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

LNAI 4264

Algorithmic Learning Theory

17th International Conference, ALT 2006
Barcelona, Spain, October 2006
Proceedings

Springer

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

Algorithmic Learning Theory

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

Springer

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

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11894841 06/3142 5 4 3 2 1 0

Lecture Notes in Artificial Intelligence 4264

Edited by J. G. Carbonell and J. Siekmann

Subseries of Lecture Notes in Computer Science

Lecture Notes in Artificial Intelligence (LNAI)

Vol. 4264: J.L. Balcázar, P.M. Long, F. Stephan (Eds.), Algorithmic Learning Theory. XIII, 393 pages. 2006.

Vol. 4248: S. Staab, V. Svátek (Eds.), Engineering Knowledge in the Age of the Semantic Web. XIV, 400 pages. 2006.

Vol. 4223: L. Wang, L. Jiao, G. Shi, X. Li, J. Liu (Eds.), Fuzzy Systems and Knowledge Discovery. XXVIII, 1335 pages. 2006.

Vol. 4213: J. Fürnkranz, T. Scheffer, M. Spiliopoulou (Eds.), Knowledge Discovery in Databases: PKDD 2006. XXII, 660 pages. 2006.

Vol. 4212: J. Fürnkranz, T. Scheffer, M. Spiliopoulou (Eds.), Machine Learning: ECML 2006. XXIII, 851 pages. 2006.

Vol. 4211: P. Vogt, Y. Sugita, E. Tuci, C. Nehaniv (Eds.), Symbol Grounding and Beyond. VIII, 237 pages. 2006.

Vol. 4203: F. Esposito, Z.W. Raś, D. Malerba, G. Semeraro (Eds.), Foundations of Intelligent Systems. XVIII, 767 pages. 2006.

Vol. 4201: Y. Sakakibara, S. Kobayashi, K. Sato, T. Nishino, E. Tomita (Eds.), Grammatical Inference: Algorithms and Applications. XII, 359 pages. 2006.

Vol. 4198: O. Nasraoui, O. Zaiane, M. Spiliopoulou, B. Mobasher, B. Masand, P. Yu (Eds.), Web Minding and Web Usage Analysis. IX, 177 pages. 2006.

Vol. 4196: K. Fischer, I.J. Timm, E. André, N. Zhong (Eds.), Multiagent System Technologies. X, 185 pages. 2006.

Vol. 4188: P. Sojka, I. Kopeček, K. Pala (Eds.), Text, Speech and Dialogue. XIV, 721 pages. 2006.

Vol. 4183: J. Euzenat, J. Domingue (Eds.), Artificial Intelligence: Methodology, Systems, and Applications. XIII, 291 pages. 2006.

Vol. 4180: M. Kohlhase, OMDoc – An Open Markup Format for Mathematical Documents [version 1.2]. XIX, 428 pages. 2006.

Vol. 4177: R. Marín, E. Onaindía, A. Bugarín, J. Santos (Eds.), Current Topics in Artificial Intelligence. XV, 482 pages. 2006.

Vol. 4160: M. Fisher, W.v.d. Hoek, B. Konev, A. Lisitsa (Eds.), Logics in Artificial Intelligence. XII, 516 pages. 2006.

Vol. 4155: O. Stock, M. Schaerf (Eds.), Reasoning, Action and Interaction in AI Theories and Systems. XVIII, 343 pages. 2006.

Vol. 4149: M. Klusch, M. Rovatsos, T.R. Payne (Eds.), Cooperative Information Agents X. XII, 477 pages. 2006.

Vol. 4139: T. Salakoski, F. Ginter, S. Pyysalo, T. Pahikkala, Advances in Natural Language Processing. XVI, 771 pages. 2006.

Vol. 4133: J. Gratch, M. Young, R. Aylett, D. Ballin, P. Olivier (Eds.), Intelligent Virtual Agents. XIV, 472 pages. 2006.

Vol. 4130: U. Furbach, N. Shankar (Eds.), Automated Reasoning. XV, 680 pages. 2006.

Vol. 4120: J. Calmet, T. Ida, D. Wang (Eds.), Artificial Intelligence and Symbolic Computation. XIII, 269 pages. 2006.

Vol. 4114: D.-S. Huang, K. Li, G.W. Irwin (Eds.), Computational Intelligence, Part II. XXVII, 1337 pages. 2006.

Vol. 4108: J.M. Borwein, W.M. Farmer (Eds.), Mathematical Knowledge Management. VIII, 295 pages. 2006.

Vol. 4106: T.R. Roth-Berghofer, M.H. Göker, H. A. Güvenir (Eds.), Advances in Case-Based Reasoning. XIV, 566 pages. 2006.

Vol. 4099: Q. Yang, G. Webb (Eds.), PRICAI 2006: Trends in Artificial Intelligence. XXVIII, 1263 pages. 2006.

Vol. 4095: S. Nolfi, G. Baldassarre, R. Calabretta, J.C. T. Hallam, D. Marocco, J.-A. Meyer, O. Miglino, D. Parisi (Eds.), From Animals to Animats 9. XV, 869 pages. 2006.

Vol. 4093: X. Li, O.R. Zaïane, Z. Li (Eds.), Advanced Data Mining and Applications. XXI, 1110 pages. 2006.

Vol. 4092: J. Lang, F. Lin, J. Wang (Eds.), Knowledge Science, Engineering and Management. XV, 664 pages. 2006.

Vol. 4088: Z.-Z. Shi, R. Sadananda (Eds.), Agent Computing and Multi-Agent Systems. XVII, 827 pages. 2006.

Vol. 4087: F. Schwenker, S. Marinai (Eds.), Artificial Neural Networks in Pattern Recognition. IX, 299 pages. 2006.

Vol. 4068: H. Schärfe, P. Hitzler, P. Øhrstrøm (Eds.), Conceptual Structures: Inspiration and Application. XI, 455 pages. 2006.

Vol. 4065: P. Perner (Ed.), Advances in Data Mining. XI, 592 pages. 2006.

Vol. 4062: G. Wang, J.F. Peters, A. Skowron, Y. Yao (Eds.), Rough Sets and Knowledge Technology. XX, 810 pages. 2006.

Vol. 4049: S. Parsons, N. Maudet, P. Moraitis, I. Rahwan (Eds.), Argumentation in Multi-Agent Systems. XIV, 313 pages. 2006.

Vol. 4048: L. Goble, J.-J.C.. Meyer (Eds.), Deontic Logic and Artificial Normative Systems. X, 273 pages. 2006.

Vol. 4045: D. Barker-Plummer, R. Cox, N. Swoboda (Eds.), Diagrammatic Representation and Inference. XII, 301 pages. 2006.

Vol. 4031: M. Ali, R. Dapoigny (Eds.), Advances in Applied Artificial Intelligence. XXIII, 1353 pages. 2006.

Vol. 4029: L. Rutkowski, R. Tadeusiewicz, L.A. Zadeh, J.M. Zurada (Eds.), Artificial Intelligence and Soft Computing – ICAISC 2006. XXI, 1235 pages. 2006.

Vol. 4027: H.L. Larsen, G. Pasi, D. Ortiz-Arroyo, T. Andreasen, H. Christiansen (Eds.), Flexible Query Answering Systems. XVIII, 714 pages. 2006.

Vol. 4021: E. André, L. Dybkjær, W. Minker, H. Neumann, M. Weber (Eds.), Perception and Interactive Technologies. XI, 217 pages. 2006.

Vol. 4020: A. Bredenfeld, A. Jacoff, I. Noda, Y. Takahashi (Eds.), RoboCup 2005: Robot Soccer World Cup IX. XVII, 727 pages. 2006.

Vol. 4013: L. Lamontagne, M. Marchand (Eds.), Advances in Artificial Intelligence. XIII, 564 pages. 2006.

Vol. 4012: T. Washio, A. Sakurai, K. Nakajima, H. Takeda, S. Tojo, M. Yokoo (Eds.), New Frontiers in Artificial Intelligence. XIII, 484 pages. 2006.

Vol. 4008: J.C. Augusto, C.D. Nugent (Eds.), Designing Smart Homes. XI, 183 pages. 2006.

Vol. 4005: G. Lugosi, H.U. Simon (Eds.), Learning Theory. XI, 656 pages. 2006.

Vol. 3978: B. Hnich, M. Carlsson, F. Fages, F. Rossi (Eds.), Recent Advances in Constraints. VIII, 179 pages. 2006.

Vol. 3963: O. Dikenelli, M.-P. Gleizes, A. Ricci (Eds.), Engineering Societies in the Agents World VI. XII, 303 pages. 2006.

Vol. 3960: R. Vieira, P. Quaresma, M.d.G.V. Nunes, N.J. Mamede, C. Oliveira, M.C. Dias (Eds.), Computational Processing of the Portuguese Language. XII, 274 pages. 2006.

Vol. 3955: G. Antoniou, G. Potamias, C. Spyropoulos, D. Plexousakis (Eds.), Advances in Artificial Intelligence. XVII, 611 pages. 2006.

Vol. 3949: F. A. Savacı (Ed.), Artificial Intelligence and Neural Networks. IX, 227 pages. 2006.

Vol. 3946: T.R. Roth-Berghofer, S. Schulz, D.B. Leake (Eds.), Modeling and Retrieval of Context. XI, 149 pages. 2006.

Vol. 3944: J. Quiñonero-Candela, I. Dagan, B. Magnini, F. d'Alché-Buc (Eds.), Machine Learning Challenges. XIII, 462 pages. 2006.

Vol. 3930: D.S. Yeung, Z.-Q. Liu, X.-Z. Wang, H. Yan (Eds.), Advances in Machine Learning and Cybernetics. XXI, 1110 pages. 2006.

Vol. 3918: W.K. Ng, M. Kitsuregawa, J. Li, K. Chang (Eds.), Advances in Knowledge Discovery and Data Mining. XXIV, 879 pages. 2006.

Vol. 3913: O. Boissier, J. Padget, V. Dignum, G. Lindemann, E. Matson, S. Ossowski, J.S. Sichman, J. Vázquez-Salceda (Eds.), Coordination, Organizations, Institutions, and Norms in Multi-Agent Systems. XII, 259 pages. 2006.

Vol. 3910: S.A. Brueckner, G.D.M. Serugendo, D. Hales, F. Zambonelli (Eds.), Engineering Self-Organising Systems. XII, 245 pages. 2006.

Vol. 3904: M. Baldoni, U. Endriss, A. Omicini, P. Torroni (Eds.), Declarative Agent Languages and Technologies III. XII, 245 pages. 2006.

Vol. 3900: F. Toni, P. Torroni (Eds.), Computational Logic in Multi-Agent Systems. XVII, 427 pages. 2006.

Vol. 3899: S. Frintrop, VOCUS: A Visual Attention System for Object Detection and Goal-Directed Search. XIV, 216 pages. 2006.

Vol. 3898: K. Tuyls, P.J. 't Hoen, K. Verbeeck, S. Sen (Eds.), Learning and Adaption in Multi-Agent Systems. X, 217 pages. 2006.

Vol. 3891: J.S. Sichman, L. Antunes (Eds.), Multi-Agent-Based Simulation VI. X, 191 pages. 2006.

Vol. 3890: S.G. Thompson, R. Ghanea-Hercock (Eds.), Defence Applications of Multi-Agent Systems. XII, 141 pages. 2006.

Vol. 3885: V. Torra, Y. Narukawa, A. Valls, J. Domingo-Ferrer (Eds.), Modeling Decisions for Artificial Intelligence. XII, 374 pages. 2006.

Vol. 3881: S. Gibet, N. Courty, J.-F. Kamp (Eds.), Gesture in Human-Computer Interaction and Simulation. XIII, 344 pages. 2006.

Vol. 3874: R. Missaoui, J. Schmidt (Eds.), Formal Concept Analysis. X, 309 pages. 2006.

Vol. 3873: L. Maicher, J. Park (Eds.), Charting the Topic Maps Research and Applications Landscape. VIII, 281 pages. 2006.

Vol. 3864: Y. Cai, J. Abascal (Eds.), Ambient Intelligence in Everyday Life. XII, 323 pages. 2006.

Vol. 3863: M. Kohlhase (Ed.), Mathematical Knowledge Management. XI, 405 pages. 2006.

Vol. 3862: R.H. Bordini, M. Dastani, J. Dix, A.E.F. Seghrouchni (Eds.), Programming Multi-Agent Systems. XIV, 267 pages. 2006.

Vol. 3849: I. Bloch, A. Petrosino, A.G.B. Tettamanzi (Eds.), Fuzzy Logic and Applications. XIV, 438 pages. 2006.

Vol. 3848: J.-F. Boulicaut, L. De Raedt, H. Mannila (Eds.), Constraint-Based Mining and Inductive Databases. X, 401 pages. 2006.

Vol. 3847: K.P. Jantke, A. Lunzer, N. Spyratos, Y. Tanaka (Eds.), Federation over the Web. X, 215 pages. 2006.

Vol. 3835: G. Sutcliffe, A. Voronkov (Eds.), Logic for Programming, Artificial Intelligence, and Reasoning. XIV, 744 pages. 2005.

Vol. 3830: D. Weyns, H. V.D. Parunak, F. Michel (Eds.), Environments for Multi-Agent Systems II. VIII, 291 pages. 2006.

Vol. 3817: M. Faundez-Zanuy, L. Janer, A. Esposito, A. Satue-Villar, J. Roure, V. Espinosa-Duro (Eds.), Nonlinear Analyses and Algorithms for Speech Processing. XII, 380 pages. 2006.

Vol. 3814: M. Maybury, O. Stock, W. Wahlster (Eds.), Intelligent Technologies for Interactive Entertainment. XV, 342 pages. 2005.

¥479.00元

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	Francisco Casacuberta
Amos Beimel	Alexey Chernov
Jochen Blath	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 Ryabko
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

Jose L. Balcázar, Philip M. Long, and Frank Stephan

The conference "Algorithmic Learning Theory 2006" is dedicated to studies of learning from a mathematical and algorithmic perspective. Researchers consider various abstract models of the problem of learning and investigate how the learning goal in such a setting can be formulated and achieved. These models describe ways to define

- the goal of learning,
- how the learner retrieves information about its environment,
- how to form of the learner's models of the world (in some cases).

Retrieving information is in some models is passive where the learner just views a stream of data. In other models, the learner is more active, asking questions or learning from its actions. Besides explicit formulation of hypotheses in an abstract language with respect to some indexing system, there are also more implicit methods like making predictions according to the current hypothesis on some arguments which then are evaluated with respect to their correctness and wrong predictions (coming from wrong hypotheses) incur some loss on the learner. In the following, a more detailed introduction is given to the five invited talks and then to the regular contributions.

Gunnar Rätsch works on boosting and support vector machines. His is also interested in online-learning, optimisation theory, new inference principles and new types of machine learning problems. He also applies his results to real word problems from computational biology and chemistry. In his invited talk for ALT 2006, Gunnar spoke about using boosting techniques to solve semi-infinite linear programs, which can be used to address a wide variety of learning problems, including learning to make complex predictions.

Carole Goble works on the World-Wide Web, particularly the Semantic Web and Electronic Science / Grids. As the name suggests, the Semantic Web aims to facilitate the expression and use of meaning on the World-Wide Web. Electronic Science is scientific investigation in which groups are distributed globally. In her invited lecture for DS 2006, Garole Goble presented these two areas and laid out why these two areas depend on each other.

Hans Ulrich Simon studies the complexity of learning, that is, how much of resources of various types are needed for solving theoretically formulated learning problems. In particular, he has worked on query-learning and statistical models. In his invited talk for ALT 2006, Hans Ulrich Simon described work on the learnability of function classes using statistical queries, in which an algorithm interacts with the environment by asking for estimates of probabilities. The model is motivated because previous work had shown that algorithms that obey such a restriction can be made robust against certain kinds of noise. For finite classes, Hans described connections between the complexity of learning with statistical queries

J.L. Balcázar, P.M. Long, and F. Stephan (Eds.): ALT 2006, LNAI 4264, pp. 1–9, 2006.

and the structure of the matrix of correlations between pairs of possible target functions. The structure is captured by the spectral norm of this matrix.

Padhraic Smyth works on all aspects linked to large scale databases as they are found in many applications. To extract and retrieve useful information from such large data bases is an important practical problem. For that reason, his research focusses on using large databases to build descriptive models that are both accurate and understandable. His invited talk for DS 2006 is on data-driven discovery with statistical approaches. Generative probabilistic models have already been proven a useful framework in machine learning from scientific data and the key ideas of this research include (a) representing complex stochastic phenomena using the structured language of graphical models, (b) using latent (hidden) variables for inference about unobserved phenomena and (c) leveraging Bayesian ideas for learning and predicting. Padhraic Smyth began his talk with a brief review of learning from data with hidden variables and then discussed some recent work in this area.

Andrew Y. Ng has research interests in machine learning and pattern recognition, statistical artificial intelligence, reinforcement learning and adaptive control algorithms for text and web data processing. He presented the joint invited talk of ALT 2006 and DS 2006. His talk was on algorithms for control that learn by observing the behaviors of competent agents, rather than through trial and error, the traditional reinforcement learning approach.

The Presentation for the E. M. Gold Award. The first contributed talk presented at ALT 2006 was the talk "Learning unions of $\omega(1)$-dimensional rectangles" by Alp Atici and Rocco Servedio for which the first author received the E. M. Gold Award, as the program committee felt it was the best contribution submitted to ALT 2006 which is co-authored by a student. Atici and Servedio study the learnability of unions of rectangles over $\{0, 1, \ldots, b-1\}^n$ in dependence of b and n. They give algorithms polynomial in n and $\log b$ to learn concepts which are the majority of polynomially many or the union of polylogarithmically many rectangles of dimension a bit below $\log(n \log b)$ and $\log^2(n \log b)$, respectively.

Query Learning. Query Learning is a learning model where a learner or pupil asks a teacher questions about the concept to be learned. One important component of this model is a formal query language used during the learning process; the teacher has to answer every query posed in this language correctly. In this model, the complexity of a learning problem is the maximum number of queries needed by the best learning algorithm provided that the answers of the teacher meet the given specifications; however the teacher himself can be adversary in the sense that he can make the learner to learn as slow as possible as long as he does not violate the constraints. In some settings, also probabilistic teachers are considered instead of adversarial ones.

Nader H. Bshouty and Ehab Wattad investigate the question of how to learn halfspaces with random consistent hypothesis. In their model, the learner combines in each round several randomly selected halfspaces consistent with all data seen so far by majority vote to one object and then queries the teacher whether

these objects are correct. If so, the learner has succeeded; otherwise the teacher returns a counterexample where the hypothesis and the concept to be learnt disagree. In addition to the teacher, the learner has access to a random oracle returning half spaces consistent with the counterexamples seen so far. The authors show that this algorithm needs roughly only two thirds as many queries to the teacher as the best known previous algorithm working with single halfspaces as hypotheseses space.

Matti Kääriäinen deals with the setting where the learner receives mostly unlabeled data, but can actively ask a teacher to label some of the data. Most previous work on this topic has concerned the realizable case, in which some member of a concept class achieves perfect accuracy. Kääriäinen considers the effects of relaxing this constraint in different ways on what can be proved about active learning algorithms.

Jorge Castro extends the setting of exact learning with queries into the world of quantum mechanics. He obtains counterparts of a number of results on exact learning; the new results hold for algorithms that can ask queries that exploit quantum effects.

The complexity of teaching. Learning and teaching are viewing the learning process from two sides. While learning mainly focusses on the aspect of how to extract information from the teacher, teaching focusses on the question of how to help a pupil to learn fast; in the most pessimistic models, the teacher must *force* learning. In this model it can be more interesting to consider randomized or adversarial learners than cooperative ones; a teacher and a cooperative pupil might agree on some coding which permits rapid learning success. Nevertheless, the learner should have some type of consistency constraint since otherwise the teacher cannot force the learner to update wrong hypotheses.

Frank Balbach and Thomas Zeugmann consider in their paper a setting where the learning task consists only of finitely many concepts and the learner keeps any hypothesis until it becomes inconsistent with the current datum presented by the teacher; at that moment the learner revises the hypothesis to a new one chosen from all consistent hypothesis at random with respect to the uniform distribution. The authors show that it is NP-hard to find out whether a good teacher might force the learners to learn a given polynomial-sized class in a given time with high probability. Furthermore, the choice of the sequence on which the learners would succeed is hard; as otherwise one could simulate the learners on this sequence and retrieve their expected behaviour from that knowledge.

Inductive Inference and its complexity. Inductive inference is the learning-theoretic counterpart to recursion theory and studies the learnability of classes of recursive or recursively enumerable sets in the limit. Gold's fundamental paradigm is that such a class is given and that the learner sees an infinite sequence in arbitrary, perhaps adversary order containing all the elements of the set but nothing else except perhaps pause symbols. From these data and some implicit knowledge about the class the learner makes a finite sequence of conjectures such that the last one is an index for the set to be learned, that is, an algorithm which