d-dependency structures in the relational model of data

By G. CZÉDLI

I. 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 nonempty 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:

	a_1	a_2	$\dots$	a_n
g	$g(a_1)$	$g(a_2)$	$\dots$	$g(a_n)$
$\vdots$				

($g \in R$ and $g(a_i) \in T_{a_i}$).

This table is not unique, the order of columns and that of rows are arbitrary.

The concept of functional dependency is due to E. F. CODD [2, 3]. For the 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 $A \xrightarrow[R]{f} B$ 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 $A \xrightarrow[R]{f} B$ is said to be a functional dependency.

From the above definition we can obtain three other concepts of dependency by changing the quantifiers. Particularly, the concept of d-dependency is introduced as follows:

DEFINITION. 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 in the relationship R (in notation $A \xrightarrow[R]{d} B$ or simply $A \xrightarrow{d} B$) 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.

In any relationship of a 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 we want to show the applicability of d-dependencies in this aspect. For this reason we give an example. Let

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

and let a relationship R be given in the following table:

author	title	room	bookcase	author	title	room	bookcase
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

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 $\{\text{author, title}\} \xrightarrow[R]{d} \{\text{room, bookcase}\}$. Furthermore, the book with author=title=i ($i=1, 2, \ldots, 12$) 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 reader 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 concept of d-dependency can have some practical importance. The task of obtaining information from a given data structure is closely connected with the dependencies that are present in the data stucture. So, when we list some possibly advantageous properties of using d-dependencies below, we restrict our interest to the case of obtaining information only. Suppose the user "knows" the values of attributes of a given set A of attribute names and wants to learn the values of attributes of another set B.

(1) If $A \xrightarrow{d} B$ (in a given relationship R) then the user is not assumed to know all the attribute values from A. If he knows the value of at least one attribute in A and the d-dependency $A \xrightarrow{d} B$ is also given (by a suitable family of functions δ_a $(a\in A)$, $\delta_a\colon T_a \to \prod_{b\in B} T_b$, compare with the functions $\left[\frac{i+3}{4}\right]$ and $1+3\left\{\frac{i}{3}\right\}$ in

our example), then he can find the values of attributes in B by scanning a part of R only. (The values of attributes in B can be not unique if $A \xrightarrow{f} B$ does not hold.)

(2) Suppose both $A \xrightarrow{f} B$ and $A \xrightarrow{d} B$ hold. (This was the case in our example with $A=\{\text{author, title}\}$ and $B=\{\text{room, bookcase}\}$.) Sometimes, in spite of scanning a part of R, the method of (1) can be more immediate than making use of the explicit function $\varphi\colon \prod_{a\in A} T_a \to \prod_{b\in B} T_b$ which describes the functional dependency $A \xrightarrow{f} B$, since such a function φ can be given by another table in general.

(3) One can have $A \xrightarrow{d} B$ without having $A \xrightarrow{f} B$.

(4) The user can need only at least one value of attributes in B (without knowing which one is correct). E.g., this can occur when he is interested in C, B is an intermediate step, and $B \xrightarrow[Q]{d} C$ holds in an other relationship Q.

For a given relationship R over Ω let

$$\mathscr{F}_R = \left\{(A, B)\colon\ A \subseteqq \Omega,\ B \subseteqq \Omega,\ A \xrightarrow[R]{f} B\right\}$$

and

$$\mathscr{D}_R = \left\{(A, B)\colon\ A \subseteqq \Omega,\ B \subseteqq \Omega,\ A \xrightarrow[R]{d} B\right\}.$$

$\mathscr{F}_R$ and $\mathscr{D}_R$ are called the full family of functional dependencies of R and the full family of d-dependencies of R, respectively. In [1] W. W. ARMSTRONG has given an abstract characterization of full families of functional dependencies. Our main goal here is to give an abstract characterization for full families of d-dependencies. Due to duality between the concept of functional dependency and that of d-dependency, a considerable part of Armstrong's paper [1] is dualized and used in the present paper.

II. Abstract characterization of d-dependencies

Let Ω be a finite non-empty set and let $P(\Omega)$ denote the set of all subsets of Ω. We define a partial order $\leqq$ over $P(\Omega)\times P(\Omega)$ by $(A, B)\leqq(C, D)$ iff $A\subseteqq C$ and $B\supseteqq D$. We recall a definition from Armstrong's paper [1]:

A subset $\mathscr{F}$ of $P(\Omega)\times P(\Omega)$ is called an *abstract full family of functional dependencies* over Ω if the following four axioms hold for any elements A, B, C and D in $P(\Omega)$:

(F1) $(A, A)\in\mathscr{F}$.
(F2) $(A, B)\in\mathscr{F}$ and $(B, C)\in\mathscr{F}$ imply $(A, C)\in\mathscr{F}$.
(F3) If $(A, B)\in\mathscr{F}$ and $(A, B)\leqq(C, D)$ then $(C, D)\in\mathscr{F}$.
(F4) If $(A, B)\in\mathscr{F}$ and $(C, D)\in\mathscr{F}$ then $(A\cup C, B\cup D)\in\mathscr{F}$.

Now Armstrong's abstract characterization of functional dependencies is the following:

A subset $\mathscr{F}$ of $P(\Omega)\times P(\Omega)$ is of the form $\mathscr{F}=\mathscr{F}_R$ for some relationship R over Ω iff $\mathscr{F}$ is an abstract full family of functional dependencies.

To formulate our main result the following definition is needed.

Definition. A subset $\mathscr{D}$ of $P(\Omega)\times P(\Omega)$ is called an *abstract full family of d-dependencies* if the following five axioms hold for any elements A, B, C and D in $P(\Omega)$. (The notation $X \xrightarrow{d} Y$ will be used instead of $(X, Y)\in\mathscr{D}$).

(D1) $A \xrightarrow{d} A$.

(D2) If $A \xrightarrow{d} B$ and $B \xrightarrow{d} C$ then $A \xrightarrow{d} C$.

(D3) If $A \xrightarrow{d} B$ and $(C, D)\leqq(A, B)$ then $C \xrightarrow{d} D$.

(D4) If $A \xrightarrow{d} B$ and $C \xrightarrow{d} D$ then $A\cup C \xrightarrow{d} B\cup D$.

(D5) If $A \xrightarrow{d} \emptyset$ then $A=\emptyset$.

In the main theorem below an abstract characterization of d-dependencies is given.

Theorem. *Let Ω be an arbitrary non-empty set of attribute names. Then, for any non-empty relationship R over Ω, $\mathscr{D}_R$ is an abstract full family of d-dependencies. Conversely, for any abstract full family $\mathscr{D}$ of d-dependencies over Ω there exists a nonempty relationship R over Ω such that $\mathscr{D}=\mathscr{D}_R$.*

Remark. The case $R=\emptyset$ is excluded from the Theorem. However this fact does not mean the loss of generality, since $\mathscr{D}_\emptyset$ trivially can be characterized by $\mathscr{D}_\emptyset=P(\Omega)\times P(\Omega)$.

III. The proof of the Theorem

It is a straightforward consequence of definitions that $\mathscr{D}_R$ is an abstract full family of d-dependencies.

To prove the converse several lemmas will be needed. In what follows all concepts and statements concern a fixed set $\Omega=\{a_1, \ldots, a_n\}$ of attribute names. For an abstract full family $\mathscr{D}$ of d-dependencies let us denote by $\mathscr{M}_\mathscr{D}$ the set of maximal elements of $\mathscr{D}$.

Claim 1. Let us denote $(A, B)\in\mathscr{M}_\mathscr{D}$ by $A\nearrow B$. Then $\mathscr{M}_\mathscr{D}$ has the following four properties:

(M1) For any $A\in P(\Omega)$ there exist X and Y in $P(\Omega)$ such that $(A, A)\leqq(X, Y)$ and $X\nearrow Y$;

(M2) If $A\nearrow B$, $C\nearrow D$ and $(A, B)\leqq(C, D)$, then $(A, B)=(C, D)$;

(M3) If $A\nearrow B$, $B\subseteqq C$ and $C\nearrow D$, then $A\subseteqq C$;

(M4) If $A\nearrow\emptyset$ then $A=\emptyset$;

where A, B, C and D are universally quantified over $P(\Omega)$.

Proof. M1, M2 and M4 are trivially satisfied. Suppose we have $A\nearrow B$, $B\subseteqq C$ and $C\nearrow D$. Then $B \xrightarrow{d} C$ (i.e., $(B, C)\in\mathscr{D}$) follows from D1 and D3, whence $A \xrightarrow{d} D$ follows by D2. Now D4 yields $A\cup C \xrightarrow{d} D$. The maximality of (C, D) in $\mathscr{D}$ implies $A\cup C\subseteqq C$, whence we obtain the required inclusion $A\subseteqq C$. □

Let a subset $\mathscr{M}$ of $P(\Omega)\times P(\Omega)$ be called an *m-family* if it satisfies the axioms M1, M2, M3 and M4.

Claim 2. For any m-family $\mathcal{M}$ the set

$$\mathcal{D}_{\mathcal{M}} = \{(A, B) \in P(\Omega) \times P(\Omega)\colon \text{there exists } (C, D) \in \mathcal{M} \text{ such that } (A, B) \leqq (C, D)\}$$

is an abstract full family of d-dependencies.

Proof. It is trivial that $\mathcal{D}_{\mathcal{M}}$ satisfies D1, D3 and D5. To check D2, let (A, B) and (B, C) belong to $\mathcal{D}_{\mathcal{M}}$. Then $(A, B) \leqq (A_1, B_1)$ and $(B, C) \leqq (B_2, C_2)$ hold for some (A_1, B_1) and $(B_2, C_2) \in \mathcal{M}$. From $B_1 \subseteq B \subseteq B_2$ and M3 we obtain $A_1 \subseteq B_2$. Now $(A, C) \in \mathcal{D}_{\mathcal{M}}$ follows from $(A, C) \leqq (B_2, C_2)$.

As for D4, suppose (A, B) and (C, D) are in $\mathcal{D}_{\mathcal{M}}$. Let (A_1, B_1) and (C_1, D_1) be taken from $\mathcal{M}$ such that $(A, B) \leqq (A_1, B_1)$ and $(C, D) \leqq (C_1, D_1)$. Now M1 yields the existence of an (U, V) in $\mathcal{M}$ with the property $(B_1 \cup D_1, B_1 \cup D_1) \leqq (U, V)$. Since $B_1 \subseteq U$ and $D_1 \subseteq U$, M3 applies. We obtain $A_1 \subseteq U$ and $C_1 \subseteq U$. Thus the required $(A \cup C, B \cup D) \in \mathcal{D}_{\mathcal{M}}$ follows from $(A \cup C, B \cup D) \leqq (A_1 \cup C_1, B_1 \cup D_1) \leqq \leqq (U, V)$. □

Lemma 1. For any abstract full family $\mathcal{D}$ of d-dependencies the family $\mathcal{M}_{\mathcal{D}}$ of maximal elements of $\mathcal{D}$ is an m-family. Conversely, any m-family $\mathcal{M}$ is the family of maximal elements of exactly one abstract full family

$$\mathcal{D}_{\mathcal{M}} = \{(A, B) \in P(\Omega) \times P(\Omega)\colon (A, B) \leqq (C, D) \text{ for some } (C, D) \in \mathcal{M}\}$$

of d-dependencies.

Proof. D3 yields that any abstract full family $\mathcal{D}$ of d-dependencies is uniquely determined by $\mathcal{M}_{\mathcal{D}}$. The rest has already been proved in Claims 1 and 2. □

Now we could deal with m-families instead of abstract full families by Lemma 1. However, the concept of m-families is still complicated to our purposes. Surprisingly, certain semilattices will be suitable to characterize both abstract full families and m-families. For the sake of brevity, 0—1 subsemilattices of $P(\Omega)$ will be called *d-semilattices*. I.e., $\mathcal{S}$ is a d-semilattice over Ω iff it is a subset of $P(\Omega)$ containing $\emptyset$, Ω and the intersection of any two of its elements. The following statements will show the significance of d-semilattices. First, for an m-family $\mathcal{M}$, $\mathcal{S}_{\mathcal{M}}$, the semilattice according to $\mathcal{M}$, is defined by

$$\mathcal{S}_{\mathcal{M}} = \{A\colon A \subseteq \Omega \quad \text{and} \quad (A, B) \in \mathcal{M} \text{ for some } B\}.$$

Similarly, for an abstract full family $\mathcal{D}$, $\mathcal{S}_{\mathcal{D}}$ is defined by $\mathcal{S}_{\mathcal{M}_{\mathcal{D}}}$.

Claim 3. $\mathcal{S}_{\mathcal{M}}$ and $\mathcal{S}_{\mathcal{D}}$ are d-semilattices for any abstract full family $\mathcal{D}$ of d-dependencies and m-family $\mathcal{M}$.

Proof. It is enough to check that $\mathcal{S}_{\mathcal{M}}$ is a d-semilattice. From M1 we conclude that $\Omega \in \mathcal{S}_{\mathcal{M}}$. M1 and M4 yields that $\emptyset \in \mathcal{S}_{\mathcal{M}}$. Suppose A and B are in $\mathcal{S}_{\mathcal{M}}$, and let C, D be chosen so that $(A, C) \in \mathcal{M}$ and $(B, D) \in \mathcal{M}$. By M1 a pair $(U, V) \in \mathcal{M}$ is obtained such that $(A \cap B, A \cap B) \leqq (U, V)$. Since $V \subseteq A$ and $V \subseteq B$, M3 applies and we obtain $U \subseteq A$ and $U \subseteq B$. Hence $A \cap B = U$ implies $A \cap B \in \mathcal{S}_{\mathcal{M}}$. □

Claim 4. For any d-semilattice $\mathcal{S}$ over Ω the family

$$\mathcal{D}_{\mathcal{S}} = \{(A, B) \in P(\Omega) \times P(\Omega)\colon \text{ for any } X \in \mathcal{S}\ B \subseteq X \text{ implies } A \subseteq X\}$$

is an abstract full family of d-dependencies.

The *proof* is straightforward and so it will be omitted.

Lemma 2. For any abstract full family $\mathcal{D}$ of d-dependencies $\mathcal{S}_{\mathcal{D}}$ is a d-semilattice. Conversely, any d-semilattice $\mathcal{S}$ coincides with $\mathcal{S}_{\mathcal{D}}$ for exactly one abstract full family $\mathcal{D}$ of d-dependencies, namely for $\mathcal{D} = \mathcal{D}_{\mathcal{S}}$.

Proof. We have already proved that $\mathcal{D}_{\mathcal{S}}$ is an abstract full family of d-dependencies and $\mathcal{S}_{\mathcal{D}}$ is a d-semilattice. First we show that $\mathcal{S} = \mathcal{S}_{\mathcal{D}_{\mathcal{S}}}$. Suppose $A \in \mathcal{S}$ and choose $B \in P(\Omega)$, $B \subseteq A$, such that B is minimal with respect to the property $(A, B) \in \mathcal{D}_{\mathcal{S}}$. In order to show that (A, B) is a maximal element in $\mathcal{D}_{\mathcal{S}}$, we assume that $(A, B) < (C, D) \in \mathcal{D}_{\mathcal{S}}$. Then $A \subset C$ because of the coice of B, and we have $(A, B) < (C, B) \leqq (C, D)$. Hence we obtain $(C, B) \in \mathcal{D}_{\mathcal{S}}$. Now $B \subseteq A \in \mathcal{S}$ and the definition of $\mathcal{D}_{\mathcal{S}}$ yield $C \subseteq A$, which is a contradiction. Therefore (A, B) is maximal in $\mathcal{D}_{\mathcal{S}}$ and so $A \in \mathcal{S}_{\mathcal{D}_{\mathcal{S}}}$.

To show the converse inclusion, suppose $A \in \mathcal{S}_{\mathcal{D}_{\mathcal{S}}}$. Then (A, B) is maximal in $\mathcal{D}_{\mathcal{S}}$ for some B. Let $\mathcal{H}$ denote the set $\{X\colon X \in P(\Omega) \text{ and } A \subset X\}$. Since $(X, B) \notin \mathcal{D}_{\mathcal{S}}$ $(X \in \mathcal{H})$, we can assign an element $U_X \in \mathcal{S}$ such that $B \subseteq U_X$ and $X \nsubseteq U_X$. Since $\mathcal{S}$ is a finite semilattice, $H = \bigcap \{U_X\colon X \in \mathcal{H}\}$ belongs to $\mathcal{S}$. Now $B \subseteq H$ and $(A, B) \in \mathcal{S}$ implies $A \subseteq H$. If we had $H \in \mathcal{H}$ then $H \nsubseteq U_H$ would contradict $H = \bigcap \{U_X\colon X \in \mathcal{H}\} \subseteq U_H$. Consequently, $A \not\subset H$ and so $A = H \in \mathcal{S}$. The equality $\mathcal{S} = \mathcal{S}_{\mathcal{D}_{\mathcal{S}}}$ has been shown.

For the uniqueness of $\mathcal{D}$ we assume that $\mathcal{S} = \mathcal{S}_{\mathcal{D}_1} = \mathcal{S}_{\mathcal{D}_2}$. We denote $\mathcal{M}_{\mathcal{D}_i}$ by $\mathcal{M}_i$. Suppose (A, B) belongs to $\mathcal{D}_1$. We can choose elements (A_i, B_i) from $\mathcal{M}_i$ $(i = 1, 2)$ such that $(A, B) \leqq (A_1, B_1)$ and $(B, B) \leqq (A_2, B_2)$. Since $A_2 \in \mathcal{S}$, $(A_2, C) \in \mathcal{M}_1$ for a suitable C. Now M3 yields $A_1 \subseteq A_2$, which implies $(A, B) \leqq \leqq (A_2, B_2)$. By D3 we obtain $(A, B) \in \mathcal{D}_2$. We have shown the inclusion $\mathcal{D}_1 \subseteq \mathcal{D}_2$, while $\mathcal{D}_2 \subseteq \mathcal{D}_1$ follows similarly. □

A map $\varphi\colon P(\Omega) \to P(\Omega)$ is called a closure operator (on Ω) if for any $X, Y \in P(\Omega)$, $X \subseteq Y$,

$$X \subseteq X\varphi = (X\varphi)\varphi$$

and

$$X\varphi \subseteq Y\varphi$$

hold. For any d-semilattice $\mathcal{S}$ we define a closure operator $\varphi_{\mathcal{S}}$ by the following way:

$$X\varphi_{\mathcal{S}} = \bigcap \{Y\colon Y \in \mathcal{S} \text{ and } X \subseteq Y\}.$$

It is easy to see that, for any $X \in P(\Omega)$, $X\varphi_{\mathcal{S}} \in \mathcal{S}$. Moreover $X \in \mathcal{S}$ iff $X = X\varphi_{\mathcal{S}}$

Lemma 3. Let $\mathcal{S}$ be a d-semilattice and $X \in P(\Omega)$. Then $X\varphi_{\mathcal{S}} = \{a\colon (\{a\}, X) \in \mathcal{D}_{\mathcal{S}}\}$.

Proof. Let U denote the right-hand side of the above equality, and let $\mathcal{D} = \mathcal{D}_{\mathcal{S}}$. D1 and D3 yield $X \subseteq U$ and $(X, U) \in \mathcal{D}$. We have $(U, X) \in \mathcal{D}$ by D4. Let A be a minimal (with respect to $\subseteq$) subset of X for which $(U, A) \in \mathcal{D}$. We claim that

$(U, A) \in \mathcal{M}_{\mathcal{D}}$. To show this let the opposite case, $(U, A) < (V, B) \in \mathcal{D}$, be assumed. By the choice of A we have $U \subset V$ and $(U, A) < (V, A) \leq (V, B)$. We obtain

$(V, A) \in \mathcal{D}$ by D3,

$(A, X) \in \mathcal{D}$ by D1 and D3,

and

$(V, X) \in \mathcal{D}$ by D2.

Now $(\{v\}, X) \in \mathcal{D}$ for any $v \in V$ by D3. This means $V \subseteq U$, which contradicts $U \subset V$. Thus we have shown $(U, A) \in \mathcal{M}_{\mathcal{D}}$. Therefore $U \in \mathcal{S}_{\mathcal{D}}$, whence $U \in \mathcal{S}$ by Lemma 2.

Now $X_{\varphi_{\mathcal{S}}} \subseteq U$ follows from $X \subseteq U$ and $U \in \mathcal{S}$. To make the proof complete we have to show that if $X \subseteq C \in \mathcal{S}$ then $U \subseteq C$. Suppose $X \subseteq C \in \mathcal{S}$ and choose an element $D \in P(\Omega)$ such that (C, D) is a maximal element in $\mathcal{D}$. Since (U, A) is also a maximal element in $\mathcal{D}$ and $A \subseteq X \subseteq C$, we obtain $U \subseteq C$ from M3. □

Corollary. Let $\mathcal{S}$ be a *d*-semilattice and let $X \in P(\Omega)$. Then $X \in \mathcal{S}$ iff $(\{a\}, X) \in \mathcal{D}_{\mathcal{S}}$ for $a \in X$ only.

The concept of *d*-semilattices is already simple and worth being connected with relationships. A natural connection is given in the following definition.

Definition. Let R be a relationship. We define $\mathcal{S}_R$, the *d*-semilattice associated with R, to be $\mathcal{S}_{\mathcal{D}_R}$.

Now, by Lemma 2, we have only to prove that for any *d*-semilattice $\mathcal{S}$ there exists a relationship R such that $\mathcal{S} = \mathcal{S}_R$. The simplest case is settled in the following

Claim 5. Let $\mathcal{S} = \{\emptyset, \Omega\}$. Then $\mathcal{S} = \mathcal{S}_R$ for any one-element relationship R.

The *proof*, which is trivial by definitions, will be omitted.

For $A \in P(\Omega)$ we define an at most three-element *d*-semilattice $\mathcal{T}_A$ to be $\{\emptyset, A, \Omega\}$. A relatively simple case is handled in the following

Lemma 4. Let $A \in P(\Omega)$, $A \neq \emptyset$ and $A \neq \Omega$. Then $\mathcal{T}_A = \mathcal{S}_R$ where $R = \{g, h\}$ is a two-element relationship defined by

$g(a) = 1$ for all $a \in \Omega$,

$h(a) = 2$ for $a \in A$,

$h(a) = 1$ for $a \in \Omega \setminus A$.

Proof. The relationship R can be visualized by the following table:

	A			Ω/A		
	a	b	...	c	d	...
g	1	1		1	1	
h	2	2		1	1	

Since $(\{x\}, A) \notin \mathcal{D}_R$ for $x \notin A$, we have $A \in \mathcal{S}_R$ by the Corollary. Hence $\mathcal{T}_A \subseteq \mathcal{S}_R$. Now suppose $X \in P(\Omega)$, $X \neq A$, $X \neq \emptyset$, $X \neq \Omega$. If $A \setminus X$ is non-empty, say $u \in A \setminus X$, then $(\{u\}, X) \in \mathcal{D}_R$. Hence $u \in X\varphi_{\mathcal{S}_R} \setminus X$. If $A \setminus X$ is empty, i.e. $A \subset X \subset \Omega$, then $(\{v\}, X) \in \mathcal{D}_R$ and $v \in X\varphi_{\mathcal{S}_R} \setminus X$ for any $v \in \Omega \setminus X$. In both cases $X \neq X\varphi_{\mathcal{S}_R}$, whence $X \notin \mathcal{S}_R$. Therefore $\mathcal{T}_A = \mathcal{S}_R$. □

For d-semilattices $\mathcal{S}_i$ $(i\in I)$ we define $\sum_{i\in I}\mathcal{S}_i$, the sum of $\mathcal{S}_i$, to be the smallest d-semilattice containing $\mathcal{S}_i$ for all $i\in I$. It is easy to check that:

Claim 6. Let $\mathcal{S}_i$ $(i\in I,\ I$ finite) be d-semilattices over Ω. Then the following equality holds:

$$\sum_{i\in I}\mathcal{S}_i=\Big\{\bigcap_{i\in I}A_i\colon A_i\in\mathcal{S}_i\Big\}.$$

Now we introduce an addition concept for relationships, which will be in a close connection with the addition of d-semilattices.

Definition. Let R_i $(i\in I,\ I$ finite) be non-empty relationships over Ω, where $R_i\subseteq\prod_{b\in\Omega}T_{b,i}$. For $i\in I$ and $f\in R_i$ we define $f^i\in\prod_{b\in\Omega}(\{i\}\times T_{b,i})$ by $f^i(b)=(i,f(b))$ $(b\in\Omega)$. Set $R'_i=\{f^i\colon f\in R_i\}$. Then $\sum_{i\in I}R_i$, the sum of R_i, is defined to be $\bigcup_{i\in I}R'_i$.

Roughly saying, we obtain $\sum_{i\in I}R_i$ if we make R_i $(i\in I)$ pairwise disjoint by indices and take their disjoint union.

A crucial step of our proof is

Lemma 5. Let R_i $(i\in I,\ I$ finite) be arbitrary relationships over Ω. Let $\sum_{i\in I}R_i$ be denoted by R. Then $\mathcal{S}_R=\sum_{i\in I}\mathcal{S}_{R_i}$.

Proof. Let $\mathcal{D}_R$, $\mathcal{D}_{R_i}$, $\mathcal{S}_R$ and $\mathcal{S}_{R_i}$ $(i\in I)$ be denoted by $\mathcal{D}$, $\mathcal{D}_i$, $\mathcal{S}$ and $\mathcal{S}_i$, respectively. First we show that for any $(A,B)\in P(\Omega)\times P(\Omega)$

$$(A,B)\in\mathcal{D}\quad\text{iff}\quad(A,B)\in\mathcal{D}_i\ \text{for all}\ i\in I. \tag{1}$$

Suppose $(A,B)\in\mathcal{D}$. Let $g,h\in R_i$ such that $g(a)=h(a)$ for some $a\in A$. Then $g^i(a)=h^i(a)$ as well, whence $g^i(b)=h^i(b)$ and so $g(b)=h(b)$ for some $b\in B$. I.e., $(A,B)\in\mathcal{D}_i$. Conversely, let $(A,B)\in\mathcal{D}_i$ for all $i\in I$. Suppose $g^i,h^j\in R$ and $g^i(a)=h^j(a)$. Then $(i,g(s))=(j,h(a))$ implies $i=j$ and $g(a)=h(a)$ in R_i. Therefore there exists $b\in B$ such that $g(b)=h(b)$, from which we obtain $g^i(b)=(i,g(b))=(j,h(b))=h^j(b)$. Thus (1) has been shown.

Now let us assume that $A\in\mathcal{S}$. We compute by Lemma 3, Corollary 1 and (1) as follows:

$$A=A\varphi_{\mathcal{S}}=\{a\colon(\{a\},A)\in\mathcal{D}\}=\{a\colon(\{a\},A)\in\mathcal{D}_i\ \text{for}\ i\in I\}=$$
$$=\bigcap_{i\in I}\{a\colon(\{a\},A)\in\mathcal{D}_i\}=\bigcap_{i\in I}A\varphi_{\mathcal{S}_i}.$$

Therefore $A\in\sum_{i\in I}\mathcal{S}_i$ by Claim 6. We have obtained that $\mathcal{S}\subseteq\sum_{i\in I}\mathcal{S}_i$. To prove the converse inclusion let $A\in\mathcal{S}_i$ and suppose $A\notin\mathcal{S}$. Then there exists an $a\in\Omega$ such that $a\in A\varphi_{\mathcal{S}}\setminus A$. We have $(\{a\},A)\in\mathcal{D}$ by Lemma 3 and $(\{a\},A)\in\mathcal{D}_i$ by (1). But $(\{a\},A)\in\mathcal{D}_i$ implies $a\in A\varphi_{\mathcal{S}_i}=A$, which is a contradiction. Hence $A\in\mathcal{S}$ and therefore $\mathcal{S}_i\subseteq\mathcal{S}$. Finally, $\mathcal{S}_i\subseteq\mathcal{S}$ $(i\in I)$ implies $\sum_{i\in I}\mathcal{S}_i\subseteq\mathcal{S}$, which completes the proof.

Now we can prove

Lemma 6. For any d-semilattice $\mathcal{S}$ there exists a relationship R such that $\mathcal{S}=\mathcal{S}_R$.

Proof. If $\mathcal{S}$ has at least three elements then Lemmas 4 and 5 together with the equality

$$\mathcal{S} = \sum_{\substack{A\in\mathcal{S} \\ A\neq\emptyset,\Omega}} \mathcal{T}_A$$

imply our statement. The rest is included in Claim 5. □

Finally, Lemmas 2 and 6 complete the proof of Theorem.

Abstract

The concept of d-dependencies in relationships is introduced. An axiomatic description of d-dependencies in an arbitrary relationship is presented.

BOLYAI INSTITUTE
A. JÓZSEF UNIVERSITY
ARADI VÉRTANUK TERE 1.
SZEGED, HUNGARY
H—6720

References

[1] ARMSTRONG, W. W., Dependency structures of data base relationships, *Information Processing* 74, North Holland, 1974, pp. 580—583.
[2] CODD, E. F., A relational model of data for large shared data banks, *Comm. ACM*, v. 13, 1970, pp. 377—387.
[3] CODD, E. F., Further normalization of the data base relational model, Courant Inst. Comp. Sci. Symp. 6, *Data Base Systems*, Prentice-Hall, Englewood Cliffs, N. J., 1971, pp. 33—64.

(Received Sept. 17, 1979)