Representation of Semantic Mappings

Results from the Breakout Session

Heiner Stuckenschmidt[1], Mike Uschold[2]

[1]Vrije Universiteit Amsterdam
[2]Boeing Phantom Works, Seattle

The aim of this breakout session was to chart the landscape of existing approaches for representing mappings between heterogeneous models, identify common ideas and formulate research questions to be addressed in the future. In the session, the discussion mainly concerned three aspects: The nature of mappings, existing proposals for mappings and open research questions.

1 The Nature of Mappings

When talking about the nature of mappings it quickly turned out that the nature of mappings is not understood well enough to come up with a definition of what a mapping is. Instead, a number of questions about the nature of mappings were formulated that illustrate the design space of mapping approaches.

What do Mappings define ? The first question to be answered about a mapping is the way it establishes a relation between two models. The basic distinction that can be made here is between operational and declarative specifications of relations. Operational approaches describe the procedure of translating elements from one model into the other. An example of an operational specification is the use of re-write rules. Declarative specifications define constraints on the joint interpretation of the two models. This is normally done in terms of logical axioms that contain elements from different models.

What do Mappings preserve ? It is normally assumed that mappings preserve the 'meaning' of of the two models in the sense that the semantic relation between the intended interpretations of connected elements is the one specified in the mapping. A problem with this assumption is that it is virtually impossible to verify this property. Instead, there are a number of verifiable formal properties that mappings can be required to satisfy. Examples of such formal properties are the satisfiability of the overall model, preservation of possible inferences or the preservation of answers to queries. Often such properties can only be stated relative to a given application context, such as a set of queries to be answered or a set of tasks to be solved.

What do Mappings connect ? There are many different kinds of models that we might want to connect with semantic mappings (database schemas, thesauri, ontologies, ...) and even more languages for representing these models. Given this

Dagstuhl Seminar Proceedings 04391
Semantic Interoperability and Integration
http://drops.dagstuhl.de/opus/volltexte/2005/53

variety, it is not clear what mappings should actually connect. The most general approach is to define mappings between arbitrary statements in the languages used in the connected models. This is not always feasible as complexity issues might make it necessary to restrict the kinds of statements that can be mapped to a subset of the language. A general question is also whether it makes sense at all to make mappings language dependent. An alternative is to specify mappings independent of a particular logic – this can be achieved by defining mappings between the signatures of models – or even independent of the representation syntax by mapping between identifiers independent of what kind of element they represent.

How are Mappings organized ? The final question is how mappings are organized. They can either be part of a given model or specified independently. In this case the question is how to distinguish between mappings and other elements in the models. Mappings can be uni- or bidirectional. Further it has to be defined whether a set of mappings is normative or whether it is possible to have different sets of mappings according to different applications, viewpoints or different matchers.

2 Available Tools and Technologies

Another question is about the availability of mapping technology that is available for use today. This question is not only of great practical relevance, but is also necessary to identify open problem and set a research agenda.

2.1 OWL and Relatives

An obvious choice is to first look at current developments in the area of the semantic web, where mapping between semantic models is one of the most pressing problems. The Web Ontology Language (OWL) indeed contains a number of elements that can be used to define semantic relations between elements from different models. It is possible, for example to state that two instances relations, or concepts are the same or different or to define a concept to be more specific or more general than a concept in another model. This can be sufficient for some purposes, but there are also shortcomings of the direct use of OWL constructs for defining mappings. One of the short comings is that OWL constructs are limited to defining abstract relations, but they cannot be used to define transformations on data. This gap will probably be closed by the introduction of the semantic web rule language which is currently under development. Another problem with the way OWL treats mappings in that it takes an 'all or nothing' approach to the import of external axioms. In particular, mapped concepts are only treated a simple names unless the corresponding ontology is imported. Importing an ontology, however, leaves no control over which statements are important and which not. This can be a problem when trying to map models with a high degree of heterogeneity because it is likely that the overall model will be inconsistent.

The recently proposed C-OWL language offers a solution to this problem by introducing explicit mappings between OWL models. The mappings are stored in separate files thus keeping the mapped ontologies independent. A special semantics for mapping rules allows the selective use of external knowledge and even allows to use information from ontologies that are inconsistent with one another. Another shortcoming that is currently not addressed in any of the approaches is the inability to define the strength of mappings.

2.2 Other Approaches

Over the excitement connected with Semantic Web technologies, the wide range of other existing mapping technologies is in danger of being ignored. Nevertheless, there are a number of mapping languages that have been used in application. Many of these languages have their origin in the database community. Examples of such languages are:

- datalog
- F-Logic
- $\mathcal{DLR}$

Besides these database-related technologies there are a number of approaches that use general first order logic axioms or other expressive knowledge representation languages such as KIF or LOOM. We can observe that all of these approaches are more or less subsets of first order logic, however, we have to acknowledge that mappings do not necessarily have to be specified in logic. There are also transformation languages such as XSLT are available for specifying semantic mappings.

3 Open Problems and Research Directions

While languages for representing semantic models have been studied intensively in the past, work on mappings is in a very preliminary stage. Existing proposals for mapping languages have a number of serious limitations. We identified a number of challenging topics that should addressed in future research.

Language Heterogeneity Despite the tendency of standardizing languages for representing semantic models, one of the most pressing problems is still the translation between models encoded in different languages. The challenge is to provide translations with guaranteed formal properties. First results have been achieved in this direction using theory of institutions that describes different logics in a uniform way.

The Nature of Semantic Relations Most existing mapping approach use a very limited set of semantic relations that can hold between elements from different

models. In particular implication and equivalence are frequently used. Many realistic setting, however, demand for richer relations such as inconsistency, effect-cause relations or overlap. Very limited work exists on approaches for measuring the degree of relatedness specified by a mapping. This is in particular important when mappings are created by automatic mapping tools. A very specific problem with respect to semantic relations is the definition of semantic relations between models that describe the domain of interest at different levels of abstraction.

Mappings as first Class Citizen A general observation about the state of the art of mapping representations is that mappings are not yet considered to be first class entities in semantic models. While most approaches agree on elements such as concepts relations and instances, mappings are not yet an agreed element of semantic modelling. Our conclusion is that this attitude has to change. In order to make advancements in mapping technologies, we need specialized tools that support the creation and maintenance of mappings. Important operations on mappings such as reasoning about, retrieving and composing mappings are currently not supported.

A Framework for Comparing Mappings A very concrete research task is to design a common framework for comparing existing mapping approaches. It is probably not possible to find a framework that us sufficiently concrete and still covers all possible approaches. A promising restriction is to only consider declarative mappings. As mentioned above, most of these approaches are related to first order logic, we can hope to find a logical framework that captures most existing approaches.