Mach Learn (2008) 73: 1–2
DOI 10.1007/s10994-008-5078-2

Guest editors' introduction: special issue on inductive logic programming (ILP-2007)

Hendrik Blockeel · Jude Shavlik · Prasad Tadepalli

Published online: 25 July 2008

The 17th International Conference on Inductive Logic Programming was held in Corvallis, Oregon, June 19 to 21, 2007, collocated with the ICML-2007, the 24th International Conference on Machine Learning. The conference program featured several invited talks, plenary paper presentations, poster presentations, and, jointly with ICML, a panel discussion on the future of structured machine learning. Much of the presented work has been included in the book *Inductive Logic Programming: 17th International Conference, ILP 2007, Corvallis, OR, USA, June 19–21, 2007, Revised Selected Papers*, published by Springer Verlag as Volume 4894 of the *Lecture Notes in Artificial Intelligence* series.

Like previous special issues on ILP conferences, this issue contains a small selection of articles describing work presented at the conference. In contrast to most previous years, the issue does not contain extended versions of papers that have already been included in the proceedings. Instead, authors were invited to extend their paper into a journal article and submit it to this journal, as an alternative to having the full conference paper included in the proceedings. Thus eliminating concerns about possible redundancy between the conference and journal publications, the conference chairs hoped to create a faster journal publication track than was otherwise possible.

This special issue contains four articles. We selected three of these from the regular papers presented at the conference, based on their reviews and presentations. Reflecting the current emphasis of the field, all these papers address combining relational representations

H. Blockeel
Department of Computer Science, Katholieke Universiteit Leuven, Leuven, Belgium
e-mail: hendrik.blockeel@cs.kuleuven.be

J. Shavlik
Department of Computer Science, University of Wisconsin, Madison, WI, USA
e-mail: shavlik@cs.wisc.edu

P. Tadepalli (✉)
Computer Science Department, Oregon State University, Corvallis, OR, USA
e-mail: tadepall@cs.orst.edu

with probabilistic models and statistical inference. To give a broader introduction to this developing area and summarize the future directions, we invited the panelists to write a paper elaborating their panel presentations. All articles including the panel paper were peer-reviewed in their journal form before being accepted.

In "Structured Machine Learning: the Next Ten Years", Dietterich, Domingos, Getoor, Muggleton and Tadepalli present their views on the history, present, and near future of Inductive Logic Programming, and more broadly, Structured Machine Learning, an umbrella term for Inductive Logic Programming and closely related fields such as Statistical Relational Learning. The authors provide a stimulating high-level overview of the current status of the field and identify many exciting and challenging topics for further research in this very active area. They also provide a valuable bibliography for newcomers to the field.

One of the central problems of working with an expressive language that includes relations and probabilities is that of intractability of learning and reasoning. This problem of intractability is addressed head-on by the paper on "Learning to Assign Degrees of Belief in Relational Domains" by Koriche. This paper adapts the learning to reason (L2R) framework of Khardon and Roth to learn to answer probabilistic queries in relational domains. In contrast with the traditional view of knowledge representation and inference in languages such as Markov logic networks, where an expressive knowledge base is created independent of its intended use, in the L2R framework, the agent learns to answer particular types of queries about the domain. This approach simplifies the knowledge representation and reasoning requirements, making it possible to design efficient learning and inference algorithms tuned to the classes of queries the learner is likely to encounter. Koriche gives polynomial-time inference algorithms for some restricted classes of decomposable quantified conjunctive queries. In contrast, probabilistic inference in Markov logic networks is #EXP-hard in the size of the formulas even in the case of finite domains.

Most work on learning probabilistic models relies on examples labeled deterministically. In many real world domains, the examples are often noisy, uncertain, and ambiguously labeled. In their paper "Learning Probabilistic Logic Models from Probabilistic Examples", Chen, Muggleton, and Santos give a possible worlds semantics to Stochastic Logic Programs (SLPs), and use it to model an application in toxicology for rats that combines the analysis of Nuclear Magnetic Resonance (NMR) data with rich background knowledge. Their results using both their approach and another modeling tool called PRISM (for PRogramming In Statistical Modeling) show that probabilistic logic models learned from probabilistic examples are significantly more accurate than those learned from deterministic examples.

Decomposing the causes of prediction error into two components—bias and variance—has proven to be an insightful way to understand the behavior of supervised learning algorithms. In their paper Neville and Jensen extend the classic bias-variance analysis to account for the impact of approximation during the inference process that commonly accompanies statistical models, especially when performing collective inference. They perform empirical studies that demonstrate that inference can be a substantial contributor to errors in prediction and also provide insight into common approaches to statistical-relational learning.

Together, the articles included in this special issue offer a representative view of contemporary research on inductive logic programming. We wish the reader a pleasant and stimulating reading.