Information Sciences 299 (2015) 312–336

Contents lists available at ScienceDirect

Information Sciences

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

An expanded double-quantitative model regarding probabilities and grades and its hierarchical double-quantitative attribute reduction

Xianyong Zhang [a,b,c,*], Duoqian Miao [b,c]

[a] College of Mathematics and Software Science, Sichuan Normal University, Chengdu 610068, PR China
[b] Department of Computer Science and Technology, Tongji University, Shanghai 201804, PR China
[c] Key Laboratory of Embedded System and Service Computing, Ministry of Education, Shanghai 201804, PR China

ARTICLE INFO

Article history:
Received 23 April 2014
Received in revised form 2 November 2014
Accepted 3 December 2014
Available online 16 December 2014

Keywords:
Rough set theory
Decision-theoretic rough set
Graded rough set
Double quantification
Granular computing
Attribute reduction

ABSTRACT

Probabilities and grades serve as relative and absolute measures, respectively. They are used to establish the decision-theoretic rough set (DTRS) and graded rough set (GRS) – two basic quantitative models. The double-quantification of probabilities and grades exhibits systematicness and completeness in view of the two-dimensional feature of the approximate space; however, double-quantitative construction becomes a problem, and double-quantitative reduction is rarely reported. Thus, this paper mainly constructs an expanded double-quantitative model by logically integrating probabilities and grades; it further studies relevant double-quantitative reduction by hierarchically preserving specific regions. (1) First, a novel model is established via the logic integration and expansion requirement, and its regional system and granular hierarchy are studied via granular computing. Thus, regional semantics is extracted via basic semantics granules. Regional calculation is realized by two algorithms, and the algorithm regarding calculation granules exhibits optimization according to algorithm analyses. (2) Second, three types of model-regional preservation reducts and their hierarchy are discussed in the two-category case. Thus, SRP-Reduct, CRP-Reduct, and APP-Reduct are studied by exploring four-region preservation properties, constructing two-region classification regions, and preserving four original approximations, respectively. Furthermore, a relevant reduction hierarchy is thoroughly achieved. (3) Moreover, the model and its reduction are illustrated by two examples of decision tables. The constructional model conducts double-quantification regarding probabilities and grades; thus, it exhibits double-quantitative semantics and benignly expands DTRS-Model, GRS-Model, and Pawlak-Model. Furthermore, its hierarchical reduction reflects some double-quantitative reduction essence; thus, its reduction expands qualitative Pawlak-Reduction while guides quantitative DTRS-Reduction and GRS-Reduction.

1. Introduction

Rough set theory (RS-Theory) is an important mathematical theory for information processing. The initial Pawlak-Model [31,32] – only a qualitative model – does not perform well in quantitative problems in applications. In contrast, quantitative

* Corresponding author at: College of Mathematics and Software Science, Sichuan Normal University, Chengdu 610068, PR China.
E-mail addresses: xianyongzh@sina.com.cn (X. Zhang), miaoduoqian@163.com.cn (D. Miao).

http://dx.doi.org/10.1016/j.ins.2014.12.006
0020-0255/

models can effectively solve problems of data noise, thus holding application significance; Ref. [55] proposed a relevant framework. Thus, the probabilistic rough set (PRS) [1,12,52,71] utilizes the probability measure to proffer application merits regarding measurability, generality, and robustness. PRS involves several concrete models, including the decision-theoretic rough set (DTRS) [58] and the variable precision rough set (VPRS) [72] (which will be considered next). Herein, several PRS-Models are first introduced by the probability and threshold. The 0.5-probabilistic rough set [44] uses probability threshold 0.5; the parameterized rough set [4] uses two thresholds to describe the absolute and relative rough membership; the Bayesian rough set [40] compares the concept probability and prior probability; and the game-theoretic rough set [5] adopts the game theory to determine thresholds. In fact, PRS holds statistical advantages due to information concentration of the relative probability measure. The graded rough set (GRS) [56] also serves as a fundamental quantitative model due to the intuition and vividness of its absolute grade measure. Next, DTRS, VPRS, GRS – three basic types of quantitative models – are introduced appropriately. DTRS [58] established threshold semantics and calculation by using the Bayesian risk decision and three-way decisions. For DTRS, Refs. [53,54] analyzed three-way decisions superiority; Refs. [13,14,23,38] discussed model development and threshold calculation; Refs. [7,8,10,25,59,65,69] studied attribute reduction; Refs. [9,17,19,22,61] researched model applications (regarding regression, clustering, and semi-supervised learning); and Refs. [17,18,21,70] exploited multi-category construction. VPRS [72] exhibited error tolerance by considering relative misclassification. For VPRS, Refs. [2,6,24,26,43] studied attribute reduction, and Refs. [28,30,46,47,49] performed model applications in geological and medical fields, among others. GRS [56] was proposed by exploring relationships between rough sets and modal logics, and Refs. [20,48,57] reported some model constructions.

As is known, measures underline the applications and expansion of quantitative models; moreover, further measure fusion holds theoretical and practical significance. In fact, PRS and GRS mainly depend on the relative probability and absolute grade to implement quantitative applications and theoretical expansion. In RS-Theory, the relative ratio $\frac{|[x]\cap X|}{|[x]|}$ serves as a core measure in the approximate space; in fact, it is related to several different terms – the rough membership [33], probability [58], misclassification degree [72], and precision [64,66]. Meanwhile, internal grade $|[x]\cap X|$ and external grade $|[x]| - |[x]\cap X|$ [64,66] act as two important absolute measures. Relative and absolute measures reflect relative and absolute accuracy or fault-tolerance from two different quantitative viewpoints. Thus, both exhibit a close, supplementary, and dialectical relationship, and each one actually has its own representation virtues and application environments. Furthermore, double-quantification regarding their fusion has systematicness and feasibility. It also exhibits completeness for RS-Theory because of the two-dimensional feature of the approximate space [64,66]. Thus, the double-quantification inherits RS-Theory's quantitative essence and can adapt to complex application environments. Moreover, the double-quantification can be used to construct models to improve or expand the basic quantitative and qualitative models. Therefore, the double-quantification becomes a novel and valuable subject; however, its construction becomes a problem. At present, some fruits regarding the precision and grade measures preliminarily exist within the VPRS and GRS frameworks, where threshold system (β, k) is used. Regarding VPRS and GRS, Ref. [67] conducted a comparative study via their similarity; Refs. [3,48] investigated their transformations via broader relations; Ref. [66] constructed two extended double-quantitative models by using the Cartesian product; and Ref. [64] established a general framework of double-quantification exploration by using logical combination. In this paper, a novel double-quantitative model with benign expansion will be constructed by logically integrating the probability and grade measures. A more general threshold system (α, β, k) will be examined herein.

Attribute reduction with reducts becomes an essential subject in RS-Theory by virtue of optimization and generalization, thus underlining many practical applications. Different reduction algorithms and reduct constructions have been extensively explored in some research [11,15,41,42,60,68]. In particular, Ref. [27] investigated the hierarchy of qualitative Pawlak-Reduction. Classical reduction theory depends on the classification-positive region (C-POS); thus, Pawlak-Reduction mainly preserves C-POS due to the monotonicity of qualitative C-POS change. However, quantitative C-POS change usually exhibits non-monotonicity and uncertainty [63]. Hence, quantitative reduction accompanies some anomalies [2,43,59]; in fact, it has already transcended qualitative reduction, thus becoming a focus and a difficulty. Refs. [7,8,10,25,59,65,69] and Refs. [2,6,24,26,43] studied DTRS-Reduction and VPRS-Reduction, respectively. In particular, Ref. [65] analyzed the hierarchy of two-category DTRS-Reducts. Moreover, GRS-Reduction has become rare, and further double-quantitative reduction is nearly entirely absent. In this paper, the double-quantitative reduction of the constructional double-quantitative model will be investigated by a natural strategy of region preservation.

Against the above backgrounds, this paper mainly constructs an expanded double-quantitative model by logically integrating probabilities and grades. It further studies relevant double-quantitative reduction by hierarchically preserving specific regions. In particular, granular computing (GrC) [16,62] effectively provides a structural methodology for hierarchical information processing by emphasizing multiple granules, levels, and perspectives. Refs. [35–37,45,50,51] investigated general GrC, and Refs. [29,34,39,64,66] further considered GrC in RS-Theory. In fact, double-quantification [64,66] is closely related to GrC. For the new model from the double-quantification, GrC will be utilized to investigate its metrical integration, regional system, quantitative semantics, optimal calculation, and hierarchical reduction. Thus, three corresponding parts will be discussed.

(1) First, the novel model is constructed via the logic integration and expansion requirement; furthermore, its regional system and granular hierarchy are studied. Thus, quantitative semantics and the basic calculation of model regions are achieved via granule construction and algorithm analyses, respectively.