International Journal of Mathematical, Engineering and Management Sciences
Vol. 7, No. 2, 288-298, 2022
https://doi.org/10.33889/IJMEMS.2022.7.2.019

Some Problems Related to Reducts of Consistent Incomplete Decision Tables

Dinh Thu Khanh
Mathematic and Informatics Department,
Thang Long University, Hanoi, Vietnam.
E-mail: dinhthukhanh@gmail.com

Vu Duc Thi
Institute of Information Technology,
Vietnam Academy of Science and Technology, Hanoi, Vietnam.
E-mail: vdthi@vnu.edu.vn

Nguyen Long Giang
Institute of Information Technology,
Vietnam Academy of Science and Technology, Hanoi, Vietnam.
Corresponding author: nlgiang@ioit.ac.vn

Le Hoang Son
VNU Information Technology Institute,
Vietnam National University, Hanoi, Vietnam.
E-mail: sonlh@vnu.edu.vn

(Received on September 22, 2021; Accepted on February 11, 2022)

Abstract
In multi-criteria decision making, attribute reduction has attracted the attention of researchers for more than two decades. So far, numerous scientists have proposed algorithms to construct reducts in decision tables. However, most of the suggested algorithms are heuristic which discovers a reduction based on criteria of the attribute set. In fact, studying the properties of reducts to build efficient attribute reduction models is an urgent problem. In this research, we present some properties of reducts in incomplete decision tables by the relational database theory approach. It was found that the properties of reducts in incomplete decision tables are equivalent to properties of the Sperner-systems in the theory of relational database. By studying the properties of the Sperner-systems, the efficient attribute reduction models can be built to improve the efficiency of multi-criteria decision making systems.

Keywords- Sperner-system, Decision table, Incomplete decision table, Reduct.

1. Introduction

One of the most important problems in multi-criteria decision making systems is attribute reduction which attracted the consideration of researchers for more than two decades. The objectives of the attribute reduction process are removing unnecessary and redundant attributes and keep the attribute reduction set (called as reduct). Various methods have been introduced in order to obtain the reduct of decision tables based on Rough Set (RS) (Pawlak, 1991) or extended RS (Giang et al., 2019, 2021a; Thang et al., 2021; Ni et al., 2020; Shu et al., 2020; Zhang et al., 2020). However, these algorithms are heuristic. The best reduct is found by the best quality of classification on the attribute set. In fact, studying the properties of reduct to build efficient attribute reduction and data mining models is an urgent problem. However, results related to the attributes of reduct in decision tables are restricted.

For consistent complete decision tables (CCDTs), many researches were proposed (Thi & Giang, 2011, 2013; Demetrovics et al., 2014, 2015, 2018; Giang & Son, 2015; Giang et al., 2021a) related to the properties of reduct and inferring knowledge by relational database theory approach in recent years. As an example, Giang and Son (2015) proposed two algorithms in polynomial time: the first algorithm uses the relational database theory approach to find all reductive attributes; the second algorithm is also to find all reductive attributes by RS theory. Demetrovics et al. (2018) proposed some algorithms to decrease the quantity of objects in a CCDT by discovering some features of reduct and reductive attributes by the relational database theory approach. The proposed algorithms can be effectively applied in the data preprocessing phase to increase the performance of data mining and machine learning models. In Giang et al. (2021a, 2021b), the authors discovered some properties of reduct related to Sperner-system and state that the study of some properties on reduct is equivalent to the study of some properties on the Sperner-system. This result opens a new research direction on building efficient attribute reduction models based on research on Sperner-system.

In practical problems, decision tables often miss values in the attribute value domain, known as incomplete decision tables. On these tables, a tolerance relationship is constructed on the attribute value domain with a tolerance RS model (Kryszkiewicz, 1998). However, studies related to reduct properties of incomplete decision tables are still limited. By extending the results in the paper (Giang & Son, 2015), all reducts of consistent incomplete decision tables (CIDTs) were calculated by a new method in Demetrovics et al. (2013) with a polynomial time. Consequently, the motivation of the paper is to extend some results on the properties of reduct on incomplete decision tables to further research on building efficient attribute reduction models.

In this study, some properties of reduct in CIDTs based on the Sperner-system are proposed. The contributions of the paper include:

(1) Discover some properties of reduct in CIDTs and prove that all reducts are equivalent to a Sperner-system.
(2) Propose an algorithm to create an incomplete decision table based on a given Sperner-system.

These contributions let us to build efficient attribute reduction models by studying Sperner-system and generate data from a given Sperner-system to support machine learning algorithms in training and testing model. There are four sections in this paper. Section 2 describes the fundamental definitions related the RS theory and concepts related to relational database together with some combinational results in relational database. Section 3 presents the contributions of the paper on the study of properties of reduct obtained from CIDTs equivalent to Sperner-systems. The last section addresses the concluding and further research.

2. Background

Theory of RS and the relational tables are some basic definitions discovered in Pawlak (1991), Kryszkiewicz (1998), Demetrovics et al. (2013, 2014).

Definition 1. (Pawlak, 1991) The decision table consist four elements $DS=(U,C\cup D,V,f)$, in which $U=\{u_1,\dots,u_n\}$ is a set of objects; $C=\{c_1,...,c_n\}$ is a set of condition attribute. These two sets are finite and non-empty. D is a decision attribute set. Two sets, C and D, are separate. Let V_a is the value set of the attribute $a\in R=C\cup D$, then $V=\cup V_a$; $f:U\times R$ is the information function define as for $\forall u\in U;\ a\in R,\ f(u,a)\in V_a$.

Denote $a(x)$ is symbolized for the value a on item x for each $x\in U,\ a\in R$,. If $P=\{p_1,p_2,...,p_k\}\subseteq A$ is a part of attributes then $p_i(x)$ set is signified as $P(x)$. Consequently, if x,y in U are two objects if $p_i(x)=p_i(y)$ with $i=1,...,k$ then $P(x)=P(y)$.

Not losing the comprehensive characteristics, hypothesis D only has only one attribute d (D) can be reduced to one attribute by using an encryption (Thi & Giang, 2013). From above, our research focuses to decision tables, $DS=(U,R,V,f)$ in which $R=C\cup\{d\}$ and $\{d\}\notin C$.

Definition 2. (Pawlak, 1991) Given $DS=(U,C\cup D,V,f)$. Each subset in C, $E\subseteq C$, defines an indistinguishable relation, called equivalence relation, as follows:

$$IND(E)=\{(x,y)\in U^2 \mid \forall a\in E, f(x,a)=f(y,a)\} \tag{1}$$

The partition of U is signified by $U/E=\{E_1,E_2,...,E_m\}$. One element in U/E is named an equivalence class.

For each $P\subseteq C$ and $Y\subseteq U$, we have:

P-upper approximation of Y : $\overline{P}Y=\{y\in U \mid [y]_P\cap Y\neq\emptyset\}$

P-lower approximation of Y : $\underline{P}Y=\{y\in U \mid [y]_P\subseteq Y\}$

P-boundary region of Y : $\overline{P}Y/\underline{P}Y$

P-positive region of $\{d\}$: $POS_P(\{d\})=\cup_{Y\in U/D}(\underline{P}Y)$

If $POS_P(\{d\})=U$ or functional dependency $P\to d$ is true then DS is consistent which is for any x and y belong to U, if $P(x)=P(y)$ then $d(x)=d(y)$. In opposite, DS is inconsistent decision table.

Definition 3. (Pawlak, 1991) Let $DS=(U,R,V,f)$ be a decision table and $R=C\cup\{d\}$. If $P\subseteq C$ satisfies following conditions:

(1) $POS_C(\{d\})=POS_P(\{d\})$

(2) $\forall P'\subset P: POS_C(\{d\})\neq POS_{P'}(\{d\})$

then P is a reduct of C

According to Definition 3, in a consistent decision table, P is a reduct of DS if functional dependency $P\to d$ is true, $\forall P'\subset P,\ P'$ ½ $\{d\}$. According to Pawlak (1991), the set P in Definition 1 is named reduct based on a positive region. This reduct is termed Pawlak reduct. The set of entire reducts of DS is denoted as $ARED(C)$.

Definition 4. (Kryszkiewicz, 1998) $r = p_1, \dots, p_m$ is a relation on $R = \{a_1, ..., a_n\}$ if $\forall a_i \in R$ has D_{a_i} and $* \in D_{a_i}$ and $p_j : R \to \cup D_{a_i}$ we have $p_j(a_i) \in D_{a_i}$

Definition 5. (Kryszkiewicz, 1998) Let r be a relation on $R = \{a_1, ..., a_n\}$ and A is a subset of R. Then, $p_i \sim p_j(A)$ if each a belongs to A: $p_i(a) = p_j(a)$ or $p_i(a) = *$ or $p_j(a) = *$.

Definition 6. Let $r = \{p_1, \dots, p_m\}$ on $R = \{a_1, ..., a_n\}$. Then, $X, Y \subseteq R$ and X tolerance determines Y, denoted by $X \xrightarrow{t} Y$, if $(\forall\, p_i, p_j \in r)$ (if $p_i \sim p_j(X)$ then $p_i \sim p_j(Y)$)). Set $T_r = \{(X,Y) : X, Y \subseteq R$ and $X \xrightarrow{t} Y\}$. It is easy to check that:

(1) $(X, X) \in T_r \ \forall X \subseteq R$

(2) $(X,Y) \in T_r$ then $X \subseteq Z, W \subseteq Y$ has $(Z, W) \in T_r$

(3) $(X,Y) \in T_r, (Y, Z) \in T_r \Rightarrow (X, Z) \in T_r$

Set $X^{+} = \{x \in R : X \xrightarrow{t} \{x\}\}$

Definition 7. (Kryszkiewicz, 1998) Given $DS = (U, R, V, f)$, $R = C \cup \{d\}$. If $\exists a \in C$ and there is at least one missing value, denoted as '*', in V_a then DS is called as an incomplete decision table. This decision table is presented as $IDT = (U, R, V, f)$.

According to Definition 4, IDT is consistent if $C \xrightarrow{t} \{d\}$. We can see that if IDT is inconsistent, we can check by using a polynomial time algorithm on elements of U to eliminate the elements, making IDT consistently. After the elimination, we have the set U' such that $DS = (U', R, V, f)$ is consistent.

Definition 8. (Kryszkiewicz, 1998) Given $IDT = (U, R, V, f)$ with $R = C \cup \{d\}$ and a reduct $P \subseteq C$. If $P \xrightarrow{t} \{d\}$ and $\forall P'$ Ö P then $P' \overset{t}{\nrightarrow} \{d\}$. It means that P' is a proper subset of P such that P' does not tolerance determine d. Set $IARED(C) = \{P : P$ is reduct of $IDT\}$

Definition 9. (Demetrovics et al., 2013) Let $R = \{a_1, \dots, a_n\}$. $K = \{A_1, \dots, A_m\}$ is a Sperner - system on R if $A_i \not\subseteq A_j \ \forall i, j = 1..m$.

Definition 10. (Demetrovics et al., 2013) Assume that $K = \{A_1, \dots, A_m\}$ is a Sperner-system on R. Set $K^{-1} = \{B \subsetneq R: (A \in K \Rightarrow A \not\subseteq B$ and $B \subsetneq C\}$. K^{-1} is named the antikey of K. Obviously, K^{-1} is also a Sperner - system on R. We have $K^{-1} \subset R$ because K is a Sperner-system on R. K^{-1} is maximal set that does not include the elements of K.

Definition 11. (Demetrovics et al., 2014) Provided that $IDT = (U, R, V, f)$ is a CIDT and $R = C \cup \{d\}$, set $r = \{u_1, \dots, u_m\}$, then a Sperner-system is defined as $ARED(C) = K_d^t = \{X \subseteq C : X \xrightarrow{t} \{d\}$ and $\nexists Y : Y \xrightarrow{t} \{d\}$ and $Y \subsetneq X\}$.

Definition 12. Given $IDT = (U, R, V, f)$ and $R = C \cup \{d\}$. Let $r = U = \{u_1, \dots, u_m\}$. The equivalent set ε_r is determined by

$\varepsilon_r = \{M_{ij} : i \geq 1 \text{ and } i \leq j \leq m\}$ with $M_{ij} = \{x \in R : x(u_i) = x(u_j) \text{ or } x(u_i) = * \text{ or } x(u_j) = *\}$ (2)

The set E_d is defined by $E_d = \{A \in \varepsilon_r : A \neq R,\ d \notin X \text{ and } \nexists B \in \varepsilon_r : d \notin B \text{ and } A \subsetneq B\}$ (3)

Next, some combinational results related to the relational database which can be found in Demetrovics et al. (2018) are described following.

Algorithm 1. (Demetrovics et al., 2018) Find K^{-1} in Sperner-system K.

Input: $K = \{P_1, ..., P_n\}$ is a Sperner system over A.
Output: K^{-1}
Step 1: Set $K_1 = \{R - \{a\} : a \in P_1\}$. It is obviously that $K_1 = \{P_1\}^{-1}$
Step s+1: (*s*<*n*): Let $K_s = F_s \cup \{X_1, ..., X_{ts}\}$, where $X_1, ..., X_{ts}$ are elements of K_s containing P_{s+1} and $F_s = \{A \in K_s : P_{s+1} \not\subseteq A\}$. For $\forall\, i$ $(i = 1, ..., t_s)$, we compute $\{P_{s+1}\}^{-1}$ on X_i in an analogous way as K_1, that are the subsets maximization of X_i not including P_{s+1}. In detail, $\{P_{s+1}\}^{-1} = \{P_{s+1} - \{a\} : a \in X_i\}$. We signify them by $A_1^i, ..., A_{ri}^i$. Let: $K_{s+1} = F_s \cup \{A_s^i : A \in F_s \Rightarrow A_q^i \not\subset A, i >= 1 \text{ and } i \leq t_q, 1 \leq q \leq r\}$ Lastly, set $K^{-1} = K_n$.

Theorem 1. (Demetrovics et al., 2018) For $s(1 \leq s \leq n)$ $K_s = \{P_1, ..., P_s\}^{-1}$, that is $K_n = K^{-1}$. Obliviously K, K^{-1} are unique and it is drawn the definition of K^{-1} that the Algorithm 1 is independent of the order of $P_1, ..., P_n$. The number of elements in K_s is denoted that l_s $(1 \leq s \leq n-1)$. Then, $K_s = F_s \cup \{X_1, ..., X_{ts}\}$.

Proposition 1. (Demetrovics et al., 2018) Algorithm 1 has computation complexity as $O\left(|R|^2 \sum_{s=1}^{n-1} t_s u_s\right)$ in the worst case, where $u_s = I_s - t_s$ if $I_s > t_s$ and $u_s = 1$ if $I_s = t_s$.

3. Some New Results in Incomplete Decision Tables

Herein, some theorems related to reduct of CIDTs are introduced. We discover the equivalence properties about the reducts obtained from a CIDT with Sperner-systems. Then, an algorithm used to construct a CIDT is proposed.

Theorem 2. Suppose that $IDT = (U, R, V, f)$ is a CIDT and $R = C \cup \{d\}$, set $r = U = \{u_1, \dots, u_m\}$

From r we calculate the equivalence set:

$\varepsilon_r = \{M_{ij} :\ i, j \in \{1, 2, ..., m\},\ i \leq j\}$ where $M_{ij} = \{x \in R : x(u_i) = x(u_j) \text{ or } x(u_i) = * \text{ or } x(u_j) = *\}$

From ε_R, we define

$E_d = \{X \in \varepsilon_r : X \neq R, d \notin X \text{ and } \nexists\, Y \in \varepsilon_r : d \notin Y \text{ and } X \subsetneq Y\}$ (4)

$K_d^t = \{X \subseteq Z : X \xrightarrow{t} \{d\} \text{ and } \nexists Y : Y \xrightarrow{t} \{d\} \text{ and } Y \subsetneq X\}$ (5)

Then we have $E_d = (K_d^t)^{-1}$.

Proof: If $\forall X \in E_d$ we can see that $X = X^+$. Because if $X \subsetneq X^+$, there is an element e that $e \notin X$ but $e \in X^+$. Because X is the equivalent maximum set then $\exists i, j\ (1 \le i < j \le m)$ that $M_{ij} = X$. Based on the definition of set X^+, we have $X \xrightarrow{t} \{e\}$ and from the definition of set M_{ij} then $e \in M_{ij}$. Therefore, if $X = X^+$ and $d \notin X$ then $d \notin X^+$. This leads to $X \overset{t}{\nrightarrow} \{d\}$ (X does not tolerance determine d). Assumption that we have Y with $X \subsetneq Y$. From the definition of X, if $d \notin Y$ then $h_i \sim h_j(Y)$ is incorrect $\forall i,\ j \in \{1, 2, ..., m\}$. Moreover, according to the definition of tolerance determination, we get $Y \xrightarrow{t} R$. In the case of $d \in Y$, it is clear to see that $d \in Y^+$. Therefore, in both cases, we have $\forall Y$: $X \subsetneq Y \Rightarrow Y \xrightarrow{t} \{d\}$. Then, according to the definition of K_d^t then $Z \in K_d^t$ for $Z \subseteq Y$. Using the definition of set $(K_d^t)^{-1}$ we have $X \in (K_d^t)^{-1}$.

Oppositely, if $X \in (K_d^t)^{-1}$ then $A^+ = A$. Because if $X \subsetneq X^+$ then based on the definition of antikey set we have $Z \in K_d^t$ for $Z \subseteq X^+$, means $X^+ \xrightarrow{t} \{d\}$, leading to $X \xrightarrow{t} \{d\}$. According to the definition of $(K_d^t)^{-1}$ then A does not tolerance determine $\{d\}$ ($X \overset{t}{\nrightarrow} \{d\}$). Thus $X^+ = X$. Based on the definition of set E_d and set $(K_d^t)^{-1}$(is the set of largest sets do not tolerance determined d), means $X \in E_d$. Therefore $E_d = (K_d^t)^{-1}$.

Remark 1. It is easy to see that we calculate sets M_{ij} has polynomial time complexity with m and $|C|$. From it the finding set E_d has polynomial time complexity with m and $|C|$, too.

Example 1. Assume that $IDT = (U, R, V, f)$ is a CIDT as Table 1 where $U = \{u_1, ..., u_5\}$, $C = \{b_1, ..., b_4\}$ = {Price, Mileage, Size, Max-speed}; d is decision attribute and $R = C \cup \{d\}$.

Table 1. An incomplete decision table of car choices.

	b_1	b_2	b_3	b_4	d
u_1	H	H	*	*	P
u_2	L	*	F	L	G
u_3	L	L	C	H	P
u_4	M	H	C	H	G
$_5$	M	H	C	*	G

Note. C: Compact; F: Full; G: Good; H: High; L: Low; M: Medium; P: Poor. *Missing value.

We calculate E_d. It can be seen that: $M_{12}=\{b_2,b_3,b_4\}$, $M_{13}=\{b_3,b_4,d\}$, $M_{14}=\{b_2,b_3,b_4\}$, $M_{15}=\{b_2,b_3,b_4\}$, $M_{23}=\{b_1,b_2\}$, $M_{24}=\{b_2,d\}$, $M_{25}=\{b_2,b_4,d\}$, $M_{34}=\{b_3,b_4\}$, $M_{35}=\{b_3,b_4\}$, $M_{45}=\{b_1,b_2,b_3,b_4,d\}$.

According to the definition of E_d, there are $A_1=\{b_2,b_3,b_4\}$ and $A_2=\{b_1,b_2\}$ which satisfy the condition of E_d. Thus $E_d=\{\{b_2,b_3,b_4\},\{b_1,b_2\}\}$.

Now, the results correlated to the family of reducts equivalent properties in Sperner-systems are presented as following. First, some related results are proven below.
From the above results, the equivalence in family of reducts in incomplete consistent decision tables on Sperner-systems are proposed.

Theorem 3. Let $IDT=(U,R,V,f)$ be a CIDT with $R=C\cup\{d\}$, then $IARED(C)$ is a Sperner-system on C. Vice versa, if K is a Sperner-system on C, there is an incomplete consistent decision table $IDT=(U,R,V,f)$ satisfying $K=IARED(C)$.

Proof. Given $IDT=(U,R,V,f)$ and $R=C\cup\{d\}$. From reduct definition, $IARED(C)$ is a Sperner-system on C. Assume that $K=\{A_1,...,A_n\}$ is a Sperner-system on C. As shown in Algorithm 1, the antikey K^{-1} is constructed from K. Assume that $K^{-1}=\{B_1,...,B_n\}$. A consistent incomplete decision table $IDT=(U,R,V,f)$ is built as follow:
Suppose that $K^{-1}=\{A_1,...,A_t\}$, set $U=\{u_0,u_1,...,u_t\}$
Set $c(u_0)=0$, $\forall c\in C$ and $d(u_0)=0$

If $G=A_1\cap...\cap A_t$ then each g belongs to G and each i with $i=1,...m$ set $g(u_i)=*$.
For each i with $i=1,...t$ set $d(u_i)=i$.
For each c belongs to A_1 we set $c(u_i)=*$ and each c does not belong to A_1 set $c(u_1)=1$
For any $i=2,...t$ we set $d(u_i)=i$ and $c(u_i)=0$ if $c\in A_i$, $c(u_i)=i$ if $c\notin A_i$,
It is easy to see that for $i=2,...t$ we have $M_{1i}=A_1$.
From Algorithm 1, Theorem 2 and Definition 10 of E_d, we have $E_d=K^{-1}$.

Based on Sperner-system, antikey set and the reduct definition, we have $K^{-1}=IARED(C)$. Thus, the theorem has been proved.

As the result, the research on the family of reducts of $IDT=(U,R,V,f)$ is equivalent to the research of Sperner-systems on C with $R=C\cup\{d\}$. From the results in Theorem 2, we propose the flowchart of Algorithm 2 as follows.

Algorithm 2. Constructing a CIDT from a given Sperner-system K over C.
Figure 1 is the flowchart of Algorithm 2.

Input: Let $K=\{P_1,...,P_m\}$ be a Sperner-system on C.
Output: $IDT=(U,R,V,f)$ with $R=C\cup\{d\}$ and $K=IARED(C)$.
Step 1: From K, by using Algorithm 1 in Section 3 we construct K^{-1}; *Step 2:* Provided that $K^{-1}=\{A_1,...,A_t\}$, $U=\{u_0,u_1,...,u_t\}$ was set; For any $c\in C$, initiate $c(u_0)=0$ and $d(u_0)=0$; If $G=A_1\cap...\cap A_t$ then each g belongs to G and each i with $i=1,...m$ set $g(u_i)=*$; Set $d(u_i)=i$ $(i=\overline{1,t})$; For each c belongs to A_1 we set $c(u_1)=*$ and each c does not belong to A_1 set $c(u_1)=1$; For any $i=2,...t$ we set $d(u_i)=i$ and $c(u_i)=\begin{cases}0 & \text{if } c\in A_i\\ i & \text{if } c\notin A_i\end{cases}$ It is easy to see that for $i=2,...t$ we have $M_{1i}=A_1$;

The following corollary is obtained from Theorem 2, Theorem 3.

Next, computation complexity of Algorithm 2 is calculated. At the first step, we use Algorithm 1 to define K^{-1}. Consequently, Algorithm 2 and Algorithm 1 has the same complexity. Thus, computation complexity of Algorithm 2 is exponential in *m*.

Corollary 1. Supposed that $IDT=(U,R,V,f)$ with $R=C\cup\{d\}$ is a CIDT, then number of elements in $IARED(C)$ is less than $C_n^{[n/2]}$, $n=|C|$.

Example 2. Let $K=\{P_1,P_2,P_3,P_4\}$ be Sperner-system on $R=\{b_1,b_2,b_3,b_4,b_5,b_6,b_7\}$, $P_1=\{b_1,b_2,b_3\}$, $P_2=\{b_2,b_3,b_5\}$, $P_3=\{b_2,b_3,b_5,b_7\}$, $P_4=\{b_2,b_4,b_6,b_7\}$. We find K^{-1}. According to Algorithm 1, we have $K^{-1}=\{P_{12},P_{33},P_{34},P_{42},P_{43},P_{44}\}$, in which $P_{12}=\{b_1,b_3,b_4,b_5,b_6,b_7\}$, $P_{33}=\{b_2,b_3,b_6,b_7\}$, $P_{34}=\{b_2,b_3,b_5,b_6\}$, $P_{43}=\{b_1,b_2,b_4,b_5,b_7\}$, $P_{44}=\{b_1,b_2,b_4,b_5,b_6\}$.

According to Algorithm 2 we have the following CIDT as Table 2:

Table 2. A consistent incomplete table.

b_1	b_2	b_3	b_4	b_5	b_6	b_7	*d*
0	0	0	0	0	0	0	0
*	1	*	*	*	*	*	1
2	0	0	2	2	0	0	2
3	0	0	3	0	0	3	3
0	0	4	4	0	0	0	4
0	0	5	0	0	5	0	5
0	0	6	0	0	0	6	6

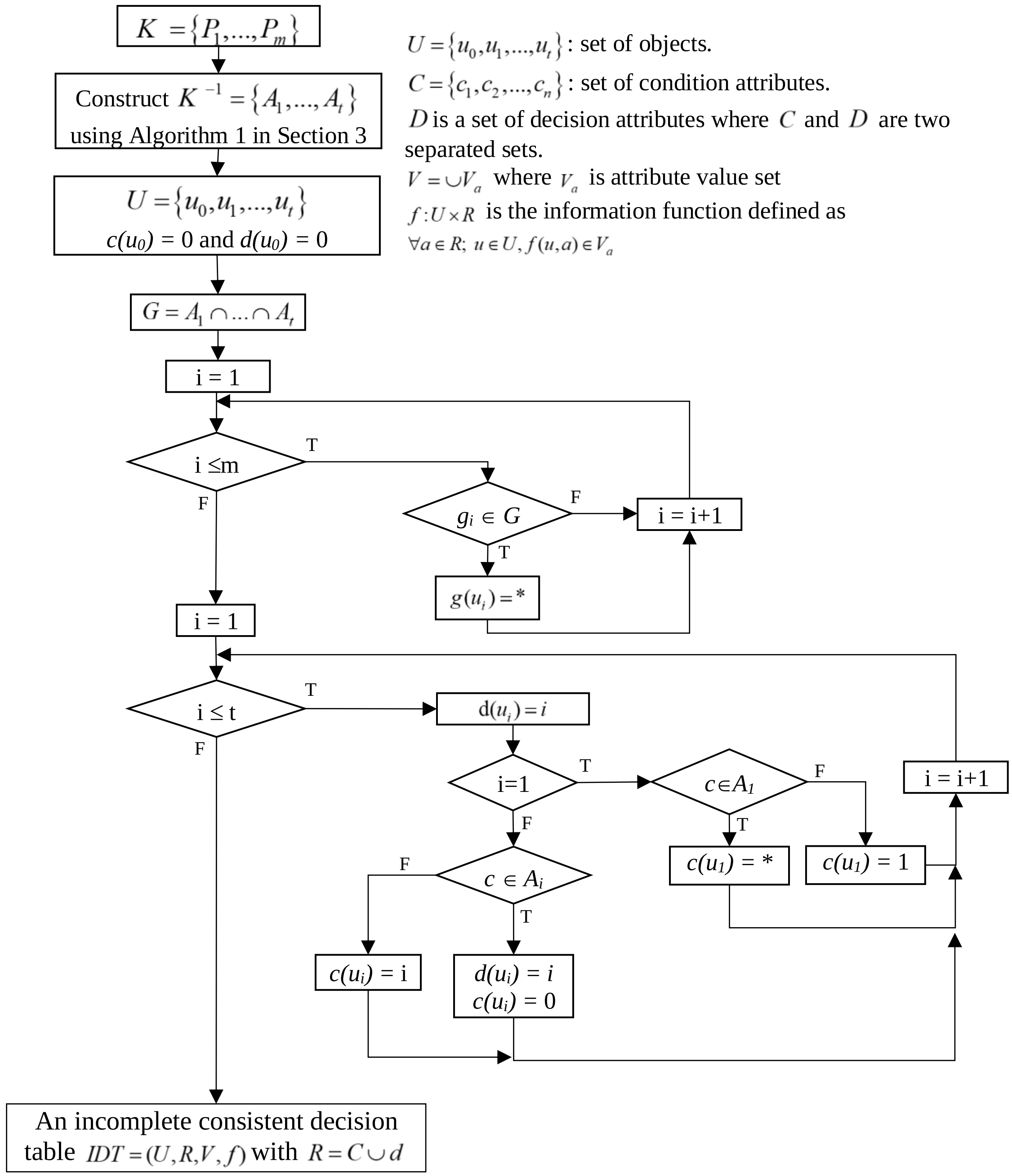

Figure 1. Flowchart of Algorithm 2.

Remark 2. It can be seen that if we change the order of elements of $K^{-1} = \{A_1, ..., A_t\}$, then we obtain some other CIDTs $IDT = (U, R, V, f)$ with $R = C \cup d$ such that $K = IARED(C)$.

Example 3. Based on Example 2, K^{-1} is built by swap the order of P_{12} and P_{33}. According to Algorithm 2, we have Table 3 as follows:

Table 3. A consistent incomplete table for Example 3.

b_1	b_2	b_3	b_4	b_5	b_6	b_7	d
0	0	0	0	0	0	0	0
1	*	*	1	1	*	*	1
0	2	0	0	0	0	0	2
3	0	0	3	0	0	3	3
0	0	4	4	0	0	0	4
0	0	5	0	0	5	0	5
0	0	6	0	0	0	6	6

4. Conclusions

With the continuous growth of the current data volume, it is urgent to research and propose attribute reduction algorithms to increase the effectiveness of multi-criteria decision making models. In this research, some properties of reducts in CIDTs are introduced. We also proved that all those reducts is equivalent to given Sperner-system. Therefore, a novel algorithm in order to build an incomplete decision table based on this Sperner-system is proposed. These results let us build efficient attribute reduction models by studying Sperner-system and generate data from given Sperner-system to support multi-criteria decision making systems in training and testing model. Further research is to study more properties on Sperner-system and reduct to propose efficient attribute reduction and multi-criteria decision making models.

Conflict of Interest

The authors declare that they have no conflict of interest regarding the publication of this article. All authors have checked and agreed to the submission.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors would like to thank the editor and anonymous reviewers for their comments that help improve the quality of this work.

References

Demetrovics, J., Quang, H. M., Thi, V. D., & Anh, N. V. (2018). An efficient nethod to reduce the size of consistent decision tables. *Acta Cybernetica*, *23*(4), 1039–1054.

Demetrovics, J., Thi, V. D., & Giang, N. L. (2013). An efficient algorithm for determining the set of all reductive attributes in incomplete decision tables. *Cybernetics and Information Technologies*, *13*(4), 118–126.

Demetrovics, J., Thi, V. D., & Giang, N. L. (2014). On finding all reducts of consistent decision tables. *Cybernetics and Information Technologies*, *14*(4), 3–10. doi: https://doi.org/10.1515/cait-2014-0001.

Demetrovics, J., Thi, V. D., Giang, N. L., & Duong, T. H. (2015). On the time complexity of the problem related to reducts of consistent decision tables. *Serdica Journal of Computing*, *9*(2), 167–176.

Giang, N. L., Janos, D., Thi, V. D., & Khoa, P. D. (2021a). Some properties related to reduct of consistent decision systems. *Cybernetics and Information Technologies*, *21*(2), 3–9. doi: https://doi.org/10.2478/cait-2021-0015.

Giang, N. L., Ngan, T. T., Tuan, T. M., Phuong, H. T., Abdel-Basset, M., de Macêdo, A. R. L., & de Albuquerque, V. H. C. (2019). Novel incremental algorithms for attribute reduction from dynamic decision tables using hybrid filter–wrapper with fuzzy partition distance. *IEEE Transactions on Fuzzy Systems*, *28*(5), 858–873.

Giang, N. L., Tuan, N. A., Ngan, T. T., Son, N. N., & Thang, N. T. (2021b). Filter-wrapper incremental algorithms for finding reduct in incomplete decision systems when adding and deleting an attribute set. *International Journal of Data Warehousing and Mining (IJDWM)*, *17*(2), 39–62.

Giang, N. L., & Son, N. H. (2015). Searching for reductive attributes in decision tables. In *Transactions on rough sets XIX* (pp. 51–64).Berlin, Heidelberg: Springer.

Kryszkiewicz, M. (1998). Rough set approach to incomplete information systems. *Information Sciences*, *112*(1–4), 39–49. doi: https://doi.org/10.1016/S0020-0255(98)10019-1.

Ni, P., Zhao, S., Wang, X., Chen, H., Li, C., & Tsang, E. C. (2020). Incremental feature selection based on fuzzy rough sets. *Information Sciences*, *536*, 185–204.

Pawlak, Z. (1991). *Rough sets: Theoretical aspects of reasoning about data* (Vol. 9). Springer Science & Business Media.

Shu, W., Qian, W., & Xie, Y. (2020). Incremental feature selection for dynamic hybrid data using neighborhood rough set. *Knowledge-Based Systems*, *194*, 105516.

Thang, N. T., Giang, N. L., Long, H. V., Tuan, N. A., Tran, T. M., & Tan, N. D. (2021). Efficient algorithms for dynamic incomplete decision systems. *International Journal of Data Warehousing and Mining (IJDWM)*, *17*(3), 44–67.

Thi, V. D., & Giang, N. L. (2011). A method to construct decision table from relation scheme. *Cybernetics and Information Technologies*, *11*(3), 32–41

Thi, V. D., & Giang, N. L. (2013). A method for extracting knowledge from decision tables in terms of functional dependencies. *Cybernetics and Information Technologies*, *13*(1), 73–82. doi: https://doi.org/10.2478/cait-2013-0007.

Zhang, C., Dai, J., & Chen, J. (2020). Knowledge granularity based incremental attribute reduction for incomplete decision systems. *International Journal of Machine Learning and Cybernetics*, *11*(5), 1141–1157. doi: https://doi.org/10.1007/s13042-020-01089-4.