International Journal of Approximate Reasoning 55 (2014) 225–237

Contents lists available at SciVerse ScienceDirect

International Journal of Approximate Reasoning

journal homepage: www.elsevier.com/locate/ijar

Multigranulation decision-theoretic rough sets

Yuhua Qian [a], Hu Zhang [b], Yanli Sang [b], Jiye Liang [a,*]

[a] Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University, Taiyuan, 030006 Shanxi, China
[b] School of Computer and Information Technology, Shanxi University, Taiyuan, 030006 Shanxi, China

ARTICLE INFO

Article history:
Available online 23 March 2013

Keywords:
Decision-theoretic rough sets
Granular computing
Multigranulation
Bayesian decision theory

ABSTRACT

The Bayesian decision-theoretic rough sets propose a framework for studying rough set approximations using probabilistic theory, which can interprete the parameters from existing forms of probabilistic approaches to rough sets. Exploring rough sets in the viewpoint of multigranulation is becoming one of desirable directions in rough set theory, in which lower/upper approximations are approximated by granular structures induced by multiple binary relations. Through combining these two ideas, the objective of this study is to develop a new multigranulation rough set model, called a multigranulation decision-theoretic rough set. Many existing multigranulation rough set models can be derived from the multigranulation decision-theoretic rough set framework.

1. Introduction

Rough set theory, originated by Pawlak [24,25], has become a well-established theory for uncertainty management in a wide variety of applications related to pattern recognition, image processing, feature selection, neural computing, conflict analysis, decision support, data mining and knowledge discovery [3,5,10,11,15,16,28–31,34,36,41,55]. In the past ten years, several extensions of the rough set model have been proposed in terms of various requirements, such as the decision-theoretic rough set model (see [51]), the variable precision rough set (VPRS) model (see [56,58]), the rough set model based on tolerance relation (see [12–14]), the Bayesian rough set model (see [37]), the Dominance-based rough set model (see [4]), game-theoretic rough set model (see [6,7]), the fuzzy rough set model and the rough fuzzy set model (see [2]).

Recently, the probabilistic rough sets have been paid close attentions [8,45,48,50,52]. A special issue on probabilistic rough sets was set up in International Journal of Approximate Reasoning, in which six relative papers were published [48]. Yao presented a new decision making method based on the decision-theoretic rough set, which is constructed by positive region, boundary region and negative region, respectively [52]. In the literature [50], the author further emphasized the superiority of three-way decisions in probabilistic rough set models. In fact, the probabilistic rough sets are developed based on the Bayesian decision principle, in which its parameters can be learned from a given decision table. Three-way decisions are most of superiorities of probabilistic rough set models. The decision-theoretic rough sets can derive various existing rough set models through setting the thresholds α and β. Since the decision-theoretic rough sets was proposed by Yao [49], it have attracted more and more concerns. Azam and Yao [1] proposed a threshold configuration mechanism for reducing the overall uncertainty of probabilistic regions in the probabilistic rough sets. Jia et al. [9] developed an optimization representation of decision-theoretic rough set model, and gave a heuristic approach and a particle swarm optimization approach for searching an attribute reduction with a minimum cost. Liu et al. [23] combined the logistic regression and the decision-theoretic rough set into a new classification approach, which can effectively reduce the misclassification rate. Yu et al. [53] applied decision-theoretic rough set model for automatically determining the number of clusters with much smaller time cost.

* Corresponding author. Tel./fax: +86 0351 7018176.
E-mail addresses: jinchengqyh@126.com (Y. Qian), zhanghu@sxu.edu.cn (H. Zhang), sangyl@sxu.edu.cn (Y. Sang), ljy@sxu.edu.cn (J. Liang).

0888-613X/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ijar.2013.03.004

In the view of granular computing (proposed by Zadeh [54]), in existing rough set models, a general concept described by a set is always characterized via the so-called upper and lower approximations under a single granulation, i.e., the concept is depicted by known knowledge induced from a single relation (such as equivalence relation, tolerance relation and reflexive relation) on the universe [17,18,51]. Conveniently, this kind of rough set models is called single granulation rough sets, just SGRS. In many circumstances, we often need to describe concurrently a target concept through multi binary relations according to a user's requirements or targets of problem solving. Based on this consideration, Qian et al. [26–28] introduced multigranulation rough set theory (MGRS) to more widely apply rough set theory in practical applications, in which lower/upper approximations are approximated by granular structures induced by multi binary relations. From the viewpoint of rough set's applications, the multigranulation rough set theory is very desirable in many real applications, such as multi-source data analysis, knowledge discovery from data with high dimensions and distributive information systems.

Since the multigranulation rough set was proposed by Qian in 2006 [26], the theoretical framework have been largely enriched, and many extended multigranulation rough set models and relative properties and applications have also been proposed and studied [27–32]. Wu and Leung [39] proposed a formal approach to granular computing with multi-scale data measured at different levels of granulations, and studied theory and applications of granular labelled partitions in multi-scale decision information systems. Tripathy et al. [38] developed an incomplete multigranulation rough sets in the context of intuitionistic fuzzy rough sets and gave some important properties of the new rough set model. Raghavan and Tripathy [33] first researched the topological properties of multigranulation rough sets. Based on the idea of multigranulation rough sets, Xu et al. [42–44] developed a variable multigranulation rough set model, a fuzzy multigranulation rough set model and an ordered multigranulation rough set model. Wu [40] extended classical multigranualtion rough sets to the version based on a fuzzy relation, and proposed a new multigranulation fuzzy rough set (MGFRS). Zhang et al. [57] defined a variable precision multigranulation rough set, in which the optimistic multigranulation rough sets and the pessimistic one can be regarded as two extreme cases. Through introducing some membership parameters, this model becomes a multigranulation rough set with dynamic adaption according to practical acquirements. Yang et al. [46,47] examined the fuzzy multigranulation rough set theory, and revealed the hierarchical structure properties of the multigranulation rough sets. Liu and Miao [21,22] established a multigranulation rough set approach in covering contexts. Liang et al. [19] presented a kind of efficient feature selection algorithms for large scale data with a multigranulation strategy. She et al. [35] explored the topological structures and the properties of multigranulation rough sets. Lin et al. [20] gave a neighborhood multigranulation rough set model for multigranulation rough data analysis in the context of hybrid data. In the murigranulation rough set theory, each of various binary relation determines a corresponding information granulation, which largely impacts the commonality between each of the granulations and the fusion among all granulations. As one of very important rough set models, the decision-theoretic rough sets (DTRS) are still not be researched in the context of multigranulation, which limits its further applications in many problems, such as multi-source data analysis, knowledge discovery from data with high dimensions and distributive information systems.

In what follows, besides those motivations mentioned in first multigranulation rough set paper (see Cases 1–3 in the literature [29]), we further emphasize the specific interest of multigranulation rough sets, which can be illustranted from the following three aspects.

- Multigranulation rough set theory is a kind of of information fusion strategies through single granulation rough sets. Optimistic version and pessimistic version are only two simple methods in these information fusion approaches, which are used to easily introduce multigranulation ideas to rough set theory.
- In fact, there are some other fusion strategies [20,39–41,43]. For instance, in the literature [39], Xu et al. introduced a supporting characteristic function and a variable precision parameter β, called an information level, to investigate a target concept under majority granulations.
- With regard to some special information systems, such as multi-source information systems, distributive information systems and groups of intelligent agents, the classical rough sets can not be used to data mining from these information systems, but multigranulation rough sets can.

In this study, our objective is to develop a new multigranulation rough decision theory through combining the multigranulation idea and the Bayesian decisoin theory, called multigranulation decision-theoretic rough sets (MG-DTRS). We mainly give three common forms, the mean multigranulation decision-theoretic rough sets, the optimistic multigranulation decision-theoretic rough sets, and the pessimistic multigranulation decision-theoretic rough sets.

The study is organized as follows. Some basic concepts in classical rough sets and multigranulation rough sets are briefly reviewed in Section 2. In Section 3, we first analyze the loss function and the entire decision risk in the context of multigranulation. Then, we propose three multigranulation decision-theoretic rough set forms that include the mean multigranulation decision-theoretic rough sets, the optimistic multigranulation decision-theoretic rough sets, and the pessimistic multigranulation decision-theoretic rough sets. When the thresholds have a special constraint, the multigranulation decision-theoretic rough sets will produce one of various variables of multigranulation rough sets. In Section 4, we establish the relationships among multigranulation decision-theoretic rough sets (MG-DTRS), other MGRS models, single granulation decision-theoretic rough sets (SG-DTRS) and other SGRS models. Finally, Section 5 concludes this paper by bringing some remarks and discussions.