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

Algorithmic Learning Theory

LNAI 4754

18th International Conference, ALT 2007
Sendai, Japan, October 2007
Proceedings

Springer

Hutter • Servedio
Takimoto (Eds.)

LNAI
4754

Algorithmic
Learning Theory

ALT
2007

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: Applied for

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 10th 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 & 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: Air Force Office of Scientific Research (AFOSR); 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; Research Institute of Electrical Communication (RIEC), Tohoku University; New Horizons in Computing, MEXT Grant-in-Aid for Scientific Research on Priority Areas; and 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; Division of Computer Science, Hokkaido University for providing the web-page and online submission system; and 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 service in maintaining the ALT web pages and the ALT submission system,

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

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	(Australian National University, Australia) (Chair)
Rocco Servedio	(Columbia University, USA) (Chair)
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

Editors' Introduction

Marcus Hutter, Rocco A. Servedio, and Eiji Takimoto

Philosophers have pondered the phenomenon of learning for millennia; scientists and psychologists have studied learning for more than a century. But the analysis of learning as a *computational* and *algorithmic* phenomenon is much more recent, going back only a few decades. Learning theory is now an active research area that incorporates ideas, problems, and techniques from a wide range of disciplines including statistics, artificial intelligence, information theory, pattern recognition, and theoretical computer science. Learning theory has many robust connections with more applied research in machine learning and has made significant contributions to the development of applied systems and to fields such as electronic commerce and computational biology.

Since learning is a complex and multi-faceted phenomenon, it should come as no surprise that a wide range of different theoretical models of learning have been developed and analyzed. This diversity in the field is well reflected in the topics addressed by the invited speakers to ALT 2007 and DS 2007, and by the range of different research topics that have been covered by the contributors to this volume in their papers. The research reported here ranges over areas such as unsupervised learning, inductive inference, complexity and learning, boosting and reinforcement learning, query learning models, grammatical inference, online learning and defensive forecasting, and kernel methods. In this introduction we give an overview first of the five invited talks of ALT 2007 and DS 2007 and then of the regular contributions in this volume. We have grouped the papers under different headings to highlight certain similarities in subject matter or approach, but many papers span more than one area and other alternative groupings are certainly possible; the taxonomy we offer is by no means absolute.

Avrim Blum works on learning theory, online algorithms, approximation algorithms, and algorithmic game theory. His interests within learning theory include similarity functions and clustering, semi-supervised learning and co-training, online learning algorithms, kernels, preference elicitation and query learning, noise-tolerant learning, and attribute-efficient learning. In his invited talk for ALT 2007, Avrim spoke about developing a theory of similarity functions for learning and clustering problems. Some of the aims of this work are to provide new insights into what makes kernel functions useful for learning, and to understand what are the minimal conditions on a similarity function that allow it to be useful for clustering.

Alexander Smola works on nonparametric methods for estimation, in particular kernel methods and exponential families. He studies estimation techniques including Support Vector Machines, Gaussian Processes and Conditional Random Fields, and uses these techniques on problems in bioinformatics, pattern recognition, text analysis, computer vision, network security, and optimization for parallel processing. In his invited lecture for ALT 2007, co-authored with Arthur Gretton, Le Song, and Bernhard Schölkopf, Alexander spoke about a

technique for comparing distributions without the need for density estimation as an intermediate step. The approach relies on mapping the distributions into a reproducing kernel Hilbert space, and has a range of applications that were presented in the talk.

Masaru Kitsuregawa works on data mining, high performance data warehousing, high performance disk and tape arrays, parallel database processing, data storage and the Web, and related topics. His invited lecture for DS 2007 was about "Challenges for Info-plosion."

Thomas G. Dietterich studies topics in machine learning including sequential and spatial supervised learning, transfer learning, and combining knowledge and data to learn in knowledge-rich/data-poor application problems. He works on applying machine learning to a range of problems such as ecosystem informatics, intelligent desktop assistants, and applying AI to computer games. His invited lecture for DS 2007 discussed the role that machine learning can play in ecosystem informatics; this is a field that brings together mathematical and computational tools to address fundamental scientific and application problems in the ecosystem sciences. He described two on-going research efforts in ecosystem informatics at Oregon State University: (a) the application of machine learning and computer vision for automated arthropod population counting, and (b) the application of linear Gaussian dynamic Bayesian networks for automated cleaning of data from environmental sensor networks.

Jürgen Schmidhuber has worked on a range of topics related to learning, including artificial evolution, learning agents, reinforcement learning, metalearning, universal learning algorithms, Kolmogorov complexity and algorithmic probability. This work has led to applications in areas such as finance, robotics, and optimization. In his invited lecture (joint for ALT 2007 and DS 2007), Jürgen spoke about the algorithmic nature of discovery, perceived beauty, and curiosity. Jürgen has been thinking about this topic since 1994, when he postulated that among several patterns classified as "comparable" by some subjective observer, the subjectively most beautiful is the one with the simplest (shortest) description, given the observer's particular method for encoding and memorizing it. As one example of this phenomenon, mathematicians find beauty in a simple proof with a short description in the formal language they are using.

We now turn our attention to the regular contributions contained in this volume.

Inductive Inference. Research in inductive inference follows the pioneering work of Gold, who introduced a recursion-theoretic model of "learning in the limit." In the basic inductive inference setting, a learning machine is given a sequence of (arbitrarily ordered) examples drawn from a (recursive or recursively enumerable) language L, which belongs to a known class C of possible languages. The learning machine maintains a hypothesis which may be updated after each successive element of the sequence is received; very roughly speaking, the goal is for the learning machine's hypothesis to converge to the target language after finitely many steps. Many variants of this basic scenario have been studied in inductive inference during the decades since Gold's original work.

John Case, Timo Kötzing and Todd Paddock study a setting of learning in the limit in which the time to produce the final hypothesis is derived from some ordinal which is updated step by step downwards until it reaches zero, via some "feasible" functional. Their work first proposes a definition of feasible iteration of feasible learning functionals, and then studies learning hierarchies defined in terms of these notions; both collapse results and strict hierarchies are established under suitable conditions. The paper also gives upper and lower runtime bounds for learning hierarchies related to these definitions, expressed in terms of exponential polynomials.

John Case and Samuel Moelius III study *iterative learning*. This is a variant of the Gold-style learning model described above in which each of a learner's output conjectures may depend only on the learner's current conjecture and on the current input element. Case and Moelius analyze two extensions of this iterative model which incorporate parallelism in different ways. Roughly speaking, one of their results shows that running several distinct instantiations of a single learner in parallel can actually increase the power of iterative learners. This provides an interesting contrast with many standard settings where allowing parallelism only provides an efficiency improvement. Another result deals with a "collective" learner which is composed of a collection of communicating individual learners that run in parallel.

Sanjay Jain, Frank Stephan and Nan Ye study some basic questions about how hypothesis spaces connect to the class of languages being learned in Gold-style models. Building on work by Angluin, Lange and Zeugmann, their paper introduces a comprehensive unified approach to studying learning languages in the limit relative to different hypothesis spaces. Their work distinguishes between four different types of learning as they relate to hypothesis spaces, and gives results for vacillatory and behaviorally correct learning. They further show that every behaviorally correct learnable class has a *prudent* learner, i.e., a learner using a hypothesis space such that it learns every set in the hypothesis space.

Sanjay Jain and Frank Stephan study Gold-style learning of languages in some special numberings such as Friedberg numberings, in which each set has exactly one number. They show that while explanatorily learnable classes can all be learned in some Friedberg numberings, this is not the case for either behaviorally correct learning or finite learning. They also give results on how other properties of learners, such as consistency, conservativeness, prudence, iterativeness, and non U-shaped learning, relate to Friedberg numberings and other numberings.

Complexity aspects of learning. Connections between complexity and learning have been studied from a range of different angles. Work along these lines has been done in an effort to understand the computational complexity of various learning tasks; to measure the complexity of classes of functions using parameters such as the Vapnik-Chervonenkis dimension; to study functions of interest in learning theory from a complexity-theoretic perspective; and to understand connections between Kolmogorov-style complexity and learning. All four of these aspects were explored in research presented at ALT 2007.

Vitaly Feldman, Shrenek Shah, and Neal Wadhwa analyze two previously studied variants of Angluin's exact learning model that make learning more challenging: learning from equivalence and incomplete membership queries, and learning with random persistent classification noise in membership queries. They show that under cryptographic assumptions about the computational complexity of solving various problems the former oracle is strictly stronger than the latter, by demonstrating a concept class that is polynomial-time learnable from the former oracle but is not polynomial-time learnable from the latter oracle. They also resolve an open question of Bshouty and Eiron by showing that the incomplete membership query oracle is strictly weaker than a standard perfect membership query oracle under cryptographic assumptions.

César Alonso and José Montaña study the Vapnik-Chervonenkis dimension of concept classes that are defined in terms of arithmetic operations over real numbers. Such bounds are of interest in learning theory because of the fundamental role the Vapnik-Chervonenkis dimension plays in characterizing the sample complexity required to learn concept classes. Strengthening previous results of Goldberg and Jerrum, Alonso and Montaña give upper bounds on the VC dimension of concept classes in which the membership test for whether an input belongs to a concept in the class can be performed by an arithmetic circuit of bounded depth. These new bounds are polynomial both in the depth of the circuit and in the number of parameters needed to codify the concept.

Vikraman Arvind, Johannes Köbler, and Wolfgang Lindner study the problem of properly learning k-juntas and variants of k-juntas. Their work is done from the vantage point of parameterized complexity, which is a natural setting in which to consider the junta learning problem. Among other results, they show that the consistency problem for k-juntas is $W[2]$-complete, that the class of k-juntas is fixed parameter PAC learnable given access to a $W[2]$ oracle, and that k-juntas can be fixed parameter improperly learned with equivalence queries given access to a $W[2]$ oracle. These results give considerable insight on the junta learning problem.

The goal in transfer learning is to solve new learning problems more efficiently by leveraging information that was gained in solving previous related learning problems. One challenge in this area is to clearly define the notion of "relatedness" between tasks in a rigorous yet useful way. M. M. Hassan Mahmud analyzes transfer learning from the perspective of Kolmogorov complexity. Roughly speaking, he shows that if tasks are related in a particular precise sense, then joint learning is indeed faster than separate learning. This work strengthens previous work by Bennett, Gács, Li, Vitányi and Zurek.

Online Learning. Online learning proceeds in a sequence of rounds, where in each round the learning algorithm is presented with an input x and must generate a prediction y (a bit, a real number, or something else) for the label of x. Then the learner discovers the true value of the label, and incurs some loss which depends on the prediction and the true label. The usual overall goal is to keep the total loss small, often measured relative to the optimal loss over functions from some fixed class of predictors.

Jean-Yves Audibert, Rémi Munos and Csaba Szepesvári deal with the stochastic multi-armed bandit setting. They study an Upper Confidence Bound algorithm that takes into account the empirical variance of the different arms. They give an upper bound on the expected regret of the algorithm, and also analyze the concentration of the regret; this risk analysis is of interest since it is clearly useful to know how likely the algorithm is to have regret much higher than its expected value. The risk analysis reveals some unexpected tradeoffs between logarithmic expected regret and concentration of regret.

Jussi Kujala and Tapio Elomaa also consider a multi-armed bandit setting. They show that the "Follow the Perturbed Leader" technique can be used to obtain strong regret bounds (which hold against the best choice of a fixed lever in hindsight) against adaptive adversaries in this setting. This extends previous results for FPL's performance against non-adaptive adversaries in this setting.

Vovk's Aggregating Algorithm is a method of combining hypothesis predictors from a pool of candidates. Steven Busuttil and Yuri Kalnishkan show how Vovk's Aggregating Algorithm (AA) can be applied to online linear regression in a setting where the target predictor may change with time. Previous work had only used the Aggregating Algorithm in a static setting; the paper thus sheds new light on the methods that can be used to effectively perform regression with a changing target. Busuttil and Kalnishkan also analyze a kernel version of the algorithm and prove bounds on its square loss.

Unsupervised Learning. Many of the standard problems and frameworks in learning theory fall under the category of "supervised learning" in that learning is done from labeled data. In contrast, in unsupervised learning there are no labels provided for data points; the goal, roughly speaking, is to infer some underlying structure from the unlabeled data points that are received. Typically this means clustering the unlabeled data points or learning something about a probability distribution from which the points were obtained.

Markus Maier, Matthias Hein, and Ulrike von Luxburg study a scenario in which a learning algorithm receives a sample of points from an unknown distribution which contains a number of distinct clusters. The goal in this setting is to construct a "neighborhood graph" from the sample, such that the connected component structure of the graph mirrors the cluster ancestry of the sample points. They prove bounds on the performance of the k-nearest neighbor algorithm for this problem and also give some supporting experimental results. Markus received the E. M. Gold Award for this paper, as the program committee felt that it was the most outstanding contribution to ALT 2007 which was co-authored by a student.

Kevin Chang considers an unsupervised learning scenario in which a learner is given access to a sequence of samples drawn from a mixture of uniform distributions over rectangles in d-dimensional Euclidean space. He gives a streaming algorithm which makes only a small number of passes over such a sequence, uses a small amount of memory, and constructs a high-accuracy (in terms of statistical distance) hypothesis density function for the mixture. A notable feature of the algorithm is that it can handle samples from the mixture that are presented

in any arbitrary order. This result extends earlier work of Chang and Kannan which dealt with mixtures of uniform distributions over rectangles in one or two dimensions.

Language Learning. The papers in this group deal with various notions of learning languages in the limit from positive data. Ryo Yoshinaka's paper addresses the question of what precisely is meant by the notion of efficient language learning in the limit; despite the clear intuitive importance of such a notion, there is no single accepted definition. The discussion focuses particularly on learning very simple grammars and minimal simple grammars from positive data, giving both positive and negative results on efficient learnability under various notions.

François Denis and Amaury Habrard study the problem of learning stochastic tree languages, based on a sample of trees independently drawn according to an unknown stochastic language. They extend the notion of rational stochastic languages over strings to the domain of trees. Their paper introduces a canonical representation for rational stochastic languages over trees, and uses this representation to give an efficient inference algorithm that identifies the class of rational stochastic tree languages in the limit with probability 1.

Query Learning. In query learning the learning algorithm works by making queries of various types to an oracle or teacher; this is in contrast with "passive" statistical models where the learner typically only has access to random examples and cannot ask questions. The most commonly studied types of queries are membership queries (requests for the value of the target function at specified points) and equivalence queries (requests for counterexamples to a given hypothesis). Other types of queries, such as subset queries (in which the learner asks whether the current hypothesis is a subset of the target hypothesis, and if not, receives a negative counterexample) and superset queries, are studied as well.

Sanjay Jain and Efim Kinber study a query learning framework in which the queries used are variants of the standard queries described above. In their model the learner receives the *least* negative counterexample to subset queries, and is also given a "correction" in the form of a positive example which is nearest to the negative example; they also consider similarly modified membership queries. These variants are motivated in part by considerations of human language learning, in which corrected versions of incorrect utterances are often provided as part of the learning process. Their results show that "correcting" positive examples can sometimes give significant additional power to learners.

Cristina Tîrnăucă and Timo Knuutila study query learning under a different notion of correction queries, in which the prefix of a string (the query) is "corrected" by the teacher responding with the lexicographically first suffix that yields a string in the language. They give polynomial time algorithms for pattern languages and k-reversible languages using correction queries of this sort. These results go beyond what is possible for polynomial-time algorithms using membership queries alone, and thus demonstrate the power of learning from these types of correction queries.

Lev Reyzin and Nikhil Srivastava study various problems of learning and verifying properties of hidden graphs given query access to the graphs. This setting lends itself naturally to a range of query types that are somewhat different from those described above; these include edge detection, edge counting, and shortest path queries. Reyzin and Srivastava give bounds on learning and verifying general graphs, degree-bounded graphs, and trees with these types of queries. These results extend our understanding of what these types of queries can accomplish.

Rika Okada, Satoshi Matsumoto, Tomoyuki Uchida, Yusuke Suzuki and Takayoshi Shoudai study learnability of finite unions of linear graph patterns from equivalence queries and subset queries. These types of graph patterns are useful for data mining from semi-structured data. The authors show that positive results can be achieved for learning from equivalence and subset queries (with counterexamples), and give negative results for learning from restricted subset queries (in which no counterexamples are given).

Kernel-Based Learning. A kernel function is a mapping which, given two inputs, implicitly represents each input as a vector in some (possibly high-dimensional or infinite dimensional) feature space and outputs the inner product between these two vectors. Kernel methods have received much attention in recent years in part because it is often possible to compute the value of the kernel function much more efficiently than would be possible by performing an explicit representation of the input as a vector in feature space. Kernel functions play a crucial role in Support Vector Machines and have a rich theory as well as many uses in practical systems.

Developing new kernel functions, and selecting the most appropriate kernels for particular learning tasks, is an active area of research. One difficulty in constructing kernel functions is in ensuring that they obey the condition of positive semidefiniteness. Kilho Shin and Tetsuji Kuboyama give a sufficient condition under which it is ensured that new candidate kernels constructed in a particular way from known positive semidefinite kernels will themselves be positive semidefinite and hence will indeed be legitimate kernel functions. Their work gives new insights into several kernel functions that have been studied recently such as principal-angle kernels, determinant kernels, and codon-improved kernels.

Guillaume Stempfel and Liva Ralaivola study how kernels can be used to learn data separable in the feature space except for the presence of random classification noise. They describe an algorithm which combines kernel methods, random projections, and known noise tolerant approaches for learning linear separators over finite dimensional feature spaces, and give a PAC style analysis of the algorithm. Given noisy data which is such that the noise-free version would be linearly separable with a suitable margin in the implicit feature space, their approach yields an efficient algorithm for learning even if the implicit feature space has infinitely many dimensions.

Adam Kowalczyk's paper deals with analyzing hypothesis classes that consist of linear functionals superimposed with "smooth" feature maps; these are the types of hypotheses generated by many kernel methods. The paper studies continuity of two important performance metrics, namely the error rate and

the area under the ROC (receiver operating characteristic curve), for hypotheses of this sort. Using tools from real analysis, specifically transversality theory, he shows that pointwise convergence of hypotheses implies convergence of these measures with probability 1 over the selection of the test sample from a suitable probability density.

Other Directions. Several papers presented at ALT do not fit neatly into the above categories, but as described below each of these deals with an active and interesting area of research in learning theory.

Hypothesis boosting is an approach to combining many weak classifiers, or "rules of thumb," each of which performs only slightly better than random guessing, to obtain a high-accuracy final hypothesis. Boosting algorithms have been intensively studied and play an important role in many practical applications. In his paper, Takafumi Kanamori studies how boosting can be applied to estimate conditional probabilities of output labels in a multiclass classification setting. He proposes loss functions for boosting algorithms that generalize the known margin-based loss function and shows how regularization can be introduced with an appropriate instantiation of the loss function.

Reinforcement learning is a widely studied approach to sequential decision problems that has achieved considerable success in practice. Dealing with the "curse of dimensionality," which arises from large state spaces in Markov decision processes, is a major challenge. One approach to dealing with this challenge is *state aggregation*, which is based on the idea that similar states can be grouped together into meta-states. In his paper Ronald Ortner studies pseudometrics for measuring similarity in state aggregation. He proves an upper bound on the loss incurred by working with aggregated states rather than original states and analyzes how online aggregation can be performed when the MDP is not known to the learner in advance.

In defensive forecasting, the problem studied is that of online prediction of the binary label associated with each instance in a sequence of instances. In this line of work no assumption is made that there exists a hidden function dictating the labels, and in contrast with other work in online learning there is no comparison class or "best expert" that is compared with. One well-studied parameter of algorithms in this setting is the calibration error, which roughly speaking measures the extent to which the forecasts are accurate on average. In his paper Vladimir V. V'yugin establishes a tradeoff between the calibration error and the "coarseness" of any prediction strategy by showing that if the coarseness is small then the calibration error can also not be too small. This negative result comes close to matching the bounds given in previous work by Kakade and Foster on a particular forecasting system.

July 2007

Marcus Hutter
Rocco A. Servedio
Eiji Takimoto

Author Index

Alonso, César L. 107
Arvind, Vikraman 120
Audibert, Jean-Yves 150

Blum, Avrim 9
Busuttil, Steven 181

Case, John 34, 49
Chang, Kevin L. 211

Denis, François 242
Dietterich, Thomas G. 10

Elomaa, Tapio 166

Feldman, Vitaly 94

Gretton, Arthur 13

Habrard, Amaury 242
Hein, Matthias 196
Hutter, Marcus 1

Jain, Sanjay 64, 79, 257

Kalnishkan, Yuri 181
Kanamori, Takafumi 358
Kinber, Efim 257
Kitsuregawa, Masaru 12
Knuutila, Timo 272
Köbler, Johannes 120
Kötzing, Timo 34
Kowalczyk, Adam 343
Kuboyama, Tetsuji 313
Kujala, Jussi 166

Lindner, Wolfgang 120

Mahmud, M.M. Hassan 135
Maier, Markus 196
Matsumoto, Satoshi 298
Moelius III, Samuel E. 41
Montaña, José Luis 107
Munos, Rémi 150

Okada, Rika 298
Ortner, Ronald 373

Paddock, Todd 34

Ralaivola, Liva 328
Reyzin, Lev 285

Schmidhuber, Jürgen 32
Schölkopf, Bernhard 13
Servedio, Rocco A. 1
Shah, Shrenik 94
Shin, Kilho 313
Shoudai, Takayoshi 298
Smola, Alex 13
Song, Le 13
Srivastava, Nikhil 285
Stempfel, Guillaume 328
Stephan, Frank 64, 79
Suzuki, Yusuke 298
Szepesvári, Csaba 150

Tîrnăucă, Cristina 272
Takimoto, Eiji 1

Uchida, Tomoyuki 298

von Luxburg, Ulrike 196
V'yugin, Vladimir V. 388

Wadhwa, Neal 94

Ye, Nan 64
Yoshinaka, Ryo 227

Lecture Notes in Artificial Intelligence (LNAI)

Vol. 4754: M. Hutter, R.A. Servedio, E. Takimoto (Eds.), Algorithmic Learning Theory. XI, 403 pages. 2007.

Vol. 4737: B. Berendt, A. Hotho, D. Mladenic, G. Semeraro (Eds.), From Web to Social Web: Discovering and Deploying User and Content Profiles. XI, 161 pages. 2007.

Vol. 4733: R. Basili, M.T. Pazienza (Eds.), AI*IA 2007: Artificial Intelligence and Human-Oriented Computing. XVII, 858 pages. 2007.

Vol. 4722: C. Pelachaud, J.-C. Martin, E. André, G. Chollet, K. Karpouzis, D. Pelé (Eds.), Intelligent Virtual Agents. XV, 425 pages. 2007.

Vol. 4720: B. Konev, F. Wolter (Eds.), Frontiers of Combining Systems. X, 283 pages. 2007.

Vol. 4702: J.N. Kok, J. Koronacki, R. Lopez de Mantaras, S. Matwin, D. Mladenič, A. Skowron (Eds.), Knowledge Discovery in Databases: PKDD 2007. XXIV, 640 pages. 2007.

Vol. 4701: J.N. Kok, J. Koronacki, R.L.d. Mantaras, S. Matwin, D. Mladenič, A. Skowron (Eds.), Machine Learning: ECML 2007. XXII, 809 pages. 2007.

Vol. 4696: H.-D. Burkhard, G. Lindemann, R. Verbrugge, L.Z. Varga (Eds.), Multi-Agent Systems and Applications V. XIII, 350 pages. 2007.

Vol. 4694: B. Apolloni, R.J. Howlett, L. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part III. XXIX, 1126 pages. 2007.

Vol. 4693: B. Apolloni, R.J. Howlett, L. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part II. XXXII, 1380 pages. 2007.

Vol. 4692: B. Apolloni, R.J. Howlett, L. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part I. LV, 882 pages. 2007.

Vol. 4687: P. Petta, J.P. Müller, M. Klusch, M. Georgeff (Eds.), Multiagent System Technologies. X, 207 pages. 2007.

Vol. 4682: D.-S. Huang, L. Heutte, M. Loog (Eds.), Advanced Intelligent Computing Theories and Applications. XXVII, 1373 pages. 2007.

Vol. 4676: M. Klusch, K. Hindriks, M.P. Papazoglou, L. Sterling (Eds.), Cooperative Information Agents XI. XI, 361 pages. 2007.

Vol. 4667: J. Hertzberg, M. Beetz, R. Englert (Eds.), KI 2007: Advances in Artificial Intelligence. IX, 516 pages. 2007.

Vol. 4660: S. Džeroski, J. Todorovski (Eds.), Computational Discovery of Scientific Knowledge. X, 327 pages. 2007.

Vol. 4659: V. Mařík, V. Vyatkin, A.W. Colombo (Eds.), Holonic and Multi-Agent Systems for Manufacturing. VIII, 456 pages. 2007.

Vol. 4651: F. Azevedo, P. Barahona, F. Fages, F. Rossi (Eds.), Recent Advances in Constraints. VIII, 185 pages. 2007.

Vol. 4648: F. Almeida e Costa, L.M. Rocha, E. Costa, I. Harvey, A. Coutinho (Eds.), Advances in Artificial Life. XVIII, 1215 pages. 2007.

Vol. 4635: B. Kokinov, D.C. Richardson, T.R. Roth-Berghofer, L. Vieu (Eds.), Modeling and Using Context. XIV, 574 pages. 2007.

Vol. 4632: R. Alhajj, H. Gao, X. Li, J. Li, O.R. Zaïane (Eds.), Advanced Data Mining and Applications. XV, 634 pages. 2007.

Vol. 4629: V. Matoušek, P. Mautner (Eds.), Text, Speech and Dialogue. XVII, 663 pages. 2007.

Vol. 4626: R.O. Weber, M.M. Richter (Eds.), Case-Based Reasoning Research and Development. XIII, 534 pages. 2007.

Vol. 4617: V. Torra, Y. Narukawa, Y. Yoshida (Eds.), Modeling Decisions for Artificial Intelligence. XII, 502 pages. 2007.

Vol. 4612: I. Miguel, W. Ruml (Eds.), Abstraction, Reformulation, and Approximation. XI, 418 pages. 2007.

Vol. 4604: U. Priss, S. Polovina, R. Hill (Eds.), Conceptual Structures: Knowledge Architectures for Smart Applications. XII, 514 pages. 2007.

Vol. 4603: F. Pfenning (Ed.), Automated Deduction – CADE-21. XII, 522 pages. 2007.

Vol. 4597: P. Perner (Ed.), Advances in Data Mining. XI, 353 pages. 2007.

Vol. 4594: R. Bellazzi, A. Abu-Hanna, J. Hunter (Eds.), Artificial Intelligence in Medicine. XVI, 509 pages. 2007.

Vol. 4585: M. Kryszkiewicz, J.F. Peters, H. Rybinski, A. Skowron (Eds.), Rough Sets and Intelligent Systems Paradigms. XIX, 836 pages. 2007.

Vol. 4578: F. Masulli, S. Mitra, G. Pasi (Eds.), Applications of Fuzzy Sets Theory. XVIII, 693 pages. 2007.

Vol. 4573: M. Kauers, M. Kerber, R. Miner, W. Windsteiger (Eds.), Towards Mechanized Mathematical Assistants. XIII, 407 pages. 2007.

Vol. 4571: P. Perner (Ed.), Machine Learning and Data Mining in Pattern Recognition. XIV, 913 pages. 2007.

Vol. 4570: H.G. Okuno, M. Ali (Eds.), New Trends in Applied Artificial Intelligence. XXI, 1194 pages. 2007.

Vol. 4565: D.D. Schmorrow, L.M. Reeves (Eds.), Foundations of Augmented Cognition. XIX, 450 pages. 2007.

Vol. 4562: D. Harris (Ed.), Engineering Psychology and Cognitive Ergonomics. XXIII, 879 pages. 2007.

Vol. 4548: N. Olivetti (Ed.), Automated Reasoning with Analytic Tableaux and Related Methods. X, 245 pages. 2007.

Vol. 4539: N.H. Bshouty, C. Gentile (Eds.), Learning Theory. XII, 634 pages. 2007.

Vol. 4529: P. Melin, O. Castillo, L.T. Aguilar, J. Kacprzyk, W. Pedrycz (Eds.), Foundations of Fuzzy Logic and Soft Computing. XIX, 830 pages. 2007.

Vol. 4520: M.V. Butz, O. Sigaud, G. Pezzulo, G. Baldassarre (Eds.), Anticipatory Behavior in Adaptive Learning Systems. X, 379 pages. 2007.

Vol. 4511: C. Conati, K. McCoy, G. Paliouras (Eds.), User Modeling 2007. XVI, 487 pages. 2007.

Vol. 4509: Z. Kobti, D. Wu (Eds.), Advances in Artificial Intelligence. XII, 552 pages. 2007.

Vol. 4496: N.T. Nguyen, A. Grzech, R.J. Howlett, L.C. Jain (Eds.), Agent and Multi-Agent Systems: Technologies and Applications. XXI, 1046 pages. 2007.

Vol. 4483: C. Baral, G. Brewka, J. Schlipf (Eds.), Logic Programming and Nonmonotonic Reasoning. IX, 327 pages. 2007.

Vol. 4482: A. An, J. Stefanowski, S. Ramanna, C.J. Butz, W. Pedrycz, G. Wang (Eds.), Rough Sets, Fuzzy Sets, Data Mining and Granular Computing. XIV, 585 pages. 2007.

Vol. 4481: J. Yao, P. Lingras, W.-Z. Wu, M. Szczuka, N.J. Cercone, D. Ślęzak (Eds.), Rough Sets and Knowledge Technology. XIV, 576 pages. 2007.

Vol. 4476: V. Gorodetsky, C. Zhang, V.A. Skormin, L. Cao (Eds.), Autonomous Intelligent Systems: Multi-Agents and Data Mining. XIII, 323 pages. 2007.

Vol. 4456: Y. Wang, Y.-m. Cheung, H. Liu (Eds.), Computational Intelligence and Security. XXIII, 1118 pages. 2007.

Vol. 4455: S. Muggleton, R. Otero, A. Tamaddoni-Nezhad (Eds.), Inductive Logic Programming. XII, 456 pages. 2007.

Vol. 4452: M. Fasli, O. Shehory (Eds.), Agent-Mediated Electronic Commerce. VIII, 249 pages. 2007.

Vol. 4451: T.S. Huang, A. Nijholt, M. Pantic, A. Pentland (Eds.), Artifical Intelligence for Human Computing. XVI, 359 pages. 2007.

Vol. 4441: C. Müller (Ed.), Speaker Classification. X, 309 pages. 2007.

Vol. 4438: L. Maicher, A. Sigel, L.M. Garshol (Eds.), Leveraging the Semantics of Topic Maps. X, 257 pages. 2007.

Vol. 4434: G. Lakemeyer, E. Sklar, D.G. Sorrenti, T. Takahashi (Eds.), RoboCup 2006: Robot Soccer World Cup X. XIII, 566 pages. 2007.

Vol. 4429: R. Lu, J.H. Siekmann, C. Ullrich (Eds.), Cognitive Systems. X, 161 pages. 2007.

Vol. 4428: S. Edelkamp, A. Lomuscio (Eds.), Model Checking and Artificial Intelligence. IX, 185 pages. 2007.

Vol. 4426: Z.-H. Zhou, H. Li, Q. Yang (Eds.), Advances in Knowledge Discovery and Data Mining. XXV, 1161 pages. 2007.

Vol. 4411: R.H. Bordini, M. Dastani, J. Dix, A.E.F. Seghrouchni (Eds.), Programming Multi-Agent Systems. XIV, 249 pages. 2007.

Vol. 4410: A. Branco (Ed.), Anaphora: Analysis, Algorithms and Applications. X, 191 pages. 2007.

Vol. 4399: T. Kovacs, X. Llorà, K. Takadama, P.L. Lanzi, W. Stolzmann, S.W. Wilson (Eds.), Learning Classifier Systems. XII, 345 pages. 2007.

Vol. 4390: S.O. Kuznetsov, S. Schmidt (Eds.), Formal Concept Analysis. X, 329 pages. 2007.

Vol. 4389: D. Weyns, H. Van Dyke Parunak, F. Michel (Eds.), Environments for Multi-Agent Systems III. X, 273 pages. 2007.

Vol. 4386: P. Noriega, J. Vázquez-Salceda, G. Boella, O. Boissier, V. Dignum, N. Fornara, E. Matson (Eds.), Coordination, Organizations, Institutions, and Norms in Agent Systems II. XI, 373 pages. 2007.

Vol. 4384: T. Washio, K. Satoh, H. Takeda, A. Inokuchi (Eds.), New Frontiers in Artificial Intelligence. IX, 401 pages. 2007.

Vol. 4371: K. Inoue, K. Satoh, F. Toni (Eds.), Computational Logic in Multi-Agent Systems. X, 315 pages. 2007.

Vol. 4369: M. Umeda, A. Wolf, O. Bartenstein, U. Geske, D. Seipel, O. Takata (Eds.), Declarative Programming for Knowledge Management. X, 229 pages. 2006.

Vol. 4343: C. Müller (Ed.), Speaker Classification I. X, 355 pages. 2007.

Vol. 4342: H. de Swart, E. Orłowska, G. Schmidt, M. Roubens (Eds.), Theory and Applications of Relational Structures as Knowledge Instruments II. X, 373 pages. 2006.

Vol. 4335: S.A. Brueckner, S. Hassas, M. Jelasity, D. Yamins (Eds.), Engineering Self-Organising Systems. XII, 212 pages. 2007.

Vol. 4334: B. Beckert, R. Hähnle, P.H. Schmitt (Eds.), Verification of Object-Oriented Software. XXIX, 658 pages. 2007.

Vol. 4333: U. Reimer, D. Karagiannis (Eds.), Practical Aspects of Knowledge Management. XII, 338 pages. 2006.

Vol. 4327: M. Baldoni, U. Endriss (Eds.), Declarative Agent Languages and Technologies IV. VIII, 257 pages. 2006.

Vol. 4314: C. Freksa, M. Kohlhase, K. Schill (Eds.), KI 2006: Advances in Artificial Intelligence. XII, 458 pages. 2007.

Vol. 4304: A. Sattar, B.-h. Kang (Eds.), AI 2006: Advances in Artificial Intelligence. XXVII, 1303 pages. 2006.

Vol. 4303: A. Hoffmann, B.-h. Kang, D. Richards, S. Tsumoto (Eds.), Advances in Knowledge Acquisition and Management. XI, 259 pages. 2006.

Vol. 4293: A. Gelbukh, C.A. Reyes-Garcia (Eds.), MICAI 2006: Advances in Artificial Intelligence. XXVIII, 1232 pages. 2006.

Lecture Notes in Artificial Intelligence

Subseries of Lecture Notes in Computer Science

The LNAI series reports state-of-the-art results in artificial intelligence research, development, and education, at a high level and in both printed and electronic form. Enjoying tight cooperation with the R&D community, with numerous individuals, as well as with prestigious organizations and societies, LNAI has grown into the most comprehensive artificial intelligence research forum available.

The scope of LNAI spans the whole range of artificial intelligence and intelligent information processing including interdisciplinary topics in a variety of application fields. The type of material published traditionally includes

- proceedings (published in time for the respective conference)
- post-proceedings (consisting of thoroughly revised final full papers)
- research monographs (which may be based on PhD work)

More recently, several color-cover sublines have been added featuring, beyond a collection of papers, various added-value components; these sublines include

- tutorials (textbook-like monographs or collections of lectures given at advanced courses)
- state-of-the-art surveys (offering complete and mediated coverage of a topic)
- hot topics (introducing emergent topics to the broader community)

In parallel to the printed book, each new volume is published electronically in LNCS Online.

Detailed information on LNCS can be found at
www.springer.com/lncs

Proposals for publication should be sent to
LNCS Editorial, Tiergartenstr. 17, 69121 Heidelberg, Germany
E-mail: lncs@springer.com

ISSN 0302-9743

ISBN 978-3-540-75224-0

Lecture Notes in Artificial Intelligence
Lecture Notes in Computer Science

springer.com

Hutter • Servedio
Takimoto (Eds.)

LNAI
4754

Algorithmic Learning Theory

ALT
2007