Collaborative modeling of processes and ontologies

Marco Rospocher

Fondazione Bruno Kessler (FBK)

http://dkm.fbk.eu/rospocher - rospocher@fbk.eu

BMIR Research Colloquium, Stanford, USA – September 22, 2011
Develop a theoretical and practical framework that:

Supports the integrated modeling of Processes and Ontologies;

Fosters the collaboration between domain experts and knowledge engineers.

**WHY?**

need of a comprehensive model which requires the description of both the dynamic component (processes) and the static component (ontology);

need for an agile collaboration between domain experts and knowledge engineers. Need to actively involve the domain experts in the modeling process.
The research vision - architecture

Formal representation of integrated processes and ontologies

Architecture for collaborative conceptual modeling

Theoretical framework
Outline of the presentation

Formal representation of processes and ontologies

Architecture for collaborative conceptual modeling in wikis

MoKi and some of its real usages
FORMAL REPRESENTATION OF PROCESSES AND ONTOLOGIES
Integrating processes and ontologies

Roles / Organization

Documents

Actions
Integrating processes and ontologies

Example of queries and reasoning that involves both ontological and process knowledge:

*What are the activities performed by a certain role (e.g. PC Chair)?*

*Where are documents (e.g. reviews, notifications) produced?*

*What are the activities where something is published? What are the activities where something is sent out?*

*What are the activities an author perform right before submitting something?*

Example of application that requires querying for both ontological and process knowledge:

*Managing cross-cutting concerns in business processes.*
Integrating processes and ontologies
Semantically annotated business processes are encoded into a logical knowledge base implemented in OWL.

Note: Business Process Diagrams (BPDs) are specified using the Business Process Modelling Notation (BPMN).
BPMN Ontology

Provides a formalization in OWL DL of the structural part of BPMN.

Two parts:
1. is - a taxonomy of all the BPMN elements;
2. attributes and properties which describe how to use these elements to compose a BPD.

A Start Event MUST NOT be a target for Sequence Flow

$$\text{sequence\_flow} \sqsubseteq \forall \text{has\_connecting\_obj\_target\_ref}.(\neg \text{start\_event})$$
## BPMN Ontology

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DL Expressivity</td>
<td>SHOIN(D)</td>
</tr>
<tr>
<td>Classes</td>
<td>117</td>
</tr>
<tr>
<td>Object Properties</td>
<td>123</td>
</tr>
<tr>
<td>Datatype Properties</td>
<td>48</td>
</tr>
<tr>
<td>Individuals</td>
<td>104</td>
</tr>
<tr>
<td>Class Axioms</td>
<td>463</td>
</tr>
<tr>
<td>Object Property Axioms</td>
<td>236</td>
</tr>
<tr>
<td>Datatype Property Axioms</td>
<td>96</td>
</tr>
<tr>
<td>Individual Axioms</td>
<td>250</td>
</tr>
<tr>
<td>Annotation</td>
<td>504</td>
</tr>
</tbody>
</table>
BPMN Ontology

Current version based on v1.1 of the BPMN specifications by OMG (ontology for v2.0 almost ready)

It is not intended to model the dynamic behaviour of business process diagrams.

*If there are multiple outgoing Sequence Flow then only one Gate (or the DefaultGate) SHALL be selected during performance of the Process.*

There are a few documented properties which are not represented due to expressiveness limitation imposed by Description Logics.

*The ConditionExpression attribute MUST be unique for all the outgoing Sequence Flows connected to an Inclusive Gateway*

Available for download at:

Business Domain Ontology

Represents the (specific) business domain.

Used to annotate the elements of the business process diagram.

Can be composed of:

- Top level ontologies, such as DOLCE;
- Domain-specific ontologies.
BPD Instances

Represents the specific annotated business process diagram.
BPD Instances

Represents the specific annotated business process diagram.

Create an individual for each graphical element of the business process.

$s_1, s_2, s_3, s_4, t_1, t_2, g_1, g_2$
BPD Instances

Represents the specific annotated business process diagram.

BPMN-type assertions: for every graphical element $g$ of BPMN type $T$ occurring in the process, we add the assertions $T(g)$.

```
sequence_flow(s_4)
```
**BPD Instances**

Represents the specific annotated business process diagram.

**BPMN-structural assertions:** For every connecting object \( c \), going from \( a \) to \( b \), we add assertions of the form source\((c, a)\) and target\((c, b)\).

\[
\text{has_sequence_flow_source_ref}(s_1, g_1) \\
\text{has_sequence_flow_target_ref}(s_1, t_1)
\]
BPD Instances

Represents the specific annotated business process diagram.

**Semantic assertions**: For every graphical element $g$ of the process which is annotated with $C$ (where $C$ is a complex concept expression of the domain ontology), we add the assertion $C(g)$. 
Automatic OWL A-box generation

The transformation of an annotated Business Process Diagram into an OWL A-box is performed automatically.

Available for download at:

http://selab.fbk.eu/difrancescomarino/SemanticBPM/
Return all the activities that buy products and for which there exists at least a path, consisting of sequence flows, that connects a to check product availability activity to the given activity.

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX p: <http://exampleOntology#>
SELECT ?a1 ?a2
WHERE { ?a1 rdf:type bpmn:activity.
  ?a1 rdf:type p:to_check_product_availability.
  ?a1 (bpmn:has_sequence_flow_source_ref_inv/bpmn:has_sequence_flow_target_ref)* ?a2.
  ?a2 rdf:type bpmn:activity.
  ?a2 rdf:type p:to_buy_product.}
Process Constraints

The framework also enables to define constraints for:

**correct/incorrect annotation of business process graphical elements:**

- *A BPMN activity is annotatable only with actions of the domain ontology (and not e.g., with documents);*

**valid critical patterns:**

- **containment constraints:** the activity of managing a shopping cart is a sub-process which contains an activity of removing products from the cart;

- **precedence constraints:** the activity of providing personal data is immediately preceded by an activity of reading the policy of the organization;

- **exception handling constraint:** the activity of reserving products in the On-line Shop pool has always to catch a “product unavailability” error event;

**Using DL-reasoning we can:**

Check compatibility of process constraints.
Integrating processes and ontologies

Selected publications:


*Semantically-aided business process modeling* - C. Di Francescomarino, C. Ghidini, M. Rospocher, L. Serafini, P. Tonella - International Semantic Web Conference (ISWC’09)

*Reasoning on semantically annotated processes* - C. Di Francescomarino, C. Ghidini, M. Rospocher, L. Serafini, P. Tonella - International Conference on Service Oriented Computing (ICSOC’08)

Next steps: extension to the dynamics of executions.
AN ARCHITECTURE FOR COLLABORATIVE CONCEPTUAL MODELING IN WIKIS
Why a wiki-based conceptual modeling tool?

Wikis support **collaborative** editing;

Users are quite **familiar** with viewing/editing wiki content (e.g. Wikipedia);

Only a **web-browser** is required on the client side;

Wikis provide a **shared knowledge repository** accessible by users spread all over the world;

Wikis can provide a **uniform tool/interface** for the specification of different model types (e.g. ontologies, processes, …);
An architecture for collaborative conceptual modeling in wikis

1. **One element** ↔ **One page**

   each element of the model is represented by a page in the wiki;

---

**Concept “Mountain”** ↔ **Mountain**

A **mountain** is a large **landform** that stretches above the surrounding land in a limited area usually in the form of a peak. A mountain is generally steeper than a **hill**.

The highest mountain on earth is the **Mount Everest**
2. **Unstructured and structured descriptions**

Each page contains both structured and unstructured content;

---

**Mountain**

A mountain is a large landform that stretches above the surrounding land in a limited area usually in the form of a peak. A mountain is generally steeper than a hill. The highest mountain on earth is the Mount Everest.

```
□ Landform
□ ¬Hill \(\sqcap\) ¬Plain
□ \(\forall\) madeOf(Earth \(\sqcup\) Rock)
□ \(\exists\) height. \(\geq\) 2500

Mountain(Mt.Everest)
Mountain(Mt.Kilimanjaro)
```

(unstructured content) (structured content)
An architecture for collaborative conceptual modeling in wikis

3. Different views to access the model:

different views to support different modeling actors;

A mountain is a large landform that stretches above the surrounding land in a limited area usually in the form of a peak. A mountain is generally steeper than a hill.

The highest mountain on earth is the Mount Everest.

Mountain

<table>
<thead>
<tr>
<th>is a</th>
<th>landform</th>
</tr>
</thead>
<tbody>
<tr>
<td>different from</td>
<td>hill, plain</td>
</tr>
<tr>
<td>made of</td>
<td>earth</td>
</tr>
<tr>
<td>made of</td>
<td>rock</td>
</tr>
<tr>
<td>height</td>
<td>at least 2,500m</td>
</tr>
<tr>
<td>samples</td>
<td>Mt. Everest, Mt. Kilimanjaro</td>
</tr>
</tbody>
</table>

(semi-structured view)

Mountain

⊆ Landform
⊆ ¬Hill ∩ ¬Plain
⊆ ∀madeOf(Earth ∪ Rock)
∃height. ≥ 2500
Mountain(Mt.Everest)
Mountain(Mt.Kilimanjaro)

(fully-structured view)
AND SOME OF ITS REAL USAGES
Wiki-based modeling tool;

Supports the integrated modeling of Processes and Ontologies;

Provides modeling support both for domain experts and knowledge engineers, fostering the collaboration between them;

Based on the framework presented so far.
Usages of IP FP6 EU Project [03/2006 – 02/2010]

Purpose: modeling of tasks/processes in an enterprise and of the topics related to that task (competencies)

Used by:

4 SMEs
3 Universities
several related summer schools and university courses
Purpose: build/revise an environmental ontology

Developed the new key concepts extraction functionalities
Used to automatically create part of the ontology (pollen)
eContentplus EU Project [09/2007 – 08/2010]

Purpose: build/revise an ontology of organic agriculture and agroecology

Used to foster collaboration between domain experts (FAO) and knowledge engineers

Follow-up: Organic.Lingua (FP7 Pilot Tipe B EU project [36 months])

Extend MoKi to multilingua models and interface
Italian national project [01/2010 – 12/2011]

Purpose: model processes for analysis/revision and dematerialization

Used by 7 Italian regions:

Piemonte, Emilia Romagna 1 & 2, Puglia, Liguria, Marche, Trentino

Medium size models produced in around 2 weeks.
OncoCure

Funded by Fondazione Caritro, Trento [2007 – 2008]

Purpose: modeling breast cancer clinical protocols encoded in Asbru.

Customized version of the tool

Actively used mainly by KE

Positive feedback by the doctor who produced the clinical guidelines in “reviewing” the model created.
FBK Joint Research Project [2009 - 2013]

Purpose: modeling of nurse activities in an oncology ward.

Collaboration between “observer” and KEs for the creation of the process diagrams

Planning to integrate ontological information soon
Lessons learned

Wikis can be a powerful way to lower the entrance barrier for modeling tools and to share knowledge;

Real need to integrate processes and ontologies, and to include in processes organizational aspects taken from a formal description (ontology);

Collaboration happens and is helpful;

Need to guide domain experts by providing schemata of representations; e.g., what characterize a document?
Evaluation

Performed within ProDe project (to be presented @ ISWC2011);

Users: 14 Public Administration employees distributed across 6 teams creating different integrated models;

Research questions considered:

   RQ1: Is MoKi easy to use for domain experts?

   RQ2: Is MoKi useful for collaboratively modeling domain knowledge?

   RQ3: Are all the provided views useful or is there a ‘best’ view among the different interface views provided by MoKi for: (a) getting the model overview? (b) navigating the model? (c) creating new entities?

Analyses performed:

Quantitative analysis of the data on the usage of MoKi (editing logs, web-server logs, ...):
Evaluation Results

RQ1 (ease of use):

The users perceive the tool as more than easy to use:

• 72% of employees spent only less than two days to learn how to use tool;

• the same percentage learned it autonomously.

RQ2 (usefulness for collaborative modeling):

The users positively perceive the overall usefulness of the tool for the collaborative modeling of documents and processes:

• Correlation between the size of the subject’s team and his/her feedback about tool usefulness for collaborative purposes (esp. in team with 3+ or more users).

• Result further validated by the intensive usage of collaborative functionalities by people in large team.

RQ3 (usefulness of provided views):

All the views provided by the tool have their own usefulness.
Current & Future Works

Fully implement the formal framework for integrating processes and ontology (preliminary prototype now ready)

Develop ad-hoc templates to guide DE in modeling activities

  describing an artifact is different than describing a role

Support usage of ontology patterns

  to speed up modeling activities, and limit modeling errors

Extend key-concepts extraction functionalities

  Support extraction / identification of semantic relation in text (e.g. “isA”) between concepts
Publications and demos:


Released Open Source in July 2010 (version 1.2 – GPL2)

MoKi WebSite:

URL: http://moki.fbk.eu

On-line demos, code download, documentation, news, support…
Joint work with…

On all this stuff…:

Chiara Ghidini, Lucian Serafini.

Semantically Annotated Business Processes:

Paolo Tonella, Chiara Di Francescomarino

MoKi:

Nahid Mahbub, Gaetano Calabrese, Mauro Dragoni, Rakebul Hasan, Musawar Saeed

eHealth Applications:

Claudio Eccher

Term Extraction (for ontology building / evaluation):

Sara Tonelli, Emanuele Pianta
Thank You!

Questions?

Marco Rospocher
http://dkm.fbk.eu/rospocher
rospocher@fbk.eu