Some Problems Related to Keys and the Boyce-Codd Normal Form

Vu Duc Thi[*] and Nguyen Hoang Son[†]

Abstract

The aim of this paper is to investigate the connections between minimal keys and antikeys for special Sperner-systems by hypergraphs. The Boyce-Codd normal form and some related problems are also studied in this paper.

1 Introduction

In the relational datamodel, one of the important concepts is the functional dependency. Several types of families of functional dependencies which satisfy some conditions are known under the name of normal forms (NFs). The most desirable NF is Boyce-Codd NF (BCNF) that has been investigated in a lot of papers (see [2, 8, 9, 10]). The minimal keys and set of antikeys are interesting concepts in the relational datamodel (see, e.g., [11, 12]). A set of minimal keys and set of antikeys form Sperner-systems. Sperner-systems and sets of minimal keys are equivalent in the sense that for an arbitrary Sperner-system K a family of functional dependencies F can be constructed so that the minimal keys of F are exactly the elements of K (see [5]).

Hypergraph theory (see, e.g., [3]) is an important subfield of discrete mathematics with many relevant applications in both theoretical and applied computer science. The transversal and the minimal transversal of a hypergraph are important concepts in this theory, on one hand.

The paper is structured as follows: in the second section, some necessary definitions and results about hypergraph theory are given.

In Section 3, transformations of the notions and the results of Section 2 concerning hypergraphs to relational databases are shown. We prove that the set of all prime attributes is the set of all independent attributes of a given relation scheme. We give an effective algorithm finding a BCNF relation r such that r represents a given BCNF relation scheme s (i.e., $K_r = K_s$, where K_r and K_s are sets of all minimal keys of r and s). We aslo give an effective algorithm which from a given BCNF

[*]Institute of Information Technology, National Centre for Natural Science and Technology of Vietnam, 18 Hoang Quoc Viet, Hanoi, Vietnam

[†]Department of Mathematics, College of Sciences, Hue University, Vietnam

relation r finds a BCNF relation scheme s such that $K_r = K_s$. Section 4, we study the connections between minimal keys and antikeys for special Sperner-system by hypergraphs.

2 Basic definitions and results

In this section we start with some basic definitions and results on hypergraphs.

Definition 2.1. *Let R be a nonempty finite set and put $\mathcal{P}(R)$ for the family of all subsets of R (its power set). The family $\mathcal{H} = \{E_i : E_i \in \mathcal{P}(R), i = 1, ..., m\}$ is called a hypergraph over R if $E_i \neq \emptyset$ holds for all i (in [3] it is required that the union of E_is is R, in this paper we do not require this).*

The elements of R are called vertices, and the sets $E_1, ..., E_m$ the edges of the hypergraph $\mathcal{H}$.

A hypergraph $\mathcal{H}$ is called simple if it satisfies $\forall E_i, E_j \in \mathcal{H} : E_i \subseteq E_j \Rightarrow E_i = E_j$. It can be seen that simple hypergraphs are Sperner-systems.

One can see easily that the family $m(\mathcal{H}) = \{E_i \in \mathcal{H} : \nexists E_j \in \mathcal{H} : E_j \subset E_i\}$ is a simple hypergraph, and that $m(\mathcal{H})$ is uniquely determined by $\mathcal{H}$.

Definition 2.2. *Let $\mathcal{H}$ be a hypergraph over R. A set $T \subseteq R$ is called a transversal of $\mathcal{H}$ (sometimes it is called hitting set) if it meets all edges of $\mathcal{H}$, i.e., $\forall E \in \mathcal{H} : T \cap E \neq \emptyset$. Denote by $Trs(\mathcal{H})$ the family of all transversals of $\mathcal{H}$. A transversal T of $\mathcal{H}$ is called minimal if no proper subset T' of T is a transversal.*

The family of all minimal transversals of $\mathcal{H}$ called the transversal hypergraph of $\mathcal{H}$, and denoted by $Tr(\mathcal{H})$. Clearly, $Tr(\mathcal{H})$ is a simple hypergraph.

The following algorithm finds the family of all minimal transversals of a given hypergraph (by induction).

Algorithm 2.1. (Demetrovics and Thi [7]).
Input: Let $\mathcal{H} = \{E_1, ..., E_m\}$ be a hypergraph over R.
Output: $Tr(\mathcal{H})$.
Method:
Step 0: We set $L_1 := \{\{a\} : a \in E_1\}$. It is obvious that $L_1 = Tr(\{E_1\})$.
Step q+1: $(q < m)$ Assume that

$$L_q = S_q \cup \{B_1, ..., B_{t_q}\},$$

where $B_i \cap E_{q+1} = \emptyset, i = 1, ..., t_q$ and $S_q = \{A \in L_q : A \cap E_{q+1} \neq \emptyset\}$.

For each i $(i = 1, ..., t_q)$ constructs the set $\{B_i \cup \{b\} : b \in E_{q+1}\}$. Denote them by $A_1^i, ..., A_{r_i}^i (i = 1, ..., t_q)$. Let

$$L_{q+1} = S_q \cup \{A_p^i : A \in S_q \Rightarrow A \not\subset A_p^i, 1 \leq i \leq t_q, 1 \leq p \leq r_i\}.$$

Theorem 2.1. (Demetrovics and Thi [7]). *For every q $(1 \leq q \leq m)$ $L_q = Tr(\{E_1, ..., E_q\})$, i.e., $L_m = Tr(\mathcal{H})$.*

It can be seen that the determination of $Tr(\mathcal{H})$ based on our algorithm does not depend on the order of $E_1, ..., E_m$.

Remark 2.1. (Demetrovics and Thi [7]). Denote $L_q = S_q \cup \{B_1, ..., B_{t_q}\}$, and $l_q (1 \leq q \leq m-1)$ be the number of elements of L_q. It can be seen that the worst-case time complexity of our algorithm is

$$\mathcal{O}(|R|^2 \sum_{q=0}^{m-1} t_q u_q),$$

where $l_0 = t_0 = 1$ and

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

Clearly, in each step of our algorithm L_q is a simple hypergraph. It is known that the size of arbitrary simple hypergraph over R cannot be greater than $C_n^{[n/2]}$, where $n = |R|$. $C_n^{[n/2]}$ is asymptotically equal to $2^{n+1/2}/(\pi.n)^{1/2}$. From this, the worst-case time complexity of our algorithm cannot be more than exponential in the number of attributes. In cases for which $l_q \leq l_m (q = 1, ..., m-1)$, it is easy to see that the time complexity of our algorithm is not greater than $\mathcal{O}(|R|^2|\mathcal{H}||Tr(\mathcal{H})|^2)$. Thus, in these cases this algorithm finds $Tr(\mathcal{H})$ in polynomial time in $|R|, |\mathcal{H}|$ and $|Tr(\mathcal{H})|$. Obviously, if the number of elements of $\mathcal{H}$ is small, then this algorithm is very effective. It only requires polynomial time in $|R|$.

The above algorithm reminds that in [3], but its form seems to be more convenient for our applications.

The following proposition is obvious.

Proposition 2.1. (Demetrovics and Thi [7]). *The time complexity of finding $Tr(\mathcal{H})$ of a given hypergraph $\mathcal{H}$ is (in general) exponential in the number of elements of R.*

Proposition 2.1 is still true for a simple hypergraph.

However, if we restrict the number of edges of a hypergraph, then the time complexity of finding $Tr(\mathcal{H})$ of a given hypergraph $\mathcal{H}$ is polynomial time.

Algorithm 2.2.
Input: Let $\mathcal{H} = \{E_1, ..., E_k\}$ be a simple hypergraph over R, where k is a constant.
Output: $Tr(\mathcal{H})$.
Method:
Step 1: We construct the set

$$\mathcal{G} = \{\{e_1\} \cup ... \cup \{e_k\} : e_i \in E_i, 1 \leq i \leq k\}.$$

Step 2: Compute
$$m(\mathcal{G}) = \{E_i \in \mathcal{G} : \nexists E_j \in \mathcal{G} : E_j \subset E_i\}.$$

Step 3: Let $Tr(\mathcal{H}) = m(\mathcal{G})$.

It is obvious that $m(\mathcal{G}) = Tr(\mathcal{H})$. Furthermore, $\mathcal{G} \supseteq Tr(\mathcal{H})$, and $|\mathcal{G}| < |R|^k$. Hence, in this case Algorithm 2.2 finds $Tr(\mathcal{H})$ in polynomial time. Clearly, if k is small, then our algorithm is very effective.

Definition 2.3. *Let R be a set and $R' \subseteq R$ a subset of it. Then $\overline{R'}$ denotes $R - R'$. Let $\mathcal{H}$ be a hypergraph over R. Then $\overline{\mathcal{H}} = \{\overline{E} : E \in \mathcal{H}\}$ is called the comlemented hypergraph of $\mathcal{H}$.*

It is known [3] that if $\mathcal{H}$ is a hypergraph, then $\overline{\overline{\mathcal{H}}} = \mathcal{H}$, and $\mathcal{H}$ is simple iff $\overline{\mathcal{H}}$ is simple.

3 Boyce-Codd normal form and transversals

Definition 3.1. *Let $R = \{a_1, ..., a_n\}$ be a nonempty finite set of attributes. A functional dependency (FD) is a statement of form $X \to Y$, where $X, Y \subseteq R$. The FD $X \to Y$ holds in a relation $r = \{h_1, ..., h_m\}$ over R if*

$$(\forall h_i, h_j \in r)((\forall a \in X)(h_i(a) = h_j(a)) \Rightarrow (\forall b \in Y)(h_i(b) = h_j(b))).$$

We also say that r satisfies the FD $X \to Y$.
Let F_r be a family of all FDs that holds in r. Then $F = F_r$ satisfies
(F1) $X \to X \in F$,
(F2) $(X \to Y \in F, Y \to Z \in F) \Rightarrow (X \to Z \in F)$,
(F3) $(X \to Y \in F, X \subseteq V, W \subseteq Y) \Rightarrow (V \to W \in F)$,
(F4) $(X \to Y \in F, V \to W \in F) \Rightarrow (X \cup V \to Y \cup W \in F)$.

A family of FDs satisfying (F1) - (F4) is called a f-family over R.
Clearly, F_r is a f-family over R. It is known [1] that if F is an arbitrary f-family, then there is a relation r over R such that $F_r = F$.
Given a family F of FDs over R, there exists a unique minimal f-family F^+ that contains F. It can be seen that F^+ contains all FDs which can be derived from F by the rules (F1) - (F4).
A relation scheme s is a pair (R, F), where R is a set of attributes, and F is a set of FDs over R. Denote $X^+ = \{a \in R : X \to \{a\} \in F^+\}$. X^+ is called the closure of X over s. It is clear that, $X \to Y \in F^+$ iff $Y \subseteq X^+$.
Clearly, if $s = (R, F)$ is a relation scheme, then there is a relation r over R such that $F_r = F^+$ (see, [1]).
Let r be a relation, $s = (R, F)$ be a relation scheme over R and $A \subseteq R$. Then A is a key of r (a key of s) if $A \to R \in F_r (A \to R \in F^+)$. A is a minimal key of $r(s)$ if A is a key of $r(s)$ and any proper subset of A is not a key of $r(s)$.
Denote $K_r(K_s)$ the set of all minimal keys of $r(s)$. It can be seen that K_r, K_s are simple hypergraphs over R.

Definition 3.2. *Let $s = (R, F)$ be a relation scheme over R. We say that an attribute $a \in R$ is prime if it belongs to a minimal key of s, and nonprime otherwise. $s = (R, F)$ is in BCNF if $A \to \{a\} \notin F^+$ for $A^+ \neq R, a \notin A$.*

If a relation scheme is changed to a relation we have the definition of BCNF for relation.

Let s be a relation scheme and r a relation over R. We say that r represents s if $K_r = K_s$.

Definition 3.3. *Let r be a relation over R, and E_r the equality set of r, i.e.* $E_r = \{E_{ij} : 1 \leq i < j \leq |r|\}$, *where* $E_{ij} = \{a \in R : h_i(a) = h_j(a)\}$. *Let* $T_r = \{E_{ij} \in E_r : \nexists E_{pq} \in E_r : E_{ij} \subset E_{pq}\}$. *Then T_r is called the maximal equality system of r.*

Definition 3.4. *Let K be a simple hypergraph over R. We define the set of antikeys of K, denoted by K^{-1}, as follows:*

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

It is easy to see that K^{-1} is also a simple hypergraph over R.

In this paper, we always assume that if a simple hypergraph plays the role of the set of minimal keys (antikeys), then this simple hypergraph is not empty (does not contain R).

Definition 3.5. *Let $s = (R, F)$ be a relation scheme and r a relation over R. For every $A \subseteq R$, set $I(A) = \{a \in R : A \rightarrow \{a\} \notin F^+\}$. Then $I(A)$ is called an independent set of s. For r, put $I(A) = \{a \in R : A \rightarrow \{a\} \notin F_r\}$. Denote by I_s the family of all independent sets of s.*

Set $m(s) = \{B \in I_s : B \neq \emptyset, \nexists C \in I_s : C \subset B\}$. $m(s)$ is called the family of all minimal independent sets of s. Clearly, $m(s)$ is a simple hypergraph over R.

It can be seen that A is a key of s if and only if $I(A) = \emptyset$.

Denote by I_r and $m(r)$ the family of all independent sets and the family of all minimal independent sets of r.

The following result was discovered in [7].

Theorem 3.1. *(Demetrovics and Thi [7]). Let $s = (R, F)$ be a relation scheme over R. Then*

$$Tr(K_s) = m(s).$$

It is known [3] that if $\mathcal{H}, \mathcal{G}$ are two simple hypergraphs over R, then $\mathcal{H} = Tr(\mathcal{G})$ if and only if $\mathcal{G} = Tr(\mathcal{H})$. From this we obtain

Corollary 3.1. *Let $s = (R, F)$ be a relation scheme over R. Then*

$$K_s = Tr(m(s)).$$

Definition 3.6. *Let $s = (R, F)$ be a relation scheme over R. We say that an attribute $a \in R$ is independent if it belongs to an independent set of s, and dependent otherwise.*

Denote by D_n the set of all dependent attributes of s. Clearly, $R - D_n$ is the set of all independent attributes of s.

Lemma 3.1. *Let $\mathcal{H}$ be a simple hypergraph over R. Then*

$$\cup Tr(\mathcal{H}) = \cup \mathcal{H}.$$

Proof. Assume that $a \in \cup Tr(\mathcal{H})$. Hence, there exists a minimal transversal T of $\mathcal{H}$ such that $a \in T$. From this, we obtain $a \in E, E \in \mathcal{H}$. This means that $a \in \cup \mathcal{H}$. Consequently, $\cup Tr(\mathcal{H}) \subseteq \cup \mathcal{H}$ holds.

Conversely, if $a \in \cup \mathcal{H}$ then there is $E \in \mathcal{H}$ such that $a \in E$. From this, according to the definition of transversal hypergraph of $\mathcal{H}$ there exists $T \in Tr(\mathcal{H})$ such that $a \in T$, i.e. $a \in \cup Tr(\mathcal{H})$. Hence, $\cup \mathcal{H} \subseteq \cup Tr(\mathcal{H})$. The proof is complete. □

From Lemma 3.1 we obtain the following

Corollary 3.2. *Let $s = (R, F)$ be a relation scheme over R, $m(s)$ be a family of all independent sets of s. Then*

$$\cup Tr(m(s)) = \cup m(s).$$

Theorem 3.2. *Let $s = (R, F)$ be a relation scheme over R. Then*

$$\cup K_s = R - D_n.$$

Proof. Assume that a is an element of $R - D_n$, i.e., there exists an $I(A) \in m(s)$ such that $a \in I(A)$. Hence, $a \in \cup m(s)$. By Corollary 3.2 we attain $a \in \cup Tr(m(s))$. By Theorem 3.1 we also obtain $a \in \cup K_s$. Thus, $R - D_n \subseteq \cup K_s$.

Conversely, suppose that $a \in \cup K_s$. Thus, by Corollary 3.1 and Corollary 3.2 $a \in \cup m(s)$. Hence, there exists an $I(A) \in m(s)$ such that $a \in I(A)$, i.e., $a \in R - D_n$. Consequently, $\cup K_s \subseteq R - D_n$.

The theorem is proved.   □

Minimal keys and antikeys are related as follows:

Proposition 3.1. *Let $s = (R, F)$ be a relation scheme over R. Then*

$$K_s^{-1} = \overline{Tr(K_s)}.$$

Proof. Assume $X \in K_s^{-1}$. From Definition 3.4 we have that for every minimal key K, $K - X \neq \emptyset$, thus $\overline{X} \cap K \neq \emptyset$. Which implies that $\overline{X} \in Trs(K_s)$. On the other hand, according to the definition of antikey set, we have

$$X \cup \{a\} \supseteq K,$$

where $a \in \overline{X}$ and $K \in K_s$, which implies that $(\overline{X} - \{a\}) \cap K = \emptyset$. Consequently, $\overline{X} \in Tr(K_s)$, i.e., $X \in \overline{Tr(K_s)}$. Hence, we have $K_s^{-1} \subseteq \overline{Tr(K_s)}$.

Conversely, suppose that $Y \in Tr(K_s)$. Then $\overline{Y}$ is not superset of any minimal keys. Clearly, for all $a \in Y, Y - \{a\} \notin Trs(K_s)$, i.e. $(Y - \{a\}) \cap K = \emptyset$. This means that

$$\overline{Y} \cup \{b\} \supseteq K,$$

for all $b \in Y$. Consequently, $\overline{Tr(K_s)} \subseteq K_s^{-1}$.

The proposition is proved.   □

Remark 3.1. Let $s = (R, F)$ be a relation scheme over R. Set $Z_s = \{A^+ : A \subseteq R\}$, i.e., Z_s is the set of all closures of s. Put $T_s = \{A \in Z_s : A \neq R, \nexists B \in Z_s : A \subset B\}$. Hence, T_s is the set of all maximal elements of $Z_s - \{R\}$. By the definition of the independent set of s, we can see that $T_s = \{R - B : B \in m(s)\}$.

From Theorem 3.1, Proposition 3.1 and Remark 3.1 we have

Proposition 3.2. *Let* $s = (R, F)$ *be a relation scheme over* R. *Then*

$$\overline{Tr(K_s)} = T_s.$$

The Proposition 3.2 means that for all $A \in \overline{Tr(K_s)} : A^+ = A$ and $A \neq R$.

Remark 3.2. Let r be a relation over R. From r we compute E_r. We construct the maximal equality system T_r of r. Then we have $T_r = K_r^{-1}$ (see, e.g., [8]). Denote elements of T_r by $A_1, ..., A_t$.

Set $M_r = \{B : B \neq \emptyset, B = A_i - \{a\} : a \in R, i = 1, ..., t\}$. Denote elements of M_r by $B_1, ..., B_l$. We construct a relation $r' = \{h_0, h_1, ..., h_l\}$ as follows:

$$\text{for all } a \in R, h_0(a) = 0, \forall i = 1, ..., l$$

$$h_i(a) = \begin{cases} 0, & \text{if } a \in B_i, \\ i, & \text{otherwise.} \end{cases}$$

Clearly, r' is in BCNF and $K_r = K_{r'}$.

We give the following algorithm that from a given relation scheme s constructs a relation r such that r represents s.

Algorithm 3.1.
Input: a BCNF relation scheme $s =< R, F >$.
Output: a BCNF relation r such that $K_r = K_s$.
Method:
Step 1: From s compute K_s.
Step 2: By Algorithm 2.1 we construct the set $Tr(K_s)$.
Step 3: Compute $\overline{Tr(K_s)}$. Denote elements of $\overline{Tr(K_s)}$ by $A_1, ..., A_t$.
Step 4: Set $Q_s = \{B : B \neq \emptyset, B = A_i - \{a\} : a \in R, i = 1, 2, ..., t\}$. Denote elements of Q_s by $B_1, ..., B_l$.
Step 5: Construct a relation $r = \{h_0, h_1, ..., h_l\}$ as follows:

$$\text{for all } a \in R, h_0(a) = 0, \forall i = 1, ..., l$$

$$h_i(a) = \begin{cases} 0, & \text{if } a \in B_i, \\ i, & \text{otherwise.} \end{cases}$$

Based on Proposition 3.1, Remark 3.2 and Proposition 3.2 we have $K_r = K_s$ and r is in BCNF. It is easy to see that the time complexity of Algorithm 3.1 is exponential in the number of attributes.

Let r be a relation over R. Let $N_r = \{N_{ij} : 1 \leq i < j \leq |r|\}$, where $N_{ij} = \{a \in R : h_i(a) \neq h_j(a)\}$. Then N_r is called the nonequality set of r.

Let $M_r = \{A \in N_r : \nexists B \in N_r : B \subset A\}$. M_r is called the minimal nonequality system of r.

The following result was discovered in [7].

Theorem 3.3. (Demetrovics and Thi [7]). *Let r be a relation over R. Then $K_r = Tr(M_r)$, where M_r is the minimal nonequality system of r.*

From Theorem 3.3 we have an effective application of Theorem 3.3, which is the following algorithm finding a BCNF relation scheme s such that $K_s = K_r$ from a given relation r in BCNF.

Algorithm 3.2.
Input: Let r be a BCNF relation over R.
Output: a BCNF relation scheme $s = < R, F >$ such that $K_s = K_r$.
Method:
Step 1: From r compute N_r.
Step 2: From N_r compute the minimal nonequality system M_r.
Step 3: By Algorithm 2.1 constructs $Tr(M_r)$. Clearly, $K_r = Tr(M_r)$.
Step 4: Denoting elements of K_r by $A_1, ..., A_m$. We construct a relation scheme as follows: $s = < R, F >$, where $F = \{A_1 \rightarrow R, ..., A_m \rightarrow R\}$.

Clearly, s is in BCNF and $K_s = K_r$. The time complexity of this algorithm is the time complexity of Algorithm 2.1. In many cases this algorithm is very effective (see Remark 2.1).

4   Special Sperner-systems and transversals

The notion of saturated Sperner-system is defined in [6] as follows:

Definition 4.1. (Demetrovics [6]). *A Sperner-system K over R is saturated if for any $A \subseteq R, K \cup \{A\}$ is not a Sperner-system.*

Now we are going to give a new characterization of saturate Sperner-systems. To do this, we need the following definition:

Definition 4.2. *Let $\mathcal{H}$ and $\mathcal{G}$ be two hypergraphs over R. Then $\mathcal{H} > \mathcal{G}$ iff for every $H \in \mathcal{H}$ there exists $G \in \mathcal{G}$ such that $H \supset G$, and $\mathcal{H} < \mathcal{G}$ iff for every $H \in \mathcal{H}$ there exists $G \in \mathcal{G}$ such that $H \subset G$.*

From this definition we obtain the following:

Proposition 4.1. *Let $\mathcal{H} \neq \emptyset$ and $\mathcal{G} \neq \emptyset$ be two hypergraphs over R. Then*
   (1) $\emptyset > \mathcal{H}$ and $\emptyset < \mathcal{H}$.
   (2) $\mathcal{H} > \{\emptyset\}$.
   (3) $\{\emptyset\} < \mathcal{H}$.
   (4) $\mathcal{H} > \mathcal{G}$ (resp. $\mathcal{H} < \mathcal{G}$) does not imply $\mathcal{G} < \mathcal{H}$ (resp. $\mathcal{G} > \mathcal{H}$.
   (5) $\mathcal{H} < \{R\}$ iff $R \notin \mathcal{H}$.
   (6) $\mathcal{H} \subseteq \mathcal{G}$ does not imply $\mathcal{H} < \mathcal{G}$.

Proof.
(1) It is obvious from Definition 4.2.
(2) Since $\mathcal{H}$ is hypergraph, we have (2).
(3) By similar arguments we also have (3).
(4) We give a counterexample. Let $R = \{a, b, c\}$. Consider the hypergraphs

$$\mathcal{H} = \{\{a, b\}\}, \mathcal{G} = \{\{a\}, \{b\}, \{c\}, \{a, b, c\}\}.$$

It holds that $\mathcal{H} > \mathcal{G}$ (resp. $\mathcal{H} < \mathcal{G}$), but it does not hold that $\mathcal{G} < \mathcal{H}$ (resp. $\mathcal{G} > \mathcal{H}$).
(5) From definition of hypergraphs and Definition 4.2 we obtain (5).
(6) We give a counterexample. Let $R = \{a, b\}$. Consider the hypergraphs

$$\mathcal{H} = \{\{a\}\}, \mathcal{G} = \{\{a\}, \{b\}\}.$$

It holds that $\mathcal{H} \subseteq \mathcal{G}$, but it does not hold that $\mathcal{H} < \mathcal{G}$.
The proposition is proved. ∎

Remark 4.1. $>$ and $<$ are transitive on the hypergraphs on R.

Theorem 4.1. *Let K be a Sperner-system over R. Then K is saturated if and only if $\overline{Tr(K)} < K$.*

Proof. Let K be a saturated Sperner-system. Suppose that there exists an $A \in \overline{Tr(K)}$ such that for every $B \in K, A \not\subset B$. By Proposition 3.1 and Definition 3.4 we have $K \cup \{A\}$, a Sperner-system. Which contradicts the hypothesis that K is saturated. Consequently, $\overline{Tr(K)} < K$.
Conversely, suppose that $\overline{Tr(K)} < K$, but K is not saturated. Hence, there exists an $A \subset R$ such that $K \cup \{A\}$ is a Sperner-system. Because $R \notin K$, for every $C \in K$ we have $C \subset R$. Thus, we can construct B such that $A \subseteq B, K \cup \{B\}$ is a Sperner-system and for every $D(B \subset D)$, there exists $C \in K$ such that $D \supseteq C$. Which implies that $B \in K^{-1}$. This contradicts the hypothesis $\overline{Tr(K)} < K$, i.e., for every $A \in K^{-1}$ (because $\overline{Tr(K)} = K^{-1}$), there exists $B \in K$ such that $A \subset B$. Consequently, K is saturated. The theorem is proved. ∎

Definition 4.3. *Let K be a Sperner-system over R. We say that K is embedded if for every $A \in K$ there is a $B \in H$ such that $A \subset B$, where $H^{-1} = K$.*

From Proposition 3.1, Theorem 4.1 we have the following

Proposition 4.2. *Let K be a Sperner-system over R. Then K is saturated if and only if $\overline{Tr(K)}$ is embedded.*

From Proposition 3.1 and Proposition 4.2 the following corollary is immediate:

Corollary 4.1. *Let K be a Sperner-system over R. Then K is saturated if and only if K^{-1} is embedded.*

Corollary 4.1 was shown in [12].

Definition 4.4. *Let K be a Sperner-system over R. We say that K is inclusive if for every $A \in K$, there exists a $B \in K^{-1}$ such that $B \subset A$.*

From Proposition 3.1, Definition 4.2 and Definition 4.4, the following proposition is evident.

Proposition 4.3. *Let K be a Sperner-system over R. Then K is inclusive if and only if $K > \overline{Tr(K)}$.*

Remark 4.2. (Demetrovics [4]). If K is an arbitrary Sperner-system over R, then there is a relation scheme $s = (R, F)$ such that $K = K_s$.

Theorem 4.2. *Let K be a Sperner-system over R. Then K is inclusive if and only if $\overline{Tr(\overline{Tr(K)})} < \overline{Tr(K)}$.*

Proof. Suppose that K is an inclusive Sperner-system, but there exists an $A \in \overline{Tr(\overline{Tr(K)})}$ such that for every $B \in \overline{Tr(K)}, A \not\subset B$. Hence, $\overline{Tr(K)} \cup \{A\}$ is a Sperner-system. By Remark 4.2, for K there is a relation scheme s such that $K = K_s$. If $A^+ \subset R$ then according to Proposition 3.2 there exists $C \in \overline{Tr(K)}$ such that $A^+ \subseteq C$, which contradicts the fact that $\overline{Tr(K)} \cup \{A\}$ is a Sperner-system. Consequently, A is a key of s. It is obvious that there is a minimal key $A'(A' \subseteq A)$ such that $A' \in K$. Thus, $\overline{Tr(K)} \cup \{A'\}$ is a Sperner-system. By Proposition 4.3, this is a contradiction. Consequently, $\overline{Tr(\overline{Tr(K)})} < \overline{Tr(K)}$.

Conversely, assume that $\overline{Tr(\overline{Tr(K)})} < \overline{Tr(K)}$. By Proposition 4.2, we obtain which $\overline{Tr(K)}$ is saturated. From this, Proposition 3.2 and Proposition 4.3, we have K, an inclusive Sperner-system. The theorem is proved. □

From Theorem 4.2, Definition 4.3, Proposition 4.2 and Proposition 3.1, we have the following

Corollary 4.2. *K is an inclusive Sperner-system if and only if K^{-1} is a saturated one.*

Corollary 4.2 was shown in [12].
From Corollary 4.1 and Corollary 4.2 the following corollary is obvious:

Corollary 4.3. *Let K be a Sperner-system over R. Denote H a Sperner-system for which $H^{-1} = K$. Then the followings are equivalent:*
(1) K is saturated,
(2) K^{-1} is embedded,
(3) H is inclusive.

References

[1] Armstrong W. W., *Dependency Structure of Database Relationship*, Information Processing 74, North-Holland Pub. Co. (1974) 580-583.

[2] Beeri C., Bernstein P. A., *Computational Problems Related to the Design of Normal Form Relation Scheme*, ACM Trans. on Database Syst. **4, 1** (1979) 30-59.

[3] Berge C., *Hypergraphs: Combinatorics of Finite Sets*, North - Holland, Amsterdam (1989).

[4] Demetrovics J.,*Logical and structural investigation of relation datamodel*, MTA SZTAKI Tanulm-ányok, **114** (1980), 1-97 (in Hungarian).

[5] Demetrovics J., *On the equivalence of candidate keys with Sperner systems*, Acta Cybernetica **4**, (1979), 247-252.

[6] Demetrovics J., F üredi Z., Katona G., *A függőségek és az individumok száma közötti kapcsolat összetett adatrendszerek esetén*, Alkalmazott Matematikai Lapok **9** (1983),13-21.

[7] Demetrovics J., Thi V. D., *Describing Candidate Keys by Hypergraphs*, Computers and Artificial Intelligence **18, 2** (1999), 191-207.

[8] Demetrovics J., Thi V.D., *On the time complexity of algorithms related to Boyce-Codd normal form*, SERDICA-Bulgaricae Mathematicae Publicationes, **19**, (1993), 134-144.

[9] Demetrovics J., Thi V. D., *Some computational problems related to Boyce-Codd normal form*, Annales Univ. Sci. Budapest. Sect. Comp. **19**, (2000), 119-132.

[10] Gottlob G., Libkin L., *Investigations on Armstrong relations, denpendency inference, and excluded functional dependencies*, Acta Cybernetica Hungary **9, 4** (1990), 385-402.

[11] Lucchesi C. L., Osborn S. L., *Candidate keys for relations*, J. Comput. Syst. Scien. **17, 2** (1978), 270-279.

[12] Thi V. D., *Minimal keys and Antikeys*, Acta Cybernetica **7, 4** (1986), 361-371.

Received January, 2004