Information Sciences 334–335 (2016) 174–204

Contents lists available at ScienceDirect

Information Sciences

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

Quantitative/qualitative region-change uncertainty/certainty in attribute reduction: Comparative region-change analyses based on granular computing

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

[a] College of Mathematics and Software Science, Sichuan Normal University, Chengdu 610068, Sichuan, China
[b] Department of Computer Science and Technology, Tongji University, Shanghai 201804, China
[c] Department of Computer Science, University of Regina, Regina S4S0A2, Saskatchewan, Canada

ARTICLE INFO

Article history:
Received 14 December 2014
Revised 7 November 2015
Accepted 26 November 2015
Available online 2 December 2015

Keywords:
Rough set theory
Granular computing
Attribute reduction
Uncertainty
Probabilistic rough set
Decision-theoretic rough set

ABSTRACT

Attribute reduction is a fundamental research theme in rough sets and granular computing (GrC). Its scientific construction originally depends on the region-change law. At present, only region-change non-monotonicity/monotonicity is mined in the quantitative/qualitative model. The in-depth region-change truth and its GrC mechanism have significance, especially for follow-up attribute reduction. This paper commences probing region-change essence, mainly from a novel uncertainty/certainty viewpoint. Concretely, we make comparative region-change analyses based on GrC, by resorting to the qualitative Pawlak-Model and quantitative DTRS-Model (the decision-theoretic rough set model). (1) Knowledge-coarsening is investigated to describe attribute deletion. (2) Granule-merging and its region-distribution are studied to probe region-change functions. (3) Region-change is analyzed in Pawlak-Model to mine qualitative region-change certainty and its relevant properties. (4) Region-change is analyzed in DTRS-Model to mine quantitative region-change uncertainty and its relevant properties. (5) Comparative region-change analyses are summarized, and further experiment verification is provided. Knowledge-coarsening and granule-merging establish GrC mechanisms for extensive region-change analyses. Quantitative/qualitative region-change uncertainty/certainty and relevant principles are discovered via DTRS-Model/Pawlak-Model. By virtue of the GrC technology and comparative strategy, this study reveals region-change uncertainty/certainty to deepen region-change non-monotonicity/monotonicity; furthermore, it underlies attribute reduction, especially with regard to quantitative models.

1. Introduction

Rough set theory (RS-Theory) [32] is a fundamental uncertainty analysis theory that is designed to process uncertain, imprecise, and incomplete data information. Its basis is *the rough set model (RS-Model)*. The initial RS-Model, Pawlak-Model [31], originates from qualitative definitions for regions and thus acts as only a qualitative model. Because it lacks quantitative mechanisms, Pawlak-Model cannot effectively tackle extensive quantitative problems that concern fault-tolerance or robustness. Uncertainty measure-based quantitative models, which are partly unified by subsethood measures [54], exhibit improvements and

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

http://dx.doi.org/10.1016/j.ins.2015.11.037

applications. As a classical uncertainty measure, the probability was introduced into RS-Theory to construct *the probabilistic rough set (PRS)* [1,28,45,51,52,59,73]. At present, PRS relies on measurability, generality, and flexibility to generate mainstream models of multiple types. With regard to the probability and threshold, *the decision-theoretic rough set (DTRS)* uses a pair of thresholds [57], the variable precision rough set uses a single threshold [75], the game-theoretic rough set determines the threshold via game theory [1], the Bayesian rough set compares the prior and posterior probabilities [42], and the parameterized rough set uses two thresholds to describe rough membership [6]. Note that the probability in PRS is only a type of relative measure with a statistical concentration. In contrast, an absolute measure that involves vivid intuition also exists, such as the grade. Accordingly, the grade measure is utilized to construct the graded rough set, which is a type of quantitative model that has absoluteness [24,55]. Furthermore, the relative and absolute measures systematically construct double-quantitative models to manifest diversity and completeness [16,65–67].

DTRS, a fundamental quantitative model, is appropriately introduced for relevant use. DTRS adopts the conditional probability and Bayesian risk decision to establish threshold-quantitative semantics and three-way decisions [57]. As a result, DTRS improves on some basic RS-Models and provides a quantitative exploration platform. For relevant studies, three-way decision superiority was analyzed in [52], model development and threshold calculation were studied in [5,16–19,37,38,43,70], attribute reduction was discussed in [10,14,29,58,68,69,71], and model applications on clustering and regression were researched in [12,20,25,60]. In particular, three-way decisions have been promoted to the three-way decision theory with extensive research [5,7,13,18–20,26,61,63,74].

Attribute reduction serves as an essential subject in RS-Theory. It utilizes optimization and generalization to exhibit effective applications in knowledge discovery and data mining [11,27,29,30,35,44–47,58,66]. In the classical pattern, attribute reduction is accompanied by the RS-Model and, thus, depends on *the region-change (Rg-Change)* law that is mined in the RS-Model. In the qualitative Pawlak-Model, *the classification-positive region (Cl-POS)* change has monotonicity. This benign feature, which essentially originates from general Rg-Change monotonicity, naturally causes the Cl-POS preservation criterion to construct Pawlak-Reduction [31,32]. In the quantitative model, Cl-POS change exhibits non-monotonicity, which has been verified [10,29,30,45,46,58,66]. This new phenomenon, which in fact comes from Rg-Change non-monotonicity, is not sufficient to inspire a reasonable reduction criterion. Therefore, quantitative reduction transcends qualitative Pawlak-Reduction, and some reduction anomalies emerge [30,46,58]. At present, the quantitative reduction construction becomes not only a focus but also a difficulty. As far as quantitative DTRS-Reduction is concerned, basic reducts with regard to the Cl-POS preservation, Cl-POS measure, minimum cost, region distribution, and structure hierarchy are constructed in [71], [14], [10], [29], and [68,69], respectively; moreover, general reducts are established via multiple measures [58].

The Rg-Change law in the RS-Model underlies the attribute reduction, especially in the usual regional approach. Rg-Change non-monotonicity/monotonicity is currently acquired in the quantitative/qualitative model. For the quantitative models, quantitative expansion brings some complexity that transcends the qualitative model. In particular, Rg-Change non-monotonicity brings some reduction confusions and hampers further quantitative reduction. What is the Rg-Change truth hidden behind the non-monotonicity phenomenon? What is the relevant formation mechanism? Both of these questions become important for the quantitative models' applications, especially for quantitative reduction development. However, there are rarely in-depth reports on the quantitative Rg-Change essence, especially from an uncertainty viewpoint. Based on an extensive survey, we have discovered that Rg-Change uncertainty acts as the objective truth and broad context for Rg-Change non-monotonicity. Thereby, Rg-Change uncertainty analyses become a novel and valuable approach, and relevant uncertainty mechanisms underlie quantitative reduction. Note that uncertainty is an essential feature and a critical subject in intelligent information processing. In RS-Theory, some uncertainty aspects were explored in [1,4,8,21,36,39,40,53,64], and the relevant uncertainty was specifically measured by information entropy [28,29,44,45]. In contrast, for the qualitative model, we discover that Rg-Change certainty becomes general to underlie Rg-Change monotonicity; thus, relevant Rg-Change certainty analyses are required as well.

For uncertainty/certainty analyses, *Granular computing (GrC)* [23,62] is one of the most effective technologies, and GrC-based uncertainty/certainty analyses have become a challenging focus. In fact, GrC is a powerful structural methodology for effectively processing hierarchical information. It highlights trialistic characteristics with regard to multiple granules, levels, and perspectives. GrC exhibits extensive research [13,15,33,34,47,49,50], which is especially based on information granulation. In particular, GrC is addressed concretely in RS-Theory [9,17,21,22,37,38,41,48,56,66]. Note that attribute reduction closely adheres to GrC. Attribute reduction mainly depends on knowledge granulation, which is a type of hierarchical transformation of knowledge structures, and knowledge granulation leads to structural coarsening, granular merging, and further regional change. Therefore, knowledge granulation becomes the root cause of presentative uncertainty/certainty, and the relevant uncertainty/certainty mainly depends on information transformation among different knowledge-granular hierarchies. In short, attribute reduction depends on knowledge granulation to closely follow GrC, while the GrC approach can thoroughly analyze attribute reduction, especially its Rg-Change uncertainty/certainty.

Against the above background, this paper begins to probe the Rg-Change essence in attribute reduction, mainly by adopting a novel uncertainty/certainty viewpoint. To reveal the relevant mechanisms and rules, qualitative Pawlak-Model and quantitative DTRS-Model are concretely utilized to generate comparative Rg-Change (certainty and uncertainty) analyses based on GrC. Five gradual parts are produced, as follows: (1) Knowledge-coarsening is studied to describe attribute deletion. (2) Granule-merging and its region-distribution are studied to probe Rg-Change functions. (3) Rg-Change is analyzed in the Pawlak-Model to mine qualitative Rg-Change certainty and its relevant properties. (4) Rg-Change is analyzed in the DTRS-Model to mine quantitative Rg-Change uncertainty and its relevant properties. (5) Comparative Rg-Change analyses are summarized, and further