Serdica

Bulgariacae mathematicae publicationes

Сердика

Българско математическо списание

SERDICA — Bulgaricae mathematicae publicationes
19 (1993) 134-144

Mathematical Structures - Informatics - Mathematical Modelling
Papers dedicated to Academician L. Iliev on the Occasion of his Eighthieth Birthday

ON THE TIME COMPLEXITY OF ALGORITHMS RELATED TO BOYCE-CODD NORMAL FORM[†]

JANOS DEMETROVICS, VU DUC THI

Dedicated to Academician Ljubomir Iliev on the Occasion of his Eightieth Birthday

ABSTRACT. The functional dependency (FD) was introduced by E. F. Codd. Several types of families of FDs which satisfy some conditions are known under the name normal forms (NFs). The most desirable NF is Boyce-Codd NF (BCNF) that has been investigated in a lot of paper. It is shown [4] that every set of attributes with an associated set of FDs has a decomposition into third NF which has the loss-less-join property and preserves FDs. However, for BCNF this does not always exists. The key is an interesting concept in the relational datamodel. In this paper we investigate the time complexity of algorithms and problems concerning sets of all minimal keys of relation schemes and relation in BCNF class. We give two algorithms that find a BCNF relation r such that r represents a given BCNF relation scheme (i.e. $K_r = K_s$, where K_r and K_s are sets of all minimal keys of r and s). This paper also gives two algorithms which from a given BCNF relation find a BCNF relation scheme such that $K_r = K_s$. We estimate these algorithms. We prove that in BCNF class the time complexity of problem that finds a BCNF relation representing a given BCNF relation scheme s is exponential in the size of s. Conversly, we show that the complexity of finding a BCNF relation scheme s from a given BCNF realtion r such that r represents s also is exponential in the number of attributes. The concept of Armstrong relation for FD was introduced by R. Fagin. In database theory it is studied by many researchers. It is known that in BCNF class a relation r represents a relation scheme s iff r is an Armstrong relation of s. Consequently, our four algorithms and two problems are still true when r is an Armstrong relation of s.

1. Introduction. Let us give some necessary definitions and results that are used in next section.

† Research supported by Hungarian Foundation for Scientific Research Grant 2575

Definition 1.1. *Let $R = \{a_1, \ldots, a_n\}$ be a nonempty finite set of attributes, $r = \{h_1, \ldots, h_m\}$ be a relation over R, and $A, B \subseteq R$.*

Then we say that B functionally depends on A in r (denoted $A \xrightarrow[r]{f} B$) iff

$$(\forall h_i, h_j \in r)((\forall a \in A)(h_i(a) = h_j(a)) \Longrightarrow (\forall b \in B)(h_i(b) = h_j(b)))$$

Let $F_r = \{(A, B) : \ A, B \subseteq R, \ A \xrightarrow[r]{f} B\}$. F_r is called the full family of functional dependencies of r. Where we write (A, B) or $A \to B$ for $A \xrightarrow[r]{f} B$ when r, f are clear form the context.

Definition 1.2. *A functional dependency over R is a statement of the form $A \to B$, where $A, B \subseteq R$. The FD $A \to B$ holds in a relation r if $A \xrightarrow[r]{f} B$. We also say that r satisfies the FD $A \to B$.*

Clearly, F_r is a set of all FDs that hold in r.

Definition 1.3. *Let R be a nonempty finite set, and denote $P(R)$ its power set. Let $y \subseteq P(R) \times P(R)$. We say that y is an f-family over R iff for all A, B, C, $D, \subseteq R$*

(1) $(A, A) \in y$,

(2) $(A, B) \in y$, $(B, C) \in y \Longrightarrow (A, C) \in y$,

(3) $(A, B) \in y$, $A \subseteq C, D \subseteq B \Longrightarrow (C, D) \in y$,

(4) $(A, B) \in y$, $(C, D) \in y \Longrightarrow (A \cup C, B \cup D) \in y$.

Clearly, F_r is an f-family over R.

It is known [1] that if y is an arbitrary f-family, then there is a relation r over R such that $F_r = y$.

Definition 1.4. *A relation scheme s is a pair $< R, F >q$, where R is a set of attributes, and F is a set of FDs over R. Let F^+ be a set of all FDs that can be derived from F by the rules in definition 1.3. Denote $A^+ = \{a : A \to \{a\} \in F^+\}$. A^+ is called the closure of A over s. It is clear that $A \to B \in F^+$ iff $B \subseteq A^+$.*

Clearly, if $s =< R, F >$ is a relation scheme, then there is a relation r over R such that $F_r = F^+$ (see, [1]). Such a relation is called an Armstrong relation of s. It is obvious that all FDs of s hold in r.

Definition 1.5. *Let r be a relation, $s =< R, F >$ be a relation scheme, y be an f-family over R and $A \subseteq R$. Then A is a key of r (a key of s, a key of y) if $A \xrightarrow[r]{f} R$ ($A \to R \in F^+$, $(A, R) \in y$). A is a minimal key of $r(s, y)$ if A a key of $r(s, y)$, and any proper subset of A is not a key of $r(s, y)$. Denote $K_r, (K_s, K_y)$ the set of all minimal keys of $r(s, y)$.*

Clearly, K_r, K_s, K_y are Sperner systems over R.

Definition 1.6. *Let K be a Sperner system over R. We define the set of antikeys of K, denoted by K^{-1}, as follows:*

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

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

It is known [6] that if K is an arbitrary Sperner system over R then there is a relation scheme s such that $K_s = K$.

In this paper we always assume that if a Sperner system plays the role of the set of minimal keys (antikeys), then this Sperner system is not empty (does not contain R). We consider the comparison of two attributes as an elementary step of algorithms. Thus, if we assume that subsets of R are represented as sorted lists of attributes, then a Boolean operation on two subsets of requires at most $|R|$ elementary steps.

Definitiom 1.7. *Let $I \subseteq P(R), R \in I$, and $A, B \in I \Longrightarrow A \cap B \in I$. Let $M \subseteq P(R)$. Denote $M^+ = \{\cap M' : M' \subseteq M\}$. We say that M is a generator of I iff $M^+ = I$. Note that $R \in M^+$ but not in M, since it is the intersection of the empty collection of sets.*

Denote $N = \{A \in I : A \neq \cap\{A' \in I : A \subseteq A'\}\}$.

In [6] it is proved that N is the unique minimal generator of I. Thus, for any generator N' of I we obtain $N \subseteq N'$.

Definition 1.8. *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 $T_r = \{A \in P(R) : \exists E_{ij} = A, \nexists E_{pq} : A \subset E_{pq}\}$. Then T_r is called the maximal equality system of r.*

Definition 1.9. *Let r be a relation, and K a Sperner system over R. We say that r represents K iff $K_r = K$.*

The following theorem is known ([8]).

Theorem 1.10 *Let K be a non-empty Sperner system and r a relation over R. Then r represents K iff $K^{-1} = T_r$, where T_r is the maximal equality system of r.*

Let $s = \langle R, F \rangle$ be a relation scheme over R, K_s is a set of all minimal keys of s. Denote K_s^{-1} the set of all antikeys of s. From Theorem 1.10 we obtain the following corollary.

Corollary 1.11 *Let $s = \langle R, F \rangle$ be a relation scheme and r a relation over R. We say thet r represents s if $K_r = K_s$. Then r represents s iff $K_s^{-1} = T_r$, where T_r is the maximal equality system of r.*

In [7] we proved the following theorem.

Theorem 1.12 *Let $r = \{h_1, \ldots, h_m\}$ be a relation, and F an f-family over R. Then $F_r = F$ iff for every $A \in P(R)$*

$$H_F(A) = \begin{cases} \bigcap\limits_{A \subseteq E_{ij}} E_{ij} & \text{if } \exists E_{ij} \in E_r : A \subseteq E_{ij} \\ \\ R & \text{otherwise} \end{cases}$$

where $H_F(A) = \{a \in R : (A, \{a\}) \in F\}$ *and* E_r *is the equality set of* r.

We say that a relation scheme $s =< R, E >$ (a realtion r) is in BCNF if $\forall A \subseteq R$ either $A^+ = A$ or $A^+ = R$ ($H_{F_r}(A) = A$ or $H_{F_r}(A) = R$).

2. Results. All relation and realtion schemes are investigated in this section are in BCNF. We construct four algorithms concerning minimal keys of relations and relation schemes. We estimate these algorithms. We present two problems the worst-case time complexity of which are exponential.

Let $s =< R, F >$ be a relation scheme over R. From s we construct $Z(s) = \{X^+ : X \subseteq R\}$, and compute the minimal generator N_s of $Z(s)$. We put

$$T_s = \{A \subseteq N_s : \nexists B \in N_s : A \subset B\}.$$

It is known [1] that for a given relation scheme s there is a relation r such that r is an Armstrong relation of s. On the other hand, by Corollary 1.11 and Theorem 1.12 the following proposition is clear.

Proposition 2.1. *Let* $s =< R, F >$ *be a relation scheme over* R. *Then*

$$K_s^{-1} = T_s$$

Definition 2.2 *Let* $s =< R, F >$ *be a relation scheme. We say that* s *is a* k*-relation scheme over* R *if* $F = \{K_1 \to R, \ldots, K_m \to R\}$, *where* $\{K_1, \ldots, K_m\}$ *is a Sperner system over* R. *It is easy to see that* $K_s = \{K_1, \ldots, K_m\}$.

Clearly, if $s =< R, F >$ is in BCNF then use the algorithm for finding a minimal $F^+ = F'^+$, see [10]. Conversely, it can be seen that an arbitrary k-relation scheme is in BCNF. Consequently, we can consider a relation scheme in BCNF as a k-relation scheme.

Remark 2.3 It is known [10] that $s =< R, F >$ is in BCNF iff its minimum cover is a k-relation scheme. Consequently, the BCNF property of s is polynomially recognizable.

Let r be a relation over R. From r we compute E_r. We construct the maximal equality system T_r of r. By Theorem 1.10 we obtain $T_r = K_r^{-1}$. Denote elements of T_r by $A_1, \ldots, A_t$.

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

For all $a \in R$, $h_0(a) = 0$, for each $i = 1, \ldots, s$ $h_i(a) = 0$ if $a \in B_i$, in the converse case we set $h_i(a) = i$.

By [10] r' is in BCNF and $K_r = K_{r'}$ (1). It is easy to see that M_r and r' are constructed in polynomial time in the size of r.

Set $H_{F_r}(A) = \{a \in R : (A, \{a\}) \in F_r\}$, $Z_{F_r} = \{A : H_{F_r}(A) = A\}$. Denote by N_{F_r} the minimal generator of Z_{F_r}.

Based on definition of BCNF relation and from (1) we can see that a relation r is in BCNF iff $N_{F_r} = N_{F_r'}$. Because for an arbitrary relation r $N_{F_r'}$ is computed in polynomial time, the BCNF property of r can be tested in polynomial time.

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

Algorithm 2.4
Input: a BCNF relation scheme $s = < R, F >$
Output: a BCNF relation r such that $K_r = K_s$
Step 1: From s we construct $Z(s) = \{X^+ \ : \ X \subseteq R\}$
Step 2: Compute the minimal generator N_s of $Z(s)$
Step 3: Compute $T_s = \{A \in N_s \ : \ \nexists B \in N_s \ : \ A \subset B\}$. Denote elements of T_s by $A_1, \ldots, A_t$.
Step 4: Set $Q_s = \{A_i - a \ : \ a \in R, \ i = 1, \ldots, t\}$. Denote elements of Q_s by $B_1, \ldots, B_l$.
Step 5: Construct a relation $r = \{h_0, h_1, \ldots, h_l\}$ as follows:
For all $a \in R$, $h_0(a) = 0$, for each $i = 1, \ldots, l$ $h_i(a) = 0$ if $a \in B_i$, in the converse case we set $h_i(a) = i$.

Based on Proposition 2.1 and Remark 2.3 we obtain $K_r = K_s$ and r is in BCNF. It is easy to see that the time complexity of Algorithm 2.4 is exponential in the number of attributes.

It is known [15] that there is an algorithm that finds a set of all antikeys from a given Sperner system.

Algorithm 2.5 (Finding a set of antikeys)
Input: Let $K = \{B_1, \ldots, B_m\}$ be a Sperner system over R
Output: K^{-1}
Step 1: We set $K_1 = \{R - \{a\} \ : \ a \in B_1\}$. It is obviouse that $K_1 = \{B_1\}^{-1}$
Step $q + 1$: $(q < m)$. We assume that $K_q = F_q \cup \{X_1, \ldots, X_{t_q}\}$, where $X_1, \ldots, X_{t_q}$ containing B_{q+1} and $F_q = \{A \in K_q \ : \ B_{q+1} \not\subseteq A\}$. For all i $(i = 1, \ldots, t_q)$ we construct the antikeys of $\{B_{q+1}\}$ on X_i in an analogouse way as K_1. Denote them by $A^i_1, \ldots, A^i_{r_i}$ $(i = 1, \ldots, t_q)$. Let

$$K_{q+1} = F_q \cup \{A^i_p : A \in F_q \Longrightarrow A^i_q \not\subset A, \ 1 \leq i \leq t_q, \ 1 \leq p \leq r_i\}.$$

We set $K^{-1} = K_m$.

Theorem 2.6 [15] *For every q $(1 \leq q \leq m)$, $K_q = \{B_1, \ldots, B_q\}^{-1}$, i.e. $K_m = K^{-1}$.*

It can be seen that K and K^{-1} are uniquely determined by one another and the determination of K^{-1} based on our algorithm does not depend on the order of

$B_1, \ldots, B_m$. Denote $K_q = F_q \cup \{X_1, \ldots, X_{t_q}\}$ and l_q $(1 \leq q \leq m-1)$ is the number of elements of K_q.

Proposition 2.7 [15] *The worst-case time complexity of Algorithm 2.5 is*

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

where

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

Clearly, in each step of our algorithm K_q is a Sperner system. In cases for which $l_q \leq l_m$ $(q = 1, \ldots, m-1)$ it is easy to see that the time complexity of our algorithm is not greater than $O(|R|^2|K||K^{-1}|^2)$. Thus, in these cases Algorithm 2.5 finds K^{-1} in polynomial time in $|R|, |K|$, and $|K^{-1}|$. It can be seen that if the number of elements of K is small then Algorithm 2.5 is very effective. It only requires polynomial time in $|R|$.

By Algorithm 2.5 we construct a second algorithm that finds a relation such that this relation represents a given relation scheme. By Remark 2.3 it is simple that we can consider an arbitrary relation scheme in BCNF as a k-relation scheme.

Algorithm 2.8

Input: $s = < R, F = \{K_1 \to R, \ldots, K_m \to R\} >$ be a k-relation scheme

Output: a BCNF relation r such that $K_r = K_s$

Step 1: From $K = \{K_1, \ldots, K_m\}$ we construct $K^{-1} = \{B_1, \ldots, B_t\}$ by Algorithm 2.5

Step 2: Set $M = \{B_i - a \ : \ a \in R, i = 1, \ldots, t\}$

Step 3: Denote elements of M by $A_1, \ldots, A_l$, construct a relation $r = \{h_0, h_1, \ldots, h_l\}$ as follows: For all $a \in R \ : \ h_0(a) = 0$. For $i = 1, \ldots, l$ we set $h_i(a) = 0$ if $a \in A_i$, in the converse case $h_i(a) = i$.

By Remark 2.3, Corollary 1.11 we obtain $K_r = K_s$, and r is a BCNF relation.

Clearly, set M and relation r are constructed in polynomial time in the size of K^{-1}. Consequently, the time complexity of this algorithm is $O(|R|^3 \sum_{q=1}^{m-1} t_q u_q)$, meaning of t_q, u_q see Proposition 2.7. In many cases this algorithm requires polynomial time in the size of s (see Proposition 2.7).

Now we construct two algorithms which find a BCNF relation scheme such that a given BCNF relation represents this relation scheme.

Given a relation scheme $s = < R, F >$, we say that a functional dependency $A \to B \in F$ is redundant if either $A = B$ or there is $C \to B \in F$ such that $C \subsetneq A$.

Algorithm 2.9

Input: a BCNF relation $r = \{h_1, \dots, h_m\}$ over R.

Output: a BCNF relation scheme $s =< R, F >$ such that $K_r = K_s$.

Step 1: Find the equality set $E_r = \{E_{ij} \ : \ 1 \leq i < j \leq m\}$

Step 2: Find the maximal equality system T_r. Denote elements of T_r by $A_1, \dots, A_t$.

Step 3: For each $B \subseteq R$ there is not A_i such that $B \subseteq A_i$, we set $B \to R$. Denote T the set of all such functional dependencies.

Step 4: Set $F = T - Q$, where $Q = \{X \to Y \in T \ : \ X \to Y$ is a redundant functional dependency$\}$.

Clearly, by Theorem 1.12, Algorithm 2.9 finds a k-relation s such that a given BCNF relation r represents s, and r is an Armstrong relation of s (i.e. $F_r = F^+$). It can be seen that the time complexity of Algorithm 2.9 is exponential in the number of attributes.

Algorithm 2.10 [8] (Finding a minimal key from a set of antikey)

Input: Let K be a Sperner system, H a Sperner system, and $C = \{b_1, \dots, b_m\} \subseteq R$ such that $H^{-1} = K$ and $\exists B \in K \ : \ B \subseteq C$.

Output: $D \in H$

Step 1: Set $T(0) = C$

Step $i+1$: Set $T = T(i) - b_{i+1}$

$$T(i+1) = \begin{cases} T & \text{if } \forall B \in K \ : \ T \not\subseteq B \\ & \\ T(i) & \text{otherwise} \end{cases}$$

We set $D = T(m)$.

Lemma 2.11 [8] *If K is a set of antikeys, then $T(m) \in H$*

Lemma 2.12 [8] *Let H be a Sperner system over R, and $H^{-1} = \{B_1, \dots, B_m\}$ be a set of antikeys of H, $T \subseteq H$. Then $T \subset H$, $T \neq \emptyset$ if and only if there is a $B \subseteq R$ such that $B \in T^{-1}$, $B \not\subseteq B_i$ $(\forall i : 1 \leq i \leq m)$.*

Based on Lemma 2.12 and from Algorithm 2.10 we have the following algorithm.

Algorithm 2.13 (Finding a set of minimal key from a set of antikeys)

Input: Let $K = \{B_1, \dots, B_k\}$ be a Sperner system over R.

Output: H such that $H^{-1} = K$.

Step 1: By Algorithm 2.10 we outpute an A_1, set $K(1) = A_1$.

Step $i+1$: If there exists a $B \in K_i^{-1}$ such that $B \not\subseteq B_j$ $(\forall j : 1 \leq j \leq m)$, then by Algorithm 2.10 we compute an A_{i+1}, where $A_{i+1} \in H$, $A_{i+1} \subseteq B$. Set $K(i+1) = K(i) \cup A_{i+1}$. In the converse case we set $H = K(i)$.

Proposition 2.14 [16] *The time complexity of Algorithm* 2.13 *is*

$$O(n(\sum_{q=1}^{m-1}(kl_q + nt_q u_q) + k^2 + n)),$$

where $|R| = n$, $|K| = k$, $|H| = m$, *meaning of* l_q, t_q, u_q *see Proposition 2.7.*

Clearly, in case for which $l_q \le k$ $(\forall q: 1 \le q \le m-1)$ the time complexity of our algorithm is $O(|R|^2|K|^2|H|)$. It is easy to that in these case Algorithm 2.13 finds the set of minimal keys in polynomial time in the size of R, K, H. If $|H|$ is polynomial in $|R|$ and $|K|$, then our algorithm is effective. It can be seen that if the number of elements of H is small then Algorithm 2.13 is very effective.

Algorithm 2.15
Input: Let r be a BCNF relation over R
Output: a BCNF relation scheme s such that $K_s = K_r$
Step 1: From r compute E_r
Step 2: From E_r compute the maximal equality system T_r
Step 3: By Algorithm 2.13 we construct a set of all minimal keys H of r
Step 4: Denoting elements of H by $A_1, \dots, A_m$ we construct a relation scheme as follows: $s = < R, F >$, where $F = \{A_1 \to R, \dots, A_m \to R\}$.

Based on Theorem 1.10, Algorithm 2.13 and Definition 2.2 we have $K_s = K_r$. It is clear that the time complexity of this algorithm is the time complexity of Algorithm 2.13. In many cases this algorithm is very effective (see Proposition 2.14).

Theorem 2.16 [14] *Let* K *is a Sperner system over* R. *Denote* $s(K) = \min\{m : |r| = m, K_r = K\}$. *Then* $(2|K^{-1}|)^{1/2} \le s(K) \le |K^{-1}| + 1$.

Remark 2.17. Let us take a partition $R = X_1 \cup, \dots, \cup X_m \cup W$, where $|R| = n$, $m = [n/3]$, and $|X_i| = 3$ $(1 \le i \le m)$.

We set

$H = \{B : |B| = 2, B \subseteq X_i \text{ for some } i\}$ if $|W| = 0$,

$H = \{B : |B| = 2, B \subseteq X_i \text{ for some } i : 1 \le i \le m-1 \text{ or } B \subseteq X_m\}$ if $|W| = 1$,

$H = \{B : |B| = 2, B \subseteq X_i \text{ for some } i : 1 \le i \le m \text{ or } B = W\}$ if $|W| = 2$,

It is easy to see that

$H^{-1} = \{A : |A \cap X_i| = 1, \forall i\}$ if $|W| = 0$,

$H^{-1} = \{A : |A \cap X_i| = 1, (1 \le i \le m-1) \text{ and } |A \cap (X_m \cup W)| = 1\}$ if $|W| = 1$,

$H^{-1} = \{A : |A \cap X_i| = 1, (1 \le i \le m) \text{ and } |A \cap W| = 1\}$ if $|W| = 2$,

If set $K = {H^{-1}}^{-1}$, i.e. H^{-1} is a set of minimal keys of K, then we have

$K = \{C : |C| = n-3, C \cap X_i = \emptyset \text{ for some } i\}$ if $|W| = 0$,

$K = \{C : |C| = n-3, C \cap X_i = \emptyset \text{ for some } i \ (1 \le i \le m-1) \text{ or } |C| = n-4, C \cap (X_m \cup W) = \emptyset\}$ if $|W| = 1$,

$K = \{C : |C| = n-3, C \cap X_i = \emptyset$ for some i $(1 \le i \le m)$ or $|C| = n-2, C \cap W = \emptyset\}$ if $|W| = 2$.

It is clear that $n - 1 \le |H| \le n + 2$, $3^{[n/4]} < |H^{-1}|$, $|K| \le m + 1$. Besed on this partition, Theorem 2.16, and Algorithm 2.8, 2.9 we obtain the following theorems.

Theorem 2.18. *In* BCNF *class of relations and relation schemes, the time complexity of finding a relation r from a given relation scheme s such that r represent s is exponential in the size of s.*

Proof. We have to prove that:

(1) There is an algorithm finding a BCNF relation r from a given BCNF relation scheme s such that r represent s and the time complexity of this algorithm is exponential time in the size of s.

(2) There exists a BCNF relation scheme $s = < R, F >$ such that the number of rows of any BCNF relation representing s is exponential in the size of s.

For (1): We have Algorithm 2.8.

For (2): According to Theorem 2.16 we have $(2|K^{-1}|)^{1/2} \le s(K)$. We construct a k-relation scheme $s = < R, F >$, where $F = \{B \to R : B \in H\}$. It is obvious that $H^{-1} = K_s^{-1}$. Hence, $(2^{1/2}3^{[n/8]}) \le s(K_s)$ holds. It can be seen that BCNF relation r that is constructed in Algorithm 2.8 has the number of rows at most $|U||H^{-1}| + 1$. Thus, we always can construct a BCNF relation scheme s such that the number of rows of any BCNF relation representing s is exponential in the size of s. The proof is complete.

Theorem 2.19. *In BCNF class of relation and relation schemes over R, the time complexity of finding a relation scheme s from a given relation r such that $K_r = K_s$ is exponential in the number of attributes.*

Proof. It is clear that the worst-case time complexity of Algorithm 2.9 is exponential in the size of R. In Remark 2.17 we have $|K| \le m + 1$. We set $M = \{C - a : \forall a, C : a \in R, C \in K\}$. Denote elements of M by $C_1, \ldots, C_t$. Construct a relation $r = \{h_0, h_1, \ldots, h_t\}$ as follows: For all $a \in R$ $h_0(a) = 0$, for $i = 1, \ldots, t$, $h_i(a) = 0$ if $a \in C_i$, in the converse case $h_i(a) = i$. Clearly, $|r| \le (m + 1)|R| + 1$ holds. We construct a relation scheme $s = < R, F >$ with $F = \{A \to R : A \in H^{-1}\}$. It is obvious that $3^{[n/4]} < |F|$, and $K_r = K_s$. Clearly, a minimum cover of any BCNF relation scheme is a k-relation scheme. Thus, we always can construct a BCNF relation r in which the number of rows of r is at most $(m+1)|R| + 1$ but for any BCNF relation scheme $s = < R, F >$ such that $K_r = K_s$, the number of elements of F is exponential in the number of attributes. Our proof is complete.

It is known that in BCNF class a relation r represents a relation scheme s iff r is an Armstrong relation of s. Consequently, our four algorithms and two problems are still true when r is an Armstrong relation of s.

Acknowledgments. The authors are grateful to Dr. Uhrin Bela for useful comments to the first version of the manuscript.

REFERENCES

[1] W. W. ARMSTRONG, Dependency Structures of Database Relationships, Information Processing 74, Holland publ. Co. 1974, 580-583.

[2] C. BEERI, M. DOWD, R. FAGIN, R. STAMAN, On the Structure of Armstrong relations for Functional Dependenies, *J. ACM.* **31** (1984), 30-46.

[3] C. BEERI, P. A. BERNSTEIN, Computational problems related to the design of normal form relational schemas, *ACM Trans. on Database syst.* **4** (1979), 30-59.

[4] P. A. BERNSTEIN, Synthesizing third normal form relations from functional dependencies, *ACM Trans. Database System 1* **1** (1976), 227-298.

[5] G. BUROSCH, J. DEMETROVICS, G. O. H. KATONA, The poset of closure as a model of changing database, Order 4, 1987, 227-298.

[6] J. DEMETROVICS, Relacios adatmodell logikai es structuralis vizsgalata *MTA-SZTAKI Tanulmanyok* **114** (1980), 1-97 (in Hungarian).

[7] J. DEMETROVICS, V. D. THI, Some results about functional dependencies, *Acta Cybernetica Hungary*, **VIII/3** (1988), 273-278.

[8] J. DEMETROVICS, V. D. THI, Relations and minimal keys, *Acta Cybernetica Hungary* **VIII/3** (1988), 279-285.

[9] J. DEMETROVICS, V. D. THI, On key in the Relational Datamodel, *EIK* **24** (1988), 515-519.

[10] G. GOTTLOB, L. LIBKIN, Investigation on Armstrong relations, dependency inference, and excluded functional dependencies, *Acta Cybernetica Hungary*, **9** (1990), 385-402.

[11] C. L. LUCCHETI, S. L. OSBORN, Candidate keys for relation, *J. Comp. Syst. Scien.* **17** (1978), 270-279.

[12] H. MANNILA, K. J. RAIHA, Design by example: An Application of Armstrong relations, *J. Comp. Syst. Scien.* **33** (1986), 126-141.

[13] H. MANNILA, K. J. RAIHA, Algorithms for inferring functional dependencies, Proc. of the thirteenth international conference on very large data bases, Brighton, September 1987.

[14] V. D. THI, Funkcional fuggoseggel kapcsolatos nehany kombinatorikai jellegu vizsgalat a relacios adatmodellben, *MTA-SZTAKI Tanulmanyok* **191** (1986), 1-157 (in Hungarian).

[15] V. D. THI, Minimal key and Antikeys, *Acta Cybernetica Hungary* **7** (1986), 361-371.

[16] V. D. THI, Relacios adatmodell antikulcsairol, *Alkalmazott Matematikai Lapok* **12** (1986), 111-124 (in Hungarian).

Computer and Automation Institute
Hungarian Academy of Sciences
P.O.Box 63, Budapest, H-1502
HUNGARY