An Ontology-Based Data Mediation Framework for Semantic Environments

Adrian Mocan, University of Innsbruck, Austria
Emilia Cimpian, University of Innsbruck, Austria

ABSTRACT

In a semantic environment data is described by ontologies and ontology mapping has become a crucial aspect in solving the heterogeneity problems of semantically described data. This means that alignments between ontologies have to be created, most probably during design-time, and used in various run-time processes. Such alignments describe a set of mappings between the source and target ontologies, where the mappings show how instance data from one ontology can be expressed in terms of another ontology. In this article we propose a formal model for creation of mappings and we explore how such a model maps onto a design-time graphical tool that can be used in creating alignments between ontologies. In the other direction, we investigate how such a model helps in expressing the mappings in a logical language, based on the semantic relationships identified using the graphical tool.

Keywords: conceptual model; computer-mediated communication; data integration, data semantics; ontologies

INTRODUCTION

Semantic Web technologies have become more and more popular during the last decade. Based on simple and appealing ideas (common understanding of data, common formats for representing these data), Semantic Web aims at providing a framework that would allow information sharing across the Web in a manner understandable not only by humans, but also by the machines.

The agreed upon format for representing this information are the ontologies, but the representation of data using ontologies cannot guarantee the homogeneity and consistency of information. Even if the ontologies are supposed to be a formal explicit specification of a shared conceptualization (Gruber, 1993), they are usually developed in isolation, and shared between well-defined boundaries. This leads to the development of different conceptualizations of the same domain, different ontologies modeling the same aspects but in different manners.

In this context, ontology mapping has become a crucial aspect in solving heterogeneity
problems between semantically described data. The benefits of using ontologies, especially in heterogeneous environments where more than one ontology is used, can only be realized if this process is both correct and efficient. A trend is to provide graphical tools capable of creating alignments during design-time in a (semi-)automatic manner (Ehrig, Staab & Sure, 2005; Mocan & Cimpian, 2005; Noy & Munsen, 2003; Silva & Rocha, 2003). These alignments consist of mapping rules, frequently described as statements in a logical language, and can be executed for performing the actual mediation when needed. One of the main challenges is to fully isolate the domain expert (who is indispensable if 100% accuracy is required) from the burdens of logics using a graphical tool, and in the same time to be able to create complex, complete and correct mappings between the ontologies.

We consider that it is absolutely necessary to formally describe the mappings creation process and to link it with the instruments available in the graphical tool and with a mapping representation formalism that can be used later during run-time. This allows the capturing of the actions performed by the human user in a meaningful way with respect to the visualized ontology structure and then to associate the results of these actions (mappings) with concrete statements in a mapping language (mapping rules).

The article proposes a set of strategies and enhancements to the classical approach towards ontology mapping and run-time mediators. Additionally it proposes a formal model that unifies the conceptual models of the design-time and run-time tools, improving and making more explicit the process of translating the domain expert inputs placed in graphical interface in logical formalism that is to be executed by the run-time tool.

The article structure is as follows: the next section presents the context and motivation for the work. The third section introduces the model we propose expressed using First-Order Logic (Genesereth & Nilson, 1987). The fourth section describes how this model can be applied to WSMO (Feier, Polleres, Roman, Domingue, Stollberg & Fensel, 2005) ontologies, while the fifth section presents the creation of mapping rules; the prototype that implements and applies the proposed formal model is described in the sixth section. Following this related work and conclusions are presented.

CONTEXT AND MOTIVATION

The work described in this article has been carried out in the Web Service Execution Environment (WSMX) working group, whose scope is to build a framework that enables discovery, selection, mediation, invocation and interoperation of Semantic Web Services (Mocan, Moran, Cimpian & Zaremba, 2006). Web Services are semantically described using ontologies, but as they are generally developed in isolation, heterogeneity problems appear between the underlying ontologies. Without resolving these problems the communication (data exchange) between different Semantic Web Services cannot take place.

The data mediation process in WSMX includes two phases: a design-time and a run-time phase. The mismatches between the ontologies are identified at design-time, while the found semantic relationships are expressed as mapping rules; these mapping rules are used at run-time to transform the data passing through the system. The run-time phase can be completely automated, while the design-time phase remains semi-automatic, requiring the inputs of a domain expert.

For the design-time phase a semi-automatic ontology mapping tool was developed that allows the user to create alignments between ontologies and to make these alignments available for the run-time process. There has been much research in the area of graphical mapping tools, for example, Noy and Munsen (2003) and Silva and Rocha (2003), however we believe there are many challenges still to be addressed. In particular, our focus has been on providing the user with proper support (e.g., suggestions and guidance), and in defining strategies that hide the burden of logical languages that are
Related Content

Neural Networks for Content-Based Image Retrieval
[www.igi-global.com/chapter/neural-networks-content-based-image/28930?camid=4v1a](www.igi-global.com/chapter/neural-networks-content-based-image/28930?camid=4v1a)

Study on Ontology Ranking Models Based on the Ensemble Learning
[www.igi-global.com/article/study-on-ontology-ranking-models-based-on-the-ensemble-learning/203696?camid=4v1a](www.igi-global.com/article/study-on-ontology-ranking-models-based-on-the-ensemble-learning/203696?camid=4v1a)

ACRONYM: Context Metrics for Linking People to User-Generated Media Content
[www.igi-global.com/chapter/acronym-context-metrics-linking-people/76177?camid=4v1a](www.igi-global.com/chapter/acronym-context-metrics-linking-people/76177?camid=4v1a)
Context-Aware Presentation of Linked Data on Mobile
[www.igi-global.com/article/context-aware-presentation-of-linked-data-on-mobile/129762?camid=4v1a](www.igi-global.com/article/context-aware-presentation-of-linked-data-on-mobile/129762?camid=4v1a)