Semantic mapping: ontology-based vs. model-based approach
Alternative or complementary approaches?

K. Arnarsdóttir[1], A.-J. Berre[2], A. Hahn[3], M. Missikoff[4], F. Taglino[4]

[1] runa@heimaey.com (kristrun.arnarsdottir@aetat.no)
[2] SINTEF ICT, P. O. Box 124 Blindern, N-0314 Oslo, Norway, arne.j.berre@sintef.no
[3] University of Oldenburg, Ammerländer Heerstr. 118, D-26111 Oldenburg, Germany,
[4] CNR-IASI, LEKS. Viale Manzoni, 30 – 00185 Rome, Italy

Abstract: Model transformation and semantic mapping are enabling technologies for new, advanced solutions to address enterprise software interoperability. Such technlogies are nowdays tackled by two different disciplines: software architectures, with the MDA approach, and semantic interoperability, with an ontology-based approach. This paper briefly compares the model-based (MB) and the ontology-based (OB) approaches and draws some preliminary conclusions on similarities and differences, in the perspective of combining these approaches in a value-adding way.

Keywords: model mapping, semantic mapping, ontology

1. Introduction

Interoperability is one of the major challenges to be addressed in achieving efficient software application cooperation, within and among enterprises. Today, the integration costs for enterprise applications cooperation are still extremely high, because of different business processes, data organization, application interfaces that need to be reconciled, typically with great manual (and therefore error prone) intervention. This problem has been addressed independently by MDA and ontology-based approaches.

The Model Driven Architecture (MDA) proposed by the Object Management Group (OMG)[1] uses platform-independent models (PIMs) 6 as the context for identifying relations between different applications. Transformation is a central concept in MDA to address how to convert one model into another model of the same system, and further into executable code. Today's de facto standard for creating software models, in the context of an object-oriented approach for system design, is the Unified Modelling Language (UML) 9. MDA and UML 2.0 provide technologies

[1] OMG public website: http://www.omg.org/

to handle meta models, constraints etc. which can be used for semantic enrichment and model transformation.

Today, ontology technologies have reached a good level of maturity and their applications to industrial relevant problems are proliferating. Ontologies are the key elements of the Semantic Web. The notion of the Semantic Web is led by W3C[2] and defined to be a "common framework allowing data to be shared and reused across application, enterprise and community boundaries" 12. Web Ontology Language (OWL) 10 is the recommended standard for building ontologies in the context of the Semantic Web. Ontologies support semantic mapping construction by providing explicitly defined meaning of the information to be exchanged.

In this paper, we look at how the two different approaches, model-based and ontology-based, can be used to address semantic mapping. We contrast them to verify if they are two alternative approaches or whether these can be combined in a value-adding way. We conclude this short paper by indicating a few similarities and differences, in the light of devising a hybrid approach as a hypothesis for further research.

2. Comparing the two approaches to semantic mapping

Model-based and ontology-based approaches can be viewed as two solutions for addressing semantic interoperability. Both approaches start addressing two conceptual models where semantically related concepts are to be identified. The main difference resides in the fact that model-driven approach aims at finding the semantic mapping directly starting from the two models, say PIM A and PIM B, deriving then the PIM2PIM mapping. Conversely, the ontology-based approach does it indirectly, by means of a Reference Ontology. Therefore, in the latter case, the semantic mapping is obtained by the composition of two partial PIM2ONT and ONT2PIM mappings. This difference in expressing targets for semantic mapping is sketchily illustrated in **Fig. 1**. Note that the model-based also might use mapping through an intermediate model, if a standard model for the domain exists.

In Athena 5, a large European IST Integrate project, the two different technologies have been applied to support model mapping. Semantic mapping involves the application of an ontology. Current literature does not provide detailed description regarding how this is to be done, as pointed out by 13 and 3. In Athena, a solution as been proposed, based on semantic annotation (A* tool), reconciliation rules generation (Argos tool), and a reconciliation execution engine (Ares). Parallely, in Athena, also a model-based approach has been proposed, based on a graphic tool (Semaphore) aimed at supporting the user in specifying the mappings and XSLT based transformation rules.

[2] W3C: World Wide Consortium - public website: http://www.w3.org/

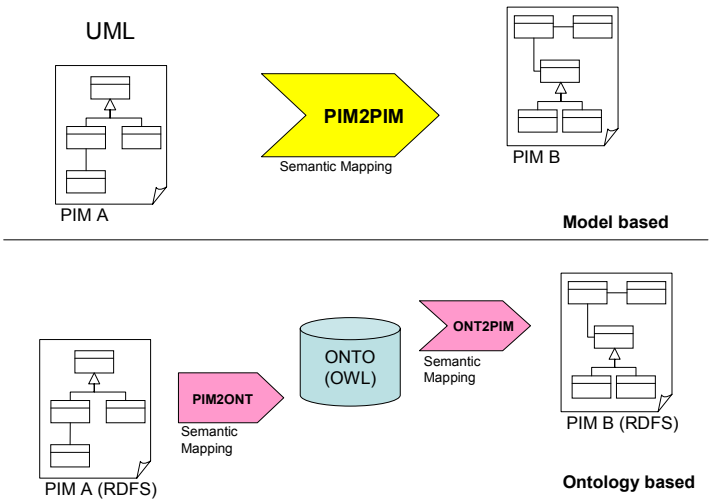

Fig. 1. The model based and the Ontology based semantic mapping approaches

In the ontology-based approach, OWL is used to model a Reference Ontology, while RDFS is used to model schemas to be mapped. Model to ontology mappings are defined with a layered Semantic Annotation (SA) approach. Rule based languages such as Jena-rules 7 are used when creating a set of executable reconciliation rules, starting from SA expressions. The reconciliation engine is based on the inference engine provided by the Jena platform, and the data interchange messages, input to the rules, are in RDF format.

In model-based approach, UML is used to express conceptual models. A MOF compliant language is defined as part of the solution in order to capture relationships between data elements. Transformation languages are to be used to create executable rules, and transformation techniques can be used in the process of detailing the information needed, converting form more abstract MOF compliant languages to more formal ones.

In order to compare the two different approaches, relative strengths have to be identified, within the Semantic mapping context and the related tasks.

Fig. 2 illustrates the different tasks involved in the interoperability solving process.

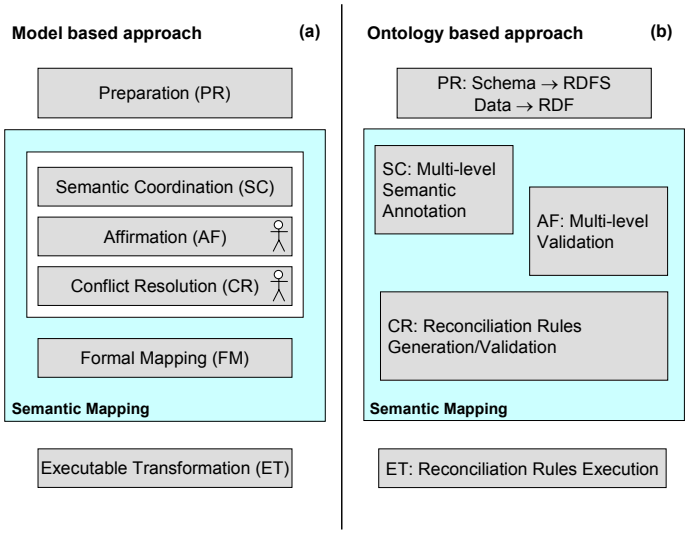

Fig. 2. Semantic mapping phased approach: (a) Model based; (b) Ontology based

Following is an explanation of the different tasks:

- **Preparation** (PR) is about preparing the data sources for semantic mapping. In case of the model-based approach, the context used is PIM and the task involves converting source and target for semantic mapping to this context. This includes applying reverse engineering technology, which in MDA context is referred to as Architecture Driven Modernization (ADM)[3]. The preparation phase within the ontology-based approach involves converting data sources to RDF(S) format, in particular, schemas into RDFS and actual data into RDF.
- **Semantic Coordination** (SC) is about automatically detecting evidence for semantic relationships using available algorithms similar to what described in 8 and 2. The term chosen here is influenced by 1. In the MB approach, this phase concerns the identification of mappings by directly contrasting the two PIM models. This includes application of auxiliary resources such as WordNet to aid in the process of identifying relationships. In the OB approach, this phase consists in the Semantic Annotation of the resources to be reconciled (according to the multi-level approach developed in the A* tool) by contrasting them with the Reference Ontology.
- **Affirmation** (AF) is about confirming or rejecting the automatically identified relationships from previous tasks as well as identifying new ones. In the MB approach, this phase consists in a final validation of the previous identified mappings, while in the OB there is a stepwise validation activity, performed at each of the four levels of the A* annotation method.
- **Conflict Resolution** (CR) is about refining mapping relationships with additional information needed to resolve conflicts/clashes. In the MB approach, this is

[3] Architecture-Driven Modernization Task Force, http://www.omg.org/adm/

achieved within the graphical mapping tool. In the OB approach, this is performed in the first phase of the reconciliation rules definition, implemented by the Argos tool of Athena.

- **Formal Mappings** (FM) is about transforming previously captured information into formal machine interpretable expressions. In the MB approach, mappings are represented through QVT 14 (Query View Transformation) expressions, while in the OB approach, they are represented by means of Jena2 rules, generated and managed by Argos.
- **Executable Transformation** (ET) is not a human task but represents the application of the final output of the semantic mapping process, which is a set of executable rules needed to convert between different messages, having different structures and labelling information. In the MB approach this task is achieved by the UMT-QVT, while in the OB approach it is achieved by the Ares engine, part of the Athena Semantic Framework.

3. Conclusions

From the above analysis, necessarily sketchy, it emerges that the two analyzed approaches are similar in the global process deployed to solve the interoperability problem, but the specific steps appear to be quite different. The main difference that emerges is represented by the possibility of a direct mapping in the case of model-based approach, while the ontology-based approach requires the development of a Reference Ontology (see Fig. 1). Another important difference is represented by the two layering approaches adopted. The MB approach proposes a layered approach in the system modeling phase, based on the three MDA layers, CIM, PIM, PSM, with a special emphasis on the intermediate PIM level. Conversely, the OB approach proposes a 4 layers approach, in the semantic mapping phase, based on 4 different levels of mappings expressiveness: terminological, path, structural, and semantic mappings. In the OB approach, the Affirmation phase (called: validation) is performed interleaved with each step of the Semantic Coordination phase: i.e., annotation and validation are tightly interwoven; conversely, in the MB approach, the AF phase is unique, strictly following the SC phase (see **Fig. 2**). A final difference is the representation framework. In the MB approach representation is mainly diagrammatic, strongly rooted in the UML and MOF standards, with the use of QVT or XSLT for the transformation mechanisms and XMI for the exchanged resources. In the OB approach, the representation of the resources and the mapping are mainly based on Semantic Web standards: RDF(S) for the exchanged resources, OWL for the Reference Ontology, and Jena-rules for the reconciliation mechanisms.

In a future, more extended work, we intend to analyse more in details the above steps, by using a specific use case drawn from the Athena scenarios. The idea is to fully develop the same case independently with the two approaches, then we will analyse the results and will try to propose a unified approach where the most effective solutions from both sides will be extracted and merged together.

4. References

1. P. Bouquet, L. Serafini and S. Zanobini, *Semantic Coordination: A New Approach and an Application*, Lecture Notes in Computer Science, 2870/2003 (2003), pp. 130-145.
2. L. v. Elst and M. Kiesel, *Generating and Integrating Evidence for Ontology Mappings*, in E. Motta, N. Shadbolt, A. Stutt and N. Gibbins, eds., EKAW 2004: Engineering Knowledge in the Age of the Semantic Web. 14th International Conference., Lecture Notes in Computer Science. Springer-Verlag, Whittlrebury Hall, UK, 2004, pp. 15-29. http://www.dfki.uni-kl.de/~elst/papers/ekaw04_LvE_MK_final.pdf
3. A. Gómez-Pérez, M. Fernández-López and O. Corcho, *Ontological Engineering with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*, Springer-Verlag, London, 2004.
4. IDA, *IDA e-Procurement Protocol XML Schemas Initiatives. IDA working document*, European Communities, 2004. http://xml.coverpages.org/IDA-OrderInvoice18083.pdf
5. IST-507849, ATHENA *General Description*, ATHENA Consortium, 2004. http://www.athena-ip.org/components/com_docman/dl2.php?archive=0&file=QXRoZW5hX0dlbmVyYWxfRGVzY3JpcHRpb25fdjEwLnBkZg==
6. A. Kleppe, J. Warmer and W. Bast, *MDA Explained. The Model Driven Architecture: Practice and Promise*, Pearson Education Inc., Boston, MA, 2003.
7. OpenSource, *Jena - A Semantic Web Framework for Java - Jena 2 Inference support*, SourceForge.net, 2003. http://jena.sourceforge.net/inference/
8. E. Rahm and P. A. Bernstein, *A survey of approaches to automatic schema matching*, VLDB, 10 (2001), pp. 334-350.
9. J. Rumbaugh, I. Jacobson and G. Booch, *The Unified Modelling Language Reference Manual, Second Edition*, Pearson Education, Inc., Boston, 2005.
10. W3C, *OWL Web Ontology Language Overview*, in F. v. Harmelen and D. McGuinnes, eds., W3C Recommendation, W3C, 2004. http://www.w3.org/TR/owl-features/
11. W3C, *Resource Description Framework (RDF): Concepts and abstract syntax*, in G. Kline and J. J. Caroll, eds., W3C Recommendation, W3C, 2004. http://www.w3.org/TR/rdf-concepts/
12. W3C, Semantic Web, 2005. http://www.w3.org/2001/sw/
13. T. Wache, V. T., U. Visser, H. Stuckemschmidt, G. Schuster, H. Neumann and S. Hübner, *Ontology-Based Integration of Information - A Survey of Existing Approaches*, IJCAI-01 Workshop - Ontologies and Information Sharing, Published on CEUR-WS: 16-Aug-2001, Seattle, Washington, USA, 2001, pp. 108-118. http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-47/wache.pdf
14. *MOF QVT Final Adopted Specification*, http://www.omg.org/docs/ptc/05-11-01.pdf