

The limits and possibilities of combining Description Logics and Datalog

Riccardo Rosati
Dipartimento di Informatica e Sistemistica
Università di Roma “La Sapienza”
Via Salaria 113, 00198 Roma, Italy
email: rosati@dis.uniroma1.it

Abstract

Description Logics are currently the most used formalisms for building ontologies, and have been proposed as standard languages for the specification of ontologies in the Semantic Web. The problem of adding rules to Description Logics is currently a hot research topic, due to the interest of Semantic Web applications towards the integration of rule-based systems with ontologies. Most of the approaches in this field concern the study of description logic knowledge bases augmented with rules expressed in Datalog and its nonmonotonic extensions. In this talk we present a set of computational results which identify, from the viewpoint of the expressive abilities of the two formalisms, minimal combinations of Description Logics and (nonmonotonic) Datalog in which reasoning is undecidable. Then, based on the above results, we briefly survey some recent proposals for overcoming such expressive limitations.

Ontologies and Description Logics

Description Logics (DLs) are currently the most used formalisms for building ontologies, and have been proposed as standard languages for the specification of ontologies in the Semantic Web [9]. In particular, the OWL family of languages, based on Description Logics, has been explicitly designed to this purpose.

DLs are a family of knowledge representation formalisms based on first-order logic (FOL). In fact, almost all DLs coincide with decidable fragments of function-free first-order logic with equality, and the language of a DL can be seen as a restricted FOL language over unary and binary predicates and with a controlled form of quantification (actually, DLs are equipped with a special, variable-free syntax). Notably, DLs have been designed to optimize the trade-off between expressive abilities and complexity of reasoning, hence the computational properties of DLs have been extensively studied [2].

Limitations of current ontology formalisms

However, the experience in building practical applications has revealed several shortcomings of OWL and, in general, of Description Logics. In particular, the typical expressiveness of DLs does not allow for addressing the following aspects:

- the possibility of defining predicates of arbitrary arity (not just unary and binary)
- the use of variable quantification beyond the tree-like structure of DL concepts (many DLs actually correspond to subsets of the *two-variable fragment* of first-order logic)
- the possibility of formulating expressive queries over DL knowledge bases (beyond concept subsumption and instance checking)
- the possibility of formalizing various forms of closed-world reasoning over DL knowledge bases
- more generally, the possibility of expressing forms of nonmonotonic knowledge, like default rules [10]

The issue of how to overcome these limitations of OWL and DLs is currently receiving a lot of attention in the Semantic Web community [1]. In this respect, we observe that some of the above mentioned representational abilities which are missing in DLs require *nonmonotonicity* of the underlying logical formalism. This is in contrast with the well-known monotonic nature of classical first-order logic, which corresponds to the following property: if a theory T entails a conclusion ϕ , then, for every formula ψ , the theory $T \cup \{\psi\}$ entails ϕ . Such a property does not hold anymore in the presence of closed-world knowledge and default knowledge [10, 3].

This implies that the attempt to extend the expressive abilities of DLs, in order to fully overcome the above limitations, requires to leave the realm of classical first-order logic, and to look at nonmonotonic logic.

Rule-based knowledge representation

Almost all the kinds of knowledge that cannot be formally addressed in a classical, first-order logic setting have a “rule-like” form, i.e., can be expressed by statements of the form “if the precondition ψ holds then the conclusion ϕ holds”, where the precondition and the conclusion are logical properties.

However, such a piece of knowledge cannot simply be formalized through the standard material implication of classical logic: in other words, it is not possible to capture the intended meaning of the above statement by an implication in classical first-order logic of the form $\psi \rightarrow \phi$.

In this respect, a very relevant role is played by research in logic programming. In fact, logic program rules are implications with a non-standard semantics. And, in the context of ontologies, nonmonotonic extensions of logic programming are of particular interest [3].

Therefore, rule-based formalisms grounded in logic programming have repeatedly been proposed as a possible solution to overcome the above limitations, so adding a rule layer on top of OWL is nowadays seen as the most important task in the development of the Semantic Web language stack. The Rule Interchange Format (RIF) working group of the World Wide Web Consortium (W3C) is currently working on standardizing such a language [1].

Most of the proposals in this field focus on logic programs expressed in Datalog (and its nonmonotonic extensions) [4]. With respect to DLs, Datalog allows for using predicates of arbitrary arity, the explicit use of variables, and the ability of expressing more powerful queries. Moreover, its nonmonotonic features (in particular, the negation-as-failure operator *not*) allow for expressing default rules and forms of closed-world reasoning.

Integrating DLs and rules: the computational issue

Many semantic and computational problems have emerged in the combination of DLs and rule-based representation formalisms. Among them, we concentrate on the computational issue that arises when combining DLs and rules. Indeed, as shown by the first studies in this field [7], decidability (and complexity) of reasoning is a crucial issue in systems combining DL knowledge bases and Datalog rules. In fact, in general this combination does not preserve decidability, i.e., starting from a DL knowledge base in which reasoning is decidable and a set of rules in which reasoning is decidable, reasoning in the knowledge base obtained by integrating these two components may not be a decidable problem.

In this talk we present a set of computational results which identify, from the viewpoint of the expressive abilities of the two formalisms, minimal combinations of Description Logics and (nonmonotonic) Datalog in which rea-

soning is undecidable.

More precisely, we both resume known results on decidability and complexity of combining DLs and rules [7, 5, 6, 8, 11, 12] and present new computational results on reasoning in DLs with rules, to the aim of identifying the aspects of the DL language and the rule language that are critical for decidability and effectiveness of reasoning.

In particular, we single out (i) some expressive abilities of the DL language, (ii) some expressive abilities of the rule language, and (iii) different “degrees of integration” between the two languages. Then, we identify minimal combinations of the above three aspects which lead to undecidability of reasoning in a system integrating a Datalog rule base and a DL knowledge base. We believe that such an analysis is very important in order to identify effective formalisms for the integration of ontologies and rules.

Finally, we briefly survey recent proposals for overcoming some of the expressive limitations pointed out by the above analysis.

References

- [1] Rule interchange format working group charter. <http://www.w3.org/2005/rules/wg/charter>.
- [2] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, and P. F. Patel-Schneider, editors. *The Description Logic Handbook: Theory, Implementation and Applications*. Cambridge University Press, 2003.
- [3] C. Baral and M. Gelfond. Logic programming and knowledge representation. *J. of Logic Programming*, 19–20:73–148, 1994.
- [4] T. Eiter, G. Gottlob, and H. Mannilla. Disjunctive Datalog. *ACM Trans. on Database Systems*, 22(3):364–418, 1997.
- [5] T. Eiter, T. Lukasiewicz, R. Schindlauer, and H. Tompits. Combining answer set programming with description logics for the semantic web. In *Proc. of the 9th Int. Conf. on the Principles of Knowledge Representation and Reasoning (KR 2004)*, pages 141–151, 2004.
- [6] I. Horrocks and P. F. Patel-Schneider. A proposal for an OWL rules language. In *Proc. of the 13th Int. World Wide Web Conf. (WWW 2004)*, pages 723–731, 2004.
- [7] A. Y. Levy and M.-C. Rousset. Combining Horn rules and description logics in CARIN. *Artificial Intelligence*, 104(1–2):165–209, 1998.
- [8] B. Motik, U. Sattler, and R. Studer. Query answering for OWL-DL with rules. *J. of Web Semantics*, 3(1):41–60, 2005.
- [9] P. F. Patel-Schneider, P. J. Hayes, I. Horrocks, and F. van Harmelen. OWL web ontology language; semantics and abstract syntax. W3C candidate recommendation, <http://www.w3.org/tr/owl-semantics/>, november 2002.
- [10] R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.
- [11] R. Rosati. On the decidability and complexity of integrating ontologies and rules. *J. of Web Semantics*, 3(1):61–73, 2005.
- [12] R. Rosati. DL+log: tight integration of Description Logics and disjunctive Datalog. In *Proc. of the 10th Int. Conf. on the Principles of Knowledge Representation and Reasoning (KR 2006)*, pages 68–78, 2006.