SEMANTIC TECHNOLOGIES FOR WORKFLOW AND INFORMATION MANAGEMENT

Wolfgang Marquardt
AVT – Process Systems Engineering & Center for Computational Engineering Science
RWTH Aachen University, Aachen, Germany

Introduction

Product and process design processes (not only but also) in the process industries are of a cooperative nature involving different departments in a single or in more than one enterprise typically at geographically distributed sites. They use and produce a multitude of documents with many interdependencies and overlaps stored in very different electronic formats.

There are many software tools are used to support the project team. Some of them are of a domain specific nature (i.e. process simulators, CAD or CAE systems, etc.) and others are of a generic type and independent of the requirements of chemical engineering design processes (i.e. word processors, project management systems, etc.). An efficient support of chemical engineering design processes requires coherent information technology support across the lifecycle of the project. Existing software tools for dedicated tasks in the project lifecycle have to be enhanced and linked to evolve into a coherent design support system which not only offers user-interface, data and control integration but which rather integrates the distributed, collaborative, and concurrent design process carried out by interdisciplinary teams in different organizations.

To date, no satisfactory solution is available in industry. Therefore, an enormous potential in productivity increase of design teams and consequently in cost reduction and quality improvement exists. Leveraging such potential constitutes a tremendous economical opportunity in particular for enterprises in high wage countries such as Europe, Japan, the US and Canada.

This problem has been subject of the Collaborative Research Center IMPROVE [1], a long-term collaborative research project at RWTH Aachen University, Germany, which has been carried out since 1997 for a period of about 12 years in close collaboration between computer science and chemical engineering research groups. The findings of this project have been summarized in a recently published research monograph [2].

This brief paper addresses the problem from the perspective of semantic technologies. It gives an overview on our recent research which aims at applying and benchmarking ontological modeling and associated semantic technologies to selected key problems arising during chemical engineering design processes. We do extensively refer to our own papers which contain a lot of references to the work of other research groups in the field for the sake of brevity.

Understanding the Work Process

Real industrial work processes are complex and consequently difficult to plan, document, improve and reuse. A fundamental understanding of these work processes is considered to be a prerequisite for their reengineering and for the development of effective information technology based support systems. To this end we have developed a simple and intuitive graphical notation, called C3, which is used to develop a coarse-grained representation of a collaborative work process including the various activities, the actors and their roles, the information used and produced during the activities, and the resources (software tools, experiments etc.) used. An overview on the modeling language and its application to a number of industrial case studies is presented in [3, 4]. The modeling procedure has shown its power in such diverse applications as modeling of product design processes, conceptual process design processes, work processes in detail engineering as well as operational processes.

While the C3 modeling methodology is very useful for the participative modeling of work processes by the design team itself, it lacks the detail and degree of formality to be useful for monitoring the design process, for reasoning on its status, for the detection of inconsistencies of incomplete tasks or for the conceptual design and construction of software tools to support the design process. Therefore, formal information modeling by means of ontologies [5] and their processing with semantic software technologies is a logical next step.

Information Modeling

Information modeling has a long tradition in engineering design processes. Most notably, the STEP [6] modeling activity has been carried out for many years, not only but also in the process engineering domain. ISO 15026 [7] is a prominent example. These formats are intended to be used as data exchange formats in order to integrate software tools used during the engineering design process, to build data warehouses for the integration of existing data stores, or to even implement comprehensive engineering design databases which hold all project-relevant information on the designed process. Ontologies have been used just recently to represent the core concepts of ISO 15026 as an upper ontology [8].

These formats have not gained sufficient attention. Hence, the most prevalent problems of information integration, such as consistency or a desired degree of inconsistency along the project lifecycle, the exchange of information between tools and teams during the design process, the dependencies between documents, or more precisely, between the contents of the documents on different levels of granularity, or even the homogenization of different proprietary data formats has not been successfully addressed by this approach.

One reason of the lack of acceptance seems to be an insufficient capturing of the semantics of the information modeled. Hence, we are exploring the use of more expressive modeling approaches to facilitate the capturing of the detailed semantics of the modeled information items. Ontologies are employed to not only represent the result of the design process ("the product") but also the work process leading this result ("the actions") including the resources employed ("the tools"). The requirements in and the different strategies for application domain modeling are thoroughly discussed in [9], where our own work is also put into the perspective of existing approaches.

OntoCAPE – a Comprehensive Process Engineering Ontology

Object-oriented information modeling has a long tradition in our research group dating back to the early 1990 [10] where information modeling has been addressed in the context of developing advanced software environments for mathematical modeling [11]. OntoCAPE is based on the concepts of general systems theory. While its focus has been largely on the representation of the products of the design process, it also includes concepts to integrate the design process knowledge with the product. Its architectural design is based on abstraction and modularization which guarantees transparency and extensibility. OntoCAPE is represented in OWL [12]. A detailed description of the design principles or even a presentation of its major concepts is beyond the scope of this brief paper. We rather refer to the literature. In particular, an introduction and an overview on OntoCAPE is given in [13, 14], while a series of technical reports available from the author [15-19] present a detailed exposition of the ontology. A comprehensive presentation of the ontology and its design principles will be available in a forth-coming book which will be published early 2009 [20]. OntoCAPE is available under an open source license model at the website of the author (http://www.avt.rwth-aachen.de/AVT/index.php?id=486).

We want to stress that OntoCAPE is a full-fledged ontology which is in contrast to current practice where the term ontology is used in case ontological languages are used for knowledge representation but where all other requirements are not fulfilled [20]. Two types of ontology-like structures can be identified, which can be referred to as *pseudo ontologies* and *lightweight ontologies*. A pseudo-ontology is understood to be a part of a software system formulated in an ontology language, but has not been designed for reuse across various applications which we consider mandatory for a full-fledged ontology. Another differentiating factor is the semantic richness. Often, the so-called lightweight ontologies do not define the semantics of the concepts through axiomatic definitions, but rather utilize only a subset of the available modeling elements such as classes and instances but not relation properties, local constraints, or other forms of axioms. In this sense, OntoCAPE is considered a full-fledged ontology, to our knowledge the most comprehensive if not the only heavyweight ontology in the area of process engineering.

Recent Applications of OntoCAPE

OntoCAPE has been used in a number of large-scale academic and industrial case studies to validate the modeling approach and to demonstrate the applicability and the potential of semantic technologies in process engineering applications.

In the COGents project, funded by the European Union, an early version of OntoCAPE has been used to implement an agent-based system to search and select mathematical models on the web to match a given specification [21]. The ontology has been used to formalize the communication between the agents of the COGents system and to support the match making between a specification and the properties of a certain model. A process data warehouse has been designed and prototypically implemented to demonstrate the use of ontologies for information management in design processes [22] including the capturing of the activities of the designers, the resources used and the design products developed during the design process. Furthermore, an experimental mathematical modeling system ModKit+ has been designed and prototypically implemented using OntoCAPE [23, 24].

An industrial case study is currently in progress where OntoCAPE is used to integrate heterogeneous data in an industrial development process spanning different major tasks. In this project OntoCAPE is used to design a homogenization layer on top of XML-based representations of the content of a number of software tools. OntoStudio, a product of Ontoprise, is used to implement reasoning functionality. The case study is provided by Evonik Degussa. A more detailed description of the scope of this project is given in [25] while first results are included in [26].

Extending OntoCAPE Towards Work Process Modeling

While ontological technologies have been mainly used to represent, integrate and retrieve information on the result of the design process which is used or produced during its various steps, the methodology can be carried over to the formal representation of design products (or documents), the design processes and its resources. First steps towards the development of formal models using semantic modeling and ontological technologies are presented in [27].

This work is currently extended in order to extend OntoCAPE with process modeling capability. In particular, the processes to be modeled are not only confined to product and process design processes but also to operational processes to be applied to the plant. The objective of a current collaborative project aims at the development of such an extension of the ontology and on software tools which support the modeler and which allow an export of the ontological model on either the product and process design or the operational process to another software application which interprets or enacts the process model. Examples include discrete-event simulators to compute performance measures of the design process or a control software based on sequential function charts which implements the operational process of a certain unit. This project is carried out in close cooperation between our research group, Air Products, BASF SE, Bayer Technology Services and Siemens AG. A more detailed description of the project can be found in [28].

Final Comments

Our research on the use of semantic technologies in process engineering has clearly revealed the enormous potential of ontology-based information modeling and of the reasoning capabilities of semantic software technologies which are developed in the context of the semantic web. A properly chosen architecture of the ontology empowers the chemical engineer to extend and modify the ontology by means of high level modeling tools himself. The direct access to the knowledge facilitates maintenance and extension of the information model and the knowledge bases. The reasoning capabilities of modern semantic tools facilitate checking the logical consistency of the information model and the implemented knowledge. We not only have validated the concepts on academic toy problems, but we have implemented prototypical systems of future design support functionality and we are currently validating the methodology, OntoCAPE and the semantic technologies in industrial projects.

There are however also some drawbacks. In particular, the development of an extensible and widely usable ontology is by no means straightforward. Though we believe that OntoCPAE constitutes a very good foundation for a generic and widely usable ontology for process engineering applications, we still see a lot of room for improvement and for extension of the information model and the knowledge bases. Furthermore, the semantic technologies, in particular the ontology editors and reasoning tools are still under development. Performance is an issue for the latter, while usability by application-domain rather than knowledge engineering experts is the key issue for the former. The most important bottleneck at the moment is the lack of few best practice applications and demonstrators in an industrial setting and the lack of detailed application models to be used for the variety of tasks during the lifecycle of a design project or even of the plant itself.

Due to the immense workload required to come up with useful and comprehensive ontologies for process engineering applications, it is highly desirable, that academics and industry join forces in order to develop, maintain and gradually extend a process engineering ontology for process engineering. OntoCAPE is definitely a perfect starting point for such an undertaken.

Acknowledgements

The continuing financial support of Deutsche Forschungsgemeinschaft (DFG) is gratefully appreciated. The author would like to thank his colleagues in computer science, in particular M. Nagl, for many stimulating and fruitful discussions, and his students, in particular R. Hai, J. Morbach, M. Theißen, and A. Wiesner for their contributions to the success of this project and for their support in compiling this paper and its presentation in the conference.

References

1. Marquardt, W., and M. Nagl, (2004), "Workflow and information centered support of design processes - the IMPROVE perspective", *Comp. Chem. Engg*. 29**,** pp**.** 65-82.
2. Nagl, M. and W. Marquardt, eds. (2008): *Collaborative and Distributed Chemical Engineering Design Processes: From Understanding to Substantial Support.* Springer, Berlin.
3. Theißen, M., Hai R. and W. Marquardt (2008), "Design process modeling in chemical engineering", *Journal of Computing and Information Science in Engineering* 8, 011007.
4. Theißen, M., Hai, R., Morbach, J., Schneider, R. and W. Marquardt (2008), *"Scenario-based analysis of industrial work processes"*, in M. Nagl and W. Marquardt (2008), Chapt. 5.1.
5. Uschold, M. and M. Grüninger (1996), "Ontologies: principles, methods and applications," *Knowl. Eng. Rev.* 11 (2), pp. 93–136.
6. ISO10303, STandard for the Exchange of Product data
7. ISO15926, Industrial Automation Systems and Integration – Integration of Life-Cycle Data for Process Plants Including Oil and Gas Production Facilities
8. Batres, R., West, M., Leal, D., Price, D., Masaki, K., Shimada, Y., Fuchino, T. and Y. Naka (2007), "An upper ontology based on ISO 15926", Comp. Chem. Eng. 31 (5/6), 519–534.
9. Morbach, J., Theißen, M., Hai, R. and W. Marquardt (2008), "An Introduction to Application Domain Modeling," in M. Nagl and W. Marquardt (2008).
10. Marquardt, W., Gerstlauer, A., Gilles, E.D. (1993), „Modeling and Representation of Complex Objects: a Chemical Engineering Perspective", 6th Int. Conf. on Industrial and Engineering Applications to Artificial Intelligence and Expert Systems, Edinburgh, Scotland, pp. 219-228.
11. Bogusch, R., Lohmann, B. and W. Marquardt (2001), "Computer-Aided Process Modeling with MODKIT, *Comput. Chem. Eng*g. 25, pp. 963-995.
12. W3C (2002), *"The OWL representation ontology"*, Online available at www.w3.org/2002/ 07/owl, accessed October 2007.
13. Morbach, J., Yang, A. and W. Marquardt (2007), "OntoCAPE – a large-scale ontology for chemical process engineering," *Engineering Applications of Artificial Intelligence* 20:2, pp. 147–161.
14. Morbach, J., Bayer, B., Yang, A., and W. Marquardt (2008), "Product Data Models", in M. Nagl and W. Marquardt (2008), Chapter 2.2.
15. Morbach, J., Bayer, B., Wiesner, A., Yang, A. and W. Marquardt (2008), *„OntoCAPE 2.0 – The Upper Level"*, Technical Report (LPT-2008-25), Lehrstuhl für Prozesstechnik, RWTH Aachen University.
16. Morbach, J., Wiesner, A. and W. Marquardt (2008), "*OntoCAPE 2.0 – The Meta Model"*, Technical Report (LPT-2008-24), Lehrstuhl für Prozesstechnik, RWTH Aachen University.
17. Morbach, J., Yang, A., Wiesner, A. and W. Marquardt (2008), *„OntoCAPE 2.0 – Supporting Concepts"*, Technical Report (LPT-2008-26), Lehrstuhl für Prozesstechnik, RWTH Aachen University.
18. Morbach, J., Yang, A. and W. Marquardt (2008), *"OntoCAPE 2.0 – Materials*", Technical Report (LPT-2008-27), Lehrstuhl für Prozesstechnik, RWTH Aachen University.
19. Morbach, J., Yang, A. and W. Marquardt (2008), *"OntoCAPE 2.0 – Mathematical Models"*,

Technical Report (LPT-2008-28), Lehrstuhl für Prozesstechnik, RWTH Aachen University.

20. Marquardt, W., Morbach, J., Wiesner, A. and A. Yang (2009), *“OntoCAPE – A (Re)Usable Ontology for Chemical Process Engineering”*, Springer, Berlin, in preparation.
21. Yang, A., Braunschweig, B., Fraga, E.S., Guessoum, Z., Marquardt, W., Nadjemi, O., Paen, D., Piñol, D., Roux, P., Sama, S., Serra, M., and I. Stalker (2007), “A multi-agent system to facilitate component-based process modelling and design”, *Comp. Chem. Engg.* 32, pp. 2290-2305.
22. Brandt, S.C., Morbach, J., Miatidis, M., Theißen, M., Jarke, M., and W. Marquardt (2008), “An ontology-based approach to knowledge management in design processes”, *Comp. Chem. Engg.* 32, pp. 320–342.
23. Yang, A., Morbach, J. and W. Marquardt (2004*), “From conceptualization to model generation: the roles of ontologies in process modeling”*, Proc. Sixth International Conference on Foundations of Computer-Aided Process Design, pp. 591-594.
24. Yang, A. and W. Marquardt (2004), *“An ontology-based approach to conceptual process modelling”,* Proc. of ESCAPE-14 (European Symposium on Computer Aided Process Engineering), Lisbon, Portugal, pp. 1159-1164.
25. Morbach, J. and W. Marquardt (2008), “Ontology-based integration and management of distributed design data”, in M. Nagl and W. Marquardt (2008), pp. 647–655.
26. Wiesner, A., Morbach, J. and W. Marquardt (2008), “Semantic data integration for process engineering design data”, Proceedings of the 10th International Conference on Enterprise Information System - ICEIS, pp. 190-195.
27. Eggersmann, M., Kausch, B., Luczak, H, Marquardt, W., Schlick, C., Schneider, N., Schneider, R. and M. Theißen (2008), „Work process models“, in M. Nagl and W. Marquardt (2008), Chapt. 2.4.
28. Theißen, M., Hai, R. and W. Marquardt (2008), “Computer-assisted work process modeling in chemical engineering”, in Nagl, M. and W. Marquardt (2008), pp. 656–666.