

An Improved Attribute Reduction Algorithm based on Granular Computing

X. Tang, L. Shu

Xiao Tang*

1. University of Electronic Science and Technology of China,
School of Mathematical Sciences,
Chengdu, 611731, China

2. College of Mathematics and Software Science,
Sichuan Normal University,
Chengdu, 610066, China

*Corresponding author: 80651177@163.com

Lan Shu

University of Electronic Science and Technology of China,
School of Mathematical Sciences,
Chengdu, 611731, China
lanshu@uestc.edu.cn

Abstract: Granular computing is a new intelligent computing method based on problem solving, information processing and pattern classification. Granular computing based attribute reduction method is an important application of Granular computing. These algorithms are mostly based on reduction core. However, some information systems may have no reduction core, especially in the actual application data. For this case, those algorithms are powerless. In this paper, an improved reduction algorithm based on granular computing is proposed. The algorithm is validated by the experimental result.

Keywords: attribute reduction, granular computing, rough set, attribute significance.

1 Introduction

Granular computing is a method for analysis of multi-layer granular structure based on problem solving, pattern classification and information processing. It's also a newly cross discipline among rough set theory, fuzzy set theory, data mining and artificial intelligence. With less than 20 years' development, granular computing has already made remarkable achievements and great contribution to the field of computer science [1, 2]. Through rapid development of society and continuous progress in science and technology, a variety of data is increasing gradually, and then we entered the so-called "Big Data Time". The main goal of data mining is to find potential, desired and useful knowledge from those big data. Rough set theory is an efficient mathematical tool to deal with imprecise, incomplete and inconsistent data. It has already made great strides in its theory and has been widely used in practical application.

Attribute reduction is the main content of rough set theory. The core task of attribute reduction is that dimensionality and storage space may be reduced under the condition of maintaining classification capacity, so as to improve the efficiency of system classification [3, 4]. Therefore, it is not only the hot spot of intelligence computing, but also the important task of information processing.

In 1979, professor L.A. Zadeh discussed the theory of fuzzy information granulation in his paper "Fuzzy Sets and Information Granularity", and first proposed the concept of information granulation. Then, professor J.R. Hobbs of Stanford University introduced granularity theory

in his paper "Granularity" published on International Joint Conference on Artificial Intelligence held in Los Angeles [5]. The granularity theory is presented firstly. The idea of granularity theory is that the bigger, whole, unresolved questions can be broken into several smaller ones by granulating, and these small questions can be combined into the bigger, whole questions. In 1990, the Chinese scholar Zhang Bo and Zhang Ling proposed the theory of quotient space based on problem solving [6]. They thought that human beings can analyze the same problem from different granulation, and make an easy conversion in different knowledge granularity. If people can formalize the analysis process to make the computer possess the ability, it will greatly improve the development of artificial intelligence. Furthermore, professor L.A. Zadeh raised the theory of Computing with Words in his paper "Fuzzy Logic=Computing with Words", and thus the fuzzy granularity theory was born [7]. This theory is to do fuzzy reasoning and judgments by using natural language, so as to realize the fuzzy intelligent control. In the same year, when Professor L.Y. Lin visited in Professor Zadeh's Key Laboratory of UC-Berkeley University, he presented the subject "Granular Computing" and got approval from Zadeh, it marked the birth of granular computing.

Professor Miao, et al. [8] gave the definition of knowledge granularity and knowledge discernibility in fuzzy set theory model, and pointed out the relationship between knowledge granularity and knowledge discernibility: the smaller knowledge granularity is, the stronger knowledge distinguishable ability is; on the contrary, the bigger knowledge granularity is, the weaker knowledge distinguishable ability is. Reference [9] defined the concept of the difference of granularity and granularity entropy on the basis of fuzzy set's algebraic method and information theory approach, and proposes attribute reduction algorithm based on granular computing. Reference [10] presented attribute reduction algorithm based on Granular Computing, using the equivalent relation in rough set to construct granule, and attribute significance is regarded as heuristic information. Reference [11] put forward attribute reduction method based on model of granular computing in information systems. Reference [12] proposed an improvement of attribute reduction algorithm based on Granular Computing. This algorithm is to get attribute core using discernibility matrix, and then make attribute reduction based on attribute significance as heuristic information. Reference [13] proposed an incomplete order decision table reduction algorithm based on granular computing.

These reduction methods based on granular computing are mainly first to calculate reduction core of system, then get reduction based on core. However, in practical application, some information systems may have no reduction core. In this case, this paper proposes an improved reduction algorithm based on attribute significance of granular computing, and numerical experiments show the effectiveness of the algorithm.

2 Basic Concepts of Rough Sets

2.1 Rough sets

Let a quadruple $S = (U, A, V, f)$ be an information systems (IS), in which $U = \{x_1, x_2, \dots, x_n\}$ is a non-empty finite set called the domain of discourse; $A = \{a_1, a_2, \dots, a_m\}$ is a non-empty and finite set of attributes; V is a set of attribute values domain, $V = \bigcup_{a \in A} V_a$; $f : U \times A \rightarrow V$ is a mapping, each attribute of the object in the domain of discourse by the mapping has a corresponding information value, i.e. $\forall a \in A, x \in U, f(x, a) \in V_a$. If the attributes set A is composed of condition attributes set C and decision attributes set D , the quadruple $S = (U, A, V, f)$ is also called decision information system (*DIS*). The information system, also known as knowledge representation system, is the main expression of knowledge of rough sets. It is simply expressed in (U, A) . If P is a subset of attributes set A , each subset $P \subseteq A$ determines a binary indistin-

guishable relation $IND(P)$, $IND(P) = \{(u, v) \in U \times U \mid \forall a \in P, a(u) = a(v)\}$. A set $X \subseteq U$ represent a concept and partition included by $IND(P)$ is called a knowledge base and denoted by $U/IND(P)$. In particular, the partition $U/IND(P) = \{Y_1, Y_2, \dots, Y_k\}$ is the knowledge base of decision classes.

A knowledge base (U, R) is also called an approximation space, where U is the domain of discourse and R is an equivalence relation on U . Let $X \subseteq U$ and $R \subseteq A$, the sets $\underline{R}(X) = \{x \in U \mid [x]_R \subseteq X\}$ and $\bar{R}(X) = \{x \in U \mid [x]_R \cap X \neq \emptyset\}$ are respectively called lower approximation set and upper approximation set. Where $[x]_R$ refers to an equivalence class of $IND(P) = \cap IND(R)$ determined by element x . If $\underline{R}(x) = \bar{R}(x)$, then X is called a definable set on U ; If $\underline{R}(x) \neq \bar{R}(x)$, then X is called a rough set on U .

If $IND(R) \neq IND(R - |a|)$, then a is indispensable in the set R , otherwise a is dispensable. If every $a \in R$ is indispensable, then R is called independent. Let $Q \subseteq P$, that is, Q is a subset of P , if Q is independent and $IND(Q) = IND(P)$, then Q is a reduction of P , denoted as $Q = red(P)$. The union set of indispensable attribute in the set A is called a core set, denoted as $core(P)$, $core(P) = \cap red(P)$.

2.2 Knowledge granulation and partition

Definition 1[8]. Let (U, R) be an approximation space, $P \in R$ is an equivalence relation on U , called knowledge. The approximation space is also called knowledge base. The equivalence class $[x]_P = \{x \in U, (x_i, x_j) \in P\}$ is called knowledge granule. The quotient set $U/P = \{[x]_P \mid x \in U\}$ is called a P -granularity partition. The granularity of knowledge is defined as $GD(P)$,

$$GD(P) = \frac{|P|}{|U \times U|} = \frac{|P|}{|U|^2} \quad (1)$$

Where $|P|$ denotes the cardinality of the set $P \subseteq U \times U$.

The granularity of knowledge P can express its distinguishable ability. For $\forall u, v \in U$, if $(u, v) \in P$, then they belong to the same equivalence class, i.e. they are indistinguishable. The knowledge P 's discernibility could be defined as $Dis(P)$, $Dis(P) = 1 - GD(P)$. In general, the greater the granularity is, the weaker the distinguishable ability will be, vice versa.

Theorem 1[8]. Let $P \in R$ be a knowledge of knowledge base $K = (U, R)$, if $U/P = \{X_1, X_2, \dots, X_n\}$, then

$$GD(P) = \left(\sum_{i=1}^n |X_i|^2 \right) / |U|^2. \quad (2)$$

Property 1. Let $P, Q \in R$ be an equivalence relations on U , $U/P = \{X_1, X_2, \dots, X_n\}$, $U/Q = \{Y_1, Y_2, \dots, Y_n\}$, if $P = Q$, then $GD(P) = GD(Q)$ and $Dis(P) = Dis(Q)$; if $P \prec Q$, $GD(P) < GD(Q)$ and $Dis(Q) < Dis(P)$.

proof. (a) If $P = Q$, then $m = n$, $X_i = Y_i$, so $GD(P) = GD(Q)$, $Dis(Q) = Dis(P)$.

(b) If $P \prec Q$, then $|P| < |Q|$, so $GD(P) < GD(Q)$. Since $Dis(P) = 1 - GD(P)$, we could observe that $Dis(Q) < Dis(P)$.

Property 2. Let $P \in R$ be an equivalence relation on U , $U/P = \{X_1, X_2, \dots, X_n\}$, if the equivalence relation P divides from knowledge granules in U/R , then $GD(P) \leq GD(R)$, $Dis(R) \leq Dis(P)$.

Proof. We suppose that the knowledge granule X_i from U/R is divided into two knowledge granules X_{i1} and X_{i2} , that is $X_i = X_{i1} \cup X_{i2}$ and $X_{i1} \cap X_{i2} = \emptyset$, $U/P = \{X_1, X_2, \dots, X_{i-1}, X_{i1}, X_{i2}, X_{i+1}, \dots, X_n\}$, so

$$GD(R) = \left(\sum_{j=1}^n |X_j|^2 \right) / |U|^2$$

$$\begin{aligned}
 &= \left(\sum_{j=1}^{i-1} |X_j|^2 \right) / |U|^2 + |X_i|^2 / |U|^2 + \left(\sum_{j=i+1}^n |X_j|^2 \right) / |U|^2 \\
 &= \left(\sum_{j=1}^{i-1} |X_j|^2 \right) / |U|^2 + [|X_{i1}| + |X_{i2}|]^2 / |U|^2 + \left(\sum_{j=i+1}^n |X_j|^2 \right) / |U|^2 \\
 &\geq \left(\sum_{j=1}^{i-1} |X_j|^2 \right) / |U|^2 + [|X_{i1}|^2 + |X_{i2}|^2] / |U|^2 + \left(\sum_{j=i+1}^n |X_j|^2 \right) / |U|^2 \\
 &= GD(P),
 \end{aligned}$$

$$Dis(R) = 1 - GD(R) \leq 1 - GD(P) = Dis(P).$$

Property 3. Let (U, R) be a knowledge base and $P \in R$ be an equivalence relation on U , $U/R = \{X_1, X_2, \dots, X_n\}$, Q is the union of knowledge granules in U/R , then $GD(R) \leq GD(Q)$, $Dis(Q) \leq Dis(R)$.

Proof. We suppose that the knowledge granule X_k is the union of X_i and X_{i+1} , then $U/Q = \{X_1, X_2, \dots, X_{i-1}, X_k, X_{i+2}, \dots, X_n\}$, so

$$\begin{aligned}
 GD(R) &= \left(\sum_{j=1}^n |X_j|^2 \right) / |U|^2 \\
 &= \left(\sum_{j=1}^{i-1} |X_j|^2 \right) / |U|^2 + |X_i|^2 / |U|^2 + |X_{i+1}|^2 / |U|^2 + \left(\sum_{j=i+2}^n |X_j|^2 \right) / |U|^2 \\
 &\leq \left(\sum_{j=1}^{i-1} |X_j|^2 \right) / |U|^2 + [|X_i| + |X_{i+1}|]^2 / |U|^2 + \left(\sum_{j=i+2}^n |X_j|^2 \right) / |U|^2 \\
 &= \left(\sum_{j=1}^{i-1} |X_j|^2 \right) / |U|^2 + |X_k|^2 / |U|^2 + \left(\sum_{j=i+2}^n |X_j|^2 \right) / |U|^2 \\
 &= GD(Q),
 \end{aligned}$$

$$Dis(Q) \leq Dis(R).$$

Property 4. Let $S = (U, A, V, f)$ be an information system, $P, Q \subseteq A$,

- (1) If $P \Rightarrow Q$, then $GD(P) \leq GD(Q)$;
- (2) If $P \Leftrightarrow Q$, then $GD(P) = GD(Q)$.

Proof. (1) If $P \Rightarrow Q$, then $IND(P) \subseteq IND(Q)$, this is $|IND(P)| \leq |IND(Q)|$. On the other hand, $GD(P) = GD(IND(P)) = |IND(P)| / |U|^2$, and $GD(Q) = GD(IND(Q)) = |IND(Q)| / |U|^2$, so $GD(P) \leq GD(Q)$.

(2) If $P \Leftrightarrow Q$, then $P \Rightarrow Q$ and $Q \Rightarrow P$. By (1), we could see that $GD(P) \leq GD(Q)$ and $GD(Q) \leq GD(P)$, so $GD(P) = GD(Q)$.

Property 5. Let $S = (U, A, V, f)$ be an information system, $P, Q \subseteq A$,

- (1) If $P \Rightarrow Q$, then $Dis(P) \geq Dis(Q)$;
- (2) If $P \Leftrightarrow Q$, then $Dis(P) = Dis(Q)$.

Proof. It follows immediately from Definition 1 and Property 4.

Deduction 1. Let $S = (U, A, V, f)$ be an information system, if $P \subseteq Q \subseteq A$, then $GD(Q) \leq GD(P)$ and $Dis(Q) \geq Dis(P)$.

Remark. Deduction 1 illustrates that for the subset of A , when the attribute number increased, the knowledge granularity is reduced, thus, the discernibility is increased.

3 Attribute reduction algorithm based on attribute significance

Definition 2[8]. Let $S = (U, A, V, f)$ be an information system, the attribute significance

could be defined as $Sig_{A-\{a\}}(a)$,

$$Sig_{A-\{a\}}(a) = GD(A - \{a\}) - GD(A). \quad (3)$$

Remark. In an information system $S = (U, A, V, f)$, the attribute significance of each attribute $a \in A$ could be measured by knowledge granularity.

Definition 3[8]. Let $S = (U, A, V, f)$ be an information system, C is a subset of A , $C \subseteq A$, for $\forall a \in A - C$, the attribute significance of attribute a relative to attribute set C could be defined as $Sig_C(a)$,

$$Sig_C(a) = GD(C) - GD(C \cup \{a\}). \quad (4)$$

Remark. Definition 3 illustrates that the attribute significance of attribute a relative to attribute set C could be measured by change of the knowledge granularity. When a attribute is added to attribute set C , C 's knowledge granularity may change. If C 's knowledge granularity change, then attribute a is indispensable.

Definition 4[8]. Let $S = (U, A, V, f)$ be an information system, $a \in A$, if $GD(A - \{a\}) = GD(A)$, then attribute a is dispensable, otherwise, attribute a is indispensable. If every $a \in A$ is indispensable, then A is called independent.

Definition 5[8]. Let $S = (U, A, V, f)$ be an information system, $P \subseteq A$, if P is independent and $GD(P) = GD(A)$, then P is a reduction of A , denoted as $red(A)$. The union set of indispensable attribute in the set A is called a core set, denoted as $core(P)$, $core(P) = \cap red(P)$.

Property 6. Attribute a is indispensable, if and only if

$$Sig_{A-\{a\}}(a) > 0. \quad (5)$$

Proof. \Rightarrow If attribute a is indispensable, then $GD(A - \{a\}) \neq GD(A)$. As we know that $GD(A - \{a\}) \geq GD(A)$, so $Sig_{A-\{a\}}(a) = GD(A - \{a\}) - GD(A) > 0$.

\Leftarrow Obviously.

Property 7. $Core(A) = \cup \{a \in A \mid Sig_{A-\{a\}}(a) > 0\}$.

Proof: It follows immediately from Definition 3 and Property 6.

Remark. The attribute significance from the perspective of knowledge granularity provides a method of attribute reduction: We could judge the significance of attribute a by discussing whether $GD(A - \{a\})$ is equal to $GD(A)$. If $GD(A - \{a\}) = GD(A)$ then a is dispensable, otherwise a is indispensable. Thus we could obtain the reduction core $Core(A)$. Next calculate the significance of the rest attribute relative to $Core(A)$. If $GD(Core(A) \cup a) = GD(A)$, then the set $Core(A) \cup a$ is the reduction of the information system, where $a = \{a \in A - Core(A) \mid \max Sig_{Core(A)}(a)\}$.

Algorithm 1:

Input: An information system $S = (U, A, V, f)$, where $U = \{x_1, x_2, \dots, x_n\}$, $A = \{a_1, a_2, \dots, a_m\}$.

Output: $red(A)$ and $Core(A)$ // the sets of reductions and core.

Step 1: For $i = 1, i \leq n, ++i; j = 1, j \leq m, ++j$ begin, calculate $GD(A)$ // the knowledge granularity of attribute set A .

Step 2: Calculate $Sig_{A-\{a\}}(a)$ // the significance of attribute $a \in A$.

Step 3: Calculate $Core(A)$, $Core(A) = \{a \in A \mid Sig_{A-\{a\}}(a) > 0\}$.

Step 4: If $GD(Core(A)) = GD(A)$, output $red(A) = Core(A)$, end.

If $GD(Core(A)) > GD(A)$, turn next.

Step 5: Calculate $\max_{b \in B} Sig_{Core(A)}(b)$, $B = A - Core(A)$ // the significance of attribute $b \in B = A - Core(A)$ for $Core(A)$.

Step 6: If $GD(Core(A) \cup a) = GD(A)$, output $red(A) = Core(A) \cup b$ and $Core(A)$, end.

If $GD(Core(A) \cup b) > GD(A)$, repeat step 5,

Step 7: Calculate $\max_{c \in C} Sig_{Core(A) \cup b}(b) \cup (Core(A) \cup b)$ // the significance of attribute $c \in$

$C = A - Core(A) \cup b$ for $Core(A) \cup b$,

Step 8: For $j = 1$ to $|C|$, repeat step 7 until the knowledge granularity is equal to $GD(A)$

Step 9: Output $red(A) = Core(A) \cup b \cup c \cup \dots$ and $Core(A)$.

Example 1. Let $S = (U, A, V, f)$ be an information system (Table 1). There are 6 objects and 4 attributes, where $U = \{x_1, x_2, x_3, x_4, x_5\}$, $A = \{a_1, a_2, a_3, a_4\}$. Calculate the reduction of the system.

Table1. An information system

U	a_1	a_2	a_3	a_4
x_1	1	0	2	2
x_2	0	1	1	1
x_3	2	0	0	1
x_4	1	1	0	2
x_5	2	2	0	0
x_6	2	1	1	1

It is easy to calculate that

$$U/A = \{x_1, x_2, x_3, x_4, x_5, x_6\}, GD(A) = \left(\sum_{i=1}^n |X_i|^2 \right) / |U|^2 = \frac{6}{36} = \frac{1}{6},$$

$$U/A - \{a_1\} = \{x_1, \{x_2, x_6\}, x_3, x_4, x_5\}, GD(A - \{a_1\}) = \frac{8}{36},$$

$$Sig_{A-\{a_1\}}(a_1) = GD(A - \{a_1\}) - GD(A) = \frac{8}{36} - \frac{6}{36} = \frac{2}{36},$$

$$U/A - \{a_2\} = \{x_1, x_2, x_3, x_4, x_5, x_6\}, GD(A - \{a_2\}) = \frac{6}{36},$$

$$Sig_{A-\{a_2\}}(a_2) = GD(A - \{a_2\}) - GD(A) = \frac{6}{36} - \frac{6}{36} = 0,$$

$$U/A - \{a_3\} = \{x_1, x_2, x_3, x_4, x_5, x_6\}, GD(A - \{a_3\}) = \frac{6}{36},$$

$$Sig_{A-\{a_3\}}(a_3) = GD(A - \{a_3\}) - GD(A) = \frac{6}{36} - \frac{6}{36} = 0,$$

$$U/A - \{a_4\} = \{x_1, x_2, x_3, x_4, x_5, x_6\}, GD(A - \{a_4\}) = \frac{6}{36},$$

$$Sig_{A-\{a_4\}}(a_4) = GD(A - \{a_4\}) - GD(A) = \frac{6}{36} - \frac{6}{36} = 0.$$

Based on the result above, we can see that the core of the system is

$$Core(A) = \{a \in A \mid Sig_{A-\{a\}}(a) > 0\} = \{a_1\}, GD(Core(A)) = \frac{14}{36} > \frac{6}{36} = GD(A).$$

Let a set $B = A - Core(A) = \{a_2, a_3, a_4\}$, calculate the significance of the rest attribute relative to $Core(A)$:

$$Sig_{Core(A)}(a_2) = GD(Core(A)) - GD(Core(A) \cup a_2) = \frac{14}{36} - \frac{6}{36} = \frac{8}{36},$$

$$Sig_{Core(A)}(a_3) = GD(Core(A)) - GD(Core(A) \cup a_3) = \frac{14}{36} - \frac{8}{36} = \frac{6}{36},$$

$$Sig_{Core(A)}(a_4) = GD(Core(A)) - GD(Core(A) \cup a_4) = \frac{14}{36} - \frac{10}{36} = \frac{4}{36}.$$

Since $\max_{b \in B} Sig_{Core(A)}(b) = \frac{8}{36}$, $U/Core(A) \cup a_2 = \{x_1, x_2, x_3, x_4, x_5, x_6\}$, $GD(Core(A) \cup a_2) = GD(A)$, we can see that $Core(A) \cup a_2 = \{a_1, a_2\}$ is the reduction, that is $red(A) = \{a_1, a_2\}$, $Core(A) = \{a_1\}$.

4 An improved reduction algorithm

The reduction of an information system is not the only, some may have more than one reductions. But the reduction results may not be able to get reduction core, especially in the actual application data. For this case, Algorithm 1 is powerless. Now we will improve the algorithm, which can deal with the system with reduction core and no reduction core.

In an information system $S = (U, A, V, f)$, if $Sig_{A-\{a\}}(a) = 0$, then a is dispensable. For $\forall a \in A$, if $Sig_{A-\{a\}}(a) = 0$, then the system has no reduction core.

Definition 6. Let $S = (U, A, V, f)$ be an information system, P is a subset of A , $P \subseteq A$, the attribute significance of P relative to A could be defined as $Sig_{A-P}(P)$,

$$Sig_{A-\{P\}}(P) = GD(A - P) - GD(A). \quad (6)$$

Particularly, if $P = A$, $Sig_{\emptyset}(P)$ is represented as $Sig(P)$, and $Sig(P) = Sig_{\emptyset}(P) = GD(\emptyset) - GD(A) = 1 - GD(P) = Dis(P)$. Where $GD(\emptyset) = 1$ (since $U/IND(\emptyset) = \{U\}$).

Algorithm 2:

Input: An information system $S = (U, A, V, f)$, where $U = \{x_1, x_2, \dots, x_n\}$, $A = \{a_1, a_2, \dots, a_m\}$.

Output: $red(A)$ and $Core(A)$ // the sets of reductions and core.

Step 1: For $i = 1, i \leq n, ++i; j = 1, j \leq m, ++j$ begin, calculate $GD(A)$. // the knowledge granularity of attribute set A

Step 2: Calculate $Sig_{A-\{a\}}(a)$. // the significance of attribute $a \in A$

If $Sig_{A-\{a\}}(a) \neq 0$ // the system has reduction core, turn step 8;

If $Sig_{A-\{a\}}(a) = 0$ // the system has no reduction core, next;

Step 3: Calculate $Sig_{A-\{a_i, a_j\}}(a_i, a_j)$, $1 \leq i \neq j \leq m$ // the significance of the combination of any two attributes in A ,

Step 4: For $i = 1, i \leq n, ++i; j = 1, j \leq m, ++j$, find out

$$red(A) = \left\{ (a_i, a_j) \mid \max_{1 \leq i \neq j \leq m} Sig_{A-\{a_i, a_j\}}(a_i, a_j) \right\}. // \text{the suboptimal reduction of the system}$$

Step 5: If $GD(red(A)) = GD(A)$, then output the reduction $red(A) = red(A)$, end. If $GD(red(A)) > GD(A)$, turn next;

Step 6: Calculate $\max_{b \in B} Sig_{red(A)}(b)$, $B = A - red(A)$. // the significance of attribute $b \in B = A - red(A)$ for $red(A)$

Step 7: If $GD(red(A) \cup b) = GD(A)$, then output $red(A) = red(A) \cup b$, end.

If $GD(red(A) \cup b) > GD(A)$, repeat step 6 until $GD(red(A) \cup b) = GD(A)$, output $red(A)$, end.

Step 8: Calculate $Core(A) = \{a \in A \mid Sig_{A-\{a\}}(a) > 0\}$.

Step 9: If $GD(Core(A)) = GD(A)$, output $red(A) = Core(A)$, end.

If $GD(Core(A)) > GD(A)$, turn next;

Step10: Calculate $\max_{b \in B} Sig_{Core(A)}(b)$, $B = A - Core(A)$ // the significance of attribute $b \in B = A - Core(A)$ for $Core(A)$

Step11: If $GD(Core(A) \cup b) = GD(A)$, output $red(A) = Core(A) \cup b$ and $Core(A)$, end.

If $GD(Core(A) \cup b) > GD(A)$, repeat step 10.

Step12: Calculate $\max_{c \in C} Sig_{Core(A) \cup b}(b) \cup (Core(A) \cup b)$ // the significance of attribute $c \in C = A - Core(A) \cup b$ for $Core(A) \cup b$.

Step13: Output $red(A) = Core(A) \cup b \cup c \cup \dots$ and $Core(A)$, end.

Example 2. Let $S = (U, A, V, f)$ be an information system (Table 2). There are 5 objects and 4 attributes, where $U = \{x_1, x_2, x_3, x_4, x_5\}$, $A = \{a_1, a_2, a_3, a_4\}$. Calculate the reduction of the system.

Table2. An information system

U	a_1	a_2	a_3	a_4
x_1	1	0	2	1
x_2	2	1	2	1
x_3	0	1	2	0
x_4	2	0	1	1
x_5	1	2	2	0

It is easy to calculate that

$$U/A = \{x_1, x_2, x_3, x_4, x_5\}, GD(A) = \left(\sum_{i=1}^n |X_i|^2 \right) / |U|^2 = \frac{5}{25} = \frac{1}{5},$$

$$U/A - \{a_1\} = \{x_1, x_2, x_3, x_4, x_5\}, GD(A - \{a_1\}) = \frac{1}{5},$$

$$Sig_{A-\{a_1\}}(a_1) = GD(A - \{a_1\}) - GD(A) = 0;$$

$$U/A - \{a_2\} = \{x_1, x_2, x_3, x_4, x_5\}, GD(A - \{a_2\}) = \frac{1}{5},$$

$$Sig_{A-\{a_2\}}(a_2) = GD(A - \{a_2\}) - GD(A) = 0;$$

$$U/A - \{a_3\} = \{x_1, x_2, x_3, x_4, x_5\}, GD(A - \{a_3\}) = \frac{1}{5},$$

$$Sig_{A-\{a_3\}}(a_3) = GD(A - \{a_3\}) - GD(A) = 0;$$

$$U/A - \{a_4\} = \{x_1, x_2, x_3, x_4, x_5\}, GD(A - \{a_4\}) = \frac{1}{5},$$

$$Sig_{A-\{a_4\}}(a_4) = GD(A - \{a_4\}) - GD(A) = 0.$$

Based on the result above, we can see that the system has no core. Next we calculate the significance of attribute combination $Sig_{A-\{a_i, a_j\}}(a_i, a_j)$, $1 \leq i \neq j \leq 4$:

$$GD(A - \{a_1, a_2\}) = \frac{9}{25}, Sig_{A-\{a_1, a_2\}}(a_1, a_2) = GD(A - \{a_1, a_2\}) - GD(A) = \frac{4}{25};$$

$$GD(A - \{a_1, a_3\}) = \frac{7}{25}, Sig_{A-\{a_1, a_3\}}(a_1, a_3) = GD(A - \{a_1, a_3\}) - GD(A) = \frac{2}{25};$$

$$GD(A - \{a_1, a_4\}) = \frac{7}{25}, Sig_{A-\{a_1, a_4\}}(a_1, a_4) = GD(A - \{a_1, a_4\}) - GD(A) = \frac{2}{25};$$

$$GD(A - \{a_2, a_3\}) = \frac{5}{25}, Sig_{A-\{a_2, a_3\}}(a_2, a_3) = GD(A - \{a_2, a_3\}) - GD(A) = 0;$$

$$GD(A - \{a_2, a_4\}) = \frac{7}{25}, Sig_{A-\{a_2, a_4\}}(a_2, a_4) = GD(A - \{a_2, a_4\}) - GD(A) = \frac{2}{25};$$

$$GD(A - \{a_3, a_4\}) = \frac{5}{25}, Sig_{A-\{a_3, a_4\}}(a_3, a_4) = GD(A - \{a_3, a_4\}) - GD(A) = 0.$$

We can see that $GD(a_1, a_2) = GD(A)$, so $red(A) = \{a_1, a_2\}$ is the reduction of the system.

5 Results and Discussion

The idea of Algorithm 1 is listed as follow: At first, work out the reduction core $Core(A)$ by finding out the set $\{a \in A | Sig_{A-\{a\}}(a) > 0\}$. Then calculate the significance of the rest attribute relative to $Core(A)$. If $GD(Core(A) \cup a) = GD(A)$, then the set $Core(A) \cup a$ is the reduction of the information system. The time complexity of the algorithm is $T \approx O(|C|^3 \cdot |U|^2)$. The precondition of Algorithm 1 is working out reduction core. However, some information systems may have no reduction core. In this case, Algorithm 1 is powerless. In this paper, an improved algorithm is proposed. In an information system with no reduction core, the suboptimal reduction $red(A)$ replaces the reduction core. The improved algorithm does not increase in time complexity. It can deal with the system with reduction core and no reduction core.

Acknowledgment

We would like to thank all the referees for their valuable comments in improving this paper. This work is supported by The Foundation of Sichuan Normal University (Grant No. 13KYL15).

Bibliography

- [1] Zhang W.X.; Wu W.Z.; Liang J.Y.(2001); The Theory and Method of Rough set, *Beijing: Science Press, China*.
- [2] Zeng H.L.(2004); Intelligent Computing, *Chongqing: Chongqing University Press, China*, 1-89.
- [3] Miao D.Q.; Wang G.Y.; Liu Q.; Lin T.Y.; Yao Y.Y.(2007); Granular Computing: Past, Present and Future Prospects, *Beijing: Science Press, China*.
- [4] Wang G.Y.; Li D.Y.; Yao Y.Y., et al(2012); Cloud Model and Granular Computing, *Beijing: Science Press, China*.
- [5] Hobbs J.R.(1985); Granularity, *In Proceedings of the 9th international Joint Conference on Artificial Intelligence. (IJCAI) Los Angeles*, 432-435.
- [6] Zhang B.; Zhang L.(1990); Theory and Application of Problem Solving. *Beijing : Tsinghua University Press ,China*.
- [7] Zadeh L.A.(1996); Fuzzy Logic=Computing with Words. *IEEE Transactions on Fuzzy Systems*, 4(2): 103-111.
- [8] Miao D.Q.; Fan S.D.(2002); The calculation of knowledge granulation and its application. *Systems Engineering- Theory & Practice*, 22(1): 48-56.
- [9] Yan L.H.; Han X.(2008); Attribute reduction based on granular computing. *Computer Applications and Software*, 25(4):239-240.
- [10] Zhao M.; Luo K.; Qin Z.(2008); Attribute reduction algorithm based on granular computing. *Computer Engineering and Applications*, 44(30): 157-159.
- [11] Feng L.; Liu Z.P.; Fang D.(2010); An approach on attribute reduction based on model of granular computing in information systems. *Journal of Chongqing University of Posts and Telecommunications*, 22(5):652-655.
- [12] Wang H.X.; Cheng Y.H.(2010); Application and improvement of attribute reduction arithmetic based on granular computing. *Microcomputer Information*, 26(5): 33-35.
- [13] Shi J.L.; Du G.Y.; Xiong D.L.(2012); An incomplete order decision table reduction algorithm based on granular computing. *Computer Applications and Software*, 29(10): 113-116.