Reasoning under Uncertainty with Log-Linear Description Logics

Mathias Niepert

KR & KM Research Group,
Universität Mannheim
Mannheim, Germany
{mathias}@informatik.uni-mannheim.de

Abstract. The position paper provides a brief summary of log-linear description logics and their applications. We compile a list of five requirements that we believe a probabilistic description logic should have to be useful in practice. We demonstrate the ways in which log-linear description logics answer to these requirements.

1 Introduction

Uncertainty is pervasive in the real world and reasoning in its presence one of the most pressing challenges in the development of intelligent systems. It is therefore hard to imagine how the Semantic Web could succeed without the ability to represent and reason under uncertainty. Nevertheless, purely logical approaches to knowledge representation and reasoning such as description logics have proven useful in providing the formal backbone of the Semantic Web. There is not only a large body of important work on the logical and algorithmic properties of such languages but also highly optimized tools that are successfully employed in meaningful applications. Still, the need to model uncertainty persists. Two prominent examples where the processing of uncertainty is crucial are (a) data integration (schema and instance alignment) and (b) ontology learning. In both cases, algorithms usually generate confidence values for particular axioms. In ontology matching, for instance, string similarity measures are often used to find confidence values for equivalence axioms between concepts and properties, respectively.

There have been attempts to combine logic and probability in various ways. Resulting approaches are probabilistic formalism for description logics [4, 5, 2, 6, 8] and, more generally, statistical relational languages [3]. The former are important theoretical contributions but have not been adopted by practitioners. We believe this is primarily due to the computational complexity of probabilistic inference, the rather involved way of expressing uncertainties syntactically, and the lack of implementations. Statistical relational approaches, on the other hand, have been successfully applied to numerous real-world problems but they do not explicitly take into account the notion of coherency and consistency which is crucial in the context of the Semantic Web.

Name	Syntax	Semantics
top	$\top$	$\Delta^{\mathcal{I}}$
bottom	$\bot$	$\emptyset$
nominal	$\{a\}$	$\{a^{\mathcal{I}}\}$
conjunction	$C \sqcap D$	$C^{\mathcal{I}} \cap D^{\mathcal{I}}$
existential restriction	$\exists r.C$	$\{x \in \Delta^{\mathcal{I}} \vert \exists y \in \Delta^{\mathcal{I}} : (x, y) \in r^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\}$
GCI	$C \sqsubseteq D$	$C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$
RI	$r_1 \circ ... \circ r_k \sqsubseteq r$	$r_1^{\mathcal{I}} \circ ... \circ r_k^{\mathcal{I}} \subseteq r^{\mathcal{I}}$

Table 1. The description logic $\mathcal{EL}^{++}$ without nominals and concrete domains.

Based on these observations (and biases), and with the more concrete applications of ontology learning and matching in mind, we have compiled the following wish list for a probabilistic description logic.

1. The system must be usable by individuals knowledgeable only in Semantic Web languages and tools such as OWL and Protégé;
2. It must be possible to express uncertainty in form of *degrees of confidence* (real-valued weights) and not necessarily in form of precise probabilities. Real-world problems such as ontology matching and learning require this;
3. The user should not have to worry about inconsistent and incoherent input to the probabilistic reasoner. All types of inconsistencies are handled by the probabilistic reasoner and not the user;
4. Two types of queries should be supported under uncertainty: (a) The "most probable ontology" query and (b) the probability of (conjunctions) of axioms query; and
5. The worst-case complexity should not exceed that of probabilistic graphical models such as Markov and Bayesian networks. While inference in these models is generally NP-hard, numerous highly efficient algorithms exist and can be employed in the context of probabilistic DLs.

These five requirements are captured by log-linear description logics [7]. We provide a brief overview of log-linear description logics and discuss how this family of probabilistic logics answers to the outlined requirements.

2 Log-Linear Description Logics

Log-linear description logics integrate description logics with probabilistic log-linear models. Detailed technical and empirical results are available [7] and are mostly omitted in this position paper. The syntax of log-linear description logics is taken from the underlying description logic. However, it is possible to assign real-valued weights to axioms. Here, we focus on the log-linear description logic based on $\mathcal{EL}^{++}$ [1] without concrete domains (see Table 1) which we denote as

$\mathcal{EL}^{++}$-LL. $\mathcal{EL}^{++}$ captures the expressivity of numerous ontologies in the biomedical sciences and other domains, and it is the description logic on which the web ontology language profile OWL 2 EL is based. More formally, a $\mathcal{EL}^{++}$-LL ontology $\mathcal{C} = (\mathcal{C}^{\mathsf{D}}, \mathcal{C}^{\mathsf{U}})$ is a pair consisting of a *deterministic* $\mathcal{EL}^{++}$ CBox (set of axioms) $\mathcal{C}^{\mathsf{D}}$ and an *uncertain* CBox $\mathcal{C}^{\mathsf{U}} = \{(c, w_c)\}$ which is a set of pairs (c, w_c) with each c being a $\mathcal{EL}^{++}$ axiom and w a real-valued weight assigned to c. While the *deterministic* CBox contains axioms that are known to be true the *uncertain* CBox contains axioms for which we only have a *degree of confidence*. Every axiom can either be part of the deterministic or the uncertain CBox but not of both.

The semantics of log-linear DLs is based on joint probability distributions over *coherent* $\mathcal{EL}^{++}$ CBoxes and similar to that of Markov logic [9]. The weights of the axioms determine the log-linear probability distribution. For a $\mathcal{EL}^{++}$-LL CBox $(\mathcal{C}^{\mathsf{D}}, \mathcal{C}^{\mathsf{U}})$ and a $\mathcal{EL}^{++}$ CBox $\mathcal{C}'$ over the same set of basic concept descriptions and role names, we have that

$$P(\mathcal{C}') = \begin{cases} \frac{1}{Z} \exp\left(\sum_{\{(c,w_c) \in \mathcal{C}^{\mathsf{U}} : \mathcal{C}' \models c\}} w_c\right) & \text{if } \mathcal{C}' \text{ is } \textit{coherent} \\ & \text{and } \mathcal{C}' \models \mathcal{C}^{\mathsf{D}}; \\ 0 & \text{otherwise} \end{cases}$$

where Z is the normalization constant of the log-linear probability distribution.

The semantics of the log-linear description logic leads to probability distributions one would expect under the open world semantics of description logics.

Example 1. Let $\mathsf{Student}$ and $\mathsf{Professor}$ be two classes and let $\mathcal{C}^{\mathsf{D}} = \emptyset$ and $\mathcal{C}^{\mathsf{U}} = \{\langle \mathsf{Student} \sqsubseteq \mathsf{Professor}, 0.5\rangle, \langle \mathsf{Student} \sqcap \mathsf{Professor} \sqsubseteq \bot, 0.5\rangle\}$. Then[1], $P(\{\mathsf{Student} \sqsubseteq \mathsf{Professor}, \mathsf{Student} \sqcap \mathsf{Professor} \sqsubseteq \bot\}) = 0$, $P(\{\mathsf{Student} \sqsubseteq \mathsf{Professor}\}) = Z^{-1}\exp(0.5)$, $P(\{\mathsf{Student} \sqsubseteq \mathsf{Professor}, \mathsf{Professor} \sqsubseteq \mathsf{Student}\}) = Z^{-1}\exp(0.5)$, $P(\{\mathsf{Student} \sqcap \mathsf{Professor} \sqsubseteq \bot\}) = Z^{-1}\exp(0.5)$, $P(\{\mathsf{Professor} \sqsubseteq \mathsf{Student}\}) = Z^{-1}\exp(0)$, and $P(\emptyset) = Z^{-1}\exp(0)$ with $Z = 3\exp(0.5) + 2\exp(0)$.

We distinguish two types of probabilistic queries. The maximum a-posteriori (MAP) query: *"Given a $\mathcal{EL}^{++}$-LL CBox, what is a most probable coherent $\mathcal{EL}^{++}$ CBox over the same concept and role names?"*; and the conditional probability query: *"Given a $\mathcal{EL}^{++}$-LL CBox, what is the probability of a conjunction of axioms?"* We believe that the first type of query is useful since it infers the most probable *coherent* ontology from one that contains axioms with confidence values. The MAP query, therefore, has immediate applications in ontology learning and matching.

Probabilistic inference in log-linear description logics seems daunting at first, considering the combinatorial complexity of the problem. It turns out, however, that both the MAP and the conditional probability query can be computed efficiently for ontologies with thousands of known and uncertain axioms [7]. The worst-case complexity of both queries is equivalent to the worst-case complexity of the analogous queries in Markov and Bayesian networks (requirement 5).

[1] We omit trivial axioms that are present in every classified CBox such as $\mathsf{Student} \sqsubseteq \top$ and $\mathsf{Student} \sqsubseteq \mathsf{Student}$.

3 Log-Linear Description Logics in Practice

ELOG is a log-linear description logic reasoner developed at the University of Mannheim. A detailed description, the source code, and example ontologies are available at its webpage[2]. ELOG directly loads ontologies expressed in OWL 2 EL. The assignment of confidence values to axioms is made with the annotation property "confidence." Consider the following example ontology zoo.owl:

```
SubClassOf(
  Annotation(<http://URI/ontology#confidence> "0.5"^^xsd:double)
  <http://zoo/Penguin>
  <http://zoo/Bird>
)
DisjointClasses(
  <http://zoo/Bird>
  <http://zoo/Mammal>
)
```

Here, the subclass axiom is assigned the confidence value 0.5 and the disjointness axiom is considered true since it is not annotated. Therefore, the subclass axiom is part of the uncertain CBox and the disjointness axiom is part of the deterministic CBox. Considering that annotations can simply be added with popular ontology editors such as Protégé or using the OWL API[3], log-linear description logics fulfill requirements 1 and 2. In addition, the annotated axioms do not have to be consistent or coherent in any way because ELOG computes the probabilistic queries with respect to the joint probability distribution over *coherent* ontologies. Thus, ELOG also fulfills requirement 3.

References

1. Baader, F., Brandt, S., Lutz, C.: Pushing the $\mathcal{EL}$ envelope. In: Proceedings of the International Joint Conference on Artificial Intelligence (2005)
2. Costa, P.: Bayesian semantics for the Semantic Web. Ph.D. thesis, George Mason University (2005)
3. Getoor, L., Taskar, B.: Introduction to Statistical Relational Learning. MIT Press (2007)
4. Jaeger, M.: Probabilistic reasoning in terminological logics. In: Proceedings of the Conference on the Principles of Knowledge Representation and Reasoning (1994)
5. Koller, D., Levy, A., Pfeffer, A.: P-classic: A tractable probabilistic description logic. In: Proceedings of the 14th AAAI Conference on Artificial Intelligence (1997)
6. Lukasiewicz, T., Straccia, U.: Managing uncertainty and vagueness in description logics for the semantic web. J. of Web Sem. 6 (2008)
7. Niepert, M., Noessner, J., Stuckenschmidt, H.: Log-Linear Description Logics. In: Proceedings of the International Joint Conference on Artificial Intelligence (2011)
8. Predoiu, L., Stuckenschmidt, H.: Probabilistic models for the semantic web. In: The Semantic Web for Knowledge and Data Management (2008)
9. Richardson, M., Domingos, P.: Markov logic networks. Machine Learning (2006)

[2] http://code.google.com/p/elog-reasoner/

[3] http://owlapi.sourceforge.net/