

# Nonmonotonic Onto-Relational Learning

Francesca A. Lisi and Floriana Esposito

Dipartimento di Informatica, Università degli Studi di Bari  
Via E. Orabona 4, 70125 Bari, Italy  
{lisi, esposito}@di.uniba.it

**Abstract.** In this paper we carry on the work on Onto-Relational Learning by investigating the impact of having disjunctive DATALOG with default negation either in the language of hypotheses or in the language for the background theory. The inclusion of nonmonotonic features strengthens the ability of our ILP framework to deal with incomplete knowledge by performing some form of commonsense reasoning. One such ability can turn out to be useful in application domains, such as the Semantic Web, which require that kind of reasoning.

## 1 Motivation

An increasing amount of conceptual knowledge is being made available in the form of *ontologies* [8] mostly specified with languages based on *Description Logics* (DLs) [1]. The problem of adding rules to DLs is currently a hot research topic in Knowledge Representation (KR), due to the interest of Semantic Web applications towards the integration of rule systems with ontologies testified by the activity of the W3C Rule Interchange Format (RIF) working group<sup>1</sup> and of the 'Web Reasoning and Rule Systems' conference series<sup>2</sup>. Practically all the approaches in this field concern the study of DL knowledge bases (KBs) augmented with rules expressed in DATALOG [4] and its nonmonotonic (NM) extensions such as disjunctive DATALOG with default negation ( $\text{DATALOG}^{-\vee}$ ) [6]. Many technical problems arise in this kind of KR systems. In particular, the full interaction between a DL KB and a DATALOG program easily leads to semantic and computational problems related to the simultaneous presence of knowledge interpreted under the *Open World Assumption* (OWA) and knowledge interpreted under the *Closed World Assumption* (CWA) [13]. The KR framework  $\mathcal{DL}+\text{LOG}$  allows for the *tight* integration of DLs and  $\text{DATALOG}^{-\vee}$ , through a *weak* safeness condition for variables in rules [14].

In [11] we have laid the foundations of an extension of Relational Learning, called Onto-Relational Learning, to account for ontologies. In that work we have proposed to adapt generalized subsumption [3] to a decidable instantiation of  $\mathcal{DL}+\text{LOG}$  obtained by integrating the DL *SHIQ* [10] and positive DATALOG [4]. The resulting hypothesis space can be searched by means of refinement operators

<sup>1</sup> [http://www.w3.org/2005/rules/wiki/RIF\\_Working\\_Group](http://www.w3.org/2005/rules/wiki/RIF_Working_Group)

<sup>2</sup> <http://www.rr-conference.org/>

either top-down or bottom-up. In order to define a coverage relation we have assumed the ILP setting of learning from interpretations. Both the coverage relation and the generality relation boil down to query answering in  $\mathcal{DL}+\text{LOG}$ . These ingredients for Onto-Relational Learning do not depend on the scope of induction and are still valid for any other decidable instantiation of  $\mathcal{DL}+\text{LOG}$ , provided that positive DATALOG is still considered. In this paper we carry on the work initiated in [11] by investigating the impact of having  $\text{DATALOG}^{-\forall}$  either in the language of hypotheses or in the language for the background theory. The inclusion of the NM features of  $\mathcal{DL}+\text{LOG}$  *full* will strengthen the ability of our ILP framework to deal with incomplete knowledge by performing some form of commonsense reasoning. One such ability can turn out to be useful in application domains, such as the Semantic Web, which require that kind of reasoning.

The paper is organized as follows. Section 2 introduces the KR framework of  $\mathcal{DL}+\text{LOG}$ . Section 3 sketches an ILP algorithm for learning  $\mathcal{DL}+\text{LOG}$  rules. Section 4 concludes the paper with final remarks.

## 2 $\mathcal{DL}+\text{log}$ : Integrating DLs and Disjunctive Datalog

Description Logics (DLs) are a family of decidable First Order Logic (FOL) fragments that allow for the specification of knowledge in terms of classes (*concepts*), binary relations between classes (*roles*), and instances (*individuals*) [2]. Complex concepts can be defined from atomic concepts and roles by means of constructors such as atomic negation ( $\neg$ ), concept conjunction ( $\sqcap$ ), value restriction ( $\forall$ ), and limited existential restriction ( $\exists$ ) - just to mention the basic ones. The DL *SHIQ* was the starting point for the design of the ontology language OWL for the Semantic Web [9]. A DL KB can state both is-a relations between concepts (*axioms*) and instance-of relations between individuals (resp. couples of individuals) and concepts (resp. roles) (*assertions*). Concepts and axioms form the so-called TBox whereas individuals and assertions form the so-called ABox<sup>3</sup>. A *SHIQ* KB encompasses also a RBox, i.e. axioms defining hierarchies over roles. The semantics of DLs can be defined through a mapping to FOL. Thus, coherently with the OWA that holds in FOL semantics, a DL KB represents all its models. The main reasoning task for a DL KB is the *consistency check* that is performed by applying decision procedures based on tableau calculus.

Disjunctive DATALOG ( $\text{DATALOG}^{-\forall}$ ) is a variant of DATALOG that admits disjunction in the rules' heads and default negation [6]. The presence of disjunction in the rules' heads because it makes  $\text{DATALOG}^{-\forall}$  inherently nonmonotonic, i.e. new information can invalidate previous conclusions. Among the many alternatives, one widely accepted semantics for  $\text{DATALOG}^{-\forall}$  is the extension to the disjunctive case of the *stable model semantics* originally conceived for normal logic programs (i.e. logic programs with default negation) [7]. According to

---

<sup>3</sup> When a DL-based ontology language is adopted, an ontology is nothing else than a TBox eventually coupled with a RBox. If the ontology is populated, it corresponds to a whole DL KB, i.e. encompassing also an ABox.

this semantics, a DATALOG<sup>-v</sup> program may have several alternative models (but possibly none), each corresponding to a possible view of the reality.

The hybrid KR framework of  $\mathcal{DL}+\text{LOG}$  allows for the tight integration of DLs and DATALOG<sup>-v</sup> [14]. More precisely, it allows a DL KB to be extended with DATALOG<sup>-v</sup> rules of the form:

$$r_1(\mathbf{Y}_1), \dots, r_m(\mathbf{Y}_m), s_1(\mathbf{Z}_1), \dots, s_k(\mathbf{Z}_k), \text{not } u_1(\mathbf{W}_1), \dots, \text{not } u_h(\mathbf{W}_h) \leftarrow p_1(\mathbf{X}_1) \vee \dots \vee p_n(\mathbf{X}_n) \quad (1)$$

where  $n, m, k, h \geq 0$ , each  $p_i(\mathbf{X}_i)$ ,  $r_j(\mathbf{Y}_j)$ ,  $s_l(\mathbf{Z}_l)$ ,  $u_k(\mathbf{W}_k)$  is an atom and each  $p_i$  is either a DL-predicate or a DATALOG predicate, each  $r_j$ ,  $u_k$  is a DATALOG predicate, each  $s_l$  is a DL-predicate. Peculiar to  $\mathcal{DL}+\text{LOG}$  is the condition of *weak safeness*: Every head variable of a rule must appear in at least one of the atoms  $r_1(\mathbf{Y}_1), \dots, r_m(\mathbf{Y}_m)$ . It allows to overcome the main representational limits of the approaches based on the DL-safeness condition, e.g. the possibility of expressing conjunctive queries (CQ) and unions of conjunctive queries (UCQ)<sup>4</sup>, by keeping the integration scheme still decidable. For  $\mathcal{DL}+\text{log}$  a FOL semantics and a NM semantics have been defined. The FOL semantics does not distinguish between head atoms and negated body atoms. Thus, the form (1) is equivalent to:

$$p_1(\mathbf{X}_1) \vee \dots \vee p_n(\mathbf{X}_n) \vee u_1(\mathbf{W}_1), \dots, u_h(\mathbf{W}_h) \leftarrow r_1(\mathbf{Y}_1), \dots, r_m(\mathbf{Y}_m), s_1(\mathbf{Z}_1), \dots, s_k(\mathbf{Z}_k) \quad (2)$$

The NM semantics is based on the stable model semantics of DATALOG<sup>-v</sup>. According to it, DL-predicates are still interpreted under OWA, while DATALOG predicates are interpreted under CWA. Notice that, under both semantics, entailment can be reduced to satisfiability and, analogously, that CQ answering can be reduced to satisfiability. The NMSAT- $\mathcal{DL}+\text{LOG}$  algorithm has been provided for checking only the NM-satisfiability of finite  $\mathcal{DL}+\text{LOG}$  KBs because FOL-satisfiability can always be reduced (in linear time) to NM-satisfiability by rewriting rules from the form (1) to the form (2). It is shown that the decidability of reasoning in  $\mathcal{DL}+\text{LOG}$ , thus of ground query answering, depends on the decidability of the Boolean CQ/UCQ containment problem in  $\mathcal{DL}$ . Currently, *SHIQ*+LOG is one of the most expressive decidable instantiations of  $\mathcal{DL}+\text{LOG}$ .

### 3 Learning $\mathcal{DL} + \text{log}$ Rules

We face the problem of inducing an integrity theory  $\mathcal{H}$  for an extensional DATALOG database  $\Pi_F$  by exploiting a background theory  $\mathcal{K}$  which consists of a  $\mathcal{DL}$  KB  $\Sigma$  and the intensional DATALOG database  $\Pi_R$ . Since the scope of induction is description and a  $\mathcal{DL}+\text{LOG}$  KB may be incomplete, this learning task can be considered as a case of characteristic induction from entailment. We assume that

<sup>4</sup> A *Boolean UCQ* over a predicate alphabet  $P$  is a FOL sentence of the form  $\exists \mathbf{X}. \text{conj}_1(\mathbf{X}) \vee \dots \vee \text{conj}_n(\mathbf{X})$ , where  $\mathbf{X}$  is a tuple of variable or constant symbols and each  $\text{conj}_i(\mathbf{X})$  is a set of atoms whose predicates and arguments are in  $P$  and  $\mathbf{X}$  respectively. A *Boolean CQ* corresponds to a Boolean UCQ for  $n = 1$ .

```

NMDISC- $\mathcal{DL}+\text{log}$ ( $\mathcal{L}, \mathcal{K}, \Pi_F$ )
1.  $\mathcal{H} \leftarrow \emptyset$ 
2.  $\mathcal{Q} \leftarrow \{ \square \}$ 
3. while  $\mathcal{Q} \neq \emptyset$  do
4.    $\mathcal{Q} \leftarrow \mathcal{Q} \setminus \{R\}$ ;
5.   if NMSAT- $\mathcal{DL}+\text{LOG}$ ( $\mathcal{K} \cup \Pi_F \cup \mathcal{H} \cup \{R\}$ )
6.     then  $\mathcal{H} \leftarrow \mathcal{H} \cup \{R\}$ 
7.     else  $\mathcal{Q} \leftarrow \mathcal{Q} \cup \{R' \in \mathcal{L} \mid R' \in \rho^{-\vee}(R)\}$ 
8.   endif
9. endwhile
return  $\mathcal{H}$ 

```

**Fig. 1.** Main procedure of NMDISC- $\mathcal{DL}+\text{LOG}$

$\Pi_F$  and  $\Sigma$  shares a common set of constants. The integrity theory  $\mathcal{H}$  is therefore a set of  $\mathcal{DL}+\text{LOG}$  rules. It must be induced so that  $\mathcal{B} = (\Sigma, \Pi_F \cup \Pi_R \cup \mathcal{H})$  is a NM-satisfiable  $\mathcal{DL}+\text{LOG}$  KB. In the following we sketch the ingredients for an ILP system, named NMDISC- $\mathcal{DL}+\text{LOG}$ , able to discover  $\mathcal{DL}+\text{LOG}$  integrity theories on the basis of NMSAT- $\mathcal{DL}+\text{LOG}$ .

The algorithm in Figure 1 defines the main procedure of NMDISC- $\mathcal{DL}+\text{LOG}$ : it starts from an empty theory  $\mathcal{H}$  (1), and a queue  $\mathcal{Q}$  containing only the empty clause (2). It then applies a search process (3) where each element  $R$  is deleted from the queue  $\mathcal{Q}$  (4), and tested for satisfaction w.r.t. the data  $\Pi_F$  by taking into account the background theory  $\mathcal{K}$  and the current integrity theory  $\mathcal{H}$  (5)<sup>5</sup>. If the rule  $R$  satisfies the database (6), it is added to the theory (7). If the rule violates the database, its refinements according to  $\mathcal{L}$  are considered (8). The search process terminates when  $\mathcal{Q}$  becomes empty (9). Note that the algorithm does not specify the search strategy. In order to get a minimal theory (i.e., without redundant clauses), a pruning step and a post-processing phase can be added to NMDISC- $\mathcal{DL}+\text{LOG}$  by further calling NMSAT- $\mathcal{DL}+\text{LOG}$ <sup>6</sup>.

The language  $\mathcal{L}$  of hypotheses allows for generating  $\mathcal{DL}+\text{LOG}$  rules. Since we assume the database  $\Pi_F$  to be correct, a rule  $R$  must be modified to make it satisfiable by  $\Pi_F$  by either (i) strenghtening  $body(R)$  or (ii) weakening  $head(R)$ . The refinement operator  $\rho^{-\vee}$  combines the two refinement operators presented in [11] by applying the rules:

- $\langle AddDataLit \rangle$ ,  $\langle AddOntoLit \rangle$ , and  $\langle SpecOntoLit \rangle$  on  $body(R)$ ;
- $\langle AddDataLit \rangle$ ,  $\langle AddOntoLit \rangle$ , and  $\langle GenOntoLit \rangle$  on  $head(R)$

Note that, since we are working under NM-semantics, the use of  $\langle AddDataLit \rangle$  for adding negated DATALOG atoms to the body of rules is not the same as for adding DATALOG atoms to the head.

<sup>5</sup> The NM-satisfiability test includes also the current induced theory in order to deal with the nonmonotonicity of induction in the normal ILP setting.

<sup>6</sup> Based on the following consequence of the Deduction Theorem in FOL: Given a KB  $\mathcal{B}$  and a rule  $R$  in  $\mathcal{DL}+\text{LOG}$ , we have that  $\mathcal{B} \models R$  iff  $\mathcal{B} \wedge \neg R$  is unsatisfiable.

*Example 1.* Let us consider a  $\mathcal{DL}+\text{LOG}$  KB  $\mathcal{B}$  (adapted from [14]) which integrates the following  $\mathcal{DL}$  KB  $\Sigma$  (ontology about persons)

```

PERSON  $\sqsubseteq$   $\exists$  FATHER-.MALE
MALE  $\sqsubseteq$  PERSON
FEMALE  $\sqsubseteq$  PERSON
FEMALE  $\sqsubseteq$   $\neg$ MALE
MALE(Bob)
PERSON(Mary)
PERSON(Paul)

```

and the following  $\text{DATALOG}^{-\vee}$  program  $\Pi$  (database about students):

```

FEMALE(X)  $\leftarrow$  girl(X)
MALE(X)  $\leftarrow$  boy(X)
boy(Paul)
girl(Mary)
enrolled(Paul,c1)
enrolled(Mary,c1)
enrolled(Mary,c2)
enrolled(Bob,c3)

```

containing also rules linking the database to the ontology. The following rules belonging to the language  $\mathcal{L}$  built upon the alphabet  $\{\text{boy}(\_), \text{girl}(\_), \text{not boy}(\_), \text{not girl}(\_), \text{enrolled}(\_, c1), \text{enrolled}(\_, c2), \text{enrolled}(\_, c3), \text{PERSON}(\_)\}$ :

```

 $\leftarrow$  enrolled(X,c1), PERSON(X)
boy(X)  $\vee$  girl(X)  $\leftarrow$  enrolled(X,c1)
 $\leftarrow$  enrolled(X,c2), not boy(X)
 $\leftarrow$  enrolled(X,c3), MALE(X)

```

are an excerpt of the discovered integrity theory.

As for the hypothesis ordering, note that generalized subsumption - chosen in [11] for adaptation - applies only to definite clauses (yet, with some trick to normal clauses). Conversely, relative subsumption [12] can cope with arbitrary clauses and admit an arbitrary finite set of clauses as the background theory. So it is suitable for extension to  $\mathcal{DL}+\text{LOG}$  rules.

## 4 Conclusions and Future Work

In this paper we have carried on the work on Onto-Relational Learning by considering the problem of learning  $\mathcal{DL}+\text{LOG}$  rules to be used as integrity theory for a database whose schema is represented also by means of an ontology. The main procedure of  $\text{NMDISC-}\mathcal{DL}+\text{LOG}$  is inspired by [5] as for the scope of induction and the algorithm scheme but differs from it in several points, notably the adoption of (i) relative subsumption instead of  $\theta$ -subsumption, (ii) stable model semantics instead of completion semantics, and (iii) learning from entailment instead of learning from interpretations, to deal properly with the chosen

representation formalism for both the background theory and the language of hypotheses. The NM features as well as the DL component of  $\mathcal{DL}+\text{LOG}$  allow  $\text{NMDISC-}\mathcal{DL}+\text{LOG}$  to induce very expressive integrity theories. In this paper we have therefore addressed an issue that has been brought to the attention of the database community with the advent of the Semantic Web, i.e. the issue of how ontologies (and semantics conveyed by them) can help solving typical database problems, through a better understanding of KR aspects related to databases. In the future we plan to further investigate this issue from the ILP perspective.

## References

1. F. Baader, D. Calvanese, D. McGuinness, D. Nardi, and P.F. Patel-Schneider, editors. *The Description Logic Handbook: Theory, Implementation and Applications (2nd ed.)*. Cambridge University Press, 2007.
2. A. Borgida. On the relative expressiveness of description logics and predicate logics. *Artificial Intelligence*, 82(1–2):353–367, 1996.
3. W. Buntine. Generalized subsumption and its application to induction and redundancy. *Artificial Intelligence*, 36(2):149–176, 1988.
4. S. Ceri, G. Gottlob, and L. Tanca. What you always wanted to know about datalog (and never dared to ask). *IEEE Transactions on Knowledge and Data Engineering*, 1(1):146–166, 1989.
5. L. De Raedt and M. Bruynooghe. A theory of clausal discovery. In *IJCAI*, pages 1058–1063, 1993.
6. T. Eiter, G. Gottlob, and H. Mannila. Disjunctive DATALOG. *ACM Transactions on Database Systems*, 22(3):364–418, 1997.
7. M. Gelfond and V. Lifschitz. Classical negation in logic programs and disjunctive databases. *New Generation Computing*, 9(3/4):365–386, 1991.
8. T. Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5:199–220, 1993.
9. I. Horrocks, P.F. Patel-Schneider, and F. van Harmelen. From *SHIQ* and RDF to OWL: The making of a web ontology language. *Journal of Web Semantics*, 1(1):7–26, 2003.
10. I. Horrocks, U. Sattler, and S. Tobies. Practical reasoning for very expressive description logics. *Logic Journal of the IGPL*, 8(3):239–263, 2000.
11. F.A. Lisi and F. Esposito. Foundations of Onto-Relational Learning. In F. Železný and N. Lavrač, editors, *Inductive Logic Programming*, volume 5194 of *Lecture Notes in Artificial Intelligence*, pages 158–175. Springer, 2008.
12. G.D. Plotkin. A further note on inductive generalization. *Machine Intelligence*, 6:101–121, 1971.
13. R. Rosati. Semantic and computational advantages of the safe integration of ontologies and rules. In F. Fages and S. Soliman, editors, *Principles and Practice of Semantic Web Reasoning*, volume 3703 of *Lecture Notes in Computer Science*, pages 50–64. Springer, 2005.
14. R. Rosati.  $\mathcal{DL}+\text{log}$ : Tight Integration of Description Logics and Disjunctive Datalog. In P. Doherty, J. Mylopoulos, and C.A. Welty, editors, *Proc. of Tenth International Conference on Principles of Knowledge Representation and Reasoning*, pages 68–78. AAAI Press, 2006.