Supporting BPMN Model Creation with Routing Patterns

Idan Wolf and Pnina Soffer

Information Systems Department, University of Haifa, A. Hushi 199
Haifa 3498838, Israel
idanwf@gmail.com, spnina@is.haifa.ac.il

Abstract. Business process modelers often struggle with appropriately representing routing situations in a model. In particular, difficulties may be encountered when using BPMN, due to its large number of constructs and the lack of ontological clarity of this language.

The paper proposes routing patterns combined with a decision guidance tool to support BPMN model creation. The use of patterns is proposed based on cognitive considerations, which are explained to provide justification to the proposed support. The set of patterns builds on an existing set of routing behaviors and operationalizes these behaviors by providing their BPMN representations. The effect of this support is tested in a study, whose findings indicate a significant effect on the quality of the produced models. The findings also indicate that the use of the guided routing patterns leads to a longer time required for modeling as compared to unsupported modeling.

Keywords: BPMN, Routing patterns, Process modeling, Empirical study

1 Introduction

Business process models play an important role in the development of business processes and information systems. The creation of a business process model requires gaining an understanding of the domain and specifying its required behavior using some process modeling language. The resulting model should be syntactically correct (correctly using the modeling language), logically and semantically correct (truthfully representing the behavior of the domain and lacking logical errors such as deadlocks), and understandable to its readers. These quality requirements make process modeling a challenging task.

A particularly challenging task in process modeling is the appropriate construction of routing structures. Routing structures include split nodes, where the thread of control is split into several threads that can be taken alternatively or in parallel, and merge nodes, where several threads are merged into a single one. Empirical evidence show that such structures are associated with difficulties both in model reading (are more difficult to understand [4, 16]) and in model construction (entail modeling errors [8, 9]). Explanations suggested for these difficulties include the existing variety of possible behaviors at these nodes [17], the need to accurately specify decision logic

[20], and the fact that as opposed to other process model elements (e.g., activity, resource), routing nodes are not directly observable in a domain but rather abstraction of possible behavior patterns across different process instances.

Routing structures form basic constructs in practically all the process modeling languages, although different languages employ different sets of constructs for this purpose. Business Process Model and Notation (BPMN) [11] is a popular modeling language, and is the current de facto standard process modeling language. BPMN has a strong expressive power (see evaluation in [7]), facilitated by a large set of constructs. However, a large number of constructs might include construct redundancy [22], leading to unclear semantics and entailing a less conclusive modeling decision making [18]. In fact, evaluations of BPMN for ontological clarity have identified such deficiencies [7, 15]. In addition, a study of the actual use of BPMN constructs has indicated that only a relatively small set of constructs are commonly used by modelers and can be considered core constructs, while many other constructs are seldom used [23]. In addition, several studies have criticized BPMN from a perspective of cognitive effectiveness [4,5,6].

BPMN includes seven kinds of gateways (actually specified using 8 symbols), which are constructs directly used for routing, but routing can also be specified using other constructs (flows, events) or combinations of them. This makes the specification of routing in BPMN a considerably challenging part of process modeling.

This paper attempts to support the modeling of routing structures in BPMN. It does so by using a combination of routing design patterns and a decision support sheet. We suggest this combination based on cognitive considerations and evaluate it by an experiment whose subjects are Information Systems students.

The paper is organized as follows. Section 2 provides a background about the cognitive process of modeling, justifying the proposition of patterns to support this process. Section 3 describes the experimental study and its findings, which are discussed in Section 4. Finally, conclusions are given in Section 5.

2 A Cognitive Perspective of Modeling

We consider the construction of a process model which represents a given domain behavior as a problem solving task and the model as the solution. Empirical observations [14] have indicated that process modeling involves three phases: comprehension, when the modeler develops an understanding of the represented domain; modeling, when this understanding is transformed into modeling constructs; and reconciliation, when model elements are reconciled, moved, and renamed, to improve appearance and understandability. These three phases are repeated in iterations, each relating to a chunk of the model. Iterative chunking has been indicated to take place in general problem solving [10], and attributed to working memory limitations. We focus on the comprehension and modeling phases, when the modeler develops domain understanding and maps it into constructs of a modeling language.

According to Newell and Simon [10], when facing a task, the problem solver formulates a mental model of the problem, and uses it to reason about the solution and to apply solution procedures. In process modeling, solution procedures entail mapping the mental model of the domain behavior into a model in the particular modeling language. According to [10] the mental model is affected by the characteristics of the task and the methods used for achieving it. Consequently, for a BPMN modeling task, the mental model might use concepts related to BPMN constructs (e.g., gateway, event), and then the appropriate BPMN constructs (e.g., a specific type of event) should be selected and combined to form a concrete process model.

According to the cognitive schema theory [3], mental models are types of cognitive schemas related to the understanding of a specific situation that serve for solving a current problem. Mental models are constructed by using lower-level cognitive schemas, called memory objects, as building blocks. Memory objects are components of human knowledge stored in long-term memory. The simplest objects are basic concepts, called p-prims; above them are integrated objects that enable people to recognize and classify patterns in the external world so they can respond with appropriate mental or physical actions. A mental model is constructed by mapping memory objects onto components of a currently faced real-world phenomenon, reorganizing and connecting them into a model of the whole situation. A complex memory object can also be an example from past experience, which is retrieved from long term memory and adapted by analogy to the current situation.

The construction of the mental model is highly affected by the available memory objects. According to the cognitive load theory [2], the burden on the limited capacity of working memory can be reduced by using schemas that allow categorizing multiple elements as a single element [12]. When the cognitive schemas used are low level and require further integration to construct a mental model, cognitive load is increased. This might lead to reduced task performance [13].

When the task is to create a BPMN model of complex routing behavior, two main difficulties arise. First, BPMN constructs are basic objects that require effort for combining them into a mental model that fits the current situation. As a result, it is likely that the mental model does not use specific constructs, but generalized and higher-level concepts (e.g., split, event). Second, the selection of a specific combination of constructs to which the mental model should map is difficult due to the construct redundancy of BPMN [7, 15]. This makes the mapping decision inconclusive and difficult [18].

To overcome these two difficulties, we suggest the use of routing patterns. First, the patterns as concepts can form objects at a suitable granularity level for effectively serving as building blocks in a mental model. As such, they help the modeler classify the situation and generalize it. Second, the mapping to specific combinations of BPMN constructs is immediate, as these are specified in the patterns. In addition, the selection of an appropriate pattern for a given situation can be supported by a structured process of alternatives evaluation, which can be guided by a series of designated questions that classify the situation.

3 Empirical Study

The empirical study was aimed at evaluating the use of routing patterns and decision guidance when constructing a BPMN model. The main question was whether the guided use of patterns yields models of a higher syntactic and semantic quality. As a baseline for comparison we addressed the use of individual BPMN constructs, which is the common set of concepts analysts possess. Moreover, while individual constructs can serve as basic memory objects, model examples can be used as composite ones to be used by analogy. We hence decided that the study should compare the guided use of patterns – composite reusable building blocks that can be easily composed at a given situation – with the combination of atomic concepts and relevant examples. While the main question related to the effectiveness of the modeling, we also posed a second question, related to modeling efficiency in terms of the time required.

3.1 Routing Patterns and decision support

To address the above questions, we have used the set of routing behaviors which was tested in the study reported by [19]. This set addresses binary split/merge situations, and includes four split types and seven merge types, including types that are not recognized as Workflow Patterns [17]. In the study reported by Soffer et. al [19], training with this set had a positive effect on the formation of mental models, reflected in understanding domain behavior from textual descriptions. This set, however, is abstract and unrelated to a modeling notation. Hence, to make is operational for BPMN modeling, we developed BPMN representations of the behavior types in the set.

The set of routing behaviors, listed in Table 1, relates to binary splits and merges, but can easily be generalized to larger cases. The representation of the types as BPMN routing patterns was developed by one of the researchers and evaluated by the other. When more than one representation was possible, the alternative representations were discussed by both researchers until a preferred option was agreed upon. Finally, all the patterns were evaluated by an independent BPMN expert. An example pattern is presented in Fig. 1.

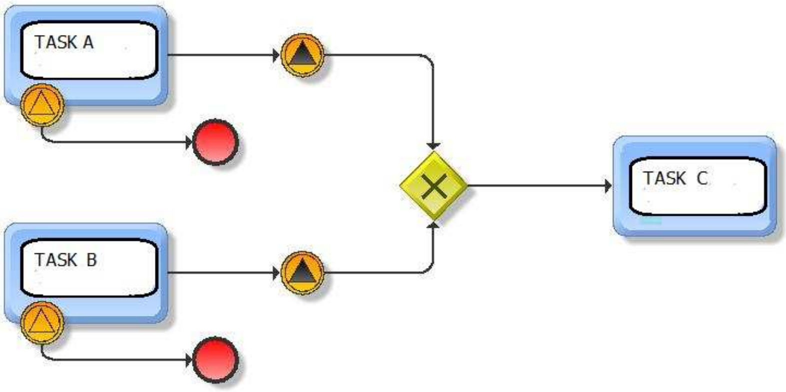

Fig. 1. BPMN routing pattern for immediate continuation with cancellation. When a task is completed (either A or B) it throws a signal event and the process continues with Task C; the signal is caught by the other task (A or B) as an interrupting event, which terminates this task.

Table 1. Set of routing behavior types for binary cases

Type	**Description**
Splits	
Exclusive (XOR)	Exactly one branch needs to be activated
Parallel (AND)	Both branches need to be activated
Inclusive (OR)	At least one branch needs to be activated
Constrained Or (COR)	A specific branch needs to be activated, the other is optional
Merges	
Immediate continuation	The process continues when the merge is reached. When both branches are active and one reaches the merge – the other proceeds independently.
Immediate continuation with cancellation	The process continues when the merge is reached. When both branches are active and one reaches the merge – the other is stopped.
Immediate continuation with asymmetric cancellation	The process continues when the merge is reached. If one branch arrives first – the other is stopped. If the other branch arrives first – the first one proceeds. In other words – if both branches are active – one always completes and the other completes only if it arrives first.
Immediate continuation with mutual blocking	The process can continue when either branch arrives at merge but is stopped when both arrive together.
Synchronization	The process can continue when both branches have arrived at the merge. When one branch arrives, continuation “waits” for the other.
Asymmetric synchronization	The process can continue only when a specific (“necessary”) branch arrives at the merge. If the other branch arrives first, the necessary one must proceed independently since continuation requires it. If the necessary branch arrives first, the other one can still proceed.
Asymmetric synchronization with cancellation	The process can continue only when a specific (“necessary”) branch arrives at the merge. If the other branch arrives first, the necessary one must proceed independently since continuation requires it. If the necessary branch arrives first, the other one is stopped.

In addition, to facilitate the selection of an appropriate pattern for a given situation, we have developed a decision tree-like guidance sheet for selecting an appropriate pattern for a given situation, as shown in Fig. 2.

3.2 Settings

The experiment was conducted with 36 Information Systems students attending a course on systems analysis. Throughout the course, the participants had already studied business process modeling using Event-driven Process Chains (EPC), Petri nets and BPMN. The participants had also experienced using the mentioned modeling notations in several projects and realistic case studies.

The students were randomly divided into two separate rooms and groups. To verify that the assignment was indeed random, we conducted an independent-samples T-test on the average homework grades achieved in the course, and found no significant difference between the groups.

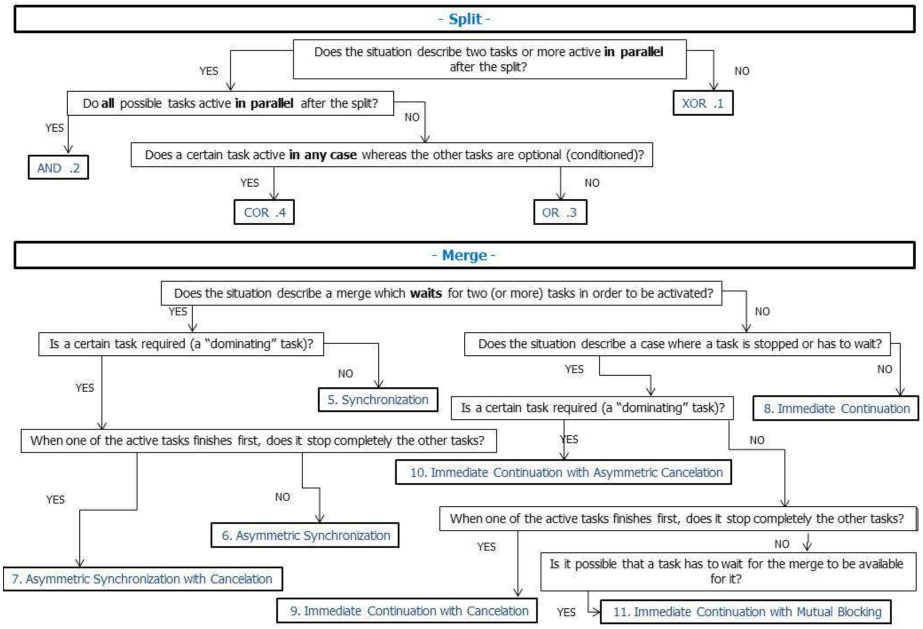

Fig. 2. A decision tree-like guidance for type selection

One group, the "Treatment" group (consisted of 17 participants), was trained with the routing patterns, and then performed the experimental assignment. The "Control" group (consisted of 19 participants) was trained with the precise meaning of individual BPMN constructs, emphasizing event and routing constructs. Both groups were shown illustrative examples; in fact, the same examples were used in the training of both groups. However, for the Treatment group these were examples of the reusable routing patterns, while for the Control group the same examples were presented as illustrating possible combinations of BPMN constructs.

The task included four descriptions of short business process situations that had to be modeled in BPMN. The business process situations focused on the dynamics of routing points in a process (merges and splits) in different domains.

3.3 Procedure

Training: Each group received one hour of training. The treatment group received, first, a short explanation on the concept of reusable patterns in process modeling. Afterwards, each pattern was presented by its business meaning, specification and an example. An example situation was analyzed using the decision support sheet. Yet, it was clarified to the subjects that the use of patterns is not mandatory and they can decide differently.

The training provided to the Control group included a reminder of the elements in BPMN. Since the notation consists of many elements, and in order to keep relevance, the reminder addressed the elements of gateways, flows, and relevant events, discussing semantic differences among them and how they can be combined. The reminder was accompanied by modeling examples, whose form was similar to the cases in the task. Moreover, the examples included the routing patterns. However, the participants were not aware that the examples include reusable patterns.

Printout: At the beginning of both training sessions, printouts were handed to the participants in both groups so they could write notes during the training and also as a reference during the task. The Treatment group received a printout consisting of: a list of BPMN elements, the BPMN routing patterns, and the decision sheet. The control group received a printout consisting of a list of BPMN elements.

Task performance: The task was performed right after the training. No time limit was set. When the students completed the task they submitted their work; the submission time was recorded for every student. As an incentive for good performance, a bonus of 3 points to the total course grade was promised to the five best performing students in each one of the groups.

Task Materials: the task materials included four situation descriptions that had to be modeled in BPMN. In total, the participants had to compose 10 routing structures in their task models (5 splits and 5 merges). The situations were selected to include routing behaviors which cannot be represented by single constructs, and correspond to our routing patterns. Table 2 presents the task cases and their corresponding routing patterns.

Table 2: Task cases and corresponding patterns

Case	Splits	Merges
1	AND, AND	Immediate continuation with cancellation; Synchronization
2	AND	Asymmetric Synchronization
3	AND	Immediate continuation with mutual blocking
4	COR	Asymmetric synchronization with cancellation

In order to avoid a learning curve effect, we created four different versions of the assignment. Each version had a different order of the cases, so participants received different versions of the assignment.

3.4 Measurement and Hypotheses

The dependent variables were the performance score for the modeling assignment and the time taken to complete the assignment. We hypothesized a difference between the Treatment group and the Control group in these two variables. Accordingly:

H_{1a}: The performance scores for subjects in the Treatment group will be different than those of subjects in the Control group.

H_{1b}: The times for performing the task for subjects in the Treatment group will be different than those of subjects in the Control group.

As shown in Table 2, the four cases included 10 routing structures. Each case was scored based on the following scheme:

- 0 points were given for mostly inappropriate representation, syntactically and semantically.
- 1 point was given for partially appropriate representation.
- 2 points were given for fully appropriate representation.

The grading was done separately by the two researchers and discussed in cases of disagreement until consensus was reached. The time taken to complete the assignment was measured in minutes and recorded upon submission of the assignment.

3.5 Analysis and Findings

The results obtained for performance scores as well as time are presented in Table 3. To test whether the observed differences between the groups' results are statistically significant, we have used an independent sample T-test for both variables, after verifying that they were Normally distributed.

Table 3: Results: performance score and time

Variables / Groups		**N**	**Min**	**Max**	**Mean**	**Std. Deviation**	**Sig.**
Performance score	Control	19	10	18	14.47	2.653	0.007*
	Treatment	17	12	20	16.71	2.544	
Task time (minutes)	Control	19	22	50	34.95	8.423	0.000*
	Treatment	17	34	66	49.18	10.513	

As both tests yielded significant results (see p values in Table 3), we can make the following conclusions. Considering the performance scores hypothesis, H_{1a} can be accepted. Furthermore, the difference in the performance scores is significantly in favor of the Treatment group, thus our conclusion is that the guided use of the routing patterns has a positive effect on the quality of the model. Considering the task times hypothesis, H_{1b} can be accepted. Furthermore, the task performance times are significantly higher for the Treatment group, indicating that the guided pattern-based modeling process is longer than when they are not used. This can be concluded for a novice population, like the participants of the study.

4 Discussion

The findings of the reported study indicate that a guided use of routing patterns can yield BPMN models of higher semantic and syntactic quality than a modeling process that does not use such patterns. The study compared a treatment group, using a set of patterns and a decision-support sheet, with a control group that served as a proxy to the "ordinary" modeling process – using the constructs of the modeling language and some experience-based examples. While the use of individual BPMN constructs as basic concepts is rather straightforward, the use of examples requires careful attention for several reasons.

First, the use of worked examples as a learning approach has been extensively studied (e.g., [1]) and found effective for strengthening problem solving capabilities with a focus on structural aspects. This seems to be in contrast to our findings. However, the example-based learning approach devotes much attention to how the examples are presented to the learner. In particular, it is stressed that examples should be presented in the context of problem classification. In our study, such context existed for the treatment group and not for the control group. Following this, the use of examples by the control group is not in line with the example-based educational approach and is not expected to yield similar learning effects.

Second, the immediacy and short period of time between seeing the example and performing the task is important when interpreting the results of the study. According to [21] events that occurred recently are easier to recall, and this might bias the judgment of their appropriateness as a basis for decisions at the current situation. In our study it is likely that having recently seen relevant examples made it easy for the subjects to recall and use them. Indeed, many of the models created by the control group attempted to adapt these examples to the given situations. As a result, the scores of the control group were generally high, although still significantly lower than those of the treatment group. It is plausible to believe that a longer delay between the training and the task performance would have made the relevant examples harder to recall, and result in models that are less similar to the examples in the control group. Accordingly, the difference in the scores of the control and the treatment group might have been larger. Furthermore, an interesting experimental setting for future research would introduce other, less relevant, examples during the time between the training and the task performance. These might then be easier to recall than the previously given examples, creating bias and reducing the quality of the produced models.

Another interesting finding is the difference in performance time between the groups. It appears that while supporting a systematic and effective modeling process, our guided patterns slowed this process significantly. This is not surprising, since a structured cognitive process that evaluates alternatives and selects an appropriate one should take longer than a quick retrieval and adaptation of an example. Furthermore, the longer time can also be explained by the fact that in our case the subjects were using printed material (i.e., "paper objects" rather than memory objects). Going over this material took time and slowed the modeling process.

It should be noted that while the reported study addressed modeling in BPMN, the set of routing behavior has served as abstract concepts in the study reported in [19], detached from any modeling language, and compared against a subset of the workflow patterns 17]. No decision-support sheet was used there, and yet the mere training with this set of conceptual routing behaviors was found to support domain understanding. This encourages the development of similar patterns for other modeling languages as well (subject to their expressive power limitations).

Also note that we did not ask the subjects about their perceptions regarding usefulness of the patterns, ease of use, and mental effort required for performing the task. These can be addressed in future studies.

5 Conclusion

Empirical evidence accumulated over time indicates that business process modelers often struggle with appropriately representing routing situations. In particular, difficulties may be encountered when using BPMN, due to its large number of constructs, the numerous possible combinations of these constructs, and the lack of ontological clarity of this language.

The paper proposes routing patterns combined with a decision guidance tool to support BPMN model creation. Cognitive considerations justify our prediction that the guided use of patterns would constitute an appropriate modeling support. These relate to the formation of a mental problem representation, where the patterns can serve for classifying the situation, and to immediately transforming the mental model into BPMN. The set of patterns builds on an existing set of routing behaviors and operationalizes these behaviors by providing their BPMN representations.

We have conducted an empirical study to evaluate the effect of the proposed support on modeling routing situations in BPMN. The results of the study indicate that the proposed support significantly improves the quality of the models, but increases the modeling time. These findings imply a potential contribution of embedding similar routing patterns and decision guidance into modeling tools that are used in practice.

However, the experiment used novice subjects in a learning environment, and its findings are limited to similar settings. Furthermore, additional and deeper understanding is still required, especially with respect to repeated application of this modeling support over time. Questions such as what would be the prolonged effect of providing such modeling support, would the respective decision criteria be internalized and become automatically used by modelers or abandoned with time are still unanswered. These should still be addressed by future research.

References

1. Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D.: Learning from examples: Instructional principles from the worked examples research. *Review of educational research*, *70*(2), (2000) 181-214.
2. Chandler, P., & Sweller, J.: Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8, (1991) 293-332.
3. Derry S.D. : Cognitive Schema Theory in the Constructivist Debate, Educational Psychologist 31(3/4), (1996) 163-174
4. Figl, K., J. Mendling, Mark Strembeck, Jan Recker.: On the Cognitive Effectiveness of Routing Symbols in Process Modeling Languages. *Business Information Systems (BIS)*, Berlin, Lecture Notes in Business Information Processing (2010).
5. Figl, K., Recker, J., & Mendling, J. A study on the effects of routing symbol design on process model comprehension. *Decision Support Systems*,*54*(2), (2013) 1104-1118.
6. Genon, N., Heymans, P., & Amyot, D.: Analysing the Cognitive Effectiveness of the BPMN 2.0 Visual Notation. Software Language Engineering: Third International Conference, (2011) 377-396.

7. Indulska, J. R. M.: How good is BPMN really? Insights from theory and practice. In *14th European conference on information systems* (2006).
8. Mendling, J., Moser, M., Neumann, G., Verbeek, H.M.W., Dongen van, B.F., and Aalst van der, W.M.P.: Faulty EPCs in the SAP Reference Model, in: S. Dustdar, J.L. Fiadeiroand A. Sheth (eds.), *Proceedings of the 4th International Conference Business Process Management (BPM 2006)*, LCNS 4102, Berlin: Springer-Verlag, (2006) 451-457.
9. Mendling, J., Verbeek, H.M.W., Dongen van, B.F., Aalst van der, W.M.P. , and Neumann, G.: Detection and Prediction of Errors in EPCs of the SAP Reference Model, *Data and Knowledge Engineering* (64), (2008) 312-329.
10. Newell A, Simon HA. *Human Problem Solving*. Englewood Cliffs, NJ: Prentice Hall (1972)
11. OMG. Business Process Model and Notation (BPMN) version 2.0. http://www.omg.org/spec/BPMN/2.0: OMG (2011).
12. Paas F, J. E. Tuovinen, H. Tabbers, and P. W. M. V. Gerven.: Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. Educational Psychologist, 38(1) (2003) 63-71.
13. Paas, F. A. Renkl and J. Sweller, Cognitive Load Theory: Instructional Implications of the Interaction between Information Structures and Cognitive Architecture, *Instructional Science* **32,** (2004) 1–8.
14. Pinggera, J., Soffer, P., Fahland, D., Weidlich, M., Zugal, S., Weber, B. and Mendling, J.: Styles in business process modeling: an exploration and a model. *Software & Systems Modeling*, (2013) 1-26.
15. Recker, J., Rosemann, M., Green, P., Indulska, M.,: Do ontological deficiencies in modeling grammars matter? *MIS Quarterly* 35(1), (2011) 57-79.
16. Reijers H.A, Mendling J.: A Study into the Factors that Influence the Understandability of Business Process Models, *IEEE Transactions On Systems, Man, And Cybernetics – Part A,* 41(3), (2011) 449-462.
17. Russell, N. C., ter Hofstede, A.H.M., Aalst van der, W.M.P. , and Mulyar, N.. *Workflow Control-Flow Patterns: A Revised View*, BPM Center Report BPM-06-22, BPMcenter.org (2006).
18. Soffer P. and Hadar I., 2007, Applying Ontology-Based Rules to Conceptual Modeling: A Reflection on Modeling Decision Making, European Journal of Information Systems 16(4), pp. 599-611
19. Soffer P., Kaner M., and Wand Y., Towards Understanding the Process of Process Modeling: Theoretical and Empirical Considerations, In BPM 2011 Workshops Proceedings, LNBIP 99, Springer-Verlag, (2012) 357-369
20. Soffer, P., Wand, Y., and Kaner, M.: Semantic analysis of flow patterns in business process modeling, in: G. Alonso, P. Dadam and M. Rosemann (eds.), *Proceedings of the 5th International Conference Business Process Management (BPM 2007)*, LNCS 4714, Berlin: Springer-Verlag, (2007) 400-407.
21. Tversky, A., and Kahneman, D.: Judgment under uncertainty: Heuristics and biases. *science*, *185*(4157), (1974) 1124-1131.
22. Wand, Y., and Weber, R.: On the Ontological Expressiveness of Information Systems Analysis and Design Grammars, *J. of Information Systems* (3), (1993) 217-237.
23. Zur Muehlen, M., & Recker, J.. How much language is enough? Theoretical and practical use of the business process modeling notation. In *Advanced information systems engineering* Springer Berlin Heidelberg (2008) 465-479.