Advances on Rough Sets and Knowledge Technology

Preface

_____

This special issue is dedicated to the 2012 Joint Rough Set Symposium (JRS 2012) held in Chengdu, China, Aug. 17-20, 2012. JRS 2012 is a joint conference of the 8th International Conference on Rough Sets and Current Trends in Computing (RSCTC 2012) and the 7th International Conference on Rough Sets and Knowledge Technology (RSKT 2012). After peer review, the editor-in-chief and guest editors accept 8 revised and extended papers of JRS 2012 to reflect the current development of rough sets and knowledge technology.

The paper "A Multifaceted Analysis of Probabilistic Three-way Decisions", by Xiaofei Deng, Yiyu Yao, analyzes probabilistic three-way decisions/classifications from several perspectives. Through identifying an inadequacy of two-way decisions with respect to controlling the levels of various decision errors, they examine the motivations and advantages of three-way decisions. Then, a general framework is shown for computing the required thresholds of a three-way decision model as an optimization problem. Two special cases, e.g., a decision-theoretic rough set model and an information-theoretic rough set model are investigated. Finally, a heuristic algorithm is developed based on gradient-descent approaches for finding the required thresholds. It is validated that three-way decisions are a theoretically sound and practically applicable decision-making strategy to deal with insufficient uncertain information.

The paper "Application of Rough Set Theory to Prediction of Antimicrobial Activity of Bis-Quaternary Imidazolium Chlorides", by Łukasz Pałkowski, Jerzy Błaszczyński, Jerzy Krysiński, Roman Słowiński, Andrzej Skrzypczak, Jan Błaszczak, Eugenia Gospodarek and Joanna Wróblewska, investigates relationships between chemical structure, surface active properties and antibacterial activity of 70 bis-quaternary imidazolium chlorides. An information system is constructed including seven condition attributes which describe chemical structure and properties of imidazolium chlorides and a decision attribute corresponds to antimicrobial properties. A priori unknown rules exhibiting monotonicity relationships in the data are discovered by application of dominance-based rough set approach. Prognostic models of new compounds with favorable antimicrobial properties may be created by the discovered strong decision rules. It is also showed that the dependencies between chemical structure, physicochemical properties, and biological activity of analyzed molecules can be explained by the discovered rules.

The paper "Preorder Information Based Attributes Weights Learning in Multi-attribute Decision Making", by Baoli Wang, Jiye Liang and Yuhua Qian, presents a granular computing based data-driven

weights learning method for processing the multi-attribute decision making problems. The implied pre-order relations under four evaluation forms are analyzed and their corresponding preorder granular structures are constructed. Then a fuzzy measure of an attribute set by the similarity degree between a special preorder pairs of granular structures is defined. Two illustrated examples are provided to show the feasibility and effectiveness of the proposed method.

The paper "Data-driven Valued Tolerance Relation Based on the Extended Rough Set", by Guoyin Wang, Lihe Guan, Weizhi Wu and Feng Hu, aims to deal with two problems in the valued tolerance relation based rough sets, namely, how to define a reasonable tolerance degree and how to obtain a reasonable threshold from incomplete information systems. A data-driven valued tolerance relation based rough set model is presented. An objective method for calculating the tolerance degree is proposed, which is on the basis of the statistical characteristics of attribute values in incomplete information systems, and considers the effect of the number of the same known values between objects. Then, an auto-selection algorithm of the threshold is presented, which does not require any prior domain knowledge. Experimental results indicate the proposed model can achieve better and more stable classification results than other extended models of the classical rough set theory.

The paper "An Analysis of Probabilistic Approximations for Rule Induction from Incomplete Data Sets", by Patrick G. Clark, Jerzy W. Grzymala-Busse and Zdzislaw S. Hippe, evaluates whether the probabilistic approximations should be used in rule induction from incomplete data. Experiments are designed by using six standard data sets, where four of them are incomplete and two of them have missing attribute values that are randomly replaced. Two interpretations of missing attribute values: lost values and "do not care" conditions and three definitions of approximations: singleton, subset and concept are employed in these six data sets. It is shown that the proper probabilistic approximations are not very useful for rule induction from incomplete data sets. It is also indicated that the best approximation is dependent on the testing data sets.

The paper "Fuzzy Rough Decision Trees", by Shuang An, Qinghua Hu, Hong Shi and Jianwu Dang, develops a fuzzy rough decision tree algorithm to address classification problems described with symbolic, real-valued or fuzzy features. In this algorithm, the fuzzy dependency is used to select splitting nodes, the pureness which is calculated with lower approximation of fuzzy rough sets determines splits of branches, and the pureness of branches or the number of nodes decides the stopping criterion. By the proposed algorithm, a classification tree can be directly generated without discretization or fuzzification of continuous attributes. Experimental results show its effectiveness compared with several classical decision tree algorithms.

The paper "Incremental Maintenance of Rough Fuzzy Set Approximations under the Variation of Object Set", by Anping Zeng, Tianrui Li, Junbo Zhang and Hongmei Chen, discusses the dynamic approach for maintenance of rough fuzzy set approximations when the object set in fuzzy decision systems evolves over time. Principles for updating lower and upper approximations under rough fuzzy set are firstly presented. The corresponding algorithms for dynamic maintenance of approximations are developed in the cases of immigration and emigration of multiple objects. Experiments on 6 UCI data sets are conducted to show the effectiveness of the proposed methods.

The paper "Interactive Method for Semantic Document Indexing Based on Explicit Semantic Analysis", by Wojciech Świeboda, Adam Krasuski, Hung Son Nguyen and Andrzej Janusz, introduces a general framework incorporating semantic indexing and search of texts within scientific document repositories. A semantic interpreter for automatic tagging of textual data is interactively updated based on feedback from the users to improve quality of the tags that it produces. The Explicit Semantic Analysis

(ESA) method is used to index the document corpus. The relatedness between words and concepts are measured by an external knowledge base. Then meaningful concepts are assigned to given texts by those assessments. The weights expressing relations between particular words and concepts are showed to be improved by interaction with users or by employment of expert knowledge. The feasibility of the proposed approach is validated by experiments on a document corpus acquired from the PubMed Central repository.

Finally, we wish to thank all the authors for their contribution. We wish to express our gratitude to all the referees for their careful, insightful and constructive reviews that led to further improvement of the articles. We would like to give special thanks to the Editor-in-Chief of the Journal, Prof. Damian Niwiński, for accepting to publish this Special Issue and for his help throughout the publication process.

Acknowledgement

Guest Editors

Tianrui Li
Southwest Jiaotong University
Chengdu 610031, China
trli@swjtu.edu.cn

Hongmei Chen
Southwest Jiaotong University
Chengdu 610031, China
hmchen@swjtu.edu.cn

JingTao Yao
University of Regina
Regina, Saskatchewan, S4S 0A2, Canada
jtyao@cs.uregina.ca

Hung Son Nguyen
The University of Warsaw
Banacha 2, 02-097, Warsaw, Poland
son@mimuw.edu.pl