

Hierarchical Granular Computing Theory and its Application

Shoubai Xiao

Jiangxi University of Technology, Nanchang 330098, China

Keywords: Granular Computing; Hierarchical; Rough sets; Fuzzy sets; Quotient space

Abstract. Granular Computing is a new method of simulating human thinking and solving complex problems in the current field of computational intelligence research. It covers theories, methods and techniques of all relevant granularity, which is the powerful tool of studying complex problem solving, massive data mining and fuzzy information processing and so on. The main idea of granular computing approach is to solve problems at different levels of granularity, reflecting the intelligence in human problem solving process to a great extent. With the deepening of granular computing research work, different theoretical models of granular computing have been acquired from different angles, the major granular computing model includes theoretical model of fuzzy set, theory model of rough set and theory model of commercial space. This paper analyzes the theoretical basis of existing granular computing model and centers on granular computing theory study in the hierarchical order, and has conducted systematic analysis and research for the construction of hierarchical knowledge granularity space, the uncertainty of hierarchical knowledge granularity spatial structure, the uncertainty of rough sets under hierarchical knowledge granular space and knowledge acquisition based on hierarchical knowledge granularity and so on.

Introduction

There has been more than 20 years since the granular computing has been proposed, and it has attracted great attention from researchers in recent years. Granular computing has become a new field of research in academic world, IEEE Computational Intelligence Society established a special group of Granular Computing (Task Force on GrC) in 2004, and began to convene IEEE International Conference on Granular Computing in 2005. The First IEEE International Conference on Granular Computing has been successfully held in China in 2005 and it has been later held once a year. The First National Conference on Granular Computing (CG Shen) was held in 2007 in Taiyuan, and it has been later held once a year. The success of these meetings has promoted the sustainable development of granular computing research.

Granular computing is a new computing paradigm and concept of information processing, covering all the granularity-related theories, methods, techniques and tools, which has been applied for describing and handling uncertain, vague, incomplete and massive information and providing a problem solving method based on relationship between granule and granule. The research purpose of granular computing is to find an approximate solution rather than the best solution to simplify the problem, achieve robustness of problem solving and reduce solving costs. Its implementation method is to achieve interaction between the different levels of granularity by selecting the appropriate level of granularity, so as to solve problems that can't be solved or are difficult to be solved. In recent years, granular computing research work has attracted widespread concern in many international and

domestic researchers. It is believed that granular computing theory and methods will be gradually improved and will be applied to the relevant fields successfully through joint efforts.

On the basis of Lin' study, Yao has conducted a detailed study for granular computing in combination with neighborhood system and has published a series of research results and has applied it to knowledge mining and other areas to establish the inclusion relation between IF-THEN rules and the granularity sets, and has also proposed to solve consistent classification problems by virtue of the grid structure formed by the division, so as to provide new methods and ideas for data mining. According to rough set theory, Yao has discussed the applications of granular computing in machine learning, data analysis, data mining, rule extraction, intelligent data processing and granule logic and so on. Yao puts forward three kind of views about granular computing:

- (1) From the perspective of philosophy, granular computing is a structured thinking method;
- (2) From the perspective of application, granular computing is a common and structured method of problem solving;
- (3) From the perspective of computation, granular computing is a typical example of information processing.

The article has put forward a kind of measuring method based on the rough set ambiguity of information entropy, and it is found that ambiguity of rough sets experiences monotone decreasing with the decreasing of knowledge granularity. Uncertainty measures in rough sets include roughness of rough sets, rough entropy, ambiguity and fuzzy entropy and other methods; under the same knowledge granularity space, these measurements can better characterize the uncertainty of rough set, but in hierarchical knowledge granularity space, these measurements vary with knowledge granularity, which may not be in line with people's cognitive rules.

When the knowledge granule in positive region and negative region are finely divided particles of knowledge, or when knowledge granule in the boundary domain are divided "proportionally", rough set ambiguity proposed in the article will not change, while the ambiguity of rough sets will decrease strictly with the decreasing of ambiguity of knowledge granule in the rough sets in other cases, which is in line with people's cognitive rules. These findings make up the deficiency of uncertainty measurement research of the existing rough sets.

Rough set theory model

Polish scholar, Pawlak has put forward the rough set theory in the 1980s, which is based on the following assumptions: Human intelligence (knowledge) is a kind of classification ability. This assumption may not be perfect, but very refined. On this basis, he proposed the concept could be represented by a subset in domain of discourse, so a subset is given in domain of discourse, or a partition is given (so-called partition refers to that set X is divided into pairwise and disjoint subsets and the merge of all these subsets cluster equals to X). Given that a partition in set X is equivalent to

an equivalence relation R in X . X / R means all equivalence classes exported by R in U , $[x]_R$ represents the equivalence class of element x , where $x \in U$ is called as knowledge base(x, R) in domain of discourse by Pawlak. General concept X (X is a subset of u) can be represented by knowledge in knowledge base. As for sets that can't be represented by union sets in elements of (x, R), the concept of topological kernel and closures can be used to introduce lower and upper approximation of concepts, which are defined as:

Lower approximation set: $R(X) = \{x \in U \mid [x]_R \subset X\}$

Lower approximation set: $R(X) = \{x \in U \mid [x]_R \cap X \neq \emptyset\}$

Boundary region: $BN_R(X) = R^-(X) - R_-(X)$

Positive region: $POS_R(X) = R_-(X)$

Negative region: $NEG_R(X) = U - R^-(X)$

In the case of $R_-(X) \neq R^-(X)$, it is said X is rough. The main features of rough set theory are as follows:

- (1) there is no need to provide priori knowledge in addition to the desired dataset for solving problems;
- (2) it is an effective tool for data analysis;
- (3) can express and deal with incomplete information;
- (4) it is a soft computing method.

Concept of soft computing is put forward by the founder of fuzzy sets, Zadeh. The traditional method of calculating is the so-called hard computing, which applies with precise and fixed algorithm to express and solve problems; and the guiding principle of soft computing is to allow the use of imprecise, uncertain and partially true approximate solution to replace exact solution of the original problem and establish the solution with strong robustness and low cost, so as to better coordinate with the reality system.

Classical rough set theory is mainly aimed at a comprehensive information system, namely all attribute values for each sample of the process object is already known. But incomplete information system exists widely in daily data, the classical rough set theory can't deal with these incomplete information systems caused by the omission of information or defect directly. In order to ensure the rough set theory to apply to the processing for incomplete information system, at present, there are two main ways: the first method is to make up incomplete data, this method is difficult and inaccurate; the second is to expand the classical rough set model so that the extended rough set model can deal with incomplete information systems. Rough set theory can be expanded from at least three aspects:

- (1) generalization of equivalence relations;
- (2) the structure and extension of the knowledge representation methods based on knowledge granule;
- (3) Algebraic method of rough set.

Structural analysis for hierarchical quotient space

Granular computing model based on fuzzy sets (computing with words) theory, the granular computing model based on rough set theory and the granular computing model based on quotient space can reflect that human beings have the capacity to deal with complex problems at different levels of granularity on a certain extent. To be able to analyze and deal with the problem at different levels, a hierarchical structure of the problem space should be established at first, namely

hierarchical knowledge granularity spatial structure, which can be referred to as the hierarchical structure, and the nature of the hierarchical structure will determine the problem solving efficiency directly.

Dealing with problems in the hierarchical order is a common approach. For example, when an experienced master checks the machine's fault, he would divide the machine into several parts and then narrow failures down to one part, and then further decompose this section, and continue to repeat this process until he could find the failure. The efficiency of this method is much higher. The problems are solved in different levels and then gathered to get the solution for the original problems, which is the hierarchical problem solving process. We regard problem space as domain of discourse, the problem space abstraction process is the gradual granulation process, and we can finally get a hierarchical knowledge granularity space structure. The quality of this structure depends on the granulation method, which directly determines the efficiency of problem solving. Hierarchical structure of problem space is set up; the solution efficiency is high; on the other hand, solving efficiency will be reduced.

Quotient space theory is one of the main theoretical approaches of granular computing; fuzzy quotient space theory is a combination of fuzzy set theory and the theory of quotient space. On this basis, the arbitrary threshold corresponding to fuzzy quotient space has been studied; the recursive establishment from a hierarchical structure to fuzzy equivalence relation has been achieved through constructing normalized isosceles distance function hierarchically, and the decomposition and synthesis of the different levels of quotient space by virtue of intersection operation and union operation of fuzzy equivalence relation.

It is assumed that X should be a regular set, $X \times X$ represents the Cartesian product. It is assumed that R should be a fuzzy relation on $X \times X$, if it satisfies:

$$(1) \quad \forall x \in X, R(x, x) = 1;$$

$$(2) \quad \forall x, y \in X, R(x, y) = R(y, x)$$

$$(3) \quad \forall x, y, z \in X, R(x, z) \geq \sup_{y \in X} \min R(x, y), R(y, z)$$

R would be called the fuzzy equivalence relation on X .

It is assumed that R should be a fuzzy relation on x , R_λ would be called as the section relation of R , where R_λ is defined as the following:

$$R_\lambda = \{ (x, y) \mid R(x, y) \geq \lambda \}$$

R_λ would be a regular equivalence relation on X .

It is assumed that the equivalence relation on X is R , the corresponding quotient space is written as $X(\lambda), \{X(\lambda) \mid 0 \leq \lambda \leq 1\}$ could be called as a hierarchical structure on X .

Uncertainty measurement of hierarchical quotient space

The uncertainty of knowledge includes knowledge space uncertainty and the concept of uncertainty. Nature of hierarchical granular computing method is to granulize problem space into a

hierarchical knowledge granularity space, and then find solution on different levels of granularity in the top-to-down order (access to knowledge), then compound the solution of the original problem in bottom-up order.

Different people can acquire different fuzzy similarity relationship according to their subjective awareness and evaluation criteria, and then can achieve the fuzzy equivalence relations and quotient space hierarchical structure and give their classification. The classification result may be the same, may be different, what are the causes? How is the quality, knowledge content and uncertainty of classification result (hierarchical quotient space structure)? Whether the influence of different hierarchical quotient space to the classification is the same? These questions have not been answered well. Starting from the general fuzzy relations, fuzzy similarity relation, fuzzy equivalence relations, and the relations between hierarchical quotient space structure and information entropy sequences have been researched to further reveal the essential characteristics of the hierarchical quotient space structure.

For ease of discussion, it is necessary to provide the basic concept at first. It is assumed that X should be a regular finite set. It is assumed the fuzzy similarity relation matrix should be square matrix,

then it can be represented as $\forall m \geq n, \hat{R} = \tilde{R}^m$.

It is assumed that R should be the fuzzy equivalence relation on x , the set $\pi_x(R)$ can be defined as the following:

$$\pi_x(R) = \{X(\lambda) \mid 0 \leq \lambda \leq 1\}$$

$\pi_x(R)$ can be regarded as the corresponding hierarchical quotient space of fuzzy equivalence relation R (hierarchical structure, for example: set $X = \{x_1, x_2, x_3, x_4, x_5\}$, fuzzy equivalence relation R_t (matrix representation) in the set can be:

$$R_t = \begin{bmatrix} & 10 & 04 & 08 & 05 & 05 \\ & 04 & 10 & 04 & 04 & 04 \\ & 08 & 04 & 10 & 05 & 05 \\ 05 & 04 & 05 & 10 & 06 & \\ 05 & 04 & 05 & 06 & 10 & \end{bmatrix}$$

Corresponding quotient space sequence is as follows:

$$(X) = \left\{ \left\{ \begin{matrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{matrix} \right\} \mid 0 \leq \lambda_1 \leq 0.4 \right\}$$

$$(X) = \left\{ \left\{ \begin{matrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{matrix} \right\} \mid 0.4 \leq \lambda_1 \leq 0.5 \right\}$$

$$(X) = \left\{ \left\{ \begin{matrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{matrix} \right\} \mid 0.5 \leq \lambda_1 \leq 0.6 \right\}$$

$$(X) = \left\{ \left\{ \begin{matrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{matrix} \right\} \mid 0.6 \leq \lambda_1 \leq 0.8 \right\}$$

$$(\mathcal{X}) = \left\{ \left\{ x_1, x_2, x_3, x_4, x_5 \right\} \mid \lambda_1 \leq 1 \right\}$$

Therefore, corresponding quotient space of fuzzy equivalence relation R_t is as follows:

$$\pi_x(\mathcal{X}) = \left\{ \left(\lambda_1, X(\lambda_2), X(\lambda_3), X(\lambda_4), X(\lambda_5) \right) \right\}$$

Conclusions

Granular computing is a new concept and new method in the field of artificial intelligence, which covers all theories, methods, techniques and tools relevant to granularity, mainly used for uncertain, imprecise and incomplete information processing, large scale massive data mining and complex problem solving. The essential characteristics of granular computing are to find good approximate solutions by the choice of appropriate granularity. Fuzzy set theory, rough set theory and the theory of quotient space are the three main basic theories of granular computing. Structured definition of fuzzy sets, topology of knowledge granularity space and quotient space attributed in rough set theory can reflect the core idea of structural description, such structure is a kind of hierarchical structure, which plays a key role in granularity in problem space, and its complexity and structural efficiency can directly determine the problem solving and solution accuracy. On the basis of the existing granule computing theory research, the granular computing model has been abstracted to hierarchical granular computing model in the article, and a series of studies have been conducted for such granule computing model.

To sum up, the main research contents and results of this paper includes the following two aspects:

(1). Although the fuzzy equivalence relations can be established in the quotient space on the premise that threshold value is 1 through normalized isosceles distance function, however, as for the quotient space of any threshold value, how to build a normalized isosceles distance function to establish fuzzy equivalence relation is a very important issue. On the basis of the fuzzy quotient space theory research, the article has proposed approaches to establish normalized isosceles distance function in order to achieve the mutual transformation between hierarchical quotient space structure and fuzzy equivalence relation and get the correspondence between the intersection operation and union operation of fuzzy equivalence relations and the decomposition and synthesis of quotient space. These results further have improved fuzzy quotient space granular computing theory.

(2). Uncertainty of problems includes uncertainty of knowledge space and uncertainty reflected by concepts in the knowledge space. The article has put forward the uncertainty measurement methods for hierarchical quotient space on the basis of hierarchical knowledge granularity space, which has provided a new way of thinking for the uncertainty of measurement granule computing model. The article has proposed to use the information sequence to measure the hierarchical sequence quotient space uncertainty, it is found that the information sequence of hierarchical knowledge quotient space will increase strictly with the decreasing of knowledge granule; the article has also proposed the definition of hierarchical quotient space and proved the judging theorem of hierarchical quotient space, namely, the same division sequence and the same information sequence are the sufficient and necessary conditions for isomorphic hierarchical quotient space; and the article has also illustrated the information sequence can better characterize the uncertainty of hierarchical quotient space by analyzing the inherent relation between fuzzy similarity relation, fuzzy equivalence internal

association and information sequence. These research results have laid a good foundation for the uncertainty of the measurement hierarchical quotient space.

Acknowledgement

The article has been funded by Jiangxi Science and Technology Planning Project: systematic study from Jiangxi intellectual property transformation and trade center and Jiangxi University of Technology [No. ZR13ZD03].

References

- [1] Zhang Qian, Zhang Ling, Problem Solving Theory and Application (M), Beijing: Tsinghua University Press, 1990.
- [2] Hobbs J R. Granularity [C]. Proceedings of the 9th International Joint Conference on Artificial Intelligence, 1985:432—435.
- [3] Lin T Y Neighborhood Systems and Relational Database[C]. In Proceedings of CSC'88, 1988:725.
- [4] Lin T Y Granular Computing on Binary Relations I: Data Mining and Neighborhood Systems, 11: Rough Set Representations and Belief Functions [M]. Rough Sets in Knowledge Discovery Skowron A and Polkowski L (Eds.), Physica—Verlag, 1998:107—140.
- [5] Lin T Y Data Mining: Granular Computing Approach Methodologies for Knowledge Discovery and Data Mining [C]. Proceedings of PAKDD'99, LNAI1574, 1999:24—33.
- [6] Lin T Y Granular Computing :Fuzzy Logic and Rough Sets[M]. In Zadeh and KaePrzyk editors, Computing with Words in Information Intelligent Systems, Physica-Verlag, 1999:183—200.
- [7] Lin T Y Data Mining and Machine Oriented Modeling: A Granular Computing Approach[J]. Journal of Applied Intelligence, 2000, 13(2):113—124.
- [8] Lin T Y Granular Computing :Structures, Representations, Applications and Future Directions[C]. Proceedings of 9th International Conference, RSFDGrC2003, Lecture Notes on Artificial Intelligence LNAI2639, Springer-Verlag, 2003:16—24.
- [9] Lin T Y Granular Computing: A Problem Solving Paradigm[C]. The Proceedings of the 2005 IEEE International Conference on Fuzzy Systems, 2005:132—137.
- [10] Lin T Y, Zadeh L A. Foreword to the Special Section on Granular Computing [J]. IEEE Transactions on Fuzzy Systems 2008, 16(2):282—284.