Comparison of Granular Computing Models in a Set-Theoretic Framework

Ying Wang and Duoqian Miao

Department of Computer Science and Technology,
Tongji University, Shanghai 201804, P.R. China
wangy@tongji.edu.cn, miaoduoqian@163.com

Abstract. Many granular computing models have been proposed. A set-theoretic framework for constructing granules is easy to understand. A granule is a subset of a universal set, and a granular structure is a family of subsets of the universal set. By comparing set-based granular structures, the relationships and differences among rough set modal, hierarchical multi-dimensional data model and multi-granulation rough set model are discussed in this paper.

Keywords: Rough sets, Granular computing, Granule construction, Multi-granulation, Problem solving.

1 Introduction

Granular computing (GrC) is originated to mimic human's ability to perceive the world and to solve the problem under multiple levels of granularity for reasons for example, imprecise data, simplicity, cost, and so on [1-7]. It is a general computation theory that effectively uses granules, such as classes, clusters, subsets, groups, and intervals, to build an efficient computational model for complex applications with huge amounts of data, information and knowledge. The most apparent characteristics of GrC are representation and reasoning using granules at different levels of abstraction, and switching between them in problem solving [8,9].

The two principle issues of GrC are construction of granules to represent a problem and inference with granules in problem solving [10,11,3]. The granulation of a problem concerns granules construction and a hierarchical granular structure formation by a family of granules. Inference with granules relates to choosing an appropriate level or levels of granularity and reasoning in a granular space.

Rough sets (RS) provide one of the useful computing models of GrC. A partition model in GrC can be easily derived from it [17]. In constructing set-based granular structures, we assume that a granule is a subset of a universal set, and a granular structure is constructed based on the standard set-inclusion relation on a family of subsets of the universal set [12,13]. Many GrC models have been proposed. In this paper, we compare three set-based models, namely, rough sets

F.M. Zanzotto et al. (Eds.): BI 2012, LNCS 7670, pp. 332–337, 2012.

model proposed by Pawlak [19], hierarchical multi-dimensional data model proposed by Feng and Miao [14,15], and multi-granulation rough sets by Qian and Liang [16]. The relations and differences of these models are investigated.

2 Set-Based Granular Structure in Rough Sets

In rough set theory, a universe is divided into indiscernible classes by an equivalence relation. The granules and granular structure can be defined as the equivalence classes [13].

2.1 Set-Based Definition of Granules

Definition 1. *Let U a finite nonempty universal set. A subset $g \in 2^U$ is defined as a granule, where 2^U is the power set of U.*

Definition 2. *For $g, g' \in 2^U$, if $g \subseteq g'$, g is a sub-granule of g' and g' a super-granule of g.*

Under the partial order $\subseteq$, the empty set ϕ is the smallest granule and the universe U is the largest granule. When constructing a granular structure, we may consider a family G of subsets of U and an order relation on G.

Definition 3. *Suppose $G \subseteq 2^U$ is a nonempty family of subsets of U. The poset $(G, \subseteq)$ is a granular structure, where $\subseteq$ is the set-inclusion relation.*

2.2 Set-Based Granule Construction and Granular Structure

Rough set theory gives a partition-based model of GrC [17]. The theory concerns the analysis of data given in a tabular form. Formally, an information system or information table is represented by:

$$IS = (U, A, V, f) = (U, A, \{V_a | a \in A\}, \{f_a | U \times a \longrightarrow V_a\})$$

where
U : is a finite nonempty set of objects;
A : is a finite nonempty set of attributes;
V_a : is a nonempty set of values for an attribute $a \in A$,ie. $V = Va$;
f_a: is an information function.

A basic notion of rough sets is a granulation of the sets of objects based on their descriptions. For an attribute subset $P \subseteq A$, one can define an equivalence relation E_P on U as:

$$xE_Py \Leftrightarrow \forall a \in P, f_a(x) = f_a(y)$$

The equivalence class containing x is denoted by $[x]_{Ep} = \{y | xE_Py\}$, and is a granule. The granular partition level constructed by equivalence relation E_P is consist of equivalence classes $\{[x]_{E_p}\}$. It can be denoted by $GLP(E_P)$ or

$GLP\{[x]_{E_p}\}$. Both $GLP(E_P)$ and $GLP\{[x]_{E_p}\}$ mean the granule level governed by equivalence relation E_P, but the later indicates it by equivalence classes.

In order to understand the RS from a view of GrC, more concepts are emphasized as follows.

Original granule level is the granular partition level $GLP(E_A)$, where A is a set including all attributes , is the finest nonzero granule level in the information table $IS = (U, A, V, f)$.

A granular space is the granule levels which a problem is resolved. Single-granulation granular space and multi-granulation granular space is distinguished by its numbers of granule level.

Basic granules and definable granules are related to granule levels. The granules $[x]_{E_p} = \{y|xE_Py\}$, constructed by equivalence relation E_P, are basic granules of $GLP(E_P)$. A set $X = \cup\{[x]_{E_p}|x \in U\}$, a union of basic granules of $GLP(E_P)$, is a single-granulation definable granule in the level $GLP(E_P)$. A multi-granulation definable granule is a union of basic granules from multiple granule levels.

Basic granules are different from definable granules. The basic ones can act as primary units to compose definable granules. A basic granule is the finest granule and cannot be decomposed in its granule level, but may be decomposed into basic granules of finer granule levels. A basic granule in the coarser granule level may be a definable granule in the finer level. A basic granule in the coarser granule level is a super-granule of some basic granules in the finer level, and a basic granule in the finer level is a sub-granule of a basic granule in the coarser granule level.

In information system $IS = (U, A, V, f)$, let attribute sets $P_i \subseteq A$, $i = \{1, 2, ..., N\}$. If $P_1 \subset P_2 \subset ... \subset P_N \subseteq A$, then $GLP(E_{P_N}) \prec ... \prec GLP(E_{P_2}) \prec GLP(E_{P_1})$. The basic granules of $GLP(E_{P_i})$ are nested.

The above step-finer levels of granule make up a granular structure. This provides an effective method to obtain a granular structure. In rough sets, the attribute sets which satisfy set inclusion relation $P_1 \subset P_2 \subset ... \subset P_N \subseteq A$, can be derived by adding attributes gradually.

From the above statement, the following conclusions can be drawn: (1) A granular space is consist of basic granules. Basic granules are different from definable granules. (2) A granular structure can be derived according to increasing or decreasing data description number, i.e. attribute number. (3) Processing data in different attribute description means in different granulation level. (4) The concepts of single-granulation and multi-granulation are proposed definitely.

2.3 Reduction from View of GrC

The knowledge reduction is a primary concept in rough sets. Reducing of knowledge consists of removing of superfluous partitions (equivalence relations) or /and superfluous basic categories in the knowledge base in such a way that the set of elementary categories in the knowledge base is preserved [19]. This procedure enables us to eliminate all unnecessary knowledge from the knowledge base, preserving only the part of the knowledge which is really useful.

From the view of GrC, the reduction in RS is the process to change the granular space from the original granule level to the coarsest level of problem solving.

Choosing a granular space from all possible granular structures is a huge work. Suppose a information table $IS = (U, A, V, f)$ have N attributes, attribute sets P_i containing M attributes is the subset of A, total possible number of granule levels $GLP(P_i)$ is C_N^M. The reduction of an IS is a NP-hard problem, a heuristic method is needed to find a reduction.

Let an attribute set $R \subset A$. If R is a reduction of A, after reducing, the actual problem-solving granule level jumps from $GLP(E_A)$ to $GLP(E_R)$, namely a granular space jumps. The problem can be resolved in this level, such as rules minning. In this sense, rough sets are considered as a single-granulation tool.

3 Granules in Hierarchical Multi-dimensional Data Model

The data in rough sets are given in a tabular form. If the attribute in the table is capable of generalizing to more abstracting levels, then one table should be changed to more tables. Every table provides one single-granulation granular space, so hierarchical multi-dimensional data model [14] contributes a multi-granulation granular space. The problem solving can be proceeded in multi-granulation, for example, hierarchical decision rules mining [15].

The discussions in section 2 are suitable for hierarchical multi-dimensional data model. With reference to that, the following definitions are obvious.

Given a information table $IS = (U, A, V, f)$ and concept hierarchical tree of all attributes. The original granule level of IS becomes the (0, 0, . . . ,0)th information table, with all attribute values at leaves level of their concept hierarchies respectively. The (0, 0, . . . ,0)th information table can be denoted as follows:

$$IS_{00\ldots0} = (U_{00\ldots0}, A, V^{00\ldots0}, f_{00\ldots0})$$

When the attribute A_i is generalized to k_i concept hierarchy, the $(k_1, k_2, ..., k_m)th$ information table is referred to as:

$$IS_{k_1k_2\ldots k_m} = (U_{k_1k_2\ldots k_m}, A, V^{k_1k_2\ldots k_m}, f_{k_1k_2\ldots k_m})$$

For more detail, see the papers [14,15].

In $IS_{k_1k_2\ldots k_m}$, let an attribute set $P \subset A$, R is a reduction of A.

Definition 4. *The basic granules in $GLP(E_P^{k_1k_2\ldots k_m})$ of $IS_{k_1k_2\ldots k_m}$ are denoted as:*

$$[x]_{E_P}^{k_1k_2\ldots k_m} = \{y | xE_P^{k_1k_2\ldots k_m} y\} \Leftrightarrow \forall a_i \in P, f_{k_i}(x) = f_{k_i}(y)$$

Definition 5. *The original granule level of $IS_{k_1k_2\ldots k_m}$ is denoted as $GLP(E_A^{k_1k_2\ldots k_m})$. The problem-solving granule level of it is denoted as $GLP(E_R^{k_1k_2\ldots k_m})$*

When the attribute A_i is generalizes from k_i -1 to k_i concept hierarchy, the basic granules get coarser. Some definable granules in granulation level k_i -1 are also changed to basic granules in granulation level k_i. The original granule level of the according information table is also changed, and gets coarser. A lattice is formatted by all the original granule levels of $IS_{k_1 k_2 \ldots k_m}$. This is proved in the literature [14].

In this section, the hierarchical multi-dimensional data model produces a multi-granulation granular space by generalizing the attributes. So the problem can be resolved in multi-granulation granular space and the results derived may be more flexible and efficiency than that in rough sets.

4 Granule Construction in Multi-granulation Rough Sets

Rough sets are mainly concerned with the approximation of sets described by a single binary relation on the universe. Qian et al. extend Pawlak's rough sets to a multi-granulation rough set model (MGRS), where the set approximations are defined by using multi-equivalence relation on the universe [16].

Definition 6. *Let P, Q be two partitions on the universe U, and $X \subseteq U$. The lower approximation and upper approximation of X in U are defined as:*

$$\underline{X}_{P+Q}=\{x : P(x) \subseteq X, \text{ or } Q(x) \subseteq X\} \text{ and } \overline{X}^{P+Q} =\sim \underline{(\sim X)}_{P+Q}$$

In the view of granular computing, the granules in MGRS are as follows: $[x]_{E_{P+Q}}= [x]_{E_P}$ or $[x]_{E_Q}$. Acorrding to statement in section 2, the granules in MGRS constitute a granular covering level, is denoted as $GLC(E_{P+Q})$. The granular space in MGRS is a multi-granulation granular covering level.

In the view of GrC, the relationships and differences among rough set model, hierarchical multi-dimensional data model and multi-granulation model can be discovered from their granular spaces. The granular space in rough sets is a single-granulation partition level, in hierarchical multi-dimensional data model is multi-granulation partition levels, and in MGRS is a two-granulation cover level.

A two-granulation cover level in MGRS is consisit of two single-granulation partition level in rough sets. The former has more granules and can produce more precise result[16] than any of its partition level in rough sets.

5 Conclusions

From the set-based granules construction, the hierarchical RS on multi-dimensional data model and MGRS model provide new granule levels according to rough sets by different methods. Many concepts are proposed in this paper definitely. The relations and differences of these models are investigated.

The more GrC models proposed the more granule levels occured. In actual applications, how to choose the most appropriate one becomes a challenge problem and is greatly problem-oriented. This will be the most important problem in the future.

Acknowledgments. This work was supported by National Natural Science Fund of China (Nos. 60970061, 61075056, 61103067).

References

1. Zedah, L.A.: Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Sets and Systems 90, 111–127 (1997)
2. Bargiela, A., Pedryz, W.: Granular Computing: An Introduction. Kluwer Academic Publishers, Boston (2002)
3. Bargiela, A., Pedryz, W.: The roots of Granular Computing. In: IEEE International Conference on Granular Computing, pp. 806–809 (2006)
4. Yao, Y.: Human-inspired granular computing. In: Yao, J.T. (ed.) Novel Developments in Granular Computing: Applications for Advanced Human Reasoning and Soft Computation, pp. 1–15. Information Science Reference, Herskey (2010)
5. Bargiela, A., Pedrycz, W.: Toward a theory of granular computing for human-centered information processing. IEEE Transactions on Fuzzy Systems 16(2), 320–330 (2008)
6. Pedrycz, W.: Granular computing and human-centricity in computational intelligence. In: IEEE International Conference on Cognitive Informatics (ICCI), p. 2 (2010)
7. Skowron, A., Wasilewski, P.: An Introduction to Perception Based Computing. In: Kim, T.-h., Lee, Y.-h., Kang, B.-H., Ślęzak, D. (eds.) FGIT 2010. LNCS, vol. 6485, pp. 12–25. Springer, Heidelberg (2010)
8. Hobbs, J.R.: Granularity. In: Proceedings of the 9th International Joint Conference on Artificial Intelligence, pp. 432–435 (1985)
9. Zhang, L., Zhang, B.: The Theory and Application for problem solving, 2nd edn. TsingHua University Publisher, Beijing (2007) (in Chinese)
10. Yao, Y.: Granular computing. Computer Science (Ji Suan Ji Ke Xue) 31, 1–5 (2004)
11. Yao, Y.: Perspectives of granular computing. In: Proceedings of 2005 IEEE International Conference on Granular Computing, vol. 1, pp. 85–90 (2005)
12. Yao, Y., Miao, D., Zhang, N., Xu, F.: Set-Theoretic Models of Granular Structures. In: Yu, J., Greco, S., Lingras, P., Wang, G., Skowron, A. (eds.) RSKT 2010. LNCS, vol. 6401, pp. 94–101. Springer, Heidelberg (2010)
13. Yao, Y., Miao, D.Q., Zhan, N., Xu, F.F.: Set-theoretic Approaches to Granular Computing. J. Fundamenta Informaticae 115, 247–264 (2012)
14. Feng, Q.: Research on Methods of Granular Computing based on Multidimensional Data Model, Shanghai, Tongji University (2009) (dissertation in Chinese)
15. Feng, Q., Miao, D., Cheng, Y.: Hierarchical decision rules mining. Expert Systems with Applications 37(3), 2081–2091 (2010)
16. Qian, Y.H., Liang, J.Y., Yao, Y., et al.: MGRS: A multi-granulation rough set. J. Information Sciences. 180, 949–970 (2010)
17. Yao, Y.: A Partition Model of Granular Computing. In: Peters, J.F., Skowron, A., Grzymała-Busse, J.W., Kostek, B.z., Świniarski, R.W., Szczuka, M.S. (eds.) Transactions on Rough Sets I. LNCS, vol. 3100, pp. 232–253. Springer, Heidelberg (2004)
18. Feng, Q., Miao, D.: Structured Prior Knowledge and Granular Structures. In: Zhong, N., Li, K., Lu, S., Chen, L. (eds.) BI 2009. LNCS, vol. 5819, pp. 115–126. Springer, Heidelberg (2009)
19. Pawlak, Z.: Rough Sets, Theoretical Aspects of Reasoning about Data. Kluwer Academic Publishers, Dordrecht (1991)