



Semantics Web and Ontology Learning

Kaninda Musumbu

► To cite this version:

Kaninda Musumbu. Semantics Web and Ontology Learning. 3rd International Conference On Application Of Information And Communication Technology And Statistics In Economy And Education, Dec 2013, Sofia, Bulgaria. pp.s7-4. hal-00965607

HAL Id: hal-00965607

<https://hal.science/hal-00965607>

Submitted on 28 Mar 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Semantics Web and Ontology Learning

K. Musumbu

LaBRI (UMR 5800 du CNRS),
351, cours de la Libération, F-33.405 TALENCE Cedex
e-mail: musumbu@labri.fr

Abstract

The Semantics Web is a vision for the future of the Web in which information is given explicit meaning, making it easier for machines to automatically process and integrate information available on the Web. An ontology defines the terms used to describe and represent an area of knowledge. Ontologies are used by people, databases, and applications that need to share domain information (a domain is just a specific subject area or area of knowledge, like medicine, tool manufacturing, real estate, automobile repair, financial management, etc). The amount of data caused by the success of Internet is demanding methodologies and tools to automatically extract unknown and potentially useful knowledge out of it, generating structured representations with that knowledge. The main goal of this paper is to present comprehensive introduction into MDA-based ontology development. It will provide an introduction to the field of the Semantics Web and ontology engineering, a description of several UML- and metamodeling- based solutions and tools for ontology development.

Keywords: Business rules, knowledge based systems, model driven architecture, reasoning, semantics web, ontology.

1 Introduction

Certain commonly used terms had multiple meanings all equally valid but which we had not differentiated adequately, leading to much confusion. To avoid this confusion one must define an ontology, which included semiformal definitions of all key terms and their relationships. The term ontology means a specification of a conceptualization. It is used to describe the highlevel structure of the software in the most direct and most expressive manner possible. The Semantic Web endows the network of disparate information that is currently accessible on the Internet with meaning. Because this meaning can be gleaned and processed automatically by software, the Semantic Web opens up the exciting and awe-inducing possibility of a unified global intelligence accessible to all. The e-learning is the use of the new multimedia technologies of the Internet to improve the quality of the learning by facilitating on one hand the access to resources and to services, on the other hand the exchanges and the collaboration in distance. The main goal of this paper is to present comprehensive introduction into MDA-based ontology development. It will provide an introduction to the field of the Semantics Web and ontology engineering, a description of several UML- and metamodeling-based solutions and tools for ontology development. The objectif is to show how the ontologies can be useful in the domain of e-learning. This paper will be divided in three parts. The first covers the basis of both main topics ontology, semantic web and standards. The second part, explain the Model Driven Architecture. The last is the central one, it starts with a review of several approaches and aim to bridge the gap between ontology development and software engineering methodologies.

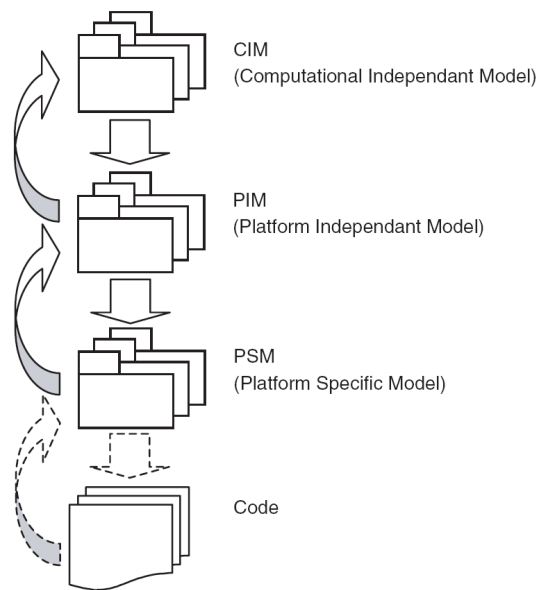


Figure 1: Global view of the Model Driven Architecture approach

2 Semantic Web

The Semantic Web has the potential for semantically richer representations of things (e.g., Web pages, applications, and persons) and their relations on the Web, and thus should provide us with more intelligent services. In MDA, an instance of MOF [10] is used for representing models but our works are only concerned by UML models. For adding semantics in UML models we can use:

- *UML profile*: UML can be used for modeling many domains. The problem with this is that UML models are so generic that it is impossible to know either it is object application, a meta-model, a model, a database structures or anything else only by looking at it [9]. For adding precision, the OMG has standardized the concept of UML profile [11]. A UML profile is a set of techniques and mechanisms allowing to adapt UML in a particular and specific domain. UML profile can be used in any UML model and do not modify the structure of the meta-model. UML profiles are stereotypes or labels which can be injected into models. After having stuck labels on models, we can make inference using it. As we can see, doing this can solve our problem of semantics lack on model in a low level, but this is not exploitable by machines because there is no notion of logic and taxonomy and semantic is not formally defined.
- *Object Constraint Language*: In UML it was not possible to define the body of an operation (or a method) so the OCL [12] was standardized by OMG for doing it. OCL allows expressing any kind of constraints on UML models. For example, we can express constraints like: “before renting a car you must be sure that it is OK”. Well, OCL seems to be a good solution for our problem but it is not the case. The first problem with OCL is that he does not support side effect operations and the second is that he does not offer automatic inference for machines.
- *Action Semantic*: remember that the main constraint with OCL was that he only supports no side effects operations. To solve this constraint, the OMG standardize Action Semantic [13]. Well, now we have a formalism being able to express any kind of operations and constraints but it is not enough. This formalism is complicated to use [9], was not created while thinking to machine comprehension and self-use, and do not have a textual formalism.

As we can see, none of the UML “techniques” is suitable for adding semantics in models. In an other side a new domain of computer is growing more and more: semantic web. The aim of the semantic web is to make the web both comprehensible by humans and machines [14]. A part

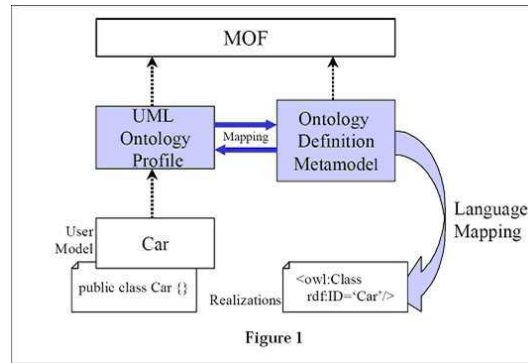


Figure 2: ODM principle

of semantic web is about ontology and reasoning. Modeling concept defined by ontologies can be used to model the concepts in a domain, the relationships between them, and the properties that can be used to describe instances of those concepts [15]. In addition, the Web Ontology Language (OWL)[24] supports the inclusion of certain types of constraint in ontology, allowing new information to be deduced when combining instance data with these logic's description [15]. At this point our dilemma was how can we use MDA models and Semantic Web? Ontology Definition Meta Model was the response to our need.

Moreover, careful selection of a knowledge representation may simplify problem solving, whereas an unfortunate selection may lead to difficulties or even failure to find a solution. Complex problems require a combination of several different representations. In order to be practical, every knowledge representation technique needs a notation. A formal notation often used for representing knowledge is first-order predicate calculus. Of course, it is not suitable for all types of knowledge. The notation of description logics has also proven expressive enough to represent knowledge formally. More recently, the trend has been to use XML syntax for representing knowledge structures.

3 Ontologies and their application

There is an increasing awareness among researchers and developers that ontologies are not just for knowledge-based systems, but for all software systems. The major application fields for ontologies nowadays include knowledge management, e-learning, ecommerce, and integration of Web resources, intranet documents, and databases. Knowledge sharing and reuse through automatic exchange of Web documents among applications and agents is possible only if the documents contain ontologically encoded information. One view of ontology representation languages is that they can be used as metadata languages ontologies engineered with such languages can provide vocabularies for metadata to facilitate the management, discovery, and retrieval of resources on the Web. A metamodel is an explicit model of the constructs and rules needed to build specific models within a domain of interest. This characterizes a valid metamodel as an ontology, since such constructs and rules represent entities in a domain and their relationships i.e., a metamodel is an ontology used by modelers.

The MDA and its four-layer architecture provide a solid basis for defining the metamodels of any modeling language, and thus a language for modeling ontologies based on the MOF. Such a language can employ MDA tool support for modeling, model management, and interoperability with other MOF-based metamodels. The current software tools do not implement many of the fundamental MDA concepts. However, we can expect that most of these tools, which are presently oriented towards UML and the modeling layer (M1), will be improved and equipped with MDA support in the coming years.

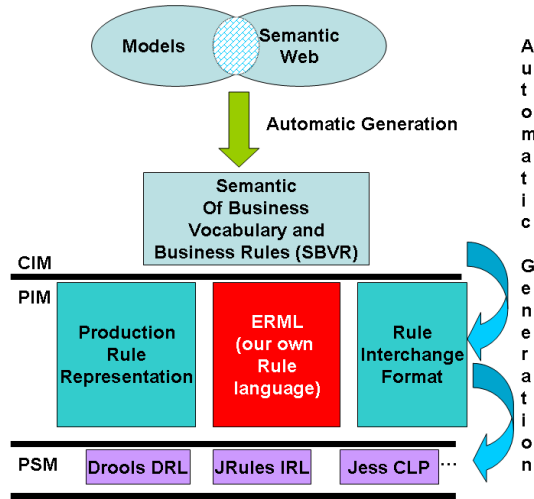


Figure 3: Our approach throughout the MDA layers

4 Ontologies Learning

Ontology learning (ontology extraction, ontology generation, or ontology acquisition) is the automatic or semi-automatic creation of ontologies, including extracting the corresponding domain's terms and the relationships between those concepts from a corpus of natural language text, and encoding them with an ontology language for easy retrieval. As building ontologies manually is extremely labor-intensive and time consuming, there is great motivation to automate the process. Typically, the process starts by extracting terms and concepts or noun phrases from plain text using linguistic processors such as part-of-speech tagging and phrase chunking. Then statistical or symbolic techniques are used to extract relation signatures. ontologies have to be structured optimally against the knowledge base they represent. The structure of an ontology needs to be changed continuously so that it is an accurate representation of the underlying domain. Recently, an automated method was introduced for engineering ontologies in life sciences such as Gene Ontology (GO),[7] one of the most successful and widely used biomedical ontology.[8] Based on information theory, it restructures ontologies so that the levels represent the desired specificity of the concepts. Similar information theoretic approaches have also been used for optimal partition of Gene Ontology.[9] Given the mathematical nature of such engineering algorithms, these optimizations can be automated to produce a principled and scalable architecture to restructure ontologies such as GO.

5 Conclusion

Ontology engineering aims to make explicit the knowledge contained within software applications, and within enterprises and business procedures for a particular domain. Ontology engineering offers a direction towards solving the interoperability problems brought about by semantic obstacles, such as the obstacles related to the definitions of business terms and software classes. However, real-world applications and their development are rarely clearly bordered in their scope; that's why these environments are not enough. It is, therefore, necessary to integrate applications that are used for intelligent systems development into mainstream software platforms. This topic is going to gain more and more attention with the development of the Semantic Web [Berners-Lee et al., 2001] and with increased integration of intelligent techniques in common information systems.

References

- [1] Barbara von Halle. *Business Rules Applied*. John Wiley & Sons, New York, USA, 2002.

- [2] Ronald G. Ross. *Principles of the Business Rule Approach*. Addison-Wesley, Boston, USA, 2003.
- [3] The Object Management Group OMG. *Semantics of Business Vocabulary and Business Rules (SBVR)*. OMG Specification, March 2006.
- [4] W3C. *Rule Interchange Format (RIF)*. W3C Workgroup, 2005.
- [5] The Object Management Group OMG. *Production Rule Representation (PRR) RFP*. OMG Request For Proposal (br/2003-09-03), 2003.
- [6] IBM T.J. Watson Research Center. *CommonRules project*. Intelligent Agents project (1994-97), 1997.
- [7] RuleML. *The RuleML initiative*.
- [8] The Object Management Group OMG. *Model Driven Architecture Guide Version 1.0.1*. OMG Specification, June 2003.
- [9] Xavier Blanc. *MDA en action*. Eyrolles, France, 2005.
- [10] The Object Management Group OMG. *Meta Object Facility (MOF) Specification Version 1.4*. OMG Specification (formal/02-04-03), April 2002.
- [11] The Object Management Group. *Unified Modeling Language: Superstructure*. OMG Specification, February 2004.
- [12] The Object Management Group OMG. *UML 2.0 OCL Specification*. OMG Specification, October 2003.
- [13] The Action Semantics Consortium. *Action semantics for the uml*. OMG Specification (ad/2001-03-01), March 2001.
- [14] Thomas B. Passin. *Explorer's guide to the Semantic Web*. Manning Publications Co, Greenwich, UK, 2004.
- [15] Stephen Cranefield and Jin Pan. *Bridging the Gap Between the Model-Driven Architecture and Ontology Engineering. Proc. of AOSE 2004 Workshop*, 2004.
- [16] Dragan Gaëvič, Dragan Djurić, and Vladan Devedžić. *Model Driven Architecture and Ontology Development*. Springer-Verlag, Berlin, DE, 2006.
- [17] Gruber, Thomas. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal Human-Computer Studies* Vol. 43, Issues 5-6, November 1995..
- [18] The Object Management Group OMG. *Request For Proposal for Ontology Definition Metamodel*. OMG Request For Proposal, March 2003.
- [19] Kenneth Baclawski, Mieczyslaw K. Kokar, Paul A. Kogut, Lewis Hart, Jeffrey Smith, William S. Holmes III, Jerzy Letkowski, and Michael L. Aronson. *Extending UML to Support Ontology Engineering for the Semantic Web. Lecture Notes in Computer Science*, 2185:342+, 2001.
- [20] Dragan Djuric, Dragan Gasevic, and Vladan Devedzic. *Ontology Modeling and MDA. Journal of Object Technology*, 4(1):109–128, 2005.
- [21] The Object Management Group OMG, IBM, and Sandpiper Software. *Ontology Definition Metamodel*. OMG Specification, June 2006.
- [22] H. Knublauch. *Ontology-Driven Software Development in the Context of the Semantic Web: An Example Scenario with Protege/OWL. 1st International Workshop on the Model-Driven Semantic Web (MDSW2004)*, 2004.

- [23] Harold Boley, Said Tabet, and Gerd Wagner. *Design Rationale for RuleML: A Markup Language for Semantic Web Rules*. In *SWWS*, pages 381–401, 2001.
- [24] Stephen J. Mellor, Anthony N. Clark, and Takao Futagami. *Guest Editors' Introduction: Model-Driven Development*. *IEEE Software*, 20(5):14–18, 2003.
- [25] Raphael Volz. *Web Ontology Reasoning with Logic Databases*. PhD thesis, Universität Karlsruhe (TH), Universität Karlsruhe (TH), Institut AIFB, D-76128 Karlsruhe, 2004.
- [26] The Object Management Group OMG. *MOF 2.0/XMI Mapping Specification, v2.1*. OMG Specification (formal/05-09-01), 2005.
- [27] Java Community Process(JCP). *Java Metadata Interface (JMI)*. Sun Java Specification Request (JSR 40), 2002.
- [28] Mouhamed Diouf, Kaninda Musumbu, and Sofian Maabout. *Standard Business Rules Language: why and how? The 2006 International Conference on Artificial Intelligence*, June 2006.