Journal of Machine Learning Research 4 (2003) 413-414 Published 8/03

Introduction to the Special Issue on Inductive Logic Programming

James Cussens JC@CS.YORK.AC.UK
Alan M. Frisch FRISCH@CS.YORK.AC.UK
Artificial Intelligence Group
Department of Computer Science
University of York, Heslington
York YO10 5DD, United Kingdom

This special issue has its origins in the Tenth International Conference on Inductive Logic Programming (ILP-2000), which was held in London during July 2000. Authors who presented at ILP-2000 were invited to submit revised and expanded papers to this special issue, but the call was also extended to those who had presented ILP papers at other conferences in that year. The four papers in this special issue were selected from the eleven papers that were submitted.

Since its inception around the early 1990s, the ILP community has striven to make progress on three fronts: theory, implementation, and applications. This special issue continues this tradition with a paper on each of these topics plus a position paper.

The position paper arose out of a plenary talk delivered by David Page to a joint session of ILP-2000 and a large collocated conference—the International Conference on Computational Logic (CL-2000). The talk successfully inspired researchers in ILP as well as those in other areas of computational logic.

The subsequent position paper by David Page and Ashwin Srinivasan, which appears here, is also accessible to researchers outside of ILP and for all readers it provides a carefully judged view of the field. Entitled a "A Short Look Back and a Longer Look Forward," the paper provides just that. The look back briefly reviews several of the most successful applications of ILP, which it uses to illustrate some of the strengths of ILP. The look forward identifies and discusses at length five shortcomings of current ILP technology that need to be overcome (with assistance from researchers in other areas) to extend the utility of ILP.

Theoretical work in ILP and, indeed, much of computer science is typically deductive; from the definitions of computational structures one mathematically proves that the structures have certain properties. The theoretical work reported in the paper by Marco Botta, Attilio Giordana, Lorenza Saitta and Michèle Sebag takes a different form, which is more familiar in the empirical sciences, where theories that explain observed phenomena are presented. In particular, their work explains several phenomena in the behaviour of relational learning systems by considering the location of a problem instance with respect to a particular phase transition. Whereas the difficulty of solving an instance of a learning problem with a relational learner has traditionally been considered as a function of the size of the instance, this paper argues that it is better considered as a function of its location with respect to the phase transition.

The paper by Vitor Santos Costa and his eight co-authors focuses squarely on implementation. In the introduction the paper notes that: "Like many other algorithms in the field of machine learning, ILP algorithms construct 'hypotheses' for data by performing a search through a large space. Such a search typically involves generating and then testing the quality of candidates." Their paper concentrates on the crucially important but under-analysed problem of speeding up the evaluation of candidate rules in an ILP search. This issue is thoroughly analysed both theoretically and empirically, using Srinivasan's Aleph algorithm.

Vincent Claveau, Pascale Sébillot, Cécile Fabre and Pierrette Bouillon address an application in natural language processing. They use ILP—specifically the Aleph algorithm—to automatically extract, from a morpho-syntactically and semantically tagged corpus, noun-verb pairs whose elements are semantically linked. Such pairs allow a match, for example, between a query *disk store* and the text *to sell disks* and thus may be used to improve information retrieval. Two particularly important issues are: "the learning efficiency that is required to be able to deal with all the available contextual information, and [the production of] linguistically meaningful rules."

We would like to thank all those who have contributed to this special issue: the authors, the reviewers and Leslie Pack Kaelbling and David Cohn at JMLR.