TPLP **22** (3): 496–498, 2022. © The Author(s), 2021. Published by Cambridge University Press 496
doi:10.1017/S1471068421000041 First published online 10 March 2021

Book review

Probabilistic Semantic Web Reasoning and Learning, by Riccardo Zese

The book is published in the "Studies on the Semantic Web" series by IOS Press and it is an extended, revised version of Riccardo Zese's PhD thesis (Zese 2016). It introduces an extension of standard Description Logics by a distributed semantics called DISPONTE. Algorithms and implementations for reasoning in such extensions are also provided. The book contains 27 chapters that are organized in 6 parts.

Part I is the *Introduction* of the book and it is devoted to the presentation of the basic notions, including a brief exposition of the Description Logics and the motivations of the work. The structure of the book is also provided.

Part II of the book on *Description Logics* completes the presentation of the state of the art for Description Logics. The family of Description Logics, reasoning services as well as some examples are provided. For the aim of the book, strong Description Logics underlying the OWL language are considered. This allows the author to drive from the very beginning the user in the context of the Semantic Web.

Part III on *A Probabilistic Semantics for Description Logics* is dedicated to the logical system at the base of the contribution of this book, namely the probabilistic extension of Description Logics based on the DISPONTE semantics.

Probabilistic extensions of Description Logics, allowing one to label inclusions (and facts) with degrees representing probabilities, have been introduced by Riguzzi and colleagues (Riguzzi *et al.* 2015a,b). In this approach, called DISPONTE, the integration of probabilistic information with Description Logic is based on the distribution semantics for probabilistic logic programs (Sato 1995). The basic idea is to label inclusions of the TBox, as well as facts of the ABox, with a real number in the interval [0,1], representing their probabilities, assuming that axioms are independent from each other. The resulting Knowledge Base (KB) defines a probability distribution over *worlds*: roughly speaking, a world is obtained by assuming for each axiom of the KB, whether it should be considered to be true or false in that world. The distribution is further extended to queries, and the probability of the entailment of a query is obtained by marginalizing the joint distribution of the query and the worlds. As an example, consider the following version of a KB inspired by the people and pets ontology (Riguzzi *et al.* 2015a):

$$0.3 \quad :: \quad \exists hasAnimal.Pet \sqsubseteq NatureLover, \tag{A}$$
$$0.6 \quad :: \quad Cat \sqsubseteq Pet, \tag{B}$$
$$0.9 \quad :: \quad Cat(tom), \tag{C}$$
$$hasAnimal(kevin, tom). \tag{D}$$

Inclusion (A) expresses that individuals that own a pet are nature lovers with a 30% probability, whereas inclusion (B) states that cats are pets with probability 60%. The ABox fact (C) represents that Tom is a cat with probability 90%. Inclusions (A), (B)

and (C) are *probabilistic* axioms, whereas (D) is an axiom, that must always hold. The KB has eight possible worlds:

$\{(A,0),\ (B,0),\ (C,0)\}$ $\{(A,0),\ (B,0),\ (C,1)\}$ $\{(A,0),\ (B,1),\ (C,0)\}$
$\{(A,0),\ (B,1),\ (C,1)\}$ $\{(A,1),\ (B,0),\ (C,0)\}$ $\{(A,1),\ (B,0),\ (C,1)\}$
$\{(A,1),\ (B,1),\ (C,0)\}$ $\{(A,1),\ (B,1),\ (C,1)\}$

representing all possible combinations of assuming each probabilistic axiom to be either true or false. For instance, the world $\{(A,1),(B,0),(C,1)\}$ represents the situation in which (A) and (C) hold, that is, $\exists hasAnimal.Pet \sqsubseteq NatureLover$ and $Cat(tom)$, whereas (B) does not. The query $NatureLover(kevin)$ is true only in the world with (A), (B), and (C) all true, whereas it is false in all the other ones. The probability of such a query is $0.3 \times 0.6 \times 0.9 = 0.162$.

Part IV on *Inference in Probabilistic DLs* is devoted to the main contribution of the author's work: the algorithms for reasoning in the probabilistic extensions of Description Logics. After an introduction of splitting algorithms and binary decision diagrams, the algorithms BUNDLE, TRILL, and TRILL^P are introduced.

BUNDLE is an algorithm whose aim is to compute the probability of queries from a probabilistic (DISPONTE) KB, focusing on the basic Description Logic $\mathcal{ALC}$. BUNDLE exploits the existing reasoner Pellet (Sirin *et al.* 2007) that is able to return explanations for queries. The explanations are encoded in a binary decision diagram from which the probability of the query is computed.

TRILL contains three different systems that can answer queries by means of a tableau algorithm. They return the probability of the query, which represents the degree of belief of its truth. It is implemented in Prolog, which is able to handle the choice of the rule to apply by means of its built-in depth-first search mechanism and backtracking.

Moreover, the system TRILL^P returns a pinpointing formula using the approach proposed by Peñaloza (2008). The set of explanations for a query or its probability is computed, also considering Description Logics stronger than $\mathcal{ALC}$.

Statistics and experimental results are also provided, witnessing that the proposed approach is able to deal with large ontologies as well.

Part V on *Learning in Probabilistic DLs* focuses on learning a probabilistic Description Logic KB. Both the probabilities and the structure of the KB can be obtained by abstracting data. Two approaches, EDGE and LEAP, are introduced. Experimental results are discussed and they show that the proposed approach is promising.

Part VI, *Summary and Future Work*, concludes the work with some pointers to future issues and a summary of the contributions of the book. A complete and coherent bibliography is also provided.

Criticisms and conclusions. I have no hesitations in saying that this book represents a significant contribution in the field of probabilistic reasoning in the Semantic Web. I have only some minor criticisms about this book, that I try to summarize in the following:

(i) The text has some typos as well as grammatical errors.
(ii) Concerning the organization of the book, in my opinion, some chapters are too short to be considered as such, for instance, those summarizing Description Logics and Ontologies (Chapters 7 and 8 contain two pages each). In this respect, I believe that Part II could be included in Part I, essentially devoted to motivations and background material.

https://doi.org/10.1017/S1471068421000041 Published online by Cambridge University Press

(iii) Often, formal definitions are introduced before intuitions and examples. This does not help the reader to immediately become aware of the basic concepts underlying the proposed approach.
(iv) I was disappointed by the lack of examples in the description of the reasoning algorithms: I believe that some simple examples should have been provided.

The book is, in general, very well written, it has a clear structure and it is well organized. The background and the state of the art of the topic are described in detail, allowing also not expert readers to understand the matter.

Probabilistic extensions of Description Logics underlying the languages for the Semantic Web are introduced, ranging from the formal basic definitions to the algorithms for reasoning about probabilistic ontologies.

My overall opinion on this book is very good. The book may become a milestone in this field of research and it will be a key reference for future research in reasoning about uncertainty in ontologies. Several works in the recent literature have already been inspired to the approach presented in this book, in particular, those concerning probabilistic extensions of Description Logics of typicality (Pozzato 2019), also applied to model commonsense reasoning (Lieto and Pozzato 2020).

References

Lieto, A. and Pozzato, G. L. 2020. A description logic framework for commonsense conceptual combination integrating typicality, probabilities and cognitive heuristics. *Journal of Experimental and Theoretical Artificial Intelligence 32*, 5, 769–804.

Peñaloza, R. 2008. Automata-based pinpointing for DLs. In *DL2008*, CEUR Workshop Proceedings 353, F. Baader, C. Lutz and B. Motik, Eds. CEUR-WS.org.

Pozzato, G. L. 2019. Typicalities and probabilities of exceptions in nonmotonic description logics. *International Journal of Approximate Reasoning 107*, 81–100.

Riguzzi, F., Bellodi, E., Lamma, E. and Zese, R. 2015a. Probabilistic description logics under the distribution semantics. *Semantic Web 6*, 5, 477–501.

Riguzzi, F., Bellodi, E., Lamma, E. and Zese, R. 2015b. Reasoning with probabilistic ontologies. In *Proceedings of the 24th IJCAI*, Q. Yang and M. Wooldridge, Eds. AAAI Press, 4310–4316.

Sato, T. 1995. A statistical learning method for logic programs with distribution semantics. In *12th International Conference on Logic Programming*, L. Sterling, Ed. MIT Press, 715–729.

Sirin, E., Parsia, B., Grau, B. C., Kalyanpur, A. and Katz, Y. 2007. Pellet: A practical OWL-DL reasoner. *Journal of Web Semantics 5*, 2, 51–53.

Zese, R. 2016. *Probabilistic Reasoning and Learning for the Semantic Web*. Ph.D. thesis, Università degli Studi di Ferrara, Institutional Research Information System UNIFE. URL: `http://hdl.handle.net/11392/2403378`.

GIAN LUCA POZZATO
Dipartimento di Informatica, Università degli Studi di Torino, c.so Svizzera 185, 10149 Torino, Italy
(*e-mail:* gianluca.pozzato@unito.it)

https://doi.org/10.1017/S1471068421000041 Published online by Cambridge University Press