Minimal keys and antikeys

By V. D. Thi

§ 1. Introduction

The relational model, defined by E. F. Codd [3] is one of the most investigated data base models of the last years. Many papers have appeared concerning combinatorial characterization of functional dependencies, systems of minimal keys and antikeys. A set of minimal keys and a set of antikeys form Sperner-systems. Sperner-systems and sets of minimal keys are equivalent in the sense that for an arbitrary Sperner-system S a family of functional dependencies F can be constructed so that the minimal keys of F are exactly the elements of S (cf. [4]).

In the present paper we propose some combinational algorithms to determine antikeys and minimal keys. In the second part of the paper, we are going to study connections between minimal keys and antikeys for special Sperner-systems.

We start with some necessary definitions.

Definition 1.1. Let Ω be a finite set, and denote $P(\Omega)$ its power set. The mapping F: $P(\Omega) \to P(\Omega)$ is called a closure operation over Ω if, for every $A, B \subseteq \Omega$,

(1) $A \subseteq F(A)$ (extensivity),

(2) $A \subseteq B$ implies $F(A) \subseteq F(B)$ (monotonity),

(3) $F(A) = F(F(A))$ (idempotency).

In few cases Ω is represented by the set $\{1, \ldots, n\}$ or by the set of columns of an $m \times n$ matrix $\mathbf{M}$. If we use the second representation, a special closure operation $F_{\mathbf{M}}$ can be defined over the set of the columns of $\mathbf{M}$:

The i-th column of $\mathbf{M}$ belongs to $F_{\mathbf{M}}(A)$ if and only if for any two rows of $\mathbf{M}$ which are identical on A they are equal on the i-th column, too.

It is easy to see, that $F_{\mathbf{M}}(A)$ is a closure operation. It is known (see [1]) that any closure operation F over a finite set Ω can be represented by an appropriate matrix $\mathbf{M}$, that is we can choose $\mathbf{M}$ and represent Ω by the set of the columns of $\mathbf{M}$ so that F coincides with $F_{\mathbf{M}}$.

Definition 1.2. Let F be a closure operation over Ω, and $A \subseteq \Omega$. We say that

— A is a key of F, if $F(A) = \Omega$.

— A is a minimal key of F, if A is a key of F and for any $B \subseteq A$, $F(B) = \Omega$ implies $B = A$, i.e. no proper subset of A is a key of F.

Let us denote by K_F the set of all minimal keys of F. It is clear that K_F forms a Sperner-system.

If K is a Sperner-system over Ω, let us define $S(K)$ as $S(K)=\min\{m\colon K=K_{F_M}\colon M$ is an $m\times n$ matrix representation of $\Omega\}$. For a Sperner-system K, we can define the set of antikeys, denoted by K^{-1}, as follows:

$$K^{-1}=\{A\subset\Omega\colon (B\in K)\Rightarrow(B\nsubseteq A) \quad\text{and}\quad (A\subset C)\Rightarrow(\exists B\in K)(B\subseteq C)\}.$$

It is easy to see that K^{-1} is the set of subsets of Ω, which does not contain the elements of K and which is maximal for this property. They are the maximal non-keys. Clearly, K^{-1} is also a Sperner-system.

In this paper we assume that Sperner-systems playing the role of the set of minimal keys (antikeys) are not empty (do not contain the full set Ω).

§ 2. Connection between minimal keys and antikeys

The following important result was proved in [1], [5]:

Remark 2.1. If K is an arbitrary Sperner-system, then there exists a closure operation F, for which $K=K_F$ and a closure operation F', for which $K=K_{F'}^{-1}$.

Let us given an arbitrary Sperner-system $K=\{B_1, \ldots, B_m\}$ over Ω. We are now going to construct the set of antikeys K^{-1}. Let us follow the algorithm described below:

Let $K_1=\{\Omega\setminus\{a\}\colon a\in B_1\}$. It is easy to see that $K_1=\{B_1\}^{-1}$.

Let us suppose that we have constructed $K_q=\{B_1, \ldots, B_q\}^{-1}$ for $q<m$. We assume that $X_1, \ldots, X_p$ are the elements of K_q containing B_{q+1}. So $K_q=F_q\cup\{X_1, \ldots, X_p\}$, where $F_q=\{A\in K_q\colon B_{q+1}\nsubseteq A\}$. For all i $(i=1, \ldots, p)$, we construct the antikeys of $\{B_{i+1}\}$ on X_i in the analogous way as K_1, which are the maximal subsets of X_i not containing B_{q+1}. We denote them by $A_1^i, \ldots, A_{\tau_i}^i$ $(i=1, \ldots, p)$.

Let

$$K_{q+1}=F_q\cup\{A_t^i\colon A\in F_q\Rightarrow A_t^i\not\subset A,\ 1\leq i\leq p,\ 1\leq t\leq\tau_i\}.$$

We have to prove, that $K_{q+1}=\{B_1, \ldots, B_{q+1}\}^{-1}$. For this using the inductive hypothesis $K_q=\{B_1, \ldots, B_q\}^{-1}$ we show that

a) if $A\in K_{q+1}$ then A is the subset of Ω not containing B_t $(t=1, \ldots, q+1)$ and being maximal for this property, i.e. $A\in\{B_1, \ldots, B_{q+1}\}^{-1}$,

b) every $A\subseteq\Omega$ not containing the elements B_t $(t=1, \ldots, q+1)$ and being maximal for this property is an element of K_{q+1}. First we prove the validity of (a). Let $A\in K_{q+1}$. If $A\in F_q$ then A does not contain the elements B_t $(t=1, \ldots, q)$ and A is maximal for this property and at the same time $B_{q+1}\nsubseteq A$. Consequently, A is a maximal subset of Ω not containing B_t $(t=1, \ldots, q+1)$.

Let $A\in K_{q+1}\setminus F_q$. It is clear that there is an A_t^i $(1\leq i\leq p$ and $1\leq t\leq\tau_i)$ such that $A=A_t^i$. Our construction shows that $B_l\nsubseteq A_t^i$ for all l $(l=1, \ldots, q+1)$. Because A_t^i is an antikey of $\{B_{q+1}\}$ for X_i we obtain $A_t^i=X_i\setminus\{b\}$ for some $b\in B_{q+1}$. It is obvious that $B_{q+1}\subseteq A_t^i\cup\{b\}$. If $a\in\Omega\setminus X_i$ then, by the inductive hypothesis, for $A_t^i\cup\{a,b\}=X_i\cup\{a\}$ there exists B_s $(s=1, \ldots, q)$ such that $B_s\subseteq A_t^i\cup\{a,b\}$. X_t does not contain $B_1, \ldots, B_q$ by $X_i\in K_q$. Hence $a\in B_s$. If $B_s\setminus\{a\}\subseteq A_t^i$ then $B_s\subseteq A_t^i\cup\{a\}$. For every B_s $(1\leq s\leq q)$ with $B_s\subseteq X_i\cup\{a\}$ and $B_s\nsubseteq A_t^i$ we have $b\in B_s$. Hence $B_s\setminus\{a,b\}\subseteq A_t^i$. Consequently, there exists an $A_1\in F_q$ such that

$A_t^i \subset A_1$. This contradicts $A \in K_{q+1} \setminus F_q$. So there is a B_s $(1 \leqq s \leqq q)$ such that $B_s \subseteqq A_t^i \cup \{a\}$.

Next we turn to the proof of (b). Suppose that A is the maximal subset of Ω not containing B_t $(1 \leqq t \leqq q+1)$. By the inductive hypothesis, there is a $Y \in K_q$ such that $A \subseteqq Y$.

The first case: If $B_{q+1} \nsubseteqq Y$ then Y does not contain $B_1, \ldots, B_{q+1}$. Because A is the maximal subset of Ω not containing B_t $(1 \leqq t \leqq q+1)$ we obtain $A = Y$. $B_{q+1} \nsubseteqq Y$ implies $A \in F_q$. Consequently, we have $A \in K_{q+1}$.

The second case: If $B_{q+1} \subseteqq Y$ then $Y = X_i$ holds for some i in $\{1, \ldots, p\}$ and $A \subseteqq A_t^i$ holds for some t in $\{1, \ldots, \tau_i\}$. If there exists an $A_1 \in F_q$ such that $A_t^i \subset A_1$, then we also have $A \subset A_1$. By the definition of F_q it is clear that A_1 does not contain $B_1, \ldots, B_{q+1}$. This contradicts the definition of A. Hence $A_t^i \in K_{q+1}$. It is easy to see that A_t^i does not contain $B_1, \ldots, B_{q+1}$. By the definition of A we obtain $A = A_t^i$, i.e. $K_{q+1} = \{B_1, \ldots, B_{q+1}\}^{-1}$.

By the above proof it is clear that $K_m = \{B_1, \ldots, B_m\}^{-1}$. Thus we have

Theorem 2.2. $K_m = K^{-1}$.

Because K and K^{-1} are uniquely determined by each other, the determination of K^{-1} based on our algorithm does not depend on the order of $B_1, \ldots, B_m$.

Now we assume that the elementary step being counted is the comparison of two attribute names. Consequently, if we assume that subsets of Ω are represented as sorted lists of attribute names, then a Boolean operation on two subsets of Ω requires at most $|\Omega|$ elementary steps.

Let $K_o = \{\Omega\}$. According to the construction of our algorithm we have $K_q = F_q \cup \{X_1, \ldots, X_{t_q}\}$, where $1 \leqq q \leqq m-1$. Denote l_q the number of elements of K_q. It is clear that for constructing K_{q+1} the worst-case time of algorithm is $O(n^2(l_q - t_q)t_q)$ if $t_q < l_q$ and $O(n^2 t_q)$ if $l_q = t_q$. Consequently, the total time spent by the algorithm in the worst cases is

$$O\left(n^2 \sum_{q=1}^{m-1} t_q u_q\right), \quad \text{where} \quad |\Omega| = n,$$

$$u_q = \begin{cases} l_q - t_q & \text{if} \quad l_q > t_q, \\ 1 & \text{if} \quad l_q = t_q. \end{cases}$$

It is obvious that, if $F_q = \emptyset$, then $l_q = t_q$.

It can be seen that when there are only a few minimal keys (that is m is small) our algorithm is very effective, it does not requires exponential time in $|\Omega|$. In cases for which $l_q \leqq l_m$ $(\forall q\colon 1 \leqq q \leqq m-1)$ it is obvious that our algorithm requires a number of elementary operations which is not greater than $O(n^2|K||K^{-1}|^2)$. Thus, in these cases our algorithm finds K^{-1} in polynomial time in $|\Omega|$, $|K|$ and $|K^{-1}|$.

After Theorem 2.12 we shall give an example to show that our algorithm requires exponential time in $|\Omega|$. On the other hand K_q in each step of our algorithm is obviously a Sperner-system. It is known ([4]) that the size of arbitrary Sperner-system over Ω can not be greater than $\binom{n}{[n/2]}$, where $n = |\Omega|$. $\binom{n}{[n/2]}$ is asymptotically equal to

$\frac{2^{n+1/2}}{(\pi\cdot n)^{1/2}}$. Consequently, the worst-case time of our algorithm can not be more than exponential in the number of attributes.

Let $K^{-1}=\{A_1, ..., A_t\}$ be a set of antikeys. Let $R=\{h_0, h_1, ..., h_t\}$ be a relation over Ω given as follows: for all $a\in\Omega$, $h(a)=0$

$$\text{for } i \ (1 \leqq i \leqq t), \quad h_i(a) = \begin{cases} 0 & \text{if} \quad a\in A_i, \\ i & \text{if} \quad a\in\Omega\setminus A_i. \end{cases}$$

If we consider R as a matrix, then R represents K (see [5]). Thus, based on our algorithm, for an arbitrarily given Sperner-system K, we can construct a matrix which represents K.

Example 2.3. Let $\Omega=\{1, 2, 3, 4, 5, 6\}$ and $K=\{(2, 3, 4), (1, 4)\}$. According to the above algorithm we have $K_1=\{(1, 3, 4, 5, 6), (1, 2, 4, 5, 6)\}\cup F_1$, where $F_1=\{(1, 2, 3, 5, 6)\}$, and $K_2=\{(3, 4, 5, 6), (2, 4, 5, 6), (1, 2, 3, 5, 6)\}$. It is obvious that $K^{-1}=K_2$.

We consider the following matrix:

The attributes:

$$\begin{array}{cc} & \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \mathbf{M} = & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \end{pmatrix} \end{array}$$

It is clear that $\mathbf{M}$ represents K.

Now we describe the "reverse" algorithm: for given Sperner-system considered as the set of antikeys we construct its origin. The following definitions are necessary for us.

Let F be a closure operation over Ω. Set

$$Z(F) = \{A \subseteqq \Omega\colon F(A) = A\}$$

$$\text{and} \quad T(F) = \{A \subset \Omega\colon A\in Z(F) \quad \text{and} \quad A\subset B \Rightarrow F(B) = \Omega.$$

The elements of $Z(F)$ are called closed sets. It is clear that $T(F)$ is the family of maximal closed sets (except Ω). Now we prove the following lemma.

Lemma 2.4. Let F be a closure operation over Ω, and K_F the set of minimal keys of F. Then $K_F^{-1}=T(F)$.

Proof. Let A be an arbitrary antikey and suppose that $A\subset F(A)$. Hence $F(F(A))=F(A)=\Omega$. Consequently, A is a key. This contradicts $\forall B\in K_F\colon B\nsubseteq A$. If there is an A' such that $A\subset A'$ and $A'\in Z(F)\setminus\{\Omega\}$, then A' is a key. This contradicts $A'\subset\Omega$.

On the other hand, if A is a maximal closed set and there is a B $(B\in K_F)$ such that $B\subseteqq A$, then $F(A)=\Omega$, which conflicts with the fact that $A\subset\Omega$. If $A\subset D(D\subseteqq\Omega)$, then it can be seen that $F(D)=\Omega$ (because A is the maximal closed set). Consequently, A is an antikey. The lemma is proved.

Now we construct an algorithm for finding a minimal key.

Let H be a Sperner-system and $\Omega\notin H$. We take a B $(B\in H)$ and an $a\in\Omega\setminus B$. We suppose that $B=\{b_1, ..., b_m\}$. Let $G=\{B_t\in H\colon a\notin B_t\}$ and $T_0=B\cup\{a\}$. define

$$T_{q+1}=\begin{cases}T_q\setminus\{b_{q+1}\} & \text{if} \quad \forall B_i\in H\setminus G\colon T_q\setminus\{b_{q+1}\}\not\subseteq B_i,\\ T_q & \text{otherwise.}\end{cases}$$

Theorem 2.5. If H is a set of antikeys, then $\{T_0, T_1, ..., T_m\}$ are the keys and T_m is a minimal key.

Proof. By Remark 2.1 there exists a closure F such that $H=K_F^{-1}$. We prove the theorem by the induction. It is clear that T_0 is a key. If T_q and $T_{q+1}=T_q$, then it is obvious that T_{q+1} is a key. If $T_{q+1}=T_q\setminus\{b_{q+1}\}$ and $F(T_{q+1})\neq\Omega$ then, by Lemma 2.4, there is a $B_t\in H$ such that $F(T_{q+1})\subseteq B_t$. Hence $T_{q+1}\subseteq B_t$, which conflicts with the fact $\forall B_t\in H\colon T_{q+1}\not\subseteq B_t$. Consequently, T_{q+1} is a key.

Now suppose that A is a proper subset of T_m. If $a\notin A$, then, clearly, $F(A)\neq\Omega$. If $a\in A$, then there exists a $b_q\in B$ such that $b_q\in T_m\setminus A$ $(1\leqq q)$. By the given algorithm there exists a $B_t\in H\setminus G$ such that $T_{q-1}\setminus\{b_q\}\subseteq B_t$. We obtain $A\subseteq T_m\setminus\{b_q\}\subseteq\subseteq T_{q-1}\{b_q\}\subseteq B_t$ by $T_m\subseteq T_q$ $(0\leqq q\leqq m-1)$. Hence $F(A)\neq\Omega$. Consequently, T_m is a minimal key. The theorem is proved.

Remark 2.6. Theorem 2.5 is also true if $T_0=\{b_1, ..., b_m\}$ is an arbitrary key. At this time define

$$T_{q+1}=\begin{cases}T_q\setminus\{b_{q+1}\} & \text{if} \quad \forall B_i\in H\colon T_q\setminus\{b_{q+1}\}\not\subseteq B_i,\\ T_q & \text{otherwise.}\end{cases}$$

— It is clear that the worst-case time of the algorithm is $O(n^2\cdot|H|)$, where $n=|\Omega|$, $|H|$ is the number of elements of H.

— It is best to choose B such that $|B|$ is minimal.

— If there is a B such that $\forall B_t\in H\setminus\{B\}\colon B_t\cap B=\emptyset$ and $a\in \bigcup\limits_{B_t\in H\setminus\{B\}} B_t$ then $a\cup b$ is a minimal key $(\forall b\in B)$.

— If $(\Omega\setminus\bigcup\limits_{B_t\in H} B_t)\neq\emptyset$, then $a\in\Omega\setminus\bigcup\limits_{B_t\in H} B_t$ is a minimal key.

— Let $Y=\bigcup\limits_{B_t\in H} B_t$ $(B_t\neq B)$. If $B\setminus Y\neq\emptyset$, then it is best to choose $T_0= =(B\cap Y)\cup\{a\}\cup\{b\}$, where $b\in B\setminus Y$.

Remark 2.7. Let H be a Sperner-system $(\Omega\notin H)$ and $A\subset\Omega$. We can give an algorithm (which is analogous to the above one) to decide whether A is a key or not. If A is a key, then this algorithm finds an A' such that $A'\subseteq A$ and A' is a minimal key.

Remark 2.8. In the paper [5] the equality sets of the relation are defined as follows: Let $\mathbf{R}=\{h_1, ..., h_m\}$ be a relation over Ω. For $i\neq j$, we denote by E_{ij} the set $\{a\in\Omega\colon h_i(a)=h_j(a)\}$, where $1\leqq i\leqq m$, $1\leqq j\leqq m$. Now we define $\mathbf{M}= =\{E_{ij}\colon \overline{\exists} E_{pq}$ such that $E_{ij}\subset E_{pq}\}$. Practically, it is possible that there are some E_{ij} which are equal to each other. We choose one E_{ij} from $\mathbf{M}$. According to Lemma 2.4 it can be seen that $\mathbf{M}$ is the set of antikeys of $K_{F_\mathbf{R}}$ (we consider $\mathbf{R}$ as a matrix).

Example 2.9. Let $\Omega=\{1, 2, 3, 4, 5, 6\}$ and $\mathbf{R}$ be the following relation:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 2 & 0 & 0 & 1 & 2 & 2 \\ 0 & 1 & 2 & 2 & 0 & 3 \\ 3 & 2 & 1 & 0 & 3 & 0 \end{pmatrix}$$

It can be seen that $\mathbf{M}=\{(1, 2), (3, 4, 5), (4, 6)\}$, where $E_{14}=\{1, 2\}$, $E_{15}=\{4, 6\}$ and $E_{25}=\{3, 4, 5\}$. By Theorem 2.5 and Remark 2.6, it is clear that $\{1, 3\}$, $\{1, 4\}$, $\{1, 5\}$, $\{1, 6\}$, $\{2, 3\}$, $\{2, 4\}$, $\{2, 5\}$, $\{2, 6\}$ are the minimal keys. We use the algorithm (Theorem 2.5) with $T_0=\{3, 4, 6\}$ and $T_0=\{4, 5, 6\}$, then it can be seen that $\{3, 6\}$ and $\{5, 6\}$ are minimal keys. Thus, based on this algorithm for an arbitrarily given relation $\mathbf{R}$ we can find a minimal key of $\mathbf{R}$.

Let K be an arbitrary Sperner-system. The following theorem has been proved in [2].

Theorem 2.10.

$$\binom{S(K)}{2} \geqq |K^{-1}| \geqq S(K)-1.$$

Denote by $\binom{\Omega}{k}$ the family of all k-element subsets of Ω. Let $F_k(n)=\max\left\{S(K): K\subseteqq\binom{\Omega}{k},\ |\Omega|=n\right\}$.

Theorem 2.11. ([6])

$$F_k(n) \geqq \sqrt{2}\binom{2k-2}{k-1}^{1/2\cdot[n/(2k-2)]}.$$

We define the function $f_{2k-1}\colon N\to N$ for $2k-1\leqq n$ by

$$f_{2k-1}(n)=\begin{cases} \binom{2k-1}{k-1}^{n/(2k-1)} & \text{if } n\equiv 0 \pmod{2k-1}, \\ \binom{2k-1}{k-1}^{[n/(2k-1)]-1}\times\binom{2k-1+p}{k-1} & \text{if } n\equiv p \pmod{2k-1} \text{ and } 1\leqq p\leqq k-1, \\ \binom{2k-1}{k-1}^{[n/(2k-1)]}\times\binom{p}{k-1} & \text{if } n\equiv p \pmod{2k-1} \text{ and } k\leqq p\leqq 2k-2, \end{cases}$$

and the function f_{2k-2} for $2k-2\leqq n$ by

$$f_{2k-2}(n)=\begin{cases} \binom{2k-2}{k-1}^{n/(2k-2)} & \text{if } n\equiv 0 \pmod{2k-2}, \\ \binom{2k-2}{k-1}^{[n/(2k-2)]-1}\times\binom{2k-2+p}{k-1} & \text{if } n\equiv p \pmod{2k-2} \text{ and } 1\leqq p\leqq k-1, \\ \binom{2k-2}{k-1}^{[n/(2k-2)]}\times\binom{p}{k-1} & \text{if } n\equiv p \pmod{2k-2} \text{ and } k\leqq p\leqq 2k-3, \end{cases}$$

where N denotes the set of natural numbers. Let us take a partition $\Omega = X_1 \cup \ldots \cup X_m \cup W$, where $m = \left[\frac{n}{2k-1}\right]$ and $|X_i| = 2k-1$ $(1 \leqq i \leqq m)$.
Let

$$K = \{B\colon |B| = k, B \subseteqq X_i, \forall i\} \quad \text{if} \quad |W| = 0.$$

$$K = \{B\colon |B| = k, B \subseteqq X_i \ (1 \leqq i \leqq m-1) \text{ and } B \subseteqq X_m \cup W\} \quad \text{if} \quad 1 \leqq |W| \leqq k-1.$$

$$K = \{B\colon |B| = k, B \subseteqq X_i \ (1 \leqq i \leqq m) \text{ and } B \subseteqq W\} \quad \text{if} \quad k \leqq |W| \leqq 2k-2.$$

It is clear that

$$K^{-1} = \{A\colon |A \cap X_i| = k-1, \forall i\} \quad \text{if} \quad |W| = 0.$$

$$K^{-1} = \{A\colon |A \cap X_i| = k-1 \ (1 \leqq i \leqq m-1) \text{ and } |A \cap (X_m \cup W)| = k-1\}$$

$$\text{if} \quad 1 \leqq |W| \leqq k-1.$$

$$K^{-1} = \{A\colon |A \cap X_i| = k-1 \ (1 \leqq i \leqq m) \text{ and } |A \cap W| = k-1\}$$

$$\text{if} \quad k \leqq |W| \leqq 2k-2.$$

It can be seen that $f_{2k-1}(n) = |K^{-1}|$. If we take a partition: $\Omega = X_1 \cup \ldots \cup X_m \cup W$; where $m = \left[\frac{n}{2k-2}\right]$ and $|X_i| = 2k-2$, in an analogous way we

$$K = \{B\colon |B| = k, B \subseteqq X_i, \forall i\} \quad \text{if} \quad |W| = 0.$$

$$K = \{B\colon |B| = k, B \subseteqq X_i \ (1 \leqq i \leqq m-1) \text{ and } B \subseteqq X_m \cup W\} \quad \text{if} \quad 1 \leqq |W| \leqq k-1.$$

$$K = \{B\colon |B| = k, B \subseteqq X_i \ (1 \leqq i \leqq m) \text{ and } B \subseteqq W\} \quad \text{if} \quad k \leqq |W| \leqq 2k-3.$$

It is clear that

$$f_{2k-2}(n) = |K^{-1}| \quad \text{and} \quad f_{2k-2}(n) \geqq \binom{2k-2}{k-1}^{[n/(2k-2)]}.$$

Theorem 2.12. Let $\Omega = \{1, \ldots, n\}$.
If $n \equiv 0 \pmod{(2k-2)(2k-1)}$, then $f_{2k-1}(n) > f_{2k-2}(n)$. For a fixed k,

$$\lim_{n \to \infty} \frac{f_{2k-1}^{(n)}}{f_{2k-2}^{(n)}} = \infty.$$

Proof. If $k=2$, then it is easy to prove that $\forall n\colon f_3(n) \geqq f_2(n)$. If $n=6$ or $n \geqq 8$, then $f_3(n) > f_2(n)$. Let

$$F = \frac{\binom{2k-1}{k-1}^{n/(2k-1)}}{\binom{2k-2}{k-1}^{n/(2k-2)}} = \frac{\left(\frac{2k-1}{k}\right)^{n/(2k-1)}}{\binom{2k-2}{k-1}^{n/(2k-2)(2k-1)}}.$$

It is known that $n!=\sqrt{2\pi n}\left(\frac{n}{e}\right)^n \times e^{\theta_n/(12n)}$, where $0<\theta_n<1$. So

$$F \geqq \frac{\left(\frac{2k-1}{k}\right)^{n/(2k-1)}}{2^{n/(2k-1)}\cdot\left(\frac{e^{\theta_n/(12(2k-2))}}{\sqrt{\pi(k-1)}}\right)^{n/(2k-2)(2k-1)}} \geqq \frac{\left(1-\frac{1}{2k}\right)^{n/(2k-1)}}{\left(\frac{e^{1/(24(k-1))}}{\sqrt{\pi(k-2)}}\right)^{n/(2k-2)(2k-1)}} = E.$$

For this E we obtain, that

$$T = \ln E = \frac{n}{2k-1}\left(\ln\left(1-\frac{1}{2k}\right)+\frac{1}{2k-2}\left(\frac{1}{2}\ln\left(\pi(k-1)\right)-\frac{1}{24(k-1)}\right)\right)$$

and by

$$\left|\ln\left(1-\frac{1}{2k}\right)\right| \leqq \frac{1}{2k-1}$$

we have

$$T \geqq \frac{n}{2k-1}\left(\frac{1}{2k-2}\left(\frac{1}{2}\ln\left(\pi(k-1)\right)-\frac{1}{24(k-1)}\right)-\frac{1}{2k-1}\right).$$

It can be seen that if $k=3$, then

$$\frac{1}{2k-2}\left(\frac{1}{2}\ln\left(\pi(k-1)\right)-\frac{1}{24(k-1)}\right)-\frac{1}{2k-1} > 0$$

and, for every $k \geqq 4$,

$$\frac{1}{2}\ln\left(\pi(k-1)\right)-\frac{1}{24(k-1)} > 1.$$

Hence

$$\frac{1}{2k-2}\left(\frac{1}{2}\ln\left(\pi(k-1)\right)-\frac{1}{24(k-1)}\right)-\frac{1}{2k-1} > 0.$$

Consequently, if $n \equiv 0 \pmod{(2k-2)(2k-1)}$, then $f_{2k-1}(n) > f_{2k-2}(n)$. Now let n be an arbitrary natural number. It can be seen that, for a fixed k, there exists a number $M>0$ such that

$$\frac{\binom{2k-1+p}{k-1}}{\binom{2k-1}{k-1}^{1+(p/(2k-1))}} < M, \quad \frac{\binom{p}{k-1}}{\binom{2k-1}{k-1}^{p/(2k-1)}} < M,$$

$$\frac{\binom{2k-2+p}{k-1}}{\binom{2k-2}{k-1}^{1+(p/(2k-2))}} < M \quad \text{and} \quad \frac{\binom{p}{k-1}}{\binom{2k-2}{k-1}^{p/(2k-2)}} < M.$$

Hence $\ln E \underset{n\to\infty}{\to} \infty$. Consequently, $F \underset{n\to\infty}{\to} \infty$. Thus,

$$\frac{f_{2k-1}(n)}{f_{2k-2}(n)} \underset{n\to\infty}{\to} \infty.$$

(It is easy to see that $k=2$ is also true.) The theorem is proved.

As a consequence of Theorem 2.12 and Theorem 2.10 we have

Corollary 2.13

$$F_k(n) \geqq \sqrt{2 f_{2k-1}(n)}.$$

Example 2.14. In Theorem 2.12 let $k=2$. Then we have $n-1 \leqq |K| \leqq n+2$ and $3^{(n/4)} < f_3(n)$, where $n=|\Omega|$., i.e. $3^{(n/4)} < |K^{-1}|$. Thus, we always can construct an example, in which the number of K (minimal keys) is not greater than $n+2$, but the size of K^{-1} (antikeys) is exponential in the number of attributes.

§ 3. Some special Sperner-systems

In this section we investigate connections between the minimal keys and antikeys for some special Sperner-systems.

The notion of saturated Sperner-system is defined in [7], as follows:

A Sperner-system K over Ω is saturated if for any $A \subseteqq \Omega$, $K \cup \{A\}$ is not a Sperner-system.

An important result in [7] has been proved; if K is a saturated Sperner-system then $K=K_F$ uniquely determines F, where F is a closure operation.

Now we investigate some special Sperner-systems which are strictly connected with saturated Sperner-systems.

We consider the following example.

Example 3.1. Let $\Omega=\{1, 2, 3, 4, 5, 6\}$ and $N=\{(1, 2), (3, 4), (5, 6)\}$ be a Sperner-system. It can be seen that $N^{-1}=\{(1, 3, 5), (1, 3, 6), (1, 4, 5), (1, 4, 6), (2, 3, 5), (2, 3, 6), (2, 4, 5), (2, 4, 6)\}$. Let $K=N \cup N^{-1}$. It is clear that K is saturated. We use the algorithm which finds a set of antikeys. Then $K^{-1}=\{(1, 3), (1, 4), (1, 5), (1, 6), (2, 3), (2, 4), (2, 5), (2, 6), (3, 5), (3, 6), (4, 5), (4, 6)\}$.

By the fact that $K^{-1} U \{1, 2\}$ is a Sperner-system it is obvious that K^{-1} is not saturated. Thus, we have

Corollary 3.2. There is a K so that K is saturated and K^{-1} is not saturated.

Now we define the following notion.

Definition 3.3. Let K be a Sperner-system over Ω. We say that K is embedded, if for every $A \in K$ there is a $B \in H$ such that $A \subset B$, where $H^{-1}=K$. We have

Theorem 3.4. Let K be a Sperner-system over Ω. K is saturated if and only if K^{-1} is embedded.

Proof. Let K be a saturated Sperner-system. According to the definition of K^{-1} it is clear that K^{-1} is embedded. Assume that K^{-1} is an embedded Sperner-system, but K is not saturated. Consequently, for K there exists an $A\subset\Omega$ such that $K\cup\{A\}$ is a Sperner-system. It can be seen that, for every $C\in K$, we have $C\subset\Omega$ (because of $\Omega\notin K$). Hence we can construct B such that $A\subseteq B$, $K\cup\{B\}$ is a Sperner-system and, for every $B'(B\subset B')$, there is a $C\in K$ with $C\subseteq B'$. It can be seen that $B\in K^{-1}$. This contradicts the fact that K^{-1} is embedded. The proof is complete.

Now we define an inclusive Sperner-system.

Definition 3.5. Let K be a Sperner-system over Ω. We say that K is inclusive, if for every $A\in K$, there exists a $B\in K^{-1}$ such that $B\subset A$. We have

Theorem 3.6. K is an inclusive Sperner-system if and only if K^{-1} is a saturated one.

Proof. Now, assume that K is an inclusive Sperner-system but K^{-1} is not saturated. By the definition of K^{-1}, there is a $B\in(K^{-1})^{-1}$ such that $K^{-1}\cup\{B\}$ is a Sperner-system. By Remark 2.1, for K there is a closure operation F such that $K=K_F$. If $F(B)\subset\Omega$, then by Lemma 2.4 there exists an $A\in K^{-1}$ with $F(B)\subseteq A$ (the set of antikeys is family of the maximal closed sets), which conflicts with the fact that $K^{-1}\cup\{B\}$ is a Sperner-system. Consequently, B is a key. If we use the algorithm which finds a minimal key in Theorem 2.5, then it can be seen that there exists a $B'(B'\subseteq B)$ such that $B'\in K$, and it is clear that $K^{-1}\cup\{B'\}$ is a Sperner-system. This contradicts the definition of K. Thus, K^{-1} is saturated.

On the other hand by the definition of K^{-1} and by the assumption that K^{-1} is saturated it is clear that K is an inclusive Sperner-system. The theorem is proved. Now, we have the following corollary by Theorem 3.4 and Theorem 3.6.

Corollary 3.7. Let K be a Sperner-system over Ω. Denote H a Sperner-system, for which $H^{-1}=K$. The following facts are equivalent:

(1) K is saturated,
(2) K^{-1} is embedded,
(3) H is inclusive.

Proposition 3.8. There exists a Sperner-system K such that

(1) K is saturated, but K^{-1} is not saturated.
(2) K is saturated, but H is not saturated.
(3) K is embedded, but K^{-1} is not embedded.
(4) K is embedded, but H is not embedded.
(5) K is inclusive, but K^{-1} is not inclusive.
(6) K is inclusive, but H is not inclusive,

where H denotes a Sperner-system for which $H^{-1}=K$.

Proof. From Example 3.1 we have (1). By Theorem 3.4, $(K^{-1})^{-1}$ is not embedded in this example. Hence we have (3). By Theorem 3.6, in Example 3.1 H is inclusive, where $H^{-1}=K$. Now, we suppose that, if K is inclusive, then the set of antikeys of K is also inclusive. Consequently, in Example 3.1, H is inclusive, and K is an inclusive Sperner-system. From Theorem 3.6, K^{-1} is saturated. This constradicts the fact that K^{-1} in Example 3.1 is not a saturated Sperner-system. Hence we have (5). (2) can be proved as follows: Let K be a Sperner-system. Let $K^1=K$ and, for $n\geqq 2$, define

K^n by the equality $(K^n)^{-1}=K^{n-1}$. We know that the number of the Sperner-systems over Ω is finite (at most $2^{2^{|\Omega|}}$). On the other hand, K and K^{-1} are determined uniquely by each other. Consequently, there exists a number m $(2 \leqq m \leqq 2^{2^{|\Omega|}})$ such that $K^m=K$ and $K^{m-1}=K^{-1}$. If we suppose that K is saturated, then H is also saturated, where $H^{-1}=K$. This means that for every p with $2 \leqq p < m$, K^p is also saturated. This contradicts Corollary 3.2. Thus, there is a Sperner-system K such that K is saturated, but H is not saturated. By similar arguments we have also (4) and (6). The proposition is proved.

Acknowledgement

The author would like to take this opportunity to express deep gratitude to Professor J. Demetrovics and L. Hannák for their help, valuable comments and suggestions.

COMPUTER AND AUTOMATION INSTITUTE
HUNGARIAN ACADEMY OF SCIENCES
KENDE U. 13—17.
BUDAPEST, HUNGARY
H—1502

References

[1] ARMSTRONG, W. W., Dependency Structures of Data Base Relationships, Information Processing 74, North-Holland Publ. Co. (1974) 580—583.
[2] BÉKÉSSY, A., DEMETROVICS J., HANNÁK L., KATONA G. O. H., FRANKL P., On the number of maximal dependencies in data relation of fixed order. Discrete Math., 30 (1980) 83—88.
[3] CODD, E. F., Relational model of data for large shared data banks. Communications of the ACM, 13, (1970) 377—384.
[4] DEMETROVICS, J., On the equivalence of candidate keys with Sperner systems. Acta Cybernetica 4 (1979) 247—252.
[5] DEMETROVICS, J., Relációs adatmodell logikai és strukturális vizsgálata. MTA SZTAKI Tanulmányok, Budapest, 114 (1980) 1—97.
[6] DEMETROVICS J., FÜREDI Z., KATONA G. O. H., Minimum matrix representation of closure operations. Preprint of the mathematical institute of the Hungarian academy of sciences, Budapest, 12 (1983) 1—22.
[7] DEMETROVICS J., FÜREDI Z., KATONA G., A függőségek és az individumok száma közötti kapcsolat összetett adatrendszerek esetén. Alkalmazott Matematikai Lapok 9 (1983) 13—21.

(Received April 2, 1985.)