

Statistical Relational Learning: An Inductive Logic Programming Perspective

Luc De Raedt

Institute for Computer Science, Machine Learning Lab,
Albert-Ludwigs-University, Georges-Köhler-Allee, Gebäude 079,
D-79110, Freiburg i. Brg., Germany
deraedt@informatik.uni-freiburg.de

In the past few years there has been a lot of work lying at the intersection of probability theory, logic programming and machine learning [14,18,13,9,6,1,11]. This work is known under the names of statistical relational learning [7,5], probabilistic logic learning [4], or probabilistic inductive logic programming. Whereas most of the existing works have started from a probabilistic learning perspective and extended probabilistic formalisms with relational aspects, I shall take a different perspective, in which I shall start from inductive logic programming and study how inductive logic programming formalisms, settings and techniques can be extended to deal with probabilistic issues. This tradition has already contributed a rich variety of valuable formalisms and techniques, including probabilistic Horn abduction by David Poole, PRISMs by Sato, stochastic logic programs by Muggleton [13] and Cussens [2], Bayesian logic programs [10,8] by Kersting and De Raedt, and Logical Hidden Markov Models [11].

The main contribution of this talk is the introduction of three probabilistic inductive logic programming settings which are derived from the learning from entailment, from interpretations and from proofs settings of the field of inductive logic programming [3]. Each of these settings contributes different notions of probabilistic logic representations, examples and probability distributions. The first setting, probabilistic learning from entailment, is incorporated in the well-known PRISM system [19] and Cussens's Failure Adjusted Maximisation approach to parameter estimation in stochastic logic programs [2]. A novel system that was recently developed and that fits this paradigm is the nFOIL system [12]. It combines key principles of the well-known inductive logic programming system FOIL [15] with the naïve Bayes' approach. In probabilistic learning from entailment, examples are ground facts that should be probabilistically entailed by the target logic program. The second setting, probabilistic learning from interpretations, is incorporated in Bayesian logic programs [10,8], which integrate Bayesian networks with logic programs. This setting is also adopted by [6]. Examples in this setting are Herbrand interpretations that should be a probabilistic model for the target theory. The third setting, learning from proofs [17], is novel. It is motivated by the learning of stochastic context free grammars from tree banks. In this setting, examples are proof trees that should be probabilistically provable from the unknown stochastic logic programs. The sketched settings (and their

instances presented) are by no means the only possible settings for probabilistic inductive logic programming, but still – I hope – provide useful insights into the state-of-the-art of this exciting field.

For a full survey of statistical relational learning or probabilistic inductive logic programming, the author would like to refer to [4], and for more details on the probabilistic inductive logic programming settings to [16], where a longer and earlier version of this contribution can be found.

Acknowledgements

This is joint work with Kristian Kersting. The author would also like to thank Niels Landwehr and Sunna Torge for interesting collaborations on nFOIL and the learning of SLPs, respectively. This work is part of the EU IST FET project APRIL II (Application of Probabilistic Inductive Logic Programming II).

References

1. C. R. Anderson, P. Domingos, and D. S. Weld. Relational Markov Models and their Application to Adaptive Web Navigation. In D. Hand, D. Keim, O. R. Zaïne, and R. Goebel, editors, *Proceedings of the Eighth International Conference on Knowledge Discovery and Data Mining (KDD-02)*, pages 143–152, Edmonton, Canada, 2002. ACM Press.
2. J. Cussens. Loglinear models for first-order probabilistic reasoning. In K. B. Laskey and H. Prade, editors, *Proceedings of the Fifteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI-99)*, pages 126–133, Stockholm, Sweden, 1999. Morgan Kaufmann.
3. L. De Raedt. Logical settings for concept-learning. *Artificial Intelligence*, 95(1):197–201, 1997.
4. L. De Raedt and K. Kersting. Probabilistic Logic Learning. *ACM-SIGKDD Explorations: Special Issue on Multi-Relational Data Mining*, 5(1):31–48, 2003.
5. T. Dietterich, L. Getoor, and K. Murphy, editors. *Working Notes of the ICML-2004 Workshop on Statistical Relational Learning and its Connections to Other Fields (SRL-04)*, 2004.
6. N. Friedman, L. Getoor, D. Koller, and A. Pfeffer. Learning probabilistic relational models. In T. Dean, editor, *Proceedings of the Sixteenth International Joint Conferences on Artificial Intelligence (IJCAI-99)*, pages 1300–1309, Stockholm, Sweden, 1999. Morgan Kaufmann.
7. L. Getoor and D. Jensen, editors. *Working Notes of the IJCAI-2003 Workshop on Learning Statistical Models from Relational Data (SRL-03)*, 2003.
8. K. Kersting and L. De Raedt. Adaptive Bayesian Logic Programs. In C. Rouveirol and M. Sebag, editors, *Proceedings of the Eleventh Conference on Inductive Logic Programming (ILP-01)*, volume 2157 of *LNCS*, Strasbourg, France, 2001. Springer.
9. K. Kersting and L. De Raedt. Bayesian logic programs. Technical Report 151, University of Freiburg, Institute for Computer Science, April 2001.
10. K. Kersting and L. De Raedt. Towards Combining Inductive Logic Programming and Bayesian Networks. In C. Rouveirol and M. Sebag, editors, *Proceedings of the Eleventh Conference on Inductive Logic Programming (ILP-01)*, volume 2157 of *LNCS*, Strasbourg, France, 2001. Springer.

11. K. Kersting, T. Raiko, S. Kramer, and L. De Raedt. Towards discovering structural signatures of protein folds based on logical hidden markov models. In R. B. Altman, A. K. Dunker, L. Hunter, T. A. Jung, and T. E. Klein, editors, *Proceedings of the Pacific Symposium on Biocomputing*, pages 192 – 203, Kauai, Hawaii, USA, 2003. World Scientific.
12. N. Landwehr, K. Kersting, and L. De Raedt. nfoil: Integrating naive bayes and foil. In *Proceedings of the 20th National Conference on Artificial Intelligence*. AAAI Press, 2005.
13. S. H Muggleton. Stochastic logic programs. In L. De Raedt, editor, *Advances in Inductive Logic Programming*. IOS Press, 1996.
14. D. Poole. Probabilistic Horn abduction and Bayesian networks. *Artificial Intelligence*, 64:81–129, 1993.
15. J. R. Quinlan and R. M. Cameron-Jones. Induction of logic programs:FOIL and related systems. *New Generation Computing*, pages 287–312, 1995.
16. L. De Raedt and K. Kersting. Probabilistic inductive logic programming. In *Proceedings of the 15th International Conference on Algorithmic Learning Theory*. Springer, 2004.
17. L. De Raedt, K. Kersting, and S. Torge. Towards learning stochastic logic programs from proof-banks. In *Proceedings of the 20th National Conference on Artificial Intelligence*. AAAI Press, 2005.
18. T. Sato. A Statistical Learning Method for Logic Programs with Distribution Semantics. In L. Sterling, editor, *Proceedings of the Twelfth International Conference on Logic Programming (ICLP-1995)*, pages 715 – 729, Tokyo, Japan, 1995. MIT Press.
19. T. Sato and Y. Kameya. Parameter learning of logic programs for symbolic-statistical modeling. *Journal of Artificial Intelligence Research*, 15:391–454, 2001.