

Lecture Notes in Artificial Intelligence

2375

Subseries of Lecture Notes in Computer Science

Edited by J. G. Carbonell, and J. Siekmann

Lecture Notes in Computer Science

Edited by G. Goos, J. Hartmanis, and J. van Leeuwen

Springer

Berlin

Heidelberg

New York

Barcelona

Hong Kong

London

Milan

Paris

Tokyo

Jyrki Kivinen Robert H. Sloan (Eds.)

Computational Learning Theory

15th Annual Conference on Computational Learning Theory,
COLT 2002
Sydney, Australia, July 8-10, 2002
Proceedings



Springer

Series Editors

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

Volume Editors

Jyrki Kivinen

Australian National University
Research School of Information Sciences and Engineering
Canberra ACT 0200, Australia
E-mail: Jyrki.Kivinen@anu.edu.au

Robert H. Sloan

University of Illinois at Chicago, Computer Science Department
851 S. Morgan St., Chicago, IL 60607, USA
E-mail: sloan@cs.uic.edu

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Computational learning theory : proceedings / 15th Annual Conference on
Computational Learning Theory, COLT 2002, Sydney, Australia, July 8 - 10,
2002. Jyrki Kivinen ; Robert H. Sloan (ed.). - Berlin ; Heidelberg ; New
York ; Barcelona ; Hong Kong ; London ; Milan ; Paris ; Tokyo : Springer,
2002

(Lecture notes in computer science ; Vol. 2375 : Lecture notes in
artificial intelligence)
ISBN 3-540-43836-X

CR Subject Classification (1998): I.2.6, I.2.3, F.4.1, F.1.1, F.2

ISSN 0302-9743

ISBN 3-540-43836-X Springer-Verlag Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer-Verlag Berlin Heidelberg New York
a member of BertelsmannSpringer Science+Business Media GmbH

<http://www.springer.de>

© Springer-Verlag Berlin Heidelberg 2002
Printed in Germany

Typesetting: Camera-ready by author, data conversion by DA-TeX Gerd Blumenstein
Printed on acid-free paper SPIN 10870415 06/3142 5 4 3 2 1 0

Preface

This volume contains papers presented at the Fifteenth Annual Conference on Computational Learning Theory (COLT 2002) held on the main campus of the University of New South Wales in Sydney, Australia from July 8 to 10, 2002. Naturally, these are papers in the field of computational learning theory, a research field devoted to studying the design and analysis of algorithms for making predictions about the future based on past experiences, with an emphasis on rigorous mathematical analysis.

COLT 2002 was co-located with the Nineteenth International Conference on Machine Learning (ICML 2002) and with the Twelfth International Conference on Inductive Logic Programming (ILP 2002).

Note that COLT 2002 was the first conference to take place after the full merger of the Annual Conference on Computational Learning Theory with the European Conference on Computational Learning Theory. (In 2001 a joint conference consisting of the 5th European Conference on Computational Learning Theory and the 14th Annual Conference on Computational Learning Theory was held; the last independent European Conference on Computational Learning Theory was held in 1999.)

The technical program of COLT 2002 contained 26 papers selected from 55 submissions. In addition, Christos Papadimitriou (University of California at Berkeley) was invited to give a keynote lecture and to contribute an abstract of his lecture to these proceedings.

The Mark Fulk Award is presented annually for the best paper coauthored by a student. This year's award was won by Sandra Zilles for the paper "Merging Uniform Inductive Learners."

April 2002

Jyrki Kivinen
Robert H. Sloan

Thanks and Acknowledgments

We gratefully thank all the individuals and organizations responsible for the success of the conference.

Program Committee

We especially want to thank the program committee: Dana Angluin (Yale), Javed Aslam (Dartmouth), Peter Bartlett (BIOwulf Technologies), Shai Ben-David (Technion), John Case (Univ. of Delaware), Peter Grünwald (CWI), Ralf Herbrich (Microsoft Research), Mark Herbster (University College London), Gábor Lugosi (Pompeu Fabra University), Ron Meir (Technion), Shahar Mendelson (Australian National Univ.), Michael Schmitt (Ruhr-Universität Bochum), Rocco Servedio (Harvard), and Santosh Vempala (MIT).

We also acknowledge the creators of the CyberChair software for making a software package that helped the committee do its work.

Local Arrangements, Co-located Conferences Support

Special thanks go to our conference chair Arun Sharma and local arrangements chair Eric Martin (both at Univ. of New South Wales) for setting up COLT 2002 in Sydney. Rochelle McDonald and Sue Lewis provided administrative support. Claude Sammut in his role as conference chair of ICML and program co-chair of ILP ensured smooth coordination with the two co-located conferences.

COLT Community

For keeping the COLT series going, we thank the COLT steering committee, and especially Chair John Shawe-Taylor and Treasurer John Case for all their hard work. We also thank Stephen Kwek for maintaining the COLT web site at <http://www.learningtheory.org>.

Sponsoring Institution

School of Computer Science and Engineering, University of New South Wales, Australia

Referees

Peter Auer
Andrew Barto
Stephane Boucheron
Olivier Bousquet
Nicolò Cesa-Bianchi
Tapio Elomaa
Ran El-Yaniv
Allan Erskine
Henning Fernau
Jürgen Forster
Dean Foster
Claudio Gentile
Judy Goldsmith
Thore Graepel

Lisa Hellerstein
Daniel Herrmann
Colin de la Higuera
Sean Holden
Marcus Hutter
Sanjay Jain
Yuri Kalnishkan
Makoto Kanazawa
Satoshi Kobayashi
Vladimir Koltchinskii
Matti Kääriäinen
Wee Sun Lee
Shie Mannor
Ryan O'Donnell

Alain Pajor
Gunnar Rätsch
Robert Schapire
John Shawe-Taylor
Takeshi Shinohara
David Shmoys
Yoram Singer
Carl Smith
Frank Stephan
György Turán
Paul Vitányi
Manfred Warmuth
Jon A. Wellner
Robert C. Williamson

Table of Contents

Statistical Learning Theory

Agnostic Learning Nonconvex Function Classes	1
<i>Shahar Mendelson and Robert C. Williamson</i>	
Entropy, Combinatorial Dimensions and Random Averages	14
<i>Shahar Mendelson and Roman Vershynin</i>	
Geometric Parameters of Kernel Machines	29
<i>Shahar Mendelson</i>	
Localized Rademacher Complexities	44
<i>Peter L. Bartlett, Olivier Bousquet, and Shahar Mendelson</i>	
Some Local Measures of Complexity of Convex Hulls and Generalization Bounds	59
<i>Olivier Bousquet, Vladimir Koltchinskii, and Dmitriy Panchenko</i>	

Online Learning

Path Kernels and Multiplicative Updates	74
<i>Eiji Takimoto and Manfred K. Warmuth</i>	
Predictive Complexity and Information	90
<i>Michael V. Vyugin and Vladimir V. Vyugin</i>	
Mixability and the Existence of Weak Complexities	105
<i>Yuri Kalnishkan and Michael V. Vyugin</i>	
A Second-Order Perceptron Algorithm	121
<i>Nicolò Cesa-Bianchi, Alex Conconi, and Claudio Gentile</i>	
Tracking Linear-Threshold Concepts with Winnow	138
<i>Chris Mesterharm</i>	

Inductive Inference

Learning Tree Languages from Text	153
<i>Henning Fernau</i>	
Polynomial Time Inductive Inference of Ordered Tree Patterns with Internal Structured Variables from Positive Data	169
<i>Yusuke Suzuki, Ryuta Akanuma, Takayoshi Shoudai, Tetsuhiro Miyahara, and Tomoyuki Uchida</i>	

Inferring Deterministic Linear Languages 185
Colin de la Higuera and Jose Oncina

Merging Uniform Inductive Learners 201
Sandra Zilles

The Speed Prior: A New Simplicity Measure
 Yielding Near-Optimal Computable Predictions 216
Jürgen Schmidhuber

PAC Learning

New Lower Bounds for Statistical Query Learning 229
Ke Yang

Exploring Learnability between Exact and PAC 244
Nader H. Bshouty, Jeffrey C. Jackson, and Christino Tamon

PAC Bounds for Multi-armed Bandit and Markov Decision Processes 255
Eyal Even-Dar, Shie Mannor, and Yishay Mansour

Bounds for the Minimum Disagreement Problem
 with Applications to Learning Theory 271
Nader H. Bshouty and Lynn Burroughs

On the Proper Learning of Axis Parallel Concepts 287
Nader H. Bshouty and Lynn Burroughs

Boosting

A Consistent Strategy for Boosting Algorithms 303
Gábor Lugosi and Nicolas Vayatis

The Consistency of Greedy Algorithms for Classification 319
Shie Mannor, Ron Meir, and Tong Zhang

Maximizing the Margin with Boosting 334
Gunnar Rätsch and Manfred K. Warmuth

Other Learning Paradigms

Performance Guarantees for Hierarchical Clustering 351
Sanjoy Dasgupta

Self-Optimizing and Pareto-Optimal Policies
 in General Environments Based on Bayes-Mixtures 364
Marcus Hutter

Prediction and Dimension 380
Lance Fortnow and Jack H. Lutz

Invited Talk

Learning the Internet396
Christos Papadimitriou

Author Index397