Ontology-Based Assistance for Semi-Formal Process Modeling

Michael Fellmann, Novica Zarvić, Annelene Sudau and Lennard Nobbe

Information Management and Information Systems
Osnabrück University
Katharinenstr. 3
49074 Osnabrück, Germany
{michael.fellmann|novica.zarvic|asudau|lnobbe}@uni-osnabrueck.de

Abstract: The construction of semi-formal process models is a difficult task due to the complexity of today’s business processes. In this contribution, we devise an ontology-based assistant for semi-formal process modeling. The assistant is intended to support the modeler in creating high quality process models with less time and effort by reusing knowledge captured in a reference ontology. At first, we derive functionalities for such an assistant. We then show how the approach can be implemented and integrated into an existing modeling tool. Finally, we describe an experimental evaluation of the approach.

1 Introduction

The construction of semi-formal process models is challenging due to the high complexity of today’s business processes which do not impose an immanent solution for their representation. Therefore, process modeling can be considered as a form of problem solving rather than the simple mapping of known facts in a graphical model. Current tools have some deficiencies in supporting this problem solving, which can be characterized as assessment problems and reuse problems. Assessment problems can be summarized as questions regarding the fitness of the model for the intended purpose. They are hardly answered by existing modeling tools since they do not “understand” what is being modeled. Understanding in the sense of reasoning about the contents of a model e.g. in order to inform the user about modeling progress or correctness would require formalized domain knowledge incorporated into the modeling tool. Although such knowledge exists (e.g. in the form of ontologies), it is currently not embedded in the tools. Reuse problems arise during the creation of models if known solutions of business problems already achieved are not reused. Moreover, the comparison with processes of other companies is a useful and inspiring source of insight (especially spanning different sectors of industries) [APQC10] and it is not leveraged from existing process modeling tools.
In this contribution, we propose the extension of process modeling tools with an ontology-based assistance addressing the two problems introduced above. First we shortly outline some related work (section 2), before we discuss our research method (section 3) that belongs to the Design Science paradigm [HMP+04]. Then, we derive required functions that the assistant should implement in order to address the two problem areas (section 4). We then show how the approach can be implemented and integrated into the graphical user interface of an existing modeling tool and report on our ongoing development activities (section 5). Finally, we present an experiment-based evaluation approach (section 6). Finally we conclude the article (section 7). To the best of our knowledge, this is the first ontology-based approach exclusively geared towards advising the user in business problems of process modeling.

2 Related Work

The fundamental idea of our approach is to use an ontology in conjunction with assistance functionalities to support the user in modeling. Therefore, research regarding modeling support leveraging knowledge representations in a broad sense is related to our work. Such research can be found in several areas. The reference-based model construction aims at leveraging reference models to construct new models. They capture domain knowledge and hence serve as a knowledge representation. For model adaptation and reuse, several methods have been described [FL02; BB06]. Although methods such as configuration [De06; RA07; DK07a] are amenable to IT-support, they are not supported by most existing modeling tools [DK07b] and are still subject to research. Recently, reuse-based model construction is discussed in business process management as a more holistic approach in comparison to reference modeling in the sense of spanning different abstraction levels from process oriented systems. The main idea is to leverage existing process assets encompassing not only conceptual models, but also the accompanied software artifacts [FTS+11; TZD11]. Another research area focuses on query-based support for model construction where the goal is to find models or model fragments which can be used to complete the model under construction. [Ko07] describes an approach and a tool to determine the similarity of the model currently under construction compared to models which already exist in a repository. Analogous to reuse-based model construction and in contrast to reference-based model construction, no separate knowledge representation is required. In pattern-based model construction, the availability of patterns is essential. The idea of patterns originates from the discipline of architecture and has been introduced as analysis patterns in conceptual modeling by Fowler [Fo96]. They represent knowledge in the form of abstract templates used to apply well-known solutions to similar problems. In literature, different types of patterns have been suggested such as control flow [AHK+03], resource [RHE+05], service interaction, [BDH05], data [RHE+05] and activity patterns [TRI09]. Up to now, research has mainly focused on pattern detection [LIT+09; DHL+09], although a few approaches support modeling [TRC+08b] and a tool ProWAP has been developed [TRC+08a]. In comparison to other construction techniques such as reference-based model construction, patterns are more abstract in nature and usually require a thorough interpretation by the modeler.
In the area of case-based reasoning support for BPM the main idea is combining case-based reasoning and business process modeling in order to leverage past process executions when specifying a new process. Processes that have been executed (also referred to as cases or templates) are stored in a knowledge base and are used to create new processes or workflows [MZM04]. The approaches aim at improving flexibility and reuse [YCW06] increasing adaptiveness [PZW+08] and can act as a knowledge management technique for business process redesign [MMR03].

In order to position our approach of ontology-based assistance relative to the other areas of research, two criteria can be used: the amount of formalized knowledge and the complexity and sophistication of supporting tools. Regarding the amount of formalized knowledge, we make use of a comprehensive representation of business knowledge collected in the MIT Process Handbook [MCH03]. It comprises approx. 5000 business functions and processes which are described regarding their inputs, outputs, resources, specializations, generalizations, exceptions, related functions and other information (see http://ccs.mit.edu/ph/ to explore the ontology and http://ccs.mit.edu/ophi/ for the Open Process Handbook version). Regarding the amount of formalized knowledge, our approach is similar to automated model construction or case-based reasoning approaches which also rely on comprehensive knowledge bases. However, one decisive difference is that our approach does not require setting up a knowledge base with semantically described building blocks or recorded past process executions.

In contrast, our approach works “out of the box” since it leverages an existing reference ontology and is intended to assist and advice the modeler during modeling. With this paradigm of assistance, we do not want to suggest complete models or to improve models based on past executions. Consequently, the complexity of the tool which has to be developed is lower than the complexity of tools in the two aforementioned categories (cf. Table 1). In essence, we do not want to replace any of the other approaches belonging to other areas of research. We rather seek to offer new forms of additional advice and assistance which should make model creation more convenient by reusing the knowledge contained in a reference ontology.

Table 1: Comparison of research areas w.r.t. formalized knowledge and tool complexity

<table>
<thead>
<tr>
<th>Research area</th>
<th>FK</th>
<th>CT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference-based model construction</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Reuse-based model construction</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Query-based support for model construction</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Pattern-based model construction</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Automated, planning-based model construction</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Case-based reasoning support for BPM</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Ontology-based assistance for model construction (our approach)</td>
<td>●</td>
<td>○</td>
</tr>
</tbody>
</table>

Legend:
FK = usage of formalized business knowledge, CT = complexity of supporting tools;
● = high, ○ = medium, ○ = low
3 Research Method

The research approach taken in this work can be allocated to the design science paradigm [HMP+04, MS95]. Kuechler and Vaishnavi briefly describe design science research as the “construction of an information technology artifact and its evaluation” [KV11]. We address both mentioned issues – construction and evaluation – in this paper. With respect to the construction of the artifact, we first derived the functionalities of the ontology-based assistance by conceptually eliciting the requirements with the help of use cases. On the basis of the collected requirements the assistance functionalities have been implemented. This step implied the choice of a suitable reference ontology, the possibility to query the ontology in an adequate way, as well as the presentation of the query results in a modeling environment. With respect to the evaluation issue, we have first chosen a suitable design evaluation method. According to Hevner et al. [HMP+04] several methods can be used for evaluating designed artifacts. In this respect we have decided to perform an experiment in order to be able to study relevant characteristics of our artefact. Therefore we have first set up an experimental setting at our university, and analysed the results. Figure 1 conceptually depicts the described activities that have been performed in this work, where a clear distinction between the construction and evaluation activities can be seen.

4 Assistance Functionality

The functionalities of the ontology-based assistance can be derived by considering the following broadly described use cases. (U1) The modeler has to select a process from the MIT Process Handbook ontology which is the basis of the process she wants to create. This is required to enable the other assistance functionalities. (U2) To support modeling, the assistance function has to be capable of suggesting model elements. (U3) To reuse business knowledge and expertise contained in the ontology, the assistance function has to be capable of providing information either upon user request or in an unobtrusive and pro-active way during modeling. (U4) To assess the model, the assistance function has to provide a verification feature for verifying the model during modeling and when it is finished. These four use cases are described in further detail below and in Figure 2.
U1: Suggestions for the base-process. Before modeling with the assistance functionality, the modeler has to select a base-process (e.g. “Hiring process”). Therefore, the modeler has to be able to search the ontology for an adequate base-process e.g. by using keywords or advanced browsing methods. Furthermore, linguistic procedures can be applied to enhance the quality of the search results.
**U2: Suggestions of elements.** Elements which can be inserted in the process model should be suggested based on the information available about the base-process or the elements already present in the model. Using the information about the base process, activities and resources can be suggested by detecting missing elements which are known to be part of the base process but which are missing. Also, suggestions based on the inputs, outputs and exceptions of elements already present in the model can be generated. In this way, the tool can suggest e.g. to insert additional activities for error handling if the current activity the modeler has inserted is known to produce exceptions of a specific type. For each of the suggested activities, the assistance functionality should also offer more specific and/or more generic alternatives and it should display related activities, as such information is captured in the MIT Process Handbook. If the modeler inserts a model element that has been suggested by the ontology-based modeling assistance, this element will be annotated with the ontological semantics which is a prerequisite of use case U4.

**U3: Providing information about processes.** The modeler should be provided with information about a process that is contained in or can be computed using the ontology. Such information encompasses e.g. the links the process has with goals and resources or the progress of model creation. The latter can be calculated from the number of process parts in the ontology and the number of those parts which are already included in the model. Moreover, a trade-off matrix (which is also part of the MIT Process Handbook) can be used to display advantages and disadvantages of alternative process designs.

**U4: Model verification.** Using the element suggestion functionality of use case U3 results in most (or even all) of the model elements being semantically annotated. Consequently, the model elements are backed with ontological (and hence machine processable) semantics enabling semantic verification functionalities such as checking for unhandled exceptions or incomplete process models.

### 5 Implementation of the Assistance Functionalities

#### 5.1 Management of the Ontology Data

A key challenge in the development of software relying on ontologies is how to guarantee the availability as well as quality of knowledge contained in the ontology. In our research, we decided to use the MIT Process Handbook since it has obvious advantages.

First, its contents have been collected and formulated in scientific projects, which means that a certain quality level can be imputed. Second, it is comprehensive as it contains a substantial number of processes (currently approx. 8000), and thus standard processes are covered. Third, it has been translated into the Web Ontology Language (OWL) in a project at the University of Zurich so that no effort in “ontologizing” the knowledge was required. Fourth, it offers a very regular and clean structure (cf. Figure 3) making it amenable to implement the assistance functionalities on top of it.
In the OWL version of the Process Handbook, every concept matches one OWL file. Regarding the expressivity, the subset of the MIT Process Handbook we used for implementing our demonstration system had the expressivity of the description logic SHIN. For our demonstration system, we used the Jena Framework combined with an in-memory approach (in the future, we plan to migrate to the OWL-API and OWL-DB).

5.2 Querying of the Ontology

To query the ontology, we use the query language SPARQL (www.w3.org/TR/rdf-sparql-query/) which is the de-facto query language in the Semantic Web, standardized by the W3C and widely supported in triple stores and from various APIs associated with OWL.

In the following, we give an example of a SPARQL query which retrieves the parts of a process whereby the has-part relation of the MIT Process Handbook scheme (cf. Figure 3) is used. The query is generated inside the modeling tool and transmitted via HTTP to the server. The results of the query (ID, name and description of a respective process part) are used inside the modeling tool to suggest model elements (cf. use case U2 and section 5.3). The term `<BaseProcessID>` is a variable which is replaced by the ID of the current base-process which is known to the modeling tool since the user first has to select the base-process (cf. use case U1).

```
SELECT DISTINCT ?id ?name ?desc WHERE {
  <BaseProcessID> has-part ?id .
  ?id rdfs:label ?name.
  ?id rdfs:comment ?desc.
}
```
5.3 Presentation of the query-results in a modeling environment

Since it is easy to enhance through Add-Ins, we use Microsoft Visio 2010 for our demonstration system. The Add-In currently uses the template for event-driven process chains EPC. It catches events of inserted elements and matches their type against EPC shapes for the decision whether the assistant should show up or not. In the use case described by the query in section 5.2, the modeler would have modeled a few steps of a process and insert the next activity, so a screen of the application would appear.

![Assistant Tool](image)

Figure 8: Appearance of the assistant tool while a process is created in MS Visio

The activities which are *suggested automatically* can be selected and applied to the model. Moreover, the modeler can navigate through the MIT Process Handbook by using the structure of specializations and generalizations or related processes. It is also possible to search the ontology for keywords if the modeler is not satisfied with the results.

6 Experimental Evaluation

6.1 Experimental setting

The goal of the experiment that has been performed was to get a first impression how assistance-based modeling is perceived by the participants while performing a predefined task. More specifically, it targeted at determining previously unknown problems as well as the technical functionality. For doing so, several scenario-based tasks have been formulated. This was accomplished by means of a browser interface, which provides a search function for processes, where the exact phrase can be typed in, the maximum number of results to be displayed as well as which exact elements and fields should be considered. If a process is selected, it is possible to obtain further information on the process. For instance, a full description is displayed or it can be seen whether the process has relation of type specialization or generalization respectively.
The experiment took place at our university and most participating modelers (n=6) have been post-graduate students at Master or PhD level. After the experiment, the participants have been interviewed in an in-depth interview with no predefined structure in order to additionally capture participants' ideas and feelings about the assistance system.

For being able to assess the results of the experiment in an objective manner, the results had to be mapped against a predefined sample solution. By following this approach three sets of elements can be distinguished. The first set contains the elements of the predefined sample solution, the second set contains the elements that have been provided by the modeler, and finally the third set of elements, which the modeler used from the MIT process handbook. The Following figure 5 shows some situations that may occur.

![Diagram](image)

**Figure 9: Relations between sets**

On the left side we can see that the modeler has used some but not all elements from the process handbook and from the sample solution, because there are elements not covered by those sets. In Figure 5b on the right side we can see that all the elements used by the modeler are part of the sample solution, which again is part of the elements from the process handbook. This situation would – of course – mean that the elements used by the modeler are also all inside the lexical knowledge in the process handbook. This is not the case in Figure 4a and means that the elements from the handbook neither cover all elements in the sample solution nor from the modelers set.

6.2 Evaluation Results

The tasks on which the participants of the experiment worked on contained to model three different processes with and without support of the before presented assistant tool, which are: (i) hiring and new employment of a worker, (ii) the process of distributing an arbitrary product assortment, and (iii) the creation and placing advertisements for own products. The assessment of the results (cf. Figure 6) was centered on the difference of *elements* by modeling with or without support of the assistant tool, which explains the
Figure 10: Assessing the relationship between inserted model elements by manual modeling and assisted modeling

Figure 11: Assessing the time needed to create models with assistance and without assistance
relationship depicted in the diagram. Models that have been created by using the assistance function in most cases had benefit in the way that elements have been added to the model (red colour), which were beforehand absent while modeling manually (blue colour) without ontology-based assistance. However, the results of our evaluation show that one of the biggest problems is the degree of abstraction or level of detail respectively. The MIT process handbook contains process steps that are formulated in a rather generic way and not very detailed. This leads to the fact that the handbook was not very useful to the modeler, because it suggested only a small set of elements during process modeling.

The second criterion was to determine also differences in the time needed to create a process model (cf. Figure 7). Even though we have chosen parts of the handbook that seem to be more complete for our experiment, we had to notice that the information contained in the handbook for some areas seems to be incomplete by including also empty processes without any further information. In a future version of our system we will tackle this problem by integrating additional sources of process knowledge such as the Process Classification Framework (PCF) which is a collection of approx. 1000 business activities grouped on four levels. Despite the fact that our present prototype system has some shortcomings the participant’s judgment has been positive. In the interviews that we conducted after the experiment, the participants emphasized that the system has helped them to grasp the essence of the process to be modeled – that is what to model and what to omit. One participant put it like this: “If there had been suggestions, they helped me to focus on the right abstraction level and on what’s really important. It somehow increased my confidence in what I’m modeling and therefore I liked it very much. It feels like the lane-assist in my car!”.

7 Conclusion

We provided a first insight in our ongoing activity of developing an ontology-based assistance for semi-formal process modeling. The combination of a reference ontology and a modeling application provides a very promising and effective way to facilitate the creation of process models even for inexperienced modelers. It allows for assessing the model contents e.g. regarding the completeness of a model in contrast to a base-process selected from the ontology. It also leverages the business knowledge collected and represented in comprehensive ontologies such as the MIT Process Handbook.

In this contribution, we have derived functionalities for such an assistant based on use cases. We then showed how the approach can be implemented by elaborating on the storage and retrieval of ontological content. We thereby described the integration of query results into the graphical user interface of an existing modeling tool, which is the basis of our demonstration system. Finally, we provided a first evaluation of the approach describing an experiment.
In our future work, we plan to augment our demonstration system by using more comprehensive suggestions including also industry-specific suggestions provided e.g. by the industry-specific versions of the Process Classification Framework that have been released recently. Also, regarding the process of recommendation, we plan to incorporate concepts of recommender systems such as top-k rating or content-based recommendation strategies.

Acknowledgements: The authors are thankful to the European Regional Development Fund (ERDF) for supporting the project SEMPHIS-(Semantische Modellierung prozessorientierter Informationssysteme) under number ZW 6- 80131109.

References

<table>
<thead>
<tr>
<th>Reference</th>
<th>Title</th>
<th>Contributors</th>
<th>Details</th>
</tr>
</thead>
</table>
