Information Sciences xxx (2013) xxx-xxx

Contents lists available at ScienceDirect

# Information Sciences

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

#### Quantitative information architecture, granular computing and 3 rough set models in the double-quantitative approximation л space of precision and grade 5

### 8 01 Xianyong Zhang<sup>a,b,c,\*</sup>, Duogian Miao<sup>a,b</sup>

9 <sup>a</sup> Department of Computer Science and Technology, Tongii University, Shanghai 201804, PR China

10 <sup>b</sup> The Key Laboratory of Embedded System and Service Computing, Ministry of Education, Shanghai 201804, PR China

11 <sup>c</sup> College of Mathematics and Software Science, Sichuan Normal University, Chengdu 610068, PR China

#### ARTICLE INFO

- 16 Article history:
- 17 Received 14 May 2012
- 18 Received in revised form 3 September 2013
- 19 Accepted 7 September 2013
- 20 Available online xxxx
- 21 Keywords:
- 22 Rough set theory
- 23 Granular computing
- 24 Double quantification
- 25 Quantitative information architecture
- 26 Variable precision rough set
- 27 28 Graded rough set

#### ABSTRACT

Because precision and grade act as fundamental quantitative information in approximation 30 31 space, they are used in relative and absolute quantifications, respectively. At present, the double quantification regarding precision and grade is a novel and valuable subject, but 32 33 quantitative information fusion has become a key problem. Thus, this paper constructs the double-quantitative approximation space of precision and grade (PG-Approx-Space) 34 and tackles the fusion problem using normal logical operations. It further conducts 35 36 double-quantification studies on granular computing and rough set models. (1) First, for quantitative information organization and storage, we construct space and plane forms 37 of PG-Approx-Space using the Cartesian product, and for quantitative information extrac-38 tion and fusion, we establish semantics construction and semantics granules of PG-Approx-39 Space. (2) Second, by granular computing, we investigate three primary granular issues: 40 quantitative semantics, complete system and optimal calculation. Accordingly, six types 41 of fundamental granules are proposed based on the semantic, microscopic and macro-42 scopic descriptions; their semantics, forms, structures, calculations and relationships are 43 studied, and the granular hierarchical structure is achieved. (3) Finally, we investigate 44 rough set models in PG-Approx-Space. Accordingly, model regions are proposed by devel-45 oping the classical regions, model expansion is systematically analyzed, some models are 46 constructed as their structures are obtained, and a concrete model is provided. Based on 47 48 the quantitative information architecture, this paper systematically conducts and investigates double quantification and establishes a fundamental and general exploration 49 framework. 50

© 2013 Published by Elsevier Inc.

51 52

#### 1. Introduction 54

53

1.1. Pawlak-Model, VPRS-Model, GRS-Model 55

Rough set (RS) theory, first proposed by Pawlak [23], is a data analysis theory as well as a mathematical tool for dealing 56 with vague, inconsistent and incomplete information. Yao [51] provides this theory with two explanations: the set-oriented 57 and operator-oriented views. As a relatively new methodology on soft computing, RS-Theory has been extensively empha-58 59 sized in recent years, and its effectiveness has been confirmed by successful applications in many science and engineering

Q2 \* Corresponding author at: Department of Computer Science and Technology, Tongji University, Shanghai 201804, PR China. Tel.: +86 13658071900. Q1

E-mail addresses: xianyongzh@sina.com.cn (X. Zhang), miaoduoqian@163.com.cn (D. Miao).

0020-0255/\$ - see front matter © 2013 Published by Elsevier Inc. http://dx.doi.org/10.1016/j.ins.2013.09.020

Q1 Please cite this article in press as: X. Zhang, D. Miao, Quantitative information architecture, granular computing and rough set models in the double-quantitative approximation space of precision and grade, Inform. Sci. (2013), http://dx.doi.org/10.1016/j.ins.2013.09.020

6 7

12 13

3 5

1

### INS 10301 20 September 2013

**O1** 2

60

61

69

70 72

86

87

118

### ARTICLE IN PRESS

X. Zhang, D. Miao/Information Sciences xxx (2013) xxx–xxx

fields. RS-Theory has become one of the focuses of research in artificial intelligence theory and its applications, and some relevant results have been reported in [4,9,13,32,44,59].

RA

*U* is the finite and non-empty universe, while *R* denotes an equivalence relation.  $x \in U$ , equivalence class  $[x]_R$  is also called the atom granule, while a granule is generally defined as the union of certain equivalence classes, and  $U/R = \{[x]_R : x \in U\}$ represents knowledge. Thus, (U, R) constitutes the approximation space (*Approx-Space*). Suppose  $A \subseteq U, A$  represents a specific set generally called a concept, and |A| refers to the cardinal number. In the Pawlak-Model [23], the upper and lower approximations are defined by

$$\overline{R}A = \bigcup \{ [x]_R : [x]_R \cap A \neq \phi \}, \quad \underline{R}A = \bigcup \{ [x]_R : [x]_R \subseteq A \}.$$

$$\tag{1}$$

$$posRA = RA$$
,  $negRA = \sim \overline{R}A$ ,  $bnRA = \overline{R}A -$ 

(2)

denote the positive region, negative region, and boundary region, respectively. In particular, the classified three-way regions have the specific qualitative semantics and reflect certainty and uncertainty, thus underlying knowledge discovery and data mining. Moreover,  $\overline{R}$  and  $\underline{R}$  are the corresponding approximation operators. As the Pawlak-Model is only a qualitative model, it has some limitations, such as no fault-tolerance mechanisms. Thus, the Pawlak-Model needs improving, while the quantitative RS model has great value.

Uncertainty is a main feature of artificial intelligence, while probability acts as an important tool to describe uncertainty. 78 Thus, probability was introduced into RS-Theory to produce the probabilistic rough set [46,49,68]. The probabilistic rough 79 80 set has many outstanding merits, such as measurability, generality and flexibility, and it functions well in many concrete models, such as the 0.5-probabilistic rough set [38], variable precision rough set (VPRS) [14,69], decision-theoretic rough 81 set [50,54], game-theoretic rough set [2,7], parameterized rough set [6,26] and Bayesian rough set [36,58]. Meanwhile, 82 the graded rough set (GRS) [52], also a typical quantitative model, has some complementarity with the probabilistic rough 83 84 set. The VPRS and GRS are actually two fundamental types of quantitative models with both the fault-tolerance capability 85 and extended feature, and as a result, they serve as this paper's main background models.

Ziarko [69] proposed the VPRS by introducing the relative degree of misclassification:  $c([x]_R, A) = 1 - |[x]_R \cap A|/|[x]_R|$ . With a threshold of  $\beta \in [0, 0.5)$ , the upper and lower approximations become

$$\overline{R}_{\beta}A = \bigcup \{ [x]_{R} : c([x]_{R}, A) < 1 - \beta \}, \quad \underline{R}_{\beta}A = \bigcup \{ [x]_{R} : c([x]_{R}, A) \leq \beta \};$$

$$(3)$$

$$posR_{\beta}A = \underline{R}_{\beta}A, \quad negR_{\beta}A = \sim \overline{R}_{\beta}A, \quad bnR_{\beta}A = \overline{R}_{\beta}A - \underline{R}_{\beta}A$$
(4)

denote the positive region, negative region, and boundary region, respectively;  $\overline{R}_{\beta}$  and  $\underline{R}_{\beta}$  are the corresponding approximation operators. For largely solving the data noise problem, the VPRS has great significance for data acquisition and information analyses; moreover, it expands the Pawlak-Model. The VPRS is applied in many studies and applications, such as the attribute reduction and rule extraction in [1,3,8,10,11,19,37,57,66] and the practical results in geological and medical fields in [20,22,39–41,45].

 $p([x]_R, A) = 1 - c([x]_R, A) = |[x]_R \cap A| / |[x]_R|$  is the core measure in Approx-Space. It also corresponds to both the conditional 99 probability in the decision-theoretic rough set [12,15,17,31,55,65] and the rough membership proposed by Pawlak and 100 Skowron [24], and it further develops into the inclusion degree. Greco et al. [6] presented a generalized model of the VPRS 101 102 by using the absolute and relative rough membership, while Singh and Dey [33] used the rough membership for document categorization. Refs. [19,30,43] used the inclusion degree to extensively explore measures, reasoning and applications on 103 uncertainty. In this paper,  $p([x]_{e}, A)$  is specifically called **precision** of  $[x]_{e}$  with respect to A, while the threshold  $\beta \in [0, 0.5)$ 104 serves as the precision parameter. Note that precision is extracted from the VPRS but has already been generally promoted. 105 106 Thus,  $\overline{R}_{\beta}A = \bigcup \{ [x]_{R} : p([x]_{R}, A) > \beta \}$  refers to union of the equivalence classes whose precision with respect to A is greater than  $\beta$ .  $\underline{R}_{\beta}A = \cup \{[\mathbf{x}]_{R}, B] \neq 1 - \beta\}$  refers to union of the equivalence classes whose precision with respect to A is not smaller than 107  $1 - \beta$ . Similar to these quantitative descriptions, the precision description of threshold  $\beta$  is called **precision semantics**. Fur-108 thermore, the surplus three-way regions have their own precision semantics. In fact, precision semantics reflect the relative 109 110 fault-description feature and determine the VPRS application function.

Yao and Lin [52] explored the relationships between rough sets and modal logics and proposed the GRS by utilizing the graded modal logics. The GRS basically describes the absolute quantitative information about knowledge and concepts and expands the Pawlak-Model. The discussion of the GRS mainly focuses on the model construction, such as those in [16,42,53], while its background (the graded modal logic) has fruitful studies. **N** denotes the natural number set and threshold  $k \in \mathbf{N}$ . In the GRS, the upper and lower approximations become

$$\overline{R}_{k}A = \cup\{[x]_{R} : |[x]_{R} \cap A| > k\}, \quad \underline{R}_{k}A = \cup\{[x]_{R} : |[x]_{R} \cap A| \leqslant k\};$$

$$(5)$$

119  $\overline{R}_k$  and  $\underline{R}_k$  are the corresponding approximation operators. Measures  $|[x]_R \cap A|$  and  $|[x]_R| - |[x]_R \cap A|$  reflect the absolute number 120 of  $[x]_R$  elements inside and outside A, respectively. In this paper,  $\overline{g}([x]_R, A) = |[x]_R \cap A|$  and  $\underline{g}([x]_R, A) = |[x]_R | - |[x]_R \cap A|$  are called 121 **internal grade** and **external grade** of  $[x]_R$  with respect to A, respectively, while threshold k serves as the grade parameter. 122 Internal grade and external grade originate from the GRS, but after extraction and promotion, they have already transcended 123 the GRS and, in fact, describe another group of applicable measures in Approx-Space. Thus,  $\overline{R}_k A = \cup \{[x]_R : \overline{g}([x]_R, A) > k\}$  refers 124 to union of the equivalence classes whose internal grade with respect to A is greater than k;  $\underline{R}_k A = \cup \{[x]_R : g([x]_R, A) \le k\}$  refers to

01

### ARTICLE IN PRESS

### 20 September 2013

#### X. Zhang, D. Miao/Information Sciences xxx (2013) xxx-xxx

3

(7)

union of the equivalence classes whose external grade with respect to A is not greater than k. Similarly, the grade description of the threshold k is called **grade semantics**, while grade semantics reflect the absolute fault-description feature and also underlie the model applications. Usually,  $\underline{R}_k A \subseteq \overline{R}_k A$  does not hold, and there are rarely reports about the GRS classified regions, except [62].

#### 129 1.2. Double quantification

 $(r - \frac{\varepsilon_A}{2})$ 

The relative and absolute errors, two basic notions, have already existed, such as in the measurement field. Here,  $c([x]_R, A)$ 130 131 and  $g([x]_R, A)$  are the relative and absolute errors about knowledge R and set A, respectively. On the other hand,  $p([x]_R, A)$  and  $\overline{g}(|x|_{\overline{p}},A)$  can refer to the relative overlap rate and absolute overlap number, respectively. Thus, as precision and grade reflect 132 the relative and absolute quantitative information of Approx-Space, respectively, they become fundamental measures. By 133 134 incorporating precision and grade, the VPRS and the GRS further contribute to the relative and absolute quantification, thus 135 having relative and absolute fault-tolerance capabilities, respectively. The relative and absolute quantitative information 136 serve as two distinctive objective sides for the Approx-Space description, and each side has its own representative virtues 137 and application environments. Thus, each one needs to be emphasized rather than neglected. Ref. [62] provides several 138 examples to illustrate the importance of both. Herein, we provide another relevant example about measurement.

**Example 1.** Suppose *r* and *m* are the real value and measured value in a measurement process, respectively. As is well known,  $\varepsilon_R = (m - r)/r$  and  $\varepsilon_A = m - r$  denote the relative and absolute errors, respectively, and they both reflect some reliability of the measured result *m*. For the measurement evaluation, each one may be singly utilized. Thus,  $\varepsilon_R = 10\%$  actually reflects certain measurement information, but it is concerned with only one part. Suppose  $\varepsilon_{R1} = 10\%$  and  $\varepsilon_{R2} = 30\%$ . Is the first measurement more precise than the latter? Moreover, consider  $\varepsilon_{R1} = \varepsilon_{R2}$ . The answers are difficult because of information incompleteness. However, the double-error awareness will result in a new situation. Generally, suppose  $\varepsilon_R \neq 0$ , then

$$\begin{cases} \varepsilon_R = \frac{m-r}{r} \\ \varepsilon_A = m - r \end{cases}$$
(6)

150

162

$$\left\{ egin{array}{ll} T = arepsilon_{e_R} \ m = rac{1 + arepsilon_R}{arepsilon_{e_R}} arepsilon_A \end{array} 
ight.$$

Systems (r, m) and  $(\varepsilon_{R}, \varepsilon_{A})$  are equivalent, while formulas (6) and (7) become the system transformation formulas based on 151 mutual deduction. Therefore, system  $(\varepsilon_R, \varepsilon_A)$  is also complete when compared to system (r, m). Thus,  $\varepsilon_R$  and  $\varepsilon_A$  have com-152 153 pleteness, they can completely determine system (r, m) and they can fully reflect the accurate measurement information, 154 which is shown by formula (7). For assumptive two measurement projects,  $\varepsilon_{R1} = 10\%$ ,  $\varepsilon_{A1} = 1000$  (m),  $\varepsilon_{R2} = 30\%$ , 155  $\varepsilon_{A2} = 30$  (m), then by formula (7), we obtain  $r_1 = 10,000$  (m),  $m_1 = 11,000$  (m) while  $r_2 = 100$  (m),  $m_2 = 130$  (m). Though  $\varepsilon_{R2} = 30\% > 10\% = \varepsilon_{R1}$ , measurement 2 may be better because its absolute loss (30 m) is considerably lower than 156 measurement 1's (1000 meters). This assessment result particularly exhibits the system characteristic of the double-indexes. 157 Moreover, the single system  $\varepsilon_R$  (or  $\varepsilon_A$ ) is incomplete; thus, if  $\varepsilon_{R1} = \varepsilon_{R2}$ , then we must resort to additional information 158 159 regarding  $\varepsilon_{A1}$  and  $\varepsilon_{A2}$ . In fact, 160

$$\varepsilon_{R} = \frac{\varepsilon_{A}}{r}, \quad \frac{\partial \varepsilon_{R}}{\partial r} = -\frac{\varepsilon_{A}}{r^{2}}, \quad \frac{\partial \varepsilon_{R}}{\partial \varepsilon_{A}} = \frac{1}{r}, \quad d\varepsilon_{R} = -\frac{\varepsilon_{A}}{r^{2}}dr + \frac{1}{r}d\varepsilon_{A}.$$
(8)

Formula (8) shows that the relationship between  $\varepsilon_R$  and  $\varepsilon_A$  is non-linear and that  $\varepsilon_A$  can relatively impact both  $\varepsilon_R$  and its change; the similar opposite also holds. Thus,  $\varepsilon_R$  and  $\varepsilon_A$ , in fact, have a close mutual relationship, and their complementarity exists. For the dimension, the measurement system (r, m) is two-dimensional, while the double-error system  $(\varepsilon_R, \varepsilon_A)$  is usually two-dimensional as well. However, as the single system  $\varepsilon_R$  (or  $\varepsilon_A$ ) is only one-dimensional, it has less accuracy.

167 In Approx-Space, because precision and grade are the relative and absolute measures and are related to the relative and 168 absolute errors, they have a close and dialectical relationship. According to this relative or absolute measure, the VPRS and 169 GRS only make the relative or absolute (single) quantification. In fact, the core quantitative system of Approx-Space and the composite system of precision and grade are significantly similar to the measurement system (r, m) and the errors system 170  $(\varepsilon_R, \varepsilon_A)$ , respectively. Section 2 will provide some novel studies and important results of the mathematical mechanism 171 analyses, and it will demonstrate that precision and grade have a non-linear relationship. Section 2 will also discuss the 172 quantitative complementarity and quantitative completeness with respect to precision and grade. In Approx-Space, the 173 174 description with the three measures (i.e., precision, internal grade and external grade) is called double quantification, simply noted 175 as **D-Quantification**. D-Quantification serves as a general and systematic notion with both relative and absolute quantitative 176 information. Based on Section 2, the D-Quantification system has quantitative completeness for the core quantification system 177 of Approx-Space. However, when considering its three-dimensional application form, it also exhibits both quantitative 178 expansibility and quantitative inclusiveness. Compared to the single quantification, the extended D-Quantification ap-179 proaches both the inherent quantitative essence of Approx-Space and the high actual demand of RS-Theory. Accordingly, 180 D-Quantification provides more quantitative information and stronger descriptive abilities, thus implying better measurable 181 results. Furthermore, it has double fault tolerance capabilities, allowing it to adapt to more complex environments, and it can

01 4

### **ARTICLE IN PRESS**

#### 20 September 2013

#### X. Zhang, D. Miao / Information Sciences xxx (2013) xxx-xxx

construct more application models with completeness and diversity, thus improving and promoting the existing models,
 such as the VPRS and GRS. In summary, D-Quantification has great significance.

184 Regarding computational complexity, D-Quantification usually has the same feasibility level with respect to the single 185 guantification. Thus, D-Quantification is reasonable. Meanwhile, the VPRS and the GRS are used for the relative and absolute 186 quantification, respectively, and they have many close similarities and relationships, such as the similar basic properties for the approximation operators. Reference [62] conducted a comparative study of the two models and confirmed their clear 187 relationships and mutual transformations. Moreover, Refs. [5,42] investigated their formal transformations based on broader 188 189 relations. Because D-Quantification is found to have a stable foundation, it is a feasible application. At present, there are few 190 studies about D-Quantification, except for a series of relevant literatures linked to us, such as [61,63,64]. In fact, Refs. 191 [61,63,64] reflected our initial and elementary work on some partial and basic combined quantification, and they emphatically investigated some preliminary elements and specific aspects of the combination of precision and grade, such as the 192 193 single operation on AND OR NOT, several direct combination operators, several vivid models and concrete algorithms. In particular, D-Quantification has distinctive systematization and prominent generalization features and is a novel, important but 194 195 complex and abstract subject. Therefore, our previous single, vivid, concrete works have established the important foundations for the in-depth D-Quantification exploration. 196

#### 197 1.3. Granular computing

198 Granular computing (GrC), first proposed by Zadeh [56], has emerged as one of the fastest growing information processing paradigms in computational intelligence and human-centric systems. GrC is often loosely defined as an umbrella term to 199 cover any theories, methodologies, techniques or tools that make use of granules in complex problem solving [47]. The 200 201 GrC research has attracted many researchers and practitioners. For example, Pedrycz et al. investigated many fundamental 202 aspects of information granules in intelligent systems in [27–29]. In particular, Yao [47] proposed the triarchic model of GrC, 203 i.e., the philosophy of structured thinking, the methodology of structured problem solving, and the computation of struc-204 tured information processing. In essence, GrC reflects the thoughts on multiple granules, multiple levels and multiple views and concretely provides a methodology for information processing. Moreover, granules are fundamental notions, and 205 granulation is one of the key issues, while the granular hierarchical structure provides describing and solving methods of 206 corresponding hierarchical problems. 207

As RS-Theory acts as a typical concrete model for GrC, some aspects on rough GrC were extensively investigated. For example, Pawlak and Skowron [25] provided rough set perspectives of GrC, while from a rough computing perspective, Skowron and Stepaniuk [34] explored the formation of granules with different criteria. Yao [48] studied both the rough set approximation and information granulation. From the topological view of GrC, Zhu [67] explored covering-based rough sets, and Liu et al. [18] explored GrC from a rough logic aspect. In particular, based on Approx-Space, Skowron et al. [35] investigated some important issues on modeling rough GrC, and Pal et al. [21] investigated the image object extraction within the framework of both rough sets and GrC.

#### 215 1.4. Thoughts and outline

216 Because D-Quantification has the quantitative completeness and quantitative expansibility/inclusiveness, it has become a 217 novel, valuable, reasonable and feasible subject. For scientific exploration, we first must construct the existing space for D-Quantification, and for this purpose, the basic Approx-Space is worth promoting. Thus, it becomes both fundamental and 218 interesting to construct a novel double-quantitative space with wonderful and in-depth structures from which to investigate 219 220 D-Quantification. In this paper, we make a preliminary attempt at this task. Herein, Approx-Space with the three measures is specifically called the double-quantitative approximation space of precision and grade, simply noted as **PG-Approx-Space**. There-221 222 fore, PG-Approx-Space becomes a basic supported structure for D-Quantification. Meanwhile, as PG-Approx-Space now seems rather vacuous, there is an urgent need for the rigorous structural forms and quantitative information combination 223 224 connotation. Our main work is to create and investigate D-Quantification in PG-Approx-Space, noting that PG-Approx-Space 225 correspondingly becomes full.

226 First, how to create the D-Quantification becomes a primary problem, i.e., the quantitative information fusion becomes a key problem. By using the Cartesian product combination on only two measures of precision and grade, we conduct some 227 relevant investigation in [60]. In particular, logical AND OR NOT operations serve as usual combination methods and fusion 228 technologies in the classical logic, and they closely correspond to the RS-Theory set essence. Refs. [61,63,64] have separately 229 but effectively utilized the logical integration. Thus, we will systematically utilize AND OR NOT fusion methods for D-Quan-230 231 tification integration. Second, D-Quantification has the three measures, a fact that follows the basic GrC idea with respect to 232 multiple views, and thus the relevant GrC studies become necessary and feasible. In fact, GrC can provide a perfect research 233 scheme in view of the complexity of D-Quantification. Using GrC, we will investigate three primary granular issues: the 234 quantitative semantics, the complete system and optimal calculation. Third, the RS model underlies the practical 235 applications of RS-Theory, while the VPRS and GRS act as only a single quantification RS-Model. Accordingly, the double-236 quantitative RS model is worth studying, and we will discuss model expansion and model construction in PG-Approx-Space. 237 In short, we will create D-Quantification works in PG-Approx-Space using three fundamental and systematic techniques: 238 quantitative information architecture, GrC and the RS-Model.

5

X. Zhang, D. Miao/Information Sciences xxx (2013) xxx-xxx

239 Based on mathematical mechanism analyses, Sections 1.2, 2.1, 2.2 strengthen the justification for the D-Quantification 240 methodology as well as its novelty and explore PG-Approx-Space mathematical forms. We then systematically utilize 241 AND OR NOT fusion technologies to establish the semantics construction and semantics granules, which underlies D-Quan-242 tification and endows PG-Approx-Space with the information fusion connotation. Furthermore, we finally conduct D-Ouan-243 tification studies on GrC and the RS-Model. Therefore, this paper involves four concrete contents with respect to PG-Approx-244 Space: (1) the basic mathematical forms, (2) the semantics construction and semantics granules, (3) the multiple granules 245 and their features and relationships, (4) the model expansion and model construction. Furthermore, this paper exploits a dis-246 tinctive D-Quantification characteristic and has the following relevant contributions. (1) The novel mathematical forms of 247 PG-Approx-Space are well organized and store the double-quantitative information. (2) The semantics construction and semantics granules systematically utilize the normal AND OR NOT technologies to complete quantitative information extrac-248 tion and fusion, and the relevant system construction underlies the D-Quantification exploration. (3) GrC effectively exhibits 249 250 the rich granules and their relevant semantics, calculations and hierarchy of D-Quantification. (4) The RS-Model is systematically studied in PG-Approx-Space, including model expansion and model construction. Therefore, based on the quantita-251 252 tive information architecture, this paper systematically conducts and investigates D-Quantification and particularly establishes a fundamental and general exploration framework. 253

The rest of this paper is organized as follows. Regarding PG-Approx-Space, Section 2 constructs the space and plane forms 254 to organize and store quantitative information. Section 3 explores the semantics construction and semantics granules to ex-255 256 tract and fuse the quantitative information. Granulation is then mainly conducted using the semantics and the microscopic and macroscopic descriptions, and multiple fundamental granules are proposed and studied. Section 4.1 studies basic 257 semantics (BS) and BS-Granules; Sections 4.2 and 4.3 refer to B-Granules and C-Granules, respectively; Section 5.1 explores 258 259 M-Regions, BM-Regions, BMC-Granules; furthermore, Section 5.2 presents the granular hierarchical structure. For the RS-260 Model in PG-Approx-Space, Section 5.1 proposes the model regions and model definition by developing the traditional no-261 tions; Section 6.1 studies model expansion, constructs some models and further develops their structures; Section 6.2 pro-262 vides a concrete model. Finally, Section 7 concludes this paper.

#### 263 **2. Basic mathematical forms of PG-Approx-Space**

By analyzing some mathematical mechanisms on precision and grade, this section constructs the basic mathematical forms of PG-Approx-Space, including the space and plane forms. Both forms exhibit the quantitative information structures, rigorously organize and effectively store the quantitative information, and basically describe and construct PG-Approx-Space. That is, they provide the mathematical basis and concrete technology for the later in-depth D-Quantification explorations. These studies, in fact, concern the basic granular discussion of PG-Approx-Space because the equivalence classes are atom-granules and their structures are ingeniously constructed herein.

#### 270 2.1. Space form

Precision, internal grade and external grade can be utilized to measure the relationship between  $[x]_R$  and A. Furthermore, if A is treated as the set parameter while  $[x]_R$  changes in knowledge, then the three measures can act as variables. Suppose **R** denotes the real number set.  $\forall [x]_R \in U/R$ , suppose  $X = \underline{g}([x]_R, A), Y = \overline{g}([x]_R, A), Z = p([x]_R, A)$ . Thus, the three measures are transposed to three variables X, Y, Z.

**Definition 2.1.** The three-dimensional coordinate system formed by *X*, *Y* and *Z* is called *the measure space on precision, internal grade and external grade* and is simply noted as **Measure Space**.

In Definition 2.1, by being viewed as three coordinate variables, the three measures are used to establish a three-dimensional coordinate system (i.e., Measure Space). In fact, the space construction uses the Cartesian product of the relative and absolute measures, and this idea becomes scientific because D-Quantification corresponds to multiple views.

Suppose  $f_1 : U/R \to \mathbf{R}^3$ ,  $f_1([x]_R) = (X, Y, Z)$ .  $f_1$  moves each equivalence class into three-dimensional Measure Space. In other words, equivalence classes construct the three-dimensional organization using the three measures. From the three measures perspective, Approx-Space exhibits a three-dimensional form. Moreover, PG-Approx-Space, as Approx-Space with the three measures, is also described. We next investigate the relationship between precision and grade and further exhibit the exact space form of PG-Approx-Space.

#### 285 Proposition 2.2.

286 (1) 
$$Z = \frac{Y}{X+Y}$$
,  $\frac{\partial Z}{\partial X} = -\frac{Y}{(X+Y)^2}$ ,  $\frac{\partial Z}{\partial Y} = \frac{X}{(X+Y)^2}$ ,  $dZ = -\frac{Y}{(X+Y)^2} dX + \frac{X}{(X+Y)^2} dY$ .  
287 (2)  $X = Y \frac{1-Z}{Z}$ ,  $\frac{\partial X}{\partial Y} = \frac{1-Z}{Z}$ ,  $\frac{\partial X}{\partial Z} = -\frac{Y}{Z^2}$ ,  $dX = \frac{1-Z}{Z} dY - \frac{Y}{Z^2} dZ$ .  
289 (2)  $X = X \frac{Z}{Z}$ ,  $\frac{\partial Y}{\partial Y} = \frac{Z}{Z}$ ,  $\frac{\partial Y}{\partial Z} = -\frac{Y}{Z^2}$ ,  $dX = \frac{1-Z}{Z} dY - \frac{Y}{Z^2} dZ$ .

(3) 
$$Y = X \frac{Z}{1-Z}, \quad \frac{\partial Y}{\partial X} = \frac{Z}{1-Z}, \quad \frac{\partial Y}{\partial Z} = \frac{X}{(1-Z)^2}, \quad dY = \frac{Z}{1-Z} dX + \frac{X}{(1-Z)^2} dZ.$$
  
291

Q1 Please cite this article in press as: X. Zhang, D. Miao, Quantitative information architecture, granular computing and rough set models in the double-quantitative approximation space of precision and grade, Inform. Sci. (2013), http://dx.doi.org/10.1016/j.ins.2013.09.020

Q1

Q1 6

307

309 310 311

312

X. Zhang, D. Miao/Information Sciences xxx (2013) xxx-xxx

Proposition 2.2 provides the mutual relationships among the three variables, where the three items are equivalent. For item (1), equality Z = Y/(X + Y) is the non-linear relationship between precision and grade. With respect to the measure variables, Z = Y/(X + Y) further becomes a binary function, and by continuity expansion, we also provide the partial derivatives and total differential formula. Accordingly, internal grade and external grade influence precision and its change. In a word, precision and grade have a non-linear relationship and quantitative complementarity.

For the three indexes, only one constraint condition exists, and it further forms a specific binary function in the Measure Plane, i.e., Z = Y/(X + Y). Thus, all equivalence classes are endowed with a novel structure. Atom-granules are discretely distributed on the surface Z = Y/(X + Y) (here,  $X, Y, Z \ge 0$ ) in Measure Space. In other words, the space surface Z = Y/(X + Y)reflects the PG-Approx-Space. In particular, Fig. 1 shows the surface figure. Thus, PG-Approx-Space becomes well-organized and in-depth, its space form is microscopically provided, and it scientifically stores and objectively exhibits the double-quantitative information of Approx-Space.

### 308 Proposition 2.3.

(1)  $\frac{\partial Z}{\partial X} < 0, \quad \frac{\partial Z}{\partial Y} > 0.$ 

(2) If X = Y, then Z = 0.5.

113 Item (1) in Proposition 2.3 shows that precision becomes monotonic decreasing and monotonic increasing with respect to external grade and internal grade, respectively, while Item (2) shows that precision becomes 0.5 when external grade and internal grade are equal. Fig. 1 also reflects the two mathematical properties for Z = Y/(X + Y).

According to the precision parameter, the VPRS approximations correspond to two cut-sets of the surface Z = Y/(X + Y) in the direction of Z-axis. Similarly, according to the grade parameter, the GRS upper and lower approximations correspond to two cut-sets of the surface Z = Y/(X + Y) in the directions of the Y-axis and X-axis, respectively. Moreover, other regions can be similarly analyzed. Accordingly, as the space surface also includes both the VPRS and GRS descriptions, it has a strong application feature. In fact, Measure Space provides a perfect description tool, particularly for the measure applications.

Observation 2.4. The D-Quantification system ( $\underline{g}([x]_R, A), \overline{g}([x]_R, A), p([x]_R, A)$ ) and PG-Approx-Space are both threedimensional.

The D-Quantification system has three views, while PG-Approx-Space corresponds to a three-dimensional surface. Thus, Observation 2.4 provides the natural dimension conclusion.

#### 325 2.2. Plane form

The space form is mainly constructed based on precision, internal grade and external grade. However, one and only one independent equality exists among the three measures, and this point implies the possible plane description of PG-Approx-Space. For this purpose, we first construct the two-dimensional mathematical form of the core quantitative Approx-Space and further analyze PG-Approx-Space.

For Approx-Space (U, R) and set A, there are two and only two direct and core measures:  $|[x]_R|$  and  $|[x]_R \cap A|$ . Other application measures, such as precision and grade, can be formed based on the first two core measures. Thus,  $|[x]_R|$  and  $|[x]_R \cap A|$ .





### INS 10301 20 September 2013

### ARTICLE IN PRESS X. Zhang, D. Miao/Information Sciences xxx (2013) xxx-xxx

Q1

serve as the core quantitative indexes and can reflect the Approx-Space quantitative essence. For this study, both measures will be utilized to similarly construct a two-dimensional plane.  $\forall [x]_R \in U/R$ , suppose  $V = V([x]_R, A) = |[x]_R|, Y = \overline{g}([x]_R, A)$  $= |[x]_R \cap A|$ .

**Definition 2.5.** The two-dimensional coordinate system formed by *V* and *Y* is called *the measure plane on the cardinal number and internal grade* and is simply noted as **Measure Plane**.

In Definition 2.5, by being viewed as two coordinate variables, the two core measures  $|[x]_R|$  and  $|[x]_R \cap A|$  are used to establish a two-dimensional coordinate system (i.e., Measure Plane). Similarly, the plane construction utilizes the two-dimensional Cartesian product, which corresponds to the two views of the core indexes. In fact, there are no decisive relationships between  $|[x]_R|$  and  $|[x]_R \cap A|$ , though  $|[x]_R \cap A| \leq |[x]_R|$ . In other words, variables *V* and *Y* are almost completely independent. We define  $f_2$  as  $U/R \rightarrow N^2$ ,  $f_2([x]_R) = (V, Y)$ . Accordingly, all atom-granules are discretely distributed in the region  $0 \leq Y \leq V$  in the Measure Plane, and the essential two-dimensional form of the core quantitative Approx-Space is exhibited.

Observation 2.6. The core quantification system  $(|[x]_R|, |[x]_R \cap A|)$  and core quantitative Approx-Space are both twodimensional.

The system  $(|[x]_R|, |[x]_R \cap A|)$  naturally has the two-dimensional feature and serves as a core quantitative system for the system (U, R, A) on Approx-Space. In other words, Approx-Space has the two-dimensional essence with respect to quantification. In particular, the Measure Plane provides a vivid explanation.

**Theorem 2.7.** The two systems  $(g([x]_R, A), \overline{g}([x]_R, A), p([x]_R, A))$  and  $(|[x]_R|, |[x]_R \cap A|)$  are equivalent, i.e., they can mutually deduce.

**Proof.** For simplicity, we use the relevant variable symbols, i.e., *X*, *Y*, *Z*, *V*. As  $Z = 0 \iff Y = 0 \Rightarrow [x]_R \subseteq \sim A$ , this special case is clear. As a result, only the usual case  $Z \neq 0$  (i.e.,  $Y \neq 0$ ) requires analysis and the following relationships hold.

$$\begin{cases} X = V - Y \\ Y = Y \\ Z = Y/V \end{cases}$$
(9)

355

354

352

$$\begin{cases} V = X + Y = Y/Z \ (Z \neq 0) \\ Y = Y \end{cases}$$
(10)

Formula (9) shows that the three usual measures can be represented by the two core measures  $|[x]_R|$  and  $|[x]_R \cap A|$ , while formula (10) reflects the opposite. Thus, systems (X, Y, Z) and (V, Y) are equivalent.  $\Box$ 

**Theorem 2.8.** System  $(\underline{g}([x]_R, A), \overline{g}([x]_R, A))$  is complete with respect to system  $(|[x]_R, |, |[x]_R \cap A|)$ .

Theorem 2.7 shows the equivalence (i.e., the mutual decisiveness and representation) between the D-Quantification system and the core quantification system. In particular, formulas (9) and (10) provide the system transformation basis. Furthermore, Theorem 2.8 reflects the D-Quantification completeness for the core quantification system. Using mechanism analyses, we establish a key D-Quantification foundation, that is, quantitative completeness, and we highlights this paper's motivation.

**Observation 2.9.** For the relative and absolute (single) quantification, the VPRS with  $p([x]_R, A)$  serves as a one-dimensional system, while the GRS with  $\underline{g}([x]_R, A), \overline{g}([x]_R, A)$  serves as a two-dimensional system. Moreover, the probabilistic rough set also corresponds to a one-dimensional system.

369 Observations 2.4, 2.6, 2.9 describe the dimension results of the three basic systems, respectively, i.e., the D-Quantification system, core quantification system, single quantification system. Based on these results, we can identify two main D-Quantifi-370 cation features from the dimension perspective, i.e., quantitative expansibility and quantitative inclusiveness. For 371 372 Approx-Space, almost all explorations originate from its core quantification, and its core quantitative form is objectively two-dimensional. For practical applications, D-Quantification is promoted to three-dimensions by the three usual measures. 373 374 According to Theorem 2.8, the D-Quantification system has quantitative completeness on the two-dimensional core system 375  $(|[x]_{R}|, |[x]_{R} \cap A|)$ . On the contrary, the probabilistic rough set (including the VPRS) extracts only one main measure for its 376 applications, and accordingly, it degenerates into a one-dimensional system due to projection and dimension reduction. 377 However, this point also results in less accuracy. As D-Quantification has quantitative completeness, it can completely reflect 378 the inherent quantitative essence of Approx-Space. Furthermore, as D-Quantification has quantitative expansibility, it also 379 has stronger applicability, particularly regarding the relative and absolute quantitative descriptions. Moreover, because D-Quantification has quantitative inclusiveness, it can extract the low dimensional parts, such as the precision system, which 380 381 is connected with only a one-dimensional projection.

20 September 2013

## ARTICLE IN PRESS

Q1 8

X. Zhang, D. Miao/Information Sciences xxx (2013) xxx-xxx

382 Based on the above analyses, PG-Approx-Space has quantitative completeness, quantitative expansibility and 383 quantitative inclusiveness for the core quantitative Approx-Space. In other words, PG-Approx-Space demonstrates a 384 complete, expanded and applicable structure. In fact, PG-Approx-Space has extracted, included and stored the core 385 quantitative information of Approx-Space, and also important is the fact that PG-Approx-Space is highly applicable 386 for the usual measures. Accordingly, PG-Approx-Space has important research value. In the later sections, we study 387 the semantics construction, GrC, and RS-Model in PG-Approx-Space. For the RS-Model, both the double-quantitative 388 and degenerate models can be directly investigated in PG-Approx-Space. In other words, PG-Approx-Space also 389 establishes a general framework for RS-Model explorations. In particular, it is relevant that the three basic models 390 (Pawlak-Model, VPRS-Model, GRS-Model) also exist in PG-Approx-Space, though they demonstrate project extraction and dimension reduction. 391

In fact, Measure Plane, which originates from the core quantitative system, can also describe D-Quantification well and 392 393 can further provide a simple and direct plane form of PG-Approx-Space. According to formula (9), precision and grade have their own mathematical/geometric meanings in Measure Plane.  $p([x]_R, A) = Z = Y/V$  means that the precision of an equiva-394 lence class becomes the slope of the line that passes through the origin and atom-granular point.  $\overline{g}([x]_R, A) = Y$  means that the 395 internal grade of an equivalence class becomes the distance from the atom-granular point to V-axis.  $g([x]_R, A) = X = V - Y$  means 396 that the external grade of an equivalence class becomes the difference between the two distances from the atom-granular point to 397 Y-axis and V-axis. Moreover, all the VPRS and GRS regions also have their own mathematical/geometric meanings. The VPRS 398 399 upper and lower approximations with threshold  $\beta$  correspond to the regions of  $Y > \beta V$  and  $Y \ge (1 - \beta)V$ , respectively. The GRS upper and lower approximations with threshold k correspond to the regions of Y > k and  $Y \ge V - k$ , respectively. 400 Other regions can be similarly analyzed. Thus, we also use Measure Plane to creatively explain the existing fundamental 401 402 notions.

403 In summary, the two-dimensional system  $(|[x]_R|, |[x]_R \cap A|)$  serves as the core quantitative system of Approx-Space, and its 404 Measure Plane also has fundamental value to PG-Approx-Space, which has been verified based on the mathematical/geomet-405 ric meanings of both the measures and regions. According to the core information, all equivalence classes are distributed in 406 Measure Plane. On the contrary, utilizing Measure Plane can effectively extract the three measures. Thus, PG-Approx-Space also becomes well-organized and in-depth. Meanwhile, Measure Plane provides a basic plane form of PG-Approx-Space and 407 also becomes an important tool to explore D-Quantification, particularly for the GrC study. Moreover, the system 408  $(|[x]_p|, |[x]_p \cap A|)$  has calculation directness and optimization, and therefore, it is also used to examine the granular calcula-409 410 tions later.

#### 411 **3. Semantics construction and semantics granules**

Using the basic mathematical forms for Measure Space and Measure Plane, PG-Approx-Space has rigorously organized 412 413 and stored the relative and absolute quantitative information. The next task is to extract and apply the quantitative information on precision and grade. In practice, atom-granules are extracted by quantitative information and are further merged 414 415 into application regions. In other words, PG-Approx-Space provides the accurate three-measure descriptions of atom-granules located in the microcosmic bottom layer, while the practical application requires information fusion and region extrac-416 417 tion located in the macrocosmic high layer. For this application purpose, both quantitative information extraction and fusion 418 become a fundamental work in PG-Approx-Space. Furthermore, the quantitative semantics have the dual function of extrac-419 tion and fusion, and thus play a critical role. In fact, quantitative semantics involve both measure connotation and region extension. Thus, this section first explores the semantics construction principle of PG-Approx-Space and further defines 420 and studies the relevant semantics and semantics granules. 421

422 The VPRS case is first analyzed. Precision information is mainly extracted by the threshold criteria. The precision semantics are formed, and the bearing granule can be obtained. In the VPRS, precision descriptions  $p(|x|_p, A) > \beta$  and 423  $p([x]_R, A) \ge 1 - \beta$  serve as the initial precision semantics and are embodied by the relevant granules  $\cup \{[x]_R : p([x]_R, A) > \beta\}$ 424 and  $\cup \{[x]_R : p([x]_R, A) \ge 1 - \beta\}$ , i.e., the VPRS upper and lower approximations. In application, the classified three-way regions 425 are specifically required. In the set operation view, the three-way regions are mainly obtained by the intersection union com-426 427 plementary operations of the approximations, which is also shown by formula (4). In the precision semantics perspective,  $posR_{\beta}A = \cup\{[x]_{R}: p([x]_{R}, A) \ge 1 - \beta\}, negR_{\beta}A = \cup\{[x]_{R}: p([x]_{R}, A) \le \beta\}, bnR_{\beta}A = \cup\{[x]_{R}: \beta < p([x]_{R}, A) < 1 - \beta\}.$  Thus, we dis-428 cover that the three-way precision semantics originate from the logical AND OR NOT operations of  $p([x]_R, A) > \beta$  and 429  $p([\mathbf{x}]_{\mathbb{R}}, A) \ge 1 - \beta$ . For example,  $\beta < p([\mathbf{x}]_{\mathbb{R}}, A) < 1 - \beta$  means  $[p([\mathbf{x}]_{\mathbb{R}}, A) > \beta] \land \neg [p([\mathbf{x}]_{\mathbb{R}}, A) \ge 1 - \beta]$ . In short, precision semantics 430 of approximations serve as the initial fused basis, and the regions are actually the bearing granules of composite precision 431 semantics obtained by the logical AND OR NOT fusion of the initial precision semantics. 432

433 According to precision, the above example clearly reflects three basic aspects: (1) the initial quantitative information 434 extraction on the threshold, (2) the quantitative information fusion on the logical AND OR NOT operations, and (3) the nat-435 ural granules merging by the semantics action. Moreover, it also shows a concept's two dialectical sides: the semantics con-436 notation and set extension. Thus, the quantitative information extraction and fusion on precision and grade can be similarly 437 constructed in PG-Approx-Space. In other words, for D-Quantification applications, PG-Approx-Space may use thresholds  $\beta$ 438 and k to extract quantitative information from the initial approximations, may use logical AND OR NOT operations to fuse 439 quantitative information, and can further merge atom-granules into regions according to the fused quantitative description. 440 However, the three measures must be simultaneously considered.

Q1

X. Zhang, D. Miao/Information Sciences xxx (2013) xxx-xxx

**Definition 3.1. Semantics Construction Principle** of PG-Approx-Space refers to the two following rules regarding quantitative information extraction and fusion. (1) The initial measure descriptions of thresholds follow the VPRS and GRS approximations forms, i.e., only four types of initial precision or grade semantics exist:  $p([x]_R, A) > \beta, p([x]_R, A) \ge 1 - \beta$ ,  $\overline{g}([x]_R, A) > k, \overline{g}([x]_R, A) \le k$ . (2) The quantitative information fusion is mainly performed with the finite use of the classical logical AND OR NOT operations.

The semantics construction principle is established by adhering to the above details regarding the VPRS analyses. Based on the VPRS and GRS approximations, rule (1) defines the initial quantitative semantics such that the measures descriptions of thresholds become natural for the background models, and the quantitative expansibility/inclusiveness is harmoniously embodied. Rule (2) defines the quantitative information fusion technologies according to the logical AND OR NOT operations. According to this fundamental principle, the semantics and semantics granules can be naturally defined and deeply studied in PG-Approx-Space.

452 **Definition 3.2.** Regarding the semantics construction principle, the measures description of thresholds  $\beta$  and k is called 453 **semantics**. Moreover, the semantics set is noted as  $S = \{s\}$ .

The semantics in PG-Approx-Space simulate and include/extend the quantitative semantics of the regions in the VPRS-Model, GRS-Model and Pawlak-Model. In particular, the semantics construction principle provides a concrete and operational specification for the semantics construction. For example,  $s_* = [p([x]_R, A) > \beta] \land \neg[p([x]_R, A) \ge 1 - \beta] \land [\overline{g}([x]_R, A) > k] \land \neg[g([x]_R, A) \le k]$  is the semantics, and it is actually  $s_* = [\beta < p([x]_R, A) < 1 - \beta] \land [\overline{g}([x]_R, A) > k] \land [\overline{g}([x]_R, A) > k]$ ,  $\wedge [g([x]_R, A) > k] \land [\overline{g}([x]_R, A) > k]$ ; however,  $p([x]_R, A) \ge \beta$  is not semantics because it cannot be obtained using the semantics construction principle.

According to the propositional logic method, the inductive definition of semantics can be equivalently provided by applying rules (1)(2) of the semantics construction principle.

461 **Definition 3.3.** In PG-Approx-Space, semantics is defined by the following rules:

462 (1)  $p([x]_R, A) > \beta, p([x]_R, A) \ge 1 - \beta, \overline{g}([x]_R, A) > k, g([x]_R, A) \le k$  are (initial) semantics;

- (2) if  $s_1, s_2, s$  are semantics, then  $s_1 \wedge s_2, s_1 \vee s_2, \neg s$  are semantics;
- 464 (3) and the result obtained by only finite use of (1)(2) is semantics.

465

Definition 3.3 reflects the propositional logic feature of semantics construction. Rule (1) reflects the quantitative informa tion (initial) extraction and corresponds to four atomic propositions. Rule (2) reflects the quantitative information fusion and
 corresponds to the (complete) logical connectives. Rule (3) reflects the action finiteness. Thus, semantics becomes the prop ositional formula.

#### **Theorem 3.4.** Semantics system $(S, \land, \lor, \neg)$ corresponds to the propositional logic.

Theorem 3.4 deeply exhibits the logic essence of the semantics system in PG-Approx-Space. Thus, semantics has a strong basis in logic, and the propositional logic theory provides effective descriptions. For example, semantics has several forms, such as the above *s*, with the three levels for the three measures, the precision semantics with the one level for only one measure, and others with multiple levels according to a superficial form. Thus, the semantics is not necessarily the threelevel form. From the logic perspective, however, the normal form of the three-level form has great significance for the semantics system, which will be discussed in Section 4.1.

477 **Definition 3.5.** In PG-Approx-Space, the granule completely bearing certain semantics is called the corresponding 478 **semantics granule**. In other words, if  $\forall s \in S$ , then  $sg = \bigcup\{[x]_R : s\}$  is the semantics granule related to s. Moreover, the 479 semantics granule set is noted as  $SG = \{sg\}$ .

480 The semantics granule includes/extends the region notion of the VPRS-Model, GRS-Model and Pawlak-Model, and thus 481 underlies the PG-Approx-Space descriptions and D-Quantification applications. In particular, there is a one-to-one correspondence between the semantics and semantics granule. In fact, semantics acts as the semantics granular connotation 482 483 and define the set extension, while on the contrary, the semantics granule serves as the exact extension embodiment of 484 the semantics connotation. Moreover, the semantics granule construction and acquisition also reflect the granule merging process based on the fused quantitative information, i.e., atom granules that satisfy certain fused semantics conditions 485 merge into the semantics granule. According to the definition, the semantics granule can be determined and computed 486 by the semantics connotation. For example, for the above semantics s., the corresponding semantics granule is 487 488  $sg_* = \cup \{[x]_R : [\beta < p([x]_R, A) < 1 - \beta] \land [\overline{g}([x]_R, A) > k] \land [g([x]_R, A) > k]\}; \text{ thus, } sg_* \text{ is well-defined, merged and calculable.}$ 

Finally, we establish an in-depth isomorphism theory to probe the relationship between the semantics system and the semantics granule system and further give two essential mathematical structures for the semantics granule system.

491 **Theorem 3.6.** (System Isomorphism Theory). The semantics system  $(S, \land, \lor, \neg)$  and semantics granule system  $(SG, \cap, \cup, \sim)$  are 492 isomorphic.

### 20 September 2013

INS 10301

# ARTICLE IN PRESS

Q1 10

496

X. Zhang, D. Miao/Information Sciences xxx (2013) xxx-xxx

493 **Proof.** Suppose  $\varphi: S \to SG, \forall s \in S, \varphi(s) = sg = \cup\{[x]_R:s\}$ , then  $\varphi$  is well-defined.  $\varphi$  is obviously a one-to-one mapping. 494  $\forall s_1, s_2, s \in S, \varphi(s_1 \land s_2) = \cup\{[x]_R: s_1 \land s_2\} = (\cup\{[x]_R:s_1\}) \cap (\cup\{[x]_R:s_2\}) = \varphi(s_1) \cap \varphi(s_2)$ . Similarly,  $\varphi(s_1 \lor s_2) = \varphi(s_1) \cup \varphi(s_2)$ ; 495  $\varphi(\neg s) = \cup\{[x]_R:\neg s\} = \sim \varphi(s)$ . Thus, the two systems are isomorphic.  $\Box$ 

In fact, the set operations closely correspond to the logical operations, thus clarifying Theorem 3.6. Moreover, its proof reflects the equivalence between the granular merging based on the whole fused semantics and the set operation based on the partial semantics granules. More deeply, Theorem 3.6 reflects the harmony between the quantitative information fusion and application region calculation, i.e., the logical AND OR NOT operations in the semantic connotation level correspond to the set intersection union complementary operations in the granular extension level, and the logical connectives  $\rightarrow$ ,  $\leftrightarrow$ correspond to the set relationships  $\subset$  ,=, respectively.

For the semantics granule, the semantics construction principle and system isomorphism theory reflect the D-Quantification essential construction based on the VPRS and GRS approximations and finite set operations; hence, the PG-Approx-Space study highly and reasonably aligns with RS-Theory. Thus, the D-Quantification issue regarding quantitative information fusion has been perfectly resolved, and the fusion technologies with the normal operations become reasonable and necessary, particularly from the operator-oriented view [51] (this view holds as RS-Theory has not changed the classical operations and has developed the classical set theory only by adding the approximation operators).

Theorem 3.7. System Structure TheoryThe semantics granule system  $SG = \{sg\}$  constructs both a topology and a  $\sigma$  algebra on U, while the semantics system  $S = \{s\}$  also corresponds to a topology and a  $\sigma$  algebra.

511

Theorem 3.7 reflects the perfect mathematical structure feature of the semantics granule system. Thus, the research on the universe can mainly depend on the semantics granule system to develop the structural description and probabilistic measurement. In fact, this theorem not only underlies the D-Quantification applications but also provides scientific evidence for the region applications of RS-Theory. According to the inclusiveness, we partly provide the sub-topology and  $\sigma$  sub-algebra of the VPRS, i.e., { $\phi, U, \overline{R}_{\beta}A, \underline{P}_{\beta}A, negR_{\beta}A, negR_{\beta}A, negR_{\beta}A, ...$ }, which is only a subset of  $SG = \{sg\}$ . Moreover, { $posR_{\beta}A, negR_{\beta}A, negR_{\beta}A, bnR_{\beta}A$ } serves as a corresponding topological base.

For D-Quantification, this section deeply investigates quantitative information extraction and fusion to scientifically construct the semantics and construct the semantics granules with perfect mathematical structures in the application layer. The research results are in accordance with the RS-Theory set essence, and establish the scientific application framework of D-Quantification. Accordingly, the surplus sections explore some detailed contents, such as the semantic normal form, multiple granules and the RS-Model.

#### 523 **4. Basic semantics and three types of concrete granules**

#### 4.1. Basic semantics (BS) and basic semantics granule (BS-Granule)

Section 3 provides the semantics and semantics granules. Furthermore, this section focuses on the disjunctive normal form of semantics and the corresponding granules. According to logic theory, semantics must be represented by the three-level conjunctive form. From another perspective, the three-level elementary forms have basis and diversity and can reflect the exact three-measure description and PG-Approx-Space feature. Thus, the relevant decomposition becomes feasible and valuable. For simplicity, some explicit descriptions are directly used. Suppose  $p = p([x]_R, A), \overline{g} = \overline{g}([x]_R, A), g = g([x]_R, A)$ .

530 Theorem 4.1. In PG-Approx-Space, there are 10 and only 10 types of basic semantics on the three measures, and Table 1 shows the 531 relevant results.

**Proof.** According to the semantics construction principle, only finite types of semantics exist, and the basic types of the three measures can be constructed. There are three complete precision intervals:  $[0, \beta]$ ,  $(\beta, 1 - \beta)$ ,  $[1 - \beta, 1]$ . Similarly, there are two complete intervals of internal grade and external grade, respectively, i.e., [0, k],  $(k, +\infty)$ . Thus, there are 12 complete combinations of the three measures, and only two do not exist because of the mathematical properties in Proposition 2.3, i.e.,  $(1)\overline{g} > k$ ,  $g \le k$ ,  $p \le \beta$ ;  $(2)\overline{g} \le k$ , g > k,  $p \ge 1 - \beta$ .  $\Box$ 

Table 1	
---------	--

BS/BS-Granules.	

BS/BS-Granule	Precision semantics	Grade semantics	BS/BS-Granule	Precision semantics	Grade semantics
(1)	$p\leqslant eta$	$\overline{g} > k, \ \underline{g} > k$	(6)	$p\leqslant eta$	$\overline{g} \leqslant k, \ \underline{g} \leqslant k$
(2)	$eta$	$\overline{g} > k, \ \underline{g} > k$	(7)	$eta$	$\overline{g} \leqslant k, \ \underline{g} \leqslant k$
(3)	$p \ge 1 - \beta$	$\overline{g} > k, \ \underline{g} > k$	(8)	$p \geqslant 1 - eta$	$\overline{g} \leqslant k, \ \underline{g} \leqslant k$
(4)	$\beta$	$\overline{g} > k, \ \underline{g} \leqslant k$	(9)	$p\leqslant eta$	$\overline{g} \leqslant k, \ \underline{g} > k$
(5)	$p \geqslant 1 - \beta$	$\overline{g} > k, \ \underline{g} \leqslant k$	(10)	$eta$	$\overline{g}\leqslant k, \ \underline{g}>k$

Q1

### INS 10301 20 September 2013

01

557

567

### ARTICLE IN PRESS

11

X. Zhang, D. Miao/Information Sciences xxx (2013) xxx-xxx

**Definition 4.2.** The 10 types of basic semantics using the three measures are simply called basic semantics and are noted as 537 BS. Furthermore, the corresponding semantics granules are called basic semantics granules and are noted as BS-Granules. 538 As the semantics and semantics granule have strong similarities, the BS and BS-Granule are also very similar except for 539 540 the semantics essence and set nature. Table 1 provides the 10 types of BS/BS-Granules where BS serves as the main form. For 541 example, BS (2) is that precision with respect to A is in  $(\beta, 1 - \beta)$ , and both internal grade and external grade with respect to A are 542 greater than k. BS-Granule (2) has the exact merging result on atom-granules with BS 2), i.e.,  $\cup \{[x]_R, A) \in (\beta, 1-\beta),$  $\overline{g}([x]_{\mathbb{R}},A) > k, g([x]_{\mathbb{R}},A) > k\}$ , and it is actually the previous semantics granule  $s_{\mathbb{R}}$  (in Section 3). BS-Granule (2), as the set 543 544 extension, has the semantics connotation of BS (2). Obviously, the BS/BS-Granule has the direct fusion description of the 545 three measures with both the ratio and the number information. In other words, we can visually observe the fusion of 546 the relative and absolute information on precision and grade. In particular, BS and BS-Granules deeply reflect the double fault description features as well, and the double fault types are basic and complete. 547

In the propositional logic view, BS originates from the principal disjunctive normal form on  $SG = \{sg\}$ . Because of the four atomic propositions, there are 16 types of minterms with the four-level form; however,  $p > \beta$  and  $p \ge 1 - \beta$  have the correlation, and  $p, \overline{g}, \overline{g}$  have a decisive relationship. Therefore, only 10 types of minterms exist. BS takes the three-level elementary conjunctive form and is actually the minterm. Moreover, BS-Granules serve as the supporting sets of the semantics minterms.

**Proposition 4.3.** In the semantics logic system  $(S, \land, \lor, \neg, True, False)$ , suppose  $BS_0, \ldots, BS_{10}$  are the 10 types of BS, then  $BS_0 \lor \ldots \lor BS_{10} = True$  and  $BS_i \land BS_i = False$ , where  $i, j \in \{1, \ldots, 10\}$  and  $i \neq j$ .

**Theorem 4.4.** BS MeritRegarding the semantics construction principle, the non-separable BS serves as the finest/smallest semantics and can represent any semantics using the disjunction form.

Theorem 4.4 becomes clear based on the semantics mechanism and the principal disjunctive normal form. Therefore, the in-depth BS significance for semantics is reflected. BS has the basic representation feature and strong construction ability, thus providing a unified description framework for semantics. Accordingly, the semantic extraction connected with BS has important value. According to the system isomorphism theory (Theorem 3.6), granular semantics on BS corresponds to the granular relationship with BS-Granules. In other words, granular decomposition on BS-Granules can completely determine its semantics on BS and is therefore an important GrC issue in PG-Approx-Space.

Theorem 4.5. BS-Granules Merit10 BS-Granules, as a partition, act as a base for the topology on the semantics granule set (i.e., topology  $SG = \{sg\}$ ). An arbitrary semantics granule is the merging of BS-Granules, and its semantics is the disjunction of the BS extracted by the composed BS-Granules.

Based on Theorem 4.5, for a semantics granule, its semantics can be easily extracted if the relationships between it and the BS-Granules are clear such that the extraction problem of granular semantics completely changes into the study of granular relationships with respect to BS-Granules. With respect to a semantics granule, the relationships between it and BS-Granules must be given, while its semantics may be omitted. Accordingly, the research focus changes from a discussion of semantics to a study of granules where BS-Granules play a core role.

573 BS-Granules related to BS can represent semantics granules and can quickly extract semantics. Thus, BS-Granules estab-574 lish an important foundation for semantics granules calculations and applications, and thus become a fundamental type of 575 granule with a core position in PG-Approx-Space.

In Measure Space, BS-Granules correspond to the classified spaces of the surface Z = Y/(X + Y) with respect to the four planes: X = k, Y = k,  $Z = \beta$ ,  $Z = 1 - \beta$ . Fig. 2 shows the 10 BS-Granules in a homeomorphic plane on Measure Space.



Fig. 2. 10 BS-Granules in a homeomorphic plane on Measure Space.

Q1 Please cite this article in press as: X. Zhang, D. Miao, Quantitative information architecture, granular computing and rough set models in the double-quantitative approximation space of precision and grade, Inform. Sci. (2013), http://dx.doi.org/10.1016/j.ins.2013.09.020

Q1 12

X. Zhang, D. Miao / Information Sciences xxx (2013) xxx-xxx



Fig. 3. BS-Granules in Measure Plane.

In Measure Plane, BS-Granules correspond to the classified regions with respect to the four lines: Y = k, Y = V - k,  $Y = \beta V$ ,  $Y = (1 - \beta)V$ . As  $\beta < 0.5 < 1 - \beta$ , "Z = 0.5 when X = Y" (in Proposition 2.3) concludes, the intersection point of Y = k and Y = V - k must be located in the range between  $Y = \beta V$  and  $Y = (1 - \beta)V$ . This conclusion ensures the stability of the qualitative analyses of BS-Granules. In other words, the BS-Granular geometric form and structural system do not depend on the usual thresholds. On the contrary, the concrete case fully reflects the qualitative results. Fig. 3 shows the distribution of BS-Granules (where k = 10 and  $\beta = 0.25$ ), and they may be correspondingly labeled.

The space/plane form and structures of BS-Granules reflect the combining and coarsening processes of atom-granules in PG-Approx-Space. Furthermore, the BS-Granules exhibit basic structures and objective forms on the double-quantitative semantics in PG-Approx-Space. BS-Granules can be computed by the definition, and they can also be easily constructed/computed by B-Granules or C-Granules, which are proposed herein.

#### 588 4.2. Basic granules (B-Granules)

BS and BS-Granules are related to semantics and originate from the three application measures. However, the two core measures  $|[x]_R|$  and  $|[x]_R \cap A|$  have calculation directness and optimization. Thus, a basic granulation of BS-Granules with respect to the core measures will be conducted here.



#### 20 Septe

01

#### X. Zhang, D. Miao/Information Sciences xxx (2013) xxx-xxx

592 For semantics granules, there are only finite types of descriptions of  $|[x]_R|$  and  $|[x]_R \cap A|$  according to the most accurate con-593 dition. To produce these descriptions, BS-Granules only need fine and complete decomposing on the two measures. Thus, we provide and analyze them in Measure Plane. As is obtained in Section 4.1, BS-Granules are formed by the four lines: 594 595 Y = k, Y = V - k,  $Y = \beta V$ ,  $Y = (1 - \beta)V$ . Furthermore, there are only five intersection points (except the origin) determined 596 by the four lines, so the exhaustive descriptions on  $|[x]_R|$  and  $|[x]_R \cap A|$  can be produced by all the vertical and horizontal lines that pass through the five points, i.e., three vertical lines  $V = k/(1-\beta)$ , V = 2k,  $V = k/\beta$  and three horizontal lines 597  $Y = \beta k/(1-\beta)$ , Y = k,  $Y = (1-\beta)k/\beta$ . Accordingly, a complete network with 16 regions is formed, which corresponds to 598 599 the simplest form of  $||x|_R|$  and  $||x|_R \cap A|$ . Furthermore, considering BS-Granules or the initial four lines, the new granules 600 emerge (only 30 types exist in theory). Fig. 4 provides the qualitative result where k = 10 and  $\beta = 0.4$ .

601 **Definition 4.6.** The granules obtained by both BS-Granules and the full descriptions on  $|[x]_R|$ ,  $|[x]_R \cap A|$  are called basic 602 granules (**B-Granules**).

#### 603 **Proposition 4.7.** B-Granules are the decompositions of BS-Granules, while BS-Granules are the merging of B-Granules.

The granulation of B-Granules is actually conducted by both BS-Granules and the network of the six line subdivision. Thus, the relationships between BS-Granules and B-Granules become clear. In another view, B-Granules correspond to the subdivision regions of the nine lines, which accords with both BS-Granules and the desire of the exhaustive expression of  $|[x]_{R}|$  and  $|[x]_{R} \cap A|$ .

The 30 B-Granules can be accurately described, and the concrete relationships between them and the 10 BS-Granules also become clear. For example, BS-Granule (2),

612  $\cup \{ [\mathbf{x}]_{R} : p([\mathbf{x}]_{R}, A) \in (\beta, 1 - \beta), \overline{\mathbf{g}}([\mathbf{x}]_{R}, A) > k, g([\mathbf{x}]_{R}, A) > k \},$ 

613 is actually decomposed into three B-Granules such that

$$\begin{split} & \cup \{ [x]_R : |[x]_R| \in (2k, k/\beta], |[x]_R \cap A| \in (k, |[x]_R| - k) \}, \\ & \cup \{ [x]_R : |[x]_R| > k/\beta, |[x]_R \cap A| \in (\beta |[x]_R|, (1-\beta)k/\beta] \}, \end{split}$$

$$\cup \{ [x]_R : |[x]_R| > k/\beta, |[x]_R \cap A| \in ((1-\beta)k/\beta, (1-\beta)|[x]_R|) \}.$$

<sup>617</sup> Here, the B-Granular descriptions exhibit the simplest form, where  $|[x]_R|$  is described first and  $|[x]_R \cap A|$  is described second. <sup>618</sup> According to  $\beta$  and k, the parameters become constants in the first process while several variables may inevitably exist in the <sup>619</sup> second.

Theorem 4.8. B-Granules MeritAccording to  $|[x]_R|$  and  $|[x]_R \cap A|$ , B-Granules – as a partition – serve as basic granules for semantics granules, and any semantics granule can be constructed accordingly by them.

622

610

614

616

 $|[x]_R|$  and  $|[x]_R \cap A|$  act as the finest/smallest information in Approx-Space. Based on BS-Granules, B-Granules have adequately used  $|[x]_R|$  and  $|[x]_R \cap A|$ , which accords with the granulation process. Thus, Theorem 4.8 reflects the fundamental position of B-Granules. Because of adequately utilizing the core data information, B-Granules become a basic type of granule, particularly in the description and calculation aspects of PG-Approx-Space. Moreover, as B-Granules can be directly computed, the studies on both BS-Granules and macroscopic regions become operable and easy by them.

#### 628 4.3. Calculation granules (C-Granules)

For semantics granules, their decomposition on BS-Granules can be implemented, but the process is not simple. However, it is relatively easy to describe the process by  $|[x]_R|$  and  $|[x]_R \cap A|$ . Accordingly, their concrete calculations can be successfully realized by B-Granules. This method is supported by B-Granules and has strong operability and good space complexity. However, the time complexity can be further decreased because of the relatively large number of B-Granules. Therefore, we propose a special type of granule for optimal calculations by a type of half decomposition on  $|[x]_R|$  and  $|[x]_R \cap A|$ . For BS-Granules and B-Granules, the new granules are actually located in the middle layer and have stronger operability and computability.

#### 636 **Proposition 4.9.**

$$\cup\{[x]_{R}: p([x]_{R}, A) \in (\beta, 1-\beta), \overline{g}([x]_{R}, A) > k, \underline{g}([x]_{R}, A) > k\} = (\cup\{[x]_{R}: |[x]_{R}| \in (2k, k/\beta], |[x]_{R} \cap A| \in (k, |[x]_{R}| - k)\}$$

640

642 
$$\cup (\cup \{ [x]_R : |[x]_R| > k/\beta, |[x]_R \cap A| \in (\beta |[x]_R|, (1-\beta)|[x]_R|) \}.$$

643

#### Q1 14

666 667 669

670 **672** 

Table 2	
C-Granules.	

X. Zhang, D	. Miao / Information	Sciences xxx	(2013) xxx-xxx
-------------	----------------------	--------------	----------------

C-Granule	$ [x]_R $	$ [x]_R \cap A $	BS-Granule
(1)	$\leq k/(1-eta)$	$< [\mathbf{x}]_{R} -k$	(9)
(2)	$\leq k/(1-eta)$	$[ [x]_R  - k, \beta [x]_R ]$	(6)
(3)	$\leq k/(1-eta)$	$(\beta   [x]_R  , (1 - \beta)   [x]_R  )$	(7)
(4)	$\leq k/(1-eta)$	$[(1-\beta) [\mathbf{x}]_{\mathbf{R}} ,\mathbf{k}]$	(8)
(5)	$\leq k/(1-\beta)$	> k	(5)
(6)	(k/(1-eta),2k]	$\leq \beta  [\mathbf{x}]_R $	(9)
(7)	(k/(1-eta),2k]	$(\beta [x]_R ,  [x]_R  - k)$	(10)
(8)	(k/(1-eta),2k]	$[ [x]_R  - k, k]$	(7)
(9)	(k/(1-eta),2k]	$(k, (1-\beta) [x]_R )$	(4)
(10)	(k/(1-eta),2k]	$\geq (1-\beta) [\mathbf{x}]_{R} $	(5)
(11)	$(2k, k/\beta]$	$\leq \beta  [\mathbf{x}]_R $	(9)
(12)	$(2k, k/\beta]$	$(\beta [\mathbf{x}]_{R} , k]$	(10)
(13)	$(2k, k/\beta]$	$(k,  [x]_R  - k)$	(2)
(14)	$(2k, k/\beta]$	$[ [x]_{R}  - k, (1 - \beta) [x]_{R} )$	(4)
(15)	$(2k, k/\beta]$	$\geq (1-\beta) [\mathbf{x}]_R $	(5)
(16)	> k/eta	$\leq k$	(9)
(17)	> k/eta	$(k, \beta [x]_R ]$	(1)
(18)	> k/eta	$(\beta   [x]_R  , (1 - \beta)   [x]_R  )$	(2)
(19)	> k/eta	$[(1 - \beta) [x]_R ,  [x]_R  - k)$	(3)
(20)	>k/eta	$\geqslant  [x]_R  - k$	(5)

 $\textbf{Proof.} \ \ G = \cup \{[x]_{\mathcal{R}} : p([x]_{\mathcal{R}}, A) \in (\beta, 1 - \beta), \overline{g}([x]_{\mathcal{R}}, A) > k, g([x]_{\mathcal{R}}, A) > k\} = \cup \{[x]_{\mathcal{R}} : |[x]_{\mathcal{R}} \cap A| > k, |[x] >_{\mathcal{R}} \cap A| < |[x]_{\mathcal{R}}| - k, |[x] \ . \ In \ G, k \in [\alpha, \beta] \}$ 644  $<|[x]_R \cap A| < |[x]_R| - k$ ; thus,  $|[x]_R| > 2k$ , a necessary range, is obtained. Furthermore, the relationships of the relevant param-645 eters are discussed as follows. (1) If  $\beta |[x]_R| \leq k$ , i.e.,  $|[x]_R| \in (2k, k/\beta]$ , then  $(1 - \beta) |[x]_R| \geq |[x]_R - k$ ,  $|[x]_R \cap A| \in (k, |[x]_R| - k)$ ; (2) if 646  $\beta|[\mathbf{x}]_R| > k$ , i.e.,  $|[\mathbf{x}]_R| > k/\beta$ , then  $(1 - \beta)|[\mathbf{x}]_R| < |[\mathbf{x}]_R| - k$ ,  $|[\mathbf{x}]_R \cap A| \in (\beta|[\mathbf{x}]_R|, (1 - \beta)|[\mathbf{x}]_R|)$ . Thus, the accurate descriptions of the 647 two subsets appear as  $G_1 = \cup \{[x]_R : |[x]_R| \in (2k, k/\beta], |[x]_R \cap A| \in (k, |[x]_R| - k)\}, G_2 = \cup \{[x]_R : |[x]_R| > k/\beta, |[x]_R \cap A| \in (\beta | [x]_R|, k/\beta)\}$ 648  $(1 - \beta)|[x]_R|)$ . Moreover,  $G = G_1 \cup G_2$  and  $G_1 \cap G_2 = \phi$ .  $\Box$  In Proposition 4.9, G is actually BS-Granule (2), and BS-Granule 649 (2) is decomposed into two sub-granules according to the concrete conditions of  $|[x]_R|$  and  $|[x]_R \cap A|$ . The two sub-granules 650 have a simple form about  $|[x]_R|$  but a complex one about  $|[x]_R \cap A|$ . Proposition 4.9 and its proof provide an operational scheme 651 of the calculation decomposition of BS-Granules, which mainly concerns the mathematical deduction with respect to  $|[x]_p|$ 652 and  $|[x]_R \cap A|$ . In Measure Plane, the two sub-granules correspond to the subdivision regions of BS-Granule (2) by the vertical 653 line such that  $V = k/\beta$ . Generally, all BS-Granules can be similarly decomposed, which corresponds to the division by the 654 three vertical lines such that  $V = k/(1 - \beta)$ , V = 2k,  $V = k/\beta$ . As a result, a new type of granule is produced. 655

**Definition 4.10.** The sub-granules of BS-Granules that are completely decomposed by the relationships between  $|[x]_R|$  and  $k/(1-\beta), 2k, k/\beta$  are called calculation granules (**C-Granules**).

The mathematical forms and structures of C-Granules correspond to both BS-Granules (i.e., the four line subdivision) and the three vertical line subdivision, or the seven lines subdivision. Thus, the relationships between C-Granules and BS-Granules/B-Granules become clear. These results are shown in the relevant figures, such as Fig. 4.

#### 661 **Proposition 4.11.** C-Granules are the decompositions of BS-Granules or the merging of B-Granules.

For C-Granules, Proposition 4.9 shows the decomposition process of BS-Granules. Here, a merging process of B-Granules
 is also illustrated as BS-Granule (2) is the merging of three B-Granules:

$$\cup\{[x]_R: |[x]_R| \in (2k, k/\beta], |[x]_R \cap A| \in (k, |[x]_R| - k)\}$$

$$\cup \{ [x]_{R} : |[x]_{R}| > k/\beta, |[x]_{R} \cap A| \in (\beta |[x]_{R}|, (1-\beta)k/\beta] \},$$

$$\cup \{ [x]_{R} : |[x]_{R}| > k/\beta, |[x]_{R} \cap A| \in ((1-\beta)k/\beta, (1-\beta)|[x]_{R}|) \}$$

The first one is a C-Granule, and the latter two are merged into a C-Granule, where  $|[x]_R \cap A| \in (\beta | [x]_R |, (1 - \beta)k/\beta]$  and  $|[x]_R \cap A| \in ((1 - \beta)k/\beta, (1 - \beta)|[x]_R|)$  are integrated into a new condition such that  $|[x]_R \cap A| \in (\beta | [x]_R |, (1 - \beta)|[x]_R|)$ . This C-Granule is  $G_2$  (in Proposition 4.9 and its proof).

In theory, only 20 types of C-Granules exist. Table 2 provides a list of these C-Granules as well as the relationships between the C-Granules and BS-Granules. For example, C-Granule (9) is  $\cup \{[x]_R : |[x]_R] \in (k/(1-\beta), 2k], |[x]_R \cap A| \in (k, (1-\beta)|[x]_R])\}$ . Meanwhile, two examples are provided to show the granular relationships. C-Granules (9) and (14) are the decomposition of BS-Granule (4), while BS-Granule (5) is the merging of C-Granules (5)(10)(15)(20). Based on the granulation or figure, the relationships between C-Granules and B-Granules become clear and can be shown in Table 2 if the B-Granules have the relevant labels.

Q1

X. Zhang, D. Miao/Information Sciences xxx (2013) xxx-xxx

C-Granules originate from the mathematical deduction of BS-Granules. Their description first utilizes the relationships 682 683 between  $|[x]_{R}|$  and the three constant parameters  $k/(1-\beta)$ ,  $2k, k/\beta$ , and second, it utilizes the relationships between  $|[x]_R \cap A|$  and the four parameters k,  $|[x]_R| - k$ ,  $\beta |[x]_R|$ ,  $(1 - \beta) |[x]_R|$ . In contrast to the complete decomposition in the B-Granules 684 685 study, both this deduction and description correspond to the half decomposition on  $|[x]_R|$  and  $|[x]_R \cap A|$ . However, in the B-Granules descriptions,  $|[x]_R|$  needs the same three parameters  $k/(1 - \beta)$ , 2k,  $k/\beta$ , while  $|[x]_R \cap A|$  needs seven parameters. Thus, 686 by ordering only seven parameters rather than 10, concrete granule calculations by C-Granules usually have more advanta-687 688 ges with respect to the time complexity. Moreover, the half decomposition is also easier for the mathematical deduction. 689 Therefore, C-Granules have significance for granular optimal calculations.

#### 690 5. Model regions and granular hierarchical structure

#### 691 5.1. model regions

In the Pawlak-Model, the classified three regions (the positive region, negative region, and boundary region) represent qualitative semantics and underlie practical applications. In the VPRS-Model,  $\underline{R}_{\beta}A \subseteq \overline{R}_{\beta}A$ , and the corresponding regions also exist. However, in the GRS-Model, the regions become complex because  $\underline{R}_kA \subseteq \overline{R}_kA$  in the general case. Here, the traditional regions are first developed in PG-Approx-Space, and the relevant notions underlie D-Quantification applications such that the RS-Model is generally introduced into PG-Approx-Space. Furthermore, we will produce a basic region system, and perform the semantic extraction and concrete calculation.

698 **Definition 5.1.**  $\forall \overline{R}_{PG}A, \underline{R}_{PG}A \in SG$  define the upper and lower approximations, respectively. Furthermore, 699

**704** 705

707

726

$$posR_{PG}A = \overline{R}_{PG}A \cap \underline{R}_{PG}A, negR_{PG}A = \sim (\overline{R}_{PG}A \cup \underline{R}_{PG}A),$$

 $ubnR_{PG}A = \overline{R}_{PG}A - R_{PG}A, lbnR_{PG}A = R_{PG}A - \overline{R}_{PG}A,$ 

 $bnR_{PG}A = \overline{R}_{PG}A \bigtriangleup R_{PG}A = ubnR_{PG}A \cup lbnR_{PG}A$ 

denote the positive region, negative region, upper boundary region, lower boundary region, boundary region, respectively. The model regions (noted as **M-Regions**) are defined as a general designation for the seven regions, i.e., the above five notions and the upper and lower approximations.  $(U, \overline{R}_{PG}, \underline{R}_{PG})$  denotes RS-Model (i.e., **Model**) in PG-Approx-Space, where  $\overline{R}_{PG}$ and  $\underline{R}_{PG}$  imply the relevant approximation operators.

712 In fact, we directly select two arbitrary semantics granules from SG to act as the approximations and further define the 713 model with the region system. This method becomes general because of the deleting of certain constraint conditions on the 714 approximations. The approximations proposed in Definition 5.1 adhere to the generalization property, and they can include 715 and simulate the usual and original approximations from the macroscopic set system perspective. Because of the complex relationships between the upper and lower approximations, we propose the positive and negative regions, upper and lower 716 boundary regions, and boundary region for the RS-Model in PG-Approx-Space. In fact, they all develop the traditional re-717 gions. In the Pawlak-Model, as  $posRA = \underline{R}A = \overline{R}A \cap \underline{R}A$ , the positive region serves as the intersection set of the approximations. 718 With respect to the other two notions,  $negRA = \overline{RA} = \overline{RA} = \overline{RA} \cup RA$  and  $bnRA = \overline{RA} - RA = (\overline{RA} - RA) \cup (RA - \overline{RA})$ , and thus 719 similar analysis results appear. In the double-quantitative environment, in particular, the boundary region is divided into 720 721 two new notions: the upper and lower boundary regions. Moreover, the approximation operators become an abstract notion because of the abstract approximation notion. 722

## 723 **Proposition 5.2.**

### $\overline{R}_{PG}A = posR_{PG}A \cup ubnR_{PG}A, \underline{R}_{PG}A = posR_{PG}A \cup lbnR_{PG}A.$

Definition 5.1 and Proposition 5.2 show the system feature of the seven M-Regions. In RS-Theory, the classified regions act as fun damental notions. According to this idea, four fundamental M-Regions emerge.

Definition 5.3. The classified four-regions (i.e., the positive region, negative region, upper boundary region, lower boundary region) are called basic model regions and are noted as BM-Regions.

The structures and relationships of both BM-Regions and M-Regions become clear based on Definition 5.1 and Proposition
 5.2. Additionally, the four BM-Regions are obviously classified. Fig. 5 reflects these results and the two complete systems.
 Compared to the usual three-region classification, the four-region classification becomes a new and basic feature of the usual
 RS-Model in PG-Approx-Space, a distinctive result that is mainly attributed to the D-Quantification completeness.

More deeply, M-Regions/BM-Regions have concrete semantics and correspond to a coarsening sub-system of the semantics granule system (SG,  $\cap$ ,  $\cup$ ,  $\sim$ ). The seven M-Regions become a part of a sub-topology of the topology  $SG = \{sg\}$  in Theorem 3.7, and the four BM-Regions serve as a corresponding topological base, which is similar to the BS-Granules base in Theorem 4.5. Accordingly, M. Pogiones and PM. Pogiones have important relationships with companying and PS. Granules have

4.5. Accordingly, M-Regions and BM-Regions have important relationships with semantics granules and BS-Granules. How-

### 20 September 2013

### **ARTICLE IN PRESS**

X. Zhang, D. Miao / Information Sciences xxx (2013) xxx-xxx



Fig. 5. BM-Regions and M-Regions.

ran ever, when compared to semantics granules and BS-Granules, M-Regions and BM-Regions actually result in the main and refined construction for applications.

#### 741 Theorem 5.4. M-Regions/BM-Regions are merging of BS-Granules, B-Granules and C-Granules.

As semantics granules, M-Regions/BM-Regions have semantics and accordingly, they directly correspond to model appli-742 cations. As a result, both their semantic extraction and optimal calculations become important problems for which there are 743 744 two main thoughts. (1) Their calculation formulas can be directly provided by the approximations. If the approximation description refers to precision and grade, then the semantic extraction requires logic representation, which corresponds 745 to the decomposition of BS-Granules. If the description refers to  $|[x]_R|$  and  $|[x]_R \cap A|$ , then the approximations may be first 746 decomposed into B-Granules or C-Granules, and all M-Regions further obtain their description formulas and semantics by 747 748 B-Granules or C-Granules. (2) On the other hand, M-Regions can be constructed by BM-Regions. If BM-Regions obtain their decomposition on B-Granules or C-Granules, then we can completely and easily solve the problems about the concrete cal-749 750 culation and semantic extraction of M-Regions/BM-Regions. Thus, the second method is adopted, and only C-Granules are 751 used for optimal calculations.

**Definition 5.5.** For the given RS-Model  $(U, \overline{R}_{PG}, \underline{R}_{PG})$ , C-Granules included in each BM-Region can be merged into some new granules by integrating the conditions between  $|[x]_R|$ ,  $|[x]_R \cap A|$  and the relevant parameters. These new granules are called calculation granules of basic model regions and are noted as **BMC-Granules**.

Similar to C-Granules, BMC-Granules are actually produced in two consistent ways. The definition originates from the merging of C-Granules by the similarity of conditions, and the other method originates from the mathematical deduction of BM-Regions on  $|[x]_R|$  and  $|[x]_R \cap A|$  (which is also a type of half decomposition).

### 758 **Theorem 5.6.** BMC-Granules are the merging of C-Granules/B-Granules and the decomposition of BM-Regions/M-Regions.

Theorem 5.6 reflects the important position of BMC-Granules. BMC-Granules act as an important bridge between macroscopic regions and microscopic C-Granules/B-Granules, and they have the optimal calculation function for M-Regions/BM-Regions. Moreover, M-Regions/BM-Regions semantics can also be obtained by tracking the relationships between BMC-Granules and C-Granules/B-Granules and BS-Granules. We provide the corresponding algorithm of M-Regions by BMC-Granules.

764 **Algorithm 1.** M-Regions Algorithm (based on BMC-Granules)

#### Input:

Approx-Space (U, R), concept *A*, thresholds  $\beta$ , *k*;

#### output:

M-Regions in  $(U, \overline{R}_{PG}A, \underline{R}_{PG}A)$ ; 1: Compute BMC-Granules;

- 2: Construct four BM-Regions;
- 3: Obtain  $\overline{R}_{PG}A$ ,  $\underline{R}_{PG}A$ ,  $bnR_{PG}A$ .
- 4: return all M-Regions.

#### 773 774

775

All M-Regions, BM-Regions and BMC-Granules depend on models and will be illustrated using a concrete model in Section 6.2.

Q1 Please cite this article in press as: X. Zhang, D. Miao, Quantitative information architecture, granular computing and rough set models in the double-quantitative approximation space of precision and grade, Inform. Sci. (2013), http://dx.doi.org/10.1016/j.ins.2013.09.020

**Q1** 16

01

### **ARTICLE IN PRESS**

17



Fig. 6. Granular hierarchical structure in PG-Approx-Space.

#### 5.2. Granular hierarchical structure

We can now summarize the granular hierarchical structure in PG-Approx-Space. Not including atom-granules and semantics granules, there are six types of granules, i.e., BS-Granules, B-Granules, C-Granules, M-Regions, BM-Regions and BMC-Granules. There are usually three main GrC strategies, i.e., top-down, bottom-up and middle-out. On the whole, we mainly adopt the bottom-up GrC strategy in PG-Approx-Space.

- (1) Basic mathematical forms of PG-Approx-Space mainly concern the granular structures of atom-granules in PG Approx-Space. Therefore, as they correspond to only Approx-Space, they are at a basic/atomic level.
- (2) BS-Granules originate from BS (which is connected with the semantic normal form). Furthermore, B-Granules and C-Granules originate from the complete and half decomposition of BS-Granules on  $|[x]_R|$  and  $|[x]_R \cap A|$ , respectively. Therefore, as all three types of granules refer to the thresholds, they are at a microscopic level. Moreover, the granulation method is also adopted here because new granules are mainly produced by BS-Granular decomposition.
- (3) Both M-Regions and BM-Regions originate from model applications and macroscopic descriptions. To compute them,
   BMC-Granules are further provided based on the merging of C-Granules or the half decomposition of BM-Regions. As
   all three types of granules depend on concrete models in PG-Approx-Space, they are at a macroscopic level.

The relationships among all six types of granules are actually clear based on the initial definitions and above studies.
 Therefore, we have correspondingly constructed the granular hierarchical structure in PG-Approx-Space. Fig. 6 provides this
 result where unidirectional paths with arrows denote the merging relationships among these types of granules. Moreover,
 BS-Granules, M-Regions and BM-Regions are actually semantics granules.

In PG-Approx-Space, the granular hierarchical structure provides the describing and solving methods of corresponding granular problems, such as semantics extraction and region calculations. For example, M-Regions/BM-Regions calculations can be completely implemented by the M-Regions algorithm, which essentially corresponds to the granular hierarchical structure. Furthermore, both in-depth studies and applications in PG-Approx-Space may make full use of this structure.

#### 801 6. Rough set model

791

PG-Approx-Space is macroscopically presented and described by models, while the latter build up important foundations
 for D-Quantification applications. In PG-Approx-Space, this section systematically investigates model expansion and provides a concrete double-quantitative model to explain certain relevant notions, such as BMC-Granules.

### ARTICLE IN PRESS X. Zhang, D. Miao / Information Sciences xxx (2013) xxx-xxx

Q1 18

#### 805 6.1. Model expansion

Within the semantics granules framework, M-Regions are systematically defined, and BM-Regions support the whole M-Regions system. Thus, the models are naturally determined and are formally represented by the approximation operators. In PG-Approx-Space, models with double-quantitative semantics usually have double fault tolerance capabilities, allowing them to adapt to more complex environments, and they also improve the previous models from a specific perspective. From another perspective, the Pawlak-Model, VPRS-Model and GRS-Model are actually models in PG-Approx-Space, and thus the models in PG-Approx-Space have strong expansionary capability. In this paper, model expansion is mainly discussed with respect to PG-Approx-Space.

813 We first analyze the expansion idea of the VPRS-Model and GRS-Model with respect to the Pawlak-Model. In the VPRS-814 Model,  $\beta \ge 0 \iff \overline{R}_{\rho}A \subseteq \overline{R}_{0}A$ ,  $\underline{R}_{\rho}A \supseteq \underline{R}_{0}A$ , while  $\overline{R}_{\rho}A - \underline{R}_{\rho}A \subseteq \overline{R}A - \underline{R}A$ , i.e.,  $bnR_{\rho}A \subseteq bnRA$ . The VPRS-Model degenerates into the 815 Pawlak-Model when  $\beta = 0$ . Thus, it is by both lessening the upper approximation and enlarging the lower approximation 816 that the VPRS-Model expands the Pawlak-Model and that the VPRS boundary region is lessened for the Pawlak boundary 817 region. When the GRS-Model expands the Pawlak-Model, the same approximation change emerges, i.e., the upper approx-818 imation is lessened while the lower approximation is enlarged. In fact, this approximation change in expansion will induce 819 the relative lessening of the boundary region, which suggests a benign expansion direction.

**Definition 6.1.** Benign expansion refers to the model expansion with both the lessened upper approximation and enlarged
 lower approximation.

#### 822 **Proposition 6.2.** Both the VPRS-Model and GRS-Model are benign expansions of the Pawlak-Model.

Here, benign expansion mainly considers the approximation expansion direction. According to this notion, if the upper approximations are lessened while the lower approximations are enlarged for both the VPRS-Model and GRS-Model, then a novel model will be produced to benignly expand the two basic models as well as the Pawlak-Model. This desire can be fulfilled in PG-Approx-Space.

831

830 
$$\overline{R}_{\beta \wedge k}A = \cup\{[x]_R : p([x]_R, A) > \beta, \overline{g}([x]_R, A) > k\},\$$

833 
$$\underline{R}_{\beta \lor k} A = \cup \{ [\mathbf{x}]_{R} : p([\mathbf{x}]_{R}, A) \ge 1 - \beta, \text{ or }, \underline{g}([\mathbf{x}]_{R}, A) \leqslant k \},$$

then  $(U, \overline{R}_{\beta \wedge k}, \underline{R}_{\beta \wedge k})$  is called the AND-OR model on precision and grade, noted as the **PG-AO-Model**.

### 835 **Proposition 6.4.**

838  $\overline{R}_{\beta \wedge k} A = \overline{R}_{\beta} A \cap \overline{R}_{k} A, \underline{R}_{\beta \vee k} A = \underline{R}_{\beta} A \cup \underline{R}_{k} A.$ 

In Definition 6.3, the  $\overline{R}_{\beta\wedge k}A$ ,  $\underline{R}_{\beta\wedge k}A$  semantics are defined by the logical AND OR operations of precision and grade, respectively, such that  $\overline{R}_{\beta\wedge k}A$ ,  $\underline{R}_{\beta\wedge k}A \in SG$ . Thus, the new approximations and model are defined. According to system isomorphism theory (Theorem 3.6), Proposition 6.4 further shows the semantics granular operation essence of the new approximations with respect to the basic VPRS and GRS approximations.

843 **Theorem 6.5.** The PG-AO-Model is benign expansion on the VPRS-Model, GRS-Model and Pawlak-Model.

Proof. (1)  $\overline{R}_{\beta}A = \overline{R}_{\beta\wedge0}A$ ,  $\underline{R}_{\beta}A = \underline{R}_{\beta\wedge0}A$ ;  $\overline{R}_{\beta\wedge k}A \subseteq \overline{R}_{\beta}A$ ,  $\underline{R}_{\beta\vee k}A \supseteq \underline{R}_{\beta}A$ . (2)  $\overline{R}_{k}A = \overline{R}_{0\wedge k}A$ ,  $\underline{R}_{k}A = \underline{R}_{0\vee k}A$ ;  $\overline{R}_{\beta\wedge k}A \subseteq \overline{R}_{k}A$ ,  $\underline{R}_{\beta\vee k}A \supseteq \underline{R}_{k}A$ . (3)  $\overline{R}A = \overline{R}_{0}A = \overline{R}_{0\wedge0}A$ ,  $\underline{R}A = \underline{R}_{0}A = \underline{R}_{0\vee0}A$ ;  $\overline{R}_{\beta\wedge k}A \subseteq \overline{R}A$ ,  $\underline{R}_{\beta\vee k}A \supseteq \underline{R}A$ .  $\Box$ 

Theorem 6.5 shows the higher position of the PG-AO-Model where the three basic models (VPRS-Model, GRS-Model, Pawlak-Model) become only three special cases when k = 0,  $\beta = 0$ ,  $\beta = 0 = k$ , respectively. Moreover, this new model constructs D-Quantification on precision and grade and has the logical double-quantitative semantics.  $\overline{R}_{\beta\wedge k}A$  is union of the equivalence classes whose precision with respect to A is greater than  $\beta$  and whose internal grade with respect to A is greater than k;  $\underline{R}_{\beta\wedge k}A$  is union of the equivalence classes whose precision with respect to A is not smaller than  $1 - \beta$  or whose external grade with respect to A is not greater than k. In fact, the double-quantitative semantics deeply reflect the double fault descriptions and fault tolerance features.

The PG-AO-Model acts as an excellent model in PG-Approx-Space. Furthermore, its upper and lower approximation operators can be used to construct more concrete models. Thus, a full model structure is produced in PG-Approx-Space.

**Definition 6.6.** Single expansion refers to the model expansion where only a single approximation is expanded; double expansion refers to the model expansion where the two approximations are both expanded.

X. Zhang, D. Miao/Information Sciences xxx (2013) xxx-xxx

19



Fig. 7. Some models in PG-Approx-Space and their (single) expansion relationships.

Proposition 6.7. Both the VPRS-Model and GRS-Model are double expansion on the Pawlak-Model, while the PG-AO-Model is a
 double expansion on the VPRS-Model, GRS-Model and Pawlak-Model.

859

We now construct some new models in PG-Approx-Space by  $\overline{R}_{\beta/k}$ ,  $\underline{R}_{\beta/k}$ , and study their relationships on single expansion. Fig. 7 shows the relevant results, such that if there is a direct arrow line, then the model in the upper layer acts as a single expansion on the one in the lower layer. Furthermore, we can extract the relationships on double expansion by the corresponding connected paths in Fig. 7. The results include only the three layers: (1)  $(U, \overline{R}, \underline{R})$ , (2)  $(U, \overline{R}_{\beta}, \underline{R}_{k})$ ,  $(U, \overline{R}_{\beta}, \underline{R}_{k})$ ,  $(U, \overline{R}_{\beta, \lambda}, \underline{R}_{\beta})$ , (3) $(U, \overline{R}_{\beta, \lambda}, \underline{R}_{\beta/\lambda})$ , where the expansion also becomes a benign expansion. Only for model expansion, Fig. 7 also provides rich hierarchical results by paths.

In particular, the model dimension and dimensional expansion can be discussed. As is given in Section 2, the Pawlak-Model, VPRS-Model, GRS-Model and the double-quantitative model usually correspond to 0-dimension, 1-dimension, 2-dimensions and 3-dimensions, respectively. Moreover, model expansion can also be explored in the threshold expansion view. For example, the Pawlak-Model has no parameters, both the VPRS-Model and GRS-Model have one parameter, the PG-AO-Model has two parameters, and models with more parameters can be further constructed, such as  $(U, \overline{R}_{\beta_1}, \underline{R}_{\beta_2}),$  $(U, \overline{R}_{k_1}, \underline{R}_{k_2})$  and  $(U, \overline{R}_{\beta_1 \wedge k_1}, \underline{R}_{\beta_2 \vee k_2}).$ 

#### 872 6.2. A concrete model

As the PG-AO-Model exhibits benign expansion, double expansion and dimensional expansion, has the double-quantitative semantics and double fault tolerance capabilities, it serves as a perfect model in PG-Approx-Space. This section mainly uses the PG-AO-Model to analyze the relevant granules and their relationships, such as BMC-Granules and BM-Regions/M-Regions.

In the PG-AO-Model, the approximations are first defined. Additionally, the BM-Regions/M-Regions are determined and can be described by the BMC-Granules. In fact, BMC-Granules can be obtained by the mathematical deduction of the BM-Regions/M-Regions on  $|[x]_R|$  and  $|[x]_R \cap A|$ . Table 3 presents all BMC-Granules by the computation and proof. For example, BMC-Granule (1) is  $\cup\{[x]_R : |[x]_R| \le 2k, |[x]_R \cap A| > k\}$ . Meanwhile, Table 3 also provides the relationships between

#### Table 3

BMC-Granule	$ [x]_R $	$ [x]_R \cap A $	C-Granule	BM-Region
(1)	$\leq 2k$	> k	(5)(9)(10)	posR <sub>PG</sub> A
(2)	$\leqslant 2k$	$<  [x]_R  - k$	(1)(6)(7)	negR <sub>PG</sub> A
(3)	$\leqslant 2k$	$[ [x]_{R}  - k, k]$	(2)(3)(4)(8)	lbnR <sub>PG</sub> A
(4)	$(2k, k/\beta]$	$\geq  [x]_R  - k$	(14)(15)	posR <sub>PG</sub> A
(5)	$(2k, k/\beta]$	$\leq k$	(11)(12)	negR <sub>PG</sub> A
(6)	$(2k, k/\beta]$	$(k,  [x]_R  - k)$	(13)	ubnR <sub>PG</sub> A
(7)	> k/eta	$\geq (1-\beta) [x]_R $	(19)(20)	posR <sub>PG</sub> A
(8)	>k/eta	$\leq \beta  [x]_R $	(16)(17)	negR <sub>PG</sub> A
(9)	>k/eta	$(\beta [x]_R , (1-\beta) [x]_R )$	(18)	ubnR <sub>PG</sub> A

Q1 20

X. Zhang, D. Miao/Information Sciences xxx (2013) xxx-xxx

BMC-Granules and BM-Regions/C-Granules. For example, BMC-Granule (1) is the merging of three C-Granules, C-Granules (5)(9)(10), but it acts as only a part of the positive region. Thus, 20 C-Granules are actually merged into only nine BMC-Granules, and the nine are further merged into four BM-Regions.

### Proposition 6.8.

887

889

 $posR_{PG}A = (\cup\{[x]_R : |[x]_R| \le 2k, |[x]_R \cap A| > k\})$  $\cup (\cup\{[x]_R : |[x]_R| \in (2k, k/\beta], |[x]_R \cap A| \ge |[x]_R| - k\})$  $\cup (\cup\{[x]_R : |[x]_R| > k/\beta, |[x]_R \cap A| \ge (1 - \beta)|[x]_R|\}.$ 

Proposition 6.8 shows the description of the positive region using BMC-Granules, i.e., the positive region is actually the merging of three BMC-Granules, BMC-Granules (1)(4)(7).

Here, BMC-Granules are mainly obtained by the half decomposition of the BM-Regions/M-Regions. The definition of BMC-Granules shows that they also correspond to the merging of C-Granules in BM-Regions, a point that is appropriately explained by the positive region herein. By calculations of the B-Granules and C-Granules, the positive region consists of seven C-Granules: C-Granules (5)(9)(10)(14)(15)(19)(20), but the seven are actually merged into only three classes, the BMC-Granules (1)(4)(7) in Propositions 6.8. and 6.9 and its proof illustrate only one merging process.

899
 897 Proposition 6.9. BMC-Granule (1) is the merging of three C-Granules: C-Granules (5)(9)(10).

Proof. Suppose  $G_5 = \bigcup \{[x]_R : |[x]_R| \le k/(1-\beta), |[x]_R \cap A| > k\}, G_9 = \bigcup \{[x]_R : |[x]_R| \in (k/(1-\beta), 2k], |[x]_R \cap A| \in (k, (1-\beta)|[x]_R|)\}, G_{10} = \bigcup \{[x]_R : |[x]_R| \in (k/(1-\beta), 2k], |[x]_R \cap A| \ge (1-\beta)|[x]_R|\}.$  Based on the conditions in the set descriptions,  $G_9 \cup G_{10} = \bigcup \{[x]_R : |[x]_R| \in (k/(1-\beta), 2k], |[x]_R \cap A| > k\} = G^*;$  then  $G_5 \cup G^* = \bigcup \{[x]_R : |[x]_R| \le 2k, |[x]_R \cap A| > k\} = G$ .  $G_5, G_9, G_{10}, G$  are actually C-Granules (5)(9)(10) and BMC-Granule (1), respectively. Therefore, the merging result is proved.  $\Box$ 

Based on BMC-Granules, the M-Regions algorithm has been completed in the PG-AO-Model. We now explain the semantic extraction by the positive region analysis. The positive region is the merging of BMC-Granules (1)(4)(7), which are composed of C-Granules (5)(9)(10)(14)(15)(19)(20). Thus, the positive region semantics is the disjunction of the BS on the three BMC-Granules or the seven C-Granules.

**Proposition 6.10.** The positive region semantics is the disjunction of BS (3)(4)(5).

911 **Proof.** This result can be proved by the granular relationships or by Table 2. □.

Other M-Regions semantics can be similarly extracted. Moreover, the approximations semantics accord with their direct logical double-semantics in Section 6.1.

### 915 **7. Conclusion**

912

916 D-Quantification exhibits the quantitative completeness and quantitative expansibility/inclusiveness and is novel and valuable, while PG-Approx-Space provides the basic supported space. This paper first constructs the space and plane 917 918 forms of PG-Approx-Space to rigorously organize and effectively store the quantitative information. Then, the semantics construction principle is specifically established to fulfill the quantitative information extraction and fusion, while the 919 920 system isomorphism theory shows that the measure fusion accords with the set operation. Moreover, the system structure theory underlies the model applications in RS-Theory. Accordingly, the semantics and semantics granules become 921 922 the core contents in PG-Approx-Space, and BS and BS-Granules are particularly achieved. According to GrC, all six types of granules (BS-Granules, B-Granules, C-Granules, M-Regions, BM-Regions, BMC-Granules) become fundamental for the 923 semantic, microscopic and macroscopic descriptions, and furthermore, they and their granular hierarchical structure dee-924 ply describe D-Quantification. For the RS-Model in PG-Approx-Space, the model regions and model expansion and the 925 concrete models and their structures build up important bases for D-Quantification applications. In summary, based 926 on the quantitative information architecture, this paper systematically conducts and investigates D-Quantification and 927 particularly establishes a fundamental and general exploration framework, and the results underlie D-Quantification 928 929 studies and applications.

Moreover, this paper conducts the fundamental measures studies on PG-Approx-Space based on the knowledge and concept, where the concept is viewed as only the set parameter. Furthermore, both the knowledge and concept systems are
 worth deeply exploring in PG-Approx-Space, while attribute reduction and rule extraction with D-Quantification become
 the next in-depth work.

#### 934 Acknowledgements

The authors thank both the editors and the anonymous referees for their valuable suggestions, which substantially improved this paper. This work was supported by the National Science Foundation of China (61203285 and 61273304), China

20 September 2013 01

X. Zhang, D. Miao/Information Sciences xxx (2013) xxx-xxx

- 21
- Postdoctoral Science Foundation Funded Project (2013T60464 and 2012M520930), and Shanghai Postdoctoral Scientific 937 938 Program (13R21416300).

#### 939 References

947

948

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

1000

1001

1002

- 940 [1] A. An, N. Shan, C. Chan, N. Cercone, W. Ziarko, Discovering rules for water demand prediction: an enhanced rough-set approach, Engineering 941 Application of Artifical Intelligence 9 (1996) 645-653.
- 942 N. Azam, J.T. Yao, Analyzing uncertainties of probabilistic rough set regions with game-theoretic rough sets, in: International Journal of Approximate 943 Reasoning. http://dx.doi.org/10.1016/j.ijar.2013.03.015.
- 944 [3] M. Beynon, Reducts within the variable precision rough sets model: a further investigation, European Journal of Operational Research 134 (2001) 592-945 605 946
  - [4] J.H. Dai, Rough set approach to incomplete numerical data, Information Sciences 241 (2013) 43-57.
  - Z.T. Gong, B.Z. Sun, Y.B. Shao, D.G. Chen, Q. He, Variable precision rough set model based on general relation, in: Proceedings of 2004 International Conference on Machine Learning and Cybernetics, 2004, pp. 2490-2494.
- 949 [6] S. Greco, B. Matarazzo, R. Słowiński, Parameterized rough set model using rough membership and bayesian confirmation measures, International 950 Journal of Approximate Reasoning 49 (2) (2008) 285-300. 951
  - J.P. Herbert, J.T. Yao, Game-theoretic rough sets, Fundamenta Informaticae 108 (2011) 267-286.
  - T.P. Hong, T.T. Wang, S.L. Wang, Mining fuzzy β-certain and β-possible rules from quantitative data based on the variable precision rough-set model, Expert Systems with Applications 32 (2007) 223-232.
  - [9] B. Huang, H.X. Li, D.K. Wei, Dominance-based rough set model in intuitionistic fuzzy information systems, Knowledge-Based Systems 28 (2012) 115-123
  - [10] K.Y. Huang, T.H. Chang, T.C. Chang, Determination of the threshold value  $\beta$  of variable precision rough set by fuzzy algorithms, International Journal of Approximate Reasoning 52 (7) (2011) 1056-1072.
  - M. Inuiguchi, Y. Yoshioka, Y. Kusunoki, Variable-precision dominance-based rough set approach and attribute reduction, International Journal of Approximate Reasoning 50 (8) (2009) 1199-1214.
  - [12] X.Y. Jia, W.H. Liao, Z.M. Tang, L. Shang, Minimum cost attribute reduction in decision-theoretic rough set models, Information Sciences 219 (10) (2013) 151-167.
  - [13] X.P. Kang, D.Y. Li, S.G. Wang, K.S. Qu, Rough set model based on formal concept analysis, Information Sciences 222 (2013) 611-625.
  - J.K. Katzberg, W. Ziarko, Variable precision rough sets with asymmetric bounds, in: Proc. of the International Workshop on Rough Sets and Knowledge [14] Discovery, Banff, Alberta, Canada, 1993, pp. 163-190.
  - [15] D.C. Liang, D. Liu, W. Pedrycz, P. Hu, Triangular fuzzy decision-theoretic rough sets, International Journal of Approximate Reasoning 54 (8) (2013) 1087-1106.
  - C.H. Liu, D.O. Miao, N. Zhang, C. Gao, Graded rough set model based on two universes and its properties. Knowledge-Based Systems 33 (2012) 65-72. [16] [17] D. Liu, T.R. Li, D.C. Liang, Incorporating logistic regression to decision-theoretic rough sets for classifications, in: International Journal of Approximate Reasoning. http://dx.doi.org/10.1016/j.ijar.2013.02.013.
  - [18] Q. Liu, H. Sun, Y. Wang, Granulations based on semantics of rough logical formulas and its reasoning, in: Proceedings of the 11th International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing (RSFDGrC '07), 2007, pp. 419-426.
  - [19] J.S. Mi, W.Z. Wu, W.X. Zhang, Approaches to knowledge reduction based on variable precision rough set model, Information Sciences 159 (2004) 255-272.
- 974 [20] M. Ningler, G. Stockmanns, G. Schneider, H.D. Kochs, E. Kochs, Adapted variable precision rough set approach for EEG analysis, Artificial Intelligence in 975 Medicine 47 (2009) 239-261. 976
  - [21] S.K. Pal, B.U. Shankar, P. Mitra, Granular computing, rough entropy and object extraction, Pattern Recognition Letters 26 (16) (2005) 2509-2517.
  - [22] X. Pan, S.Q. Zhang, H.Q. Zhang, X.D. Na, X.F. Li, A variable precision rough set approach to the remote sensing land use/cover classification, Computers and Geosciences 36 (12) (2010) 1466-1473.
  - Z. Pawlak, Rough sets, International Journal of Computer and Information Sciences 11 (1982) 341-356. [23]
  - [24] Z. Pawlak, A. Skowron, Rough membership functions, in: R.R. Yager, M. Fedrizzi, J. Kacprzyk (Eds.), Advances in the Dempster-Shafer Theory of Evidence, Wiley, New York, 1994, pp. 251-271.
  - [25] Z. Pawlak, A. Skowron, Rough sets: some extensions, Information Sciences 177 (1) (2007) 28-40.
  - [26] Z. Pawlak, A. Skowron, Rudiments of rough sets, Information Sciences 177 (1) (2007) 3-27. [27] W. Pedrycz, Allocation of information granularity in optimization and decision-making models: towards building the foundations of Granular Computing, European Journal of Operational Research 232 (1) (2014) 137-145.
  - [28] W. Pedrycz, Granular Computing: Analysis and Design of Intelligent Systems, CRC Press, Francis Taylor, Boca Raton, 2013.
  - [29] W. Pedrycz, W. Homenda, Building the fundamentals of granular computing: a principle of justifiable granularity, Applied Soft Computing 13 (10) (2013) 4209-4218.
  - [30] Y.H. Qian, J.Y. Liang, C.Y. Dang, Consistency measure, inclusion degree and fuzzy measure in decision tables, Fuzzy Sets and Systems 159 (18) (2008) 2353-2377.
  - [31] Y.H. Qian, H. Zhang, Y.L. Sang, J.Y. Liang, Multigranulation decision-theoretic rough sets, in: International Journal of Approximate Reasoning. http:// dx.doi.org/10.1016/j.ijar.2013.03.004.
  - [32] K.Y. Qin, Z. Pei, J.L. Yang, Y. Xu, Approximation operators on complete completely distributive lattices, Information Sciences 247 (2013) 123-130.

[33] S. Singh, L. Dey, A new customized document categorization scheme using rough membership, Applied Soft Computing 5 (4) (2005) 373–390.

- [34] A. Skowron, J. Stepaniuk, Modeling of high quality granules, in: Proceedings of the International Conference on Rough Sets and Intelligent Systems Paradigms (RSEISP '07), 2007, pp. 300-309.
- [35] A. Skowron, J. Stepaniuk, R. Swiniarski, Modeling rough granular computing based on approximation spaces, Information Sciences 184 (1) (2012) 20-998 43. 999
  - [36] D. Slezak, W. Ziarko, The investigation of the bayesian rough set model, International Journal of Approximate Reasoning 40 (1-2) (2005) 81-91.
  - [37] J.Y. Wang, J. Zhou, Research of reduct features in the variable precision rough set model, Neurocomputing 72 (2009) 2643–2648.
  - [38] S.K.M. Wong, W. Ziarko, Comparison of the probabilistic approximate classification and the fuzzy set model, Fuzzy Sets and Systems 21 (1987) 357-362.
- 1003 [39] F. Xie, Y. Lin, W.W. Ren, Optimizing model for land use/land cover retrieval from remote sensing imagery based on variable precision rough sets, 1004 Ecological Modelling 222 (2) (2011) 232-240.
- 1005 [40] G. Xie, W.Y. Yue, S.Y. Wang, K.K. Lai, Dynamic risk management in petroleum project investment based on a variable precision rough set model, 1006 Technological Forecasting and Social Change 77 (6) (2010) 891-901.
- 1007 [41] G. Xie, J.L. Zhang, K.K. Lai, L. Yu, Variable precision rough set for group decision-making: an application, International Journal of Approximate 1008 Reasoning 49 (2008) 331-343. 1009
- W.H. Xu, S.H. Liu, Q.R. Wang, W.X. Zhang, The first type of graded rough set based on rough membership function, in: 2010 Seventh International [42] 1010 Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Yantai, China, 2010, pp. 1922–1926. 1011
- [43] Z.B. Xu, J.Y. Liang, C.Y. Dang, K.S. Chin, Inclusion degree: a perspective on measures for rough set data analysis, Information Sciences 141 (3-4) (2002) 1012 227-236.

#### **20 September 2013** O1 22

#### X. Zhang, D. Miao / Information Sciences xxx (2013) xxx-xxx

- 1013 [44] X.B. Yang, Y.S. Qi, X.N. Song, J.Y. Yang, Test cost sensitive multigranulation rough set: Model and minimal cost selection, Information Sciences. http:// dx.doi.org/10.1016/j.ins.2013.06.057.
- 1015 [45] I.T.R. Yanto, P. Vitasari, T. Herawan, M.M. Deris, Applying variable precision rough set model for clustering student suffering study's anxiety, Expert 1016 Systems with Applications 39 (1) (2012) 452–459.
- 1017 [46] J.T. Yao, Y.Y. Yao, W. Ziarko, Probabilistic rough sets: approximations, decision-makings, and applications, International Journal of Approximate 1018 Reasoning 49 (2) (2008) 253–254.
- 1019 [47] Y.Y. Yao, Human-inspired granular computing, in: J.T. Yao (Ed.), Novel Developments in Granular Computing: Applications for Advanced Human 1020 Reasoning and Soft Computation, Information Science Reference, Herskey, PA, 2010, pp. 1–15.
- 1021 [48] Y.Y. Yao, Information granulation and rough set approximation, International Journal of Intelligent Systems 16 (1) (2001) 87–104.
- 1022 [49] Y.Y. Yao, Probabilistic rough set approximations, International Journal of Approximate Reasoning 49 (2) (2008) 255–271.
- 1023 [50] Y.Y. Yao, The superiority of three-way decision in probabilistic rough set models, Information Sciences 181 (2011) 1080–1096.
- 1024 [51] Y.Y. Yao, Two views of the theory of rough sets in finite universes, International Journal of Approximate Reasoning 15 (4) (1996) 291–317.
- 1025 [52] Y.Y. Yao, T.Y. Lin, Generalization of rough sets using modal logics, Intelligent Automation and Soft Computing 2 (2) (1996) 103–120.
- [53] Y.Y. Yao, T.Y. Lin, Graded rough set approximations based on nested neighborhood systems, in: H.J. Zimmermann (Ed.), Proceedings of 5th European
   Congress on intelligent techniques and Soft computing, EUFIT'97, vol. 1, Verlag Mainz, Aachen, 1997, pp. 196–200.
- 1028 [54] 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 1029 Symposium on Methodologies for Intelligent Systems, North-Holland, New York, 1990, pp. 17–25.
- [55] H. Yu, Z.G. Liu, G.Y. Wang, An automatic method to determine the number of clusters using decision-theoretic rough set, International Journal of Approximate Reasoning, http://dx.doi.org/10.1016/j.ijar.2013.03.018.
- 1032 [56] LA. Zadeh, Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems 90 (2) 1033 (1997) 111-127.
- 1034 [57] H.Y. Zhang, Y. Leung, L. Zhou, Variable-precision-dominance-based rough set approach to interval-valued information systems, Information Sciences 244 (2013) 75–91.
- 1036 [58] H.Y. Zhang, J. Zhou, D.Q. Miao, C. Gao, Bayesian rough set model: a further investigation, International Journal of Approximate Reasoning 53 (2012) 541-557.
- 1038 [59] X.H. Zhang, B. Zhou, P. Li, A general frame for intuitionistic fuzzy rough sets, Information Sciences 216 (2012) 34-49.
- [60] X.Y. Zhang, D.Q. Miao, Two basic double-quantitative rough set models of precision and grade and their investigation using granular computing, International Journal of Approximate Reasoning 54 (8) (2013) 1130–1148.
- 1041 [61] X.Y. Zhang, Z.W. Mo, F. Xiong, Rough set model based on logical difference operation of precision and grade and its algorithms, Systems Engineering-Theory & Practice 31 (12) (2011) 2394–2399 (in Chinese).
- 1043 [62] X.Y. Zhang, Z.W. Mo, F. Xiong, W. Cheng, Comparative study of variable precision rough set model and graded rough set model, International Journal of 1044 Approximate Reasoning 53 (1) (2012) 104-116.
- 1045 [63] X.Y. Zhang, F. Xiong, Z.W. Mo, Rough set model based on logical OR operation of precision and grade, Pattern Recognition and Artificial Intelligence 22 (5) (2009) 697-703 (in Chinese).
- [64] X.Y. Zhang, F. Xiong, Z.W. Mo, W. Cheng, Logical AND operation model of variable precision lower approximation operator and grade upper approximation operator, Journal of Computer Applications 30 (8) (2010) 1991–1994 (in Chinese).
- 1049 [65] B. Zhou, Multi-class decision-theoretic rough sets, International Journal of Approximate Reasoning. http://dx.doi.org/10.1016/j.ijar.2013.04.006.
- 1050 [66] J. Zhou, D.Q. Miao,  $\beta$ -Interval attribute reduction in variable precision rough set model, Soft Computing 15 (8) (2011) 1643–1656.
- 1051 [67] W. Zhu, Topological approaches to covering rough sets, Information Sciences 177 (6) (2007) 1499–1508.
- 1052 [68] W. Ziarko, Probabilistic approach to rough sets, International Journal of Approximate Reasoning 49 (2) (2008) 272–284.
- 1053 [69] W. Ziarko, Variable precision rough set model, Journal of Computer and System Sciences 46 (1993) 39–59.
- 1054