University of São Paulo
Institute of Mathematics and Statistics
Bachelor of Computer Science

Tractable Probabilistic Description Logic

Andrew Ijano Lopes

Final Essay

mac 499 — Capstone Project

Program: Computer Science
Advisor: Prof. Dr. Marcelo Finger

São Paulo
January 20th, 2021

Tractable Probabilistic Description Logic

Andrew Ijano Lopes

This is the original version of the
capstone project report prepared by
the candidate Andrew Ijano Lopes, as
submitted to the Examining Committee.

Resumo

Andrew Ijano Lopes. **Lógica de Descrição Probabilística Tratável: *Algorithms and Implementation***. Monografia (Bacharelado). Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2020.

Elemento obrigatório, constituído de uma sequência de frases concisas e objetivas, em forma de texto. Deve apresentar os objetivos, métodos empregados, resultados e conclusões. O resumo deve ser redigido em parágrafo único, conter no máximo 500 palavras e ser seguido dos termos representativos do conteúdo do trabalho (palavras-chave). Deve ser precedido da referência do documento.

Palavras-chave: Lógicas de descrição. Palavra-chave2. Palavra-chave3.

Abstract

Andrew Ijano Lopes. **Tractable Probabilistic Description Logic**. Capstone Project Report (Bachelor). Institute of Mathematics and Statistics, University of São Paulo, São Paulo, 2020.

Elemento obrigatório, elaborado com as mesmas características do resumo em língua portuguesa. De acordo com o Regimento da Pós-Graduação da USP (Artigo 99), deve ser redigido em inglês para fins de divulgação. É uma boa ideia usar o sítio www.grammarly.com na preparação de textos em inglês.

Keywords: Description logics. Keyword2. Keyword3.

Lista de Abreviaturas

DL	Description Logic

Lista de Símbolos

List of Figures

List of Tables

List of Programs

Contents

Chapter 1

Introduction

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 uncertain knowledge.

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 satisfiability*.

Moreover, recent studies show that it is necessary to focus on a fragment of $\mathcal{EL}^{++}$

for obtain a tractable probabilistic reasoning, which will be called Graphic $\mathcal{EL}^{++}$ ($\mathcal{GEL}^{++}$) (Finger, n.d.). Therefore, this fragment allows axioms to 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 SATisfiability* for $\mathcal{GEL}^{++}$ (Max $\mathcal{GEL}^{++}$-SAT) and, as a result, a tractable probabilistic description logic.

Then, the objective of this project is to propose and implement tractable algorithms for weighted partial Max-SAT and Probabilistic SAT for a fragment of $\mathcal{EL}^{++}$ description logic.

In this paper, we describe the implementation of these algorithms[1] and is organized as follows: Section **??** highlights related results in the literature. The basic definition of $\mathcal{GEL}^{++}$ with its algorithms for MaxSAT and PSAT are described in Section **??** and followed by Section **??**, which presents details about the implementation and its experimental evaluation.

[1] Available at https://github.com/AndrewIjano/pgel-sat

Chapter 2

Background

2.1 The description logic $\mathcal{EL}^{++}$

2.2 Graphic $\mathcal{EL}^{++}$ ($\mathcal{GEL}^{++}$)

2.3 MaxSAT for $\mathcal{GEL}^{++}$

2.4 Probabilistic $\mathcal{GEL}^{++}$

2.5 Related Work

Chapter 3

Development

3.1 OWL parser

3.2 Knowledge Base

3.3 GEL-MaxSAT

3.4 Linear solver

3.5 PGEL-SAT reasoner

Chapter 4

Experiments

Chapter 5

Results

Chapter 6

Conclusion

References

[ANDERSEN and PRETOLANI 2001] Kim Allan ANDERSEN and Daniele PRETOLANI. "Easy cases of probabilistic satisfiability". In: *Annals of Mathematics and Artificial Intelligence* 33.1 (2001), pp. 69–91 (cit. on p. 1).

[BAADER *et al.* 2005] Franz BAADER, Sebastian BRANDT, and Carsten LUTZ. "Pushing the EL Envelope". In: *Proceedings of the 19th International Joint Conference on Artificial Intelligence*. IJCAI'05. 2005, pp. 364–369 (cit. on p. 1).

[FINGER 2019] Marcelo FINGER. "Extending EL++ with Linear Constraints on the Probability of Axioms". In: *Description Logic, Theory Combination, and All That. Essays Dedicated to Franz Baader on the Occasion of His 60th Birthday*. Ed. by Carsten LUTZ, Uli SATTLER, Cesare TINELLI, Anni-Yasmin TURHAN, and Frank WOLTER. Vol. LLNCS 11560. Theoretical Computer Science and General Issues. Springer International Publishing, 2019. ISBN: 978-3-030-22101-0. DOI: 10.1007/978-3-030-22102-7 (cit. on p. 1).

[FINGER n.d.] Marcelo FINGER. "Tractable Max-SAT in Graphic Description Logics". In preparation (cit. on p. 2).