

Lecture Notes in Artificial Intelligence 2157

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

Céline Rouveirol Michèle Sebag (Eds.)

Inductive Logic Programming

11th International Conference, ILP 2001
Strasbourg, France, September 9-11, 2001
Proceedings



Springer

Series Editors

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

Volume Editors

Céline Rouveirol
Université Paris Sud, LRI
bât. 490, 91405 Orsay Cedex, France
E-mail: Celine.Rouveirol@lri.fr

Michèle Sebag
Ecole Polytechnique, LMS
91128 Palaiseau, France
E-mail: Michele.Sebag@polytechnique.fr

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Inductive logic programming : 11th international conference ; proceedings /
ILP 2001, Strasbourg, France, September 9 - 11, 2001. Céline Rouveirol ;
Michèle Sebag (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Hong
Kong ; London ; Milan ; Paris ; Tokyo : Springer, 2001
(Lecture notes in computer science ; Vol. 2157 : Lecture notes in
artificial intelligence)
ISBN 3-540-42538-1

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

ISBN 3-540-42538-1 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 2001
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Steingräber Satztechnik GmbH, Heidelberg
Printed on acid-free paper SPIN: 10840266 06/3142 5 4 3 2 1 0

Preface

The 11th international conference on Inductive Logic Programming, ILP 2001, was held in Strasbourg, France, September 9–11, 2001. ILP 2001 was co-located with the 3rd international workshop on Logic, Learning, and Language (LLL 2001), and nearly co-located with the joint 12th European Conference on Machine Learning (ECML 2001) and 5th European conference on Principles and Practice of Knowledge Discovery in Databases (PKDD 2001).

Continuing a series of international conferences devoted to Inductive Logic Programming and Relational Learning, ILP 2001 is the central annual event for researchers interested in learning structured knowledge from structured examples and background knowledge.

One recent one major challenge for ILP has been to contribute to the exponential emergence of Data Mining, and to address the handling of multi-relational databases. On the one hand, ILP has developed a body of theoretical results and algorithmic strategies for exploring relational data, essentially but not exclusively from a supervised learning viewpoint. These results are directly relevant to an efficient exploration of multi-relational databases.

On the other hand, Data Mining might require specific relational strategies to be developed, especially with regard to the scalability issue. The near-colocation of ILP 2001 with ECML 2001-PKDD 2001 was an incentive to increase cross-fertilization between the ILP relational *savoir-faire* and the new problems and learning goals addressed and to be addressed in Data Mining.

Thirty-seven papers were submitted to ILP, among which twenty-one were selected and appear in these proceedings. Several – non-disjoint – trends can be observed, along an admittedly subjective clustering.

On the theoretical side, a new mode of inference is proposed by K. Inoue, analog to the open-ended mode of Bayesian reasoning (where the frontier between induction and abduction wanes). New learning refinement operators are proposed by L. Badea, while R. Otero investigates negation-handling settings. Rule stretching (M. Eineborg and H. Boström) can also be considered a new inductive-deductive operator.

Several hybrid frameworks are proposed, either bridging the gap between ILP and other learning paradigms, e.g. Bayesian inference (K. Kersting and L. De Raedt), Neural Nets (R. Babilio, G. Zaverucha, and V. C. Barbosa) or Feature Selection (T. Ozabaki and K. Furukawa) – or exploiting other search paradigms, e.g. Constraint Satisfaction (J. Maloberti) or Genetic Algorithms (A. Braud and C. Vrain), to address particular ILP tasks.

Among the tasks addressed, changes of representation take an increasing importance, ranging from propositionalization (A. Braud and C. Vrain, already

mentioned) and construction of structural features (S. Kramer), to aggregation-based transformations (M.-A. Krogel and S. Wrobel).

Committee-based and statistical machine learning interestingly pervade ILP through boosting-based approaches (S. Hoche and S. Wrobel), positive-only learning (F. Zelezny), and transductive inference (M. Eineborg and H. Bostrom, already mentioned). Last but not least, an efficient relational cross-validation procedure is proposed by J. Struyf and H. Blockeel.

The application papers deserve a special mention as they demonstrate when and how relational representations can make the difference. Language-related applications range from Natural Language (M. Nepil) to XML documents (A. Yamamoto, K. Ito, A. Ishino, and H. Arimura), and shell logs (N. Jacobs and H. Blockeel). Bio-informatics offers many challenging relational problems (A. Karwath and R. D. King), in the spirit of the founding ILP application, i.e. the mutagenesis problem¹. Other applications are concerned with medical control (R. Quiniou, M.-O. Cordier, G. Carrault, and F. Wang) and spatial data mining (D. Malerba and F. A. Lisi).

The invited talks, one joint conference with LLL, given by D. Roth, Univ of Illinois, USA and one by H. T. T. Toivonen, Nokia, Finland, described the challenges a in two of the hottest fields for Machine Learning and ILP: Natural Language, and the Genome².

We wish to thank all researchers who submitted their papers to ILP 2001, all external referees whose kind help was very welcome, and the members of the Program Committee for their commitment to making ILP 2001 an open and lively high scientific venue.

July 2001

Céline Rouveirol and Michèle Sebag
Program Chairs

¹ R.D. King, A. Srinivasan, and M.J.E. Sternberg, *Relating chemical activity to structure: an examination of ILP successes*, New Gen. Comput., 13, 1995.

² Available at: <http://www.lri.fr/ilp2001/>

Organization

Local chair: Nicolas Lachiche (Univ. Robert Schuman -
LSIT, Strasbourg, France)
Web chair: Fabien Torre (Univ. Lille 3, France)

Program Committee

H. Blockeel, Belgium	A. Frisch, UK	B. Pfahringer, NZ
J.-F. Boulicaut, France	K. Furukawa, Japan	L. Saitta, Italy
H. Boström, Sweden	J-U. Kietz, Switzerland	C. Sammut, Australia
I. Bratko, Slovenia	N. Lavrač, Slovenia	A. Srinivasan, UK
J. Cussens, UK	J. Lloyd, Australia	P. Tadepalli, USA
L. Dehaspe, Belgium	S. Kramer, Germany	S. Wrobel, Germany
S. Džeroski, Slovenia	S. Matwin, Canada	A. Yamamoto, Japan
F. Esposito, Italy	S. Muggleton, UK	
P. Flach, UK	D. Page, USA	

External Referees

A. Borgida, Ch. Giraud-Carrier, J.-F. Puget, L. de Raedt, V. Ventos, C. Vrain.

Sponsoring Institutions

ILPNet2,
IUT Robert Schuman,
Université Robert Schuman,
Université Louis Pasteur,
Conseil général du Bas-Rhin
Région Alsace,
and the city of Strasbourg.

Table of Contents

A Refinement Operator for Theories	1
<i>Liviu Badea</i>	
Learning Logic Programs with Neural Networks	15
<i>Rodrigo Basilio, Gerson Zaverucha, and Valmir C. Barbosa</i>	
A Genetic Algorithm for Propositionalization	27
<i>Agnès Braud and Christel Vrain</i>	
Classifying Uncovered Examples by Rule Stretching	41
<i>Martin Eineborg and Henrik Boström</i>	
Relational Learning Using Constrained Confidence-Rated Boosting	51
<i>Susanne Hoche and Stefan Wrobel</i>	
Induction, Abduction, and Consequence-Finding	65
<i>Katsumi Inoue</i>	
From Shell Logs to Shell Scripts	80
<i>Nico Jacobs and Hendrik Blockeel</i>	
An Automated ILP Server in the Field of Bioinformatics	91
<i>Andreas Karwath and Ross D. King</i>	
Adaptive Bayesian Logic Programs	104
<i>Kristian Kersting and Luc De Raedt</i>	
Towards Combining Inductive Logic Programming with Bayesian Networks	118
<i>Kristian Kersting and Luc De Raedt</i>	
Demand-Driven Construction of Structural Features in ILP	132
<i>Stefan Kramer</i>	
Transformation-Based Learning Using Multirelational Aggregation	142
<i>Mark-A. Krogel and Stefan Wrobel</i>	
Discovering Associations between Spatial Objects : An ILP Application . . .	156
<i>Donato Malerba and Francesca A. Lisi</i>	
θ -Subsumption in a Constraint Satisfaction Perspective	164
<i>Jérôme Maloberti and Michèle Sebag</i>	
Learning to Parse from a Treebank: Combining TBL and ILP	179
<i>Miloslav Nepil</i>	

Induction of Stable Models 193
Ramón P. Otero

Application of Pruning Techniques for Propositional Learning to Progol . 206
Tomonobu Ozaki and Koichi Furukawa

Application of ILP to Cardiac Arrhythmia Characterization
for Chronicle Recognition 220
René Quiniou, Marie-Odile Cordier, Guy Carrault, and Feng Wang

Efficient Cross-Validation in ILP 228
Jan Struyf and Hendrik Blockeel

Modelling Semi-structured Documents with Hedges
for Deduction and Induction..... 240
Akihiro Yamamoto, Kimihito Ito, Akira Ishino, and Hiroki Arimura

Learning Functions from Imperfect Positive Data 248
Filip Železný

Author Index 261