Serdica J. Computing **9** (2015), No. 2, 167–176

Serdica
Journal of Computing

Bulgarian Academy of Sciences
Institute of Mathematics and Informatics

ON THE TIME COMPLEXITY OF THE PROBLEM RELATED TO REDUCTS OF CONSISTENT DECISION TABLES*

Janos Demetrovics, Vu Duc Thi, Nguyen Long Giang, Tran Huy Duong

Abstract. In recent years, rough set approach computing issues concerning reducts of decision tables have attracted the attention of many researchers. In this paper, we present the time complexity of an algorithm computing reducts of decision tables by relational database approach. Let $DS = (U, C \cup \{d\})$ be a consistent decision table, we say that $A \subseteq C$ is a relative reduct of DS if A contains a reduct of DS. Let $s = \langle C \cup \{d\}, F \rangle$ be a relation schema on the attribute set $C \cup \{d\}$, we say that $A \subseteq C$ is a relative minimal set of the attribute d if A contains a minimal set of d. Let Q_d be the family of all relative reducts of DS, and P_d be the family of all relative minimal sets of the attribute d on s. We prove that the problem whether $Q_d \subseteq P_d$ is co-NP-complete. However, the problem whether $P_d \subseteq Q_d$ is in P.

ACM Computing Classification System (1998): I.5.2, I.2.6.

Key words: decision table, reduct, relation, relation schema, minimal set, time complexity.

*This work is sponsored by a research grant from Vietnam National University, Hanoi (QG.15.41).

1. Introduction. Attribute reduction is an important problem in the preprocessing of the data mining process. The aim of attribute reduction is to remove redundant attributes in order to improve the performance of data mining algorithms. In decision tables, attribute reduction is the process of finding a reduct which keeps the classification ability of decision tables. For consistent decision tables, paper [2] shows that the time complexity of the problem to find all reducts is exponential in the number of attributes in the worst case.

In this paper, we present our research on the time complexity of algorithms concerning reducts of consistent decision tables. Let $DS = (U, C \cup \{d\})$ be a consistent decision table where U is the set of objects and C is the set of conditional attributes. We say that $A \subseteq C$ is a relative reduct of DS if A contains a reduct of DS. Let $s = \langle C \cup \{d\}, F \rangle$ be a relation schema on the attribute set $C \cup \{d\}$. We say that $A \subseteq C$ is a relative minimal set of the attribute d if A contains a minimal set of d. Let Q_d be the family of all relative reducts of DS, and P_d be the family of all relative minimal sets of the attribute d on s. In this paper, we prove that the problem whether $Q_d \subseteq P_d$ is co-NP-complete. However, the problem whether $P_d \subseteq Q_d$ is polynomial time.

Now, we present some basic concepts about information systems, decision table, reduct in rough set theory [6] and some concepts in relational databases [1, 2]. Firstly, we summarize some basic concepts in rough set theory [6].

Definition 1.1. *An information system is $IS = (U, A)$ in which U is a finite and non-empty set of objects; A is a finite and non-empty set of attributes. Each attribute $a \in A$ determines a map: $a : U \to V_a$ where V_a is the value range of attribute $a \in A$.*

For any $u \in U$, $a \in A$ we will denote the value of attribute a on object u by $a(u)$. If $B = \{b_1, b_2, \ldots, b_k\} \subseteq A$ is a subset of attributes then the set of $b_i(u)$ is denoted as $B(u)$. Therefore, if u and v are two objects in U then $B(u) = B(v)$ if and only if $b_i(u) = b_i(v)$ for any $i = 1, \ldots, k$.

Definition 1.2. *A decision table is an information system $DS = (U, C \cup D)$ where $A = C \cup D$, C is the set of condition attributes, D is the set of decision attributes and $C \cap D = \emptyset$.*

Without loss of generality, suppose that D consists of the only one decision attribute d. Therefore, we consider the decision table $DS = (U, C \cup d)$ where $\{d\} \notin C$. A decision table DS is consistent if and only if the functional dependency $C \to \{d\}$ is true, which means that for any $u, v \in U$, if $C(u) = C(v)$ then $d(u) = d(v)$. Conversely, DS is inconsistent.

Definition 1.3. *Let $DS = (U, C \cup d)$ be a consistent decision table and an attribute $R \subseteq C$. R is called a reduct of DS if:*

1) *For any $u, v \in U$, if $R(u) = R(v)$ then $d(u) = d(v)$.*

2) *For any $E \subset R$, there exists $u, v \in U$ such that $E(u) = E(v)$ and $d(u) \neq d(v)$.*

The above reduct is called Pawlak reduct. Let $RED(C)$ is the set of all reducts of C.

In what follows, we introduce some basic concepts in relational database theory.

Let $R = \{a_1, \ldots, a_n\}$ be a finite set of attributes and let $D(a_i)$ be the set of all possible values of attribute a_i; a relation r over R is the set of tuples $\{h_1, \ldots, h_m\}$ where $h_j : R \to \underset{a_i \in R}{\cup} D(a_i), 1 \leq j \leq m$ is a function that $h_j(a_i) \in D(a_i)$.

Let $r = \{h_1, \ldots, h_m\}$ be a relation over $R = \{a_1, \ldots, a_m\}$. Any pair of attribute sets $A, B \subseteq R$ is called the functional dependency (FD for short) over R, and denoted by $A \to B$, if and only if

$$(\forall h_i, h_j \in r)\left((\forall a \in A)(h_i(a) = h_j(a)) \Rightarrow (\forall b \in B)(h_i(b) = h_j(b))\right).$$

The set $F_r = \{(A, B) : A, B \subseteq R, A \to B\}$ is called the full family of functional dependencies in r. Let $P(R)$ be the power set of attribute set R. A family $F = P(R) \times P(R)$ is called an f-family over R if and only if for all subsets of attributes $A, B, C, D \subseteq R$ the following properties hold:

(1) $(A, A) \in F$.

(2) $(A, B) \in F,\ (B, C) \in F \Rightarrow (A, C) \in F$.

(3) $(A, B) \in F,\ A \subseteq C, D \subseteq B \Rightarrow (C, D) \in F$.

(4) $(A, B) \in F,\ (C, D) \in F \Rightarrow (A \cup C, B \cup D) \in F$.

Clearly F_r is an f-family over R. It is also known [1] that if F is an f-family over R then there is a relation r such that $F_r = F$. Let us denote by F^+ the set of all FDs, which can be derived from F by using the rules (1)–(4).

A pair $s = (R, F)$, where R is a set of attributes and F is a set of FDs on R, is called the relation scheme. For any $A \subseteq R$, the set $A^+ = \left\{a : A \to \{a\} \in F^+\right\}$ is called the closure of A on s. It is clear that $A \to B \in F^+$ if and only if $B \subseteq A^+$. Similarly, $A_r^+ = \left\{a : A \to \{a\} \in F^+\right\}$ is called the closure of A on relation r.

Let $s = (R, F)$ be a relation scheme over R and $a \in R$. The set

$$\mathcal{K}_a^s = \{A \subseteq R : A \to \{a\}, B : (B \to \{a\})(B \subset A)\}$$

$\mathcal{K}_a^s$ is called the family of minimal sets of the attribute a over s. Similarly, the set

$$\mathcal{K}_a^r = \{A \subseteq R : A \to \{a\}, B \subseteq R : (B \to \{a\})(B \subset A)\}$$

$\mathcal{K}_a^r$ is called the family of minimal sets of the attribute a over r.

Recall that a family $\mathcal{K} \subseteq P(R)$ is a Sperner system on R if for any A, $B \in \mathcal{K}$ implies $A \not\subset B$. It is clear that $\mathcal{K}_a^r$, $\mathcal{K}_a^s$ are Sperner systems over R. Let $\mathcal{K}$ be a Sperner system. We defined the set $\mathcal{K}^{-1}$ as follows:

$$\mathcal{K}^{-1} = \{A \subset R : (B \in \mathcal{K}) \Rightarrow (B \not\subset A)\} \text{ and if } (A \subset C) \Rightarrow (\exists B \in \mathcal{K})(B \subseteq C).$$

It is easy to see that $\mathcal{K}^{-1}$ is a Sperner system on R, too. If $\mathcal{K}$ is a Sperner system over R as the set of all minimal keys of relation r (or relation scheme s) then $\mathcal{K}^{-1}$ is the set of subsets of R which does not contain the elements of $\mathcal{K}$ and which is maximal for this property. $\mathcal{K}^{-1}$ is called antikeys. If $\mathcal{K}$ is a Sperner system over R as the family of minimal sets of the attribute a over r (or s), in other words $\mathcal{K} = \mathcal{K}_a^r$ (or $\mathcal{K} = \mathcal{K}_a^s$), then $\mathcal{K}^{-1} = \{\mathcal{K}_a^r\}^{-1}$ (or $\mathcal{K}^{-1} = \{\mathcal{K}_a^s\}^{-1}$) is the family of maximal subsets of R which are not the family of minimal sets of the attribute a, defined as [1]:

$$\{\mathcal{K}_a^r\}^{-1} = \left\{A \subseteq R : A \to \{a\} \notin F_r^+, A \subset B \Rightarrow B \to \{a\} \in F_r^+\right\},$$

$$\{\mathcal{K}_a^s\}^{-1} = \left\{A \subseteq R : A \to \{a\} \notin F^+, A \subset B \Rightarrow B \to \{a\} \in F^+\right\}.$$

2. Results. Firstly, we prove that the problem whether $Q_d \subseteq P_d$ is co-NP-complete where Q_d is the family of all relative reducts of the consistent decision table $DS = (U, C \cup d)$, P_d is the family of all relative minimal sets of the attribute d on the relation scheme $s = \langle C \cup \{d\}, F\rangle$.

Lemma 2.1 ([2]). *Let $DS = (U, C \cup d)$ be a consistent decision table where $C = \{c_1, c_2, \ldots, c_n\}$, $U = \{u_1, u_2, \ldots, u_m\}$. Let us consider $r = \{u_1, u_2, \ldots, u_m\}$ on the attribute set $R = C \cup \{d\}$. We set*

$\mathcal{E}_r = \{E_{\mathrm{ij}} : 1 \le i < j \le m\}$ where $E_{\mathrm{ij}} = \{a \in R : a(u_i) = a(u_j)\}$

$\mathcal{M}_d = \{A \in \mathcal{E}_r : d \notin A, B \in \mathcal{E}_r : d \notin B, A \subset B\}$

Then $\mathcal{M}_d = (\mathcal{K}_d^r)^{-1}$, where $\mathcal{K}_d^r$ is the family of all minimal sets of the attribute d on relation r.

Lemma 2.2. *Let $DS = (U, C \cup d)$ be a consistent decision table, then $(\mathcal{K}_d^r)^{-1}$ is a Sperner system over C. Conversely, if $\mathcal{K}$ is a Sperner system over*

C then there exists a consistent decision table $DS = (U, C \cup d)$ such that $\mathcal{K} = (\mathcal{K}_d^r)^{-1}$.

Proof. According to the definition,

$$\{\mathcal{K}_a^r\}^{-1} = \left\{A \subseteq R : A \to \{a\} \notin F_r^+, A \subset B \Rightarrow B \to \{a\} \in F_r^+\right\}.$$

It is obvious that $(\mathcal{K}_d^r)^{-1}$ is a Sperner system. Conversely, If $\mathcal{K}$ is a Sperner system over C, suppose that $\mathcal{K} = \{A_1, \ldots, A_m\}$, we construct a decision table $DS = (U, C \cup d)$ as follows:

We set $U = \{u_0, u_1, \ldots, u_m\}$, $R = C \cup \{d\}$.

1) For all $c \in C$, we set $u_0\,(c) = 0$. Set $u_0\,(d) = 0$

2) For all i $(i = 1, \ldots, m)$, we set $u_i\,(c) = 0$ if $c \in A_i$; $u_i\,(c) = i$ otherwise. Set $u_i\,(d) = i$ for all i $(i = 1, \ldots, m)$.

We set $\mathcal{E}_r = \{E_{ij} : 1 \leq i < j \leq m\}$ where $E_{ij} = \{a \in R : a\,(u_i) = a\,(u_j)\}$.

We set $\mathcal{M}_d = \{A \in \mathcal{E}_r : d \notin A, B \in \mathcal{E}_r : d \notin B, A \subset B\}$.

We can see that $\mathcal{M}_d = \{A_1, \ldots, A_m\}$. According to Lemma 2.1 we have $\mathcal{M}_d = (\mathcal{K}_d^r)^{-1}$. Consequently, $\mathcal{K} = (\mathcal{K}_d^r)^{-1}$. □

Lemma 2.3. *Suppose that $\mathcal{K} = \{K_1, K_2, \ldots, K_t\}$ is a Sperner system over C. We construct the relation scheme $s_d = \langle R, F\rangle$, where $R = C \cup \{d\}$, $F = \{K_i \to \{d\}\,,\ i = 1, \ldots, t\}$. Then $\mathcal{K} = \mathcal{K}_d^s - \{d\}$.*

Proof. (1) For any $K \in \mathcal{K}$ we have $K \to \{d\}\,, K \neq \{d\}$ and $\mathcal{K}$ is a Sperner system, so there is no $K' \subset K$ such that $K' \to \{d\}$. Consequently, K is a minimal set of the attribute d on s_d, meaning that $K \in \mathcal{K}_d^s - \{d\}$.

(2) Conversely, for any $K \in \mathcal{K}_d^s - \{d\}$ we have $K \to \{d\}\,, K \neq \{d\}$ and there is no $K' \subset K$ such that $K' \to \{d\}$.

It is easy to see that for any $K_i \in \mathcal{K}$ we have $K \not\subset K_i$, because if $K \subset K_i$ then $K \to \{d\}$ is not true according to the definiton of relation schema s_d. Moreover, for any $K_i \in \mathcal{K}$ we have $K_i \not\subset K$, because if $K_i \subset K$ then K is not a minimal set of the attribute d over s_d. Consequently, we can conclude that $T = \{K_1, K_2, \ldots, K_t\}$ is a Sperner system over C. According to the definition of closure of the set K on relation schema s_d we have $K^+ = K$. That is $K \to \{d\}$ is not true. This is in contradiction with the assumption that K is a minimal set of the attribute d on s_d. So $K \in \mathcal{K}$ where $\mathcal{K} = \{K_1, K_2, \ldots, K_t\}$.

From (1) and (2) we have $\mathcal{K} = \mathcal{K}_d^s - \{d\}$. □

It is known that problem A is co-NP-complete if and only if the problem negative A is NP-complete.

It is known that the below problem Subset Delimiter Complementarity (SDC) [4] is co-NP-complete:

Let T be a finite set and there are two families $P = \{P_1, P_2, \ldots, P_n\}$, $Q = \{Q_1, Q_2, \ldots, Q_m\}$ of subsets of T. The problem of determining that for any $A \subseteq T$ there exists P_i such that $P_i \subseteq A$ or Q_j such that $A \subseteq Q_j$ for any $i = 1, \ldots, n$, $j = 1, \ldots, m$ is co-NP-complete. The paper [4] shows that if Q is a Sperner system then the problem SDC is still co-NP-complete.

Now, we present a co-NP-complete problem concerning reducts of decision tables.

Definition 2.1. *Let $DS = (U, C \cup d)$ be a consistent decision table; we say that $A \subseteq C$ is a relative reduct of DS if A contains a reduct of DS. Let $s = \langle C \cup \{d\}, F \rangle$ be a relation schema; we say that $A \subseteq C$ is a relative minimal set of the attribute d if A contains a minimal set B of d, where $B \in \mathcal{K}_d^s - \{d\}$. Let Q_d be the family of all relative reducts of DS, and P_d be the family of all relative minimal sets of the attribute d on s.*

Theorem 2.1. *Let $DS = (U, C \cup d)$ be a consistent decision table and $s = \langle C \cup \{d\}, F \rangle$ be a relation schema. The $Q_d \subseteq P_d$ problem is co-NP-complete.*

Proof. For any $A \subseteq C \cup \{d\}$, we have a polynomial algorithm to check whether A is a relative reduct of DS or not. Based on the algorithm to find closure A_+ [1] and the definition of relative minimal set of the attribute d on relation schema s (Definition 2.1), it is easy to construct a polynomial algorithm to check whether A is a relative minimal set of d on s or not. Then we choose an arbitrary subset $A \subseteq C \cup \{d\}$ such that A is a relative reduct of DS but A is not a relative minimal set of d on s. Obviously, this algorithm is a polynomial nondeterministic algorithm. Consequently, our problem is co-NP. Now let us consider the problem in [4] for the set T and the family $P = \{P_1, P_2, \ldots, P_n\}$, $Q = \{Q_1, Q_2, \ldots, Q_m\}$, where Q is a Sperner system on T.

We denote $P^1 = \{P_i \in P \,|\, \nexists P_j \in P : P_i \subset P_j, 1 \le i \le n, 1 \le i \le n\}$.

It is clear that P^1 is the set of all maximal elements of P and P^1 is a Sperner system on T. We can calculate P^1 from P by a polynomial algorithm with the number of P and T. It is easy to see that $\{T, P^1, Q\}$ is an equivalent instant of $\{T, P, Q\}$. So, we can assume that P is a Sperner system on T. We can prove that the problem SDC is transformed to our problem by a polynomial algorithm.

Set $R = T \cup \{d\}$, $s = \langle R, F\rangle$, where $F = \{P_1 \to \{d\}, \ldots, P_n \to \{d\}\}$.

We construct the decision table $DS = (U, C \cup d)$ as follows: set $U = \{u_0, u_1, \ldots, u_m\}$.

– For any $c \in T$, set $c(u_0) = 0$ and $d(u_0) = 0$. – For any $c \in T$, $i = 1, \ldots, m$, set $c(u_i) = 0$ if $c \in Q_i$, conversely $c(u_i) = i$. Set $d(u_i) = i$

It is clear that s is constructed in polynomial time according to $|T|$ and $|P|$, $|Q|$.

According to Lemma 2.3 we have $P = \mathcal{K}_d^s - \{d\}$. By the definition of relative minimal set A of d on the relation schema s, there exists P_i such that $P_i \subseteq A$, where $1 \le i \le n$. According to Lemma 2.2 we have $Q = (\mathcal{K}_d^r)^{-1}$, so A is a relative reduct of DS if and only if for any $i = 1, \ldots, m$ we have $A \not\subset Q_i$.

If Q_d is the family of all relative reducts of DS and P_d is the family of all relative minimal sets of the attribute d on s, we have $Q_d \subseteq P_d$ if and only if for any $A \subseteq T$ and $i = 1, \ldots, n$, $A \not\subset Q_i$ there exists P_j such that $P_j \subseteq A$. Consequently, we can conclude that the problem SDC is tranformed to our problem by a polynomial algorithm. Theorem 2.1 is proved. □

Now, we prove that the $P_d \subseteq Q_d$ problem is polynomial time.

Definiton 2.2. *Let* $R = \{a_1, \ldots, a_n\}$. *Set* $P(R) = \{A : A \subseteq R\}$ *and* $I \subseteq P(R)$. *Then* I *is called a meet-semilattice if* $R \in I$ *and* $A, B \in I \Rightarrow A \cap B \in I$.

Suppose that $M \in P(R)$. Set $M^+ = \{\cap M^1 : M^1 \subseteq M\}$, then we say that M is generator of I if $M^+ = I$.

Note that $R \in M^+$ but R is not a element of M. We denote:

$N_I = \{A \in I : A \neq \cap \{A^1 \in I \wedge A \subset A^1\}\}$

In [3], J. Demetrovics shows that N_I is the only minimal generator of I. This means that for any generator N^1 of I we have $N_I \subseteq N^1$).

Lemma 2.4. *Let* $s = \langle R, F\rangle$ *be a relation schema. Set* $Z_s = \{A^+ : A \subseteq R\}$. N_s *is a minimal generator of* Z_s. *Set*

$$M_s = \{A : A \in N_s, d \notin A, B \in N_s : d \notin B \wedge A \subset B\}.$$

Then we have $(\mathcal{K}_d^r)^{-1} = M_s$.

P r o o f. According to the method to construct the decision table in Theorem 2.1 we have $E_{1i} = A_{i-1}$ where $2 \le i \le t+1$ and $E_{ij} = A_{1i} \cap A_{1j}$ where $2 \le i < j \le t+1$, so $E_{ij} = E_{1i} \cap E_{1j}$ or $E_{ij} \subset E_{1i}, E_{ij} \subset E_{1j}$ where

$2 \le i < j \le t+1$. Therefore, the set $M = \{E_{1i} : 1 \le i \le t+1\}$ has the property $\forall A \in M$: there is no $B \in M$ such that $A \subset B$. According to the definition of maximal equality system M_r over r, we have $M_s = M_r = \{E_{1i} : 2 \le i \le t+1\}$. Hence, $M_s = M_d = (K_d^s)^{-1} = \{A_1, A_2, \ldots, A_t\}$ (I). In next content, we prove $M_s = (K_d^r)^{-1}$ where K_d^r is the family of all minimal sets of the attribute d over r.

i) For $A \in M_s$ we have $A^+ = A$, and A does not contain d so A^+ does not contain d, hence $A \to \{d\} \notin F^+$. Moreover, if there is a B such that $A \subset B$, according to the method to calculate the closure of an attribute set over a relation we have $B^+ = R$ and B^+ contains d, or $B \to \{d\} \in F^+$. So $(K_d^r)^{-1} = \mathrm{MAX}\left(F_r^+, d\right)$ where

$$\mathrm{MAX}\left(F_r^+, d\right) = \left\{A \subseteq R : A \to \{d\} \notin F^+, A \subset B \Rightarrow B \to \{d\} \in F^+\right\}$$

so we conclude $A \in (K_d^r)^{-1}$.

ii) Conversely, if $A \in (K_d^r)^{-1}$ then obviously $A \neq R$. If there is a B such that $A \subset B$ and $A \to B$, then by the definition of antikeys we have $B \to d$ and $A \to d$. This is a contradiction. So there does not exist a B such that $A \subset B$ and $A \to B$, that is $A^+ = A$ holds. Moreover, also according to the definition of antikeys, if there exists $B^1 \neq R$ such that $A \subset B^1$, then $B^1 \to d$ or $d \subset B^{1+}$. Therefore, A is the maximal set which satisfies $A = A^+$ and A does not contain d (1). On the other hand, over the relation r constructed, for any $B \in M_s$ we have $B \neq R$, $B = B^+$ and B does not contain d. If there is a D such that $B \subset D$ then $D^+ = R$ or $d \subset D^+$. Therefore, M_r is the set of all maximal sets B which satisfies $B = B^+$ and B does not contain d (2). From (1) and (2) we can conclude $A \in M_s$.

From i) and ii) we obtain $M_s = (K_d^r)^{-1}$. □

Lemma 2.5. *Let $DS = (U, C \cup d)$ be a consistent decision table and let $s = \langle C \cup \{d\}, F\rangle$ be a relation schema. Suppose that $\mathcal{M}_d = \{M_1, \ldots, M_t\}$. Then we have $P_d \subseteq Q_d$ if and only if $M_i^+ \neq R$ for any $1 \le i \le t$.*

Proof. (1) Suppose that $M_i^+ \neq R$ for any $1 \le i \le t$. Set $(\mathcal{K}_d^s)^{-1} = \{A_1, \ldots, A_p\}$. According to the definition of $(\mathcal{K}_d^s)^{-1}$, if $B \in P_d$ then $B \not\subseteq A_j$ for any $1 \le j \le p$. According to Lemma 2.4 and for any $M_i \in \mathcal{M}_d$, $1 \le i \le t$ and $M_i^+ \neq R$, there exists A_j, $1 \le j \le p$ such that $M_i \subseteq A_j$. Then, $B \not\subseteq M_i$, $1 \le i \le t$. According to Lemma 2.1 $\mathcal{M}_d = (\mathcal{K}_d^r)^{-1}$ we have $B \in Q_d$.

(2) Conversely, suppose that $P_d \subseteq Q_d$. If $i, 1 \le i \le t$ and $M_i^+ = R$ then $M_i \in P_d$. Then $M_i \notin Q_d$. This is in contrast with $P_d \subseteq Q_d$. Consequently, we have $M_i^+ \neq R$ for any $1 \le i \le t$.

From (1) and (2), the lemma is proved. □

From Lemma 2.5 we have an algorithm to check whether $P_d \subseteq Q_d$.

Algorithm 2.1. Determine whether $P_d \subseteq Q_d$.

Input: The consistent decision table $DS = (U, C \cup d)$, the relation schema $s = \langle C \cup \{d\}, F\rangle$.

Output: Determine whether $P_d \subseteq Q_d$ or not.

Step 1: From DS calculate $\mathcal{M}_d = \{M_1, \ldots, M_t\}$;

Step 2: Calculate M_i^+ where $1 \leq i \leq t$. If $\forall i$, $M_i^+ \neq R$ then $P_d \subseteq Q_d$. Conversely, $P_d \not\subseteq Q_d$.

From Lemma 2.5 and the time to calculate closure, to calculate $\mathcal{M}_d$ we have Lemma 2.6:

Lemma 2.6. Let $DS = (U, C \cup d)$ be a consistent decision table and let $s = \langle C \cup \{d\}, F\rangle$ be a relation schema. The time complexity of Algorithm 2.1 is polynomial in the size of DS.

4. Conclusions. This paper presented the time complexity of algorithms concerning reducts of consistent decision tables. Let $DS = (U, C \cup d)$ be a consistent decision table and $s = \langle C \cup \{d\}, F\rangle$ be a relation schema on the attribute set $C \cup \{d\}$. Let Q_d be the family of all relative reducts of DS and P_d be the family of all relative minimal sets of the attribute d on s. The problem whether $Q_d \subseteq P_d$ is co-NP-complete and the problem whether $P_d \subseteq Q_d$ is polynomial time. Besides, this paper also proposes a polynomial time algorithm to check $P_d \subseteq Q_d$.

REFERENCES

[1] Demetrovics J., V. D. Thi. Some remarks on generating Armstrong and inferring functional dependencies relation. *Acta Cybernetica*, **12** (1995), 167–180.

[2] Demetrovics J., V. D. Thi, N. L. Giang. On Finding All Reducts of Consistent Decision Tables. *Cybernetics and Information Technologies*, **14** (2014), No 4, 3–10.

[3] Demetrovics J. Logical and structural study of relational data model. *MTA SZTAKI Tanulmányok*, Budapest, **114** (1980), 1–94. (in Hungarian)

[4] Gottlob G., L. Libkin. Investigation on Armstrong relations dependency inference and excluded functional dependencies. *Acta Cybernetica*, **9** (1990), No 4, 385–402.

[5] Giang N. L., V. D. Thi. Some Problems Concerning Condition Attributes and Reducts in Decision Tables. In: Proceeding of the Fifth National Symposium "Fundamental and Applied Information Technology Research" (FAIR), Dong Nai, Vietnam, 2011, 142–152.

[6] Pawlak Z. Rough sets: Theoretical Aspects of Reasoning About Data. Kluwer Academic Publishers, 1991.

Demetrovics Janos
Institute for Computer and Control (SZTAKI)
Hungarian Academy of Sciences
e-mail: `demetrovics@sztaki.mta.hu`

Vu Duc Thi
Institute of Information Technology
VNU, Vietnam
e-mail: `vdthi@vnu.edu.vn`

Nguyen Long Giang
Institute of Information Technology
VAST, Vietnam
e-mail: `nlgiang@ioit.ac.vn`

Tran Huy Duong
Institute of Information Technology
VAST, Vietnam
e-mail: `huyduong@ioit.ac.vn`

Received December 18, 2015
Final Accepted February 1, 2016