Granular Computing: Practices, Theories and Future Directions

Tsau Young (T. Y.) Lin
Department of Computer Science, San Jose State University
San Jose, California 95192, USA
tylin@cs.sjsu.edu

Article Outline

Glossary

All terms are explained in classical sets, but implicitly, we are assuming all terms and assertions do include fuzzified versions (if fuzzifiable).

Granulation : Granulation is an operation or a process of forming granules, with a granule being a collection of objects (points) that are drawn together by some constraints, such as indistinguishability, similarity or functionality.

Granular Structure : Granular Structure is the collection of granules, in which the internal structure of each granule is visible as a sub-structure. Informally speaking, granular structure is a collection of white box granules.

Quotient Structure : A quotient structure is the mathematical structure of the collection of granules, in which each granule is regarded as an element(point) of a set, but the interactions among granules are preserved. Informally speaking, quotient structure is a collection of black box granules.

The collection of $\{... -2, 0, 2, ...\}$ and $\{... -3, 1, 3, ...\}$ is a granular structure. Let E be the first subset (even integers) and O be second subset(the odd integers.) We write the two subsets by [E] and [O], when we think of them as points. Then the collection of [E] and [O] is the quotient structure.

Neighborhood System (Local Granular Model abb. Local GrC Model) : A domain of interests (a classical set) U is called the universe. To each point p in the universe, a family of subsets is assigned. Such a family (could be empty) and each such a subset is called a neighborhood system NS(p) at p and a neighborhood at p respectively. The collection β of such a family at every point of the universe is called a neighborhood system NS(U) of the universe. Neighborhood and neighborhood system are pre-GrC language; in granular computing, they are called granule and the granular structure respectively. The pair (U, β) is called a local granular model, since each granule is associated with some points.

Topological Neighborhood System : A neighborhood system is called a topological neighborhood system, if it satisfies the axioms of topology.

Binary Neighborhood System(Binary Granular Model; Binary GrC Model) : A binary neighborhood system is a neighborhood system defined by a binary relation R. A (right) neighborhood is defined as follows: $B(p) = \{x \mid (p, x) \in R\}$. The collection B of $B(p)$ at each p is the (right) binary neighborhood system. Similarly, we can define left version: A left neighborhood system L is defined by the $L(p) = \{x \mid (x, p) \in R\}$ at every point p. Note that the right and left neighborhood systems determine each other. The pair (U, β) is called a binary granular model, where β is right(left) neighborhood system or R.

pre-Topology : Pre-topology is a general term referring to the neighborhood system, which includes topological neighborhood system and binary neighborhood system as special cases. Technically, it is equivalent to neighborhood system.

Bag : A bag is similar to a set, but allows an element to appear more than once, For example $\{1, 2, 1, 2, 1\}$ is a bag, but not a set. If a bag contains n elements, we may say it is a n-bag. For example, previous bag is 5-bag.

Relational Structure (Relational Granular Model; Relational GrC Model) : A family of classical sets is called a universe and denoted by $\mathcal{U}$. A Cartesian product of a n-bag of $\mathcal{U}$ is called a n-product set. A n-ary relation is a subset of a n-product set. A collection β of n-ary relations (n could vary) is called a relational structure. The pair (U, β) is called a Relational GrC Model. Note that this relational structure, except the size, is similar to the relational structure (without functions) of the First Order Logic; Relational GrC Model permits n to be any cardinal number.

Partial Covering (Global Granular Model; Global GrC Model) : Let U be a classical set, called the universe. Let $\beta = \{F^1, F^2, \ldots\}$ be a family of subsets. Such a β is a partial covering, and (full)covering, if the union of β is the whole universe. The pair (U, β) is called Global Granular Model (Global GrC Model).

Equivalence Relation : A binary relation $\mathcal{R}$ is called an equivalence relation, if it has the following properties: Let u, v and w be elements of U.

reflexive: u $\mathcal{R}$ u
symmetric: u $\mathcal{R}$ v implies v $\mathcal{R}$ u
transitive: u $\mathcal{R}$ v and u $\mathcal{R}$ w implies u $\mathcal{R}$ w

Partition : A partition $\mathcal{P}$ of a classical set U is a collection of subsets that are mutually disjoint and their union is U. Each subset is called an equivalence class. This name is derived from the fact that partition is equivalent to the following equivalence relation: We define u $\mathcal{R}$ v, if and only if u and v belongs to the same equivalence class. Such $\mathcal{R}$ is the equivalence relation corresponding to the partition $\mathcal{P}$. Note that a partition is a special type of granular structure, so an equivalence class is a special granule.

1 Definition of the Subject and Its Importance

Granular Computing (GrC) is still in its inception stage, we use motivation as its statements of importance.

How Important is Granular Computing? Granulation seems to be a natural methodology deeply rooted in human thinking. Many daily "things" are routinely granulated into sub"things;" human body has been granulated into head, neck, and so forth. The notion is intrinsically fuzzy, vague and imprecise. Formalization is difficult, mathematicians idealized/simplified it into the notion of partitions (=equivalence relations), and have developed it into a fundamental part of mathematics, e.g., congruence in Euclidean geometry, quotient structures (groups, rings, etc) in algebra, the concept of "a. e." (almost every where) in analysis. Nevertheless, the notion of partitions (see glossary), which absolutely does not permit any overlapping among its granules, seems to be too restrictive for real world problems. Even in natural science, classification *does* permit small degree of overlapping; there are beings that are both appropriate subjects of zoology and botany. So a more general theory, namely, Granular Computing (GrC) is needed:

What is Granular Computing(GrC)? It has been a changing paradigm. We believe in incremental developments. Its development may be similar, though not as glorious, to that of classical geometry. Specific geometries, such as, Euclidean, hyperbolic, elliptic geometries, had appeared before the the unified theory was proposed in Kleins Erlangen program. Except some details, we are on the final line; in GrC2008 keynote, we (this section editor) verbally had proposed to regard the category theory based model (Eighth GrC Model) as the final GrC model; see Section 4.

Granular Computing is a recent label coined by Lin and Zadeh (see Introduction) to denote a set of common, even ancient, concepts and practices. The subject will be presented from various angles: What are the target concept (defined by examples), key constituents, and the current interpretations or semantic views? How far has it been formalized? What are the important applications?

1) The "Definition" of Target Concept: The key concept in Granular Computing is the concept of granulation, which has been "defined" implicitly by a set of intuitive examples (see Section 3 for more), including Zadeh's intuitive view. Here is a list of representative examples. In the following, we will call the domain of interests the universe and denoted it by U.

1. Many daily things have been routinely granulated into "sub"things; human body is granulated into head, neck, and etc. The notion is intrinsically fuzzy, vague and imprecise. A formal model just had been proposed by this writer in the keynote of GrC2008. Its final form should appear soon in the International Journal of Granular Computing, Rough Sets, and Intelligent Systems.
2. The space and time has been granulated intuitively into infinitesimal granules; a circle was viewed as a polygon with infinitesimal sides. The idea was known to Zeno (490BC, implicitly in his paradox), Archimedes (287-212BC) and etc. It led to the invention of calculus, topology, and non-standard analysis; we have modeled it in the First GrC Model.
3. The simplest kind of granulation is partition (see glossary). Its algebraic correspondence, equivalence relation, has played an important role in Euclidean Geometry.(300 BC)
4. Heisenberg uncertainty principle states that, in general, neither the momentum nor the position of a particle can be determined simultaneously with arbitrary great precision. In other words, a precise measurement of the momentum can only determine

a "neighborhood"(granule) of positions and vice versa. The idea is abstract into Third GrC Model.

5. A collection of fuzzy sets (in fuzzy control) or functions (e.g. Radial-Basis-Functions) which has the universal approximation property is useful granular structure in a function space or a set of fuzzy sets. The idea is modeled in Sixth GrC Model.
6. A committee in a human society is a granule in a social network. Observe that each member may play different roles. By viewing the collection of roles as a relational schema, a committee is a tuple, not necessary a subset. The idea is modeled in Fifth GrC Model and Second GrC Model.
7. A mathematical proof or computer program often contains some lemmas or subprograms. These lemmas or subprograms are granules. These are conceptual examples. They are modeled in Seventh GrC Model for computable domains, and in Ninth GrC Model for general cases.
8. Computers or clusters of computers in Grid/Cloud computing are granules. These are hardware examples. This also belongs to Seventh GrC Model.
9. Zadehs informal definition [39]: "information granulation involves partitioning a class of objects(points) into granules, with a granule being a clump of objects (points) which are drawn together by indistinguishability, similarity or functionality."

‘

2) A Category Theory based Formal model (Eighth GrC Models) is proposed to be *the* Formal Model for GrC. It realizes all classical examples given above; see Section 4. By specifying the abstract category to various ones, eight common models have been explained in this article. Two models are illustrated at the end of this section;

3) The Key Constituents: Two Operators, Three Semantic Views, and Four structures.

3.1) Two Operators: information granulation and integration. In granulating a problem, a dual action, namely, integrating the solutions of sub-problems is triggered. Here, we highlight some unusual points; see Section 6.

1. Recursively granulating a problem may take Np hard time to terminate.. Dynamic programming technique has been used in the classical cases.
2. Two distinct problems may have the same granulation. Divide/granulate and conquer may have deeper meanings. The EXTENSION Functor of Homological Algebra of Commutative group is illustrated. This is an unexplored area.

3.2) Three Semantics Views: Granules may be interpreted from

1. Uncertainty Theory: A granule is a unit of lacking precise knowledge.
2. Knowledge Engineering: A granule is a unit of Basic Knowledge (Information).

3. How-to-Solve/Compute-it: A granule is a sub-problem or software unit. It is a special type of basic knowledge.

Each view may have its own GrC theory; For example, Concept Approximations are useful in Second View, while Information Hiding is in Third View; see Section 7.

3.3) Four Structures: Granular, Quotient, Knowledge and Linguistic Structures:

1. Granular Structure (GrS): It is the collection of all granules. In the case of partition, GrS is the collection of the equivalence classes.

2. Quotient Structure (QS): If each granule is abstracted into a point and the intersections of granules are abstract to the interactions of points, then such a collection of points is called quotient structure. In the case of partition, the quotient structure is a classical set, called quotient set.

   Informally, GrS is a collection of white boxes (the content of granules are visible), while the quotient set is a collection of black boxes (the contents of granules are hided). The process of abstracting granular structure into quotient structure is called information hiding; see Examples in Section 5.

3. Knowledge structure: By giving each granule (point) in the quotient structure a meaningful symbol then the named quotient structure is called knowledge structure. The knowledge structure provides an intuitive view of the quotient structure; the symbols and interaction among symbols are in sync with the granules (points) and interactions among granules(points).

   In the case of n partitions (equivalence relations), the knowledge structure can be arranged into a n-column relational table. In the case of n binary relations, the table has been called binary information table, granular table or topological table [16], [19].

4. Linguistic structure: By giving each granule in the granular structure a word that reflects its meaning. The interactions among these words are reflected implicitly in precisiated natural language. (in knowledge structure, the interactions among symbols are explicitly reflected from the quotient structure) The linguistic structure is the domain of computing with words

5) Applications to Computer Security, Web Technology, and Complex Data.

1. Discretionary Access Control Model (DAC) [8], [18]. This structure has been captured in Third GrC Model. Based on it, information flows on DAC can be analyzed; this has been considered a very "difficult" area. As a consequence the Aggressive Chinese Wall Security Policy can be enforced, namely, the system can guarantee that a company's data will never flow into "enemy" hands, where "enemy" is a granule of companies that are in conflict.

2nd GrC	generalized to	5th GrC
U	$\longrightarrow$	$\mathcal{U} = \{U_j^h, h, j, = 1, 2, \ldots\}$
$F^1 \subseteq U$	$\longrightarrow$	$R^1 \subseteq U_1^1 \times U_2^1 \times \ldots$
$\ldots$	$\ldots$	$\ldots$
$F^j \subseteq U$	$\longrightarrow$	$R^j \subseteq U_1^j \times U_2^j \times \ldots$
$\ldots$	$\ldots$	$\ldots$

Table 1: Generalize 2nd to 5th GrC Models

2. Documents can be clustered into a simplicial complex (a common structure in Combinatorial Topology [34]) of keywords and co-occurring keywordsets. The set of keywords can be regarded as a set of vertices, and the collection of co-occurring keywords (within a small neighborhood) is a set of simplexes. Together, they form a simplicial complex [21], [23]. Simplexes are granules; a simplicial complex is a Second GrC model.
3. Granular computing has been used to solve the modeling problem of complex architectures [25]. It uses Fifth GrC model.

Illustration of Working Formal Models: The most general model is expressed in category theory. However, for easiness, we explain the Second GrC model first.

1) *Second GrC Model.* Let U be a classical set, called the universe. Let $\beta = \{F^1, F^2, \ldots\}$ be a family of subsets. Then the pair (U, β), called Global GrC Model or the Second GrC Model. The β, some time, is called Partial Covering(PCov)

A *granule* can be intuitively defined as a clump of objects that are drawn together by the constraints X, where X can be indistinguishability, similarity or functionality and etc [39], [15]. In Second GrC model, a granule is a set, namely, we have implicitly assumed that the constraints are uniform. In general, each object may receive distinct constraints, so in Fifth GrC model, a granule is a tuple, not necessarily a set: One can regard the constraints as a schema (of a relational database), and a granule is a tuple under such a schema. In this case the collection of granules are tuples from various relations.

To understand the next model, Table 1, that explains the generalization process, may be helpful.

2) *Fifth GrC Model*

1. Let $\mathcal{U} = \{U_j^h, h, j, = 1, 2, \ldots\}$ be a given family of classical sets, called the universe. Note that distinct indices do not imply the sets are distinct.
2. Let $U_1^j \times U_2^j \times \ldots; j = 1, 2, \ldots$ be a family of Cartesian products of various lengths.
3. Recall that an n-ary relation is a subset $R^j \subseteq U_1^j \times U_2^j \times \ldots U_n^j$.
4. Let $\beta = \{R^1, R^2, \ldots\}$ be a given family of n-ary relations for various n.

Then the pair $(\mathcal{U}, \beta)$, called Relational GrC Model, is the formal definition of Fifth GrC Model

2 Introduction

What is Granular Computing (GrC)? The best approach is to trace how the intuitions have been evolved. For this purpose, let us recall the event. In the academic year 1996-97, when Lin had his sabbatical leave at Berkeley, Zadeh suggested granular mathematics (GrM) to be his research area. To limit the scope, Lin proposed the term granular computing [40]. So, at the beginning, roughly GrC is the computable part of granular mathematics.

What is Granular Mathematics (GrM)? Zadeh in his 1979 paper [38] had implicitly explained his view. Here, we take a simpler view: It is a new mathematics, in which "points" are replaced by or associated with "granules." We call this process granulation, and the collection of granules the granular structure

What is a Granule? This is the main topic. There is an obvious candidate, namely, an equivalence class of a *partition* that is an ancient notion in mathematics. In general, a granule can be a crisp/fuzzy subset, a function, an algorithm, a random variable(measurable function), a generalized functions and etc.

Traditionally, "How to solve it " [31] cannot be any part of formal mathematics, however, "how to compute" is an integral part of computing. So GrC has included the mathematical/computational problem solving practices.

3 Classical Examples of Granulation

- Ancient Examples

E1 Granulation of human body: the granules are head, neck, body, hand and etc.

Many daily things are routinely granulated into "sub"things, probably since ancient time. Human body is granulated into head, neck, and etc. Currently, there are no flawless formal models to capture such intuitive concept yet. The obvious one does not work well: One can easily write down a fuzzy membership function to represent a head, a neck or a body. However, the likelihood of any two persons to write down the same membership function for the same granule (e.g., head) is extremely low. Hence we need a much more subtle theory. A new proposal on qualitative fuzzy set theory is in preparation by this author, this example can be realized.

E2 Granulation of the Space and Time

The space and time has been granulated intuitively into infinitesimal granules by early scientists; this notion was knwon to Zeno, Archimedes, and etc. This intuitive notion led to the invention of calculus by Newton and Leibniz. However, its formalizations, theory of limit (18th century), topology (early 20 century [28]) and non-standard analysis (mid 20th century [32]) are relatively recent. This ancient example, inspired two models, Local GrC model (First GrC Model) and Global GrC Model(Second GrC Model).

E3 Partition

The simplest kind of granulation is partition (see glossary). Its algebraic correspondence, equivalence relation, has played an important role in Euclidean Geometry.(300 BC)

- Modern Examples

E4 Local Granules of Uncertainty - from Quantum Mechanics

Heisenberg uncertainty principle states that, in general, neither the momentum nor the position of a particle can be determined simultaneously with arbitrary great precision. In other words, a point of the momentum (a precise measurement) can determine only a "neighborhood" of positions and vice versa. The idea is simplified into Third GrC Model (Binary GrC Model).

E5 Local Granules of Basic Knowledge - Discretionary Access Control Models in Computer Security

This example is a common data structure in many computer systems. For each user, there is a set of users(friends) who can access his files, or a set of users(foe), who cannot access his files; this set is called explicitly denied access list.

Examples [E4] and [E5] are "serious" examples. The idea is modeled in third GrC Model. Mathematically, it is equivalent to a binary relation. Geometrically, a binary relation is a graph, or network.

E6 Global Granules of Knowledge - Simplicial Complexes.

A simplicial complex consists of two objects: one is a finite set of vertices, Another one is a family of subsets, called simplexes, of vertices that satisfy the closed condition, namely a subset of a simplex is a simplex. It is a common mathematical structure in combinatorial topology. Currently, it is finding its way to web technology [21].

- Further Examples

Next, we give examples of non-commutative granules, which is a generalization of binary relation (Binary GrC Model; Third GrC Model).

E7 A committee in a human society (a set of human beings) is a granule.

A committee in a human society (a set of human beings) is a granule. Observe that each member may play different roles, so the committee may not consist of homogenous members; so the members cannot exchange their roles. By viewing the collection of roles as a relational schema, a committee is a tuple. This idea is modeled in Fifth GrC Model. Observe that different types of committees have different schema. The set of committees under the same schema forms a relation. A collection of n-nary relations for various n in a society can be viewed as a granulation of the society.

In these examples, granules can be fuzzy sets. Since a fuzzy set is characterized by a membership function, so we have generalized the idea further

E8 A granule can be a function, a random variable (a measurable function) or even a generalized function (such as Dirac delta function).

This class of examples led to Sixth GrC model.

Traditionally, "How to solve it " [31] has not been any part of formal mathematics, however, "how to compute" is an integral part of computing. So GrC includes some mathematical/computational practices.

E9 A granule can be a lemma in a mathematical proof, a subprogram in a program, or a machine/cluster of machines in a grid/clouding computing environment. Formally, within the computable domains, a granule is a sub-Turing machine. This is modeled in Seventh GrC Model. For general cases, they are included in the Ninth GrC Model.

E10 Computers or clusters of computers in Grid/Cloud coputing are granules. These are hardware examples. This also belongs to Seventh GrC Model.

4 Formal Models of Granulation

Based on these examples, nine models are discussed. The category theory based model (Eighth GrC Model) has been proposed by this writer in the keynote of GrC2008 as *the formal model of GrC*. Rest of the eight models are basically "convenient models" in the sense that they can be derived from the most general model, but for convenient, they are modeled independently.

4.1 First GrC Models and Ancient Examples

This model is derived from the Ancient Example E2, which probably is the most lively example in the early notion of granules. It led to the invention of calculus by Newton and Leibniz. Actually the idea was much more ancient; it was in the mind of Archimedes, Zeno, and etc. Yet the solutions were in modern time. Two formalizations had emerged. One was the concept of limit(19th century) that led to the notion of topology (early 20th century). The other one was the non-standard analysis (mid 20th century), which formally realized the original intuition.

These developments conclude that the following concepts are "equivalent" (not in the formal mathematical sense, as the first one is merely an intuitive notion):

1. The ancient intuitive notion of infinitesimal granule.

2. The formal infinitesimal granule in the non standard analysis.

3. Topological Neighborhood System (TNS) in standard world.

In other words, the ancient intuition of infinitesimal granules (with the required properties) is realized, not by a set, but by a family of subsets, that satisfies the axioms of topology. Nevertheless, in this paper a (modern) granule will refer to a neighborhood, but, not to the whole neighborhood system.

The notion of topology can be defined in two ways:

1. A topology τ is a family of subsets, called open sets, that satisfies the (global version) axioms of topology.

2. A topology, called topological neighborhood system (TNS), is an assignment that associates each point p a family of subsets, TNS(p), that satisfies the (local version) axioms of topology

   .

These two definitions lead us to First and Second GrC Models (Local and Global GrC Models). In this sub-section, we will focus on the First GrC Model: Let U and V be two classical sets. Let NS be a mapping, called neighborhood system(NS),

$$NS : V \longrightarrow 2^{(P(U))},$$

where $P(X)$ is the family of all crisp/fuzzy subsets of X. 2^Y is the family of all crisp subsets of Y, where $Y = P(U)$. In other words, NS associates each point p in V, a family $NS(p)$ of crisp/fuzzy subsets of U. Such a subset is called a neighborhood (granule) at p, and $NS(p)$ is called a neighborhood system at p.

Definition 1 *First GrC Model: The 3-tuple* (V, U, β) *is called* **Local GrC Model**, *where* β *is a neighborhood system (NS). If* $V = U$, *the 3-tuple is reduced to a pair* (U, β). *In addition, if we require NS to satisfy the topological axioms, then it becomes a TNS.*

Some Intuition behind the NS:

The following arguments are adopt from my pre-GrC paper [14]

(1) The Meaning of "Near"

The notion of near is rather difficult to formalize. Let us examine the following two examples.

1. Is Santa Monica "near" Los Angels? Answers could vary. For local residents, who have cars, answers are often "yes." For visitors, who have no cars, answers may be "no."

2. Is 1.73 "near" $\sqrt{3}$? Again answers vary; it depends on what should be the appropriate tolerance radius.

Intrinsically "near" is a subjective judgment. One might wonder whether there is a scientific theory for such subjective judgments? Mathematicians have offered a nice solution. They simply include all contexts into its formalism. Here is the formalism of the second question: Given the radius of an acceptable error, say, radius of errors 1/100 (a given context)

Is 1.73 "near" $\sqrt{3}$?

With the agreement 1/100 is acceptable, then 1.73 is near $\sqrt{3}$! In this case, all possible contexts are ϵ that represents all positive real numbers. Similarly, if a neighborhood system has been assigned to each city in Los Angeles area. For example, based on car driving, public transportation, walking and etc, we assign a neighborhood to *each* city for *each* context. Under such a concept of neighborhood system, we could have a definite answer for Example 1. So a proper formulation for such a question is:

Assuming that we are taking public transportations (a given context), Is p near q?

Therefore, a neighborhood system is a good infrastructure for addressing the concept of "near"!

These analysis leads to the following conclusions.

1. In Modeling, a neighborhood system is a good infrastructure for providing all possible contexts.

2. Under this model, in an application, selecting a context means selecting *a fixed* neighborhood as a unit of tolerance(uncertainty).

Now, under this concept, we will re-examine previous examples

Example 1 *If we have chosen "driving half an hour" as acceptable distance, then Santa Monica is "near" Los Angels.*

Example 2 *Let the collection of ϵ-neighborhoods be the neighborhood system for the real numbers R; Then(R, ϵ-neighborhoods) is a First GrC Model, where ϵ could take any real value Now, we re-state the previous example using this First GrC Model*

1. *Assuming we have agreed ϵ = 1/100 is acceptable, then 1.73 is "near" $\sqrt{3}$*

2. *But, if we have only agreed ϵ = 1/1000 then 1.73 is not "near" $\sqrt{3}$*

3. *Next let us consider a deeper question*

   Is the sequence 1, 1/2, 1/3, .., 1/n,... "near" zero?

   Then, it is possible, we can have a "yes" answer for all contexts: For any given context, namely, $\epsilon > 0$, there is a number N = [1/ϵ] + 1, such that, for all $n > N$, $1/n$ is "near" zero, where [1/ϵ] denotes the biggest integer $\leq 1/\epsilon$.

For readers who familiar with the standard (ϵ, δ)-definition of limit can spot the origin of neighborhood systems. Such a context free (all possible contexts) answer is precisely the classical notion of limits, $lim_{n\to\infty} 1/n = 0$. Using our language, we may say that limit is the context free answers of "near"

Perhaps we should also point out here that there is no context free answers for the question whether two points are "near."

Brief pre-GrC historical notes:

1. In 1988-89, Lin generalized TNS to the Neighborhood Systems(NS) by simply dropping the (local version) axioms of topology [7], [9] and apply it to approximate retrievals. Each neighborhood was treated as a unit of uncertainty.

2. In the same year (1989), Lin also examined a non-reflexive and symmetric binary relation (conflict of interests) for computer security from the view of NS [8].

3. Abstractly, Lin imposed NS structure on the attribute domains for approximate retrieval. Taking this view, we should mentioned that earlier D. Hsiao imposed equivalence relations on the domain for access precision in early 1970 [3], [36]. In 1980, S. Ginsburg and R. Hull had imposed partial ordering on attribute domains [4], [5].

4. In much earlier, NS was studied in [33] as a generalization of topology. Note that however, there are fundamental differences, for example, the concept of closures are different. The term pre-topology also has been used for referring NS and TNS.

5. In early GrC period, Lin, by mapping the NS onto Zadeh's intuitive definition, used NS as his first mathematical GrC model [15], [16], [17].

4.2 Second GrC Models and Modern Examples

As in the previous case, by dropping the global axioms of topology, we have Second GrC model.

Definition 2 *Second GrC Model: The Pair* (U, β), *where* β *is a family of subsets of* U, *is called* **Global GrC Model**. *The* β, *some time, is referred to as a partial covering(PCov).*

Note that Second GrC model is a special case of First GrC model: If we regard the sub-collection of all members of the partial covering β, that contains p, as a neighborhood system at p, then this Second GrC model is an example of First GrC model.

The modern example, simplicial complexes, is an important example of such a model: A simplicial complex consists of a set of vertices and a family of subsets, called simplexes, that satisfies the closed condition [34]

[Digression] Perhaps, it is worthwhile to note that

- the closed condition of simplicial complex is the apriori principle in association (rules) mining.

This observation play an important role in document clustering [23].

4.3 Third and Fourth GrC Models and Modern Examples

In this section, we will build a new model that realized modern example [E4] and [E5]. Recall that [E4] concludes that, a precise measure of the momentum can only determine a (probabilistic) "neighborhood" of positions; and [E5] concludes that in computer security, the Discretionary Access Control Model (DAC) assigns to each user p a family of users, Y_i, $i = 1, \ldots$, who can access p's data. In other words, each p is assigned a granule of friends.

To formalize these examples, let U and V be two classical sets. Each $p \in V$ is assigned a subset, $B(p)$, of "basic knowledge" (a set of friends or a "neighborhood" of positions).

$$p \longrightarrow B(p) = \{Y_i,\ i = 1, \ldots\} \subseteq U$$

Such a set $B(p)$ is called a (right) binary neighborhood and the collection $\{B(p) \mid \forall p \in V\}$ is called the binary neighborhood system (BNS).

Definition 3 *Third GrC Model: The 3-tuple (U, V, β), where β is a BNS, is called a* **Binary GrC Model**. *If $U = V$, then the 3-tuple is reduced a pair (U, β).*

Observe that BNS is equivalent to a binary relation(BR):

$$BR = \{(p, Y) \mid Y \in B(p) \text{ and } p \in V\}.$$

Conversely, a binary relation defines a (right) BNS as follows:

$$p \longrightarrow B(p) = \{Y \mid (p, Y) \in BR\}$$

So both modern examples give rise to BNS, which was called a binary granular structure in [15]. We would like to note that based on this (right) BNS, the (left) BNS can also be defined:

$$D(p) = \{Y \mid p \in B(Y)\} \text{ for all } p \in V\}.$$

Note that BNS is a special case of NS, namely, it is the case when the collection NS(p) is a singleton B(p). So the Third GrC Model is a special case of First GrC Model.

The algebraic notion, binary relations, in computer science, is often represented geometrically as graphs, networks, forest and etc. So Third GrC Model has captured most of the mathematical structure in computer science.

Next, instead of a single binary relation, we consider the case: β is a set of binary relations. It was called a [binary] knowledge base [15]. Such a collection naturally defines a NS.

Definition 4 *Fourth GrC Model: the Pair (U, β), where β is a set of binary relations, is called Multi-Binary GrC Model. This model is most useful in data bases; hence it has been called Binary Granular Data Model(BGDM), in the case of equivalence relations, it is called Granular Data Model(GDM)*

Observe that a Fourth GrC Model can be converted, say by a mapping $\mathcal{G}$, to a First Model. Conversely, a First GrC Model induces, say by $\mathcal{F}$, to a Fourth Model. So First and Fourth models are equivalent, but not naturally, namely, $\mathcal{G}$ and $\mathcal{F}$ are not inverse to each other.

4.4 Models for Further Examples

As we have observed in Section 1 that the collection of n objects that are "drawn together" is, not necessary a subset, but is a tuple in an n-ary relation. For example, if the universe is a human society then a group of people may be drawn into a committee with distinct roles, such as the chair, vice chair, secretary, treasurer, and etc. As every member has different role, they can not be swapped around. So the committee is not a set; it is a tuple under the schema that consists of distinct roles.

Definition 5 *Fifth GrC Model:*

1. *Let $\mathcal{U} = \{U_j^h, h, j, = 1, 2, \ldots\}$ be a given family of classical sets, called the universe. Note that distinct indices do not imply the sets are distinct.*
2. *Let $U_1^j \times U_2^j \times \ldots$ be a family of Cartesian products of various length.*
3. *Recall that an n-ary relation is a subset $R^j \subseteq U_1^j \times U_2^j \times \ldots U_n^j$.*
4. *Let $\beta = \{R^1, R^2, \ldots\}$ be a given family of n-ary relations for various n.*

The pair $(\mathcal{U}, \beta)$, called Relational GrC Model, is a formal definition of Fifth GrC Model

Note that this granular structure is the relational structure (without functions) in the First Order Logic, if n only varies through finite cardinal number

For next two models, we will use the language of category theory in next sub-section. We may note that we have not committed ourselves to every specific details yet.

Definition 6 *Sixth GrC Model is in the categories of functions, random variables and even generalized functions.*

Fuzzy sets are described by membership functions, so granules can be regarded as membership functions; note the First to Fifth GrC Models include fuzzy sets. Hence, we consider further generalizations: granules are functions, random variables (measurable functions) generalized functions (e.g. Dirac delta functions) .

In the case, a granule is a function, we may require that the granular structure (the collection of granules) has the universal approximation property, namely, any function in the universe can be approximated by the functions in the collections. The membership functions selected in fuzzy controls do have such properties. In neural networks, the functions generated by the activation functions also have such property [29]

In the case of probability/measure theory, quantum mechanics may be a good guiding example.

Definition 7 *Seventh GrC Model is in the category of Turing machines.*

For examples, a collection of lemmas in mathematical proof (mechanizable), a set of subprograms in a computer program, or a computer or cluster of computers in grid/cloud computing are granules in the model.

Definition 8 *Ninth GrC Model is in the category of qualitative fuzzy sets.*

This model was proposed after the Eighth GrC model. It has not been published in printing form yet. The idea is similar to the model that we have called it sofset(this is not a typo) [13]. It associates to each "real world" fuzzy set, a collection of membership functions; please watch for new development.

4.5 Category Theory Based Models

Now we generalize the category of sets to general categories. It is somewhat a surprise that this is basically the same as the category of relational databases [11]. In other words, the abstract structures of data and knowledge are similar. After analysis, it seems reasonable, because in GrC approach, the basic unit of knowledge is a granule of data.

Let us set up some language for Category Theory. A category consists of

1. A class of objects, and
2. A set Mor(X, Y) of morphisms for every ordered pair of objects X and Y, which satisfies certain properties. For this paper, the formal details are not important; we only need the language loosely.

Here are some examples.

1. The Category of Sets: The objects are classical sets. The morphisms are the maps.
2. The Category of Sets with binary relations as morhisms: The objects are classical sets. The morphisms are binary relations. This is the Category of Entity Relationships Models.
3. The Category of Power Sets: The object U_X is the power set $P(X)$ of a classical set X. Let U_Y be another object, where Y is another classical set. The morphisms are the maps, $P(f) : U_X \longrightarrow U_Y$ that are induced by maps $f : X \longrightarrow Y$.

Let CAT be a given category.

Definition 9 *Category Theory Based GrC Model:*

1. $\mathcal{C} = \{C_j^h, h, j, = 1, 2, \ldots\}$ *is a family of objects in the Category CAT.*
2. *There are families (which are bags; see glossary) of Cartesian products,* $C_1^j \times C_2^j \times \ldots$ *of objects,* $j = 1, 2, \ldots$ *of various lengths. They are called product objects.*
3. *An n-ary relation object* R^j *is a sub-object of the product object* $C_1^j \times C_2^j \times \ldots C_n^j$.
4. $\beta = \{R^1, R^2, \ldots\}$ *be a family of n-ary relations (n could vary).*

The pair $(\mathcal{C}, \beta)$, *called Categorical GrC Model (Eighth GrC model), is the formal model of granulation.*

By specifying the general category to various special cases, we have all models: By specifying the category to be the category of sets, we have Fifth GrC model. Further by limiting n to 2, we have the First GrC Model, and Fourth GrC Model. By assuming the symmetry for all n-ary relations, we have the Second GrC Model. By restricting the number of relations to be one and n=2, we have the Third GrC Model. For Sixth and Seventh, further researches are needed.

4.6 Overview of Early GrC Models

Schematically we summarize the relationships, based on the Granular Structures, of early GrC Models as follows: (the diagram will be different, if it is based on their approximation spaces)
"$\Rightarrow\Leftarrow$ " is a two way generalization but they are not inverse to each other.
"$\Rightarrow$, $\Uparrow$, and $\Downarrow$ " are one way generalizations.
"GM " means GrC Models and RST means Rough Set Model.

$$RST \Longrightarrow \begin{bmatrix} \textit{Global GM} & \Longrightarrow & \textit{Local GM} \\ \Downarrow & & \\ \textit{Multi Binary GM (BGDM)} & \begin{matrix}\Leftarrow \\ \Rightarrow\end{matrix} & \textit{Local GM} \\ \Uparrow & & \Downarrow \\ \textit{Binary GM} & \Rightarrow & \textit{Relation GM} \end{bmatrix} \Rightarrow \textit{Category GM}$$

5 Granular and Related Structures

Four structures are related to each other, they are all briefly explained in Section 1. We will focus on quotient structure here.

A granule can be examined from three states: We will use the following example to illustrate the idea. Let U be the set $(Z, +)$ of the integers as an additive group, Let E and O be the set of even and odd integers respectively.

1. Isolated State (Internal State): Let us consider the case E is in the isolated state, that is, independent from the universe U. In this case, only the internal structure of E is available to us: we can only know that the number 2 is playing the role of the identity in an additive group, but is *not* aware of the existence of 1 in the outside of E. So in isolated state, E is the additive group $(Z, +)$. Some time, this state may also be called internal state as only the internal structure is involved,

2. Embedded State (Conceptual State): In this state, E is a subgroup of U. We not only know that the number 2 is playing the role of the identity in E and we also *do know* that 2 is an even integer in U.

[Digress] In category theory, a subobject (in this case a subgroup) is often represented by a pair, (a group, the mapping of the group into the subgroup) ,

$$(Z, +)) \longrightarrow_{homomorphism} E \subseteq (Z, +)$$

The pair (the first $(Z, +), \longrightarrow_{homomorphism}$) is the subobject (subgroup). The first component $(Z, +)$ represents the internal structure of (or isolated state of) even integers, the middle E is the embedded state of even integers. Since this state involve with every aspect of the concept of a granule, one may call it conceptual state.

3. Quotient State (External State): Note that the quotient structure Q consists of two elements, namely, the two subset of even and odd integers (E and O, respectively). But the two subsets appear as two elements, $[E]$, and $[O]$ in Q. Note the bracketed symbols emphatically denote their roles as points(elements) of a set; no contents of subsets are visible. The interactions between the two subsets induce interactions between these two elements. So Q is the integer mod 2 (as additive group). In notations,

$$Q = (Z_2, +) \equiv \{[0]_2, [1]_2\}$$

as an additive group. So the quotient state of E is the element $[0]_2$ in Q. Since the quotient state only involves the external relationships with other granules, this state may also called external state.

With these preparations, we can define the concepts of granular structure(GrS), quotient structure(QS) and knowledge structures (KS): GrS is the collection of granules in embedded states. The quotient structure is the structure of granules in quotient states, namely, each granule is abstracted to an element(point) and the interactions among granules are abstracted to the interactions among elements (points).

Among these four structures only the quotient structure is a difficult concept. Here, we will illustrate only the quotient structure. We will start from the simplest case, namely, the GrS is a partition. By definition, a partition is a collection of equivalence classes that are mutually disjoint and their union is the whole universe. Hence, by abstracting each equivalence class to a point, we have a set of points that have no interactions. This implies:

Proposition 1 *The quotient structure of a partition is a classical set.*

Next, we need a lemma.

Lemma 1 *The collection of complete inverse image of a given mapping $f : U \longrightarrow V$ defines a partition. Namely, the collection $\{f^{-1}(f(p)) \mid p \in U\}$ forms a partition.*

So NS and BNS induce partitions on U. These partitions have been called the derived partition (or derived equivalence relation); see [15].

Let $B : U \longrightarrow P(U)$ and $NS : U \longrightarrow 2^{(P(U))}$ be a binary neighborhood system (BNS) and neighborhood system (NS) respectively. Then (U, NS) and (U, B) are First and Third GrC Models respectivley. By the previous Lemma, the complete inverse images of the two mappings,

B and NS, are partitions (equivalence relations). The two partitions, denoted by E_B and E_{NS}, induce two quotient sets U/E_B, U/E_{NS} respectively. Moreover, the pre-topologies, BNS and NS, of U naturally induce the respective pre-topologies on the quotient sets U/E_B and U/E_{NS}, respectively [9],[12].

Proposition 2 *Such pre-topological spaces U/E_B (BNS-space) and U/E_{NS} (NS-space) are the quotient structures of Binary and Local GrC Models (Third and First GrC Models) respectively.*

This example indicates that there are additional pre-topological structures on the information tables [19] [1].

Let (U, β) be a Global GrC Model, where $\beta = \{F^1, F^2, \ldots\}$ be a partial covering, We will use $[F^{j_1}], [F^{j_2}] \ldots$ to denote a set of points (when we think of $F^1, F^2, \ldots$ as points).

Proposition 3 *The granular structure, β, generates a semi-group $S(\beta)$ under set theoretical intersection. Then the quotient structure of $(U, S(\beta))$ is a semi-group generated by $[F^{j_1}], [F^{j_2}] \ldots$ under "intersection" $[F^{j_1}] \circ [F^{j_2}] = [F^{j_1} \cap F^{j_2}]$.*

This example indicates that there are additional algebraic structures on the information tables [24].

We observe that each quotient structure is not easy to determine. It depends on what are the mathematical structure under consideration. Let us recall some works from pure mathematics. Let the universe U be the ring Z of integers. In ring theory, the collection β of all prime ideals is the center of attentions. Let p be a prime number, then the prime ideal consists of following set

$$\{\ldots, -2p, -p, 0, p, 2p, \ldots\},$$

together with some algebraic structure.

By regarding each prime ideal as a point, The collection of prime ideals is a set, often denoted by Spec(Z) in algebraic geometry. The structure of prime ideals turns Spec(Z) into a topological space under Zariski topology.

Example 3 *Two granular and quotient structures*

1. *The quotient structure of (U, β), where β is the collection of the prime ideals, is a topological space Spec(Z) [27].*

2. *The quotient structure of $(U, S(\beta))$,where $S(\beta)$ is the semi-group generated by the intersection, is isomorphic to the semigroup of positive integers.*

In RST, the approximation spaces of the first and second examples are different. However in GrC, the approximations (under the knowledge engineering view) of the two examples are the same. In GrC, granules represent known basic units of knowledge (known concepts), hence the intersections of known concepts are known concepts. So in GrC we take all possible intersections of granules to approximate unknown concepts. This example clearly indicate the effective-ness of this view.

6 Integration - A Dual of Granulation

Classical Divide(partitioning) and Conquer is extended to Granulate and Conquer. In other words, we are considering the cases, in which the sub-problems may not be logically independent to each other. So the standard thinking may not work, for example

- ∘ It may take Np-hard time to granulate a problem, from top to bottom.

This issue is not new. Some cases have been addressed in the topic of "dynamic programming" in data structure courses. We have used "topological divide" to solve the cases in Third GrC model [20].

In this section, we will explore the following problem: For convenience, we will use the term, sub-structure, to denote (1) the collection of the internal structures of each granule and (2) the quotient structure. Nowe, we raise the following questions:

- Can we find more than one universes (granular models) that have the same sub-structure?
- Equivalently, Could two distinct problems, after granulation, have the same sub-structure?

In homological algebra, this is called the extension problem; in computer science, we feel that integration is a better name.

We will illustrate the concept in the following two cases

1) Simple Integration: no information structure.

2) Integration with information structure.

6.1 Simple Examples of Integration

This is truly unexplored area we will explain the case when granulations are partitions.

6.1.1 A Simple Partition

We will illustrate the key concepts by simple examples. Let the universe U be the *set* of all integers, $Z = \{\ldots,\ 1,\ 0,\ 1, \ldots\}$ and has the following granular structure, namely, a given collection of two granules

$$\{\ \ \{... - 2, 0, 2, ...\}, \{... - 3, 1, 3, ...\}\ \ \}.$$

We name the granules E and O (the even and odd integers). Using the languages of Section 5, we say E and O are in the embedded state. We will use $[E]$ and $[O]$ to denote the quotient state. So the set $Q = \{[E], [O]\}$ is the quotient set. Intuitively, Q is a set of black boxes; the internals of black boxes are invisible.

Now we will summarize the important structures of U so that integration can be formulated: Let $Int(X)$ denote the internal structure of a granule X, that is, the structure of X in isolation (forgetting itself as a subset of U).

1. $A = Int(\{... - 2, 0, 2, ...\})$ is the set of integers Z.
2. $B = Int(\{... - 3, 1, 3, ...\})$ is the set of integers Z.
3. Two copies of integers, A and B, are mapped to even integers and odd integers respectively.
4. Q is the quotient set that consists of $[E]$ and $[O]$ as two elements(points).

Schematically we are given the following situation:

$$Z = \left\{ \begin{array}{lcl} A = Int(E) & \longrightarrow & E \\ B = Int(O) & \longrightarrow & O \end{array} \right\} (\subseteq)\ U \ \longrightarrow Q = \{[E], [O]\}$$

6.1.2 Integrations on Partitions

From the point of view of problem solving,

- The quotient structure represents the recipe (a set of higher level instructions) of integrating the sub-solutions. It is the "Main Program" that integrates the returns of sub-program calls.

Here a sub-solution means the solution of a sub-problem that has been solved in isolation. Now the integration problem can be formulated as follows:

1. We only know that two copies of integers, A and B, are mapped to the two subsets of unknown universe.
2. The images of the mappings form a partition in the unknown universe U
3. But we do know its quotient structure $Q = \{[E], [O]\}$, namely a set of two elements, where $[E]$ represents the point that are abstracted from the image E of A in U; similarly $[O]$ is that of B.

Schematically we can summarized it as follows: the unknown U has two unknown subsets, Image A and Image B, that form a partition on unknown U, but with known quotient set, which is equivalent to the integers mod 2.

$$Z = \left\{ \begin{array}{lcl} A & \longrightarrow & ImageA \\ B & \longrightarrow & ImageB \end{array} \right\} \ (\subseteq) UnknownU \ \longrightarrow Q = Z_2$$

- **Can we construct the unknown universe? And is it unique?**

The answer is "yes" and is unique. It is the Cartesian product $Q \times Z$. Observe that from set theoretical point of view, $Q \times Z$ is equivalent to Z, so the constructed one is the old friend.

6.2 Integration on Partition with Information Structure

Let us consider a second view on the set of integers. But this time, the universe carries additional information, namely, the additive structure of integers, $(Z, +)$. This universe is denoted by $(U, +)$. Then

1. Int(E) is the *additive group* $(Z, +)$ of integers, and Int(O) is a *set* Z of integers.
2. The quotient structure $(Q, +) = \{[E], [O]; +\}$ is an additive group: [E]+[E]=[E], [E]+[O]=[O]+[E]=[O], [O]+[O]=[E]. This (Q, +) is often called the integer mod 2, and denoted by $(Z_2, +)$.

Again, we are given a similar situation

$$\left\{ \begin{array}{lll} (Z,+) = Int(E) & \longrightarrow_{(homomorphism)} & E \\ Z = Int(O) & \longrightarrow_{(map)} & O \end{array} \right\} \quad \subseteq (U,+) \longrightarrow (Q,+)$$

The upper arrow

$$(Z,+) = Int(E) \longrightarrow_{(homomorphism)} (U,+) \longrightarrow (Q,+)$$

is very similar to a short exact sequence(SES) in homological algebra.

Now, we will extend the SES to granules:

$$Int(GrS) \longrightarrow_{map} U \longrightarrow QS$$

where (1) Int(GrS) is the collection of the internal structures of all granules, (2)its image in U (under the map $\longrightarrow_{map}$) is GrS, the granular structure in U, and (3) QS denote the quotient structure of GrS. The map, $\longrightarrow_{map}$, is on-to-one within each granule, but two granules may have overlapping images in U. We will call this a *granular exact sequence.*

Can $(U, +)$ be reconstructed back?

The answer is more than YES; there are *two* solutions ! The two solutions are: $(Z_2 \times Z, +)$ and $(Z, +)$. They are not equivalent as additive groups. This fact can be expressed by the extension functor, namely, $EXT(Z, Z_2) \neq 0$.

The important question is

Could such a functor be extended to formal granular models?

The answer may be "yes." The granular exact sequences shall play a similar role as short exact sequences. There are some preliminary results in [22]; for example, the fact that the knowledge representation of symmetric binary relation is complete implies that the integration is unique.

6.3 Integrations on non-Partition Case

We will consider a simple example using the concept of granular exact sequence. Let $U = \{a, b, c\}$, let β be the set of all subsets of U. So the internal structure of the granules are: (we will use n-granule to denote a granule of n elements) One copy of 3-granule; three copies of 2-granule, and three copies of 1-granule. This is a Global GrC Model. The quotient structure is a collection of 7 elements(to find the quotient structure is not easy, but easy to verify). These elements are related by partial ordering, called "a face of". We will use the geometry to represent this quotient structure; it is a triangle spanned by **i**(1,0,0), **j**=(0,1,0), **k**=(0,0,1). The triangle can be viewed geometrically (a simplicial complex of closed triangle) as ONE open triangle, THREE open segments, and THREE vertices. It is a partial ordered set of 7 elements.

Now the granular exact sequence(three steps) can be described as follows:

1. 3-granule $\{g_1^3, g_2^3, g_3^3\} \longrightarrow$ a subset of unknow universe $\longrightarrow$ open triangle $\Delta(\mathbf{ijk})$
2. first 2-granule $\{g_1^2, g_2^2\} \longrightarrow$ a subset of unknow universe $\longrightarrow$ first open segement $\Delta(\mathbf{ij})$
3. second 2-granule $\{g_3^2, g_4^2\} \longrightarrow$ a subset of unknow universe $\longrightarrow$ second open segement$\Delta(\mathbf{ik})$
4. third 2-granule $\{g_5^2, g_6^2\} \longrightarrow$ a subset of unknow universe $\longrightarrow$ third open segement$\Delta(\mathbf{jk})$
5. first 1-granule $\{g_1^1\} \longrightarrow$ a subset of unknow universe $\longrightarrow$ first vertex $\Delta(\mathbf{i})$
6. second 1-granule $\{g_2^1\} \longrightarrow$ a subset of unknow universe $\longrightarrow$ second vertex $\Delta(\mathbf{j})$
7. third 1-granule $\{g_3^1\} \longrightarrow$ a subset of unknow universe $\longrightarrow$ third vertex $\Delta(\mathbf{k})$

We will use $I(g_j^i)$ to denote the image in the unknown universe. We do know many of distinct g_j^i are mapped to the same point in the quotient structure. Here are the mappings

1. 3-granule $\{I(g_1^3), I(g_2^3), I(g_3^3)\} \longrightarrow \Delta(\mathbf{ijk})$
2. first 2-granule $\{I(g_1^2), I(g_2^2)\} \longrightarrow \Delta(\mathbf{ij})$
3. second 2-granule $\{I(g_3^2), I(g_4^2)\} \longrightarrow \Delta(\mathbf{ik})$
4. third 2-granule $\{I(g_5^2), I(g_6^2)\} \longrightarrow \Delta(\mathbf{jk})$
5. first 1-granule $\{I(g_1^1)\} \longrightarrow \Delta(\mathbf{i})$
6. second 1-granule $\{I(g_2^1)\} \longrightarrow \Delta(\mathbf{j})$
7. third 1-granule $\{I(g_3^1)\} \longrightarrow \Delta(\mathbf{k})$

From the maps above, we may make some identifications: Observe that from the maps, 1st, 2nd, 5th, we can identify, without losing generality, that $I(g_1^3) = I(g_1^2) = I(g_1^1)$. By similar arguments, we have

1. $I(g_1^3) = I(g_1^2) = I(g_3^2) = I(g_1^1) \longrightarrow \Delta(\mathbf{i})$

2. $I(g_2^3) = I(g_2^2) = I(g_5^2) = I(g_2^1) \longrightarrow \Delta(\mathbf{j})$
3. $I(g_3^3) = I(g_4^2) = I(g_6^2) = I(g_3^1) \longrightarrow \Delta(\mathbf{k})$

So the 12 points are actually three points; they will be denoted by $\{I(g_1^3), I(g_2^3), I(g_3^3)\}$. Thus the granular structure of the unknown universe is: One 3-granule $\{I(g_1^3), I(g_2^3), I(g_3^3)\}$, three of its 2-subgranules $\{I(g_1^3), I(g_2^3)\}$, $\{I(g_1^3), I(g_3^3)\}$, $\{I(g_2^3), I(g_3^3)\}$ and three of 1-subgranules $\{I(g_1^3)\}$, $\{I(g_2^3)\}$, $\{I(g_3^3)\}$. So we have re-captured the unknown U and β - This is the integration. The key question is: Is this U unique, the answer is no; we will skip it here.

7 Semantic Views

Granules may be interpreted from three views.

1. Uncertainty Theory: A granule is a unit of lacking precise knowledge. Both L.A. Zadeh and T. Y. Lin, who coined the label, started from the uncertainty theory. Lin took his neighborhood system as a system of uncertainty. Zadeh has a grand project [40].
2. Knowledge Engineering: A granule is a unit of basic knowledge (Information). In their book, D. Stanat and D. McAllister state "Knowledge varies in sophistication from simple classification to ..." [35]. A classification is a partition, so an equivalence class is a (unit of) basic knowledge; this view is also promoted by Pawlak [30]. So we believe: a granule is a (unit of) basic knowledge.
3. How-to-Solve/Compute-it: A granule is a sub-problem or software unit. It is a special type of basic knowledge.

Each view may have its own GrC theory: Some fundamental operators are:

1) *Information Hiding*: It is a transformation of granular structures into quotient structures (see glossary).

A quotient structure is the mathematical structure of the collection of granules, in which each granule is regarded as an element(point), and the interactions among granules are transformed into the interactions among elements. For example, in group theory a quotient group is a collection of cosets, in which each coset is regarded as an element and the multiplication of cosets is abstracted into the multiplication of elements in the quotient group. Using software engineering language, a granule in quotient structure is a black box, while in granular structure, it is a while box. These are easy case, for granulations that have non-empty overlappings, the quotient structure may not be easy to determine; see Section 5.

2) *Information Integration*: see Section 6.

3) *Knowledge Representation*: This amounts to give each point in the quotient structure (Section 5) a meaningful name; we will call it naming map. So knowledge representation is a composition of information hiding (a map from granular structure to quotient structure) and the naming map.

Knowledge representation is another way to discover new knowledge by organizing the knowledge structure in appropriate fashion, such as, an information table (= a relation in database theory). In GrC, the table often has additional algebraic and/or topological structures [19], [20].

4) *Concept Approximation*- Among three semantic views, knowledge engineering view is the most suitable view for this operation. Under this view, concept approximation is to express approximately the unknown concepts (arbitrary subsets of the universe) in terms of known basic knowledges (granules).

It is reasonable to regard that "and" ($\cap$) or "or" ($\cup$) operations of two basic units of known knowledge is also a known knowledge. So, we take finite intersection and any number of union as acceptable knowledge operations; the last one is derived from the topological spaces. However, we believe a negation of a known knowledge is not necessary a piece of known knowledge; so negation is not an acceptable operation.

Note that this approximation is different from *Rough Set Approximations*, which, including its generalizations, are based on the sole operation "or" ($\cup$). In other words, RST does not regard the "and" of two known concepts is also a known concept. So strictly speaking, rough set approximation is not a concept approximation.

GrC has many models; each model has slightly different approximation theory. First, we will explain

Second GrC Model (Global GrC Model): Let $\mathcal{C}_1$ be a given collection. Let $\mathcal{G}_1$ be the collection of all possible finite intersections of $\mathcal{C}_1$. Then by definition, the pair (U, β) is a Second GrC Model (Global GrC Model), where $\beta = \mathcal{C}_1$. Let G be a variable that varies through the collection $\mathcal{G}_1$, then we define

Definition 10 *Three approximations*

1. *Upper approximation:*
   $C[X] = \overline{\beta}[X] = \{p : \ \forall\ G, \text{ such that, } \ p \ \in G\ \&\ G \cap X \neq \emptyset\}$.

2. *Lower approximation:*
   $I[X] = \underline{\beta}[X] = \{p : \ \exists\ a\ G, \text{ such that, } p \ \in G\ \&\ G \subseteq X\}$.

3. *Closed set based upper approximation: [33] used closed closure operator. It applies closure operator repeatedly (transfinitely many steps) until the results stop growing. The space is called Frechet(V)-space or (V)-space.*
   $Cl[X] = X \cup C[X] \cup C[C[X]] \cup C[C[C[X]]] \ldots$ *(transfinite). For such a closure, it is a closed set*

Theorem 1 *The concept approximation space of Global GrC Model (Second GrC Model) is a topological space.*

We should also note that under the rough set approximation, this model is not a topological space.

Next, let us consider the approximation theory of

First GrC Model (Local GrC Model): Before, we proceed, let us examine the classical case, the topological space (U, τ). A subset $N(p) \subseteq U$ is a neighborhood of p, if $N(p)$ contains an open set that contains p. The union of all such open sets is the interior points of $N(p)$, which is the largest open set in $N(p)$; let us denote it by $O(p)$. Observe that $O(p)$ consists of every point that regards $N(p)$ as its neighborhood.

Now, we will generalize this idea to First GrC Model. Let NS(p) be the neighborhood system at p. Let $G(p)$ be the collection of all finite intersections of all neighborhoods in $NS(p)$. Let G be a variable that varies through $G(p)$.

Definition 11 *With such a G, the previous equations given above do define the appropriate notions of $C[X]$, $I[X]$, $Cl[X]$ for First, Third and Fourth GrC Models.*

We should caution here that in the Third GrC Model, there is at most one neighborhood at each point, so there is no "true" intersection.

Let $N(p)$ represent an arbitrary neighborhood of $NS(p)$ Let $C_N(p)$, called the center set of N(p), consists of all those points that have $N(p)$ as its neighborhood. (Note $C_N(p)$ is the generalization of $O(p)$ in TNS).

Now we will observe some harder question: Do the intersections of neighborhoods at distinct points belong to some $G(p)$?

Proposition 4 *Theorem of Intersections in NS*

1. *$N(p) \cap N(q)$ is in G(p)=G(q), iff $C_N(p) \cap C_N(q) \neq \emptyset$.*
2. *$N(p) \cap N(q)$ is not in any $G(p)$ $\forall$ p, iff $C_N(p) \cap C_N(q) = \emptyset$.*

If we regard $N(p)$ as a known basic knowledge, then we should define the knowledge operations: Let $\circ$ be the "and" operation of the basic knowledge (a neighborhood). For technical reasons, the $\emptyset$ is regard as a piece of the given basic knowledge.

Definition 12 *New operations in NS*

1. *$N(p) \circ N(q) = N(p) \cap N(q)$, iff $C_N(p) \cap C_N(q) \neq \emptyset$.*
2. *$N(p) \circ N(q) = \emptyset$, iff $C_N(p) \cap C_N(q) = \emptyset$.*

Observe that BNS is a special cases of NS. So we have:

Definition 13 *Let B be a BNS, then*

1. *$B(p) \circ B(q) = B(p) = B(q)$, iff $C_B(p) = C_B(q)$.*
2. *$B(p) \circ B(q) = \emptyset$, iff $C_B(p) \cap C_B(q) = \emptyset$. Note that $B(p) \cap B(q)$ may not empty, but it is not a neighborhood of any point.*

Observe that in Binary GrC Model, two basic knowledges are either the same or the set theoretical intersection does not represent any basic known knowledge.

5) Higher Order Concept Approximations

In Fifth GrC model, we consider the relations (subsets of product space) as basic knowledge. Any subset in a product space is a new unknown concept. We will illustrate the idea in the following case: U^j is either a copy of V or U. Moreover, in each product space, there is at most one copy of V, but no restrictions on the number of copies of U. If a Cartesian product has no V component, it is called U-product space. If there is one and only copy of V, it is called a product space with unique V.

1. u and u^1 is said to be directly related, if u and u^1 are in the same tuple (of a relation in β), where $u \in U$ and u^1 could be an element of U or V.
2. u and u^2 is said to be indirectly related, if there is a finite sequence $u_i, i = 1, 2, \ldots, t$ such that (1) u_i and u_{i+1} are directly related for every i, and (2) $u = u_1$ and $u^2 = u_t$.
3. An element $u \in U$ is said to be v-related ($v \in V$), if u and v are directly or indirectly related.
4. v-neighborhood, U_v, consists of all the $u \in U$ that is v-related.

In such a relational GrC model (U^j with unique V) induces a map:

$$B : V \longrightarrow 2^U; v \longrightarrow U_v,$$

Such a map defines a binary neighborhood system(BNS), where U_v is a v-neighborhood in U, and hence induces a *binary GrC model* (U, V, B). Next, we will consider the case $U = V$, and definde

Definition 14 *The high order approximations of Fifth GrC model is the approximations based on the v-neighborhood system.*

6) Concept Approximations in other Categories

In a category of functions, we will be interested in those granular structures that have universal approximation property. For a category of Turing machines (algorithms), it is still unclear as how to define the concept approximations.

8 Future Directions

Granular Computing is still in its inception stage; possible directions are wide open. Here we will focus only on those issues that are touched in this article.

1) Developments of Categories

In this paper, a category based model is proposed as *the* Formal Model for GrC. It can be specialized into various models to realize all the classical examples, including the first example, the granulation of human body. We should note that the claim on the realization of the first example is not in printing form yet. However, this author feels that it is important to inform the readers that is occurring.

The key to realize the first example is based on the category of qualitative fuzzy sets or sofsets(this is not a typo); please watch for new development. The categories of functions, random variables(measurable functions) and Turing machines are also need to be developed.

2) Developments of Granular Structures

Given a granular structure, we associate it with four structures (including itself). Among them quotient and knowledge structures are mathematical consequences of granular structure (if it is given mathematically). However the linguistic structure is not a mathematical formalism, but is a natural language formulation. In this paper, there is no report on this direction. We urge the readers to read Zadeh's article.

3) Imported Concepts

For information integration (this may correspond to Zadeh's term, "organization"), we have illustrated the idea imported from homological algebra. It is unclear if we have imported the correct thinking; but it does point out essential problems in granulate and conquer.

4) GrC and RST

RST has been served as the "model" of GrC developments. So there are a lots of similarity, here, we would like to caution the readers that, there are fundamental differences. For example, the fundamental views of uncertainty are quite different; Pawlak used "unable to specify" as the base of uncertainty, while GrC regard a granule as a unit of uncertainty (such as uncertainty in quantum mechanics) Also the approximation theories are different. Of course, there are other differences; we skip.

5) GrC, Databases and Data Mining

As we have pointed out that the categorical structures of databases and GrC are similar; at the same time, we need to point out the differences in semantics. Nevertheless, we are looking forward to the transfer of database technology to GrC. For data mining, please see database section on the articles by this author on "deductive data mining using GrC," and "mining decision rules using RST."

6) GrC and Fuzzy Logic

Most of expositions have been based on classical sets (and fuzzifiable concepts) For more intrinsic fuzzy view, we strongly recommend the readers to read Zadeh's article.

7) GrC and Clouding Computing

Theoretically, cloud computing can be related to the GrC on the category of Turing machines. We expect some strong interactions in near future.

As we have observed that GrC is deeply rooted in human thinking, we expect GrC will have many interactions with wide variety of areas.

References

[1] Chiang,I.J. Lin, T. Y., Liu, Y (2005): Table Representations of Granulations Revisited. In: 10th International Conference, RSFDGrC 2005, Regina, Canada, August 31 - September 3, 2005;728-737 Proceedings, Part I. Lecture Notes in Computer Science 3641 Springer 2005;728-737. ISBN 3-540-28653-5

[2] Dubois D, Prade H (1992): Putting Rough Sets and Fuzzy Sets Together. In: Slowinski R (ed) Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory, Kluwer Academic Publishers, DorDrecht, The Neitherland, 1992, 203-232.

[3] Hsiao, D. and Harary, F. (1970): A Formal System for Information Retrieval from Files. *Communications of the ACM*, 13(2), Feb. 1970; Corrigenda, 13(4), April 1970. 67-73.

[4] Seymour Ginsburg, Richard Hull (1981): Ordered Attribute Domains in the Relational Model. XP2 Workshop on Relational Database Theory June 22-24, 1981; The Pennsylvania State University, PA, USA.

[5] Seymour Ginsburg, Richard Hull(1983): Order Dependency in the Relational Model. Theor. Comput. Sci. 26: 149-195.

[6] Baliga,P,. Lin, T. Y.: Kolmogorov complexity based automata modeling for intrusion detection. 2005 IEEE International Conference on Granular Computing, Beijing, China, July 25-27, 2005, 387-392. IEEE 2005, ISBN 0-7803-9017-2

[7] Lin, T.Y. (1988): Neighborhood systems and relational database. In: Proceedings of the Sixteenth ACM Annual Conference on Computer Science, Atlanta, Georgia, USA, February 23-25, 1988, p 725. ACM 1988,ISBN 0-89791-260-8

[8] Lin, T. Y. (1989a): Chinese Wall Security Policy–An Aggressive Model, Proceedings of the Fifth Aerospace Computer Security Application Conference,Tuscon, Arizona, December 4-8, 1989, 286-293.

[9] Lin, T.Y. (1989b): Neighborhood systems and approximation in database and knowledge base systems. In *Proceedings of the Fourth International Symposium on Methodologies of Intelligent Systems (Poster Session)*, Charlotte, North Carolina, Oct 12, 1989, pages 75–86. (Available to the public from the National Technical Information Services, U. S. Departmetnof Commerce, 5285 Port Royal Rd., Springfield, VA 22161. NTIS price codesprited copy: A11 Micorfiche A01. Available to DOE and DOE contractors from the office od scientific and technical information P. O. Box 62, Oak Ridge, TN 37831; prices available from (615) 576-8401,)

[10] Lin, T. Y., Huang, K. J. Liu, Q. and Chen, W. (1990): Rough Sets, Neighborhood Systems and Approximation, Proceedings of the Fifth International Symposium on Methodologies of Intelligent Systems, Selected Papers, Knoxville, Tennessee, October 25-27, 1990,130-141. Library ofCongressCatalog Card Number:90-84690

[11] Lin, T. Y. (1990): Relational Data Models and Category Theory (Abstract). In: CSC '90, Proceedings of the ACM 18th Annual Computer Science Conference on Cooperation, February 20-22, 1990, p. 424. Sheraton Washington Hotel, Washington, DC, USA. ACM, 1990

[12] Lin, T. Y.(1992): Topological and Fuzzy Rough Sets, in: Decision Support by Experience - Application of the Rough Sets Theory, R. Slowinski (ed.), Kluwer Academic Publishers, DorDrecht, The Neitherland, 1992, 287-304.

[13] Lin, T. Y.(1996): A Set Theory for Soft Computing. In: Proceedings of 1996 IEEE International Conference on Fuzzy Systems,New Orleans, Louisiana, September 8-11, 1996, 1140-1146.

[14] Lin, T. Y.(1997): Neighborhood Systems -A Qualitative Theory for Fuzzy and Rough Sets, Advances in Machine Intelligence and Soft Computing, Volume IV. Ed. Paul Wang, 1997, Duke University, North Carolina, 132-155. ISBN:0-9643454-3-3

[15] Lin, T.Y. (1998a): Granular Computing on binary relations I: data mining and neighborhood systems. In Skoworn, A. and Polkowski, L., editors, *Rough Sets In Knowledge Discovery*, pages 107–121. Physica-Verlag, Heidelberg

[16] Lin, T.Y. (1998b) Granular Computing on Binary Relations II: Rough set representations and belief functions. In Skoworn, A. and Polkowski, L., editors, *Rough Sets In Knowledge Discovery*, pages 121–140. Physica-Verlag, Heidelberg

[17] Lin, T.Y. (1999): Granular Computing: Fuzzy logic and rough sets. In Zadeh, L. and Kacprzyk, J., editors, *Computing with Words in Information/Intelligent Systems*, pages 183–200. Physica-Verlag, Heidelberg.

[18] Lin, T. Y. (2003): Chinese Wall Security Policy Models: Information Flows and Confining Trojan Horses. In: Sabrina De Capitani di Vimercati, Indrakshi Ray, Indrajit Ray (Eds.): Data and Applications Security XVII: Status and Prospects, IFIP TC-11 WG 11.3

Seventeenth Annual Working Conference on Data and Application Security, August4-6, 2003, Estes Park, Colorado, USA. Kluwer Boston/Dordrecht/London 2004, 275-287. ISBN 1-4020-8069-7.

[19] Lin,T. Y. Liau, C. J (2005): Granular Computing and Rough Sets. In: Oded Maimon, Lior Rokach (Eds.): The Data Mining and Knowledge Discovery Handbook. Springer New York, NY, USA TS32005, 535-561. ISBN 0-387-24435-2.

[20] Lin, T. Y.(2005): Divide and Conquer in Granular Computing,: Topological Partition, In: the Proceeding of the 2005 North American Fuzzy Information Processing Society Annual Conference: Computing for Real World Applications, June 22-25, 2005, Ann Arbor, Michigan, USA

[21] Lin, T. Y., Chiang, I. J. (2005): A simplicial complex, a hypergraph, structure in the latent semantic space of document clustering. Int. J. Approx. Reasoning 40(1-2): 55-80 (2005)

[22] Lin, T. Y. (2006a): A Roadmap from Rough Set Theory to Granular Computing. In: Rough Sets and Knowledge Technology, First International Conference, RSKT 2006, Chongquing, China, July 24-26, 2006, Proceedings. Lecture Notes in Computer Science 4062 Springer 2006, 33-41. ISBN 3-540-36297-5

[23] Lin, T. Y., Sutojo,A., Hsu, J. D. (2006b): Concept Analysis and Web Clustering using Combinatorial Topology. In: Workshops Proceedings of the 6th IEEE International Conference on Data Mining (ICDM 2006), 18-22 December 2006, Hong Kong, China. IEEE Computer Society 2006, ISBN 0-7695-2702-7; 412-416

[24] Lin, T. Y. (2006): Granular Computing on Partitions, Coverings, and Neighborhod Systems. In: the Journal of Nanchang Institute of Technology Vol 25 No 2, 1-7. (Special Issue as The Proceeding of The International Forum on Theory of GrC from Rough Set Prospective.

[25] Liu, Y., Lin, T. Y., Xu, C.,Zhang,Q., Huang, L., He, P. (2008) Modeling Complex Architectures based on Granular Computing on Ontology. Submitted to IEEE TRANSACTIONS ON FUZZY SYSTEMS.

[26] Maclane S.(1975): Homology. Reprint of the 1975 edition. Classics in Mathematics. Springer-Verlag, Berlin, 1995. x+422 pp. ISBN 3-540-58662-8

[27] Mumford, D.: Introduction ot Algebriac Geometry, The Red Book of Varieties and Schemes, mimeographed notes from the Harvard Mathematics Department 1967, reprinted in Lecture Notes in Mathematics 1348, Springer-Verlag Berlin, Heidelberg , 1988.

[28] Munkres, J. (2000): Topology (2nd Edition), PRENTICE-HALL, 2000

[29] Park, J. W., Sandberg, I. W.,(1991): Universal Approximation Using Radial-Basis-Function Networks. Neural Computation 3, 246-257.

[30] Pawlak, Z. (1991): Rough sets. Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, DorDrecht, The Neitherland, 1991 ISBN 0-7923-1472-7

[31] G. Polya (1957): How to Solve It, 2nd ed., Princeton University Press, Princeton, New Jersey, USA 1957, ISBN 0-691-08097-6.

[32] Robinson, A. (1966): Non-standard Analysis North-Holland, Amsterdam, 1966

[33] Sierpenski, W. (trans by Krieger, C.) (1952): General Topology(Mathematical Exposition No 7), University of Toronto Press 1952.

[34] Spanier, E. H (1966): Algebraic Topology, McGraw Hill, 1966 (Paperback, Sprngr, Dec 6, 1994). Originally was published by McGraw, now @ 1966 Springer-Verlag, New York Inc]

[35] Stanat, D. and McAllister, D.:Discrete Mathematics in Computer Science, Prentice-Hall, , Englewood Cliffs, N. J. 1977.

[36] Wong, E., Chiang, T. C. (1971): Canonical Structure in Attribute Based File Organization. *Communications of the ACM* 14(9): 593-597 (1971)

[37] Zimmerman, H. (1991): *Fuzzy Set Theory –and its Applications.* Kluwer Acdamic Publisher DorDrecht, The Neitherland,.

[38] Zadeh, L.A.(1979): Fuzzy sets and Information Granularity, in: M. Gupta, R. Ragade, and R. Yager (Eds.), Advances in Fuzzy Set Theory and Applications, North-Holland, Amsterdam, 3-18, 1979.

[39] Zadeh,L. A. (1996): The Key Roles of Information Granulation and Fuzzy logic in Human Reasoning. In: *1996 IEEE International Conference on Fuzzy Systems*, New Orleans, Louisiana, September 8-11, 1, 1996.

[40] Zadeh, L.A. (1998): Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/ intelligent systems, Soft Computing, 2, 23-25.

[41] Zadeh, L.A. (1997): Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems 90, 111-127, 1997.