On the functional dependency and some generalizations of it

By J. DEMETROVICS and GY. GYEPESI

§ 0. Introduction

According to E. F. CODD [6] a relation is a matrix without two identical rows. Rows correspond to data records and columns to the attributes that are to be stored of a data item. He also introduced [7] the concept of *functional dependency:* a set of columns depends on another if fixing the values in a row taken on the first determines those on the second.

Other concepts of his are the *key* (a set of attributes on which all depend) and the *candidate key* (a minimal key).

Candidate keys clearly do not contain each other [10].

The possible mathematical structure of functional dependencies was first investigated by W. W. ARMSTRONG [1]. Among others he found that this structure is determined by the *maximal dependencies* (those which have maximal attribute subsets depending on minimal ones) and even by the dependent sides of the maximal dependencies. We also heavily use these "maximal dependent subsets of attributes" as technical tools.

Different kinds of functional dependency have also been introduced [3], [11], [13], [14], and axiomatized, usually in systems similar to those investigated by Armstrong.

The harder problems of the topic are usually of combinatorial nature (see [4], [5], [9], [15]).

In this paper in § 1 we give the formal definition of the functional, dual, strong and weak dependencies and give new axioms for full *f*- *d*- and *s*-families.

In § 2 we show the analogy and differences among the dependencies of different types and give an axiom for full *w*-families.

In § 3 we deal with a question stated in [9].

Before starting § 1 we make some remarks concerning the practice:

The functional, dual, strong and weak dependencies studied in this paper are those restrictions which allow the characterization of a relation by restrictions of it to certain proper subsets of the attribute set.

Certain dependencies of a relational data base are known by its designer. We call these *initial dependencies.* In general initial dependencies imply new dependencies. W. W. ARMSTRONG [2] has developed a method to find the dependencies implied by a given set of initial functional dependencies. He also gave a characterization of the sets of initial dependencies that imply all the dependencies of a given full f-family and are of minimal cardinality. This characterization has a logical nature; we give a combinatorial equivalent of it.

We use the following notational conventions: Ω denotes the set of attributes, $P(\Omega)$ denotes its power set. If g is a function with X as its domain and $Z \subseteq X$ then $g \restriction Z$ denotes the function which has domain Z and for any $z \in Z$ $g(z) = g \restriction Z(z)$. $\subset$ means strict inclusion.

§ 1. Old and new axioms

We start with the definitions of functional, dual, strong and weak dependencies based on [1] and [8].

Definition 1.1. Let A, B be subsets of Ω and let R be a relation over Ω. Then we say that B

(i) *functionally;*
(ii) *dually;*
(iii) *strongly;*
(iv) *weakly*

depends on A in R if

(i) $(\forall g, h \in R)(g \restriction A = h \restriction A \Rightarrow g \restriction B = h \restriction B)$;

(ii) $(\forall g, h \in R)\big((\exists a \in A)(g(a) = h(a)) \Rightarrow (\exists b \in B)(g(b) = h(b))\big)$;

(iii) $(\forall g, h \in R)\big((\exists a \in A)(g(a) = h(a)) \Rightarrow g \restriction B = h \restriction B\big)$;

(iv) $(\forall g, h \in R)\big(g \restriction A = h \restriction A \Rightarrow (\exists b \in B)(g(b) = h(b))\big)$;

holds respectively and denote these by $A \overset{f}{\underset{R}{\rightarrow}} B$, $A \overset{d}{\underset{R}{\rightarrow}} B$, $A \overset{s}{\underset{R}{\rightarrow}} B$, $A \overset{w}{\underset{R}{\rightarrow}} B$ corresponding to the type of the denoted dependency.

The following example [8] illustrates the effect of the dual dependency.

EXAMPLE. Let $\Omega = \{\text{author, title, hall, shelf}\}$. Let we have a library with eighteen books, three halls and three shelves in every hall; one shelf holds two books. Let the relation R containing the data of the library given by the following table:

author	*title*	*hall*	*shelf*	*author*	*title*	*hall*	*shelf*
1	1	1	2	10	10	3	2
2	2	1	3	11	11	3	3
3	3	1	1	12	12	3	1
4	4	1	2	1	4	1	1
5	5	2	3	5	8	3	3
6	6	2	1	4	1	1	3
7	7	2	2	7	10	3	2
8	8	2	3	6	10	2	2
9	9	3	1	6	9	2	1

Thus {author, title} $\xrightarrow[R]{d}$ {hall, shelf} holds, and for $i=1, \ldots, 12$ the book by author i and entitled i is on the $\left(1+3\cdot\left\{\frac{i}{3}\right\}\right)$-th shelf of the $\left[\frac{i+3}{4}\right]$-th hall ($[x]$ denotes the integer part and $\{x\}$ the fraction part of x). The reader, knowing the author or the title of the required book, may find it without examining the whole library: for example if i is the author of the book, then it is enough to look the $\left[\frac{i+3}{4}\right]$-th hall, and the $\left(1+3\cdot\left\{\frac{i}{3}\right\}\right)$-th shelves of the other two halls.

In R {author, title} $\xrightarrow[R]{f}$ {hall, shelf} holds too, but to store this functional dependency is equivalent to store the table of R; the {author, title} $\xrightarrow[R]{d}$ {hall, shelf} dependency is more effective.

If R is a relation over Ω, $\mathcal{Y}\in\{\mathcal{F}, \mathcal{D}, \mathcal{S}, \mathcal{W}\}$ and $y\in\{f, d, s, w\}$ corresponds to $\mathcal{Y}$, then we write

$$\mathcal{Y}_R = \{(A, B)\colon A \xrightarrow[R]{y} B\}.$$

We call the sets which have the form $\mathcal{Y}_R$ *full y-families*, where y corresponds to $\mathcal{Y}$.

In order to investigate the various dependencies the first step is the axiomatization of full y-families for $y\in\{f, d, s, w\}$. In [1] there is a system of axioms for full f-family and in [8] there are for full d- and s-families. For the sake of completeness we reproduce them here.

Let $\mathcal{Y}\subseteq P(\Omega)\times P(\Omega)$. Then we say that $\mathcal{Y}$ satisfies the $\mathfrak{f}$-axioms, if for all $A, B, C, D\subseteqq\Omega$

(F1) $(A, A)\in\mathcal{Y}$;

(F2) $(A, B)\in\mathcal{Y},\quad (B, C)\in\mathcal{Y}\Rightarrow(A, C)\in\mathcal{Y}$;

(F3) $(A, B)\in\mathcal{Y},\quad C\supseteqq A,\quad D\subseteqq B\Rightarrow(C, D)\in\mathcal{Y}$;

(F4) $(A, B)\in\mathcal{Y},\quad (C, D)\in\mathcal{Y}\Rightarrow(A\cup C, B\cup D)\in\mathcal{Y}$.

$\mathcal{Y}$ satisfies the ϑ-axioms if for all $A, B, C, D\subseteqq\Omega$

(D1) $(A, A)\in\mathcal{Y}$;

(D2) $(A, B)\in\mathcal{Y},\quad (B, C)\in\mathcal{Y}\Rightarrow(A, C)\in\mathcal{Y}$;

(D3) $(A, B)\in\mathcal{Y},\quad C\subseteqq A,\quad B\subseteqq D\Rightarrow(C, D)\in\mathcal{Y}$;

(D4) $(A, B)\in\mathcal{Y},\quad (C, D)\in\mathcal{Y}\Rightarrow(A\cup C, B\cup D)\in\mathcal{Y}$;

(D5) $(A, \emptyset)\in\mathcal{Y}\Rightarrow A=\emptyset$.

$\mathcal{Y}$ satisfies the γ-axioms if for all $A, B, C, D\subseteqq\Omega$ and for any $a\in\Omega$

(S1) $(\{a\}, \{a\})\in\mathcal{Y}$;

(S2) $(A, B)\in\mathcal{Y},\quad (B, C)\in\mathcal{Y},\quad B\neq\emptyset\Rightarrow(A, C)\in\mathcal{Y}$;

(S3) $(A, B)\in\mathcal{Y},\quad C\subseteqq A,\quad D\subseteqq B\Rightarrow(C, D)\in\mathcal{Y}$;

(S4) $(A, B)\in\mathcal{Y},\quad (C, D)\in\mathcal{Y}\Rightarrow(A\cap C, B\cup D)\in\mathcal{Y};$

(S5) $(A, B)\in\mathcal{Y},\quad (C, D)\in\mathcal{Y}\Rightarrow(A\cup C, B\cap D)\in\mathcal{Y}.$

We need the following technical lemma.

Lemma 1.1. Let $\mathcal{F}\subseteq P(\Omega)\times P(\Omega)$ be such that $(X, Y)\in\mathcal{F}$ and $Y\neq\emptyset$ imply $X\neq\emptyset$. Then $\mathcal{F}$ satisfies the $\mathfrak{f}$-axioms iff $\mathcal{D}=\{(A, B)\colon (B, A)\in\mathcal{F}\}$ satisfies the ϑ-axioms.

Proof. Trivial by the $\mathfrak{f}$- and ϑ-axioms. (D5) makes necessary the assumption that $(X, Y)\in\mathcal{F}$ and $Y\neq\emptyset$ imply $X\neq\emptyset$. □

REMARK. The assumption $((X, Y)\in\mathcal{F}$ and $Y\neq\emptyset$ imply $X\neq\emptyset)$ in Lemma 1.1 is not an important restriction: if $\mathcal{F}$ satisfies the $\mathfrak{f}$-axioms let $\mathcal{F}'=\mathcal{F}\setminus\{(\emptyset, X)\colon X\neq\emptyset\}$. Then $\mathcal{F}'$ obviously satisfies the $\mathfrak{f}$-axioms and the critical assumption as well, and we have that $X\neq\emptyset$ implies $(X, Y)\in\mathcal{F}\Leftrightarrow(X, Y)\in\mathcal{F}'$.

In the following we give new axioms instead of the $\mathfrak{f}$- ϑ- and γ-axioms and give an axiom that characterizes the *weak full w-families* which is such a full w-family that whenever (X, Y) is an element of the family then X is not void.

F-axiom. Let $\mathcal{F}\subseteq P(\Omega)\times P(\Omega)$. Then we say that *$\mathcal{F}$ satisfies the F-axiom* if for any $(X, Y)\in P(\Omega)\times P(\Omega)\setminus\mathcal{F}$ there is an $E\subseteq\Omega$ such that

(i) $X\subseteq E$ and $Y\nsubseteq E$;

(ii) if $(X', Y')\in\mathcal{F}$ and $X'\subseteq E$ then $Y'\subseteq E$.

D-axiom. Let $\mathcal{D}\subseteq P(\Omega)\times P(\Omega)$. Then we say that *$\mathcal{D}$ satisfies the D-axiom* if for any $(X, Y)\in P(\Omega)\times P(\Omega)\setminus\mathcal{D}$ there is an $E\subseteq\Omega$ such that

(i) $X\cap E\neq\emptyset$ and $Y\cap E=\emptyset$;

(ii) if $(X', Y')\in\mathcal{D}$ and $X'\cap E\neq\emptyset$ then $Y'\cap E\neq\emptyset$.

S-axiom. Let $\mathcal{S}\subseteq P(\Omega)\times P(\Omega)$. Then we say that *$\mathcal{S}$ satisfies the S-axiom* if for any $(X, Y)\in P(\Omega)\times P(\Omega)\setminus\mathcal{S}$ there is an $E\subseteq\Omega$ such that

(i) $X\cap E\neq\emptyset$ and $Y\nsubseteq E$

(ii) if $(X', Y')\in\mathcal{S}$ and $X'\cap E\neq\emptyset$ then $Y'\subseteq E$.

W-axiom. Let $\mathcal{W}\subseteq P(\Omega)\times P(\Omega)$. Then we say that *$\mathcal{W}$ satisfies the W-axiom* if for any $(X, Y)\in P(\Omega)\times P(\Omega)\setminus\mathcal{W}$ there is an $E\subseteq\Omega$ such that

(i) $X\subseteq E$ and $Y\cap E=\emptyset$;

(ii) if $(X', Y')\in\mathcal{W}$ and $X'\subseteq E$ then $Y'\cap E\neq\emptyset$.

Theorem 1.1. (i) Let $\mathcal{F}\subseteq P(\Omega)\times P(\Omega)$. Then $\mathcal{F}$ satisfies the $\mathfrak{f}$-axioms iff $\mathcal{F}$ satisfies the F-axiom.

(ii) Let $\mathcal{D}\subseteq P(\Omega)\times P(\Omega)$. Then $\mathcal{D}$ satisfies the ϑ-axioms iff $\mathcal{D}$ satisfies the D-axiom.

(iii) Let $\mathcal{S}\subseteq P(\Omega)\times P(\Omega)$. Then $\mathcal{S}$ satisfies the γ-axioms iff $\mathcal{S}$ satisfies the S-axiom.

Proof. (i) Suppose that $\mathcal{F}$ satisfies the F-axiom. Then

(F1) If $(A, A)\notin\mathcal{F}$ then there is an $E\subseteq\Omega$ such that $A\subseteq E$ and $A\nsubseteq E$ which is a contradiction.

(F2) If $(A, B) \in \mathscr{F}$, $(B, C) \in \mathscr{F}$ and $(A, C) \notin \mathscr{F}$ then there is an $E \subseteq \Omega$ such that $A \subseteq E$ and $C \nsubseteq E$. Furthermore $(A, B) \in \mathscr{F}$, $A \subseteq E$ imply $B \subseteq E$, and using $(B, C) \in \mathscr{F}$, $C \subseteq E$ which is a contradiction.

(F3) If $(A, B) \in \mathscr{F}$, $A' \supseteq A$, $B' \subseteq B$ and $(A', B') \notin \mathscr{F}$ then there is an $E \subseteq \Omega$ such that $A' \subseteq E$ and $B' \nsubseteq E$ and $(A, B) \in \mathscr{F}$, $A \subseteq E$ imply that $B \subseteq E$. Thus, by $B' \subseteq B$, $B' \subseteq E$ which is again a contradiction.

(F4) If $(A, B) \in \mathscr{F}$, $(C, D) \in \mathscr{F}$ and $(A \cup C, B \cup D) \notin \mathscr{F}$ then there is an $E \subseteq \Omega$ such that $A \cup C \subseteq E$ and $B \cup D \nsubseteq E$; e.g. $B \nsubseteq E$. But $(A, B) \in \mathscr{F}$ and $A \subseteq E$ imply that $B \subseteq E$, which is a contradiction.

Suppose now that $\mathscr{F}$ satisfies the $\mathfrak{f}$-axioms. Let $(A, B) \in P(\Omega) \times P(\Omega) \setminus \mathscr{F}$.

Claim. There is an $E \supseteq A$ such that $(E, B) \in P(\Omega) \times P(\Omega) \setminus \mathscr{F}$ and $E' \supset E$ implies $(E', B) \in \mathscr{F}$.

$(\Omega, \Omega) \in \mathscr{F}$ by (F1). Thus, by (F3), $(\Omega, B) \in \mathscr{F}$ holds. $A \subseteq \Omega$ and $(A, B) \in P(\Omega) \times \times P(\Omega) \setminus \mathscr{F}$, consequently there is an $E \subseteq \Omega$ which is maximal w.r. to the properties $(E, B) \in \mathscr{F}$ and $E \supseteq A$.

This E clearly satisfies the restrictions of the Claim.

Let $E \supseteq A$ which is guaranteed by the Claim. We state that E satisfies (i) and (ii) of the F-axiom. Namely, by the choice of E, $A \subseteq E$ holds. By (F1) and (F3), $B \subseteq E$ implies $(E, B) \in \mathscr{F}$. Thus we have $B \nsubseteq E$.

Let $(C, D) \in \mathscr{F}$ and $C \subseteq E$. $D \nsubseteq E$ implies $E' = D \cup E \supset E$, and, by the maximality of E, $(E', B) \in \mathscr{F}$ holds. Since $(E, D) \in \mathscr{F}$, by (F3), and $(E, E) \in \mathscr{F}$ by (F1), we have $(E, E') \in \mathscr{F}$. Now $(E, E') \in \mathscr{F}$, and $(E', B) \in \mathscr{F}$ and (F2) imply that $(E, B) \in \mathscr{F}$, which is a contradiction.

(ii) Let $\mathscr{F} = \{(A, B) : (B, A) \in \mathscr{D}\}$. Then, by Lemma 1.1, $\mathscr{F}$ satisfies the $\mathfrak{f}$-axioms iff $\mathscr{D}$ satisfies the ϑ-axioms. Hence, by (i), it is enough to show that $\mathscr{F}$ satisfies the F-axiom iff $\mathscr{D}$ satisfies the D-axiom.

Suppose that $\mathscr{F}$ satisfies the F-axiom. For $(A, B) \in P(\Omega) \times P(\Omega) \setminus \mathscr{F}$ let $E(A, B)$ be a subset of Ω such that $A \subseteq E(A, B)$, $B \nsubseteq E(A, B)$ and if both $(A', B') \in \mathscr{F}$ and $A' \subseteq E(A, B)$ hold, then $B' \subseteq E(A, B)$. By the F-axiom such an $E(A, B)$ exists. By the definition of $\mathscr{F}$ whenever $(A, B) \in P(\Omega) \times P(\Omega)$ then $(A, B) \in P(\Omega) \times P(\Omega) \setminus \mathscr{F}$ iff $(B, A) \in P(\Omega) \times P(\Omega) \setminus \mathscr{D}$.

Now it is easy to check that for $(B, A) \in P(\Omega) \times P(\Omega) \setminus \mathscr{D}$, $\Omega \setminus E(A, B)$ satisfies the D-axiom.

If $\mathscr{D}$ satisfies the D-axiom, then $\mathscr{F}$ satisfies the F-axiom; this can be shown by the same argument.

(iii) Suppose that $\mathscr{S}$ satisfies the S-axiom. Then the proof of the fact that $\mathscr{S}$ satisfies the γ-axioms is an easy modification of the proof of (i). We deal with but (S1) and (S2).

(S1) if $(\{a\}, \{a\}) \notin \mathscr{S}$, then there is an $E \subseteq \Omega$ such that $\{a\} \cap E \neq \emptyset$ and $\{a\} \cap (\Omega - E) \neq \emptyset$ which contradicts $|\{a\}| = 1$.

(S2) if $(A, B) \in \mathscr{S}$, $(B, C) \in \mathscr{S}$, $B \neq \emptyset$ and $(A, C) \notin \mathscr{S}$ then there is an $E \subseteq \Omega$ such that

(a) $A \cap E \neq \emptyset$

(b) $C \cap (\Omega \setminus E) \neq \emptyset$ and

(c) $(F, D) \in \mathscr{S}$, $F \cap E \neq \emptyset$ imply that $D \subseteq E$.

(a) and (c) imply $B \subseteq E$. By $B \neq \emptyset$ we have $B \cap E \neq \emptyset$. Hence, by $(B, C) \in \mathscr{S}$, $C \subseteq E$ holds which is a contradiction.

Suppose now that $\mathscr{S}$ satisfies the γ-axioms. Let $(A, B)\in P(\Omega)\times P(\Omega)\setminus\mathscr{S}$

Claim. There is an $a\in A$ and an $E\subseteq\Omega$ such that
(a) $a\in E$;
(b) $(\{a\}, E)\in\mathscr{S}$ and
(c) $E'\supset E$ implies that $(\{a\}, E')\notin\mathscr{S}$.

If for any $a\in A$ we have $(\{a\}, B)\in\mathscr{S}$ then $(A, B)\in\mathscr{S}$ by the repeated application of (S5).[1] Hence there is an $a\in A$ such that $(\{a\}, B)\notin\mathscr{S}$. Now if for every $b\in B$ $(\{a\}, \{b\})\in\mathscr{S}$ holds then by the repeated application of (S4) we have $(\{a\}, B)\in\mathscr{S}$. Thus there is a $b\in B$ such that $(\{a\}, \{b\})\notin\mathscr{S}$.

By (S1) and (S3) there is an $E\subseteq\Omega$ such that $a\in E$, $(\{a\}, E)\in\mathscr{S}$ and E is maximal w.r. to this property. This E is appropriate for the Claim.

Let $E\subseteq\Omega$ and $a\in A$ guaranteed by the Claim. Then by (S3) we have $b\notin E$. Hence $A\cap E\neq\emptyset$ and $B\cap(\Omega\setminus E)\neq\emptyset$. Now let $(C, D)\in\mathscr{S}$ such that $C\cap E\neq\emptyset$; let $c\in C\cap E$. Suppose that $D\cap(\Omega\setminus E)\neq\emptyset$; let $d\in D\cap(\Omega\setminus E)$. By (S3) we have $(\{c\}, \{d\})\in\mathscr{S}$ and by (S1) we have $(\{c\}, \{c\})\in\mathscr{S}$. $(\{a\}, E)\in\mathscr{S}$ implies that $(\{a, c\}, \{c\})\in\mathscr{S}$, by (S5). Hence (S3) implies that $(\{a\}, \{c\})\in\mathscr{S}$. Now $(\{a\}, \{c\})\in\mathscr{S}$, $(\{c\}, \{d\})\in\mathscr{S}$ and (S2) imply that $(\{a\}, \{d\})\in\mathscr{S}$. Thus by (S4) we have $(\{a\}, E\cup\{d\})\in\mathscr{S}$ which is a contradiction as $E'=E\cup\{d\}\supset E$.

Consequently the E guaranteed by the Claim demonstrates that $\mathscr{S}$ satisfies the S-axiom. □

§ 2. The equality-set

Definition 2.1. Let R be a relation over Ω. We define *the equality-set of* R, $\mathscr{E}_R$ as follows: For $h, g\in R$ let $E(h, g)=\{a\in\Omega\colon h(a)=g(a)\}$ and let $\mathscr{E}_R=\{E(h, g)\colon h, g\in R \text{ and } h\neq g\}$.

Definition 2.2. Let $\mathscr{A}$ be a-set system. Then $\mathscr{A}$ *is a Δ-system* if for any $A, B, C, D\in\mathscr{A}$, $A\neq B$ and $C\neq D$ implies that $A\cap B=C\cap D$.

Remark. It is easy to see that $\mathscr{A}$ is a Δ-system iff for any $A, B\in\mathscr{A}$, $A\neq B$ implies that $A\cap B=\bigcap\mathscr{A}$.

Theorem 2.1 (i) Let R be a relation over Ω and let h, f, g different elements of R. Then $E(h, g)$, $E(h, f)$, $E(g, f)$ form a Δ-system.

(ii) Let $\mathscr{E}=\{E_{i,j}\colon 1\leqq i<j\leqq k\}$ such that for each $1\leqq i<j<l\leqq k$, $\{E_{i,j}, E_{i,l}, E_{j,l}\}$ is a Δ-system. Then there is a relation R over Ω with $\mathscr{E}_R=\mathscr{E}$.

Proof. By symmetry it is enough to prove that $a\in E(h, g)\cap E(h, f)$ implies $a\in E(g, f)$. But $a\in E(h, g)\cap E(h, f)$ means that $g(a)=h(a)=f(a)$, hence $a\in E(g f)$.

We construct by induction the rows $h_1, \dots, h_k$ of R. Let $h_1(a)=0$ for each $a\in\Omega$, and assume that $n<k$ and the rows $h_1, \dots, h_n$ have been defined s.t. for each $1\leqq i<j\leqq n$, $E(h_i, h_j)=E_{i,j}$ holds. We construct h_{n+1} as follows:

$$h_{n+1}(a)=\begin{cases} h_i(a), & \text{if } a\in E_{i,n+1} \text{ for some } 1\leqq i\leqq n;\\ \max\big(h_i(b)\colon b\in\Omega \ \&\ 1\leqq i\leqq n\big)+1 & \text{else.}\end{cases}$$

[1] $A\neq\emptyset$, since $\forall b\in B$ $(\{b\},\{b\})\in\mathscr{S}$ by (S1), hence $(B, B)\in\mathscr{S}$ by (S4) and (S5). But if $A=\emptyset$, then $(A, B)\in\mathscr{S}$ by (S3).

Then

(a) h_{n+1} is well-defined.

To prove this we have to show that $a \in E_{i,n+1} \cap E_{j,n+1}$ implies $h_i(a) = h_j(a)$. But this is obvious because $E_{i,j}$, $E_{i,n+1}$, $E_{j,n+1}$ form a Δ-system and the induction hypothesis holds for $i, j \leqq n$.

(b) if $1 \leqq i \leqq n$ and $a \notin E_{i,n+1}$ then $h_i(a) \neq h_{n+1}(a)$.

Suppose first that $a \in E_{j,n+1}$ for some $1 \leqq j < n+1$. Then, by (a) and the definition of h_{n+1}, $h_{n+1}(a) = h_j(a)$ holds. Furthermore $a \notin E_{i,j}$ because $\{E_{i,j}, E_{j,n+1}, E_{i,n+1}\}$ is a Δ-system. Thus the induction hypothesis implies $h_i(a) \neq h_j(a)$, that is $h_i(a) \neq h_{n+1}(a)$.

If $a \notin \bigcup_{1 \leqq j \leqq n} E_{j,n+1}$ then we have $h_{n+1}(a) \neq h_i(a)$ by the definition of h_{n+1}. This completes the proof of (b).

Now by (a) and (b) it is clear that for $1 \leqq i \leqq n$, $E(h_i, h_{n+1}) = E_{i,n+1}$ and hence the induction step works. Let $R = \{h_1, \dots, h_k\}$. Then $\mathscr{E}_R = \mathscr{E}$ obviously holds. □

After Theorem 2.1 there is a natural way to axiomatize full families of dependencies of any type. This follows next:

F′-axiom. Let $\mathscr{F} \subseteq P(\Omega) \times P(\Omega)$. Then $\mathscr{F}$ *satisfies the F′-axiom* if there is a natural number k and an indexed set of subsets of Ω, $\{E_{i,j}: 1 \leqq i < j \leqq k\}$ such that

(i) If $(X, Y) \in P(\Omega) \times P(\Omega) \setminus \mathscr{F}$ then there are $1 \leqq i < j \leqq k$ such that $X \subseteq E_{i,j}$ and $Y \nsubseteq E_{i,j}$.

(ii) If $(X, Y) \in \mathscr{F}$, $1 \leqq i < j \leqq k$ and $X \subseteq E_{i,j}$ then $Y \subseteq E_{i,j}$.

(iii) For any $1 \leqq i < j < l \leqq k$, $\{E_{i,j}, E_{i,l}, E_{j,l}\}$ is a Δ-system.

D′-axiom. Let $\mathscr{D} \subseteq P(\Omega) \times P(\Omega)$. Then $\mathscr{D}$ *satisfies the D′-axiom* if there is a natural number k and an indexed set of subsets of Ω, $\{E_{i,j}: 1 \leqq i < j \leqq k\}$ such that

(i) If $(X, Y) \in P(\Omega) \times P(\Omega) \setminus \mathscr{D}$ then there are $1 \leqq i < j \leqq k$ such that $X \cap E_{i,j} \neq \emptyset$ and $Y \cap E_{i,j} = \emptyset$.

(ii) If $(X, Y) \in \mathscr{D}$, $1 \leqq i < j \leqq k$ and $X \cap E_{i,j} \neq \emptyset$ then $Y \subseteq E_{i,j} \neq \emptyset$.

(iii) The same as (iii) of the F'-axiom.

S′-axiom. Let $\mathscr{S} \subseteq P(\Omega) \times P(\Omega)$. Then $\mathscr{S}$ *satisfies the S′-axiom* if there is a natural number k and an indexed set of subsets of Ω, $\{E_{i,j}: 1 \leqq i < j \leqq k\}$ such that

(i) If $(X, Y) \in P(\Omega) \times P(\Omega) \setminus \mathscr{S}$ then there are $1 \leqq i < j \leqq k$ such that $X \cap E_{i,j} \neq \emptyset$ and $Y \nsubseteq E_{i,j}$.

(ii) If $(X, Y) \in \mathscr{S}$, $1 \leqq i < j \leqq k$ and $X \cap E_{i,j} \neq \emptyset$ then $Y \subseteq E_{i,j}$.

(iii) The same as (iii) of the F'-axiom.

W′-axiom. Let $\mathscr{W} \subseteq P(\Omega) \times P(\Omega)$. Then $\mathscr{W}$ *satisfies the W′-axiom* if there is a natural number k and an indexed set of subsets of Ω, $\{E_{i,j}: 1 \leqq i < j \leqq k\}$ such that

(i) If $(X, Y) \in P(\Omega) \times P(\Omega) \setminus \mathscr{W}$ then there are $1 \leqq i < j \leqq k$ such that $X \subseteq E_{i,j}$ and $Y \cap E_{i,j} = \emptyset$.

(ii) If $(X, Y) \in \mathscr{W}$, $1 \leqq i < j \leqq k$ and $X \subseteq E_{i,j}$ then $Y \cap E_{i,j} \neq \emptyset$.

(iii) The same as (iii) of the F'-axiom.

REMARK. Observe that the $E_{i,j}$-s in the F'-axiom are maximal dependent sets, i.e. if $(X, Y)\in\mathcal{F}$ and $X\subseteq E_{i,j}$ then $Y\subseteq E_{i,j}$.

Theorem 2.2. (i) Let $\mathcal{Y}\subseteq P(\Omega)\times P(\Omega)$ and $Y\in\{F, D, S\}$. Then $\mathcal{Y}$ satisfies the Y-axiom iff $\mathcal{Y}$ satisfies the Y'-axiom.

(ii) Let Ω be a finite set, $|\Omega|\geqq 3$. Then there is a $\mathcal{W}\subseteq P(\Omega)\times P(\Omega)$ such that $\mathcal{W}$ satisfies the W-axiom and $\mathcal{W}$ does not satisfy the W'-axiom.

Proof. (i) Let first $Y=F$ and suppose that $\mathcal{Y}$ satisfies the F-axiom. Write $\mathcal{Y}=\mathcal{F}$. For any $(X, Y)\in P(\Omega)\times P(\Omega)\setminus\mathcal{F}$ take an $E(X, Y)\subseteq\Omega$ guaranteed by the F-axiom. List these $E(X, Y)$-s as $E_2, \dots, E_k$ (the indices begin with 2). For $1<j\leqq k$ let $E_{1,j}=E_j$ and for $1<i<j\leqq k$ let $E_{i,j}=E_i\cap E_j$. We claim that $\{E_{i,j}: 1\leqq i<j\leqq k\}$ demonstrates that $\mathcal{F}$ satisfies the F'-axiom. The requirement (i) of the F'-axiom holds by $\{E_2, \dots, E_k\}\subseteq\{E_{i,j}: 1\leqq i<j\leqq k\}$. We left to the reader to check that (ii) holds too. To prove (iii) of the F'-axiom let $1\leqq i<j<l\leqq k$.

We distinguish two cases:

(a) $i=1$. Then $E_{i,j}=E_j$; $E_{i,l}=E_l$ and $E_{j,l}=E_j\cap E_l$. Thus the intersection of any two members of $\{E_{i,j};\ E_{i,l};\ E_{j,l}\}$ is $E_j\cap E_l$. This means that $\{E_{i,j};\ E_{i,l};\ E_{j,l}\}$ is Δ-system.

(b) $1<i$. Then $E_{i,j}=E_i\cap E_j$; $E_{i,l}=E_i\cap E_l$ and $E_{j,l}=E_j\cap E_l$. Thus the intersection of any two members of $\{E_{i,j};\ E_{i,l};\ E_{j,l}\}$ is $E_i\cap E_j\cap E_l$. This means that $\{E_{i,j};\ E_{i,l};\ E_{j,l}\}$ is a Δ-system.

If $\mathcal{Y}$ satisfies the F'-axiom then $\mathcal{Y}$ obviously satisfies the F-axiom.

Now let $Y=D$ and suppose that $\mathcal{Y}$ satisfies the D-axiom. Write $\mathcal{Y}=\mathcal{D}$. For any $(X, Y)\in P(\Omega)\times P(\Omega)\setminus\mathcal{D}$ take an $E(X, Y)\subseteq\Omega$ guaranteed by the D-axiom. List these $E(X, Y)$-s as $E_1, \dots, E_k$. For $1\leqq i\leqq k$ let $E_{2i-1,2i}=E_i$ and if $1\leqq i<j\leqq 2k$ and $E_{i,j}$ is still undefined then let $E_{i,j}=\emptyset$. It is easy to see that $\{E_{i,j}: 1\leqq i<j\leqq 2k\}$ shows the D'-axiom to hold for $\mathcal{D}$. If $\mathcal{Y}$ satisfies the D'-axiom then it trivially satisfies the D-axiom.

The case $Y=S$ is an easy modification of the proof in the case $Y=F$.

(ii) For the sake of simplicity suppose that $\Omega=\{a, b, c\}$. (In the general case pick two different elements a and b of Ω. The role of $\{c\}$ will played by $\Omega\setminus\{a, b\}$). Let $\mathcal{W}=\{(A, B)\in P(\Omega)\times P(\Omega): A\subseteq\{a\}\Rightarrow a\in B$ and $A\subseteq\{b\}\Rightarrow b\in B\}$. Then $\mathcal{W}$ satisfies the W-axiom while if $(A, B)\in P(\Omega)\times P(\Omega)\setminus\mathcal{W}$ then either ($A\subseteq\{a\}$ and $a\notin B$) or ($A\subseteq\{b\}$ and $b\notin B$). For (A, B), $E=\{a\}$ taken in the 1st case and $E=\{b\}$ in the 2nd one shows the W-axiom to hold.

We claim that $\mathcal{W}$ does not satisfy the W'-axiom. Suppose indirectly that $\mathcal{E}=\{E_{i,j}: 1\leqq i<j\leqq k\}$ is a system that shows the W'-axiom to hold for $\mathcal{W}$.

Then

(1) $\{a\}\in\mathcal{E}$ and $\{b\}\in\mathcal{E}$ while $(\{a\}, \Omega\setminus\{a\})\in P(\Omega)\times P(\Omega)\setminus\mathcal{W}$ and $(\{b\}, \Omega\setminus\{b\})\in P(\Omega)\times P(\Omega)\setminus\mathcal{W}$ hold.

(2) $\emptyset\notin\mathcal{E}$ and $\{c\}\notin\mathcal{E}$ while $(\emptyset, \Omega)\in\mathcal{W}$ and $(\{c\}, \Omega\setminus\{c\})\in\mathcal{W}$ hold.

By (1), $\{a\}\in\mathcal{E}$ and $\{b\}\in\mathcal{E}$, that is $\{a\}=E_{i,j}$ and $\{b\}=E_{l,m}$ for some $1\leqq i, j, l, m\leqq k$. We distinguish two cases:

(a) $i=l$. Then $\{a\}=E_{i,j}$ and $\{b\}=E_{i,m}$. $\{E_{i,j};\ E_{i,m};\ E_{j,m}\}$ is a Δ-system, consequently either $E_{j,m}=\emptyset$ or $E_{j,m}=\{c\}$. Both cases contradict (2).

(b) $|\{i, j, l, m\}|=4$. We may suppose that $(i, j)=(1, 2)$ and $(l, m)=(3, 4)$ while we are interested in $\{E_{i,j};\ E_{i,l};\ E_{i,m};\ E_{i,l};\ E_{j,m};\ E_{l,m}\}$. What is $E_{2,3}$?

By (2) $E_{2,3} \neq \emptyset$ and $E_{2,3} \neq \{c\}$. The cases $E_{2,3} = \{a\}$ or $\{b\}$ arise to the case (a). $E_{2,3} \neq \{b, c\}$ while $E_{1,2}$; $E_{2,3}$; $E_{1,3}$ form a Δ-system and thus $E_{2,3} = \{b, c\}$ implies $E_{1,3} = \emptyset$, contradicting 2. Similarly $E_{2,3} \neq \{a, c\}$. Thus $\{a, b\} \subseteq E_{2,3}$. The possibilities for $E_{1,3}$ are the same as for $E_{2,3}$ that is $\{a, b\} \subseteq E_{1,3}$. But then $b \in E_{1,3} \cap E_{2,3}$ and $b \notin E_{1,2}$, contradicting $\{E_{1,2}; E_{2,3}; E_{1,3}\}$'s being a Δ-system.

The proof is complete. □

REMARK. Theorem 2.2 demonstrates the difference between the weak dependency and the rest.

Theorem 2.3. Let $\mathcal{Y} \subseteq P(\Omega) \times P(\Omega)$ satisfy the Y'-axiom for some $Y \in \{F, D, S, W\}$. Then there is a relation R over Ω with $\mathcal{Y} = \mathcal{Y}_R$. Conversely, if R is a relation over Ω then $\mathcal{Y}_R$ satisfies the Y'-axiom.

Proof. Let $\mathcal{E} = \{E_{i,j}: 1 \leqq i < j \leqq k\}$ show that $\mathcal{Y}$ satisfies the Y'-axiom. Then the requirement (iii) of the Y'-axiom and Theorem 2.1 (ii) imply that there is a relation R over Ω such that $\mathcal{E}_R = \mathcal{E}$. By the Y'-axiom it is obvious that $\mathcal{Y} = \mathcal{Y}_R$.

Conversely, if R is a relation over Ω, then writing $R = \{h_1, \ldots, h_k\}$, $E_{i,j} = E(h_i, h_j)$; $\{E_{i,j}: 1 \leqq i < j \leqq k\}$ shows that $\mathcal{Y}_R$ satisfies the Y'-axiom. □

§ 3. Combinatorial results

Definition 3.1. Let $\mathcal{F}$ be a full f-family and let $A \subseteq \Omega$. Then A is a *candidate key for* $\mathcal{F}$ if $(A, \Omega) \in \mathcal{F}$ and for any $A' \subset A$ $(A', \Omega) \notin \mathcal{F}$ holds. Let R be a relation over Ω, then the set of candidate keys of R is the set of candidate keys of $\mathcal{F}_R$.

Let $\mathcal{C}$ denote the set of candidate keys of $\mathcal{F}$. Then $\mathcal{C}$ is a Sperner system, i.e. $(\forall A, B \in \mathcal{C})\ (A \subseteq B \Rightarrow A = B)$. We deal with the following question of [9]:

(*) Let $r(n)$ denote the smallest integer for which *any* Sperner system $C \subseteq P(\Omega)$ is the set of candidate keys of a suitable relation over the n-element set Ω with at most $r(n)$ rows. What can be said about $r(n)$?

In [9] it is shown that for any Sperner system there is a relation with this system as its set of candidate keys and that

$$\sqrt{2\binom{n}{[n/2]}} \leqq r(n) \leqq 2 \cdot \binom{n}{[n/2]}.$$

We give sharper estimations for $r(n)$.

Theorem 3.1. $\frac{1}{n^2}\binom{n}{[n/2]} < r(n) \leqq \binom{n}{[n/2]} + 1.$

Proof. First we prove the upper bound. Let $\mathcal{C} \subseteq P(\Omega)$ be a Sperner system. Let $\mathcal{B}$ consist of the maximal sets that do not contain any members of $\mathcal{C}$. Let $B_2, \ldots, B_k$ be the members of $\mathcal{B}$. For $1 < j \leqq k$ let $E_{1,j} = B_j$ and for $1 < i < j \leqq k$ let $E_{i,j} = B_i \cap B_j$. Then $\{E_{i,j}: 1 \leqq i < j \leqq k\}$ satisfies the requirements of the Theorem 2.1 (ii), hence there is a relation R over Ω with k rows such that $\mathcal{E}_R = \{E_{i,j}: 1 \leqq i < j \leqq k\}$. Then obviously $\mathcal{C}$ is the set of candidate keys of R. It is trivial that $\mathscr{B}$ is a Sperner system, and thus $|\mathscr{B}| \leqq \binom{n}{[n/2]}$ that is $k \leqq \binom{n}{[n/2]}+1$. The rest of the proof is due to L. RÓNYAI. We start with two trivial observations.

1. Let R be a relation over Ω with r rows. Then there is a relation R' over Ω such that R' uses no more than r symbols and $\mathscr{E}_R = \mathscr{E}_{R'}$.

2. Let R be a relation over Ω with r rows and let $r' > r$. Then there is a relation R' over Ω with r' rows such that $\mathscr{E}_R = \mathscr{E}_{R'}$.

By 1. and 2. the number of Sperner systems which may be represented as sets of candidate keys of relations with r rows is no more then $r^{r \cdot n}$. Hence

$$r(n)^{r(n) \cdot n} > 2^{\binom{n}{[n/2]}}$$

which implies

$$r(n) > \frac{1}{n^2}\binom{n}{[n/2]}. \quad \square$$

It is natural to ask the following analogon of (∗):

Let $R(n)$ denote the smallest integer for which *any* full family $\mathscr{F} \subseteq P(\Omega) \times P(\Omega)$ is the set of functional dependencies of a suitable relation over the n-element set Ω with at most $R(n)$ rows. What can be said about $R(n)$?

By the proof of Theorem 2.2 (i) it is obvious that $R(n) \leqq$(the maximal number of subsets of Ω such that the intersection of any two of them is not a third). Thus, by a theorem of D. KLEITMAN [12], $R(n) \leqq c \cdot \binom{n}{[n/2]}$ where $c = 3/2$. Z. FÜREDI and J. PACH have shown, that this number is less then $(1 + (c \cdot \log n)/n)\binom{n}{[n/2]}$.

Lastly we give the combinatorial characterization — according to § 0 — of the sets which are of minimal cardinality with respect to the property that they imply all the dependencies of a given full f-family.

We need some definitions and a lemma.

Definition 3.2. Let $\mathscr{M} \subseteq P(\Omega)$.

(i) We say that $\mathscr{M}$ *has the intersection property* if for any $\mathscr{M}' \subseteq \mathscr{M}$, $\bigcap \mathscr{M}' \in \mathscr{M}$ holds.

(ii) An $M \in \mathscr{M}$ *is irreducible* if $M \neq \bigcap \{M' \in \mathscr{M}:\ M \subset M'\}$ (recall that $\subset$ means strict inclusion).

(iii) An $\mathscr{N} \subseteq \mathscr{M}$ *generates* $\mathscr{M}$ if $\mathscr{M} = \{\bigcap \mathscr{N}':\ \mathscr{N}' \subseteq \mathscr{N}\}$.

Lemma 3.1. Let $\mathscr{M}$ have the intersection property and let $\mathscr{N} = \{M \in \mathscr{M}:\ M$ is irreducible$\}$. Then an $\mathscr{N}' \subseteq \mathscr{M}$ generates $\mathscr{M}$ iff $\mathscr{N} \subseteq \mathscr{N}'$.

Proof. The following proof is standard in lattice theory. If $\mathscr{N}'$ generates $\mathscr{M}$, then $\mathscr{N} \subseteq \mathscr{N}'$ is obvious. For the converse we have to prove that $\mathscr{N}$ generates $\mathscr{M}$. Suppose indirectly that there is an $X \in \mathscr{M} \setminus \mathscr{N}$ such that $X \neq \bigcap \{Y:\ Y \in \mathscr{N}\ \&\ X \subset Y\}$. Let X be of minimal cardinality with respect to this property. $X \notin \mathscr{N}$ means that $X = \bigcap \{Y:\ Y \in \mathscr{M}\ \&\ X \subset Y\}$, hence $X \subset Y$ implies that there is an $\mathscr{N}_y \subseteq \mathscr{N}$ such that $Y = \bigcap \mathscr{N}_y$. Let $\mathscr{N}_x = \bigcup \{\mathscr{N}_y:\ X \subset Y$ and $Y \in \mathscr{M}\}$. Then $\mathscr{N}_x \subseteq \mathscr{N}$ and $X = \bigcap \mathscr{N}_x$ which is a contradiction. $\square$

REMARK. Observe that the proofs of the Theorems in [2] are essentially our proof of Lemma 3.1.

Corollary. If $\mathcal{M}$ has the intersection property then there is exactly one $\mathcal{N} \subseteqq \mathcal{M}$ which generates $\mathcal{M}$ and has minimal cardinality.

Theorem 3.2. Let $\mathcal{F}$ be a full f-family, let $\mathcal{B}$ be the set of maximal dependent sets for $\mathcal{F}$ and let $\mathcal{C}$ be the set which generates $\mathcal{B}$ and has minimal cardinality (in [1] it is shown that $\mathcal{B}$ has the intersection property).

Then for any $\mathcal{F}' \subseteqq \mathcal{F}$ we have the following: $\mathcal{F}'$ implies all the dependencies of $\mathcal{F}$ and $\mathcal{F}'$ has minimal cardinality with respect to this property if and only if for any $C \in \mathcal{C}$ there is an $A_c \subseteqq \Omega$ such that $\mathcal{F}' = \{(A_c, C) \colon C \in \mathcal{C}\}$.

We left the easy proof of the Theorem to the reader. We think that it is interesting to compare Theorem 3.2 with the Theorem on pp. 16 of [2].

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, pp. 580—583.
[2] ARMSTRONG, W. W., On the generation of dependency structures of relational data bases, *Publication* 272, Université de Montréal, 1977.
[3] BEERI, C., R. FAGIN, J. H. HOWARD, A complete axiomatization for functional and multivalued dependencies in database relations, *Proc. ACM SIGMOD* Int. Conf. an Management of Data, Toronto, 1977, pp. 47—61.
[4] BÉKÉSSY, A., J. DEMETROVICS, Contribution to the theory of data base relations, *Discrete Math.* v. 27, 1979, pp. 1—10.
[5] BÉKÉSSY, A., J. DEMETROVICS, L. HANNÁK, P. FRANKL, G. KATONA, On the number of maximal dependencies in a data base relation of fixed order, *Discrete Math.*, v. 29, 1980, pp. 83—88.
[6] CODD, E. F., A relational model of data for large shared data banks, *Comm. ACM*, v. 13, 1970, pp. 377—387.
[7] CODD, E. F., Further normalization of the data base relational model, Courant Computer Science Symposia 6 Data Base System, Prentice Hall, Englewood Cliffs, N. J., 1971, pp. 33—64.
[8] CZÉDLI, G., Függőségek relációs adatbázis modellben, *Alkalmaz. Mat. Lapok*, 1980.
[9] DEMETROVICS, J., Candidate keys and antichains, *SIAM J. on Algebraic and Discrete Methods*, v. 1, 1980, p. 92.
[10] DEMETROVICS, J., On the equivalence of candidate keys with Sperner systems, *Acta Cybernet.*, v. 4, 1979, pp. 247—252.
[11] FAGIN, R., Multivalued dependencies and a new normal form for relational data-bases, *ACM Trans. Database Sys.*, v. 2, 1977, pp. 262—278.
[12] KLEITMAN, D., On a problem of Erdős, *Proc. Amer. Math. Soc.*, v. 18, 1966, pp. 139—141.
[13] MENDELZON, A. O., On axiomatizing multivalued dependencies in relational databases, *J. Assoc. Comput. Sci.*, v. 26, 1979, pp. 37—44.
[14] RISSANEN, J., Independent components of relations, *ACM Trans. Database Sys.*, v. 2, 1977, pp. 317—325.
[15] YU, C. T., D. T. JOHNSON, On the complexity of finding the set of candidate keys for a given set of functional dependencies, *Inform. Process. Lett.*, v. 5, 1976, pp. 100—101.

(Received Jan. 19, 1981)