Application of Granular Computing in Knowledge Reduction

Lai Wei[1] and Duoqian Miao[2]

[1] Department of Computer Science and Technology, Tongji University
Shanghai, 200092, P.R. China
weily105@gmail.com

[2] Department of Computer Science and Technology, Tongji University
Shanghai, 200092, P.R. China
miaoduoqian@163.com

Abstract. Skowron's discernibility matrix is one of representative approaches in computing relative core and relative reducts, while redundant information is also involved. To decrease the complexity of computation, the idea of granular computing is applied to lower the rank of discernibility matrix. In addition, the absorptivity based on bit-vector computation is proposed to simplify computation of relative core and relative reducts.

Keywords: Rough set, discernibility matrix, granular computing, absorptivity.

1 Introduction

Reduction of Knowledge[1], one of crucial parts in rough set theory[5], plays a very important role in the fields of knowledge discovery[4], decision analysis[6], clustering analysis[2] and so on. The knowledge having been simplified can decrease the complexity of computing and improve the adaptability of knowledge in certain extent.

Information Granulation[3] is helpful to problem solving. Observing things on different levels of granularities, one can acquire various levels of knowledge, as well as inherent knowledge structures, and then choose what he needs, which can improve the efficiency of algorithm in reduction of knowledge.

One approach about reduction of knowledge proposed by Skowron, is named discernibility matrix[7], a representative way in computing all the reduction of attributes in knowledge representation system. However, it is on the basis of objects. As the number of objects increases, the computing process of the approach is unimaginable, which, in fact, contains lots of redundant information.

In this paper, we use the idea of granular computing to eliminate the redundant information in discernibility matrix. Thus, the workload is diminished, and the space of storage is saved. In addition, a new kind of method called absorptivity, based on bit-vector, is also proposed to decrease the computing complexity and could be easily operated by computer. An example is presented at the end of the paper.

G. Wang et al. (Eds.): RSKT 2006, LNAI 4062, pp. 357–362, 2006.

2 Basic Conception

Decision table is a kind of important knowledge representation system. Most decision problems can be expressed by it. So now we describe a decision table to expatiate on the application in reduction of attributes by means of the idea of granular computing. If we discuss the corresponding problems in information table, we only need to weaken the relative core and relative reducts.

Definition 1. Decision Discernibility Matrix[7]**.** Let $DT = (U, C \bigcup D, V, f)$ is a decision table, where U is any nonempty finite set called a universe,$U = \{x_1, x_2, \cdots, x_n\}$. Then we define

$$M_{n\times n} = (c_{ij})_{n\times n} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ * & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & c_{nn} \end{bmatrix}$$

as a *decision discernibility matrix*, where for $\forall i, j = 1, 2, \cdots, n$

$$c_{ij} = \begin{cases} \{a|(a \in C) \vee (f_a(x_i) \neq f_a(x_j))\}, & f_D(x_i) \neq f_D(x_j); \\ \emptyset, & f_D(x_i) \neq f_D(x_j) \wedge f_C(x_i) = f_C(x_j); \\ -, & f_D(x_i) = f_D(x_j). \end{cases} \tag{1}$$

The definition of the discernibility matrix is very familiar to us, so the meaning of c_{ij} would not be explained here. We just recite several necessary propositions.

Property 1. In a consistent decision table, the relative D core is equal to the set which is composed by all the simple attribute (single attribute), namely

$$CORE_C(D) = \{a|(a \in C) \wedge (\exists c_{ij}, ((c_{ij} \in M_{n\times n}) \wedge (c_{ij} = \{a\})))\}. \tag{2}$$

Property 2. Let $\forall B \subseteq C$,if satisfies the two conditions below: (1) For $\forall c_{ij} \in M_{n\times n}$, when $c_{ij} \neq \emptyset, c_{ij} \neq -$, always gets $B \bigcap c_{ij} \neq \emptyset$.(2) If B is relative independent to D , then B is a relative reduct of the decision table.

From the upper statement, we can get the relative core and relative reducts. But in fact, many elements in the original discernibility matrix are redundant. It is unnecessary to compare with the objects which are in the same equivalent classes, because the value of c_{ij} obtained in the discernibility matrix is either "$-$" or $\emptyset$. It occupies much storage space and increases the complexity of computing.

Definition 2. Discernibility Matrix Based on Information Granule. Let $DT = (U, C \bigcup D, V, f)$ be a decision table, $U/IND(C) = \{E_i | \forall E_i = [u]_{IND(c)}, 1 \leq i \leq m\}$, where the universe U is a nonempty finite set, $U = \{x_1, x_2, \cdots, x_n\}$, then we define

$$M^G_{m\times m} = (r^G_{ij})_{n\times n} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mm} \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ * & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & r_{mm} \end{bmatrix}$$

as the *discernibility matrix based on the information granule*, where for

$\forall i, j = 1, 2, \cdots, m.$

$$r^G_{ij} = \begin{cases} \{a|(a \in C) \vee (f_a(E_i) \neq f_a(E_j))\}, & f_D(E_i) \neq f_D(E_j); \\ \emptyset, & f_D(E_i) \neq f_D(E_j) \wedge f_C(E_i) = f_C(E_j) \\ -, & f_D(E_i) = f_D(E_j). \end{cases} \tag{3}$$

In common situation,$m \ll n$, so we can reduce the rank of discernibility matrix by the ideal of granular computing, then the work of computation can be decreased. The approach is also suitable to the information table.

We just need replace the objects x_i, x_j with E_i, E_j, then the definition of the discernible function (Boolean function) based on the information granule can be obtained. Also we can prove easily:

1) The relative D core is equal to the set which is composed of all the simple attribute (single attribute);
2) If B is a relative reduct of a decision table, it satisfies
(1) For $\forall r^G_{ij} \in M^G_{m\times m}$, when $r^G_{ij} \neq \emptyset, r^G_{ij} \neq -$, always gets $B \bigcap r^G_{ij} \neq \emptyset$;
(2) If B is relative independent to D.

3 Absorptivity Based on Bit-Vector

Generally speaking, the process of obtaining core and reducts by discernibility matrix always converses to find the minimal disjunction normal form. If the number of the items is huge, the cost will be very large. Therefore we propose an approach called *absorptivity* based on a bit operation in binary system to simplify computing and save storage space. The unnecessary elements will be deleted through *absorptivity.* The measure is benefit to operate on a large scale of data or information, and can improve the efficiency of attributes reduction algorithm.

Definition 3. Absorptivity Based on Bit-Vector. Given a discernibility matrix based on information granule of a decision table, for any $r^G_{ij} \in M_{m\times m}$, $v^G_{ij} = \{\bullet, \bullet, \cdots, \bullet\}$ represents a vector with the dimension of $Card(C)$, where every component is either "1" or "0". If $a_i \in C \wedge a_i \in r^G_{ij}$, then we let the ith be "1", else be "0"; $\tau = \sum \bullet$ represents the rank of v_{ij}, we define

1) $\forall a_l \in r^G_{ij}, a_l \notin r^G_{ks}(i \neq k \vee j \neq s)$. It means v^G_{ij} and v^G_{ks} are independent to each other, and can not replace by each other, namely the values of their

corresponding components are not all "1" in vector expression. We record the two different vectors;
2) $\exists a_l \in r_{ij}^G, a_l \in r_{ij}^G (i \neq k \vee j \neq s)$. It means v_{ij}^G and v_{ks}^G are relative to each other. If $\tau_{ij} < \tau ks$, then we replace v_{ks}^G with v_{ij}^G, it says v_{ij}^G absorbs v_{ks}^G; else v_{ij}^G is replaced by v_{ks}^G, it says v_{ks}^G absorbs v_{ij}^G; If $\tau_{ij} = \tau_{ks}$, we say they can replace by each other, or absorb each other; If we append a criterion function(sort function) to select attributes, we can choose the priority attributes judged by the function.

Through the *absorptivity*, we can find the relative core and relative reducts easily. But there are some tips need to be noticed.

1) If we just need to get the relative core, then when v_{ij}^G and v_{ks}^G are relative to each other and $\tau_{ij} = \tau_{ks}$, we need not record the different vector, the vector we get at last is the relative core.
2) If we need to find all the relative reducts, then when v_{ij}^G and v_{ks}^G are relative to each other and $\tau_{ij} = \tau_{ks}$, we can not drop the different vector. Because we know the relative reduct is not unique, the different vectors also contain the information about one relative reduct. At last all the vectors recorded consist of a matrix, then we choose the "1" in different rows and different columns, and get all the relative reducts.

In the next paragraph, we illustrate how to use the *absorptivity* in detail.

4 Analysis

Example: We make an expatiation on the approach through a decision table(see the bibliography[8]).
Solve: We get the relative core and the relative reduct of the decision table by discernibility matrix of Skowron, then

$$M_{31\times 31} = \begin{bmatrix} - & & & & \\ - & - & & & \\ \{a,d\} & \{a\} & - & & \\ \vdots & \vdots & \vdots & \ddots & \\ \{a,b,c,d\} & \{a,b,c,d\} & \ldots\ldots & - & \\ \{a,b,c,d\} & \{a,b,c,d\} & \ldots\ldots & - & - \end{bmatrix}$$

According to the discernible function of discernibility matrix and traditional absorptivity of logic operation, we get the conjunction normal formal. Then the minimal disjunction normal formal of the conjunction normal formal of discernible function can be obtained by logic operation:

$$L_\wedge(M) = \bigwedge_{c_{ij}\neq\emptyset \wedge c_{ij}\neq -} (c_{ij}) = L_\vee(M) = (a \wedge c \wedge d).$$

Then the relative core and relative reduction of this decision table are $\{a, c, d\}$, that is $RED_D(C) = CORE_D(C) = \{\{a, c, d\}\}$.

In common situation, the relative core is unique, while the relative reduct is not. The simplify process of logic operation upper will cause "combination explode" as the cardinal number of universe increases. Therefore, we improve the efficiency of this algorithm by the ideal of granular computing and *absorptivity* based on bit-vector. Then we can get the discernibility matrix based on information granule, according to definition 2:

$$M_{13\times 13} = \begin{bmatrix} - & & & & \\ - & - & & & \\ - & - & - & & \\ \{a,b,d\} & \{a,d\} & \ldots & - & \\ \{a,b\} & \{a\} & \ldots & - & \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ \{a,c\} & \{a,b,c\} & \ldots\ldots\ldots & - & \\ \{a,b,c\} & \{a,b,c\} & \ldots\ldots\ldots & - & - \end{bmatrix}$$

$m = 13 \leq n = 31$.

We can see the rank of discernibility matrix has been largely decreased.

Then we try to get the relative core and relative reducts through *absorptivity* and the operation of database.

From the matrix, we first get a vector $(1,1,0,1)$. Then we use a table to represent, namely

a	b	c	d	rank
1	1	0	1	3

Following we add the vector $(1,1,0,0)$, according to *absorptivity* we replace the vector with $(1,1,0,0)$, so the table will be:

a	b	c	d	rank
1	1	0	0	2

Add the vector $(1,0,1,0)$, then this vector and the upper vector can replace with each other according to *absorptivity*, but we do not get the criterion function, so it is not necessary to replace. But it can not be dropped, it must be recorded in the table, then we get:

a	b	c	d	rank
1	1	0	0	2
1	0	1	0	2

Repeat the upper process until the difference of all information granules have been compared, we get the table finally:

Table 1

a	b	c	d	rank
1	0	0	0	1
0	0	1	0	1
0	0	0	1	1

According to the definition of *absorptivity*, we select the "1" in different columns and different rows, then constitute a vector $(1, 0, 1, 1)$, represent the relative reduct $\{a, c, d\}$. Moreover, we can see the rank of every vector is 1. That means the attributes which these vectors represent are all single attribute in the discernibility matrix, consequently we can get the relative core $\{a, c, d\}$.

5 Conclusion

Reduction of knowledge is the kernel problem in rough set theory. Skrown's discernibility matrix is a kind of effective approach to seek for the relative core and all the relative reducts in knowledge representation system, but we find that there is lots of redundant information in operation which is unnecessary. As a result, the idea of granular computing is used to lower the rank of discernibility matrix. What's more, the computing of logic conjunction and disjunction operation is so much complicated that the new absorptivity based on bit-vector is proposed. It is different from the absorptivity in logic operation, and can simplify computation greatly. At last, we give an example to analyze the results to support our ideas.

Acknowledgment

The paper is supported by grant No: 60175016;604750197 from the National Science Foundation of China.

References

1. Andrew Stranieri, John Zeleznikow.: Knowledge Discovery for Decision Support in Law. In: Proceedings of the twenty first international conference on Information systems, Brisbane (2000) 635-639.
2. Andy Podgurski, Charles Yang.: Partition testing, stratified sampling, and cluster analysis. In: Proceedings of the 1st ACM SIGSOFT symposium on Foundations of software engineering SIGSOFT '93, New York (1993) 169–181.
3. Li, D.G., Miao, D.Q., Zhang, D.X., Zhang, H.Y.: An Overview of Granular Computing. *Computer Science*. 9 (2005) 1–12.
4. Toshinori, Munakata.: Knowledge Discovery. *Communications of the ACM*. **11** (1999) 26–29.
5. Pawlak, Z.:Rough sets. *International Journal of Information and Computer Science*. **11** (1982) 341–356.
6. Skowron, A., Rauszer, C.: The discernibility matrices and functions in information system. *Intelligent Decision Support-Handbook of Application and Advantages of the Rough Sets Theory*. (1991) 331–362.
7. Wang, G.Y.: *Rough Sets and KDD*. Xian Traffic University publishing company, Xian(2001).
8. Zhang, W.X., Wu, W.Z., Liang, J.Y., Li, D.Y.: *The Theory and Approach in Rough Sets*. Science and Technology publisher,Beijing(2001).