Lecture Notes in Artificial Intelligence 1967

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
Singapore
Tokyo

Setsuo Arikawa Shinichi Morishita (Eds.)

Discovery Science

Third International Conference, DS 2000
Kyoto, Japan, December 4-6, 2000
Proceedings

Springer

Series Editors

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

Volume Editors

Setsuo Arikawa
Kyushu University, Faculty of Information Science and
Electrical Engineering, Department of Informatics
6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan
E-mail: arikawa@i.kyushu-u.ac.jp

Shinichi Morishita
University of Tokyo, Faculty of Science
Department of Information Science
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
E-mail: moris@is.s.u-tokyo.ac.jp

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Discovery science : third international conference ; proceedings / DS 2000, Kyoto, Japan, December 4 - 6, 2000 / Setsuo Arikawa ; Shinichi Morishita (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Hong Kong ; London ; Milan ; Paris ; Singapore ; Tokyo : Springer, 2000
(Lecture notes in computer science ; Vol. 1967 : Lecture notes in artificial intelligence)
ISBN 3-540-41352-9

CR Subject Classification (1998): I.2, H.2.8, H.3, J.1, J.2

ISBN 3-540-41352-9 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
© Springer-Verlag Berlin Heidelberg 2000
Printed in Germany

Typesetting: Camera-ready by author
Printed on acid-free paper SPIN: 10781098 06/3142 5 4 3 2 1 0

Preface

This volume contains 3 invited papers, 15 regular papers, and 22 poster papers that were selected for presentation at the Third International Conference on Discovery Science (DS 2000), which was held 4-6 December 2000 in Kyoto. The Program Committee selected the contributed papers from 48 submissions.

Three distinguished researchers accepted our invitation to present talks: Jeffrey D. Ullman (Stanford University), Joseph Y. Halpern (Cornell University), and Masami Hagiya (University of Tokyo).

The Program Committee would like to thank all those who submitted papers for consideration and the invited speakers. I would like to thank the Program Committee members, the Local Arrangements Committee members, and the Steering Committee members for their splendid and hard work. Finally, special thanks go to the PC Assistant Shoko Suzuki for her assistance in the development of web pages and the preparation of these proceedings.

September 2000 Shinichi Morishita

Organization

Discovery Science 2000 is organized as part of the activities of the Discovery Science Project sponsored by Grant-in-Aid for Scientific Research in the Priority Area from the Ministry of Education, Science, Sports and Culture (MESSC) of Japan, in cooperation with the Japanese Society for Artificial Intelligence, and with SIG of Data Mining, Japan Society for Software Science and Technology.

Conference Chair

Setsuo Arikawa	Kyushu University

Program Committee

Shinichi Morishita (Chair)	University of Tokyo
Hiroki Arimura	Kyushu University
Roberto Bayardo	IBM Almaden Research Center
Charles Elkan	University of California, San Diego
Peter A. Flach	University of Bristol
Koichi Furukawa	Keio University
Randy Goebel	University of Alberta
Yukiyoshi Kameyama	Kyoto University
Naoki Katoh	Kyoto University
Ross D. King	University of Wales
Nada Lavrac	Jozef Stefan Institute
Katharina Morik	University of Dortmund
Hiroshi Motoda	Osaka University
Raymond T. Ng	University of British Columbia
Koichi Niijima	Kyushu University
Claude Sammut	University of New South Wales
Taisuke Sato	Tokyo Institute of Technology
Etsuya Shibayama	Tokyo Institute of Technology
Kyuseok Shim	Bell Laboratories
Ayumi Shinohara	Kyushu University
Carl H. Smith	University of Maryland
Osamu Watanabe	Tokyo Institute of Technology
Limsoon Wong	Kent Ridge Digital Labs

Local Arrangements Committee

Masahiko Sato (Chair)	Kyoto University
Yukiyoshi Kameyama	Kyoto University
Hiroyuki Kawano	Kyoto University
Sadao Kurohashi	Kyoto University
Ayumi Shinohara	Kyushu University

Steering Committee

Setsuo Arikawa (Chair)	Kyushu University
Yasumasa Kanada	University of Tokyo
Akira Maruoka	Tohoku University
Satoru Miyano	University of Tokyo
Masahiko Sato	Kyoto University
Taisuke Sato	Tokyo Institute of Technology

Table of Contents

Invited Papers

A Survey of Association-Rule Mining 1
Jeffrey D. Ullman

Degrees of Belief, Random Worlds, and Maximum Entropy 15
Joseph Y. Halpern

Discovery and Deduction 17
Masami Hagiya and Koichi Takahashi

Regular Papers

Integrating Information Visualization and Retrieval for Discovering Internet Sources 38
Hayato Ohwada and Fumio Mizoguchi

A Unifying Approach to HTML Wrapper Representation and Learning . . . 50
Gunter Grieser, Klaus P. Jantke, Steffen Lange, and Bernd Thomas

Discovery of Web Communities Based on the Co-occurrence of References . 65
Tsuyoshi Murata

Clustering and Visualization of Large Protein Sequence Databases by Means of an Extension of the Self-Organizing Map 76
Panu Somervuo and Teuvo Kohonen

A Simple Greedy Algorithm for Finding Functional Relations: Efficient Implementation and Average Case Analysis 86
Tatsuya Akutsu, Satoru Miyano, and Satoru Kuhara

Graph-Based Induction for General Graph Structured Data and Its Application to Chemical Compound Data 99
Takashi Matsuda, Tadashi Horiuchi, Hiroshi Motoda, and Takashi Washio

Discovering Characteristic Expressions from Literary Works: A New Text Analysis Method beyond N-Gram Statistics and KWIC 112
Masayuki Takeda, Tetsuya Matsumoto, Tomoko Fukuda, and Ichirō Nanri

Classifying Scenarios Using Belief Decision Trees 127
Zied Elouedi and Khaled Mellouli

A Practical Algorithm to Find the Best Subsequence Patterns 141
Masahiro Hirao, Hiromasa Hoshino, Ayumi Shinohara, Masayuki Takeda, and Setsuo Arikawa

On-Line Estimation of Hidden Markov Model Parameters 155
Jun Mizuno, Tatsuya Watanabe, Kazuya Ueki, Kazuyuki Amano, Eiji Takimoto, and Akira Maruoka

Computationally Efficient Heuristics for If-Then Rule Extraction from Freed-Forward Neural Networks 170
Hyeoncheol Kim

Language Learning with a Neighbor System 183
Yasuhito Mukouchi and Masako Sato

Application of Multivariate Maxwellian Mixture Model to Plasma Velocity Distribution Function 197
Genta Ueno, Nagatomo Nakamura, Tomoyuki Higuchi, Takashi Tsuchiya, Shinobu Machida, and Tohru Araki

Knowledge Discovery from fMRI Brain Images by Logical Regression Analysis 212
Hiroshi Tsukimoto, Mitsuru Kakimoto, Chie Morita, Yoshiaki Kikuchi, Eiko Hatakeyama, and Yoshifumi Miyazaki

Human Discovery Processes Based on Searching Experiments in Virtual Psychological Research Environment 225
Kazuhisa Miwa

Poster Papers

Prediction of Binding Affinities for Protein-ligand Complexes with Neural Network Models 240
Vladimir V. Krepets and Natalya V. Belkina

Automatic and Accurate Determination of the Onset Time of the Quasi-periodic Oscillation 242
Tomoyuki Higuchi

The Role of Choice in Discovery 247
Judith Devaney

Search for New Methods for Assignment of Complex Molecular Spectra . . . 252
Takehiko Tanaka

Computational Analysis for Discovery on the Plasma Waves Observed by Scientific Satellites 255
Yoshiya Kasahara, Ryotaro Niitsu, Yousuke Akimoto, and Toru Sato

Direction Finding of the Waves in Plasma Using Energy Function 258
Yoshitaka Goto, Yoshiya Kasahara, and Toru Sato

Coping the Challenge of Mutagenes Discovery with GUHA+/- for Windows 261
Premysl Zak and Jaroslava Halova

Discovering Interpretable Rules that Explain Customers' Brand Choice Behavior 263
Yukinobu Hamuro, Naoki Katoh, and Katsutoshi Yada

Mining for 4ft Association Rules 268
Jan Rauch and Milan Šimůnek

Rule Discovery Technique Using Genetic Programming Combined with Apriori Algorithm 273
Ayahiko Niimi and Eiichiro Tazaki

Discovery of M-of-N Concepts for Classification 278
Nobuhiro Yugami

Issues in Organizing a Successful Knowledge Discovery Contest 282
Einoshin Suzuki

Knowledge Integration of Rule Mining and Schema Discovering 285
Kohei Maruyama and Kuniaki Uehara

Discovery of Correlation from Multi-stream of Human Motion 290
Mitsuomi Shimada and Kuniaki Uehara

An Appropriate Abstraction for Constructing a Compact Decision Tree . . . 295
Yoshimitsu Kudoh and Makoto Haraguchi

Extraction Positive and Negative Keywords for Web Communities 299
Daisuke Ikeda and Sachio Hirokawa

Nonequilibrium Thermodynamics from Time Series Data Analysis 304
Hiroshi H. Hasegawa, Takashi Washio, Yukari Ishimiya, and Takeshi Saito

Automatic Determination Algorithm for the Optimum Number of States in NL-HMnet 306
Motoyuki Suzuki and Shozo Makino

Comparative Study of Automatic Acquisition Methods of Image Processing Procedures 311
Toshihiro Hamada, Akinobu Shimizu, Toyofumi Saito, Jun-ichi Hasegawa, and Jun-ichiro Toriwaki

Extraction of Authors' Characteristics from Japanese Modern Sentences via N-gram Distribution 315
Tsukasa Matsuura and Yasumasa Kanada

Combination Retrieval for Creating Knowledge from Sparse Document Collection 320
Naohiro Matsumura and Yukio Ohsawa

Discovery of Nominally Conditioned Polynomials Using Neural Networks, Vector Quantizers and Decision Trees 325
Kazumi Saito and Ryohei Nakano

Author Index 331