Elektronische Informationsverarbeitung und Kybernetik EIK 17 (1981) 2/3, 103 –112 (Journal of Information Processing and Cybernetics)

On Dependencies in the Relational Model of Data

By Gábor Czédli

Abstract. Functional dependencies in the relational model of data were introduced by Codd. It was Armstrong who gave an abstract characterization for this concept. In the present paper three other types of dependencies are characterized in a similar abstract way. Simultaneous characterizations for these concepts are also given.

1. Introduction

The use of the relational model of data structures proposed by E. F. Codd [2, 3] is a promising mathematical tool for handling data. In this model the user's data are represented by relationships.

For definition, let Ω be a finite non-empty set, and for each $b \in \Omega$ let T_b be a non-empty set associated with b. The elements of Ω are called attribute names and T_b is said to be the domain of b. Now a relationship over Ω is defined to be any finite subset of $\prod_{b \in \Omega} T_b$. A relationship R over $\Omega = \{a_1, \dots, a_n\}$ can be represented by a two-

dimensional table in which the columns correspond to attribute names and rows correspond to the elements of R:

 $(g \in R \text{ and } g(a_i) \in T_{a_i})$. This table is not unique, the order of columns and that of rows are arbitrary.

In the relational model a data base is a system of finitely many time-varying relationships.

The concept of functional dependency is due to Codd [2, 3]. For definition, let A and B be subsets of Ω and let R be a relationship over Ω . We say that B functionally depends on A in R (in notation $A \xrightarrow{f} B$ or simply $A \xrightarrow{f} B$) if for all $g, h \in R$

$$(\forall a \in A) (g(a) = h(a)) \Rightarrow (\forall b \in B) (g(b) = h(b))$$

is satisfied. The link $A \xrightarrow{f} B$ is said to be a functional dependency (or briefly f-dependency).

From the above definition we can obtain three other concepts of dependencies by changing the quantifiers.

Definition ([4, 5]). Let A and B be subsets of Ω and let R be a relationship over Ω . B is said to be d-dependent on A if for any $g, h \in R$

$$(\exists a \in A) (g(a) = h(a)) \Rightarrow (\exists b \in B) (g(b) = h(b))$$

holds. B is said to be strongly dependent (or briefly s-dependent) on A if for any g, $h \in R$

$$(\exists a \in A) (g(a) = h(a)) \Rightarrow (\forall b \in B) (g(b) = h(b))$$

holds. B is said to be weakly dependent (or w-dependent) on A if for any $g, h \in R$

$$(\forall a \in A) (g(a) = h(a)) \Rightarrow (\exists b \in B) (g(b) = h(b))$$

holds. For $z \in \{f, d, s, w\}$ let $A \xrightarrow{z} B$ mean that B is z-dependent on A in the relationship R.

In order to expound what made us to deal with these concepts let us consider the following example. Let

 $Q = \{\text{author, title, room, bookcase}\}\$

Table 1

author	title	room	bookcase
1	1	1	2
2	2	1	3
3	3	1	1
4	4	1	2
5	5	2	3
6	6	2	1
7	7	$egin{array}{c} 2 \\ 2 \end{array}$	2
8	8	2	3
9	9	3	1
10	10	3	2
11	. 11	3	3
12	12	3	1
1	4	1	1
5	8	3	3
4	1	1	3
7	10	3	2
6	10	2	2
6	9	2	1

and let a relationship R be given in Table 1. For the sake of visibility we can think R is a library in which eighteen books are stocked. The library consists of three rooms, each room has three bookcases, and only two books can go in each bookcase. The library is organized so that {author, title} $\frac{d}{R}$ {room, bookcase}. Furthermore, the book with author = title = i (i = 1, 2, ..., 12) is in the $\left[\frac{i+3}{4}\right]$ -th room in the $\left(1+3\left\{\frac{i}{3}\right\}\right)$ -th bookcase. (Here [x] denotes the largest integer not greater than x and $\{x\} = x - [x]$.) A visitor who knows that either the title or the author of a particular book is, say, i can find the book by scanning the $\left[\frac{i+3}{4}\right]$ -th room and the $\left(1+3\left\{\frac{i}{3}\right\}\right)$ -th bookcases only.

Now in connection with this example we try to express why the concepts of fd-, s-, and w-dependencies can have some practical importance. The final purpose of any data bank system is providing the user with actual information. In the general case the user "knows" (all or at least one of) the values of attributes of a given set A of attribute names and wants to learn the values, all or at least one, of attributes of another set B. The user can succeed in obtaining the information he wants if and only if B is dependent on A in the given relationship. I.e., dependencies are in a close connection with user's activities.

In any time-varying data structure at a particular moment of time there are dependencies. Some of them may be fortuitous or unimportant, but it is reasonable to require that at least certain dependencies be present at any time. Organizing the data structure and some of the user's activities can be based on these constant dependencies. In case of functional dependencies this has been shown in Codd's papers [2, 3].

Now the following five reasons have been collected to show the advantage of using more types of dependencies besides the functional one.

- (1) The user can happen to know only at least one but not all the values of attributes of A in the "life". Just think of the visitor of the library in our example. If A is a set of several attributes of a criminal, say $A = \{\text{length, age, citizenship, } \ldots \}$ and R is a relationship of a criminal data bank then a detective also can be such a user at the beginning of his investigation. Now the d- and s-dependencies are related just to this situation.
- (2) Sometimes the user can need only the value of at least one attribute of a given set B. E.g., if the user is interested in C, B is an intermediate step and $B \xrightarrow{s} C$ holds in another relationship Q.
- (3) There can be other types of dependencies between A and B even if there is no functional one between them.
- (4) Sometimes the information supply can be accelerated by describing a particular dependency with functions. The only requirement tailored to those functions is that they should be computed easily or stored in relatively small tables. For instance, in our example, the dependency

{author, title}
$$\xrightarrow{\mathbf{d}}$$
 {room, bookcase}

is described with the functions $\left[\frac{i+3}{4}\right]$ and $1+3\left\{\frac{i}{3}\right\}$. Owing to these functions the user, looking for a particular book, need not scan the whole table of R, he has to scan only a part of it.

(5) {author, title} $\frac{\mathbf{f}}{R}$ {room, bookcase} also holds in our example. Consequently, there exists a function which describes this dependency. But the table of this function is the table of R itself and so scanning the whole table cannot be avoided in this way. However, based on d-dependency, we have seen that scanning the whole table is not necessary. I.e., sometimes it is not the functional dependency which yields the most economic way of information supply.

Given a relationship R over Ω , for any $z \in \{f, d, s, w\}$ let \mathcal{I}_R denote the set $\{(A, B): A, B \subseteq \Omega \text{ and } A \xrightarrow{z} B\}$. \mathcal{I}_R is called the z-family of R. Similarly, let $\mathcal{I}_R^+ = \{(A, B): A \neq \emptyset \text{ and } (A, B) \in \mathcal{I}_R\}$, which is called the z⁺-family of R. Let $\mathfrak{P}(\Omega)$ and $\mathfrak{P}^+(\Omega)$ denote the set of all subsets and all non-empty subsets of Ω , respectively.

Armstrong's famous theorem [1] characterized all subsets of $\mathfrak{P}(\Omega) \times \mathfrak{P}(\Omega)$ which coincide with \mathcal{F}_R for some relationship R over Ω . Similar characterizations for \mathcal{D}_R and \mathcal{F}_R were proved in [4, 5]. Since the empty set usually has no practical importance as a component of a dependency, we can be content with characterizations for \mathcal{F}_R^+ as

well. For any $z \in \{f, d, w, s\}$ \mathcal{A}_{\emptyset} and $\mathcal{A}_{\emptyset}^+$ can be characterized trivially, thus, without loss of generality, the case $R = \emptyset$ will be left out of our considerations.

Our main result, Theorem 3.1, will characterize \mathcal{W}_R^+ . From this Theorem a characterization for \mathcal{F}_R^+ (which is a weakened form of Armstrong's result) and that for \mathcal{D}_R and \mathcal{F}_R will be deduced. Although these characterizations for \mathcal{F}_R^+ , \mathcal{D}_R and \mathcal{F}_R have been known, their proofs in this paper are relatively short and entirely new. Finally, we are interested in more than one of the families \mathcal{F}_R^+ , \mathcal{D}_R , \mathcal{W}_R^+ and \mathcal{F}_R that are associated with the same relationship R. Such simultaneous characterizations will be presented in the last section.

2. Connections between dependencies

In this section we are interested in connections between \mathcal{F}_R , \mathcal{D}_R , \mathcal{F}_R and \mathcal{W}_R for a fixed relationship R.

Claim 2.1 ([5]). Let R be a relationship over Ω and let A, $B \subseteq \Omega$. Then we have

$$A \xrightarrow{\mathbf{f}} B \quad iff \quad (\forall b \in B) \ (A \xrightarrow{\mathbf{w}} \{b\});$$

$$A \xrightarrow{\mathbf{d}} B \quad iff \quad (\forall a \in A) \ (\{a\} \xrightarrow{\mathbf{w}} B);$$

$$A \xrightarrow{\mathbf{s}} B \quad iff \quad (\forall b \in B) \ (A \xrightarrow{\mathbf{d}} \{b\});$$

$$A \xrightarrow{\mathbf{s}} B \quad iff \quad (\forall a \in A) \ (\{a\} \xrightarrow{\mathbf{f}} B) \ . \ \Box$$

This claim is an immediate consequence of definitions, thus the proof is omitted. We have obtained that \mathcal{W}_R uniquely determines \mathcal{F}_R , \mathcal{F}_R^+ , \mathcal{D}_R and \mathcal{F}_R . Simple examples show that \mathcal{D}_R does not determine \mathcal{F}_R or \mathcal{F}_R^+ , and vice versa.

3. A characterization of w+-families

We intend to prove the following

Theorem 3.1. Let Ω be a finite non-empty set. Then for any subset \mathscr{W} of $\mathfrak{P}^+(\Omega) \times \mathfrak{P}(\Omega)$ the following two conditions are equivalent.

- (i) There exists a non-empty relationship R over Ω such that $\mathcal{W} = \mathcal{W}_R^+$;
- (ii) W has the following four properties:

(W1) $(A, A) \in \mathcal{U}$ for any $A \in \mathfrak{P}^+(\Omega)$,

(W2) If $(A, B) \in \mathcal{W}$, $A \subseteq A'$ and $B \subseteq B'$, then $(A', B') \in \mathcal{W}$,

(W3) If
$$A, B \in \mathfrak{P}^+(\Omega)$$
 and, denoting $\Omega \setminus Y$ by \overline{Y} for any $Y \in \mathfrak{P}(\Omega)$, $(\forall X \in \mathfrak{P}(\Omega))$ $(A \subseteq X \subseteq \overline{B} \text{ implies } (X, \overline{X}) \in \mathcal{W})$ holds, then $(A, B) \in \mathcal{W}$, $(W4) \ \mathcal{W} \subseteq \mathfrak{P}^+(\Omega) \times \mathfrak{P}^+(\Omega)$.

Proof. Suppose (i) is satisfied. It is trivial that (W1), (W2) and (W4) hold. Suppose that (W3) is not satisfied. Then there exists a pair $(A, B) \in \mathcal{W} = \mathcal{W}_R^+$ such that

$$(\forall X \in \mathfrak{P}(\Omega)) (A \subseteq X \subseteq \overline{B} \text{ implies } (X, \overline{X}) \in \mathscr{W}_R^+)$$

holds. Since $(A, B) \in \mathcal{W}_R^+$, there are $g, h \in R$ such that g(a) = h(a) for all $a \in A$ and $g(b) \neq h(b)$ for all $b \in B$. Let us set $X = \{a \in \Omega : g(a) = h(a)\}$. Then $A \subseteq X \subseteq \overline{B}$ and $(X, \overline{X}) \in \mathcal{W}_R^+$ yield a contradiction. We have proved that (i) implies (ii).

In order to prove the converse we need some preliminaries. Since the w⁺-family of an at most two-element relationship can be described easily, for a given \mathcal{U} we shall construct a suitable relationship R from at least two-element ones. That is why an addition concept for relationships will be introduced in the following.

Definition. Let R_i ($i \in I$, I finite, non-empty) be non-empty relationships over Q. Then $\sum_{i \in I} R_i$, the sum of R_i , is defined to be the relationship

$$\{(i,g):i\in I,g\in R_i \text{ and } (i,g)\ (a)=(i,g(a)) \text{ holds for all } a\in\Omega\}$$
.

Roughly saying, $\sum_{i \in I} R_i$ is the "disjoint" union of R_i .

Lemma 3.2. Suppose R is the sum of the relationships R_i , $i \in I$, where I is a finite non-empty index set. Then $\mathcal{W}_R^+=\bigcap_{i\in I}\mathcal{W}_{R_i}^+$.

The proof is straightforward, so it is left to the reader.

For any subset X of Ω , $\emptyset + X + \Omega$, we define a two-element relationship $R_X = \{g, h\}$ in the following way: g(a) = 1 for any $a \in \Omega$, h(a) = 1 for $a \in X$ and h(a) = 0 for $a \in \overline{X}$.

Lemma 3.3. Let \mathscr{U} be a subset of $\mathfrak{P}^+(\Omega) \times \mathfrak{P}^+(\Omega)$ satisfying (W1), (W2), (W3) and (W4). Suppose $(X, \overline{X}) \in \mathscr{U}$ for some $X \in \mathfrak{P}^+(\Omega)$, $X \neq \Omega$. Then $\mathscr{U} \subseteq \mathscr{U}_{R_X}^+$.

Proof. Suppose $(C, D) \in \mathcal{W}$ but $(C, D) \notin \mathcal{W}_{R_X}$. Then $C \subseteq X$ and $D \subseteq \overline{X}$. Therefore $(X, \overline{X}) \in \mathcal{W}$ by (W2), a contradiction. \square

Now suppose \mathcal{W} is a subset of $\mathfrak{P}^+(\Omega) \times \mathfrak{P}^+(\Omega)$ satisfying (W1), (W2), (W3) and (W4). If $(X,\overline{X}) \in \mathcal{W}$ holds for any $X \in \mathfrak{P}^+(\Omega)$, $X \neq \Omega$, then $W = \mathfrak{P}^+(\Omega) \times \mathfrak{P}^+(\Omega)$ follows from (W3) and (W4). Thus $\mathcal{W} = \mathcal{W}_R^+$ stands for any one-element relationship R. If $(X,\overline{X}) \in W$ does not hold for all $X \in \mathfrak{P}^+(\Omega) \setminus \{\Omega\}$ we can define a relationship R as follows:

$$R = \sum_{\substack{\varnothing \subset X \subset \Omega \\ (X, \overline{X}) \notin \mathscr{W}}} R_X.$$

We obtain $\mathscr{U} \subseteq \mathscr{U}_R^+$ from Lemmas 3.2 and 3.3. Suppose $A, B \in \mathfrak{P}^+(\Omega)$ and $(A, B) \not\models \mathscr{U}$. Then, by (W3), there is an $Y, A \subseteq Y \subseteq \overline{B}$ such that $(Y, \overline{Y}) \notin \mathscr{U}$. Then

$$(A,B)\in \mathscr{W}_{R_{\pmb{x}}}^+ \supseteq \bigcap_{\substack{\varnothing\subset X\subset \Omega\\ (X,\widetilde{X})\in \mathscr{W}}} \mathscr{W}_{R_{\pmb{x}}}^+ = \mathscr{W}_R^+.$$

Hence $\mathscr{U} = \mathscr{U}_R^+$ completing the proof of Theorem 3.1. \square

So far we do not know any analogous characterization for w-families. To make the difficulties perceptible we mention that for $\Omega = \{a, b\}$ and

$$\mathscr{W} = \{(\varnothing, \Omega), (\{a\}, \{a\}), (\{a\}, \Omega), (\{b\}, \{b\}), (\{b\}, \Omega), (\Omega, \{a\}), (\Omega, \{b\}), (\Omega, \Omega)\}$$

(W1), (W2) and (W3) hold. However, as an easy consideration shows, there is no relationship R such that $\mathcal{W} = \mathcal{W}_R$.

4. d-families and f+-families

The following theorem gives a characterization for d-families:

Theorem 4.1 ([4, 5]). For any subset \mathcal{D} of $\mathfrak{P}(\Omega) \times \mathfrak{P}(\Omega)$ the following two conditions are equivalent:

- (i) There exists a non-empty relationship R over Ω such that $\mathcal{D} = \mathcal{D}_R$;
- (ii) D satisfies the following five properties:
 - (D1) $(A, A) \in \mathcal{D}$ for all $A \in \mathfrak{P}(\Omega)$,
 - (D2) For any $A, B, C \in \mathfrak{P}(\Omega)$, $(A, B) \in \mathcal{D}$ and $(B, C) \in \mathcal{D}$ imply $(A, C) \in \mathcal{D}$,
 - (D3) If $(A, B) \in \mathcal{D}$, $C \subseteq A$ and $B \subseteq D \subseteq \Omega$, then $(C, D) \in \mathcal{D}$,
 - (D4) If $(A, B) \in \mathcal{D}$ and $(C, D) \in \mathcal{D}$ then $(A \cup C, B \cup D) \in \mathcal{D}$,
 - (D5) $(A, \emptyset) \in \mathcal{D}$ implies $A = \emptyset$.

Proof. It is trivial that (i) implies (ii). To prove the converse let \mathcal{D} be a subset of $\mathfrak{P}(\Omega) \times \mathfrak{P}(\Omega)$ which satisfies (D1), ..., (D5). Let us define

$$\begin{split} \mathscr{W}_0 &= \{(A,B) \in \mathfrak{P}^+(\Omega) \times \mathfrak{P}^+(\Omega) \colon \varnothing \neq X \subseteq A \text{ and } \varnothing \neq Y \subseteq B \\ \text{ for some } (X,Y) \in \mathcal{D}\} \text{ ,} \\ \mathscr{W}_1 &= \{(A,B) \in \mathfrak{P}^+(\Omega) \times \mathfrak{P}^+(\Omega) \colon \text{ for any } X \in \mathfrak{P}(\Omega), A \subseteq X \subseteq \overline{B} \\ \text{ implies } (X,\overline{X}) \in \mathscr{W}_0\} \text{ ,} \\ \mathscr{W} &= \mathscr{W}_0 \cup \mathscr{W}_1 \text{ .} \end{split}$$

We claim that there is a relationship R such that $\mathscr{U}=\mathscr{U}_R^+$. By Theorem 3.1, it is sufficient to check $(W1),\ldots,(W4)$. (W1) follows from (D1), while (W2) and (W4) are trivial. Observe that $(X,\overline{X})\in\mathscr{U}$ iff $(X,\overline{X})\in\mathscr{U}_0$, whence (W3) follows.

It remains to prove that $\mathcal{D} = \mathcal{D}_R$. As a consequence of (D1), (D3) and (D4), we have that, for any $A, B \in \mathfrak{P}(\Omega)$, $(A, B) \in \mathcal{D}$ iff $(\{a\}, B) \in \mathcal{D}$ for all $a \in A$. Since \mathcal{D}_R also satisfies (D1), ..., (D5), it is sufficient to show that, for any $a \in \Omega$ and $B \in \mathfrak{P}^+(\Omega)$, $(\{a\}, B) \in \mathcal{D}$ iff $(\{a\}, B) \in \mathcal{D}_R$. Suppose $(\{a\}, B) \in \mathcal{D}$. Then $(\{a\}, B) \in \mathscr{W}_0 \subseteq \mathscr{W} = \mathscr{W}_R$, whence Claim 2.1 yields $(\{a\}, B) \in \mathcal{D}_R$.

Conversely, let us assume that though $(\{a\}, B) \in \mathcal{D}_R$, $(\{a\}, B) \notin \mathcal{D}$. Then we have $(\{a\}, B) \in \mathcal{W}$ by Claim 2.1. It follows from (D3) that $(\{a\}, B) \in \mathcal{W}_1$. Set $B_1 = \{x \in \mathcal{Q} : (\{x\}, B) \in \mathcal{D}\}$. Now $a \notin B_1$ and, by (D1) and (D3), $B \subseteq B_1$. We obtain $(\overline{B}_1, B_1) \in \mathcal{W}_0$ from $\{a\} \subseteq \overline{B}_1 \subseteq \overline{B}$ and $(\{a\}, B) \in \mathcal{W}_1$. Therefore $(X, Y) \in \mathcal{D}$ for some $\mathcal{O} \neq X \subseteq \overline{B}_1$ and $\mathcal{O} \neq Y \subseteq B_1$. Let us choose an element q from X, then (D3) yields $(\{q\}, B_1) \in \mathcal{D}$. Since, by (D4), $(B_1, B) = (\bigcup_{x \in B_1} \{x\}, \bigcup_{x \in B_1} B) \in \mathcal{D}$, (D2) yields $(\{q\}, B) \in \mathcal{D}$. Hence $q \in B_1$, which contradicts $q \in X \subseteq \overline{B}_1$. Theorem 4.1 has been proved. \square

The characterization of f⁺-families is very similar to that of d-families. The following theorem is a little bit weakened form of *Armstrong*'s result [1].

ŢΧ

Theorem 4.2. For any subset \mathcal{F} of $\mathfrak{P}^+(\Omega) \neq \mathfrak{P}(\Omega)$ the following two conditions are equivalent:

- (i) There exists a non-empty relationship R over Ω such that $\mathcal{F} = \mathcal{F}_R^+$;
- (ii) F satisfies the following four properties:
 - (F1) $(A, A) \in \mathcal{F}$ for all $A \in \mathfrak{B}^+(\Omega)$,
 - (F2) Whenever $(A, B) \in \mathcal{F}$ and $(B, C) \in \mathcal{F}$ then $(A, C) \in \mathcal{F}$,
 - (F3) If $(A, B) \in \mathcal{F}$, $A \subseteq C \subseteq \Omega$ and $D \subseteq B$, then $(C, D) \in \mathcal{F}$,
 - (F4) If $(A, B) \in \mathcal{F}$ and $(C, D) \in \mathcal{F}$ then $(A \cup C, B \cup D) \in \mathcal{F}$.

Proof. Evidently, (i) implies (ii). To prove the converse let \mathcal{F} be a subset of $\mathfrak{P}^+(\Omega) \times \mathfrak{P}(\Omega)$ which satisfies (F1), ..., (F4). Let us define

$$\begin{split} \mathscr{W}_0 &= \{(A,B) \in \mathfrak{P}^+(\Omega) \times \mathfrak{P}^+(\Omega) \colon \varnothing \neq X \subseteq A \text{ and } \varnothing \neq Y \subseteq B \\ & \text{hold for some } (X,Y) \in \mathcal{F}\} \ , \\ \mathscr{W}_1 &= \{(A,B) \in \mathfrak{P}^+(\Omega) \times \mathfrak{P}^+(\Omega) \colon \text{for any } X \in \mathfrak{P}(\Omega), A \subseteq X \subseteq \overline{B} \\ & \text{implies } (X,\overline{X}) \in \mathscr{W}_0\} \ , \\ \mathscr{W} &= \mathscr{W}_0 \cup \mathscr{W}_1 \ . \end{split}$$

It is easy to check that $\mathscr W$ satisfies $(W1), \ldots, (W4)$. Hence, by Theorem 3.1, we can choose a non-empty relationship R such that $\mathscr W=\mathscr W_R^+$. So it suffices to show that $\mathscr F=\mathscr F_R^+$. From properties (F3), (F4) and (F1) it is not hard to check that, for $A\in \mathfrak B^+(\Omega)$ and $B\in \mathfrak B(\Omega)$, $(A,B)\in \mathcal F$ iff for all $b\in B$ $(A,\{b\})\in \mathcal F$. Since $\mathscr F_R^+$ also satisfies the properties $(F1),\ldots,(F4)$, we have only to show that, for $A\in \mathfrak B^+(\Omega)$,

 $b \in \Omega$, $(A, \{b\}) \in \mathcal{F}$ iff $(A, \{b\}) \in \mathcal{F}_R^+$. Suppose $(A, \{b\}) \in \mathcal{F}$, then $(A, \{b\}) \in \mathcal{W}_0 \subseteq \mathcal{W}$, and so $(A, \{b\}) \in \mathcal{F}_R^+$ follows from Claim 2.1.

Now let us assume that though $(A, \{b\}) \in \mathcal{F}_R^+$, $(A, \{b\}) \in \mathcal{F}$. Set $A_1 = \{x \in \Omega : A_1 = \{x \in \Omega : A_2 = \{x \in \Omega : A_2$ $(A, \{x\}) \in \mathcal{F}$. Then $b \in A_1$. Moreover, $A \subseteq A_1$ by (F1) and (F3). Claim 2.1 yields $(A, \{x\}) \in \mathcal{F}$. $\{b\}$) $\in \mathscr{U}$. Since $A \subseteq A_1 \subseteq \{b\}$ and $(A, \{b\}) \in \mathscr{U}_1$ follows from (F3), we obtain $(A_1, \{b\}) \in \mathscr{U}_1$ $A_1 \in \mathcal{W}_0$. Hence $(X, Y) \in \mathcal{F}$ holds for some $\emptyset \neq X \subseteq A_1$ and $\emptyset \neq Y \subseteq A_1$. Fix an element $y \in Y$, then $(A_1, \{y\}) \in \mathcal{F}$ follows from (F3). Since $(A, \{a\}) \in \widetilde{\mathcal{F}}$ stands for any $a \in A_1$, (F4) yields $(A, A_1) \in \mathcal{F}$. From (F2) we conclude $(A, \{y\}) \in \mathcal{F}$, i.e. $y \in A_1$. This is a contradiction, completing the proof. \square

5. s-families

For s-families we have the following characterization.

Theorem 5.1 ([5]). For any subset \mathscr{S} of $\mathfrak{P}(\Omega) \times \mathfrak{P}(\Omega)$ the following two conditions are equivalent:

- (i) There is a non-empty relationship R over Ω such that $\mathscr{S} = \mathscr{S}_R$;
- (ii) I satisfies the following five properties:
 - (S1) Whenever $(A, B) \in \mathcal{F}, C \subseteq A \text{ and } D \subseteq B \text{ then } (C, D) \in \mathcal{F},$
 - (S2) For any $A, B, C \in \mathfrak{P}(\Omega)$ if $(A, B) \in \mathcal{F}, (B, C) \in \mathcal{F}$ and $B \neq \emptyset$ then $(A, C) \in \mathcal{F}$,
 - (S3) If $(A, B) \in \mathcal{F}$ and $(C, D) \in \mathcal{F}$ then $(A \cap C, B \cup D) \in \mathcal{F}$,
 - (S4) If $(A, B) \in \mathcal{F}$ and $(C, D) \in \mathcal{F}$ then $(A \cup C, B \cap D) \in \mathcal{F}$,
 - (S5) For all $a \in \Omega$: $(\{a\}, \{a\}) \in \mathcal{F}$.

Proof. The present proof, being based on Theorem 4.1, is quite different from that in [5]. It is easy to see that (i) implies (ii). To prove the converse, let \mathcal{F} be supposed a subset of $\mathfrak{P}(\Omega) \times \mathfrak{P}(\Omega)$ satisfying (S1), ..., (S5). Let us define $\mathcal{D}_0 = \{(A, B) \in \mathfrak{P}(\Omega) \times \mathfrak{P}(\Omega) : \text{there exist } k \geq 1 \text{ and, for } i = 0, 1, \dots, k - 1, (X_i, i) \}$

$$(Y_i) \in \mathcal{S}$$
 such that $A \subseteq \bigcup_{i < k} X_i, \bigcup_{i < k} Y_i \subseteq B$, and, for

$$i < k, Y_i \neq \emptyset \} \cup$$

$$\cup \{(\emptyset, \emptyset)\}$$

and

$$\mathcal{D} = \{ (A, B) \in \mathfrak{P}(\Omega) \times \mathfrak{P}(\Omega) : \text{there exist } k \geq 1 \text{ and, for } i \leq k, X_i \subseteq \Omega \text{ such that } A = X_0, B = X_k \text{ and, for } i < k, (X_i, X_{i+1}) \in \mathcal{D}_0 \}.$$

We claim that \mathcal{D} is a d-family, i.e. $\mathcal{D} = \mathcal{D}_R$ for some relationship R. By Theorem 4.1 we have only to check properties $(D1), \ldots, (D5)$. \mathcal{D} evidently satisfies (D2). We obtain from (S5) that (D1) holds in \mathcal{D}_0 and \mathcal{D} . Evidently, \mathcal{D}_0 satisfies properties (D3), (D4) and (D5). Hence if $(X_i, X_{i+1}) \in \mathcal{D}_0$ for i < k and $C \subseteq X_0$, $D \supseteq X_k$, then (C, X_1) , $(X_1, X_2), \ldots, (X_{k-1}, D)$ all belong to \mathcal{D}_0 . Therefore (D3) holds in \mathcal{D} . Similarly, if $m \leq k$, $(X_i, X_{i+1}) \in \mathcal{D}_0$ for i < k and $(Y_i, Y_{i+1}) \in \mathcal{D}_0$ for i < m then, by letting Y_i be equal to Y_m for $j=m+1,\ldots,k$, we have $(X_i \cup Y_i,X_{i+1} \cup Y_{i+1}) \in \mathcal{D}_0$ for i < k. Consequently, \mathcal{D} satisfies (D4). Finally, \mathcal{D} satisfies (D5) since so does \mathcal{D}_0 .

Now we can choose a non-empty relationship R such that $\mathcal{D} = \mathcal{D}_R$. It suffices to show that $\mathscr{S} = \mathscr{S}_R$. Properties (S1), (S3) (and, in case $B = \emptyset$, (S4) and (S5)) yield that, for any $A, B \in \mathfrak{P}(\Omega)$, $(A, B) \in \mathcal{F}$ iff for all $b \in B$ $(A, \{b\}) \in \mathcal{F}$. Since \mathcal{F}_R also satisfies (S1), ..., (S5), it is sufficient to prove that $(A, \{b\}) \in \mathcal{S}$ is equivalent to $(A, \{b\}) \in \mathscr{S}_R$. Let us assume that $(A, \{b\}) \in \mathscr{S}$, then $(A, \{b\}) \in \mathscr{D}_0 \subseteq \mathscr{D}$. Hence $(A, \{b\}) \in \mathscr{D}_0 \subseteq \mathscr{D}$. $\{b\}\in \mathscr{F}_R$ follows from Claim 2.1.

Conversely, let $(A, \{b\}) \in \mathcal{F}_R$ be supposed. Claim 2.1 yields $(A, \{b\}) \in \mathcal{D}$. We need the following observation:

(*) For arbitrary $U \subseteq \Omega$ and $v \in \Omega$, $(U, \{v\}) \in \mathcal{D}_0$ implies $(U, \{v\}) \in \mathcal{F}$.

Indeed, if $U \subseteq \bigcup_{i < k} X_i$, $\bigcup_{i < k} Y_i \subseteq \{v\}$ and, for i < k, $(X_i, Y_i) \in \mathcal{F}$ and $Y_i \neq \emptyset$, then $Y_i = \{v\}$ holds for all i. From (S3) we obtain $(\bigcup_{i < k} X_i, \{v\}) \in \mathcal{F}$. Hence (S1) yields $(U, \{v\}) \in \mathcal{F}$.

Since $(A, \{b\}) \in \mathcal{D}$, there exist a positive integer k, which will be supposed minimal, and $Z_0, \ldots, Z_k \in \mathfrak{P}(\mathcal{Q})$ such that $Z_0 = A, Z_k = \{b\}$ and $(Z_i, Z_{i+1}) \in \mathcal{D}_0$ for all i < k. For $A = \emptyset$ (S1) and (S5) immediately imply $(A, \{b\}) \in \mathcal{F}$. So $A \neq \emptyset$ can be supposed. Since \mathcal{D}_0 has property (D5), $Z_i \neq \emptyset$ holds for each $i \leq k$. Let us assume that $k \geq 2$. Let c be an arbitrary element of Z_{k-2} . Since \mathcal{D}_0 satisfies (D3), $(\{c\}, Z_{k-1}) \in \mathcal{D}_0$. Therefore there are m and, for i < m, $(X_i, Y_i) \in \mathcal{F}$ such that $\{c\} \subseteq \bigcup_{i < m} X_i, \bigcup_{i < m} Y_i \subseteq Z_{k-1}$

and $Y_i \longleftrightarrow \emptyset$ for i < m. Let us fix an index j, j < m, such that $c \in X_j$. From (S1) we obtain $(\{c\}, Y_j) \in \mathcal{S}$. Since $(Z_{k-1}, \{b\}) \in \mathcal{S}$ holds by (\times) , (S1) yields $(Y_j, \{b\}) \in \mathcal{S}$, too. Hence (S2) applies and we obtain $(\{c\}, \{b\}) \in \mathcal{S}$. Since c was an arbitrary element of Z_{k-2} , (S4) implies $(Z_{k-2}, \{b\}) \in \mathcal{S}$. This contradicts the minimality of k.

Therefore k=1 and $(A, \{b\}) \in \mathcal{F}$ immediately follows from (X). The proof is complete. \square

6. Simultaneous characterizations

A subset \mathcal{U}^+ of $\mathfrak{P}(\Omega) \times \mathfrak{P}(\Omega)$ is called an (abstract) w⁺-family if it satisfies condition (ii) of Theorem 3.1.

Clearly, \mathcal{W}^+ is a w⁺-family iff $\mathcal{W}^+ = \mathcal{W}_R^+$ for some non-empty relationship R. Similarly, a subset \mathcal{D} (\mathcal{F}^+ , \mathcal{F} , resp.) of $\mathfrak{P}(\Omega) \times \mathfrak{P}(\Omega)$ is called an (abstract) d-family (f⁺-family, s-family, resp.) if it satisfies (ii) of Theorem 4.1 (4.2, 5.1, resp.).

Definition. Let \mathcal{W}^+ , \mathcal{D} , \mathcal{F} and F^+ be an abstract w⁺-family, d-family, s-family and f⁺-family over Ω , respectively. Let us define

$$\begin{array}{l} \mathcal{D}(\mathcal{W}^+) = \{(A,\,B) \in \mathfrak{P}(\Omega) \times \mathfrak{P}(\Omega) \colon \text{ for each } a \in A, \, (\{a\},\,B) \in \mathcal{W}^+\} \ , \\ \mathcal{F}^+(\mathcal{W}^+) = \{(A,\,B) \in \mathfrak{P}^+(\Omega) \times \mathfrak{P}(\Omega) \colon \text{ for each } b \in B, \, (A,\,\{b\}) \in \mathcal{W}^+\} \ , \\ \mathcal{S}(\mathcal{F}^+) = \{(A,\,B) \in \mathfrak{P}(\Omega) \times \mathfrak{P}(\Omega) \colon \text{ for each } a \in A, \, (\{a\},\,B) \in \mathcal{F}^+\} \ , \\ \mathcal{S}(\mathcal{D}) = \{(A,\,B) \in \mathfrak{P}(\Omega) \times \mathfrak{P}(\Omega) \colon \text{ for each } b \in B, \, (A,\,\{b\}) \in \mathcal{D}\} \ . \end{array}$$

From Claim 2.1 and Theorem 3.1, 4.1 and 4.2 we can conclude that:

Claim 6.1. $\mathcal{D}(\mathcal{W}^+)$ and $\mathcal{F}^+(\mathcal{W}^+)$ are a d-family and an f^+ -family, respectively, while $\mathcal{S}(\mathcal{F}^+)$ and $\mathcal{S}(\mathcal{D})$ are s-families. Moreover, for any non-empty relationship R we have $\mathcal{D}(\mathcal{W}_R^+) = \mathcal{D}_R$, $\mathcal{F}^+(\mathcal{W}_R^+) = \mathcal{F}_R^+$ and $\mathcal{S}(\mathcal{F}_R^+) = \mathcal{S}(\mathcal{D}_R) = \mathcal{F}_R$. \square

Definition. Given a subset \mathcal{K} of $\mathfrak{P}(\Omega) \times \mathfrak{P}(\Omega)$, let us define $\mathcal{W}^+(\mathcal{K})$ by

$$\begin{split} \mathscr{W}^+(\mathscr{K}) &= \{ (A,B) \in \mathfrak{P}^+(\varOmega) \times \mathfrak{P}^+(\varOmega) \colon \text{for any } X \in \mathfrak{P}(\varOmega), \, A \subseteq X \subseteq \overline{B} \\ &\text{implies } (X,\overline{X}) \in \mathscr{W}_0(\mathscr{K}) \} \, \cup \, \mathscr{W}_0(\mathscr{K}) \end{split}$$

where

$$\label{eq:w0} \begin{split} \mathscr{W}_0(\mathcal{K}) = \{(A,B) \in \mathfrak{P}^+(\varOmega) \times \mathfrak{P}^+(\varOmega) \colon \varnothing \neq X \subseteq A \ \text{ and } \ \varnothing \neq Y \subseteq B \\ \text{ hold for some } (X,Y) \in \mathcal{K}\} \ . \end{split}$$

Claim 6.2. For any d-family \mathcal{D} and f^+ -family \mathcal{F}^+ $\mathcal{W}^+(\mathcal{D})$ and $\mathcal{W}^+(\mathcal{F}^+)$ are w^+ -families, and we have $\mathcal{D}(\mathcal{W}^+(\mathcal{D})) = \mathcal{D}$ and $\mathcal{F}^+(\mathcal{W}^+(\mathcal{F}^+)) = \mathcal{F}^+$.

Proof. Let \mathcal{D} be a d-family. In the proof of Theorem 4.1 we have shown that $\mathcal{W}^+(\mathcal{D})$ is a w⁺-family and, choosing a non-empty relationship R such that $\mathcal{W}^+(\mathcal{D}) =$

90

On Dependencies in the Relational Model of Data

 $=\mathcal{W}_{R}^{+},\ \mathcal{D}_{R}=\mathcal{D}.$ Hence we have $\mathcal{D}=\mathcal{D}_{R}=\mathcal{D}(\mathcal{W}_{R}^{+})\mathcal{D}-(\mathcal{U}+(\mathcal{D}))$ by Claim 6.1. For an f⁺-family \mathcal{F}^{+} in the proof of Theorem 4.2 we have shown that $\mathcal{W}^{+}(\mathcal{F}^{+})$ is a w⁺-family. and $\mathcal{F}_{R}^{+}=\mathcal{F}^{+}$ holds for any non-empty relationship R satisfying $\mathcal{W}_{R}^{+}=\mathcal{W}^{+}(\mathcal{F}^{+})$ Therefore $\mathcal{F}^{+}=\mathcal{F}_{R}^{+}=\mathcal{F}^{+}(\mathcal{U}_{R}^{+})=\mathcal{F}^{+}(\mathcal{W}_{R}^{+}(\mathcal{F}^{+}))$. \square

Given a finite non-empty set Ω , let $\mathcal{L}(\Omega)$ denote the set of w⁺-families over Ω . Since the intersection of w⁺-families is a w⁺-family and $\mathcal{L}(\Omega)$ has a largest element, namely $\mathfrak{P}^+(\Omega) \times \mathfrak{P}^+(\Omega)$, $\mathcal{L}(\Omega) = (\mathcal{L}(\Omega); \wedge, \vee)$ is a lattice with operations \wedge and \vee defined by

$$egin{aligned} & \mathscr{W}_1 \wedge \mathscr{W}_2 = \mathscr{W}_1 \cap \mathscr{W}_2 \,, \\ & \mathscr{W}_1 \vee \mathscr{W}_2 = \, \bigcap \, \left\{ \mathscr{U} \in \mathscr{L}(\Omega) \colon \mathscr{W}_1 \subseteq \mathscr{W} \, \text{ and } \, \mathscr{W}_2 \subseteq \mathscr{W} \right\} \,. \end{aligned}$$

Another description of the join in $\mathcal{L}(\Omega)$ is given in the following

Lemma 6.3. For arbitrary \mathscr{W}_1 , $\mathscr{W}_2 \in \mathscr{L}(\Omega)$ we have $\mathscr{W}_1 \vee \mathscr{W}_2 = \{(A, B) \in \mathfrak{P}^+(\Omega) \times \mathfrak{P}^+(\Omega) : \text{ for any } X, A \subseteq X \subseteq \overline{B} \text{ implies } (X, \overline{X}) \in \mathscr{W}_1 \cup \mathscr{W}_2 \}.$

Proof. Let \mathscr{W} stand for the right hand side of this equality. (W2) yields $\mathscr{W}_i \subseteq \mathscr{W}$ for i = 1, 2. If \mathscr{U} is a w⁺-family containing \mathscr{U}_1 and \mathscr{U}_2 , then $\mathscr{W} \subseteq \mathscr{U}$ follows from (W3). Finally, it is easy to check that \mathscr{W} is a w⁺-family itself. \square

Now all kinds of simultaneous characterizations can be based on the following

Theorem 6.4. Let \mathcal{D} and \mathcal{F}^+ be a d-family and an f⁺-family over Ω , respectively. Then the following two conditions are equivalent:

- (i) There exists a non-empty relationship R such that $\mathcal{D} = \mathcal{D}_R$ and $\mathcal{F}^+ = \mathcal{F}_R^+$;
- (ii) $\mathcal{D}(\mathcal{W}^+(\mathcal{D}) \vee \mathcal{W}^+(\mathcal{F}^+)) = \mathcal{D}$ and $\mathcal{F}^+(\mathcal{W}^+(\mathcal{D}) \vee \mathcal{W}^+(\mathcal{F}^+)) = \mathcal{F}^+$.

Proof. Let us first assume that (ii) holds, and denote $\mathscr{U}^+(\mathcal{D}) \vee \mathscr{U}^+(\mathcal{F}^+)$ by \mathscr{U}^+ . By Theorem 3.1 we can choose a non-empty relationship R such that $\mathscr{U}^+ = \mathscr{U}_R^+$. Then we have $\mathcal{D}_R = \mathcal{D}(\mathscr{U}_R^+) = \mathcal{D}(\mathscr{U}^+) = \mathcal{D}$ and $\mathscr{F}_R^+ = \mathscr{F}^+(\mathscr{U}_R^+) = \mathscr{F}^+(\mathscr{U}^+) = \mathscr{F}^+$, indeed.

To prove the converse we need the following two observations:

- (1) Let \mathcal{D} be a d-family. If for a w⁺-family \mathcal{U} we have $\mathcal{D} = \mathcal{D}(\mathcal{U})$ then $\mathcal{U}^+(\mathcal{D}) \subseteq \mathcal{U}$.
- (2) Let \mathcal{F}^+ be an f^+ -family. If for a w^+ -family \mathcal{U} we have $\mathcal{F}^+ = \mathcal{F}^+(\mathcal{U})$ then $\mathcal{W}^+(\mathcal{F}) \subseteq \mathcal{U}$.

Suppose we have w⁺-families \mathcal{U} and \mathcal{V} such that $\mathcal{D} = \mathcal{D}(\mathcal{U})$ and $\mathcal{F}^+ = \mathcal{F}^+(\mathcal{V})$. To show $\mathcal{W}^+(\mathcal{D}) \subseteq \mathcal{U}$ and $\mathcal{W}^+(\mathcal{F}^+) \subseteq \mathcal{V}$, by (W 3) it is sufficient to prove that $\mathcal{W}_0(\mathcal{D}) \subseteq \mathcal{U}$ and $\mathcal{W}_0(\mathcal{F}^+) \subseteq \mathcal{V}$. üuppose $(A, B) \in \mathcal{W}_0(\mathcal{D})$, say $\emptyset + X \subseteq A$ and $\emptyset + Y \subseteq B$ hold for $(X, Y) \in \mathcal{D}$. Since $\mathcal{D} = \mathcal{D}(\mathcal{U})$, $(\{x\}, Y) \in \mathcal{U}$ holds for any $x \in X$. Hence $(A, B) \in \mathcal{U}$ follows from (W 2). Similarly, if $(A, B) \in \mathcal{W}_0(\mathcal{F}^+)$ then $\emptyset + X \subseteq A$ and $\emptyset + Y \subseteq B$ hold for some $(X, Y) \in \mathcal{F}^+$. Since $\mathcal{F}^+ = \mathcal{F}^+(\mathcal{V})$, $(X, \{y\}) \in \mathcal{V}$ stands for any $y \in Y$. Thus $(A, B) \in \mathcal{V}$ follows from (F2). Observations (1) and (2) have been shown. \square

Now let (i) be assumed. Since Claim 6.1 implies $\mathcal{D}(\mathcal{W}_R^+) = \mathcal{D}_R = \mathcal{D}$ and $\mathcal{F}^+(\mathcal{W}_R^+) = \mathcal{F}_R^+ = \mathcal{F}^+$, our observations (1) and (2) yield $\mathcal{W}^+(\mathcal{D}) \subseteq \mathcal{W}_R^+$ and $\mathcal{W}^+(\mathcal{F}^+) \subseteq \mathcal{W}_R^+$. Moreover, for any two w⁺-families \mathcal{U} and \mathcal{V} , $\mathcal{U} \subseteq \mathcal{V}$, evidently we have $\mathcal{D}(\mathcal{U}) \subseteq \mathcal{D}(\mathcal{V})$ and $\mathcal{F}^+(\mathcal{U}) \subseteq \mathcal{F}^+(\mathcal{V})$. Therefore, by making use of Claim 6.1, we can compute as follows:

$$\begin{split} \mathcal{D} &= \mathcal{D}\big(\mathcal{W}^+(\mathcal{D})\big) \subseteq \mathcal{D}\big(\mathcal{W}^+(\mathcal{D}) \vee \mathcal{W}^+(\mathcal{F}^+)\big) \subseteq \mathcal{D}(\mathcal{W}_R^+) \stackrel{\cdot}{=} \mathcal{D}_R = \mathcal{D} \;, \\ \mathcal{F}^+ &= \mathcal{F}^+\big(\mathcal{W}^+(\mathcal{F}^+)\big) \subseteq \mathcal{F}^+\big(\mathcal{W}^+(\mathcal{D}) \vee \mathcal{W}^+(\mathcal{F}^+)\big) \subseteq \mathcal{F}^+(\mathcal{W}_R^+) = \mathcal{F}_R^+ = \mathcal{F}^+ \;. \end{split}$$

We have obtained (ii), which was to be proved.

Corollary 6.5. Let \mathcal{D} , \mathcal{F}^+ , \mathcal{F} and \mathcal{W}^+ stand for abstract d-, f⁺-, s- and w⁺-families over Ω , respectively.

- (a) There exists a non-empty relationship R such that $\mathcal{D} = \mathcal{D}_R$, $\mathcal{F}^+ = \mathcal{F}_R^+$ and $\mathcal{F} = \mathcal{F}_R$ if and only if $\mathcal{F} = \mathcal{F}(\mathcal{D})$ and (ii) of Theorem 6.4 holds;
- (b) There exists a non-empty relationship R such that $\mathcal{W}^+ = \mathcal{W}_R^+$, $\mathcal{D} = \mathcal{D}_R$ and $\mathcal{F}^+ = \mathcal{F}_R^+$ iff $\mathcal{D} = \mathcal{D}(\mathcal{W}^+)$ and $\mathcal{F}^+ = \mathcal{F}^+(\mathcal{W}^+)$ hold;
- (c) There exists a non-empty relationship R such that $\mathcal{D} = \mathcal{D}_R$, $\mathcal{F}^+ = \mathcal{F}_R^+$, $\mathcal{F} = \mathcal{F}_R$ and $\mathcal{W}^+ = \mathcal{W}_R^+$ iff $\mathcal{D} = \mathcal{D}(\mathcal{W}^+)$, $\mathcal{F}^+ = \mathcal{F}^+(\mathcal{W}^+)$ and $\mathcal{F} = \mathcal{F}(\mathcal{D})$.

This corollary follows easily from Theorem 6.4 and Claim 6.1. Some cases which are excluded from this corollary can be handled similarly.

References

- [1] Armstrong, W. W., Dependency Structures of Data Base Relationships. In: Information Processing 74; North-Holland Publ. Comp., 1974; pp. 580-583.
- [2] Codd, E. F., A Relational Model of Data for Large Shared Data Banks. Comm. ACM 13 (1970) 6, 377-387.
- [3] Codd, E. F., Further Normalization of the Data Base Relational Model. In: Courant Inst. Comp. Sci. Symp. 6, Data Base Systems; Prentice-Hall, Englewood Cliffs 1971; pp. 33-64.
- [4] Czédli, G., d-dependency structures in the relational model of data. Acta Cybernetica 5 (1980), 49-57.
- [5] Czédli, G., Dependencies in the relational model of data (in Hungarian). Alkalmazott Matematikai Lapok (Budapest) 6 (1980) (in print).

Kurzfassung

Die vorliegende Arbeit schließt an die von Armstrong gegebenen Charakterisierungen des Coddschen Konzepts der funktionalen Abhängigkeiten in relationalen Datenmodellen an. Es werden drei weitere Typen von Abhängigkeiten untersucht und abstrakt charakterisiert. Auch simultane Charakterisierungen für diese Abhängigkeiten sind angegeben.

Резюме

 ∂ . Φ . $Ko\partial\partial$ ввел понятие функциональной зависимости в реляционных моделях баз данных. B. B. Apмстрон дал абстрактную характеристику этого понятия. В настоящей работе аналогично характеризуются другие типы зависимостей. Исследуются также их симультанные характеристики.

(Received: first version March 27, 1980, extended version September 29, 1980)

Author's address
Dr. G. Czédli
Bolyai Institute
József Attila University
6720 Szeged
Aradi vértanuk tere 1
Hungary