Sanjay Jain
Hans Ulrich Simon
Etsuji Tomita (Eds.)

LNAI 3734

Algorithmic Learning Theory

16th International Conference, ALT 2005
Singapore, October 2005
Proceedings

Springer

Sanjay Jain Hans Ulrich Simon
Etsuji Tomita (Eds.)

Algorithmic Learning Theory

16th International Conference, ALT 2005
Singapore, October 8-11, 2005
Proceedings

Springer

Series Editors

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

Volume Editors

Sanjay Jain
National University of Singapore
School of Computing, Department of Computer Science
117543 Singapore
E-mail: sanjay@comp.nus.edu.sg

Hans Ulrich Simon
Ruhr-Universität Bochum
Fakultät für Mathematik, Lehrstuhl Mathematik und Informatik
44780 Bochum, Germany
E-mail: simon@lmi.ruhr-uni-bochum.de

Etsuji Tomita
The University of Electro-Communications
Department of Information and Communication Engineering
Chofugaoka 1–5–1, Chofu, Tokyo 182–8585, Japan
E-mail: tomita@ice.uec.ac.jp

Library of Congress Control Number: 2005933042

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

ISSN 0302-9743
ISBN-10 3-540-29242-X Springer Berlin Heidelberg New York
ISBN-13 978-3-540-29242-5 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springeronline.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11564089 06/3142 5 4 3 2 1 0

Lecture Notes in Artificial Intelligence 3734
Edited by J. G. Carbonell and J. Siekmann

Subseries of Lecture Notes in Computer Science

Lecture Notes in Artificial Intelligence (LNAI)

Vol. 3734: S. Jain, H.U. Simon, E. Tomita (Eds.), Algorithmic Learning Theory. XII, 491 pages. 2005.

Vol. 3720: J. Gama, R. Camacho, P. Brazdil, A. Jorge, L. Torgo (Eds.), Machine Learning: ECML 2005. XXXI, 770 pages. 2005.

Vol. 3717: B. Gramlich (Ed.), Frontiers of Combining Systems. X, 321 pages. 2005.

Vol. 3702: B. Beckert (Ed.), Automated Reasoning with Analytic Tableaux and Related Methods. XIII, 343 pages. 2005.

Vol. 3698: U. Furbach (Ed.), KI 2005: Advances in Artificial Intelligence. XIII, 409 pages. 2005.

Vol. 3690: M. Pěchouček, P. Petta, L.Z. Varga (Eds.), Multi-Agent Systems and Applications IV. XVII, 667 pages. 2005.

Vol. 3684: R. Khosla, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part IV. LXXIX, 933 pages. 2005.

Vol. 3683: R. Khosla, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part III. LXXX, 1397 pages. 2005.

Vol. 3682: R. Khosla, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part II. LXXIX, 1371 pages. 2005.

Vol. 3681: R. Khosla, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part I. LXXX, 1319 pages. 2005.

Vol. 3673: S. Bandini, S. Manzoni (Eds.), AI*IA 2005: Advances in Artificial Intelligence. XIV, 614 pages. 2005.

Vol. 3662: C. Baral, G. Greco, N. Leone, G. Terracina (Eds.), Logic Programming and Nonmonotonic Reasoning. XIII, 454 pages. 2005.

Vol. 3661: T. Panayiotopoulos, J. Gratch, R. Aylett, D. Ballin, P. Olivier, T. Rist (Eds.), Intelligent Virtual Agents. XIII, 506 pages. 2005.

Vol. 3658: V. Matoušek, P. Mautner, T. Pavelka (Eds.), Text, Speech and Dialogue. XV, 460 pages. 2005.

Vol. 3651: R. Dale, K.-F. Wong, J. Su, O.Y. Kwong (Eds.), Natural Language Processing – IJCNLP 2005. XXI, 1031 pages. 2005.

Vol. 3642: D. Ślezak, J. Yao, J.F. Peters, W. Ziarko, X. Hu (Eds.), Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, Part II. XXIII, 738 pages. 2005.

Vol. 3641: D. Ślezak, G. Wang, M. Szczuka, I. Düntsch, Y. Yao (Eds.), Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, Part I. XXIV, 742 pages. 2005.

Vol. 3632: R. Nieuwenhuis (Ed.), Automated Deduction – CADE-20. XIII, 459 pages. 2005.

Vol. 3630: M.S. Capcarrere, A.A. Freitas, P.J. Bentley, C.G. Johnson, J. Timmis (Eds.), Advances in Artificial Life. XIX, 949 pages. 2005.

Vol. 3626: B. Ganter, G. Stumme, R. Wille (Eds.), Formal Concept Analysis. X, 349 pages. 2005.

Vol. 3625: S. Kramer, B. Pfahringer (Eds.), Inductive Logic Programming. XIII, 427 pages. 2005.

Vol. 3620: H. Muñoz-Avila, F. Ricci (Eds.), Case-Based Reasoning Research and Development. XV, 654 pages. 2005.

Vol. 3614: L. Wang, Y. Jin (Eds.), Fuzzy Systems and Knowledge Discovery, Part II. XLI, 1314 pages. 2005.

Vol. 3613: L. Wang, Y. Jin (Eds.), Fuzzy Systems and Knowledge Discovery, Part I. XLI, 1334 pages. 2005.

Vol. 3607: J.-D. Zucker, L. Saitta (Eds.), Abstraction, Reformulation and Approximation. XII, 376 pages. 2005.

Vol. 3596: F. Dau, M.-L. Mugnier, G. Stumme (Eds.), Conceptual Structures: Common Semantics for Sharing Knowledge. XI, 467 pages. 2005.

Vol. 3593: V. Mařík, R. W. Brennan, M. Pěchouček (Eds.), Holonic and Multi-Agent Systems for Manufacturing. XI, 269 pages. 2005.

Vol. 3587: P. Perner, A. Imiya (Eds.), Machine Learning and Data Mining in Pattern Recognition. XVII, 695 pages. 2005.

Vol. 3584: X. Li, S. Wang, Z.Y. Dong (Eds.), Advanced Data Mining and Applications. XIX, 835 pages. 2005.

Vol. 3581: S. Miksch, J. Hunter, E. Keravnou (Eds.), Artificial Intelligence in Medicine. XVII, 547 pages. 2005.

Vol. 3577: R. Falcone, S. Barber, J. Sabater-Mir, M.P. Singh (Eds.), Trusting Agents for Trusting Electronic Societies. VIII, 235 pages. 2005.

Vol. 3575: S. Wermter, G. Palm, M. Elshaw (Eds.), Biomimetic Neural Learning for Intelligent Robots. IX, 383 pages. 2005.

Vol. 3571: L. Godo (Ed.), Symbolic and Quantitative Approaches to Reasoning with Uncertainty. XVI, 1028 pages. 2005.

Vol. 3559: P. Auer, R. Meir (Eds.), Learning Theory. XI, 692 pages. 2005.

Vol. 3558: V. Torra, Y. Narukawa, S. Miyamoto (Eds.), Modeling Decisions for Artificial Intelligence. XII, 470 pages. 2005.

Vol. 3554: A. Dey, B. Kokinov, D. Leake, R. Turner (Eds.), Modeling and Using Context. XIV, 572 pages. 2005.

Vol. 3550: T. Eymann, F. Klügl, W. Lamersdorf, M. Klusch, M.N. Huhns (Eds.), Multiagent System Technologies. XI, 246 pages. 2005.

Vol. 3539: K. Morik, J.-F. Boulicaut, A. Siebes (Eds.), Local Pattern Detection. XI, 233 pages. 2005.

Vol. 3538: L. Ardissono, P. Brna, A. Mitrovic (Eds.), User Modeling 2005. XVI, 533 pages. 2005.

Vol. 3533: M. Ali, F. Esposito (Eds.), Innovations in Applied Artificial Intelligence. XX, 858 pages. 2005.

Vol. 3528: P.S. Szczepaniak, J. Kacprzyk, A. Niewiadomski (Eds.), Advances in Web Intelligence. XVII, 513 pages. 2005.

Vol. 3518: T.B. Ho, D. Cheung, H. Liu (Eds.), Advances in Knowledge Discovery and Data Mining. XXI, 864 pages. 2005.

Vol. 3508: P. Bresciani, P. Giorgini, B. Henderson-Sellers, G. Low, M. Winikoff (Eds.), Agent-Oriented Information Systems II. X, 227 pages. 2005.

Vol. 3505: V. Gorodetsky, J. Liu, V. Skormin (Eds.), Autonomous Intelligent Systems: Agents and Data Mining. XIII, 303 pages. 2005.

Vol. 3501: B. Kégl, G. Lapalme (Eds.), Advances in Artificial Intelligence. XV, 458 pages. 2005.

Vol. 3492: P. Blache, E. Stabler, J. Busquets, R. Moot (Eds.), Logical Aspects of Computational Linguistics. X, 363 pages. 2005.

Vol. 3490: L. Bolc, Z. Michalewicz, T. Nishida (Eds.), Intelligent Media Technology for Communicative Intelligence. X, 259 pages. 2005.

Vol. 3488: M.-S. Hacid, N.V. Murray, Z.W. Raś, S. Tsumoto (Eds.), Foundations of Intelligent Systems. XIII, 700 pages. 2005.

Vol. 3487: J. Leite, P. Torroni (Eds.), Computational Logic in Multi-Agent Systems. XII, 281 pages. 2005.

Vol. 3476: J. Leite, A. Omicini, P. Torroni, P. Yolum (Eds.), Declarative Agent Languages and Technologies II. XII, 289 pages. 2005.

Vol. 3464: S.A. Brueckner, G.D.M. Serugendo, A. Karageorgos, R. Nagpal (Eds.), Engineering Self-Organising Systems. XIII, 299 pages. 2005.

Vol. 3452: F. Baader, A. Voronkov (Eds.), Logic for Programming, Artificial Intelligence, and Reasoning. XI, 562 pages. 2005.

Vol. 3451: M.-P. Gleizes, A. Omicini, F. Zambonelli (Eds.), Engineering Societies in the Agents World V. XIII, 349 pages. 2005.

Vol. 3446: T. Ishida, L. Gasser, H. Nakashima (Eds.), Massively Multi-Agent Systems I. XI, 349 pages. 2005.

Vol. 3445: G. Chollet, A. Esposito, M. Faundez-Zanuy, M. Marinaro (Eds.), Nonlinear Speech Modeling and Applications. XIII, 433 pages. 2005.

Vol. 3438: H. Christiansen, P.R. Skadhauge, J. Villadsen (Eds.), Constraint Solving and Language Processing. VIII, 205 pages. 2005.

Vol. 3430: S. Tsumoto, T. Yamaguchi, M. Numao, H. Motoda (Eds.), Active Mining. XII, 349 pages. 2005.

Vol. 3419: B. Faltings, A. Petcu, F. Fages, F. Rossi (Eds.), Recent Advances in Constraints. X, 217 pages. 2005.

Vol. 3416: M. Böhlen, J. Gamper, W. Polasek, M.A. Wimmer (Eds.), E-Government: Towards Electronic Democracy. XIII, 311 pages. 2005.

Vol. 3415: P. Davidsson, B. Logan, K. Takadama (Eds.), Multi-Agent and Multi-Agent-Based Simulation. X, 265 pages. 2005.

Vol. 3403: B. Ganter, R. Godin (Eds.), Formal Concept Analysis. XI, 419 pages. 2005.

Vol. 3398: D.-K. Baik (Ed.), Systems Modeling and Simulation: Theory and Applications. XIV, 733 pages. 2005.

Vol. 3397: T.G. Kim (Ed.), Artificial Intelligence and Simulation. XV, 711 pages. 2005.

Vol. 3396: R.M. van Eijk, M.-P. Huget, F. Dignum (Eds.), Agent Communication. X, 261 pages. 2005.

Vol. 3394: D. Kudenko, D. Kazakov, E. Alonso (Eds.), Adaptive Agents and Multi-Agent Systems II. VIII, 313 pages. 2005.

Vol. 3392: D. Seipel, M. Hanus, U. Geske, O. Bartenstein (Eds.), Applications of Declarative Programming and Knowledge Management. X, 309 pages. 2005.

Vol. 3374: D. Weyns, H. V.D. Parunak, F. Michel (Eds.), Environments for Multi-Agent Systems. X, 279 pages. 2005.

Vol. 3371: M.W. Barley, N. Kasabov (Eds.), Intelligent Agents and Multi-Agent Systems. X, 329 pages. 2005.

Vol. 3369: V. R. Benjamins, P. Casanovas, J. Breuker, A. Gangemi (Eds.), Law and the Semantic Web. XII, 249 pages. 2005.

Vol. 3366: I. Rahwan, P. Moraitis, C. Reed (Eds.), Argumentation in Multi-Agent Systems. XII, 263 pages. 2005.

Vol. 3359: G. Grieser, Y. Tanaka (Eds.), Intuitive Human Interfaces for Organizing and Accessing Intellectual Assets. XIV, 257 pages. 2005.

Vol. 3346: R.H. Bordini, M. Dastani, J. Dix, A.E.F. Seghrouchni (Eds.), Programming Multi-Agent Systems. XIV, 249 pages. 2005.

Vol. 3345: Y. Cai (Ed.), Ambient Intelligence for Scientific Discovery. XII, 311 pages. 2005.

Vol. 3343: C. Freksa, M. Knauff, B. Krieg-Brückner, B. Nebel, T. Barkowsky (Eds.), Spatial Cognition IV. XIII, 519 pages. 2005.

Vol. 3339: G.I. Webb, X. Yu (Eds.), AI 2004: Advances in Artificial Intelligence. XXII, 1272 pages. 2004.

Vol. 3336: D. Karagiannis, U. Reimer (Eds.), Practical Aspects of Knowledge Management. X, 523 pages. 2004.

Vol. 3327: Y. Shi, W. Xu, Z. Chen (Eds.), Data Mining and Knowledge Management. XIII, 263 pages. 2005.

Vol. 3315: C. Lemaître, C.A. Reyes, J.A. González (Eds.), Advances in Artificial Intelligence – IBERAMIA 2004. XX, 987 pages. 2004.

Vol. 3303: J.A. López, E. Benfenati, W. Dubitzky (Eds.), Knowledge Exploration in Life Science Informatics. X, 249 pages. 2004.

Vol. 3301: G. Kern-Isberner, W. Rödder, F. Kulmann (Eds.), Conditionals, Information, and Inference. XII, 219 pages. 2005.

Vol. 3276: D. Nardi, M. Riedmiller, C. Sammut, J. Santos-Victor (Eds.), RoboCup 2004: Robot Soccer World Cup VIII. XVIII, 678 pages. 2005.

Vol. 3275: P. Perner (Ed.), Advances in Data Mining. VIII, 173 pages. 2004.

Preface

This volume contains the papers presented at the 16th Annual International Conference on Algorithmic Learning Theory (ALT 2005), which was held in Singapore (Republic of Singapore), October 8–11, 2005. The main objective of the conference is to provide an interdisciplinary forum for the discussion of the theoretical foundations of machine learning as well as their relevance to practical applications. The conference was co-located with the 8th International Conference on Discovery Science (DS 2005). The conference was also held in conjunction with the centennial celebrations of the National University of Singapore.

The volume includes 30 technical contributions, which were selected by the program committee from 98 submissions. It also contains the ALT 2005 invited talks presented by Chih-Jen Lin (National Taiwan University, Taipei, Taiwan) on "Training Support Vector Machines via SMO-type Decomposition Methods," and by Vasant Honavar (Iowa State University, Ames, Iowa, USA) on "Algorithms and Software for Collaborative Discovery from Autonomous, Semantically Heterogeneous, Distributed, Information Sources." Furthermore, this volume includes an abstract of the joint invited talk with DS 2005 presented by Gary L. Bradshaw (Mississippi State University, Starkville, USA) on "Invention and Artificial Intelligence," and abstracts of the invited talks for DS 2005 presented by Ross D. King (The University of Wales, Aberystwyth, UK) on "The Robot Scientist Project," and by Neil Smalheiser (University of Illinois at Chicago, Chicago, USA) on "The Arrowsmith Project: 2005 Status Report." The complete versions of these papers are published in the DS 2005 proceedings (Lecture Notes in Computer Science Vol. 3735).

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 Rotem Bennet for his paper, "Learning Attribute-Efficiently with Corrupt Oracles," co-authored by Nader Bshouty.

This conference was the 16th in a series of annual conferences established in 1990. Continuation of the ALT series is supervised by its steering committee consisting of: Thomas Zeugmann (Hokkaido University, Sapporo, Japan), Chair; Arun Sharma (Queensland University of Technology, Brisbane, Australia), Co-chair; Naoki Abe (IBM Thomas J. Watson Research Center, Yorktown, USA); Klaus Peter Jantke (FIT Leipzig e.V., Germany); Roni Khardon (Tufts University, Medford, USA); Phil Long (Columbia University, New York, USA); Hiroshi Motoda (Osaka University, Japan); Akira Maruoka (Tohoku University, Sendai, Japan); Luc De Raedt (Albert-Ludwigs-University, Freiburg, Germany); Takeshi Shinohara (Kyushu Institute of Technology, Iizuka, Japan); Osamu Watanabe (Tokyo Institute of Technology, Japan).

We would like to thank all the individuals and institutions that contributed to the success of the conference: the authors for submitting papers; the invited

Organization

Conference Chair

Sanjay Jain	National University of Singapore, Republic of Singapore

Program Committee

Hans Ulrich Simon	Ruhr-Universität Bochum, Germany (Chair)
Etsuji Tomita	The University of Electro-Communications, Japan (Chair)
Dana Angluin	Yale University, USA
Hiroki Arimura	Hokkaido University, Japan
John Case	University of Delaware, USA
Nello Cristianini	University of California at Davis, USA
Victor Dalmau	Universitat Pompeu Fabra, Spain
Claudio Gentile	Università dell'Insubria, Italy
Vasant Honavar	Iowa State University, USA
Satoshi Kobayashi	The University of Electro-Communications, Japan
Phil Long	Columbia University, USA
Frank Stephan	National University of Singapore
Esko Ukkonen	University of Helsinki, Finland
Vladimir Vovk	Royal Holloway, UK
Manfred Warmuth	University of California at Santa Cruz, USA
Kenji Yamanishi	NEC Corporation, Japan
Takashi Yokomori	Waseda University, Japan
Thomas Zeugmann	Hokkaido University, Japan

Local Arrangements

Lee Wee Sun	National University of Singapore, Republic of Singapore

Subreferees

Ilkka Autio
Francis Bach
Asa Ben-Hur
Tijl de Bie
Jorge Castro
Nicolò Cesa-Bianchi
Jiang Chen
Ricard Gavaldà
Kazuhiro Hotta
Matti Kääriäinen
Jyrki Kivinen
Mikko Koivisto
Dima Kuzmin
Michael Last
Jun Liao
Huma Lodhi
Satoshi Morinaga
Sayan Muckherjee
Jaakko Peltonen
Jan Poland
Roman Rosipal
Yasubumi Sakakibara
Johan Suykens
Yasuhiro Tajima
Haruhisa Takahashi
Jun-ichi Takeuchi
Noriyuki Tanida
Jason Weston

Sponsoring Institutions

Lee Foundation, Singapore, Republic of Singapore

Asian Office of Aerospace Research and Development, US Air Force

School of Computing, National University of Singapore

Fakultät für Mathematik, Ruhr-Universität Bochum

Division of Computer Science, Hokkaido University

Institute for Theoretical Computer Science, University of Lübeck

MBZ Marketing Büro Zeugmann

Table of Contents

PAC-Learning

Query-Learning

Inductive Inference

Language Learning

Learning and Logic

Learning from Expert Advice

Online Learning

Defensive Forecasting

Teaching

Editors' Introduction

Sanjay Jain, Hans Ulrich Simon, and Etsuji Tomita

"Learning" is a complex phenomenon that is studied in different scientific disciplines. A computer program with the ability to "learn" contains mechanisms for gathering and evaluating information and, consequently, for improving its performance. Algorithmic Learning Theory provides a mathematical foundation for the study of learning programs. It is concerned with the design and analysis of learning algorithms. The analysis proceeds in a formal model such as to provide measures for the performance of a learning algorithm or for the inherent hardness of a given learning problem.

The variety of applications for algorithms that learn is reflected in the variety of formal learning models. For instance, we can distinguish between a passive mode of "learning from examples" and active modes of learning where the algorithm has more control about the information that is gathered. As for learning from examples, a further decision is whether we impose statistical assumptions on the sequence of examples or not. Furthermore, we find different success criteria in different models (like "approximate learning" versus "exact learning").

The papers collected in this volume offer a broad view on the current research in the field including studies on several learning models (such as kernel-based learning, pac-learning, query learning, inductive inference, learning from expert advice, online learning and teaching). It comes without saying that these models are getting more and more refined as a response to new challenges in computer science (like taking advantage of the large amount of real-world data that become available in digital form, or like serving the growing need for adaptive techniques in the management and control of Internet applications).

The volume is structured as follows. It first presents the invited lectures and then the regular contributions. The latter are grouped according to thematic categories. In the remainder of the introduction, we provide the reader with a rough "road map".

The invited lecture for ALT 2005 and DS 2005 by Gary Bradshaw compares Invention and Artificial Intelligence and points out some analogies. Starting from two case studies, the invention of the air-plain (1799–1909) and the invention of a model rocket by a group of high school students in rural West Virginia in the late 1950's, it is argued that general principals of invention may be applied to expedite the development of AI systems.

The invited lecture for DS 2005 by Neil R. Smalheiser gives a report on the current status of the Arrowsmith Project. The roots of this project trace back to the 1980's when Don Swanson proposed the concept of "undiscovered public knowledge" and published several examples in which two disparate literatures held complementary pieces of knowledge that, when brought together, made compelling and testable predictions about potential therapies for human disorders. In continuation of this work in the 1990's, Smalheiser and Swanson created

S. Jain, H.U. Simon, and E. Tomita (Eds.): ALT 2005, LNAI 3734, pp. 1–9, 2005.

a computer-assisted search strategy ("Arrowsmith"). The lecture reviews the development until today.

Ross D. King, presenting the second invited lecture for DS 2005, pursues the question whether it is possible to automate the scientific process. This question is of increasing importance because, in many scientific areas, data are being generated much faster than they can be effectively analyzed by humans. In his lecture, King describes a physically implemented robotic system that applies techniques from artificial intelligence to carry out cycles of scientific experimentation. This is fleshed out by describing the application of the system to complex research problems in bioinformatics.

The invited lecture for ALT 2005 by Vasant Honavar (co-authored by Doina Caragea, Jun Zhang, Jie Bao, and Jyotishman Pathak) is concerned with the problem of data-driven knowledge acquisition and decision-making. Honavar describes the hurdles represented by massive size, semantic heterogeneity, autonomy, and distributed nature of the data repositories. He introduces some of the algorithmic and statistical problems that arise in such a setting and presents algorithms for learning classifiers from distributed data that offer rigorous performance guarantees relative to their centralized counterparts. The lecture furthermore presents techniques that help to cope with the problem of semantic heterogeneity.

The second invited lecture for ALT 2005 by Chih-Jen Lin (co-authored by Pai-Hsuen Chen and Rong-En Fan) concerns the convex quadratic optimization problem that has to be solved by Support Vector Machines (SVMs). Since the underlying cost matrix is high-dimensional and dense, this cannot be done by classical methods (like, for example, "Interior Point"). Instead a decomposition method is used. It proceeds iteratively and has, in each iteration, to solve a small-dimensional subproblem. Lin focuses on decomposition methods of the SMO-type and elaborates how the implementation of the procedure for "Working Set Selection" affects the speed of convergence to the optimal solution. In a sense, Lin's lecture offers the opportunity of looking deeply inside a widely used learning machine.

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

Kernel-Based Learning: During the last decade there has been a lot of interest in learning systems that express the "similarity" between two "instances" as an inner product of vectors in an appropriate feature space. The inner product operation is often not carried out explicitly, but reduced to the evaluation of a so-called kernel function which operates on instances of the original space. This offers the opportunity to handle high-dimensional feature spaces in an efficient manner. This strategy, introduced by Vapnik and co-workers in connection with the so-called Support Vector Machine, is a theoretically well founded and very powerful method that, in the years since its introduction, has already outperformed many other learning systems in a wide variety of applications.

Gretton, Bousquet, Smola and Schölkopf continue a line of research where the kernel-based approach is used as a tool for the detection of statistical dependencies. More precisely, they provide a measure of statistical dependence between random variables based on the Hilbert-Schmidt norm of a cross covariance operator. Their method has some theoretically appealing features while being experimentally competitive to existing methods.

Kowalczyk and Chapelle study the surprising phenomenon of "anti-learning", where data sets are represented in such a way that some well-known standard learning techniques lead to hypotheses performing much worse than random guessing. While there seem to exist "real-life" data sets which are of that type, the authors consider artificial data whose analysis sheds some light on this (at first glance) counterintuitive phenomenon. They identify some abstract properties of a given positive definite kernel which inevitably lead to anti-learning (for some standard learning techniques including "large-margin classification" or "nearest neighbour"). They furthermore explain which kernel-transformations convert "anti-learning" into "learning" and which do not.

Bayesian and Statistical Models: Causal networks are graphical representations for "causality" between random variables. Like Bayesian networks or other graphical models they represent knowledge being not as strict as formal logical statements but being quite helpful for what is called "reasoning under uncertainty". Bayesian Inference leads to decisions that minimize a risk function. Bayesian learning refers to the problem of inferring the unknown parameters of a distribution (chosen from a known parameterized class of distributions). Typically the a priory distribution for the unknown parameters gives support to a wide range of parameters, whereas the a posteriori distribution is peaked around the true parameter values.

By means of statistical information we can determine causal networks only up to Markov-equivalence. An equivalence class can be represented by a chordal chain graphs that contains directed and undirected edges. Each undirected edge represents two mutually correlated variables where we cannot know whether causality (if any) goes in one or the other direction. He, Geng and Liang describe a hierarchical partitioning of a class of Markov-equivalent causal networks (given by a chordal chain graph) into finer and finer subclasses up to the point where the partition consists of the single causal networks. They prove that, at each stage of the refinement process, an equivalence class can be again represented as a chordal chain graph.

Variational Bayesian Learning results from Bayesian Learning by introducing a simplifying assumption (in case there are hidden variables) that makes the approach computationally more tractable. Empirically it is known to have good generalization performance in many applications. Watanabe and Watanabe provide some additional theoretical support by proving lower and upper bounds on the stochastic complexity in the Variational Bayesian learning of the mixture of exponential families.

As for classification tasks, the Bayes classifier chooses the label that has the maximum a posteriori probability. In order to apply the formula of Bayes, we have to know the a priori probability and (an approximation of) the class-conditional densities. Thonangi and Pudi suggest a new method for the approximation of the class-conditional densities: they use a class-conditional probability model of the data (based on certain conjunctions of binary features and the corresponding relative frequencies) and determine the density function of maximum entropy that satisfies this model.

PAC-Learning: In the PAC-learning model, the learner receives as input training examples, drawn at random according to an unknown distribution and labeled according to an unknown target function f, and returns a "hypothesis" h that (with high probability of success) is a close approximation of f. While the first papers on pac-learning focussed on binary classification problems with the probability of misclassification as an underlying pseudo-metric, there have been many extensions of the basic model since then.

Fakcharoenphol and Kijsirikul revisit the well-studied problem of converting base binary learners (coping with binary classification problems) into multi-class learners. They reconsider various classical constructions, including "one-versus-all" and the "(adaptive) directed acyclic graph approach", and come up with some new generalization error bounds.

The paper by Ryabko relates "PAC-learnability" and "computability" by showing the following result: the number of examples needed by a recursive PAC-learner to learn a computable function cannot be upper-bounded by any computable function. The result is obtained by applying a negative result on data compression based on Kolmogorov complexity.

The papers by Palmer and Goldberg and by Guttman, Vishwanathan and Williamson are both dealing with so-called Probabilistic Deterministic Finite State Automata (PDFA). A PDFA, in contrast to a DFA (its deterministic counterpart), performs random state transitions and thus represents a probability distribution over strings. In a recent paper from 2004, it was shown by Clark and Thollard that the class of PDFAs is polynomially pac-learnable where the polynomial (which bounds sample size and run-time) depends on the size of the target PDFA, a measure for the "distinguishability" of states and the expected length of strings generated from any state. Their pseudo-metric (measuring the distance between target distribution and the hypothesis) is the KL-divergence. By passing from KL-divergence to the variation distance (still useful for pattern classification tasks), Palmer and Goldberg are able to considerably simplify the proof for PAC-learnability and to remove the dependence on the expected string length. Guttman, Vishwanathan and Williamson extend the result by Clark and Thollard into another direction by generalizing the notion of "distinguishability".

Query-Learning: In the query-learning model, the learner is allowed to ask certain queries to an oracle. The learning performance is measured by the number and type of queries needed to exactly (sometimes approximately) identify an unknown target concept f (where "concept" means "binary-valued function"). Among the most popular query types are "membership queries (MQs)" (asking for the class label of an instance) and "equivalence queries (EQs)" (asking for a "counterexample" to the current hypothesis). Statistical queries can be used to get a statistical estimate for the probability of an event that involves a labeled random example. This query type plays a special role because an algorithm