

Lecture Notes in Artificial Intelligence 2842

Edited by J. G. Carbonell and J. Siekmann

Subseries of Lecture Notes in Computer Science

Ricard Gavaldà Klaus P. Jantke
Eiji Takimoto (Eds.)

Algorithmic Learning Theory

14th International Conference, ALT 2003
Sapporo, Japan, October 17-19, 2003
Proceedings

Series Editors

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

Volume Editors

Ricard Gavaldà
Technical University of Catalonia
Department of Software (LSI)
Jordi Girona Salgado 1-3, 08034 Barcelona, Spain
E-mail: gavald@lsi.upc.es

Klaus P. Jantke
Deutsches Forschungszentrum für Künstliche Intelligenz GmbH
Im Stadtwald, Geb. 43.8, 66125 Saarbrücken, Germany
E-mail: jantke@dfki.de

Eiji Takimoto
Tohoku University
Graduate School of Information Sciences
Sendai 980-8579, Japan
E-mail: t2@ecei.tohoku.ac.jp

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 Nationalbibliografie;
detailed bibliographic data is available in the Internet at <<http://dnb.ddb.de>>.

CR Subject Classification (1998): I.2.6, I.2.3, F.1, F.2, F.4.1, I.7

ISSN 0302-9743

ISBN 3-540-20291-9 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 PTP-Berlin GmbH
Printed on acid-free paper SPIN: 10963852 06/3142 5 4 3 2 1 0

Preface

This volume contains the papers presented at the 14th Annual Conference on Algorithmic Learning Theory (ALT 2003), which was held in Sapporo (Japan) during October 17–19, 2003. The main objective of the conference was to provide an interdisciplinary forum for discussing the theoretical foundations of machine learning as well as their relevance to practical applications. The conference was co-located with the 6th International Conference on Discovery Science (DS 2003).

The volume includes 19 technical contributions that were selected by the program committee from 37 submissions. It also contains the ALT 2003 invited talks presented by Naftali Tishby (Hebrew University, Israel) on “Efficient Data Representations that Preserve Information,” by Thomas Zeugmann (University of Lübeck, Germany) on “Can Learning in the Limit be Done Efficiently?,” and by Genshiro Kitagawa (Institute of Statistical Mathematics, Japan) on “Signal Extraction and Knowledge Discovery Based on Statistical Modeling” (joint invited talk with DS 2003). Furthermore, this volume includes abstracts of the invited talks for DS 2003 presented by Thomas Eiter (Vienna University of Technology, Austria) on “Abduction and the Dualization Problem” and by Akihiko Takano (National Institute of Informatics, Japan) on “Association Computation for Information Access.” The complete versions of these papers were published in the DS 2003 proceedings (Lecture Notes in Artificial Intelligence Vol. 2843).

ALT has been awarding the *E. Mark Gold Award* for the most outstanding paper by a student author since 1999. This year the award was given to Sandra Zilles for her paper “Intrinsic Complexity of Uniform Learning.”

This conference was the 14th in a series of annual conferences established in 1990. Continuation of the ALT series is supervised by its steering committee, consisting of: Thomas Zeugmann (Univ. of Lübeck, Germany), Chair, Arun Sharma (Univ. of New South Wales, Australia), Co-chair, Naoki Abe (IBM T.J. Watson Research Center, USA), Klaus Peter Jantke (DFKI, Germany), Phil Long (National Univ. of Singapore), Hiroshi Motoda (Osaka Univ., Japan), Akira Maruoka (Tohoku Univ., Japan), Luc De Raedt (Albert-Ludwigs-Univ., Germany), Takeshi Shinohara (Kyushu Institute of Technology, Japan), and Osamu Watanabe (Tokyo Institute of Technology, Japan).

We would like to thank all individuals and institutions who contributed to the success of the conference: the authors for submitting papers, the invited speakers for accepting our invitation and lending us their insight into recent developments in their research areas, as well as the sponsors for their generous financial support.

Furthermore, we would like to express our gratitude to all program committee members for their hard work in reviewing the submitted papers and participating in on-line discussions. We are also grateful to the external referees whose reviews made a considerable contribution to this process.

We are also grateful to the DS 2003 Chairs Yuzuru Tanaka (Hokkaido University, Japan), Gunter Grieser (Technical University of Darmstadt, Germany) and Akihiro Yamamoto (Hokkaido University, Japan) for their efforts in coordinating with ALT 2003, and to Makoto Haraguchi and Yoshiaki Okubo (Hokkaido University, Japan) for their excellent work on the local arrangements. Last but not least, Springer-Verlag provided excellent support in preparing this volume.

August 2003

Ricard Gavaldà
Klause P. Jantke
Eiji Takimoto

Organization

Conference Chair

Klaus P. Jantke DFKI GmbH Saarbrücken, Germany

Program Committee

Ricard Gavaldà (Co-Chair) Tech. Univ. of Catalonia, Spain
Eiji Takimoto (Co-Chair) Tohoku Univ., Japan
Hiroki Arimura Kyushu Univ., Japan
Shai Ben-David Technion, Israel
Nicolò Cesa-Bianchi Univ. di Milano, Italy
Nello Cristianini UC Davis, USA
François Denis LIF, Univ. de Provence, France
Kouichi Hirata Kyutech, Japan
Sanjay Jain Nat. Univ. Singapore, Singapore
Stephen Kwek Univ. Texas, San Antonio, USA
Phil Long Genome Inst. Singapore, Singapore
Yasubumi Sakakibara Keio Univ., Japan
Rocco Servedio Columbia Univ., USA
Hans-Ulrich Simon Ruhr-Univ. Bochum, Germany
Frank Stephan Univ. Heidelberg, Germany
Christino Tamon Clarkson Univ., USA

Local Arrangements

Makoto Haraguchi (Chair) Hokkaido Univ., Japan
Yoshiaki Okubo Hokkaido Univ., Japan

Subreferees

Kazuyuki Amano Joshua Goodman
Dana Angluin Colin de la Higuera
Tijl De Bie Hiroki Ishizaka
Laurent Brehelin Jeffrey Jackson
Christian Choffrut Satoshi Kobayashi
Pedro Delicado Jean-Yves Marion
Claudio Gentile Andrei E. Romashchenko
Rémi Gilleron Hiroshi Sakamoto
Sally Goldman Kengo Sato

VIII Organization

Dale Schuurmans
Chema Sempere
Shinichi Shimozono
Takeshi Shinohara
Robert Sloan

Lee Wee Sun
Hisao Tamaki
Marc Tommasi
Takashi Yokomori

Sponsoring Institutions

The Japanese Ministry of Education, Culture, Sports, Science and Technology
The Suginome Memorial Foundation, Japan

Table of Contents

INVITED PAPERS

Abduction and the Dualization Problem	1
<i>Thomas Eiter</i>	
Signal Extraction and Knowledge Discovery Based on Statistical Modeling	3
<i>Genshiro Kitagawa</i>	
Association Computation for Information Access	15
<i>Akihiko Takano</i>	
Efficient Data Representations That Preserve Information	16
<i>Naftali Tishby</i>	
Can Learning in the Limit Be Done Efficiently?	17
<i>Thomas Zeugmann</i>	

REGULAR CONTRIBUTIONS

Inductive Inference

Intrinsic Complexity of Uniform Learning	39
<i>Sandra Zilles</i>	
On Ordinal VC-Dimension and Some Notions of Complexity	54
<i>Eric Martin, Arun Sharma, Frank Stephan</i>	
Learning of Erasing Primitive Formal Systems from Positive Examples...	69
<i>Jin Uemura, Masako Sato</i>	
Changing the Inference Type – Keeping the Hypothesis Space	84
<i>Frank Balbach</i>	

Learning and Information Extraction

Robust Inference of Relevant Attributes	99
<i>Jan Arpe, Rüdiger Reischuk</i>	
Efficient Learning of Ordered and Unordered Tree Patterns with Contractible Variables	114
<i>Yusuke Suzuki, Takayoshi Shoudai, Satoshi Matsumoto, Tomoyuki Uchida, Tetsuhiro Miyahara</i>	

Learning with Queries

On the Learnability of Erasing Pattern Languages in the Query Model... 129 <i>Steffen Lange, Sandra Zilles</i>	129
Learning of Finite Unions of Tree Patterns with Repeated Internal Structured Variables from Queries 144 <i>Satoshi Matsumoto, Yusuke Suzuki, Takayoshi Shoudai, Tetsuhiro Miyahara, Tomoyuki Uchida</i>	144

Learning with Non-linear Optimization

Kernel Trick Embedded Gaussian Mixture Model 159 <i>Jingdong Wang, Jianguo Lee, Changshui Zhang</i>	159
Efficiently Learning the Metric with Side-Information 175 <i>Tijl De Bie, Michinari Momma, Nello Cristianini</i>	175
Learning Continuous Latent Variable Models with Bregman Divergences 190 <i>Shaojun Wang, Dale Schuurmans</i>	190
A Stochastic Gradient Descent Algorithm for Structural Risk Minimisation 205 <i>Joel Ratsaby</i>	205

Learning from Random Examples

On the Complexity of Training a Single Perceptron with Programmable Synaptic Delays 221 <i>Jiří Šíma</i>	221
Learning a Subclass of Regular Patterns in Polynomial Time 234 <i>John Case, Sanjay Jain, Rüdiger Reischuk, Frank Stephan, Thomas Zeugmann</i>	234
Identification with Probability One of Stochastic Deterministic Linear Languages 247 <i>Colin de la Higuera, Jose Oncina</i>	247

Online Prediction

Criterion of Calibration for Transductive Confidence Machine with Limited Feedback 259 <i>Ilija Nouretdinov, Vladimir Vovk</i>	259
Well-Calibrated Predictions from Online Compression Models 268 <i>Vladimir Vovk</i>	268

Transductive Confidence Machine Is Universal 283
Ilya Nouretdinov, Vladimir V'yugin, Alex Gammerman

On the Existence and Convergence of Computable Universal Priors 298
Marcus Hutter

Author Index 313

