

Lecture Notes in Artificial Intelligence

2583

Subseries of Lecture Notes in Computer Science

Edited by J. G. Carbonell and J. Siekmann

Lecture Notes in Computer Science

Edited by G. Goos, J. Hartmanis, and J. van Leeuwen

**Springer**

*Berlin*

*Heidelberg*

*New York*

*Barcelona*

*Hong Kong*

*London*

*Milan*

*Paris*

*Tokyo*

Stan Matwin Claude Sammut (Eds.)

# Inductive Logic Programming

12th International Conference, ILP 2002  
Sydney, Australia, July 9-11, 2002  
Revised Papers



Springer

## Series Editors

Jaime G. Carbonell, Carnegie Mellon University, Pittsburgh, PA, USA  
Jörg Siekmann, University of Saarland, Saarbrücken, Germany

## Volume Editors

Stan Matwin

University of Ottawa, School of Information Technology and Engineering  
800 King Edward Ave., Ottawa, ON, K1N 6N5 Canada  
E-mail: stan@site.uottawa.ca

Claude Sammut

University of New South Wales, School of Computer Science and Engineering  
Sydney, NSW 2052, Australia  
E-mail: claudes@unsw.edu.au

Cataloging-in-Publication Data applied for

A catalog record for this book is available from the Library of Congress

Bibliographic information published by Die Deutsche Bibliothek

Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliographie;  
detailed bibliographic data is available in the Internet at <<http://dnd.ddb.de>>.

CR Subject Classification (1998): I.2.3, I.2, D.1.6, F.4.1, F.2.2, H.2.8

ISSN 0302-9743

ISBN 3-540-00567-6 Springer-Verlag Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer-Verlag Berlin Heidelberg New York,  
a member of BertelsmannSpringer Science+Business Media GmbH

<http://www.springer.de>

© Springer-Verlag Berlin Heidelberg 2003  
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Olgun Computergraphik  
Printed on acid-free paper SPIN: 10872409 06/3142 5 4 3 2 1 0

# Preface

The Twelfth International Conference on Inductive Logic Programming was held in Sydney, Australia, July 9–11, 2002. The conference was colocated with two other events, the Nineteenth International Conference on Machine Learning (ICML 2002) and the Fifteenth Annual Conference on Computational Learning Theory (COLT 2002).

Started in 1991, Inductive Logic Programming is the leading annual forum for researchers working in Inductive Logic Programming and Relational Learning.

Continuing a series of international conferences devoted to Inductive Logic Programming and Relational Learning, ILP 2002 was the central event in 2002 for researchers interested in learning relational knowledge from examples.

The Program Committee, following a resolution of the Community Meeting in Strasbourg in September 2001, took upon itself the issue of the possible change of the name of the conference. Following an extended e-mail discussion, a number of proposed names were subjected to a vote. In the first stage of the vote, two names were retained for the second vote. The two names were: Inductive Logic Programming, and Relational Learning. It had been decided that a 60% vote would be needed to change the name; the result of the vote was 57% in favor of the name Relational Learning. Consequently, the name Inductive Logic Programming was kept.

In order to facilitate submissions, the Program Committee decided to extend the deadline and at the same time publish the proceedings after the conference (the extended deadline was cutting into the time needed by the publisher to produce the proceedings from the camera-ready copy). At the extended deadline there were 45 submissions to the conference, which resulted in 24 accepted papers, of which two withdrew and 22 were presented.

The conference program consisted of two invited talks, the presentations of accepted technical papers, a Work-in-Progress session, and a panel. The invited talk by John Shawe-Taylor introduced the concept of the Set Covering Machine and its relations to ILP. The second invited talk, by Sašo Džeroski, joint with ICML, reviewed the goals of ILP and showed how they fruitfully support Computational Scientific Discovery. There was a Work-in-Progress session with 10 papers, whose abstracts were published in a separate booklet. Furthermore, there was a panel “If I had to start a Ph.D. in ILP today, what would my topic be?”, chaired by S. Matwin, with the following panelists: I. Bratko, P. Flach, J. Lloyd, S. Muggleton, C. Rouveirol, P. Tadepalli, and L. Saitta.

The ILP 2002 technical program, as with previous editions of the conference, covered a broad range of topics, from implementation techniques to applications, and from kernels for structured data to associations for relational data.

ILP 2002 would not have been possible without the generous support of ILPNet2 to European participants. This is kindly acknowledged. Several volunteers assisted the Program Chairs in different phases of the organization of

the conference. We are very grateful to Erick Alphonse from Université Paris XI for his help in organizing the distribution of papers to reviewers, to Svetlana Kiritchenko from the University of Ottawa for her help in producing the camera-ready version of these proceedings, and to Michael Bain from the University of New South Wales for his help in local arrangements.

November 2002

Stan Matwin  
Claude Sammut

# Organization

Program Chairs: Stan Matwin, University of Ottawa, Canada  
Claude Sammut, University of New South Wales, Australia

## Program Committee

H. Blockeel (Belgium)	L. Holder (USA)	L. De Raedt (Germany)
J.F. Boulicaut (France)	T. Horvath (Germany)	D. Roth (USA)
I. Bratko (Slovenia)	D. Jensen (USA)	C. Rouveirol (France)
J. Cussens (UK)	R. Khardon (USA)	L. Saitta (Italy)
T. Dietterich (USA)	J.-U. Kietz (Switzerland)	M. Sebag (France)
S. Džeroski (Slovenia)	S. Kramer (Germany)	A. Srinivasan (UK)
F. Esposito (Italy)	N. Lavrač (Slovenia)	P. Tadepalli (USA)
P. Flach (UK)	R. Mooney (USA)	S. Wrobel (Germany)
J. Fürnkranz (Austria)	S. Muggleton (UK)	A. Yamamoto (Japan)
K. Furukawa (Japan)	D. Page (USA)	G. Zaverucha (Brazil)
L. Getoor (USA)	B. Pfahringer (NZ)	J.-D. Zucker (France)

# Table of Contents

## Contributed Papers

Propositionalization for Clustering Symbolic Relational Descriptions . . . . .	1
<i>Isabelle Bournaud, Mélanie Courtine, and Jean-Daniel Zucker</i>	
Efficient and Effective Induction of First Order Decision Lists . . . . .	17
<i>Mary Elaine Califf</i>	
Learning with Feature Description Logics . . . . .	32
<i>Chad M. Cumby and Dan Roth</i>	
An Empirical Evaluation of Bagging in Inductive Logic Programming . . . . .	48
<i>Inês de Castro Dutra, David Page, Vítor Santos Costa, and Jude Shavlik</i>	
Kernels for Structured Data . . . . .	66
<i>Thomas Gärtner, John W. Lloyd, and Peter A. Flach</i>	
Experimental Comparison of Graph-Based Relational Concept Learning with Inductive Logic Programming Systems . . . . .	84
<i>Jesus A. Gonzalez, Lawrence B. Holder, and Diane J. Cook</i>	
Autocorrelation and Linkage Cause Bias in Evaluation of Relational Learners . . . . .	101
<i>David Jensen and Jennifer Neville</i>	
Learnability of Description Logic Programs . . . . .	117
<i>Jörg-Uwe Kietz</i>	
1BC2: A True First-Order Bayesian Classifier . . . . .	133
<i>Nicolas Lachiche and Peter A. Flach</i>	
RSD: Relational Subgroup Discovery through First-Order Feature Construction . . . . .	149
<i>Nada Lavrač, Filip Železný, and Peter A. Flach</i>	
Mining Frequent Logical Sequences with SPIRIT-L <sup>o</sup> G . . . . .	166
<i>Cyrille Masson and François Jacquenet</i>	
Using Theory Completion to Learn a Robot Navigation Control Program . . . . .	182
<i>Steve Moyle</i>	
Learning Structure and Parameters of Stochastic Logic Programs . . . . .	198
<i>Stephen Muggleton</i>	



A Novel Approach to Machine Discovery: Genetic Programming and Stochastic Grammars .....	207
<i>Alain Ratle and Michèle Sebag</i>	
Revision of First-Order Bayesian Classifiers .....	223
<i>Kate Revoredo and Gerson Zaverucha</i>	
The Applicability to ILP of Results Concerning the Ordering of Binomial Populations .....	238
<i>Ashwin Srinivasan</i>	
Compact Representation of Knowledge Bases in ILP.....	254
<i>Jan Struyf, Jan Ramon, and Hendrik Blockeel</i>	
A Polynomial Time Matching Algorithm of Structured Ordered Tree Patterns for Data Mining from Semistructured Data .....	270
<i>Yusuke Suzuki, Kohtaro Inomae, Takayoshi Shoudai, Tetsuhiro Miyahara, and Tomoyuki Uchida</i>	
A Genetic Algorithms Approach to ILP .....	285
<i>Alireza Tamaddon-Nezhad and Stephen Muggleton</i>	
Experimental Investigation of Pruning Methods for Relational Pattern Discovery .....	301
<i>Irene Weber</i>	
Noise-Resistant Incremental Relational Learning Using Possible Worlds ...	317
<i>James Westendorp</i>	
Lattice-Search Runtime Distributions May Be Heavy-Tailed .....	333
<i>Filip Železný, Ashwin Srinivasan, and David Page</i>	
<b>Invited Talk Abstracts</b>	
Learning in Rich Representations: Inductive Logic Programming and Computational Scientific Discovery .....	346
<i>Sašo Džeroski</i>	
<b>Author Index</b> .....	351