Tractable Probabilistic Description Logic

Andrew Ijano Lopes

Supervisor: Marcelo Finger

April 16, 2020

1 Introduction and Motivation

Description logics are a family of formal knowledge representation languages, being of particular importance in providing a logical formalism for ontologies and the Semantic Web. Also, they are notable in biomedical informatics for assisting the codification of biomedical knowledge. Due to these uses, there is a great demand to find tractable (i.e., polynomial-time decidable) description logics.

One of them, the logic $\mathcal{EL}^{++}$, is one of the most expressive description logics in which the complexity of inferential reasoning is tractable (Baader et al., 2005). Even though it is expressive enough to deal with several practical applications, there was also a need to model situations in which a General Concept Inclusion Axiom is not always true, which has already been proposed in the literature (Boole, 1854).

Example 1.1. Consider a medical situation, adapted from (Finger, 2019), in which a patient may have non-specific symptoms, such as high fever, skin rash and joint pains. Also, dengue is a disease that can account for those symptoms, but not all patients present all symptoms. Such an uncertain situation is suitable for probabilistic modeling.

In a certain hospital, a patient with high fever has some probability of having dengue, but that probability is 5% larger if the patient has rash too. On the other hand, dengue is not very prevalent and is not observed in the hospital 70% of the time. If those probabilistic constraints are satisfiable, one can also ask the minimum and maximum probability that a hospital patient John, with fever and rash, is a suspect of suffering from dengue.

For classical propositional formulas, this problem, called *probabilistic satisfiability* (PSAT), has already been presented with tractable fragments (Andersen and Pretolani, 2001). On the other hand, in description logics, most studies result in intractable reasoning; moreover, by adding probabilistic reasoning capabilities to $\mathcal{EL}^{++}$, in order to model such situation, the complexity reaches NP-completeness (Finger, 2019).

To solve this problem, probabilistic constrains can be applied to axioms and its probabilistic satisfaction can be seen in a linear algebraic view. Furthermore, it can be reduced to an optimization problem, which can be solved by an adaptation of the simplex method with column generation. (Finger, 2019) Thus, it is possible to reduce the column generation problem to the *weighted partial maximum satisfatibility*.

Then, recent studies show that it is necessary to focus on a fragment of $\mathcal{EL}^{++}$ for obtain a tractable probabilistic reasoning. This fragment is called *Graphic* $\mathcal{EL}^{++}$ ($\mathcal{GEL}^{++}$) and it is defined as an $\mathcal{EL}^{++}$-fragment in which its set of axioms and *role inclusions* contains formulas in *normal form* and does not allow explicit conjunction axioms. Therefore, axioms can be seen as edges in a graph, as opposed to hyperedges in a hypergraph, which is the case of $\mathcal{EL}^{++}$. This allows the use of graph-based machinery to develop tractable algorithm for the *weighted partial Maximum SATisfatibility* for $\mathcal{GEL}^{++}$ (Max $\mathcal{GEL}^{++}$-SAT) and, as a result, a tractable probabilistic description logic.

2 Objectives

- Study and implement tractable algorithms for the problem of *weighted partial* Max $\mathcal{GEL}^{++}$-SAT.
- Study and implement algorithms for the problem of *probabilistic satisfatibility for* $\mathcal{GEL}^{++}$ (PGEL-SAT), using the Max $\mathcal{GEL}^{++}$-SAT solver as a subroutine. Thus, it is excepted to achieve a tractable algorithm for a probabilistic description logic.

Furthermore, a possible extension for this project is explore the PGEL-SAT tractability to propose a machine learning algorithm for the probabilities involved.

3 Work plan

1. Study the problem and propose algorithms to solve it;

2. Implement a Max $\mathcal{GEL}^{++}$-SAT solver using max-flow/min-cut techniques;
3. Implement a PGEL-SAT solver;
4. Implement a reader of ontologies in the Web Ontology Language (OWL);
5. Study methods to generate random probabilistic $\mathcal{GEL}^{++}$ formulas;
6. Generate statistical results about the tractability of the algorithms;
7. Write the monograph;
8. Write the poster and presentation;

Most of the study of the problem started last year and has already been done, as well as some initial implementations.

Activity	Months								
	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov
1.	x	x							
2.	x	x	x	x					
3.	x	x	x	x	x	x			
4.					x	x			
5.						x	x	x	
6.							x	x	
7.				x	x	x	x	x	x
8.								x	x

References

K. A. Andersen and D. Pretolani. Easy cases of probabilistic satisfiability. *Annals of Mathematics and Artificial Intelligence*, 33(1):69–91, 2001.

F. Baader, S. Brandt, and C. Lutz. Pushing the EL envelope. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence*, IJCAI'05, pages 364–369, 2005.

G. Boole. *An Investigation of the Laws of Thought: On which are Founded the Mathematical Theories of Logic and Probabilities*. Collected logical works. Walton and Maberly, 1854. URL https://books.google.com.br/books?id=DqwAAAAAcAAJ.

M. Finger. Extending EL++ with linear constraints on the probability of axioms. In C. Lutz, U. Sattler, C. Tinelli, A.-Y. Turhan, and F. Wolter, editors, *Description Logic, Theory Combination, and All That*, volume LLNCS 11560 of *Theoretical Computer Science and General Issues*. Springer International Publishing, 2019. ISBN 978-3-030-22101-0. doi: 10.1007/978-3-030-22102-7.