Semantic Integration : Position Statement

Michael Grüninger and **Joseph B. Kopena**
Manufacturing Systems Integration Division
National Institute of Standards and Technology
100 Bureau Drive
Gaithersburg, MD 20899-8260
gruning@cme.nist.gov, jkopena@cme.nist.gov

Overview of our Research

Much of the research in semantic integration has reduced the problem to ontology integration – two software applications can be integrated by specifying the semantic mappings between their respective ontologies. However, few applications in practice have explicitly specified ontologies, and even when they do, the ontologies are not fully axiomatized (that is, there exist intended interpretations that are not models of the axioms or there exist unintended models of the axioms). Consequently, ontology mappings are not sufficient to achieve semantic integration.

To address this problem, we adopt what we call the Ontological Stance – we model a software application as if it were an inference system with an axiomatized ontology, and use this ontology to predict the set of sentences that the inference system decides to be satisfiable. This is analogous to the intentional stance, which is the strategy of interpreting the behavior of an entity by treating it as if it were a rational agent who performs activities in accordance with some set of intentional constraints.

In our work, we focus on techniques for achieving the semantic integration of software applications directly by using ontologies as interlingua between the applications themselves. The distinguishing feature of this Interlingua architecture is the existence of a mediating ontology that is independent of the software applications' ontologies and which is used as a neutral interchange ontology. The semantic mappings between application and interlingua ontologies are manually generated and verified prior to interaction time. This process of creating the mapping between the application ontology and the interlingua ontology is identical to the process of creating a mapping directly between two application ontologies. However, it has two advantages. First, we only need to specify one mapping for each application ontology, rather than specifying a mapping for each pair of application ontologies. Second, if the application ontologies and the interlingua ontology are specified using the same logical language, then the translation can be accomplished by applying deduction to the axioms of the interlingua ontology and the ontology mappings. If these mappings have already been verified to preserve semantics between the application and interlingua ontologies, we are guaranteed that translation between the applications also preserves semantics.

The two tools that we are developing are described in our workshop paper. The Twenty Questions tool supports the semiautomatic generation of semantic mappings between the PSL Ontology and the terminology used by a software application. The Process Information Exchange Protocol compares the profiles generated by the Twenty Questions for different software applications to determine which concepts can be either fully or partially shared.

Challenges and Open Problems

What is Semantic Integration?

We still lack a precise characterization of the problem of semantic integration. In some sense, if the ontologies are using the same underlying logical language then the notion of relative interpretation is necessary for semantic integration. However, it is not sufficient – it does not capture all of our intuitions concerning partial translation and it does not distinguish between ontologies for different but overlapping domains.

Testing Semantic Mappings

Once semantic mappings have been proposed between two ontologies or software applications, we still need some methodology for evaluating the correctness and completeness of the mappings so that we can determine whether or not semantic integration has been achieved. If the ontologies are fully axiomatized, then we can provide a model-theoretic evaluation of the semantic mappings (e.g. preservation of models or submodels). However, as we observed above, most software applications do not use fully axiomatized ontologies; the best we can do in these cases is to use an empirical methodology to evaluate the semantic mappings between the terminology of the applications. Adopting the Ontological Stance, we can determine whether inferences performed by the applications are preserved by the mappings.

Implementation of Testbeds

There are several critical issues in semantic integration that can only be solved by empirical approaches. These include the expressiveness/decidability tradeoff for ontology representation languages, the evaluation of different mapping techniques, and determining whether the lack of ontology reuse is due to superficial or deep ontological commitments. We need to establish academic and industrial testbeds that consist of multiple agents and ontologies within different integration architectures, so that participants can carry out experiments to test the critical issues in semantic integration.

Authors

Michael Grüninger

Michael Gruninger is currently a Research Scientist in the Institute for Systems Research at the University of Maryland College Park and also a Guest Researcher at the National Institute for Standards and Technology (NIST). Michael was previously a Senior Research Scientist in the Enterprise Integration Laboratory of the Department of Mechanical and Industrial Engineering at the University of Toronto, where he was the project manager for numerous international projects in collaboration with industry, academia, and government.

His current research focuses on the design and formal characterization of ontologies and their application to problems in manufacturing and enterprise engineering. He is the project leader for the Process Specification Language project at NIST. He is also the project leader for ISO 18629 (Process Specification Language) within the International Standards Organization (ISO).

Joseph Kopena

Joseph Kopena is an undergraduate research assistant in Drexel University's Geometric and Intelligent Computing Laboratory. His research interests include knowledge representation and common sense reasoning in engineering domains and automated system integration through formal ontology. Contact him via joe@plan.mcs.drexel.edu.