Bridging the Paradigm Gap with Rules for OWL

Pascal Hitzler(1)*, Jürgen Angele(2), Boris Motik(3)†, Rudi Studer(1),(2),(3)

(1)AIFB, University of Karlsruhe, Germany
(2)Ontoprise GmbH, Karlsruhe, Germany
(3)FZI Research Center, Karlsruhe, Germany

1 OWL and the Need for Rules in Practice

Accelerated by the vision of the semantic web, semantic technologies have recently made significant advances. The underlying methods and paradigms are already being transferred to adjacent areas of research in artificial intelligence, knowledge management, and elsewhere. Textbooks explaining the foundations have appeared. Large national and international projects on the topic are under way.

The technology received a crucial impulse when the Web Ontology Language OWL became a W3C recommendation in 2004. It is already being perceived as a basic knowledge representation language with potential which goes far beyond the semantic web use case. At the same time, it is apparent that OWL needs to be extended with additional expressive features in order to become applicable and useful in many domains. Corresponding efforts are being pursued with frenzy by research institutions and industry. These efforts are being aided substantially by the fact that OWL is based on sound logical foundations as it can be understood as a well-understood decidable albeit very expressive fragment of first order logic.

The need for an extension of OWL by rules has been known since the beginning. Indeed, the rule-based ontology representation language F-Logic [6, 1] is being used widely where a rule-based approach appears to be more feasible. As an example, we mention the HALO project[1] by Vulcan Inc.[2] whose ultimate goal is the creation of a "digital Aristotle", an expert tutor in a wide variety of subjects, with deep reasoning abilities. The initial

*The first named author acknowledges support by the European Union under the KnowledgeWeb Network of Excellence, and by the German Federal Ministry for Education and Research, BMBF, under the SmartWeb project.

†The third named author acknowledges support by the European Union under the DIP project.

[1]http://www.projecthalo.com/

[2]http://www.vulcan.com/

effort was structured around the challenge of responding to variants of AP Chemistry questions that focused on a portion of the "Advanced Placement test: Chemistry", used in the US as a qualification test before entering university as a student. The F-Logic-based system OntoNova, developed by Ontoprise, answers questions from this AP test. System performance turned out to be much better in general than that of real students taking the exam, thus displaying the strength of semantic technologies — and of F-Logic-based approaches.

2 Requirements on a Rule Language

We believe that OWL was a major and crucial step forward in the standardization of semantic technologies. The next step towards a rule language should draw on the impulse received and integrate seamlessly with OWL.

At the same time, we see that there is a plethora of needs for expressive features for semantic web support, including non-monotonic reasoning aspects, reasoning with inconsistencies, fuzzy and probabilistic features, built-in support, etc. We believe that it would go too far at this stage to accommodate any or all of these features. However, in order to facilitate future extensions, clean semantic foundations for the rule language will be required, and proximity to paradigms for which the basic methodologies for such extensions have already been established.

Likewise, we argue that a bridge should be drawn towards F-Logic based ontology languages, which can be done naturally in the course of developing a rule extension to OWL. Basic research required for integrating OWL with F-Logic has already been done, and can be drawn upon. As we will see in the next section, this perspective is also close to the SWRL rule extension for OWL, proposed in [2], whose basic ideas are compatible with the envisaged extension of OWL.

We furthermore believe that scalability issues should be taken into consideration. The size of the web, and of real applications, is overwhelming. As it is commonly accepted that complex reasoning tasks scale badly, we believe that heuristic approaches will play a major role in the future. This aspect emphasizes the need to stay within a cleanly defined logical framework, in order to be able to draw upon corresponding research in automated reasoning.

Finally, we stress the point that decidability was a major design decision in the making of OWL, and this contributed to its success. Indeed, complete but undecidable algorithms necessarily get into infinite computation loops on some inputs. While this is an undesirable aspect in itself, it also makes extensions and integrations with other knowledge representation paradigms much more difficult. As OWL is decidable, it is not hampered in this respect. Furthermore, decidable algorithms are usually amenable for optimization, even if of high computational complexity, which made it possible to implement practical reasoners for OWL and to apply them to real-world problems. Design decisions for a rule extension for OWL should therefore take decidability considerations into account. While we believe that a decidable rule extension would be too limited, decidable fragments should at least be identifiable, and considered as belonging to the core language.

In summary, we believe that the following requirements shall be taken into consideration in the development of a rule language:

- Seamless integration with OWL.
- Bridging the gap to F-Logic.
- Extensible in many directions.
- Scalable.
- Identifiable decidable fragments.

3 Towards a Proposal

We now briefly outline our work that addresses some of the issues outlined in the previous section. In [8] we consider the so-called *DL-safe* subset of SWRL. As shown in [9], DL-safe rules also integrate seamlessly with F-Logic by virtue of a natural translation of (a large fragment of) F-Logic into DL-safe rules. Roughly speaking, DL-safe rules can have an arbitrary structure, but to achieve decidability, the applicability of such rules is restricted to objects explicitly named in the knowledge base. For example, if *Person*, *livesAt*, and *worksAt* are concepts and roles from *KB*, the following rule is not DL-safe:

$$\mathit{Homeworker}(x) \leftarrow \mathit{Person}(x), \mathit{livesAt}(x,y), \mathit{worksAt}(x,y)$$

The reason for this is that both variables x and y occur in DL-atoms, but do not occur in a body atom with a predicate outside of KB. This rule can be made DL-safe by adding special non-DL-literals $\mathcal{O}(x)$ and $\mathcal{O}(y)$ to the rule body, and by adding a fact $\mathcal{O}(a)$ for each individual a. Hence, the above rule is transformed into the following rule:

$$\mathit{Homeworker}(x) \leftarrow \mathit{Person}(x), \mathit{livesAt}(x,y), \mathit{worksAt}(x,y), \mathcal{O}(x), \mathcal{O}(y)$$

A natural question arises about whether such a restriction is still useful in practical applications. For applications that require intensional reasoning, this is obviously a severe restriction: in such applications typically only a handful of individuals is known by name. However, a significant number of applications, such as metadata management on the Semantic Web or information integration, require extensive ABox reasoning. In such applications most individuals are usually known by name anyway, so the DL-safety restriction is not very restrictive in many cases.

To deal with DL-safe rules, we have developed new reasoning and query answering algorithms. They explore novel results relating description logics and the well-known formalism of (disjunctive) datalog [4, 5, 3]. In particular, we provide an algorithm that reduces a description logic knowledge base KB to a (disjunctive) datalog program $\mathsf{DD}(KB)$ without loosing relevant consequences. The algorithm also provided complexity bounds as side results, showing that efficient reasoning with OWL ontologies extended with rules may be feasible even for ontologies with large amount of data.

We are currently implementing these algorithms in KAON2[3], a new hybrid reasoner. If desired, a demonstration of the tool can be provided at the workshop. Whereas a full performance evaluation is yet to be performed, our preliminary experiments, reported in [7], are very promising.

We certainly do not argue that DL-safe rules should be the only extension made to OWL in the creation of a rule language. We rather argue that DL-safe rules should be considered as the *decidable base* of a rule language, whose design should follow this fundamental idea in order to stay within the tradition of OWL, and to bridge the gap to the paradigms already in use. Further extensions by expressive features such as non-monotonicity, fuzzyness, etc. can then be adopted naturally from the large amount of corresponding mature research results in the logic programming community.

4 Conclusion

Due to its ever growing popularity, we believe that a language for the rules on the Semantic Web should integrate seamlessly with OWL and its decidability tradition, should bridge the gap to F-Logic-based ontology languages, and should be based on firm logical foundations which allow to address extensions and scalability issues.

To address these issues, we proposed the so-called DL-safe fragment of SWRL as a foundation and starting point. It achieves decidability by restricting the applicability of rules only to objects explicitly named in the knowledge base. For applications that require extensive ABox reasoning this is not a severe restriction, since in such applications most objects are usually known by name. For such a logic we provide a reasoning procedure based on the translation of a description logic knowledge base into disjunctive datalog rules without losing interesting consequences. We have implemented the approach and, although a full evaluation is yet to be conducted, the initial experiments are very promising. Hence, we believe that DL-safe rules should be taken as foundation for a practical Semantic Web rule language.

[3] http://kaon2.semanticweb.org/

References

[1] Jürgen Angele and Georg Lausen. Ontologies in F-logic. In Steffen Staab and Rudi Studer, editors, *Handbook on Ontologies*, pages 29–50. Springer, 2004.

[2] I. Horrocks and P. F. Patel-Schneider. A Proposal for an OWL Rules Language. In *Proc. of the Thirteenth Int. World Wide Web Conf.(WWW 2004)*. ACM, 2004.

[3] U. Hustadt, B. Motik, and U. Sattler. Reasoning in Description Logics with a Concrete Domain in the Framework of Resolution. In R. López de Mántaras and L. Saitta, editors, *Proc. of the 16th European Conf. on Artificial Intelligence (ECAI 2004)*, pages 353–357, Valencia, Spain, August 2004.

[4] U. Hustadt, B. Motik, and U. Sattler. Reducing $\mathcal{SHIQ}^-$ Description Logic to Disjunctive Datalog Programs. In D. Dubois, C. Welty, and M.-A. Williams, editors, *Proc. of the 9th Int. Conf. on Knowledge Representation and Reasoning (KR 2004)*, pages 152–162, Menlo Park, California, USA, June 2004. AAAI Press.

[5] U. Hustadt, B. Motik, and U. Sattler. A Decomposition Rule for Decision Procedures by Resolution-based Calculi. In F. Baader and A. Voronkov, editors, *Proc. of the 11th Int. Conf. on International Conference on Logic for Programming Artificial Intelligence and Reasoning (LPAR 2004)*. Springer, 2005. To appear.

[6] M. Kifer, G. Lausen, and J. Wu. Logical foundations of object-oriented and frame-based languages. *Journal of the ACM*, 42(4):741–843, 1995.

[7] B. Motik, A. Maedche, and R. Volz. Optimizing Query Answering in Description Logics using Disjunctive Deductive Databases. In *10th International Workshop on Knowledge Representation meets Databases (KRDB-2003)*, Hamburg, Germany, September 15-16 2003.

[8] B. Motik, U. Sattler, and R. Studer. Query Answering for OWL-DL with Rules. In S. A. McIlraith, D. Plexousakis, and F. van Harmelen, editors, *Proc. of the 3rd Int. Semantic Web Conf. (ISWC 2004)*, volume 3298 of *Lecture Notes in Computer Science*, pages 549–563, Hiroshima, Japan, 2004. Springer.

[9] Boris Motik. Framework for hybrid, modularized ontology representation and reasoning for semantics-based services. EU DIP project deliverable D1.3, http://dip.semanticweb.org/deliverables.html, June 2004.