\{pp_k \in \Omega \mid a_d \in AS_k, a_j \in AS_k, a_d \neq a_j\right\}. \tag{16}$$

Expression (15), in combination with expression (16), shows that the probability that both actions a_d and a_j are executed is the sum of the path probabilities p_k assigned to the process paths pp_k that contain both actions a_d and a_j.

The variance of the cash flow CF_P of a process P is (for details, see appendix D):

$$Var[CF_P] = \sum_{k=1}^{|\Omega|} E[(CF_P - E[CF_P])^2 \mid PI = k] \cdot Prob(PI = k) \tag{17}$$

$$= \sum_{k=1}^{|\Omega|} E\left[\left(CF_{pp_k} - E[CF_P]\right)^2\right] \cdot p_k \tag{18}$$

$$= -E[CF_P]^2 + \sum_{k=1}^{|\Omega|} \left(Var\left[CF_{pp_k}\right] + E\left[CF_{pp_k}\right]^2\right) \cdot p_k \tag{19}$$

$$\begin{aligned}
&= -E[CF_P]^2 + \sum_{a_d \in AS} \left(Var\left[CF_{a_d}\right] + E\left[CF_{a_d}\right]^2\right) \cdot Pr(a_d) \\
&\quad + \sum_{s=1}^{S} \left(Var\left[CF_{pa_s}\right] + E\left[CF_{pa_s}\right]^2\right) \\
&\quad + \sum_{a_d, a_j \in AS, a_d \neq a_j} \left(\rho_{a_d,a_j} \cdot \sigma_{a_d} \cdot \sigma_{a_j} + E\left[CF_{a_d}\right] E\left[CF_{a_j}\right]\right) \cdot Pr\left(a_d, a_j\right) \\
&\quad + 2 \sum_{a_d \in AS} \sum_{s=1}^{S} \left(\rho_{a_d,pa_s} \cdot \sigma_{a_d} \cdot \sigma_{pa_s} + E\left[CF_{a_d}\right] E\left[CF_{pa_s}\right]\right) \cdot Pr(a_d) \\
&\quad + 2 \sum_{s=1}^{S-1} \sum_{j=s+1}^{S} \left(\rho_{pa_s,pa_j} \cdot \sigma_{pa_s} \cdot \sigma_{pa_j} + E\left[CF_{pa_s}\right] E\left[CF_{pa_j}\right]\right).
\end{aligned} \tag{20}$$

If there are no dependencies (i.e., if all correlations are zero) and no process attributes are considered—if, for example, it is the same for different process alternatives—(20) simplifies to

$$\begin{aligned}
Var[CF_P] = -E[CF_P]^2 &+ \sum_{a_d \in AS} \left(Var\left[CF_{a_d}\right] + E\left[CF_{a_d}\right]^2\right) \cdot Pr(a_d) \\
&+ \sum_{a_d, a_j \in AS, a_d \neq a_j} E\left[CF_{a_d}\right] E\left[CF_{a_j}\right] \cdot Pr\left(a_d, a_j\right).
\end{aligned} \tag{21}$$

As expression (18) shows, the variance is the weighted average of the expected values of the squared difference between the cash flow of a certain process path and the expected value of the cash flow of the process. Although it might seem intuitive at first glance, it is not $CF_P = \sum_{k=1}^{|\Omega|} CF_{pp_k} \cdot p_k$.

Overall, with expression (14) and (20) in combination with expression (3) and (4), it is possible to determine $E[NPV]$ and $Var[NPV]$, which can then be used to calculate the rNPV with expression (2).

II.2.5 Evaluation

The evaluation of an artifact is an important step in design-oriented research, and various methods are available (Hevner et al. 2004; Peffers et al. 2008). Determining the utility of an artifact would be best achieved through a process-modeling tool that incorporates the valuation calculus and is used in a naturalistic setting with real users and real problems. However, this would be very time-consuming and resource-intensive. The evaluation framework for design science research presented in Venable et al. (2012) suggests performing the evaluation in an artificial setting. Sonnenberg and vom Brocke (2012) describe three evaluation activities (EVAL 1, EVAL 2, and EVAL 3) for such artificial settings. Each activity justifies a self-contained research contribution. We carry out all three activities to evaluate the artifact under study as follows:

EVAL 1: This activity is performed to justify the problem statement, research gap, and design objectives. This activity is conducted in sections 1 and 2.

EVAL 2: This activity validates the design specification and justifies the design tool/ methodology. While Section 4 provides mathematical proofs and logical reasoning (formal deduction), valid evaluation methods for this activity, Section 5.1 shows the results of a feature comparison to illustrate the extent to which the stated design objectives of Section 2.2 are met. Section 5.2 "show[s] analytically that [the] artifact behaves as intended for a single test case" (Sonnenberg and vom Brocke 2012, p. 395) in order to demonstrate its feasibility. We therefore rely on the example introduced in Section 3.

EVAL 3: This activity validates an instance of the artifact in an artificial setting to prove its applicability. This is done in Section 5.3 by demonstrating how the artifact helped correct the calculation logic of the commercial process-modeling tool of the CubeFour company.

In addition to these three activities, in Section 5.4, we conduct a discussion regarding a competing artifact by comparing the valuation calculus with process simulations.

II.2.5.1 Feature Comparison

Section 2.2 outlines the four requirements (design objectives) for determining the rNPV. To verify if this paper contributes meaningfully to BPM research, we compare the valuation calculus with these requirements.

(R1) *Control flow*: The valuation calculus is based on path probabilities (see appendix A for details); it is thus based on the path that a process instance takes from the start to the end of a process. For each process that fulfills assumption (A1)—if the process is correct and sound and if its possible failures are known—all process paths can be determined. Process paths define how a process instance can reach the end of the process. Since process instances consider the control flow of a process and as process paths define the way of a process instance from start to finish, we consider the control flow of a process by using process paths for the valuation calculus. Although assumption (A1) is rather general, unknown failures (which exist when no process path considers them) are not considered in the valuation calculus. In any case, known or expected failures are considered.

(R2) *Cash flows*: The valuation calculus is designed to work for additive quantities, as shown by expression (5). Since the cash flows of the actions can be added to determine the rNPV, cash flows are considered in the described valuation calculus.

(R3) *Long-term perspective*: The calculation of the rNPV is based on the NPV presented in expression (1). The NPV considers the cash flows of future periods and the time value of money, incorporating a long-term perspective into value-based BPM.

(R4) *Risk*: To consider risk in value-based BPM, we must be able to measure it. Section 4.3 describes how the variance of *NPV* can be determined, which is used to measure risk.

Overall, while requirements (R2), (R3), and (R4) are fulfilled straightforwardly, some minor limitations regarding the control flow exist, as stated above (R1). However, assuming that we only consider correct and sound processes is feasible. Thus, we reduce the research gap considerably.

II.2.5.2 Illustrative Example (continued)

Let us again consider the payroll process *PR* introduced in Section 3 to demonstrate the feasibility of the valuation calculus. As illustrated in Section 4, determining the expected value and variance of the cash flow of a process is particularly challenging. We thus focus on this calculation. We first calculate the probability of each action (for detailed results see appendix E) based on expression (10). With expression (14), we then calculate the expected value:

$$E[CF_{PR}] = E\left[CF_{a_1}\right] + E\left[CF_{a_2}\right] \cdot \sum_{i=0}^{\infty} 0.1^i + E\left[CF_{a_3}\right] \cdot \sum_{i=0}^{\infty} 0.1^i + E\left[CF_{a_4}\right] \cdot 0.1 \cdot \sum_{i=0}^{\infty} 0.1^i$$
$$= 1{,}000 + 500 \cdot \frac{10}{9} + 500 \cdot \frac{10}{9} + 5{,}000 \cdot 0.1 \cdot \frac{10}{9} = 2{,}666.67.$$

For the variance of CF_{PR}, we first calculate the probability $Pr\left(a_d, a_j\right)$ with expression (15) before determining the variance of CF_{PR} with expression (21); we do not consider any dependencies (for detailed results, see appendix F):

$$Var[CF_{PR}] = -(E[CF_{PR}]^2) + (0 + 1{,}000^2) \cdot 1 + (0 + 500^2) \cdot \frac{10}{9} + (0 + 500^2) \cdot \frac{10}{9}$$

$$+ (0 + 5{,}000^2) \cdot 0.1 \cdot \frac{10}{9} + 2 \cdot \Big[500 \cdot 1{,}000 \cdot \frac{10}{9} + 500 \cdot 500 \cdot \frac{10}{81}$$

$$+500 \cdot 1{,}000 \cdot \frac{10}{9} + 500 \cdot 500 \cdot \frac{110}{81} + 500 \cdot 500 \cdot \frac{10}{81}$$

$$+5{,}000 \cdot 1{,}000 \cdot \frac{1}{9} + 5{,}000 \cdot 500 \cdot \frac{20}{81} + 5{,}000 \cdot 500 \cdot \frac{20}{81} + 5{,}000 \cdot 5{,}000 \cdot \frac{1}{81}\Big]$$

$$= 2{,}108.19^2.$$

In our example, there are no cash flows for the process as a whole CF_{pa_s} (S = 0). Thus, the sums in expression (20) that include CF_{pa_s} are zero. As a result, the payroll process PR has an expected cash outflow of 2,666.67, with a variance of $2{,}108.19^2$. These numbers can also be calculated for the process alternative PR' in order to enable a comparison between process alternatives. The payroll process alternative PR' has an expected cash outflow of 2,470.59 with a variance of $2{,}506.05^2$. In this case PR has a higher expected cash outflow than PR', though the variance is lower, indicating a lower risk. We thus cannot decide if PR' improves PR. However, if we assume further parameters with expressions (3) and (4), we can calculate $E[NPV]$ and $Var[NPV]$. In a last step, we can incorporate these values with expression (2), which results in one value for PR and one value for PR' for comparison.

Although, as mentioned, there is likely not an infinite number of process paths, this example shows that it is possible to consider such a case.

II.2.5.3 The Case of CubeFour

The following is a case presentation describing how the insights in Section 4 helped correct the calculation capabilities of the "cube4process" process-modeling tool used by CubeFour. Although the capabilities of cube4process were already more advanced than those of most other tools, we were able to help improve these capabilities using our valuation calculus.

Cube4process enables its users to not only model processes but also add financial information, such as the (expected) cash flow of an action's execution. This information can be added to every action. The probabilities of each transition within the process model can be added as well, which can then be used to determine the path probabilities (see appendix A for details). With this information, the tool provides the expected cash flow of the process analytically. The tool supports the basic control flow patterns XOR-split, XOR-join, AND-split, and AND-join (van der Aalst et al. 2003) as well as loops (with some minor exceptions). The tool is also intended to support OR-splits. However, after reviewing the tool based on the mathematical insights in this paper, it was discovered that OR-splits, in particular, add extra complexity to the determination of the expected cash flow, as described below.

Consider the process seen in Figure 3. After an OR-split, the process continues with, depending on the transition conditions, only one transition, any combination of two transitions, or even all three transitions. The transitions are not mutually exclusive, as with a XOR-split. This is why the transition probabilities in Figure 3 do not add up to 1. Thus, in 60% of the process instances, action *B* is executed after action *A*. Action *C* is executed after *A* in 50% of the process instances, and *D* after *A* in 10% of the process instances.

Using cube4process, the process in Figure 3 is modeled as presented in Figure 4 (Task 1:= action *A*, Task 2:= action *B*; Task 3:= action *C*, Task 4:= action *D*, and Task 5:= action *E*). Below each action, one can see the additional information regarding the cash flows of the action's execution. The first number gives the minimal cash flow of an execution, the second number is the average cash flow, and the third number is the maximal cash flow. The fourth number is the minimal cash flow of the whole process from the start until after the execution of the action. The fifth number is the corresponding average cash flow, and the sixth number is the maximal cash flow.

The information regarding the transition probabilities is important for reaching the correct determination of the expected value because this will determine the probability that an

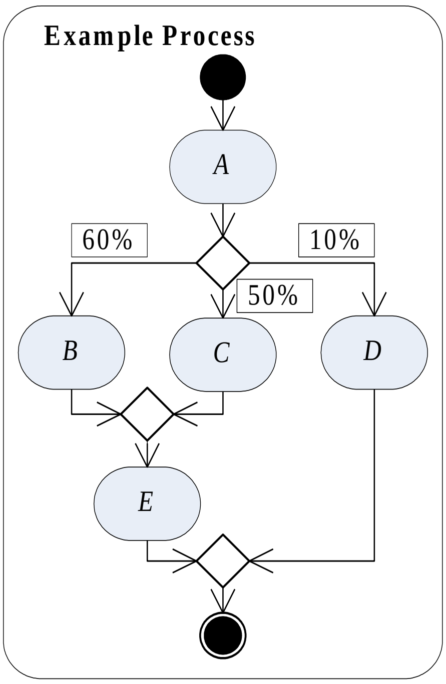


Figure 3. Example process with OR-split

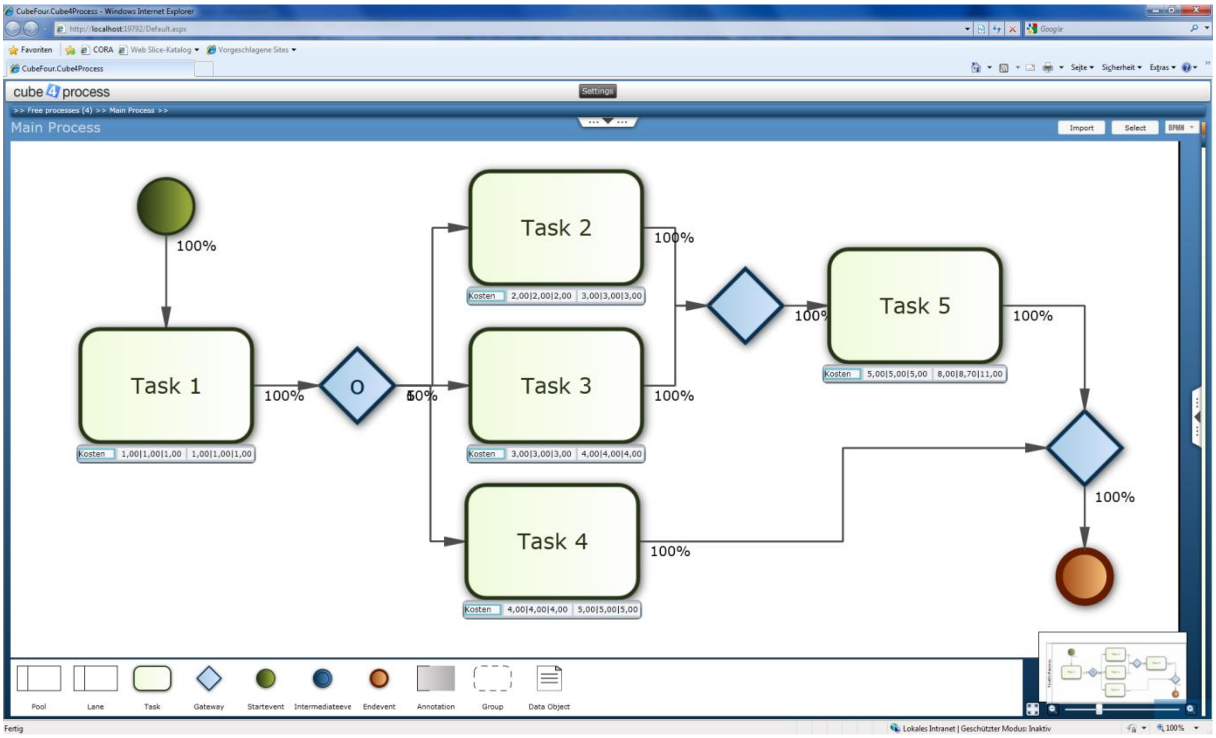


Figure 4. The process in Figure 3 modeled with cube4process

action will be executed when the process is executed. It is easy to see that $Pr(A) = 1, Pr(B) = 0.6, Pr(C) = 0.5$, and $Pr(D) = 0.1$. However, what is the probability that action E will be executed? Figure 5 provides an overview of the determination of cube4process about the probabilities and the expected value before the correction through the mathematical insights by this paper. The probability that E will be executed is given as 0.8 and the expected value as 8.1. CubeFour used the addition law of probability for this calculation. The tool made the following calculation:

$$Pr(E) = Pr(B) + Pr(C) - Pr(B) \cdot Pr(C) = 0.6 + 0.5 - 0.6 \cdot 0.5 = 0.8.$$

Here, it is implicit that each action is an event; thus, the calculation is based on a probability space whose sample space Ω is the set of all the actions of a process. However, it can be shown that a process cannot be modeled as a probability space based on actions as the events. Hence, as the calculation is not based on a valid probability space, it cannot be guaranteed to provide correct results. This holds true for all control flows and can best be illustrated by a process that contains an OR-split, which is why this construct is the focus of this section.

After considering the valuation calculus of this paper, all calculations, if implemented correctly, will lead to valid results because, in definition 2, this paper provides a valid

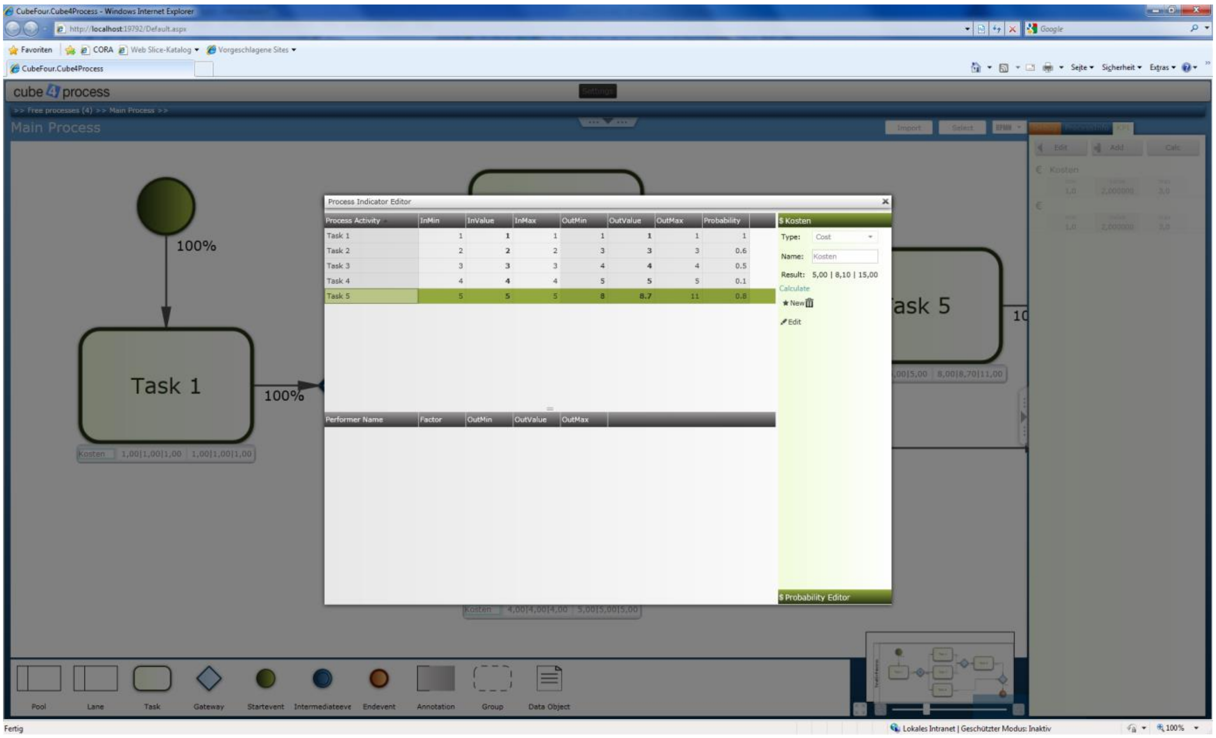


Figure 5. Results of the analytical determination of the probabilities and the expected value

probability space that provides the foundation for a correct calculation of the probability. Here, the probability that an action will be executed is given by expression (10) with

$$Pr(a_d) = \sum_{pp_k \in PP_{a_d}} p_k \ .$$

The probability that action a_d will be executed is the sum of the path probabilities of the process paths in which the action takes part. However, as we briefly illustrate below, it is impossible to calculate the probability that action E will be executed with the given information using this valid method. The given transition probabilities 60%, 50%, and 10% do not give enough information to enable a determination of the path probabilities and thus the probability that an action will be executed. This is because, for example, the 60% indicates only that action B is executed in 60% of the process instances, but does not indicate in how many of these process instances action C or D is also executed, information necessary for determining the probability of each process path. The problem is illustrated by the two examples of path probabilities in Table 2 and Table 3. First, let us assume that the path probabilities are given according to the values in Table 2.

process path	pp_1	pp_2	pp_3	pp_4	pp_5	pp_6	pp_7
actions	*A,B,E*	*A,C,E*	*A,D*	*A,B,C,E*	*A,B,D,E*	*A,C,D,E*	*A,B,C,D,E*
path probability p_k	0.43	0.37	0.02	0.1	0.05	0.01	0.02

Table 2. Actions and path probabilities of all process paths

Then it is:

$$Pr(B) = p_1 + p_4 + p_5 + p_7 = 0.43 + 0.1 + 0.05 + 0.02 = 0.6,$$

$$Pr(C) = p_2 + p_4 + p_6 + p_7 = 0.37 + 0.1 + 0.01 + 0.02 = 0.5,$$

$$Pr(D) = p_3 + p_5 + p_6 + p_7 = 0.02 + 0.05 + 0.01 + 0.02 = 0.1,$$

$$Pr(E) = p_1 + p_2 + p_4 + p_5 + p_6 + p_7 = 0.43 + 0.37 + 0.1 + 0.05 + 0.01 + 0.02 = 0.98.$$

Let us assume that the path probabilities would be slightly different according to the values in Table 3.

process path	pp_1	pp_2	pp_3	pp_4	pp_5	pp_6	pp_7
actions	*A,B,E*	*A,C,E*	*A,D*	*A,B,C,E*	*A,B,D,E*	*A,C,D,E*	*A,B,C,D,E*
path probability p_k	0.43	0.36	0.03	0.11	0.04	0.01	0.02

Table 3. Actions and slightly changed path probabilities of all process paths

Then it still is

$$Pr(B) = p_1 + p_4 + p_5 + p_7 = 0.43 + 0.11 + 0.04 + 0.02 = 0.6,$$

$$Pr(C) = p_2 + p_4 + p_6 + p_7 = 0.36 + 0.11 + 0.01 + 0.02 = 0.5,$$

$$Pr(D) = p_3 + p_5 + p_6 + p_7 = 0.03 + 0.04 + 0.01 + 0.02 = 0.1.$$

However, it is

$$Pr(E) = p_1 + p_2 + p_4 + p_5 + p_6 + p_7 = 0.43 + 0.36 + 0.11 + 0.04 + 0.01 + 0.02 \\ = 0.97.$$

Thus, although we do not change the information provided in Figure 3 because the probabilities of action *B*, *C*, and *D* do not change, the probability of action *E* changes, which also changes the expected value of the process. Therefore, the transition probabilities seem insufficient for considering OR-splits during the calculation of the expected value. Additional information about the probability for the combination of the actions after an OR-split is required.

The developers of cube4process were given an insight into the mathematical foundation of processes. As a result, CubeFour was able to correct the calculation of their tool, providing a mathematically sound calculation of the expected value and creating a valuable asset for use in process improvement projects.

II.2.5.4 Comparison with Process Simulations

Section 4 describes the focus placed on the expected value of a process cash flow and its variance because these are central to the determination of the expected value and variance of a process' NPV. In Section 4.3, we show how they can be calculated. However, they could also be determined via process simulations, which are thus a competing artifact. In Table 4, we therefore compare our valuation calculus with process simulations to determine the expected value of a process cash flow and its variance using the criteria we consider most distinctive.

	Process simulation (PS)	Valuation Calculus
Expres-siveness	A PS can explicitly consider various factors such as time, costs, and resource restrictions. However, if the PS aims to determine a monetary value for a process, then the question arises how factors like resource restrictions are transformed into monetary values.	The presented valuation calculus takes on a value-oriented/value-based perspective. Thus, factors like time and resource restrictions have to be transformed into cash flows to be considered. While this might be possible with some factors, it is challenging with others.
Process com-plexity	A PS is able to handle processes with a very complex control flow. However, increasing complexity increases the runtime of a PS.	When implemented by a tool, the determination of the rNPV may be impossible for processes with a very complex control flow, though theoretically possible according to our valuation calculus, or the runtime for the calculation may be very high, even higher than with a PS.

	Process simulation (PS)	Valuation Calculus
Information needed	The structure of the process, the transition probabilities, and the probability distribution of CF_{a_d} and CF_{pa_s}.	The structure of the process, the transition probabilities, and only the expected value and the variance of CF_{a_d} and CF_{pa_s}.
Precision of results	A PS delivers imprecise results (Sun et al. 2006), which means that the calculation cannot be repeated in a manner that leads to the same result with every run (Pearn et al. 1998). It is a technique that can approximate the expected value and variance, but it cannot provide the correct value (van Hee and Reijers 2000). However, the more extensive the PS, the higher its precision.	The valuation calculus provides precise results.
Sensitivity analysis	A PS supports “what-if” analysis (van der Aalst 2001) to determine how the result of a process changes if, for example, one factor is changed at a time. Because of its lack of precision, however, the extent to which a changed result is due solely to the changed factor cannot be precisely determined. A change in a result could be due to the imprecision of the PS.	If a process is modified, the rNPV can be calculated again, which allows for the determination of whether a process improved from a value-based perspective to the process change. Thus, a “what-if” analysis is possible with the valuation calculus as well. This analysis is precise and thus indicates if the change in the result is due solely to the change of the process.

Table 4. Comparison of process simulation and the analytical approach of this paper

Overall, we consider process simulations to be advantageous in their expressiveness and their treatment of processes with complex control flows. However, we consider this paper’s approach to be advantageous in terms of required information and its precision in determining the expected value and variance. Particularly beneficial is the fact that, because we do not need to know the whole probability distribution of CF_{a_d} and CF_{pa_s}, the presented valuation calculus might encourage a broader use in practice.

II.2.6 Conclusion and Outlook

Process measures are important instruments for analyzing processes and deciding on process changes. For the decision making-oriented branch of value-based BPM, the rNPV of a process is an important process measure. However, current research on value-based BPM provides the rNPV on only the corporate level. Thus, this paper connects the corporate level with the operational process level, providing a valuation calculus that considers the control flow of processes. This paper contributes to value-based BPM in the following ways:

1. This paper develops its valuation calculus such that the rNPV of a process can be calculated, bringing value-based BPM to the operational process level and allowing it to be implemented via process-modeling tools. A modeling tool with such calculation capabilities is a valuable asset to any process manager who needs to decide among various process alternatives while considering the principles of value-based management.
2. The paper provides a theoretical foundation for more formal research in BPM, while already making a valuable contribution to practice, as seen in Section 5.3. The valuation calculus has helped improve the calculation capabilities of a commercial process-modeling tool currently being developed by CubeFour.
3. Finally, since the paper's formalism in calculating the expected value and variance is based on the fact that the cash flow of a process path is the sum of the cash flows of the actions in that path, this formalism is usable not only for cash flows but for any kind of additive quantity, such as costs, energy, or used material.

Despite the contributions of this paper to BPM research and practice, it has limitations that point to possibilities for future study:

1. The first limitation regarding the calculation of the rNPV is assumption (A3), that there are no dependencies among process instances, as made in other works (Bolsinger et al. 2011; Davamanirajan et al. 2006). A more general version of the valuation calculus could consider these dependencies via correlation coefficients. However, this version would make the presentation of the valuation calculus overly complex. The presented simplification eases the communication of the valuation calculus significantly.

2. Another limitation regarding the applicability of the valuation calculus is the availability of necessary data. Along with the need to determine the expected value of the action cash flow and its variance is the need to determine the path probabilities. Doing so requires information regarding the transition probabilities from action to action. The transition probabilities could be estimated by an expert (Hubbard 2007) or by analyzing process log files (zur Muehlen and Shapiro 2010) using, for example, a process-mining framework like ProM (Rubin et al. 2007). Furthermore, tapping the full potential of the variance requires that the correlations be determined. Gathering these data is possible, particularly when process-mining techniques are used, but it is not easy.

3. Value-based BPM is based on monetary values and uses cash flows as the common denominator. This common denominator allows a comparison among various process alternatives. However, different performance dimensions are typically used in BPM, such as time, cost, quality, and flexibility (Reijers and Liman Mansar 2005). While costs are already a monetary value, the other dimensions need to be monetized for the presented valuation calculus. Of the other dimensions, time, in combination with wages, can most readily be transformed into monetary values. While quality and flexibility are more challenging, some papers focus on the transformation of flexibility into monetary values (Braunwarth and Ullrich 2010; Neuhuber et al. 2013). Thus, although having a common denominator is an advantage, much more research on converting other BPM goals/non-monetary dimensions into monetary values is required. It will then be possible for value-based BPM to exploit its full potential as a comprehensive framework for BPM decisions by supporting the improvement of processes through a monetary-centered view of BPM.

4. As discussed in Section 5.4, process simulations can probably be used more conveniently with more processes than can an implemented version of the valuation calculus, as processes can be very complex. For complex processes with a high number of process paths, the expected value and variance must be calculated automatically because manual calculation would be very time-consuming. Algorithms are thus needed to determine the path sets and path probabilities. These have not been sufficiently explored. Some algorithms can calculate path sets (Byers and Waterman 1984) using depth-first search. However, as these algorithms are not specifically for processes, they do not consider all control flow patterns nor calculate path

probabilities. The depth-first search is widely used and well-studied (Sedgewick and Schidlowsky 2003). Thus, a depth-first search algorithm can be used to get all path sets and calculate the path probabilities while considering the control flow patterns. However, the runtime of such algorithms could be high for complex processes.

5. Finally, since processes can be very complex, a more formal and extensive assessment than that given in Section 5.1 is needed to determine the extent of the valuation calculus' validity for different kinds of processes. Processes are complex not only from a control flow perspective but also from, for example, resource, data, time, and function perspectives (van der Aalst 2013). Such different perspectives need to be subject to further research.

II.2.7 References

Abdellaoui M, Bleichrodt H, l'Haridon O, Paraschiv C (2013) Is there one unifying concept of utility? An experimental comparison of utility under risk and utility over time. Manage Sci 59(9):2153-2169. doi: 10.1287/mnsc.1120.1690

Andersen S, Harrison GW, Lau MI, Rutström EE (2008) Eliciting risk and time preferences. Econometrica 76(3):583-618. doi: 10.1111/j.1468-0262.2008.00848.x

Bai X, Krishnan R, Padman R, Wang HJ (2013) On Risk Management with Information Flows in Business Processes. Inform Syst Res 24(3):731-749. doi: 10.1287/isre.1120.0450

Bai X, Padman R, Krishnan R (2007) A Risk Management Approach to Business Process Design. In: Rivard S, Webster J (eds) Proceedings of the 27th International Conference on Information Systems (ICIS), Montreal, Canada, Paper 28

Bamberg G, Spremann K (1981) Implications of Constant Risk Aversion. Math Method Oper Res 25(7):205-224

Becker J, Kahn D (2005) Der Prozess im Fokus. In: Becker J, Kugeler M, Rosemann M (eds) Prozessmanagement - Ein Leitfaden zur prozessorientierten Organisationsgestaltung. Springer-Verlag GmbH, Berlin, Germany

Beer M, Fridgen G, Müller H, Wolf T (2013) Benefits Quantification in IT Projects. In: Alt R, Franczyk B (eds) Proceedings of the 11th International Conference on Wirtschaftsinformatik (WI), Leipzig, Germany, pp 707-720

Bernoulli D (1954) Exposition of a New Theory on the Measurement of Risk. Econometrica 22(1):23-36

Bolsinger M, Bewernik M, Buhl HU (2011) Value-Based Process Improvement. In: Tuunainen VK, Rossi M, Nandhakumar J (eds) Proceedings of the 19th European Conference on Information Systems (ECIS), Helsinki, Finland, Paper 21

Braunwarth K, Ullrich C (2010) Valuating Business Process Flexibility achieved through an alternative Execution Path. In: Johnson R, De Villiers C, Kruger N, Bartmann D (eds) Proceedings of the 18th European Conference on Information Systems (ECIS), Pretoria, South Africa, Paper 24

Braunwarth K, Kaiser M, Müller A (2010) Economic Evaluation and Optimization of the Degree of Automation in Insurance Processes. Bus Inf Syst Eng 2(1):29-39. doi: 10.1007/s12599-009-0088-6

Buhl HU, Röglinger M, Stöckl S, Braunwarth K (2011) Value Orientation in Process Management - Research Gap and Contribution to Economically Well-founded Decisions in Process Management. Bus Inf Syst Eng 3(3):163-172

Byers T, Waterman M (1984) Determining All Optimal and Near-Optimal Solutions When Solving Shortest Path Problems by Dynamic Programming. Oper Res 32(6):1381-1384

Coenenberg AG, Salfeld R (2007) Wertorientierte Unternehmensführung: Vom Strategieentwurf zur Implementierung. Schäffer-Poeschel, Stuttgart, Germany

Copeland TE, Koller T, Murrin J (1990) Valuation: measuring and managing the value of companies. Wiley, New York, NY, USA

Copeland TE, Weston JF, Shastri K (2005) Financial Theory and Corporate Policy. Pearson Education Inc., Boston, MA, USA

Danielson MG, Heck JL, Shaffer DR (2008) Shareholder theory - How opponents and proponents both get it wrong. J Appl Finance 18(2):62-66

Datar S, Kulp SC, Lambert RA (2001) Balancing performance measures. J Account Res 39(1):75-92

Davamanirajan P, Kauffman RJ, Kriebel CH, Mukhopadhyay T (2006) Systems Design, Process Performance, and Economic Outcomes in International Banking. J Manage Inform Syst 23(2):65-90

Devaraj S, Kohli R (2001) The IT payoff: measuring the business value of information technology investments. Financial Times/Prentice Hall

Feiler PH, Humphrey WS (1993) Software process development and enactment: Concepts and definitions. In: Proceedings of the 2nd International Conference on the Software Process. Continuous Software Process Improvement. Berlin, Germany, pp 28-40

Feller W (1971) An Introduction to Probability Theory and Its Applications, Vol. 2. John Wiley & Sons, Inc., New York, NY, USA

Franz P, Kirchmer M, Rosemann M (2011) Value-Driven Business Process Management – Impact and Benefits. http://www.accenture.com/SiteCollectionDocuments/PDF/Accenture-Value-Driven-Business-Process-Management.pdf. Accessed 2014-02-21

Freund RJ (1956) The Introduction of Risk into a Programming Model. Econometrica 24(3):253-263

Fridgen G, Müller H (2009) Risk/Cost Valuation of Fixed Price IT Outsourcing in a Portfolio Context. In: Nunamaker J, Currie W, Chen H, Slaughter S (eds) Proceedings of the 30th International Conference on Information Systems (ICIS), Phoenix, AZ, USA

Friedman M, Savage LJ (1948) The Utility Analysis of Choices Involving Risk. J Polit Econ 56(4):279-304

Gartner (2013) Hunting and Harvesting in a Digital World: The 2013 CIO Agenda.

Gibbons R (2005) Incentives between Firms (And within). Manage Sci 51(1):2-17

González LS, Rubio FG, González FR, Velthuis MP (2010) Measurement in business processes: a systematic review. BPMJ 16(1):114-134. doi: 10.1108/14637151011017976

Grob HL (1993) Capital Budgeting with Financial Plans: An Introduction. Gabler, Wiesbaden, Germany

Gulledge TR, Hirschmann P, Scheer A (1997) Value-Based Management of Inter-Organizational Business Processes. In: Krallmann H (ed) Proceedings of the 3rd International Conference on Wirtschaftsinformatik (WI), Berlin, Germany, pp 73-98

Hevner AR, March ST, Park J, Ram S (2004) Design Science in Information Systems Research. MISQ 28(1):75-105

Hillier FS (1963) The Derivation of Probabilistic Information for the Evaluation of Risky Investments. Manage Sci 9(3):443-457

Hollingsworth DC, WfMC (2003) The WfMC Glossary. In: Fischer L (ed) Workflow Handbook 2003. Future Strategies Inc., Lighthouse Point, Florida, USA, pp 263-299

Hubbard DW (2007) How to Measure Anything: Finding the Value of Intangibles in Business. John Wiley & Sons, Inc., New York, NY, USA

Ittner CD, Larcker DF (2001) Assessing empirical research in managerial accounting: a value-based management perspective. J Account Econ 32(1-3):349-410

Kasanen E, Trigeorgis L (1994) A market utility approach to investment valuation. Eur J Oper Res 74(2):294-309

Keeney RL (1994) Creativity in decision making with value-focused thinking. Sloan Manage Rev 35(4):33-41

Koller T, Goedhart M, Wessels D (2010) Valuation: Measuring and Managing the Value of Companies. John Wiley, New Jersey, USA

Kruschwitz L, Husmann S (2010) Finanzierung und Investition. Oldenburg Verlag, Munich, Germany

Kruschwitz L, Löffler A (2003) Semi-subjektive Bewertung. Z Betriebswirt 73(12):1335-1345

Kueng P, Kawalek P (1997) Goal-based business process models: creation and evaluation. BPMJ 3(1):17-38

Linderman K, McKone-Sweet KE, Anderson JC (2005) An integrated systems approach to process control and maintenance. Eur J Oper Res 164(2):324-340

Longley-Cook AG (1998) Risk-Adjusted Economic Value Analysis. N Am Actuar J 2(1):87-100

Meredith JR, Raturi A, Amoako-Gyampah K, Kaplan B (1989) Alternative Research Paradigms in Operations. J Oper Manage 8(4):297-326

Mosteller F, Nogee P (1951) An experimental measurement of utility. J Polit Econ 59:371-404

Neiger D, Churilov L (2004a) Goal-Oriented Business Process Engineering Revisited: a Unifying Perspective. In: Cordeiro J, Felipe J (eds) Proceedings of the 1st International Workshop on Computer Supported Activity Coordination (CSAC) in conjunction with the 6th International Conference on Enterprise Information Systems (ICEIS), Porto, Portugal, pp 149-163

Neiger D, Churilov L (2004b) Goal-Oriented Business Process Modeling with EPCs and Value-Focused Thinking. Business Process Management (LNCS) 3080:98-115

Neiger D, Churilov L, zur Muehlen M, Rosemann M (2006) Integrating Risks in Business Process Models with Value Focused Process Engineering. In: Ljungberg J, Andersson M (eds) Proceedings of the 14th European Conference on Information Systems (ECIS), Gothenburg, Sweden, pp 1606-1615

Neuhuber LCN, Krause F, Roeglinger M (2013) Flexibilization Of Service Processes: Toward An Economic Optimization Model. In: Brinkkemper S, Batenburg R, van Hillegersberg J, van Heck E, Spiekermann S, Connolly R (eds) Proceedings of the 21st European Conference on Information Systems (ECIS), Utrecht, The Netherlands, Paper 66

Object Management Group (2011) UML Superstructure Specification v 2.4.1. http://www.omg.org/spec/UML/2.4.1/Superstructure/PDF. Accessed 2014-02-21

Pearn W, Lin G, Chen K (1998) Distributional and inferential properties of the process accuracy and process precision indices. Commun Stat-Theor M 27(4):985-1000

Peffers K, Tuunanen T, Rothenberger MA, Chatterjee S (2008) A Design Science Research Methodology for Information Systems Research. J Manage Inf Syst 24(3):45-77

Rappaport A (1986) Creating Shareholder Value: The New Standard for Business Performance. Free Press, New York, NY, USA

Reijers HA, Liman Mansar S (2005) Best practices in business process redesign: an overview and qualitative evaluation of successful redesign heuristics. Omega 33(4):283-306. doi: 10.1016/j.omega.2004.04.012

Rotaru K, Wilkin C, Churilov L, Neiger D, Ceglowski A (2011) Formalizing process-based risk with Value-Focused Process Engineering. Inf Syst E-Bus Manage 9(4):447-474. doi: 10.1007/s10257-009-0125-5

Rubin V, Günther CW, van der Aalst WMP, Kindler E, van Dongen BF, Schäfer W (2007) Process mining framework for software processes. In: Wang Q, Pfahl D, Raffo DM (eds) Proceedings of the International Conference on Software Process: Software Process Dynamics and Agility (ICSP), Minneapolis, MN, USA, pp 169-181

Sampath P, Wirsing M (2011) Evaluation of Cost Based Best Practices in Business Processes. In: Halpin T, Nurcan S, Krogstie J (eds) Proceedings of the 12th International Conference on Business Process Modeling, Development, and Support (BPMDS) and 16th International Conference on Exploring Modelling Methods for Systems Analysis and Design (EMMSAD) held at the 23rd International Conference on Advanced Information Systems Engineering (CAiSE): Enterprise, Business-Process and Information Systems Modeling, London, United Kingdom, pp 61-74

Sedgewick R, Schidlowsky M (2003) Algorithms in Java, Part 5: Graph Algorithms. Addison-Wesley Longman Publishing Co., Inc, Boston, MA, USA

Sen S, Raghu T (2013) Interdependencies in IT Infrastructure Services: Analyzing Service Processes for Optimal Incentive Design. Inform Syst Res 24(3):822-841. doi: 10.1287/isre.2013.0475

Sidorova A, Isik O (2010) Business process research: a cross-disciplinary review. BPMJ 16(4):566-597. doi: 10.1108/14637151011065928

Sonnenberg C, vom Brocke J (2012) Evaluations in the Science of the Artificial–Reconsidering the Build-Evaluate Pattern in Design Science Research. In: Peffers K, Rothenberger M, Kuechler B (eds) Proceedings of the 7th International Conference on Design Science Research in Information Systems: Advances in Theory and Practice (DESRIST), Las Vegas, NV, USA, pp 381-397

Stewart GB, Stern JM (1991) The quest for value: the EVA management guide. HarperBusiness, New York, NY, USA

Sun SX, Zhao JL, Nunamaker JF, Sheng ORL (2006) Formulating the Data-Flow Perspective for Business Process Management. Inform Syst Res 17(4):374-391

Swalm RO (1966) Utility Theory - Insights into Risk Taking. Harvard Bus Rev 44(6):123-136

Thome R, Müller T, Vogeler K (2011) Zukunftsthema Geschäftsprozessmanagement. http://www.pwc.de/de_DE/de/prozessoptimierung/assets/PwC-GPM-Studie.pdf. Accessed 2014-02-21

Tregear R (2012) Practical Process: Measuring Processes. http://www.bptrends.com/publicationfiles/11-06-2012-Practical%20Process_Measuring%20Processes-Tregear.pdf. Accessed 2014-02-21 van der Aalst WMP (2013) Business Process Management: A Comprehensive Survey. ISRN Software Eng, vol. 2013, Article ID 507984, 37 pages. doi: 10.1155/2013/507984

van der Aalst WMP, van Hee KM, ter Hofstede AH, Sidorova N, Verbeek H, Voorhoeve M, Wynn MT (2011) Soundness of workflow nets: classification, decidability, and analysis. Form Asp Comput 23(3):333-363

van der Aalst WMP (2001) Re-engineering knock-out processes. Decis Support Syst 30(4):451-468. doi: 10.1016/S0167-9236(00)00136-6

van der Aalst WMP, ter Hofstede AH, Kiepuszewski B, Barros AP (2003) Workflow patterns. Distrib Parallel Dat 14(3):5-51

van Hee K, Reijers H (2000) Using Formal Analysis Techniques in Business Process Redesign. In: van der Aalst WMP, Desel J, Oberweis A (eds) Business Process Management - Models, Techniques, and Empirical Studies. Springer-Verlag, Berlin Heidelberg, Germany, pp 142-160

Venable J, Pries-Heje J, Baskerville R (2012) A comprehensive framework for evaluation in design science research. In: Peffers K, Rothenberger M, Kuechler B (eds) Proceedings of the 7th International Conference on Design Science Research in Information Systems: Advances in Theory and Practice (DESRIST), Las Vegas, NV, USA, pp 423-438

Vergidis K, Tiwari A, Majeed B (2008) Business Process Analysis and Optimization: Beyond Reengineering. IEEE T Syst Man Cy C 38(1):69-82. doi: 10.1109/TSMCC.2007.905812

vom Brocke J, Becker J, Braccini AM, Butleris R, Hofreiter B, Kapočius K, De Marco M, Schmidt G, Seidel S, Simons A, Skopal T, Stein A, Stieglitz S, Suomi R, Vossen G, Winter R, Wrycza S (2011a) Current and future issues in BPM research: a European perspective from the ERCIS meeting 2010. CAIS 28(1):393-414

vom Brocke J, Sonnenberg C, Baumoel U (2011b) Linking Accounting and Process-aware Information Systems–Towards a Generalized Information Model For Process-oriented Accounting. In: Tuunainen VK, Rossi M, Nandhakumar J (eds) Proceedings of the 19th European Conference on Information Systems (ECIS), Helsinki, Finland, Paper 23

vom Brocke J, Recker JC, Mendling J (2010) Value-oriented Process Modeling: Integrating Financial Perspectives into Business Process Re-design. BPMJ 16(2):333-356

vom Brocke J, Sonnenberg C, Simons A (2009) Wertorientiertes Prozessmanagement: State-of-the-Art und zukünftiger Forschungsbedarf. In: Hansen HR, Karagiannis D, Fill H (eds) Proceedings of the 9th International Conference on Wirtschaftsinformatik (WI), Vienna, Austria, pp 253-262

Wolf C, Harmon P (2012) The State of Business Process Management 2012. http://www.bptrends.com/bpt/wp-content/surveys/2012-_BPT%20SURVEY-3-12-12-CW-PH.pdf. Accessed 2014-02-21

Wynn MT, Low WZ, Nauta W (2013) A framework for cost-aware process management: generation of accurate and timely management accounting cost reports. In: Ferrarotti F, Grossmann G (eds) Proceedings of the Ninth Asia-Pacific Conference on Conceptual Modelling (APCCM), Adelaide, Australia, pp 79-88

Young SD, O'Byrne SF (2001) EVA and value-based management: a practical guide to implementation. McGraw-Hill, New York, NY, USA

Zimmermann S, Katzmarzik A, Kundisch D (2008) IT Sourcing Portfolio Management for IT Service Providers - A Risk/Cost Perspective. In: Boland R, Limayem M, Pentland B (eds) Proceedings of 29th International Conference on Information Systems (ICIS), Paris, France

zur Muehlen M, Shapiro R (2010) Business Process Analytics. In: vom Brocke J, Rosemann M (eds) Handbook on Business Process Management 2: Strategic Alignment, Governance, People and Culture. Springer, Berlin, Germany, pp 137-158

II.2.8 Appendix

II.2.8.A Determination of Path Probabilities

To determine the expected value of a process cash flow and its variance, it is essential to determine the path probabilities p_k. This is presented in the following. During a process improvement project, a process is presented as a process model with a process-modeling tool. With the help of this formal presentation, it is possible to formally describe, how a path probability p_k is determined. In order to do so, the process model (as defined in Hollingsworth and WfMC [2003, p. 266]) of a process *P* is defined as a graph *G*.

The process model of a process *P* is a graph, because a process model is a set of nodes (*vertices*) that are interconnected by arrows (*edges*) (Gibbons 1985). The set of vertices is denoted by *V* and the set of edges by *E* and we write *G = (V, E)*. Because the edges are arrows, a process is a directed graph (Gibbons 1985). More precisely, we assume that a process model of a process *P* is defined as a graph *G* as followed:

(D1) A process model of a process *P* is a directed graph *G = (V, E)* with one root vertex a_0 and one final vertex a_{D+1}, toward which all edges are directed. It is *V* the set of vertices and *E* the set of edges.

(D2) The set *V* consists of the set of actions *A* united with the set *RC* of the routing constructs (van der Aalst et al. 2003) to denote control flow patterns of *P*, a_0 and a_{D+1}, i.e., $V \coloneqq A \cup RC \cup a_0 \cup a_{D+1}$.

(D3) *A* contains all *D* actions of *P*, numbered from 1 to *D*.

(D4) *RC* is the set of the routing constructs to denote the control flow patterns, e.g., XOR-split, XOR-join, AND-split and AND-join (van der Aalst et al. 2003). Each element has one distinct index. For example, in Figure 2 (left process) these vertices are $\textit{XOR-join}_1$, $\textit{AND-split}_2$, $\textit{AND-join}_3$ and $\textit{XOR-split}_4$.

(D5) The edge-set *E* contains all the directed edges between the vertices. The directed edge (v_i, v_j, p_{ij}) is a member of the set *E* if and only if there is an arrow between vertex $v_i \in V$ and $v_j \in V$, pointing from v_i to v_j and having a probability for this transition (Hollingsworth and WfMC 2003, p. 282) of p_{ij}, with $0 < p_{ij} \leq 1$, as weight. Each vertex in *A* has exactly one edge pointing toward it and exactly one edge pointing away from it.

The actions and routing constructs of a process path pp_k plus a_0 and a_{D+1} form (in a first step) a path multiset PS_k, whose elements are out of *V*. The fact that it is a multiset is important to consider loops, as then the same vertices of *G* can occur several times. Each vertex v_i in PS_k that occurs more than once (in a second step) is given an index $n \in \mathbb{N}$ in the form $v_i^{(1)}, v_i^{(2)}, \dots, v_i^{(n)}, \dots$. The index indicates the number of the iteration of a loop that the vertex is assigned to. This is to distinguish the vertices from one another because each of them is from different iterations that have different probabilities of being executed. In the left process in Figure 2, there are the path sets

$PS_1 = \{a_0, a_1, XOR\text{-}join_1^{(1)}, AND\text{-}split_2^{(1)}, a_2^{(1)}, a_3^{(1)}, AND\text{-}join_3^{(1)}, XOR\text{-}split_4^{(1)}, a_5\}$,

$PS_2 = \{a_0, a_1, XOR\text{-}join_1^{(1)}, AND\text{-}split_2^{(1)}, a_2^{(1)}, a_3^{(1)}, AND\text{-}join_3^{(1)}, XOR\text{-}split_4^{(1)}, a_4^{(1)}, XOR\text{-}join_1^{(2)}, AND\text{-}split_2^{(2)}, a_2^{(2)}, a_3^{(2)}, AND\text{-}join_3^{(2)}, XOR\text{-}split_4^{(2)}, a_5\}$

and so on, with $v_1 := a_0$, $v_2 := a_1$, $v_3^{(1)} := XOR\text{-}join_1^{(1)}$, $v_3^{(2)} := XOR\text{-}join_1^{(2)}$, …, $v_4^{(1)} := AND\text{-}split_2^{(1)}$, $v_4^{(2)} := AND\text{-}split_2^{(2)}$, …, $v_5^{(1)} := a_2^{(1)}$, $v_5^{(2)} := a_2^{(2)}$, …, $v_6^{(1)} := a_3^{(1)}$, $v_6^{(2)} := a_3^{(2)}$, …, $v_7^{(1)} := AND\text{-}join_3^{(1)}$, $v_7^{(2)} := AND\text{-}join_3^{(2)}$, …, $v_8^{(1)} := XOR\text{-}split_4^{(1)}$, $v_8^{(2)} := XOR\text{-}split_4^{(2)}$, …, $v_9^{(1)} := a_4^{(1)}$, …, and $v_{10} := a_5$.

Every process path pp_k is executed with a certain *path probability* p_k that is the product of the transition probabilities of process path pp_k:

$$p_k = \prod_{v_i^{(m)}, v_j^{(n)} \in PS_k} p_{i^{(m)}j^{(n)}} \qquad \text{for all } p_{i^{(m)}j^{(n)}} > 0. \tag{22}$$

The transition probability $p_{i^{(m)}j^{(n)}}$ that $v_i^{(m)}$ is followed by $v_j^{(n)}$ can be estimated and is fixed. These transition probabilities could be estimated by an expert (Hubbard 2007) or by analyzing process log files (zur Muehlen and Shapiro 2010) using, for example, a process-mining framework like ProM (Rubin et al. 2007). In the left process in Figure 2, for example, for the process path pp_1 there are the (non-zero) transition probabilities $p_{12} = 1$, $p_{23^{(1)}} = 1$, $p_{3^{(1)}4^{(1)}} = 1$, $p_{4^{(1)}5^{(1)}} = 1$, $p_{4^{(1)}6^{(1)}} = 1$, $p_{5^{(1)}7^{(1)}} = 1$, $p_{6^{(1)}7^{(1)}} = 1$, $p_{7^{(1)}8^{(1)}} = 1$ and $p_{8^{(1)},10} = 0.9$. All other transition probabilities are zero. Then it is

$$p_1 = \prod_{v_i^{(m)}, v_j^{(n)} \in PS_1} p_{i^{(m)}j^{(n)}}$$

$$= \underbrace{1}_{a_0\ to\ a_1} \cdot \underbrace{1}_{a_1\ to\ XOR-join_1} \cdot \underbrace{1}_{XOR-join_1\ to\ AND-split_2} \cdot \underbrace{1}_{AND-split_2\ to\ a_2} \cdot \underbrace{1}_{AND-split_2\ to\ a_3}$$

$$\cdot \underbrace{1}_{a_2\ to\ AND-join_3} \cdot \underbrace{1}_{a_3\ to\ AND-join_3} \cdot \underbrace{1}_{AND-join_3\ to\ XOR-split_4} \cdot \underbrace{0.9}_{XOR-split_4\ to\ a_5}$$

$$= 0.9$$

and

$p_2 = \prod_{v_i^{(m)}, v_j^{(n)} \in PS_2} p_{i^{(m)} j^{(n)}} = 0.09$, etc.

Expression (22) is not true in the event that a process model contains an OR-split (van der Aalst et al. 2003). This fact is important in Section 5.3, when showing how this valuation calculus helped to improve the calculation capabilities of a process-modeling tool. However, every OR-split can formally be transformed into a composition of XOR-splits and AND-splits, which allows the use of expression (22). Otherwise, the path probabilities need to be estimated.

II.2.8.B Process-Probability-Space

In probability theory, "a probability space is a triple $(\Omega, \mathcal{F}, \mathrm{P})$ of a sample space Ω, a [sigma]-algebra $\mathcal{F}$ and a probability measure **P** on $\mathcal{F}$" (Feller 1971, p. 116). The sample space Ω is the set of all possibilities that the object under consideration can take; it is thus the set of all possible process paths, as these represent all possibilities of a process execution. A sigma-algebra has properties such that:

(i) "If a set A is in $\mathcal{F}$ so is its complement $[A^C = \Omega \backslash A]$.

(ii) If $\{A_n\}$ is any countable collection of sets in $\mathcal{F}$, then also their union $\bigcup A_n$ and intersection $\bigcap A_n$belong to $\mathcal{F}$" (Feller 1971, p. 112).

That the sigma-algebra in definition 2 is the power set of the set of all process paths means that (i) and (ii) are fulfilled.

"A probability measure **P** on a [sigma]-algebra $\mathcal{F}$ of sets in Ω is a function assigning a value $\mathrm{P}\{A\} \geq 0$ to each set A in $\mathcal{F}$ such that $\mathbf{P}\{\Omega\} = 1$ and that for every countable collection of non-overlapping sets A_n in $\mathcal{F}$ [it is] $\mathbf{P}\{\bigcup A_n\} = \sum_n \mathbf{P}\{A_n\}$" (Feller 1971, p. 115).

All process paths are mutually exclusive, and they represent all possibilities how a process can be executed. Every process path pp_k is executed with a certain *path probability* $p_k > 0$. Given that there is exactly one process path taken if a process is executed and that they are mutual exclusive, the probabilities p_k sum up to 1, fulfilling $\mathbf{P}\{\Omega\} = 1$. The property

$\mathbf{P}\{\bigcup A_n\} = \sum_n \mathbf{P}\{A_n\}$ also holds for every countable collection of non-overlapping sets A_n in $\mathcal{F}$ since $\mathcal{F}$ is the power set of Ω.

II.2.8.C Expected Value of the Process Cash Flow

Let the probability that an action $a_d \in AS$, with $AS := \bigcup_{k=1}^{|\Omega|} AS_k$, is executed when executing a process be

$$Pr(a_d) := \mathbf{P}_P\{PP_{a_d}\} = \sum_{k=1}^{|\Omega|} p_k \cdot \mathbb{I}_{AS_k}(a_d)$$

with the indicator function

$$\mathbb{I}_{AS_k}(a_d) = \begin{cases} 1, & a_d \in AS_k \\ 0, & a_d \notin AS_k \end{cases}$$

and the set PP_{a_d} of process paths in which the action a_d is

$$PP_{a_d} = \{pp_k \in \Omega | a_d \in AS_k\}.$$

Then it is:

$$E[CF_P] = \sum_{k=1}^{|\Omega|} E[CF_P|PI = k] \cdot Prob(PI = k) = \sum_{k=1}^{|\Omega|} E[CF_{pp_k}] \cdot p_k$$

$$= \sum_{k=1}^{|\Omega|} \left(p_k \cdot E\left[\sum_{a_d \in AS_k} CF_{a_d} + \sum_{s=1}^{S} CF_{pa_s} \right] \right)$$

$$= \sum_{k=1}^{|\Omega|} \left(p_k \cdot \left(\sum_{a_d \in AS_k} E[CF_{a_d}] + \sum_{s=1}^{S} E[CF_{pa_s}] \right) \right)$$

$$= \sum_{k=1}^{|\Omega|} \left(\sum_{a_d \in AS_k} p_k \cdot E[CF_{a_d}] \right) + \sum_{k=1}^{|\Omega|} \left(p_k \cdot \sum_{s=1}^{S} E[CF_{pa_s}] \right)$$

$$= \sum_{k=1}^{|\Omega|} \left(\sum_{a_d \in AS_k} \mathbb{I}_{AS_k}(a_d) \cdot p_k \cdot E[CF_{a_d}] \right) + \sum_{s=1}^{S} E[CF_{pa_s}] \cdot \underbrace{\sum_{k=1}^{|\Omega|} p_k}_{=1}$$

$$= \sum_{k=1}^{|\Omega|} \left(\sum_{a_d \in AS} \mathbb{I}_{AS_k}(a_d) \cdot p_k \cdot E[CF_{a_d}] \right) + \sum_{s=1}^{S} E[CF_{pa_s}]$$

$$= \sum_{a_d \in AS} \left(\sum_{k=1}^{|\Omega|} \mathbb{I}_{AS_k}(a_d) \cdot p_k \cdot E\left[CF_{a_d}\right] \right) + \sum_{s=1}^{S} E\left[CF_{pa_s}\right]$$

$$= \sum_{a_d \in AS} E\left[CF_{a_d}\right] \left(\sum_{k=1}^{|\Omega|} p_k \cdot \mathbb{I}_{AS_k}(a_d) \right) + \sum_{s=1}^{S} E\left[CF_{pa_s}\right]$$

$$= \sum_{a_d \in AS} E\left[CF_{a_d}\right] \cdot Pr(a_d) + \sum_{s=1}^{S} E\left[CF_{pa_s}\right]$$

II.2.8.D Variance of the Process Cash Flow

In the following first step, it is shown that $Var[CF_P] = -E[CF_P]^2 + \sum_{k=1}^{|\Omega|} p_k \cdot E\left[CF_{pp_k}^2\right]$ in two ways. The first way is similar to the beginning of the calculation for the expected value in appendix C. The second way is more detailed and includes $\sum_{k=1}^{|\Omega|} E\left[\left(CF_{pp_k} - E[CF_P]\right)^2\right] \cdot p_k$, a more intuitive expression for $Var[CF_P]$. This is why both ways are presented.

Way 1

$$Var[CF_P] = E[CF_P^2] - E[CF_P]^2 = -E[CF_P]^2 + \sum_{k=1}^{|\Omega|} E[CF_P^2 | PI = k] \cdot Prob(PI = k)$$

$$= -E[CF_P]^2 + \sum_{k=1}^{|\Omega|} p_k \cdot E\left[CF_{pp_k}^2\right]$$

Way 2

$$Var[CF_P] = E[(CF_P - E[CF_P])^2]$$

$$= \sum_{k=1}^{|\Omega|} E[(CF_P - E[CF_P])^2 | PI = k] \cdot Prob(PI = k)$$

$$= \boldsymbol{\sum_{k=1}^{|\Omega|} E\left[\left(CF_{pp_k} - E[CF_P]\right)^2\right] \cdot p_k}$$

$$= \sum_{k=1}^{|\Omega|} p_k \cdot E\left[CF_{pp_k}^2 - 2 \cdot CF_{pp_k}^2 \cdot E[CF_P] + E[CF_P]^2\right]$$

$$= \sum_{k=1}^{|\Omega|} p_k \cdot E\left[CF_{pp_k}^2\right] - 2 \cdot E[CF_P] \underbrace{\sum_{k=1}^{|\Omega|} p_k \cdot E\left[CF_{pp_k}\right]}_{=E[FQ_P]} + E[CF_P]^2 \underbrace{\sum_{k=1}^{|\Omega|} p_k}_{=1}$$

$$= -2 \cdot E[CF_P]^2 + E[CF_P]^2 + \sum_{k=1}^{|\Omega|} p_k \cdot E\left[CF_{pp_k}^2\right] = -E[CF_P]^2 + \sum_{k=1}^{|\Omega|} p_k \cdot E\left[CF_{pp_k}^2\right]$$

In the following second step, it is shown how $Var[CF_P]$ can be calculated only by using the expected values and variances of the cash flows of the actions of a process.

Let the probability that both actions $a_d \in AS$ and $a_j \in AS$, with $AS \coloneqq \bigcup_{k=1}^{|\Omega|} AS_k$, are executed when executing a process be

$$Pr(a_d, a_j) := \mathbf{P}_P\left\{PP_{a_d,a_j}\right\} = \sum_{k=1}^{|\Omega|} p_k \cdot \mathbb{I}_{AS_k}(a_d) \cdot \mathbb{I}_{AS_k}(a_j)$$

with the set PP_{a_d,a_j} of process paths, which contains the action a_d as well as the action a_j:

$$PP_{a_d,a_j} = \left\{pp_k \in \Omega | a_d \in AS_k, a_j \in AS_k\right\}.$$

Then it is:

$$Var[CF_P] \overset{first\ step}{\widehat{=}} -E[CF_P]^2 + \sum_{k=1}^{|\Omega|} p_k \cdot E\left[CF_{pp_k}^2\right]$$

$$= -E[CF_P]^2 + \sum_{k=1}^{|\Omega|} p_k \cdot E\left[\left(\sum_{a_d \in AS_k} CF_{a_d} + \sum_{s=1}^{S} CF_{pa_s}\right)^2\right]$$

$$= -E[CF_P]^2 + \sum_{k=1}^{|\Omega|} p_k \cdot E\left[\left(\sum_{a_d \in AS} CF_{a_d} \cdot \mathbb{I}_{AS_k}(a_d) + \sum_{s=1}^{S} CF_{pa_s}\right)^2\right]$$

$$= -E[CF_P]^2 + \sum_{k=1}^{|\Omega|} p_k$$

$$\cdot E\left[\left(\sum_{a_d \in AS} CF_{a_d} \cdot \mathbb{I}_{AS_k}(a_d)\right)^2 + 2\left(\sum_{a_d \in AS} CF_{a_d} \cdot \mathbb{I}_{AS_k}(a_d)\right)\left(\sum_{s=1}^{S} CF_{pa_s}\right)\right.$$

$$\left. + \left(\sum_{s=1}^{S} CF_{pa_s}\right)^2\right]$$

$$
\begin{aligned}
&= -E[CF_P]^2 + \sum_{k=1}^{|\Omega|} p_k \\
&\quad \cdot E\left[\sum_{a_d \in AS} CF_{a_d}^2 \cdot \mathbb{I}_{AS_k}(a_d) + \sum_{a_d, a_j \in AS, a_d \neq a_j} CF_{a_d} \cdot CF_{a_j} \cdot \mathbb{I}_{AS_k}(a_d) \cdot \mathbb{I}_{AS_k}(a_j)\right. \\
&\quad \left. + 2 \sum_{a_d \in AS} \sum_{s=1}^{S} CF_{a_d} \cdot \mathbb{I}_{AS_k}(a_d) \cdot CF_{pa_s} + \sum_{s=1}^{S} CF_{pa_s}^2 + 2 \sum_{s=1}^{S-1} \sum_{j=s+1}^{S} CF_{pa_s} \cdot CF_{pa_j}\right]
\end{aligned}
$$

$$
\begin{aligned}
&= -E[CF_P]^2 + \sum_{k=1}^{|\Omega|} p_k \left(\sum_{a_d \in AS} E\left[CF_{a_d}^2\right] \cdot \mathbb{I}_{AS_k}(a_d) + \sum_{s=1}^{S} E\left[CF_{pa_s}^2\right] \right. \\
&\quad + \sum_{a_d, a_j \in AS, a_d \neq a_j} E\left[CF_{a_d} \cdot CF_{a_j}\right] \cdot \mathbb{I}_{AS_k}(a_d) \cdot \mathbb{I}_{AS_k}(a_j) \\
&\quad \left. + 2 \sum_{a_d \in AS} \sum_{s=1}^{S} E\left[CF_{a_d} \cdot CF_{pa_s}\right] \cdot \mathbb{I}_{AS_k}(a_d) + 2 \sum_{s=1}^{S-1} \sum_{j=s+1}^{S} E\left[CF_{pa_s} \cdot CF_{pa_j}\right] \right)
\end{aligned}
$$

$$
\begin{aligned}
&= -E[CF_P]^2 + \sum_{a_d \in AS} E\left[CF_{a_d}^2\right] \cdot \left(\sum_{k=1}^{|\Omega|} p_k \cdot \mathbb{I}_{AS_k}(a_d) \right) + \sum_{s=1}^{S} E\left[CF_{pa_s}^2\right] \cdot \underbrace{\sum_{k=1}^{|\Omega|} p_k}_{=1} \\
&\quad + \sum_{a_d, a_j \in AS, a_d \neq a_j} E\left[CF_{a_d} \cdot CF_{a_j}\right] \cdot \left(\sum_{k=1}^{|\Omega|} p_k \cdot \mathbb{I}_{AS_k}(a_d) \cdot \mathbb{I}_{AS_k}(a_j) \right) \\
&\quad + 2 \sum_{a_d \in AS} \sum_{s=1}^{S} E\left[CF_{a_d} \cdot CF_{pa_s}\right] \cdot \left(\sum_{k=1}^{|\Omega|} p_k \cdot \mathbb{I}_{AS_k}(a_d) \right) \\
&\quad + 2 \sum_{s=1}^{S-1} \sum_{j=s+1}^{S} E\left[CF_{pa_s} \cdot CF_{pa_j}\right] \cdot \underbrace{\sum_{k=1}^{|\Omega|} p_k}_{=1}
\end{aligned}
$$

$$
\begin{aligned}
&= -E[CF_P]^2 + \sum_{a_d \in AS} E\left[CF_{a_d}^2\right] \cdot Pr(a_d) + \sum_{s=1}^{S} E\left[CF_{pa_s}^2\right] \\
&\quad + \sum_{a_d, a_j \in AS, a_d \neq a_j} E\left[CF_{a_d} \cdot CF_{a_j}\right] \cdot Pr(a_d, a_j) \\
&\quad + 2 \sum_{a_d \in AS} \sum_{s=1}^{S} E\left[CF_{a_d} \cdot CF_{pa_s}\right] \cdot Pr(a_d) + 2 \sum_{s=1}^{S-1} \sum_{j=s+1}^{S} E\left[CF_{pa_s} \cdot CF_{pa_j}\right]
\end{aligned}
$$

$$= -E[CF_P]^2 + \sum_{a_d \in AS} \left(Var\left[CF_{a_d}\right] + E\left[CF_{a_d}\right]^2\right) \cdot Pr(a_d) + \sum_{s=1}^{S} \left(Var\left[CF_{pa_s}\right] + E\left[CF_{pa_s}\right]^2\right)$$

$$+ \sum_{a_d, a_j \in AS, a_d \neq a_j} \left(Cov\left[CF_{a_d}, CF_{a_j}\right] + E\left[CF_{a_d}\right] E\left[CF_{a_j}\right]\right) \cdot Pr\left(a_d, a_j\right)$$

$$+ 2 \sum_{a_d \in AS} \sum_{s=1}^{S} \left(Cov\left[CF_{a_d}, CF_{pa_s}\right] + E\left[CF_{a_d}\right] E\left[CF_{pa_s}\right]\right) \cdot Pr(a_d)$$

$$+ 2 \sum_{s=1}^{S-1} \sum_{j=s+1}^{S} \left(Cov\left[CF_{pa_s}, CF_{pa_j}\right] + E\left[CF_{pa_s}\right] E\left[CF_{pa_j}\right]\right)$$

$$= -E[CF_P]^2 + \sum_{a_d \in AS} \left(Var\left[CF_{a_d}\right] + E\left[CF_{a_d}\right]^2\right) \cdot Pr(a_d) + \sum_{s=1}^{S} \left(Var\left[CF_{pa_s}\right] + E\left[CF_{pa_s}\right]^2\right)$$

$$+ \sum_{a_d, a_j \in AS, a_d \neq a_j} \left(\rho_{a_d, a_j} \cdot \sigma_{a_d} \cdot \sigma_{a_j} + E\left[CF_{a_d}\right] E\left[CF_{a_j}\right]\right) \cdot Pr\left(a_d, a_j\right)$$

$$+ 2 \sum_{a_d \in AS} \sum_{s=1}^{S} \left(\rho_{a_d, pa_s} \cdot \sigma_{a_d} \cdot \sigma_{pa_s} + E\left[CF_{a_d}\right] E\left[CF_{pa_s}\right]\right) \cdot Pr(a_d)$$

$$+ 2 \sum_{s=1}^{S-1} \sum_{j=s+1}^{S} \left(\rho_{pa_s, pa_j} \cdot \sigma_{pa_s} \cdot \sigma_{pa_j} + E\left[CF_{pa_s}\right] E\left[CF_{pa_j}\right]\right)$$

II.2.8.E Probability of Each Action in Process PR

In order to determine the expected value of CF_{PR} we first need to determine the probability of each action. This is:

$$Pr(a_1) = 0.9 + 0.09 + 0.009 + \cdots = 0.9 \cdot \sum_{i=0}^{\infty} 0.1^i = 1,$$

$$Pr\left(a_2^{(1)}\right) = 0.9 + 0.09 + 0.009 + \cdots = 0.9 \cdot \sum_{i=0}^{\infty} 0.1^i = 1,$$

$$Pr\left(a_2^{(2)}\right) = 0.09 + 0.009 + 0.0009 + \cdots = 0.09 \cdot \sum_{i=0}^{\infty} 0.1^i = 0.1,$$

$$\ldots,$$

$$Pr\left(a_3^{(1)}\right) = 0.9 + 0.09 + 0.009 + \cdots = 0.9 \cdot \sum_{i=0}^{\infty} 0.1^i = 1,$$

$$Pr\left(a_3^{(2)}\right) = 0.09 + 0.009 + 0.0009 + \cdots = 0.09 \cdot \sum_{i=0}^{\infty} 0.1^i = 0.1,$$

$$\ldots,$$

$$Pr\left(a_4^{(1)}\right) = 0.09 + 0.009 + 0.0009 + \cdots = 0.09 \cdot \sum_{i=0}^{\infty} 0.1^i = 0.1,$$

$$\ldots .$$

Thus, it is for example

$$\sum_{i=1}^{\infty} Pr\left(a_2^{(i)}\right) = \sum_{i=0}^{\infty} 0.1^i = \frac{1}{1-0.1} = \frac{10}{9},$$

which is multiplied with $E\left[CF_{a_2}\right]$ since it is $E\left[CF_{a_2^{(i)}}\right] = E\left[CF_{a_2}\right]$ for all $i \in \mathbb{N}$.

II.2.8.F Details to determine the variance of $\boldsymbol{CF_{PR}}$

In order to determine the variance of CF_{PR} with expression (21) it is necessary to calculate $\sum_{a_d,a_j \in AS, a_d \neq a_j} E\left[CF_{a_d}\right] E\left[CF_{a_j}\right] \cdot Pr\left(a_d, a_j\right)$. Hence, we need to determine the probabilities $Pr\left(a_d, a_j\right)$. According to expression (15) the process paths and the respective path probabilities need to be calculated. For example there are the process paths

pp_1: $a_1, a_2^{(1)}, a_3^{(1)}$; pp_2: $a_1, a_2^{(1)}, a_3^{(1)}, a_4^{(1)}, a_2^{(2)}, a_3^{(2)}$;

pp_3: $a_1, a_2^{(1)}, a_3^{(1)}, a_4^{(1)}, a_2^{(2)}, a_3^{(2)}, a_4^{(2)}, a_2^{(3)}, a_3^{(3)}$;

pp_4: $a_1, a_2^{(1)}, a_3^{(1)}, a_4^{(1)}, a_2^{(2)}, a_3^{(2)}, a_4^{(2)}, a_2^{(3)}, a_3^{(3)}, a_4^{(3)}, a_2^{(4)}, a_3^{(4)}$, and

pp_5: $a_1, a_2^{(1)}, a_3^{(1)}, a_4^{(1)}, a_2^{(2)}, a_3^{(2)}, a_4^{(2)}, a_2^{(3)}, a_3^{(3)}, a_4^{(3)}, a_2^{(4)}, a_3^{(4)}, a_4^{(4)}, a_2^{(5)}, a_3^{(5)}$,

with $p_1 = 0.9$; $p_2 = 0.09$; $p_3 = 0.009$; $p_4 = 0.0009$, and $p_5 = 0.00009$. Considering this five paths Table 5 shows the probabilities $Pr\left(a_d, a_j\right)$. For example, the cell in row $a_2^{(1)}$ and column a_1 gives $Pr\left(a_2^{(1)}, a_1\right)$. Due to the fact that $Pr\left(a_d, a_j\right) = Pr\left(a_j, a_d\right)$ it is enough to determine values of the lower triangular table. Since it is $a_d \neq a_j$ in expression (21) and $Pr(a_d, a_d) = Pr(a_d)$ the values on the diagonal do not need to be determined. The process has potentially an infinite number of paths, which means that this table does not contain all relevant probabilities. However, it displays the structure how the values change, which makes it easy to consider all probabilities $Pr\left(a_d, a_j\right)$.

	a_1	$a_2^{(1)}$	$a_2^{(2)}$	$a_2^{(3)}$	$a_2^{(4)}$	$a_2^{(5)}$	$a_3^{(1)}$	$a_3^{(2)}$	$a_3^{(3)}$	$a_3^{(4)}$	$a_3^{(5)}$	$a_4^{(1)}$	$a_4^{(2)}$	$a_4^{(3)}$	$a_4^{(4)}$
a_1															
$a_2^{(1)}$	1,0000														
$a_2^{(2)}$	0,1000	0,1000													
$a_2^{(3)}$	0,0100	0,0100	0,0100												
$a_2^{(4)}$	0,0010	0,0010	0,0010	0,0010											
$a_2^{(5)}$	0,0001	0,0001	0,0001	0,0001	0,0001										
$a_3^{(1)}$	1,0000	1,0000	0,1000	0,0100	0,0010	0,0001									
$a_3^{(2)}$	0,1000	0,1000	0,1000	0,0100	0,0010	0,0001	0,1000								
$a_3^{(3)}$	0,0100	0,0100	0,0100	0,0100	0,0010	0,0001	0,0100	0,0100							
$a_3^{(4)}$	0,0010	0,0010	0,0010	0,0010	0,0010	0,0001	0,0010	0,0010	0,0010						
$a_3^{(5)}$	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001					
$a_4^{(1)}$	0,1000	0,1000	0,1000	0,0100	0,0010	0,0001	0,1000	0,1000	0,0100	0,0010	0,0001				
$a_4^{(2)}$	0,0100	0,0100	0,0100	0,0100	0,0010	0,0001	0,0100	0,0100	0,0100	0,0010	0,0001	0,0100			
$a_4^{(3)}$	0,0010	0,0010	0,0010	0,0010	0,0010	0,0001	0,0010	0,0010	0,0010	0,0010	0,0001	0,0010	0,0010		
$a_4^{(4)}$	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	

Table 5. Probabilities $\boldsymbol{Pr(a_d, a_j)}$ in Process PR

In Table 5, the values $Pr(a_d, a_j)$ for the same actions a_d and a_j are encircled. For example, the values in the cells of rows $a_3^{(1)}$ to $a_3^{(5)}$ and column $a_2^{(1)}$ to $a_2^{(5)}$contain the values for $Pr(a_d, a_j)$ considering the appearance of the actions a_2 and a_3 in the process paths pp_1 to pp_5. All of these values have to be considered when calculating $E[CF_{a_d}]E\left[CF_{a_j}\right] \cdot Pr(a_d, a_j)$ in expression (21) for the actions a_2 and a_3. The different colors show areas with the same structure of the values, to know how to use the formula for a geometric series. With this it is possible to determine $\sum_{a_d, a_j \in AS, a_d \neq a_j} E[CF_{a_d}]E\left[CF_{a_j}\right] \cdot Pr(a_d, a_j)$ in expression (21).

Overall it is

$$\sum_{a_d,a_j\in AS,a_d\neq a_j} E[CF_{a_d}]E\left[CF_{a_j}\right]\cdot Pr(a_d,a_j)$$

$$= \underbrace{2}_{due\ to\ Pr(a_d,a_j)=Pr(a_j,a_d)} \cdot \left[\underbrace{\sum_{i=1}^{\infty} E\left[CF_{a_2^{(i)}}\right]E[CF_{a_1}]\cdot Pr\left(a_2^{(i)},a_1\right)}_{grey\ dashed}\right.$$

$$+\underbrace{\sum_{i=2}^{\infty}\sum_{j=1}^{i-1} E\left[CF_{a_2^{(i)}}\right]E\left[CF_{a_2^{(j)}}\right]\cdot Pr\left(a_2^{(i)},a_2^{(j)}\right)}_{dark\ grey}$$

$$+\underbrace{\sum_{i=1}^{\infty} E\left[CF_{a_3^{(i)}}\right]E[CF_{a_1}]\cdot Pr\left(a_3^{(i)},a_1\right)}_{grey\ dashed}+\underbrace{\sum_{i=1}^{\infty} E\left[CF_{a_3^{(i)}}\right]E\left[CF_{a_2^{(1)}}\right]\cdot Pr\left(a_3^{(i)},a_2^{(1)}\right)}_{grey\ dashed}$$

$$+\underbrace{\sum_{i=2}^{\infty}\sum_{j=2}^{i} E\left[CF_{a_3^{(i)}}\right]E\left[CF_{a_2^{(j)}}\right]\cdot Pr\left(a_3^{(i)},a_2^{(j)}\right)}_{dark\ grey}$$

$$+\underbrace{\sum_{i=1}^{\infty}\sum_{j=i+1}^{\infty} E\left[CF_{a_3^{(i)}}\right]E\left[CF_{a_2^{(j)}}\right]\cdot Pr\left(a_3^{(i)},a_2^{(j)}\right)}_{black}$$

$$+\underbrace{\sum_{i=2}^{\infty}\sum_{j=1}^{i-1} E\left[CF_{a_3^{(i)}}\right]E\left[CF_{a_3^{(j)}}\right]\cdot Pr\left(a_3^{(i)},a_3^{(j)}\right)}_{dark\ grey}$$

$$+\underbrace{\sum_{i=1}^{\infty} E\left[CF_{a_4^{(i)}}\right]E[CF_{a_1}]\cdot Pr\left(a_4^{(i)},a_1\right)}_{white}+\underbrace{\sum_{i=1}^{\infty} E\left[CF_{a_4^{(i)}}\right]E\left[CF_{a_2^{(1)}}\right]\cdot Pr\left(a_4^{(i)},a_2^{(1)}\right)}_{white}$$

$$+\underbrace{\sum_{i=1}^{\infty} E\left[CF_{a_4^{(i)}}\right]E\left[CF_{a_2^{(2)}}\right]\cdot Pr\left(a_4^{(i)},a_2^{(2)}\right)}_{white}+\underbrace{\sum_{i=2}^{\infty}\sum_{j=3}^{i+1} E\left[CF_{a_4^{(i)}}\right]E\left[CF_{a_2^{(j)}}\right]\cdot Pr\left(a_4^{(i)},a_2^{(j)}\right)}_{midium\ dark\ grey}$$

$$+\underbrace{\sum_{i=1}^{\infty}\sum_{j=i+2}^{\infty} E\left[CF_{a_4^{(i)}}\right]E\left[CF_{a_2^{(j)}}\right]\cdot Pr\left(a_4^{(i)},a_2^{(j)}\right)}_{light\ grey}+\underbrace{\sum_{i=1}^{\infty} E\left[CF_{a_4^{(i)}}\right]E\left[CF_{a_3^{(1)}}\right]\cdot Pr\left(a_4^{(i)},a_3^{(1)}\right)}_{white}$$

$$+\underbrace{\sum_{i=1}^{\infty} E\left[CF_{a_4^{(i)}}\right]E\left[CF_{a_3^{(2)}}\right]\cdot Pr\left(a_4^{(i)},a_3^{(2)}\right)}_{white}+\underbrace{\sum_{i=2}^{\infty}\sum_{j=3}^{i+1} E\left[CF_{a_4^{(i)}}\right]E\left[CF_{a_3^{(j)}}\right]\cdot Pr\left(a_4^{(i)},a_3^{(j)}\right)}_{midium\ dark\ grey}$$

$$\left.+\underbrace{\sum_{i=1}^{\infty}\sum_{j=i+2}^{\infty} E\left[CF_{a_4^{(i)}}\right]E\left[CF_{a_3^{(j)}}\right]\cdot Pr\left(a_4^{(i)},a_3^{(j)}\right)}_{light\ grey}+\underbrace{\sum_{i=2}^{\infty}\sum_{j=1}^{i-1} E\left[CF_{a_4^{(i)}}\right]E\left[CF_{a_4^{(j)}}\right]\cdot Pr\left(a_4^{(i)},a_4^{(j)}\right)}_{midium\ dark\ grey}\right]$$

$$= 2 \cdot \Bigg[\underbrace{E[CF_{a_2}]E[CF_{a_1}]\sum_{i=1}^{\infty} Pr\Big(a_2^{(i)}, a_1\Big)}_{grey\ dashed} + \underbrace{E[CF_{a_2}]E[CF_{a_2}]\sum_{i=2}^{\infty}\sum_{j=1}^{i-1} Pr\left(a_2^{(i)}, a_2^{(j)}\right)}_{dark\ grey}$$

$$+ \underbrace{E[CF_{a_3}]E[CF_{a_1}]\sum_{i=1}^{\infty} Pr\Big(a_3^{(i)}, a_1\Big)}_{grey\ dashed} + \underbrace{E[CF_{a_3}]E[CF_{a_2}]\sum_{i=1}^{\infty} Pr\Big(a_3^{(i)}, a_2^{(1)}\Big)}_{grey\ dashed}$$

$$+ \underbrace{E[CF_{a_3}]E[CF_{a_2}]\sum_{i=2}^{\infty}\sum_{j=2}^{i} Pr\left(a_3^{(i)}, a_2^{(j)}\right)}_{dark\ grey}$$

$$+ \underbrace{E[CF_{a_3}]E[CF_{a_2}]\sum_{i=1}^{\infty}\sum_{j=i+1}^{\infty} Pr\left(a_3^{(i)}, a_2^{(j)}\right)}_{black}$$

$$+ \underbrace{E[CF_{a_3}]E[CF_{a_3}]\sum_{i=2}^{\infty}\sum_{j=1}^{i-1} Pr\left(a_3^{(i)}, a_3^{(j)}\right)}_{dark\ grey}$$

$$+ \underbrace{E[CF_{a_4}]E[CF_{a_1}]\sum_{i=1}^{\infty} Pr\Big(a_4^{(i)}, a_1\Big)}_{white} + \underbrace{E[CF_{a_4}]E[CF_{a_2}]\sum_{i=1}^{\infty} Pr\Big(a_4^{(i)}, a_2^{(1)}\Big)}_{white}$$

$$+ \underbrace{E[CF_{a_4}]E[CF_{a_2}]\sum_{i=1}^{\infty} Pr\Big(a_4^{(i)}, a_2^{(2)}\Big)}_{white} + \underbrace{E[CF_{a_4}]E[CF_{a_2}]\sum_{i=2}^{\infty}\sum_{j=3}^{i+1} Pr\left(a_4^{(i)}, a_2^{(j)}\right)}_{midium\ dark\ grey}$$

$$+ \underbrace{E[CF_{a_4}]E[CF_{a_2}]\sum_{i=1}^{\infty}\sum_{j=i+2}^{\infty} Pr\left(a_4^{(i)}, a_2^{(j)}\right)}_{light\ grey} + \underbrace{E[CF_{a_4}]E[CF_{a_3}]\sum_{i=1}^{\infty} Pr\Big(a_4^{(i)}, a_3^{(1)}\Big)}_{white}$$

$$+ \underbrace{E[CF_{a_4}]E[CF_{a_3}]\sum_{i=1}^{\infty} Pr\Big(a_4^{(i)}, a_3^{(2)}\Big)}_{white} + \underbrace{E[CF_{a_4}]E[CF_{a_3}]\sum_{i=2}^{\infty}\sum_{j=3}^{i+1} Pr\left(a_4^{(i)}, a_3^{(j)}\right)}_{midium\ dark\ grey}$$

$$+ \underbrace{E[CF_{a_4}]E[CF_{a_3}]\sum_{i=1}^{\infty}\sum_{j=i+2}^{\infty} Pr\left(a_4^{(i)}, a_3^{(j)}\right)}_{light\ grey} + \underbrace{E[CF_{a_4}]E[CF_{a_4}]\sum_{i=2}^{\infty}\sum_{j=1}^{i-1} Pr\left(a_4^{(i)}, a_4^{(j)}\right)}_{midium\ dark\ grey} \Bigg]$$

$$= 2 \cdot \Bigg[\underbrace{E[CF_{a_2}]E[CF_{a_1}]\sum_{i=1}^{\infty} Pr\Big(a_2^{(i)}, a_1\Big)}_{grey\ dashed} + \underbrace{E[CF_{a_2}]E[CF_{a_2}]\sum_{i=2}^{\infty}\sum_{j=1}^{i-1} Pr\left(a_2^{(i)}, a_2^{(j)}\right)}_{dark\ grey}$$

$$+\underbrace{E[CF_{a_3}]E[CF_{a_1}]\sum_{i=1}^{\infty}Pr\left(a_3^{(i)},a_1\right)}_{grey\ dashed}$$

$$+E[CF_{a_3}]E[CF_{a_2}]\left(\underbrace{\sum_{i=1}^{\infty}Pr\left(a_3^{(i)},a_2^{(1)}\right)}_{grey\ dashed}+\underbrace{\sum_{i=2}^{\infty}\sum_{j=2}^{i}Pr\left(a_3^{(i)},a_2^{(j)}\right)}_{dark\ grey}\right.$$

$$\left.+\underbrace{\sum_{i=1}^{\infty}\sum_{j=i+1}^{\infty}Pr\left(a_3^{(i)},a_2^{(j)}\right)}_{black}\right)+\underbrace{E[CF_{a_3}]E[CF_{a_3}]\sum_{i=2}^{\infty}\sum_{j=1}^{i-1}Pr\left(a_3^{(i)},a_3^{(j)}\right)}_{dark\ grey}$$

$$+\underbrace{E[CF_{a_4}]E[CF_{a_1}]\sum_{i=1}^{\infty}Pr\left(a_4^{(i)},a_1\right)}_{white}$$

$$+E[CF_{a_4}]E[CF_{a_2}]\left(\underbrace{\sum_{i=1}^{\infty}Pr\left(a_4^{(i)},a_2^{(1)}\right)}_{white}+\underbrace{\sum_{i=1}^{\infty}Pr\left(a_4^{(i)},a_2^{(2)}\right)}_{white}\right.$$

$$\left.+\underbrace{\sum_{i=2}^{\infty}\sum_{j=3}^{i+1}Pr\left(a_4^{(i)},a_2^{(j)}\right)}_{midium\ dark\ grey}+\underbrace{\sum_{i=1}^{\infty}\sum_{j=i+2}^{\infty}Pr\left(a_4^{(i)},a_2^{(j)}\right)}_{light\ grey}\right)$$

$$+E[CF_{a_4}]E[CF_{a_3}]\left(\underbrace{\sum_{i=1}^{\infty}Pr\left(a_4^{(i)},a_3^{(1)}\right)}_{white}+\underbrace{\sum_{i=1}^{\infty}Pr\left(a_4^{(i)},a_3^{(2)}\right)}_{white}\right.$$

$$+\underbrace{\sum_{i=2}^{\infty}\sum_{j=3}^{i+1}Pr\left(a_4^{(i)},a_3^{(j)}\right)}_{midium\ dark\ grey}$$

$$\left.\left.+\underbrace{\sum_{i=1}^{\infty}\sum_{j=i+2}^{\infty}Pr\left(a_4^{(i)},a_3^{(j)}\right)}_{light\ grey}\right)+\underbrace{E[CF_{a_4}]E[CF_{a_4}]\sum_{i=2}^{\infty}\sum_{j=1}^{i-1}Pr\left(a_4^{(i)},a_4^{(j)}\right)}_{midium\ dark\ grey}\right]$$

$$=2\cdot\left[\underbrace{E[CF_{a_2}]E[CF_{a_1}]\cdot\frac{10}{9}}_{grey\ dashed}+\underbrace{E[CF_{a_2}]E[CF_{a_2}]\cdot\frac{1}{9}\underbrace{\sum_{i=0}^{\infty}0.1^i}_{\frac{10}{9}}}_{dark\ grey}\right.$$

$$+\underbrace{E\left[CF_{a_3}\right]E\left[CF_{a_1}\right]\cdot\frac{10}{9}}_{grey\ dashed}+E\left[CF_{a_3}\right]E\left[CF_{a_2}\right]\left(\underbrace{\frac{10}{9}}_{grey\ dashed}+\underbrace{\frac{1}{9}\sum_{i=0}^{\infty}0.1^{i}}_{dark\ grey}+\underbrace{\frac{1}{9}\sum_{i=0}^{\infty}0.1^{i}}_{black}\right)$$

$$+\underbrace{E\left[CF_{a_3}\right]E\left[CF_{a_3}\right]\cdot\frac{1}{9}\sum_{i=0}^{\infty}0.1^{i}}_{dark\ grey}$$

$$+\underbrace{E\left[CF_{a_4}\right]E\left[CF_{a_1}\right]\cdot\frac{1}{9}}_{white}+E\left[CF_{a_4}\right]E\left[CF_{a_2}\right]\left(\underbrace{\frac{1}{9}}_{white}+\underbrace{\frac{1}{9}}_{white}+\underbrace{\frac{1}{90}\sum_{i=0}^{\infty}0.1^{i}}_{midium\ dark\ grey}+\underbrace{\frac{1}{90}\sum_{i=0}^{\infty}0.1^{i}}_{light\ grey}\right)$$

$$+E\left[CF_{a_4}\right]E\left[CF_{a_3}\right]\left(\underbrace{\frac{1}{9}}_{white}+\underbrace{\frac{1}{9}}_{white}+\underbrace{\frac{1}{90}\sum_{i=0}^{\infty}0.1^{i}}_{midium\ dark\ grey}\right.$$

$$\left.+\underbrace{\frac{1}{90}\sum_{i=0}^{\infty}0.1^{i}}_{light\ grey}\right)+\underbrace{E\left[CF_{a_4}\right]E\left[CF_{a_4}\right]\cdot\frac{1}{90}\sum_{i=0}^{\infty}0.1^{i}}_{midium\ dark\ grey}\Bigg]$$

$$=2\cdot\Bigg[\underbrace{500\cdot 1{,}000\cdot\frac{10}{9}}_{grey\ dashed}+\underbrace{500\cdot 500\cdot\frac{10}{81}}_{dark\ grey}$$

$$+\underbrace{500\cdot 1{,}000\cdot\frac{10}{9}}_{grey\ dashed}+500\cdot 500\left(\underbrace{\frac{10}{9}}_{grey\ dashed}+\underbrace{\frac{10}{81}}_{dark\ grey}+\underbrace{\frac{10}{81}}_{black}\right)+\underbrace{500\cdot 500\cdot\frac{10}{81}}_{dark\ grey}$$

$$+\underbrace{5{,}000\cdot 1{,}000\cdot\frac{1}{9}}_{white}+5{,}000\cdot 500\left(\underbrace{\frac{1}{9}}_{white}+\underbrace{\frac{1}{9}}_{white}+\underbrace{\frac{1}{81}}_{midium\ dark\ grey}+\underbrace{\frac{1}{81}}_{light\ grey}\right)+5{,}000$$

$$\cdot 500\left(\underbrace{\frac{1}{9}}_{white}+\underbrace{\frac{1}{9}}_{white}+\underbrace{\frac{1}{81}}_{midium\ dark\ grey}+\underbrace{\frac{1}{81}}_{light\ grey}\right)+\underbrace{5{,}000\cdot 5{,}000\cdot\frac{1}{81}}_{midium\ dark\ grey}\Bigg]$$

$$=2\cdot\Bigg[500\cdot 1{,}000\cdot\frac{10}{9}+500\cdot 500\cdot\frac{10}{81}$$

$$+500\cdot 1{,}000\cdot\frac{10}{9}+500\cdot 500\cdot\frac{110}{81}+500\cdot 500\cdot\frac{10}{81}$$

$$+5{,}000\cdot 1{,}000\cdot\frac{1}{9}+5{,}000\cdot 500\cdot\frac{20}{81}+5{,}000\cdot 500\cdot\frac{20}{81}+5{,}000\cdot 5{,}000\cdot\frac{1}{81}\Bigg]$$

II.2.9 References Appendix

Feller W (1971) An Introduction to Probability Theory and Its Applications, Vol. 2. John Wiley & Sons, Inc., New York, NY, USA

Gibbons A (1985) Algorithmic Graph Theory. Cambridge University Press, Melbourne, Australia

Hollingsworth DC, WfMC (2003) The WfMC Glossary. In: Fischer L (ed) Workflow Handbook 2003. Future Strategies Inc., Lighthouse Point, Florida, USA, pp 263-299

Hubbard DW (2007) How to Measure Anything: Finding the Value of Intangibles in Business. John Wiley & Sons, Inc., New York, NY, USA

Rubin V, Günther CW, van der Aalst WMP, Kindler E, van Dongen BF, Schäfer W (2007) Process mining framework for software processes. In: Wang Q, Pfahl D, Raffo DM (eds) Proceedings of the International Conference on Software Process: Software Process Dynamics and Agility (ICSP), Minneapolis, MN, USA, pp 169-181

van der Aalst WMP, ter Hofstede AH, Kiepuszewski B, Barros AP (2003) Workflow patterns. Distrib Parallel Dat 14(3):5-51

zur Muehlen M, Shapiro R (2010) Business Process Analytics. In: vom Brocke J, Rosemann M (eds) Handbook on Business Process Management 2: Strategic Alignment, Governance, People and Culture. Springer, Berlin, Germany, pp 137-158

II.3 Research Paper 3: "Process Improvement through Economically Driven Routing of Instances"

Authors:	Manuel Bolsinger[a], Anna Elsäßer[a], Caroline Helm[a], Maximilian Röglinger[a] [a] FIM Research Center, Department of Information Systems Engineering & Financial Management (Prof. Dr. Hans Ulrich Buhl), University of Augsburg, Germany manuel.bolsinger@fim-rc.de anna.elsaesser@fim-rc.de caroline.helm@fim-rc.de maximilian.roeglinger@fim-rc.de
Resubmitted (major revision) to:	Business Process Management Journal

Structured Abstract:

Purpose *Process improvement is a fundamental activity of the BPM lifecycle. However, practitioners still lack concrete guidance and adequate objectives for process improvement. Moreover, improvement projects typically tie up considerable amounts of capital and are very risky. Thus, more guidance is needed on how to derive concrete recommendations for process improvement in a goal-oriented manner, which do not require huge investments.*

Design/methodology/approach *We propose a decision model that determines along which paths the instances of a process should be routed to maximize the value contribution of the process. To do so, the decision model requires a process model and a set of historical process instances as inputs.*

Findings *The decision model builds on the idea that only the parameters of the process, i.e., the values according to which it is decided on which path an instance traverses the process, can be modified, without altering the structure of the process. The decision model determines the parameter setting that maximizes the value contribution of the process, which is measured in terms of expected cash flow of the process. When determining the optimal parameter setting, the decision model considers that different instances and paths have different cash flow effects.*

Practical implications *We prototypically implemented the decision model and report on the insights from a demonstration example that is based on the order verification process of an IT distributor.*

Originality/value *The decision model complements existing approaches to process improvement as it reveals additional improvement potential by focusing on the decision points in a process without altering the structure of the process. The decision model also enables identifying an optimal parameter setting, as a concrete recommendation for process improvement, in line with the principles of value-based BPM.*

II.3.1 Introduction

Process orientation is a widely adopted paradigm of organizational design and proved to be a source of strong corporate performance (Skrinjar et al., 2008, Kohlbacher and Reijers, 2013, Hammer, 2010). Thus, process improvement is a fundamental activity of the business process management (BPM) lifecycle (Zellner, 2011, Dumas et al., 2013). Both the increasing number of scientific publications and the high priority on CIO agendas are indicative of the attention that process improvement receives from scholars and practitioners alike (Sidorova and Isik, 2010, Luftman et al., 2012). However, despite a long tradition and an extensive body of knowledge, approaches to process improvement still are in high demand (van der Aalst, 2013).

In the literature, approaches to process analysis and improvement are divided into data- and model-based approaches (van der Aalst, 2013). Model-based approaches can be further structured based on the type of process model that they rely on (Vergidis et al., 2008). The most prominent class of model-based approaches builds on diagrammatic process models and the experience of process experts. Despite their popularity, approaches that belong to this class have been criticized for providing little guidance, being susceptible to subjective biases, concealing how improvement ideas are generated, and requiring a great amount of manual effort (Reijers and Liman Mansar, 2005, Zellner, 2011, Vergidis et al., 2008). Moreover, although there are different means of integrating objectives into process models, the effect of improvement ideas is typically estimated in terms of gut feeling, plausibility considerations, or qualitative criteria (Buhl et al., 2011, Neiger et al., 2006). To reduce subjective influences and to make the derivation of improvement ideas more transparent, researchers began collecting best practices and inferring reusable patterns (Reijers and Liman Mansar, 2007, Zellner, 2013). Nevertheless, there are hardly any approaches that provide concrete ideas for process redesign (van der Aalst, 2013). To mitigate the required manual effort and to provide more substantial guidance, another class of model-based improvement approaches is considered useful—the class of approaches that build on mathematical process models. Such approaches allow for automated process validation, verification, and quantitative optimization regarding the extent to which process improvement ideas align with predefined objectives (Hofacker and Vetschera, 2001). Approaches based on mathematical process models, however, present their own drawbacks because they are typically geared toward a restricted set of application domains and sometimes apply to sequential processes only (Vergidis et al., 2008). Moreover, the objectives used for quantitative optimization usually refer to technical criteria, such as quality, availability or cycle time, that are not necessarily aligned with economic objectives (Buhl et al., 2011, vom Brocke et al., 2010). Under the notion of value-based BPM, some recent approaches provide economically well-founded decision support by assessing the impact of process improvement in terms of its contribution to the company value. Most of these approaches, however, abstract from the control flow perspective (e.g. Afflerbach et al., 2013, Forstner et al., 2014). Other approaches economically analyze the control flow of a process and help compare alternative process designs, but do not provide recommendations for improvement (e.g. Bolsinger, 2014, vom Brocke et al., 2010).

In sum, there are the following research gaps: (1) practitioners who set out to improve a process still lack guidance, (2) improvement approaches that build on mathematical process models and could provide guidance are very complex and restricted with regard to their applicability, and (3) improvement ideas are typically evaluated using non-economic objectives whereas economically well-founded approaches do not provide recommendations on a control flow level of detail. Moreover, the majority of improvement projects that result from the approaches mentioned above require considerable re-engineering efforts, i.e., they tie up large amounts of capital and are very risky (Devaraj and Kohli, 2002). Therefore, companies should not only conduct extensive re-engineering projects, but also strive for continuous improvement (Trkman, 2010). In this paper, we take on a continuous improvement perspective and address the following research question derived from the research gaps: *How can concrete recommendations for process improvement be derived that do not require extensive re-engineering projects and align with economic objectives?*

As an answer to this question, we propose a decision model that complements existing process improvement approaches by focusing on the routing decisions that are made in a process model, without altering the structure of the process. Extending the previous work on value-based BPM, the decision model determines along which paths the instances of a process should be routed to maximize the value contribution of the process. The value contribution of a process is measured in terms of its expected cash flow, a figure that complies with the paradigm of value-based BPM. As the decision model takes a process model and a set of historical instances as input, it bridges the gap between the data- and model-based process improvement approaches. It also addresses the "Improve model" BPM use case as well as the interface between BPM and operations management (vom Brocke et al., 2011, van der Aalst, 2013).

We proceed as follows: First, we sketch the theoretical background related to process design and value-based BPM. We then present the decision model and report on a demonstration example that is based on the order verification process of an IT distributor. We conclude by summarizing the study's results, discussing its limitations, and suggesting topics for further research.

II.3.2 Theoretical background

II.3.2.1 Foundations of business process design

Processes can be categorized into core, support, and management processes (Harmon, 2010). Core processes are "collection[s] of inter-related events, activities, and decision points that involve a number of actors and objects, and that collectively lead to an outcome that is of value to at least one customer" (Dumas et al., 2013, p. 5). This recent definition, which explicitly refers to decision points, is in line with other widely-adopted definitions such as those proposed by Hammer and Champy (1993) or Davenport (1993). Support processes ensure that the core processes continue to function. Management processes entail the planning, organization, communication, monitoring, and controlling of activities. We focus on core processes and refer to them as processes.

Process models play a fundamental role in relation to improvement activities because they are the form in which processes are typically documented (Recker, 2007). Therefore, process modeling has been subject to extensive research. Numerous modeling languages are available to create process models for different purposes (Recker et al., 2009). Vergidis et al. (2008) distinguish diagrammatic, mathematical, and execution-oriented process models. Diagrammatic models help visualize and communicate processes. Mathematical models allow for formal validation, verification, and optimization. Execution-oriented models enable automated process execution. Popular languages include BPEL, BPMN, EPCs, Petri Nets, UML activity diagrams, or YAWL. The classification proposed by Vergidis et al. (2008) is not disjoint, i.e., languages may refer to more than one model type (e.g., Petri Nets). In this paper, we use activity diagrams from the UML specification to visualize processes (Object Management Group, 2011). As input for the decision model, we transform activity diagrams into mathematical process models based on graph theory.

From the graph theory perspective, process models are directed graphs that, in a very general sense, consist of nodes and edges (Biswal, 2005). Nodes refer to events, activities, and decision points. They can be annotated with information such as cycle times, financials, or required resources. The integration of financial information into process models provides the foundation for an economic valuation of improvement ideas and process design alternatives (vom Brocke et al., 2010). Edges, modeled as a relation on the set of nodes, indicate how nodes are connected. Just like nodes, edges can be annotated with information such as conditions and transition probabilities. Transition probabilities indicate the probability of getting from one node to another directly linked node (Mitzenmacher and Upfal, 2005). We define a sequence of directly linked nodes as a path. The nodes that represent decision points (decision nodes) are of particular interest when analyzing the routing of instances because they define the paths of a process. The transition probability of an edge that connects a decision node with a subsequent node is typically smaller than 100%. In the decision model, we rely on paths, transition probabilities, and additional information annotated to nodes.

II.3.2.2 Value orientation in business process management

Value-based BPM is a paradigm in which all process-related activities and decisions are valued according to their contribution to the company value. As such, value-based BPM applies the principles of value-based management to process decision-making. Because value-based BPM has been evolving into an established paradigm of process decision-making, we use it as a foundation for deriving the objective function of the decision model (Buhl et al., 2011, Kirchmer, 2008).

Value-based management, as a substantiation and extension of the shareholder value concept, sets the maximization of the company value as the primary objective of all business activities (Koller et al., 2010). The company value is determined based on future cash flow (Rappaport, 1986). Value-based management builds on the work of Rappaport (1986), Copeland et al. (1990), and Stewart and Stern (1991). It can only be claimed to be implemented if all corporate activities are aligned with the objective of maximizing the company value. Therefore, companies must not only be able to quantify the company value on the corporate level, but also the value contribution of individual activities or decisions. To comply with the principles of value-based management, decisions must be based on cash flow, consider risks, and incorporate the time-value of money (Buhl et al., 2011).

With regard to value-based BPM, annotating cash flow to the elements of a process model is an essential preparatory step (vom Brocke et al., 2010). Analogous to activity-based costing, outflows are allocated to the actions of a process whenever possible. Outflows that cannot be reasonably allocated to actions are attributed to the process itself. The outflows caused by an action may be estimated based on the actual use of equipment and resources such as IT services or people. Depending on the extent to which an action is automated and how many resources are available, the outflows may vary according to the workload (Braunwarth et al., 2010). Just like the second group of outflows, the inflows of a process are linked with the process itself and not with single actions.

II.3.3 Decision model

II.3.3.1 Basic idea

The model of a process encompasses all paths that an instance may take (van der Aalst, 2013). Which path an instance takes depends on the values that this instance takes for decision-relevant process characteristics (e.g. rating index) and on how the related decision nodes are parameterized, i.e., how it is determined which outgoing edge an instance takes. Figure 1 illustrates this interplay for a single decision node. In our case, all decision nodes have two outgoing edges, each annotated with a condition. Considering a single decision node, the conditions of both outgoing edges complement each other such that they split the domain of the related process characteristic into two disjoint intervals (e.g. rating index ≤ 100 and rating index > 100). The value according to which the domain is split is called the parameter of the decision node (100 in the example of the rating index). That is, if an instance reaches a decision node and the value that this instance takes for the related process characteristic is smaller than or equal to the parameter of this decision node, the instance continues with one of the two outgoing edges. Otherwise, the instance continues with the other outgoing edge. The path an instance takes through the process model results from the values that the instance takes for all decision-relevant process characteristics as well as on all parameters, i.e., the parameter setting. What is interesting from a decision-making perspective is that the optimal parameter for a decision node cannot be determined in isolation, but depends, among others, on all parameters located upstream in the process, the control flow of the process, and the cash flow effects of all instances.

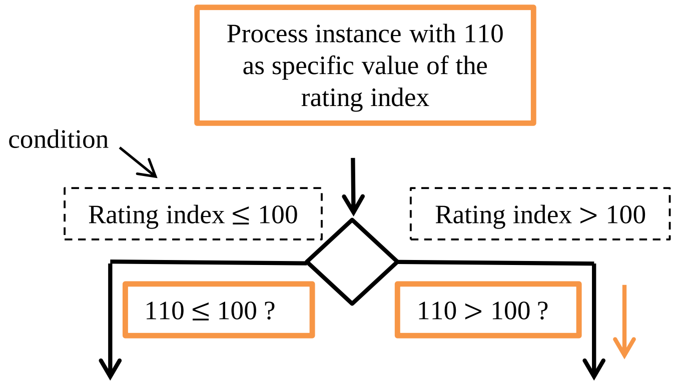


Figure 1. Interplay of decision nodes, process characteristics, and parameters.

To improve the process under investigation, the decision model allows the parameters to be modified without changing the structure of the process. The decision model determines the parameter setting that maximizes the value contribution of the process in terms of its expected cash flow. Therefore, it also takes a representative set of historical instances as input and identifies the parameter setting that would have maximized the cash flow of the historical instances. When estimating the expected cash flow, the decision model takes into account that different instances and paths have different cash flow effects. For example, each instance causes different inflows and the outflows for executing an action depend on the overall workload of this action. Moreover, the parameter setting influences with which probability process risks for a specific instance occur.

II.3.3.2 General setting and objective function

The process model has to be available as a graph $G = (V, E)$, where V and E denote the set of nodes and edges, respectively. There are numerous paths $z \in Z$ through the process that depend on how the decision nodes $V^{\mathrm{DN}} \subset V$ of the process model are interconnected. We make the following assumption regarding the process model:

(A1) The process model has a single starting point (e.g. an initial node or a signal accept state) as well as multiple final, action, and decision (merge) nodes. Each decision node $k \in V^{\mathrm{DN}}$ relates to a metrically scaled process characteristic X_k and has two outgoing edges. Each outgoing edge has a condition that connects the process characteristic with a metrically scaled parameter y_k. We consider only those decision nodes in V^{DN} whose parameter can be chosen freely. The paths of the process split into completion and cancellation paths. On completion paths, the outcome of the process is successfully sold to a customer. On cancellation paths, the process is deliberately terminated by the company before its outcome is sold.

The restriction toward metrically scaled process characteristics and parameters implies that yes/no decisions and decisions based on nominally scaled process characteristics are excluded. The reason is that, for such decisions, the path an instance takes only depends on the properties of the instance and cannot be influenced from the outside. However, assumption (A1) is less restrictive than it seems at first glance as numerous process models can be transformed such that they comply with assumption (A1). For example, parallelizations can often be treated as a single action from a valuation perspective. Loops can be transformed into a finite set of paths if the maximum amount of repetitions is known. We are able to deal with decisions based on ordinally scaled process characteristics because the values of such characteristics can be transformed into scalars without affecting the underpinning preference relation. Finally, we also cover decision nodes with more than two outgoing edges as such decision nodes can be transformed into several consecutive decision nodes with two outgoing edges each.

The decision model also uses a representative set of historical instances I as a reference for economic analysis. Such a set may have been extracted from process event logs (van der Aalst, 2011). Each instance takes an arbitrary value for the process characteristics X_k. Thus, the process characteristics are random variables. More precisely, because the parameter setting $\vec{y}$ determines which instances take which path, the process characteristics follow conditional distributions. All instances that take a completion path for a given parameter setting $\vec{y}$ are referred to as completed instances and included in the set $I_{\vec{y}}^{\text{comp}} \subseteq I$. All other instances are called cancelled instances.

To identify the optimal parameter setting in line with the paradigm of value-based BPM, we make the following assumption regarding the objective function:

(A2) Decision makers aim to maximize the value contribution of the process. The value contribution is measured in terms of the expected cash flow, which splits into expected inflows and outflows and depends on the parameter setting $\vec{y}$. The objective function is given in Formula (1).

$$\text{MAX: } CF(\vec{y}) = Inflows(\vec{y}) - Outflows(\vec{y}) \tag{1}$$

II.3.3.3 Concretization of the objective function

II.3.3.3.1. Cash outflows

As for the cash outflows, we distinguish between outflows that are associated with the actions of the process (e.g. outflows for resource consumption, IT services, and people) and outflows that cannot be reasonably assigned to an action (e.g. outflows for the procurement of trading goods). In this section, we model the outflows that are associated with the actions of the process. We deal with the other group of outflows by multiplying the inflows with an instance-specific profit margin λ_i (Robinson et al., 2012).

The expected outflows of the process can be decomposed in three steps: First, the expected outflows on the process level equal the sum of the expected outflows per path $CF_z^{\text{out,path}}$. Second, the expected outflows of a distinct path result from multiplying the expected outflows of all actions on that path with the number of all instances $N_{z,\vec{y}}^{\text{path}}$ that take this path for a given parameter setting $\vec{y}$. Third, the expected outflows of an action $CF_k^{\text{out,action}}$ vary with the number of instances $N_{k,\vec{y}}^{\text{action}}$ that pass this action across all paths for a given parameter setting $\vec{y}$. The expected outflows of an action occur each time this action is executed. The second and third decomposition steps are shown in Formula (2).

$$CF_z^{\text{out,path}} = \sum_{(k,\cdot)\in E_z^{\text{path}}} CF_k^{\text{out,action}}\left(N_{k,\vec{y}}^{\text{action}}\right) \cdot N_{z,\vec{y}}^{\text{path}} \tag{2}$$

where $E_z^{\text{path}} \subset E$ is the set of all edges on path $z \in Z$

As indicated in the theoretical background, the extent to which the expected outflows of a distinct action vary with the workload of that action depends, among others, on how many resources are available. We consider workload-dependent outflows and limited resources here because it is known that resources cannot be changed significantly in the short term (Betts et al., 2000). We assume:

(A3) The resources available to execute the process are fixed. The capacity of these resources is limited and cannot be changed significantly in the short term.

The number of instances $N_{z,\vec{y}}^{\text{path}}$ that take a distinct path for a given parameter setting results from multiplying the number of all instances $|I|$ with the transition probabilities s_{kw} of all edges on that path. The number of instances $N_{k,\vec{y}}^{\text{action}}$ that pass a distinct action for a given parameter setting is obtained by adding up the number of instances of those paths that include this action. To check whether a distinct path z contains a distinct action k, we use an indicator function that returns 1 if path z contains action k and 0 otherwise. This leads to Formula (3) and Formula (4).

$$N_{z,\vec{y}}^{\text{path}} = |I| \cdot \prod_{(k,w)\in E_z^{\text{path}}} s_{kw} \tag{3}$$

$$N_{k,\vec{y}}^{\text{action}} = \sum_{z\in Z} 1_{E_z^{\text{path}}}\big((k,\cdot)\big) \cdot N_{z,\vec{y}}^{\text{path}} \tag{4}$$

Analogous to the fact that the process characteristics are random variables that, in general, follow conditional distributions, the transition probabilities can be interpreted as conditional probabilities. The reason is that the transition probabilities s_{kw} do not only depend on the local parameter y_k, but also on the parameters located upstream in the process as well as on the distributions of the related process characteristics. Formula (5) provides a formalization.

$$s_{kw} = F_{kw}(y_k) = P\big(X_k \lesseqqgtr_{kw} y_k \big| \{X_a \lesseqqgtr_{ab} y_a \big| (a,b) \in E_k^{\text{DN}}\}\big) \tag{5}$$

$$\text{with } 0 \leq F_{kw}(y_k) \leq 1 \text{ and } \forall\, k \in V^{\text{DN}}\colon \textstyle\sum_{j=1}^{|V|} F_{kj}(y_k) = 1$$

The operator $\lesseqqgtr_{ab}$ returns the relational operator, i.e., ≤ or >, of the edge that connects node a with node b. The set E_k^{DN} contains the outgoing edges of all decision nodes that lie on any path to the decision node k, i.e., the relevant outgoing edges from decision nodes located upstream in the process. Overall, the expected outflows on the process level are modeled as shown in Formula (6).

$$Outflows(\vec{y}) = \sum_{z\in Z} \sum_{(k,\cdot)\in E_z^{\text{path}}} CF_k^{\text{out,action}}\big(N_{k,\vec{y}}^{\text{action}}\big) \cdot N_{z,\vec{y}}^{\text{path}} \tag{6}$$

II.3.3.3.2. Cash inflows

Positive inflows occur for completed instances only. For all cancelled instances, the inflows are 0. In practice, the inflows associated with a completed instance are beset with risks that occur with a distinct probability. For example, the inflows may fail if the customer cancels an order or is insolvent. Moreover, the inflows may be reduced after the selling transaction has been completed if the process outcome requires rework due to bad quality. As the outflows of an instance occur regardless of whether the inflow risks occur and as the probability of such risks depends at least in parts on the routing decisions made in the process, it is reasonable to consider the average amount of such risks R_i (e.g. reduction of inflows in case of a cancelled order or additional rework) and their probability of occurrence $p_{i,\vec{y}}$ when determining the optimal parameter setting.

To estimate the probability of occurrence, we use a pragmatic approach based on risk levels (Featherstone et al., 2006, Korpela et al., 2002). Thereby, we draw from the financial management knowledge base where, under the name of scoring systems, there are similar approaches that use key factors and qualitative scores to predict probabilities of default (Mays, 2001, Trueck and Rachev, 2009, West, 2000). In our approach, risk levels represent scores and process characteristics represent key factors. We define a risk level $r_{k,i}$ as a scalar that is derived from the value that a completed instance i takes for the process characteristic X_k. The mapping functions that connect the domains of the process characteristics with the risk levels must be determined outside the decision model. We assume:

(A4) Each process instance i features one risk level $r_{k,i}$ per process characteristic X_k. The probability $p_{i,\vec{y}}$ with which inflow-related process risks occur can be estimated based on risk levels as shown in Formula (7).

$$p_{i,\vec{y}} = \frac{\sum_{k \in V^{\mathrm{DN}}} r_{k,i}}{\sum_{k \in V^{\mathrm{DN}}} r_k^{\max}} \cdot \eta_i \qquad \text{with } r_{k,i} \in \{0,1,\dots,r_k^{\max}\} \text{ and } \eta_i \in [0;1] \tag{7}$$

The probability of occurrence, as modeled in Formula (7), is an adjusted average risk level. The first factor represents the relative risk level of a completed instance with respect to the maximum risk level over all process characteristics. Simplifying, the effects of the process characteristics on the probability of occurrence are assumed to be independent (Featherstone et al., 2006). We therefore sum up the risk levels $r_{k,i}$ and divide them by the sum of the highest risk level $r_k^{\max}$ per process characteristic. Process characteristics may have different highest risk levels to account for a differently strong impact on the probability of occurrence. The first factor of Formula (7) can be easily adapted to reflect different circumstances: First, if a process characteristic that does not affect the probability of occurrence, we can exclude it by assigning a risk level of 0 to all values from its domain and to the highest risk level. Second, if characteristics other than those included in the set of process characteristics influence the probability of occurrence, the set V^{DN} can be extended accordingly.

The second factor of Formula (7) allows the probability of occurrence to be adjusted based on information that is not included in the process characteristics, but that depends on the overall path an instance takes. For example, an instance may traverse the process on a completion path in which the customer has intense personal contact with company employees. Such contact may significantly improve the payment behavior of the customer and/or reduce the probability that the customer will cancel his order, two circumstances that positively influence the probability of occurrence. Thus, we use an instance-specific reduction factor η_i that depends on the path an instance *i* takes through the process.

Formula (8) shows how the expected inflows at the process level are calculated when considering the profit margin λ_i and the average amount of the inflow-related process risks R_i. Because positive inflows occur for completed instances only, we consider instances from $I_{\vec{y}}^{\text{comp}}$.

$$Inflows(\vec{y}) = \sum_{i \in I_{\vec{y}}^{\text{comp}}} \left[\lambda_i \cdot CF_i^{\text{in}} - \frac{\sum_{k \in V^{\text{DN}}} r_{k,i}}{\sum_{k \in V^{\text{DN}}} r_k^{\text{max}}} \cdot \eta_i \cdot R_i \right] \tag{8}$$

II.3.3.3.3. Objective function

Taking the expected inflows and outflows together, the expected cash flow at the process level can be determined as shown in Formula (9). This function is the objective function of the decision model.

$$\begin{aligned} MAX: CF(\vec{y}) = & \sum_{i \in I_{\vec{y}}^{\text{comp}}} \left[\lambda_i \cdot CF_i^{\text{in}} - \frac{\sum_{k \in V^{\text{DN}}} r_{k,i}}{\sum_{k \in V^{\text{DN}}} r_k^{\text{max}}} \cdot \eta_i \cdot R_i \right] \\ & - \sum_{z \in Z} \sum_{(k,\cdot) \in E_z^{\text{path}}} CF_k^{\text{out,action}}\left(N_{k,\vec{y}}^{\text{action}}\right) \cdot N_{z,\vec{y}}^{\text{path}} \end{aligned} \tag{9}$$

II.3.4 Demonstration

We now demonstrate how the decision model can be applied. To do so, we first introduce the order verification process of an IT distributor to which we apply the decision model. We then operationalize the decision model. After presenting and interpreting the optimization results, we conduct a scenario analysis with respect to the reduction factor. For reasons of confidentiality, the identity of the IT distributor will not be disclosed. Moreover, all input data had to be anonymized.

II.3.4.1 Sample process

The aim of the order verification process was to decide which incoming orders should be accepted or rejected. The process begins when an order arrives. First, the data of the customer making the order is sent to the trade credit insurance. Then, the company checks the rating index of the customer (X_2), which is already stored in the company's customer relationship management (CRM) system. If the rating index is too low, the order is rejected. This includes that, in addition to ordinary customer data, the data already transmitted to the trade credit insurance for the request is deleted. Otherwise, after getting the amount of the trade credit insurance, the company validates the difference between the order size and the guarantee amount provided by the trade credit insurance (X_5). In the sample process, it was assumed that the guarantee amount provided by the trade credit insurance would cover all negative economic effects related to the failure of a specific instance (e.g. additional inventory, depreciation, custom-built outcome). If the difference is too high, the order is rejected. Otherwise, the company checks how many orders the customer made in the past. If the number of historical orders (X_7) is sufficiently high, the order is accepted. Otherwise, the order is assigned to an internal assessor who takes a look at it and calls the customer to briefly clarify open issues. Then, the assessor may grant a company limit (X_9), i.e., a type of internal insurance, to the customer in line with the company's risk management guidelines. Depending on the height of the company limit, the order is accepted or rejected. If accepted, the reason for completion is stored in the company's CRM system and a notification is sent to the customer. Otherwise, the order is rejected.

The process model is shown in Figure 2. Each decision node is labeled with its respective process characteristic. In order to refer to distinct nodes and transition probabilities, the decision and action nodes are indexed. The index is ordered in ascending order, line-by-line, from left to right. Overall, the process has five paths, two of which are completion paths and three of which are cancellation paths.

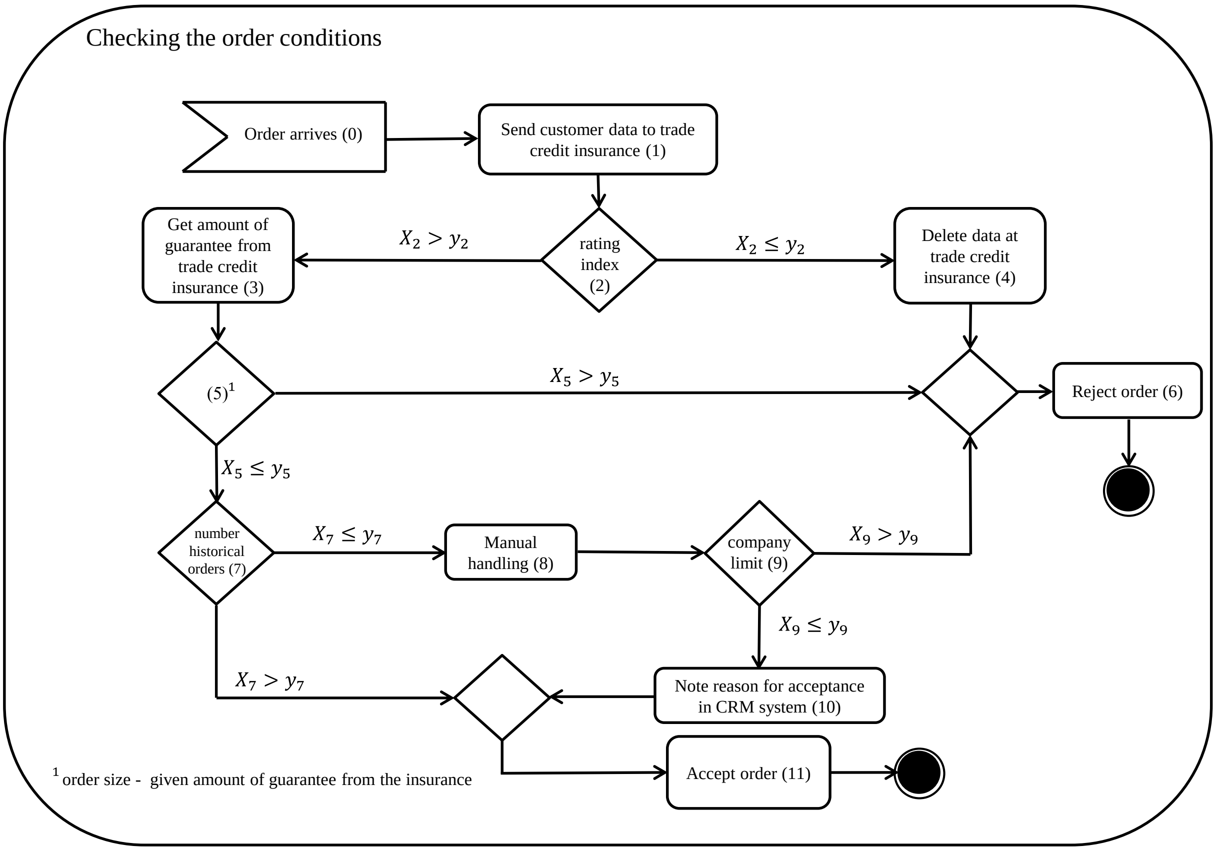


Figure 2. Sample IT distributor process for accepting and rejecting orders.

II.3.4.2 Operationalization

Before we can apply the decision model, we have to operationalize its components and collect data. Table 1 shows the data used to apply the decision model. As we had access to some parts of the required real-world data (e.g. value ranges of some process characteristics and a small set of process instances), we completed the data by generating a set of 600 process instances. We thereby considered the logical relation among the process characteristics as well as the given real-world data to derive a realistic probability distribution per process characteristic. For example, instances with a high rating index received a higher amount of guarantee.

II.3.4.2.1. Expected cash outflows per action

We use piecewise-defined functions to operationalize the expected outflows per action and to reflect that these outflows vary according to the workload associated with an action. Alternatively, we could have used continuous functions that monotonically increase with the number of instances that an action has to process. Because, to our knowledge, such functions are difficult to estimate based on real-world data, we decided in favor of piecewise-defined functions. Moreover, we distinguish three workload situations, i.e., normal, moderately excessive, and peak workloads (Braunwarth et al., 2010). Within each workload situation, the expected outflows were assumed to be constant.

Each part of an outflow function refers to a distinct workload situation. In the case of a normal workload, the resources assigned to an action can handle the amount of instances. The expected outflows only depend on how much equipment is used and the extent to which IT services and people are involved. In the case of a moderately excessive workload, the amount of instances exceeds the regular capacity but can be managed through hours of overtime and additional IT service capacity that is available in the short term. We calculate the respective expected outflows by multiplying the outflows from the normal workload situation with an overhead factor (Vanderbeck, 2013). If peak workload is reached, an action is unable to process further instances. That is, the process cannot be executed with the available resources. To avoid peak situations, we set the expected outflows to a prohibitively high value that far exceeds the outflows of the other workload situations.

Table 2 shows the capacity ranges, expected outflows, and overhead factors we used for the demonstration example. For example, the action “reject order" operates at a normal workload if less than 250 instances occur in a distinct period. In this case, an execution costs 3€. The additional workload can be handled for up to 500 instances. Each execution then costs 4.50€, which amounts to 150% of normal workload outflows. Peak workload is reached for more than 500 instances. Because such amounts cannot be processed, we set the expected outflows per execution to 1,000,000€.

II.3.4.2.2. Risk levels and reduction factor

The probability with which the inflow-related risks occur for a distinct completed instance depends on its risk levels and the reduction factor (Formula 7). In the example at hand, the most important inflow-related process risk was that the inflows completely fail because the customer cancels an order or is insolvent. Therefore, the amount of risk of a distinct completed instance was set to the inflows of that instance multiplied with the profit margin of that instance, i.e., $R_i = \lambda_i \cdot CF_i^{\text{in}}$. Moreover, all process characteristics were assumed to equally influence the probability of occurrence. We therefore decided that each risk level ranges from 1 to 10. Table 3 shows the mapping functions used here. For example, a customer with a rating index of 200 has a risk level of 10, whereas a customer with a rating index of 800 has a risk level of 3.

Regarding the reduction factor, the “manual handling” action was said to reduce the probability of occurrence by 50%, because this action is the only point in the order verification process in which the company is in touch with the customer. The personal contact was assumed to improve the relationship between the customer and the company such that the customer is much less likely to cancel the order. However, the cancellation risk is independent from the liquidity risk that is covered by the risk levels in this example. Therefore, the probability of occurrence is only reduced by 50%. Consequently, all instances that traverse the process on a completion path that includes "manual handling" have a reduction factor of 0.5. All other instances have a reduction factor of 1.

In a concrete industry setting, the mapping functions and the reduction factor can be estimated by combining process data from the previous years with the assessment of subject matter experts. The mapping functions can also be determined using process characteristics if the historical process data also contains information about which instances were cancelled.

II.3.4.3 Optimization results and interpretation

To determine the optimal parameter setting, we implemented a prototype in Microsoft Excel that was specifically geared to the order verification process. We choose Microsoft Excel because it is used in almost all organizations worldwide and because our decision model could be easily implemented in Microsoft Excel. Both properties are very well in line with our objective of providing concrete and low-threshold recommendations for process improvement. We applied a two-step solution procedure: In the first step, we conducted an exhaustive enumeration with an increment of 10 for each parameter. This increment was chosen because it was the smallest increment that returned the optimization results in reasonable time. An increment of 10 also fitted the value ranges of our process characteristics very well. With the smallest value range being about 50 for the number of historical orders (X_7), a larger increment would have been too coarse-grained. The parameter setting that maximized the expected cash flow in the first step served as an input for the second step, in which we applied the evolutionary algorithm implemented in Excel’s Solver component (Yu and Gen, 2010). Using the best parameter setting from the first step as starting point here was reasonable as the evolutionary algorithm implemented in Microsoft Excel does not necessarily return the optimal solution for a randomly chosen starting point (Rocco et al., 2000). Even with this two-step procedure, it took about five hours on a regular workstation to approximate the optimal parameter setting based on 193,000 possible parameter settings. If we had applied an exhaustive enumeration with an increment of 1, we would have had to calculate the cash flow effects of about 1.93 billion parameter settings. As results for such a problem size cannot be determined within reasonable time, we decided in favor of the two-step solution procedure.

The optimization results are shown in Table 4. The optimal parameter setting includes a rating index of 540, a 280€ difference between the order size and the guarantee amount, 18 historical orders, and a 147€ company limit. This parameter setting leads to an expected cash flow of 3,711€ and an average probability of occurrence[3] of 30%.

The results can be illustrated by looking at the conditional distributions of the process characteristics. Figure 3 shows a histogram for each process characteristic. In each histogram, we structured the domains of definition into classes. The class that contains the optimal parameter is highlighted. The distribution of the rating index (X_2) shown in Figure 3a equals the distribution shown in Table 1 because all 600 instances arrive at the related decision node. Only 403 instances reach the decision node that refers to the difference between the order size and the guarantee amount (X_5). This is reasonable because, for the optimal parameter setting, all instances with a rating index below 540 (y_2) are rejected. Thus, the conditional distribution shown in Figure 3b differs from the distribution shown in Table 1. The next decision node refers to the number of historical orders (X_7) and is reached by 341 instances. Despite this reduction in the amount of instances, the distribution shown in Figure 3c has almost the same shape as the unconditional distribution from Table 1, except that there are much fewer instances per class. Only 248 instances reach the decision node that refers to the company limit (X_9). It is remarkable that compared to the unconditional distribution shown in Table 1, the conditional distribution here contains much fewer instances in the classes referring to a lower company limit, whereas the amount of instances in the classes referring to a higher company limit are almost identical.

As can be seen when analyzing the conditional distributions, modifying a parameter can, ceteris paribus, impact the optimization results to different extents. If a parameter is shifted in the direction of those classes that contain many instances, the optimization results are much more sensitive than when a parameter is shifted in the direction of those classes that contain only a few or no instances. For example, an increase in the parameter that refers to the difference between the order size and the guarantee amount (y_5) has almost no impact on the optimization results. Reducing this parameter, however, has a strong impact on the

---

[3] The average probability of occurrence is calculated as follows: $p_{\vec{y}} = \sum_{i \in I_{\vec{y}}^{\text{comp}}} p_{i,\vec{y}} / \left| I_{\vec{y}}^{\text{comp}} \right|$.

optimization results because many more instances change their path. It may also happen that the optimization results are initially not sensitive to changes in a parameter but then suddenly change intensely. For example, this occurs if we decrease the parameter that refers to the company limit (y_9).

The optimization results have the following implications. First, the company receives a parameter setting that was determined in line with the control flow of the process under investigation as well as with the characteristics of a representative set of historical instances. Given all input data, the optimal parameter setting reflects the best solution that can be obtained for the process in terms of its expected cash flow without changing the process structure. Second, the optimal parameter setting can be easily implemented without conducting extensive improvement projects. However, one has to keep in mind that the optimal parameter setting is determined based on historical instances. The instances that occur in the next period will certainly differ from those contained in the reference set. Thus, we advise that the decision model be applied repeatedly while continuously updating the set of historical instances.

Based on this analysis, we can summarize the following insights into the behavior of the decision model: First, the impact of modifying a parameter depends on how the values that the instances take for the respective process characteristic are distributed around the current parameter value. The more instances have a value close to the current parameter value, the higher the impact of modifying the parameter. Second, the impact of modifying a parameter cannot be analyzed in isolation. Rather, there are dependencies among the parameters due to the control flow. The more decision nodes located downstream in the process, the higher the impact of modifying a parameter. Third, the impact of modifying a parameter does not only depend on the control flow and the distribution of the process characteristics, but also on the interplay between inflows and outflows, the risk levels, and the reduction factor.

Table 1. Data used to apply the decision model to the sample process.

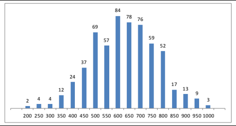
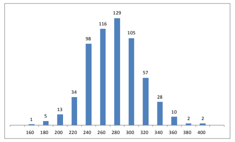
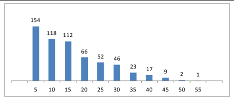
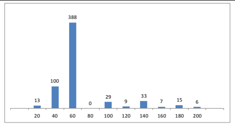
Attribute	Description	Value range and distribution (for parameters) from the set of historical process instances	Possible source(s)
Expected cash inflows (CF_i^{in})	The size of an incoming order measured in monetary units.	Minimum: 262€ Maximum: 447€ Mean value: 349€	Directly associated with the incoming order; can be retrieved from the company's order management and/or enterprise resource planning system.
Expected cash outflows per action ($CF_k^{\text{out,action}}$)	A piecewise defined function that varies with the workload of the action, i.e., the number of instances that pass the action.	See section 4.2.1 and Table 2 for more details	Can be retrieved from the company's workflow management system and enterprise resource planning system, the controlling and human resources department, as well as the capacity management department.
Rating index (X_2)	The rating index of the customer measured in terms of a scalar. The higher the rating index, the better.	Minimum: 194 Maximum: 975 Mean value: 604 	Provided by an external rating agency and is already stored in the company's CRM system.
Guarantee amount	The guarantee amount the trade credit insurance provides for a specific order. It depends on the rating index and is a percentage of the order size.	Minimum: 17€ Maximum: 186€ Mean value: 83€	Provided by external trade credit insurance.
Difference between the order size and the guarantee amount (X_5)	The difference results from subtracting the guarantee amount provided of trade credit insurances from the order size.	Minimum: 157€ Maximum: 397€ Mean value: 266€ 	-
Number of historical orders (X_7)	The number of historical orders made by the customer who relates to the instance under consideration.	Minimum: 0 Maximum: 52 Mean value: 14 	Can be retrieved from the company's CRM system.
Company limit (X_9)	The company limit is kind of internal insurance that may complement the guarantee amount provided by the external trade credit insurance.	Minimum: 0€ Maximum: 198€ Mean value: 58€ 	Has to be determined in line with the company's internal risk management guideline. In our example, customers with a high rating index and many historical orders get a higher company limit because they seem to be reliable.
Profit margin (λ_i)	The profit margin indicates which fraction of the cash inflows have to be used for covering cash outflows that cannot be assigned to distinct process actions.	0.2 (for all instances)	Can be determined based on the products included in the order and the cost information stored in the company's enterprise resource planning system.
Risk level ($r_{k,i}$)	See section 3.3.1 for more details.	See section 4.2.2 and Table 3 for more details.	Has to be estimated by consulting subject matter experts and/or based on historical data.
Reduction factor (η_i)	See section 3.3.1 for more details.	0.5 for all instances that traverse the process on a path that includes the action "manual handling", 1 otherwise (see section 4.2.2 for more details)	Has to be estimated by consulting subject matter experts.

Table 2. Expected cash outflow functions for each action.

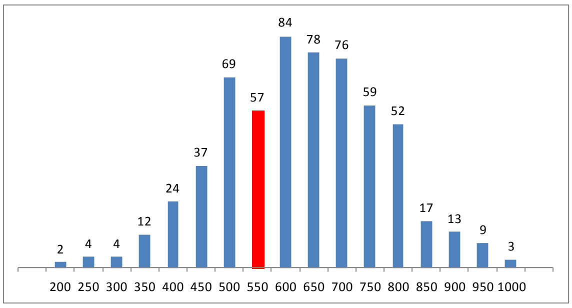
Action		Workload situation		
		Normal workload	**Moderately excessive workload**	**Peak workload**
Send customer data to trade credit insurance	capacity ranges			
	$CF_1^{\text{out,action}}$	constant 3		
Get guarantee amount from trade credit insurance	capacity ranges	$N_{3,\vec{y}}^{\text{action}} \leq 200$	$200 < N_{3,\vec{y}}^{\text{action}} \leq 500$	$N_{3,\vec{y}}^{\text{action}} > 500$
	$CF_3^{\text{out,action}}$	4€	10€ (= 4€ · 2.5)	1,000,000€
Delete data at trade credit insurance	capacity ranges	$N_{4,\vec{y}}^{\text{action}} \leq 200$	$200 < N_{4,\vec{y}}^{\text{action}} \leq 400$	$N_{4,\vec{y}}^{\text{action}} > 400$
	$CF_4^{\text{out,action}}$	7€	12.25€ (= 7€ · 1.75)	1,000,000€
Manual handling	capacity ranges	$N_{8,\vec{y}}^{\text{action}} \leq 250$	$250 < N_{8,\vec{y}}^{\text{action}} \leq 400$	$N_{8,\vec{y}}^{\text{action}} > 400$
	$CF_8^{\text{out,action}}$	2€	6€ (= 2€ · 3)	1,000,000€
Note reason for acceptance in CRM system	capacity ranges	$N_{10,\vec{y}}^{\text{action}} \leq 150$	$150 < N_{10,\vec{y}}^{\text{action}} \leq 500$	$N_{10,\vec{y}}^{\text{action}} > 500$
	$CF_{10}^{\text{out,action}}$	4€	6€ (= 4€ · 1.5)	1,000,000€
Accept order	capacity ranges	$N_{11,\vec{y}}^{\text{action}} \leq 321$	$321 < N_{11,\vec{y}}^{\text{action}} \leq 400$	$N_{11,\vec{y}}^{\text{action}} > 400$
	$CF_{11}^{\text{out,action}}$	4€	20€ (= 4€ · 5)	1,000,000€
Reject order	capacity ranges	$N_{6,\vec{y}}^{\text{action}} \leq 250$	$250 < N_{6,\vec{y}}^{\text{action}} \leq 500$	$N_{6,\vec{y}}^{\text{action}} > 500$
	$CF_6^{\text{out,action}}$	3€	4.5€ (= 3€ · 1.5)	1,000,000€

Table 3. Mapping functions for the risk levels.

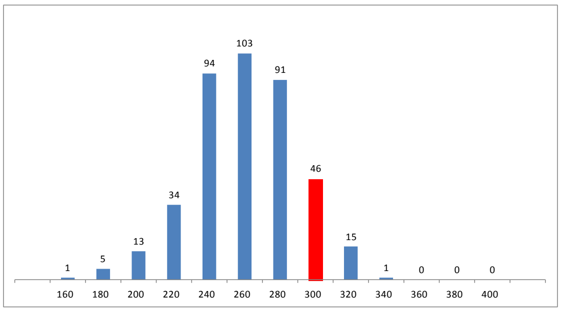
Process characteristic	Value Range	Risk level	Process characteristic	Value Range	Risk level
Rating index (X_2)	$x_{i,2} < 280$	10	**Difference between the order size and the guarantee amount (X_5)**	$x_{i,5} < 170$	1
	$280 \leq x_{i,2} < 350$	8		$170 \leq x_{i,5} < 200$	3
	$350 \leq x_{i,2} < 400$	7		$200 \leq x_{i,5} < 235$	4
	$400 \leq x_{i,2} < 600$	6		$235 \leq x_{i,5} < 265$	5
	$600 \leq x_{i,2} < 650$	5		$265 \leq x_{i,5} < 300$	6
	$650 \leq x_{i,2} < 780$	4		$300 \leq x_{i,5} < 350$	7
	$780 \leq x_{i,2} < 820$	3		$x_{i,5} \geq 350$	10
	$820 \leq x_{i,2} < 880$	2			
	$x_{i,2} \geq 880$	1			
Number of historical orders (X_7)	$x_{i,7} < 5$	10	**Company limit (X_9)**	$x_{i,9} < 15$	10
	$5 \leq x_{i,7} < 13$	6		$15 \leq x_{i,9} < 28$	8
	$13 \leq x_{i,7} < 22$	3		$28 \leq x_{i,9} < 45$	6
	$22 \leq x_{i,7} < 28$	2		$45 \leq x_{i,9} < 68$	5
	$x_{i,7} \geq 28$	1		$68 \leq x_{i,9} < 100$	4
				$100 \leq x_{i,9} < 137$	3
				$137 \leq x_{i,9} < 165$	2
				$x_{i,9} \geq 165$	1

Table 4. Optimization results.

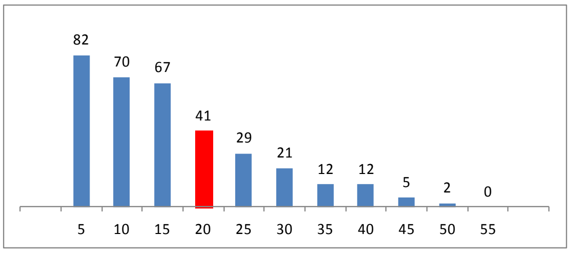
Description	Optimal result
Overall results	
Expected cash flow (CF)	3,711€
Average probability of occurrence ($p_{i,\bar{y}}$)	30%
Parameter setting with respect to the process characteristics	
Rating index (y_2)	540
Difference between the order size and the guarantee amount (y_5)	280€
Number of historical orders (y_7)	18
Company limit (y_9)	147€
Transition probabilities	
$s_{2,4} = P(X_2 \leq y_2)$	33%
$s_{5,7} = P(X_5 \leq y_5 \| X_2 > y_2)$	85%
$s_{7,8} = P(X_7 \leq y_7 \| X_5 \leq y_5, X_2 > y_2)$	73%
$s_{9,10} = P(X_9 \leq y_9 \| X_7 \leq y_7, X_4 \leq y_4, X_2 > y_2)$	92%



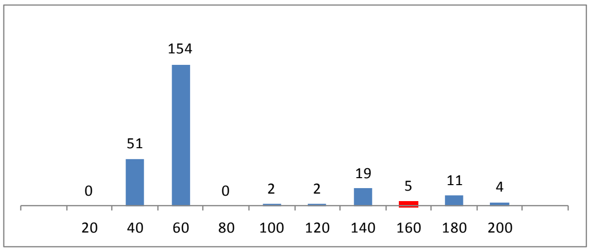
a) Distribution of the rating index (X_2). (600 instances)



b) Conditional distribution of the difference between the order size and the guarantee amount (X_5). (403 instances)



c) Conditional distribution of the number of historical orders (X_7). (341 instances)



d) Conditional distribution of the company limit (X_9). (248 instances)

Figure 3. Conditional distributions of the process characteristics for the optimal parameter setting.

II.3.4.4 Scenario analysis with respect to the reduction factor

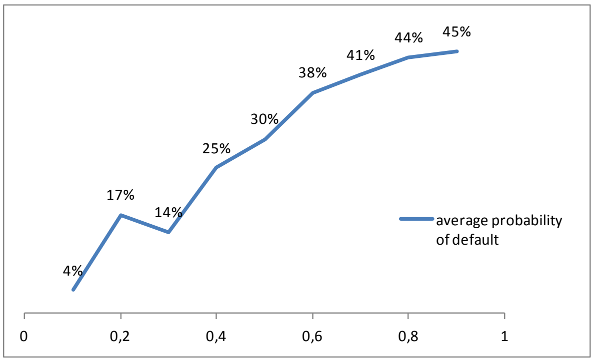
When applying the decision model, the probability of occurrence, which consists of the risk levels and the reduction factor (Formula 7), is the most difficult component to estimate. While the risk levels have the advantage of being aggregated into an average value such that estimation errors may cancel one another out, the reduction factor has to be determined stand-alone. We therefore use a scenario analysis to investigate how different reduction factor values influence the optimization results and the behavior of the decision model.

To determine the optimal parameter setting in section 4.3, we used a reduction factor of 0.5 for all instances that traverse the process on a path that includes the "manual handling" action. We used a reduction factor of 1 for all other instances. For the scenario analysis, we vary the reduction factor between 0.1 and 0.9, with an increment of 0.1, for all instances on paths that include the "manual handling" action. The reduction factor for all other instances remains 1. Figure 4 and Table 5 show the optimization results for all scenarios.

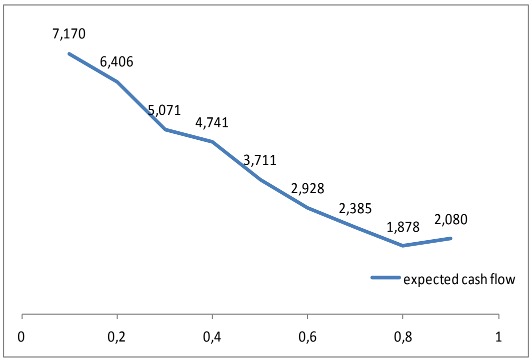
We first consider the average probability of occurrence and the expected cash flow. The average probability of occurrence ranges from 4% to 45% and increases with the reduction factor (Figure 4a). This behavior is reasonable because the reduction factor directly influences the probability of occurrence of each completed instance. Seemingly deviating from Formula (7), the average probability of occurrence increases only approximately proportionally to the reduction factor. The reason is that the average probability of occurrence depends on the concrete set of completed instances, whereas Formula (7) focuses on single instances. The expected cash flow ranges from 7,170€ to 2,080€. It decreases with an increase in the reduction factor (Figure 4b). The rationale behind this behavior is that, in the case of a higher reduction factor, the inflow-related risks have a higher probability of occurrence, a circumstance that causes the expected inflows to decrease.

With regard to the optimal parameter setting, the parameters for the rating index (y_2) and the difference between the order size and the guarantee amount (y_5) remain almost constant (Figure 4c). Only the parameters for the number of historical orders (y_7) and the company limit (y_9) vary with the reduction factor. This is reasonable because the decision nodes that relate to parameters y_2 and y_5 are not directly linked with the "manual handling" action. Parameter y_5 only increases if the reduction factor is 0.9. That is, the parameter value jumps up to 333€, whereas it is about 280€ in all other scenarios. As, for all instances that reach the decision node of parameter y_5, the difference between the order size and the guarantee amount is lower than 333€, all instances are forwarded to the decision node that relates to the number of historical orders ($s_{5,7} = 100\%$). The majority of these instances are accepted because their respective customers have a high number of historical orders ($1 - s_{7,8} = 75\%$) or low values for the company limit ($s_{9,10} = 16\%$). Because the concrete instances that are accepted in this scenario show superior cash flow effects than the instances that are accepted for a reduction factor of 0.8 (e.g. higher order size, less outflows for execution), the expected cash flow at the process level increases although the average probability of occurrence remains almost constant.

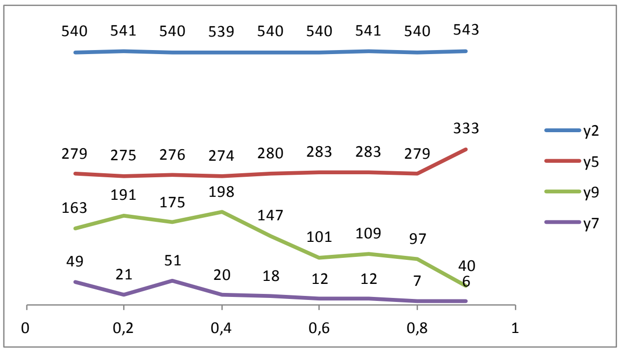
The parameter that relates to the number of historical orders (y_7) decreases for a reduction factor of 0.3 and beyond (Figure 4c). If the reduction factor is 0.3, the transition probability $s_{7,8}$ is 100% such that all instances that reach the decision node are routed to the "manual handling" action. The reason is that a small reduction factor has an extensive impact on the individual probability of occurrence. Moreover, the outflows per execution are quite low for the "manual handling" action even in the case of a high workload. Correspondingly, for reduction factors of 0.4 and higher, progressively fewer instances are routed via the "manual handling" action. As a result, an increasing number of instances that reach the decision node that relates to y_7 are directly accepted. The down and up for small reduction factors can be explained as follows: Because a small reduction factor implies a large reduction in the average probability of occurrence, all instances are routed via the "manual handling" action regardless of their number of historical orders. Therefore, parameter y_7 is high in such cases. It is notable that a decrease in y_7 for a reduction factor of 0.2 occurs at the same time when the parameter that relates to the company limit (y_9) increases. That is, in this case, the amount of historical orders required for immediate acceptance is smaller because the check of the company limit is more restrictive.



a) Relation between the average probability of occurrence and the value of the reduction factor.



b) Relation between the expected cash flow at the process level and the value of the reduction factor.



c) Development of the parameters in dependence of the value for the reduction factor.

Figure 4. Scenario analysis with respect to the reduction factor.

The parameter that relates to the company limit (y_9) is the one with the highest variation. This is reasonable because this parameter refers to the bottom-most decision node of the process and thus depends on the parameters of all other three decision nodes, among others. Until a reduction factor of 0.5, the parameter is so high that almost all instances that reach the respective decision node are accepted. The reason for this behavior is that, in such cases, it makes no sense to reject an instance after it has passed the "manual handling" action and features a low probability of occurrence regarding the inflow-related risks. For higher reduction factors, parameter y_9 is almost constantly decreasing. That is, parameter y_9 begins to decrease from about the same reduction factor value as the parameter that relates to the number of historical orders (y_7). Overall, the reduction factor has a strong enough impact on the probability of occurrence that it can still be more reasonable to reject an instance after it has passed the "manual handling" action than to accept a higher probability of occurrence. The lower the effect of the reduction factor, the more restrictive parameter y_9 is. Ultimately, only $s_{9,10} = 16\%$ of all instances that reach the decision node related to the company limit are accepted.

Overall, we conclude the following: First, different reduction factor values lead to different average probability of occurrence, expected cash flow, and parameter settings. Moreover, small estimation errors, as can be seen in Figure 4c, do not greatly affect the optimization results. Therefore, it was reasonable to include the reduction factor to account for a different type of information about the process. Nevertheless, the reduction factor should be determined with due care because large estimation errors impact the optimal parameter setting. Second, despite numerous interdependent effects, we were able to generate reasonable explanations regarding how the decision model behaves in different scenarios.

This result corroborates the notion that the decision model behaves plausibly and supports the making of rational process improvement decisions.

II.3.5 Conclusion and outlook

In this paper, we raised the question of how to derive concrete recommendations for process improvement that do not require extensive re-engineering projects and align with economic objectives. As an answer to this question, we proposed a decision model that indicates along which paths the instances of a process should be routed to maximize the value contribution of the process. Taking a process model and a set of historical instances as input, the decision model determines an optimal parameter setting. Each parameter refers to a distinct decision point in the process model and indicates on which outgoing edge an incoming instance continues its way through the process. When determining the optimal parameter setting, the decision model considers that different instances and paths have different cash flow effects. For example, each instance causes inflows of a certain amount and the outflows related to executing each action may depend on the overall workload of that action. Moreover, the parameter setting influences with which probability process risks for a specific instance occur. We also reported on the insights gained from applying the decision model to a demonstration example that is based on the order verification process of an IT distributor. The decision model is beset with limitations that stimulate further research:

1. The decision model considers a single process, which must also be one of the company's core processes. In the real world, however, we typically find networks of inter-connected core, support, and management processes. For future research, it seems worthwhile to extend the decision model such that it applies to numerous processes and also incorporates dependencies among process. Process improvement may then lead to results in which companies accept negative expected cash flow for support processes if this enables much higher expected cash flow for the corresponding core processes.

2. The decision model assumes that the set of historical instances is representative for the upcoming periods. It determines the parameter setting that maximizes the expected cash flow of a single period, which, in this case, is equivalent to maximizing the net present value of the expected cash flow of multiple identical periods. Moreover, the decision model may be applied repeatedly while updating the set of historical instances to adjust the parameter setting. Nevertheless, some long-term effects of process improvement are neglected so far. Future research should address this shortcoming by striving for an integrated short-term/long-term decision calculus. Such a decision calculus may, on the

one hand, be based on the net present value of the optimal expected cash flow of consecutive periods. On the other hand, such a decision calculus should incorporate inter-temporal dependencies. For example, the parameter setting of one period may depend on the parameter settings that have been chosen in the previous periods: customers may behave differently depending on past parameter settings in terms of their probabilities of cancelling orders or the probability of requesting the outcome of a process in a distinct period.

3. When we conducted the scenario analysis for the demonstration example, it was evident that, even for the comparatively small sample process, it was difficult to analyze how the different components of the decision model interact, how the optimal results can be validated, and how robust the results of the decision model are. The reason for this is that numerous components have to be considered in such analyses, such as the control flow structure of the process under investigation, the nesting of the decision points, the instance-specific cash inflows and probabilities of default, the reduction factor, and the workload-dependent outflows related to process execution. Although we were able to come up with reasonable explanations for the sample process, this may no longer be possible for more complex processes. To strengthen the analytic capabilities as well as insights into how the decision model behaves and the components mentioned above interact, further research should apply the decision model to more complex processes in the context of a multiple case study.

4. A multiple case study, as just recommended, would also benefit from a well-elaborated software tool that implements the decision model. Such a software tool should be able to handle processes of a higher complexity than the sample process as well as automatically process multiple sets of historical process instances and models. The software tool should also have an interface with common workflow management systems such that the workflow specifications of improved processes (if available) can be updated automatically. Finally, when implementing a software tool, future research should analyze whether the solution procedure we applied to improve the sample process can be enhanced. Ultimately, the potential of the decision model can only be tapped in industry if it disposes of an appropriate tool support and can be efficiently applied to many different processes.

Description	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5*	Scenario 6	Scenario 7	Scenario 8	Scenario 9
Reduction factor (η_i)	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
Overall results									
Expected cash flow (CF)	7,170€	6,406€	5,071€	4,741€	3,711€	2,928€	2,385€	1,878€	2,080€
Average probability of occurrence ($p_{\bar{y}}$)	4%	17%	14%	25%	30%	38%	41%	44%	45%
Parameter setting with respect to the process characteristics									
Rating index (y_2)	540	541	540	539	540	540	541	540	534
Difference between the order size and the guarantee amount (y_5)	279€	275€	276€	274€	280€	283€	283€	279€	333€
Number of historical orders (y_7)	49	21	51	20	18	12	12	7	6
Company limit (y_9)	163€	191€	175€	198€	147€	101€	109€	97€	40€
Transition probabilities									
$s_{2,4} = P(X_2 \leq y_2)$	33%	33%	33%	33%	33%	33%	33%	33%	32%
$s_{5,7} = P(X_5 \leq y_5 \mid X_2 > y_2)$	85%	81%	82%	80%	85%	87%	87%	85%	100%
$s_{7,8} = P(X_7 \leq y_7 \mid X_5 \leq y_5, X_2 > y_2)$	100%	76%	100%	76%	73%	49%	49%	33%	25%
$s_{9,10} = P(X_9 \leq y_9 \mid X_7 \leq y_7, X_4 \leq y_4, X_2 > y_2)$	95%	99%	98%	100%	92%	84%	85%	84%	16%

Table 5. Scenario analysis with respect to the reduction factor.

* Optimal solution presented in section 4.3.

II.3.6 Literature

Afflerbach, P., Kastner, G., Krause, F., and Röglinger, M. (2013), "An Optimization Model for Valuating Process Flexibility", Proceedings of the 34th International Conference on Information Systems, ICIS, Milan, Italy 2013.

Betts, A., Meadows, M. and Walley, P. (2000), "Call centre capacity management", Interntional Journal of Service Industry Management, Vol. 11 No. 2, pp. 185-196.

Biswal, P.C. (2005), Discrete Mathematics and Graph Theory, Prentice-Hall of India, New Delhi.

Bolsinger, M. (2014), "Bringing value-based business process management to the operational process level, appears in Information Systems and e-Business Management.

Braunwarth, K., Kaiser, M. and Müller, A. (2010), "Economic Evaluation and Optimization of the Degree of Automation in Insurance Processes", Business & Information Systems Engineering, Vol. 2 No. 1, pp. 29-39.

Buhl, H.U., Röglinger, M., Stöckl, S. and Braunwarth, K. (2011), "Value Orientation in Process Management - Research Gap and Contribution to Economically Well-founded Decisions in Process Management", Business & Information Systems Engineering, Vol. 3 No. 3, pp. 163-172.

Copeland, T.E., Koller, T. and Murrin, J. (1990), Valuation: measuring and managing the value of companies, Wiley, New York.

Davenport, T.H. (1993), Process Innovation: Reengineering work through Information Technology, Harvard Business School Press, Boston.

Devaraj, S. and Kohli, R. (2002), The IT Payoff: Measuring the Business Value of Information Technology Investments, Prentice Hall, Upper Saddle River.

Dumas, M., La Rosa, M., Mendling, J. and Reijers, H. (2013), Fundamentals of Business Process Management, Springer, Heidelberg.

Featherstone, A., Roessler, L. and Barry, P. (2006), "Determining the Probability of Default and Risk-Rating Class for Loans in the Seventh Farm Credit District Portfolio", Review of Agricultural Economics, Vol. 28 No. 1, pp. 4-23.

Forstner, E., Kamprath, N., and Röglinger, M. (2014), "Capability development with process maturity models - Decision framework and economic analysis", Journal of Decision Systems, Vol. 23 No. 2, pp.127-150.

Hammer, M. (2010), "What is Business Process Management?" in vom Brocke, J. and Rosemann, M. (Eds.), Handbook on Business Process Management 1 - Introduction, Methods and Information Systems, Springer, Heidelberg, pp. 3-16.

Hammer, M. and Champy, J. (1993), Reengineering the corporation: a manifesto for business revolution, Harper Business, New York.

Harmon, P. (2010), Business Process Change, Morgan Kaufmann, Burlington.

Hofacker, I. and Vetschera, R. (2001), "Algorithmical approaches to business process design", Computers & Operations Research, Vol. 28 No. 13, pp. 1253-1275.

Kirchmer, M. (2008), High Performance Through Process Excellence: From Strategy to Execution with Business Process Management, Springer Verlag, Pennsylvania.

Kohlbacher, M. and Reijers, H. (2013), "The effects of process-oriented organizational design on firm performance", Business Process Management Journal, Vol. 19 No. 2, pp. 245-262.

Koller, T., Goedhart, M. and Wessels, D. (2010), Valuation: Measuring and Managing the Value of Companies, John Wiley, New Jersey.

Korpela, J., Kyläheiko, K., Lehmusvaara, A. and Tuominen, M. (2002), "An analytic approach to production capacity allocation and supply chain design", International Journal of Production Economics, Vol. 78 No. 2, pp. 187-195.

Luftman, J., Zadeh, H.S., Derksen, B., Santana, M., Rigoni, E.H. and Huang, Z. (2012), "Key information technology and management issues 2011-2012: an international study", Journal of Information Technology, Vol. 27 No. 3, pp. 198-212.

Mays, E. (2001), Handbook of Credit Scoring, The Glenlake Publishing Company, Chicago.

Mitzenmacher, M. and Upfal, E. (2005), Probability and Computing: Randomized Algorithms and Probabilistic Analysis, Cambridge University Press, New York.

Neiger, D., Churilov, L., zur Muehlen, M. and Rosemann, M. (2006), "Integrating Risks in Business Process Models with Value Focused Process Engineering", in Ljungberg, J. and Andersson, M. (Eds.), Proceedings of the 14th European Conference of Information Systems, ECIS 2006, AIS, Göteborg, Sweden 2006, pp. 1606-1615.

Object Management Group (2011), "OMG Unified Modeling Language (OMG UML) Superstructure Version 2.4.1", available at:

http://www.omg.org/spec/UML/2.4.1/Superstructure/PDF (accessed: 04 February 2014).

Rappaport, A. (1986), Creating Shareholder Value: The New Standard for Business Performance, Free Press, New York.

Recker, J. (2007), "Understanding Quality in Process Modelling: Towards a Holistic Perspective", Australasian Journal of Information Systems, Vol. 2 No. 14, pp. 43-63.

Recker, J., Rosemann, M., Indulska, M. and Green, P. (2009), "Business Process Modeling - A Comparative Analysis", Journal of the Association for Information Systems, Vol. 10 No. 4, pp. 333-363.

Reijers, H.A. and Liman Mansar, S. (2007), "Best practices in business process redesign: use and impact", Business Process Management Journal, Vol. 13 No. 2, pp. 193-213.

Reijers, H.A. and Liman Mansar, S. (2005), "Best practices in business process redesign: an overview and qualitative evaluation of successful redesign heuristics", Omega, Vol. 33 No. 4, pp. 283-306.

Robinson, T.R., Henry, E., Pirie, W.L. and Broihahn, M. (2012), International Financial Statement Analysis, CFA Institute Investment Series, Hoboken.

Rocco, C.M., Miller, A.J., Moreno, J.A., and Carrasquero, N. (2000), “A cellular evolutionary approach applied to reliability optimization of complex systems”, Proceedings of Annual Reliability and Maintainability Symposium, 2000, Los Angeles, California 2000, pp. 210-215.

Sidorova, A. and Isik, O. (2010), "Business process research: a cross-disciplinary review", Business Process Management Journal, Vol. 16 No. 4, pp. 566-597.

Skrinjar, R., Bosilj-Vuksic, V. and Indihar-Stemberger, M. (2008), "The impact of business process orientation on financial and non-financial performance", Business Process Management Journal, Vol. 14 No. 5, pp. 738-754.

Stewart, G.B. and Stern, J.M. (1991), The quest for value: the EVA management guide, HarperBusiness, New York.

Trkman, P. (2010), "The critical success factors of business process management", International Journal of Information Management, Vol. 30 No. 2, pp. 125-134.

Trueck, S. and Rachev, S.T. (2009), Rating Based Modeling of Credit Risk: Theory and Application of Migration Matrices, Academic, Amsterdam.

van der Aalst, W.M.P. (2013), "Business Process Management: A Comprehensive Survey", ISRN Software Engineering, Vol. 2013, Article ID 507984.

van der Aalst, W.M.P. (2011), Process Mining. Discovery, Conformance and Enhancement of Business Processes, Springer, Berlin.

Vanderbeck, E.J. (2013), Principles of Cost Accounting, South-Western Cengage Learning, Mason.

Vergidis, K., Tiwari, A. and Majeed, B. (2008), "Business Process Analysis and Optimization: Beyond Reengineering", IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews, Vol. 38 No. 1, pp. 69-82.

vom Brocke, J., Recker, J. and Mendling, J. (2010), "Value-oriented process modeling: integrating financial perspectives into business process re-design", Business Process Management Journal, Vol. 16 No. 2, pp. 333-356.

vom Brocke, J., Becker, J., Braccini, A.M., Butleris, R., Hofreiter, B., Kapočius, K., De Marco, M., Schmidt, G., Seidel, S., Simons, A., Skopal, T., Stein, A., Stieglitz, S., Suomi, R., Vossen, G., Winter, R. and Wrycza, S. (2011), "Current and future issues in BPM research: a European perspective from the ERCIS meeting 2010", Communications of the AIS (CAIS), Vol. 28 No. 1, pp. 393-414.

West, D. (2000), "Neural network credit scoring models", Computers & Operations Research, Vol. 27 No. 11-12, pp. 1131-1152.

Yu, X. and Gen, M. (2010), Introduction to Evolutionary Algorithms, Springer, London.

Zellner, G. (2011), "A structured evaluation of business process improvement approaches", Business Process Management Journal, Vol. 17 No. 2, pp. 203-237.

Zellner, G. (2013), "Towards a framework for identifying business process redesign patterns", Business Process Management Journal, Vol. 19 No. 4, pp. 600-623.

III Automated Process Design

III.1 Research Paper 4: "Automated Planning of Process Models: The Construction of Exclusive Choices"

Authors:	Prof. Dr. Bernd Heinrich[a], Manuel Bolsinger[b], Marc-Andre Bewernik[c], [a] Chair for Information Systems II, University of Regensburg, Germany bernd.heinrich@wiwi.uni-regensburg.de [b] FIM Research Center, Department of Information Systems Engineering & Financial Management (Prof. Dr. Hans Ulrich Buhl), University of Augsburg, Germany manuel.bolsigner@fim-rc.de [c] SOFORT AG, Gauting, Germany m.bewernik@sofort.com
Published in:	Proceedings of the 30th International Conference on Information Systems, Phoenix, Arizona, USA, Paper 184

Abstract:

In our competitive world, companies need to adapt their processes quickly in order to react, for instance, to changing customer demands. Process models, as means to support process management, are nowadays often created and adapted in a time-consuming, widely used manual manner. Semantic Business Process Management in combination with planning approaches can alleviate this drawback by enabling an automated planning of process models. This paper describes the drawbacks that existing planning algorithms have related to the creation of process models. Therefore, we introduce - based on the design science paradigm - an innovative algorithm (method) that is suitable for the planning of process models focusing on the construction of the control flow pattern exclusive choice. Demonstrating the feasibility and the effectiveness of our method, we implemented our approach as a prototype. Finally, we evaluate the algorithm in terms of different properties like termination and its applicability within a real-use situation.

III.1.1 Introduction

In order to describe the increasingly complex processes within and across enterprises as well as for communication and training purposes, process modeling has proven to be an important instrument. A number of process modeling techniques have been developed in the past decades, including modeling languages like Event-driven Process Chains (EPC) or UML activity diagrams. However, process modeling and optimization are still time-consuming in practice, if new process models need to be designed or existing ones need to be adapted to changing requirements (Becker and Kahn 2003; Borges et al. 2005; Ma and Leymann 2008). Hornung et al. (2007) wrote, for instance, that "Manual process modeling is a time-consuming task and thus increases the total amount of modeling time.". Nevertheless, changing customer needs etc. make it necessary to maintain and adjust process models frequently. Even if reference process models are used, they have to be changed due to the new requirements as well, which again is time-consuming and costly. In addition, in many domains there are no reference process models at all that can be used. With this in mind, a fast and under economic considerations reasonable construction or adaption of process models is often difficult. For that reason, many process management departments in companies also need to deal with the criticism to cause too high costs compared to their benefit (e.g. Recker et al. 2005). A semantic annotation of process models, as envisioned in the research area Semantic Business Process Management (SBPM) can alleviate this drawback (Betz et al. 2006; Brockmans et al. 2006; Hepp et al. 2005; Hepp and Dumitri 2007; Thomas and Fellmann 2007) in combination with existing AI planning techniques (Bertoli et al. 2006; Hoffmann and Brafman 2005) by enabling an automated planning of process models (Heinrich et al. 2008; Henneberger et al. 2008).

In order to plan process models, not only a sequence of actions – the atomic elements of a process – but also other control structures (see Van der Aalst et al. 2003), which are provided by modeling languages and describe the control flow of a process, have to be constructed automatically. Thus, a fundamental challenge for the planning of process models is to consider such control structures, which can be reduced to a few fundamental control flow patterns. As the control flow pattern *exclusive choice* (besides e.g. the *simple merge*, and the *parallel split*) is one of the basic patterns for designing process models, this paper will examine its *automated construction*. The task of an automated construction of process models can be understood as a planning problem (Ghallab et al. 2004) with the objective to arrange the single process components, i.e. the actions, in an appropriate order.

To this end, we introduce – based on the design science paradigm – a technical definition of our planning domain and a novel algorithm (method) for the automated construction of exclusive choices within process models (cp. Figure 1). Therefore, we need special definitions provided by an abstract representation language. The main contributions are as follows[1]:

❶ The construction of exclusive choices is based on the preconditions and effects of actions[2]. In addition, planning has to consider that various, maybe large data types e.g. double (Biron and Malhotra 2004) can be assigned to them and that some actions accept only certain ranges of values of a data type (so called restrictions). We tackle these challenges by means of the representation of possibly infinite *sets of world states*, so *called belief states*.

❷ These belief states are presented in an *abstract representation language* which makes it possible to construct exclusive choices independently from well-known representation languages.

❸ This enables the definition of a planning domain (conditional deterministic belief-state-transition system) and a planning problem which allow us to design the necessary conditions (for instance, in UML they are called guards) – needed for the planning of the control flow pattern exclusive choice – automatically by an algorithm.

Considering the guidelines for conducting design science research by Hevner et al. (2004), we organize the paper as follows: The second section specifies both the problem context for which the new approach is relevant and the requirements that must be met in order to plan the exclusive choice in process models whereupon the related work is discussed. The third section introduces an abstract representation language and shows how belief states are represented. This is followed by a description of the planning model, our planning domain. Section five answers the key research question of how the exclusive choice control flow

[1] The semantic annotation of actions is analyzed in a step before the planning of control structures and is therefore not the focus of this paper (for a more detailed description of how the semantic annotation is used, we refer to Heinrich et al. 2008).

[2] The term *preconditions* denotes everything an action needs in order to be performed, including input parameters; the term *effects* denotes everything an action provides after it was performed, including output parameters, as it is used in AI planning.

pattern can be constructed automatically by providing the necessary algorithm. The penultimate section is dedicated to an evaluation of our approach. Here, we demonstrate the feasibility and the implementability of our approach by means of a prototype (instantiation) and illustrate its application within a real-use situation (here, our instantiation has the character of a working example (see also Gregor and Jones 2007, p. 323)). Furthermore, we mathematically evaluated the presented approach in terms of different properties like termination. Finally, the last section summarizes our considerations and provides an outlook on future steps.

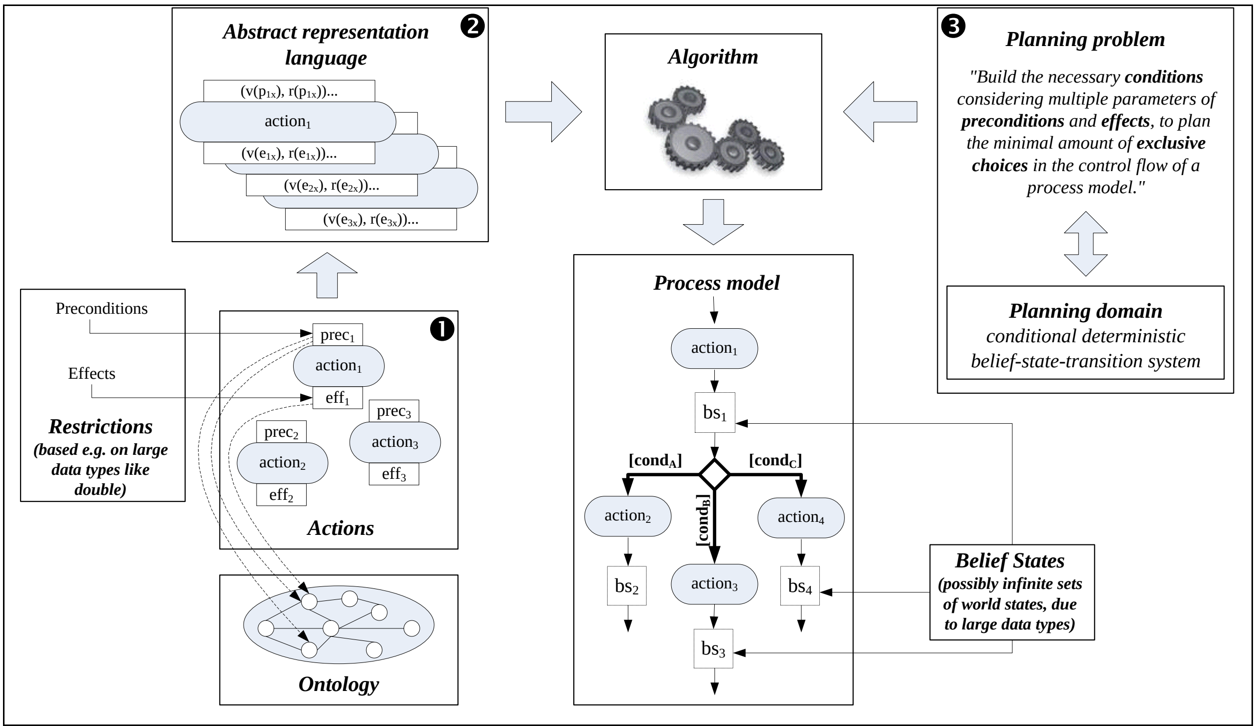


Figure 1. The Planning Approach to Construct the Control Flow Pattern Exclusive Choice

III.1.2 Problem Context, Requirements and Related Work

At first, we present the problem context of this paper with regard to the research strand of SBPM. After this, we discuss the types of processes, which seem to be appropriate for the use of automated process planning in practice. That is followed by aligning the problem context to the planning of exclusive choices, before we formulate the requirements for the planning domain and the algorithm, which are based on literature. At the same time, these requirements are the source for the subsequent analysis of the related work, to define the need for research.

III.1.2.1 Problem Context

Based on the awareness that process modeling is time-consuming, the emerging research strand of SBPM tries to reach a higher level of automation in the creation and adaptation of process models by means of their semantic annotation as well as planning algorithms. For a more detailed description of SBPM see e.g. Hepp et al. (2005), Hepp and Dumitri (2007) and Thomas and Fellmann (2007). This is the field of study in which we envision the automated planning of process models (which is understood as a plan in the following sections). We propose that if actions, which can be – in accordance to Hepp and Dumitri (2007) – stored in a process library, are semantically annotated, it becomes possible to create process models automatically for a given problem description (see also Henneberger et al. 2008). The annotation of an action includes a semantic annotation of the *preconditions* needed for it to be performed and the *effects* provided after it has been performed. A description for a planning problem comprises, besides the actions, of an initial state representing the overall process input and a set of goals representing the desired process output. Given such a problem description, a planner is expected to build feasible process models (e.g. Heinrich et al. 2008; Henneberger et al. 2008). During the overall planning, the semantic annotation of the corresponding elements of the problem is analyzed by an inference mechanism to identify the existing dependencies between initial state, actions and goal states. Since this semantic reasoning can be done prior to the planning of the control flow patterns (see Heinrich et al. 2008), we do not address it further in this paper.

As discussed in the introduction, automated planning is thought to contribute to process modeling to design and adapt process models faster. However, the question of which processes are adequate to apply a process planning algorithm to and which are not, needs to be clarified. Such a classification is required to specify the business problem context as well as the boundaries within which the approach is expected to be applied. Processes are classified in literature related to different criteria (e.g. Marjanovic 2005; Weske et al. 2004). Here, criteria like degree of process repetition, frequency of process redesign and process value seem promising. According to the papers mentioned before, a professional *manual* process design is suggested (in contrast to not doing such a process design) especially for repetitive processes that need to be (re)designed repeatedly and which are of high value for a firm. This is justified primarily by economic reasons, i.e. any high initial costs of analyzing and documenting the problem domain (for which the process is constructed) as well as the costs of a manual adaption of an existing process are worthwhile, if the analysis and documentation, once made, can be used again during another redesign. The classification by these authors can be transferred in a sense to automated planning of process models. Here, too, high initial costs occur to analyze and annotate the process actions (preconditions and effects) and to implement the planning algorithm. Another similarity to *manual* process design is that the annotations and the implemented algorithm can later be reused during further (re)design projects, which reduces the costs and time of the design. Therefore, we will focus first on repetitive processes that need to be (re)designed repeatedly and which are of high value. Such processes seem to be the first choice for applying automated planning approaches. This boundary of the problem context is expected to be relevant, since usually there are many processes in companies that belong to this process class. Thus, we have to evaluate later, if under the defined conditions – especially repetitive processes that need to be (re)designed repeatedly – an automated planning of process models is useful.

In our context, we focus on the, so far, *unsolved* issue of planning the control flow pattern exclusive choice in process models. As suggested by Hevner et al. (2004), we decompose the problem of planning a whole process model into subproblems such as the planning of exclusive choices in order to address this subproblem in depth. The planning domain and algorithm have to cope with a number of requirements, which are presented in the following subsection of this paper before related work is discussed.

III.1.2.2 Requirements

The planning domain must meet general requirements (see also Bertoli et al. 2006; Constantinescu et al. 2004; Meyer and Kuropka 2005; Meyer and Weske 2006; Pathak et al. 2006). These requirements are also the basis for the analysis of existing approaches and AI planning techniques in the next subsection. The following requirements need to be considered:

(R1) *Preconditions and effects and the data types of their parameters*: It must be possible to assign various data types (e.g. defined in Biron and Malhotra 2004) to parameters of preconditions and effects of actions and consider them within the planning model. Some actions may require restricted ranges of values for a data type. Therefore, restrictions have to be considered as well. In other words, it should be possible to specify the range of values an action accepts for a precondition and the range of values it produces for an effect. Since the restrictions of consecutive actions may not match completely, the planning algorithm needs to consider all possible ranges in the process model. In a given world state according to an individual process execution, every parameter can be represented by a unique value, i.e., allowing ranges of values is equivalent to allowing a possibly infinite set of world states.

(R2) *Planning independently of a concrete representation language*: In order to use the planning approach independently of a concrete representation language, a general and formal framework (planning domain and algorithm) has to be built by the use of an abstract representation language.

(R3) *Planning of exclusive choices*: An algorithm should be able to plan control structures and especially exclusive choices in process models automatically. Van der Aalst et al. (2003) provides a comprehensive overview of the various control structures that may be part of a process model. Moreover, he analyzes how these structures are represented in different process modeling languages such as UML activity diagrams. Thus, following these findings, a planning domain needs to consider conditions for constructing control structures, such as *exclusive choice*.

III.1.2.3 Related Work

The planning problem described in this paper can be characterized as a nondeterministic planning problem with initial state uncertainty. It is nondeterministic because we abstract from an individual process execution. Therefore the realizations of parameter values (and thus the world state) are not determined at the moment of planning (Ghallab et al. 2004). Likewise parameters are not fully determined in the initial state either, which is frequently called initial state uncertainty (R1) (Bonet and Geffner 2001). Although there are already algorithms that can cope with nondeterminism and initial state uncertainty (e.g. Bertoli et al. 2006; Bonet and Geffner 2001), these approaches do not reach out for the planning of process models because of their limited capabilities concerning the construction of control structures (R3). Other algorithms, such as the one presented in Bertoli et al. (2006), progress from an initial belief state to one of the goal belief states. It builds a search tree, to find all possible paths beginning with the initial belief state, branching on conditions. The approaches of such conditional planning (e.g. Bertoli et al. 2006; Hoffmann and Brafman 2005) do not fit our problem. They encode so called observations, which are points in the plan, where it is necessary to determine, if some logical expression is valid or not in order to proceed further in the plan. They encode the observations separately in the form of observation variables and observation actions, respectively, making them part of their planning domain. In practice this means, that there are observations in the domain description, which makes such an observation the only point in the plan, where it might branch (e.g. in order to construct an exclusive choice). Thus, exclusive choices are possible (see also the conditional planners such as Bonet and Geffner 2001; Hoffmann and Brafman 2005) but are "hard-coded" in the domain (e.g. in terms of sensing actions) and additionally restricted to Boolean variables. In process modeling, which is our problem context, such observations are not given as this would simplify planning process models to a great extent. In practice this means, that a planning approach must determine the conditions by itself without having the observations given in advance (R3). Finally, there needs to be a way of representing belief states, in order to map one belief state into another, which is a key problem in belief space according to Geffner (2002) (R1). One promising approach seems to be the use of Binary Decision Diagrams as proposed in Bertoli et al. (2006). Another way is the representation of a belief state implicitly by an initial belief state together with a sequence of actions that leads to this belief state as done in Hoffmann and Brafman (2006). Other planners, for example in Bonet and Geffner (2000), enumerate all states of the world explicitly that may occur after applying an action. Because of large data types (e.g. double), which are essential for the problem context of process modeling, these approaches are not suitable.

As the composition of (semantic) web services forms a similar issue as the automated planning of process models we additionally want to briefly mention some planning and rule-based approaches that have already been adapted for the composition of (semantic) web services (e.g. Liang and Su 2005, Pistore et al. 2005, Weigand et al. 2008, Wu et al. 2003). However, none of these approaches meet the above-mentioned requirements for process planning, especially nondeterministic planning, the construction of exclusive choices, and the ability to handle different data types. Yet, these issues are necessary for the planning of process models.

III.1.3 Abstract Representation Language

In this section we present a formal definition of belief states, making it possible to explicitly represent an infinite state space independently of a concrete representation language. This is the foundation for both the description of our planning model and the development of our algorithm in order to meet the requirements (R1) to (R3). With the abstract representation language it is possible to represent possibly infinite sets of world states quite easily.

Furthermore, it is a rather intuitive way – from a process modeling perspective – to express certain preconditions and effects[3]. We do not use other representation languages like set-theoretic representation or state-variable representation (see Ghallab et al. 2004), since we do not have a classical planning problem and we want to build a planning model that is independent of a concrete representation language[4]. In this way, it is a kind of specialized language to describe our problem considered in a new way (March and Smith 1995). When talking about process models, the annotation of their actions includes a specification of the preconditions and the effects. We define a parameter of the preconditions and effects as *belief state tuple* that consists of the parameter name and a set of values, all of which can be assigned to the parameter in a specific world state (according to an individual process execution). Thus the name of a parameter is also understood as a variable that can take on all the values in the set of values. The data type of a parameter is the predefined domain of a belief state tuple.

Definition 1 *(Belief state tuple)*. A *belief state tuple p* is a tuple of a *belief state variable* $v(p)$ and a subset $r(p)$ of its predefined domain $dom(p)$, which we will write as $p := (v(p),r(p))$. It is $v(p) \in r(p)$ in a specific world state. When talking about belief states, $v(p)$ is the symbol of the belief state variable. The set $r(p) \subseteq dom(p)$ is called the *belief-state-variable restriction* (abbr.: restriction) of $v(p)$, which contains the values that can be assigned to $v(p)$ in a possible world state. If $r(p)=\varnothing$ then, the belief state variable does not exist (anymore), allowing the deletion of a belief state variable.

According to this definition, each belief state variable $v(p)$ has a predefined data type (e.g. double) specifying the predefined domain $dom(p)$. Additionally, restrictions $r(p)$ can be defined for each belief state variable $v(p)$. A restriction can either be described by logical expressions (e.g. $u \geq 4 \wedge u \leq 5$) defining a set of values or an explicit enumeration of values (e.g. $u \in \{4,5\}$) for a specific belief state variable (e.g. u).

Example 2. The set $bs_1 = \{ (u,[4,5]), (v,\{T\}), (w,\{T\}), (y,\{F\}), (z,[5,6]) \}$ represents a set of belief state tuples. The restriction of u is an interval of the double data type, whereas four

[3] As mentioned above, these preconditions and effects need to be semantically annotated. This is realized by linking their parameters already during the specification of the actions with classes of an OWL 2 ontology (Motik et al. 2008). These semantics are analyzed by an inference step prior to the planning of control flows and its patterns (see Heinrich et al. (2008)), so that for the further paper it is sufficient to demonstrate the approach on a syntactic basis.

[4] However, if we restrict all of the atoms and belief state variables to be ground, then from a theoretical point of view our abstract representation language has equivalent expressiveness than the languages mentioned.

and five are part of the interval. The domains of the belief state variables might be predefined as *dom(u)=dom(z)=double, dom(v)=dom(w)=dom(y)=boolean.*

With the definition of a belief state tuple we have one cornerstone of the planning model, presented in the next section. By the use of these belief state tuples we can explicitly represent belief states and explain in what sense we understand them as possibly infinite sets of world states by the help of definition 4 in combination with definition 3.

Definition 3 ($\sqsubseteq$). Let $A = \{u_1, \ldots, u_k\}$ and $B = \{w_1, \ldots, w_m\}$ be two finite sets of belief state tuples, then: $A \sqsubseteq B :\Leftrightarrow \forall u \in A\ \exists w \in B\colon v(u)=v(w) \wedge r(u) \subseteq r(w) \wedge |r(u)|=1$.

Definition 4 (*Belief state and world state*). Let $BST = \{p_1, \ldots, p_n\}$ be a finite set of *belief state tuples*. A *belief state bs* is a subset of *BST*, containing every belief state variable one time at the most. A world state *s* is a member of the belief state *bs*, in the context that it is $s \sqsubseteq bs$.

The set *bs* is constituted by the restrictions that currently apply to a set of belief state variables. Similar to Petrick and Bacchus (2002), the set *bs* thus could be interpreted as a kind of knowledge base capturing the knowledge about available belief state variables. The set *bs* therefore describes different conceivable world states that may occur during plan execution. It needs to be distinguished from the world state, which generally refers to an individual situation at process execution time. According to literature (e.g. Bertoli et al. 2006; Bonet and Geffner 2000; Hoffmann and Brafman 2005; Hoffmann and Brafman 2006), a set of world states is called belief state. Since the set *bs* is a set of world states in the context of definition 4, we will follow this wording and refer to the set *bs* as a *belief state*.

This way of representing a set of world states is one starting point to define and solve our planning problem. Furthermore, we can explicitly represent belief states in a rather intuitive way. Additionally, with this abstract representation language (R2), a belief state can be a possibly infinite set of world states (R1). Hence, a world state *s* is an instance of a belief state *bs*. In the following section this allows us to describe a concrete transition function, which is needed for the planning model.

III.1.4 Planning Model

As stated above, our approach can be seen as a nondeterministic planning problem with initial state uncertainty. Our approach is inspired by the framework given in Bertoli et al.

(2006), but we will describe our domain in the abstract representation language specified before. Furthermore, we will handle nondeterminism in a different way. This is necessary because former approaches include observations in the domain description. In the context of process modeling these observations are not known in advance. To solve this problem, our idea is to automatically create sets of conditions, where a plan can branch. This is based on the presented requirements for the automated planning of the exclusive choice control structure (R3), which leads to conditional plans.

In this section we describe our search process (related to the guidelines for conducting design science research) and start with a nondeterministic state-transition system and its definition. Then we modify it by using the introduced abstract representation language of the previous section. Instead of states we use belief states in the transition system, which makes it possible to change the transition function in a first step to be deterministic regarding belief states. As a result of that change, we define a *deterministic belief-state-transition system*, making it possible to cope with (R1) and (R2). In a second step we extend the transition function by what we call *conditions*, in order to build branches in a plan. As a consequence, we describe a *conditional deterministic belief-state-transition system* to handle our planning problem at the end of this section.

III.1.4.1 Nondeterministic State-Transition System – The Starting Point

When being confronted with a nondeterministic planning problem, it is common to use a nondeterministic planning domain. In general, "a nondeterministic state-transition system is defined in terms of its states, its actions, and of a transition function that describes how (the execution of) an action leads from one state to possibly many states" (Bertoli et al. 2006). We use this as a working definition of a nondeterministic state-transition system. More formally a state-transition system and (non-)determinism in state space are defined in Bertoli et al. (2006) as follows.

Definition 5 (*Nondeterministic state-transition system*). "A nondeterministic state-transition system is a tuple $\Sigma = (S, A, R)$, where

- S is a finite set of *states*,
- A is a finite set of *actions*,

- and $R : S \times A \rightarrow 2^S$ is the transition function. The transition function associates to each state $s \in S$ and to each action $a \in A$ the set $R(s, a) \subseteq S$ of next states."

Definition 6 ((*Non-)determinism in state space*). "An action a is *applicable* in a state s ([…]) iff $|R(s, a)| > 0$; it is *deterministic* (*nondeterministic*) in s iff $|R(s, a)| = 1$ ($|R(s, a)| > 1$). If a is applicable in s, then $R(s, a)$ is the set of states that can be reached from s by performing a."

As mentioned, we take this nondeterministic state-transition system as a starting point, and simply define it in a different way by using the introduced abstract representation language of the previous section.

III.1.4.2 Deterministic Belief-State-Transition System – The First Step

So far, we defined a nondeterministic state-transition system and what we understand as a belief state. With this, we define a deterministic belief-state-transition system and (non-)determinism in belief space. Similar to the working definition of the nondeterministic state-transition system, we formulate a working definition of a so called *deterministic belief-state-transition system* which is defined in terms of its belief states (sets of states), its actions, and of a transition function that describes how (the execution of) an action leads from one belief state to one and only one belief state. More formally:

Definition 7 *(Deterministic belief-state-transition system)*. Let $BST = \{p_1, \ldots, p_n\}$ be a finite set of belief state tuples. A deterministic belief-state-transition system is a tuple $\Sigma_d = (BS, A, \gamma_d)$, such that:

- $BS \subseteq 2^{BST}$ is a finite set of belief states, i.e., each belief state $bs \in BS$ is a subset of BST.

- A is a finite set of actions. Each action $a \in A$ is a triple consisting of the action name and two subsets of BST, which we will write as $a := (name(a), precond(a), effects(a))$. The **set** $precond(a) \subseteq BST$ are the preconditions of a and the **set** $effects(a) \subseteq BST$ are the effects of a.

- An action a is *applicable* in a belief state bs (denoted with *applicable(a, bs)*) iff $bs \sqsubseteq precond(a)$ ($\sqsubseteq$ as defined in definition 9). This phrases a sufficient condition that needs to be met so that an action a can actually be performed in a belief state bs, since a can be performed in all the possible world states of bs. All belief state variables in $precond(a)$ are available in bs. At the same time, the restriction of each belief state variable in bs is a subset of the restriction required by a belief state variable in *precond(a)*. In other words, an action is applicable iff the action can be performed in each world state $s \sqsubseteq bs$.

- The transition function is $\gamma_d : BS \times A \rightarrow 2^{BS}$ with $\gamma_d(bs,a) = \{ (bs \setminus \{ (v_{bs},r_{bs}) \in bs \mid v_{bs} = v_{effects}, (v_{effects},r_{effects}) \in effects(a) \}) \cup effects(a) \}$ if $a \in A$ is applicable in $bs \in BS$, and undefined otherwise.
- 2^{BS} is closed under γ_d, i.e., if $bs \in BS$, then for every action a that is applicable in bs, $\gamma_d(bs, a) \in 2^{BS}$.

Definition 8 ((*Non-)determinism in belief space*). An action a is *deterministic* (*nondeterministic*) in bs iff $|\gamma_d(bs, a)| = 1$ ($|\gamma_d(bs, a)| > 1$). If a is applicable in bs, then $\gamma_d(bs, a)$ is the set of belief states that can be reached from bs by performing a.

Definition 9 ($\sqsubseteq$). Let $A = \{u_1, ..., u_k\}$ and $B = \{w_1, ..., w_m\}$ be two finite sets of belief state tuples, then: $A \sqsubseteq B :\Leftrightarrow \forall w \in B\ \exists u \in A: v(w)=v(u) \wedge r(u) \subseteq r(w)$.

Example 10. Let $(a_1, precond(a_1):=\{ (u,[3,6]), (v, \{T,F\}), (w,\{T\}), (y,\{F\}) \}, effects(a_1):= \{ (u,[2,7]), (v,\{F\}), (x,[1,4]) \})$ be an action and let $bs_1 = \{ (u,[4,5]), (v,\{T\}), (w,\{T\}), (y,\{F\}), (z,[5,6]) \}$ be a belief state. Here a_1 is applicable in the belief state bs_1, since $bs_1 \sqsubseteq precond(a_1)$, because for each belief state variable among the preconditions of a_1 there is a belief state variable in bs_1 and it is $[4,5]_{u\_bs1} \subseteq [3,6]_{u\_a1}$, $\{T\}_{v\_bs1} \subseteq \{T,F\}_{v\_a1}$, $\{T\}_{w\_bs1} \subseteq \{T\}_{w\_a1}$ and $\{F\}_{y\_bs1} \subseteq \{F\}_{y\_a1}$[5]. As a result it is $\gamma_d(bs_1,a_1)=\{ \{ (u,[2,7]), (v,\{F\}), (w,\{T\}), (x,[1,4]), (y,\{F\}), (z,[5,6]) \} \}$.

We compare definition 5 and definition 7 to show, that definition 7 just extends definition 5, but is basically another way to define a nondeterministic state-transition system. Our transition system is called a belief-state-transition system, since it is not based on states, but on belief states. A nondeterministic state-transition system can be written as a deterministic belief-state-transition system, because:

- Both transition systems have a finite set of sets of world states, $|S|<\infty$ and $|BS|<\infty$. The nondeterministic state-transition system has a finite set of world states S. A single world state $s \in S$ can be understood as a set of world states, having only one element. Due to that, the nondeterministic state-transition system has a finite set of sets of world states. The

[5] The indices are just given to note, to which belief state variable the restriction belongs to, for example $[4,5]_{u\_bs1}$ indicates that the interval of numbers $[4,5]$ is the restriction of u in bs_1.

deterministic belief-state-transition system is based on a finite set of belief states *BS*, and therefore has a finite set of sets of world states, too.

- Both transition systems have a finite set of actions.
- Both transition systems have a transition function that associates to each set of world states and to each action, a set of next world states.

According to definition 8, the belief-state-transition system is called deterministic, since it is $|\gamma_d(bs, a)| = 1$. This deterministic belief-state-transition system allows transitions from a set of world states, with more than one element, to a set of next world states, which is made possible through the use of belief state tuples. Here, the nondeterministic state-transition system is extended, since it only allows transitions from one world state to a set of next world states. Due to the definition of the belief states tuples, another extension is that a belief state is a possibly infinite set of world states. These two extensions tackle the problem of representing and updating belief states in large state spaces. With this novel model definition, it is now – in contrast to former works – possible to cope with (R1) and (R2).

As the next example demonstrates, the transition function γ_d might leave out transitions that are possible in the process modeling context, due to the fact that a transition takes place only when an action is applicable in a belief state.

Example 11. Let $(a_1,\ \{ (u,[3,6]), (v,\{T, F\}), (w,\{T\}), (y,\{F\}) \},\ \{ (u,[2,7]), (v,\{F\}), (x,[1,4]) \})$ be an action and let $bs_2 = \{ (u,[1,5]), (v,\{T\}), (w,\{T\}), (y,\{T,F\}), (z,[5,6]) \}$ be a belief state. Here a_1 is not applicable in the belief state bs_2, since $bs_2 \nsubseteq precond(a_1)$, because it is $[1,5]_{u\_bs2} \nsubseteq [3,6]_{u\_a1}$ and $\{T,F\}_{y\_bs2} \nsubseteq \{F\}_{y\_a1}$. The transition function $\gamma_d(bs_2,a_1)$ would be considered as not defined, although it would be defined if, for example, $(u,\{4\})$ and $(y,\{T\})$ hold in an individual process execution (a certain world state of bs_2). Thus, it is necessary to consider branches with conditions in the constructed process model (see (R3)).

III.1.4.3 Conditions – Enabling the Second Step

As a result of example 11, we realize the need to generalize the transition function to allow the performing of what we call *partly applicable* actions in a belief state. In a next step, we therefore extend the transition function γ_d of definition 7 by so called *conditions*, which are comparable to the routing constraints in Sun et al. (2006).

Definition 12 *(Partly applicable).* An action a is *partly applicable* in a belief state bs (denoted with *partly_applicable(a, bs))* iff:

$$\forall u \in precond(a)\ \exists w \in bs: v(u)=v(w)\ \wedge\ (r(u) \cap r(w) \neq \emptyset).$$

Definition 12 describes a necessary condition that needs to be met so that an action a can actually be performed in a belief state bs. All belief state variables in *precond(a)* are available in bs and the restriction of each belief state variable (i.e. the set of possible values) does not contradict the restriction required by a for that belief state variable. However, there may still be situations (certain world states of bs) where performing a is not possible due to the restrictions. If we allow partly applicable actions, our transition function needs to be able to handle actions that are partly applicable in a belief state.

Example 13. Let $(a_1,\ \{ (u,[3,6]), (v,\{T,F\}), (w,\{T\}), (y,\{F\}) \},\ \{ (u,[2,7]), (v,\{F\}), (x,[1,4]) \})$ be an action and let $bs_2 = \{ (u,[1,5]), (v,\{T\}), (w,\{T\}), (y,\{T,F\}), (z,[5,6]) \}$ be a belief state. Here a_1 is partly applicable in the belief state bs_2, because it is $[1,5]_{u\_bs2} \cap [3,6]_{u\_a1} = [3,5] \neq \emptyset$, $\{T\}_{v\_bs2} \cap \{T,F\}_{v\_a1} = \{T\} \neq \emptyset$, $\{T\}_{w\_bs2} \cap \{T\}_{w\_a1} = \{T\} \neq \emptyset$ and $\{T,F\}_{y\_bs2} \cap \{F\}_{y\_a1} = \{F\} \neq \emptyset$.

If we have a belief state bs, then there might exist a nonempty set of actions A_{p_a} with actions that are partly applicable in bs. In an individual process execution, for every belief state variable in bs a specific value can be observed for a certain world state $s \in bs$. In this world state s, it might be possible to perform all actions in A_{p_a} or just the actions in a subset $A_{perform} \subseteq A_{p\_a}$. This means, for every belief state variable of a belief state we have to discover for which observations (constellation of parameter values) an action is in $A_{perform}$. In other words, we need to detect, when it is possible to perform an action and consider it in a process model and when this is not possible. Therefore, we need to find a *set of conditions* under which an action can always be performed, that is, under which set of conditions an action is applicable in bs (R3). With these definitions, it is possible to plan exclusive choices, since we now know the conditions and the corresponding actions to construct branches in a process model. Former works (e.g. Bertoli et al. 2006) do not consider sets of conditions (sets of sets of observations). This makes our notation more expressive.

Definition 14 (*Condition*). A *condition* q is a tuple of a *condition variable* $v(q)$ and a subset $r(q) \neq \emptyset$ of its predefined domain *dom(q)*, which we will write as $q := (v(q),r(q))$. It is $v(q) \in r(q)$ in a certain world state. When talking about belief states, then $v(q)$ is the symbol of the condition variable. The set $r(q) \subseteq dom(q)$ is called the *condition restriction* of $v(q)$, which is the set of values that might be assigned to $v(q)$ in a possible world state.

The restriction $r(q)$ of a condition q is a set of possible values that might be observed in a world state for one condition variable $v(q)$. A set of conditions c is built for every action that is partly applicable in a belief state bs, which are the actions in A_{p_a}. This set of conditions c might be different for every action in A_{p_a}, and it is – in contrast to other planning problems – not provided prior to planning a process model. It needs to be determined automatically (R3). In a certain world state, we can then perform these actions, where all conditions are *fulfilled*, i.e., for every condition in c there is an observed value in the world state, and the observed value of the condition variable is a member of the restriction of the condition for this condition variable. This way it is known in each state, which actions can be performed, or as we call it, can be *executed*. An action can be *executed* in a certain world state of a belief state, either if the action is applicable in the belief state ($c=\emptyset$) or if the action is both partly applicable in the belief state and all of its conditions are fulfilled.

Example 15. Let $(a_1,\ \{ (u,[3,6]), (v,\{T,F\}), (w,\{T\}), (y,\{F\}) \},\ \{ (u,[2,7]), (v,\{F\}), (x,[1,4]) \})$ be an action and let $bs_2 = \{ (u,[1,5]), (v,\{T\}), (w,\{T\}), (y,\{T,F\}), (z,[5,6]) \}$ be a belief state. As shown in example 13, a_1 is partly applicable in bs_2. Since a_1 is partly applicable, we need to find the conditions that have to be fulfilled so that the action can be executed. We can do that by looking for the reason, why a_1 is not applicable in bs_2. It is not applicable, because it is $[1,5]_{u\_bs2} \not\subseteq [3,6]_{u\_a1}$ and $\{T,F\}_{y\_bs2} \not\subseteq \{F\}_{y\_a1}$, as presented in example 11. If we could restrict these restrictions even more, then a_1 would be applicable in bs_2. This is exactly what we do with the set of conditions. If $(u,[3,5])$ and $(y,\{F\})$ holds then a_1 would be applicable in bs_2, which makes $c_1 = \{ (u,[3,5]), (y,\{F\}) \}$ the set of conditions that need to be fulfilled to execute a_1 in a certain world state of bs_2.

As mentioned, the set of conditions might be different for every action that is partly applicable. In order to assign a set of conditions c to a belief state bs and an action a we define a *condition function*. This function associates to each belief state bs and each action a a set of possible conditions.

Definition 16 *(Condition function)***.** Let $\Sigma_d = (BS, A, \gamma_d)$ be a deterministic belief-state-transition system. Let C be a possibly infinite set of conditions. A condition function over

BS and *C* is a function $\theta : BS \times A \rightarrow 2^C$ (denoted with CONDITIONFUNCTION(*bs*, *a*)), which associates to each belief state *bs* and each action *a* the set of possible conditions $\theta(bs, a) \subseteq C$.

Therefore, not only an action influences the transition from one belief state to another one, but also the conditions under which this action can be executed. We remark that in practice the conditions are not given additionally to the domain in any way, but need to be created by an algorithm (R3).

III.1.4.4 Planning Domain and Planning Problem – The Second Step

The previous discussions lead to a *conditional deterministic belief-state-transition system*, which we consider to be our *planning domain*.

Definition 17 *(Planning domain)*. Let $BST = \{p_1, \ldots, p_n\}$ be a finite set of belief state tuples. Our *planning domain* on *BST* is a conditional deterministic belief-state-transition system $\Sigma_{cd} = (BS, A, \theta, \gamma_{cd})$, such that:

- $BS \subseteq 2^{BST}$ is a finite set of belief states.
- *A* is a finite set of actions.
- $\theta : BS \times A \rightarrow 2^C$ is a condition function over *BS* and *A*, with the set of conditions $C = \bigcup_{i=1}^{n} \left\{ \left(v(p_i), r\right) \middle| \, r \in 2^{r(p_i)} \setminus \varnothing \right\}$.
- The transition function is $\gamma_{cd} : BS \times 2^C \times A \rightarrow 2^{BS}$ with $\gamma_{cd}(bs,c,a) = \{ ((((bs \setminus \{ (v_{bs},r_{bs}) \in bs \mid v_{bs} = v_c, (v_c,r_c) \in c \}) \cup c) \setminus \{(v_c,r_c) \in c \mid v_c = v_{effects}, (v_{effects},r_{effects}) \in effects(a)\}) \setminus \{(v_{bs},r_{bs}) \in bs \mid v_{bs} = v_{effects}, (v_{effects},r_{effects}) \in effects(a) \}) \cup effects(a) \}$ if $a \in A$ is partly applicable in $bs \in BS$ and $c \in 2^C$ is a set of conditions for *a*, and undefined otherwise.
- 2^{BS} is closed under γ_{cd}, i.e., if $bs \in BS$, then for every action *a* that is partly applicable in *bs*, and for every set of conditions $c \in 2^C$ that need to be considered, $\gamma_{cd}(bs,c,a) \in 2^{BS}$.

In contrast to former approaches like Bertoli et al. (2006), our conditions are not part of the domain, since this is not realistic at all in the context of process modeling. Now, they can be automatically derived from the domain, through the condition function and the set *C*, satisfying (R3).

Example 18. Let $(a_1, \{ (u,[3,6]), (v,\{T,F\}), (w,\{T\}), (y,\{F\}) \}, \{ (u,[2,7]), (v,\{F\}), (x,[1,4]) \})$ be an action and let $bs_2 = \{ (u,[1,5]), (v,\{T\}), (w,\{T\}), (y,\{T,F\}), (z,[5,6]) \}$ be a belief state. Here a_1 is partly applicable in bs_2, because it is $[1,5]_{u\_bs2} \cap [3,6]_{u\_a1} = [3,5] \neq \emptyset$, $\{T\}_{v\_bs2} \cap \{T,F\}_{v\_a1} = \{T\} \neq \emptyset$, $\{T\}_{w\_bs2} \cap \{T\}_{w\_a1} = \{T\} \neq \emptyset$ and $\{T,F\}_{y\_bs2} \cap \{F\}_{y\_a1} = \{F\} \neq \emptyset$. Let $c_1 = \{ (u,[3,5]), (y,\{F\}) \}$ be the set of conditions, that need to be fulfilled to execute a_1 in bs_2, as discovered in example 15. As a result we have $\gamma_{cd}(bs_2,c_1,a_1) =\{ \{ (u,[2,7]), (v,\{F\}), (w,\{T\}), (x,[1,4]), (y,\{F\}), (z,[5,6]) \} \}$.

In practice, for each belief state all partly applicable actions are determined, then the result of the condition function is calculated for each action, and at the end a new belief state is created as described by γ_{cd}. In summary, our planning problem is defined as follows.

Definition 19 *(Planning problem)*. Our *planning problem* is a triple $P = (\Sigma_{cd}, bs, BS_g)$, where:

- $\Sigma_{cd} = (BS, A, \theta, \gamma_{cd})$ is a planning domain.
- $bs \neq \emptyset$, the belief state prior to the exclusive choice, is a member of BS.
- $A_{p\_a} \subseteq A$ is a set of all actions that are partly applicable in bs.
- $BS_g \subseteq 2^{BS}$ is a set of belief states called *goal belief states* that are required to exist after the exclusive choice. The set of goal belief states is:

$$BS_g = \bigcup_{i=1}^{m} \left\{ \gamma_{cd}\left(bs, \theta(bs, a_i), a_i\right) \in 2^{BS} \;\middle|\; a_i \in A_{p\_a} \right\}, m = \left|A_{p\_a}\right|.$$

The planning problem states, that given the planning domain and the belief state bs, each goal belief state must be constructed in order to solve the problem.

III.1.5 Algorithm

The focus of this section is on the algorithm (method) that constructs the conditions and the branches, which is the realization of the condition function. Therefore, we use another algorithm as a starting point (like in Bertoli et al. 2006) that progresses from an initial belief state to a goal belief state. It builds a search tree, to find all the possible paths starting in the initial belief state. We enhance such an existing approach with our algorithm and its ability to identify the required conditions. The resulting search tree ST is a graph of a set of nodes *Nodes(ST)*, which are the belief states, and a set of labeled arcs *Arcs(ST)*. We label the arcs with both, an action and the corresponding conditions that need to be fulfilled to execute this action in the belief state.

Our focus is on how the condition function θ: $BS \times A \rightarrow 2^C$ can be realized, in other words, how the set of conditions can be built. First, we describe how the EXTENDTREE primitive, presented in Bertoli et al. (2006), can be modified to include the condition function (Figure 3 - appendix). Second, we introduce the CONDITIONFUNCTION primitive, see Figure 4, which builds the set of conditions, thus being a realization of our condition function θ. Finally, we present the PARTITION subroutine in Figure 5. This recursive subroutine creates disjoint partitions of the restrictions of certain belief state tuples in a belief state, which are needed to build the set of conditions.

The EXTENDTREE primitive receives the current search tree *ST* and a node, which is a belief state, where the tree can be extended. For every partly applicable action, including also the applicable actions, a set of conditions and the resulting next node are built. This next node is added to the set of nodes *Nodes(ST)* of the current search tree. A new arc is constructed which includes the action and the set of conditions as labels, and then added to the set of arcs *Arcs(ST)*. We go into detail on line 10 of the EXTENDTREE primitive at the end of this section.

An arc has two labels, an action and the set of conditions that need to be fulfilled in order to execute the action in a node *bs* to reach *bs'*. If an action is partly applicable, then the CONDITIONFUNCTION primitive, see Figure 4, needs to be executed. The primitive gets a node *bs* and an action *a*. If *a* is applicable in *bs* then, there is no need for conditions and the empty set is returned. On the other hand, if *a* is just partly applicable (and not applicable), then a set of conditions c_a is created for *a*. The lines 6-10 create A_{p_a}, the set of all partly applicable actions in *bs*. The set bs_p is the set of all belief state tuples in *bs*, where the belief state variable is also in the preconditions of *a* (line 11). The rest of the primitive is then performed for every belief state tuple in bs_p.

We take one element (v_{bs},r_{bs}) of bs_p. The set *Partition* of nonempty sets is a partition of r_{bs}. The elements of *Partition* are pairwise disjoint and the elements of *Partition* cover r_{bs}. The Partition subroutine creates the set *Partition*. In line 15, the set c_{part} is built, which contains those elements of *Partition* that are a subset of r_w, being the restriction of a belief state tuple v_{bs} in the preconditions of *a*. As a next step, the elements of c_{part} are joined to form c_p. This set c_p is the condition for v_{bs} that, possibly among others, needs to be fulfilled in order to execute *a*. In the end, c_p is joined with c_a to construct the set of conditions, which is returned, after lines 13-21 are carried out for every element of bs_p. The

CONDITIONFUNCTION primitive creates the minimal amount of conditions, with the minimal sets of observations (due to lack of space, we omit the proof). This leads to a minimal quantity of exclusive choices in the process model, which is advantageous for its presentation and layout.

In line 14 of the CONDITIONFUNCTION primitive the PARTITION subroutine is carried out, which is defined recursively. This subroutine starts with two sets. The set r is a restriction r_{bs} of a belief state tuple (v_{bs},r_{bs}). The second set R is a collection of restrictions. It is the set of all restrictions for v_{bs} that are part of the preconditions of each action in A_{p_a}. This is done in order to partition r_{bs} in pairwise disjoint sets so that the set c_p in the CONDITIONFUNCTION primitive can be constructed. It is possible that there are subsets of r_{bs}, which are not covered by any precondition of the partly applicable actions. That is to say, there are subsets of r_{bs}, where the intersection of these subsets and the union of the sets in R are empty. To handle this case, we always add an arc in the EXTENDTREE primitive (line 10) that leads to the termination of the plan if the values in the union of these subsets are observed in the world state at execution time.

As mentioned, the focus of this section is on the algorithm that plans the conditions and the branches. The conditions are not given additionally to the domain in any way, but are created by the algorithm, as required by (R3).

III.1.6 Evaluation

The presented algorithm was implemented prototypically as part of the open source process modeling tool AgilPro. The algorithm and the prototypical implementation (instantiation) were evaluated as shown in this section.

a) Analysis of the algorithm properties: We mathematically evaluated the algorithm in terms of completeness, minimality, termination and computational complexity regarding time. It could be shown that the approach creates complete results and a minimal number of exclusive choices in a process model for a given problem. Furthermore, the algorithm terminates. Considering the computational complexity regarding time, it can be shown that, for example, the runtime increases subproportionaly when the initial belief state or the goal belief states are extended. Due to lack of space we only show that the algorithm terminates:

Theorem 20. Given our planning domain, the execution of the algorithm EXTENDTREE terminates.

Proof. Termination is proved by showing that each iteration of every for-loop in the algorithm terminates, and that the number of iterations is finite. If *a* is applicable in *bs*, then the CONDITIONFUNCTION terminates (lines 2-3). The else-case (lines 4-23) is complex.

Before starting with the CONDITIONFUNCTION, we show that the set *Partition* in line 14 is finite. The set *Partition* is built by the subroutine PARTITION, presented in Figure 5. The set R of the PARTITION subroutine equals to the set $\{r_u \mid v_{bs}=v_u,\ (v_u,r_u) \in precond(a_{p\_a}),\ a_{p\_a} \in A_{p\_a}\}$ when it is invoked the first time. The set $precond(a_{p\_a})$ is finite for each $a_{p\_a} \in A_{p\_a}$, due to the fact that it is a subset of *BST*. The set of all actions A is finite, and so is A_{p_a}, because it is a subset of A. As $precond(a_{p\_a})$ and A_{p_a} are finite, the set R is finite, when PARTITION is invoked the first time. Every line of the PARTITION subroutine, besides 8 and 11, terminates, simply because there are just set operations. The subroutine is invoked recursively only a finite number of times. This is due to the following three facts:

- R is finite.
- The number of elements in R decreases by one every time the subroutine is invoked recursively.
- The subroutine is invoked only when there is more than one element in R, representing a lower bound.

In other words, the subroutine is invoked recursively just a finite number of times, because the set R is finite, the number of elements is strictly decreasing and there is a lower bound. Thus, the PARTITION subroutine terminates and the set *solution* is finite. The set *solution* is returned to the CONDITIONFUNCTION, constituting the finite set *Partition*.

With the set *Partition* being finite, it is possible to show that the CONDITIONFUNCTION terminates. Lines 7-9 terminate, because the verification whether a_p is partly applicable in *bs* or not terminates and the procedure of uniting A_{p_a} with a_p terminates. The for-loop of lines 6-10 terminates, since A is finite. We show that lines 14-20 terminate. The set c_{part} is a subset of *Partition*, which makes c_{part} finite. Line 18 terminates, because the procedure of uniting *part* with c_p terminates. Since c_{part} is finite, lines 17-19 terminate. The operation of uniting c_a with c_p terminates. To show, that the number of iterations of lines 14-20 is finite, we show that bs_p is a finite set of belief state tuples. The set *BST* of all belief state tuples is finite. The node *bs* is a subset of *BST*, which makes it finite as well. The set bs_p is a subset of *bs*, and is therefore finite. Thus, the for-loop of lines 13-21 terminates, and so does the CONDITIONFUNCTION. Since the set *A* is finite the EXTENDTREE primitive terminates. *q.e.d.*

b) Analysis of the implementation: Besides the manual analysis of the source code (structured walk through) by persons other than the programmers, we made a series of tests using the JUnit Framework, including runs with extreme values, JUnit regression tests and unit tests. The implemented algorithms did not show any defects at the end of the test phase.

c) *Formal evaluation of the results:* It can be shown that the constructed exclusive choice is syntactically correct. As stated in Sadiq and Orlowska (2000), an exclusive choice needs to be "exclusive and complete", which means that in each process instance exactly one of the alternative partly applicable actions is executed or the process ends. This is guaranteed due to the facts that the conditions of the partly applicable actions are pairwise disjoint and that these conditions cover the whole domains of the respective belief states. Further forms of evaluation, as for example the proof that no data-flow anomalies exist (Sun et al. 2006) and that the resulting process model fulfills the soundness property as proposed by van der Aalst (2000), are not applicable for an exclusive choice itself (we conducted such evaluations in other papers considering the overall planned process model).

d) Defined Requirements: We presented an abstract representation language in order to explicitly represent a possibly infinite set of world states in the form of belief states, providing an intuitive formalism (from a process modeling perspective) for the planning problem. With this language, we addressed the requirements (R1) and (R2). In our approach, belief states are defined as sets of feasible values of belief state variables and thus implicitly describe sets of conceivable world states. On this basis we constructed a conditional deterministic belief-state-transition system, what we considered to be our planning model. In this planning model we had a concrete representation of our transition function, which also accounts for sets of conditions (cp. (R3)).

e) Operational evaluation of the results: Hevner et al. (2004) stressed that an artifact must be evaluated with respect to the practical utility provided. Since competing artifacts do not exist in our case, a comparison related to efficiency is not possible at all. This in mind, we analyzed the practical applicability in different real-use situations ("proof of construction"). One of those situations – which we consider in the following – is an example taken from the security-order-management of a European financial services provider. Here, processes had to be redesigned in the past due to new products, new regulations or changing organizational requirements (like to outsource parts of a process to external service providers). These processes have to be not only (re)designed repeatedly but are also repetitive processes which are of high value for the firm. We analyzed such previous redesigns and studied two aspects in detail. Firstly, would it be possible at all to apply our algorithm in these situations and to which extent match the results of the automated planning with manually built exclusive choice patterns? And secondly, does the application of the algorithm make sense regarding economical aspects, i.e. do we benefit from the automated planning of process models compared to a manual design and what costs do result from the artifact's application?

Considering the first aspect, it can be said, that the algorithm constructed not only the exclusive choices resulting from the manual redesign, but also additional feasible solutions. For instance, the resulting process was planned for the execution of security-orders where several steps including check routines had to be modeled. For brevity, we only present a small part of the whole process, where the security-order data is entered, the order itself is already validated and we now need to decide which check routine should be used (see Figure 2). The resulting belief state *bs* contains, among others, the tupel *orderAmount* that reaches from *0* to *250,000* Euro and the tuple *orderState* with the values *valid* and *invalid*. We now illustrate the input of the PARTITION subroutine, the result of that subroutine, and how this result is used in the CONDITIONFUNCTION primitive.

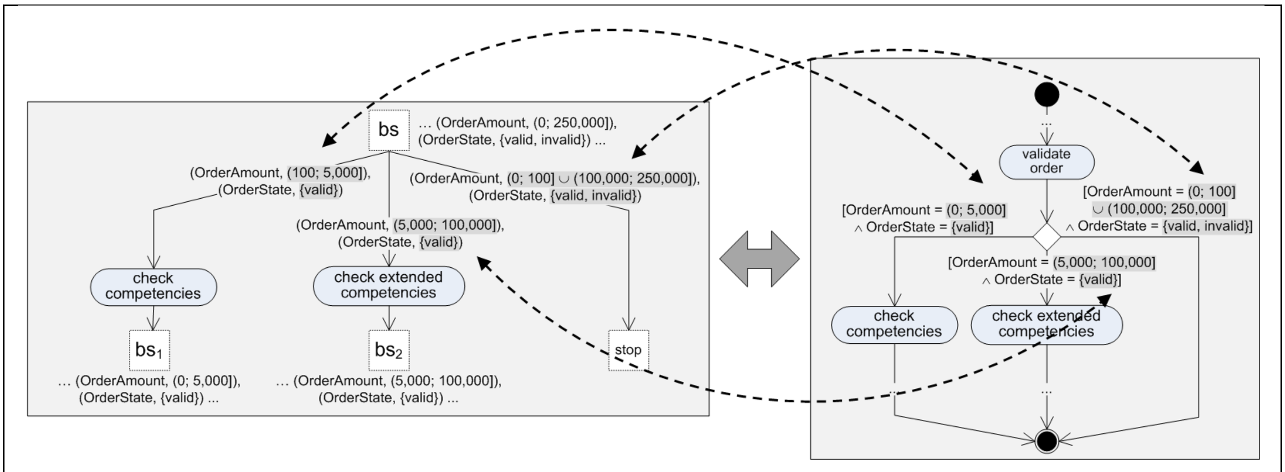


Figure 2. Transformation of a Search Tree ST (Left Side) to an UML Activity Diagram

The partly applicable actions in *bs* are *checkCompetencies* with *(orderAmount, (100; 5,000]), (orderState, {valid})* and *checkExtendedCompetencies* with *(orderAmount, (5,000; 100,000]), (orderState, {valid})*. For the creation of the respective conditions (only shown for the action *checkCompetencies*), the subroutine PARTITION gets *r = [0; 250,000]* and *R = { (0; 5,000], (5,000; 100,000] }* in its first execution and provides *solution = { (100; 5,000], (5,000; 100,000], (0; 100] ∪ (100,000; 250,000] }*. The second invocation of PARTITION (*r = {valid; invalid}* and *R = {valid}*) returns *solution = { {valid}; {invalid} }*. With these partitions, the CONDITIONFUNCTION determines for each of the partly applicable actions, the necessary condition that is *(orderAmount, (100; 5,000]), (orderState, {valid})* for *checkCompetencies* and *(orderAmount, (5,000; 100,000]), (orderState, {valid})* for *checkExtendedCompetencies*. The third condition *(orderAmount, (0; 100] ∪ (100,000; 250,000]), (orderState, {valid; invalid})* that needs to be considered does not lead to an executable action but to the end of the process as there is no executable action for it.

The assessment showed the practical applicability of the algorithm in some real-use situations. However, what about the efficiency of such applications? Here, we have to differentiate: To conduct the algorithm, an initial annotation of actions according to their preconditions and effects is necessary. These one-time annotation costs are limited, if a firm already uses a process modeling tool featuring a XML interface. Such an interface can be used in order to export actions to AgilPro. The financial services provider we consider here used, for instance, the ARIS toolset, which allowed us to export a huge number of actions. In the field of the security-order-management about 200 different actions including their preconditions and effects were imported from the ARIS toolset and afterwards checked (semantically). This review comprised for example the check for completeness of the action’s preconditions and effects and on a semantic level if the same parameter names of the preconditions and effects were used for different concepts as well as if different parameter names were used for the same concepts. In some cases, preconditions and effects had to be completed or corrected. Such a review, which is to some extent also necessary during manual process design, ought to evidently be performed more accurately for automated process planning. Annotating ten actions with averagely five preconditions and five effects in about half an hour, the net time for the annotation of actions in our AgilPro tool is comparatively small. For this purpose, the relevant actions have to be defined and labeled before their preconditions and effects are defined and semantically annotated. There, each precondition and effect is specified by choosing the mapping concept from the ontology and determining the appropriate restriction. All in all, the initial costs of automated planning were about 20% higher than the previous costs of manual design.

However, the resulting annotation costs need to be put into perspective, if a redesign of the processes considered is necessary more than just one time (see also the specified problem context in the second section). In future redesigns of the processes the *initial* costs are lower, because the annotations as well as the deployed algorithm can be reused. This even holds for the common case, where annotated actions (e.g. an action "*scan document*") can be reused in further processes (like for example in the loan department). Therefore, it can be seen, that especially under the condition of a higher frequency of process redesigns, an automated planning of processes models leads in sum to slightly higher costs compared to a manual redesign.

Furthermore, in our above mentioned real-use situation considering the security-order-management, a set of feasible process models was generated within one day by means of AgilPro. This is an obvious advantage compared to a manual design, which took more than one week. Also, the planner generated not only the manual designed process model, but other feasible solutions as well. Some of these process solutions need, for instance, less staff capacity of the financial services provider than the manual designed process. This leads to lower process costs of about 2.5% in average for each process run (calculation basis: process *'execution of security-orders'*). Since the security-order-management process is a highly repetitive process, the higher initial costs of automatic planning can be amortized in our case within half a year. And, according to our projects with firms, these circumstances are not unique in practice. But, in any other case, an individual analysis is really necessary to assess if an automated planning is useful. In addition, we have to evaluate another point regarding economical aspects: Several actions – as described above – are not used within only one process, but are reused in other processes, too. Thus, if an action is used several times, its annotation costs can be allocated to all redesigns of those processes using the considered action. However, such economical analysis cannot be done for only a few real-use situations but in medium- or long-term studies. Therefore, the generalizability of our evaluation of the practical utility is limited. Nonetheless, at this time, this limitation is unavoidable since such an iterative study is time-intensive, mitigating the possibility of conducting multiple real-use situations simultaneously. We anticipate that these new cases will support the relevance identified in this paper.

III.1.7 Conclusion

In this paper, we described how control structures can be planned automatically within the research strand of Semantic Business Process Management. Related to the guidelines for conducting design science research by Hevner et al. (2004) we can summarize as follows: Our key *artifact* is a method in terms of an algorithm for planning exclusive choices within a process model. We regard this as an important step to automate and hence to support the task of designing a process model. Both, the algorithm and our planning problem are formally noted and can therefore be well-defined and mathematically evaluated. Based on statements in literature (e.g. Borges et al. 2005; Ma and Leymann 2008) and on our own project expectations that manual process modeling is cost-intensive and very time-consuming, we describe our *problem context*. Here, our artifact is thought to contribute to process modeling to design and adapt process models faster and to be useful regarding economical aspects. Since such a statement cannot hold for every process type, we concentrate on repetitive processes that need to be (re)designed repeatedly. Considering the real-use situations, in which we applied the algorithm, we found that this focus is reasonable. The *evaluation* was done on the one hand by mathematical methods, but not in comparison with competing artifacts, since the artifact solves a heretofore unsolved problem. On the other hand, we studied literature and derived key requirements that a planning algorithm should meet in our problem context. These requirements were not only the guideline when developing our artifact. Moreover the artifact was evaluated against the defined requirements. Additionally, we evaluated the algorithm in real-use situations with respect to its applicability and the practical utility provided. This is appropriate since it analyzes the planning algorithm “in depth in business” (Hevner et al. 2004). Nevertheless, considering such economical analysis, future work is needed and intended to support the assessment and justification of an automated planning. The paper points out, that existing algorithms have some serious disadvantages, especially regarding the requirements derived from literature. So, they are not appropriate to our context. The *research contribution* of our approach is to avoid these problems and meet the defined requirements. Therefore, our paper fills a gap in science and practice. To support a *rigorous* definition of our artifact, we represented it formally based on an also formally denoted planning domain. Such a technical representation assists a mathematical evaluation of the algorithm, too. The *search process* is on the one hand directed by the requirements. On the other hand, we describe this process beginning with the abstract representation language and its advantages. Furthermore, we show in detail which steps are necessary to develop our artifact (see section planning model). Regarding the *communication* of our results, we choose a more technical, rigorous presentation, because we want to convincingly demonstrate that our artifact can be realized and implemented. However, we also tried to attract a managerial audience by means of the illustrated business problem context as well as the economic aspects of automated planning process models. Further work is proposed on the question of how other control flow patterns, like arbitrary cycles, can be considered. For this, the designed algorithm is a reliable basis.

III.1.8 Appendix

```
procedure ExtendTree(bs, ST)
  forall a ∈ A
    if partly_applicable(a, bs) then
      c := ConditionFunction(bs, a)
      bs' := γcd(bs, c, a)
            Nodes(ST) := Nodes(ST) ∪ {bs'}
            Arcs(ST) := Arcs(ST) ∪ ⟨bs, a, c, bs'⟩
    endif
  endfor
  Arcs(ST) := Arcs(ST) ∪ ⟨bs, else, stop⟩
end
```

Figure 9. Node Expansion Primitive

```
function CONDITIONFUNCTION(bs, a)
   if applicable(a, bs) then
      return ∅
   else
      A_p_a := ∅
      forall a_p ∈ A
         if partly_applicable(a_p, bs) then
            A_p_a := A_p_a ∪ a_p
         endif
      endfor
      bs_p := {(v_bs,r_bs) ∈ bs | v_bs = v_pre, (v_pre,r_pre) ∈ precond(a)}
      c_a := ∅
      forall (v_bs,r_bs) ∈ bs_p
         Partition := PARTITION(r_bs,{r_u|v_bs=v_u,(v_u,r_u)∈ precond(a_p_a),a_p_a∈ A_p_a})
         c_part := {part ∈ Partition | part ⊆ r_w, v_bs = v_w,(v_w,r_w)∈ precond(a)}
         c_p := ∅
         forall part∈ c_part
            c_p := c_p ∪ part
         endfor
         c_a := c_a ∪ c_p
      endfor
      return c_a
   endif
end
```

Figure 10. Condition Primitive

```
function PARTITION(r, R)
   nondeterministically choose rest ∈ R
   diff := r \ rest
   intersection := r ∩ rest
   solution = ∅
   if |R|>1 then
      if diff ≠ ∅ then
         solution := PARTITION(diff, R \ rest)
      endif
      if intersection ≠ ∅ then
         solution := solution ∪ PARTITION(intersection, R \ rest)
      endif
   else
      solution := {diff, intersection} \ {∅}
   endif
   return solution
end
```

Figure 11. Subroutine of the Condition Function

III.1.9 References

Becker, J., and Kahn, D. 2003. "The process in focus," in Process Management. A Guide for the Design of Business Processes, J. Becker, M. Kugeler and M. Rosemann (eds.), Berlin: Springer, pp. 1-12.

Bertoli, P., Cimatti, A., Roveri, M., and Traverso, P. 2006. "Strong planning under partial observability," Artificial Intelligence (170:4), April, pp. 337-384.

Betz, S., Klink, S., Koschmider, A., and Oberweis, A. 2006. "Automatic user support for business process modeling," in Proceedings of the 3rd European Semantic Web Conference (ESWC 2006), K. Hinkelmann, D. Karagiannis, N. Stojanovic and G. Wagner (eds.), June 2006, pp. 1-12.

Biron, P. V., and Malhotra, A. 2004. "XML Schema Part 2: Datatypes Second Edition," 04/30/2009, (http://www.w3.org/TR/xmlschema-2/).

Bonet, B., and Geffner, H. 2000. "Planning with Incomplete Information as Heuristic Search in Belief Space," in Proceedings of the 5th International Conference on Artificial Intelligence Planning Systems (AIPS 2000), S. Chien, S. Kambhampati and C. A. Knoblock (eds.), April 2000, pp. 52-61.

Bonet, B., and Geffner, H. 2001. "GPT: A Tool for Planning with Uncertainty and Partial Information," in Proceedings of the 17th International Joint Conference on Artificial Intelligence (IJCAI 2001), B. Nebel (ed), August 2001, pp. 82-87.

Borges, M. R. S., Pino, J. A., and Valle, C. 2005. "Support for decision implementation and follow-up," European Journal of Operational Research (160:2), January, pp. 336-352.

Brockmans, S., Ehrig, M., Koschmider, A., Oberweis, A., and Studer, R. 2006. "Semantic alignment of business processes," in Proceedings of the 8th International Conference on Enterprise Information Systems (ICEIS 2006), Y. Manolopoulos, J. Filipe, P. Constantopoulos and J. Cordeiro (eds.), May 2006, pp. 191-196.

Constantinescu, I., Faltings, B., and Binder, W. 2004. "Large scale, type-compatible service composition," in Proceedings of the IEEE International Conference on Web Services (ICWS 04), July 2004, pp. 506-513.

Geffner, H. 2002. "Perspectives on Artificial Intelligence Planning," in Proceedings of the Eighteenth National Conference on Artificial Intelligence (AAAI-02), July 2002, pp. 1013-1023.

Ghallab, M., Nau, D., and Traverso, P. 2004. Automated Planning - Theory & Practice, San Francisco: Morgan Kaufmann.

Gregor, S., and Jones, D. 2007. "The Anatomy of a Design Theory," Journal of the Association of Information Systems (8:5), May, pp. 312-335.

Heinrich, B., Bewernik, M., Henneberger, M., Krammer, A., and Lautenbacher, F. 2008. "SEMPA - A Semantic Business Process Management Approach for the Planning of Process Models," Business & Information Systems Engineering (formerly WIRTSCHAFTSINFORMATIK) (50:6), pp. 445-460 (in German).

Henneberger, M., Heinrich, B., Bauer, B., and Lautenbacher, F. 2008. "Semantic-Based Planning of Process Models," in Proceedings of the Multikonferenz Wirtschaftsinformatik (MKWI) 2008, M. Bichler, T. Hess, H. Krcmar, U. Lechner, F. Matthes, A. Picot, B. Speitkamp and P. Wolf (eds.), February 2008, pp. 1677-1689.

Hepp, M., and Dumitri, R. 2007. "An ontology framework for semantic business process management," in Proceedings of the 8th International Conference on Business Informatics (WI 2007): eOrganisation: Service-, Prozess-, Market-Engineering, A. Oberweis, C. Weinhardt, H. Gimpel, A. Koschmider, V. Pankratius and B. Schnizler (eds.), February 2007, pp. 423-440.

Hepp, M., Leymann, F., Domingue, J., Wahler, A., and Fensel, D. 2005. "Semantic Business Process Management: A Vision Towards Using Semantic Web Services for Business Process Management," in Proceedings of the IEEE International Conference on e-Business Engineering (ICEBE 2005), W. Tsai, J. Chung and M. Younas (eds.), October 2005, pp. 535-540.

Hevner, A. R., March, S. T., Park, J., and Ram, S. 2004. "Design Science in Information Systems Research," MIS Quarterly (28:1), March, pp. 75-105.

Hoffmann, J., and Brafman, R. I. 2006. "Conformant planning via heuristic forward search: A new approach," Artificial Intelligence (170:6), May, pp. 507-541.

Hoffmann, J., and Brafman, R. I. 2005. "Contingent Planning via Heuristic Forward Search with Implicit Belief States," in Proceedings of the 15th International Conference on Automated Planning and Scheduling (ICAPS 05), S. Biundo, K. L. Myers and K. Rajan (eds.), June 2005, pp. 71-80.

Hornung, T., Koschmider, A., and Oberweis, A. 2007. "Rule-based Autocompletion of Business Process Models," in CAiSE Forum 2007, Conference on Advanced Information Systems Engineering, June 2007.

Liang, Q., and Su, S. 2005. "AND/OR Graph and Search Algorithm for Discovering Composite Web Services," in International Journal of Web Services Research (2:4), October, pp. 48-67.

Ma, Z., and Leymann, F. 2008. "A Lifecycle Model for Using Process Fragment in Business Process Modeling," Proceedings of the 9th Workshop on Business Process Modeling, Development, and Support (BPDMS 2008) in Conjunction with the 20th International Conference on Advanced Information Systems Engineering (CAiSE'08), June 2008, pp. 1-9.

March, S. T., and Smith, G. F. 1995. "Design and natural science research on information technology," Decision Support Systems (15:4), December, pp. 251-266.

Marjanovic, O. 2005. "Towards IS supported coordination in emergent business processes," Business Process Management Journal (11:5), pp. 476-487.

Meyer, H., and Kuropka, D. 2005. "Requirements for Service Composition," Technische Berichte Nr. 11 des Hasso-Plattner-Instituts fuer Softwaretechnik an der Universität Potsdam.

Meyer, H., and Weske, M. 2006. "Automated Service Composition using Heuristic Search," in Proceedings of the 4th International Conference on Business Process Management (BPM 2006), S. Dustdar, J. L. Fiadeiro and A. Sheth (eds), September 2006, pp. 81-96.

Motik, B., Patel-Schneider, P. F., and Horrocks, I. 2008. "OWL 2 Web Ontology Language: Structural Specification and Functional-Style Syntax," W3C working draft, April at http://www.w3.org/TR/owl2-syntax/, last called: 2009-09-07.

Pathak, J., Basu, S., and Honavar, V. 2006. "Modeling Web Services by Iterative Reformulation of Functional and Non-Functional Requirements," in Proceedings of the fourth International Conference on Service Oriented Computing (ICSOC 2006), A. Dan and W. Lamersdorf (eds.), December 2006, pp. 314-326.

Petrick, R. P. A., and Bacchus, F. 2002. "A Knowledge-Based Approach to Planning with Incomplete Information and Sensing," in Proceedings of the Sixth International Conference on AI Planning & Scheduling (AIPS 2002), M. Ghallab, J. Hertzberg and P. Traverso (eds.), April 2002, pp. 212-221.

Pistore, M., Traverso, P., Bertoli, P., and Marconi, A. 2005. "Automated Synthesis of Composite BPEL4WS Web Services," in Proceedings of the IEEE International Conference on Web Services 2005 (ICWS2005), July, pp. 293-301.

Recker, J., Rosemann, M., and van der Aalst, W. M. P. 2005. "On the User Perception of Configurable Reference Process Models - Initial Insights," in Proceedings of the 16th Australasian Conference on Information Systems (ACIS'2005), B. Campbell, J. Underwood and D. Bunker (eds.), November 2005, pp. 1-10.

Sadiq, W., and Orlowska, E. 2000. "Analyzing Process Models Using Graph Reduction Techniques," in Information Systems (25:2), April, pp. 117-134.

Sun, S. X., Zhao, J. L., Nunamaker, J. F., and Sheng, O. R. L. 2006. "Formulating the Data-Flow Perspective for Business Process Management," Information Systems Research (17:4), December, pp. 374-391.

Thomas, O., and Fellmann, M. 2007. "Semantic business process management: ontology-based process modeling using event-driven process chains," International Journal of Interoperability in Business Information Systems (2:1), pp. 29-43.

van der Aalst, W. M. P. 2000. "Workflow Verification: Finding Control-Flow Errors Using Petri-Net-based Techniques," in Business Process Management: Models, Techniques, and Empirical Studies, W. M. P. van der Aalst, J. Desel and A. Oberweis (eds.), Berlin: Springer, pp. 161-183.

van der Aalst, W. M. P., ter Hofstede, A. H. M., Kiepuszewski, B., and Barros, A. P. 2003. "Workflow patterns," Distributed and Parallel Databases (14:3), July, pp. 5-51.

Weigand, H., van den Heuvel, W.-J., and Hiel, M. 2008. "Rule-based service composition and service-oriented business rule management," in Proceedings of the International Workshop on Regulations Modelling and Deployment (ReMoD'08), J. Vanthienen and S. Hoppenbrouwers (eds.), June 2008, pp. 1-12.

Weske, M., van der Aalst, W. M. P., and Verbeek, H. M. W. 2004. "Advances in business process management," Data & Knowledge Engineering (50:1), July 2004, pp. 1-8.

Wu, D., Sirin, E., Hendler, J., Nau, D., and Parsia, B. 2003. "Automatic Web Services Composition Using SHOP2," in Proceedings of Planning for Web Services Workshop in ICAPS 2003, June 2003.

IV Results and Future Research

In this section, the key findings of the doctoral thesis (Section IV.1) and the potential for future research are presented (Section IV.2).

IV.1 Results

The main objective of this doctoral thesis is to contribute to the field of BPM by focusing in particular on the design phase of the BPM lifecycle from a business perspective as well as from an IT perspective. After emphasizing the importance of the design phase of the BPM lifecycle, the doctoral thesis focuses on a value-based and automated process design. Regarding value-based process design, the research papers focus on introducing the principles of value-based management to the operational process level, connecting the business model with the processes from a business perspective. Regarding automated process design, the research paper focuses on contributing to an IT-enabled process design, taking on an IT perspective on BPM. In the following, the key findings of the research papers that are included in this doctoral thesis are presented. At the end, future research opportunities are discussed.

In Section II.1, the first result is to make process improvement decisions at the process level that are in the interest of a company as a whole. This means, at the process level it is possible to decide if a process alternative should be realized or if the existing process should remain unchanged. For this, the principles of value-based management are introduced in more detail to BPM, so that the improvement decisions at the process level are in line with the value creation goal of a company. In order to decide at the process level in a value-based manner, it is presented how a company can be regarded as a portfolio of processes and that the value of a single process is reflected by its (uncertain) net present value. This way it is shown in what way a potential change in process value, due to a potential improvement, influences the company value. One principle of value-based management is the inclusion of risk within the valuation of a process, which is due to the uncertainty of the net present value of a process. It is shown how the risk that a process contributes to a company's overall risk can be measured and how it can be combined with the expected return of a process, which is the expected net present value of a process, to determine the process value. Moreover, when combining the expected return of a process and the risk of the process, the risk attitude of a company/person in charge can be considered as well, which is in line with decision theory.

In particular, it is shown that it is not necessary to determine the absolute value of an existing process and of a process alternative to determine if the process alternative should be implemented, but that it is enough to know the difference in the expected returns and the difference in the risks in order to decide if the process alternative should be realized. All of this sets the stage for the improvement of processes from a business perspective and connects the business model with the processes of a company.

While Section II.1 adds to BPM in formally showing how the value of a process is connected to the value of a company, defining the value contribution of a process to the company, Section II.2 presents in detail how the absolute value of a process is determined. For this, the value of a process is formalized as a risk-adjusted expected net present value (rNPV). In a next step the rNPV is further detailed to the single execution of a process, the process instance. In doing so, single process executions are connected with the company value. Such level of detail enables the valuation calculus of Section II.2 to incorporate the structure of processes, which is achieved by modeling a process as a probability space, the so-called process-probability-space. This formalism allows the mathematically sound calculation of the expected value and the variance of the process cash flow. To this point, in BPM, these values were determined via simulations, which can be advantageous, but also have some drawbacks, as discussed in Section II.2. Furthermore, the given valuation calculus already provided helpful insights for the CubeFour company to correct the calculation logic of their commercial process-modeling tool. These insights also showed a challenge when calculating the expected value and the variance of the process cash flow of a process that includes the OR-split workflow pattern. Moreover, it even points to the possibility that ordinary process simulation tools provide a wrong expected value and variance of the process cash flow, and any other additive quantity, like time, as certain additional information, which is needed to consider OR-splits, is not requested in these process simulation tools.

Section II.1 and Section II.2 connect the business model layer with the process layer in a conceptual way from a business perspective. This means, the company value is connected with the cash flows that are caused by process instances. While the rNPV can be used to valuate processes and to determine if a process alternative adds more value to a company than another process alternative or the existing process, it does not give concrete suggestions to process managers on how to build process alternatives that have a higher rNPV than the existing process. Section II.3 therefore provides concrete guidance on how to improve a process in a value-based manner. For this, the expected value of the process cash flow is used as an objective function that needs to be maximized. This expected value depends on the probability that a certain process path is executed by a process instance. This probability in turn depends on how a process instance is routed through a process. The routing of a process instance is mainly influenced by the conditions at the decision points of a process. Section II.3 shows how to set the parameters of these conditions in order to maximize the expected process cash flow. For this, the conditions of process models are transferred into a formal objective function. Thus, concrete guidance is given to process managers on how to increase the value of a process. In addition to giving guidance for process managers, the process improvement does not require a large re-engineering project but only a minimally invasive project that changes the condition parameters. This reduces the risk of improvement failure.

Summarizing Section II, it can be concluded that a value-based process design can be beneficial for companies in order to align their processes to the business model from a business perspective, contributing to the value of the company.

While Section II focuses on the goal of the process design, Section III focuses on the duration of the design phase as well as on the quality of process models that are commonly used during a design phase to visualize and specify processes. The main contribution of Section III to the body of BPM knowledge is an algorithm to create the control flow pattern “exclusive choice”. When incorporating this algorithm in approaches that aim to create whole process models in an automated manner, then more processes can be modeled automatically. This reduces the time of the design phase and increases the quality of process models. In order to develop this algorithm, the so-called abstract representation language is developed initially to establish the necessary constructs for the planning model. The planning model, which includes the definition of the planning domain and the planning problem, is based on existing literature from the area of artificial intelligence. In two steps, this existing planning model is extended to cope with the properties of the pattern “exclusive choice”, mainly with the conditions associated with this pattern. This planning model allows the development of the algorithm.

Taking the major results of the research papers within Sections II and III together, this doctoral thesis contributes to the existing literature in BPM research by its particular focus on a value-based and automated process design. Most notably, it complements previous research by explicitly connecting the company value with the process value and by providing the valuation calculus to determine the process value. In addition, it presents an algorithm to automate the modeling of processes. However, despite the presented results, there remain challenges, which offer starting points for future research.

IV.2 Future Research

In the following, potential aspects for future research are highlighted for all research papers that are included in this doctoral thesis. In Section II, regarding value-based process design, there are the following aspects for further research:

1. When determining the value that a process contributes to a company, the rNPV is used. The presented valuation calculus can be used in particular for processes that are executed independently from other processes. However, this independency is most likely not given in a real-world setting. Thus, there is a need for further research on the types of dependencies among processes and on how this can be incorporated into the valuation of processes. One possibility might be the use of correlation coefficients.
2. Value-based process design is based on monetary values. It uses cash flows as the common denominator on which the focus should be placed. However, typically in BPM the dimensions to evaluate processes are time, costs, quality, and flexibility. While costs are already a monetary value and time can be monetized in a straightforward manner, for example via wages, it is more difficult for the other dimensions. If these dimensions, and others, like customer satisfaction, can be properly transformed in monetary values, then value-based BPM can provide a framework that can support decisions in line with the value creation goal of a company.
3. The valuation calculus to determine the rNPV of a process considers the structure of a process. However, the structure of a process can be very complex, which may render the manual determination of the rNPV impossible, even though it is theoretically possible. Hence, there needs to be research on algorithms that can determine the rNPV. Only then it is possible to determine the rNPV for complex processes, even though the runtime of the algorithms might be high.
4. The presented valuation calculus focuses in particular on the structure of a process. Thus, mainly the control flow perspective on a process is considered, potentially leaving out, for example, a data and function perspective. Such perspectives can add more complexity

to a process that the valuation calculus may not be able to cope with. Thus, it needs to be examined if the valuation calculus is sufficient regarding these perspectives or if it needs to be extended.

5. Section II.3 shows how concrete guidance can be given on the improvement of processes while using the expected process cash flow as objective function. However, this does not consider the risk associated with this cash flow. In addition, this does not consider multiple periods, which are part of the rNPV as objective function. Therefore, future research needs to consider the rNPV.
6. The introduced decision model in Section II.3 is designed to be usable for a variety of processes. In Section II.3 the decision model is applied to the provided sample process, which is successfully achieved for the given case. However, while applying the model, it seemed that the application of the model might be more difficult for more complex processes. Therefore, the decision model should be applied to more cases, in particular with regard to more complex processes, to study the behavior of the model in more detail and to validate its applicability for improving processes.
7. Similar to the need for an algorithm to determine the rNPV, a software tool is needed that implements the decision model of Section II.3. The software tool should be able to handle complex processes. If there is an interface with a workflow management system, this tool could analyze the data in real time, continuously improving a process. This could also lead to more insights into the model, showing possibilities to further enhance the decision model.

While these are possibilities for future research in the area of value-based process design, there are the following possibilities for future research in the area of automated process design, discussed in Section III.

1. Although there is work that valuates manual against automated process design from a business perspective, showing evidence that there is value in automated process design, there is need for more economic analyses to justify automated process design. In particular, this is true when designing processes automatically that are not highly repetitive in nature. If there is also evidence that justifies the use for other processes as well, the area of application would increase.
2. Section III adds to the research in automated process design in a very specialized way. The focus is on one workflow pattern only. Although this shows the complexity of an

automated process modeling of a whole process, because already one pattern is complex to consider, there is research needed on other workflow patterns. Only if all workflow patterns can be considered will a fully automated process modeling be achieved.

Taken together, the research papers presented in this doctoral thesis contribute to BPM in that the principles of value-based management are introduced in detail to BPM and, in that, one particular part of process models can be created automatically. Though this doctoral thesis cannot answer all questions and challenges regarding a value-based and automated process design, it complements previous work in this area. As BPM, and in particular process improvement, is expected to continue to play a prominent role, the hope is that this doctoral thesis can provide researchers and companies with helpful insights in this area of BPM to face the challenges of an ever-changing environment.