# Lecture Notes in Artificial Intelligence4264Edited by J. G. Carbonell and J. Siekmann

Subseries of Lecture Notes in Computer Science

José L. Balcázar Philip M. Long Frank Stephan (Eds.)

# Algorithmic Learning Theory

17th International Conference, ALT 2006 Barcelona, Spain, October 7-10, 2006 Proceedings



Series Editors

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

Volume Editors

José L. Balcázar Universitat Politecnica de Catalunya, Dept. Llenguatges i Sistemes Informatics c/ Jordi Girona, 1-3, 08034 Barcelona, Spain E-mail: balqui@lsi.upc.edu

Philip M. Long Google 1600 Amphitheatre Parkway, Mountain View, CA 94043, USA E-mail: plong@google.com

Frank Stephan National University of Singapore, Depts. of Mathematics and Computer Science 2 Science Drive 2, Singapore 117543, Singapore E-mail: fstephan@comp.nus.edu.sg

Library of Congress Control Number: 2006933733

CR Subject Classification (1998): I.2.6, I.2.3, F.1, F.2, F.4, I.7

LNCS Sublibrary: SL 7 - Artificial Intelligence

| ISSN    | 0302-9743   |
|---------|---|
| ISBN-10 | 3-540-46649-5 Springer Berlin Heidelberg New York     |
| ISBN-13 | 978-3-540-46649-9 Springer 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. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springer.com

© Springer-Verlag Berlin Heidelberg 2006 Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, IndiaPrinted on acid-free paperSPIN: 1189484106/31425 4 3 2 1 0

## Preface

This volume contains the papers presented at the 17th Annual Internation Conference on Algorithmic Learning Theory (ALT 2006) which was held in Barcelona (Catalunya, Spain), October 7–10, 2006. The conference was organized with support from the PASCAL Network within the framework of PASCAL Dialogues 2006, which comprised three conferences:

Learning 2006 provided a forum for interdisciplinary study and discussion of the different aspects of learning and took place October 2–5, 2006 on the campus of Vilanova i La Geltrú.

ALT 2006 was dedicated to the theoretical foundations of machine learning and took place in the rooms of the Institute of Catalan Studies in Barcelona. ALT provides a forum for high-quality talks with a strong theoretical background and scientific interchange in areas such as query models, on-line learning, inductive inference, algorithmic forecasting, boosting, support vector machines, kernel methods, reinforcement learning and statistical learning models.

DS 2006 was the 9th International Conference on Discovery Science and focused on the development and analysis of methods for intelligent data analysis, knowledge discovery and machine learning, as well as their application to scientific knowledge discovery; as is already tradition, it was collocated and held in parallel with Algorithmic Learning Theory.

In addition to these three conferences, the European Workshop on Curricular Issues in Learning Theory initiated as the first regular meeting the Curriculum Development Programme of the PASCAL Network taking place on October 11, 2006.

The volume includes 24 contributions which the Programme Committee selected out of 53 submissions. It also contains descriptions of the five invited talks of ALT and DS; longer versions of the DS papers are available in the proceedings of DS 2006. These invited talks were presented to the audience of both conferences in joint sessions.

- Gunnar Rätsch (Friedrich Miescher Labor, Max Planck Gesellschaft, Tübingen, Germany): "Solving Semi-Infinite Linear Programs Using Boosting-Like Methods" (invited speaker for ALT 2006)
- Carole Goble (The University of Manchester, UK): "Putting Semantics into e-Science and the Grid" (invited speaker for DS 2006)
- Hans Ulrich Simon (Ruhr-Universität Bochum, Germany): "The Usage of the Spectral Norm in Learning Theory: Some Selected Topics" (invited speaker for ALT 2006)
- Padhraic Smyth (University of California at Irvine, USA): "Data-Driven Discovery Using Probabilistic Hidden Variable Models" (invited speaker for DS 2006)

Andrew Ng (Stanford University, USA): "Reinforcement Learning and Apprenticeship Learning for Robotic Control" (invited speaker jointly of ALT 2006 and DS 2006)

Since 1999, ALT has been awarding the E. M. Gold Award for the most outstanding contribution by a student. This year the award was given to Alp Atici for his paper "Learning Unions of  $\omega(1)$ -Dimensional Rectangles," co-authored by Rocco A. Servedio. We would like to thank Google for sponsoring the E. M. Gold Award.

Algorithmic Learning Theory 2006 was the 17th in a series of annual conferences established in Japan in 1990. A second root is the conference series Analogical and Inductive Inference previously held in 1986, 1989, 1992 which merged with the conference series ALT after a collocation in the year 1994. From then on, ALT became an international conference series, which kept its strong links to Japan but was also regularly held at overseas destinations including Australia, Germany, Italy, Singapore, Spain and the USA.

Continuation of ALT 2006 was supervised by its Steering Committee consisting of Naoki Abe (IBM Thomas J. Watson Research Center, Yorktown, USA), Shai Ben-David (University of Waterloo, Canada), Roni Khardon (Tufts University, Medford, USA), Steffen Lange (FH Darmstadt, Germany), Philip M. Long (Google, Mountain View, USA), Hiroshi Motoda (Osaka University, Japan), Akira Maruoka (Tohoku University, Sendai, Japan), Takeshi Shinohara (Kyushu Institute of Technology, Iizuka, Japan), Osamu Watanabe (Tokyo Institute of Technology, Japan), Arun Sharma (Queensland University of Technology, Brisbane, Australia – Co-chair), Frank Stephan (National University of Singapore, Republic of Singapore) and Thomas Zeugmann (Hokkaido University, Japan – Chair).

We would in particular like to thank Thomas Zeugmann for his continuous support of the ALT conference series and in particular for running the ALT Web page and the ALT submission system which he programmed together with Frank Balbach and Jan Poland. Thomas Zeugmann assisted us in many questions with respect to running the conference and to preparing the proceedings.

The ALT 2006 conference was made possible by the financial and administrative support of the PASCAL network, which organized this meeting together with others in the framework of PASCAL Dialogues 2006. Furthermore, we acknowledge the support of Google by financing the E. M. Gold Award (the corresponding award at Discovery Science 2006 was sponsored by Yahoo). We are grateful for the dedication of the host, the Universitat Politécnica de Catalunya (UPC), who organized the conference with much dedication and contributed to ALT in many ways. We want to express our gratitude to the Local Arrangements Chair Ricard Gavaldà and all other colleagues from the UPC, UPF and UB, who put so much time into making ALT 2006 to a success. Here we want also acknowledge the local sponsor Idescat, Statistical Institute of Catalonia. Furthermore, the Institute for Theoretical Computer Science of the University of Lübeck as well as the Division of Computer Science, Hokkaido University, Sapporo, supported ALT 2006. The conference series ALT was this year collocated with the series Discovery Science as in many previous years. We are greatful for this continuous collaboration and would like in particular to thank the conference Chair Klaus P. Jantke and the Programme Committee Chairs Nada Lavrac and Ljupco Todorovski of Discovery Science 2006.

We also want to thank the Programme Committee and the subreferees (both listed on the next pages) for their hard work in selecting a good programme for ALT 2006. Reviewing papers and checking the correctness of results are demanding in time and skills and we very much appreciated this contribution to the conference.

Last but not least we also want to thank the authors for choosing ALT 2006 as a forum to report on their research.

August 2006

Jose L. Balcázar Philip M. Long Frank Stephan

## Organization

#### **Conference Chair**

Jose L. Balcázar

Universitat Politécnica de Catalunya, Barcelona, Spain

#### **Program Committee**

| Shai Ben-David      | University of Waterloo                      |
|---------------------|---|
| Olivier Bousquet    | Pertinence                                  |
| Nader Bshouty       | Technion                                    |
| Nicolò Cesa-Bianchi | Universitá degli Studi di Milano            |
| Henning Fernau      | University of Hertfordshire                 |
| Bill Gasarch        | University of Maryland                      |
| Sally Goldman       | Washington University in St. Louis          |
| Kouichi Hirata      | Kyushu Institute of Technology, Iizuka      |
| Marcus Hutter       | IDSIA                                       |
| Efim Kinber         | Sacred Heart University                     |
| Philip M. Long      | Google (Co-chair)                           |
| Shie Mannor         | McGill University                           |
| Eric Martin         | The University of New South Wales           |
| Partha Niyogi       | University of Chicago                       |
| Steffen Lange       | Fachhochschule Darmstadt                    |
| Hans-Ulrich Simon   | Ruhr-Universität Bochum                     |
| Frank Stephan       | National University of Singapore (Co-chair) |
| Etsuji Tomita       | The University of Electro-Communications    |
| Sandra Zilles       | DFKI  |

#### Local Arrangements

Ricard Gavaldà

Universitat Politécnica de Catalunya, Barcelona, Spain

### Subreferees

Hiroki Arimura Amos Beimel Jochen Blath Francisco Casacuberta Alexey Chernov Alexander Clark Frank Drewes Vitaly Feldman Dmitry Gavinsky **Robert Glaubius** Gunter Grieser Colin de la Higuera Hiroki Ishizaka Yuri Kalnishkan Roni Khardon Adam Klivans Shigenobu Kobayashi Shane Legg Hanna Mazzawi Tetsuhiro Miyahara Sayan Mukherjee Thomas Nichols Jan Poland

Michael Richter Sebastien Roch Joseph Romano Daniel Reidenbach Cynthia Rudin Daniil Rvabko Hiroshi Sakamoto Rocco Servedio Takeshi Shinohara William D. Smart Yasuhiro Tajima Eiji Takimoto Franck Thollard Vladimir Vovk Mitsuo Wakatsuki Osamu Watanabe

### **Sponsoring Institutions**

Spanish Ministry of Science Google Pascal Network of Excellence PASCAL Dialogues 2006 Universitat Politècnica de Calalunya Idescat, Statistical Institute of Catalonia Institut für Theoretische Informatik, Universität Lübeck Division of Computer Science, Hokkaido University

# Table of Contents

| Editors' Introduction<br>Jose L. Balcázar, Philip M. Long, Frank Stephan  | 1  |
|---|----|
| Invited Contributions   |    |
| Solving Semi-infinite Linear Programs Using Boosting-Like Methods<br>Gunnar Rätsch                                  | 10 |
| e-Science and the Semantic Web: A Symbiotic Relationship<br>Carole Goble, Oscar Corcho, Pinar Alper, David De Roure | 12 |
| Spectral Norm in Learning Theory: Some Selected Topics<br>Hans Ulrich Simon   | 13 |
| Data-Driven Discovery Using Probabilistic Hidden Variable Models<br>Padhraic Smyth                                  | 28 |
| Reinforcement Learning and Apprenticeship Learning for Robotic<br>Control<br>Andrew Y. Ng                           | 29 |
| Regular Contributions   |    |
| Learning Unions of $\omega(1)$ -Dimensional Rectangles<br>Alp Atici, Rocco A. Servedio                              | 32 |
| On Exact Learning Halfspaces with Random Consistent Hypothesis<br>Oracle  | 48 |
| Active Learning in the Non-realizable Case<br>Matti Kääriäinen  | 63 |
| How Many Query Superpositions Are Needed to Learn?<br>Jorge Castro  | 78 |
| Teaching Memoryless Randomized Learners Without Feedback<br>Frank J. Balbach, Thomas Zeugmann                       | 93 |

| The Complexity of Learning $SUBSEQ(A)$<br>Stephen Fenner, William Gasarch   | 109 |
|---|-----|
| Mind Change Complexity of Inferring Unbounded Unions of Pattern<br>Languages from Positive Data<br>Matthew de Brecht, Akihiro Yamamoto                        | 124 |
| Learning and Extending Sublanguages<br>Sanjay Jain, Efim Kinber   | 139 |
| Iterative Learning from Positive Data and Negative Counterexamples<br>Sanjay Jain, Efim Kinber  | 154 |
| Towards a Better Understanding of Incremental Learning<br>Sanjay Jain, Steffen Lange, Sandra Zilles   | 169 |
| On Exact Learning from Random Walk<br>Nader H. Bshouty, Iddo Bentov   | 184 |
| Risk-Sensitive Online Learning<br>Eyal Even-Dar, Michael Kearns, Jennifer Wortman   | 199 |
| Leading Strategies in Competitive On-Line Prediction<br>Vladimir Vovk   | 214 |
| Hannan Consistency in On-Line Learning in Case of Unbounded Losses<br>Under Partial Monitoring<br>Chamy Allenberg, Peter Auer, László Györfi, György Ottucsák | 229 |
| General Discounting Versus Average Reward<br>Marcus Hutter  | 244 |
| The Missing Consistency Theorem for Bayesian Learning: Stochastic<br>Model Selection<br>Jan Poland  | 259 |
| Is There an Elegant Universal Theory of Prediction? Shane Legg  | 274 |
| Learning Linearly Separable Languages<br>Leonid Kontorovich, Corinna Cortes, Mehryar Mohri  | 288 |
| Smooth Boosting Using an Information-Based Criterion<br>Kohei Hatano  | 304 |

| Large-Margin Thresholded Ensembles for Ordinal Regression: Theory<br>and Practice                | 319 |
|--|-----|
| Asymptotic Learnability of Reinforcement Problems with Arbitrary<br>Dependence                   | 334 |
| Probabilistic Generalization of Simple Grammars and Its Application<br>to Reinforcement Learning | 348 |
| Unsupervised Slow Subspace-Learning from Stationary Processes<br>Andreas Maurer                  | 363 |
| Learning-Related Complexity of Linear Ranking Functions<br>Atsuyoshi Nakamura                    | 378 |
| Author Index   | 393 |