Contextualized RDF Importing

Jie Bao, Li Ding, and Deborah L. McGuinness

Tetherless World Constellation
Rensselaer Polytechnic Institute
Troy, NY, 12180-3590, USA
{baojie,dingl, dlm}@cs.rpi.edu

Abstract. RDF is a key enabling technology for linked data. However, currently RDF lacks a mechanism to connect data from different documents as well to address the contextual differences in these documents. We propose to introduce `rdf:imports` for context-aware integration of RDF documents.

1 Introduction

While RDF has been used as a major supportive technology for linked data, RDF per se does not support data linking and reuse. An RDF document may refer another RDF document by prefix declaration and the use of external resources. However, RDF does not specify how knowledge in the external document or resource should be used in the referring document. For example, when an RDF document uses `foaf:Person`, it is not required that the knowledge in the FOAF (Friend of a Friend) ontology[1] should be used, e.g., that `foaf:Person` is a subclass of `foaf:Agent`. Whether knowledge related to external resources is used is up to the applications that use the referring RDF document. This may lead to the risk of *losing semantics* associated with a vocabulary thus resulting in unintended use of the vocabulary.

OWL [9] introduced `owl:imports` for transferring knowledge from one OWL document to another OWL document. However, `owl:imports` follows a "copy + paste" approach such that axioms in all documents in the importing transitive closure are taken together as a union, and the referring document is only interpreted in that union. This leads to the issue of *loss of context*. For example, if a document imports bussiness reports from two companies published on the Semantic Web (e.g., in the XBRL format and then translated into the OWL/RDF format), simply putting their data together in a union regardless of their contexts may lead to misuse, e.g., "Q1 2010" of one company may correspond to "Q4 2009" of the other company due to their accounting differences, thus data of "Q1 2010" from the two companies actually has different contextual meanings.

For another example (adapted from [4]), the equivalence relationship represented by `owl:sameAs` is often context-dependent, and is accurate only in the

[1] http://www.foaf-project.org/

context of one application. Li Ding has two FOAF profiles, one at Stanford University and one at RPI. The more recent RPI FOAF profile indicates that he holds a job title of "Research Scientist". However, if we connect the two URIs using `owl:sameAs`, an OWL reasoner can infer that Li Ding holds the position of "Research Scientist" at Stanford University, which has never been the case.

C-OWL [2] aims to address the context-insensitive issue of OWL. However, it is limited to creating vocabulary mappings between ontologies, thus can't provide a general solution to knowledge reuse in RDF.

In this paper we propose to introduce `rdf:imports` and several other context modeling predicates to enable contextualized knowledge reuse in RDF. This proposal is based on our previous work on a general context modeling framework for Semantic Web [1], which is in turn an extension of the McCarthy-style context modeling in AI and Semantic Web [8, 5]. We believe the proposal may bring several benefits include the following:

- It enables knowledge reuse in RDF documents without suffering the loss of contexts as `owl:imports` would cause;
- It allows us to use context relation statements to selectively control the transfer and non-transfer of knowledge between contexts, thus can avoid many "out of context" misuse of knowledge;
- It is compatible with existing RDF semantics as an extension of the current specification.

We assume the reader's familiarity of the RDF syntax and semantics [6, 7].

2 RDF Importing and Contexts

An RDF document *may* use the predicate `rdf:imports` to reuse knowledge in other RDF document. Different from `owl:imports`, the RDF importing declaration does not always transfer knowledge from the importee document to the importer document. Rather, it creates a *citation* relation between the two documents.

Optionally, contextual relations may be declared between the two documents to enable selective knowledge transfer. An RDF document, or a part of an RDF document (e.g., a named graph [3] or even a single triple), *may* has an optional declaration of one or multiple associated contexts. The context is identified by an IRI which *may* (by default) corresponds to the IRI of the document or the named graph. The RDF document, or a context document dereferenced at the context's own IRI (if it is different from the RDF document's IRI), may declare relations between this context and other contexts.

Some commonly useful relations between contexts include (but not limited to) the follows. We assume that C1 and C2 are two contexts, R1 and R2 are two RDF documents that are associated to C1 and C2, respectively, and that R2 imports R1.

- Compatibility: If C1 is compatible with C2, then statements in R1 can be interpreted in R2 (while they are not necessarily true in R2). For instance,

C1 may be based on the Open World Assumption, while C2 is based on the Closed World Assumption (therefore `rdfs:domain` and `rdfs:range` have different meanings in C1 and C2)[2].
- Incompatibility: If C1 is incompatible with C2, then R1 can not be merged with R2. This relation is useful to prevent misuse of ontology out of context. For example, we may declare that Li Ding's two FOAF profiles are using two incompatible contexts.
- Extension: If C2 extends C1, then statements in R1 are also true in R2[3]. For example, the context of U.S. criminal laws extends the context of U.S. Constitution.
- Embedding ("part of"): If C1 is embedded in C2, then R2 "quotes" R1. Embedding not necessarily mean transfer of knowledge. For example, a company's Q1 2010 report context may be part of its 2010 annual report context.

3 Semantics

The semantics of RDF importing extends the model-theoretic RDF Semantics [6] in the following aspects:

- Every context C has its own local domain of universe of I_C. All RDF documents that share the same context also share the same domain of universe in their interpretations. On the other hand, RDF documents that are in different contexts will not share the same domain of universe.
- RDF(S) vocabulary and axioms are interpreted locally in the associated contexts. For instance, the `rdf:type` relation in a context C will be satisfied by the semantic conditions:

  x is in ICEXT_C(y) if and only if <x,y> is in $\mathrm{IEXT}_C(\mathrm{I}_C$(`rdf:type`))

  where subscript $_C$ indicates that the corresponding mappings are only in the domain of universe I_C.
- Relations between contexts establish semantic conditions between contexts. For instance, if C2 extends C1, then if a semantic condition holds in I_{C1}, it must also hold in I_{C2}. If C2 is incompatible with C1, then all semantics conditions in I_{C1} are ignored in I_{C2}.
- If R2 (with context C2) imports R1 (with context C1), then whether semantic conditions held in I_{C1} are required in I_{C2} depends on the relations between C1 and C2.

In particular, a blank node associated with a context is interpreted as an unnamed object in the local domain of the context. For instance,

```
_:x rdf:label "Turkey"
```

[2] This is a well known problem that RDF axioms are often "mistakenly" interpreted as integrity constraints, cf. [10].

[3] Using this notion, we can regard that by using `owl:imports`, we have the assumption that every context is an extension of every other context.

in the context of countries will be interpreted as an object in the domain of countries, and even if it is used together with an ontology in the context of animals, `_:x` won't be interpreted in the domain of animals.

When there is no context declaration for an RDF document, a default universal context is used which has a universal domain of universe. The proposed semantics will be reduced to the usual RDF semantics for this case.

4 Conclusions

In this position paper, we propose to introduce a new RDF predicate, `rdf:imports`, to enable contextualized knowledge reuse among RDF documents as needed in many applications, e.g., linked data. We give, at a high level, an outline of the syntax and semantic framework of the proposal, and show that it is useful for several representative examples.

We purposely leave out the design of a concrete syntax for context representation. We also do not discuss properties of RDF contexts and importing relations, e.g., provenances. These are left as future work.

References

1. J. Bao, J. Tao, and D. L. McGuinness. Context representation for the semantic web. In *Web Science Conference*, 2010.
2. P. Bouquet, F. Giunchiglia, F. van Harmelen, L. Serafini, and H. Stuckenschmidt. C-OWL: Contextualizing ontologies. In D. Fensel, K. P. Sycara, and J. Mylopoulos, editors, *International Semantic Web Conference*, volume 2870 of *Lecture Notes in Computer Science*, pages 164–179. Springer, 2003.
3. J. J. Carroll, C. Bizer, P. J. Hayes, and P. Stickler. Named graphs, provenance and trust. In *WWW*, pages 613–622, 2005.
4. L. Ding, J. Shinavier, T. Finin, and D. McGuinness. An empirical study of owl:sameas use in linked data. In *Web Science Conference*, 2010.
5. R. V. Guha, R. McCool, and R. Fikes. Contexts for the semantic web. In S. A. McIlraith, D. Plexousakis, and F. van Harmelen, editors, *International Semantic Web Conference*, volume 3298 of *Lecture Notes in Computer Science*, pages 32–46. Springer, 2004.
6. P. Hayes, editor. *RDF Semantics*. W3C Recommendation. World Wide Web Consortium, February 2004.
7. G. Klyne and J. J. Carroll. Resource description framework (RDF): Concepts and abstract syntax. World Wide Web Consortium, Recommendation REC-rdf-concepts-20040210, February 2004.
8. J. McCarthy. Notes on formalizing context. In *IJCAI*, pages 555–562, 1993.
9. P. Patel-Schneider, P.Hayes, and I. Horrocks. Web Ontlogy Language (OWL) Abstract Syntax and Semantics. http://www.w3.org/TR/owl-semantics/, February 2004.
10. J. Tao, E. Sirin, J. Bao, and D. L. McGuinness. Integrity Constraints in OWL. In *AAAI*, page In Press, 2010.