Machine Learning, 57, 203–204, 2004


# Guest Editorial

STAN MATWIN stan@site.uottawa.ca
*SITE—University of Ottawa, Ottawa, Ontario K1N 6N5, Canada*

This issue of the Machine Learning journal is devoted to Inductive Logic Programming (ILP). There is no need to introduce ILP or its goals here, as there are three recent special issues of the Machine Learning journal on this topic: vol. 43 Numbers 1/2, vol. 44 Number 3 and recently vol. 55 Number 2. Similarly to those past special issues, the current one originated in an ILP meeting, i.e. the 12th International Conference on ILP in Sydney, Australia, in July 2002. Following the call for papers for the special issue, 15 submissions were received. As a result of a very rigorous and thorough refereeing process in which each paper received three reviews, four papers were accepted and appear here. All submissions, and particularly the accepted ones, are considerably re-worked and extended with respect to their initial versions that appeared in the ILP 2002 Proceedings.

The paper by Thomas Gartner, John Lloyd and Peter Flach discusses the concept of kernels for structured data. This is a necessary step to bring the highly successful kernel-based approaches (e.g. Support Vector Machines) into the realm of ILP. Specifically, the authors show how to build "good" kernels for structured data expressed in typed higher-order logic (Lloyd, 2001), where a good kernel encodes semantic properties of the domain and provides a mapping in which the data becomes linearly separable. The approach here is syntax-driven, i.e. a structured data definition is decomposed in its constructors, and a kernel for the data is assembled from the kernels for its constructors. The paper proves formally that so defined kernels are correct, i.e. they are positive definite on all basic terms of the same type. The paper then shows empirically that so defined kernels are useful. Experiments with several real-life domains (the musk and the diterpene structure prediction tasks, and a new relational clustering task) report performance comparable (for musk) or better (for diterpene)than the best performance reported in the literature.

The paper "Naïve Bayesian Classification of Structured Data" by Peter Flach and Nicolas Lachiche addresses the problem of combining ILP and Bayesian learning. Two approaches in which Bayesian learning is applied at the relational level are presented. In the first approach, change of representation (propositionalisation) is performed. The original relational problem is re-represented in terms of new, logical attributes. After the fixed (for a given problem) set of new attributes is evaluated, the problem is now represented at the propositional level, and can be handled by a standard Naïve Bayes learner.

The second approach considers an upgrade of Naïve Bayesian learning by evaluating parameterized probability distributions in the space of structured objects. This is accomplished by recursively analyzing the structure of data treated as collections. Their probability distribution is obtained by taking into account probability distributions of the elements of the collection. Both approaches are compared empirically on synthetic and real-life datasets.

The authors report that the latter approach tends to perform better, given enough training data.

The paper "Integrating Guidance into Relational Reinforcement Learning" by Kurt Driessens and Saso Dzeroski discusses the challenges of reinforcement learning at the relational level. Following the approach of Dzeroski, De Raedt, and Blockeel (1998) instead of storing Q values in a table, a Q function is learned by a relational learner. Two such learners are experimented with: TG and RLL. The problem of sparseness of awards is particularly acute in Relational Reinforcement Learning setting as the expressive relational representation results in state spaces exponentially larger than spaces obtained when propositional learners learn the Q function. One approach to address this sparseness problem is to use guidance (i.e. an informed policy for the choice of the action in a state) to speed up the convergence of reinforcement learning. The guidance represents the knowledge of the problem states and actions. This paper investigates experimentally the effect of supplying guidance. Results in three domains (games Tetris and Digger, and the blocks world) indicate, among others, that performance of relational reinforcement learning improves with guidance, in terms of both the quality of the solution and the speed of convergence.

The paper "Compact Representation of Knowledge Bases in Inductive Logic Programming" by Jan Struyf, Jan Ramon, Maurice Bruynooghe, Sofie Verbaeten and Hendrik Blockeel addresses some important ILP representation and implementation problems. Specifically, the authors investigate techniques resulting in more compact representation and storage of examples by ILP systems. The goal is to reduce the duplication of the same knowledge represented in many examples. The paper defines a meta-language which dynamically constructs theories (example sets). The authors then experiment with three methods of example representation: the classical "monolithic" approach when all the examples are one logic program, the learning from interpretations approach where each example is represented by a separate theory, and the "generative" approach presented in this paper. The authors then introduce a mechanism based on the so called knowledge base graph, which detects references from one theory to another, and avoids duplication of the knowledge shared by both theories. Empirical results obtained from three domains indicate reduced memory requirements.

Interest in this Special Issue and the number of submissions received are strong evidence of the liveliness and dynamism of Inductive Logic Programming. The papers appearing here, and their strong connectedness to the subfields of machine Learning that attract most interest today—kernel methods, probabilistic learning, reinforcement learning—bear evidence of fertile cross-breeding between ILP and the other subfields of Machine Learning.

Last but not least, the papers bring in an interesting mix of application areas beyond the standard mutagenesis benchmark standard in ILP—spatial clustering, drug design, chemical structure analysis, and real-life computer games.

## References

Dzeroski, S., De Raedt, L., & Blockeel, H. (1998). Relational reinforcement learning. In D. Page (Ed.). *Proceedings of the 8th International Conference on Inductive Logic Programming*, Vol. 1446 of Lecture Notes in Artificial Intelligence (pp. 11–22). Springer-Verlag.

Lloyd, J. (2001). Knowledge Representation, Computation, and Learning in Higher-order Logic. Available at http://csl.anu.edu.au/~jwl