KAON2 — A Scalable Reasoning Tool for the Semantic Web

Boris Motik
FZI, University of Karlsruhe
76131 Karlsruhe, Germany
motik@fzi.de

Rudi Studer
AIFB, University of Karlsruhe
FZI, University of Karlsruhe
Ontoprise GmbH
76131 Karlsruhe, Germany
studer@aifb.uni-karlsruhe.de
studer@fzi.de

Abstract

Practical reasoning systems for standard ontology languages, such as OWL-DL, are crucial building blocks of the Semantic Web. To enable practical reasoning with reasonably large ontologies, we have developed a set of novel reasoning algorithms, and have implemented them in a new reasoning system KAON2. These algorithms reduce a $\mathcal{SHIQ}(\mathbf{D})$ knowledge base to a disjunctive datalog program, without losing any ground consequences, which allows applying known deductive database optimization techniques, such as magic sets, to DL query answering. Furthermore, the *data complexity* of our algorithms is in NP and, if disjunction is not used, even in P, which is better than the previously considered exponential *combined complexity*. Therefore, we believe our algorithms to be suitable for applications relying mainly on ABox reasoning. Although a complete evaluation has not yet been completed, our initial performance results are very promising.

1 Introduction

The applications of the Semantic Web technologies, such as e.g. metadata management or information integration, require reasoning over ontologies containing large amounts of data. Hence, development of reasoning systems capable of efficiently dealing with ontologies containing lots of axioms about instances is very important for the success of the Semantic Web.

KAON2[1] is a tool developed with the goal of fulfilling the above requirement. It is a full-fledged reasoning tool for various fragments of the Web Ontology Language (OWL), the standard ontology language for the Semantic Web. It is based on the $\mathcal{SHIQ}(\mathbf{D})$ description logic, which provides the logical foundation of OWL.

Furthermore, KAON2 also supports the so-called *DL-safe fragment* [Motik *et al.*, 2004] of the Semantic Web Rule Language (SWRL) [Horrocks and Patel-Schneider, 2004]. The DL $\mathcal{SHIQ}(\mathbf{D})$ and function-free rules are integrated as usual, by allowing concepts and roles to occur in rules as unary and binary predicates, respectively. Contrary to previous proposals, we allow concepts and roles to occur in rule heads; however, to achieve decidability, we require each variable in the rule to occur in a body literal with a predicate outside of the DL knowledge base. DL-safe rules provide means to circumvent certain expressivity drawbacks of OWL-DL without losing decidability of reasoning.

KAON2 contains an API for programmatic manipulation of ontologies and rules, which supports both RDF and XML formats of OWL. Furthermore, the inferencing modules of KAON2 have been designed to allow easy integration with external data sources, such as relational databases.

[1] Free for non-commercial use; the home-page and download available at http://kaon2.semanticweb.org/.

2 Reasoning in KAON2

Traditionally, tableaux algorithms [Baader and Sattler, 2001] have been used as the main mechanisms for reasoning over OWL ontologies. In KAON2, we explore a completely new approach, based on the relationship between description logics and disjunctive datalog [Hustadt *et al.*, 2004]. More concretely, given a $\mathcal{SHIQ}(\mathbf{D})$ knowledge base KB, our algorithms derive a disjunctive datalog program $\mathsf{DD}(KB)$ which entails the same set of ground consequences as KB. In this way, query answering in KB is reduced to query answering in $\mathsf{DD}(KB)$.

This work was motivated by the prospects of reusing well-known optimization techniques from the field of deductive databases, such as join order optimizations and the (disjunctive) magic sets [Cumbo *et al.*,]. These techniques have proven themselves to be very useful in practice for dealing with databases containing large amounts of data. Join order optimizations are the key technique enabling a database to choose the data access path promising the best performance based on the data statistics. Furthermore, magic sets allow to identify the subset of the database relevant to the query, and thus derive only relevant facts.

The computational complexity of OWL-Lite is ExpTime-complete, which is often taken as a serious drawback for practical applications. However, our algorithms show that *data complexity* — the complexity of reasoning under assumption that the size of data is much larger than the size of the schema — of $\mathcal{SHIQ}(\mathbf{D})$ (and therefore of OWL-Lite as well) is non-deterministically polynomial. Whereas still intractable, this is a significant improvement. Furthermore, if disjunctions are

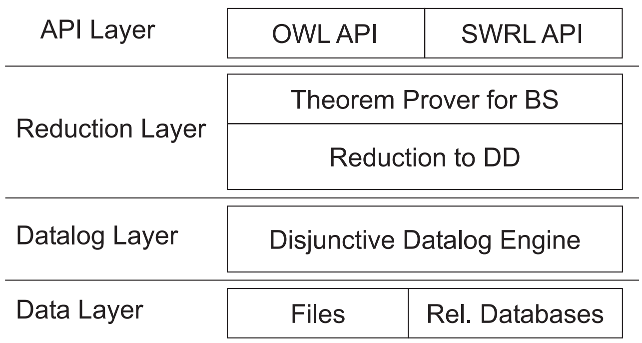

Figure 1: KAON2 Architecture

used sparsely (as it is often the case in practice), the data complexity of our algorithms is polynomial. Therefore, we believe that practical reasoning with large knowledge bases is feasible, and they theoretically confirm the validity of our approach.

The reasoning algorithms of KAON2 can be combined with DL-safe rules in a straightforward manner, by simply appending the rules to the program DD(KB). Furthermore, the implemented algorithms exhibit the principle of graceful degradation, where a performance penalty is paid only for primitives actually used.

3 Architecture

The architecture of KAON2 is shown in Figure 1. The users of KAON2 interact with the systems by means of OWL and SWRL APIs, which enable the programmatic manipulation of ontologies.

The algorithms from Section 2 are implemented in the reduction layer of KAON2 by means of a theorem prover for *basic superposition* [Bachmair *et al.*, 1995], a clausal calculus optimized for theorem proving with equality.

Reasoning with the programs obtained by the transformation is realized by means of an engine for disjunctive datalog. This engine implements several well-optimizations, such as magic sets.

Finally, KAON2 can reuse data stemming from OWL ontologies, but also from relational databases. In this way one can provide an "ontological view" over existing data.

4 Conclusion

Since KAON2 is a new tool, an extensive performance comparison of the new reasoning algorithms has not yet been been completed. However, in [Motik *et al.*, 2003] we reported on our preliminary performance tests, where we compare the performance of query answering in description logic using disjunctive datalog with the state-of-the-art reasoning system RACER [Haarslev and Möller, 2001]. These results are very encouraging, suggesting performance improvements of one, or sometimes even two orders of magnitude. Due to these results and the results about data complexity, we believe that KAON2 can be used as reasoning system for building advanced Semantic Web applications.

Acknowledgements

This work was partially funded by the EU IST project DIP 507483.

References

[Baader and Sattler, 2001] F. Baader and U. Sattler. An Overview of Tableau Algorithms for Description Logics. *Studia Logica*, 69:5–40, 2001.

[Bachmair *et al.*, 1995] L. Bachmair, H. Ganzinger, C. Lynch, and W. Snyder. Basic Paramodulation. *Information and Computation*, 121(2):172–192, 1995.

[Cumbo *et al.*,] C. Cumbo, W. Faber, G. Greco, and N. Leone. Enhancing the magic-set method for disjunctive datalog programs. In B. Demoen and V. Lifschitz, editors, *Proceedings of the 20th International Conference on Logic Programming (ICLP'04)*, number 3132 in Lecture Notes in Computer Science, pages 371–385. Springer.

[Haarslev and Möller, 2001] V. Haarslev and R. Möller. RACER System Description. In *1st Int. Joint Conf. on Automated Reasoning (IJCAR-01)*, pages 701–706. Springer, 2001.

[Horrocks and Patel-Schneider, 2004] 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.

[Hustadt *et al.*, 2004] 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.

[Motik *et al.*, 2003] 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.

[Motik *et al.*, 2004] 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.