Reports of Enlarged Session of the
Seminar of I. Vekua Institute
of Applied Mathematics
Volume 23, 2009

INTEGRATION OF ONTOLOGIES WITH RULE SYSTEMS

Rukhaia M.

Abstract. Work around Semantic Web is growing. There is a lot of work about integration of Ontology layer with Rule layer, but there is not a straightforward solution of this problem because of a various obstacles. For example, naively adding rules to ontologies raises undecidability issues.

In this paper we will gave a brief overview of the Semantic Web components, such as RDF, OWL and Rule systems. Then we survey existing approaches to the problem of combining rule languages with ontology languages for the Semantic Web. We focus on the languages based on logic and on the reasoning in such languages.

Keywords and phrases: Semantic Web, ontologies, rule systems.

AMS subject classification (2000): 03B70; 68T15.

1. Introduction. The Semantic Web is a web of data that is application-independent, composeable, classified, and part of a larger information system - ontology. Explicit representation of the semantics of data, programs, pages, and other web resources will enable us a knowledge-based web that will provide a qualitatively new level of service. Automated services will improve in their capacity to assist humans in achieving their goals by understanding more or the content on the web, and thus providing more accurate filtering, categorizing, and searching of these information sources. The main ideas behind are to add a machine-readable meaning to web pages, to use ontologies for a precise definition of shared terms in web resources, to make use of knowledge representation technology for automated reasoning from web resources, and to apply cooperative agent technology for processing the information of the Web.

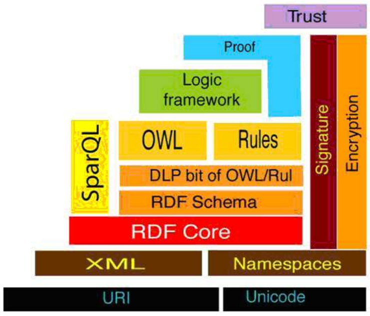

Fig.1. Semantic Web Stack[1]

The Semantic Web is represented in hierarchical layers (see Figure 1), as was proposed by Tim Berners-Lee in [9]. The work around Semantic Web is growing for getting the layers of its architecture ready. Resource Description Framework (RDF) and Web

---

[1] This picture is taken from T.B. Lee's presentation (see [9])

Ontology Language (OWL) layers has reached a certain level of maturity, they are W3C recommendations. Now main problem is integration of these layers with rules layer. There are several proposals in this direction and the objective of this document is to provide a survey of these developments.

The Resource Description Framework (RDF) is an XML-based knowledge representation language for the Semantic Web. RDF data-model building block is a $\langle subject, Predicate, object\rangle$ triple, called a statement, expressing that some resource (*subject*) is related to another resource or a value (*object*) through the property (*Predicate*).

The RDF Vocabulary Description Language (RDF Schema) is a an extension of RDF. It introduces the notions of class and property and provides mechanisms for specifying class hierarchies, property hierarchies and for defining domains and ranges of properties.

The Web Ontology Language (OWL) is build on RDF and RDF Schema. It contains three sublanguages, OWL Lite, OWL DL and OWL Full. OWL DL is based on Description Logic and it consist of *classes* and *properties*, also called *concepts* and *roles* respectively. They are considered as unary and binary predicates and interpreted as relations.

The rule languages for the Semantic Web are based on different kinds of logics and have thus well defined declarative semantics, supported by well-developed reasoning algorithms. The simplest language of this kind consists of Horn clauses not including function symbols other than constants. There are a lots of rule systems developed for semantic web, such as RuleML, SWRL, F-Logic, RIF and so on. Integration of a rule language with an ontology language requires definition of a new language, its syntax and semantics and development of reasoning algorithms for the new language. In [3] existing proposals for integration ontology languages with rule languages is classified as *hybrid* (strict separation between the rule predicates and ontology predicates) and *homogeneous* (both ontology and rule languages are embedded in one language without distinction between the rule and ontology predicates) approaches.

2. Reasoning with Rules and Ontologies. Reasoning with rules and ontologies is nowadays main research subject in the Semantic Web. The open-world semantics of OWL is very expressive, but there are several important modeling problems, described in [2], that are hard or impossible to solve using OWL, but can be easily solved by logic programming. These problems are:

Higher Relational Expressivity. OWL provides a rich set of primitives for expressing concepts, but the only axioms OWL can express are of a tree-structure. Many real-world applications require modeling general relational structures, and decidable rule-based formalism such as function-free Horn rules can do this.

Polyadic Predicates. Concepts and roles of OWL corresponds to unary and binary predicates, but in practice many relationships require predicates with arity more than two.

Closed-World Reasoning. Let consider an OWL knowledge base O. Due to the open-world semantics of OWL, we can use O to answer positive but not negative queries. Answering queries about negative information in an intuitive way requires some form of closed-world reasoning.

The meaning of closed-world reasoning and open-world reasoning are opposite of each other. In the closed-world reasoning, the knowledge base is assumed to be complete, and what is not currently known to be true is false. In contrast, with the open world assumption the knowledge base is not assumed to contain all information, and the answer of questions that is not known to be true explicitly is unknown. In other words, open-word reasoning is monotonic, that means that adding new knowledge to knowledge base does not falsify previous knowledge. This is not case in close-world reasoning, it is nonmonotonic.

One more difference between OWL and logic programming is an unique name assumption (UNA). UNA is a feature of nonmonotonic formalisms, that assumes that distinct constants mean different things, but OWL does not employ UNA. In OWL explicit statements is needed to define that two objects are different.

Integrity Constraints. In OWL domain and range restrictions constrain the type of objects that can be related by a role. Also, participation restrictions specify that certain objects have relationships to other objects, but it is well-known that integrity constraints can not be realized in OWL within first-order logic.

Modeling Exceptions. Exceptions are everywhere in the real world. It is not possible to model exceptions in OWL. To enable exception modeling, one must apply nonmonotonic formalism, usually involving some form of default negation.

Several shortcomings of the OWL, such as inability to model integrity constraints or perform closed-world querying, should be overcome by rule-based formalisms grounded in logic programming. In [4], SWRL was proposed as standard, but it was criticized because of undecidability and inability of expressing nonmonotonic reasoning tasks. In [6], Description Logic Programs (DLP), and in [7] hybrid MKNF knowledge bases was proposed as a suitable rules languages. The RIF working group of the W3C is currently working on standardizing such a rule language.

2.1. Description Logic Programs. Description Logic Programs are a straightforward intersection of DLs and LP. It enables on one side to build rules on top of ontologies and on the other side to build ontologies on top of rules in a little extent. In [6] description logic programs are defined by means of Description Horn Logic (DHL) ontology, while in [5] it is defined by means of dl-atom and answer set semantics.

The syntax of an ordinary logic program is a set of rules of the form:

$$H \leftarrow B_1 \wedge \dots B_n \wedge \sim B_{m+1} \wedge \dots \wedge \sim B_m,$$

where H, B_i are atomic formulaes and are called head and body respectively. $\sim$ stand for negation a failure. A definite LP rule is an ordinary LP rule without negation as failure. Horn rule is a Horn clause in which exactly one literal is positive. It has the same syntax as definite LP rule. An LP or Horn rule is equality-free if there is no equality predicate in it, and is datalog if there is no logical functions in it.

The semantics of an ordinary LP is a set of facts, entailed by the LP. Let def-LP be a definite equality-free datalog LP, and def-Horn be the corresponding definite equality-free datalog Horn fragment of FOL. The conclusion set of def-LP coincides with the minimal Herbrand model of def-Horn.

There is transformation from DL axioms to def-Horn. But there is some kind of axioms, that cannot b transformed into def-Horn rules. For example, cardinality restrictions correspond to assertions of variable equality and inequality in FOL, and this is not allowed in the def-Horn framework.

some DL constructors (conjunction and universal restriction) can be mapped to the heads of rules whenever they occur on the right-hand side of an inclusion axiom, while some DL constructors (conjunction, disjunction and existential restriction) can be mapped to the bodies of rules whenever they occur on the left-hand side of an inclusion axiom. Let denote by L_h (respectively L_b) classes from which can be mapped into the head (respectively body) of LP rules. And denote by L intersection of L_h and L_b. Then, there is a recursive mapping function T, which takes a DL axiom of the form $C \sqsubseteq D$, where C is an L_b-class and D is an L_h-class, and maps it into an LP rule of the form $A \leftarrow B$.

Description Horn Logic (DHL) ontology is a set of DHL axioms of the form $C \sqsubseteq D$, $A \equiv B$, $\top \sqsubseteq \forall P.D$, $\top \sqsubseteq \forall P^{-}.D$, $P \sqsubseteq Q$, $P \equiv Q$, $P \equiv Q^{-}$, $P^{+} \sqsubseteq P$, $a : D$, and $\langle a, b\rangle : P$, where C is an L_b-class, D is an L_h-class, A, B are L-classes, P, Q are properties and a, b are individuals. Applying mapping T to the DHL ontology O, gives def-Horn ruleset H, that has the same set of models and entailed conclusions as O. Using translation T, inferencing in the DHL fragment of DL can be reduced to inferencing in LP.

A def-LP is a Description Logic Program (DLP) when it is the LP-correspondent of some DHL ruleset. Semantically, a DLP is a weakening of def-Horn along the dimension of entailment power. In the conclusions it permits only facts. DLP can be viewed as an expressive subset of DL. For more detailed information, we refer reader to [6].

2.1.1 DL-Programs. Description logic programs (dl-programs) consist of a description knowledge base L and a finite set of description logic rules (dl-rules) P which may contain queries to L, possibly default negated in their bodies. In [5], dl-rules was defined as rules of the form:

$$a \leftarrow b_1, \dots, b_m, \mathit{not}\, b_{m+1}, \dots, \mathit{not}\, b_n,$$

where a is a classical literal and any b_i is either a classical literal or a dl-atom. This was extended later in [1], and they allow disjunction in the head of the rule, i.e. rules of the form:

$$a_1 \vee \cdots \vee a_k \leftarrow b_1, \dots, b_m, \mathit{not}\, b_{m+1}, \dots, \mathit{not}\, b_n,$$

where any a_i is a classical literal and any b_j is either a classical literal or a dl-atom. A dl-atom is defined as an expression of the form:

$$DL[S_1 op_1 p_1, \dots, S_m op_m p_m; Q](t), \qquad m \geq 0,$$

where each S_i is either a concept or a role, $op_i \in \{\uplus, \cup\!\!\!-, \cap\!\!\!-\}$, p_i is a unary or binary predicate symbol, called input predicate symbol, and $Q(t)$ is a dl-query. A dl-query is either a concept inclusion axiom F or its negation $\neg F$; or either of the form $C(t)$ or $\neg C(t)$, where C is a concept and t is a term; or of the form $R(t_1, t_2)$ or $\neg R(t_1, t_2)$, where R is a role and t_1, t_2 are terms.

Using dl-programs, closed-world reasoning may be easily expressed on top of an knowledge base L, which can be queried through suitable dl-atoms. If given a concept C, then its negated version $\bar{C}$ can be defined by the following dl-rule $\bar{C} \leftarrow$ $not\, DL[C](X)$. It is known, that closed-world assumption sometimes can lead to inconsistent conclusions, but for example, minimal model reasoning can avoid this problem. This extension can be easily implemented in the framework of dl-programs, by means of a suitable encoding that computes minimal models of a base L. Building minimal models of L corresponds to concluding as much negative facts as possible while keeping consistency.

2.2. Hybrid MKNF Knowledge Bases. Hybrid MKNF knowledge bases, as was proposed in [7], integrates description logic (DL) with disjunctive logic programs and negation as failure. It provides exactly the same consequences as DL and LP, respectively, if the other component is empty and allows users to switch between open and closed-world approaches on arbitrary predicates from DL and LP. It is based on the logic of Minimal Knowledge and Negation as Failure (MKNF) with several modifications.

MKNF is an extension of the first-order logic with modal operators **K** and **not**. A first-order atom $P(t_1, \dots, t_n)$ is an MKNF formula where P is a predicate and t_i are either constants or variables. If ϕ and ϕ_1 are an MKNF formulas then $\neg\phi$, $\exists x : \phi$, $\phi \wedge \phi_1$, $\phi \subset \phi_1$, $\mathbf{K}\ \phi$ and $\mathbf{not}\ \phi$ are MKNF formulas too. $\mathbf{K}\ \phi$ and $\mathbf{not}\ \phi$ are modal **K**-atom and **not**-atom respectively. $\phi[t/x]$ is the formula obtained from ϕ by replacing all free occurrences of the variable x with term t. An MKNF formula is a sentence if it does not contain free variables, ground if it does not contain variables and positive if it does not contain **not** operator.

Apart from the constants occurring in the formulas the signature contains a countably infinite supply of constants not occurring in the formulas. The Herbrand Universe of such a signature is denoted by $\triangle$. The signature contains the equality predicate $\approx$ which is interpreted as congruence relation on $\triangle$. An MKNF structure is a triple (I, M, N) where I is an Herbrand first-order interpretation over $\triangle$ and M and N are nonempty sets of Herbrand first-order interpretations over $\triangle$.

A hybrid MKNF knowledge base K consists of a knowledge base O in any decidable description logic D and a set P of MKNF rules of the following form:

$$\mathbf{K}H_1 \vee \cdots \vee \mathbf{K}\ H_n \leftarrow \mathbf{K}\ A_1, \dots, \mathbf{K}A_n,\ \mathbf{not}\ B_1, \dots, \mathbf{not}\ B_m,$$

where H_i, A_i and B_i are first-order atoms of the form $P(t_1, \dots, t_n)$. The $\mathbf{K}H_i$, $\mathbf{K}A_i$ and $\mathbf{not}B_i$ are called the rule head, positive body and negative body respectively. The semantics of a hybrid MKNF knowledge base is defined by translating it into MKNF formula.

A DL-atom is a first-order function-free atom $P(t_1, \dots, t_n)$ such that P is either $\approx$ or it occurs in a DL knowledge base O. All other atoms are non-DL-atoms. Every extension of DLs with rules are undecidable in general case. So hybrid MKNF knowledge bases are also undecidable in general case, but decidability is obtained using DL-safety concept. An MKNF rule is DL-safe if each variable in the rule occurs in a non-DL-atom of the form $\mathbf{K}A$ in the rule body. With DL-safe rules hybrid MKNF knowledge bases have same complexity as the corresponding fragment of logic programming.

In [8] was defined well-founded semantics for hybrid MKNF knowledge bases restricted to nondisjunctive MKNF rules. An MKNF rule is nondisjunctive if it has only one atom in rule head. In [8] was shown that well-founded semantics is a sound approximation of the semantics of [7], computational complexity is strictly lower and retains the property of being faithful wrt. the well-founded semantics of LPs. Better complexity is achieved by having only one three-valued model which is semantically weaker, but bottom-up computable.

3. Conclusions. This document surveys proposals for extending the Semantic Web ontology layer with rules. For the lack of space we have very briefly described basic Semantic Web components. Then we discussed the proposals addressing integration of OWL with rules. The DLP approach defines an intersection of the Description Logic underlying OWL and Horn clauses, thus making possible re-use of existing reasoners. On the other hand, hybrid MKNF knowledge bases approach, which integrates an arbitrary description logic with disjunctive logic programs and negation as failure. This integration is faithful in the sense that it provides exactly the same consequences as DL and LP, respectively, if the other component is empty and allows the user to freely switch between open-world and closed-world views on arbitrary predicates from DL and LP.

REFERENCES

1. Eiter T., Ianni G., Polleres A., Schindlauer R., Tompits H. Reasoning with rules and ontologies. *In Reasoning Web 2006, Lecture Notes in Computer Science,* **4126**, *Springer*, (2006), 93-127.

2. Motik B., Horrocks I., Rosati R., Sattler U. Can OWL and logic programming live together ever after? *In Proc. of the 2006 International Semantic Web Conference (ISWC 2006)*, 2006.

3. Antoniou G., Damasio C.V., Grosof B., Horrocks I., Kifer M., Maluszynski J., Patel-Schneider P.F. Combining Rules and Ontologies. A survey. *Rewerse* 2005.

4. W3C 2004. SWRL: A Semantic Web Rule Language Combining OWL and RuleML. *W3C Member Submission 21 May 2004. Available at http://www.w3.org/Submission/SWRL/.*

5. Eiter T., Lukasiewicz T., Schindlauer R., Tompits H. 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), Morgan Kaufmann, Los Altos*, (2004), 141-151.

6. Grosof B.N., Volz R., Horrocks I., Decker S. Description Logic Programs: Combining Logic Programs with Description Logic. *In Proc. WWW2003, Budapest, Hungary*, 2003.

7. Motik B., Rosati R. A faithful integration of description logics with logic programming. *In Proceedings of the Twentieth International Joint Conference on Artificial Intelligence (IJCAI-07), Hyderabad, India, AAAI Press*, (2007), 477-482.

8. Knorr M., Alferes J.J., Hitzler P. A Well-founded Semantics for Hybrid MKNF Knowledge Bases. *In Proceedings of the 2007 International Workshop on Description Logics (DL-2007), Brixen-Bressanone, Italy*, 2007.

9. Tim Berners-Lee. Web for real people, 2005. *Available at http://www.w3.org/2005/Talks/0511-keynote-tbl/*

Received 13.05.2009; revised 15.09.2009; accepted 22.11.2009.

Author's address:

M. Rukhaia
New University of Lisbon
Department of Information Technology
Faculty of Science and Technology
2829-516 Caparica
Portugal
E-mail: m.rukhaia@fct.unl.pt