

MULTIVALUED DEPENDENCIES IN RELATIONAL
DATABASES AND BOOLEAN ALGEBRAS

D. A. Simovici

To the memory of my Professor Mendel Haimovici

I. Introduction. In a previous paper [3] we have explored the possibility of developing an abstract theory of integrity constraints in relational databases. We have concentrated on functional dependencies and we estimate that this approach is useful for at least the following two reasons:

i) There is a significant variety of objects which can be described by the same abstract device (FD systems) as closed sets of functional dependencies.

ii) The abstract theory of functional dependencies can be used for identifying an entire new class of intractable problems in the theory of relational databases [6]. In fact, we have developed a procedure allowing a "wholesale" transportation of NP-complete problems from graph theory into relational databases, which has allowed us to retrieve classical NP-complete problems and to identify new ones [6].

The aim of this paper is to construct an abstract theory of multivalued dependencies in relational databases. Due to space limitations we shall not present the entire motivation of our results. We hope that the reader who is familiar with a standard reference (such as [5]) will realize the significance of these results. We intend to discuss elsewhere the implications of these results for intractability aspects of multivalued dependencies.

II. Trivial Multivalued Dependencies and Their Subalgebra. The reason for considering the subalgebra $\text{Triv}(x)$ generated by an element of a Boolean algebra

$B = (B, \{\vee, \wedge, \bar{}, 0, 1\})$ is for modelling the trivial multivalued dependencies of the form $x \twoheadrightarrow y$.

Define $\text{Triv}(x)$ as

$$\text{Triv}(x) = \{u \mid u \in B, u \leq x \text{ or } u = \bar{x} \vee v, v \leq x\}.$$

Proposition 1. $\text{Triv}(x)$ is a subalgebra of the Boolean algebra B for $x \in B$.

Proof. It is clear that we have $0, 1 \in \text{Triv}(x)$. Therefore, we need to prove only that $\text{Triv}(x)$ is closed with respect to " \wedge " and " $\bar{}$ ".

Let $u \in \text{Triv}(x)$. If $u \leq x$ then $\bar{u} = \bar{x} \vee (x \wedge \bar{u})$, hence $\bar{u} \in \text{Triv}(x)$ because $x \wedge \bar{u} \leq x$. If $u = \bar{x} \vee v$ then $\bar{u} = x \wedge \bar{v} \leq x$, hence in both cases $u \in \text{Triv}(x)$.

Suppose now that $u_1, u_2 \in \text{Triv}(x)$. We have the following four cases:

- i) $u_1 \leq x$ and $u_2 \leq x$. In this situation $u_1 \wedge u_2 \leq x$, hence $u_1 \wedge u_2 \in \text{Triv}(x)$.
- ii) $u_1 \leq x$ and $u_2 = \bar{x} \vee v_2, v_2 \leq x$, which gives $u_1 \wedge u_2 \leq x \wedge (\bar{x} \wedge v_2) = x \wedge v_2 \leq x$ giving again $u_1 \wedge u_2 \in \text{Triv}(x)$.
- iii) $u_1 = \bar{x} \vee v_1, v_1 \leq x$ and $v_2 \leq x$. This case is similar to ii).
- iv) $u_1 = \bar{x} \vee v_1, v_1 \leq x$ and $u_2 = \bar{x} \vee v_2, v_2 \leq x$. In this situation $u_1 \wedge u_2 = (\bar{x} \vee v_1) \wedge (\bar{x} \vee v_2) = \bar{x} \vee (v_1 \wedge v_2) \in \text{Triv}(x)$ since $v_1 \wedge v_2 \leq x$. ■

Proposition 2. The mapping $\text{Triv}: B \rightarrow \text{SALG}(B)$ is monotonic, where $\text{SALG}(B)$ is the collection of subalgebras of the Boolean algebra B .

Proof. Suppose that $y \leq z, y, z \in B$ and let $u \in \text{Triv}(y)$. If $u \leq y$ then, a fortiori, $u \leq z$.

If $u = \bar{y} \vee v$ and $v \leq y$ we can write $u = (\bar{y} \vee v) \wedge (z \vee \bar{z}) = [(\bar{y} \vee v) \wedge z] \vee [(\bar{y} \vee v) \wedge \bar{z}] = (\bar{y} \wedge z) \vee (v \wedge z) \vee \bar{z} = (\bar{y} \wedge z) \vee v \vee \bar{z}$. Since $v \vee (\bar{y} \wedge z) \leq z$ we have $u \in \text{Triv}(z)$, hence $\text{Triv}(y) \subseteq \text{Triv}(z)$. ■

III. MVD Systems. Let $B = (B, \{\vee, \wedge, \bar{}, 0, 1\})$ be a Boolean algebra. The central definition of this Note is give below:

Definition 1. An MVD system on the Boolean algebra B is a pair (B, μ) , where

B is a Boolean algebra and $\mu \subseteq B \times B$ is a relation on B satisfying the following properties:

MVD0) $y \in \text{Triv}(x)$ implies $(x,y) \in \mu$.

MVD1) If $(x,y) \in \mu$ then $(x, \bar{x} \wedge \bar{y}) \in \mu$.

MVD2) $(x,y) \in \mu$ and $v \leq w$ imply $(x \vee w, y \vee v) \in \mu$.

MVD3) If $(x,y) \in \mu, (y,z) \in \mu$ then $(x, z \wedge \bar{y}) \in \mu$.

Clearly, MVD0-3) model the axioms of multivalued dependencies in relational databases.

The mapping $\Psi_\mu : B \rightarrow 2^B$ is given by :

$$\Psi_\mu(x) = \{y \mid (x,y) \in \mu\}.$$

Proposition 3. $\Psi_\mu(x)$ is a subalgebra of B for $x \in B$.

Proof. From MVD0 we have $0, 1 \in \Psi_\mu(x)$. Suppose that $y, z \in \Psi_\mu(x)$, hence $(x, y) \in \mu$ and $(x, z) \in \mu$. By applying MVD2 to $(x, z) \in \mu$ we obtain $(x \vee y, z) \in \mu$. Also, we have $(x, x \vee y) \in \mu$ (from $(x, y) \in \mu$) and, using MVD3 we obtain $(x, z \wedge \overline{(x \vee y)}) \in \mu$, which in turn, gives $(x, x \vee [z \wedge \overline{(x \vee y)}]) \in \mu$. By applying MVD2 to $(x, z) \in \mu$ we have $(x \vee [z \wedge \overline{(x \vee y)}], z) \in \mu$. Using MVD3, it is possible to write $(x, z \wedge \overline{\{x \vee [z \wedge \overline{(x \vee y)}]\}}) \in \mu$, hence

$$(x, z \wedge \bar{x} \wedge (\bar{z} \vee x \vee y)) = (x, z \wedge \bar{x} \wedge y) \in \mu.$$

Another application of MVD2 gives:

$$(x \vee (x \wedge y \wedge z), (\bar{x} \wedge y \wedge z) \vee (x \wedge y \wedge z)) = (x, y \wedge z) \in \mu,$$

which shows that $y \wedge z \in \Psi_\mu(x)$.

To prove that $\Psi_\mu(x)$ is closed with respect to complement let us take $y \in \Psi_\mu(x)$. Applying MVD1 gives $(x, \bar{x} \wedge \bar{y}) \in \mu$. Using MVD2 we can write

$$(x \vee (x \wedge \bar{y}), (\bar{x} \wedge \bar{y}) \vee (x \wedge \bar{y})) = (x, \bar{y}) \in \mu,$$

giving $y \in \Psi_\mu(x)$. This shows that $\Psi_\mu(x)$ is a subalgebra of B . ■

According to this Proposition Ψ_μ maps B on the subalgebras of the Boolean

algebra \mathcal{B} , $\text{SALG}(\mathcal{B})$.

Proposition 4. $\Psi_\mu: \mathcal{B} \rightarrow \text{SALG}(\mathcal{B})$ is a monotonic mapping for every MVD system (\mathcal{B}, μ) .

Proof. Suppose that $x \leq y$ and $z \in \Psi_\mu(x)$, hence $(x, z) \in \mu$. Applying MVD2 with $w = y$ and $v = 0$ we obtain $(y, z) \in \mu$, hence $z \in \Psi_\mu(y)$, which shows that $\Psi_\mu(x) \subseteq \Psi_\mu(y)$. ■

We also notice another important property of Ψ_μ . If $y \in \Psi_\mu(x)$ and $z \in \Psi_\mu(y)$ then $z \wedge \bar{y} \in \Psi_\mu(x)$ by MVD3. These are characteristic properties for Ψ_μ as follows from

Proposition 5. Let $\mathcal{B} = (\mathcal{B}, \{\vee, \wedge, \bar{}, 0, 1\})$ be a Boolean algebra and let Ψ be a monotonic mapping $\Psi: \mathcal{B} \rightarrow \text{SALG}(\mathcal{B})$ satisfying the following properties:

- i) $\text{Triv}(x) \subseteq \Psi_\mu(x)$ for $x \in \mathcal{B}$;
- ii) $y \in \Psi_\mu(x)$, $z \in \Psi_\mu(y)$ imply $z \wedge \bar{y} \in \Psi_\mu(x)$.

The mapping Ψ defines an MVD system (\mathcal{B}, μ) , where $\mu = \{(x, y) \mid x, y \in \mathcal{B}, y \in \Psi(x)\}$. Since $\Psi(x)$ is a subalgebra we have $\bar{y} \in \Psi(x)$. Also, $\bar{x} \in \text{Triv}(x) \subseteq \Psi(x)$ and this gives $\bar{x} \wedge \bar{y} \in \Psi(x)$, that is $(x, \bar{x} \wedge \bar{y}) \in \mu$. We proved MVD1.

Assume now $(x, y) \in \mu$, $v \leq w$. Due to the monotonicity of Ψ we have $y \in \Psi(x \vee w)$. On another hand $v \leq x \vee w$, hence $v \in \text{Triv}(x \vee w) \subseteq \Psi(x \vee w)$.

This, in turn, gives $y \vee v \in \Psi(x \vee w)$, hence $(x \vee w, y \vee v) \in \mu$.

We notice that MVD3 follows directly from the second property of Ψ . ■

According to the last Proposition an MVD system on a Boolean algebra \mathcal{B} is perfectly determined by the monotonic mapping Ψ .

IV. Abstract Tables and MVD Systems. The notion of abstract table or, in short, table was introduced by us in [3] as a counterpart of the notion of relation as used in the realm of relational databases. We shall focus here on tables

on the disjunctive semilattice (B, \vee) which is naturally derived from the Boolean algebra $B = (B, \{\vee, \wedge, \neg, 0, 1\})$.

Let D be a non-empty set. For $t, s : B \rightarrow D$ define the equalizer of t and s as:

$$E(t, s) = \{ x \mid x \in B, t(x) = s(x) \}.$$

We shall use below the notion of ideal of a semilattice (S, \cdot) ; this is a subset I of S such that $x \cdot y \in I$ if and only if both x and y belong to I .

Let τ be a set of mappings $\tau \subseteq \{t \mid t : B \rightarrow D\}$. We shall consider the semilattice of equivalences on τ , $(Eq(\tau), \cap)$.

Proposition 6. The mapping $\beta_\tau : B \rightarrow Eq(\tau)$ given by $\beta_\tau(x) = \{(t_1, t_2) \mid t_1, t_2 \in \tau, x \in E(t_1, t_2)\}$ is a morphism between the semilattices (B, \vee) and $(Eq(\tau), \cap)$ if and only if for any $t, s \in \tau$, $E(t, s)$ is an ideal of (B, \vee) .

Proof. Consider $t, s \in \tau$ and suppose that $E(t, s)$ is an ideal and let $t, s \in \beta_\tau(x \vee x')$. One has $x \vee x' \in E(t_1, t_2)$ and this implies $x \in E(t_1, t_2)$ and $x' \in E(t_1, t_2)$. Therefore, $\beta_\tau(x \vee x') \subseteq \beta_\tau(x) \cap \beta_\tau(x')$. Conversely, $(t_1, t_2) \in \beta_\tau(x) \cap \beta_\tau(x')$ implies $x, x' \in E(t_1, t_2)$, hence $x \vee x' \in E(t_1, t_2)$. Thus $(t_1, t_2) \in \beta_\tau(x \vee x')$ and this shows that

$$\beta_\tau(x) \cap \beta_\tau(x') = \beta_\tau(x \vee x').$$

Assume now that β_τ is a morphism and let $x, x' \in E(t_1, t_2)$. One has $(t_1, t_2) \in \beta_\tau(x)$ and $(t_1, t_2) \in \beta_\tau(x')$, hence $(t_1, t_2) \in \beta_\tau(x) \cap \beta_\tau(x')$, giving $x \vee x' \in E(t_1, t_2)$. It is easy to see that if $x \vee x' \in E(t_1, t_2)$ then both x and x' belong to $E(t_1, t_2)$, which shows that $E(t_1, t_2)$ is an ideal.

Definition 2. A BD-table is a set of mappings $\tau \subseteq \{t \mid t : B \rightarrow D\}$, where (B, \vee) is the disjunctive semilattice of a Boolean algebra and $\beta_\tau : B \rightarrow Eq(\tau)$ is a morphism between the semilattices (B, \vee) and $(Eq(\tau), \cap)$.

We shall say that a BD-table τ satisfies a pair $(x, y) \in B \times B$ if for any

$t_1, t_2 \in \tau$ if $x \in E(t_1, t_2)$ then there is a t in τ such that $x \vee y \in E(t_1, t)$ and $x \vee \bar{y} \in E(t, t_2)$. This is equivalent to saying that that $\beta_\tau(x) \subseteq \beta_\tau(x \vee y) \circ \beta_\tau(x \vee \bar{y}) = \beta_\tau(x \vee \bar{y}) \circ \beta_\tau(x \vee y)$.

Let $C = (C, \{\vee, \wedge, \bar{}, 0, 1\})$ be a Boolean algebra and suppose that $\phi : C \rightarrow B$ is a morphism between the semilattices (C, \vee) and (B, \vee) . If τ is a BD-table then the set $\tau\phi = \{t\phi \mid t \in \tau\}$ is a CD-table as we proved in [3].

Let $B = (B, \{\vee, \wedge, \bar{}, 0, 1\})$ be a Boolean algebra. The ideal generated in this algebra by its element b is the set

$$(b) = \{x \mid x \in B, x \leq b\}.$$

This set can be organized itself as a Boolean algebra with respect to the same operations " \vee " and " \wedge " as the original algebra B and the relative complement $\kappa_b : B \rightarrow B$ defined by $\kappa_b(x) = \bar{x} \wedge b$. We have thus the Boolean algebra $B_b = ((b), \{\vee, \wedge, \kappa_b, 0, 1\})$.

Let b, c be two elements of B for which $b \vee c = 1$. Clearly, every $z \in B$ can be written as $z = z_1 \vee z_2$, where $z_1 \leq b$ and $z_2 \leq c$.

Consider now the (b) D-table τ and the (c) D-table σ .

Definition 3. The tuples t, s are joinable (where $t \in \tau$ and $s \in \sigma$) if $(b \wedge c) \subseteq E(t, s)$. Their join is the tuple $r : B \rightarrow D$ given by

$$r(x) = \begin{cases} t(x), & \text{if } x \in (b), \\ s(x), & \text{if } x \in (c) \end{cases}$$

In view of the joinability condition the definition of r is a correct one.

Proposition 7. The set of tuples ρ resulting by joining all joinable tuples of τ and σ is a BD-table.

Proof. Consider the tuples $r_1 : B \rightarrow D$ and $r_2 : B \rightarrow D$ resulting from joining the tuples t_1, s_1 and t_2, s_2 , respectively, where $t_1, t_2 \in \tau$ and $s_1, s_2 \in \sigma$.

Assume that $r_1(x \vee y) = r_2(x \vee y)$. If $x \vee y \in (b]$ then $x \leq b$, $y \leq b$ and this imply $t_1(x) = t_2(x)$ and $t_2(y) = t_2(y)$, giving $r_1(x) = r_2(x)$ and $r_1(y) = r_2(y)$. The case when $x \vee y \in (c]$ is similar.

Suppose now that $r_1(x) = r_2(x)$ and $r_1(y) = r_2(y)$. The single non-trivial case is when $x \leq b$ and $y \leq c$. We obtain immediately that $t_1(x) = t_2(x)$ and $s_1(y) = s_2(y)$.

To prove that $r_1(x \vee y) = r_2(x \vee y)$ we have to examine the following three cases:

- i) $x \vee y \in (b] - (c]$,
- ii) $x \vee y \in (c] - (b]$ and
- iii) $x \vee y \in (b] \cap (c] = (b \wedge c]$.

In the first situation $t_1(x \vee y) = t_2(x \vee y)$ and this implies $t_1(x) = t_2(x)$, because $x \leq b$. Also, since $x \vee y \leq b$ and $y \leq c$ we have $y = y \wedge (x \vee y) \leq b \wedge c$ and this gives $t_1(y) = s_1(y) = s_2(y) = t_2(y)$. We leave to the reader the discussion of the other cases.

Since $E(r_1, r_2)$ is an ideal it follows that ρ is indeed a BD-table. ■

We shall denote the join of τ and σ by $\tau \bowtie \sigma$.

Proposition 8. For any BD-table consider the inclusion mappings $\iota_b: (b] \rightarrow B$ and $\iota_c: (c] \rightarrow B$, where $b \vee c = 1$ and $\iota_b(x) = x$ for $x \leq b$, $\iota_c(y) = y$ for $y \leq c$. We have the inclusion:

$$\tau \subseteq (\tau \iota_b) \bowtie (\tau \iota_c).$$

Proof. Let t be a tuple of the BD-table τ ; the corresponding tuples of $\tau \iota_b$ and $\tau \iota_c$ are defined by $t_1(x) = t(x)$ if $x \leq b$ and $t_2(y) = t(y)$ if $y \leq c$. Clearly, if $x \leq b \wedge c$ then $t_1(x) = t_2(x) = t(x)$, hence t_1 and t_2 are joinable and their join is t . This gives the desired inclusion. ■

Proposition 9. If a BD-table τ satisfies a pair $(x, y) \in B \times B$ then

$$\tau = (\tau_{1_b}) \bowtie (\tau_{1_c}),$$

where $b = x \vee y$ and $c = x \vee \bar{y}$.

Proof. Due to the previous Proposition we have to prove only that

$$(\tau_{1_b}) \times (\tau_{1_c}) \subseteq \tau$$

Suppose that $r \in (\tau_{1_b}) \times (\tau_{1_c})$, where r results from joining the tuples $t \in \tau_{1_b}$ and $s \in \tau_{1_c}$, that is

$$r(x) = \begin{cases} t(x), & \text{if } x \in (b) \\ s(x), & \text{if } x \in (c) \end{cases}$$

Since $b \wedge c = (x \vee y) \wedge (x \vee \bar{y}) = x$ it follows that $x \in E(t, s)$. Due to the fact that τ satisfies (x, y) we obtain the existence of the tuple w in τ such that $b \in E(t, w)$ and $c \in E(w, s)$, which shows that $w = r$. ■

Any BD-table defines naturally an MVD system as it is stated in

Proposition 10. The set of pairs μ_τ satisfied by a BD-table τ forms an MVD system (B, μ_τ) .

Proof. We have to verify the axioms MVD0-4) for μ_τ .

MVD0). Let $y \in \text{Triv}(x)$. We have either $y \leq x$ or $y = \bar{x} \vee z$, where $z \leq x$.

In the first case $\beta_\tau(x \vee y) \circ \beta_\tau(x \vee \bar{y}) = \beta_\tau(x) \circ \beta_\tau(1) \supseteq \beta_\tau(x)$.

In the second situation, we have

$$\beta_\tau(x \vee y) \circ \beta_\tau(x \vee \bar{y}) = \beta_\tau(1) \circ \beta_\tau(x \vee (x \wedge \bar{z})) = \beta_\tau(1) \circ \beta_\tau(x) \supseteq \beta_\tau(x),$$

which proves that $y \in \text{Triv}(x)$ implies $(x, y) \in \mu_\tau$.

MVD1) If $(x, y) \in \mu_\tau$ then $\beta_\tau(x) \subseteq \beta_\tau(x \vee y) \circ \beta_\tau(x \vee \bar{y})$. If $z = \bar{x} \wedge \bar{y}$ then $\beta_\tau(x \vee z) \circ \beta_\tau(x \vee \bar{z}) = \beta_\tau(x \vee (\bar{x} \wedge \bar{y})) \circ \beta_\tau(x \vee (x \vee y)) = \beta_\tau(x \vee \bar{y}) \circ \beta_\tau(x \vee y) \supseteq \beta_\tau(x)$, hence $(x, z) \in \mu_\tau$.

MVD2). For $x_1 = x \vee w$, $y_1 = y \vee v$ and $v \leq w$ we have $x_1 \vee y_1 = x \vee w \vee y \vee v$

$= x \vee y \vee w$ and $x_1 \vee \bar{y}_1 = x \vee w \vee (\bar{y} \wedge \bar{v}) = x \vee \bar{y} \vee w$ and we have:

$$\begin{aligned} & \beta_\tau(x_1 \vee y_1) \circ \beta_\tau(x_1 \vee \bar{y}_1) = \beta_\tau(x \vee y \vee w) \circ \beta_\tau(x \vee \bar{y} \vee w) = \\ & = [\beta_\tau(x \vee y) \cap \beta_\tau(w)] \circ [\beta_\tau(x \vee \bar{y}) \cap \beta_\tau(w)]. \end{aligned}$$

Since $\beta_\tau(x \vee y) \supseteq \beta_\tau(x)$ and also $\beta_\tau(x \vee \bar{y}) \supseteq \beta_\tau(x)$ we obtain that

$$\beta_\tau(x_1 \vee y_1) \circ \beta_\tau(x_1 \vee \bar{y}_1) \supseteq \beta_\tau(x \vee w) = \beta_\tau(x_1),$$

which shows that MVD2 is satisfied.

To prove that MVD3 is satisfied we shall use a different approach.

Suppose that τ satisfies (x, y) and (y, z) . We need to prove that τ satisfies

$(x, z \wedge \bar{y})$. To this end, let us consider two tuples t_1 and t_2 such that $x \in E(t_1, t_2)$. we have to prove that there is a tuple w such that

$$x \vee (z \wedge \bar{y}) \in E(t_1, w)$$

and

$$x \vee (\bar{z} \vee y) \in E(w, t_2).$$

Since τ satisfies the pair (x, y) we may conclude that there is $r \in \tau$ such that

$$x \vee y \in E(t_2, r) \text{ and } x \vee \bar{y} \in E(r, t_1).$$

Now, we shall use the fact that τ satisfies (y, z) ; clearly, by MVD2 it also satisfies $(x \vee y, z)$, hence there is $w \in \tau$ such that

$$x \vee y \vee z \in E(r, w)$$

and

$$x \vee y \vee \bar{z} \in E(w, t_2).$$

Since $x \vee (z \wedge \bar{y}) \leq x \vee (z \vee y)$ we have, clearly the existence of the desired tuple w . ■

V. Conclusions. We have justified the idea that it is possible to reconstruct the theory of multivalued dependencies in an abstract manner by using the weak notion of table, as defined above. More aspects of this abstractions

should be investigated; for instance, one should look into the relationship between functional dependencies and multivalued dependencies, using perhaps the tools provided by the FD-systems defined in [3].

R e f e r e n c e s

- [1]. Beerl, C., Fagin, R., and Howard, J.H. A complete axiomatization for functional and multivalued dependencies in relational databases. Proc. ACM SIGMOD Conf. Manage. of Data, Toronto, Aug.1977, pp.47-61.
 - [2]. Fagin, R.: Multivalued dependencies and a new normal form for relational databases, ACM Trans. Database Syst.2,3(Sept. 1977), pp. 262-278.
 - [3]. Jurgensen, H., Simovici, D.A.: Towards an Abstract Theory of Dependency Constraints in Relational Databases, TR 1/85, University of Mass., Boston, 1985.
 - [4]. Mendelzon, A.O.: On Axiomatizing Multivalued Dependencies in Relational Databases, J. of ACM, vol. 26, No.1, 1979, pp.37-44.
 - [5]. Maier, D. : The Theory of Relational Databases, Computer Science Press, Rockville, Maryland, 1983.
 - [6]. Simovici, D.A., Reischer, C.: Graph Theory as a Source of Difficult Problems for Relational Databases, to appear in Congressus Numerantium, 1986.
-