Information Sciences 213 (2012) 1–13

Contents lists available at SciVerse ScienceDirect

Information Sciences

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

A measurement theory view on the granularity of partitions

Yiyu Yao [a,*], Liquan Zhao [a,b]

[a] Department of Computer Science, University of Regina, Regina, Saskatchewan, Canada S4S 0A2
[b] Province Key Laboratory of Electronic Business, Nanjing University of Finance and Economics, Nanjing, Jiangsu 210003, China

ARTICLE INFO

Article history:
Received 15 December 2009
Received in revised form 25 March 2012
Accepted 25 May 2012
Available online 7 June 2012

Keywords:
Expected granularity
Granular computing
Granularity
Hartley entropy
Information theory
Partition lattice

ABSTRACT

Measurement of granularity is one of the foundational issues in granular computing. This paper investigates a class of measures of granularity of partitions. The granularity of a set is defined by a strictly monotonic increasing transformation of the cardinality of the set. The granularity of a partition is defined as the expected granularity of all blocks of the partition with respect to the probability distribution defined by the partition. Many existing measures of granularity are instances of the proposed class. New measures of granularity of partitions are also introduced.

1. Introduction

Granular computing concerns problem solving and information processing at multiple levels of granularity [1,39,55,59,60,62]. A partition is perhaps the simplest granulation scheme and hence has attracted attentions from many researchers [52,58]. A partition of a universal set consists of a family of nonempty and pair-wise disjoint sets whose union is the universe. Each block in a partition may be viewed as a granule by ignoring the differences between elements in the block. Two examples of partition based granular computing theories are rough set theory [33,34,37] and quotient space theory of problem solving [63].

A crucial issue of granular computing is the search for an appropriate level of granularity by ignoring unimportant and irrelevant details. This requires the measurement of granularity. In a partition based model, it is the measurement of granularity of a partition. From a measurement-theoretic point of view [43], a measure of granularity of partitions is a quantification of some intuitive, qualitative relationships between partitions. Studies on measuring granularity of partitions mainly focus on two relations. A partition can be equivalently defined by an equivalence relation (i.e., a reflexive, symmetric, and transitive relation). Based on the set–inclusion relationship between the corresponding equivalence relations, one can define a partial order, called the refinement–coarsening relation, or the specialization–generalization relation, on the set of all partitions [34,58]. Wierman [50] introduced an equivalence relation on the set of all partitions of a set based on the notion of size–isomorphic. If two partitions have the same number of blocks and there exists a bijection that connects blocks of the same sizes, the two partitions are said to be size–isomorphic. It seems reasonable to require that any measure of granularity of partitions must reflect the refinement–coarsening relationship [7,30,31,50,56], namely, a refined partition has a lower granularity. Sometimes, the measure also needs to preserve the size–isomorphic relationship [23,25,50,66], namely, all equivalent partitions have the same level of granularity.

* Corresponding author.
E-mail addresses: yyao@cs.uregina.ca (Y. Yao), zhaoliquan2004@163.com (L. Zhao).

0020-0255/$ - see front matter
http://dx.doi.org/10.1016/j.ins.2012.05.021

From a measurement-theoretic perspective, this paper makes new contributions to the study of the granularity of partitions. We have three specific objectives that have not been considered in existing studies. First, we will establish a measurement-theoretic basis for a study of granularity. Intuitive understanding and properties of *finer than* relations are introduced on subsets of a universal set and on partitions of the universe, respectively. The results from the measurement theory enable us to establish the conditions (i.e., axioms) for the measurements of the *finer than* relations. Second, we introduce a new class of measures of granularity of partitions based on the expected granularity of blocks in a partition. The new measure reveals the inherent relationships between the granularity of blocks in a partition and the granularity of the partition. Third, we show that many existing measures are instances of the proposed class.

The rest of the paper is organized as follows. Section 2 reviews results from measurement theory to establish a theoretical basis for a study of granularity. Ordinal measurements of granularity are considered. Section 3 examines, in general, measures of granularity of a set and measures of granularity of a partition. Section 4 proposes a new class of measures of granularity of partitions, namely, granularity of a partition is the expected granularity of its blocks. Section 5 provides a critical review of existing studies on measures of granularity of partitions, introduces new measures, and comments on applications.

2. Measurement-theoretic foundations

To lay down a solid foundation for the measurement of granularity of partitions, we draw results from measurement theory. Pertinent concepts are reviewed based on books by Krantz et al. [17], Roberts [43], Fishburn [11], and French [12]. The ordinal measurements of granularity are examined.

2.1. An overview

When measuring an attribute of a class of objects or events, we may associate numbers with the individual objects so that the properties of the attribute are faithfully represented as numerical properties [17]. The properties are usually described by certain qualitative relations and operations. Formally, measurement may be studied based on homomorphisms between an empirical system and a numerical system.

A *relational system* (*structure*) is a set together with one or more relations (operations) on that set [43]. That is, a relational system is an ordered $(p+q+1)$-tuple $\mathcal{A} = (O, R_1, \ldots, R_p, \circ_1, \ldots, \circ_q)$, where O is a set, $R_1, \ldots, R_p$ are (not necessarily binary) relations on O, and $\circ_1, \ldots, \circ_q$ are binary operations on O. We call a relational system a numerical relational system if O is the set (or a subset) of real numbers. Operations can be considered as a special kind of relations. For convenience, we separate them from other relations. For modeling measurement, we start with an observed or empirical system $\mathcal{A}$ and seek a mapping into a numerical relational system $\mathcal{B}$ that preserves or faithfully reflects all the properties of the relations and operations in $\mathcal{A}$. Consider two relational systems, an empirical (a qualitative) system $\mathcal{A} = (O, R_1, \ldots, R_p, \circ_1, \ldots, \circ_q)$, and a numerical system $\mathcal{B} = (V, R'_1, \ldots, R'_p, \circ'_1, \ldots, \circ'_q)$. A function $f : O \rightarrow V$ is called a *homomorphism* from $\mathcal{A}$ to $\mathcal{B}$ if, for all $a_1, \ldots, a_{r_i} \in \mathcal{A}$,

$$R_i(a_1, \ldots, a_{r_i}) \Longleftrightarrow R'_i(f(a_1), \ldots, f(a_{r_i})), \quad i = 1, \ldots, p,$$

and for all $a, b \in \mathcal{A}$,

$$f(a \circ_j b) = f(a) \circ'_j f(b), \quad j = 1, \ldots, q.$$

That is, through a homomorphism *f*, the properties of the empirical system are truthfully reflected in the numerical system. Thus, *f* provides a measurement of the empirical system.

Consider a simple empirical system $(O, \prec, \circ)$, where O is a set of objects, $\prec$ is an ordering relation on O and $\circ$ is a binary relation on O. The numerical relation system is $(\Re, <, +)$, where $\Re$ is the set of real numbers, $<$ is the usual "smaller than" relation and $+$ is the arithmetic operation of addition. A numerical assignment $\phi(\cdot)$ is a homomorphism which maps O into $\Re$, $\prec$ into $<$, and $\circ$ into $+$ in such a way that $<$ preserves the properties of $\prec$, and $+$ preserves the properties of $\circ$ as stated by the conditions:

$$a \prec b \Longleftrightarrow \phi(a) < \phi(b),$$
$$\phi(a \circ b) = \phi(a) + \phi(b).$$

Measurement of length or weight of objects may be considered as an example using the empirical system $(O, \prec, \circ)$. For measuring length, we have a set of straight, rigid rods. The relation $\prec$ is interpreted as a "shorter than" relation. For two rods a and $b, a \prec b$ if we place them side by side with one of the endpoints coinciding and find that a does not extend to reach, or beyond, b on the opposite endpoint; operation $\circ$ denotes the operation of "concatenation" of two rods, that is, $a \circ b$ is obtained by joining a and b end-to-end along a straight line. For measuring weight, $\prec$ denotes the relation "lighter than" and $\circ$ denotes the operation of the "combination" of two objects. In addition to the basic ordering and operation, one needs to consider more properties for measuring length or weight [43].

There are three fundamental issues in measurement theory [12,17,43]. Suppose we are seeking a quantitative representation of an empirical system. The first step, naturally, is to define the relations and operations to be represented. We must describe the valid use of these relations and operations. The consistency properties to be preserved are known as *axioms*. The set of axioms characterizing the empirical system should be complete in the sense that every consistency property that we demand is either in the list or deducible from those in the list. The next task is to choose a numerical system. The final step is to construct an appropriate homomorphism. A *representation theorem* asserts that if a given empirical system satisfies certain axioms, then a homomorphism into the chosen numerical system can be constructed. A homomorphism into the set of real numbers is called a *scale*. The next question concerns the uniqueness of the scale. A uniqueness theorem is generally obtained by identifying a set of *admissible transformations*. If $\phi(\cdot)$ is a scale representing an empirical system and if $\lambda(\cdot)$ is an admissible transformation, then $\lambda(\phi(\cdot))$ is also a scale representing the same empirical system.

If the truth (falsity) of a numerical statement involving a scale remains unchanged under all admissible transformations, we say that it is quantitatively meaningful. A numerical statement may be quantitatively meaningful, but qualitatively meaningless. In order for a quantitative statement to be meaningful, it must reflect or model a meaningful statement in the empirical system.

Düntsch and Gediga [8] adopt Gigerenzer's approach for studying data models. A data model consists of a domain of interest, an empirical system, a representation system (e.g., a numerical or graphical system), and a researcher. The role of the researcher is to choose the domain of investigation, the data sample, and, based on them, to study relationships between the domain, the empirical system, and the representation system. Their study is related to measurement theory and offers another way to investigate measures of granularity.

2.2. Ordinal measurements of granularity

When applying measurement theory to measuring granularity of sets (i.e., granules) or partitions (i.e., families of sets), we must formulate a qualitative "finer than" relation, investigate its properties, and construct a homomorphism that truthfully reflects these properties. Ordinal measurements [11,43] of granularity are discussed. Their further applications and connections to existing studies are given in the next section.

Suppose O is a finite and nonempty set. Elements of O may be interpreted as granules or partitions whose granularity we want to measure. We assume that one can state certain relationships between some elements of O regarding their granularity. Formally, one can describe a strict "finer than" relation as a binary relation on O: for $a, b \in O$,

$$a \prec b \iff a \text{ is finer than } b. \tag{1}$$

In the absence of strict finer than relationship, i.e., if both $\neg(a \prec b)$ and $\neg(b \prec a)$ hold, a is said to be indifferent to b. An indifference relation $\sim$ on O can be defined as follows:

$$a \sim b \iff (\neg(a \prec b), \neg(b \prec a)) \tag{2}$$

By definition, $\sim$ is symmetric, namely, $a \sim b \Rightarrow b \sim a$. One can define another "finer than or equal relation" $\preceq$ as the union of $\prec$ and $\sim$, that is, $a \preceq b \iff (a \prec b \text{ or } a \sim b)$. Other connections between the three relations are given by $a \prec b \iff (a \preceq b, \neg(b \preceq a))$ and $a \sim b \iff (a \preceq b, b \preceq a)$. One may use either relation $\prec$ or $\preceq$. In this paper, we follow the notational systems of Fishburn [11].

From a measurement-theoretic point of view, it is important to identify the desired properties of $\prec$ and determine if such a relation can be measured by using a particular scale [12,17,43]. For this purpose, one may consider the following set of axioms: for a, b, $c \in O$,

$$\begin{aligned}
&\text{Irreflexivity}: \ \neg(a \prec a),\\
&\text{Asymmetry}: \ a \prec b \Rightarrow \neg(b \prec a),\\
&\text{Transitivity}: \ (a \prec b, b \prec c) \Rightarrow a \prec c,\\
&\text{Negative transitivity}: \ (\neg(a \prec b), \neg(b \prec c)) \Rightarrow \neg(a \prec c).
\end{aligned}$$

The asymmetry axiom requires that one cannot say a is finer than b and b is finer than a simultaneously. Asymmetry implies irreflexivity. The negative transitivity axioms states that if a is not finer than b, and b is not finer than c, one should accept that a is not finer than c. An asymmetric and negatively transitive relation is called a *weak order*. If a relation $\prec$ is a weak order, it is transitive. An irreflexive and transitive relation is called a *strict partial order*. A strict partial order is asymmetric, but may not be negatively transitive. A weak order is a special case of strict partial orders.

Each of these axioms seems reasonable for modeling the *finer than* relation $\prec$ of granularity. Depending on the choices of axioms, one may model $\prec$ as either a weak order or a strict partial order.

A few additional properties of a weak order are summarized in the following lemma [11].

Lemma 1. *Suppose a relation $\prec$ on a finite nonempty set O is a weak order. Then,*

(a) *the relation $\sim$ is an equivalence relation;*
(b) *exactly one of $a \prec b, b \prec a$ and $a \sim b$ holds for every $a, b \in O$;*
(c) *the relation $\prec'$ on $O/\sim$ defined by*

$$A \prec' B \iff a \prec b \text{ for some } a \in A \text{ and } b \in B, \tag{3}$$

is a linear order, where A and B are elements of $O/\sim$.

The results of Lemma 1 provide a better understanding of the structure of a weak order. A linear order is a weak order in which any two different elements are comparable. This lemma implies that if $\prec$ is a weak order, the indifference relation $\sim$ divides the set of objects into disjoint subsets. Furthermore, for any two equivalence classes A and B of $\sim$, either $A \prec' B$ or $B \prec' A$ holds. It is possible to arrange the elements into levels so that elements in a lower level are finer than elements in a higher level, and elements in the same level are indifferent.

In the measurement-theoretic terminology, the requirement of a weak order indeed suggests that the "finer than" relation can be measured on an ordinal scale, as shown by the following representation theorem [11,17,43].

Theorem 1. *Suppose O is a finite nonempty set and $\prec$ a binary relation on O. There exists a real-valued function $u : O \rightarrow \Re$ satisfying the condition,*

$$a \prec b \iff u(a) < u(b), \tag{4}$$

if and only if $\prec$ is a weak order. Moreover, $u' : O \rightarrow \Re$ is a another real-valued function with the same property if and only if there is a strictly monotonic increasing function $f : \Re \rightarrow \Re$ such that for all $a \in O$,

$$u'(a) = f(u(a)), \tag{5}$$

that is, u is an ordinal scale.

The numbers $u(a), u(b), \ldots$ as ordered by $<$ faithfully reflect the order of $a, b, \ldots$ under $\prec$. The function u is referred to as an order-preserving utility function. The condition (4) implies that for all $a, b \in O, a \sim b \iff u(a) = u(b)$. Thus, u truthfully preserves all three cases of Lemma 1b. By condition (4), we can make two-way inferences based on the values of u. For example, if a is finer than b, we can conclude that $u(a) < u(b)$. Conversely, if $u(a) < u(b)$, we can conclude that a is a finer that b. According to Theorem 1, the axioms of a weak order are the necessary and sufficient conditions that allow ordinal measurement.

For a strict partial order, $\sim$ is not necessarily an equivalence relation. However, a new relation $\approx$ define as,

$$a \approx b \iff (a \sim c \iff b \sim c, \text{ for all } c \in O) \tag{6}$$

is an equivalence relation [11]. Structures of a strict partial order are explained by the following lemma [11].

Lemma 2. *Suppose a relation $\prec$ on a finite nonempty set O is a strict partial order. Then,*

(a) *the relation $\approx$ is an equivalence relation;*
(b) *exactly one of $a \prec b, b \prec a, a \approx b, (a \sim b, \neg(a \approx b))$ holds for every $a, b \in O$;*
(c) *the relation $\prec''$ on $O/\approx$ defined by*

$$A \prec'' B \iff a \prec b \text{ for some } a \in A \quad \text{and} \quad b \in B, \tag{7}$$

is a strict partial order, where A and B are elements of $D/\approx$.

By Theorem 1, if a strict partial order is not a weak order, it is impossible to find a homomorphism $u : O \rightarrow \Re$ that allows us to make two-way inferences. In general, as shown by the following theorem [11], one can always make one-way inferences for the family of strict partial orders.

Theorem 2. *Suppose O is a finite nonempty set and $\prec$ a binary relation on O. If $\prec$ is a strict partial order, then there exists a real-valued function $u : O \rightarrow \Re$ such that for all $x, y \in O$,*

$$\begin{aligned} &a \prec b \Rightarrow u(a) < u(b), \\ &a \approx b \Rightarrow u(a) = u(b). \end{aligned} \tag{8}$$

For a strict partial order, a function u satisfying conditions in the Theorem 2 only partially truthfully preserves the first three cases of Lemma 2b, as indicated by one-way implications. In addition, for the last case, $(a \sim b, \neg(a \approx b))$, one may have any one of $u(a) = u(b), u(a) < u(b)$ and $u(b) < u(a)$.

The two theorems state the existence of a measurement and their constructive proofs provide a method to construct such a measurement. For concrete examples of relational systems and their measurement, see, for example, [12,17,43]. In the next section, we study a special class of functions that have an appealing intuitive interpretation as measures of granularity.

3. Measures of granularity of a partition

A partition is a family of subsets of a universe. This suggests that the granularity of a partition must depend on the granularity of the blocks in the partition. A measure of granularity of a partition is proposed based on this connection.

3.1. Granularity of a set

In granular computing, there are many ways to define a granule. We consider a special type of granules defined in a set-theoretic setting. For a finite nonempty set of objects U, a subset of U is a granule and the powerset 2^U is the set of all granules formed by the elements from U.

To build a measurement-theoretic model for interpreting the granularity of sets, we need first to construct a relational system over 2^U. In this paper, we only consider binary relations on 2^U for characterizing the granularity of sets. Suppose $\prec_s$ is a binary relation on 2^U, representing a *finer than* relationship between granules. As a first step, we need to model the relation $\prec_s$ based on our intuitive interpretation of the *finer than* relationship and investigate its properties.

Consider two granules $A, B \in 2^U$ with $A \subset B$, that is, A is a proper subset of B. It seems reasonable to claim that A is *finer than* B regarding their granularity. This immediately leads to modeling granularity by set inclusion. It can be easily verified that $\subset$ is irreflexive and transitive. By Theorem 2, there exists a function $m : 2^U \to \Re$ such that

$$A \subset B \Rightarrow m(A) < m(B). \tag{9}$$

Since $\subset$ is not negatively transitive, by Theorem 1, there does not exist a function satisfying a two-way implication. That is, the relation system $(2^U, \subset)$ is not suitable for us to make two-way inferences.

To obtain a measure of granularity for making two-way inferences, according to Theorem 1, we need at least a weak order $\prec_s$. At the same time, $\prec_s$ must be an extension of $\subset$ so that the qualitative interpretation of the granularity by $\subset$ is preserved, i.e., $\subset \subseteq \prec_s$. For this purpose, we need to introduce a qualitative notion of equivalence of sets with respect to granularity.

Definition 1. A binary "finer than" relation $\prec_s \subseteq 2^U \times 2^U$ on 2^U is defined by: $A \prec_s B$ if and only if there exists an injection from A to B and there does not exist a bijection from A to B, for $A, B \in 2^U$. A binary "g-equivalent to" relation $\equiv_s$ is defined by: $A \equiv_s B$ if there exists a bijection from A to B.

By definitions of an injection and a bijection, it follows,

$$\begin{aligned} &A \prec_s B \iff |A| < |B|, \\ &A \equiv_s B \iff |A| = |B|, \end{aligned} \tag{10}$$

where $|\cdot|$ denotes the cardinality of a set. In other words, the cardinality of a set is a measurement of both relations $\prec_s$ and $\equiv_s$. The relation $\prec_s$ is indeed a weak order satisfying the condition $A \subset B \Rightarrow A \prec_s B$. The relation $\sim_s$ of $\prec_s$ coincides with $\equiv_s$. An equivalence class of $\sim_s$ is the family of subsets of U with the same cardinality. According to Theorem 1, modeling granularity in terms of the relational system $(2^U, \prec_s)$ guarantees the existence of a measure for two-way implication and, moreover, such a measure is uniquely defined up to a strictly monotonic increasing transformation of the cardinality of sets.

A disadvantage of the system $(2^U, \prec_s)$ is that it does not explicitly use the relation $\subset$ and the g-equivalence relation $\equiv_s$. The following, easy-to-prove, lemma suggests that we can in fact use the relational system $(2^U, \subset, \equiv_s)$ for modeling granularity.

Lemma 3. *There is a function $m : 2^U \to \Re$ satisfying the condition:*

$$\text{(a)} \quad A \prec_s B \iff m(A) < m(B), \tag{11}$$

if and only if m satisfies the following two conditions: for all $A, B \in 2^U$,

$$\begin{aligned} &(\mathrm{b}_1) \quad A \subset B \Rightarrow m(A) < m(B), \\ &(\mathrm{b}_2) \quad A \equiv_s B \iff m(A) = m(B). \end{aligned} \tag{12}$$

According to the results of this lemma, we give a definition of a measure of granularity of a set.

Definition 2. Suppose U is finite and nonempty universe. A function $m : 2^U \to \Re$ is called a measure of granularity of a set if it satisfies the following conditions: for all $A, B \in 2^U$,

(i) (nonnegativity) $m(A) \geqslant 0$;
(ii) (monotonicity) $A \subset B \Rightarrow m(A) < m(B)$;
(iii) (size invariant) $A \equiv_s B \iff m(A) = m(B)$.

The soundness of the definition is guaranteed by Theorem 1 and Lemma 3 on the existence of such a measure. More specifically, a measure of granularity of a set is a strictly monotonic increasing transformation of the cardinality of the set. The nonnegativity may be viewed as a normalization condition. The monotonicity simply captures our intuition that a subset should have a lower granularity value. The size invariant property suggests that the granularity of a granule, i.e., a subset of U, is only determined by the number of elements in the granule and independent of which elements are in the granule.

Conditions (i) and (ii) are related to properties used to study nonspecificity of a granule [4]. One may establish a connection between granularity, specificity and nonspecificity of a granule. It is expected that a more nonspecific granule is large in size and hence has a large value of granularity.

3.2. Partitions and partition lattice

Granulation of a universe typically involves the division of the universe into a family of subsets, where each subset is a granule. A partition of a universe is a simple, non-overlapping granulation. The family of all partitions can be partially ordered to produce a partition lattice.

Definition 3. Suppose U is a finite nonempty set. A family of nonempty subsets of $U, \pi = \{A_1, A_2, \ldots, A_k\}$, is called a partition of U if $A_i \cap A_j = \emptyset (i \neq j)$ and $\bigcup_{i=1}^{k} A_i = U$. Each subset A_i is called a block of the partition π.

In this paper, we only consider a finite universe U. Consequently, any partition of the universe is also finite. An alternate and equivalent way to define a partition is through an equivalence relation on U, i.e., a reflective, symmetric and transitive relation on U. Let $R \subseteq U \times U$ denote an equivalence relation on U, where $U \times U$ is the Cartesian product of U and U. The equivalence class containing an element x is given by $[x]_R = \{y | y \in U, xRy\}$. There is a one-to-one correspondence between equivalence relations and partitions. For an equivalence relation R, the corresponding partition is denoted by $U/R = \{[x]_R | x \in U\}$.

One can define a partial order on partitions based on set inclusion of their blocks.

Definition 4. A partition π is a refinement of another partition π', or equivalently, π' is a coarsening of π, denoted by $\pi \sqsubseteq \pi'$, if every block of π is contained in some block of π', or equivalently, each block of π' is a union of some blocks of π. If $\pi \sqsubseteq \pi'$ and $\pi \neq \pi'$, or equivalently $\neg(\pi' \sqsubseteq \pi), \pi$ is a proper refinement of π' and we write $\pi \sqsubset \pi'$.

In terms of equivalence relations, we can have an equivalent definition of the refinement order. That is, $U/R \sqsubseteq U/R'$ if and only if $R \subseteq R'$. The family of partitions Π equipped with the order $\sqsubseteq$ define a lattice $(\Pi, \sqsubseteq)$ called the partition lattice. The lattice-theoretic operations on partitions are defined based on $\sqsubseteq$. Given two partitions π and π', their meet $\pi \wedge \pi'$ is the largest partition which is a refinement of both π and π', and their join $\pi \vee \pi'$ is the smallest partition which is a coarsening of both π and π'. The meet has all nonempty intersections of a block from π and a block from π' as its blocks. The blocks of join are the smallest subsets which are exactly a union of blocks from both π and π'. In terms of equivalence relations, given two equivalence relations R and R', the meet of U/R and U/R' is defined by the equivalence relation $R \cap R'$, and the join is defined by the equivalence relation $(R \cup R')^*$, i.e., the transitive closure of relation $R \cup R'$.

The minimum element Π_0 of the lattice $(\Pi, \sqsubseteq)$ is the finest partition, and the maximum element Π_1 of the lattice is the coarsest partition. They are given by:

$$\begin{aligned} \Pi_0 &= \{\{x\} | x \in U\}, \\ \Pi_1 &= \{U\}. \end{aligned} \tag{13}$$

For any partition $\pi \in \Pi$, we have $\Pi_0 \sqsubseteq \pi \sqsubseteq \Pi_1$. The partition lattice is useful for problem solving and reasoning. One may search the partition lattice for a suitable level of granularity [58,63].

3.3. Granularity of a partition

To have a qualitative model for the granularity of a partition, we consider a relational system with a *finer than* relation $\prec_p$ and a g-equivalence relation on the set of all partitions Π. Several attempts have been made by Wierman [50], Liang and Qian [24], Zhao et al. [64], and Zhu [65,66] to define measures of granularity based on a set of axioms regarding relations on the set of all partitions. Two of the axioms concern, respectively, a refinement–coarsening relation and an equivalence relation defined by size–isomorphic [50]. Zhu [66] presented a review of recent results on this topic. Several sets of axioms were investigated.

Considering that a partition is a set of subsets of the universe U, the problem can be similarly formulated as we did in the case of the granularity of a set. By definition, the proper refinement–coarsening relation $\sqsubset$ is a strict partial order. According to Theorem 2, there exists a measure on Π that allow us to draw one-way inference. For the existence of a measure for two-way inference, we need a weak order $\prec_p$ that contains $\sqsubset$, namely, $\sqsubset \subseteq \prec_p$. According to the cardinality of a partition, we can define a *finer than* relation $\prec_p$.

Definition 5. Suppose Π is the set of all partitions of a finite nonempty universe U. A binary *finer than* relation $\prec_p \subseteq \Pi \times \Pi$ on Π is defined by: $\pi \prec_p \pi'$ if and only if there exists an injection from π' to π and there does not exist a bijection from π' to π.

It can be easily verified that $\prec_p$ is an extension of $\sqsubset$, namely, $\sqsubset \subseteq \prec_p$ and $\prec_p$ is a weak order. According to Theorem 1, the relation $\prec_p$ can be measured based on the cardinality of a partition, that is, $\pi \prec_p \pi'$ if and only if $|\pi'| < |\pi|$. That is, the relation $\prec_p$ is characterized based on the converse of the relation defined by the cardinality of partitions. This is intuitively reasonable. A block in a finer partition is subset of a block in a coarser partition. Thus, a finer partition should divide the universe into more number of blocks than a coarser partition.

Liang and Qian [24] and Zhu [65] considered another relation $\prec'_p$ by using the cardinality of all blocks of a partition. It can be verified that $\sqsubset \subseteq \prec_{p'} \subseteq \prec_p$. In this paper, we use $\prec_p$, as an extension of $\sqsubset$, to show the existence of a measure of $\sqsubset$. It is sufficient to use this simpler relation. One may need to use another relation when measuring granularity of a covering of a universe.

The equivalence relation $\pi \sim_p \pi' \iff (\neg(\pi \prec_p \pi'), \neg(\pi' \prec_p \pi))$ groups partitions with the same number of blocks together. The condition for defining this equivalence of partitions may seem to be too weak. This may be illustrated by using a simple example. Suppose $\pi = \{\{a\}, \{b, c, d, e, f, g, h\}\}$ and $\pi' = \{\{a, b, c, d\}, \{e, f, g, h\}\}$ are two partitions on a universe $U = \{a, b, c, d, e, f, g, h\}$. According to $\sim_p$, we have that $\pi \sim_p \pi'$, namely, π and π' have same value with respect to their granularity. However, it is not difficult to convince ourselves that in some situations they are of different granularities. Many studies consider another g-equivalence relations that are weaker than $\sim_p$. A well studied example is based on the notion of size–isomorphic of partitions [23,50,65].

Definition 6. Two partitions π and π' are size–isomorphic, written $\pi \equiv_p \pi'$, if and only if there exists a bijection $f : \pi \to \pi'$ satisfying the condition $|A| = |f(A)|$ for all $A \in \pi$.

The existence of a bijection requires that $|\pi| = |\pi'|$. By definition, $\equiv_p \subseteq \sim_p$. Thus, a function that truthfully reflects $\sim_p$ would truthfully reflect $\equiv_p$. By Theorem 1, there exists a measure of granularity on Π that truthfully reflects both $\prec_p$ and $\sim_p$. It follows that the same function would enable us to draw some conclusion about $\sqsubset$ and $\equiv_p$, as summarized by the following, easy-to-prove, lemma.

Lemma 4. *If a function $G : \Pi \to \Re$ satisfies the condition, for $\pi, \pi' \in \Pi$,*

(a) $\pi \prec_p \pi' \iff G(\pi) < G(\pi')$,

then G satisfies the conditions, for all for $\pi, \pi' \in \Pi$,

(b_1) $\pi \sqsubset \pi' \Rightarrow G(\pi) < G(\pi')$,

(b_2) $\pi \equiv_p \pi' \Rightarrow G(\pi) = G(\pi')$.

The reverse of lemma is not true, that is, from (b_1) and (b_2) we cannot conclude (a). This suggests that a measure defined based on the relational system $(\Pi, \sqsubset, \equiv_p)$ is not a measure of the relational system $(\Pi, \prec_p)$.

According to Theorem 2, when the system $(\Pi, \sqsubset, \equiv_p)$ is used, there exists a measure for one-way implication. The relations $\sqsubset$ and $\equiv_p$ can be easily explained, i.e., $\sqsubset$ is related to the partial ordering $\sqsubseteq$ of the partition lattice and $\equiv_p$ is interpretable through size–isomorphic. In addition, $\equiv_p$ is a natural generalization from size–isomorphic of sets (i.e., granules) to size–isomorphic of partitions (i.e., families of granules). Therefore, we define a measure of granularity of a partition based on the less restrictive system $(\Pi, \sqsubset, \equiv_p)$.

Definition 7. Suppose Π is the set of all partitions of a finite nonempty universe U. A function $G : \Pi \to \Re$ is called a measure of granularity of a partition if it satisfies the following conditions: for all $\pi, \pi' \in \Pi$,

(I) (nonnegativity) $G(\pi) \geqslant 0$;

(II) (monotonicity) $\pi \sqsubset \pi' \Rightarrow G(\pi) < G(\pi')$;

(III) (size invariant) $\pi \equiv_p \pi' \Rightarrow G(\pi) = G(\pi')$.

The three conditions correspond naturally to the three conditions for defining a measure of granularity of a granule. Theorem 2 guarantees the existence of such a measure of the granularity of partitions. It follows easily from the definition that $G(\pi)$ has the minimal value if and only if $\pi = \Pi_0$, $G(\pi)$ has the maximum value if and only if $\pi = \Pi_1$, and $G(\Pi_0) \leqslant G(\pi) \leqslant G(\Pi_1)$ for $\pi \in \Pi$. From the discussion on measurement theory and the relationships between $\sqsubset$ and $\prec_p$ and between $\equiv_p$ and $\sim_p$, such a measure always exists.

Unlike the case of measures of granularity of sets, namely, a measure of granularity of a set is a strictly monotonic increasing transformation of the cardinality of the set, we are not able to characterize the family of measures defined by Definition 7. Nevertheless, we can introduce a special class of measures of granularity of partitions.

4. Granularity of a partition as the expected granularity of its blocks

Consider a partition $\pi = \{A_1, \ldots, A_k\}$ of a finite and nonempty universe U. One may associate it with a probability distribution [18]:

$$P_\pi = (p(A_1), \ldots, p(A_k)) = \left(\frac{|A_1|}{|U|}, \ldots, \frac{|A_k|}{|U|}\right). \tag{14}$$

Suppose $m : 2^U \to \Re$ is a measure of granularity of subsets of U. According to the probability distribution P_π, one can compute the *expected granularity* of blocks in π.

Definition 8. Suppose $\pi = \{A_1, \ldots, A_k\}$ is a partition of a finite nonempty universe U and $m : 2^U \to \Re$ is a measure of granularity of subsets of U, namely, m satisfies (i)–(iii) of Definition 2. The expected granularity of blocks of π is defined as:

$$EG_m(\pi) = \mathbf{E}_{P_\pi}(m(\cdot)) = \sum_{i=1}^{k} m(A_i) p(A_i), \tag{15}$$

where $P_\pi = (p(A_1), \ldots, p(A_k))$ is the probability distribution defined by π and $\mathbf{E}_{P_\pi}(\cdot)$ is the mathematical expectation with respect to distribution P_π.

Intuitively, the expected granularity seems to be a good candidate of a measure of granularity of a partition, as it pools together granularity of all blocks in the partition. The next theorem shows that *EG* is indeed a measure of granularity of a partition as defined by Definition 7.

Theorem 3. *The measure $EG_m : \Pi \to \Re$ is a measure of granularity of a partition, that is, EG_m satisfies conditions (I)–(III).*

Proof. Property (I) easily follows from the definition of expected granularity and the fact that *m* satisfies the nonnegativity property (i). Property (II) can be proved based on the definition of the refinement relation $\sqsubset$ and the property (ii) of *m*. Suppose $\pi \sqsubset \pi'$ holds for two partitions π and π'. This means that every block of π' is a union of one or more blocks of π and at least one block of π' is the union of at least two blocks from π. By the fact that *U* is a finite universe, there exists a finite sequence of partitions $\pi = \pi_1 \sqsubset \pi_2 \sqsubset \ldots \sqsubset \pi_n = \pi'$ such that exactly one block of π_{j+1} is the union of two blocks from π_j for $j = 1, \ldots, (n-1)$ and $n \geqslant 2$. We want to show that $EG_m(\pi_j) < EG_m(\pi_{j+1})$. Without loss of generality, suppose a block of π_{j+1} is obtained by the union of two blocks A_{j1} and A_{j2} of π_j, that is, $\pi_j = \{A_{j1}, A_{j2}, \ldots, A_{jk}\}, k \geqslant 2$ and $\pi_{j+1} = \{A_{j1} \cup A_{j2}, \ldots, A_{jk}\}$. According to the definition of EG_m and monotonicity of *m*, we have:

$$\begin{aligned} EG_m(\pi_j) &= \mathbf{E}_{P_{\pi_j}}(m(\cdot)) = \sum_{i=1}^{k} m(A_{ji})p(A_{ji}) = m(A_{j1})p(A_{j1}) + m(A_{j2})p(A_{j2}) + \sum_{i=3}^{k} m(A_{ji})p(A_{ji}) \\ &< m(A_{j1} \cup A_{j2})p(A_{j1}) + m(A_{j1} \cup A_{j2})p(A_{j2}) + \sum_{i=3}^{k} m(A_{ji})p(A_{ji}) = m(A_{j1} \cup A_{j1})(p(A_{j2}) + p(A_{j2})) + \sum_{i=3}^{k} m(A_i)p(A_i) \\ &= m(A_{j1} \cup A_{j2})p(A_{j1} \cup A_{j2}) + \sum_{i=3}^{k} m(A_{ji})p(A_{ji}) = \mathbf{E}_{P_{\pi_{j+1}}}(m(\cdot)) = EG_m(\pi_{j+1}). \end{aligned}$$

It immediately follows that $EG_m(\pi) < EG_m(\pi')$, namely, property (II) holds. Now suppose $\pi \equiv_p \pi'$ holds between two partitions π and π'. By definition of $\equiv$, there exists a bijection $f : \pi \to \pi'$ such that $|A| = |f(A)|$ for all $A \in \pi$. This means that π and π' have the same number of blocks and the corresponding blocks defined by the bijection *f* have the same number of elements. Based on the definition of EG_m, we have $EG_m(\pi) = EG_m(\pi')$. By summarizing these results, we can conclude that EG_m satisfies conditions (I)–(III) and hence is a measure of granularity of partitions. □

Given a strictly monotonic increasing function *m* of the cardinality of sets, one can define a measure of the granularity of partitions. The expected granularity defines a family of measures called *EG* class.

5. A critical review of measures of granularity of partitions

The theory of rough sets is one of the most extensively studied granular computing models based on partitions [33,34,58]. A central notion is knowledge granularity induced by partitions. However, the granularity based interpretation was not explicitly used in the literature before early 1990s. In the Foreword to Pawlak's seminal book [34], Dubois and Prade [6] interpreted rough set theory based on the granularity of knowledge. This interpretation was considered briefly by Pawlak [35], Pawlak and Słowinski [38], and made more elaborate by Pawlak [36]. The issue of measuring granularity of partitions started to receive more attention in rough set theory in the late 1990s with the publication of several influential papers (for example, see [2,7,25,30,31,50,56]). This section reviews studies that explicitly consider measures of granularity of partitions. They may be divided broadly into two classes, namely, information-theoretic measures and interaction based measures.

5.1. Information-theoretic measures

Shannon entropy and Hartley entropy have been widely used to measure the structuredness of attributes in databases and the nonspecificity of a finite set [16,18,57]. Many authors used them to design measures of granularity of partitions [2,7,9,23,25,48,50,56].

Consider a partition $\pi = \{A_1, \ldots, A_k\}$ with the associated probability distribution $P_\pi = (p(A_1), \ldots, p(A_k))$ and $p(A_i) = |A_i|/|U|$. For this probability distribution, its Shannon entropy [44], $H(P_\pi)$ or simply $H(\pi)$, is given in terms of mathematical expectation by:

$$H(\pi) = \mathbf{E}_{P_\pi}(-\log(\cdot)) = \sum_{i=1}^{k}(-\log p(A_i))p(A_i) = -\sum_{i=1}^{k} \frac{|A_i|}{|U|} \log \frac{|A_i|}{|U|}, \tag{16}$$

where $-\log p(A_i)$ is interpreted as a measure of information provided by an event A_i with probability $p(A_i)$. Shannon entropy is the expected information generated by a probability distribution. For each block A_i, its Hartley entropy [13,16] is given by:

$$H_0(A_i) = \log |A_i|. \tag{17}$$

The Hartley entropy is the amount of uncertainty associated with a finite set of possible alternatives, namely, the nonspecificity inherent in the set [16]. It is in fact a measure of granularity of a set. A set that is more nonspecific has a higher granularity. Shannon entropy can be expressed in terms of Hartley entropy as follows [16,50,56]:

$$H(\pi) = -\sum_{i=1}^{m} \frac{|A_i|}{|U|} \log \frac{|A_i|}{|U|} = \log|U| - \sum_{i=1}^{m} (\log|A_i|) \frac{|A_i|}{|U|} = H_0(U) - \mathbf{E}_{P_\pi}(H_0(\cdot)) = H_0(U) - G(\pi). \tag{18}$$

That is, Shannon entropy is the Hartley entropy of the universe minus the expected Hartley entropy of all blocks in a partition. However, as pointed out by Klir and Golger [16], there are semantic differences about the two entropies. While Hartley entropy is a measure of nonspecificity of a finite set, Shannon entropy is a measure of information induced by a probability distribution.

Miao and Wang [31] and Düntsch and Gediga [7,9] used Shannon entropy $H(\pi)$ of a partition as a measure of its roughness or granularity. They showed that $H(\pi)$ is inversely related to the refinement–coarsening relation. Wierman [50] provided a systematic and axiomatic study and explicitly called Shannon entropy of a partition a "granularity measure." In addition, he showed that $H(\pi)$ also reflects the equivalence relation defined by size–isomorphic. As $\log|\cdot|$ is a measure of granularity of a set, the second term in Eq. (18) of the Shannon entropy,

$$G(\pi) = \sum_{i=1}^{m} (\log|A_i|) \frac{|A_i|}{|U|}, \tag{19}$$

is in fact an example of the proposed *EG* class of measures. Beaubouef et al. [2] defined rough entropy of a rough set approximation as a product of Pawlak roughness measure [34] and $G(\pi)$ by implicitly treating $G(\pi)$ as a measure of granularity of a partition. Motivated by Klir and Golger's interpretation [16] of Hartley entropy, Yao [56] called $G(\pi)$ a measure of granularity of a partition. Liang and Shi [25] called $G(\pi)$ the rough entropy of π, and Bianucci et al. [3] called it co-entropy. Holschke et al. [14] used the measure $G(\pi)$ to interpret a factor in effective business process model reuse. Li et al. [19] used $G(\pi)$ to study granularity in panweighted field of pansystems. Following the works of Beaubouef et al. [2] and Wierman [50], Liang and his associates [22,23,25–27,48], Mi et al. [29], Qian et al. [41,42], and Zhu and Wen [67] introduced additional information-theoretic measures of granularity of a partition.

5.2. Interaction based measures

Interaction based measures are based on counting the number of interacting pairs of elements of a universal set under a partition. A partition π uniquely defines an equivalence relation E_π, namely, $xE_\pi y \Longleftrightarrow x$ and y are in the same block of π. More specifically, the equivalence relation can be expressed as $E_\pi = \bigcup_{i=1}^{k}(A_i \times A_i)$. The set–inclusion relation of equivalence relations defines the refinement–coarsening relation of the corresponding partitions $\pi \sqsubset \pi' \Longleftrightarrow E_\pi \subset E_{\pi'}$. Hence, we have $\pi \sqsubset \pi' \Rightarrow |E_\pi| < |E_{\pi'}|$. It suggests that the cardinality of an equivalence relation can be used as a measure of granularity of a partition. Each pair in the equivalence relation may be counted as one interaction; the size of the equivalence relation represents the total number of interactions.

Miao and Fan [30] first introduced the following interaction based measure of granularity of a partition,

$$G_m(\pi) = \frac{|E_\pi|}{|U|^2} = \frac{\sum_{i=1}^{k}|A_i|^2}{|U|^2} = \frac{\sum_{i=1}^{k}|A_i \times A_i|}{|U \times U|}, \tag{20}$$

where E_π is the equivalence relation corresponding to a partition π. The value $|A_i \times A_i|$ is the number of interacting pairs in the block A_i, the summation is the total number of interacting pairs induced by a partition, and $|U \times U|$ is the number of interacting pairs induced by the coarsest partition corresponding to the equivalence relation $U \times U$. The measure G_m may be interpreted as a normalized cardinality of an equivalence relation. The minimum value $1/|U|$ is obtained from the finest partition Π_0 and the maximum value 1 is obtained from the coarsest partition Π_1.

We can show that G_m is an example of the *EG* class of measures:

$$G_m(\pi) = \frac{\sum_{i=1}^{k}|A_i \times A_i|}{|U \times U|} = \sum_{i=1}^{k} \frac{|A_i \times A_i|}{|U \times U|} = \sum_{i=1}^{k} \frac{|A_i|}{|U|}\frac{|A_i|}{|U|} = \sum_{i=1}^{k} \frac{|A_i|}{|U|} p(A_i) = \mathbf{E}_{P_\pi}(m_m(\cdot)), \tag{21}$$

where

$$m_m(A) = \frac{|A|}{|U|}, \tag{22}$$

is a measure of granularity of a set. Feng et al. [10] introduced another interaction based measure that is also an example of the *EG* class. By taking the cardinality of a set as a measure of its granularity, they defined a measure of granularity of a partition as:

$$G_f(\pi) = \mathbf{E}_{P_\pi}(|\cdot|) = \sum_{i=1}^{k}|A_i|p(A_i) = \sum_{i=1}^{k}|A_i|\frac{|A_i|}{|U|} = \frac{1}{|U|}\sum_{i=1}^{k}|A_i|^2 = \frac{1}{|U|}|E_\pi|. \tag{23}$$

It follows that $G_f(\pi) = |U|G_m(\pi)$.

Liang et al. [23] introduced a new entropy of a partition as:

$$H_l(\pi) = \sum_{i=1}^{k} \frac{|A_i|}{|U|}\frac{|A_i^c|}{|U|} = \sum_{i=1}^{k} \frac{|A_i|}{|U|}\left(1 - \frac{|A_i|}{|U|}\right). \tag{24}$$

It was later called a complementary entropy [22]. Similar to Shannon entropy, we easily express H_l in the form of mathematical expectation:

$$H_l(\pi) = \sum_{i=1}^{k} \left(1 - \frac{|A_i|}{|U|}\right)\frac{|A_i|}{|U|} = \sum_{i=1}^{k} \left(1 - \frac{|A_i|}{|U|}\right)p(A_i) = \mathbf{E}_{P_\pi}(1 - |\cdot|/|U|). \tag{25}$$

Liang and Shi [25] expressed the following relationship between H_l and G_m, namely,

$$H_l(\pi) = \sum_{i=1}^{k} \frac{|A_i|}{|U|}\left(1 - \frac{|A_i|}{|U|}\right) = \sum_{i=1}^{k} \frac{|A_i|}{|U|} - \sum_{i=1}^{k} \frac{|A_i|^2}{|U|^2} = 1 - G_m(\pi), \tag{26}$$

and called $G_m(\pi)$ the "knowledge granulation" of π. This connection immediately provides a non-interaction based interpretation of H_l as follows:

$$H_l(\pi) = 1 - G_m(\pi) = 1 - \frac{|E_\pi|}{|U \times U|} = \frac{|U \times U| - |E_\pi|}{|U \times U|} = \frac{|U \times U - E_\pi|}{|U \times U|} = \frac{|E_\pi^c|}{|U \times U|}, \tag{27}$$

where E_π^c denotes the complement of relation E_π and the value $|E_\pi^c|$ is indeed the number of non-interacting pairs of elements induced by a partition π.

When computing the number of interactions in a block $A_i \in \pi$, the number $|A_i|^2$ takes into consideration of self-interaction (i.e., the pair $(x,x), x \in A_i$) and double counts the interaction of two objects (i.e., both pairs (x,y) and $(y,x), x,y \in A_i$). To remove such effects, Qian and Liang [40] suggested to use the number of combinations $\binom{|A_i|}{2} = |A_i|(|A_i| - 1)/2$ and introduced a measure of "combination granulation" of π as:

$$G_q(\pi) = \sum_{i=1}^{k} \left(\frac{\binom{|A_i|}{2}}{\binom{|U|}{2}}\right)\frac{|A_i|}{|U|} = \sum_{i=1}^{k} \left(\frac{(|A_i|(|A_i| - 1))/2}{(|U|(|U| - 1))/2}\right)\frac{|A_i|}{|U|} = \sum_{i=1}^{k} \left(\frac{|A_i|(|A_i| - 1)}{|U|(|U| - 1)}\right)\frac{|A_i|}{|U|} = \mathbf{E}_{P_\pi}(m_q(\cdot)), \tag{28}$$

where

$$m_q(A) = \frac{|A|(|A| - 1)}{|U|(|U| - 1)} \tag{29}$$

is a measure of granularity of a set A. Thus, G_q is an example of the proposed EG class.

Corresponding to G_q, Qian and Liang [40] introduced the notion of combination entropy of π as

$$H_q(\pi) = \sum_{i=1}^{k} \left(1 - \frac{\binom{|A|}{2}}{\binom{|U|}{2}}\right)\frac{|A_i|}{|U|} = \mathbf{E}_{P_\pi}\left(1 - \frac{\binom{|\cdot|}{2}}{\binom{|U|}{2}}\right). \tag{30}$$

The combination granulation G_q can be expressed by H_q as,

$$G_q(\pi) = 1 - H_q(\pi). \tag{31}$$

Unfortunately, unlike G_m and H_l, G_q and H_q cannot be simply expressed in terms of the cardinality of equivalence relations E_π and its complement E_π^c, respectively.

Another interaction based measure was proposed by Xu et al. [53] in terms of connectivity among nodes in a graph induced by an equivalence relation. Many authors studied interaction based measures and extended them to non-equivalence relations [19,24,26,27,41,48,54]. The representation of G_m in terms of the cardinality of an equivalence relation offers an advantage when one studies measures of granularity of a covering induced by a non-equivalence relation. For example, one may simply replace E_π in G_m by an arbitrary binary relation. Liang et al. [26] suggested another expression of G_m as:

$$G_m(\pi) = \frac{1}{|U \times U|}\sum_{i=1}^{|U|} |[x_i]_{E_\pi}|, \tag{32}$$

where $[x]_{E_\pi}$ denote the equivalence class containing x. To generalize this expression, we can simply replace the equivalence class $[x_i]_{E_\pi}$ by a neighborhood of x_i induced by a binary relation. Liang et al. [26] and Wang et al. [48] generalized this expression for a family of tolerance classes induced by a tolerance relation in an incomplete information table. Xu et al. [54] generalized this expression to measure granularity of a family of subsets induced by a dominance relation.

5.3. Cardinality of a partition

Consider the following function defined based on the cardinality of sets: for $\emptyset \neq A \subseteq U$,

$$m_1(A) = |U|\left(1 - \frac{1}{|A|}\right). \tag{33}$$

We define $m_1(\emptyset) = 0$. The first term, i.e., the cardinality of the universe U, is a constant independent of the set A. The quantity $1/|A|$ is the inverse of the cardinality of A; it is a strictly monotonic decreasing function of the cardinality with the minimum value $1/|U|$ and the maximum value 1. The second term $1 - 1/|A|$ may be viewed as a negation of the quantity $1/|A|$ and is a strictly monotonic increasing function of the cardinality. The function m_1 is a measure of granularity of a set.

With respect to the measure m_1, according to Definition 8 we have:

$$G_1(\pi) = EG_{m_1}(\pi) = \mathbf{E}_{P_\pi}(m_1(\cdot)) = \sum_{i=1}^{k} m_1(A_i)p(A_i) = \sum_{i=1}^{k}\left(|U|\left(1 - \frac{1}{|A_i|}\right)\right)\frac{|A_i|}{|U|} = \sum_{i=1}^{k}(|A_i| - 1) = |U| - k. \tag{34}$$

The cardinality of the universe $|U|$ is in fact the number of blocks in the finest partition Π_0, and k is the number of the blocks in π. The measure G_1 may be viewed as inverse function of the cardinality of a partition. It can be easily verified that $G_1(\pi) = 0$ if and only if $\pi = \Pi_0$ and $G_1(\pi) = n - 1$ if and only if $\pi = \Pi_1$. Moreover, G_1 truthfully reflects the relation $\prec_p$, that is, $\pi \prec_p \pi' \iff G_1(\pi) < G_1(\pi')$.

The measure G_1 is determined by the cardinality of a partition, and may be one of the simplest measures of granularity of partitions. Such a measure was implicitly used in some studies [27,64,66].

5.4. A parameterized family of measures of granularity

We consider a family of granularity measures of a set defined by a pair of parameters (a, b) with $a > 0$ and $b > 0$:

$$m_{(a,b)}(A) = a|A|^b. \tag{35}$$

The corresponding measure of granularity of a partition is given by:

$$G_{(a,b)}(\pi) = EG_{(a,b)}(\pi) = \mathbf{E}_{P_\pi}(m_{(a,b)}(\cdot)) = \sum_{i=1}^{k} m_{(a,b)}(A_i)p(A_i) = \sum_{i=1}^{k}(a|A_i|^b)\frac{|A_i|}{|U|} = \frac{a}{|U|}\sum_{i=1}^{k}|A_i|^{b+1}. \tag{36}$$

By combining members of this family, it is possible to obtain new measures of granularity.

All three interaction based measures G_m, G_f and G_q can be expressed through members of this parameterized family. More specifically, we have:

$$G_m(\pi) = \sum_{i=1}^{k}\frac{|A_i|}{|U|}\frac{|A_i|}{|U|} = G_{(1/|U|,1)}(\pi); \tag{37}$$

$$G_f(\pi) = \sum_{i=1}^{k}|A_i|\frac{|A_i|}{|U|} = G_{(1,1)}(\pi); \tag{38}$$

$$\begin{aligned} G_q(\pi) &= \sum_{i=1}^{k}\left(\frac{|A_i|(|A_i| - 1)}{|U|(|U| - 1)}\right)\frac{|A_i|}{|U|} = \sum_{i=1}^{k}\left(\frac{|A_i|^2}{|U|(|U| - 1)}\right)\frac{|A_i|}{|U|} - \sum_{i=1}^{k}\left(\frac{|A_i|}{|U|(|U| - 1)}\right)\frac{|A_i|}{|U|} \\ &= G_{(1/(|U|(|U|-1)),2)}(\pi) - G_{(1/(|U|(|U|-1)),1)}(\pi). \end{aligned} \tag{39}$$

They provide another interesting interpretation of the three interaction based measures.

5.5. Applications of measures of granularity

One class of applications of measures of granularity is to modify the Pawlak measure of roughness of rough set approximations. Beaubouef et al. [2] argued that a measure of roughness must be monotonic with respect to the refinement–coarsening relation of partitions. Rough set approximations obtained from a finer partition must have a lower value of roughness than ones obtained from a courser partition. Unfortunately, Pawlak measure of roughness does not satisfy the monotonicity. To resolve this problem, several authors suggested to multiple the Pawlak measure by a measure of granularity of partitions [2,27,66]. However, the meaningfulness of such a composite measure needs to be further explored. A summary and critical analysis of studies on this topic can be found in [61].

Another class of applications concerns the evaluation of significance of attributes for classification and rule induction [22,31,36,49]. As a measure of uncertainty, Shannon entropy has been used in rough set theory ever since its introduction and are continually being used, see for example, Wong et al. [51], Ślęzak [45–47]. To a large degree, this group of studies draw results from machine learning and databases [56,57]. More information on applications of measures of granularity for evaluating attributes can be found in [22,57].

The results of this paper have significant implications to applications of measures of granularity. The class of measures *EG* contains commonly used information-theoretic and interaction based measures. They all agree on relations $\prec_p$ and $\equiv_p$ and differ only for partitions for which neither relationship holds. If $\prec_p$ and $\equiv_p$ are the only relations to preserve, all measures are the same. Therefore, instead of proposing and examining individual measures, it may be more meaningful to study a class of measures. Alternatively, one may introduce and study sub-classes of measures. In addition to providing a unified framework, this paper also introduces new measures of granularity. When applying measures of granularity in rough set theory, it may be more fruitful to combine them with other measures of uncertainty [5,15,20,21,28,32].

6. Conclusion

The extensive studies on the measurement of granularity call for a solid foundation and a unified framework. We approach this problem from a measurement-theoretic point of view. Any measure of granularity should reflect our intuitive understanding of a *finer than* relationship. Two levels of granularity are considered. The granularity of a subset of a universal set depends on its size. A subset should have a lower granularity than its supersets. The granularity of a partition depends on both the number of the blocks in the partition and the sizes of the blocks. A partition should have a lower granularity than its coarsening partitions. A new class of measures is proposed by considering the expected granularity of blocks in a partition. Many existing measures are shown to be instances of the proposed class.

The results presented in this paper are closely related to an axiomatic framework proposed by Zhu [65], where a set of axioms is given to guarantee the existence of a unique measure of granularity. The proposed class of granularity measures is in the form of mathematical expectation. It will be interesting to investigate a set of axioms that characterizes uniquely this class.

Acknowledgements

The work is partially supported by a Discovery Grant from NSERC Canada, Basic Research Program of Education Office of Jiangsu Province, China, Grant No. 07KJA52004, and Natural Science Foundation of Jiangsu Province, China, Grant No. BK2009349. The authors thank anonymous reviewers for their constructive comments.

References

[1] A. Bargiela, W. Pedrycz, Granular Computing: An Introduction, Kluwer Academic Publishers, Boston, 2002.
[2] T. Beaubouef, F.E. Petry, G. Arora, Information-theoretic measures of uncertainty for rough sets and rough relational databases, Information Sciences 109 (1998) 185–195.
[3] D. Bianucci, G. Cattaneo, D. Ciucci, Entropies and co-entropies for incomplete information systems, in: Proceedings of the Second International Conference on Rough Sets and Knowledge Technology, LNCS(LNAI) 4481, 2007, pp. 84–92.
[4] S. Bodjanova, Granulation of a fuzzy set: nonspecificity, Information Sciences 177 (2007) 4430–4444.
[5] J.H. Dai, Q. Xu, Approximations and uncertainty measures in incomplete information systems, Information Sciences (2012), http://dx.doi.org/10.1016/j.ins.2012.02.032.
[6] D. Dubois, H. Prade, Foreword, in: Z. Pawlak, Rough Sets, Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Dordrecht, 1991, pp. ix–xi.
[7] I. Düntsch, G. Gediga, Uncertainty measures of rough set prediction, Artificial Intelligence 106 (1998) 109–137.
[8] I. Düntsch, G. Gediga, Rough Set Data Analysis: A Road to Non-invasive Knowledge Discovery, Methodos Publishers (UK), 2000.
[9] I. Düntsch, G. Gediga, Roughian: rough information analysis, International Journal of Intelligent Systems 16 (2001) 121–147.
[10] Q.R. Feng, D.Q. Miao, J. Zhou, Y. Cheng, A novel measure of knowledge granularity in rough sets, International Journal of Granular Computing, Rough Sets and Intelligent Systems 1 (2010) 233–251.
[11] P.C. Fishburn, Utility Theory for Decision Making, John Wiley and Sons, New York, 1970.
[12] S. French, Decision Theory: An Introduction to the Mathematics of Rationality, John Wiley and Sons, New York, 1986.
[13] R.V.L. Hartley, Transmission of information, Bell System Technical Journal 7 (1928) 535–564.
[14] O. Holschke, J. Rake, O. Levina, Granularity as a cognitive factor in the effectiveness of business process model reuse, in: Proceedings of the Seventh International Conference on Business Process Management, LNCS 5701, 2009, pp. 245–260.
[15] V.N. Huynh, Y. Nakamori, A roughness measure for fuzzy sets, Information Sciences 173 (2005) 255–275.
[16] G.J. Klir, T.A. Golger, Fuzzy Sets, Uncertainty, and Information, Prentice Hall, Englewood Cliffs, NJ, 1988.
[17] D.H. Krantz, R.D. Luce, P. Suppes, A. Tversky, Foundations of measurement, Additive and Polynomial Representations, vol. 1, Academic Press, New York, 1971.
[18] T.T. Lee, An information-theoretic analysis of relational databases. Part I: data dependencies and information metric, IEEE Transactions on Software Engineering 13 (1987) 1049–1061.
[19] H.X. Li, H. Lin, Y.L. Li, A.P. Wu, Granularity research based on panweighted field of pansystems, Kybernetes 37 (2008) 1438–1443.
[20] Y.L. Li, J.F. Tang, K.S. Chin, Y. Han, X.G. Luo, A rough set approach for estimating correlation measures in quality function deployment, Information Sciences 189 (2012) 126–142.
[21] Y.L. Li, J.F. Tang, K.S. Chin, X.G. Luo, Y. Han, Rough set-based approach for modeling relationship measures in product planning, Information Sciences 193 (2012) 199–217.
[22] J.Y. Liang, Uncertainty and feature selection in rough set theory, in: Proceedings of the Sixth International Conference on Rough Sets and Knowledge Technology, LNCS(LNAI) 6954, 2011, pp. 8–15.
[23] J.Y. Liang, K.S. Chin, C.Y. Dang, R.C.M. Yam, A new method for measuring uncertainty and fuzziness in rough set theory, International Journal of General Systems 31 (2002) 331–342.
[24] J.Y. Liang, Y.H. Qian, Axiomatic approach of knowledge granulation in information system, in: Proceedings Nineteenth Australian Joint Conference on Artificial Intelligence, LNCS(LNAI) 4304, 2006, pp. 1074–1078.

[25] J.Y. Liang, Z. Shi, The information entropy, rough entropy and knowledge granulation in rough set theory, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 12 (2004) 37–46.
[26] J.Y. Liang, Z. Shi, D. Li, M.J. Wierman, Information entropy rough entropy and knowledge granulation in incomplete information systems, International Journal of General Systems 35 (2006) 641–654.
[27] J.Y. Liang, J.H. Wang, Y.H. Qian, A new measure of uncertainty based on knowledge granulation for rough sets, Information Sciences 179 (2009) 458–470.
[28] D. Liu, T.R. Li, D. Ruan, Probabilistic model criteria with decision-theoretic rough sets, Information Sciences 181 (2011) 3709–3722.
[29] J.S. Mi, X.M. Li, H.Y. Zhao, T. Feng, Information-theoretic measure of uncertainty in generalized fuzzy rough sets, in: Proceedings of Eleventh International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing, LNCS(LNAI) 4482, 2007, pp. 63–70.
[30] D.Q. Miao, S.D. Fan, The calculation of knowledge granulation and its application, Systems Engineering Theory and Practice 22 (2002) 48–56 (in Chinese).
[31] D.Q. Miao, J. Wang, On the relationships between information entropy and roughness of knowledge in rough set theory, Pattern Recognition and Artificial Intelligence 11 (1998) 34–40 (in Chinese).
[32] F. Min, H.P. He, Y.H. Qian, W. Zhu, Test-cost-sensitive attribute reduction, Information Sciences 181 (2011) 4928–4942.
[33] Z. Pawlak, Rough sets, International Journal of Computer and Information Sciences 11 (1982) 341–356.
[34] Z. Pawlak, Rough Sets, Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Dordrecht, 1991.
[35] Z. Pawlak, Preface, in: R. Słowinski (Ed.), Intelligent Decision Support, Handbook of Applications and Advances of the Rough Sets Theory, Kluwer Academic Publishers, Dordrecht, 1992, pp. ix–xi.
[36] Z. Pawlak, Granularity of knowledge, indiscernibility and rough sets, in: Proceedings of the 1998 IEEE International Conference on Fuzzy Systems, 1998, pp. 106–110.
[37] Z. Pawlak, A. Skowron, Rudiments of rough sets, Information Sciences 177 (2007) 3–27.
[38] Z. Pawlak, R. Słowinski, Rough set approach to multi-attribute decision analysis, European Journal of Operational Research 72 (1994) 443–459.
[39] W. Pedrycz, A. Skowron, V. Kreinovich (Eds.), Handbook of Granular Computing, John Wiley and Sons, Chichester, 2008.
[40] Y.H. Qian, J.Y. Liang, Combination entropy and combination granulation in rough set theory, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 16 (2008) 179–193.
[41] Y.H. Qian, J.Y. Liang, C.Y. Dang, Knowledge structure knowledge granulation and knowledge distance in a knowledge base, International Journal of Approximate Reasoning 50 (2009) 174–188.
[42] Y.H. Qian, J.Y. Liang, C.Y. Dang, Converse approximation and rule extraction from decision tables in rough set theory, Computers and Mathematics with Applications 55 (2008) 1754–1765.
[43] F.S. Roberts, Measurement Theory, Addison-Wesley, Reading, Massachusetts, 1979.
[44] C.E. Shannon, A mathematical theory of communication I, II, Bell System Technical Journal 27 (1948) 379–423, 623–656.
[45] D. Ślęzak, Approximate entropy reducts, Fundamenta Informaticae 53 (2002) 365–390.
[46] D. Ślęzak, Degrees of conditional (in)dependence: a framework for approximate Bayesian networks and examples related to the rough set-based feature selection, Information Sciences 179 (2009) 197–209.
[47] D. Ślęzak, Rough sets and functional dependencies in data: foundations of association reducts, in: LNCS Transaction on Computational Science V, LNCS 5540, 2009, pp. 182–205.
[48] J.H. Wang, J.Y. Liang, Y.H. Qian, C.Y. Dang, Uncertainty measure of rough sets based on a knowledge granulation for incomplete information systems, International Journal of Uncertainty Fuzziness and Knowledge-Based Systems 16 (2008) 233–234.
[49] W. Wei, J.Y. Liang, Y.H. Qian, A comparative study of rough sets for hybrid data, Information Sciences 190 (2012) 1–16.
[50] M.J. Wierman, Measuring uncertainty in rough set theory, International Journal of General Systems 28 (1999) 283–297.
[51] S.K.M. Wong, W. Ziarko, R.L. Ye, Comparison of rough-set and statistical methods in inductive learning, International Journal of Man-Machine Studies 25 (1986) 53–72.
[52] W.Z. Wu, Y. Leung, Theory and applications of granular labelled partitions in multi-scale decision tables, Information Sciences 181 (2011) 3878–3897.
[53] B.W. Xu, Y.M. Zhou, H.M. Lu, An improved accuracy measure for rough sets, Journal of Computer and System Sciences 71 (2005) 163–173.
[54] W.H. Xu, X.Y. Zhang, W.X. Zhang, Knowledge granulation knowledge entropy an knowledge uncertainty measure in ordered information systems, Applied Soft Computing 9 (2009) 1244–1251.
[55] J.T. Yao, A ten-year review of granular computing, in: Proceedings of the 2007 IEEE Internationational Conference on Granular Computing, 2007, pp. 734–739.
[56] Y.Y. Yao, Probabilistic approaches to rough sets, Expert System 20 (2003) 287–297.
[57] Y.Y. Yao, Information-theoretic measures for knowledge discovery and data mining, in: Karmeshu (Ed.), Entropy Measures, Maximum Entropy and Emerging Applications, Springer, Berlin, 2003, pp. 115–136.
[58] Y.Y. Yao, A partition model of granular computing, in: LNCS Transactions on Rough Sets I, LNCS 3100, 2004, pp. 232–253.
[59] Y.Y. Yao, The art of granular computing, in: Proceeding of International Conference on Rough Sets and Intelligent Systems Paradigms, LNCS(LNAI) 4585, 2007, pp. 101–112.
[60] Y.Y. Yao, A unified framework of granular computing, in: W. Pedrycz, A. Skowron, V. Kreinovich (Eds.), Handbook of Granular Computing, John Wiley and Sons, Chichester, 2008, pp. 401–410.
[61] Y.Y. Yao, Notes on rough set approximations and associated measures, Journal of Zhejiang Ocean University (Natural Science) 29 (2010) 399–410.
[62] L.Z. Zadeh, Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems 90 (1997) 111–127.
[63] B. Zhang, L. Zhang, Theory and Application of Problem Solving, North-Holland, Amsterdam, 1992.
[64] M.Q. Zhao, Q. Yang, D.Z. Gao, Axiomatic definition of knowledge granularity and its construction method, in: Proceedings of the Third International Conference on Rough Sets and Knowledge Technology, LNCS(LNAI) 5009, 2008, pp. 348–354.
[65] P. Zhu, An improved axiomatic definition of information granulation, Fundamenta Informaticae 119 (2012) 231–247.
[66] P. Zhu, An axiomatic approach to roughness measure of rough sets, Fundamenta Informaticae 109 (2011) 463–480.
[67] P. Zhu, Q.Y. Wen, Information-theoretic measures associated with rough set approximations, Information Sciences, 2012, http://dx.doi.org/10.1016/j.ins.2012.05.014.