Semantic Resolution for E-Commerce

Yun Peng
University of Maryland Baltimore County
Department of CSEE
Baltimore, MD 21250, USA
(410)455-3816
ypeng@umbc.edu

Nenad Ivezic
National Institute of Standards and Technology
100 Bureau Drive
Gaithersburg, MD 20899, USA
(301) 975-3536
nivezic@nist.gov

ABSTRACT

In this position statement, we describe our on-going research project on resolving semantic differences for multi-agent systems in electronic commerce. The approach we are taking can be characterized by the following: 1) the semantic differences between heterogeneous agents are resolved at runtime and in an incremental fashion; 2) concepts and their classes are represented as frames based on daml+oil semantic markup language; and 3) the resolution is viewed as an abductive reasoning process and thus necessarily involves approximation and default reasoning.

Categories and Subject Descriptors

I.2.4. [**Artificial Intelligence**]: Knowledge Representation Formalisms And Methods – *frames and scripts, representation languages, semantic networks.*

General Terms

Design, Standardization, Languages.

Keywords

Ontology, semantic difference, software agents, agent communication languages, electronic commerce.

1. INTRODUCTION

Understanding the meaning of messages exchanged between software agents has long been recognized as one of the key problems to realizing multi-agent systems (MAS). Forcing all agents to use a common vocabulary defined in one or more shared ontologies is an oversimplified solution especially when these agents are designed and deployed independently of each other. This is particularly the case for agent applications in E-Commerce whose business operation environment can be seen as 1) a huge, ***open*** marketplace (accommodating many companies, each capable of entering and leaving the market freely); 2) involving ***dynamic*** relationships (with partnerships formed and dissolved easily and frequently); and 3) containing ***heterogeneous*** representations of individual agents for different enterprises. It is therefore practically infeasible and unwise to restrict all agents to using the same vocabulary (and with the same semantic understanding of the terms in the vocabulary) for exchanging information. Semantic differences should be allowed and be resolved when they arise during agent interaction.

Research work on ontology attempts to provide semantics to information exchanged over the Internet [1,2,3].The most significant recent development in this direction is the joint effort between W3C's Semantic Web Activity and the DARPA Agent Markup Language Project and the resulting daml+oil specifications [1,3]. Although the technologies developed in this effort are primarily aimed at making Web pages understandable by programs, they may serve as a basis (both as language standard and theoretical foundation) for resolving semantic differences between heterogeneous agents. However, additional methodology and mechanisms need to be developed if semantic resolution is to be done dynamically at runtime (through agent interaction). This is the primary objective of our ongoing project on semantic resolution for MAS applications in E-Commerce jointly by The Laboratory for Advanced Information Technology at UMBC and Internet Commerce for Manufacturing at NIST.

2. A SIMPLE E-COMMERCE SCENARIO

Consider the following simple E-Commerce scenario of RFQ (Request For Quote) involving two agents, the buyer A1 and seller A2. Both A1 and A2 share a common ontology ONT-0, which gives semantics of some very basic terms (e.g., those describing business transactions such as RFQ and generic part names such as CPU, memory, etc.). Each agent has its own specialized ontology (i.e., ONT-1 that gives the semantics or data model of products to order for A1, and ONT-2 that gives data model for the product catalog for A2). We assume that terms in ONT-1 and ONT-2 are defined on top of ONT-0, and thus can be reduced to terms in ONT-0.

Suppose A1 sends an RFQ to A2 for a number of "High_Power_Laptop" products, which is a term defined in ONT-1. Before A2 can determine a quote, it needs to understand what A1 means by "High_Power_Laptop" and if a term similar to that is in its catalog. We use phrase "***Semantic Resolution***" to reference the process of identifying the meaning of terms defined in different ontologies and matching them semantically.

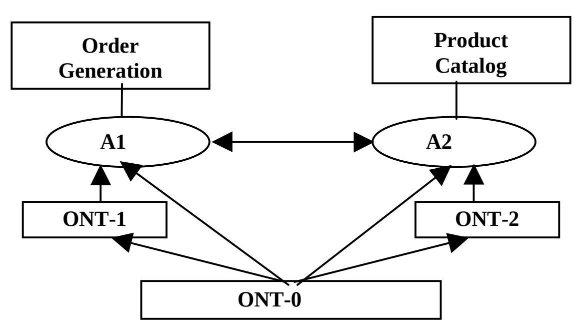

Figure 1. A simple RFQ scenario involving two agents

3. APPROACH

We define the ontologies following daml+oil convention in which each term/concept is defined as a class in a frame-like structure. Each ontology forms an ISA hierarchy according to super and sub-class relations and class-instance relations [1], and is associated with a unique namespace and URI. For example, the term *High_Power_Laptop* in ONT-1 (namespace *ns1*) can be defined as a frame below (omitting the details of daml+oil)

<*"ns1:High_Power_Laptop"*
<*SubClassOf "ns1:Laptop"*>
<*ns0:CPU "ns1:High_End_Pentium"*>
<*ns0:Memory "256M"*>
<*ns0:Disk "SCSI"*> ... >

The semantic resolution process consists of the following steps, each of which involves its own research issues.

Semantic Querying: This step involves obtaining definition of a term from an ontology owned by a different agent. For A2 to understand *High_Power_Laptop* in the example RFQ, it needs to understand all terms used to define this term in ONT-1 (including both slot name and filler name of each slot). The definition of this term can be obtained by querying A1 via some agent communication language (ACL). Since A2 only understands terms in ONT-0 or ONT-2, it would not understand *ns1:Laptop* nor *ns1:High_End_Pentium*. The process proceeds to query the definitions of these terms, and it may propagate to any number of classes in ONT-1, either vertically (for super-classes) or laterally (crossing to other classes). Vertical query propagation is particularly interesting because 1) some properties used to define a class are inherited from its super-classes; and 2) different ontologies may structure their ISA hierarchies differently.

Research issues include 1) How to structure the many ontologies on the Semantic Web so as to guarantee desired behavior for semantic querying? 2) How to control the query proliferation so that the queries will not continue endlessly while guaranteeing relevant semantic information is obtained? and 3) What requirements would this process impose on the existing ACLs?

Semantic Mapping. After "Semantic Querying" step, A2 can build a frame *ns2:High_Power_Laptop*, an image of *ns1:High_Power_Laptop* that is defined by a list of slots/fillers it understands. Then the "Semantic Mapping" step attempts to match this frame with those defined in ONT-2. The mapping, also called "re-classification", can be done by some logic-based reasoning. For example, as suggested in [5], the term *ns2:High_Power_Laptop* matches another term *ns2:X* if *ns2:X* logically follows from the properties of *ns2:High_Power_Laptop*. We may also use some description logic system to reason if *ns2:High_Power_Laptop* can be subsumed by *ns2:X*, taking advantages of DL's efficient procedures for subsumption and classification operations [2]. However, this "naïve approach" [4] often works only when an exact match can be found in the target ontology. It is often the case that two semantically similar terms (from two ontologies) agree on many properties but disagree on others, partly due to the fact that many of these properties are default rather than necessary condition of that class. Moreover, the number of slots in definitions of similar terms may be different, possibly due to different levels of granularity of their respective ISA hierarchies. Therefore, approximate/partial match becomes a necessary part of the reasoning process for semantic mapping. We will investigate the applicability of both qualitative (e.g., rough classification [5]) and quantitative (e.g., Bayesian and Fuzzy Classification) methods for approximate mapping. We will also investigate what is required to incorporate both representation and reasoning with approximate matching in ontology definition languages (e.g., daml+oil).

Hypothesize-and-Test. Semantic resolution, should not be viewed as a pure deductive reasoning, but rather a reasoning very close to *abduction*. Like other types of abductive reasoning, a target term identified during semantic mapping is not a logical consequence but a hypothesis; there may be more than one target terms that match the source term (either with the same or different degree of similarity); and a hypothesis is more plausible if it is more similar to the source term. As an abductive reasoning, the semantic resolution shall be conducted as a cycle of hypothesize-and-test. In the "hypothesize" phase, the agent generates and ranks hypothetical target terms (as described in "Semantic Mapping" step). In the "test" phase, the agent generates queries (as described in the "Semantic Querying" step) to test the plausibility of current hypotheses. The answers to these queries expand the semantic understanding of the source term, and help to differentiate existing hypotheses and possibly lead to the formation of new hypotheses in the next cycle. A significant research topic is possible human involvement when automatic semantic resolution fails. Another significant research issue is the definition of metrics of similarity and processes to determine semantic matches among the similar terms.

4. FINAL REMARKS

We believe that the joint Semantic Web Activity at W3C and the DAML effort at DARPA will have far-reaching influence on the agents R&D and application efforts. However, the agent research community must keep in mind that the full potential of the agent approach will be delivered only if the agents truly are able to leave in open, dynamic, heterogeneous environments. Therefore, the semantic resolution issue needs to be kept at the forefront of the research agenda. Because of the difficult nature of this issue, a solid, sustained research effort will be needed that spans multitude of efforts from software engineering to fundamental research. A compounding issue is that the E-Commerce players are likely to see the semantic agent-based solutions as a distinct business opportunity. Managing these real-life issues in the context of the cutting-edge, high-impact R&D work presents a significant challenge.

5. REFERENCES

[1] DAML home page, http://www.daml.org.

[2] Decker, S. et al, Knowledge representation on the Web. Proceedings of the 2000 International Workshop on Description Logics, Aachen, Germany, 8/2000., 453-469.

[3] Horrocks, I., et al, The Ontology Inference Layer OIL, http://www.ontoknowledge.org/oil/TR/oil.long.html.

[4] Stuckenschmidt, H., Uing OIL for semantic information Integration. Proceedings of the ECAI workshop on Ontologies and PSMs 2000

[5] Stuckenschmidt,, H. and Visser. U., Semantic translation based on approximate re-classification. Proceedings of the Workshop "Semantic Approximation, Granularity and Vagueness, KR'00