Towards Ontological Context Mediation for Semantic Web Database Integration: Translating COIN Ontologies Into OWL

Sumit Bhansali, Benjamin N. Grosof, Stuart E. Madnick
Massachusetts Institute of Technology, Sloan School of Management
50 Memorial Drive, Cambridge, MA 02142, USA
{bhansali, bgrosof, smadnick}@ mit.edu
www.mit.edu/~bhansali, ebusiness.mit.edu/bgrosof, web.mit.edu/smadnick/www/home.html

Abstract

The COntext INterchange (COIN) approach to information integration [Goh *et al*, 1999] [Bressan *et* al, 2000] [Firat *et* al, 2002] uses ontological mappings and enables powerful context-sensitive query mediation for semantic integration of knowledge across multiple heterogeneous database sources. Its existing applications include financial analysis for the financial services industry, as well as airfare and car-rental aggregators for the travel industry [Firat, 2003]. COIN's original development preceded the Semantic Web. In this paper, we address how to combine the best of COIN's capabilities with those of standardized Semantic Web ontologies and rules, i.e., OWL [W3C, 2004] and RuleML [RuleML, 2003] [Grosof, 2004]. As a substantial first step, we give (in summary form) a translation of COIN's ontology representation into OWL description logic, thereby enabling COIN to utilize OWL ontologies. We further identify, at a high level, how to use RuleML logic program rules, together with the ontologies, to perform the kinds of reasoning that the overall COIN approach does, including mapping ontologies and mediating queries. Challenges for future work include treatment of equational ontologies, constraint solving, and abduction.

1 COIN

The core of the COntext INterchange (COIN) approach is a context mediator that rewrites queries coming from a user (referred to as a receiver) context into a context-sensitive mediated query that addresses differences in meanings between the receiver and the source(s). The mediation procedure automatically determines the semantic conflicts between the source and receiver contexts and resolves them using appropriate conversion functions. The automatic detection and resolution of conflicts is made possible by the use of the COIN *domain model* that expresses the ontology of the subject matter or the domain of discourse along with the context knowledge and mapping to the data source(s), as depicted in Fig 1.

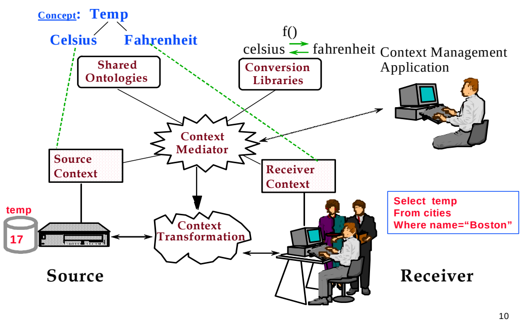

Figure 1. COIN Context-Sensitive Mediation

The COIN domain model consists of semantic types, attributes, modifiers, and elevation axioms. Semantic types are vocabulary terms that are used to capture the semantics of the data stored in the different data sources. Each semantic type may be associated with one or more COIN attributes that describe its structural properties. A semantic type may also have a special attribute called a modifier that can assume different values in different contexts. The semantic types are arranged in a relationship hierarchy using "is-a" links. The "is-a" links define inheritance relationships. If semantic type B "is-a" semantic type A, then B inherits all the attributes and modifiers of A. The overall COIN ontological model also specifies the different source and receiver contexts. Each of the source relations is also defined. The elevation axioms map the source relations to contexts and specify the mapping of the relation columns to semantic types in that context. The semantic types, attributes, modifiers, context and source specifications, and elevation axioms together compose the COIN ontological knowledge base.

2 Example of a simple COIN ontology

For presentation purposes, we use a simple COIN weather ontology as an example. The ontology has three semantic types - Location, PhysicalMeasurements, and Temperature. Temperature has PhysicalMeasurements as its parent semantic type. Location has the attributes city, state, and country. Temperature has the modifier tempUnit that assumes different values in different contexts. There are two contexts - c_usa (for USA context) and c_uk (for UK context). The modifier tempUnit has the string value "celsius" in context c_uk and the value “fahrenheit” in context c_usa. There are two source relations, weatheruk and weatherusa, that are mapped to contexts c_uk and c_usa, respectively. Both the weatheruk and weatherusa relations have 4 columns - city, state, country, and temperature. The elevation axioms map the city, state, and country columns to the semantic type “Location” and the temperature column to the semantic type “Temperature”.

3 COIN Ontological Model in OWL

Next, we describe our new translation of the COIN ontological model into OWL. This translation into OWL can also be viewed as a formulation, or representation, in OWL of the COIN ontological model. Each COIN Semantic Type is translated into a corresponding class in OWL. Each COIN attribute is mapped into a corresponding property in OWL. Each COIN is-a inheritance link between a pair of semantic types is translated into a subClassOf axiom in OWL. The is-a relationship in COIN means that the attributes of A are inherited by B. In the OWL translation, specifying B is a subclass of A entails that A’s properties are inherited by B. In our example ontology, the semantic type “Location” is mapped to the OWL class “Location” and the semantic type “Temperature” is mapped to the OWL class “Temperature”.

Each COIN context is translated into an instance of class COINContext in OWL. For example, the context c_uk (for UK context) is specified using the following syntax -

```
<owl:Class rdf:ID="COINContext"/>
<COINContext rdf:ID="c_uk"/>
```

Modifiers, like attributes, are translated into OWL properties. The domain of these properties is a COIN Semantic Type; however, the range is a value-pair of the type (context, string value). For example, if tempUnit is a modifier for temperature, the value-pair for the modifier is (c_uk, "celsius"). Modifiers may have multiple value-pairs, one pair for each context. For example, another value-pair for tempUnit would be (c_usa, "fahrenheit").

COIN source relations are translated into OWL Classes. The columns of the relations as well as the elevation axioms are translated into OWL properties.

The OWL language features employed in the above translation (of COIN ontologies) are simply the following RDF Schema (RDFS) features:

rdfs:subClassOf, rdfs:domain, rdfs:range

as well as the Class, Individual, and rdf:property. (In addition, the overall COIN data model uses large amounts of property and class instance assertions/axioms to represent the contents of the source databases.) Overall, the above translation/formulation of the COIN ontologies thus only requires the RDFS expressive subset of OWL; this is an expressive subset of OWL Lite and also of Description Logic Programs (DLP) [Grosof *et al*, 2003].

Optionally and less critically, one may add cardinality integrity constraint information to the translated ontology, using OWL’s owl:minCardinality 1 construct (which is in OWL Lite but not DLP) and/or owl:oneOf construct (which is not in OWL Lite).

4 COIN Reasoning, Rules, and RuleML

Though it is useful to represent COIN ontologies in OWL, OWL is not well-suited to do the other aspects of reasoning that are performed in the COIN approach: mapping between different ontologies, mapping between database information expressed in different ontological contexts, formulating mediated queries that combine information across multiple ontological contexts, and answering such mediated queries. For each of these reasoning tasks, logic program rules are the knowledge representation of choice. RuleML is the emerging Semantic Web rules standard, based on logic programs, and can represent the DLP subset of OWL. The COIN approach furthermore uses abductive reasoning with logic programs, along with constraint handling rules, to implement aspects of the mediation. Challenges for future work thus include how to treat abduction and constraint solving, as well as equational ontologies – which are important in financial domains [Firat *et al*, 2002] – in terms of the Semantic Web.

A longer version of this paper, detailing the translation in section 3, is available on the authors’ webpages.

Acknowledgements

This work has been supported, in part, by the MIT-MUST project, the Singapore-MIT Alliance, and a DARPA Agent Markup Language (DAML) award.

References

[Bressan *et al*, 2000] S. Bressan, C. Goh, N. Levina, S. Madnick, A. Shah, M. Siegel, "Context Knowledge Representation and Reasoning in the Context Interchange System", Applied Intelligence: Journal of Artificial Intelligence, Neutral Networks, and Complex Problem-Solving Technologies, 12(2), pp. 165-179, 2000.

[Firat, 2003] A. Firat, "Information Integration using Contextual Knowledge and Ontology Merging," PhD Thesis, MIT, 2003.

[Firat *et al*, 2002] A. Firat, S.E. Madnick, B. Grosof, "Financial Information Integration In the Presence of Equational Ontological Conflicts", Proc. Wksh. on Information Technology and Systems (WITS), pp. 211-216, 2002, .

[Goh *et al*, 1999] C.H. Goh, S. Bressan, S. Madnick, M. Siegel, "Context Interchange: New Features and Formalisms for the Intelligent Integration of Information", ACM Trans. on Information Systems,13(3), pp. 270-293, 1999.

[Grosof *et al*, 2003] B. Grosof, I. Horrocks, R. Volz, and S. Decker, “Description Logic Programs: Combining Logic Programs with Description Logic”. Proc. Intl. Conf. on the World Wide Web (WWW-2003), 2003.

[Grosof, 2004] B. Grosof, “Representing E-Commerce Rules Via Situated Courteous Logic Programs in RuleML”, Electronic Commerce Research and Applications, 3(1), pp. 2-20, Spring 2004.

[RuleML, 2003] Rule Markup Language specification, http://www.ruleml.org. Version V0.8, 2003.

[W3C, 2004] OWL Web Ontology Language Overview. http://www.w3.org/TR/owl-features/, Feb. 2004.