Logical Expressiveness of Semantic Web Languages for Bibliographic Modeling

Karen M. Wickett
wickett2@illinois.edu

Graduate School of Library and Information Science — University of Illinois at Urbana-Champaign

Libraries and the Semantic Web

Semantic Web approaches promise powerful new functionality for combining bibliographic information with knowledge about the world. The possibilities for using ontologies to perform bibliographic inferencing tasks include:

- Automatic classification based on related descriptive information.
- Collection description and management via a framework of relationships between item-level and collection-level metadata.
- Enhanced search and discovery tools that incorporate context in the form of historical or geographical information.

In order to capitalize on this potential, we must understand:

- The expressive requirements for modeling relevant bibliographic information.
- The kinds of knowledge representation possible within Semantic Web languages.

Modeling a Bibliographic Relationship

The Collection/Item Metadata Relationships (CIMR) framework is intended to support inferences from information about a collection to information about individual items in the collection.

For instance, we might want to be able to reason according to following information:

> if a collection has the value z for Owner,
>
> then each item in the collection has the value z for Owner.

CIMR axiom categories are defined in first order logic with formulas which characterize metadata attributes as binary predicates. The example above falls into the CIMR *attribute/value propagation* category and can be represented in first order logic this way:

$$\forall x \forall y \forall z\ [(\mathrm{IsGatheredInto}(x,y)\ \&\ \mathrm{Owner}(y,z)) \supset \mathrm{Owner}(x,z)]$$

Axioms such as these support automated reasoning that can enhance description, browsing and analysis in the bibliographic domain.

The problem

The architects of the CIMR framework suggest using the Web Ontology Language (OWL) to encode categories and inference rules.

Its high expressivity and known problems with decidability may give the impression that the most expressive level of OWL, OWL Full, can represent anything that can be said with first order logic,

However, this is not the case.

Even OWL Full cannot straightfowardly represent an axiom such as
if a collection has some value z for Owner,
then each item in the collection has the value z for Owner.

OWL supports some inferences such as those based on a taxonomic hierarchy but it cannot represent axioms such as the one above, which assert that something has a property in virtue of an arbitrary relationship to another object with that property.

This problem is a known consequence of restrictions that are designed to ensure the decidability of the language.

One solution

In order to incorporate propagation rules into a Semantic Web system for bibliographic services, one may use a formalism designed around expressing rules, such as the Semantic Web Rule Language (SWRL).

The attribute/value propagation rule for Owner expressed in XML Concrete Syntax for SWRL [5]:

```
<ruleml:imp>
  <ruleml:_body>
    <swrlx:individualPropertyAtom swrlx:property="&imlsdcc;isGatheredInto">
      <ruleml:var>x</ruleml:var>
      <ruleml:var>y</ruleml:var>
    </swrlx:individualPropertyAtom>
    <swrlx:individualPropertyAtom swrlx:property="&dccap;owner">
      <ruleml:var>y</ruleml:var>
      <ruleml:var>z</ruleml:var>
    </swrlx:individualPropertyAtom>
  </ruleml:_body>
  <ruleml:_head>
    <swrlx:individualPropertyAtom swrlx:property="&dccap;owner">
      <ruleml:var>x</ruleml:var>
      <ruleml:var>z</ruleml:var>
    </swrlx:individualPropertyAtom>
  </ruleml:_head>
</ruleml:imp>
```

(Rule languages operate on a basis drawn from logic programming. The consequent of a conditional rule is called the “head” of the rule and the antecedent is called the “body”.)

Other solutions may come from extensions within the proposed OWL 2 specification, which has additional support for axioms of this kind. See also Seidenberg and Rector (2006).

In summary, the ontology languages RDFS and OWL are not sufficiently expressive to deliver the full range of inferences anticipated for bibliographic applications. The declaration of additional inference rules will be required.

References

[1] K. Coyle. Meaning, Technology and the Semantic Web. *Journal of Academic Librarianship*, 34(3):263-264, 2008.

[2] S. Gradmann. rdfs:frbr -- Towards an Implemenation Model for Library Catalogs Using Semantic Web Technology. *Cataloging & Classification Quarterly*, 39(3-4):63-75, 2005.

[3] A. Renear et al. Collection/Item Metadata Relationships. In *Proceedings of the International Conference on Dublin Core and Metadata Applications*, 2008.

[4] I. Horrocks et al. From SHIQ and RDF to OWL: the making of a Web Ontology Language. *Web Semantics: Science, Services and Agents on the World Wide Web*, 1(1):7-26, 2003.

[5] http://www.w3.org/Submission/SWRL/

[6] J. Seidenberg and A. Rector. Representing Transitive Propagation in OWL. *Lecture Notes in Computer Science*, 4215:255-266, 2006.

Acknowledgements

Partially supported by a 2007 IMLS NLG grant hosted by the GSLIS Center for Informatics Research in Science and Scholarship (CIRSS).

I am very grateful for comments from Allen Renear and Dave Dubin.