# Lecture Notes in Artificial Intelligence 11103

## Subseries of Lecture Notes in Computer Science

#### **LNAI Series Editors**

Randy Goebel
University of Alberta, Edmonton, Canada
Yuzuru Tanaka
Hokkaido University, Sapporo, Japan
Wolfgang Wahlster
DFKI and Saarland University, Saarbrücken, Germany

#### LNAI Founding Series Editor

Joerg Siekmann

DFKI and Saarland University, Saarbrücken, Germany

More information about this series at http://www.springer.com/series/1244

Hung Son Nguyen · Quang-Thuy Ha Tianrui Li · Małgorzata Przybyła-Kasperek (Eds.)

# Rough Sets

International Joint Conference, IJCRS 2018 Quy Nhon, Vietnam, August 20–24, 2018 Proceedings



Editors
Hung Son Nguyen
University of Warsaw
Warsaw
Poland

Quang-Thuy Ha 
Faculty of Information Technology
Vietnam National University
Hanoi
Vietnam

Tianrui Li School of Information Science Southwest Jiaotong University Chengdu
China

Małgorzata Przybyła-Kasperek Dinstitute of Computer Science University of Silesia Sosnowiec Poland

ISSN 0302-9743 ISSN 1611-3349 (electronic) Lecture Notes in Artificial Intelligence ISBN 978-3-319-99367-6 ISBN 978-3-319-99368-3 (eBook) https://doi.org/10.1007/978-3-319-99368-3

Library of Congress Control Number: 2018951242

LNCS Sublibrary: SL7 - Artificial Intelligence

#### © Springer Nature Switzerland AG 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

#### **Preface**

The proceedings of the 2018 International Joint Conference on Rough Sets (IJCRS 2018) contain the results of the meeting of the International Rough Set Society held at the International Centre for Interdisciplinary Science and Education (ICISE) and the University of Quy Nhon in Quy Nhon, Vietnam, during August 2018.

Conferences in the IJCRS series are held annually and comprise four main tracks relating the topic rough sets to other topical paradigms: rough sets and data analysis covered by the RSCTC conference series from 1998, rough sets and granular computing covered by the RSFDGrC conference series since 1999, rough sets and knowledge technology covered by the RSKT conference series since 2006, and rough sets and intelligent systems covered by the RSEISP conference series since 2007. Owing to the gradual emergence of hybrid paradigms involving rough sets, it was deemed necessary to organize Joint Rough Set Symposiums, first in Toronto, Canada, in 2007, followed by symposiums in Chengdu, China in 2012, Halifax, Canada, 2013, Granada and Madrid, Spain, 2014, Tianjin, China, 2015, where the acronym IJCRS was proposed, continuing with the IJCRS 2016 conference in Santiago de Chile and IJCRS 2017 in Olsztyn, Poland.

The IJCRS conferences aim at bringing together experts from universities and research centers as well as from industry representing fields of research in which theoretical and applicational aspects of rough set theory already find or may potentially find usage. They also become a place for researchers who want to present their ideas to the rough set community, or for those who would like to learn about rough sets and find out if they can be useful for their problems.

This year's conference, IJCRS 2018, celebrated the 20th anniversary of the first international conference on rough sets called RSCTC, which was organized by Lech Polkowski and Andrzej Skowron during June 22–26, 1998, in Warsaw, Poland. On this occasion, we listened to a retrospective talk delivered by Andrzej Skowron, who summarized the successes of this field and showed directions for further research and development.

IJCRS 2018 attracted 61 submissions (not including invited contributions), which underwent a rigorous reviewing process. Each accepted full-length paper was evaluated by three to five experts on average. The present volume contains 45 full-length regular and workshop submissions, which were accepted by the Program Committee, as well as six invited articles.

The conference program included five keynotes and plenary talks, a fellow talk, eight parallel sessions, a tutorial, the 6th International Workshop on Three-way Decisions, Uncertainty, and Granular Computing, and a panel discussion on rough sets and data science.

The chairs of the Organizing Committee also prepared the best paper award and the best student paper award. From all research papers submitted, the Program Committee

nominated five papers as finalists for the award and, based on the final presentations during the conference, selected the winners.

We would like to express our gratitude to all the authors for submitting papers to IJCRS 2018, as well as to the members of the Program Committee for organizing this year's attractive program.

We also gratefully thank our sponsors: Vietnam National University in Ho Chi Minh City, for providing the technical support and human resources for the conference; the University of Quy Nhon, for sponsoring the reception and the conference facilities during the first day and the last day; Ton Duc Thang University, for sponsoring the pre-conference workshops on rough sets and data mining.

The conference would not have been successful without support received from distinguished individuals and organizations. We express our gratitude to the IJCRS 2018 honorary chairs, Andrzej Skowron, Huynh Thanh Dat, and Do Ngoc My, for their great leadership. We appreciate the help of Dinh Thuc Nguyen, Nguyen Tien Trung, Quang Vinh Lam, Quang Thai Thuan, Thanh Tran Thien, Luong Thi Hong Cam, Giang Thuy Minh, Phung Thai Thien Trang, Dao Thi Hong Le, Hung Nguyen-Manh, and all other representatives of Vietnam National University in Ho Chi Minh City and Quy Nhon University, who were involved in the conference organization. We would also like to thank Marcin Szelag, Sinh Hoa Nguyen, and Dang Phuoc Huy, who supported the conference as tutorial, workshop, and special session chairs. We acknowledge the significant help from Khuong Nguyen-An, Tran Thanh Hai, Ly Tran Thai Hoc, and Marcin Szczuka provided at various stages of the conference publicity, website, and material preparation.

We are grateful to Tu Bao Ho, Hamido Fujita, Hong Yu, Andrzej Skowron, Piero Pagliani, and Mohua Banerjee for delivering excellent keynote and plenary talks and fellow talks. We thank Dominik Ślęzak and Arkadiusz Wojna for the tutorial. We are thankful to Hong Ye, Mohua Banerjee, Mihir Chakraborty, Bay Vo, and Le Thi Thuy Loan for the organization of workshops and special sessions.

Special thanks go to Alfred Hofmann of Springer, for accepting to publish the proceedings of IJCRS 2018 in the LNCS/LNAI series, and to Anna Kramer for her help with the proceedings. We are grateful to Springer for the grant of 1,000 Euro for the best paper award winners. We would also like to acknowledge the use of EasyChair, a great conference management system.

We hope that the reader will find all the papers in the proceedings interesting and stimulating.

August 2018

Hung Son Nguyen Quang-Thuy Ha Tianrui Li Małgorzata Przybyła-Kasperek

## **Organization**

#### **Honorary Chairs**

Andrzej Skowron University of Warsaw, Poland

Thanh Dat Huynh VNU-HCMC, Vietnam

Ngoc My Do Ouy Nhon University, Vietnam

#### **General Chairs**

Davide Ciucci University of Milano-Bicocca, Italy

Dan Thu Tran VNU-HCMC, Vietnam

#### **Organizing Committee Chairs**

Dinh Thuc Nguyen VNU-HCMC, Vietnam

Tien Trung Nguyen Quy Nhon University, Vietnam

Quang Vinh Lam VNU-HCMC, Vietnam

Quang Thai Thuan Quy Nhon University, Vietnam Thanh Thien Tran Quy Nhon University, Vietnam

#### **Program Committee**

## **Program Committee Chairs**

Hung Son Nguyen University of Warsaw, Poland

Quang-Thuy Ha College of Technology, VNU-Hanoi, Vietnam Tianrui Li Southwest Jiaotong University, Chengdu, China

Małgorzata University of Silesia, Poland

Przybyła-Kasperek

### Workshop, Special Sessions, and Tutorial Chairs

Marcin Szeląg Poznań University of Technology, Poland Sinh Hoa Nguyen Polish-Japanese Academy of IT, Poland

Phuoc Huy Dang Dalat University, Vietnam

## **Program Committee**

Mani A. Calcutta University, India

Piotr Artiemjew University of Warmia and Mazury, Poland Jaume Baixeries Universitat Politecnica de Catalunya, Spain

#### VIII Organization

Mohua Banerjee Indian Institute of Technology Kanpur, India

Jan Bazan University of Rzeszów, Poland

Rafael Bello Universidad Central de Las Villas, Cuba

Nizar Bouguila Concordia University, Canada

Jerzy Baszczyski Poznań University of Technology, Poland

Mihir Chakraborty Jadavpur University, India

Shampa Chakraverty Netaji Subhas Institute of Technology, India

Chien-Chung Chan University of Akron, USA

Mu-Chen Chen National Chiao Tung University, Taiwan Costin-Gabriel Chiru Technical University of Bucharest, Romania

Victor Codocedo INSA Lyon, France

Chris Cornelis University of Granada, Spain
Zoltan Erno Csajbok University of Debrecen, Hungary
Jianhua Dai Hunan Normal University, China

Rafal Deja WSB, Poland

Dayong Deng Zhejiang Normal University, China

Thierry Denoeux Université de Technologie de Compiegne, France

Fernando Diaz University of Valladolid, Spain

Pawel Drozda University of Warmia and Mazury, Poland

Didier Dubois IRIT/RPDMP, France
Ivo Dntsch Brock University, Canada

Zied Elouedi Institut Superieur de Gestion de Tunis, Tunisia Rafael Falcon Larus Technologies Corporation, Canada Victor Flores Universidad Catolica del Norte, Chile

Wojciech Froelich University of Silesia, Poland University of Insubria, Italy

Piotr Gny Polish-Japanese Academy of IT, Poland

Anna Gomolinska University of Białystok, Poland Salvatore Greco University of Catania, Italy

Rafal Gruszczynski Nicolaus Copernicus University in Toruń, Poland

Jerzy Grzymala-Busse University of Kansas, USA

Bineet Gupta Shri RamSwaroop Memorial University, India

Christopher Henry
Christopher Hinde
Christopher Hinde
Oinghua Hu
University of Winnipeg, Canada
Loughborough University, UK
Tianjin University, China

Van Nam Huynh JAIST, Japan

Dmitry Ignatov National Research University HSE, Russia

Masahiro Inuiguchi Osaka University, Japan
Ryszard Janicki McMaster University, Canada
Richard Jensen Aberystwyth University, UK

Xiuyi Jia Nanjing University of Science and Technology, China

Michal Kepski University of Rzeszów, Poland

Md. Aquil Khan Indian Institute of Technology Indore, India

Yoo-Sung Kim Inha University, South Korea

Marzena Kryszkiewicz Warsaw University of Technology, Poland Yasuo Kudo Muroran Institute of Technology, Japan Yoshifumi Kusunoki Osaka University, Japan

Sergei O. Kuznetsov National Research University HSE, Russia

Xuan Viet LeQuy Nhon University, VietnamHuaxiong LiNanjing University, ChinaJiye LiangShanxi University, China

Churn-Jung Liau Academia Sinica, Taipei, Taiwan
Tsau Young Lin San Jose State University, USA
Pawan Lingras Saint Mary's University, Canada
Caihui Liu Gannan Normal University, China

Guilong Liu Beijing Language and Culture University, China

Pradipta Maji Indian Statistical Institute, India Benedetto Matarazzo University of Catania, Italy Jess Medina University of Cadiz, Spain

Ernestina Menasalvas Universidad Politecnica de Madrid, Spain Claudio Meneses Universidad Catolica del Norte, Chile Marcin Michalak Silesian University of Technology, Poland

Tams Mihlydek University of Debrecen, Hungary
Fan Min Southwest Petroleum University, China

Pabitra Mitra Indian Institute of Technology Kharagpur, India

Sadaaki Miyamoto University of Tsukuba, Japan

Mikhail Moshkov KAUST, Saudi Arabia

Michinori Nakata Josai International University, Japan

Amedeo Napoli Inria, France

Hoang Son Nguyen Hue University, Vietnam

Loan T. T. Nguyen TDTU, Vietnam

Long Giang Nguyen Institute of Information Technology, VAST, Vietnam

M. C. Nicoletti FACCAMP and UFSCar, Brazil
Vilem Novak University of Ostrava, Czech Republic

Agnieszka University of Silesia, Poland

Nowak-Brzezińska

Piero Pagliani Research Group on Knowledge and Information, Italy

Sankar Pal Indian Statistical Institute, India Krzysztof Pancerz University of Rzeszów, Poland

Vladimir Parkhomenko SPbPU, Russia

Andrei Paun University of Bucharest, Romania Witold Pedrycz University of Alberta, Canada

Tatiana Penkova Institute of Computational Modelling SB RAS, Russia

Georg Peters Munich University of Applied Sciences

and Australian Catholic University, Germany

Alberto Pettorossi Università di Roma Tor Vergata, Italy

Jonas Poelmans Clarida Technologies, UK

Lech Polkowski Polish-Japanese Academy of IT, Poland

Henri Prade IRIT - CNRS, France

Mohamed Quafafou Aix-Marseille University, France

Elisabeth Rakus-Andersson Blekinge Institute of Technology, Sweden

Sheela Ramanna University of Winnipeg, Canada

Zbigniew Ras University of North Carolina at Charlotte, USA Grzegorz Rozenberg Leiden University, The Netherlands

Henryk Rybiski Warsaw University of Technology, Poland

Rzeszów University, Poland Woiciech Rzasa

Hiroshi Sakai Kyushu Institute of Technology, Japan Guido Santos Universittsklinikum Erlangen, Germany

Loughborough University, UK Gerald Schaefer

Institute of Computing Technology Chinese Academy Zhongzhi Shi

of Sciences, China

Marek Sikora Silesian University of Technology, Poland

Bruno Simões Vicomtech-IK4, Spain

Roman Słowinski Poznan University of Technology, Poland

University of Leeds, UK

Bialystok University of Technology, Poland

University of Rzeszów, Poland Zbigniew Suraj

Indian Institute of Technology Jodhpur, India Paul Sushmita

Security On Demand, Poland University of Warsaw, Poland Marcin Szczuka University of Warsaw, Poland

AGH University of Science and Technology Ryszard Tadeusiewicz

> in Kraków, Poland VIT University, India

North Carolina A&T State University, USA Federal Research Center for Computer Science

and Control, RAS, Russia

Ho Chi Minh City University of Technology, Vietnam

University of Silesia, Poland Chongqing University of Posts and Telecommunications, China Poznan University of Technology, Poland

Security On Demand, Poland

Maria Curie-Skłodowska University, Poland

Zhejiang Ocean University, China Southwest Jiaotong University, China

University of Regina, Canada University of Regina, Canada Fuzhou University, China Chongqing University of Posts

and Telecommunications, China Université Paris Est. France

Yantai University, China Chongqing University of Posts and Telecommunications, China University of Regina, Canada

Sam Houston State University, USA University of Regina, Canada University of Silesia, Poland

Bing Zhou Wojciech Ziarko Beata Zielosko

John Stell

Jaroslaw Stepaniuk

Piotr Synak Andrzej Szałas

Bala Krushna Tripathy Li-Shiang Tsay Dmitry Vinogradov

Bay Vo

Alicja Wakulicz-Deja Guoyin Wang

Szymon Wilk

Arkadiusz Wojna Marcin Wolski

Wei-Zhi Wu Yan Yang

Jingtao Yao Yiyu Yao Dongyi Ye Hong Yu

Mahdi Zargayouna

Nan Zhang Qinghua Zhang

Yan Zhang

#### **Additional Reviewers**

Azam, Nouman Benítez Caballero, María José

Bui, Huong Chen, Chun-Hao Czołombitko, Michał Jankowski, Dariusz

Le, Tuong Li, Jinhai Mai, Son Nguyen, Dan Nguyen, Duy Ham Nguyen, Hoang Son Nguyen, Van Du Nguyen, Viet Hung Pham, Thi-Ngan Ramírez Poussa, Eloisa

Shah, Ekta Son, Le Hoang Su, Ja-Hwung Vluymans, Sarah

# Introducing Histogram Functions into a Granular Approximate Database Engine (Industry Talk)

Dominik Ślęzak<sup>1</sup> and Arkadiusz Wojna<sup>2</sup>

<sup>1</sup> Institute of Informatics, University of Warsaw, Poland
<sup>2</sup> Security On-Demand, USA/Poland

**Abstract.** We discuss an approximate database engine that we started designing at Infobright, and now we continue its development for Security On-Demand (SOD). At SOD, it is used in everyday data analytics, allowing for fast approximate execution of ad-hoc queries over tens of billions of data rows [1]. In our engine, queries are run against collections of histograms that represent domains of single columns over groupings of consecutively loaded data rows (so-called packrows). Query execution process corresponds to transformation of such granulated summaries of the input data into summaries reflecting query results [2].

We compare our algorithms that generate histogram descriptions of the original data with data quantization methods that are widely used in data mining. We also introduce a new idea of extending SQL with function hist(a) that produces quantized representation of column a by means of merging a's histograms corresponding to particular packrows into a unified a's histogram over the whole data. We refer to our recent works on summary-based data visualization [3] and machine learning [4] in order to illustrate several scenarios of utilizing hist in practice.

**Keywords:** Big data analytics · Data granulation · Data quantization

#### References

- Ślęzak, D., Chądzyńska-Krasowska, A., Holland, J., Synak, P., Glick, R., Perkowski, M.: Scalable cyber-security analytics with a new summary-based approximate query engine. In: Proceedings of BigData, pp. 1840–1849 (2017)
- Ślęzak, D., Glick, R., Betliński, P., Synak, P.: A new approximate query engine based on intelligent capture and fast transformations of granulated data summaries. J. Intell. Inf. Syst. 50(2), 385–414 (2018)
- 3. Chądzyńska-Krasowska, A., Stawicki, S., Ślęzak, D.: A metadata diagnostic framework for a new approximate query engine working with granulated data summaries. In: Polkowski, L., et al. (eds.) IJCRS 2017. LNCS, vol. 10313, pp. 623–643. Springer, Cham
- Ślęzak, D., Borkowski, J., Chądzyńska-Krasowska, A.: Ranking mutual information dependencies in a summary-based approximate analytics framework. In: Proceedings of HPCS (2018)

# **Contents**

Subjective Analysis of Price Herd Using Dominance Rough Set Induction:  Case Study of Solar Companies	1
Three-Way Decisions and Three-Way Clustering	13
Some Foundational Aspects of Rough Sets Rendering  Its Wide Applicability	29
What's in a Relation? Logical Structures of Modes of Granulation	46
Multi-granularity Attribute Reduction	61
Tolerance Methods in Graph Clustering: Application to Community  Detection in Social Networks	73
Similarity Based Rough Sets with Annotation	88
Multidimensional Data Analysis for Evaluating the Natural and Anthropogenic Safety (in the Case of Krasnoyarsk Territory)	101
A Metaphor for Rough Set Theory: Modular Arithmetic	110
A Method for Boundary Processing in Three-Way Decisions Based on Hierarchical Feature Representation  Jie Chen, Yang Xu, Shu Zhao, Yuanting Yan, Yanping Zhang, Weiwei Li, Qianqian Wang, and Xiangyang Wang	123
Covering-Based Optimistic-Pessimistic Multigranulation  Decision-Theoretic Rough Sets	137

by STRIM: In a Simulation Model for Data Generation and Verification of Induced Rules	148
Rseslib 3: Open Source Library of Rough Set and Machine  Learning Methods	162
Composite Sequential Three-Way Decisions	177
NDER Attribute Reduction via an Ensemble Approach	187
Considerations on Rule Induction Methods by the Conventional Rough Set Theory from a View of STRIM  Tetsuro Saeki, Jiwei Fei, and Yuichi Kato	202
Multi-label Online Streaming Feature Selection Based on Spectral Granulation and Mutual Information	215
Bipolar Queries with Dialogue: Rough Set Semantics	229
Approximation by Filter Functions	243
A Test Cost Sensitive Heuristic Attribute Reduction Algorithm for Partially Labeled Data	257
Logic on Similarity Based Rough Sets	270
Attribute Reduction Algorithms for Relation Systems on Two Universal Sets	284
Toward Optimization of Reasoning Using Generalized Fuzzy Petri Nets Zbigniew Suraj	294
Sequent Calculi for Varieties of Topological Quasi-Boolean Algebras	309

Conten	ts	XVII
Rule Induction Based on Indiscernible Classes from Rough Sets in Information Tables with Continuous Values		323
Contextual Probability Estimation from Data Samples – A Generalisation Hui Wang and Bowen Wang	n	337
Application of Greedy Heuristics for Feature Characterisation and Selection: A Case Study in Stylometric Domain		350
An Optimization View on Intuitionistic Fuzzy Three-Way Decisions Jiubing Liu, Xianzhong Zhou, Huaxiong Li, Bing Huang, Libo Zhang, and Xiuyi Jia		363
External Indices for Rough Clustering		378
Application of the Pairwise Comparison Matrices into a Dispersed Decision-Making System With Pawlak's Conflict Model		392
Exploring GTRS Based Recommender Systems with Users of Different Rating Patterns		405
Boundary Region Reduction for Relation Systems		418
A Method to Determine the Number of Clusters Based on Multi-validity Index		427
Algebras from Semiconcepts in Rough Set Theory		440
Introducing Dynamic Structures of Rough Sets. The Case of Text Processing: Anaphoric Co-reference in Texts in Natural Language Wojciech Budzisz and Lech T. Polkowski		455
Reduct Calculation and Discretization of Numeric Attributes in Entity Attribute Value Model		464
Medical Diagnosis from Images with Intuitionistic Fuzzy Distance Measures		479

## XVIII Contents

Rough Set Approach to Sufficient Statistics	491
A Formal Study of a Generalized Rough Set Model Based on Relative Approximations	502
Decidability in Pre-rough Algebras: Extended Abstract	511
A Conflict Analysis Model Based on Three-Way Decisions	522
Tolerance Relations and Rough Approximations in Incomplete Contexts Tong-Jun Li, Wei-Zhi Wu, and Xiao-Ping Yang	533
On Granular Rough Computing: Epsilon Homogenous Granulation  Krzysztof Ropiak and Piotr Artiemjew	546
Fuzzy Bisimulations in Fuzzy Description Logics Under the Gödel Semantics	559
An Efficient Method for Mining Clickstream Patterns	572
Transformation Semigroups for Rough Sets	584
A Sequential Three-Way Approach to Constructing a Co-association  Matrix in Consensus Clustering	599
Fuzzy Partition Distance Based Attribute Reduction in Decision Tables Van Thien Nguyen, Long Giang Nguyen, and Nhu Son Nguyen	614
Dynamic and Discernibility Characteristics of Different Attribute  Reduction Criteria	628
A New Trace Clustering Algorithm Based on Context in Process Mining Hong-Nhung Bui, Tri-Thanh Nguyen, Thi-Cham Nguyen, and Quang-Thuy Ha	644
Author Index	659