Lecture Notes in Artificial Intelligence 4754

Edited by J. G. Carbonell and J. Siekmann

Subseries of Lecture Notes in Computer Science

Marcus Hutter Rocco A. Servedio
Eiji Takimoto (Eds.)

Algorithmic Learning Theory

18th International Conference, ALT 2007
Sendai, Japan, October 1-4, 2007
Proceedings

Springer

Series Editors

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

Volume Editors

Marcus Hutter
RSISE @ ANU and SML @ NICTA
Canberra, ACT, 0200, Australia
E-mail: marcus@hutter1.net

Rocco A. Servedio
Columbia University
New York, NY, USA
E-mail: rocco@cs.columbia.edu

Eiji Takimoto
Tohoku University
Sendai 980-8579, Japan
E-mail: t2@maruoka.ecei.tohoku.ac.jp

Library of Congress Control Number: 2007935310

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 3-540-75224-2 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-75224-0 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 2007
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12164653 06/3180 5 4 3 2 1 0

Preface

This volume contains the papers presented at the 18th International Conference on Algorithmic Learning Theory (ALT 2007), which was held in Sendai (Japan) during October 1–4, 2007. The main objective of the conference was to provide an interdisciplinary 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, complexity and learning, reinforcement learning, unsupervised learning and grammatical inference. The conference was co-located with the Tenth International Conference on Discovery Science (DS 2007).

This volume includes 25 technical contributions that were selected from 50 submissions by the Program Committee. 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 2007. These invited talks were presented to the audience of both conferences in joint sessions.

- Avrim Blum (Carnegie Mellon University, Pittsburgh, USA): "A Theory of Similarity Functions for Learning and Clustering" (invited speaker for ALT 2007)
- Thomas G. Dietterich (Oregon State University, Corvallis, Oregon, USA): "Machine Learning in Ecosystem Informatics" (invited speaker for DS 2007)
- Masaru Kitsuregawa (The University of Tokyo, Tokyo, Japan): "Challenge for Info-plosion" (invited speaker for DS 2007)
- Alex Smola (National ICT Australia / ANU, Canberra, Australia): "A Hilbert Space Embedding for Distributions" (invited speaker for ALT 2007)
- Jürgen Schmidhuber (IDSIA, Lugano, Switzerland): "Simple Algorithmic Principles of Discovery, Subjective Beauty, Selective Attention, Curiosity and Creativity" (joint invited speaker for ALT 2007 and DS 2007)

Since 1999, ALT has been awarding the *E. Mark Gold* Award for the most outstanding paper by a student author. This year the award was given to Markus Maier for his paper "Cluster Identification in Nearest-Neighbor Graphs," coauthored by Matthias Hein and Ulrike von Luxburg. We thank Google for sponsoring the E.M. Gold Award.

ALT 2007 was the 18th in a series of annual conferences established in Japan in 1990. Another ancestor of ALT 2007 is the conference series Analogical and Inductive Inference, held in 1986, 1989, and 1992, which merged with the ALT conference series after a collocation in 1994. ALT subsequently became an international conference series which has kept its strong links to Japan but has also regularly been held at overseas destinations including Australia, Germany, Italy, Singapore, Spain and the USA.

Continuation of the ALT series is supervised by its Steering Committee, consisting of: Thomas Zeugmann (Hokkaido University, Japan) Chair, Steffen Lange

(FH Darmstadt, Germany) Publicity Chair, Naoki Abe (IBM Thomas J. Watson Research Center, Yorktown, USA), Shai Ben-David (University of Waterloo, Canada), Marcus Hutter (Australian National University, Canberra, Australia), Roni Khardon (Tufts University, Medford, USA), Phil Long (Google, Mountain View, USA), Akira Maruoka (Ishinomaki Senshu University, Japan), Rocco Servedio (Columbia University, New York, USA), Takeshi Shinohara (Kyushu Institute of Technology, Iizuka, Japan), Frank Stephan (National University of Singapore, Republic of Singapore), Einoshin Suzuki (Kyushu University, Fukuoka, Japan), and Osamu Watanabe (Tokyo Institute of Technology, Japan).

We would like to thank all of the individuals and institutions who contributed to the success of the conference: the authors for submitting papers, and the invited speakers for accepting our invitation and lending us their insight into recent developments in their research areas. We wish to thank the following sponsors for their generous financial support: the Air Force Office of Scientific Research (AFOSR); the Asian Office of Aerospace Research and Development (AOARD)[1]; Google for sponsoring the E.M.Gold Award; Graduate School of Information Sciences (GSIS), Tohoku University for providing secretarial assistance and equipment as well; the Research Institute of Electrical Communication (RIEC), Tohoku University; New Horizons in Computing, MEXT Grant-in-Aid for Scientific Research on Priority Areas; and the Semi-Structured Data Mining Project, MEXT Grant-in-Aid for Specially Promoted Research.

We are also grateful for the Technical Group on Computation (COMP) of the Institute of Electronics, Information and Communication Engineers (IEICE) for its technical sponsorship; the Division of Computer Science, Hokkaido University for providing the Web page and online submission system; and the Institute for Theoretical Computer Science, University of Lübeck where Frank Balbach developed a part of the online submission system.

We thank the Local Arrangements Chair Akira Ishino (Tohoku University, Japan) for his great assistance in making the conference a success in many ways. We thank Vincent Corruble for making the beautiful poster. We thank Springer for its continuous support in the preparation of this volume.

We would also like to thank all Program Committee members for their hard work in reviewing the submitted papers and participating in on-line discussions. We thank the external referees whose reviews made a substantial contribution to the process of selecting papers for ALT 2007.

We are grateful to the Discovery Science conference for its ongoing collaboration with ALT. In particular we would like to thank the Conference Chair Ayumi Shinohara (Tohoku University, Japan) and the Program Committee Chairs Vincent Corruble (UPMC, Paris, France) and Masayuki Takeda (Kyushu University, Japan) for their cooperation and support.

Finally, we would like to express special thanks to Thomas Zeugmann for his continuous support of the ALT conference series and in particular for his great

[1] AFOSR/AOARD support is not intended to express or imply endorsement by the U.S.Federal Government.

service in maintaining the ALT Web pages and the ALT submission system, which he programmed together with Frank Balbach and Jan Poland. Thomas Zeugmann assisted us in many ways by answering countless questions related to running the conference and preparing the proceedings.

July 2007

Marcus Hutter
Rocco A. Servedio
Eiji Takimoto

Organization

Conference Chair

Eiji Takimoto	Tohoku University, Japan

Program Committee

Marcus Hutter (Chair)	Australian National University, Australia
Rocco Servedio (Chair)	Columbia University, USA
Marta Arias	Columbia University, USA
Nader Bshouty	Technion, Israel
Sanjoy Dasgupta	University of California, San Diego, USA
Ran El-Yaniv	Technion, Israel
Claudio Gentile	Universita' dell'Insubria, Italy
Lisa Hellerstein	Polytechnic University, USA
Sanjay Jain	National University of Singapore, Singapore
Yuri Kalnishkan	Royal Holloway University of London, UK
Jyrki Kivinen	University of Helsinki, Finland
Adam Klivans	University of Texas, Austin, USA
Phil Long	Google, USA
Jan Poland	Hokkaido University, Japan
Daniil Ryabko	IDSIA, Switzerland
Frank Stephan	National University of Singapore, Singapore
Christino Tamon	Clarkson University, USA
Vladimir Vovk	Royal Holloway University of London, UK
Osamu Watanabe	Tokyo Institute of Technology, Japan
Bob Williamson	Australian National University, Australia

Local Arrangements

Akira Ishino	Tohoku University, Japan

Subreferees

Douglas Aberdeen
Ron Begleiter
John Case
Jiang Chen
Alexey Chernov
Alex Clarck
Yoav Freund
Alex Fukunaga
Bill Gasarch
Ricard Gavaldá
Paul Goldberg
Simon Guenter

Jeff C. Jackson
Efim Kinber
Gregory Kucherov
Steffen Lange
Shane Legg
Shie Mannor
Hanna Mazzawi
Dmitry Pechyony
Jim Royer
Tamer Salman
Shai Shalev-Shwartz
Takeshi Shinohara
Etsuji Tomita
György Turán
Rolf Wiehagen
Yair Wiener
Ryo Yoshinaka
Sandra Zilles

Sponsoring Institutions

Air Force Office of Scientific Research (AFOSR)
Asian Office of Aerospace Research and Development (AOARD)
Computation, IEICE of Japan
Google
GSIS, Tohoku University
New Horizons in Computing (NHC)
RIEC, Tohoku University
Semi-Structured Data Mining Project
Division of Computer Science, Hokkaido University
Institute for Theoretical Computer Science, University at Lübeck

Table of Contents

Editors' Introduction 1
Marcus Hutter, Rocco A. Servedio, and Eiji Takimoto

Invited Papers

A Theory of Similarity Functions for Learning and Clustering 9
Avrim Blum

Machine Learning in Ecosystem Informatics 10
Thomas G. Dietterich

Challenge for Info-plosion 12
Masaru Kitsuregawa

A Hilbert Space Embedding for Distributions 13
Alex Smola, Arthur Gretton, Le Song, and Bernhard Schölkopf

Simple Algorithmic Principles of Discovery, Subjective Beauty, Selective Attention, Curiosity and Creativity 32
Jürgen Schmidhuber

Invited Papers

Inductive Inference

Feasible Iteration of Feasible Learning Functionals 34
John Case, Timo Kötzing, and Todd Paddock

Parallelism Increases Iterative Learning Power 49
John Case and Samuel E. Moelius III

Prescribed Learning of R.E. Classes 64
Sanjay Jain, Frank Stephan, and Nan Ye

Learning in Friedberg Numberings 79
Sanjay Jain and Frank Stephan

Complexity Aspects of Learning

Separating Models of Learning with Faulty Teachers 94
Vitaly Feldman, Shrenik Shah, and Neal Wadhwa

Vapnik-Chervonenkis Dimension of Parallel Arithmetic Computations 107
César L. Alonso and José Luis Montaña

Parameterized Learnability of k-Juntas and Related Problems 120
Vikraman Arvind, Johannes Köbler, and Wolfgang Lindner

On Universal Transfer Learning 135
M.M. Hassan Mahmud

Online Learning

Tuning Bandit Algorithms in Stochastic Environments 150
Jean-Yves Audibert, Rémi Munos, and Csaba Szepesvári

Following the Perturbed Leader to Gamble at Multi-armed Bandits 166
Jussi Kujala and Tapio Elomaa

Online Regression Competitive with Changing Predictors 181
Steven Busuttil and Yuri Kalnishkan

Unsupervised Learning

Cluster Identification in Nearest-Neighbor Graphs 196
Markus Maier, Matthias Hein, and Ulrike von Luxburg

Multiple Pass Streaming Algorithms for Learning Mixtures of Distributions in $\mathbb{R}^d$ 211
Kevin L. Chang

Language Learning

Learning Efficiency of Very Simple Grammars from Positive Data 227
Ryo Yoshinaka

Learning Rational Stochastic Tree Languages 242
François Denis and Amaury Habrard

Query Learning

One-Shot Learners Using Negative Counterexamples and Nearest Positive Examples 257
Sanjay Jain and Efim Kinber

Polynomial Time Algorithms for Learning k-Reversible Languages and Pattern Languages with Correction Queries 272
Cristina Tîrnăucă and Timo Knuutila

Learning and Verifying Graphs Using Queries with a Focus on Edge Counting 285
Lev Reyzin and Nikhil Srivastava

Exact Learning of Finite Unions of Graph Patterns from Queries 298
Rika Okada, Satoshi Matsumoto, Tomoyuki Uchida, Yusuke Suzuki, and Takayoshi Shoudai

Kernel-Based Learning

Polynomial Summaries of Positive Semidefinite Kernels 313
Kilho Shin and Tetsuji Kuboyama

Learning Kernel Perceptrons on Noisy Data Using Random Projections 328
Guillaume Stempfel and Liva Ralaivola

Continuity of Performance Metrics for Thin Feature Maps 343
Adam Kowalczyk

Other Directions

Multiclass Boosting Algorithms for Shrinkage Estimators of Class Probability 358
Takafumi Kanamori

Pseudometrics for State Aggregation in Average Reward Markov Decision Processes 373
Ronald Ortner

On Calibration Error of Randomized Forecasting Algorithms 388
Vladimir V. V'yugin

Author Index 403