

Review of Doctoral Thesis

Thesis Title: Process Mediation Framework for Semantic Web Services

Thesis Author: RNDr. Roman Vaculín
Dept. of Theoretical Computer Science and Mathematical Logic
Faculty of Mathematics and Physics
Charles University
Prague, Czech Republic

Review Author: Mgr. Martin Nečaský, Ph.D.
Dept. of Software Engineering
Faculty of Mathematics and Physics
Charles University
Prague, Czech Republic

Review

Introduction

Current Web Services technologies, e.g. SOAP, WSDL, UDDI, BPEL, etc., unlocked a promising way to implement information systems in the shared, opened, dynamic and quickly evolving Web environment. An information system can be decomposed to individual independent (not necessarily software) components, called web services, that communicate with each other by exchanging XML messages. A web service is separated to two parts: interface and implementation. Only the interface is made public while the implementation is hidden. This allows web services to be loosely coupled which is one of their most important features.

However, as the academia and practice have shown, the Web Services technologies have a limited operational range. This is because communicating parties are required to have a common understanding of their business areas. More precisely, this comprises standardization of interfaces of web services and business processes. However, it seems today that such standardization is almost impossible. Usually, there are the following mismatches:

1. mismatches on the data level: incompatible XML formats are used for communication by individual parties, and
2. mismatches on the business process level: incompatible sequences of actions that are needed to be performed by individual parties.

These two main shortfalls could be overcome by a new generation of technologies called Semantic Web Services. These technologies extend current Web Services technologies by specification of semantics in form of ontologies and mapping syntactical descriptions of web services to ontologies. Roughly speaking, the data mismatches problem can be solved by sharing the same ontology or mapping ontologies on each other. The business process mismatches can be solved by process mediation. Both solutions seem to be more effective on the semantic level.

Thesis Overview and Critique

The presented thesis focuses on both problems. The first problem is simplified by an assumption of a common ontology shared by all communicating parties. There can also be a situation that different parties use different incompatible ontologies. This case is considered by the author and solved by citing existing works in the area of ontology matching. The author also simplified the problem only to the case with two parties: requestor and provider.

The thesis is structured as follows. In chapter 1, the author introduces the problem of process mediation and some related problems. In chapter 2, preliminary technologies exploited by the authors later in the thesis are introduced. In chapters 3-6, the author proposed his own original solutions in several following areas related to Semantic Web Services technologies. In chapter 7, the author presented related work and he concluded in chapter 8.

The author achieved many contributions to various problems related to Semantic Web Services technologies. The most important ones are summarized below:

- **Process Mediation** (chapter 3)
 - *Problem:* Requestor and provider want to collaborate in a business process but both require performing different actions to achieve their common goal.
 - *Thesis contributions:*
 - Abstract Process Mediation Framework (APMF) that identifies key aspects of process mediation and puts a complete formal basis for the solution.
 - Various complex scenarios have been covered by (APMF) where activities can not be matched one-to-one.
 - Algorithms for static and run-time mediation.
- **Data Mediators Discovery** (chapter 4)
 - *Problem:* It can be necessary to mediate data from one representation to another.
 - *Thesis contributions:*
 - Formalization of service advertisements.
 - Top-K algorithm returning the best K data mediators for a give request which makes the solution applicable in practice.
- **Semantic Monitoring of Web Services** (chapter 5)
 - *Problem:* Various events are emitted by communicating parties. It is necessary to detect the events and react on them. Events can be simple or complex, i.e. composed of other events.
 - *Thesis contributions:*
 - Shifting the problem of event detection from the syntactical level to the semantic level
 - Formalization of semantic events. Event detection algebra.
 - Algorithm for (complex) events detection.
- **Recovery of OWL-S Workflows** (chapter 6)
 - *Problem:* As Semantic Web Services will be deployed in very dynamic environments a mechanism for fault handling and recovery is necessary.
 - *Thesis contributions:*
 - Formal model for fault handling and recovery.
 - Complete specification of recovery actions.

Positive aspects of the thesis:

- 1) The proposed solutions are well founded on a formal basis. Moreover, they also have a practical impact.
- 2) Proposed solutions are evaluated on real data and scenarios.
- 3) Identifying not-yet-identified problems: Semantic Monitoring and Recovery.
- 4) Solutions are complete, clear and directly applicable.

Negative aspects of the thesis:

- 1) *Lack of examples.* There should be more examples demonstrating exploited technologies and introduced solutions to improve the readability of the thesis.
- 2) *Critical evaluation of the proposed solutions in Conclusions.* As current Web Services technologies have not yet achieved the expected level of application in the opened Web environment, there raises a philosophical question whether Semantic Web Services will do so. It is possible that the complexity of semantic technologies (e.g. description logic, specifying preconditions and post conditions of web services) will block their wide spread in the community of web developers.

Questions:

- 1) Could you please discuss the negative aspects, mainly 2)?
- 2) You mentioned the problem of data mediation which is crucial for your solutions to work in real systems. I think that data mediation can not be fully managed automatically but an interaction with a domain expert is necessary. In that case process mediation can not be fully solved in run-time. Do you agree with me? If yes, do you think that a possible way is a standardization of commonly shared ontologies? Don't we get back to current (Non-Semantic) Web Services technologies?
- 3) Benefits of Semantic Web Services over Web Services are known in academia. This is not so clear in case of semantic (complex) event detection. There are solutions known as Complex Event Processing (CEP) systems (HP, IBM, ...) that also allow detection of complex events. Could you please specify benefits of your semantic approach?

Conclusion

The text of the reviewed thesis is clear and the language level is high. The work includes own novel contributions into the area of Semantic Web Services. The presented results were published in 12 peer-reviewed publications. Therefore, the author **proved his ability of stand-alone scientific work**. The thesis conforms to doctoral level and **should be accepted by the committee**.

In Prague, 19th August 2009



Martin Nečaský