

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Friedemann Mattern

ETH Zurich, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

University of Dortmund, Germany

Madhu Sudan

Massachusetts Institute of Technology, MA, USA

Demetri Terzopoulos

New York University, NY, USA

Doug Tygar

University of California, Berkeley, CA, USA

Moshe Y. Vardi

Rice University, Houston, TX, USA

Gerhard Weikum

Max-Planck Institute of Computer Science, Saarbruecken, Germany

James F. Peters Andrzej Skowron
Jerzy W. Grzymała-Busse Bożena Kostek
Roman W. Świniarski Marcin S. Szczuka (Eds.)

Transactions on Rough Sets I

Editors-in-Chief

James F. Peters

University of Manitoba, Department of Electrical and Computer Engineering
Manitoba, Winnipeg, Manitoba R3T 5V6 Canada
E-mail: jfpeters@ee.umanitoba.ca

Andrzej Skowron

University of Warsaw, Institute of Mathematics
Banacha 2, 02-097 Warsaw, Poland
E-mail: skowron@mimuw.edu.pl

Volume Editors

Jerzy W. Grzymała-Busse

University of Kansas, Department of Electrical Engineering and Computer Science
3014 Eaton Hall 1520 W. 15th St., #2001 Lawrence, KS 66045-7621, USA
E-mail: Jerzy@ku.edu

Bożena Kostek

Gdansk University of Technology
Faculty of Electronics, Telecommunications and Informatics
Multimedia Systems Department, Narutowicza 11/12, 80-952 Gdansk, Poland
E-mail: bozenka@sound.eti.pg.gda.pl

Roman W. Świniarski

San Diego State University, Department of Computer Science
5500 Campanile Drive, San Diego, CA 92182-7720, USA
E-mail: rswiniar@sciences.sdsu.edu

Marcin S. Szczuka

Warsaw University, Institute of Mathematics
Banacha 2, 02-097 Warsaw, Poland
E-mail: szczuka@mimuw.edu.pl

Library of Congress Control Number: 2004108444

CR Subject Classification (1998): F.4.1, F.1, I.2, H.2.8, I.5.1, I.4

ISSN 0302-9743

ISBN 3-540-22374-6 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 to prosecution under the German Copyright Law.

Springer-Verlag is a part of Springer Science+Business Media

springeronline.com

© Springer-Verlag Berlin Heidelberg 2004
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Olgun Computergrafik
Printed on acid-free paper SPIN: 11011873 06/3142 5 4 3 2 1 0

Preface

We would like to present, with great pleasure, the first volume of a new journal, Transactions on Rough Sets. This journal, part of the new journal subline in the Springer-Verlag series Lecture Notes in Computer Science, is devoted to the entire spectrum of rough set related issues, starting from logical and mathematical foundations of rough sets, through all aspects of rough set theory and its applications, data mining, knowledge discovery and intelligent information processing, to relations between rough sets and other approaches to uncertainty, vagueness, and incompleteness, such as fuzzy sets, theory of evidence, etc.

The first, pioneering papers on rough sets, written by the originator of the idea, Professor Zdzisław Pawlak, were published in the early 1980s. We are proud to dedicate this volume to our mentor, Professor Zdzisław Pawlak, who kindly enriched this volume with his contribution on philosophical, logical, and mathematical foundations of rough set theory. In his paper Professor Pawlak shows all over again the underlying ideas of rough set theory as well as its relations with Bayes' theorem, conflict analysis, flow graphs, decision networks, and decision rules.

After an overview and introductory article written by Professor Pawlak, the ten following papers represent and focus on rough set theory-related areas. Some papers provide an extension of rough set theory towards analysis of very large data sets, real data tables, data sets with missing values and rough non-deterministic information. Other theory-based papers deal with variable precision fuzzy-rough sets, consistency measure conflict profiles, and layered learning for concept synthesis. In addition, a paper on generalization of rough sets and rule extraction provides two different interpretations of rough sets. The last paper of this group addresses a partition model of granular computing.

Other topics with a more application-orientated view are covered by the following eight articles of this first volume of Transactions on Rough Sets. They can be categorized into the following groups:

- music processing,
- rough set theory applied to software design models and inductive learning programming,
- environmental engineering models,
- medical data processing,
- pattern recognition and classification.

These papers exemplify analysis and exploration of complex data sets from various domains. They provide useful insight into analyzed problems, showing for example how to compute decision rules from incomplete data. We believe that readers of this volume will better appreciate rough set theory-related trends after reading the case studies.

Many scientists and institutions have contributed to the creation and the success of the rough set community. We are very thankful to everybody within the International Rough Set Society who supported the idea of creating a new LNCS journal subline – the Transactions on Rough Sets. It would not have been possible without Professors Peters' and Skowron's invaluable initiative, thus we are especially grateful to them. We believe that this very first issue will be followed by many others, reporting new developments in the rough set domain. This issue would not have been possible without the great efforts of many anonymously acting reviewers. Here, we would like to express our sincere thanks to all of them.

Finally, we would like to express our gratitude to the LNCS editorial staff of Springer-Verlag, in particular Alfred Hofmann, Ursula Barth and Christine Günther, who supported us in a very professional way.

Throughout preparation of this volume the Editors have been supported by various research programs and funds; Jerzy Grzymała-Busse has been supported by NSF award 9972843, Bożena Kostek has been supported by the grant 4T11D01422 from the Polish Ministry for Scientific Research and Information Technology, Roman Świniarski has received support from the “*Adaptive Data Mining and Knowledge Discovery Methods for Distributed Data*” grant, awarded by Lockheed-Martin, and Marcin Szczuka and Roman Świniarski have been supported by the grant 3T11C00226 from the Polish Ministry for Scientific Research and Information Technology.

April 2004

Jerzy W. Grzymała-Busse
Bożena Kostek
Roman Świniarski
Marcin Szczuka

LNCS Transactions on Rough Sets

This journal subline has as its principal aim the fostering of professional exchanges between scientists and practitioners who are interested in the foundations and applications of rough sets. Topics include foundations and applications of rough sets as well as foundations and applications of hybrid methods combining rough sets with other approaches important for the development of intelligent systems.

The journal includes high-quality research articles accepted for publication on the basis of thorough peer reviews. Dissertations and monographs up to 250 pages that include new research results can also be considered as regular papers. Extended and revised versions of selected papers from conferences can also be included in regular or special issues of the journal.

Honorary Editor: Zdzisław Pawlak
Editors-in-Chief: James F. Peters, Andrzej Skowron

Editorial Board

| | |
|---------------------|--------------------|
| M. Beynon | M. do C. Nicoletti |
| G. Cattaneo | H.S. Nguyen |
| A. Czyżewski | S.K. Pal |
| J.S. Deogun | L. Polkowski |
| D. Dubois | H. Prade |
| I. Duentzsch | S. Ramanna |
| S. Greco | R. Słowiński |
| J.W. Grzymała-Busse | J. Stepaniuk |
| M. Inuiguchi | R. Świniarski |
| J. Jrvinen | Z. Suraj |
| D. Kim | M. Szczuka |
| J. Komorowski | S. Tsumoto |
| C.J. Liao | G. Wang |
| T.Y. Lin | Y. Yao |
| E. Menasalvas | N. Zhong |
| M. Moshkov | W. Ziarko |
| T. Murai | |

Table of Contents

Rough Sets – Introduction

| | |
|-------------------------------------|---|
| Some Issues on Rough Sets | 1 |
| <i>Zdzisław Pawlak</i> | |

Rough Sets – Theory

| | |
|--|----|
| Learning Rules from Very Large Databases Using Rough Multisets | 59 |
| <i>Chien-Chung Chan</i> | |

| | |
|---|----|
| Data with Missing Attribute Values: Generalization of Indiscernibility Relation and Rule Induction | 78 |
| <i>Jerzy W. Grzymala-Busse</i> | |

| | |
|---|----|
| Generalizations of Rough Sets and Rule Extraction | 96 |
| <i>Masahiro Inuiguchi</i> | |

| | |
|---|-----|
| Towards Scalable Algorithms for Discovering Rough Set Reducts | 120 |
| <i>Marzena Kryszkiewicz and Katarzyna Cichoń</i> | |

| | |
|--|-----|
| Variable Precision Fuzzy Rough Sets | 144 |
| <i>Alicja Mieszkowicz-Rolka and Leszek Rolka</i> | |

| | |
|---|-----|
| Greedy Algorithm of Decision Tree Construction for Real Data Tables | 161 |
| <i>Mikhail Ju. Moshkov</i> | |

| | |
|--|-----|
| Consistency Measures for Conflict Profiles | 169 |
| <i>Ngoc Thanh Nguyen and Michal Malowiecki</i> | |

| | |
|---|-----|
| Layered Learning for Concept Synthesis | 187 |
| <i>Sinh Hoa Nguyen, Jan Bazan, Andrzej Skowron, and Hung Son Nguyen</i> | |

| | |
|--|-----|
| Basic Algorithms and Tools for Rough Non-deterministic Information Analysis | 209 |
| <i>Hiroshi Sakai and Akimichi Okuma</i> | |

| | |
|---|-----|
| A Partition Model of Granular Computing | 232 |
| <i>Yiyu Yao</i> | |

Rough Sets – Applications

| | |
|---|-----|
| Musical Phrase Representation and Recognition by Means of Neural Networks and Rough Sets | 254 |
| <i>Andrzej Czyzewski, Marek Szczerba, and Bożena Kostek</i> | |

Processing of Musical Metadata Employing Pawlak’s Flow Graphs 279
Bożena Kostek and Andrzej Czyżewski

Data Decomposition and Decision Rule Joining for Classification of Data
with Missing Values 299
Rafał Latkowski and Michał Mikołajczyk

Rough Sets and Relational Learning 321
R.S. Milton, V. Uma Maheswari, and Arul Siromoney

Approximation Space for Software Models 338
James F. Peters and Sheela Ramanna

Application of Rough Sets to Environmental Engineering Models 356
Robert H. Warren, Julia A. Johnson, and Gordon H. Huang

Rough Set Theory and Decision Rules in Data Analysis
of Breast Cancer Patients 375
*Jerzy Zatuski, Renata Szoszkiewicz, Jerzy Krysiński,
and Jerzy Stefanowski*

Independent Component Analysis, Principal Component Analysis
and Rough Sets in Face Recognition 392
Roman W. Świniarski and Andrzej Skowron

Author Index 405