Research of Knowledge Reduction Based on New Conditional Entropy

Jiucheng Xu and Lin Sun

College of Computer and Information Technology
Henan Normal University, Xinxiang Henan 453007, China
xjch3701@sina.com, slinok@126.com

Abstract. Although knowledge reduction for a decision table based on discernibility function can be used widely in data classification, there are also many disadvantages needed discussing detailedly on knowledge acquisition. To make some improvement for them, firstly, the concept of a decision table simplified was put forward for removing redundant data. Then based on knowledge granulation and conditional information entropy, the definition of a new conditional entropy, which could reflect the change of decision ability objectively and equivalently and present the concepts and operations in an inconsistent decision table simplified, was given by separating the consistent objects from the inconsistent objects. Furthermore, many propositions and properties for reduction with an inequality were proposed, and a complete knowledge reduction method was implemented. Finally, the experimental results with UCI data sets show that the proposed method of knowledge reduction is an effective technique to deal with complex data sets, and can simplify the structure and improve the efficiency of data classification.

Keywords: Rough set, decision table, knowledge reduction, knowledge granulation, conditional entropy.

1 Introduction

Rough set theory [1] has been applied to many areas successfully including data mining, pattern recognition, machine learning and so on. Knowledge reduction for data classification is always one of its most important topics. In recent years, there are many proposed algorithms of attribute reduction whose mechanisms can be classified as information entropy, position region, discernibility matrix and function, and genetic algorithms. As follows, we briefly review some relevant literatures. To evaluate uncertainty of a system, the concept of entropy was introduced by Shannon [2]. Shannon's information entropy was introduced to search reducts in classical rough set model, and the relationships of the definitions of rough reduction in algebra view and information view were presented in [3].Then in [4], two novel heuristic reduction algorithms have been proposed, whose worst time complexity is $O(|C||U|^2) + O(|U|^3)$ and $O(|C|^2|U|) + O(|C||U|^3)$ respectively, where C is a conditional attributes set, and U is a universe of objects. In [4], the algorithm with the time complexity $O(|C|^3|U|^2)$ could be run in parallel mode to

P. Wen et al. (Eds.): RSKT 2009, LNCS 5589, pp. 144–151, 2009.

compute all keys concurrently. In [6], based on mutual information, a heuristic reduction algorithm with the time complexity $O(|C||U|^2) + O(|U|^3)$ was proposed. Based on the indiscernibility relation and positive region, a complete algorithm for the reduction of attributes with time complexity $O(|C|^2|U|\log|U|)$ and space complexity $O(|C||U|)$ was introduced in [7]. Recently, a new and relatively reasonable formula was designed for an efficient attribute reduction algorithm, whose worst time complexity was cut down to $\mathrm{Max}(O(|C||U|), O(|C|^2|U/C|))$in [8]. However, many reduction algorithms have still their own limitations in [9].

A primary use of rough set theory is to reduce the number of attributes in databases thereby improving the performance of applications in a number of aspects including speed, storage, and accuracy. Thus based on discernibility function, to obtain some disadvantages of these current classical reduction algorithms, we continue deep investigations from the algebra view and the information view, and then we draw the proposition that the classical information entropy of knowledge can't reflect the complete change of decision ability objectively, and the lower approximation of a set is not equivalent to its conditional information entropy in an inconsistent decision table. In a presented decision table simplified, to overcome the shortcomings of existing measures, we construct a new conditional entropy with separating the consistent objects from the inconsistent objects, which is suitable for evaluating the roughness and accuracy of knowledge granulation, and the judgment theorems and properties for reduction with an inequality are proposed. Then based on this entropy, the new significance of an attribute is defined and compared with that based on positive region and conditional information entropy respectively, and a complete heuristic algorithm for knowledge reduction is designed and an efficient method for computing the new conditional entropy is proposed. Finally, the experimental results with seven data sets from UCI Machine Learning Repository show that the proposed method is an effective technique to deal with complex data sets, simplify the structure of decision tables, and improve the efficiency of data classification.

2 Rough Set Theory Preliminaries

A decision table is defined as $S = (U, C, D, V, f)$. Each non-empty subset $P \subseteq C \cup D$ determines a binary discernibility relation, i.e.$IND(P)$=$\{(x, y) \in U \times U | f(x, a) = f(y, a), \forall a \in P\}$. $IND(P)$ determines a partition on U, that is $U/IND(P)$ (in short U/P), and any element $[x]_P$=$\{y|f(x, a) = f(y, a), \forall a \in P\}$ in U/P is called an equivalent class. If P, Q are two equivalence relations on U, $\forall P_i \in U/P$, and $\exists Q_j \in U/Q$, so that $P_i \subseteq Q_j$, then we call the partition U/Q coarser than the partition U/P.

Let $P \subseteq C$, $POS_P(D) = \cup\{\underline{P}Y|Y \in U/D\}$ is called the P-positive region of D, where $\underline{P}Y = \cup\{[x]_P|[x]_P \subseteq Y\}$ characterizes the P-lower approximation of Y. Thus in a decision table $S = (U, C, D, V, f)$, we have the proposition that if $POS_C(D) = U$, then this decision table is called a consistent one, else an inconsistent one. Meanwhile, the set $POS_C(D)$ is called a (positive region) consistent objects set on S, and $U - POS_C(D)$ is called an inconsistent one.

Thus if $D_0 = U - POSC(D)$, then we have $\underline{C}D_0 = D_0$, and if this decision table is a consistent one, then $\underline{C}D_0 = \emptyset$ obviously.

Let R be an equivalent relation on U, U/R=$\{R_1, R_2, \ldots, R_s\}$, then the formula of knowledge granulation of R is denoted by $G(R) = \sum_{i=1}^{s}(|R_i|/|U|)^2$.

Let P, Q be two equivalent relations on U, U/P=$\{X_1, X_2, \ldots, X_n\}$, U/Q=$\{Y_1, Y_2, \ldots, Y_m\}$, then the conditional information entropy of Q to P is defined as

$$H(Q|P) = -\sum_{i=1}^{n}(|X_i|/|U|)\sum_{j=1}^{m}(|X_i \cap Y_j|/|X_i|)\log(|X_i \cap Y_j|/|X_i|). \tag{1}$$

3 Construction of New Conditional Entropy and Knowledge Reduction Method

3.1 The Proposed New Conditional Entropy

In a decision table $S = (U, C, D, V, f)$, if $r \in P \subseteq C$, $H(D|P) = H(D|P - \{r\})$, then r is unnecessary for D in P, else necessary, and if every element in P is necessary for D, then P is independent relative to D. Hence, if $H(D|P) = H(D|C)$, and P is independent relative to D, then P is called a reduct of C relative to D in [3,4].

However, we analyze these algorithms presented from the information view detailedly, then we have found that firstly as the algorithms based on conditional information entropy are concerned, whether any condition attribute is redundant depends on whether the conditional information entropy of a decision table is changed or not, after the condition attribute is deleted. It is easily shown that the conditional information entropy generated by the consistent objects set $POS_C(D)$ is 0, then the inconsistent objects set $U - POS_C(D)$ can lead to the change of conditional information entropy in a decision table. For the new added and primary inconsistent objects in every corresponding to decision attribute classes, only through the change of their conditional probability distributions can the conditional information entropy for a decision table get change. Thus the main criterions of these algorithms on estimating decision ability only include two aspects, one is the invariability of the number of definite decision rules, the other is the invariability of the reliability of indefinite decision rules. So for these current reduction algorithms, we find that these researchers only think about the change of reliability for all decision rules after reduction.

Secondly, in actual decision application, besides the reliability of decision rules, the objects coverage of decision rules is also one of the most important standards on estimating decision ability effectively.

For solving the problems above, we introduce the property that if P, Q are two equivalent relations on U, then $U/(P \cup Q) = U/P \cap U/Q$ holds. Thus we can get the proposition that if $P \subseteq C$ is an equivalent relation on U, $\forall r \in C - P$, then $U/(P \cup \{r\}) = U/P \cap U/\{r\}$=$\cup\{X/\{r\}|X \in U/P\}$ always holds.Thus we can get the knowledge granulation of $\{P \cup D\}$.

Definition 1. Let $P \subseteq C$, $U/P = \{X_1, X_2, \ldots, X_n\}$, $U/D = \{Y_1, Y_2, \ldots, Y_m\}$, then the definition of knowledge granulation of $\{P \cup D\}$ is denoted by

$$G(P \cup D) = \sum_{i=1}^{n} \sum_{j=1}^{m} (|X_i \cap Y_j|/|U|)^2. \quad (2)$$

Hence, from (1) and (2), it is shown that $|X_i \cap Y_j|/|X_i|$ determines the reliability of any decision rule generated in a decision table, and $|X_i \cap Y_j|/|U|$ depicts the objects coverage corresponding to the decision rules. Thus if the two definitions about data classification above are combined to form a new measure, then it will help to compensate for the limitation that the researchers for the current information entropy only think the change of reliability of all decision rules.

Thirdly, from the algebra view in [1] [7] [8], we analyze that the significance of an attribute is regarded as the quantitative computation of the radix of positive region which merely depicts the subsets of certain classes on U. However, from the viewpoint of information view in [3] [4], we find that its significance of an attribute is only obtained by detaching objects in different decision classes from equivalent classes of the condition attributes subset. Hence, due to the inconsistent objects, we have the proposition that the current proposed measures which are based on positive region and conditional information entropy, used as heuristic information, are still lack in dividing U into the consistent objects sets and the inconsistent objects sets in an inconsistent decision table. Therefore, these heuristic algorithms of attribute reduction from the algebra view and the information view will not be equivalent each other in the representation of concepts and operations in an inconsistent decision table.

Because D_0 is a set of inconsistent objects, we construct that the new partition $\{\underline{C}D_0, \underline{C}D_1, \underline{C}D_2, \ldots, \underline{C}D_m\}$ is divided into the consistent and the inconsistent objects sets respective1y, and then all inconsistent objects are detached to form a whole set. If any subset of the partition $\{\underline{C}D_0, \underline{C}D_1, \underline{C}D_2, \ldots, \underline{C}D_m\}$ isn't empty, we think that this new partition must be also a decision partition on U, and if this new partition has any empty decision class $\underline{C}D_i$, then the $\underline{C}D_i$ is called a redundant subset. After all redundant subsets are deleted from the new partition, it makes no difference to the new decision partition. Hence, we have a new equivalent relation generated by the new decision partition, denoted by RD, i.e. $U/RD = \{\underline{C}D_0, \underline{C}D_1, \underline{C}D_2, \ldots, \underline{C}D_m\}$. Thus it shows that the new decision partition U/RD has not only detached the consistent objects from different decision classes on U, but also separated the consistent objects from the inconsistent objects, while the classical partition U/D is gained only through detaching objects from different decision classes corresponding to the equivalent classes. Meanwhile, because a great deal of redundant data in the large decision tables have badly disturbed the data classification, for improving the efficiency of data classification, in a decision table $S = (U, C, D, V, f)$, we propose the partition $U/C = \{[U_1]_C, [U_2]_C, \ldots, [U_n]_C\}$ to construct a compact and non-redundant set of objects, that is $U^* = \{U_1 \cup U_2 \cup \ldots U_n\}$, thus we call $S^* = (U^*, C, D, V, f)$ a decision table simplified.

Definition 2. In a decision table simplified $S^* = (U^*, C, D, V, f)$, $P \subseteq C$, $U^*/RD = \{\underline{C}D_0, \underline{C}D_1, \underline{C}D_2, \ldots, \underline{C}D_m\}$, then the new conditional entropy of equivalent relation RD with respect to P is defined as

$$HG(RD; P) = H(RD|P) + G(P \cup RD). \tag{3}$$

Proposition 1. In a decision table simplified $S^* = (U^*, C, D, V, f)$, $P \subseteq C$, the partition of P on U^* is $U^*/P = \{X_1, X_2, \ldots, X_n\}$. Assume X_p and X_q are two random equivalent classes in U^*/P, so that $\forall r \in P$ is deleted from P, then $U^*/(P - \{r\}) = \{X_1, X_2, \ldots, X_{p-1}, X_{p+1}, \ldots, X_{q-1}, X_{q+1}, \ldots, X_n, X_p \cup X_q\}$ is a new partition formed by unifying X_p and X_q to $X_p \cup X_q$ from U^*/P. If U^*/RD=$\{\underline{C}D_0, \underline{C}D_1, \underline{C}D_2, \ldots, \underline{C}D_m\}$, then $HG(RD; P) \le HG(RD; P - \{r\})$.

Proof. From (1), (2), (3), order $HG_\triangle$=$HG(RD; P - \{r\}) - HG(RD; P)$, then

$$\begin{aligned} HG_\triangle = {} & \sum_{j=0}^{m} \tfrac{|X_p \cap \underline{C}Y_j|}{|U^*|} \log \tfrac{|X_p \cap \underline{C}Y_j|}{|X_p|} + \sum_{j=0}^{m} \tfrac{|X_q \cap \underline{C}Y_j|}{|U^*|} \log \tfrac{|X_q \cap \underline{C}Y_j|}{|X_q|} \\ & - \sum_{j=0}^{m} \tfrac{|(X_p \cap \underline{C}Y_j) \cup (X_q \cap \underline{C}Y_j)|}{|U^*|} \log \tfrac{|(X_p \cap \underline{C}Y_j) \cup (X_q \cap \underline{C}Y_j)|}{|X_p \cup X_q|} \\ & - \sum_{j=0}^{m} (\tfrac{|X_p \cap \underline{C}Y_j|}{|U^*|})^2 - \sum_{j=0}^{m} (\tfrac{|X_q \cap \underline{C}Y_j|}{|U^*|})^2 + \sum_{j=0}^{m} (\tfrac{|(X_p \cap \underline{C}Y_j) \cup (X_q \cap \underline{C}Y_j)|}{|U^*|})^2. \end{aligned}$$

Order $|X_p| = x$, $|X_q| = y$, $|X_p \cap \underline{C}Y_j| = ax$, $|X_q \cap \underline{C}Y_j| = by$, it is easily known that $x > 0$, $y > 0$, $0 \le a \le 1$, $0 \le b \le 1$, then

$$HG_\triangle = \frac{1}{|U^*|} \sum_{j=0}^{m} [ax \log \frac{a(x+y)}{ax+by} + by \log \frac{b(x+y)}{ax+by} + \frac{2axby}{|U^*|}] = \frac{1}{|U^*|} \sum_{j=0}^{m} f_j,$$

where f_j is a function. It is obviously true that if $ab = 0$, get f_j=0, that is $HG(RD; P) = HG(RD; P - \{r\})$. Thus $0 < a \le 1$ and $0 < b \le 1$ shall be only considered in the following. Order $ax = \varphi$, $by = \psi$, $a/b = \theta$, obviously get $\varphi > 0$, $\psi > 0$, $\theta > 0$, and then

$$f_j = \varphi \log \frac{\varphi + \theta\psi}{\varphi + \psi} + \psi \log \frac{\varphi + \theta\psi}{\theta(\varphi + \psi)} + \frac{2\varphi\psi}{|U^*|} \Rightarrow \frac{d(f_j)}{d(\theta)} = \frac{(\theta - 1)\varphi\psi}{\theta(\varphi + \theta\psi)}.$$

Thus we find out the proposition as follows: $0 < \theta < 1 \Rightarrow \frac{d(f_j)}{d(\theta)} < 0$, $\theta = 1 \Rightarrow \frac{d(f_j)}{d(\theta)} = 0$, $1 < \theta \Rightarrow \frac{d(f_j)}{d(\theta)} > 0$. Therefore, when $\theta = 1$, namely $a = b$, then the function f_j gets the minimal $f_j = 2\varphi\psi/|U^*| > 0$. Hence, the above shows that when $\forall r \in P$ is deleted from P in a decision table, there must be $HG_\triangle \ge 0$, and then $HG(RD; P) \le HG(RD; P - \{r\})$ always holds.

It is known that the coalition of some partitions can be considered as automatically comprising more two partitions continually, due to the selection of X_p and X_q at random. Meanwhile, after $\forall r \in P$ is deleted from P in a decision table, suppose that there are many partitions for coalition, then the new partition $U^*/(P - \{r\})$, formed by the equivalent relation $\{P - \{r\}\}$ on U^*, is coarser than U^*/P. So we have the proposition that the new conditional entropy of knowledge monotonously reduces with the diminishing of the information granularity.

Lemma 1. Let $P \subseteq C$ be equivalent relations on U^*, then $\forall r \in P$ is dispensable in P with respect to D if and only if $HG(RD;P) = HG(RD;P - \{r\})$.

3.2 Implementation of Knowledge Reduction Method

Definition 3. In a decision table simplified $S^* = (U^*, C, D, V, f)$, $P \subseteq C$, then the significance of an attribute $r \in C - P$ with respect to D is defined as

$$SGF(r, P, D) = HG(RD;P) - HG(RD;P \cup \{r\}). \tag{4}$$

Note that when $P = \phi$, $SGF(r, \phi, D) = -HG(RD;\{r\})$. It is seen that if the proposed $SGF(r, P, D) = 0$, then the significance of an attribute based on positive region is also 0. If the radix of positive region fills out after adding an attribute, then the significance of an attribute based on positive region isn't 0, and $SGF(r, P, D) \neq 0$. Because $SGF(r, P, D) = 0$ can depict the data classification with separating the consistent objects from the inconsistent objects, while the significance of an attribute based on conditional information entropy is not. So we have the proposition that $SGF(r, P, D)$ can include more information.

Definition 4. In a decision table simplified $S^* = (U^*, C, D, V, f)$, $P \subseteq C$, if $HG(RD;P) \leq HG(RD;C)$, $\forall r \in P \Rightarrow HG(RD;P - \{r\}) < HG(RD;C)$, and $\forall P^* \subset P \Rightarrow HG(RD;P^*) < HG(RD;P)$, then P is a reduct of C relative to D.

It is easily seen that calculating any r to the maximum of $SGF(r, P, D)$ is in fact to calculate that corresponding to the minimum of $HG(RD;P \cup \{r\})$ what is actually to calculate corresponding partitions, core ($CORE_D(C)$), and positive region principally. Hence, to improve the efficiency of data classification, making full use of these effective measures in [8] [10], we firstly design an efficient algorithm for computing the new conditional entropy $HG(RD;P \cup \{r\})$.

Input: $S^* = (U^*, C, D, V, f)$ is a decision table simplified, $P \subseteq C$, and $r \in C - P$.
Output: U^*/C, U^*/D, U^*/P, $U^*/(P \cup \{r\})$, U^*/RD, $U^*/(RD \cup P \cup \{r\})$, and $HG(RD;P \cup \{r\})$.
(1) Calculate U^*/C, U^*/D, U^*/P, $U^*/(P \cup \{r\})$, to get $POS_C(D)$, $U^* - POS_C(D)$, and U^*/RD, $U^*/(RD \cup P \cup \{r\})$, thus $H(RD|P \cup \{r\}) = H(P \cup \{r\} \cup RD) - H(P \cup \{r\})$, and $G(P \cup \{r\} \cup RD)$.
(2) Calculate $HG(RD;P \cup \{r\}) = H(RD|P \cup \{r\}) + G(P \cup \{r\} \cup RD)$.
In succession, we can obtain the minimum relative reduction set through adding attribute step by step bottom-up.
Input: $S^* = (U^*, C, D, V, f)$ is a decision table simplified, $P \subseteq C$, and $r \in C - P$.
Output: A minimum relative reduction set P.
(1) Calculate $HG(RD;C)$, and let $P = CORE_D(C)$.
(2) If $|P| = 0$, then turn to (3). If $HG(RD;P) \leq HG(RD;C)$, then turn to (5).
(3) Select an attribute r with the minimum of $HG(RD;P \cup \{r\})$, and if r is not only, then select one with the maximum of $|U^*/(P \cup \{r\})|$, and $P = P \cup \{r\}$.
(4) If $\forall r \in P \Rightarrow HG(RD;P - \{r\}) \geq HG(RD;C)$, then turn to (3), else $\{Q = P$, $P^* = P - CORE_D(C)$, $t = |P^*|$, for $\forall i (1 \leq i \leq t) \Rightarrow \{$ if $r_i \in P^*$, $P^* =$

$P^* - \{r_i\}$, $HG(RD; P^* \cup CORE_D(C)) < HG(RD; Q)$, then $P^* = P^* \cup \{r_i\},\}$
$P = P^* \cup CORE_D(C).\}$
(5) The output P is a minimum relative attribute reduction.

It is clear that this algorithm of attribute reduction is complete, in other words, none of the attributes in P can be eliminated again without decreasing its discriminating capability, while a great many reduction algorithms are still incomplete, which can't ensure that the final reducts will be obtained. Then through analyzing, it is easily known that these reduction algorithms in [7] [10] are also complete, while those algorithms in [4-6] are not. By making full use of those feasible measures, we can easily get the time complexity of the proposed algorithm which is cut down to $O(|C|^2|U^*/C|)$, which is less than that of [4-7] [9-11].

4 Experimental Results

The experiments on PC (AMD Dual Core 2.71GHz, 2GB RAM, WINXP) under JDK1.4.2, are performed on several different real-life data sets obtained from UCI Machine Learning Repository, then we choose five algorithms such as the Algorithm A in [5], MIBARK in [6], CEBARKNC in [4], Algorithm 4 in [7], the reduction method in [11], compared with the proposed algorithm, denoted by A, B, C, D, E, F respective1y. Thus we obtain the results of reduct comparison in Table 1, where s, m, n are the numbers of objects, primal condition attributes, and after reduction respectively, and t is the time of operation.

Table 1. Comparison of Reduct Results

Data Sets	s	m	A		B		C		D		E		F	
			n	t	n	t	n	t	n	t	n	t	n	t
Iris	150	4	3	0.09	3	0.08	3	0.07	3	0.05	3	0.17	3	0.04
Liver-disorders	345	6	3	0.12	3	0.19	3	0.12	3	0.11	3	0.23	3	0.09
Tic-Tac-Toe	958	9	8	0.91	8	0.49	8	0.46	8	0.41	8	5.68	8	0.39
Voting-records	435	16	10	0.92	9	0.52	9	0.50	9	0.20	9	6.15	9	0.15
Zoo	101	17	11	0.35	11	0.30	11	0.29	10	0.11	10	3.54	10	0.06
Mushroom	8142	22	5	470.5	5	16.81	4	16.02	5	5.25	5	168.63	4	4.83
Chess End Game	3196	36	29	261.6	29	23.28	29	23.35	29	3.48	29	98.37	29	3.15

From the experimental simulation results, in a decision table simplified, based on U^*/RD, the new conditional entropy is feasible to discuss roughness of rough set, and the proposed heuristic information will make up for the limitations of estimating decision ability. Hence, the method of knowledge reduction is a promising and effective technique to deal with complex data sets, and can simplify the structure and improve the efficiency of data classification.

5 Conclusion

Attribute reduction is a basic issue in knowledge acquisition and data classification, and the uncertainty of rough set has been widely studied in rough set technology. In this paper, to reflect the change of decision ability objectively in a decision table simplified, an uncertainty measure of knowledge reduction and its propositions are established by constructing a new conditional entropy. To compensate for some shortcomings of the classical reduction algorithms, we use the new conditional entropy as heuristic information to design and implement an efficient complete knowledge reduction method, whose worst time complexity has been cut down well, and the experimental results of this method are also effective. Furthermore, the next further researches, which are characterized by insufficient and incomplete information in intelligent systems, will be to consummate the more efficient methods for knowledge reduction.

Acknowledgment. This paper was supported by the National Natural Science Foundation of China under Grant (No. 60873104) and New Century Excellence Genius Support Plan of Henan Province of China (No. 2006HANCET-19).

References

1. Pawlak, Z.: Rough Sets and Intelligent Data Analysis. International Journal of Information Sciences 147, 1–12 (2002)
2. Shannon, C.E.: The Mathematical Theory of Communication. The Bell System Technical Journal 27(3/4), 373–423 (1948)
3. Wang, G.Y., Zhao, J., An, J.J., Wu, Y.: A Comparative Study of Algebra Viewpoint and Information Viewpoint in Attribute Reduction. Fundamenta Informaticae 68(3), 289–301 (2005)
4. Wang, G.Y., Yu, H., Yang, D.C.: Decision Table Reduction Based on Conditional Information Entropy. Journal of Computers 25(7), 759–766 (2002)
5. Guan, J.W., Bell, D.A.: Rough Computational Methods for Information Systems. International Journal of Artificial Intelligences 105, 77–103 (1998)
6. Miao, D.Q., Hu, G.R.: A Heuristic Algorithm for Reduction of Knowledge. Journal of Computer Research and Development 36(6), 681–684 (1999)
7. Liu, S.H., Sheng, Q.J., Wu, B., et al.: Research on Efficient Algorithms for Rough Set Methods. Journal of Computers 26(5), 524–529 (2003)
8. Xu, Z.Y., Liu, Z.P., et al.: A Quick Attribute Reduction Algorithm with Complexity of Max($O(|C||U|)$,$O(|C|^2|U/C|)$). Journal of Computers 29(3), 391–399 (2006)
9. Liu, Q.H., Li, F., et al.: An Efficient Knowledge Reduction Algorithm Based on New Conditional Information Entropy. Control and Decision 20(8), 878–882 (2005)
10. Han, J.C., Hu, X.H., Lin, T.Y.: An Efficient Algorithm for Computing Core Attributes in Database Systems. In: Zhong, N., Raś, Z.W., Tsumoto, S., Suzuki, E. (eds.) ISMIS 2003. LNCS, vol. 2871, pp. 663–667. Springer, Heidelberg (2003)
11. Jiang, S.Y., Lu, Y.S.: Two New Reduction Definitions of Decision Table. Mini-Micro Systems 27(3), 512–515 (2006)