Contents lists available at ScienceDirect

ELSEVIER





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

Double-quantitative fusion of accuracy and importance: Systematic measure mining, benign integration construction, hierarchical 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

* Key Luboratory of Embedded System and Service Computing, Ministry of Education, Shanghai 201804, PK Chin

ARTICLE INFO

Article history: Received 4 February 2015 Revised 30 August 2015 Accepted 1 September 2015 Available online 7 September 2015

Keywords: Rough set theory Granular computing Attribute reduction Uncertainty measure Double quantification

ABSTRACT

Uncertainty measure mining and applications are fundamental, and it is possible for double-quantitative fusion to acquire benign measures via heterogeneity and complementarity. This paper investigates the doublequantitative fusion of relative accuracy and absolute importance to provide systematic measure mining, benign integration construction, and hierarchical attribute reduction. (1) First, three-way probabilities and measures are analyzed. Thus, the accuracy and importance are systematically extracted, and both are further fused into importance-accuracy (IP-Accuracy), a synthetic causality measure. (2) By sum integration, IP-Accuracy gains a bottom-top granulation construction and granular hierarchical structure. IP-Accuracy holds benign granulation monotonicity at both the knowledge concept and classification levels. (3) IP-Accuracy attribute reduction is explored based on decision tables. A hierarchical reduct system is thereby established, including qualitative/quantitative reducts, tolerant/approximate reducts, reduct hierarchies, and heuristic algorithms. Herein, the innovative tolerant and approximate reducts quantitatively approach/expand/weaken the ideal qualitative reduct. (4) Finally, a decision table example is provided for illustration. This paper performs double-quantitative fusion of causality measures to systematically mine IP-Accuracy, and this measure benignly constructs a granular computing platform and hierarchical reduct system. By resorting to a monotonous uncertainty measure, this study provides an integration-evolution strategy of granular construction for attribute reduction.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Rough set theory (RS-Theory) [35,36,59,63,70,71] represents a fundamental *granular computing (GrC)* pattern for handling uncertainty issues. The initial Pawlak-Model [35] acts only as a qualitative model, so it lacks the quantitative mechanism regarding fault-tolerance and robustness. Thus, quantitative models exhibit improvements and have applications, and they could in part be unified by the subsethood measure [61]. In particular, *the probabilistic rough set (PRS)* [1,27,28,30,51,57,58,60,78] introduces the probability uncertainty measure into RS-Theory, which forms the basis of mainstream quantitative models. PRS offers measurability, generality, and flexibility and exhibits a series of concrete models, including the *decision-theoretic rough set (DTRS)* [64],

http://dx.doi.org/10.1016/j.knosys.2015.09.001 0950-7051/© 2015 Elsevier B.V. All rights reserved. game-theoretic rough set [1,2], variable precision rough set [81], Bayesian rough set [47], and parameterized rough set [6]. With the exception of PRS, the graded rough set [25,62] depends on the grade measure to become another basic type of quantitative model.

Herein, DTRS is introduced as a model example. DTRS utilizes conditional probability and the Bayesian risk decision to establish threeway decisions and threshold-quantitative semantics [64]. As a result, DTRS improves upon some basic models and provides a quantitative exploration platform. In terms of relevant studies, three-way decisions were analyzed in [17–19,58,60,79]; model development and threshold calculation were discussed in [15,16,27,45,48]; attribute reduction was studied in [10,12,31,65,74,75]; and model applications (regarding clustering, regression, and semi-supervised learning) were addressed in [13,23,24,26,66]. In fact, three-way decisions have been expanded into three-way decision theory, and this fundamental theory has been the subject of extensive study and used in a number of

^{*} 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).



Fig. 1. Accuracy and importance in the cardinality plane.

Table 1	
Main abbreviations of this paper.	

Abbreviation	Original term
RS-Theory	Rough set theory
GrC	Granular computing
DTRS	Decision-theoretic rough set
Approx-Space	Approximate space
C-POS/C-BND	Classification-positive/boundary region
POS/NEG/BND	Set-positive/negative/boundary region
D-Table	Decision table
Kn-Coarsening	Knowledge coarsening
Gr-Merging	Granule merging
Gr-Preservation	Granule preservation
IP-Accuracy	Importance-accuracy
IIP-Accuracy	Internal importance-accuracy
Gr/Kn/Cl/Con	Granule/knowledge/classification/concept
Gr-Con/Kn-Con/	Granular concept/knowledge's concept/knowledge's
Kn-Cl	classification
MT	Monotonicity target

Table 2

Three-way probabilities and three-way measures in Bottom-System (U, R, X).

Metrical essence	Causality relevance	Metrical features	Three-way probabilities	Probability formula	Three-way measures	Measure formula
$\frac{\frac{ [x]_R \cap X }{ [x]_R }}{\frac{ [x]_R \cap X }{ X }}$	Has Has	Relativity, concentration, locality Absoluteness, directness, globality (regarding X)	Likelihood probability Posterior probability	$p(X [x]_R)$ $p([x]_R X)$	Accuracy Importance	$a_X([x]_R)$ $i_X([x]_R)$
$\frac{ [x]_R }{ U }$	Never has	Absoluteness, vividness, globality (regarding U)	Prior probability	$p([x]_R)$	Generality	$g([x]_R)$

Table 3

Granule-statistical information in Bottom-System (U, C, X).

Granule $[x]^i_{\mathcal{C}}$	Interaction cardinality $ [x]^i_{\mathcal{C}} \cap X $	Granule cardinality $ [x]_{\mathcal{C}}^{i} $	Generality $g([x]^i_{\mathcal{C}})$	Importance $i_X([x]^i_{\mathcal{C}})$	Accuracy $a_X([x]^i_{\mathcal{C}})$	IP-Accuracy $ia_X([x]^i_{\mathcal{C}})$
$[x]^1_{\mathcal{C}}$	8	16	0.16	0.200	0.5	0.1
$[x]_{\mathcal{C}}^2$	6	10	0.10	0.150	0.6	0.09
$[x]_{\mathcal{C}}^3$	6	15	0.15	0.150	0.4	0.06
$[x]^4_{\mathcal{C}}$	5	10	0.10	0.125	0.5	0.0625
$[x]^5_{\mathcal{C}}$	4	16	0.16	0.100	0.25	0.025
$[x]^6_{\mathcal{C}}$	3	3	0.03	0.075	1.0	0.075
$[x]_{\mathcal{C}}^7$	3	4	0.04	0.075	0.75	0.05625
$[x]^8_{\mathcal{C}}$	2	3	0.03	0.050	2/3	1/30
$[x]^9_{\mathcal{C}}$	2	6	0.06	0.050	1/3	1/60
$[x]_{C}^{10}$	1	1	0.01	0.025	1.0	0.025
$[x]_{C}^{11}$	0	8	0.08	0.000	0.0	0.0
$[x]_{C}^{12}$	0	8	0.08	0.000	0.0	0.0

Table 4

Knowledge's relevant uncertainty measures regarding IP-Accuracy.

Knowledge \mathcal{C}_*	Kn-Con IP-Accuracy $ia_X(\mathcal{C}_*), ia_{\neg X}(\mathcal{C}_*)$	Kn-Con IIP-Accuracy $lia_X(\mathcal{C}_*), \ lia_{-X}(\mathcal{C}_*)$	Kn-Cl IP-Accuracy $ia_{\mathcal{D}}(\mathcal{C}_*)$	Kn-Cl IIP-Accuracy $lia_{\mathcal{D}}(\mathcal{C}_*)$	IP-Accuracy discrepancy, equality rates $dr_{ia_{D}}$, $er_{ia_{D}}$	IIP-Accuracy discrepancy, equality rates dr _{lia_D} , er _{lia_D}
$\{a, b, c, d, e, f\}$	0.5437, 0.6958	0.1000, 0.2667	1.2395	0.3667	0.00%, 100.00%	0.00%, 100.00%
$\{a, b, c, d, e\}$	0.5437, 0.6958	0.1000, 0.2667	1.2395	0.3667	0.00%, 100.00%	0.00%, 100.00%
$\{a, b, c, d\}$	0.5376, 0.6917	0.1000, 0.2667	1.2293	0.3667	0.82%, 99.18%	0.00%, 100.00%
$\{a, b, c\}$	0.5366, 0.6910	0.1000, 0.2667	1.2276	0.3667	0.96%, 99.04%	0.00%, 100.00%
$\{a, b\}$	0.5193, 0.6796	0.1000, 0.0000	1.1989	0.1000	3.28%, 96.72%	72.73%, 27.27%
$\{a, c\}$	0.5050, 0.6700	0.1000, 0.2667	1.1750	0.3667	5.20%, 94.80%	0.00%, 100.00%
{ <i>b</i> , <i>c</i> }	0.5314, 0.6876	0.0000, 0.2667	1.2190	0.1000	1.65%, 98.35%	27.27%, 72.73%
<i>{a}</i>	0.4375, 0.6250	0.1000, 0.0000	1.0625	0.1000	14.28%, 85.72%	72.73%, 27.27%
$\{b\}$	0.5141, 0.6761	0.0000, 0.0000	1.1902	0.1000	3.98%, 96.02%	100.00%, 0.00%
{c}	0.4762, 0.6508	0.0000, 0.2667	1.1270	0.1000	9.08%, 90.92%	27.27%, 72.73%
ø	0.4000, 0.6000	0.0000, 0.0000	1.0000	0.0000	19.32%, 80.68%	100.00%, 0.00%



Fig. 2. Metrical and regional targets' preservation deduction.

applications [5,7,14,17–19,29,55,58,60,67,69,79]. In particular, shadowed sets offer useful insight into the three-way decision mechanism and are thus related in some way to three-way decision theory [38,40,41,80].

Note that all quantitative models depend on underlying uncertainty measures. For RS-Theory, uncertainty measures underlie to a large extent quantitative applications, so their mining and applications have become a fundamental subject [4,6,8,21,34,43,49,51,61,76,78]. In fact, ratio $\frac{||x|| \cap X|}{||x||}$ acts as the core measure in the bottom approximate space (Approx-Space), and it corresponds to rough membership [37], probability [64,72], misclassification degree [81], and precision [73,77]. The extensively used $\frac{|[x] \cap X|}{|[x]|}$ is essentially a relative measure to exhibit information concentration. In contrast, a sort of absolute measure also exists to manifest data vividness, for example, the grade measure. Relative and absolute measures adopt different quantitative views for measurement, thus underlying quantitative applications (especially those regarding the approximation/error). Usually, both hold heterogeneity and complementarity, and thus, each relies on its essential benefit to occupy its own dominant environment. Therefore, their double-quantitative integration/fusion exhibits systematicness and superiority for gaining efficient applications. For this important topic, relative probability/precision and absolute grade were utilized for the double-quantitative information architecture and model construction [15,72,73,77]. In this paper, three-way probabilities and measures in Approx-Space are analyzed from the perspective of probability statistics and application semantics, respectively;



Fig. 4. Complete lattices' order preservation and IP-Accuracy's granulation monotonicity.

then, relative accuracy and absolute importance – two causality measures – are systematically extracted to implement doublequantitative fusion. Thus, we systematically mine a synthetic causality measure, *importance-accuracy (IP-Accuracy)*, which is verified as being benignly monotonous in its further granular integration construction.

Attribute reduction holds optimization and generalization to underlie practical applications regarding knowledge discovery and data mining. Thus, it is always an essential subject in RS-Theory [3,11,20,31,42,50,72,75]. Classical qualitative Pawlak-Reduction mainly depends on the classification-positive region (C-POS) and granulation monotonicity. However, quantitative C-POS change usually exhibits non-monotonicity and further uncertainty. Thus, quantitative model-based attribute reduction transcends basic qualitative reduction, exhibiting several reduction anomalies [32,52,65]. The difficulty of this reduction approach originates from the constructional region complexity in quantitative expansion. To address this challenging topic, DTRS-Reduction was researched in [10,12,31,65,74,75], and relevant hierarchical reduction theory was established in [74,75]. Note that uncertainty measures apply to attribute reduction, especially metrical reduct construction and heuristic algorithm development. In particular, information reduction is performed in entropy theory [31,49,76]. By virtue of the metrical merit, uncertainty measure-based attribute reduction has



Fig. 3. IP-Accuracy's GrC hierarchical structure.



Fig. 5. GrC reduct hierarchy regarding IP-Accuracy, as well as hierarchical reducts of Example 2.

substantial research potential and application prospects. In contrast, the difficulty of metrical reduction technology mainly exists in semantics-based measure mining, metrical hierarchical construction, and granulation monotonicity discovery. In this paper, the synthetic causality measure IP-Accuracy first performs gradual granular integration; then, IP-Accuracy with its inherent monotonicity is fully utilized to establish a hierarchical reduct system, including qualitative/quantitative reducts, tolerant/approximate reducts, reduct hierarchies, and heuristic algorithms.

GrC [22,68] is one of the most powerful structural methodologies and can effectively process hierarchical information by utilizing its trialistic characteristics, i.e., multiple granules, levels, and perspectives. GrC has been the focus of much study as a result of information granulation [39,41,44,54,56]; in particular, GrC was also deeply concerned with RS-Theory [9,16,21,33,45,46,53,63,73,77]. In fact, RS-Theory is a fundamental GrC pattern; thus, GrC could effectively apply to RS-Theory, including double quantification and attribute reduction mentioned above. (1) Double quantification refers to the initial dual viewpoint and further multiple granules/hierarchies, thus adhering to GrC. In this sense, this paper utilizes GrC to perform double-quantitative fusion mining of causality measures. (2) Attribute reduction mainly depends on knowledge granulation. The latter is a sort of hierarchical transformation of knowledge structures, so it leads to structural coarsening, granular merging, regional uncertainty, and informational monotonicity. Thus, knowledge granulation closely follows GrC, while GrC could thoroughly probe the essence of attribute reduction. Focusing on IP-Accuracy, the objective of this paper is to make a novel bottomtop GrC construction: a hierarchical integration-evolution from micro bottom Approx-Space to macro top attribute reduction. In particular, the challenges posed by the granulation monotonicity of IP-Accuracy and the relevant equality conditions are the emphasis of this research, and they are the foundation of the hierarchical reduct system.

According to the above, this paper mainly utilizes GrC to investigate the double-quantitative fusion of relative accuracy and absolute importance, and it provides systematic measure mining, benign integration construction, and hierarchical attribute reduction. Thus, there are four relevant parts. (1) Three-way probabilities and measures are analyzed. Then, accuracy and importance are systematically extracted to perform double-quantitative analysis and fusion, and thus IP-Accuracy is produced. (2) IP-Accuracy implements a bottom-top GrC construction. The relevant granulation monotonicity and hierarchical structure are revealed.



Fig. 6. Knowledge-structural hierarchies regarding conditional attribute subsets.

(3) IP-Accuracy attribute reduction is explored based on decision tables. A hierarchical reduct system is thereby established, including qualitative/quantitative reducts, reduct hierarchies, and heuristic algorithms. (4) Finally, a decision table example is provided for relevant illustration. This paper performs double-quantitative fusion of causality measures to systematically mine IP-Accuracy, and this measure benignly constructs a GrC platform and hierarchical reduct system. By resorting to a monotonous uncertainty measure, this study provides an integration-evolution strategy of granular construction for attribute reduction. In particular, IP-Accuracy and its reducts hold benign monotonicity and statistical systematicness; accordingly, the novel tolerant and approximate reducts quantitatively approach/expand/weaken the ideal qualitative reduction by a direct and statistical mean.

This paper is organized as follows. Section 2 is devoted to IP-Accuracy's systematic mining; in Section 3, the GrC for IP-Accuracy is constructed; Section 4 establishes attribute reduction for the measure; Section 5 provides a relevant illustration, and Section 6 concludes this paper. First, the main abbreviations in this paper are listed in Table 1.

2. Importance-accuracy's systematic mining

This section performs systematic mining to acquire a benign uncertainty measure: importance-accuracy (IP-Accuracy), which informs later sections regarding GrC construction and attribute reduction.

2.1. Three-way probabilities and measures

For in-depth mining, this subsection systematically inspects Approx-Space's metrics from the probability statistics and application semantics. Thus, three-way probabilities and measures produce; after comparative analyses, two causality measures naturally emerge. Herein, the narrow *measure* only applies to three-way measures (i.e., accuracy, importance, and generality).

Universe *U* is non-empty and finite. Equivalence relation *R* corresponds to partition $U/R = \{[x]_R : x \in U\}$, and U/R (or mere *R*) means knowledge. Thus, (U, R) constitutes Approx-Space, where $[x]_R$ is the basic granule. Furthermore, given concept $X \subseteq U$, (U, R, X) is called *the bottom system (Bottom-System)*. From the causality mechanism, granule $[x]_R$ and concept X represent the factor and result, respectively.

Lemma 2.1.1. $\forall S \in 2^U$, let $p(S) = \frac{|S|}{|U|}$, then $(U, 2^U, p)$ constitutes a probability space.

Probability space $(U, 2^U, p)$ underlies Bottom-System (U, R, X), especially from the mathematical statistics perspective. For example, conditional probability $p(X|[x]_R)$ could be naturally well-defined:

$$p(X|[x]_R) = \frac{p(X \cap [x]_R)}{p([x]_R)} = \frac{|X \cap [x]_R|/|U|}{|[x]_R|/|U|} = \frac{|X \cap [x]_R|}{|[x]_R|}.$$

Furthermore, we focus on the Bayesian formula:

$$p([x]_R|X) = \frac{p([x]_R) \cdot p(X|[x]_R)}{p(X)}.$$
(1)

There are four relevant probabilities, but $p(X) = \frac{|X|}{|U|}$ becomes a constant in Bottom-System (U, R, X). Thus, only the surplus three-way probabilities are worth analyzing.

Definition 2.1.2 (Three-Way Probabilities). In Bottom-System $(U, R, X), \forall [x]_R \in U/R$,

$$\begin{cases}
p(X|[x]_R) = \frac{|X \cap [X]_R|}{|[x]_R|}, \\
p([x]_R|X) = \frac{|[x]_R \cap X|}{|X|}, \\
p([x]_R) = \frac{|[x]_R|}{|U|},
\end{cases}$$
(2)

mean the likelihood, posterior, and prior probabilities, respectively.

Three-way probabilities exhibit different probability semantics and causality connotation. Prior $p([x]_R)$ solely measures cause uncertainty of knowledge *R*. Conditional probabilities exhibit two forms. For the decisiveness mechanism, likelihood $p(X|[x]_R)$ and posterior $p([x]_R|X)$ reflect causality in the cause-to-effect and effect-to-cause directions, respectively; thus, both directly describe the correlative relationship between the granule and concept. For the descriptive style, they exhibit relativity and absoluteness, respectively. In fact, the likelihood probability depends on information concentration to express a sort of relative possibility regarding causality; in contrast, the posterior probability depends on data directness to express a sort of absolute possibility regarding causality.

Three-way probabilities originate from the theoretical statistics framework of probability space $(U, 2^U, p)$. In fact, these metrics can also be directly described from the applicable semantics perspective.

Definition 2.1.3 (Three-Way Measures). In Bottom-System $(U, R, X), \forall [x]_R \in U/R$,

$$\begin{cases} a_X([X]_R) = \frac{|[X]_R \cap X|}{|[X]_R|}, \\ i_X([X]_R) = \frac{|[X]_R \cap X|}{|X|}, \\ g([X]_R) = \frac{|[X]_R|}{|U|}, \end{cases}$$
(3)

are called granule $[x]_R$'s accuracy, importance, and generality regarding concept X, respectively.

In essence, three-way measures correspond to three mappings from Approx-Space to the one-dimensional line [0, 1], where concept X may be the mapping parameter. From discriminative perspectives, they hold inherent semantics in Bottom-System (U, R, X).

- (1) As a granular local property, accuracy describes granular relative precision for concept *X*.
- (2) As a granular global property in concept *X*, importance reflects the granular absolute contribution for *X*.
- (3) As a granular global property in universe *U*, generality reflects a sort of granular absolute scale.

Thus, accuracy highlights the relativity and locality (which are related to fundamental causality), while the latter two highlight the absoluteness and globality. Furthermore, importance's absoluteness and globality also adhere to essential causality because they depend on the nuclear intersection information; in contrast, generally's never reflect causality because they are never concerned with the basic concept or further interaction. Therefore, only accuracy and importance directly contain causality. As a result, accuracy gains extensive use [37,72,73,77,81], and its integration/fusion with importance holds significance for causality reasoning. Moreover, three-way measures exhibit a systematic relationship:

$$i_X([x]_R) = \frac{g([x]_R) \cdot a_X([x]_R)}{|X|/|U|}.$$
(4)

which accords with the Bayesian formula (Formula (1)).

In particular, the measure in Bottom-System usually has three basic semantics functions, i.e., (1) measure the granule by the concept; (2) measure the concept by the granule; (3) set up a sort of confidence for a causality reasoning rule from the granule to

concept. The three measurement clues have high coordination, and Items (1) (3) are extensively used. In the above discussions, accuracy and importance perform their granular descriptions. For the rule confidence, accuracy $a_X([x]_R)$ and importance $i_X([x]_R)$ provide a sort of relative and absolute confidence for causality rule $[x]_R \Rightarrow X$: IF $x \in [x]_R$ THEN $x \in X$, respectively.

Thus far, three-way probabilities and measures are established from the probability statistics and application semantics, respectively. Next, they are contrastively summarized in Table 2.

Along the probability clue, DTRS produces with the likelihood probability and Bayesian decision [58,60,64]. Moreover, we have investigated three-way weighted entropy and three-way attribute reduction [76], which systematically integrate and hierarchically promote three-way probabilities. In particular, $\frac{||x|_R \cap X|}{||x|_R|}$ and $\frac{||x|_R \cap X|}{|X|}$ adopt different views to hold heterogeneity and complementarity for causality inference. Therefore, their system and fusion are worth analyzing for more powerful rough reasoning. Next, the two topics are successively discussed, mainly in terms of the measure approach.

2.2. Accuracy and importance's system

This subsection analyzes accuracy and importance, two causality measures, in a mathematical plane as well as their relationship from a systematic viewpoint.

In Bottom-System (U, R, X), $|[x]_R|$ and $|[x]_R \cap X|$ act as the direct and basic cardinality; their relevant measure variables \mathcal{V} and \mathcal{W} construct a fundamental two-dimensional plane $(\mathcal{V}, \mathcal{W})$ (Fig. 1) with the triangle range. In Fig. 1, accuracy and importance correspond to the slope and contraction, respectively. Next, let $\mathcal{A} = a_X([x]_R)$, $\mathcal{I} = i_X([x]_R)$. Herein, the factotum is mainly utilized to highlight relevant variables.

Proposition 2.2.1.

(1)
$$\mathcal{A} = \frac{\mathcal{W}}{\mathcal{V}}, \frac{\partial \mathcal{A}}{\partial \mathcal{V}} = -\frac{\mathcal{W}}{\mathcal{V}^2} \le 0, \frac{\partial \mathcal{A}}{\partial \mathcal{W}} = \frac{1}{\mathcal{V}} > 0, d\mathcal{A} = -\frac{\mathcal{W}}{\mathcal{V}^2} d\mathcal{V} + \frac{1}{\mathcal{V}} d\mathcal{W}.$$

(2) $\mathcal{I} = \frac{\mathcal{W}}{|X|}, \frac{d\mathcal{I}}{d\mathcal{W}} = \frac{1}{|X|} > 0, d\mathcal{I} = \frac{1}{|X|} d\mathcal{W}.$

For system $(\mathcal{V}, \mathcal{W})$, accuracy \mathcal{A} and importance \mathcal{I} become binary and univariate functions, respectively. Thus, their derivative and differential are provided by continuous expansion; furthermore, accuracy and importance exhibit monotonicity regarding the single variable.

Theorem 2.2.2 (Quantitative Completeness). Accuracy and importance are independent. Furthermore, their application system $(\mathcal{A}, \mathcal{I})$ is two-dimensional and so becomes complete for the quantitative Approx-Space.

Proof. \mathcal{A} and \mathcal{I} have no mutual determination/derivation relationships. In fact, $\frac{\mathcal{I}}{\mathcal{A}} = \frac{|[X]_R|}{|X|}$ is not a constant for knowledge granules. Hence, accuracy and importance are independent, and their system $(\mathcal{A}, \mathcal{I})$ is two-dimensional. As a result, $(\mathcal{A}, \mathcal{I})$ becomes complete for the quantitative Approx-Space because the latter is two-dimensional and its base is related to $(\mathcal{V}, \mathcal{W})$. \Box

Because of the independence, accuracy and importance hold heterogeneity, which could also be explained by their relativity and absoluteness. Furthermore, they have quantitative completeness and complementarity for the quantitative Approx-Space. Thus, they directly constitute a system via Cartesian product integration. Clearly, the measure system $(\mathcal{A}, \mathcal{I})$ adheres to Approx-Space's inherent quantitative essence, thus implying essential effectiveness and adequate performance.

Proposition 2.2.3. $\frac{\mathcal{I}}{\mathcal{A}} = \frac{\mathcal{V}}{|X|}$.

(1)
$$\begin{aligned} \mathcal{A} &= \frac{\tau}{\mathcal{V}} |X|, \quad \frac{\partial \mathcal{A}}{\partial \mathcal{I}} = \frac{|X|}{\mathcal{V}} > 0, \quad \frac{\partial \mathcal{A}}{\partial \mathcal{V}} = -\frac{\tau}{\mathcal{V}^2} |X| \le 0, \quad d\mathcal{A} = \frac{|X|}{\mathcal{V}} d\mathcal{I} - \frac{\tau}{\mathcal{V}^2} |X| d\mathcal{V}. \end{aligned} \\ (2) \quad \mathcal{I} &= \frac{\mathcal{A}\mathcal{V}}{|X|}, \quad \frac{\partial \tau}{\partial \mathcal{A}} = \frac{\mathcal{V}}{|X|} > 0, \quad \frac{\partial \tau}{\partial \mathcal{V}} = \frac{\mathcal{A}}{|X|} > 0, \quad d\mathcal{I} = \frac{\mathcal{V}}{|X|} d\mathcal{A} + \frac{\mathcal{A}}{|X|} d\mathcal{V}. \end{aligned}$$

In Proposition 2.2.3, accuracy and importance are systematically inspected. Although they never hold mutual determination, their basic non-linear relationship is constructed by \mathcal{V} , i.e., $\frac{\mathcal{I}}{\mathcal{A}} = \frac{\mathcal{V}}{|\mathcal{X}|}$. Thus, the accuracy change depends on importance, while the importance change depends on accuracy. The change dependency becomes decisive if \mathcal{V} is unchanged.

Theorem 2.2.4 (System Equivalence). *Measure system* (A, I) *and cardinality system* (V, W) *have a non-linear transformation and thus become equivalent.*

Proof.

$$\begin{cases} \mathcal{A} = \frac{\mathcal{W}}{\mathcal{V}}, \\ \mathcal{I} = \frac{\mathcal{W}}{|X|}. \end{cases} \Leftrightarrow \begin{cases} \mathcal{V} = \frac{\mathcal{I}}{\mathcal{A}}|X|, \\ \mathcal{W} = \mathcal{I}|X|. \end{cases}$$

The above equivalence verifies the non-linear transformation between two systems $(\mathcal{A}, \mathcal{I})$ and $(\mathcal{V}, \mathcal{W})$. \Box

There are two bearing systems for accuracy and importance. The cardinality system originates from the definition of both measures and thus approaches Approx-Space; in contrast, the measure system mainly cares for both measures' inherent systematicness, thus exhibiting direct applicability. In fact, the two systems become equivalent, although both have different emphases.

In Approx-Space, relative accuracy is extensively utilized [37,72,73,77,81]. However, its direct concentration also implies its simplicity and singleness. In view of the heterogeneity and complementarity from relativity and absoluteness, double-quantification [15,72,73,77] advocates measure diversity and superiority; thus, this suggests adding or fusing absolute information on the basis of relative information. In fact, importance provides the absolute estimation regarding causality; hence, accuracy and importance's double-quantitative fusion is worth further exploration, and this is therefore the next topic.

2.3. Importance-accuracy

According to three-way probabilities and measures (Section 2.1), accuracy and importance become two causality measures. In fact, they correspond to relativity/concentration and absoluteness/directness, respectively. Based on a system analysis (Section 2.2), their heterogeneity and complementarity are further verified. Therefore, their benign fusion is valuable for causality inference. In this subsection, both are fused into a benign uncertainty measure, IP-Accuracy, by adopting the weighted product technology. Herein, IP-Accuracy is mainly described by its rule confidence role.

Definition 2.3.1 (IP-Accuracy). In Bottom-System (U, R, X), $\forall [x]_R \in U/R$,

$$\mathcal{IA} = ia_X([x]_R) = i_X([x]_R) \times a_X([x]_R)$$
(5)

is called granule $[x]_R$'s *importance-accuracy* (*IP-Accuracy*) regarding concept *X*.

Proposition 2.3.2.

$$\mathcal{IA} = \frac{\mathcal{W}^2}{|X|\mathcal{V}} = \frac{|[x]_R \cap X|^2}{|X||[x]_R|}.$$
(6)

Proposition 2.3.3 (Weight Conservation).

$$\sum_{[x]_R \in U/R} i_X([x]_R) = 1.$$
(7)

IP-Accuracy, which corresponds to a mapping from Approx-Space to [0, 1], is formally constructed by the product fusion of importance and accuracy, thus exhibiting cardinality essence (Formula (6)). In the sense of weight conservation (Formula (7)), the algebraic product could be semantically explained as a sort of weighted product. Thus, IP-Accuracy's fusion scientificity and measurement semantics could be explained as follows.

- (1) First, accuracy cannot completely represent the causality between the granule and concept. For example, for two granules with the same accuracy, the granule with greater importance necessarily plays a more vital function in the causal reasoning. Then, importance could also represent the causality by virtue of its essence. In view of metrical conservation, each granule occupies only the part importance shares regarding sum 1; thus, importance could be viewed as a sort of objective weight coefficient to be appended to accuracy. The weighted pattern effectively underlies the product fusion of importance and accuracy; as a result, the corresponding IP-Accuracy more synthetically represents the causality and is thus novel and scientific.
- (2) From the rule semantics, accuracy and importance give the causality rule only relative and absolute confidences, respectively. Furthermore, the fusional IP-Accuracy could provide a sort of synthetic confidence by integrating both the relativeness and absoluteness. For example, $a_X([x]_R) = 0.5$ endows reasoning rule $[x]_R \Rightarrow X$ with relative confidence 0.5; in contrast, $i_X([x]_R) = 0.2$ endows the reasoning rule with absolute confidence 0.2, which means that the granule $[x]_R$ holds a 20% contribution for determining *X* (regardless of $[x]_R$'s scale); furthermore, $ia_X([x]_R) = 0.2 \times 0.5 = 0.1$ could more stably assess the reliability of the rule. Therefore, IP-Accuracy holds the basic rule semantics: the comprehensive reliability for the decision rule. Moreover, by virtue of the importance weight, IP-Accuracy could represent the fusional measure semantics: the granular effective precision for concept *X*.

From the double-quantitative perspective, IP-Accuracy effectively fuses two fundamental causality measures, accuracy and importance, so it has measurement merit as a result of combining metrical absoluteness and relativeness. Moreover, the superiority of IP-Accuracy could be further verified by its granulation monotonicity, which is provided in Section 3. Thus, IP-Accuracy underlies causality-based rough reasoning, so it could be utilized to comprehensively describe reasoning rules.

In a word, IP-Accuracy becomes a systematic synthetic causality measure to describe a rule's reliability and granular availability. From the mathematical viewpoint, IP-Accuracy's causality superiority mainly benefits from its non-linear promotion in system $(\mathcal{V}, \mathcal{W})$. According to Formula (3) and Fig. 1, accuracy and importance exhibit only the pure linear cardinality form. In contrast, according to Formula (6), IP-Accuracy uses the non-linear cardinality style. Thus, this non-linear improvement implies more information connotations than IP-Accuracy contains. According to Formulas (3) and (6), the relevant causality strengthening mechanism is analyzed as follows, and it radically underlies the later benign granulation monotonicity (Section 3).

- (1) Except for $|[x]_R \cap X|$, IP-Accuracy synthetically considers both $|[x]_R|$ and |X|, rather than only one. Thus, it becomes more comprehensive than accuracy and importance, especially when considering causality.
- (2) IP-Accuracy strengthens the nuclear interaction information, i.e., $|[x]_R \cap X|$. In fact, it uses the two-order form $|[x]_R \cap X|^2$, which carries more effective information than the one-order $|[x]_R \cap X|$. Thus, IP-Accuracy becomes more powerful than accuracy and importance in terms of causality.

Finally, IP-Accuracy's derivative and differential are provided by continuous expansion.

Proposition 2.3.4.

 $\begin{array}{ll} (1) & \frac{\partial(\mathcal{I}\mathcal{A})}{\partial\mathcal{I}} = \mathcal{A} \geq \mathbf{0}, \ \frac{\partial(\mathcal{I}\mathcal{A})}{\partial\mathcal{A}} = \mathcal{I} \geq \mathbf{0}, \ d(\mathcal{I}\mathcal{A}) = \mathcal{A}d\mathcal{I} + \mathcal{I}d\mathcal{A}. \\ (2) & \frac{\partial(\mathcal{I}\mathcal{A})}{\partial\mathcal{V}} = -\frac{\mathcal{W}^2}{|X|\mathcal{V}^2} \leq \mathbf{0}, \ \ \frac{\partial(\mathcal{I}\mathcal{A})}{\partial\mathcal{W}} = \frac{2\mathcal{W}}{|X|\mathcal{V}} \geq \mathbf{0}, \ \ d(\mathcal{I}\mathcal{A}) = -\frac{\mathcal{W}^2}{|X|\mathcal{V}^2}d\mathcal{V} + \frac{2\mathcal{W}}{|X|\mathcal{V}}d\mathcal{W}. \end{array}$

Corollary 2.3.5.

$$\begin{aligned} \frac{\partial (\mathcal{I}\mathcal{A})}{\partial \mathcal{V}} &= -\frac{1}{|X|}\mathcal{A}^2 \leq 0, \ \frac{\partial (\mathcal{I}\mathcal{A})}{\partial \mathcal{W}} = \frac{2}{|X|}\mathcal{A} \geq 0\\ d(\mathcal{I}\mathcal{A}) &= -\frac{1}{|X|}\mathcal{A}^2 d\mathcal{V} + \frac{2}{|X|}\mathcal{A} d\mathcal{W}. \end{aligned}$$

According to Formulas (5) and (6), Proposition 2.3.4 provides mathematical analysis results in systems (A, T) and (V, W). Furthermore, Corollary 2.3.5 expresses IP-Accuracy's derivative and differential from the accuracy perspective; in particular, this result further reflects that IP-Accuracy contains more basic information than accuracy.

3. Importance-accuracy's granular integration construction

In Section 2, IP-Accuracy - a synthetic causality measure - is systematically mined in Approx-Space by the double-quantitative fusion of accuracy and importance. In micro Bottom-System (U, R, X), IP-Accuracy mainly depends on the concept to describe granules, so granular concept IP-Accuracy (Gr-Con IP-Accuracy) is the full name in this case. Attribute reduction usually concerns a macro top with knowledge and concept systems. To explore attribute reduction, Approx-Space needs promoting to comprehensive hierarchical structures. Thus, this section progressively performs IP-Accuracy's granular integration construction in three granular levels (i.e., the micro bottom, meso middle, and macro top). Concretely, Gr-Con IP-Accuracy will be integrated into knowledge's concept IP-Accuracy and further into knowledge's classification IP-Accuracy; as a result, the relevant GrC hierarchical structure is finally established. Herein, IP-Accuracy has three levels, whose relevant abbreviations are listed in Table 1, and it is used directly without confusions. Note that granulation monotonicity acts as a crucial criterion to evaluate an uncertainty measure in RS-Theory, especially in attribute reduction [31,49,51,76]. Thus, IP-Accuracy's granulation monotonicity and relevant equality conditions are the next research emphasis.

3.1. Granulation preliminaries

Herein, necessary granulation preliminaries are first provided. *Decision table (D-Table)* $(U, C \cup D)$ serves as a basic framework; the relevant factotum also highlights the attribute subset with knowledge. Thus, $C_* \subseteq C$ correspondingly replaces the previous knowledge *R*; moreover, let $k \ge 2$. Suppose

$$\begin{cases} U/C = \{ [x]_{C}^{i} : i = 1, ..., n \}; \\ U/D = \{ X_{j} : j = 1, ..., m \}. \end{cases}$$

 $C_2 \subseteq C_1 \subseteq C$ refers to the knowledge deduction, denoted by $C_1 \Rightarrow C_2$; \Rightarrow corresponds to a knowledge-granular coarsening process and a mathematical partial order. If $C_1 \Rightarrow C_2$, then C_2 depends on and is generally rougher than C_1 (while knowledge C_1 deduces and is generally finer than knowledge C_2); thus, $C_1 \Rightarrow C_2$ is called *knowledge coarsening (Kn-Coarsening)*.

Proposition 3.1.1. In Kn-Coarsening $C_1 \Rightarrow C_2$,

$$\begin{aligned} &\forall [x]_{\mathcal{C}_1} \in U/\mathcal{C}_1, \exists [x]_{\mathcal{C}_2} \in U/\mathcal{C}_2, \\ &\text{s.t., } [x]_{\mathcal{C}_1} \subseteq [x]_{\mathcal{C}_2}. \\ &\forall [x]_{\mathcal{C}_2} \in U/\mathcal{C}_2, \exists k \in \mathbf{N}, \text{ s.t., } [x]_{\mathcal{C}_2} = \bigcup_{t=1}^k [x]_{\mathcal{C}_1}^t. \end{aligned}$$

Proposition 3.1.2. *Kn*-Coarsening $C_1 \Rightarrow C_2$ usually includes two types of granular transformation:

$$\begin{cases} Granule merging (Gr-Merging) : [x]_{\mathcal{C}_{1}}^{1} \cup \cdots \cup [x]_{\mathcal{C}_{1}}^{k} \xrightarrow{=} [x]_{\mathcal{C}_{2}}^{1}, \\ Granule preservation (Gr-Preservation) : [x]_{\mathcal{C}_{1}}^{1} \xrightarrow{=} [x]_{\mathcal{C}_{2}}^{1}. \end{cases}$$

$$(8)$$

From the granule viewpoint, *Kn-Coarsening* becomes a sort of rough hierarchical transformation between knowledge structures. Thus, it involves two types of granular actions: *Gr-Merging* and *Gr-Preservation*. Herein, the granular superscript represents the granular ordinal number, e.g., t in $[x]_{C_1}^t$ and k in $[x]_{C_1}^k$. In fact, *Gr-Merging* becomes a radical feature of *Kn-Coarsening*, and representative $[x]_{C_1}^1 \cup \cdots \cup [x]_{C_1}^k \xrightarrow{=} [x]_{C_2}^1$ could be directly utilized for complete granulation monotonicity verification.

Definition 3.1.3 ([35,36]). Positive, negative, boundary regions (POS, NEG, BND) of *X* regarding knowledge C_* are defined by:

$$\begin{cases} \text{POS}_{C_*}(X) = \{x : [x]_{\mathcal{C}_*} \subseteq X\}, \\ \text{NEG}_{C_*}(X) = \{x : [x]_{\mathcal{C}_*} \subseteq \neg X\}, \\ \text{BND}_{C_*}(X) = \{x : [x]_{\mathcal{C}_*} \cap X, \neg X \neq \varnothing\}. \end{cases}$$
(9)

Furthermore, the *classification-positive/boundary region* (*C-POS/C-BND*) of *D* regarding knowledge C_* is defined by:

$$\begin{cases} \operatorname{POS}_{\mathcal{C}_*}(\mathcal{D}) = \bigcup_{X \in U/\mathcal{D}} \operatorname{POS}_{\mathcal{C}_*}(X), \\ \operatorname{BND}_{\mathcal{C}_*}(\mathcal{D}) = U - \operatorname{POS}_{\mathcal{C}_*}(\mathcal{D}). \end{cases}$$
(10)

Thus, the dependency degree is given by:

$$\gamma_{\mathcal{C}_*}(\mathcal{D}) = \frac{|\text{POS}_{\mathcal{C}_*}(\mathcal{D})|}{|U|},\tag{11}$$

Proposition 3.1.4. In Kn-Coarsening $C_1 \Rightarrow C_2$,

$$\begin{cases} \operatorname{POS}_{\mathcal{C}_{2}}(X) \subseteq \operatorname{POS}_{\mathcal{C}_{1}}(X), \\ \operatorname{POS}_{\mathcal{C}_{2}}(\mathcal{D}) \subseteq \operatorname{POS}_{\mathcal{C}_{1}}(\mathcal{D}), \\ \gamma_{\mathcal{C}_{2}}(\mathcal{D}) \leq \gamma_{\mathcal{C}_{1}}(\mathcal{D}). \end{cases}$$
(12)

In the qualitative Pawlak-Model, C-POS becomes the union of classification concepts' POS, while the dependency degree becomes the classical measure to evaluate the classification quality. In particular, C-POS, POS, and measure γ all have granulation monotonicity (Proposition 3.1.4).

3.2. Knowledge's concept importance-accuracy

Aiming at meso *Middle-System* (U, C, X), this subsection utilizes Gr-Con IP-Accuracy in micro Bottom-System (U, C_*, X) to integratedly construct knowledge's concept IP-Accuracy, and the corresponding granulation monotonicity is emphatically verified.

Definition 3.2.1 (Kn-Con IP-Accuracy). In Middle-System (U, C, X),

$$ia_{X}(C_{*}) = \sum_{[x]_{C_{*}} \in U/C_{*}} ia_{X}([x]_{C_{*}})$$
(13)

is called knowledge C_* 's IP-Accuracy regarding concept X, i.e., knowledge's concept IP-Accuracy (Kn-Con IP-Accuracy).

By using the arithmetic sum, Kn-Con IP-Accuracy is naturally established by integrating Gr-Con IP-Accuracy because knowledge consists of knowledge granules. In fact, Kn-Con IP-Accuracy promotes Gr-Con IP-Accuracy from the GrC viewpoint. Clearly, Kn-Con IP-Accuracy $ia_X(\mathcal{C}_*)$ holds three types of harmonious semantics functions.

- (1) Measure reasoning $C_* \Rightarrow X$ in a confidence sense.
- (2) Assess knowledge C_* by concept X, a systematic parameter.
- (3) Systematically estimate concept X by knowledge C_{*}, a sort of granule system.

Next, we make efforts to inspect IP-Accuracy's granulation monotonicity in (U, C, X). For this purpose, the relevant mathematical mechanism regarding Gr-Merging is first uncovered for IP-Accuracy.

Lemma 3.2.2.

$$\sum_{t=1}^{k} \frac{a_t^2}{b_t} \ge \frac{\left(\sum_{t=1}^{k} a_t\right)^2}{\sum_{t=1}^{k} b_t}.$$
(14)

The above inequality reaches the equal sign, if and only if (iff),

$$\frac{a_1}{b_1} = \dots = \frac{a_k}{b_k}, \text{ i.e., } \frac{a_1}{b_1} = \dots = \frac{a_k}{b_k} = \frac{\sum_{t=1}^k a_t}{\sum_{t=1}^k b_t}.$$
 (15)

Theorem 3.2.3 (Gr-Merging Monotonicity). *For Gr-Merging:* $[x]_{\mathcal{C}_1}^1 \cup \cdots \cup [x]_{\mathcal{C}_1}^k \xrightarrow{=} [x]_{\mathcal{C}_2}^1$,

$$ia_X([x]_{\mathcal{C}_2}^1) \le ia_X([x]_{\mathcal{C}_1}^1) + \dots + ia_X([x]_{\mathcal{C}_1}^k).$$
(16)

The above equality holds, iff Gr-Merging concerns the equal granular accuracy. This equality requirement could be further divided into three conditions:

$$\begin{cases} a_X([x]_{\mathcal{C}_1}^1) = \dots = a_X([x]_{\mathcal{C}_1}^k) = a_X([x]_{\mathcal{C}_2}^1) = 1; \\ a_X([x]_{\mathcal{C}_1}^1) = \dots = a_X([x]_{\mathcal{C}_1}^k) = a_X([x]_{\mathcal{C}_2}^1) = 0; \\ a_X([x]_{\mathcal{C}_1}^1) = \dots = a_X([x]_{\mathcal{C}_1}^k) = a_X([x]_{\mathcal{C}_2}^1) \in (0, 1). \end{cases}$$
(17)

Proof. According to Formulas (6) and (14),

$$\begin{split} &ia_{X}([x]_{\mathcal{C}_{1}}^{1}) + \dots + ia_{X}([x]_{\mathcal{C}_{1}}^{k}) = \frac{|[x]_{\mathcal{C}_{1}}^{1} \cap X|^{2}}{|[x]_{\mathcal{C}_{1}}^{1}| \times |X|} + \dots + \frac{|[x]_{\mathcal{C}_{1}}^{k} \cap X|^{2}}{|[x]_{\mathcal{C}_{1}}^{k}| \times |X|} \\ &\geq \frac{(|[x]_{\mathcal{C}_{1}}^{1} \cap X| + \dots + |[x]_{\mathcal{C}_{1}}^{k} \cap X|)^{2}}{|[x]_{\mathcal{C}_{1}}^{1}| \times |X| + \dots + |[x]_{\mathcal{C}_{1}}^{k} \cap X|^{2}} \\ &= \frac{|([x]_{\mathcal{C}_{1}}^{1} \cap X) \cup \dots \cup ([x]_{\mathcal{C}_{1}}^{k} \cap X)|^{2}}{|[x]_{\mathcal{C}_{1}}^{1} \cup \dots \cup [x]_{\mathcal{C}_{1}}^{k}| \times |X|} = \frac{|([x]_{\mathcal{C}_{1}}^{1} \cup \dots \cup [x]_{\mathcal{C}_{1}}^{k}) \cap X|^{2}}{|[x]_{\mathcal{C}_{2}}^{1} \cap X|^{2}} \\ &= \frac{|[x]_{\mathcal{C}_{2}}^{1} \cap X|^{2}}{|[x]_{\mathcal{C}_{2}}^{1} \cap X|^{2}} = ia_{X}([x]_{\mathcal{C}_{2}}^{1}). \end{split}$$

Thus, Formula (16) holds for G-Merging. Moreover, the equality condition (Formula (15)) refers to the equal granular accuracy, and it is further divided into G-Merging's three special cases regarding 0 and 1 (Formula (17)).

Lemma 3.2.2, whose proof is given in Appendix A, establishes an important mathematical basis. Thus, IP-Accuracy gains Gr-Merging monotonicity. Concretely, the constructional granular IP-Accuracy never increases for the merged granules' IP-Accuracy sum (Formula (16)). Next, the equality conditions (Formula (17)) are further analyzed, mainly from the set-regional perspective.

Proposition 3.2.4. *Given Gr-Merging:* $[x]_{\mathcal{C}_1}^1 \cup \cdots \cup [x]_{\mathcal{C}_1}^k \xrightarrow{=} [x]_{\mathcal{C}_2}^1$.

(1)
$$a_X([x]^1_{\mathcal{C}_1}) = \dots = a_X([x]^k_{\mathcal{C}_1}) = a_X([x]^1_{\mathcal{C}_2}) = 1,$$

 $iff[x]^1_{\mathcal{C}_1}, \dots, [x]^k_{\mathcal{C}_1}, [x]^1_{\mathcal{C}_2} \subseteq X,$
 $iff[x]^1_{\mathcal{C}_1}, \dots, [x]^k_{\mathcal{C}_1}, [x]^1_{\mathcal{C}_2} \subseteq \text{POS}_{\mathcal{C}_1}(X), \text{POS}_{\mathcal{C}_2}(X).$

(2)
$$a_X([x]_{\mathcal{C}_1}^1) = \dots = a_X([x]_{\mathcal{C}_1}^1) = a_X([x]_{\mathcal{C}_2}^1) = 0,$$

 $iff[x]_{\mathcal{C}_1}^1, \dots, [x]_{\mathcal{C}_1}^k, [x]_{\mathcal{C}_2}^1 \subseteq \neg X,$
 $iff[x]_{\mathcal{C}_1}^1, \dots, [x]_{\mathcal{C}_1}^k, [x]_{\mathcal{C}_2}^1 \subseteq \text{NEG}_{\mathcal{C}_1}(X), \text{NEG}_{\mathcal{C}_2}(X).$
(3) $lf a_X([x]_{\mathcal{C}_1}^1) = \dots = a_X([x]_{\mathcal{C}_1}^k) = a_X([x]_{\mathcal{C}_1}^1) \in (0, 1]$

According to Formula (15), the monotonicity equality could be equivalently achieved by three accuracy conditions (Formula (17)); the latter are further transformed into set-regional manifestations (Proposition 3.2.4). Concretely, Items (1) (2) mean that Gr-Merging is equivalently inside X/POS and \neg X/NEG, respectively; Item (3) shows that the third equality condition is only a strong case inside BND. These results underlie further region preservation deduction.

Theorem 3.2.5 (Gr-Preservation Invariance). For Gr-Preservation $[x]_{\mathcal{C}_1}^1 \xrightarrow{=} [x]_{\mathcal{C}_2}^1$

$$ia_X([x]^1_{\mathcal{C}_2}) = ia_X([x]^1_{\mathcal{C}_1}).$$
(18)

Theorem 3.2.6 (Monotonicity I). In Kn-Coarsening $C_1 \Rightarrow C_2$,

$$ia_X(\mathcal{C}_2) \le ia_X(\mathcal{C}_1). \tag{19}$$

 $ia_X(\mathcal{C}_2) = ia_X(\mathcal{C}_1)$, iff each Gr-Merging exists in one of the following three cases:

- (1) in POS's interior regarding both C_1 and C_2 ;
- (2) in NEG's interior regarding both C_1 and C_2 ;
- (3) in BND's interior regarding both C_1 and C_2 , but on the premise of equal granular accuracy.

Proof. Kn-Coarsening usually consists of Gr-Merging and Gr-Preservation (Proposition 3.1.2). IP-Accuracy exhibits Gr-Merging monotonicity (Theorem 3.2.3) and Gr-Preservation invariance (Theorem 3.2.5). Thus, IP-Accuracy naturally holds Kn-Coarsening monotonicity, i.e., granulation monotonicity. For the monotonicity equality condition, all existing Gr-Merging satisfy condition (17) and thus equivalently exists in the above three cases. Thus, the equality condition achieves a natural evolution from internal Gr-Merging to integral Kn-Coarsening.

Corollary 3.2.7. IP-Accuracy preservation unidirectionally deduces setregion preservation in Kn-Coarsening. In other words, for $C_1 \Rightarrow C_2$, if $ia_X(\mathcal{C}_2) = ia_X(\mathcal{C}_1)$, then

$$POS_{\mathcal{C}_2}(X) = POS_{\mathcal{C}_1}(X),$$

$$NEG_{\mathcal{C}_2}(X) = NEG_{\mathcal{C}_1}(X),$$

$$BND_{\mathcal{C}_2}(X) = BND_{\mathcal{C}_1}(X);$$
(20)

however, the opposite usually does not hold.

IP-Accuracy holds granulation monotonicity (Theorem 3.2.6); i.e., it never increases in Kn-Coarsening. In particular, monotonicity acquires its equality condition from the regional perspective; concretely, $ia_X(\mathcal{C}_2) = ia_X(\mathcal{C}_1)$ means the arbitrary and conditional Gr-Merging in POS/NEG and BND, respectively. Thus, Corollary 3.2.7 clearly provides a basic connection between the measure and region. In fact, the perfect granulation monotonicity and equality condition regarding Kn-Coarsening promote the corresponding results regarding Gr-Merging (Theorem 3.2.3); thus, they hold a strong mechanism and underlie later attribute reduction.

In practice, Kn-Coarsening's integration monotonicity could be similarly inspected for other measures. However, accuracy and importance never have a perfect Gr-Merging mechanism, benign Kn-Coarsening monotonicity, or clear monotonicity equality condition. Thus, IP-Accuracy's granulation monotonicity (Theorem 3.2.6) fully reflects IP-Accuracy's superiority, which mainly benefits from the benign weight fusion and non-linear cardinality promotion (Section 2.3).

In a regional way, POS and, further, C-POS underlie attribute reduction. To approach POS and C-POS, we finally mine Kn-Con IP-Accuracy's fundamental subpart, which also holds granulation monotonicity.

Definition 3.2.8 (Kn-Con Sub-IP-Accuracy). In Middle-System (U, C, X) where $C_* \subseteq C$,

$$\begin{cases} lia_{X}(\mathcal{C}_{*}) = \sum_{[x]_{\mathcal{C}_{*}} \subseteq X} ia_{X}([x]_{\mathcal{C}_{*}}); \\ Eia_{X}(\mathcal{C}_{*}) = \sum_{[x]_{\mathcal{C}_{*}} \subseteq \neg X} ia_{X}([x]_{\mathcal{C}_{*}}); \\ Bia_{X}(\mathcal{C}_{*}) = \sum_{[x]_{\mathcal{C}_{*}} \cap X, \neg X \neq \varnothing} ia_{X}([x]_{\mathcal{C}_{*}}). \end{cases}$$
(21)

are called knowledge C_* 's internal, external, and boundary IP-Accuracy regarding concept X, respectively. Herein, the three types of sub-IP-Accuracy are simply denoted by (Kn-Con) IIP-Accuracy, EIP-Accuracy, BIP-Accuracy, respectively.

Proposition 3.2.9.

- (1) $Iia_X(\mathcal{C}_*) = \frac{|POS_{\mathcal{C}_*}(X)|}{|X|}$. (2) $Eia_X(\mathcal{C}_*) = 0$. (3) $Bia_X(\mathcal{C}_*) = \sum_{|X|_{\mathcal{C}_*} \subseteq BND_{\mathcal{C}_*}(X)} ia_X([x]_{\mathcal{C}_*})$.
- (4) $ia_X(\mathcal{C}_*) = lia_X(\mathcal{C}_*) + Bia_X(\mathcal{C}_*).$

Herein, Kn-Con IP-Accuracy $ia_X(\mathcal{C}_*)$ is further divided into IIP-Accuracy $Iia_X(\mathcal{C}_*)$, EIP-Accuracy $Eia_X(\mathcal{C}_*)$, and BIP-Accuracy $Bia_X(\mathcal{C}_*)$. According to Proposition 3.2.9 (whose proof is given in Appendix B), EIP-Accuracy is equal to 0, IIP-Accuracy becomes $\frac{|POS_{\mathcal{C}_*}(X)|}{|X|}$ to act as a main part of IP-Accuracy, and BIP-Accuracy becomes the sum of all Gr-Con IP-Accuracy regarding concept X's BND.

Theorem 3.2.10 (Monotonicity II). In Kn-Coarsening $C_1 \Rightarrow C_2$,

$$lia_{\mathcal{C}_2}(X) \le lia_{\mathcal{C}_1}(X); \tag{22}$$

 $lia_{\mathcal{C}_2}(X) = lia_{\mathcal{C}_1}(X), \text{ iff } \operatorname{POS}_{\mathcal{C}_2}(X) = \operatorname{POS}_{\mathcal{C}_1}(X).$ (23)

Proof. According to Propositions 3.1.4 and 3.2.9, $Iia_{\mathcal{C}_2}(X) = \frac{|POS_{\mathcal{C}_2}(X)|}{|X|} \leq \frac{|POS_{\mathcal{C}_1}(X)|}{|X|} = Iia_{\mathcal{C}_1}(X)$; moreover, the monotonicity equality corresponds to condition $POS_{\mathcal{C}_2}(X) = POS_{\mathcal{C}_1}(X)$. \Box

Kn-Con IIP-Accuracy represents IP-Accuracy's internal sum regarding POS and also has granulation monotonicity. Thus, IIP-Accuracy more resembles the classical regional pattern when compared to IP-Accuracy.

3.3. Knowledge's classification importance-accuracy

Aiming at macro *Top-System* (U, C, D), this subsection further utilizes Kn-Con IP-Accuracy in meso Middle-System (U, C, X) to integratedly construct classification's IP-Accuracy, and corresponding granulation monotonicity is naturally organized. In particular, preservation deduction of relevant targets becomes a new focus to underlie the later reduct hierarchies.

Definition 3.3.1 (Kn-Cl IP-Accuracy). In Top-System (U, C, D),

$$ia_{\mathcal{D}}(\mathcal{C}_*) = \sum_{j=1}^m ia_{X_j}(\mathcal{C}_*)$$
(24)

is called knowledge C_* 's IP-Accuracy regarding classification \mathcal{D} , i.e., *knowledge's classification IP-Accuracy (Kn-Cl IP-Accuracy)*.

By using the arithmetic sum, Kn-Cl IP-Accuracy is further established by integrating Kn-Con IP-Accuracy because classification consists of concepts. In fact, Kn-Cl IP-Accuracy further promotes Kn-Con IP-Accuracy from the GrC viewpoint. Similarly, Kn-Cl IP-Accuracy $ia_{\mathcal{D}}(\mathcal{C}_*)$ also holds three harmonious semantics functions.

- (1) Measure reasoning $C_* \Rightarrow D$ in a confidence sense.
- (2) Measure knowledge C_* by classification \mathcal{D} : a systematic parameter by integrating concepts X_i .
- (3) Systematically measure classification \mathcal{D} by knowledge \mathcal{C}_* : a sort of granule system.

Theorem 3.3.2 (Monotonicity III). *In Kn-Coarsening* $C_1 \Rightarrow C_2$,

$$ia_{\mathcal{D}}(\mathcal{C}_2) \le ia_{\mathcal{D}}(\mathcal{C}_1). \tag{25}$$

 $ia_{\mathcal{D}}(\mathcal{C}_2) = ia_{\mathcal{D}}(\mathcal{C}_1)$, iff $ia_{\mathcal{C}_2}(X_j) = ia_{\mathcal{C}_1}(X_j)$ for $\forall j \in \{1, ..., m\}$, iff each *Gr*-Merging exists in one of the two conditions (where \mathcal{C}_1 and \mathcal{C}_2 are equivalently concerned):

- (1) in POS's interior of only one concept;
- (2) in BND's interior of all concepts, but on the premise of equal granular accuracy for each concept.

In particular, the two conditions imply equal accuracy G-Merging in C-POS and C-BND and are thus called classification-quantitative preservation.

Proof. First, Kn-Cl IP-Accuracy becomes Kn-Con IP-Accuracy's algebraic sum integration (regarding concepts). According to Kn-Con IP-Accuracy's monotonicity (Theorem 3.2.6), $ia_D(\mathcal{C}_2) \leq ia_D(\mathcal{C}_1)$ becomes natural. Thus, we only need to further refine the equality condition "all Gr-Merging satisfies equal accuracy regarding all concepts X_i ".

For this purpose, we next adopt a two-stage discussion strategy to identify Gr-Merging's regional range. Concretely, disjoint C-POS and C-BND are first concerned, and concepts are then considered. Herein, the regional notions mainly refer to knowledge C_1 , but they are equivalent to those regarding knowledge C_2 .

alent to those regarding knowledge C_2 . Consider Gr-Merging $[x]_{C_1}^1 \cup \cdots \cup [x]_{C_1}^k \xrightarrow{=} [x]_{C_2}^1$.

- If it is distributed in both C-POS and C-BND, then exist X_j, s.t., it has at least two granules to contain different accuracy values regarding X_j. Hence, Gr-Merging is only in either C-POS or C-BND.
- (2) Suppose it is in C-POS, the union of multiple POS. If it is distributed in different POS, we could obtain the contradiction (from granular unequal accuracy), which is similar to above. Hence, it is only in the POS of a sole concept, and thus G-Merging equal accuracy is naturally realized for all concepts.
- (3) Suppose it is in C-BND, the union of multiple BND. By virtue of the similar contradiction proof, its granules must exist in the intersection of all BND and must have the same accuracy for each concept. □

Kn-Cl IP-Accuracy holds granulation monotonicity (Theorem 3.3.2), which integratedly promotes Kn-Con IP-Accuracy's granulation monotonicity (Theorem 3.2.6). Herein, the monotonicity equality's regional mining becomes complex. By analyzing IP-Accuracy's connotation, two specific regional equality conditions regarding classification-quantitative preservation are finally established to define Gr-Merging's range. In fact, the conversion between the metrical and regional descriptions performs the GrC integration construction from the set-regional to classification-regional levels. As a result, Theorem 3.3.2's two items have summarized and promoted Theorem 3.2.6's three items, and this GrC integration can be illustrated by Theorem 3.3.2's proof.

Corollary 3.3.3. *IP-Accuracy preservation unidirectionally deduces C-*POS preservation in Kn-Coarsening. In other words, for $C_1 \Rightarrow C_2$, if $ia_{\mathcal{D}}(C_2) = ia_{\mathcal{D}}(C_1)$, then $\text{POS}_{\mathcal{C}_2}(\mathcal{D}) = \text{POS}_{\mathcal{C}_1}(\mathcal{D})$; however, the opposite usually does not hold.

Proof. IP-Accuracy preservation implies the arbitrary and conditional Gr-Merging in C-POS and C-BND, respectively (Theorem 3.3.2). In contrast, C-POS preservation implies the arbitrary Gr-Merging in both C-POS and C-BND. Thus, this corollary is proved. \Box

IP-Accuracy preservation is stronger than C-POS preservation, and this conclusion is inferred by the regional equality conditions (Theorem 3.3.2).

Corollary 3.3.4. In Kn-Coarsening, IP-Accuracy preservation unidirectionally deduces set-region preservation regarding classification's all concepts. **Proof.** Kn-Cl IP-Accuracy becomes Kn-Con IP-Accuracy's sum (regarding all concepts), and both hierarchical measures have granulation monotonicity. Thus, by the contradiction proof, the former's preservation is equivalent to the latter's preservation regarding all concepts. According to Corollary 3.2.7, for concept X_j , Kn-Con IP-Accuracy preservation unidirectionally deduces set-region preservation. Hence, Kn-Cl IP-Accuracy preservation unidirectionally deduces set-region preservation regarding classification's all concepts. \Box

Corollary 3.3.4 is derived from the monotonicity integration, as well as Kn-Con IP-Accuracy preservation's strongness (for set-region preservation). Now, there are three concerned preservation targets in the above two corollaries. In Corollary 3.3.3, IP-Accuracy preservation unidirectionally deduces C-POS preservation; then, the former unidirectionally deduces set-region preservation in Corollary 3.3.4; moreover, set-region preservation naturally unidirectionally induces C-POS preservation.

In Section 3.2, Kn-Con IP-Accuracy consists of three parts, i.e., Kn-Con IIP-Accuracy, EIP-Accuracy, and BIP-Accuracy. Note that Kn-Con IP-Accuracy is promoted to Kn-Cl IP-Accuracy. Thus, Kn-Cl IP-Accuracy naturally consists of three corresponding promotional parts. Next, the relevant definition, essence, and monotonicity are straightforwardly provided by virtue of the relevant integration and promotion.

Definition 3.3.5 (Kn-Cl Sub-IP-Accuracy). In (U, C, D) where $C_* \subseteq C$,

$$\begin{cases} lia_{\mathcal{D}}(\mathcal{C}_{*}) = \sum_{j=1}^{m} lia_{X_{j}}(\mathcal{C}_{*});\\ Eia_{\mathcal{D}}(\mathcal{C}_{*}) = \sum_{j=1}^{m} Eia_{X_{j}}(\mathcal{C}_{*});\\ Bia_{\mathcal{D}}(\mathcal{C}_{*}) = \sum_{j=1}^{m} Bia_{X_{j}}(\mathcal{C}_{*}). \end{cases}$$

$$(26)$$

are called knowledge C_* 's internal, external, boundary IP-Accuracy regarding classification D, respectively; in particular, the three types of Sub-IP-Accuracy are simply denoted as *Kn-Cl IIP-Accuracy, EIP-Accuracy, BIP-Accuracy,* respectively.

Proposition 3.3.6.

(1)
$$Iia_{\mathcal{D}}(\mathcal{C}_*) = \sum_{i=1}^m \frac{|POS_{\mathcal{C}_*}(X_j)|}{|X_i|} \le \sum_{i=1}^m |POS_{\mathcal{C}_*}(X_j)| = |POS_{\mathcal{C}_*}(\mathcal{D})|.$$

(2)
$$Eia_{\mathcal{D}}(\mathcal{C}_*) = 0.$$

(3)
$$Bia_{\mathcal{D}}(\mathcal{C}_*) = \sum_{i=1}^m \sum_{[x]_{\mathcal{C}_*} \subset BND_{\mathcal{C}_*}(X_i)} ia_{X_i}([x]_{\mathcal{C}_*}).$$

(4) $ia_{\mathcal{D}}(\mathcal{C}_*) = Iia_{\mathcal{D}}(\mathcal{C}_*) + Bia_{\mathcal{D}}(\mathcal{C}_*).$

Kn-Cl IP-Accuracy $ia_{\mathcal{D}}(\mathcal{C}_*)$ consists of IIP-Accuracy $Iia_{\mathcal{D}}(\mathcal{C}_*)$, EIP-Accuracy $Eia_{\mathcal{D}}(\mathcal{C}_*)$, *BIP-Accuracy Bia*_{$\mathcal{D}(\mathcal{C}_*)$}. Furthermore, EIP-Accuracy is equal to 0, IIP-Accuracy becomes $\sum_{j=1}^{m} \frac{|\text{POS}_{\mathcal{C}_*}(X_j)|}{|X_j|}$ to act as a main part of IP-Accuracy, and BIP-Accuracy becomes the sum of all Gr-Con IP-Accuracy regarding all BND and concepts.

Theorem 3.3.7 (Monotonicity IV). In Kn-Coarsening $C_1 \Rightarrow C_2$,

$$Iia_{\mathcal{D}}(\mathcal{C}_2) \le Iia_{\mathcal{D}}(\mathcal{C}_1);$$
(27)

$$\begin{split} lia_{\mathcal{D}}(\mathcal{C}_2) &= lia_{\mathcal{D}}(\mathcal{C}_1), \text{ iff } \mathsf{POS}_{\mathcal{C}_2}(X_j) = \mathsf{POS}_{\mathcal{C}_1}(X_j), \forall j \in \{1, \dots, m\},\\ &\text{ iff } \mathsf{POS}_{\mathcal{C}_2}(\mathcal{D}) = \mathsf{POS}_{\mathcal{C}_1}(\mathcal{D}). \end{split}$$

In particular, IIP-Accuracy preservation aims to keep both C-POS and C-BND and is thus called classification-qualitative preservation.

In view of the homogeneous structure, IIP-Accuracy preservation (i.e., classification-qualitative preservation) is equivalent to C-POS preservation (Theorem 3.3.7); however, the two features exhibit the metrical and regional styles, respectively. As a result, IIP-Accuracy preservation could be used to define a qualitative reduct (Definition 4.2.3), which is equivalent to the Pawlak-Reduct but has a metrical connotation. In fact, C-POS could also fully correspond to its dependency degree $\gamma_{C_*}(\mathcal{D})$. In fact,

$$\begin{cases} \gamma_{\mathcal{C}_*}(\mathcal{D}) = \sum_{j=1}^m \frac{|\text{POS}_{\mathcal{C}_*}(X_j)|}{|U|};\\ \text{lia}_{\mathcal{D}}(\mathcal{C}_*) = \sum_{j=1}^m \frac{|\text{POS}_{\mathcal{C}_*}(X_j)|}{|X_j|}. \end{cases}$$

Thus, $Iia_{\mathcal{D}}(\mathcal{C}_*)$ has homogeneous monotonicity about and also sharply contrasts with $\gamma_{\mathcal{C}_*}(\mathcal{D})$: C-POS's metrical form. As a result, when compared to $\gamma_{\mathcal{C}_*}(\mathcal{D})$, $Iia_{\mathcal{D}}(\mathcal{C}_*)$ contains X_j 's individual information, thus becoming rational and optimal. Therefore, $Iia_{\mathcal{D}}(\mathcal{C}_*)$ underlies the better reduction heuristic information and heuristic reduction algorithm, which could be reflected by Formula (30) and Algorithm 1's parallel generalization (for IIP-Accuracy), respectively.

Finally, we plan to summarize the relevant preservation deduction, which radically underlies the reduction target hierarchy and attribute reduction hierarchy. For this purpose, the preservation deduction of IP-Accuracy and IIP-Accuracy is first provided.

Corollary 3.3.8. *IP-Accuracy preservation unidirectionally deduces IIP-Accuracy preservation.*

Proof. IP-Accuracy consists of IIP-Accuracy and BIP-Accuracy, and both IP-Accuracy and IIP-Accuracy have granulation monotonicity. Thus, that IP-Accuracy reaches its coarsening maximum unidirectionally implies that IIP-Accuracy reaches its own coarsening maximum; this result could be verified by the contradiction proof and BIP-Accuracy factor. \Box

Herein, IP-Accuracy preservation is also stronger than IIP-Accuracy preservation (Corollary 3.3.8). In fact, the former is stronger than C-POS preservation (Corollary 3.3.3), while the latter is equivalent to C-POS preservation (Theorem 3.3.7). Thus, except for the above metrical proof, this metrical preservation relationship (between IP-Accuracy and IIP-Accuracy) could also be equivalently proved by the relevant regional method. In fact, the equivalent deduction transformation corresponds to the next preservation deduction commutativity.

For IP-Accuracy/IIP-Accuracy, we make great efforts to reveal the granulation monotonicity and also to mine the relevant equality conditions in the regional way. As a result, we establish a fundamental connection between the relevant metrical and regional preservation. These basic works underlie the relationship discussions between the novel IP-Accuracy/IIP-Accuracy reduction and classical regional reduction (Section 4). Now, it is time to summarize the deduction relationships of these preservation targets by Fig. 2.

In Fig. 2, there are two types of preservation targets. Concretely, metrical targets include IP-Accuracy and IIP-Accuracy; in contrast, regional targets include Set-Region preservation, C-POS preservation, and distributional preservation (i.e., POS or NEG or BND preservation). Thus, there are two unidirectional deduction lines, i.e., Line I: IP-Accuracy preservation \Rightarrow Set-Region preservation \Rightarrow Distributional preservation \Rightarrow C-POS preservation; and Line II: IP-Accuracy preservation \Rightarrow IIP-Accuracy preservation; moreover, there is a bidirectional deduction line, i.e., Line III: C-POS preservation \Leftrightarrow IIP-Accuracy preservation \Rightarrow Set-Region preservation \Leftrightarrow IIP-Accuracy preservation \Rightarrow Distributional preservation. In particular, Lines I and III compose a natural deduction line: IP-Accuracy preservation \Rightarrow Set-Region preservation \Rightarrow Distributional preservation \Rightarrow C-POS preservation \Rightarrow IIP-Accuracy preservation \Rightarrow Set-Region preservation \Rightarrow Distributional preservation \Rightarrow C-POS preservation \Rightarrow IIP-Accuracy preservation \Rightarrow IIP-Accuracy preservation \Rightarrow IIP-Accuracy preservation \Rightarrow Set-Region preservation \Rightarrow Distributional preservation \Rightarrow C-POS preservation \Rightarrow IIP-Accuracy preservation \Rightarrow IIP-Accuracy preservation \Rightarrow Set-Region preservation \Rightarrow IIP-Accuracy preservation \Rightarrow Set-Region preservation \Rightarrow Distributional preservation \Rightarrow C-POS preservation \Rightarrow IIP-Accuracy preservation \Rightarrow Set-Region preservation \Rightarrow Distributional preservation \Rightarrow C-POS preservation \Rightarrow IIP-Accuracy preservation \Rightarrow Set-Region preservation \Rightarrow Distributional preservation \Rightarrow C-POS preservation \Rightarrow IIP-Accuracy preservation \Rightarrow Set-Region \Rightarrow

Note that preservation targets' deduction radically underlies relevant preservation reducts' hierarchy. In particular, some reduction hierarchies were provided in [74,75] for the two-category case. In Section 4, some reduction hierarchies will be investigated, and they apply to the more general multi-category case.

3.4. Importance-accuracy's granular hierarchical structure

Gr-Con IP-Accuracy, Kn-Con IP-Accuracy/IIP-Accuracy, Kn-Cl IP-Accuracy/IIP-Accuracy are provided in Sections 2.3, 3.2, and 3.3, respectively. In this subsection, the three IP-Accuracy levels are summarized to establish the relevant GrC hierarchical structure. Herein, the systematic structural figure is first provided by Fig. 3.

Next, this hierarchical structure figure is utilized to explain IP-Accuracy's GrC construction. There are four fundamental granular systems, three GrC levels (micro bottom, meso middle, macro top), two main RS-Theory terms (Approx-Space and attribute reduction), and two GrC-hierarchical evolution lines.

- (1) Bottom-System (U, C_*, X) , at the micro level, mainly concerns the knowledge-granule and concept. Thus, Gr-Con IP-Accuracy $ia_X([x]_{C_*})$ is constructed by fusing accuracy and importance. Herein, $ia_X([x]_{C_*})$ can systematically measure the following: reasoning $[x]_{C_*} \Rightarrow X$, granule $[x]_{C_*}$ (based on concept *X*), and concept *X* (based on granule $[x]_{C_*}$).In fact, the relevant system and level correspond to Approx-Space.
- (2) Middle-System (U, C, X), at the meso level, mainly concerns the knowledge and concept. Thus, Kn-Con IP-Accuracy $ia_X(C_*)$ is constructed by integrating Gr-Con IP-Accuracy $ia_X([x]_{C_*})$ because knowledge C_* is actually a granular system. Herein, $ia_X(C_*)$ can systematically measure reasoning $C_* \Rightarrow X$, knowledge C_* (based on concept *X*), and concept *X* (based on knowledge C_*). Moreover, IIP-Accuracy $Iia_X(C_*)$ is provided to describe POS. As a result, Kn-Con IP-Accuracy and IIP-Accuracy achieve basic granulation monotonicity.
- (3) Top-System (U, C, D), at the macro level, mainly concerns knowledge and classification. Thus, Kn-Cl IP-Accuracy $ia_D(C_*)$ is naturally constructed by integrating Kn-Con IP-Accuracy $ia_X(C_*)$ because classification D is actually concepts system $\{X_1, \ldots, X_m\}$. Herein, $ia_D(C_*)$ can systematically measure reasoning $C_* \Rightarrow D$, knowledge C_* (based on classification D), and classification D (based on knowledge C_*). Moreover, Kn-Cl IIP-Accuracy $Iia_D(C_*)$ is provided to describe C-POS. Thus, Kn-Cl IP-Accuracy and IIP-Accuracy also achieve basic granulation monotonicity. In particular, D-Table attribute reduction exactly corresponds to the top system and macro level; as a result, the perfect granulation monotonicity underlies further attribute reduction.
- (4) For the above three systems, IP-Accuracy utilizes the natural sum integration to realize the valuable GrC-hierarchical promotion, i.e., $ia_X([x]_{\mathcal{C}_*}) \to ia_X(\mathcal{C}_*) \to ia_\mathcal{D}(\mathcal{C}_*)$; this evolutionary line is marked by Line I. Meanwhile, IP-Accuracy's GrC evolution could also be completed by establishing a new connection based on the fourth system. The added system (U, C_*, D) , which is also at the meso middle level, mainly concerns knowledge-granule, classification, and Cl-Gr IP-Accuracy $ia_{[x]_{C_*}}(\mathcal{D})$, as well as Cl-Gr IIP-Accuracy $Iia_{[x]_{C_*}}(\mathcal{D})$. Herein, $ia_{[x]_{C_*}}(\mathcal{D})$ can systematically measure the following: reasoning $[x]_{\mathcal{C}_*} \Rightarrow \mathcal{D}$, classification \mathcal{D} (based on granule $[x]_{\mathcal{C}_*}$), and granule $[x]_{C_*}$ (based on classification \mathcal{D}). According to next Definition 3.4.1, Cl-Gr IP-Accuracy $ia_{[x]_{C_a}}(\mathcal{D})$ integrates Con-Gr IP-Accuracy $ia_{[x]_{C_*}}(X)$ in Bottom-System and is further integrated into Cl-Kn IP-Accuracy $ia_{\mathcal{C}_*}(\mathcal{D})$ in Top-System. Thus, $ia_{[x]_{C_1}}(\mathcal{D})$ could also be used to realize the new GrChierarchical evolution, i.e., $ia_{[x]_{\mathcal{C}_*}}(X) \to ia_{[x]_{\mathcal{C}_*}}(\mathcal{D}) \to ia_{\mathcal{C}_*}(\mathcal{D});$ this promotional line is marked by Line II.

(5) Line II integrates the classification first and knowledge second, while Line I adopts the contrary order. Thus, different terms usually have different semantics emphases. However, according to Theorem 3.4.2, both lines have commutativity, which radically originates from the algebraic sum commutativity; as a result, both lines obtain the same final measure in macro Top-System. In particular, Line II could establish the set-operation property regarding the classification system, which sharply contrasts with Line I's granulation monotonicity. Therefore, with the exception of granulation monotonicity, Top-System could finally achieve more classification system properties.

In summary, we adopt a bottom-top strategy to perform IP-Accuracy's GrC construction. In particular, IP-Accuracy granulation monotonicity (Theorems 3.3.2 and 3.3.7) and relevant preservation deduction (Fig. 2) not only fully reflect our basic GrC works but also effectively underlie the later attribute reduction (Section 4).

Next, we provide the relevant definition and property for Middle-System (U, C_*, D) , which are mainly concerned in the above explanation Items (4) (5). For the new measure, IIP-Accuracy could gain the relevant construction and analysis, which are similar to IP-Accuracy's (especially in Items (4) (5)).

Definition 3.4.1.

- (1) Con-Gr IP-Accuracy $ia_{[x]_{\mathcal{C}_*}}(X_j) = ia_{X_j}([x]_{\mathcal{C}_*}).$
- (2) CI-Gr IP-Accuracy $ia_{[x]_{C_*}}(\mathcal{D}) = \sum_{j=1}^m ia_{[x]_{C_*}}(X_j);$ CI-Gr IIP-Accuracy $ia_{[x]_{C_*}}(\mathcal{D}) = \sum_{\text{POS}_{C_*}(X_j) \supseteq [x]_{C_*}} ia_{[x]_{C_*}}(X_j).$
- (3) CI-Kn IP-Accuracy $ia_{\mathcal{C}_*}(\mathcal{D}) = \sum_{[x]_{\mathcal{C}_*} \in U/\mathcal{C}_*} ia_{[x]_{\mathcal{C}_*}}(\mathcal{D});$ CI-Kn IIP-Accuracy $lia_{\mathcal{C}_*}(\mathcal{D}) = \sum_{[x]_{\mathcal{C}_*} \in U/\mathcal{C}_*} lia_{[x]_{\mathcal{C}_*}}(\mathcal{D}).$

Theorem 3.4.2.

$$\begin{cases} ia_{\mathcal{C}_*}(\mathcal{D}) = ia_{\mathcal{D}}(\mathcal{C}_*);\\ lia_{\mathcal{C}_*}(\mathcal{D}) = lia_{\mathcal{D}}(\mathcal{C}_*). \end{cases}$$
(28)

Note that IP-Accuracy's GrC hierarchical system could be compared to and comprehended by a practical assessment pattern, at least in some sense. Thus, a relevant example is finally provided to gain more thorough explanations.

Example 1. There are *m* players X_j and *n* judges $[x]_i$. For players X_j , judges $[x]_i$ are endowed with objective systematic weights \mathcal{I}_{ij} , and they also have their own subject judgement \mathcal{A}_{ij} . Thus, the weighted product $\mathcal{I}_{ij} \times \mathcal{A}_{ij}$ leads to the effective judgement, while the further weighted sum provides the synthetic assessment. Concretely, there are four systems and relevant assessments.

- (1) In system ([x]_i, X_j) (with one judge and one player), data *I*_{ij} × *A*_{ij} could be used for the judge to synthetically assess the player (based on the systematic weight) and also for the player to synthetically evaluate the judge. Thus, *A*_{ij}, *I*_{ij}, *X*_{ij} × *A*_{ij} correspond to accuracy, importance, and IP-Accuracy in Bottom-System, respectively. In particular, the former three's subjectivity, objectivity, effectiveness generally correspond to the latter three's individuality, systematicness, synthesis, respectively.
- (2) In system $(\{[x]_1, \ldots, [x]_n\}, X_j)$ (with systematic judges and one player), data $\sum_{i=1}^n \mathcal{I}_{ij} \times \mathcal{A}_{ij}$ could be used for the judges system to synthetically assess the player and also for the latter to synthetically evaluate the former. Thus, the weighted sum information corresponds to Kn-Con IP-Accuracy in Middle-System one.

- (3) In system ($[x]_i, \{X_1, \ldots, X_m\}$) (with one judge and systematic players), data $\sum_{j=1}^m \mathcal{I}_{ij} \times \mathcal{A}_{ij}$ could be used for the judge to synthetically assess the players system and also for the latter to synthetically evaluate the former. Thus, the weighted sum information corresponds to Cl-Gr IP-Accuracy in Middle-System two.
- (4) In system ({[x]₁,..., [x]_n}, {X₁,..., X_m}) (with systematic judges and systematic players), data ∑^m_{j=1}[∑ⁿ_{i=1} I_{ij} × A_{ij}] = ∑ⁿ_{i=1}[∑^m_{j=1} I_{ij} × A_{ij}] could be used for the judges system to synthetically assess the players system and also for the latter to synthetically evaluate the former. Thus, the weighted sum information corresponds to Kn-Cl/Cl-Kn IP-Accuracy in Top-System. □

4. Importance-accuracy attribute reduction

IP-Accuracy's GrC construction is hierarchically completed in Section 3. IP-Accuracy's/IIP-Accuracy's granulation monotonicity (Theorem 3.3.2/3.3.7) not only verifies IP-Accuracy's fusion superiority but also underlies further reduction construction. In particular, the monotonicity equality conditions and their equivalent regional manifestations, as well as the in-depth preservation deduction (Fig. 2), solidly underlie IP-Accuracy/IIP-Accuracy reducts and their hierarchies (regarding regional reducts).

In this section, IP-Accuracy/IIP-Accuracy and its monotonicity are utilized to finally investigate attribute reduction regarding D-Table (U, C, D). Concretely, qualitative/quantitative reducts, their hierarchies, and heuristic algorithms are systematically provided. Herein, IP-Accuracy/IIP-Accuracy is at the macro top, and constraint Cl-Kn is usually omitted. In particular, the hierarchical reduction theory has been established in [74,75]. In this section, we will give a well-researched and descriptive emphasis to the hierarchical reducts, in view of multiple reducts' diversity.

4.1. Reduct preliminaries

This subsection provides some reduct preliminaries. First, the classical Pawlak-Reduct [35,36] is reviewed; then, generalized reducts are proposed based on a monotonicity target; finally, hierarchical reducts [74,75] are introduced. Herein, let $\mathcal{B} \subseteq C$.

Lemma 4.1.1. The following two items are equivalent:

- (1) $\text{POS}_{\mathcal{B}'}(\mathcal{D}) \neq \text{POS}_{\mathcal{B}}(\mathcal{D}), \forall \mathcal{B}' \subset \mathcal{B};$
- (2) $\text{POS}_{\mathcal{B}-\{b\}}(\mathcal{D}) \neq \text{POS}_{\mathcal{B}}(\mathcal{D}), \forall b \in \mathcal{B}.$

Herein, Items (1) (2) correspond to the maximality and independence regarding C-POS preservation, respectively. They could be equivalently used for Pawlak-Reduct.

Definition 4.1.2 (Pawlak-Reduct). \mathcal{B} is Pawlak-Reduct of \mathcal{C} if it satisfies following Items (1) (2) or (1) (3) regarding C-POS preservation.

(1) $\text{POS}_{\mathcal{B}}(\mathcal{D}) = \text{POS}_{\mathcal{C}}(\mathcal{D});$

- (2) $\text{POS}_{\mathcal{B}'}(\mathcal{D}) \neq \text{POS}_{\mathcal{B}}(\mathcal{D}), \forall \mathcal{B}' \subset \mathcal{B};$
- (3) $\text{POS}_{\mathcal{B}-\{b\}}(\mathcal{D}) \neq \text{POS}_{\mathcal{B}}(\mathcal{D}), \forall b \in \mathcal{B}.$

Herein, $Core(C) = \{c \in C : POS_{C-\{c\}}(D) \neq POS_{C}(D)\}$ and Red(C) denote Pawlak-Reduct's core and set, respectively.

In view of C-POS granulation monotonicity, Pawlak-Reduct is naturally established by preserving C-POS, thus becoming a qualitative reduct. Herein, Item (1) reflects the joint sufficiency regarding C-POS preservation, and Items (2) (3) could be arbitrarily chosen for the individual necessity. The following core-reduct relationship is basic.

Proposition 4.1.3. $Core(\mathcal{C}) = \bigcap_{\mathcal{B} \in Red(\mathcal{C})} \mathcal{B}.$

Among all reduction algorithms, heuristic algorithms are appreciated in view of their high efficiency. The basic heuristic algorithm mainly utilizes the dependency degree, and

$$sig(c, \mathcal{C}_*; \mathcal{D}) = \gamma_{\mathcal{C}_*}(\mathcal{D}) - \gamma_{\mathcal{C}_* - \{c\}}(\mathcal{D})$$
(29)

becomes the corresponding reduction-heuristic information for conditional attribute *c*.

The above Pawlak-Reduct results mainly originate from the C-POS granulation monotonicity (Proposition 3.1.4). In fact, they have powerful generalization. Thus, we primarily propose a type of generalized reducts via a monotonicity target. Note that Ref. [11] generalizes multiple types of exiting attribute reducts. In contrast, the generalized reducts proposed below mainly focus on and promote fundamental monotonicity.

Definition 4.1.4. Let *MT* be an isomorphism between complete lattices $(2^{\mathcal{C}}, \subseteq)$ and (L, \leq) ; herein, if $\mathcal{C}_2 \subseteq \mathcal{C}_1 \subseteq \mathcal{C}$, then $MT(\mathcal{C}_2) \preceq MT(\mathcal{C}_1)$, while $MT(\emptyset)$, $MT(\mathcal{C})$ become the least and greatest elements in *L*, respectively. Thus, *MT* is called *a monotonicity target (MT)* regarding \mathcal{C} . For Kn-Coarsening $\mathcal{C}_1 \Rightarrow \mathcal{C}_2$, monotonicity target preservation (*MT*-*Preservation*) is defined by $MT(\mathcal{C}_2) = MT(\mathcal{C}_1)$.

Note that Ref. [74] utilizes the complete lattice to establish the structural target (ST), which is multi-dimensional. In this paper, we propose the monotonicity target (MT), which is also related to the complete lattice but is integral. For example, C-POS is MT in the mapping sense, where the relevant mapping changes \emptyset to \emptyset . Clearly, MT holds granulation monotonicity, and MT-Preservation is naturally defined. Thus, the relevant MT-Preservation reducts could be further established.

Lemma 4.1.5. The following two items are equivalent:

- (1) $MT(\mathcal{B}') \neq MT(\mathcal{B}), \forall \mathcal{B}' \subset \mathcal{B};$ (2) $MT(\mathcal{B} - \{b\}) \neq MT(\mathcal{B}), \forall b \in \mathcal{B}.$
- **Definition 4.1.6** (MT-Preservation Reduct). \mathcal{B} is MT-Preservation Reduct of \mathcal{C} if it satisfies following Items (1) (2) or (1) (3) regarding MT-Preservation.

(1) $MT(\mathcal{B}) = MT(\mathcal{C});$

- (2) $MT(\mathcal{B}') \neq MT(\mathcal{B}), \forall \mathcal{B}' \subset \mathcal{B};$
- (3) $MT(\mathcal{B} \{b\}) \neq MT(\mathcal{B}), \forall b \in \mathcal{B}.$

Herein, $Core_{MT}(C) = \{c \in C : MT(C - \{c\}) \neq MT(C)\}$ and $Red_{MT}(C)$ denote MT-Preservation Reduct's core and set, respectively.

Proposition 4.1.7. $Core_{MT}(\mathcal{C}) = \bigcap_{\mathcal{B} \in Red_{MT}(\mathcal{C})} \mathcal{B}.$

Lemma 4.1.5 and Proposition 4.1.7 are proved in Appendixes C and D, where MT granulation monotonicity plays the key role. In fact, they extend Lemma 4.1.1 and Proposition 4.1.3, respectively. Moreover, MT-Preservation Reduct extends Pawlak-Reduct. Thus, we establish a more general framework of attribute reduction by virtue of the national target with granulation monotonicity. MT-Preservation Reduct and its property apply to extensive monotonicity reducts. In particular, all of the reducts proposed later fall into the MT-Preservation Reduct category, so MT-Preservation Reduct provides a unified normal pattern. For these reducts, we mainly emphasize their target monotonicity and preservation while neglecting their natural definition equivalence and core-reduct relationship. As a result, we could more thoroughly focus on multiple reduct types and their systematic hierarchies.

Theorem 4.1.8 (Hierarchical Reduct). Strong and weak reducts are defined by the strong and weak reduction targets, respectively. The strong reduct's core and set are denoted by $Core_{strong}(C)$ and $Red_{strong}(C)$, respectively, and so are the weak symbols $Core_{weak}(C)$ and $Red_{weak}(C)$. Thus, two items hold as follows:

- (1) $Core_{strong}(\mathcal{C}) \supseteq Core_{weak}(\mathcal{C});$
- (2) $\forall \mathcal{B}_{strong} \in Red_{strong}(\mathcal{C}), \exists \mathcal{B}_{weak} \in Red_{weak}(\mathcal{C}), s.t., \mathcal{B}_{weak} \subseteq \mathcal{B}_{strong}.$

According to the hierarchical reduction theory [74,75], the strong reduct has a strong reduction target, and its core includes weak reduct's core. Moreover, in contrast to Item (2) (Theorem 4.1.8), the weak reduct could exist by transcending the inclusion relationship from the strong reduct. Moreover, denote $\mathcal{B}_{strong} \Rightarrow \mathcal{B}_{weak}$; then, relation \Rightarrow becomes a partial order with the natural transitivity, but it transcends simple set relations \subseteq and \supseteq .

4.2. Importance-accuracy qualitative reducts

This subsection utilizes IP-Accuracy and IIP-Accuracy to establish two types of qualitative reducts and further studies their relationships for qualitative Pawlak-Reduct.

Definition 4.2.1 (IP-Accuracy Reduct). \mathcal{B} is IP-Accuracy Reduct of \mathcal{C} if it satisfies following Items (1) (2) or (1) (3) regarding IP-Accuracy preservation.

- (1) $ia_{\mathcal{D}}(\mathcal{B}) = ia_{\mathcal{D}}(\mathcal{C});$
- (2) $ia_{\mathcal{D}}(\mathcal{B}') \neq ia_{\mathcal{D}}(\mathcal{B}), \forall \mathcal{B}' \subset \mathcal{B};$
- (3) $ia_{\mathcal{D}}(\mathcal{B} \{b\}) \neq ia_{\mathcal{D}}(\mathcal{B}), \forall b \in \mathcal{B}.$

Herein, $Core_{ia}(\mathcal{C}) = \{c \in \mathcal{C} : ia_{\mathcal{D}}(\mathcal{C} - \{c\}) \neq ia_{\mathcal{D}}(\mathcal{C})\}$ and $Red_{ia}(\mathcal{C})$ denote IP-Accuracy Reduct's core and set, respectively.

IP-Accuracy's granulation monotonicity and relevant equality conditions are provided in Theorem 3.3.2. In view of Gr-Merging's accuracy requirement and region feature, the objective of IP-Accuracy preservation is actually to quantitatively keep the classification. By virtue of this reduction target, IP-Accuracy Reduct is naturally established to become a qualitative reduct but exhibits a metrical style.

Theorem 4.2.2. *IP-Accuracy Reduct is stronger than Pawlak-Reduct. Thus,*

(1) $Core_{ia}(\mathcal{C}) \supseteq Core(\mathcal{C}).$

(2) $\forall \mathcal{B}_{ia} \in Red_{ia}(\mathcal{C}), \exists \mathcal{B} \in Red(\mathcal{C}), s.t., \mathcal{B} \subseteq \mathcal{B}_{ia}.$

IP-Accuracy preservation unidirectionally deduces C-POS preservation (Corollary 3.3.3 or Fig. 2), so IP-Accuracy Reduct is stronger than Pawlak-Reduct. Thus, the above hierarchical reduction description becomes clear.

Definition 4.2.3 (IIP-Accuracy Reduct). \mathcal{B} is IIP-Accuracy Reduct of \mathcal{C} if it satisfies following Items (1) (2) or (1) (3) regarding IIP-Accuracy preservation.

(1) $lia_{\mathcal{D}}(\mathcal{B}) = lia_{\mathcal{D}}(\mathcal{C});$

- (2) $lia_{\mathcal{D}}(\mathcal{B}') \neq lia_{\mathcal{D}}(\mathcal{B}), \forall \mathcal{B}' \subset \mathcal{B};$
- (3) $lia_{\mathcal{D}}(\mathcal{B} \{b\}) \neq lia_{\mathcal{D}}(\mathcal{B}), \forall b \in \mathcal{B}.$

Herein, $Core_{lia}(C) = \{c \in C : lia_{D}(C - \{c\}) \neq lia_{D}(C)\}$ and $Red_{lia}(C)$ denote IIP-Accuracy Reduct's core and set, respectively.

Theorem 4.2.4. IIP-Accuracy Reduct is equivalent to Pawlak-Reduct. Thus,

- (1) $Core_{lia}(C) = Core(C);$
- (2) $Red_{Iia}(\mathcal{C}) = Red(\mathcal{C}).$

IIP-Accuracy's granulation monotonicity and relevant equality conditions are provided in Theorem 3.3.7. In view of Gr-Merging's C-POS result, the objective of IIP-Accuracy preservation is actually to qualitatively keep the classification. By virtue of this reduction target, IIP-Accuracy Reduct is naturally established to become a qualitative reduct with a metrical style. Clearly, IIP-Accuracy Reduct is equivalent to Pawlak-Reduct, according to Theorem 3.3.7 or Fig. 2.

Theorem 4.2.5. *IP-Accuracy Reduct is stronger than IIP-Accuracy Reduct. Thus,*

- (1) $Core_{ia}(\mathcal{C}) \supseteq Core_{Iia}(\mathcal{C});$
- (2) $\forall \mathcal{B}_{ia} \in Red_{ia}(\mathcal{C}), \exists \mathcal{B}_{lia} \in Red_{lia}(\mathcal{C}), s.t., \mathcal{B}_{lia} \subseteq \mathcal{B}_{ia}.$

According to Corollary 3.3.8 or Fig. 2, Theorem 4.2.5 clearly exhibits the hierarchy between IP-Accuracy and IIP-Accuracy Reducts. Moreover, this hierarchy conclusion could be inferred by Theorems 4.2.2 and 4.2.4.

The above reduction hierarchies mainly focus on only three reduct types: IP-Accuracy Reduct, IIP-Accuracy Reduct, and Pawalk-Reduct. Note that the hierarchy radically originates from the reduction target deduction. Thus, more reduction hierarchies could be established by Fig. 2 and become later research content (Section 4.4).

Aiming at qualitative IP-Accuracy and IIP-Accuracy Reducts, heuristic reduction algorithms are next constructed. Clearly, underlying IP-Accuracy and IIP-Accuracy (and their monotonicity) could be fully utilized. Thus, IP-Accuracy and IIP-Accuracy are first used to construct two types of reduction-heuristic information:

$$\begin{cases} \operatorname{sig}_{ia}(c, \mathcal{C}_*; \mathcal{D}) = ia_{\mathcal{D}}(\mathcal{C}_*) - ia_{\mathcal{D}}(\mathcal{C}_* - \{c\});\\ \operatorname{sig}_{lia}(c, \mathcal{C}_*; \mathcal{D}) = lia_{\mathcal{D}}(\mathcal{C}_*) - lia_{\mathcal{D}}(\mathcal{C}_* - \{c\}). \end{cases}$$
(30)

Moreover, the reduction core is easily calculated to underlie attribute addition algorithms. Thus, both the heuristic information and reduction core are utilized to develop the relevant heuristic reduction algorithm. Next, the IP-Accuracy-Based heuristic algorithm for IP-Accuracy Reduct is provided, i.e., Algorithm 1, while the parallel algorithm for IIP-Accuracy Reduct could be similarly developed. **Algorithm 1.** IP-Accuracy-Based heuristic algorithm for IP-Accuracy Reduct

Input: D-Table $(U, C \cup D)$;
Output: IP-Accuracy Reduct $\mathcal{B}_{ia} \in Red_{ia}(\mathcal{C})$.
1: Compute $Core_{ia}(\mathcal{C})$.
2: $\mathcal{B}_{ia} = Core_{ia}(\mathcal{C}).$
3: while $ia_{\mathcal{D}}(\mathcal{B}_{ia}) \neq ia_{\mathcal{D}}(\mathcal{C})$ do
4: $\forall c \in C - B_{ia}$, calculate $sig_{ia}(c, B_{ia}; D)$;
choose $c_0 = \arg \max_{c \in C - B_{ia}} sig_{ia}(c, B_{ia}; D);$
let $\mathcal{B}_{ia} = \mathcal{B}_{ia} \cup \{c_0\}.$
5: end while
6: return \mathcal{B}_{ia} .

Based on the attribute addition strategy, Algorithm 1 mainly utilizes IP-Accuracy heuristic information $sig_{ia}(c, C_*; D)$ to calculate IP-Accuracy Reduct. Concretely, Step 1 gives the basic core; Step 2 chooses the core as the reduct basis; Steps 3 and 4 further seek an attribute subset to satisfy the IP-Accuracy preservation target, and the added c_0 is randomly chosen inside $C - B_{ia}$ via the highest IP-Accuracy heuristic information. As a result, Algorithm 1 could usually yield one IP-Accuracy Reduct (regardless of independence/maximality), thus becoming convergent and effective.

Herein, the two heuristic reduct algorithms (i.e., Algorithm 1 and its parallel) are correspondingly inspired by their own bearing measures. Moreover, they could alternatively be inspired by their opposite bearing measures. In other words, IP-Accuracy/IIP-Accuracy Reduct could be heuristically calculated by IIP-Accuracy/IP-Accuracy. For Pawlak-Reduct, Algorithm 1 effectively provides a heuristic algorithm because IP-Accuracy Reduct is stronger than Pawlak-Reduct (Theorem 4.2.2); in contrast, the parallel algorithm for IIP-Accuracy Reduct directly provides a heuristic algorithm because IIP-Accuracy Reduct is equivalent to Pawlak-Reduct (Theorem 4.2.4).

4.3. Importance-accuracy quantitative reducts

In Section 4.2, two qualitative reducts are proposed, and both are compared to Pawlak-Reduct. For qualitative reducts, the qualitative equality with the measure criterion becomes somewhat strict, especially in certain practical environments. In practice, there is not only calculation error but also data noise. Thus, the relevant quantitative reduct with some metrical change exhibits radical improvements and underlies extensive applications. Moreover, the quantitative reduct expansion could enlarge IP-Accuracy Reduct's application scope because IP-Accuracy preservation is relatively strict (Fig. 2). Usually, the quantitative reduct is indirectly realized by the quantitative model and its quantitative function. Herein, we alternatively resort to a new approach; i.e., the quantitative reduct is directly established by the monotonous measure and its quantitative function. Thus, this subsection mainly depends on IP-Accuracy/IIP-Accuracy to construct quantitative reducts, including the novel tolerant and approximate reducts.

Theorem 4.3.1.

(1) Define $U/\emptyset = U$ and $CKS_{\mathcal{C}} = \{U/\mathcal{C}_* : \mathcal{C}_* \subseteq C\}$. Then, $(CKS_{\mathcal{C}}, \Rightarrow)$ constitutes a complete lattice; U/\mathcal{C} and U/\emptyset serve as the least and greatest elements, respectively.

- (2) Define mapping $kn : 2^{\mathcal{C}} \to CKS_{\mathcal{C}}, \forall \mathcal{C}_* \in 2^{\mathcal{C}}, kn(\mathcal{C}_*) = U/\mathcal{C}_*.$ Then, surjective kn constructs an order-preservation mapping between complete lattices $(2^{\mathcal{C}}, \supseteq)$ and $(CKS_{\mathcal{C}}, \Rightarrow)$. Furthermore, by virtue of kn, both complete lattices exhibit a complete lattice homomorphism.
- (3) By virtue of IP-Accuracy $ia_{\mathcal{D}}(\mathcal{C}_*)$, define mapping $m_{ia} : CKS_{\mathcal{C}} \rightarrow [1, ia_{\mathcal{D}}(\mathcal{C})], \forall U/\mathcal{C}_* \in CKS_{\mathcal{C}}, m_{ia}(U/\mathcal{C}_*) = ia_{\mathcal{D}}(\mathcal{C}_*)$. Then, nonsurjective m_{ia} constructs an order-preservation mapping between complete lattices ($CKS_{\mathcal{C}}, \Rightarrow$) and ($[1, ia_{\mathcal{D}}(\mathcal{C})], \ge$).
- (4) IP-Accuracy $ia_{\mathcal{D}}(\mathcal{C}_*)$ essentially determines mapping $ia_{\mathcal{D}}: 2^{\mathcal{C}} \rightarrow [1, ia_{\mathcal{D}}(\mathcal{C})], \forall \mathcal{C}_* \in 2^{\mathcal{C}}, ia_{\mathcal{D}}(\mathcal{C}_*) = ia_{\mathcal{D}}(\mathcal{C}_*)$. Thus, $ia_{\mathcal{D}}$ becomes the composite mapping of kn and m_{ia} , i.e., $ia_{\mathcal{D}} = m_{ia} \circ kn$ and $ia_{\mathcal{D}}(\mathcal{C}_*) = (m_{ia} \circ kn)(\mathcal{C}_*) = m_{ia}(kn(\mathcal{C}_*)) = m_{ia}(U/\mathcal{C}_*) = ia_{\mathcal{D}}(\mathcal{C}_*)$. Additionally, $ia_{\mathcal{D}}$ constructs an order-preservation mapping between complete lattices $(2^{\mathcal{C}}, \supseteq)$ and $([1, ia_{\mathcal{D}}(\mathcal{C})], \geq)$. As a result, $ia_{\mathcal{D}}$ holds the monotonicity and boundedness, i.e.,

$$\begin{cases} ia_{\mathcal{D}}(\mathcal{C}_1) \ge ia_{\mathcal{D}}(\mathcal{C}_2), & \text{if } \mathcal{C} \supseteq \mathcal{C}_1 \supseteq \mathcal{C}_2; \\ ia_{\mathcal{D}}(\mathcal{C}_*) \in [1, ia_{\mathcal{D}}(\mathcal{C})], \forall \mathcal{C}_* \subseteq \mathcal{C}. \end{cases}$$

In Theorem 4.3.1, three complete lattices and their order preservation mappings are thoroughly constructed for the conditional attribute, coarsening knowledge, and IP-Accuracy measure. These algebraic results become natural but in-depth. Herein, CKS_C means the coarsening knowledge set. Moreover, extreme \emptyset and $U/\emptyset = \{U\}$ lead to

$$ia_{\varnothing}(\mathcal{D}) = \sum_{j=1}^{m} ia_{\varnothing}(X_j) = \sum_{j=1}^{m} \frac{|U \cap X_j|^2}{|U||X_j|} = \sum_{j=1}^{m} \frac{|X_j|}{|U|} = 1;$$

thus, IP-Accuracy $ia_{\mathcal{D}}(\mathcal{C}_*)$ has the lower bound 1. Next, these relevant mathematical results are exhibited in Fig. 4. There, the three complete lattices and their representative elements (including the least and greatest elements) are located at three different levels; furthermore, the three order preservation mappings connect these levels. In particular, IP-Accuracy's granulation monotonicity regarding coarsening order relation \Rightarrow is also expressed by a non-increasing mapping figure.

Theorem 4.3.2.

$$\lim_{\mathcal{C} \to \mathcal{C}} ia_{\mathcal{D}}(\mathcal{C}_*) = ia_{\mathcal{D}}(\mathcal{C}),\tag{31}$$

i.e., $\forall \varepsilon > 0, \exists \mathcal{C}_{\varepsilon}, \text{ if } \mathcal{C}_* \supseteq \mathcal{C}_{\varepsilon} \text{ (i.e., } \mathcal{C}_* \Rightarrow \mathcal{C}_{\varepsilon}), \text{ then } ia_{\mathcal{D}}(\mathcal{C}) - ia_{\mathcal{D}}(\mathcal{C}_*) < \varepsilon.$

In Theorem 4.3.2, IP-Accuracy's limit form is provided to describe IP-Accuracy's granulation monotonicity. In particular, the relevant ε - C_{ε} description could be reflected by Fig. 4. This limit result reflects a sort of generalized continuity for IP-Accuracy's granulation monotonicity, thus underlying the next quantitative reduct construction.

Definition 4.3.3. Define mappings $dr_{ia_{\mathcal{D}}}$, $er_{ia_{\mathcal{D}}}: 2^{\mathcal{C}} \to [0, 1]$, $\forall \mathcal{C}_* \in 2^{\mathcal{C}}$,

$$\begin{cases} dr_{ia_{\mathcal{D}}}(\mathcal{C}_{*}) = \frac{ia_{\mathcal{D}}(\mathcal{C}) - ia_{\mathcal{D}}(\mathcal{C}_{*})}{ia_{\mathcal{D}}(\mathcal{C})} = 1 - \frac{ia_{\mathcal{D}}(\mathcal{C}_{*})}{ia_{\mathcal{D}}(\mathcal{C})},\\ er_{ia_{\mathcal{D}}}(\mathcal{C}_{*}) = 1 - dr_{ia_{\mathcal{D}}}(\mathcal{C}_{*}) = \frac{ia_{\mathcal{D}}(\mathcal{C}_{*})}{ia_{\mathcal{D}}(\mathcal{C})}. \end{cases}$$
(32)

Herein, $dr_{ia_{\mathcal{D}}}(\mathcal{C}_*)$ and $er_{ia_{\mathcal{D}}}(\mathcal{C}_*)$ are called knowledge \mathcal{C}_* 's *discrepancy and equality rates* regarding knowledge \mathcal{C} 's IP-Accuracy, respectively.

Proposition 4.3.4.

$$\begin{cases} dr_{ia_{\mathcal{D}}}(\mathcal{C}_{*}) \in [0, 1 - \frac{1}{ia_{\mathcal{D}}(\mathcal{C})}];\\ er_{ia_{\mathcal{D}}}(\mathcal{C}_{*}) \in [\frac{1}{ia_{\mathcal{D}}(\mathcal{C})}, 1]; \end{cases}$$
(33)

 $dr_{ia_{\mathcal{D}}}(\mathcal{C}_*) = 0$, iff $ia_{\mathcal{D}}(\mathcal{C}_*) = ia_{\mathcal{D}}(\mathcal{C})$, iff $er_{ia_{\mathcal{D}}}(\mathcal{C}_*) = 1$.

Theorem 4.3.5 (Monotonicity V). Discrepancy and approximation rates hold granulation monotonicity. Concretely, in Kn-Coarsening $\mathcal{C}_1 \Rightarrow \mathcal{C}_2$,

$$\begin{cases} dr_{ia_{\mathcal{D}}}(\mathcal{C}_{2}) \geq dr_{ia_{\mathcal{D}}}(\mathcal{C}_{1});\\ er_{ia_{\mathcal{D}}}(\mathcal{C}_{2}) \leq er_{ia_{\mathcal{D}}}(\mathcal{C}_{1}); \end{cases}$$
(34)

 $dr_{ia_{\mathcal{D}}}(\mathcal{C}_2) = dr_{ia_{\mathcal{D}}}(\mathcal{C}_1)$, iff $ia_{\mathcal{D}}(\mathcal{C}_2) = ia_{\mathcal{D}}(\mathcal{C}_1)$, iff $er_{ia_{\mathcal{D}}}(\mathcal{C}_2) = er_{ia_{\mathcal{D}}}(\mathcal{C}_1)$.

Complementary discrepancy and equality rates are proposed to measure discrepancy and equality degrees between knowledge C_* 's and knowledge C's IP-Accuracy, respectively. Herein, C's P-Accuracy acts as the ideal target value, thus becoming the comparison criterion. Moreover, a controllable measure style of relative ratio is adopted to form the relevant interval range with $\frac{1}{ia_{\mathcal{D}}(\mathcal{C})}$. In particular, the relevant granulation monotonicity is provided and thus leads to the corresponding quantitative reducts as follows.

Definition 4.3.6 (IP-Accuracy Tolerant Reduct). Give IP-Accuracy tolerance rate $\delta \in [0, 1 - \frac{1}{ia_{\mathcal{D}}(\mathcal{C})}]$. \mathcal{B} is IP-Accuracy δ -Tolerant > Reduct of C if it satisfies following Items (1) (2) or (1) (3) regarding IP-Accuracy δ -tolerance preservation.

- $\begin{array}{ll} (1) & dr_{ia_{\mathcal{D}}}(\mathcal{B}) \leq \delta; \\ (2) & dr_{ia_{\mathcal{D}}}(\mathcal{B}') > \delta, \ \forall \mathcal{B}' \subset \mathcal{B}; \\ (3) & dr_{ia_{\mathcal{D}}}(\mathcal{B} \{b\}) > \delta, \ \forall b \in \mathcal{B}. \end{array}$

Herein, $Core_{ia}^{\delta}(\mathcal{C}) = \{c \in \mathcal{C} : dr_{ia_{\mathcal{D}}}(\mathcal{C} - \{c\}) > \delta\}$ and $Red_{ia}^{\delta}(\mathcal{C})$ denote IP-Accuracy δ -Tolerant Reduct's core and set, respectively.

Definition 4.3.7 (IP-Accuracy Approximate Reduct). Give IP-Accuracy approximation rate $\eta \in [\frac{1}{ia_{\mathcal{D}}(\mathcal{C})}, 1]$. \mathcal{B} is IP-Accuracy η -Approximate Reduct of C if it satisfies following Items (1) (2) or (1) (3) regarding IP-Accuracy η -approximation preservation.

- $\begin{array}{ll} (1) & er_{ia_{\mathcal{D}}}(\mathcal{B}) \geq \eta; \\ (2) & er_{ia_{\mathcal{D}}}(\mathcal{B}') < \eta, \ \forall \mathcal{B}' \subset \mathcal{B}; \\ (3) & er_{ia_{\mathcal{D}}}(\mathcal{B} \{b\}) < \eta, \ \forall b \in \mathcal{B}. \end{array}$

Herein, $Core_{ia}^{\eta}(\mathcal{C}) = \{c \in \mathcal{C} : er_{ia_{\mathcal{D}}}(\mathcal{C} - \{c\}) < \eta\}$ and $Red_{ia}^{\eta}(\mathcal{C})$ denote IP-Accuracy η -Approximation Reduct's core and set, respectively.

Tolerance rate δ first provides the tolerance threshold for objective discrepancy rate $dr_{ia_{\mathcal{D}}}(\mathcal{C}_*)$; thus, IP-Accuracy δ -tolerance preservation (i.e., $dr_{ia_{\mathcal{D}}}(\mathcal{C}_*) \in [0, \delta]$) becomes the reduction target to determine IP-Accuracy Tolerant Reduct. Similarly, equality rate $e_{ia_{\tau}}(\mathcal{C}_*)$ holds the objectivity, approximation rate η provides the subjective approximation threshold, and IP-Accuracy η approximation preservation (i.e., $er_{ia_{\mathcal{D}}}(\mathcal{C}_*) \in [\eta, 1]$) becomes the reduction target. In particular, both IP-Accuracy δ -tolerance and η -approximation preservation hold convergence for IP-Accuracy preservation: a sort of quantitative preservation of classification (Theorem 3.3.2); thus, both IP-Accuracy Tolerant and Approximate Reducts are locally enlarged by surrounding the ideal knowledge state of this sort of quantitative preservation. Based on the above analyses, both reducts hold not only the directness regarding the qualitative reduction target but also the statistical property regarding IP-Accuracy.

Note that IP-Accuracy Tolerant/Approximate Reduct adheres to MT-Preservation Reduct (Definition 4.1.6). In fact, IP-Accuracy δ tolerance also has monotonicity, which comes from the discrepancy rate's monotonicity (Theorem 4.3.5). In the lattice sense, if $C_2 \subseteq C_1$ then $(dr_{ia_D}(C_2) \in [0, \delta]) \preceq (dr_{ia_D}(C_1) \in [0, \delta])$, which accords with $dr_{ia_{\mathcal{D}}}(\mathcal{C}_2) \ge dr_{ia_{\mathcal{D}}}(\mathcal{C}_1)$ (Formula (34)). Moreover, IP-Accuracy η approximation's monotonicity could be similarly analyzed.

Theorem 4.3.8. IP-Accuracy Tolerant and Approximate Reducts become equivalent when $\delta + \eta = 1$.

The two quantitative reducts exhibit clear definition duality and hold the equivalent transformation when $\delta + \eta = 1$. Thus, IP-Accuracy Tolerant Reduct is mainly used as a representative.

Theorem 4.3.9 (Quantitative Reduct Features).

- (1) Approach: IP-Accuracy δ -Tolerant (η -Approximate) Reduct approaches IP-Accuracy Reduct, when $\delta \rightarrow 0$ ($\eta \rightarrow 1$).
- (2) Expansion: IP-Accuracy δ -Tolerant (η -Approximate) Reduct expands IP-Accuracy Reduct and could degenerate into the latter when $\delta = 0$ ($\eta = 1$).
- (3) Weakening: IP-Accuracy Tolerant (Approximate) Reduct is weaker than IP-Accuracy Reduct.

Proof. (1) Item (1) is natural and could also be verified by Theorem 4.3.2 and Fig. 4. (2) According to Monotonicity V (Theorem 4.3.5), when $\delta = 0$ ($\eta = 1$), IP-Accuracy δ -tolerance (η -approximate) preservation degenerates into IP-Accuracy preservation (i.e., $dr_{ia_{\mathcal{D}}}(\mathcal{C}_*) = 0$). (3) IP-Accuracy preservation (i.e., $dr_{ia_{\mathcal{D}}}(\mathcal{C}_*) = 0$) unidirectionally deduces IP-Accuracy δ -tolerance (η -approximate) preservation.

Corollary 4.3.10.

- (1) $Core_{ia}^{\delta}(\mathcal{C}), Core_{ia}^{\eta}(\mathcal{C}) \subseteq Core_{ia}(\mathcal{C});$
- (2) $\forall \mathcal{B}_{ia} \in \operatorname{Red}_{ia}^{\delta}(\mathcal{C}), \ \exists \mathcal{B}_{ia}^{\delta} \in \operatorname{Red}_{ia}^{\delta}(\mathcal{C}), \ \exists \mathcal{B}_{ia}^{\eta} \in \operatorname{Red}_{ia}^{\eta}(\mathcal{C}), \ s.t., \ \mathcal{B}_{ia}^{\delta}, \ \mathcal{B}_{ia}^{\eta} \subseteq \mathcal{B}_{ia}.$

IP-Accuracy Tolerant and Approximate Reducts originate from IP-Accuracy Reduct by introducing the applicable tolerance and approximation, respectively. Theorem 4.3.9 summarizes both quantitative reducts' three fundamental features regarding qualitative IP-Accuracy Reduct, i.e., the approach, expansion, and weakening; in fact, they correspond to $\delta \to 0$ ($\eta \to 1$), $\delta = 0$ ($\eta = 1$), $\delta > 0$ ($\eta < 1$), respectively. Thus, quantitative reducts also exhibit directness regarding the qualitative reduct in a statistical way. Therefore, when compared to qualitative IP-Accuracy Reduct, IP-Accuracy Tolerant and Approximate Reducts become two essential types of quantitative reducts, thus holding efficient and extensive quantitative applications. In particular, Corollary 4.3.10 endows the above weakening feature with the hierarchical description. More hierarchical results of quantitative reducts will be intensively reflected in Section 4.4.

For quantitative reducts' heuristic algorithms, a representative is provided for calculating IP-Accuracy Tolerant Reduct. Clearly, the discrepancy rate could be chosen to construct key heuristic information:

$$sig_{ia}^{\delta}(c, \mathcal{C}_*; \mathcal{D}) = dr_{ia_{\mathcal{D}}}(\mathcal{C}_*) - dr_{ia_{\mathcal{D}}}(\mathcal{C}_* - \{c\}).$$
(35)

Thus, we naturally develop the following Algorithm 2, which accords with the previous Algorithm 1. Herein, attribute's sorting and choice based on $sig_{ia}^{\delta}(c, C_*; D)$ are essentially equivalent to those based on $sig_{ia}(c, C_*; D)$, in view of the order-preservation of relevant granulation monotonicity.

Algorithm 2. IP-Accuracy discrepancy rate-based heuristic algorithm for IP-Accuracy Tolerant Reduct

```
Input: D-Table (U, C \cup D) and tolerance rate \delta;

Output: IP-Accuracy Tolerant Reduct B_{ia}^{\delta} \in Red_{ia}^{\delta}(C).

1: Compute Core_{ia}^{\delta}(C).

2: B_{ia}^{\delta} = Core_{ia}^{\delta}(C).

3: while dr_{ia_{0}}(B_{ia}^{\delta}) > \delta do

4: \forall c \in C - B_{ia}^{\delta}, calculate sig_{ia}^{\delta}(c, B_{ia}^{\delta}, D);

choose c_{0} = \arg\max_{ccC - B_{ia}^{\delta}}sig_{ia}^{\delta}(c, B_{ia}^{\delta}; D);

let B_{ia}^{\delta} = B_{ia}^{\delta} \cup \{c_{0}\}.

5: end while

6: return B_{ia}^{\delta}.
```

By simulating the above quantitative reducts regarding IP-Accuracy, we could similarly study the quantitative reducts regarding IIP-Accuracy, which more approach Pawlak-Reduct. Within the IIP-Accuracy framework, the discrepancy and equality rates become $dr_{lia}(C_*)$ and $er_{lia}(C_*)$, while the tolerance and approximation rate thresholds δ', η' are usually in the range [0, 1]. The naturally defined IIP-Accuracy Tolerant and Approximate Reducts also approach/expand/weaken IIP-Accuracy Reduct. In particular, the two quantitative reducts mainly surround the classificationqualitative preservation (Theorem 3.3.7). Finally, we provide the reduction hierarchy for the IIP-Accuracy and IP-Accuracy quantitative reducts.

Proposition 4.3.11. *IIP-Accuracy Tolerant/Approximate Reduct is usually weaker than IP-Accuracy Tolerant/Approximate Reduct, especially for appropriate tolerance/approximation rate thresholds.*

4.4. Importance-accuracy reduct hierarchies

In Section 4.2, we provide two qualitative reduct types (IP-Accuracy/IIP-Accuracy Reduct) and their hierarchies regarding qualitative Pawlak-Reduct. In Section 4.3, we provide four quantitative reduct types (IP-Accuracy/IIP-Accuracy Tolerant/Approximate Reduct) and their hierarchies regarding qualitative IP-Accuracy/IIP-Accuracy Reduct. All of these constructional reducts belong to MT-Preservation Reduct because the underlying target holds granulation monotonicity; in particular, they hold some statistical systematicness because IP-Accuracy/IIP-Accuracy holds statistical systematicness; moreover, Tolerant/Approximate Reduct holds the directness. In this section, these reduct types' hierarchy relationships are summarized. For this purpose, the above metrical and regional reduction fruits are first utilized to provide the following reduct hierarchy figure regarding IP-Accuracy, i.e., Fig. 5.

Next, Fig. 5 is utilized to put forth some explanations:

(1) All nine reduct types apply to the extensive multiplecategory case. The qualitative reducts concern five reduct types (in the bottom half), which completely correspond to the previous preservation deduction targets (Fig. 2); herein, the set-region preservation reduct and distributional preservation reduct are naturally defined. In contrast, the quantitative reducts concern four reduct types (in the top half), which are provided in Section 4.3.

- (2) Arrows represent the reduction hierarchy relation ⇒ (from reduction strengthening to weakening), which correspond to the deduction relation of reduction preservation targets. Thus, all of the reduct hierarchies become clear. In particular, IP-Accuracy Reduct and IIP-Accuracy Tolerant/Approximate Reduct become the strongest and weakest reduct types, respectively. To realize the hierarchical transformation of both, there are actually three different approaches, including IP-Accuracy Reduct ⇒ IP-Accuracy Tolerant Reduct ⇒ IIP-Accuracy Approximate Reduct.
- (3) Reduct hierarchies closely adhere to GrC. In fact, different hierarchical reducts correspond to different results regarding the granular scale, presentative precision, and applicable generalization. Thus, IP-Accuracy Reduct exhibits the finest scale, the highest formal precision, and the narrowest generalization, while IIP-Accuracy Tolerant/Approximate Reduct exhibits the relevant opposites.

According to hierarchical reduction theory [74,75], Fig. 5 could be further utilized to make hierarchical descriptions regarding the core and reduct, such as those in Corollary 4.3.10. Note that a strong reduct can provide some guidance for seeking weak reducts. Thus, the strongest IP-Accuracy Reduct is utilized to finally develop a hierarchical heuristic algorithm for searching IIP-Accuracy Approximate Reduct, which becomes Algorithm 3.

Algorithm 3. IP-Accuracy Reduct-based heuristic algorithm for IIP-Accuracy Approximate Reduct

Input: D-Table $(U, C \cup D)$, approximation rate threshold η' ;	
Output: IIP-Accuracy Approximate Reduct $\mathcal{B}_{lia}^{\eta'} \in Red_{lia}^{\eta'}(\mathcal{C})$.	
1: Give IP-Accuracy Reduct \mathcal{B}_{ia} .	
2: Yield $Core_{lia}^{\eta'}(\mathcal{C})$.	
3: Search $\mathcal{B}_{lia}^{\eta'}$ in range $Core_{lia}^{\eta'}(\mathcal{C}) \subseteq \mathcal{B}_{lia}^{\eta'} \subseteq \mathcal{B}_{ia}$ to satisfy IIP-	
Accuracy η' -approximation preservation.	
4: return $\mathcal{B}_{lia}^{\eta'}$.	

In Algorithm 3, Step 1 provides a strong IP-Accuracy Reduct for the hierarchical heuristics, Step 2 yields the core of weak IIP-Accuracy Approximate Reduct, and Step 3 performs the hierarchical heuristic search to gain the weak target: IIP-Accuracy η' approximation preservation. As a result, this algorithm is convergent and effective and could achieve all IP-Accuracy Approximate Reducts included in the given B_{ia} (regardless of independence/maximality). In particular, the hierarchical heuristic algorithm sharply contrasts with the metrical heuristic algorithm (e.g., Algorithm 2). The two algorithm types mainly differ in the heuristic formation respect; for heuristic development, the hierarchical structure and monotonous measure are utilized, respectively. Moreover, the two types of heuristic formation could be synthetically utilized to develop more efficient algorithms; for example, Algorithm 3 could be efficiently improved by introducing IIP-Accuracy into Step 3's search.

5. A decision table example

D-Table usually originates from a real-world example, and the relevant RS-Theory approach could depend on D-Table to bring

extensive practical application performance. In this section, we mainly resort to D-Table to illustrate this paper's approach. For convenience, the granule-statistical information and knowledge-structural hierarchy are provided after basic processing; thus, they are directly utilized to illustrate the three-way measures and fusional IP-Accuracy, IP-Accuracy's GrC construction and granulation monotonicity, qualitative/quantitative reducts and their hierarchies.

Example 2. For D-Table $(U, C \cup D)$, the granule-statistical information regarding knowledge *C* and concept *X* is provided in Table 3. Herein, |U| = 100, |U/C| = n = 12, |X| = 40.

Table 3 focuses on Bottom-System (U, C, X) via eight columns. Thus, these columns' data are concretely utilized to illustrate the relevant subjects. First, the 1st column provides twelve granules' serial number (i.e., *i* in $[x]_{C}^{i}$), which comes from the interaction cardinality ordering in the 2nd column. The 4th, 5th, 6th columns exhibit generality, importance, and accuracy, respectively. The threeway measures completely correspond to three-way probabilities (i.e., the prior, posterior, and likelihood probabilities); in fact, they radically originate from and equivalently correspond to the two-initial data: the interaction and granule cardinalities (in the 2nd and 3rd columns). Moreover, the 7th column provides IP-Accuracy, which embodies the corresponding data product of 5th and 6th columns.

Accuracy usually acts as the previous single causality measure, but exhibits only the relative causality. In this paper, importance with absolute causality is introduced to fuse accuracy, so IP-Accuracy is further mined to represent synthetic causality. Herein, IP-Accuracy's improvements are explained based on accuracy.

- (1) There are two specific groups of granules with the same accuracy, i.e., $[x]_C^1$, $[x]_C^4$ with accuracy 0.5 and $[x]_C^6$, $[x]_C^{10}$ with 1.0. In the usual accuracy environment, $[x]_L^1$ and $[x]_C^4$ cannot be discriminated, so they are treated equally in the same degree. However, they play different roles in the causality applications (such as rule reasoning) because $i_X([x]_C^1) > i_X([x]_C^4)$. For causality, $[x]_C^1$ actually holds more absolute contributions (based on importance) than $[x]_C^4$ and therefore should gain more systematic attentions and synthetic availability/reliability. Furthermore, the fusional IP-Accuracy could provide a synthetic causality estimation, and $[x]_C^1$ receives high reasoning trust because $ia_X([x]_C^1) > ia_X([x]_C^6)$ and $[x]_C^{10}$ also achieves a similar reasonable result.
- (2) When compared to $[x]_{C}^{7}$ and $[x]_{C}^{8}$, $[x]_{C}^{1}$ and $[x]_{C}^{2}$ depend on their higher importance and IP-Accuracy to gain higher attention and trust, though their accuracy is lower. As a result, $[x]_{C}^{1} \Rightarrow X$ and $[x]_{C}^{2} \Rightarrow X$ become more reliable than $[x]_{C}^{7} \Rightarrow X$ and $[x]_{C}^{8} \Rightarrow X$, especially from the systematic and synthetic viewpoint.

The reasonable results above reflect IP-Accuracy's development and scientificity as derived from the usual single accuracy causality. The radical reason is the double-quantitative fusion of both causality measures, where accuracy and importance exhibit heterogeneity and complementarity.

In this example, the granule-statistical information comes from Bottom-System (U, C, X) and is thus the finest one. Herein, $C = \{a, b, c, d, e, f\}, U/D = \{X, \neg X\}$; thus, $(n, m) = (12, 2), (|X|, |\neg X|) = (40, 60)$. Next, Fig. 6 provides the relevant knowledge-structural hierarchy. This hierarchy figure evolutes the initial large

D-Table and thus becomes smaller and simpler. Herein, attributes a, b, c have more complete descriptions, while attributes d, e, f provide only some auxiliary functions.

First, this hierarchy figure's construction is simply explained.

- For Kn-Coarsening from C, the twelve granules regarding C are the finest, so they establish a basis for the knowledge-structural description. In particular, they depend on their accuracy to exhibit the following ordering: (6) (10) (7) (8) (2) (1) (4) (3) (9) (5) (11) (12); thus, they are represented by twelve serial points.
- (2) Each attribute subset corresponds to a knowledge-structure, and the latter implies Gr-Merging and Gr-Preservation for the twelve granules. Thus, for each attribute subset and its knowledge-structure, small rectangles (containing serial points) mark Gr-Merging, and two vertical lines are used to define X, $\neg X$.
- (3) Knowledge's coarsening relation is represented by the arrow to connect all knowledge-structures. Thus, this hierarchy figure is formed. In particular, the bottom half has part Kn-Coarsening regarding $\{a, b, c, d, e, f\} \Rightarrow \{a, b, c, d, e\} \Rightarrow \{a, b, c, d\} \Rightarrow \{a, b, c\}$; in contrast, the top half has all Kn-Coarsening regarding $\{a, b, c\}$'s subsets.

In Fig. 6, each piece of knowledge can rapidly achieve its structural granule and hierarchical position. In fact, this hierarchy is highly related to the reduction target structure [74], so it becomes enough for the next knowledge-based GrC discussion and reduction analyses.

Then, we calculate each knowledge's IP-Accuracy, IIP-Accuracy, discrepancy and equality rates. Note that this calculation process radically reflects the relevant GrC construction, which is related to the sum integration. The computational results are shown in Table 4. Herein, we provide the presentative calculation process regarding knowledge $\{a, b, c\}$, where only five granules exist.

- (1) $ia_X(\mathcal{C}_*) = \frac{4}{40} \times \frac{4}{4} + \frac{5}{40} \times \frac{5}{7} + \frac{25}{40} \times \frac{25}{51} + \frac{6}{40} \times \frac{6}{22} + \frac{0}{40} \times \frac{0}{16} = 0.5366, \quad ia_{-X}(\mathcal{C}_*) = \frac{0}{60} \times \frac{0}{4} + \frac{2}{60} \times \frac{2}{7} + \frac{26}{60} \times \frac{26}{51} + \frac{16}{60} \times \frac{12}{22} + \frac{16}{60} \times \frac{16}{16} = 0.6910; \quad Iia_{-X}(\mathcal{C}_*) = \frac{4}{40} \times \frac{4}{4} = 0.1000, \quad Iia_{-X}(\mathcal{C}_*) = \frac{16}{50} \times \frac{16}{16} = 0.2667.$
- (2) $ia_{\mathcal{D}}^{(\mathcal{C}_*)} = ia_X(\mathcal{C}_*) + ia_{\neg X}(\mathcal{C}_*) = 0.5366 + 0.6910 = 1.2276;$ $lia_{\mathcal{D}}(\mathcal{C}_*) = lia_X(\mathcal{C}_*) + lia_{\neg X}(\mathcal{C}_*) = 0.1000 + 0.2667 = 0.3667.$
- (3) Similar to the above two steps, the initial twelve granules are utilized to obtain: $ia_{\mathcal{D}}(\mathcal{C}) = ia_X(\mathcal{C}) + ia_{\neg X}(\mathcal{C}) = 0.5437 + 0.6958 = 1.2395; \quad lia_{\mathcal{D}}(\mathcal{C}) = lia_X(\mathcal{C}) + lia_{\neg X}(\mathcal{C}) = 0.1000 + 0.2667 = 0.3667.$ The two data regarding $\mathcal{C} 1.2395$ and 0.3667 establish the comparative criterion to calculate the discrepancy and equality rates. Thus, regarding $\{a, b, c\}, dr_{ia_{\mathcal{D}}} = \frac{1.2395 1.2276}{1.2395} = 0.0096 = 0.96\%, er_{ia_{\mathcal{D}}} = 99.04\%; \quad dr_{lia_{\mathcal{D}}} = \frac{0.3667 0.3667}{0.3667} = 0.00\%, er_{lia_{\mathcal{D}}} = 100.00\%,$

Thus far, basic data processing has been completed. Next, Table 4 (as well as Fig. 6) is fully utilized to illustrate IP-Accuracy's relevant GrC terms (Section 3), including the hierarchy construction, granulation monotonicity, equality condition, and preservation deduction.

(1) Kn-Con IP-Accuracy aims to calculate all Gr-Con IP-Accuracy's sum, and this process holds an integration from the granule to knowledge. The 2nd column verifies Kn-Con IP-Accuracy's granulation monotonicity (Theorem 3.2.6), and $ia_X(\mathcal{C}_*)$ changes from maximum 0.5437 to limit 0.4000 in Kn-Coarsening. Herein, the same Kn-Con IP-Accuracy emerges in Kn-Coarsening $\{a, b, c, d, e, f\} \Rightarrow \{a, b, c, d, e\}$. The strict Kn-Coarsening has three groups of Gr-Merging,

and they all satisfy the theoretical equality conditions (Theorem 3.2.6); in particular, Gr-Merging $[x]^1_{\{a,b,c,d,e,f\}} \cup [x]^4_{\{a,b,c,d,e,f\}} \stackrel{=}{\rightarrow} [x]^1_{\{a,b,c,d,e\}}$ is in BND and C-BND, and the three granules concerned have the same accuracy, 0.5. The 3rd column could similarly verify Kn-Con IIP-Accuracy's granulation monotonicity (as well as the equality condition on POS) (Theorem 3.2.10). Thus, Middle-System (U, C, X)'s inspection is finished.

- (2) The 4th column's data become the 2nd column's doubledata sum; clearly, this fact reflects Kn-Cl IP-Accuracy's integration regarding *X*, ¬*X*. The 4th column could verify Kn-Cl IP-Accuracy's granulation monotonicity (Theorem 3.3.2). Concretely, *ia*_D(C_{*}) changes from maximum 1.2395 to limit 1; moreover, Kn-Coarsening {*a*, *b*, *c*, *d*, *e*, *f*} ⇒ {*a*, *b*, *c*, *d*, *e*} shows the theoretical equality conditions. Meanwhile, the 5th column – which comes from the 3rd column's sum – could be used to exhibit Kn-Cl IIP-Accuracy's granulation monotonicity (Theorem 3.3.7); in particular, there are more groups of Gr-Merging to illustrate the equality conditions.
- (3) Top-System (U, C, D) is continuously used to illustrate preservation deduction (Fig. 2). Only $C \Rightarrow \{a, b, c, d, e\}$ exhibits IP-Accuracy preservation. In contrast, $C \Rightarrow \{a, b, c, d, e\}, C \Rightarrow \{a, b, c, d\}, C \Rightarrow \{a, b, c\}, C \Rightarrow \{a, b, c, d\}, C \Rightarrow \{a, b, c\}, C \Rightarrow \{a, c\}$ all hold IIP-Accuracy preservation; they and only they have C-POS preservation. Thus, as is shown by this example, IP-Accuracy preservation unidirectionally deduces IIP-Accuracy preservation (Corollary 3.3.8), while the latter is equivalent to C-POS preservation (Theorem 3.3.7).

Finally, Table 4 (as well as Fig. 6) is utilized to calculate qualitative/quantitative reducts and analyze their hierarchies (Section 4).

- (1) {a, b, c, d, e} has IP-Accuracy preservation and has no proper subsets to satisfy the target; hence, it becomes the sole IP-Accuracy reduct. Only {a, b, c, d, e}, {a, b, c, d}, {a, b, c}, {a, c} exhibit IIP-Accuracy preservation; hence, the smallest {a, c} becomes the sole IIP-Accuracy reduct. For this two-category example, the three types of regional and qualitative reducts (i.e., the set-region preservation reduct, distributional preservation reduct, and Pawlak-Reduct) become equivalent, and they are further equivalent to IIP-Accuracy Reduct; hence, they also have the sole reduct {a, c}.
- (2) The 6th and 7th columns provide the discrepancy and equality rates regarding IP-Accuracy and IIP-Accuracy, respectively. Thus, the granulation monotonicity and granulation boundedness are verified for relevant rates, e.g., $dr_{ia_{T}}$ is monotonous in [0.00%, 19.32%]. Let $\delta = 2.5\% \in [0.00\%, 19.32\%]$ and $\eta = 97.5\% \in [80.68\%, 100.00\%]$. Thus, only $\{a, b, c, d, e\}$, $\{a, b, c, d\}$, $\{a, b, c\}$, $\{b, c\}$ exhibit IIP-Accuracy δ -tolerance preservation; hence, the smallest $\{b, c\}$ becomes the sole IP-Accuracy Approximate Reduct. Similarly, let $\delta' = 30\% \in [0\%, 100\%]$ and $\eta' = 70\% \in [0\%, 100\%]$; then, $\{c\}$ becomes the sole IIP-Accuracy Tolerance/Approximate Reduct.
- (3) In particular, all reduct results are appended to Fig. 5. Thus, these relevant reducts naturally verify the relevant reduct hierarchies, which are exactly and completely in Fig. 5. For example, $\{a, b, c, d, e\} \supseteq \{a, c\}, \{b, c\}$ could verify Theorems 4.2.2, 4.2.5, and Corollary 4.3.10. Finally, the hierarchical heuristic Algorithm 3 is appropriately illustrated. Concretely, the strongest IP-Accuracy Reduct $\{a, b, c, d, e\}$ could

provide the heuristic information regarding the internal inclusion to quickly search the weakest IIP-Accuracy Approximate Reduct $\{c\}$. \Box

6. Conclusions

Aiming at uncertainty measure mining and applications, this paper conducts systematic double-quantitative fusion of accuracy and importance to mine IP-Accuracy and further explores IP-Accuracy's granular integration construction and attribute reduction applications. Herein, IP-Accuracy's benign mining adopts the double-quantitative fusion strategy, while the later integration construction and attribute reduction mainly utilize the GrC technology. As a result, three-way probabilities and measures are systemically developed, causality-based IP-Accuracy is synthetically mined, a fundamental GrC platform is benignly constructed, and a hierarchical reduction system is fully established.

By resorting to a monotonous uncertainty measure, this study provides an integration-evolutionary GrC construction strategy for attribute reduction. Note that IP-Accuracy and its reduction hold benign monotonicity and statistical systematicness. As a result, the innovative tolerant and approximate reducts, in a direct and statistical way, quantitatively approach/expand/weaken the ideal qualitative reduct; thus, they efficiently and extensively apply to quantitative environments, such as those with data noise. Moreover, MT-Preservation Reduct provides a unified normal pattern for monotonicity targets, thus holding great generalization significance.

Several problems are retained for in-depth explorations.

- (1) The uncertainty measure fusion is fundamental for the uncertainty presentation and measurement. As is noted in Section 2.1, Approx-Space's metrics usually play a dual role for the probability statistics and application semantics. However, only their semantics statuses are utilized for IP-Accuracy's fusion mining and integration construction. Thus, their statistics positions and relevant uncertainty fusions are worth exploring further. In other words, the metrical mining and fusional application could proceed from the statistics viewpoint and even from the synthetic viewpoint.
- (2) IP-Accuracy becomes the central uncertainty measure. It is mined mainly by a sort of system detection for three-way measures. In fact, accuracy and importance act as main causality factors in Bottom-System, and their benign fusion benefits from the weight function and synthetic semantics. As a result, fusional IP-Accuracy holds a non-linear promotion, and its evolutionary hierarchical form holds granulation monotonicity in both Middle-System and Top-System. Thus, the surplus absolute generality is worth introducing to mine more monotonous uncertainty measures. In other words, doublequantification work needs to be carried out in an in-depth manner.
- (3) In this paper, IP-Accuracy's GrC construction becomes one of the main contributions, and relevant granulation monotonicity underlies the in-depth attribute reduction applications. To establish attribute reduction, we actually provide a mature granular integration technology – an effective bottom-up strategy – for the evolutionary GrC construction. Thus, this hierarchical promotion approach is worth generalizing for extensive monotonous uncertainty measures, including existing and potential measures. In particular,

when introduced into RS-Theory, classical entropy never has granulation monotonicity [76]; in contrast, weighted entropy could perform a granular integration strategy to construct the relevant conditional entropy and mutual information and to acquire final granulation monotonicity [76].

- (4) IP-Accuracy attribute reduction is systematically investigated; herein, the tolerant and approximate reducts – two novel quantitative reducts – have wide potential for development in theory and in application. This metrical approach holds directness and statistics, so it differs from the model-regional method; thus, it is worth comparing to the latter, especially regarding quantitative performance. In particular, the illustrative example concerns only the two basic categories. Thus, further practical verification (especially regarding extensive multiple categories) is needed for both the efficiency of the relevant granulation monotonicity, reduct hierarchy, and heuristic algorithm and applicability of the tolerant/approximate reducts. Moreover, MT-Preservation Reduct is worth exploring further.
- (5) This paper is devoted to the application of basic research, and the primary approach and theoretical contribution are thoroughly illustrated by a D-Table example based on granulestatistical information and the knowledge-structural hierarchy. In fact, offering a detailed, more real-world example could highlight the essence and relevance of our approach. In other words, it is worthwhile to verify the content of this paper with an exact real-world example, especially to illustrate the inherent effectiveness and efficiency. Furthermore, it is worth applying in a practical environment.

Based on the work in this paper and taking into account the above issues, we are engaged in follow-up research and have already made some progress, including double quantification regarding accuracy and generality, metrical integration regarding the probability distribution, uncertainty fusion regarding weighted entropy, the GrC and reduction construction regarding information entropy and mutual information. Moreover, practical applications are concerned.

Acknowledgements

The authors thank both the editors and the reviewers for their valuable suggestions, which substantially improved this paper.

This work was supported by National Science Foundation of China (61203285 and 61273304), Specialized Research Fund for Doctoral Program of Higher Education of China (20130072130004), Postdoctoral Science Foundation Funded Project of China (2013T60464), and Scientific Research Project of Sichuan Provincial Education Department of China (15ZB0028).

Appendix A. Proof of Lemma 3.2.2

Proof. This lemma is verified by inducing index *k*; moreover, Formula (15) mainly uses the simpler style, i.e., $\frac{a_1}{b_1} = \cdots = \frac{a_k}{b_k}$.

(1) When k = 2, we only prove

$$\frac{a_1^2}{b_1} + \frac{a_2^2}{b_2} \ge \frac{(a_1 + a_2)^2}{b_1 + b_2}.$$
(A.1)

By multiplying common multiple $b_1b_2(b_1 + b_2)$, this goal inequality becomes:

$$a_{1}^{2}b_{2}(b_{1}+b_{2}) + a_{2}^{2}b_{1}(b_{1}+b_{2}) \ge b_{1}b_{2}(a_{1}+a_{2})^{2},$$

i.e., $a_{1}^{2}b_{1}b_{2} + a_{1}^{2}b_{2}^{2} + a_{2}^{2}b_{1}^{2} + a_{2}^{2}b_{1}b_{2}$
$$\ge b_{1}b_{2}a_{1}^{2} + 2b_{1}b_{2}a_{1}a_{2} + b_{1}b_{2}a_{2}^{2},$$

i.e., $a_{1}^{2}b_{2}^{2} + a_{2}^{2}b_{1}^{2} \ge 2b_{1}b_{2}a_{1}a_{2},$
i.e., $(a_{1}b_{2} - a_{2}b_{1})^{2} \ge 0.$

Thus, Formula (14) holds when k = 2. Moreover, it degenerates into an equality, iff $\frac{a_1}{b_1} = \frac{a_2}{b_2}$; in other words, Formula (15) also holds.

(2) Suppose Formulas (14) and (15) hold for $k = k_0$. Thus,

$$\sum_{l=1}^{k_0} \frac{a_l^2}{b_l} \ge \frac{\left(\sum_{t=1}^{k_0} a_t\right)^2}{\sum_{t=1}^{k_0} b_t};$$
(A.2)

moreover, the relevant equality holds, iff $\frac{a_1}{b_1} = \cdots = \frac{a_{k_0}}{b_{k_0}}$. Then, consider case $k = k_0 + 1$. Thus,

$$\begin{split} \sum_{t=1}^{k_0+1} \frac{a_t^2}{b_t} &= \sum_{t=1}^{k_0} \frac{a_t^2}{b_t} + \frac{a_{k_0+1}^2}{b_{k_0+1}} \geq \frac{\left(\sum_{t=1}^{k_0} a_t\right)^2}{\sum_{t=1}^{k_0} b_t} + \frac{a_{k_0+1}^2}{b_{k_0+1}} \\ &\geq \frac{\left(\sum_{t=1}^{k_0} a_t + a_{k_0+1}\right)^2}{\sum_{t=1}^{k_0} b_t + b_{k_0+1}} = \frac{\left(\sum_{t=1}^{k_0+1} a_t\right)^2}{\sum_{t=1}^{k_0+1} b_t}. \end{split}$$

The above two inequality signs regarding $k = k_0 + 1$ mainly come from the relevant results regarding $k = k_0$ and k = 2 (i.e., Formulas (A.2) and (A.1)), respectively; thus, Formula (14) holds. Moreover, the above two quality signs are reached,

iff
$$\frac{a_1}{b_1} = \dots = \frac{a_{k_0}}{b_{k_0}}, \quad \frac{\sum_{t=1}^{n_0} a_t}{\sum_{t=1}^{k_0} b_t} = \frac{a_{k_0+1}}{b_{k_0+1}},$$

iff $\frac{a_1}{b_1} = \dots = \frac{a_{k_0}}{b_{k_0}} = \frac{a_{k_0+1}}{b_{k_0+1}};$

 $-\nu$

thus, Formula (15) holds.

Therefore, Formulas (14) and (15) hold via the above induction. \Box

Appendix B. Proof of Proposition 3.2.9

Proof.

$$\begin{aligned} lia_{X}(\mathcal{C}_{*}) &= \sum_{[x]_{\mathcal{C}_{*}} \subseteq \text{POS}_{\mathcal{C}_{*}}(X)} i_{X}([x]_{\mathcal{C}_{*}}) \times a_{X}([x]_{\mathcal{C}_{*}}) \\ &= \sum_{[x]_{\mathcal{C}_{*}} \subseteq \text{POS}_{\mathcal{C}_{*}}(X)} \frac{|[x]_{\mathcal{C}_{*}}|}{|X|} \times 1 = \frac{|\text{POS}_{\mathcal{C}_{*}}(X)|}{|X|}; \\ Eia_{X}(\mathcal{C}_{*}) &= \sum_{[x]_{\mathcal{C}_{*}} \subseteq \text{NEG}_{\mathcal{C}_{*}}(X)} i_{X}([x]_{\mathcal{C}_{*}}) \times a_{X}([x]_{\mathcal{C}_{*}}) \\ &= \sum_{[x]_{\mathcal{C}_{*}} \subseteq \text{NEG}_{\mathcal{C}_{*}}(X)} i_{X}([x]_{\mathcal{C}_{*}}) \times 0 = 0. \end{aligned}$$

Thus, Items (1) (2) (4) are proved. Moreover, Item (3) could be proved by the equivalence between $[x]_{C_*} \cap X, \neg X \neq \emptyset$ and $[x]_{C_*} \subseteq BND_{\mathcal{C}_*}(X)$. \Box

Appendix C. Proof of Lemma 4.1.5

Proof.

(1) $\forall b \in \mathcal{B}$, let $\mathcal{B}' = \mathcal{B} - \{b\} \subset \mathcal{B}$. According to Item (1), $MT(\mathcal{B} - \{b\}) \neq MT(\mathcal{B})$, i.e., Item (2) holds.

(2) $\forall \mathcal{B}' \subset \mathcal{B}$, then $\exists b \in \mathcal{B} - \mathcal{B}' \subset \mathcal{B}$, s.t., $\mathcal{B}' \subseteq \mathcal{B} - \{b\} \subset \mathcal{B}$. According to MT granulation monotonicity (Definition 4.1.4) and Item (2), $MT(\mathcal{B}') \preceq MT(\mathcal{B} - \{b\}) \prec MT(\mathcal{B})$, so $MT(\mathcal{B}') \neq MT(\mathcal{B})$, i.e., Item (1) holds. \Box

Appendix D. Proof of Proposition 4.1.7

Proof.

- (1) If $c \notin \bigcap_{\mathcal{B} \in Red_{MT}(\mathcal{C})} \mathcal{B}$, then $\exists \mathcal{B} \in Red_{MT}(\mathcal{C})$ but $c \notin \mathcal{B}$. Thus, $MT(\mathcal{B}) = MT(\mathcal{C})$ and $\mathcal{B} \subseteq \mathcal{C} - \{c\} \subset \mathcal{C}$. According to MT granulation monotonicity (Definition 4.1.4), $MT(\mathcal{C} - \{c\}) = MT(\mathcal{C})$, so $c \notin Core_{MT}(\mathcal{C})$. Hence, $Core_{MT}(\mathcal{C}) \subseteq \bigcap_{\mathcal{B} \in Red_{MT}(\mathcal{C})} \mathcal{B}$.
- (2) If $c \notin Core_{MT}(\mathcal{C})$, then $MT(\mathcal{C} \{c\}) = MT(\mathcal{C})$; hence, exist MT-Preservation Reduct \mathcal{B} in $\mathcal{C} - \{c\}$. Thus, $MT(\mathcal{B}) = MT(\mathcal{C} - \{c\})$, and \mathcal{B} has the MT independence (i.e., Lemma 4.1.5's Item (2)) in $\mathcal{C} - \{c\}$. Furthermore, $MT(\mathcal{B}) = MT(\mathcal{C})$ and \mathcal{B} has the MT independence in \mathcal{C} , so $\mathcal{B} \in Red_{MT}(\mathcal{C})$. However, $c \notin \mathcal{B}$ because $\mathcal{B} \subseteq \mathcal{C} - \{c\}$. Hence, $c \notin \bigcap_{\mathcal{B} \in Red_{MT}(\mathcal{C})} \mathcal{B}, \bigcap_{\mathcal{B} \in Red_{MT}(\mathcal{C})} \mathcal{B} \subseteq$ $Red_{MT}(\mathcal{C})$. \Box

References

- N. Azam, J.T. Yao, Analyzing uncertainties of probabilistic rough set regions with game-theoretic rough sets, Int. J. Approx. Reason. 55 (1) (2014) 142–155.
- [2] N. Azam, J.T. Yao, Interpretation of equilibria in game-theoretic rough sets, Inf. Sci. 295 (2015) 586–599.
- [3] M.E. Cornejo, J. Medina, E. Ramirez-Poussa, Attribute reduction in multi-adjoint concept lattices, Inf. Sci. 294 (2015) 41–56.
- [4] J.H. Dai, W.T. Wang, J.S. Mi, Uncertainty measurement for interval-valued information systems, Inf. Sci. 251 (2013) 63–78.
- [5] X.F. Deng, Y.Y. Yao, Decision-theoretic three-way approximations of fuzzy sets, Inf. Sci. 279 (2014) 702–715.
- [6] S. Greco, B. Matarazzo, R. Slowinski, Parameterized rough set model using rough membership and bayesian confirmation measures, Int. J. Approx. Reason. 49 (2) (2008) 285–300.
- [7] B.Q. Hu, Three-way decisions space and three-way decisions, Inf. Sci. 281 (2014) 21-52.
- [8] Q.H. Hu, L. Zhang, D.G. Chen, W. Pedrycz, D.R. Yu, Gaussian kernel based fuzzy rough sets: model, uncertainty measures and applications, Int. J. Approx. Reason. 51 (4) (2010) 453–471.
- [9] B. Huang, C.X. Guo, Y.J. Zhuang, H.X. Li, X.Z. Zhou, Intuitionistic fuzzy multigranulation rough sets, Inf. Sci. 277 (2014) 299–320.
- [10] X.Y. Jia, W.H. Liao, Z.M. Tang, L. Shang, Minimum cost attribute reduction in decision-theoretic rough set models, Inf. Sci. 219 (10) (2013) 151–167.
- [11] X.Y. Jia, L. Shang, B. Zhou, Y.Y. Yao, Generalized attribute reduct in rough set theory, Knowl.-Based Syst. (2015), doi:10.1016/j.knosys.2015.05.017.
- [12] X.Y. Jia, Z.M. Tang, W.H. Liao, L. Shang, On an optimization representation of decision-theoretic rough set model, Int. J. Approx. Reason. 55 (1) (2014) 156–166.
- [13] F. Li, M. Ye, X.D. Chen, An extension to rough c-means clustering based on decision-theoretic rough set model, Int. J. Approx. Reason. 55 (1) (2014) 116–129.
- [14] H.X. Li, L.B. Zhang, B. Huang, X.Z. Zhou, Sequential three-way decision and granulation for cost-sensitive face recognition, Knowl.-Based Syst. (2015), doi:10.1016/ j.knosys.2015.07.040.
- [15] W.T. Li, W.H. Xu, Double-quantitative decision-theoretic rough set, Inf. Sci. 316 (2015) 54–67.
- [16] W.T. Li, W.H. Xu, Multigranulation decision-theoretic rough set in ordered information system, Fundam. Inform. 139 (2015) 1–23.
- [17] D.C. Liang, D. Liu, Deriving three-way decisions from intuitionistic fuzzy decisiontheoretic rough sets, Inf. Sci. 300 (2015) 28–48.
- [18] D.C. Liang, D. Liu, Systematic studies on three-way decisions with interval-valued decision-theoretic rough sets, Inf. Sci. 276 (2014) 186–203.
- [19] D.C. Liang, W. Pedrycz, D. Liu, P. Hu, Three-way decisions based on decisiontheoretic rough sets under linguistic assessment with the aid of group decision making, Appl. Soft Comput. 29 (2015) 256–269.
- [20] J.Y. Liang, F. Wang, C.Y. Dang, Y.H. Qian, A group incremental approach to feature selection applying rough set technique, IEEE Trans. Knowl. Data Eng. 26 (2) (2014) 294–308.
- [21] J.Y. Liang, J.H. Wang, Y.H. Qian, A new measure of uncertainty based on knowledge granulation for rough sets, Inf. Sci. 179 (4) (2009) 458–470.
- [22] T.Y. Lin, Granular computing: from rough sets and neighborhood systems to information granulation and computing with words, in: European Congress on Intelligent Techniques and Soft Computing, 1997, pp. 1602–1606.

- [23] P. Lingras, M. Chen, D.Q. Miao, Rough cluster quality index based on decision theory, IEEE Trans. Knowl. Data Eng. 21 (7) (2009) 1014–1026.
- [24] P. Lingras, M. Chen, D.Q. Miao, Semi-supervised rough cost/benefit decisions, Fundam. Inform. 94 (2) (2009) 233–244.
- [25] C.H. Liu, D.Q. Miao, N. Zhang, C. Gao, Graded rough set model based on two universes and its properties, Knowl.-Based Syst. 33 (2012) 65–72.
- [26] D. Liu, T.R. Li, D.C. Liang, Incorporating logistic regression to decisiontheoretic rough sets for classifications, Int. J. Approx. Reason. 55 (1) (2014) 197–210.
- [27] D. Liu, T.R. Li, D. Ruan, Probabilistic model criteria with decision-theoretic rough sets, Inf. Sci. 181 (2011) 3709–3722.
- [28] D. Liu, T.R. Li, J.B. Zhang, Incremental updating approximations in probabilistic rough sets under the variation of attributes, Knowl.-Based Syst. 73 (2015) 81–96.
- [29] D. Liu, D.C. Liang, C.C. Wang, A novel three-way decision model based on incomplete information system, Knowl.-Based Syst. (2015), doi:10.1016/j.knosys.2015. 07.036.
- [30] W.M. Ma, B.Z. Sun, Probabilistic rough set over two universes and rough entropy, Int. J. Approx. Reason. 53 (4) (2012) 608–619.
- [31] X.A. Ma, G.Y. Wang, H. Yu, T.R. Li, Decision region distribution preservation reduction in decision-theoretic rough set model, Inf. Sci. 278 (2014) 614–640.
- [32] J.S. Mi, W.Z. Wu, W.X. Zhang, Approaches to knowledge reduction based on variable precision rough set model, Inf. Sci. 159 (3–4) (2004) 255–272.
- [33] S.K. Pal, B.U. Shankar, P. Mitra, Granular computing, rough entropy and object extraction, Pattern Recogn. Lett. 26 (16) (2005) 2509–2517.
- [34] I.K. Park, G.S. Choi, Rough set approach for clustering categorical data using information-theoretic dependency measure, Inf. Syst. 48 (2015) 289–295.
- [35] Z. Pawlak, Rough sets, Int. J. Inf. Comput. Sci. 11 (5) (1982) 341–356.
- [36] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning About Data, vol. 9, Kluwer Academic Publishers, Dordrecht, 1991.
- [37] Z. Pawlak, A. Skowron, Rough membership functions, Advances in the Dempster-Shafer Theory of Evidence, John Wiley and Sons, New York, 1994, pp. 251–271.
- [38] W. Pedrycz, From fuzzy sets to shadowed sets: interpretation and computing, Int. J. Intell. Syst. 24 (1) (2009) 48–61.
- [39] W. Pedrycz, Granular Computing: Analysis and Design of Intelligent Systems, CRC Press, Francis Taylor, Boca Raton, 2013.
- [40] W. Pedrycz, Shadowed sets: representing and processing fuzzy sets, IEEE Trans. Syst. Man Cybern. Part B: Cybern. 28 (1) (1998) 103–109.
- [41] W. Pedrycz, G. Vukovich, Granular computing with shadowed sets, Int. J. Intell. Syst. 17 (2) (2002) 173–197.
- [42] J. Qian, P. Lv, X.D. Yue, C.H. Liu, Z.J. Jing, Hierarchical attribute reduction algorithms for big data using MapReduce, Knowl.-Based Syst. 73 (2015) 18–31.
- [43] Y.H. Qian, J.Y. Liang, C.Y. Dang, Consistency measure, inclusion degree and fuzzy measure in decision tables, Fuzzy Sets Syst. 159 (18) (2008) 2353–2377.
- [44] Y.H. Qian, J.Y. Liang, W.Z. Wu, C.Y. Dang, Information granularity in fuzzy binary GrC model, IEEE Trans. Fuzzy Syst. 19 (2) (2011) 253–264.
- [45] Y.H. Qian, H. Zhang, Y.L. Sang, J.Y. Liang, Multigranulation decision-theoretic rough sets, Int. J. Approx. Reason. 55 (1) (2014) 225–237.
- [46] A. Skowron, J. Stepaniuk, R. Swiniarski, Modeling rough granular computing based on approximation spaces, Inf. Sci. 184 (1) (2012) 20–43.
- [47] D. Slezak, W. Ziarko, The investigation of the bayesian rough set model, Int. J. Approx. Reason. 40 (1–2) (2005) 81–91.
- [48] B.Z. Sun, W.M. Ma, H.Y. Zhao, Decision-theoretic rough fuzzy set model and application, Inf. Sci. 283 (2014) 180–196.
- [49] L. Sun, J.C. Xu, Y. Tian, Feature selection using rough entropy-based uncertainty measures in incomplete decision systems, Knowl.-Based Syst. 36 (2012) 206–216.
- [50] C.Z. Wang, M.W. Shao, B.Q. Sun, Q.H. Hu, An improved attribute reduction scheme with covering based rough sets, Appl. Soft Comput. 26 (2015) 235–243.
- [51] G.Y. Wang, X.A. Ma, H. Yu, Monotonic uncertainty measures for attribute reduction in probabilistic rough set model, Int. J. Approx. Reason. 59 (2015) 41–67.
- [52] J.Y. Wang, J. Zhou, Research of reduct features in the variable precision rough set model, Neurocomputing 72 (2009) 2643–2648.
- [53] W.H. Xu, J.Z. Pang, S.Q. Luo, A novel cognitive system model and approach to transformation of information granules, Int. J. Approx. Reason. 55 (3) (2014) 853– 866.
- [54] J.T. Yao, A.V. Vasilakos, W. Pedrycz, Granular computing: perspectives and challenges, IEEE Trans. Cybern. 43 (6) (2013) 1977–1989.
- [55] Y.Y. Yao, An outline of a theory of three-way decisions, in: J. Yao, Y. Yang, R. Slowinski, S. Greco, H. Li, S. Mitra, L. Polkowski (Eds.), Proceedings of the 8th International RSCTC Conference, LNCS (LNAI), vol. 7413, 2012, pp. 1–17.
- [56] Y.Y. Yao, Interpreting concept learning in cognitive informatics and granular computing, IEEE Trans. Syst. Man Cybern. Part B: Cybern. 39 (4) (2009) 855–866.
- [57] Y.Y. Yao, Probabilistic rough set approximations, Int. J. Approx. Reason. 49 (2) (2008) 255–271.
- [58] Y.Y. Yao, The superiority of three-way decision in probabilistic rough set models, Inf. Sci. 181 (2011) 1080–1096.
- [59] Y.Y. Yao, The two sides of the theory of rough sets, Knowl.-Based Syst. 80 (2015) 67–77.
- [60] Y.Y. Yao, Three-way decisions with probabilistic rough sets, Inf. Sci. 180 (2010) 341–353.
- [61] Y.Y. Yao, X.F. Deng, Quantitative rough sets based on subsethood measures, Inf. Sci. 267 (2014) 306–322.

- [62] Y.Y. Yao, T.Y. Lin, Generalization of rough sets using modal logics, Intell. Automat. Soft Comput. 2 (2) (1996) 103–120.
- [63] Y.Y. Yao, Y.H. She, Rough set models in multigranulation spaces, Inf. Sci. (2015), doi:10.1016/j.ins.2015.08.011.
- [64] Y.Y. Yao, S.K.M. Wong, P. Lingras, A decision-theoretic rough set model, in: Z.W. Ras, M. Zemankova, M.L. Emrich (Eds.), The 5th International Symposium on Methodologies for Intelligent Systems, North-Holland, New York, 1990, pp. 17–25.
- [65] Y.Y. Yao, Y. Zhao, Attribute reduction in decision-theoretic rough set models, Inf. Sci. 178 (17) (2008) 3356–3373.
- [66] H. Yu, Z.G. Liu, G.Y. Wang, An automatic method to determine the number of clusters using decision-theoretic rough set, Int. J. Approx. Reason. 55 (1) (2014) 101– 115.
- [67] H. Yu, C. Zhang, G.Y. Wang, A tree-based incremental overlapping clustering method using the three-way decision theory, Knowl.-Based Syst. (2015), doi:10. 1016/j.knosys.2015.05.028.
- [68] L.A. Zadeh, Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets Syst. 90 (1997) 111– 127.
- [69] H.Y. Zhang, S.Y. Yang, J.M. Ma, Ranking interval sets based on inclusion measures and applications to three-way decisions, Knowl.-Based Syst. (2015), doi:10.1016/ j.knosys.2015.07.025.
- [70] Q.H. Zhang, J. Wang, G.Y. Wang, H. Yu, The approximation set of a vague set in rough approximation space, Inf. Sci. 300 (2015) 1–19.
- [71] X.H. Zhang, J.H. Dai, Y.C. Yu, On the union and intersection operations of rough sets based on various approximation spaces, Inf. Sci. 292 (2015) 214–229.

- [72] X.Y. Zhang, D.Q. Miao, An expanded double-quantitative model regarding probabilities and grades and its hierarchical double-quantitative attribute reduction, Inf. Sci. 299 (2015) 312–336.
- [73] X.Y. Zhang, D.Q. Miao, Quantitative information architecture, granular computing and rough set models in the double-quantitative approximation space on precision and grade, Inf. Sci. 268 (2014) 147–168.
- [74] X.Y. Zhang, D.Q. Miao, Reduction target structure-based hierarchical attribute reduction for two-category decision-theoretic rough sets, Inf. Sci. 277 (2014) 755– 776.
- [75] X.Y. Zhang, D.Q. Miao, Region-based quantitative and hierarchical attribute reduction in the two-category decision theoretic rough set model, Knowl.-Based Syst. 71 (2014) 146–161.
- [76] X.Y. Zhang, D.Q. Miao, Three-way weighted entropies and three-way attribute reduction, in: D.Q. Miao, W. Pedrycz, D. Slezak, G. Peters, Q.H. Hu, R.Z. Wang (Eds.), Proceedings of the 9th International RSKT Conference, LNCS (LNAI), vol. 8818, 2014, pp. 707–719.
- [77] X.Y. Zhang, D.Q. Miao, Two basic double-quantitative rough set models of precision and grade and their investigation using granular computing, Int. J. Approx. Reason. 54 (8) (2013) 1130–1148.
- [78] X.R. Zhao, B.Q. Hu, Fuzzy and interval-valued fuzzy decision-theoretic rough set approaches based on fuzzy probability measure, Inf. Sci. 298 (2015) 534–554.
- [79] B. Zhou, Y.Y. Yao, J.G. Luo, Cost-sensitive three-way email spam filtering, J. Intell. Inf. Syst. 42 (1) (2014) 19–45.
 - [80] J. Zhou, W. Pedrycz, D.Q. Miao, Shadowed sets in the characterization of roughfuzzy clustering, Pattern Recogn. 44 (8) (2011) 1738–1749.
 - [81] W. Ziarko, Variable precision rough set model, J. Comput. Syst. Sci. 46 (1993) 39–59.