Proceedings of I-KNOW '05
Graz, Austria, June 29 - July 1, 2005

A Methodological Approach for Constructing Ontology-Based Reference Models in Digital Production Engineering

Mark Hefke, Peter Szulman, Adrian Trifu
FZI Forschungszentrum Informatik, Haid-u-Neu-Str. 10-14, 76131 Karlsruhe, Germany
{hefke, szulman, trifu}@fzi.de

Abstract: In the digital planning process of a manufacturing plant, several partners like OEM, prime contractor and its subcontractors are involved. Since the partners have partially overlapping views (electricity, mechanical structure, plant controlling) on the plant to be designed, they have to exchange data during their collaboration. Due to syntactical, structural and semantical differences, data integration is necessary but also complicated. Our method of resolution comprises an ontology-based reference model, which all partners map to as well as an underlying technical infrastructure. This paper focuses on the methodology for constructing an ontology-based reference model in digital production engineering.

Keywords: Semantic integration, manufacturing plant, data heterogeneity, reference model, ontology
Categories: H.3.2, H.3.3, H.4.3, H.5.3

1 Introduction

The planning process for a manufacturing plant demands the cooperation of several independent partners. As the mechanical concept planning is in general realized by the original equipment manufacturer (OEM), electrical planning and PLC programming of the control system are usually carried out by a prime contractor and its subcontractors. Achieving a fully digital, integrated planning process is hampered by the syntactical, structural and semantic heterogeneity of each partner's view on a shared amount of data. Therefore, our aim in the research project MODALE[1] is to overcome this by using a virtual data integration approach, based on a reference model of the plant, which is modelled as an ontology common to all partners of the digital planning process [Schmidgall et al. 05, Szulman et al. 05]. The key idea is to represent the semantic meaning of that part of data in each partner's data schema, that has to be exchanged with others. This will be done by building up component ontologies and mapping them in a second step onto a central reference model, in order to allow common data exchange and seamless execution of appropriate data transformations on demand. When applying ontologies in general, the meaning of data is representable by using machine understandable formal semantic descriptions. However "lightweight semantic descriptions" like XML schemas, together with XML documents become well spread nowadays promising to overcome problems resulting from data heterogeneity, but do not fulfil our requirements in this context, as they are

[1] http://www.modale.de

not able to represent complex semantic information, but only a description of a simple document structure. Furthermore, formal ontologies are necessary for executing complex queries, like e.g. "give me all geometric information from partner OEM during the process step preliminary design of the plant", or even ontology reasoning. As earlier described every partner has its own partner ontology, which represents his view on the plant and maps it onto the common reference model by the use of semantic bridges [Maedche et al 03]. In this paper, we now focus on methodological aspects of constructing and reusing such a common reference model and connecting it to existing partner ontologies for accomplishing digital planning projects. The paper is structured as follows. Section 2 gives a short presentation of our reference models and refers to our previous publications. Section 3 is dedicated to describing the integration methodology, followed by tool support in section 4. In section 5, we give a short overview of related work and conclude in section 6 by providing insight into our future work.

2 Ontology-based reference models

This section is a brief description of our ontology-based reference model approach. For a detailed presentation, please refer to [Szulman et al. 05].
Ontology-based reference models are based on two main ideas, the first is to use a central, partner-independent model of the plant or installation (a reference model), as a sort of universal exchange format. Partners that participate in the data exchange do not need one-to-one interfaces to each other, which would potentially lead to n-to-m interfaces, but instead each partner specifies one interface to the central reference model.The second main idea is to use ontologies and semantic bridges [Maedche et al. 03] in order to overcome semantic heterogeneity. Apart from the central reference model, each partner has its independent (partner) ontology, which is specific to its domain, used tools and underlying data models, etc. In order to enable data exchange, we use the MAFRA approach [Maedche et al 03] to specify semantic bridges between concepts and properties in the reference ontology (model) and partner-specific ontologies. Semantic bridges are directed, and can specify transformation rules for the instance data. As previously suggested, the purpose of a common reference model is to facilitate data and information exchange between individual partners cooperating in the digital planning process. As a natural consequence, the main focus of the reference model lies on the "intersection regions" between partners. These "regions" can be defined as the subset of all produced information that is of interest to at least two of the cooperating partners. In the course of the planning process, several types of models are used. Conceptually, they reside on three distinct levels of abstraction (see Figure 1): the instance level, the concept level and the concept bridging level.

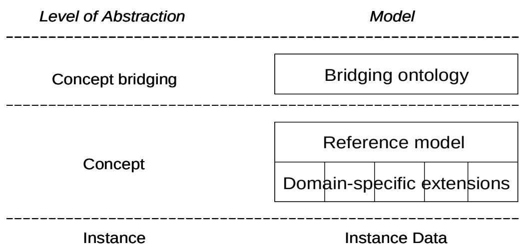

Figure 1: Levels of abstraction and their corresponding model types

The reference model (RM) describes the global uniform view of the plant using commonly accepted or standardized concepts and properties (attributes and relations) and may have domain specific extensions. Its concepts/ properties are either mapped onto or are being mapped onto those on the partner models. The RM can also be used as an integrated view on which global queries can be formulated. The RM concepts/ properties and those of the partner ontologies are linked using semantic bridges, as described in [Maedche et al 03]. A so-called bridging ontology, presented in Figure 2, is a meta-model for these bridges.

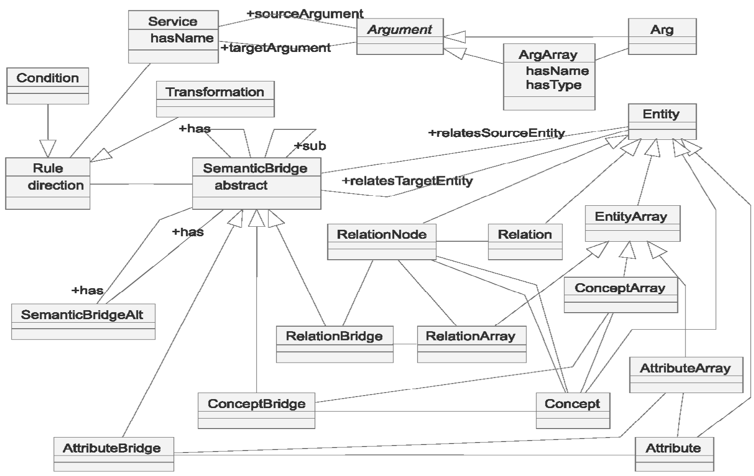

Figure 2: The bridging ontology

When concept or property instantiations are to be transferred from one partner (source) to another (destination), the two bridges are crossed (from source partner to reference model and from reference model to destination partner), and appropriate transformations are executed on the transferred instances. As an example, let us consider two partners (Partner A and B) that need to exchange geometrical data concerning a mechanical component designed by partner A. Partner A uses Tool A to design the component, and partner B uses Tool B to visualize the design. The tools use different file formats (syntax, structure) and semantics for representing and storing geometrical data and schemas: Tool A represents 2D segments as pairs of 2D Points, while Tool B represents them as vectors. The first step is to capture the two schemas in two ontologies. The two relevant ontology fragments relating to our example are shown in Figure 3. In the first ontology, a segment is characterized by an ID, start and an end point, which in turn has an ID and two coordinates. In ontology B, points are defined similarly, but there is no concept called segment. Instead we have a vector, which contains ID, modulus and an angle alpha. In order to enable the two tools to interoperate, we need to map concepts, attributes and relations from ontology A to concepts, attributes and relations in ontology B.

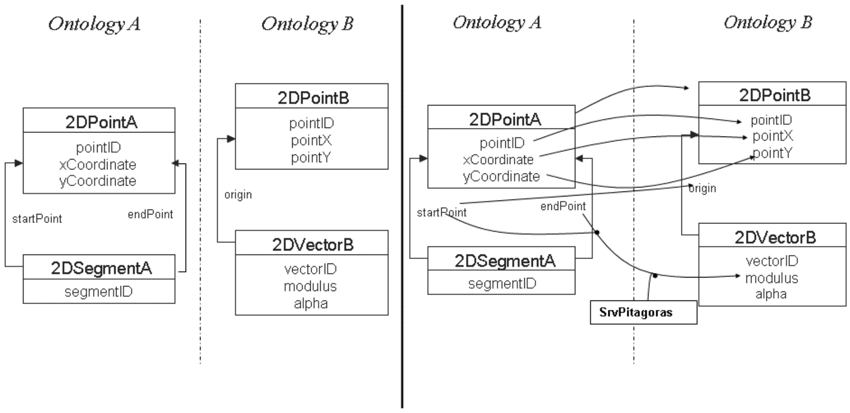

Figure 3: Two incompatible concept worlds (left), Ontology fragments connected by semantic bridges (right)

An obvious concept bridge can be instantiated between “2DPointA” and “2DPointB”. Their corresponding attributes can also be mapped without difficulty (“pointID” to “pointID”, “xCoordinate” to “pointX” and “yCoordinate” to “pointY”). Using a relation bridge, relation “startPoint” can be mapped to the relation “origin”. In order to transform segments to vectors, two attribute bridges are defined between “startPoint” and “endPoint” (belonging to 2DSegmentA) and modulus and alpha (belonging to 2DVectorB). These bridges are “2 to 1” bridges, each augmented with an appropriate transformation. For example, the transformation rule associated to the bridge for “modulus” will use a service called “SrvPitagoras” which will have to be invoked to compute the vector’s modulus, based on the coordinates of its two extremities. The two ontology fragments and the semantic bridges connecting them are shown in Figure 3. Now having presented the situation at runtime, the obvious question arises: *how do we get there?*

3 Methodology for constructing the reference model

It was clear from the very beginning that the two most important requirements for the methodology are firstly to reflect the iterative nature of the planning process, and secondly to favour the emergence of a project-independent standard, a reusable model, which we call a Standard Reference Model (SRM), as opposed to the Project-specific Reference Model (PRM). The first requirement is determined by the fact that the plant model is only complete at the end of the project Therefore the PRM must evolve continuously, as new parts are being developed and new data is generated. In order to accelerate this evolution process, an explicit desire was to derive a standard model kernel (the SRM), specific to plant manufacturing, which would be project-independent and could be reused over and over. The SRM should also be able to evolve over several projects, until it reaches a mature status. We started from a typical planning process for a digital plant as the basis for our methodology. It is important to note that we are only interested in the inter-partner interactions, and not the partner-internal processes or workflows. The integration methodology is iterative, and consists of so called *integration iterations*, which take place when the need to exchange information between partners becomes evident (e.g. when the prime contractor needs to receive the product design from the OEM, or when the prime contractor wishes to assign tasks to its subcontractors). Therefore in an integration iteration, there are always well defined information sources, information sinks and actual data which have to be transferred (see Figure 4).

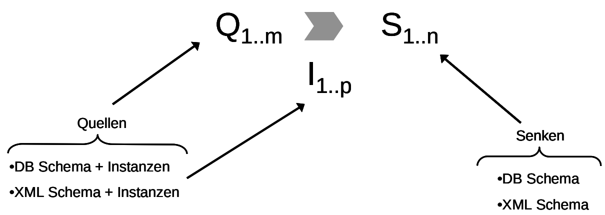

Figure 4: A representation of an integration iteration

With this in mind and given a bootstrapped PRM (initial PRM is identical to (a subset of) the SRM), the structure of a single iteration is the following:

1. ***creation of partner ontologies*****:** both source and sink partners must adapt or extend their partner ontologies with the model of the data to be exchanged. This step typically occurs once for every tool that is used, and does not need to be repeated in the future, if the tools do not change. This task is a potentially very complex modelling task. If the current reference model already describes a concept world specific for a given partner (e.g. one of the domain specific extensions of the SRM), that partner may opt for reusing that extension. This later simplifies the mapping process very much, because only one-to-one bridges with no transformation rules need to be defined between the PRM and the partner ontology. However, the partner is not restricted by the PRM in any way, and may model his own concept world as it best suits him.
2. ***semantic harmonization*****:** This is the most challenging task and consists of specifying the mappings (semantic bridges) between the local and the PRM ontologies. This is a need-driven process, that is, the information sinks have

a L-A-V (local-as-view) view on the PRM, which in turn has a L-A-V view on sources. If needed, the PRM ontology needs to be extended with appropriate concepts. In such cases the involved partners must negotiate an acceptable "standard". This standard can be either copied/derived from one of the partner-ontologies, or a partner-neutral standard (such as STEP) may be used. Once mappings are defined, data transformation rules can be specified for them. These rules describe the way in which instances from a source (or several sources), are to be transformed into instances of the information sink(s);

3. ***data exchange***: based on the two sets of mappings (from sources to PRM and from PRM to the sinks) appropriate transformations are applied to instances of the sources and resulting instances are transferred to the sinks. Selection of the source instances can be one of: *all* (all source instances are transfered), *specific* (a specific instance or set of instances is transfered), or *query-based* (the set of instances to be transferred is the result of a query, either on the PRM, or ideally on the partner ontology). References between instances on the sink side must reflect references between source instances. This is taken care of by the semantic reconciliation engine of the MAFRA framework (see [Harmonise 03]).
4. ***updating* SRM**: After the project has been finished, the SRM has to be updated according to the changes in the PRM. From a technical point of view this means merging two ontologies (the SRM and the new version of the PRM). On the other hand, this process is more than just a merging procedure but rather a standardization process.

We are currently performing a validation of the methodology with a number of industrial partners in the automotive industry. Based on their specific needs and state-of-the-art tools used in the industry, we bootstrapped a very simple initial SRM, which has the structure as shown in Figure 5.

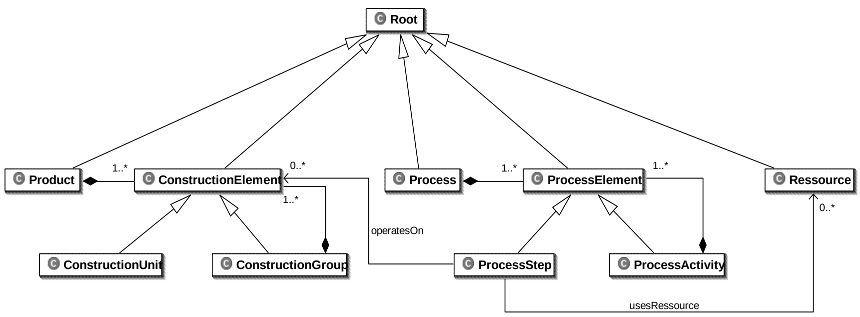

Figure 5: The initial SRM structure

The SRM basically corresponds to the high-level view that the OEM has on the *Product* (e.g. the geometry of the car to be produced), the manufacturing *Process* (e.g. in what sequence welding robots are activated), and the *Resources* that are used in the process (e.g. welding robots, command and control modules, etc).

4 Tool support

Our tool support is based on the KAON tool suite (http://kaon.semanticweb.org) as well as the MAFRA mapping framework for distributed ontologies (http://mafra-toolkit.sf.net/), both were developed at our institute. As a proof of concept, we provided a web-service based prototypic implementation of our approach (the architecture of the prototype is shown in Figure 6).

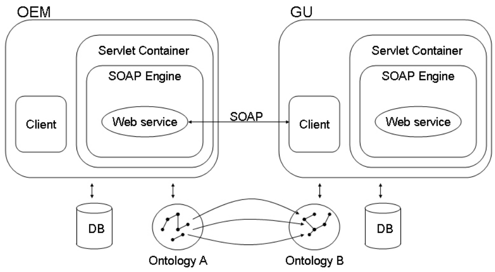

Figure 6: Architecture of the prototype

Experiences showed, that traditional ontology management tools are often too complex with regard to functionality and handling. Some partners preferred to work with their internally well-established methods and office tools, in order to describe their view on a specific domain. Therefore we e.g. developed interoperability support between the KAON ontology language and the two formats EXCEL and UML (excel2kaon/ UML2KAON).

5 Related Work

At this stage, there exist several methodological approaches and (ontology-based) infrastructures for supporting materialized or virtual integration of heterogeneous information sources, e.g. BUSTER[2] (Bremen University Semantic Translator for Enhanced Retrieval), Infomaster[3], MOMIS[4] (Mediator envirOnment for Multiple Information Sources), TSIMMIS[5] (The Stanford-IBM Manager of Multiple Information Sources), MAFRA, which provides a basis in our approach and many others. An overview and detailed SOTA-description with regard to (ontology-based) information integration approaches is also given in [P. Szulman, A. Trifu, 04]. In digital production engineering, the IDA-STEP (Integration distributed Applications on the Basis of STEP Data Models) approach, which is based on the ISO standard STEP (Product Data Representation and Exchange), supports the integration of CAD, CAM, and PDM data. The ISO standard PLIB allows the representation and exchange of part library data. However, from our point of view, our approach of providing a

[2] http://www.semantic-translation.de/
[3] http://infomaster.stanford.edu/infomaster-info.html
[4] http://dbgroup.unimo.it/Momis/
[5] http://www-db.stanford.edu/tsimmis/tsimmis.html

reusable SRM ontology for integrating heterogeneous information sources in digital production engineering, a methodology for the (tool-supported) build-up of such a reference model, as well as an underlying ontology-based integration infrastructure is rather unique in this context.

6 Summary and Future Work

In this paper, we have described a methodological approach for constructing ontology-based reference models in digital production engineering, which are the basis for virtual data and information integration between cooperating partners. In this context, we have presented a methodology for the process of setting up such a reference model and hooking up all participants/ preparing them for data exchange. As tool support, we rely on already established tools, also developed in our institute, such as KAON and MAFRA. The paper presents work in progress within the BMBF-funded project MODALE. For the future, we identified a number of open issues, which we would like to address. First, we would like to improve tool support in the modeling process by investigating ways of automatic ontology extraction from data base schema. Particularly interesting is the question of how to combine a top-down modeling approach (the way humans think) with a bottom-up approach (which results from automatic ontology extraction). Second, we wish to improve the mapping processes by using various similarity measures to infer possible semantic bridges between concepts in separate ontologies. Furthermore, we intend to integrate our mapping tool with (semi-) automatically generated data dictionaries, in order to help domain and/or modelling experts faster understand "foreign domains", during the mapping process.

References

[T. R. Gruber 93] T. R. Gruber: “Toward Principles for the Design of Ontologies Used for Knowledge Sharing”, International Workshop on Formal Ontology, Padoba, Italy.

[A. Maedche, Boris Motik, Nuno Silva, Raphael Volz, 03] A. Maedche, B. Motik, N. Silva, R. Volz: “MAFRA – A Mapping FRAmework for Distributed Ontologies”, Proceedings of the 13th European Conference on Knowledge Engineering and Knowledge Management EKAW, Madrid, Spain.

[Harmonise, 03]: “Deliverable 3.2: Semantic mapping and Reconciliation Engine subsystems”, The Harmonise Project, IST-2000-29329.

[G. Schmidgall, J. Kiefer, T. Bär, 05] G. Schmidgall, J. Kiefer, T. Bär: “Objectives of Integrated Digital Production Engineering in the Automotive Industry”, Proceedings of the 16th IFAC World Congress, Prague, Czech Republic.

[P. Szulman et al., 05] P. Szulman, M. Hefke, A. Trifu, M. Soto, D. Assmann, J. Doerr, M. Eisenbarth: “Using Ontology-Based Reference Models in Digital Production Engineering Integration”, Proceedings of the 16th IFAC World Congress, Prague, Czech Republic.

[P. Szulman, A. Trifu, 04] P. Szulman, Adrian Trifu: “MODALE, State of The Art on Information Integration”, Project deliverable, http://www.modale.de/Modale/cms/ Arbeitspakete/AP2/