On the Number of Independent Functional Dependencies*

János Demetrovics[1], Gyula O.H. Katona[2], Dezső Miklós[2], and Bernhard Thalheim[3]

[1] Computer and Automation Institute, Hungarian Academy of Sciences, Kende u. 13-17, H-1111 Hungary
dj@ilab.sztaki.hu
[2] Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Budapest P.O.B. 127, H-1364 Hungary
ohkatona@renyi.hu, dezso@renyi.hu
[3] Institute for Computer Science and Practical Mathematics, Christian-Albrechts-University, Olshausenstr. 40 D-24098 Kiel, Germany
thalheim@is.informatik.uni-kiel.de

Abstract. We will investigate the following question: what can be the maximum number of independent functional dependencies in a database of n attributes, that is the maximum cardinality of a system of dependencies which which do not follow from the Armstrong axioms and none of them can be derived from the remaining ones using the Armstrong axioms. An easy and for long time believed to be the best construction is the following: take the maximum possible number of subsets of the attributes such that none of them contains the other one (by the wellknown theorem of Sperner [8] their number is $\binom{n}{n/2}$) and let them all determine all the further values. However, we will show by a specific construction that it is possible to give more than $\binom{n}{n/2}$ independent dependencies (the construction will give $(1+\frac{1}{n^2})\binom{n}{n/2}$ of them) and — on the other hand — the upper bound is $2^n - 1$, which is roughly $\sqrt{n}\binom{n}{n/2}$.

1 Introduction

Results obtained during database design and development are evaluated on two main criteria: *completeness* of and *unambiguity* of specification. Completeness requires that all constraints that must be specified are found. Unambiguity is necessary in order to provide a reasoning system. Both criteria have found their theoretical and pragmatical solution for most of the known classes of constraints. Completeness is, however, restricted by the human ability to survey large constraint sets and to understand all possible interactions among constraints.

* The work was supported by the Hungarian National Foundation for Scientific Research grant numbers T037846 and AT048826.

Many database theory and application problems (e.g., data search optimization, database design) are substantially defined by the *complexity* of a database, i.e., the size of key, functional dependency, and minimal key systems. Most of the known algorithms, e.g., for normalization, use the set of all minimal keys or non-redundant sets of dependencies. Therefore, they are dependent on the cardinality of these sets. The maintenance complexity of a database depends on how many integrity constraints are under consideration. Therefore, if the cardinality of constraint sets is large, then maintenance becomes infeasible. (Two-tuple constraints such as functional dependencies require $O(m^2)$ two-tuple comparisons for relations with m elements.) Furthermore, they indicate whether algorithms are of interest for practical purposes, since the complexity of most known algorithms is measured by the input length. For instance, algorithms for constructing a minimal key are bound by the maximal number of minimal keys. The problem of deciding whether there is a minimal key with at most k attributes is NP-complete. The problem of deciding whether two sets of functional dependencies are equivalent is polynomial in the size of the two sets.

Therefore, the database design process may only be complete if all integrity constraints that cannot be derived by those that have already been specified have been specified. Such completeness is not harmful as long as constraint sets are small. The number of constraints may however be exponential in the number of attributes [2]. Therefore, specification of the *complete set of functional dependencies* may be a task that is `infeasible`. This problem is closely related to another well-known combinatoric problem presented by Janos Demetrovics during MFDBS'87 [9] and that is still only partially solved:

Problem 1. *How big the number of independent functional dependencies of an n-ary relation schema can be?*

Let R be a relational database model, and X denote the set of attributes. We say that (for two subsets of attributes A and B) $A \to B$, that is, B functionally depends on A, if in the database R the values of the attributes in A uniquely determine the values of the attributes in B. In case a some functional dependencies $\mathcal{F}$ given, the closure of $\mathcal{F}$, usually denoted by $\mathcal{F}^+$, is the set of all functional dependencies that may be logically derived from $\mathcal{F}$. E.g., $\mathcal{F}$ may be considered the obvious and important functional dependencies (like mother's name and address uniquely determine the name of a person) and then the closure, $\mathcal{F}^+$ is the set of all dependencies that can be deduced from $\mathcal{F}$.

The set of rules that are used to derive all functional dependencies implied by $\mathcal{F}$ were determined by Armstrong in 1974 and are called the *Armstrong axioms.* These rules are easily seen to be necessary and all other natural rules can be derived from them. They are the following:

- *reflexivity rule* if A is a set of attributes and B a subset of it, then $A \to B$.
- *augmentation rule* If $A \to B$ holds and C is an arbitrary set of attributes, then $A \cup C \to B \cup C$ holds as well.
- *transitivity rule* If $A \to B$ and $B \to C$ hold, then $A \to C$ holds as well.

Let us mention here, that though there are further natural rules of the dependencies, the above set is complete, that is $\mathcal{F}^+$ can always be derived from $\mathcal{F}$ using only the above three axioms. For example, union rule, that is, the natural fact that $A \to B$ and $A \to C$ imply $A \to B \cup C$ can be derived by augmenting $A \to C$ by A ($A \to A \cup C$), augmenting $A \to B$ by C ($A \cup C \to B \cup C$) and using transitivity for the resulting rules: $A \to A \cup C \to B \cup C$.

In this paper we will investigate the maximum possible number of independent functional dependencies of a database of n attributes. That is, the maximum of $\mathcal{F}$, where it is a system of independent, non-trivial dependencies ($A \to B$ where $B \subset A$ are not in $\mathcal{F}$) and no element of $\mathcal{F}$ can be logically derived from the other elements of $\mathcal{F}$. In this case we will call $\mathcal{F}$ independent.

Introduce the following useful notations: $[n] = \{1, 2, \ldots, n\}$. The family of all k-element subsets of $[n]$ is

$$\binom{[n]}{k}.$$

For the sake of simplicity we will denote the i^{th} attribute by i.

It is clear that $A \to C$ and $B \to C$ can not be in $\mathcal{F}$ for a pair of subsets $A \subset B$ since then $B \to C$ would be logically obtained by another given dependency ($A \to C$), reflexivity ($A \subset B$ implies $B \to A$) and transitivity ($B \to A \to C$ implies $B \to C$). On the other hand, it is easy to see if we have a system of independent subsets of the attributes (that is, none of them containing the other one) and assume that all of them imply the whole set of attributes, this system of dependencies will be independent. This leads to the natural construction of a large set of independent dependencies by taking the maximum number of incomparable subsets of attributes, which is by Sperner's theorem [8] equal to $\binom{n}{\lfloor n/2 \rfloor}$ and let the whole set of attributes depend on all of them. This would give a set of dependencies of cardinality $\binom{n}{\lfloor n/2 \rfloor}$. On the other hand, it is easy to see that if $\mathcal{F}$ only consists of dependencies $A \to B$ with $|A| = k$ for a given constant k and for every such an $A \subset X$ there is at most one element of $\mathcal{F}$ of the form $A \to B$, then $\mathcal{F}$ is independent (a more detailed version of this argument will be given in the proof of Lemma 4). That is, the above construction will give an independent set of dependencies and a lower bound of $\binom{n}{\lfloor n/2 \rfloor}$.

However, as it will be shown by the construction of the next section, this is not the best possible bound, we can enlarge it. Still the best known lower bound is of the magnitude of $\binom{n}{\lfloor n/2 \rfloor}$ (smaller than $c\binom{n}{\lfloor n/2 \rfloor}$ for any constant $c > 1$, while the best upper bound proven in the following section is $2^n - 1$, which is roughly $\sqrt{n}\binom{n}{n/2}$. Finally, the last section of the paper will contain concluding remarks, including the answer to the following question:

Problem 2. *Is the maximum number of functional dependencies the same as the maximum number of minimal keys?*

More complexity results are discussed and proven in [5, 7, 10] or in [1, 2, 3, 6].

2 Lower Estimate: A Construction

Theorem 1. *If n is an odd prime number then one can construct*

$$\left(1+\frac{1}{n^2}+o\left(\frac{1}{n^2}\right)\right)\binom{n}{\lfloor\frac{n}{2}\rfloor}$$

independent functional dependencies on an n-element set of attributes.

The proof will consist of a sequence of lemmas. We will also use the following proposition:

Proposition 1. (see [11]) *Assign to a functional dependency $A \to B$ the set of $2^n - 2^{|B|}$ Boolean vectors $\mathbf{a}= (a_1, ..., a_n)$ of the form:*

$$a_i = \begin{cases} 1\,, & \text{if } i \in A \\ 0 \text{ or } 1\,, & \text{if } i \in (B \setminus A) \text{ but not all entries } = 0 \\ 0 \text{ or } 1\,, & \text{otherwise.} \end{cases}$$

Then, a set of functional dependencies implies another functional dependency if and only if the Boolean vectors of the implied functional dependency are contained in the union of the sets of Boolean vectors of the given functional dependencies.

Lemma 2. *If n is an odd prime number, one can find $\frac{1}{n^2}\binom{n}{\frac{n+3}{2}}$ subsets $V_1, V_2, \ldots$ of size $\frac{n+3}{2}$ in the set $[n] = \{1, 2, \ldots n\}$ in such a way that $|V_i \cap V_j| < \frac{n-1}{2}$ holds.*

Proof. The method of the paper [4] is used. Consider the subsets $\{x_1, x_2, \ldots, x_{\frac{n+3}{2}}\}$ of integers satisfying $1 \leq x_i \neq x_j \leq n$ for $i \neq j$ and the equations

$$x_1 + x_2 + \cdots + x_{\frac{n+3}{2}} \equiv a \pmod n, \tag{1}$$

$$x_1 x_2 \cdot \cdots \cdot x_{\frac{n+3}{2}} \equiv b \pmod n \tag{2}$$

for some fixed integers a and b.

Suppose that two of them, say V_1 and V_2 have an intersection of size $\frac{n-1}{2}$. We may assume, without loss of generality, that the first two elements are different, that is $V_1 = \{x_1, x_2, \ldots, x_{\frac{n+3}{2}}\}$ and $V_2 = \{x_1', x_2', \ldots, x_{\frac{n+3}{2}}\}$. (1) and (2) imply $x_1 + x_2 \equiv x_1' + x_2'$ and $x_1 x_2 \equiv x_1' x_2' \pmod n$. Since the set $\{1, 2, \ldots n\}$ constitutes a field mod$n$ if n is prime, this system of equations has a unique solution, that is $x_1 = x_1', x_2 = x_2'$; the two sets are the same: $V_1 = V_2$. This contradiction proves that our sets cannot have $\frac{n-1}{2}$ common elements.

The total number of subsets of size $\frac{n+3}{2}$ is

$$\binom{n}{\frac{n+3}{2}}.$$

Each of these sets give some a and b in (1) and (2), respectively. That is, the family of all $\frac{n+3}{2}$-element sets can be divided into n^2 classes. One of these has a size at least

$$\frac{1}{n^2}\binom{n}{\frac{n+3}{2}}. \qquad \square$$

We will need the notion of the *shadow* of a family $\mathcal{A} \subset \binom{[n]}{\frac{n+1}{2}}$. It will be denoted by

$$\sigma(\mathcal{A}) = \{B : \; |B| = \frac{n-1}{2}, \quad \exists A \in \mathcal{A} : B \subset A\}.$$

A pair $\{U_1, U_2\}$ $U_i \in \binom{[n]}{\frac{n+1}{2}}$ is called *good* if $|U_1 \cap U_2| = \frac{n-1}{2}$ holds. The family $\mathcal{P} \subset \binom{[n]}{\frac{n+1}{2}}$ is a *chain* if $\mathcal{P} = \mathcal{P}_1 \cup \ldots \cup \mathcal{P}_l$ where the $\mathcal{P}_i$'s are good pairs and $\sigma(\mathcal{P}_i) \cap \sigma(\mathcal{P}_j) = \emptyset$ for $i \neq j$. The weight $w(\mathcal{P})$ of this chain is l.

Lemma 3. *There is a chain $\mathcal{P} \subset \binom{[n]}{\frac{n+1}{2}}$ chain with weight at least*

$$|\mathcal{P}| = \frac{1}{n^2}\binom{n}{\frac{n+3}{2}}.$$

Proof. Start with the family $\mathcal{V} = \{V_1, V_2, \ldots\}$ ensured by Lemma 2. In each V_i choose two different $\frac{n+1}{2}$-element subsets U_{i1} and U_{i2}. It is easy to see that this pair of subsets is good. Also, these pairs form a chain. The number of the pairs in the chain can be obtained from Lemma 2. $\square$

Lemma 4. *If $\mathcal{P} \subset \binom{[n]}{\frac{n+1}{2}}$ is a chain then the following set of functional dependencies is independent:*

$$A \to B, \quad \textit{where } |A| = \frac{n-1}{2}, \quad |B| = \frac{n+1}{2}, \quad A \subset B \in \mathcal{P}, \tag{3}$$

$$A \to A', \quad \textit{where } |A| = \frac{n-1}{2}, \quad A \notin \sigma(\mathcal{P})$$

$$\textit{and } A' \textit{ is arbitrarily chosen so that } A \subset A', \quad |A'| = \frac{n+1}{2}. \tag{4}$$

(Note that in (3) we have all dependencies $A \to B$ given by the conditions, while in (4) for every remaining A we choose only one (exactly one) A' satisfying the conditions.)

Proof. We will use Proposition 1 for the proof. According to rules (3) and (4), for every $A \subset X$ with $|A| = \frac{n-1}{2}$ there are either one (B) or two (B_1 and B_2) subsets of X of size $\frac{n+1}{2}$ with $A \to B$ or $A \to B_i$, always $A \subset B,\ B_i$. In the first case consider one of the Boolean vector $\boldsymbol{a}$ corresponding to $A \to B$:

$$a_i = \begin{cases} 1\,, \text{ if } \; i \in A \\ 0\,, \text{ if } \; i \in (B \setminus A) \\ 0\,, \text{ otherwise.} \end{cases}$$

This vector has exactly $\frac{n-1}{2}$ 1 entries, and from the definition in Proposition 1 it is also clear that all Boolean vectors corresponding to all functional dependencies given by (3) and (4) have at least $\frac{n-1}{2}$ 1 entries. Therefore, the Boolean vector $\boldsymbol{a}$ may correspond to any other dependency $A' \longrightarrow B'$ only with $A' = A$, which is not the case now, the dependency $A \longrightarrow B$ may not be deduced from the others.

If we have both $A \longrightarrow B_1$ and $A \longrightarrow B_2$, consider one of the Boolean vector $\boldsymbol{a}$ corresponding to $A \longrightarrow B_2$:

$$a_i = \begin{cases} 1\,, \text{ if } \ i \in A \\ 0\,, \text{ if } \ i \in (B_2 \setminus A) \\ 1\,, \text{ if } \ i \in (B_1 \setminus A) \\ 0\,, \text{ otherwise.} \end{cases}$$

In this case $\boldsymbol{a}$ has $\frac{n+1}{2}$ 1 entries and still all Boolean vectors corresponding to all other given functional dependencies have at least $\frac{n-1}{2}$ 0 entries. Therefore, if a Boolean vector corresponding to a given functional dependency $A' \longrightarrow B'$ is equal $\boldsymbol{a}$, the obligatory $\frac{n-1}{2}$ 1 entries must form a subset of the $\frac{n+1}{2}$ 0 entries of $\boldsymbol{a}$, or, in other words, A' must be a subset of B_1, an $\frac{n+1}{2}$ element set. However, by the construction according to (3), for all subsets C of B_1 of size $\frac{n-1}{2}$ we have $C \longrightarrow B_1$. Since B_1 is larger than C only by one element, all the corresponding Boolean vectors must have entries 1 at the positions corresponding to C and entry 0 at the only position corresponding to $B_1 \setminus C$ (this entry could be literally 0 or 1, but not all of them 1, but since it is alone, it means it should be 0). This, by Proposition 1, implies that $A \longrightarrow B_2$ may not be deduced from the set of remaining given Boolean vectors. □

Alternative proof. One can prove Lemma 4 without the use of Proposition 1, simply from the Amstrong axioms, as it follows.

Again, we will show that none of the dependencies given in the lemma can be deduced from the other ones.

Now we start with the case when the dependency $A \longrightarrow B_1$ has a pair $A \longrightarrow B_2 (B_1 \neq B_2)$ in our system. This can happen only in case of (3): $A \longrightarrow B_1$ where $\mathcal{P} = \{B_1, B_2\}, B_1 = A \cup \{b_1\}, B_2 = A \cup \{b_2\}$. We want to verify that $A \longrightarrow B_1$ cannot be deduced from the other ones.

Observe that none of the other ones, $X \longrightarrow Y$ satisfy both $X \subseteq B_2$ and $Y \nsubseteq B_2$. In other words

$$\text{either} \quad X \nsubseteq B_2 \quad \text{or} \quad Y \subseteq B_2 \tag{5}$$

holds for every dependency given by (3) and (4), different from $A \longrightarrow B_1$. Let us see that this property is preserved by the Armstrong axioms.

It is trivial in the case of the reflexivity rule, since it gives $X \longrightarrow Y$ only when $Y \subseteq X$ therefore $X \subseteq B_2$ implies $Y \subseteq B_2$.

Consider the augmentation rule. Suppose that $X \longrightarrow Y$ satisfies (5) and U is an arbitrary set. If $X \nsubseteq B_2$ then the same holds for $X \cup U$, that is $X \cup U \nsubseteq B_2$. On the other hand, if $X \subseteq B_2$ then $Y \subseteq B_2$. If $U \subseteq B_2$ also holds then $Y \cup U \subseteq B_2$, if however $U \nsubseteq B_2$ then $X \cup U \nsubseteq B_2$. The dependency $X \cup U \longrightarrow Y \cup U$ obtained by the augmentation rule also satisfies (5) in all of these cases.

Finally, suppose that $X \to Y$ and $Y \to Z$ both satisfy (5). We have to show the same for $X \to Z$. If $X \not\subseteq B_2$ then we are done. Suppose $X \subseteq B_2$ and $Y \subseteq B_2$. Then $Z \subseteq B_2$, must hold, as desired.

Since $A \to B_1$ does not have property (5), it cannot be deduced from the other dependencies.

Consider now the case when the distinguished dependency $A \to B_1$ has no pair $A \to B_2 (B_1 \neq B_2)$ in our system. This can happen both for (3) and (4). The the proof is similar to the case above. No other dependency $X \to Y$ satisfies $X \subseteq A, Y \not\subseteq A$, that is (5) holds for them if B_2 is replaced by A. □

Remark. One may think that a better construction can be made if we allow three sets B_1, B_2, B_3 of size $\frac{n+1}{2}$ with pairwise intersections $|B_1 \cap B_2| = |B_2 \cap B_3| = \frac{n-1}{2}$ where no other intersections (of these three and other sets $B_i (3 < i)$ of this size) are that big, and $A \to B_i$ holds for every subset A of B_i. Unfortunately we found a counter-example for $n > 5$.

Let $B_1 = \{1, \ldots, \frac{n+1}{2}\}, B_2 = \{2, \ldots, \frac{n+3}{2}\}, B_3 = \{3, \ldots, \frac{n+5}{2}\}$, $B_4 = \{1, \ldots, \frac{n-3}{2}, \frac{n+3}{2}, \frac{n+5}{2}\}$. It is easy to see that $|B_1 \cap B_2| = |B_2 \cap B_3| = \frac{n-1}{2}$, but $|B_1 \cap B_3|, |B_1 \cap B_4|, |B_2 \cap B_4|, |B_3 \cap B_4|$ are all smaller.

Introduce the notations $A_1 = \{2, \ldots, \frac{n+1}{2}\}, A_2 = \{3, \ldots, \frac{n+3}{2}\}, A_3 = \{2, 3, \ldots, \frac{n-3}{2}, \frac{n+3}{2}, \frac{n+5}{2}\}$. The the following chain of functional dependencies is obvious: $A_1 \to B_2 \to A_2 \to B_3 \to B_2 \cup B_3 \to A_3 \to B_4 \to \{1\}$. Hence we have $A_1 \to B_1$ without using it.

Proof of Theorem 1. Use Lemma 4 with the chain found in Lemma 3. It is easy to see that there is at least one $A \to C$ for every $\lfloor \frac{n}{2} \rfloor = \frac{n-1}{2}$-element subset A and the weight of the chain gives the number of those A being the left-hand side of exactly two dependencies. This gives the number of dependencies:

$$\binom{n}{\lfloor \frac{n}{2} \rfloor} + \frac{1}{n^2}\binom{n}{\frac{n+3}{2}} = \left(1 + \frac{1}{n^2} + o\left(\frac{1}{n^2}\right)\right)\binom{n}{\lfloor \frac{n}{2} \rfloor}.$$

□

3 Upper Estimate

Theorem 5. *For every n an upper bound for the maximum number of independent functional dependencies on an n-element set of attributes is $2^n - 1$.*

Proof. Let $\mathcal{F}$ be a set of independent functional dependencies on the set of attributes X. First, replace each dependency $A \to B$ in $\mathcal{F}$ by $A \to A \cup B$, obtaining $\mathcal{F}'$. We claim that $\mathcal{F}'$ is independent as well. It simply comes from the fact that by the reflexivity and augmentation axioms the two dependencies $A \to B$ and $A \to A \cup B$ are equivalent. Also, $|\mathcal{F}| = |\mathcal{F}'|$, since the images of dependencies in $\mathcal{F}$ will be different in $\mathcal{F}'$. Assume, on the contrary, that $A \to A \cup B$ is equal to $A \to A \cup C$ for $A \to B$ and $A \to C$ in $\mathcal{F}$. But then $A \cup B = A \cup C$, $C \subset A \cup B$ and therefore $A \to B$ implies $A \to A \cup B$ implies $A \to C$, a contradiction.

We may therefore consider only set of independent dependencies where for all $(A \to B) \in \mathcal{F}$ we have $A \subset B$. Take now the following graph G: let the vertices of the graph be all the 2^n subsets of the n attributes and for $A,\ B \subset X$ the edge (A, B) will be present in the G iff $A\ toB$ or $B \to A$ is in $\mathcal{F}$. We claim that this graph may not contain a cycle, therefore it is a forest, that is it has at most $2^n - 1$ edges, or dependencies.

Assume, on the contrary, that $A_1, A_2, A_3, \ldots, A_n = A_1$ is a cycle in G, that is for all $i = 1, \ldots n-1$ the edge $(A_i, A_{i+1}$ is present in G, meaning that either $A_i \to A_{i+1}$ or $A_{i+1} \to A_i$ is in $\mathcal{F}$. Note that for every $i = 1, \ldots n-1$ we have $A_{i+1} \to A_i$ since either this dependency is in $\mathcal{F}$ or in case of $(A_i \to A_{i+1}) \in \mathcal{F}$, we have that $A_i \subset A_{i+1}$, yielding $A_{i+1} \to A_i$ by reflexivity. Take now an i such that $A_i \to A_{i+1}$ (if we have no such a dependency, take the "reverse" of the cycle, $A_n, A_{n-1}, \ldots, A_1 = A_n$). This can be obtained from the other dependencies present in $\mathcal{F}$ by the transitivity chain $A_i \to A_{i-1} \to \cdots A_1 = A_n \to A_{n-1} \to \cdots \to A_{i+2} \to A_{i+1}$, contradicting the independency of the rules in $\mathcal{F}$. $\square$

We have an alternative proof using Proposition 1 as well.

4 Remarks, Conclusions

The main contribution of the paper is the improvement of upper and lower bounds for independent sets of functional dependencies and, thus, contributing to the solution of Problem 1.

1. The lower and upper estimates seem to be very far from each other. However, if the lower estimate is written in the form

$$c\frac{2^n}{\sqrt{n}}$$

(using the Stirling formula) then one can see that the "difference" is only a factor $\sqrt{n}$ what is negligable in comparison to 2^n. The logarithms of the lower and upper estimates are $n - \frac{1}{2}\log n$ and n.

However we strongly believe that truth is between

$$\left(1 + \frac{\alpha}{n} + o\left(\frac{1}{n}\right)\right)\binom{n}{\lfloor \frac{n}{2} \rfloor}$$

and

$$(\beta + o(1))\binom{n}{\lfloor \frac{n}{2} \rfloor},$$

where $0 < \alpha$ and $\beta \leq 2$.

2. Theorem 1 is stated only for odd prime numbers. The assumption in Lemma 2 is only technical, we strongly believe that its statement is true for other integers, too. (Perhaps with a constant less than 1 over the n^2.) We did not really try to prove this, since the truth in Theorem 1 is more, anyway. What do we obtain from Lemma 2 if it is applied for the largest prime less than n? It is known from number theory that there is a prime p satisfying $n - n^{5/8} < p \leq n$. This will lead to an estimate where $\frac{1}{n^2}$ is replaced by

$$\frac{1}{n^2 2^{n^{5/8}}}.$$

This is much weaker than the result for primes, but it is still more than the number of functional dependencies in the trivial construction.

3. One may have the feeling that keys are the real interesting objects in a dependency system. That is, the solution of any extremal problem must be a set of keys. Our theorems show that this is not the case.

More precisely, suppose that only keys are considered in our problem, that is the maximum number of independent keys is to be determined. If this set of keys is $\{A_i \to X\}$ $(1 \leq i \leq m)$, then $A_i \not\subset A_j$ must be satisfied, and therefore by Sperner's theorem $m \leq \binom{n}{\lfloor n/2 \rfloor}$. In this case the largest set of dependencies (keys) is provided by the keys $A \to X$, where $A \in \binom{[n]}{\lfloor \frac{n}{2} \rfloor}$.

Theorem 1 shows that the restriction to consideration of key systems during database design and development is an essential restriction. Systems of functional dependencies must be considered in parallel. Therefore, we derived a negative answer to Problem 2.

References

1. Beeri,C., Dowd, M., Fagin, R., and Statman, R., On the structure of Armstrong relations for functional dependencies. *Journal of ACM*, **31**, 1, January 1984, 30–46.
2. Demetrovics, J. and Katona, G.O.H., Combinatorial problems of database models. *Colloquia Mathematica Societatis Janos Bolyai 42,* Algebra, Cominatorics and Logic in Computer Science, Gÿor (Hungary), 1983, 331–352.
3. Demetrovics, J., Libkin, L., and Muchnik, I.B., Functional dependencies and the semilattice of closed classes. *Proc. MFDBS-89*, LNCS 364, 1989, 136–147.
4. Graham, R., Sloane, N., Lower bounds for constant weight codes, *IEEE Trans. Inform. Theory* **26**(1980) 37-43.
5. Katona, G.O.H. and Demetrovics, J., A survey of some combinatorial results concerning functional dependencies in relational databases. Annals of Mathematics and Artificial Intelligence, 6, 1992.
6. Mannila, H. and K.-J. Räihä, K.-J., *The design of relational databases.* Addison-Wesley, Amsterdam, 1992.
7. Seleznjev, O. and Thalheim, B., On the number of minimal keys in relational databases over nonuniform domains. *Acta Cybern.*, **8**, 267–271. 1988.
8. Sperner, E., Ein Satz über Untermengen einer endlichen Menge, *Math. Z.* **27**(1928) 544-548.
9. Thalheim, B., Open problems in relational database theory. *Bull. EATCS*, **32**:336 - 337, 1987.
10. Thalheim, B., Generalizing the entity-relationship model for database modeling. *JNGCS, 1990*, **3**, 3, 197 - 212.
11. Thalheim, B., *Dependencies in Relational Databases.* Teubner-Texte zur Mathematik, B.G. Teubner Verlagsgesellschaft, Stuttgrat - Leipzig, 1991.