Semantic Web Architecture: Stack or Two Towers?

Ian Horrocks[1], Bijan Parsia[2], Peter Patel-Schneider[3], and James Hendler[2]

[1] School of Computer Science
University of Manchester, UK
www.cs.man.ac.uk/~horrocks/
[2] Maryland Information and Network Dynamics Laboratory
University of Maryland
www.mindswap.org/~bparsia/
[3] Bell Labs Research
Murray Hill, NJ, USA
www.bell-labs.com/user/pfps/

Abstract. We discuss language architecture for the Semantic Web, and in particular different proposals for extending this architecture with a rules component. We argue that an architecture that maximises compatibility with existing languages, in particular RDF and OWL, will benefit the development of the Semantic Web, and still allow for forms of closed world assumption and negation as failure.

Up until recent times it has been widely accepted that the architecture the Semantic Web will be based on a hierarchy of languages, each language both exploiting the features and extending the capabilities of the layers below. This has been famously illustrated in Tim Berners-Lee's "Semantic Web Stack" diagram [3] (see Figure 1).

As a result of the work of the W3C Web Ontology Working Group, the "Ontology" layer has now been instantiated with the Web Ontology Language OWL [2]. Since then, attention has turned to the rules layer, and much effort has been devoted to the design of suitable rules languages. Perhaps influenced by some of this work, recently seen versions of the Semantic Web Stack diagram have illustrated a weakened version of the layering idea, with rules and ontologies (OWL) sitting side by side on top of a layer labelled as the "DLP bit of OWL/Rules" [4] (see Figure 2).

Unfortunately, this modified stack is based on some fundamental misconceptions about the semantic relationships between the various languages. In particular, the modified stack suggests that DLP [7] can be layered on top of RDFS and form a common basis for parallel Rules (presumably intended as Datalog/Logic Programming style rules) and OWL layers. This suggestion is, however, based on incorrect assumptions about the semantics of DLP. In particular, if we want Datalog style closed world semantics for Rules (in order to support Negation as Failure), as is argued by some proponents, then the resulting rules language is only a *syntactic* extension of DLP, and is *not* semantically compatible with DLP—in fact DLP is a subset of Horn rules and has standard First Order semantics.

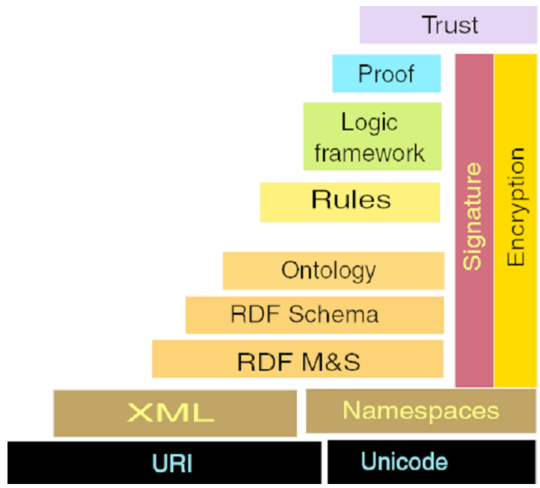

Fig. 1. Semantic Web Stack

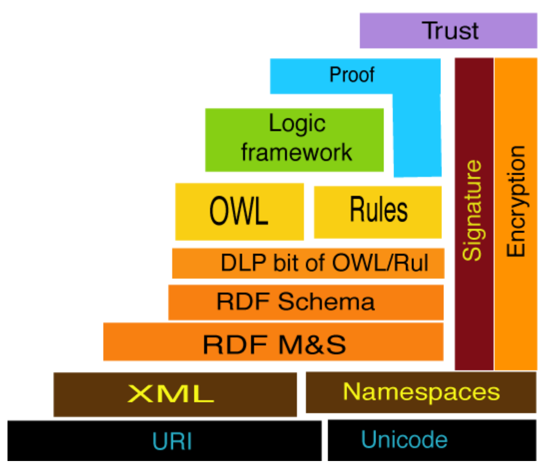

Fig. 2. Latest version of the Semantic Web Stack

Of course it is possible to treat DLP rules as having Datalog semantics (i.e., semantics based on a closed world assumption and Herbrand models [6]). In this case, however, DLP is no longer semantically compatible with OWL and so cannot be situated below OWL in the stack. In fact, when given such a semantics, DLP (and rules languages that extend DLP) are not even semantically compatible with RDF [9]. This is easy to see if we imagine querying an RDF ontology with a more expressive query language, for example one that includes counting or negation (as, for example, SQL). Given an ontology containing only a single RDF triple:

$$\langle\#\text{pat}\rangle\langle\#\text{knows}\rangle\langle\#\text{jo}\rangle.$$

the answer to a query asking if pat knows exactly one person would be "no" under RDF's open world semantics, but "yes" under the closed world semantics of Datalog.

It is thus more appropriate to view DLP with Datalog semantics as being layered directly on top of the XML layer. Datalog rules, and various extensions such as negation as failure (NAF) would then naturally layer on top of (this version of) DLP. Similarly, OWL and other First Order extensions (such as FOL or SCL [10]) would naturally layer on top of RDFS.[4] It has been suggested that the two different semantics (Datalog and First Order) could be unified in some overarching "Logic Framework", although it is an open research problem as to how this could be done.

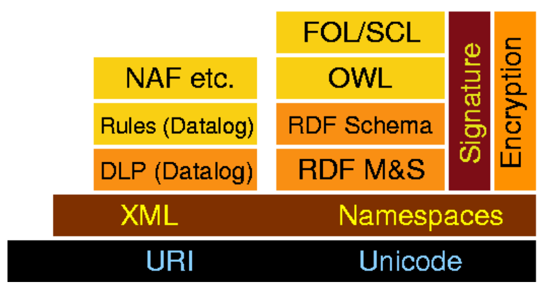

Fig. 3. Semantic Web Stack with Datalog Rules

This more precise analysis of the semantic relationships between the various languages demonstrates that the Datalog view of DLP and of rules actually leads to a stack like the one illustrated in Figure 3, where the Datalog languages and First Order languages are in two separate towers. The Proof and Trust layers have been omitted, as these are currently rather speculative, as has the overarching "Logic Framework", given that, as mentioned above, there is currently no suggestion as to what kind of logic might instantiate this layer.

An alternative view of DLP is as a subset of First Order Horn clauses (as proposed in [7]). In this case DLP can be seen simply as a subset of OWL (although more useful

[4] There is an issue with the meta-level features of RDFS, which has been resolved in OWL by having one language "species" that layers on top of the First Order subset of RDFS (i.e., OWL DL) and another language species that layers on top of the whole of RDFS (i.e., OWL Full).

lightweight OWL subsets could be imagined, e.g., based on the $\mathcal{EL}$ description logic, which covers many important use cases, and for which all key inference problems can be solved in polynomial time [1]). A First Order rules language such as SWRL can then be layered on top of OWL. More expressive languages such as full First Order Logic (First Order Predicate Calculus) would layer naturally on top of SWRL [11].

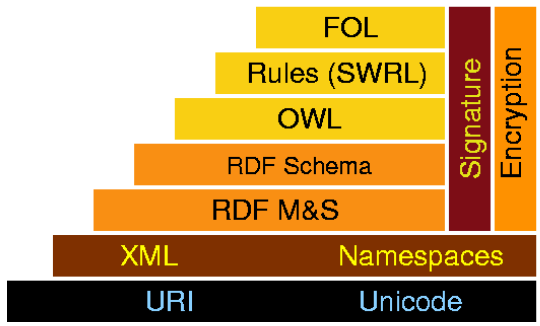

Fig. 4. Semantic Web Stack with First Order Rules

The resulting stack is illustrated in Figure 4 (the DLP/lightweight OWL subset layer has been omitted, but could be inserted between RDFS and OWL). This language architecture has many attractive features when compared to the one illustrated in Figure 3. On the one hand, rules in this framework extend existing work on both RDFS and OWL, as well as providing a foundation for further extensions within a coherent semantic framework. Features such as closed world assumption and negation as failure (NAF) can be supported by powerful query languages—queries already have a closed world flavour (because distinguished variables can only bind to named individuals), and it is natural to extend this with NAF by way of query subtraction (e.g., the answer to the query "faculty(?x) and NAF professor(?x)" can be computed by subtracting the answer to the query "professor(?x)" from the answer to the query "faculty(?x)"). These features are already supported in query languages such as SPARQL [14] and nRQL [8] (the query language implemented in the Racer system). Moreover, recent work on integrating rules with OWL suggests that future versions of this framework could include, e.g., a decidable subset of SWRL, and a principled integration of OWL and Answer Set Programming [5, 12, 13].

On the other hand, adopting Datalog rules (and DLP with Datalog semantics) would effectively establish *two* Semantic Webs, with little or no semantic interoperability between the rules based Semantic Web and the ontology based Semantic Web, even at the RDF level. These two versions of the Semantic Web would inevitably be in competition with each other, and this would make the Semantic Web much less appealing: new users would be presented with a difficult choice as to which part to choose, and in choosing would sacrifice semantic interoperability with the other part.

References

1. F. Baader, S. Brandt, and C. Lutz. Pushing the $\mathcal{EL}$ envelope. In *Proc. of the 19th Int. Joint Conf. on Artificial Intelligence (IJCAI 2005)*, 2005. To appear.
2. Sean Bechhofer, Frank van Harmelen, Jim Hendler, Ian Horrocks, Deborah L. McGuinness, Peter F. Patel-Schneider, and Lynn Andrea Stein. OWL web ontology language reference. W3C Recommendation, 10 February 2004. Available at `http://www.w3.org/TR/owl-ref/`.
3. Tim Berners-Lee. WWW past & future, 2003. Available at `http://www.w3.org/2003/Talks/0922-rsoc-tbl/`.
4. Tim Berners-Lee. Web for real people, 2005. Available at `http://www.w3.org/2005/Talks/0511-keynote-tbl/`.
5. Thomas Eiter, Thomas Lukasiewicz, Roman Schindlauer, and Hans Tompits. Combining answer set programming with description logics for the semantic web. In *Proc. of the 9th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR 2004)*, pages 141–151. Morgan Kaufmann, Los Altos, 2004.
6. Michael Gelfond and Vladimir Lifschitz. The stable model semantics for logic programming. In *Proceedings of the Fifth International Conference on Logic Programming*, pages 1070–1080. MIT Press, 1988.
7. Benjamin N. Grosof, Ian Horrocks, Raphael Volz, and Stefan Decker. Description logic programs: Combining logic programs with description logic. In *Proc. of the Twelfth International World Wide Web Conference (WWW 2003)*, pages 48–57. ACM, 2003.
8. Volker Haarslev, Ralf Möller, and Michael Wessel. Querying the Semantic Web with Racer + nRQL. In *Proc. of the KI-2004 Intl. Workshop on Applications of Description Logics (ADL'04)*, 2004.
9. Patrick Hayes. RDF model theory. W3C Recommendation, 10 February 2004. Available at `http://www.w3.org/TR/rdf-mt/`.
10. Patrick Hayes and Christopher Menzel. Simple common logic. In *W3C Workshop on Rule Languages for Interoperability*, 27–28 April 2005. Washington, D.C., USA.
11. Ian Horrocks, Peter F. Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosof, and Mike Dean. SWRL: A semantic web rule language combining owl and ruleml. W3C Member Submission, 21 May 2004. Available at `http://www.w3.org/Submission/SWRL/`.
12. Boris Motik, Ulrike Sattler, and Rudi Studer. Query answering for OWL-DL with rules. In *Proc. of the 2004 International Semantic Web Conference (ISWC 2004)*, pages 549–563, 2004.
13. Riccardo Rosati. On the decidability and complexity of integrating ontologies and rules. *J. of Web Semantics*, 2005. To appear.
14. SPARQL query language for RDF. W3C Working Draft, October 2004. Available at `http://www.w3.org/TR/rdf-sparql-query/`.