International Journal of Approximate Reasoning 98 (2018) 236–258

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Double-quantitative decision-theoretic approach to multigranulation approximate space ☆

Jianhang Yu [a], Biao Zhang [a], Minghao Chen [a,*], Weihua Xu [b]

[a] Department of Mathematics, Harbin Institute of Technology, Harbin 150001, PR China
[b] School of Mathematics and Statistics, Chongqing University of Technology, Chongqing 400054, PR China

ARTICLE INFO

Article history:
Received 4 July 2017
Received in revised form 2 May 2018
Accepted 2 May 2018
Available online 7 May 2018

Keywords:
Double-quantification
Decision-theoretic rough set
Graded rough set
Multigranulation approximate space
Bayesian decision

ABSTRACT

The decision-theoretic rough set, as a special case of probabilistic rough set, mainly utilizes conditional probability to express relative quantitative information, while the graded rough set is characterized by absolute quantitative information between the partitions and basic concept. Thus, the double-quantification integrating relative and absolute quantitative information has become a fundamental topic for model construction, especially for developing the decision-theoretic rough set. In this study, we propose a basic framework of double-quantitative decision-theoretic rough set based on Bayesian decision and graded rough set approach in multigranulation approximate space. Three pairs of double-quantitative multigranulation decision-theoretic rough set models are established, which consist of a dual of optimistic double-quantitative multigranulation decision-theoretic rough sets, pessimistic double-quantitative multigranulation decision-theoretic rough sets and mean double-quantitative multigranulation decision-theoretic rough sets. These models essentially indicate the relative and absolute information quantification. Furthermore, some essential properties of these models are addressed and the decision rules which incorporate the relative and absolute quantitative information are investigated. Finally, an illustrative case about medical diagnosis is conducted to interpret and evaluate the double-quantitative decision-theoretic approach.

1. Introduction

Rough set (RS) theory which was originated by Pawlak [24,26], is an extension of the classical set theory and could be regarded as a mathematical and soft computing tool to handle imprecision, vagueness and uncertainty in data analysis. It has become a well-established theory for uncertainty management in a wide variety of applications related to pattern recognition [37], information fusion [8], feature selection [11,12], uncertainty analysis [5,17], rule learning [15], data modeling [36], and knowledge discovery [57]. Given there are no fault tolerance mechanisms between equivalence classes and basic concept set, several proposals of generalized quantitative rough set models were developed to resolve this limitation by using a graded set inclusion. The probabilistic rough set (PRS) introduces the probability uncertainty measure into RS [25], which forms the basis of mainstream quantitative models [1,21,23,40,52,53,59]. PRS offers measurability, generality,

☆ This work is supported by the National Natural Science Foundation of China (Nos. 11771111, 11471088, 61472463, 61402064, 61772002) and the National Natural Science Foundation of CQ CSTC (No. CSTC 2015jcyjA40053).

* Corresponding author.
E-mail addresses: yujh2013@foxmail.com (J. Yu), zhangb@hit.edu.cn (B. Zhang), chenmh130264@aliyun.com (M. Chen), chxuwh@gmail.com (W. Xu).

https://doi.org/10.1016/j.ijar.2018.05.001
0888-613X/

and flexibility and exhibits a series of concrete models which consist of the decision-theoretic rough set (DTRS) [49], game-theoretic rough set [2], variable rough set [58], Bayesian rough set [39], and parameter rough set [10]. They aim at modeling data relationships expressed in terms of frequency distribution rather than in terms of a full inclusion relation. With the exception of PRS, the graded rough set (GRS) depends on two absolute measures becomes another basic type of quantitative model [50].

In the viewpoint of information quantification, the DTRS and GRS can respectively reflect relative and absolute quantitative information about the degree of overlap between equivalence classes and concept set [13,62]. The relative and absolute quantitative information are two distinctive objective sides that describe approximate space, and each has its own virtues and application environments, so that none can be neglected. Here, we illustrate three examples to highlight the significance of combining the relative quantification and absolute quantification, and the necessity of these two types of quantitative model is exhibited in different scenarios.

(1) There is a good project that needs to be invested due to the lack of funds. So decision makers prepare to attract 20 million dollars to support the project implementation, and there are two mutually exclusive investment companies Alpha and Beta as alternatives that means only one investment company will be selected. According to the financial report, Alpha has a 50 million dollars budget available for investment and plans to invest 15 million dollars in this project, while Beta has a 30 million dollars budget available for investment and prepares to invest 12 million dollars in this project. So, which one is more suitable as a partner? There is no doubt that Alpha is the preferable choice, although the relative proportion of budget is only 30%, which is lower than that 40% of Beta. Here, one focus mainly on the absolute quantitative information and believes that lower priority of the relative quantitative information.
(2) A company is ready to purchase a large quantity of products in the near future. A and B are the suppliers of this product, and the price difference between them is tiny. Therefore the chief executive officer prepares to conduct a selective examination of product quality to determine who will win the bid. The results of sample survey show that company A has 15 substandard products in the 300 samples and company B has 20 substandard products in 1000 samples. It is clearly that the number of unqualified products of A than B more $(15 < 20)$. But we still believe that B is a better candidate due to the product qualified rate is $98\% > 95\%$. As this example suggests, the higher priority should be the relative quantitative information not the absolute quantitative information.
(3) Suppose A and B are research institutes with 70 and 30 proposed projects, respectively, but only 50 projects will be approved and funded in total. How does one make a final decision on which projects to implement? If only the relative quantitative information is considered, one may conclude that A and B will achieve 35 and 15 establishment projects, respectively. However, is this division fair and reasonable in reality? It may be feasible if the two institutes with almost the same scientific research ability. However, if the research level of A is much higher than B, then A should obtain more than 35 projects and B should receive less than 15 projects. In practice, the absolute quantitative information that the number of approved projects is a pivotal index. Therefore, an ideal evaluation must employ rational combinations of the two evaluation indexes.

The relative and absolute measures adopt different quantitative views for measurement, thus underlying quantitative applications. Usually, both hold heterogeneity and complementarity, and thus, each relies on its essential benefit to occupy its own dominant environment. In recent decades, a lot of research interests are attracted by the double-quantitative fusion of relative quantitative information and absolute quantitative information. Zhang developed a comparative study of relative quantitative rough set model and absolute quantitative rough set model [62], then he systematically researched the issues of double-quantitative fusion [63–66]. Based on these achievements, Li constructed two double-quantitative decision-theoretic rough set models [13], Fan studied this issue based on logical conjunction and logical disjunction operation [6] and Fang proposed another kind of double relative quantitative decision-theoretic rough set models, which essentially indicate the relative and absolute quantification [7].

The granular computing (GrC), another powerful tool in artificial intelligence and data processing, which is a term coined jointly by Zadeh [60,61] and Lin [18]. Bargiela [3,4] and Pedrycz [28–30] conducted a series of systematic studies on GrC and many constructive achievements were obtained. In the view of GrC, a general concept described by a set is always characterized via the so-called upper and lower approximations under a single granulation, namely, the concept is depicted by knowledge induced from a single binary relation on the universe of discourse [51]. In many practical circumstances, we need to describe concurrently a target concept through multi binary relations according to users' requirements and targets of problem solving. Based on this thought, Qian et al. first investigated multigranulation rough set (MGRS) theory to more widely apply rough set theory [31], and introduced the incomplete multigranulation rough set [32]. Since the MGRS was established, the theoretic framework have been largely enriched, and many generalized MGRS models and their applications have also been investigated [33,34]. Wu extended classical MGRS to a novel version based on a fuzzy binary relation [41]. She explored the topological structures and the essential properties of MGRS [38], and Yang revealed the hierarchical structures properties of the MGRS [48]. Furthermore, Wu and Leung proposed a formal approach to granular computing with multi-scale data measured at different levels of granulations [42], Lin applied this method to information fusion by combining with evidence theory [19]. Prior to this study, we have expanded the classical MGRS model to a generalized formal [44], and developed the MGRS approach in fuzzy tolerance approximation space [45] and ordered information system [46,55,56],