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

LNAI 4754

Algorithmic Learning Theory

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

Springer

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

Algorithmic Learning Theory

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

Springer

Series Editors

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

Volume Editors

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

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

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

Library of Congress Control Number: 2007935310

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

LNCS Sublibrary: SL 7 – Artificial Intelligence

ISSN 0302-9743
ISBN-10 3-540-75224-2 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-75224-0 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springer.com

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

Lecture Notes in Artificial Intelligence 4754
Edited by J. G. Carbonell and J. Siekmann

Subseries of Lecture Notes in Computer Science

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.

¥603.00元

Preface

This volume contains the papers presented at the 18th International Conference on Algorithmic Learning Theory (ALT 2007), which was held in Sendai (Japan) during October 1–4, 2007. The main objective of the conference was to provide an interdisciplinary forum for high-quality talks with a strong theoretical background and scientific interchange in areas such as query models, on-line learning, inductive inference, algorithmic forecasting, boosting, support vector machines, kernel methods, complexity and learning, reinforcement learning, unsupervised learning and grammatical inference. The conference was co-located with the Tenth International Conference on Discovery Science (DS 2007).

This volume includes 25 technical contributions that were selected from 50 submissions by the Program Committee. It also contains descriptions of the five invited talks of ALT and DS; longer versions of the DS papers are available in the proceedings of DS 2007. These invited talks were presented to the audience of both conferences in joint sessions.

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

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

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

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

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

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

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

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

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

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

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

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

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

July 2007

Marcus Hutter
Rocco A. Servedio
Eiji Takimoto

Table of Contents

Online Learning

Unsupervised Learning

Language Learning

Query Learning

Kernel-Based Learning

Other Directions

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

M. Hutter, R.A. Servedio, and E. Takimoto (Eds.): ALT 2007, LNAI 4754, pp. 1–8, 2007.

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