

Mohand-Said Hacid
Neil V. Murray
Zbigniew W. Raś
Shusaku Tsumoto (Eds.)

LNAI 3488

Foundations of Intelligent Systems

15th International Symposium, ISMIS 2005
Saratoga Springs, NY, USA, May 2005
Proceedings



Springer

TP18-53
M592
2005

Mohand-Said Hacid Neil V. Murray
Zbigniew W. Raś Shusaku Tsumoto (Eds.)

Foundations of Intelligent Systems

15th International Symposium, ISMIS 2005
Saratoga Springs, NY, USA, May 25-28, 2005
Proceedings



E200501341



Springer

Volume Editors

Mohand-Said Hacid
Université Claude Bernard Lyon 1
LIRIS - UFR d'Informatique
69622 Villeurbanne Cedex, France
E-mail: mshacid@bat710.univ-lyon.fr

Neil V. Murray
University at Albany, SUNY
Department of Computer Science
Albany, NY 12222, USA
E-mail: nvm@cs.albany.edu

Zbigniew W. Raś
University of North Carolina
Department of Computer Science
Charlotte, NC 28223, USA
E-mail: ras@uncc.edu

Shusaku Tsumoto
Shimane Medical University
School of Medicine, Department of Medical Informatics
89-1 Enya-cho, Izumo 693-8501, Japan
E-mail: tsumoto@computer.org

Library of Congress Control Number: 2005925177

CR Subject Classification (1998): I.2, H.3, H.2.8, H.4, H.5.2, K.4.4, F.1

ISSN	0302-9743
ISBN-10	3-540-25878-7 Springer Berlin Heidelberg New York
ISBN-13	978-3-540-25878-0 Springer 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. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media
springeronline.com

© Springer-Verlag Berlin Heidelberg 2005
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11425274 06/3142 5 4 3 2 1 0

Lecture Notes in Artificial Intelligence 3488

Edited by J. G. Carbonell and J. Siekmann

Subseries of Lecture Notes in Computer Science

Preface

This volume contains the papers selected for presentation at the 15th International Symposium on Methodologies for Intelligent Systems, ISMIS 2005, held in Saratoga Springs, New York, 25–28 May, 2005. The symposium was organized by SUNY at Albany. It was sponsored by the Army Research Office and by several units of the University at Albany including its Division for Research, College of Arts and Sciences, Department of Computer Science, and Institute for Informatics, Logics, and Security Studies (formerly the Institute for Programming and Logics). ISMIS is a conference series that was started in 1986 in Knoxville, Tennessee. Since then it has been held in Charlotte (North Carolina), Knoxville (Tennessee), Turin (Italy), Trondheim (Norway), Warsaw (Poland), Zakopane (Poland), Lyon (France), and Maebashi City (Japan).

The Program Committee selected the following major areas for ISMIS 2005: intelligent information systems, knowledge discovery and data mining, knowledge information and integration, knowledge representation, logic for artificial intelligence, soft computing, Web intelligence, Web services, and papers dealing with applications of intelligent systems in complex/novel domains.

The contributed papers were selected from almost 200 full draft papers by the Program Committee members: Troels Andreassen, Peter Baumgartner, Boualemn Benatallah, Salima Benbernou, Veronique Benzaken, Petr Berka, Elisa Bertino, Alan Biermann, Jacques Calmet, Sandra Carberry, Juan Carlos Cubero, Luigia Carlucci Aiello, Shu-Ching Chen, Christine Collet, Agnieszka Dardzinska, Ian Davidson, Robert Demolombe, Jitender Deogun, Jon Doyle, Tapio Elomaa, Attilio Giordana, Jerzy Grzymala-Busse, Mirsad Hadzikadic, Reiner Haehnle, Janusz Kacprzyk, Vipul Kashyap, Jan Komorowski, Jacek Koronacki, Tsau Young Lin, Donato Malerba, David Maluf, Davide Martinenghi, Stan Matwin, Natasha Noy, Werner Nutt, James Peters, Jean-Marc Petit, Vijay Raghavan, Jan Rauch, Gilbert Ritschard, Erik Rosenthal, Marie-Christine Rousset, Nahid Shahmehri, Andrzej Skowron, Dominik Slezak, Nicolas Spyrtos, V.S. Subrahmanian, Einoshin Suzuki, Domenico Talia, Yuzuru Tanaka, Farouk Toumani, Athena Vakali, Takashi Washio, Alicja Wieczorkowska, Xindong Wu, Xintao Wu, Ronald Yager, Yiyu Yao, and Saygin Yucel.

The list of additional reviewers: Johan Aberg, Marie Agier, Lefteris Angelis, Annalisa Appice, Khalid Belhajjame, Margherita Berardi, Patrick Bosc, Henrik Bulskov, Franck Capello, Costantina Caruso, Michelangelo Ceci, Min Chen, Ioan Chisalita, Laurence Cholvy, Dario Colazzo, Carmela Comito, Antonio Congiusta, Frédéric Cuppens, Anusch Daemi, Fabien De Marchi, Elizabeth Diaz, Devdatt Dubhashi, Peter Gammie, Arnaud Giacometti, Martin Giese, Paul Goutam, Gianluigi Greco, Samira Hammiche, Thomas Herault, Seung Hyun Im, Kimihito Ito, Mikhail Jiline, Kristofer Johannisson, Fabrice Jouanot, Matti Kääriäinen, Svetlana Kiritchenko, Rasmus Knappe, Lotfi Lakhal, Patrick Lambrix, Tine Lassen, Anne Laurent, Dominique Laurent, Alexandre Lefebvre, João Fernando Lima Alcantara, Gabriela Lindemann, Yann Loyer, James Lu, Perry Mar, Carlo Mastroianni, Mirjam Minor, Dagmar Monett,

Dheerendranath Mundluru, Oliver Obst, Raffale Perego, Aarne Ranta, Brigitte Safar, Lorenza Saitta, Eric San Juan, Kay Schröter, Michele Sebag, Biren Shah, Laurent Simon, Giandomenico Spezzano, Diemo Urbig, Thomas Vestskov Terney, Li-Shiang Tsay, Phan-Luong Viet, Richard Waldinger, Christoph Wernhard, Ying Xie, Jianhua Yan, Ke Yin, José Luis Zechinelli Martini, Na Zhao, and Xiaosi Zhou.

We wish to express our thanks to all ISMIS 2005 reviewers and to Alan Biermann, Jaime Carbonell, and Salvatore Stolfo who presented invited talks at the symposium. We express our appreciation to the sponsors of the symposium and to all who submitted papers for presentation and publication in the proceedings. Our sincere thanks go Floriana Esposito, David Hislop, Robert Meersman, Hiroshi Motoda, Raghu Ramakrishnan, Zbigniew Ras (Chair), Lorenza Saitta, Maria Zemankova, Djamel Zighed, and Ning Zhong who served as members of the ISMIS 2005 Steering Committee, and to both Adrian Tanasescu who was responsible for the website and submission site management and Lynne Casper who worked on all aspects of local arrangements. Also, our thanks are due to Alfred Hofmann of Springer for his continuous support.

February, 2005

M.-S. Hacid, N.V. Murray, Z.W. Raś, S. Tsumoto

Lecture Notes in Artificial Intelligence (LNAI)

- 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. 3488: M.-S. Hacid, N.V. Murray, Z.W. Raś, S. Tsumoto (Eds.), *Foundations of Intelligent Systems*. XIII, 700 pages. 2005.
- Vol. 3452: F. Baader, A. Voronkov (Eds.), *Logic for Programming, Artificial Intelligence, and Reasoning*. XI, 562 pages. 2005.
- Vol. 3419: B. Faltings, A. Petcu, F. Fages, F. Rossi (Eds.), *Constraint Satisfaction and Constraint Logic Programming*. 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. 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.
- Vol. 3265: R.E. Frederking, K.B. Taylor (Eds.), *Machine Translation: From Real Users to Research*. XI, 392 pages. 2004.
- Vol. 3264: G. Paliouras, Y. Sakakibara (Eds.), *Grammatical Inference: Algorithms and Applications*. XI, 291 pages. 2004.
- Vol. 3259: J. Dix, J. Leite (Eds.), *Computational Logic in Multi-Agent Systems*. XII, 251 pages. 2004.
- Vol. 3257: E. Motta, N.R. Shadbolt, A. Stutt, N. Gibbins (Eds.), *Engineering Knowledge in the Age of the Semantic Web*. XVII, 517 pages. 2004.
- Vol. 3249: B. Buchberger, J.A. Campbell (Eds.), *Artificial Intelligence and Symbolic Computation*. X, 285 pages. 2004.
- Vol. 3248: K.-Y. Su, J. Tsujii, J.-H. Lee, O.Y. Kwong (Eds.), *Natural Language Processing – IJCNLP 2004*. XVIII, 817 pages. 2005.
- Vol. 3245: E. Suzuki, S. Arikawa (Eds.), *Discovery Science*. XIV, 430 pages. 2004.
- Vol. 3244: S. Ben-David, J. Case, A. Maruoka (Eds.), *Algorithmic Learning Theory*. XIV, 505 pages. 2004.
- Vol. 3238: S. Biundo, T. Frühwirth, G. Palm (Eds.), *KI 2004: Advances in Artificial Intelligence*. XI, 467 pages. 2004.
- Vol. 3230: J.L. Vicedo, P. Martínez-Barco, R. Muñoz, M. Saiz Noeda (Eds.), *Advances in Natural Language Processing*. XII, 488 pages. 2004.

- Vol. 3229: J.J. Alferes, J. Leite (Eds.), *Logics in Artificial Intelligence*. XIV, 744 pages. 2004.
- Vol. 3228: M.G. Hinchey, J.L. Rash, W.F. Truszkowski, C.A. Rouff (Eds.), *Formal Approaches to Agent-Based Systems*. VIII, 290 pages. 2004.
- Vol. 3215: M.G. Negoita, R.J. Howlett, L.C. Jain (Eds.), *Knowledge-Based Intelligent Information and Engineering Systems, Part III*. LVII, 906 pages. 2004.
- Vol. 3214: M.G. Negoita, R.J. Howlett, L.C. Jain (Eds.), *Knowledge-Based Intelligent Information and Engineering Systems, Part II*. LVIII, 1302 pages. 2004.
- Vol. 3213: M.G. Negoita, R.J. Howlett, L.C. Jain (Eds.), *Knowledge-Based Intelligent Information and Engineering Systems, Part I*. LVIII, 1280 pages. 2004.
- Vol. 3209: B. Berendt, A. Hotho, D. Mladenic, M. van Someren, M. Spiliopoulou, G. Stumme (Eds.), *Web Mining: From Web to Semantic Web*. IX, 201 pages. 2004.
- Vol. 3206: P. Sojka, I. Kopecek, K. Pala (Eds.), *Text, Speech and Dialogue*. XIII, 667 pages. 2004.
- Vol. 3202: J.-F. Boulicaut, F. Esposito, F. Giannotti, D. Pedreschi (Eds.), *Knowledge Discovery in Databases: PKDD 2004*. XIX, 560 pages. 2004.
- Vol. 3201: J.-F. Boulicaut, F. Esposito, F. Giannotti, D. Pedreschi (Eds.), *Machine Learning: ECML 2004*. XVIII, 580 pages. 2004.
- Vol. 3194: R. Camacho, R. King, A. Srinivasan (Eds.), *Inductive Logic Programming*. XI, 361 pages. 2004.
- Vol. 3192: C. Bussler, D. Fensel (Eds.), *Artificial Intelligence: Methodology, Systems, and Applications*. XIII, 522 pages. 2004.
- Vol. 3191: M. Klusch, S. Ossowski, V. Kashyap, R. Unland (Eds.), *Cooperative Information Agents VIII*. XI, 303 pages. 2004.
- Vol. 3187: G. Lindemann, J. Denzinger, I.J. Timm, R. Unland (Eds.), *Multiagent System Technologies*. XIII, 341 pages. 2004.
- Vol. 3176: O. Bousquet, U. von Luxburg, G. Rätsch (Eds.), *Advanced Lectures on Machine Learning*. IX, 241 pages. 2004.
- Vol. 3171: A.L.C. Bazzan, S. Labidi (Eds.), *Advances in Artificial Intelligence – SBIA 2004*. XVII, 548 pages. 2004.
- Vol. 3159: U. Visser, *Intelligent Information Integration for the Semantic Web*. XIV, 150 pages. 2004.
- Vol. 3157: C. Zhang, H. W. Guesgen, W.K. Yeap (Eds.), *PRICAI 2004: Trends in Artificial Intelligence*. XX, 1023 pages. 2004.
- Vol. 3155: P. Funk, P.A. González Calero (Eds.), *Advances in Case-Based Reasoning*. XIII, 822 pages. 2004.
- Vol. 3139: F. Iida, R. Pfeifer, L. Steels, Y. Kuniyoshi (Eds.), *Embodied Artificial Intelligence*. IX, 331 pages. 2004.
- Vol. 3131: V. Torra, Y. Narukawa (Eds.), *Modeling Decisions for Artificial Intelligence*. XI, 327 pages. 2004.
- Vol. 3127: K.E. Wolff, H.D. Pfeiffer, H.S. Delugach (Eds.), *Conceptual Structures at Work*. XI, 403 pages. 2004.
- Vol. 3123: A. Belz, R. Evans, P. Piwek (Eds.), *Natural Language Generation*. X, 219 pages. 2004.
- Vol. 3120: J. Shawe-Taylor, Y. Singer (Eds.), *Learning Theory*. X, 648 pages. 2004.
- Vol. 3097: D. Basin, M. Rusinowitch (Eds.), *Automated Reasoning*. XII, 493 pages. 2004.
- Vol. 3071: A. Omicini, P. Petta, J. Pitt (Eds.), *Engineering Societies in the Agents World*. XIII, 409 pages. 2004.
- Vol. 3070: L. Rutkowski, J. Siekmann, R. Tadeusiewicz, L.A. Zadeh (Eds.), *Artificial Intelligence and Soft Computing – ICAISC 2004*. XXV, 1208 pages. 2004.
- Vol. 3068: E. André, L. Dybkjær, W. Minker, P. Heisterkamp (Eds.), *Affective Dialogue Systems*. XII, 324 pages. 2004.
- Vol. 3067: M. Dastani, J. Dix, A. El Fallah-Seghrouchni (Eds.), *Programming Multi-Agent Systems*. X, 221 pages. 2004.
- Vol. 3066: S. Tsumoto, R. Słowiński, J. Komorowski, J.W. Grzymała-Busse (Eds.), *Rough Sets and Current Trends in Computing*. XX, 853 pages. 2004.
- Vol. 3065: A. Lomuscio, D. Nute (Eds.), *Deontic Logic in Computer Science*. X, 275 pages. 2004.
- Vol. 3060: A.Y. Tawfik, S.D. Goodwin (Eds.), *Advances in Artificial Intelligence*. XIII, 582 pages. 2004.
- Vol. 3056: H. Dai, R. Srikant, C. Zhang (Eds.), *Advances in Knowledge Discovery and Data Mining*. XIX, 713 pages. 2004.
- Vol. 3055: H. Christiansen, M.-S. Hacid, T. Andreassen, H.L. Larsen (Eds.), *Flexible Query Answering Systems*. X, 500 pages. 2004.
- Vol. 3048: P. Faratin, D.C. Parkes, J.A. Rodríguez-Aguilar, W.E. Walsh (Eds.), *Agent-Mediated Electronic Commerce V*. XI, 155 pages. 2004.
- Vol. 3040: R. Conejo, M. Urretavizcaya, J.-L. Pérez-de-la-Cruz (Eds.), *Current Topics in Artificial Intelligence*. XIV, 689 pages. 2004.
- Vol. 3035: M.A. Wimmer (Ed.), *Knowledge Management in Electronic Government*. XII, 326 pages. 2004.
- Vol. 3034: J. Favela, E. Menasalvas, E. Chávez (Eds.), *Advances in Web Intelligence*. XIII, 227 pages. 2004.
- Vol. 3030: P. Giorgini, B. Henderson-Sellers, M. Winikoff (Eds.), *Agent-Oriented Information Systems*. XIV, 207 pages. 2004.
- Vol. 3029: B. Orchard, C. Yang, M. Ali (Eds.), *Innovations in Applied Artificial Intelligence*. XXI, 1272 pages. 2004.
- Vol. 3025: G.A. Vouros, T. Panayiotopoulos (Eds.), *Methods and Applications of Artificial Intelligence*. XV, 546 pages. 2004.
- Vol. 3020: D. Polani, B. Browning, A. Bonarini, K. Yoshida (Eds.), *RoboCup 2003: Robot Soccer World Cup VII*. XVI, 767 pages. 2004.
- Vol. 3012: K. Kurumatani, S.-H. Chen, A. Ohuchi (Eds.), *Multi-Agents for Mass User Support*. X, 217 pages. 2004.
- Vol. 3010: K.R. Apt, F. Fages, F. Rossi, P. Szeredi, J. Vánca (Eds.), *Recent Advances in Constraints*. VIII, 285 pages. 2004.
- Vol. 2990: J. Leite, A. Omicini, L. Sterling, P. Torroni (Eds.), *Declarative Agent Languages and Technologies*. XII, 281 pages. 2004.
- Vol. 2980: A. Blackwell, K. Marriott, A. Shimojima (Eds.), *Diagrammatic Representation and Inference*. XV, 448 pages. 2004.

¥755.20元

Table of Contents

Invited Papers

Methodologies for Automated Telephone Answering <i>Alan W. Biermann, R. Bryce Inouye, Ashley McKenzie</i>	1
Anomaly Detection in Computer Security and an Application to File System Accesses <i>Salvatore J. Stolfo, Shlomo Hershkop, Linh H. Bui, Ryan Ferster, Ke Wang</i>	14

Regular Papers

1A Knowledge Discovery and Data Mining (1)

A Machine Text-Inspired Machine Learning Approach for Identification of Transmembrane Helix Boundaries <i>Betty Yee Man Cheng, Jaime G. Carbonell, Judith Klein-Seetharaman</i>	29
Visualization of Similarities and Dissimilarities in Rules Using Multidimensional Scaling <i>Shusaku Tsumoto, Shoji Hirano</i>	38
Learning Profiles Based on Hierarchical Hidden Markov Model <i>Ugo Galassi, Attilio Giordana, Lorenza Saitta, Maco Botta</i>	47
Statistical Independence from the Viewpoint of Linear Algebra <i>Shusaku Tsumoto</i>	56

2A Intelligent Information Systems (1)

Bitmap Index-Based Decision Trees <i>Cécile Favre, Fadila Bentayeb</i>	65
On Automatic Modeling and Use of Domain-Specific Ontologies <i>Troels Andreassen, Henrik Bulskov, Rasmus Knappe</i>	74
Using Self-Adaptable Probes for Dynamic Parameter Control of Parallel Evolutionary Algorithms <i>Xavier Bonnaire, María-Cristina Riff</i>	83

Robust Inference of Bayesian Networks Using Speciated Evolution and Ensemble

Kyung-Joong Kim, Ji-Oh Yoo, Sung-Bae Cho 92

1B Knowledge Discovery and Data Mining (2)

Mining the Semantic Web: A Logic-Based Methodology

Francesca A. Lisi, Floriana Esposito 102

Analysis of Textual Data with Multiple Classes

Shigeaki Sakurai, Chong Goh, Ryohei Orihara 112

SARM – Succinct Association Rule Mining: An Approach to Enhance Association Mining

Jitender Deogun, Liying Jiang 121

Learning the Daily Model of Network Traffic

Costantina Caruso, Donato Malerba, Davide Papagni 131

2B Intelligent Information Systems (2)

ARGUS: Rete + DBMS = Efficient Persistent Profile Matching on Large-Volume Data Streams

Chun Jin, Jaime Carbonell, Phil Hayes 142

Failing Queries in Distributed Autonomous Information System

Zbigniew W. Raś, Agnieszka Dardzińska 152

Evaluation of Two Systems on Multi-class Multi-label Document Classification

Xiao Luo, A. Nur Zincir-Heywood 161

Uncertain Knowledge Gathering: An Evolutionary Approach

Dennis Hooijmaijers, Damien Bright 170

1C Information and Knowledge Integration (1)

Duality in Knowledge Compilation Techniques

Neil V. Murray, Erik Rosenthal 182

Data Protection in Distributed Database Systems

Chun Ruan, Vijay Varadharajan 191

2C Soft Computing (1)**A Softened Formulation of Inductive Learning and Its Use for Coronary Disease Data***Janusz Kacprzyk, Grażyna Szkatula* 200**Subsystem Based Generalizations of Rough Set Approximations***Yiyu Yao, Yaohua Chen* 210**1D Clustering (1)****Model-Based Cluster Analysis for Web Users Sessions***George Pallis, Lefteris Angelis, Athena Vakali* 219**On Autonomous k -Means Clustering***Tapio Elomaa, Heidi Koivistoinen* 228**CSI: Clustered Segment Indexing for Efficient Approximate Searching on the Secondary Structure of Protein Sequences***Minkoo Seo, Sanghyun Park, Jung-Im Won* 237**Using Supervised Clustering to Enhance Classifiers***Christoph F. Eick, Nidal Zeidat* 248**2D Web Data****Modelling Good Entry Pages on the Web***Theodora Tsikrika, Mounia Lalmas* 257**A Query Expression and Processing Technique for an XML Search Engine***Wol Young Lee, Hwan Seung Yong* 266**Adapting the Object Role Modelling Method for Ontology Modelling***Peter Spyns* 276**Identifying Content Blocks from Web Documents***Sandip Debnath, Prasenjit Mitra, C. Lee Giles* 285**1E Information and Knowledge Integration (2)****Building the Data Warehouse of Frequent Itemsets in the DWFIST Approach***Rodrigo Salvador Monteiro, Geraldo Zimbrão, Holger Schwarz,
Bernhard Mitschang, Jano Moreira de Souza* 294

Normal Forms for Knowledge Compilation <i>Reiner Hähnle, Neil V. Murray, Erik Rosenthal</i>	304
2E Soft Computing (2)	
On the Approximate Division of Fuzzy Relations <i>Patrick Bosc, Olivier Pivert, Daniel Rocacher</i>	314
Efficient Learning of Pseudo-Boolean Functions from Limited Training Data <i>Guoli Ding, Jianhua Chen, Robert Lax, Peter Chen</i>	323
1F Knowledge Discovery and Data Mining (3)	
Discovering Partial Periodic Sequential Association Rules with Time Lag in Multiple Sequences for Prediction <i>Dan Li, Jitender S. Deogun</i>	332
Mining and Filtering Multi-level Spatial Association Rules with ARES <i>Annalisa Appice, Margherita Berardi, Michelangelo Ceci, Donato Malerba</i>	342
Association Reducts: A Framework for Mining Multi-attribute Dependencies <i>Dominik Ślęzak</i>	354
Frequent Pattern Mining with Preferences – Utility Functions Approach <i>Elena Braynova, Hemant Pendharkar</i>	364
2F Intelligent Information Systems (3)	
Semantic-Based Access to Digital Document Databases <i>F. Esposito, S. Ferilli, T.M.A. Basile, N. Di Mauro</i>	373
Statistical Database Modeling for Privacy Preserving Database Generation <i>Xintao Wu, Yongge Wang, Yuliang Zheng</i>	382
Scalable Inductive Learning on Partitioned Data <i>Qijun Chen, Xindong Wu, Xingquan Zhu</i>	391
Agent-Based Home Simulation and Control <i>B. De Carolis, G. Cozzolongo, S. Pizzutilo, V.L. Plantamura</i>	404

1G Logic for Artificial Intelligence

Aggregates and Preferences in Logic Programming <i>S. Greco, I. Trubitsyna, E. Zuppano</i>	413
The Chisholm Paradox and the Situation Calculus <i>Robert Demolombe, Pilar Pozos-Parra</i>	425
A Logic Approach for LTL System Modification <i>Yulin Ding, Yan Zhang</i>	435
Anticipatory Agents Based on Anticipatory Reasoning <i>Feng Shang, Jingde Cheng</i>	445

2G Applications (1)

Extracting Emotions from Music Data <i>Alicja Wieczorkowska, Piotr Synak, Rory Lewis, Zbigniew W. Raś</i>	456
An Intelligent System for Assisting Elderly People <i>Alberto Tablado, Arantza Illarramendi, Miren I. Bagüés, Jesús Bermúdez, Alfredo Goñi</i>	466
Multi-strategy Instance Selection in Mining Chronic Hepatitis Data <i>Masatoshi Jumi, Einoshin Suzuki, Muneaki Ohshima, Ning Zhong, Hideto Yokoi, Katsuhiko Takabayashi</i>	475
A Probabilistic Approach to Finding Geometric Objects in Spatial Datasets of the Milky Way <i>Jon Purnell, Malik Magdon-Ismail, Heidi Jo Newberg</i>	485

1H Applications (2)

Towards Ad-Hoc Rule Semantics for Gene Expression Data <i>Marie Agier, Jean-Marc Petit, Einoshin Suzuki</i>	494
Flexible Pattern Discovery with (Extended) Disjunctive Logic Programming <i>Luigi Palopoli, Simona Rombo, Giorgio Terracina</i>	504
Interactive SOM-Based Gene Grouping: An Approach to Gene Expression Data Analysis <i>Alicja Gruzdź, Aleksandra Ihnatowicz, Dominik Ślęzak</i>	514

Some Theoretical Properties of Mutual Information for Student Assessments
in Intelligent Tutoring Systems

Chao-Lin Liu 524

2H Intelligent Information Retrieval (1)

Cooperative Query Answering for RDF

Adrian Tanasescu 535

Intelligent Information Retrieval for Web-Based Design Data Repository

Huijae Lee, Sang Bong Yoo 544

Incremental Collaborative Filtering for Highly-Scalable Recommendation
Algorithms

*Manos Papagelis, Ioannis Rousidis, Dimitris Plexousakis,
Elias Theoharopoulos* 553

1I Clustering (2)

A Distance-Based Algorithm for Clustering Database User Sessions

Qingsong Yao, Aijun An, Xiangji Huang 562

User-Interest-Based Document Filtering via Semi-supervised Clustering

Na Tang, V. Rao Vemuri 573

A Filter Feature Selection Method for Clustering

Pierre-Emmanuel Jouve, Nicolas Nicoloyannis 583

Automatic Determination of the Number of Fuzzy Clusters Using Simulated
Annealing with Variable Representation

Sanghamitra Bandyopadhyay 594

2I Knowledge Discovery and Data Mining (4)

Experimental Analysis of the Q-Matrix Method in Knowledge Discovery

Tiffany Barnes, Donald Bitzer, Mladen Vouk 603

Clustering Time-Series Medical Databases Based on the Improved Multiscale
Matching

Shoji Hirano, Shusaku Tsumoto 612

Efficient Causal Interaction Learning with Applications in Microarray

Yong Ye, Xintao Wu 622

A Dynamic Adaptive Sampling Algorithm (DASA) for Real World Applications: Finger Print Recognition and Face Recognition
Ashwin Satyanarayana, Ian Davidson 631

1J Intelligent Information Retrieval (2)

Catching the Picospams
Matthew Chang, Chung Keung Poon 641

Personalized Peer Filtering for a Dynamic Information Push
Melanie Gnasa, Sascha Alda, Nadir Gul, Armin B. Cremers 650

Getting Computers to See Information Graphics So Users Do Not Have to
Daniel Chester, Stephanie Elzer 660

2J Knowledge Representation

A Data Model Based on Paraconsistent Intuitionistic Fuzzy Relations
Haibin Wang, Rajshekhar Sunderraman 669

Logical Data Independence Reconsidered
James J. Lu 678

Estimation of the Density of Datasets with Decision Diagrams
Ansaf Salleb, Christel Vrain 688

Author Index 699

Methodologies for Automated Telephone Answering

Alan W. Biermann, R. Bryce Inouye, and Ashley McKenzie

Department of Computer Science, Duke University,
Durham NC 27708, USA
{awb, rbi, armckenz}@cs.duke.edu

Abstract. We survey some of the approaches to dialogue representation and processing for modern telephone answering systems. We include discussions of their strong and weak points and some of the performance levels obtained by them.

1 Technologies for the Third Millennium

The advent of ubiquitous computing will bring computers and the Internet into our lives on an almost continuous basis. We will have machines and communication access in our offices, in our vehicles, in our brief cases, on our walls, and elsewhere. We will be able to interact with the world, access the huge Internet resources, do local computations, and much more. But how will we utilize these machines easily enough to really profit from their capabilities? One suggestion is that we should speak to them as we do to humans and receive responses back using natural language ([1],[2],[3]). However, there have been major obstacles to the development of spoken language dialogue machines over the decades and some of these remain a problem today. This paper will address some of the major ideas in the development of dialogue systems and the levels of success that are possible. The focus of much research activity in recent years has been in automatic telephone answering ([4],[5],[6],[7],[3]) but we will also mention projects that study other types of human-machine interaction ([8],[2]).

Our primary concerns in this paper are the overall conception of what a dialogue is and how a computer may represent it and manage it. We present some of the most common and most successful paradigms for dialogue processing. Given these formats, we then discuss related issues including the control of initiative, user modeling, error correction, and dialogue optimization.

2 Paradigms for Dialogue Management

The earliest and probably the most used model for dialogues has been the finite-state machine ([9]). The model enables the designer to systematically specify the sequential steps of a dialogue anticipating what the user may say and designing the responses that are needed. This is a conservative technology that enables

precise control of every machine action and prevents surprising behaviors that could result from unanticipated situations.

For the purposes of this paper, we will do examples from the domain of credit card company telephone answering where we have some experience ([10]). We will assume a greatly simplified version of the problem to keep the presentation short but will include enough complexity to make the points. Our model will require the system to obtain the name of the caller and an identifying account code. It also will require the system to obtain the purpose of the call. In our simplified example, we will assume there are only three possible goals: to determine the current balance on the credit card account, to report that the card has been lost and that another needs to be issued, or to request that a new address be registered for this caller.

Figure 1 shows a finite-state solution for this simplified credit card problem. The interaction begins with the machine greeting the user. We omit details of what this might be but a company would have a policy as to what to say. Perhaps it would say something like this: "Hello, this is the XYZ corporation automated credit card information system. What can I do for you today?" The figure shows the kinds of caller comments expected and the machine recognition system and parser have the tasks of recognizing and processing the incoming speech. The figure shows the next steps for the machine's response. Its output system needs an ability to convert a target message into properly enunciated speech. The continuation of the diagram shows the dialogue capabilities of the system. We omit a number of transitions in the diagram that might be included for variations in the interactions, for unusual requests, and for error correction.

An example dialogue that traverses this diagram is as follows:

System: Hello. XYZ credit card management. How can I help you?

Caller: Hello. I need to report a lost credit card.

System: Please give me your name.

Caller: William Smith.

System: Thank you. And your account code?

Caller: X13794.

System: Okay, I will cancel your existing card and issue a new one within 24 hours.

A common problem with the finite state model is that designers become exhausted with the detailed specification of every possible interaction that could occur. The number of states can explode as one accounts for misrecognitions and repeated requests, variations in the strategy of the interactions, different levels of initiative, extra politeness in special cases, and more. This leads to another model for dialogue systems, the form-filling model that assumes there is a set of slots to be filled in the interaction and that processing should evolve around the concept of filling the slots. Here the software can be designed around the slots and the computations needed to fill them. The ordering of the dialogue interactions is not implicitly set ahead of time and the designer need not think about order as much. If any interaction happens to include information to fill a slot, the information is immediately parsed and entered. If an important slot