Meaning and the Semantic Web

Bijan Parsia
University of Maryland
College Park, MD, USA
bparsia@isr.umd.edu

Peter F. Patel-Schneider
Bell Labs Research
Murray Hill, NJ, USA
pfps@research.bell-labs.com

ABSTRACT

The meaning of names (URI references) is a contentious issue in the Semantic Web. Numerous proposals have been given for how to provide meaning for names in the Semantic Web, ranging from a strict localized model-theoretic semantics to proposals for a unified single meaning. We argue that a slight expansion of the standard model-theoretic semantics for names is sufficient for the present, and can easily be augmented where necessary to allow communities of interest to strengthen this spartan theory of meaning.

Categories and Subject Descriptors: I.2.4 [Knowledge Representation Formalisms and Methods]: Representation languages
General Terms: Languages, Theory
Keywords: Semantic Web, representation, meaning

1. INTRODUCTION

The Semantic Web [1] is an extension of the World Wide Web. The major philosophical difference between the Semantic Web and the World Wide Web is that the Semantic Web is supposed to provide machine accessible meaning for its constructs whereas in the World Wide Web this meaning is provided by external mechanisms. This meaning is largely based on the meaning of names which, in the Semantic Web, are URI references [4].

The initial view of the meaning of names in the Semantic Web was that the meaning of a name was determined by the owner of the name, if there is an owner. For names that use schemes based on authorities, such as the http scheme, this owner can be easily discovered by stripping off the fragment identifier, if any, and using the standard World Wide Web mechanisms to determine the owner of the resulting URI. In this view, good practice requires the URI's owner to supply documents, accessible from the corresponding URI, which more or less express a definition of that URI. That definition is determinative, at least, in that a third party which discovered that definition through normal Web mechanisms and made use of it in reasoning with documents using that URI has, *ceteris paribus*, exercised due diligence with respect to the URI owner's definitorial authority. In this view, it is unclear whether publishing documents at the relevant URI *defines* that URI, or simply provides a respectable default for random Web agents. This view of meaning led to Section 4.3, on the authoritative definition of terms, of the 23 January 2003 version of *Resource Description Framework (RDF): Concepts and Abstract Syntax* [5]. (It is still enshrined in section 3.4, on authoritative representation metadata, of the *Architecture of the World Wide Web* [3].)

During the last call period of the above document, considerable pressure was applied against this view of meaning, mostly because it does not allow for any divergence of meaning. As a result, the current version of the document [4] does not have a section on the meaning of RDF, leaving RDF [6] with only the sparse meaning provided by its model-theoretic semantics [2]. In this account, the meaning of a name in RDF is relative to a particular RDF graph (which roughly corresponds to an RDF/XML document in the World Wide Web). Furthermore, the relativization does not strongly constrain the possible consistent interpretations of the name (and, thus, of the graph). However, there is still a need to provide a stronger meaning for names than that provided by this model-theoretic semantics, in particular, to allow for or require the use of the meaning available from other documents.

WWW2004, May 17–22, 2004, New York, New York, USA.
ACM 1-58113-912-8/04/0005.

2. THE PROPOSAL

We propose that this stronger meaning need only be defined in a local sense, as opposed to the global sense above. The meaning of an occurrence of a URI is thus determined by its context, which we take to mean the document in which it appears, plus other documents explicitly mentioned in constructs like the OWL importing mechanism.[1] For determining the meaning of a collection of documents, we propose to use only that meaning determined by the formal language specifications of the Semantic Web, currently the RDF model theory [2] and the OWL model theory [7].

Our proposal allows for divergences of meaning between different documents. A document that does not explicitly import a well-known ontology document, or, indeed any commonly-used document, can easily diverge from any portion, or indeed all, of the common meaning of any name. For example, a document could ignore the common meaning of an invoice and instead use one that has the consequence that the seller owes the buyer money.

One might think that our account of meaning thus results in complete anarchy in the Semantic Web. Even if so, we believe we have embraced only those portions of anarchy that are necessary to prevent totalitarianism, for any proposal for Semantic Web meaning that cuts off easy access to disagreements will inevitably end up stultifying the Semantic Web.

Of course we really do not have a *solution* to handling disagreement in the Semantic Web. Our proposal just makes disagreement about the meaning of terms possible. A full solution to disagreement requires much more formal machinery than we feel is appropriate at this juncture in the development of the Semantic Web. Further, no strictly formal means will be completely adequate to handle disagreement, short of a full solution to the AI problem, as determining which one of a collection of contradictory claims to believe inevitably brings in matters of trust and judgement. Our

[1] We would like to be able to include portions of documents, not just whole documents. However, there are currently no mechanisms for indicating portions of Semantic Web documents written in RDF or OWL, so this is not currently possible.

point is that it is necessary to *allow* unqualified disagreement even, or especially, at this stage of the Semantic Web.

Our proposal does, however, allow for consensus to be achieved, and in an easy fashion. All that is required to achieve consensus concerning meaning is to have the same background theory (and same external intuitions, but this is outside the Semantic Web) and our proposal makes it easy to build such consensus by placing a representation of the consensus in commonly-used documents.

Communities of interest that want to mandate a shared meaning can require the use of such consensus documents. These communities would have, of course, cut themselves off from potentially valuable dissent, but there are many cases, including electronic commerce, where some common meaning is useful or even required for progress, particularly with our poor understanding of how to build truly cognitive software systems. The need in our scheme for explicit importing of these consensus documents provides signals that a document adheres to this meaning, and these signals can be read both within and without the community.

Our proposal also makes it easy to determine (most of) at least the formal part of consensus meaning that is required by a document. The explicitly imported documents in a document (and their imported documents, and so on) provide an excellent indication of just which consensus meanings a document uses. Further the Semantic Web meaning of these consensus documents is just their formal meaning, which is easy to determine.

We freely admit that this notion of Semantic Web meaning is insufficient to capture the entirely of the meaning intended by document writers, and likewise insufficient to capture the entirely of the meaning which the behaviour of many effective software agents will act upon. There is nothing, however, in our proposal that prevents software systems from augmenting, or even replacing, the Semantic Web meaning with their own notions of meaning. Semantic Web meaning only serves as a core, common meaning for Semantic Web documents, to be used or abused as desired. As the Semantic Web evolves, some of these augmented notions of the meaning of the document may become common enough, and well understood enough, to augment or replace the core, standardized meaning. But this is an juncture where we feel strongly that (further) standardization should follow (future) practice.

To allow for software systems of differing sophistication and differing needs we augment our basic proposal above to allow Semantic Web meaning to be contingent on a set of Semantic Web languages that the system understands. Semantic Web meaning for software systems that do not implement OWL entailment would then be less powerful than Semantic Web meaning for software systems that do. There are mechanisms in the World Wide Web that can be used to support their variations on Semantic Web meaning. We would augment the importing mechanism to use content negotiation, allowing OWL-aware systems to request the OWL version of a resource identified by a URI and RDF-aware systems to request the RDF version of the same resource.

This refinement does require considerable care on the part of designers of documents, so that, for example, OWL and RDF documents at the same resource have compatible meanings. However, this is not really different in spirit from issues having to do with the relationship between JPEG and GIF documents at the same URI.

One advantage is that this allows for growth in the Semantic Web. New languages, such as a rules extension for OWL, can be added to the Semantic Web and retrofitted to previous ontologies, augmenting their Semantic Web meaning for software systems that can process the new language while still retaining the old behaviour for existing systems. It is even possible to have Semantic Web languages that are not compatible with OWL or even RDF.

This view of meaning in the Semantic Web is certainly not the one what we would like to have for all time. When the Semantic Web becomes more widespread, when information in it becomes more sophisticated, and when more powerful software systems for the Semantic Web become available our simple version of Semantic Web meaning will be inadequate. For example, it would be useful to reason about contradictory information in different documents. For example, it would be useful to reason *within the Semantic Web* that some buyer and seller do have mutually contractory views. This can be done in certain kinds of modal logics, which would thus be useful as the foundations of a more powerful theory of meaning in the Semantic Web.

In this context it would also be useful to be able to reason, again within the theory of meaning of the Semantic Web, about the rights and obligations of agents within the Semantic Web. This sort of reasoning can perhaps be supported by a theory of meaning that includes these sorts of concepts.

It would also be useful to import only specific portions of documents. A future extension of Semantic Web languages that allow portions of documents to be identified would permit this more fine grained version of importing.

Finally, it does seem a bit odd, if somewhat harmless, that the OWL importing mechanism is an *OWL* importing mechanism, rather than an RDF one. It seems harmless as this is not a large extension to RDF and, if it proves popular, may become a *de facto* standard extension to RDF, eventually to be incorporated into RDF itself. (This possible course of events echoes the migration of the DAML+OIL collections constructs into RDF.) While the need for richer imports mechanisms becomes more acute for higher layers of the Semantic Web language stack, such as rules, it seems unlikely that any would conflict in a systematic way with OWL's imports construct. So we get exactly what we want: *some* mechanism that we can use *now*, that is plausibly forward compatible with future mechanisms.

3. CONCLUSION

We have argued that meaning in the Semantic Web can be usefully determined using only a slight expansion the formal meaning provided by RDF and that this formal account need not require a common, universal notion of meaning in the Semantic Web. The only sharing mechanism needed, at least for now, is an explicit importation mechanism, similar to that provided in OWL. In this way information can be shared as appropriate, thus preventing total anarchy, but need not be, thus allowing for differences of opinion, which are needed to prevent totalitarianism and its resultant stultification and ossification.

4. REFERENCES

[1] T. Berners-Lee, J. Hendler, and O. Lassila. The semantic web. *Scientific American*, May 2001.

[2] P. Hayes. RDF model theory. W3C Recommendation, 10 Feb. 2004. http://www.w3.org/TR/rdf-mt.

[3] I. Jacobs. Architecture of the world wide web, first edition. W3C Working Draft, 9 December 2003. http://www.w3.org/TR/2003/WD-webarch-20031209.

[4] G. Klyne and J. J. Carroll. Resource description framework (RDF): Concepts and abstract syntax. W3C Rec., 10 Feb. 2004. http://www.w3.org/TR/rdf-concepts.

[5] G. Klyne and J. J. Carroll. Resource description framework (rdf): Concepts and abstract syntax. W3C Recommendation, 10 Feb. 2004. http://www.w3.org/TR/2003/WD-rdf-concepts-20030123.

[6] F. Manola and E. Miller. RDF primer. W3C Recommendation, 10 Feb. 2004. http://www.w3.org/TR/rdf-primer.

[7] P. F. Patel-Schneider, P. Hayes, and I. Horrocks. OWL web ontology language semantics and abstract syntax. W3C Recommendation, 10 Feb. 2004. http://www.w3.org/TR/owl-semantics.