Acta Cybernetica, Vol. 11, No. 3, Szeged, 1994

On Strong-Generalized Positive Boolean Dependencies*

Le Thi Thanh †

Abstract

Strong-Generalized Positive Boolean Dependencies are introduced.

Key Words and Phrases: *relation, data base, functional dependency, Boolean dependency, positive Boolean dependency, generalized positive Boolean dependency, Armstrong relation, strong generalized positive Boolean dependency.*

1 Introduction

In the theory of relational databases, connections between functional and multivalued dependencies and a certain fragment of propositional logic have been investigated in several papers.

The full family and the possible mathematical structure of functional dependencies was first axiomatized by W.W.Armstrong [1]. Different kinds of functional dependencies have also been investigated. The full family of strong dependencies has been introduced and axiomatized [5,7,8,9,14,15].

The family of Boolean dependencies is introduced [13]. In [2,3], the large subclass of positive Boolean dependencies, that is, Boolean combinations of attributes and the logical constant TRUE in which neither negation nor FALSE occur are studied. In [4], the class of equational dependencies is introduced. This class includes the class of functional dependencies as well as the Boolean dependencies, the positive Boolean dependencies and the classes of dependencies considered in [6,10].

In the papers mentioned above, the connection between dependencies and the fragment of propositional logic is built on the set of truth assignments T_R of a given relation R. Namely, for each pair of distinct tuples of R, the set T_R contains the truth assignment that maps an attribute A to TRUE if the two tuples are equal on A and to FALSE if the two tuples have different values for A.

In [11] a large class of mappings for constructing the truth assignments of relations was introduced. This class includes the equality mappings mentioned above. The class of Generalized Positive Boolean dependencies is introduced on these mappings.

In this paper we introduce a class of strong-Generalized Positive Boolean dependencies. We present a characterization of Armstrong relations for a given set of strong Generalized Positive Boolean dependencies.

*Research supported by Hungarian Foundation for Scientific Research Grant 2575.

†Computer and Automation Institute, Hungarian Academy of Sciences, H-1111 Budapest, Lágymányosi u. 11. Hungary.

The paper is structured as follows. In Section 2 we give some basic definitions. The concept of strong Generalized Positive Boolean dependencies is introduced in Section 3. In Section 4 we investigate connections between full families of strong Generalized Positive Boolean dependencies, s-semillatice and strong operations. Armstrong relation, the update problem and membership problem for strong Generalized Positive Boolean dependencies are studied in Section 5, Section 6 and Section 7.

2 Basic Definitions

We assume that the reader is familiar with the relational model of database systems and with the basic concepts of relational database theory [12,16]. In this paper we use the following notation.

Let $\mathcal{U} = \{A_1, \ldots, A_n\}$ be a set of *attributes*. Corresponding to each attribute A_i is a set d_i, $1 \le i \le n$, called the *domain* of A_i. We assume that every d_i contains at least two elements.

A *relation* R over $\mathcal{U}$ is a subset of $d_1 \times \ldots \times d_n$. Elements of R are called *tuples* and we usually denote them by u, v or t. The class of all relations over $\mathcal{U}$ is denoted by $\mathcal{R}$. For $k \ge 0$, $\mathcal{R}_k$ denotes those relations in $\mathcal{R}$ that have at most k tuples. If $R \in \mathcal{R}$, $t \in R$, $A \in \mathcal{U}$ and $X \subseteq \mathcal{U}$, then we denote by $t[A]$ the value of t for the attribute A, and by $t[X]$ the set $\{t[A] \mid A \in X\}$.

By $\mathcal{F}$ we denote the set of all *formulas* that can be constructed from $\mathcal{U}$ using the logical connectives $\wedge, \vee, \rightarrow, \neg$, and logical constants 1 (TRUE) and 0 (FALSE).

For $X = \{A_{i_1}, \ldots, A_{i_k}\} \subseteq \mathcal{U}$, $\wedge X$ denotes the formula $A_{i_1} \wedge \ldots \wedge A_{i_k}$, and $\vee X$ denotes the formula $A_{i_1} \vee \ldots \vee A_{i_k}$.

Let $\mathcal{B} = \{0, 1\}$. A *valuation* is any function $x : \mathcal{U} \rightarrow \mathcal{B}$. The notation $x = (x_1, \ldots, x_n) \in \mathcal{B}^n$ means that $x(A_i) = x_i$, $A_i \in \mathcal{U}$, $1 \le i \le n$.

If $f \in \mathcal{F}$ and $x \in \mathcal{B}^n$, then $f(x)$ denotes the truth value of f on the valuation x. For a finite subset Σ of $\mathcal{F}$ and for a valuation x in $\mathcal{B}^n$, we denote $\Sigma(x) = \wedge\{f(x) \mid f \in \Sigma\}$.

Let f be a formula in $\mathcal{F}$. We denote $T_f = \{x \in \mathcal{B}^n \mid f(x) = 1\}$. For a subset Σ of $\mathcal{F}$, we denote $T_\Sigma = \cap\{T_f \mid f \in \Sigma\}$. Then $x \in T_\Sigma$ if and only if $(\forall f \in \Sigma)\,(f(x) = 1))$.

Definition 2.1 *Let f and g be two formulas. f implies g, written $f \vdash g$, if $T_f \subseteq T_g$. f and g are equivalent, $f \equiv g$, if $T_f = T_g$. For $\Sigma, \Gamma \subseteq \mathcal{F}$, $\Sigma \vdash \Gamma$ if $T_\Sigma \subseteq T_\Gamma$, and $\Sigma \equiv \Gamma$ if $T_\Sigma = T_\Gamma$.*

Let $e = (1, \ldots, 1)$ be the valuation that consists of all 1. A formula f in $\mathcal{F}$ is *positive* if $f(e) = 1$. Let $\mathcal{F}_p$ denote all positive formulas on $\mathcal{U}$. We know that $\mathcal{F}_p$ is equivalent to the set of all formulas that can be built using the connectives $\wedge, \vee, \rightarrow$ and constant 1 [10].

For each domain d_i, $1 \le i \le n$, we consider a mapping $\alpha_i : {d_i}^2 \rightarrow \mathcal{B}$. We assume that the mappings α_i satisfy the following properties.

(i) $(\forall a \in d_i)\,(\alpha_i(a, a) = 1)$,

(ii) $(\forall a, b \in d_i)\,(\alpha_i(a, b) = \alpha_i(b, a))$, and

(iii) $(\exists a, b \in d_i)\,(\alpha_i(a, b) = 0)$.

Example 2.2 *It is easy to see that the equality mappings on d_i,*

$$\alpha_i(a, b) = \begin{cases} 1 & if\ a = b \\ 0 & otherwise \end{cases}$$

$$a, b \in d_i, \ 1 \leq i \leq n$$

satisfy the properties (i) - (iii).

Example 2.3 *Let* $\mathcal{U} = \{A, B, C\}$, *where* d_A *is the set of positive integers,* d_B *is the set of real numbers and a null-value* $\perp$, *and* d_C *is the set of words* w *on a nonempty alphabet* P, *where the length of* w *is not greater than* k, $k \geq 1$. *We define the mappings* α_A, α_B, *and* α_C *as follows.*

$$\alpha_A(a,b) = \begin{cases} 1 & \text{if both } a \text{ and } b \text{ are simultaneously odd or even numbers} \\ 0 & \text{otherwise} \end{cases}$$

$$\alpha_B(a,b) = \begin{cases} 1 & \text{if both } a \text{ and } b \text{ are simultaneously real or } \perp \\ 0 & \text{otherwise} \end{cases}$$

$$\alpha_C(a,b) = \begin{cases} 1 & \text{if both } a \text{ and } b \text{ have the same length} \\ 0 & \text{otherwise} \end{cases}$$

It is not hard to verify that the mappings α_A, α_B, *and* α_C *satisfy the properties (i) - (iii).*

Let $R \in \mathcal{R}$. For $u, v \in R$ we denote by $\alpha(u, v)$ the valuation

$$(\alpha_1(u[A_1], v[A_1]), \ldots, \alpha_n(u[A_n], v[A_n])).$$

Now for $R \in \mathcal{R}$ we denote $T_R = \{\alpha(u,v) \mid u, v \in R\}$. Note that for every u in R, $\alpha(u, u) = e$, so e is in T_R.

Definition 2.4 *Elements of* $\mathcal{F}_p$ *are called generalized positive Boolean dependencies (GPBD).*

Definition 2.5 *For* $R \in \mathcal{R}$ *and* $f \in \mathcal{F}_p$, *we say that* R *satisfies the GPBD* f, *written* $R(f)$, *if* $T_R \subseteq T_f$.

Definition 2.6 *Let* $R \in \mathcal{R}$ *and* $\Sigma \subseteq \mathcal{F}_p$, *we say that* R *satisfies the set of GPBDs* Σ, *written* $R(\Sigma)$, *if* $R(f)$ *for all* $f \in \Sigma$. *This is equivalent to* $T_R \subseteq T_\Sigma$.

For $\Sigma \subseteq \mathcal{F}_p$ and $f \in \mathcal{F}_p$, $\Sigma \models f$ means that, for all $R \in \mathcal{R}$, if $R(\Sigma)$ then $R(f)$. $\Sigma \models_2 f$ means that, for all $R \in \mathcal{R}_2$, if $R(\Sigma)$ then $R(f)$. In other words, $\Sigma \models f$ if and only if for all $R \in \mathcal{R}$, $T_R \subseteq T_\Sigma$ implies $T_R \subseteq T_f$.

For the equality mappings mentioned in Example 2.2 several classes of Boolean dependencies were investigated. Boolean dependencies were introduced in [13]. Positive Boolean dependencies are studied in [2,3]. Equational dependencies were introduced in [4]. Boolean dependencies of a special form are studied in [6,10]. These papers consider dependencies equivalent to the Boolean dependencies $\wedge X \rightarrow \wedge Y$ (functional dependency), $\wedge X \rightarrow \vee Y$ (weak dependency), $\vee X \rightarrow \wedge Y$ (strong dependency), and $\vee X \rightarrow \vee Y$ (dual dependency). In [3], the authors shown that the consequence relation for positive Boolean dependencies is the same as the consequence relation for propositional logic.

3 Strong-Generalized Positive Boolean Dependencies

Definition 3.1 *Let $R = \{t_1, \ldots, t_m\}$ be a relation over the finite set of attributes $\mathcal{U}$, and $X, Y \subseteq \mathcal{U}$. We say that GPBD $\vee X \to \wedge Y$ is strong-GPBD (for short s-GPBD) in R denoted $f_R^s(X,Y) = \vee X \xrightarrow[R]{s} \wedge Y$ or $X \xrightarrow[R]{s} Y$ or $X \xrightarrow{s} Y$ if*

$$(\forall t_i, t_j \in R)(\exists A \in X)(\alpha_A(t_i[A], t_j[A]) = 1) \longrightarrow$$
$$(\forall B \in Y)(\alpha_B(t_i[B], t_j[B]) = 1).$$

Let $S_R = \{X \xrightarrow[R]{s} Y\}$. S_R is called a full family of s-GPBDs of R.

Definition 3.2 *A s-GPBD over $\mathcal{U}$ is a statement of the form $X \xrightarrow{s} Y$, where $X, Y \subseteq \mathcal{U}$. The s-GPBD $X \xrightarrow{s} Y$ holds in a relation R if $X \xrightarrow[R]{s} Y$. We also say that R satisfies the $X \xrightarrow{s} Y$.*

We now introduce five inference s-axioms for s-GPBDs. Let $\mathcal{U}$ be a finite set of attributes, and denote by $P(\mathcal{U})$ its power set. Let $G \subseteq P(\mathcal{U}) \times P(\mathcal{U})$. We say that G is a full family of s-GPBDs over $\mathcal{U}$, if for all $X, Y, Z, W \subseteq \mathcal{U}$, and $A \in \mathcal{U}$

(S1.) $f^s(A, A) \in G$

(S2.) $f^s(X,Y) \in G, f^s(Y,Z) \in G, Y \neq \emptyset \longrightarrow f^s(X,Z) \in G$

(S3.) $f^s(X,Y) \in G, Z \subseteq X, W \subseteq Y \longrightarrow f^s(Z,W) \in G$

(S4.) $f^s(X,Y) \in G, f^s(Z,W) \in G \longrightarrow f^s(X \cup Z, Y \cap W) \in G$

(S5.) $f^s(X,Y) \in G, f^s(Z,W) \in G \longrightarrow f^s(X \cap Z, Y \cup W) \in G$

Let Σ_s be a set of s-GPBDs over $\mathcal{U}$. The closure of Σ_s, written Σ_s^+, is the smallest set containing Σ_s such that s-axioms cannot be applied to the set to yield an s-GPBD not in the set. Since Σ_s^+ must be finite, we can compute it by staring with Σ_s, applying S1, S2 and S5 and adding the derived s-GPBDs to Σ_s until no new s-GPBDs can be derived.

It can be seen [11] that there is a relation R over $\mathcal{U}$ such that $S_R = \Sigma_s^+$. Such a relation is called Armstrong relation for Σ_s.

Definition 3.3 *$X \xrightarrow{s} Y$ is a s-GPBD over $\mathcal{U}$ if X and Y are both subsets of $\mathcal{U}$. Σ_s is a set of s-GPBDs over $\mathcal{U}$ if every s-GPBD in Σ_s is s-GPBD over $\mathcal{U}$.*

Definition 3.4 *If Σ_s is a set of s-GPBDs over $\mathcal{U}$ and G is the set of all possible s-GPBDs over $\mathcal{U}$, then $\Sigma_s^- = G - \Sigma_s^+$. Σ_s^- is the exterior of Σ_s.*

If Σ_s is a set of s-GPBDs over $\mathcal{U}$ and X is a subset of $\mathcal{U}$, then there is s-GPBD $X \xrightarrow{s} Y$ in Σ_s^+ such that Y is maximal: for any other s-GPBD $X \xrightarrow{s} Z$ in Σ_s^+, $Y \supseteq Z$. This result follows from S5. Y is called the closure of X, and is denoted by X^+.

Definition 3.5 *Let Σ_s be a set of s-GPBDs over $\mathcal{U}$. $X \subseteq \mathcal{U}$, $A \in \mathcal{U}$. Then $\{A\}^+ = \{B \in \mathcal{U} \mid \{A\} \xrightarrow{s} \{B\} \in \Sigma_s^+\}$, $X^+ = \{B \in \mathcal{U} \mid X \xrightarrow{s} \{B\} \in \Sigma_s^+\}$.*

$\{A\}^+$ is called the closure of $\{A\}$.

Theorem 3.6 *Inference axioms S1 to S5 are complete.*

Proof: Given a set Σ_s of s-GPBDs over $\mathcal{U}$, for any s-GPBD $X \xrightarrow{s} Y$ in Σ_s^-. We shall axhibit a relation R that satisfies Σ_s^+ but not $X \xrightarrow{s} Y$. Hence, we can see that there are no s-GPBDs implied by Σ_s that are not derived by Σ_s. Relation R will satisfy most of the s-GPBDs in Σ_s^+, for a s-GPBD $(W \xrightarrow{s} Z)$ in Σ_s^+.

Let $\mathcal{U} = \{A_1, A_2, \ldots, A_n\}$ and let a_i, b_i, c_i be distinct elements of $dom(A_i)$, $1 \leq i \leq n$. There will be only two tuples in R, t_1 and t_2. Tuple t_1 will be $< a_1 a_2 \ldots a_n >$. Tuple t_2 is defined as

$$\forall A_i \in X^+, \alpha_{A_i}(t_1[A_i], t_2[A_i]) = 1$$

and

$$\forall A_i \notin X^+, \alpha_{A_i}(t_1[A_i], t_2[A_i]) = 0.$$

First we show that R does not satisfy $X \xrightarrow{s} Y$. From the definition of R, $\exists B \in X$ that $\alpha_B(t_1[B], t_2[B]) = 1$. Suppose $\alpha_C(t_1[C], t_2[C]) = 1$ for all $C \in Y$, and hence $Y \subseteq X^+$.

But since $(X \xrightarrow{s} X^+) \in \Sigma_s^+$, by S3, we obtain that $X \xrightarrow{s} Y$ is in Σ_s^+, a contradiction to $X \xrightarrow{s} Y$ is in Σ_s^-.

Now we show that R satisfies all the s-GPBD in Σ_s^+. Let $\{B\} \in X^+$, hence by Definition 3.5. we obtain that $\{B\}^+ = X^+$. By the definition of s-GPBDs, we have $(W \xrightarrow{s} X^+) \in \Sigma_s^+$. Since $(W \xrightarrow{s} Z) \in \Sigma_s^+$, and by S5, we obtain $(W \xrightarrow{s} (X^+ \cup Z)) \in \Sigma_s^+$, so $(X^+ \cup Z) \in W^+$. Hence $Z \subseteq X^+$, and $\alpha_C(t_1[C], t_2[C]) = 1$ for all $C \in Z$. □

4 Strong-Generalized Positive Boolean Dependencies and s-semilattice

Definition 4.1 *Let $I \subseteq P(\mathcal{U})$. We say that I is a $\cap$-semilattice over $\mathcal{U}$ if $\mathcal{U} \in I$, and $X, Y \in I \longrightarrow X \cap Y \in I$. Let $M \subseteq P(\mathcal{U})$. Denote by M^+ the set $\{\cap M' \mid M' \subseteq M\}$. Then we say M generates I if $M^+ = I$.*

Theorem 4.2 [4] *Let $I \subseteq P(\mathcal{U})$ be a $\cap$-semilattice over $\mathcal{U}$. Let $N = \{X \in I : \forall Z, W \in I : X = Z \cap W \longrightarrow X = Z$ or $X = W\}$. Then N generates I and if N' generates I, then $N \subseteq N'$. N is called the minimal generator of I (It is obvious that $\mathcal{U} \in N$).*

Definition 4.3 [15] *Let $I \subseteq P(\mathcal{U})$. We say that I is an s-semilattice over $\mathcal{U}$ if I satisfies*

(1.) I is a $\cap$-semilattice,

(2.) for all $X \subseteq N \setminus \mathcal{U}$

$$((\exists A \in X)(\forall Z \in N \setminus \mathcal{U})(X \not\subset Z) \longrightarrow (A \notin Z),$$

where N is the minimal generator of I.

Definition 4.4 [15] *The mapping $F : P(\mathcal{U}) \longrightarrow P(\mathcal{U})$ is called a strong operation over $\mathcal{U}$ if for every $A, B \in \mathcal{U}$ and $X \in P(\mathcal{U})$, the following properties hold:*

(1.) $F(\emptyset) = \mathcal{U}$,

(2.) $A \in F(\{A\})$,

(3.) $B \in F(\{A\}) \longrightarrow F(\{B\}) \subseteq F(\{A\})$,

(4.) $F(X) = \bigcap_{A \in X} F(\{A\})$.

Theorem 4.5 [15] *Let F be a strong operation over $\mathcal{U}$. Let $I_F = \{F(X) \mid X \in P(\mathcal{U})\}$. Then I_F is an s-semilattice over $\mathcal{U}$. Conversely, if I is an s-semilattice over $\mathcal{U}$, then there is exactly one strong operation F such that $I_F = I$, where $F(\emptyset) = \mathcal{U}$, and for all $A \in \mathcal{U}$*

$$F(\{A\}) = \begin{cases} \bigcap\limits_{\substack{A \in W \\ W \in N \setminus \mathcal{U}}} W & \text{if } \exists W : A \in W \ (N \text{ the minimal generator of } I), \\ U & \text{otherwise.} \end{cases}$$

Theorem 4.6 *Let $G \subseteq P(\mathcal{U}) \times P(\mathcal{U})$. G is a full family of s-GPBDs over $\mathcal{U}$. Let $(X, Y) \in P(\mathcal{U}) \times P(\mathcal{U}) \setminus G$. There is an $A \in X$, and an $E_A \subseteq \mathcal{U}$ such that*

(i.) $A \in E_A$,

(ii.) $(\{A\} \xrightarrow{s} E_A) \in G$,

(iii.) $E' \supset E_A$ *implies that* $(\{A\} \xrightarrow{s} E') \notin G$.

Proof: If for any $A \in X$ we have $(\{A\}, Y) \in G$. By S5 we have $(X, Y) \in G$. Hence there is an $A \in X$ such that $(\{A\} \xrightarrow{s} Y) \notin G$. If for every $B \in Y$, $(\{A\} \xrightarrow{s} \{B\}) \in G$ holds, then by S4 $(\{A\} \xrightarrow{s} Y) \in G$.

Thus there is a $B \in Y$ such that $(\{A\} \xrightarrow{s} \{B\}) \notin G$. By S1 and S3 there is an $E_A \subseteq \mathcal{U}$ such that $A \in E_A$, $(\{A\} \xrightarrow{s} E_A) \in G$ and E_A is maximal to this property. □

Theorem 4.7 *Let $G \subseteq P(\mathcal{U}) \times P(\mathcal{U})$. G is a full family of s-GPBDs over $\mathcal{U}$ if and only if there is a family $\{E_i : i = 1, \ldots, l; \bigcup_{i=1}^{l} E_i = \mathcal{U}\}$ of subsets of $\mathcal{U}$ such that*

(i.) for all $X \subseteq \mathcal{U}$, $(\emptyset \xrightarrow{s} X) \in G$,

(ii.) for any $X, Y \subseteq \bigcap_{E_i \cap X \neq \emptyset} \longrightarrow (X \xrightarrow{s} Y) \in G$,

(iii.) $(Z \xrightarrow{s} W) \in G$, $Z \cap E_i \neq \emptyset \longrightarrow W \subseteq E_i$.

Proof: Only if: Assume that G is a full family of s-GPBDs over $\mathcal{U}$. Then by Theorem 4.6, S1, S3, and S5 for each $A \in \mathcal{U}$ we can construct an $E_i (E_i \subseteq \mathcal{U})$ such that $(\{A\} \xrightarrow{s} E_i) \in G$, and $\forall E' \mid E_i \subset E'$ implies $(\{A\} \xrightarrow{s} E') \notin G$. By Theorem 4.6, it is obvious that $A \in E_i$ and we have n such E_i-s, where $n = \mid \mathcal{U} \mid$. Thus, we have the set $E = \{E_i : i = 1, \ldots, n; \bigcup_{i=1}^{n} E_i = \mathcal{U}\}$. Assume $X = \{A_1 A_2 \ldots A_k : A_j \in \mathcal{U}, j = 1, \ldots, k\} \neq \emptyset$ and Y_1 is a set such that $(X \xrightarrow{s} Y_1) \in G$, $\forall Y_2 : Y_1 \subset Y_2$ implies $(X \xrightarrow{s} Y_2) \notin G$. By the construction of E, we have that for each A_j there is an $E_j \in E$ such that $(\{A\} \xrightarrow{s} E_j) \in G$. By S4 we have $(\bigcap_{j=1}^{k} A_j \xrightarrow{s} \bigcap_{j=1}^{k} E_j) = (X \xrightarrow{s} \bigcap_{j=1}^{k} E_j) \in G$. By Theorem 4.6 and the definition of Y_1 we have $\bigcap_{j=1}^{k} E_j \subseteq Y_1$. By $(X \xrightarrow{s} Y_1) \in G$ and by S3, we have $(\{A_j\} \xrightarrow{s} Y_1) \in G$ for all $j (j = 1, \ldots, k)$. Thus, $Y_1 \subseteq \bigcap_{j=1}^{k} E_j$ holds. Hence, $Y_1 = \bigcap_{j=1}^{k} E_j$. It is obvious that

$$\bigcap_{E_i \cap X \neq \emptyset} E_i \subseteq \bigcap_{j=1}^{k} E_j.$$

Thus, for all

$$Y (Y \subseteq \bigcap_{E_i \cap X \neq \emptyset} E_i) : Y \subseteq Y_1.$$

Hence $(X \xrightarrow{s} Y) \in G$ holds.

If $(Z \xrightarrow{s} W) \in G$, $Z \cap E_i \neq \emptyset$. Let $A_1 \in Z \cap E_i$. Suppose that $W \cap (\mathcal{U} \setminus E_i) \neq \emptyset$. Let $D_1 \in W \cap (\mathcal{U} \setminus E_i)$.

By S3 we have $(\{A_1\} \xrightarrow{s} \{D_1\}) \in G$, and by S1 we have $(\{A_1\} \xrightarrow{s} \{A_1\}) \in G$. Let $A \in E_i$, then $(\{A\} \xrightarrow{s} E_i) \in G$ implies that $(\{A, A_1\} \xrightarrow{s} \{A_1\}) \in G$ by S5. Hence by S3 we have $(\{A\} \xrightarrow{s} \{A_1\}) \in G$. Since $(\{A\} \xrightarrow{s} \{A_1\}) \in G$, $(\{A_1\} \xrightarrow{s} \{D_1\}) \in G$ and by S2 we have $(\{A\} \xrightarrow{s} \{D_1\}) \in G$. Thus, by S4 we have $(\{A\} \xrightarrow{s} E_i \cup \{D_1\}) \in G$.

On the other hand, by Theorem 4.6 we have $(\{A\} \xrightarrow{s} E_i) \in G$ and $\forall E' : E_i \subset E'$ implies $(\{A\} \xrightarrow{s} E') \in G$. Hence $W \subseteq E_i$.

If : Assume that there is a family $\{E_i : i = 1, \ldots, l : \bigcup_{i=1}^{l} E_i = \mathcal{U}\}$ such that satisfies $(i), (ii)$ and (iii).

By Theorem 4.6 we can construct an $E_i (E_i \subseteq \mathcal{U})$ so that $\forall A \in \mathcal{U}$,

$$(\{A\} \xrightarrow{s} E_i) \in G,$$

and $\forall E' : E_i \subset E'$ implies $(\{A\} \xrightarrow{s} E') \notin G$.

It is obvious that $A \in E_i$, and easy to see that $l = n$, where $n = \mid \mathcal{U} \mid$.

Then, from (ii), easy to see that $\forall A \in \mathcal{U}$, we have $(\{A\} \xrightarrow{s} \{A\}) \in G$. Assume S5 does not hold, that is if $(X \xrightarrow{s} Y) \in G$ and $(Z \xrightarrow{s} W) \in G$ then

$$((X \cap Z) \xrightarrow{s} \cup W)) \in G. \qquad (4.7.1)$$

Suppose $X \cap Z = \emptyset$ and $Y \cup W = \mathcal{U}$. From (4.7.1), we have $(\emptyset \xrightarrow{s} \mathcal{U}) \notin G$. This contradiction to (i), so S5 holds.

Assume S4 does not hold, that is if $(X \xrightarrow{s} Y) \in G$ and $(Z \xrightarrow{s} W) \in G$, then

$$((X \cap Z) \xrightarrow{s} (Y \cup W)) \notin G. \tag{4.7.2}$$

Suppose $X \cup Z = Z'$,

$$Y \cap W = W' \subseteq \bigcap_{E_i \cap X \neq \emptyset} E_i.$$

From (4.7.2), we have $(Z' \xrightarrow{s} W') \notin G$. this contradiction to (ii), so S4 holds. From $(ii), (iii)$ it is easy to see that S2, S3 hold too. □

Theorem 4.8 *Let G be a full family of s-GPBDs over $\mathcal{U}$. We define the mapping $F_G : P(\mathcal{U}) \times P(\mathcal{U})$ as follow:*

$$F_G(X) = \{A \in \mathcal{U} \mid (X \xrightarrow{s} \{A\}) \in G\}.$$

Then F_G is a strong operation over $\mathcal{U}$. Conversely, if F is an arbitrary strong operation over $\mathcal{U}$, then there is exactly one full family of s-GPBDs G such that $F_G = F$, where

$$G = \{(X \xrightarrow{s} Y) \mid X, Y \in P(\mathcal{U}) : Y \subseteq F(X)\}.$$

Proof: 1. Assume G is a full family of s-GPBDs over $\mathcal{U}$. We show that F_G is a strong operation. Since $F_G(X) = \{A \in \mathcal{U} \mid (X \xrightarrow{s} \{A\}) \in G\}$, so

$$F_G(\{A\}) = \{B \in \mathcal{U} \mid (\{A\} \xrightarrow{s} \{B\}) \in G\}. \tag{4.8.1}$$

By S1, we have that $\forall A \in \mathcal{U}, A \in F_G(\{A\})$. By (i) in Theorem 4.7,

$$\forall C \subseteq \mathcal{U}, (\emptyset \xrightarrow{s} C) \in G.$$

So we have $F_G(\emptyset) = \mathcal{U}$. By Theorem 4.6, and by (4.8.1), we obtain that for $A \in \mathcal{U}$, $F_G(\{A\}) = E_A$. So, by (ii) in Theorem 4.6, we have for $B \in \mathcal{U}$, $(\{B\} \xrightarrow{s} F_G(\{B\})) \in G$. Thus, assume $B \in F_G(\{A\})$, and by (iii) in Theorem 4.7, we have $F_G(\{B\}) \subseteq F_G(\{A\})$.

On the other hand, from (4.8.1) and Theorem 4.6, we have for $A \in \mathcal{U}$, $(\{A\} \xrightarrow{s} F_G(\{A\})) \in G$.

Let $A \in X \subseteq \mathcal{U}$, then by S5 we obtain

$$(X \xrightarrow{s} \bigcap_{A \in X} F(\{A\})) \in G.$$

That is

$$\bigcap_{A \in X} F(\{A\}) \subseteq F_G(X).$$

By the definition of $F_G(X)$, we have $(X \xrightarrow{s} F_G(X)) \in G$. Since for $\forall A \in X, X \cap F_G(\{A\}) \neq \emptyset$, by Theorem 4.7, we obtain $F_G(X) \subseteq F_G(\{A\})$. So

$$F_G(X) \subseteq \bigcap_{A \in X} F(\{A\}).$$

Hence

$$F(X) = \bigcap_{A \in X} F(\{A\}).$$

2. Assume that F is a strong operation over $\mathcal{U}$, and $G = \{(X \xrightarrow{s} Y) \mid Y \subseteq F(X)\}$. We have to show that G is a full family of s-GPBDs. That is, we show that it satisfies $(i), (ii)$ and (iii) in Theorem 4.7.

By Theorem 4.6 and Theorem 4.7, we set

$$E = \{F(\{A\}) : A \in \mathcal{U}, n =| \mathcal{U} |\}.$$

Assume

$$\bigcap_{F(\{A\}) \cap X \neq \emptyset} F(\{A\}) \subseteq F(X).$$

Since $G = \{(X \xrightarrow{s} Y) \mid Y \subseteq F(X)\}$. So if

$$Y \subseteq \bigcap_{F(\{A\}) \cap X \neq \emptyset} F(\{A\}),$$

then it satisfies (ii) in Theorem 4.7

Assume $(V, W) \in G$, and $V \cap F(\{A\}) \neq \emptyset$. Let $B \in V \cap F(\{A\})$, so $B \in V$ and $B \in F(\{A\})$. Thus, by (iii) in the definition of strong operation $B \in F(\{A\})$ implies $F(\{B\}) \subseteq F(\{A\})$. By the definition of G, we have $W \subseteq F(V)$. By (iii) in the definition of strong operation, we have

$$F(V) = \bigcap_{D \in V} F(\{D\}).$$

Since $B \in V$, so

$$\bigcap_{D \in V} F(\{D\}) \subseteq F(\{B\}).$$

Hence $D \subseteq F(\{A\})$, i.e. it satisfies (iii) in Theorem 4.7. It is clear that $\forall A \in \mathcal{U}$, $(\emptyset \xrightarrow{s} \{A\}) \in G$. □

5 Armstrong relation for s-GPBDs

Definition 5.1 *Let Σ_s be a set of s-GPBDs on $\mathcal{U}$, and let R be a relation on $\mathcal{U}$. R exactly represents Σ_s if $S_R = \Sigma_s^+$. If R exactly represents Σ_s then we also say that R is an Armstrong relation for Σ_s.*

Definition 5.2 *Let $R = \{t_1, \ldots, t_m\}$ be a relation over $\mathcal{U}$. We set $E_{ij} = \{A \in \mathcal{U} \mid \alpha_A(t_i[A], t_j[A]) = 1\}$, and $E_R = \{E_{ij}, 1 \le i, j \le m\}$. We denote $E(A) = \bigcap_{A \in E_{ij}} E_{ij}$ if there is a such E_{ij}, in the converse case set $E(A) = \mathcal{U}$, where $A \in \mathcal{U}$. Denote $E^*_R = \{E(A) \mid A \in \mathcal{U}\}$. E^*_R is called the α-attribute-equality set of R.*

A strong relation scheme is a pair $(\mathcal{U}, \Sigma_s)$, where $\mathcal{U}$ is a set of attributes and Σ_s is a set of s-GPBGs on $\mathcal{U}$.

Definition 5.3 *Let $H = < \mathcal{U}, \Sigma_s >$ be a strong relation scheme, $X \subseteq \mathcal{U}$. We set $X^+ = \{A \in \mathcal{U} \mid (X \xrightarrow{s} \{A\}) \in \Sigma_s^+\}$. X^+ is called the closure of X. Denote $I(H) = \{X^+ \mid X \in P(\mathcal{U})\}$. It can be seen that $I(H)$ (for short $I(\Sigma_s)$) is a s-semilattice over $\mathcal{U}$. Denote by $N(H)$ (for short $N(\Sigma_s)$) the minimal generator of $I(H)$.*

It is easy to see that $N(H)$ satisfies (2) in Definition 4.3 and $X^+ \cap Y^+ = (X \cup Y)^+, X^+ = \bigcap_{A \in X} \{A\}^+$.

Theorem 5.4 *Let G be a full family of s-GPBDs, and $R = \{t_1, \ldots, t_m\}$ be a relation over $\mathcal{U}$. Then R represents G iff for each $A \in \mathcal{U}$*

$$F_G(\{A\}) = \begin{cases} \bigcap_{A \in E_{ij}} E_{ij} & \text{if } \exists E_{ij} : A \in E_{ij}, \\ \mathcal{U} & \text{otherwise.} \end{cases}$$

Where $F_G(X) = \{A \in \mathcal{U} \mid (X \xrightarrow{s} \{A\}) \in G\}$, and E_{ij} is the equality set of R.

Proof: Only if: By Theorem 4.8 $S_R = G$ if and only if $F_{S_R} = F$, where F is strong operation over $\mathcal{U}$. We have show that $F_{S_R}(\{a\}) = F_G(\{A\})$ for all $A \in \mathcal{U}$. Clearly,

$$F_{S_R}(\{A\}) = \{B \in \mathcal{U} : (\{A\} \xrightarrow{s} \{B\})\}. \tag{5.4.1}$$

According to the definition of s-GPBDs we know that for any $A \in \mathcal{U}$, and $A \neq \emptyset$ $(\{A\} \xrightarrow{s} Y)$ iff

$$(\forall t_1, t_2 \in R)\alpha_A(t_1[A], t_2[A]) = 1 \longrightarrow (\forall B \in Y)\alpha_B(t_1[B], t_2[B]) = 1.$$

Let $T = \{E_{ij} \mid A \in E_{ij}\}$. It is easy to see that if $T = \emptyset$, then $F_{S_R}(\{A\}) = \mathcal{U}$ holds. If $T \neq \emptyset$. Let

$$X = \bigcap_{A \in E_{ij}} E_{ij}.$$

If $T = E$ (E is the set of all α-attribute equality sets of R), then $(\{A\} \xrightarrow{s} X)$. If $T \subset E$, then for all $E_{ij} \in T$, we have $\alpha_A(t_1[A], t_2[A]) \neq 1$. By (5.4.1), we obtain

$$F_{S_R}(\{A\}) = \bigcap_{A \in E_{ij}} E_{ij}.$$

If: If F_G holds to (5.4.1), then we have $F_G(\{A\}) = F_{S_R}(\{A\})$. By Theorem 4.8, we obtain $F_G = F_{S_R}$.

□

Definition 5.5 *Let R be a relation, an F a strong operation over $\mathcal{U}$. We say that the relation R exectly represents F iff $F_{S_R} = F$.*

Lemma 5.6 [15] *Let F be a strong operation and R a relations over $\mathcal{U}$. Then R represents F iff for all $A \in \mathcal{U}$,*

$$F(\{A\}) = \begin{cases} \bigcap_{A \in E_{ij}} E_{ij} & \text{if } \exists E_{ij} : A \in E_{ij}, \\ U & \text{otherwise.} \end{cases}$$

Theorem 5.7 *Let Σ_s be a set of s-GPBDs on $\mathcal{U}$, and let R be a nonempty relation on $\mathcal{U}$. Then R is an Armstrong relation for Σ_s if and only if*

$$N(\Sigma_s) \subseteq E_R^* \subseteq I(\Sigma_s).$$

Proof: Only if: If R is an Armstrong relation for Σ_s, then by Definition 5.1 $S_R = \Sigma_s^+$. We set $F_{\Sigma_s^+} = X^+$ for all $X \in P(\mathcal{U})$ and

$$F_{S_R}(X) = \{A \in \mathcal{U} \mid (X \xrightarrow{s} \{A\})\}.$$

By Theorem 4.8, $S_R = \Sigma_s^+$ if and only if $F_{S_R} = F$, where F is a strong operation over $\mathcal{U}$. It follows that $F_{\Sigma_s^+} = F_{S_R}$.

By Theorem 4.5 and Definition 5.3, $I(\Sigma_s) = I_{F_{S_R}}$ and $N(\Sigma_s) = N$, where N is the minimal generator of $I_{F_{S_R}}$. In other hand, since

$$F_{S_R}(X) = \bigcap_{A \in X} F_{S_R}(\{A\})$$

for all $X \in P(\mathcal{U})$, so we have to show that $F_{S_R}(\{A\}) = E(A)$ for each $A \in \mathcal{U}$.

Clearly, $F_{S_R}(\{A\}) = \{B \in \mathcal{U} \mid (\{A\} \xrightarrow{s} \{B\})\}$. By the definition of s-GPBD, we know that for any $A \in \mathcal{U}$, $A \neq \emptyset$, $(\{A\} \xrightarrow{s} Y)$ iff

$$(\forall t_i, t_j \in R)(\alpha_A(t_i[A], t_j[A]) = 1) \longrightarrow ((\forall B \in Y)(\alpha_B(t_i[B], t_j[B]) = 1)).$$

Assume $Q = \{E_{ij} \mid A \in E_{ij}\}$. It is obvious that if $Q = \emptyset$ then $F_R(\{A\}) = \mathcal{U}$. If $Q = \emptyset$, then assume that

$$X = \bigcap_{A \in E_{ij}} E_{ij},$$

then it is obvious that $(\{A\} \xrightarrow{s} X)$ and for all $E_{ij} : E_{ij} \notin Q$,

$$(\alpha_A(t_i[A], t_j[A]) \neq 1).$$

Hence,

$$F_{S_R}(\{A\}) = \bigcap_{A \in E_{ij}} E_{ij} = E(A)$$

for all $A \in \mathcal{U}$. Therefore, by Definition 5.3, $E_R^* \subseteq I_{F_R}$.

Now we show that $N(\Sigma_s) \subseteq E_R^*$. By Definition 4.3, Theorem 4.2, and Theorem 4.5, clearly to see that $N(\Sigma_s) \subseteq E_R^*$.

If: Assume that $N(\Sigma_s) \subseteq E_R^* \subseteq I(\Sigma_s)$. Since $E_R^* \subseteq I(\Sigma_s)$, and $I(\Sigma_s) = \{X^+ : X \in P(\mathcal{U})\}$,

$$X^+ = \{A \in \mathcal{U} \mid (X \xrightarrow{s} \{A\}) \in \Sigma_s^+.$$

Thus we obtain $E_R^* = \{F_{\Sigma_s^+}(\{A\}) : A \in \mathcal{U}\}$. By above proof for each $A \in \mathcal{U}$, we have that $E(A) = F_{S_R}(\{A\})$. Hence,

$$\{F_{\Sigma_s^+}(\{A\}) : A \in \mathcal{U}\} = \{F_{S_R}(\{A\}) : A \in \mathcal{U}\}.$$

Suppose $A \in \mathcal{U}$ that $F_{\Sigma_s^+}(\{A\}) \neq F_{S_R}(\{A\})$. By Definition 4.4 and Theorem 4.5 we assume that $F_{\Sigma_s^+} = Y$, where $Y \in N(\Sigma_s)$. Since $N(\Sigma_s) \subseteq E_R^*$, so $F_{\Sigma_s^+} \in E_R^*$. Clearly to see that $F_{\Sigma_s^+}(\{A\}) = E(A)$. This is a contradiction. Therefore, we obtain that $F_{\Sigma_s^+}(\{A\}) = F_{S_R}(\{A\})$ for each $a \in \mathcal{U}$. Thus, $F_{\Sigma_s^+} = F_{S_R}$, and by Theorem 4.8, $S_R = \Sigma^+$.

□

Algorithm 5.8 *(Finding Σ_s)*

(Input :) *Given relation $R = \{t_1, \ldots, t_m\}$ over $\mathcal{U}$.*

(Output :) *Construct Σ_s, such that $S_R = \Sigma_s^+$.*

(Step 1 :) *From R we compute E_R.*

(Step 2 :) *From E_R we construct $E_R^* = \{E(A) : A \in \mathcal{U}\}$.*

(Step 3 :) *Set $\Sigma_s = (\{A\} \xrightarrow{s} E(A)) \mid A \in \mathcal{U}\}$*

Clearly, the time complexity of this algorithm is polynomial in the size of R.

Algorithm 5.9 *(Finding $\{A\}$)*

(Input :) *Given $\Sigma_s = \{(A_i \xrightarrow{s} B_i) \mid i = 1, \ldots, m\}$* and $A \in \mathcal{U}$.

(Output :) *Compute $\{A\}^+$*

(Step 1 :) *$A \in \mathcal{U}$, let $L_0 = \{A\}$*

(Step i+1 :) *If there is an $(A_i \xrightarrow{s} B_i) \in \Sigma_s$*

so that $A_j \cap X^{(i)} \neq \emptyset$ and $B \not\subseteq X^{(i)}$ then

$$X^{(i+1)} = X^i \cup \Big(\bigcup_{A_j \cap X^{(i)}} B_j \Big).$$

In the converse case we set $\{A\}^+ = X^{(t)}$.

It can be seen that the time complexity of this algorithm is polynomial in the sizes of Σ_s and $\mathcal{U}$.

6 Update Problem

In [11], the update problem is introduced for a set of GPBDs Σ. Let R be a relation that satisfies a set of GPBDs Σ and t be a tuple $d_1 \times \ldots \times d_n$. We say that t can be added to R if $R \cup \{t\}$ satisfies Σ.

Theorem 6.1 [11] *Let R be a relation satisfying a set of GPBDs Σ, and let t be a tuple in $d_1 \times \ldots \times d_n$. Then t can be added to R if and only if $(\forall u \in R)(\alpha(t, u) \in T_\Sigma)$.*

Let Σ_s be a set of s-GPBDs, $\Sigma_s = \{X_i \xrightarrow{s} Y_i\}$, where $X_i, Y_i \subseteq \mathcal{U}$. Let $M = \cup X_i, N = \cup Y_i$. By Theorem 6.1 and definition of s-GPBDs, we get the following result.

Theorem 6.2 *Let R be a relation satisfying a set of s-GPBDs Σ_s, $\Sigma_s = \{X_i \xrightarrow{s} Y_i\}$, and let t be a tuple in $d_1 \times \ldots \times d_n$. Then t can be added to R if and only if $(\forall u \in R)(\forall A \in N)(\alpha_A(t[A], u[A]) = 1)$.*

It is easy to see that, if $(\forall u \in R)(\forall A \in M)(\alpha_A(t[A], u[A]) = 0)$. Then t is added to R too.

7 Membership Problem for s-GPBDs

In [11], the membership problem for GPBDs is introduced. Given a set of GPBDs Σ and a GPBD f, decide whether $\Sigma \models f$.

From Algorithms 5.8, 5.9 and $X^+ = \cup\{A\}^+ \ A \in X$. We have the following.

Proposition 7.1 *Let Σ_s be a set of s-GPBDs on $\mathcal{U}$ and $X, Y \subseteq \mathcal{U}$. Then, there is an algorithm deciding whether that $X \xrightarrow{s} Y \in \Sigma_s^+$.*

The time complexity of this algorithm is polynomial in the sizes of Σ_s and $\mathcal{U}$.

Theorem 7.2 [11] *Let Σ be a set of GPBDs on $\mathcal{U}$, and $X, Y, Z \subseteq \mathcal{U}$. Then*

1. $\Sigma \models \wedge X \to \wedge Y \Leftrightarrow$
   $(\forall x \in T_\Sigma)(((\exists A \in X)\,(x(A) = 0)) \vee ((\forall B \in Y)\,(x(B) = 1)))$.
2. $\Sigma \models \wedge X \to \vee Y \Leftrightarrow$
   $(\forall x \in T_\Sigma)(((\exists A \in X)\,(x(A) = 0)) \vee ((\exists B \in Y)\,(x(B) = 1)))$.
3. $\Sigma \models \vee X \to \wedge Y \Leftrightarrow$
   $(\forall x \in T_\Sigma)(((\forall A \in X)\,(x(A) = 0)) \vee ((\forall B \in Y)\,(x(B) = 1)))$.
4. $\Sigma \models \vee X \to \vee Y \Leftrightarrow$
   $(\forall x \in T_\Sigma)(((\forall A \in X)\,(x(A) = 0)) \vee ((\exists B \in Y)\,(x(B) = 1)))$.
5. $\Sigma \models \wedge X \to (\wedge Y \vee \wedge Z) \Leftrightarrow (\forall x \in T_\Sigma)(((\exists A \in X)\,(x(A) = 0)) \vee$
   $(((\forall B \in Y)\,(x(B) = 1)) \vee ((\forall C \in Z)\,(x(C) = 1))))$.

Theorem 7.3 *Let Σ_s be a set of s-GPBDs on U, and $X, Y \subseteq U$. Then*

$$\Sigma_s \models \vee X \rightarrow \wedge Y$$

$$\Sigma_s \models \vee X \rightarrow \vee Y \qquad\qquad \Sigma_s \models \wedge X \rightarrow \wedge Y$$

$$\Sigma_s \models \wedge X \rightarrow \vee Y$$

Proof:
By Theorem 7.2 and definition of s-GPBDs. It is easy to see that Theorem 7.3 holds. □

References

[1] Armstrong W.W., Dependency structures of database relationships. *Information Processing 74, Holland Publ.Co.* (1974), 580-583.

[2] Berman J., Blok W.J., Generalized Boolean Dependencies. *Abstracts of AMS*, **6** (1985), 163.

[3] Berman J., Blok W.J., Positive Boolean Dependencies. *Inf. Processing Letters*, **27** (1988), 147-150.

[4] Berman J., Blok W.J., Equational Dependencies. *Manuscript*, (1990).

[5] Czédli G., Függőségek relációs adatbázis modellben. *Alkalmazott Matematikai Lapok*, **6** (1980), 131-143.

[6] Czédli G., On dependencies in the relational model of data. *J.EIK*, **17** (1981), 103-112.

[7] Demetrovics J., Relációs adatmodell logikai és structurális vizsgálata. *MTA-SZTAKI Tanulmányok, Budapest*, **114** (1980), 1-97.

[8] Demetrovics J.,Gyepesi Gy., On the functional dependencies and some generalizations of it. *Acta Cybernetica*, **5** (1981), 295-305.

[9] Demetrovics J.,Gyepesi Gy., Logical dependencies in relational database. *MTA-SZTAKI Tanulmányok, Budapest*, **133**, (1982), 59-78.

[10] Demetrovics J.,Gyepesi Gy., Some generalized type functional dependencies formalized as equality set on matrics. *Discrete Applied Mathematics*, **6** (1983), 35-47.

[11] Huy N.X., Thanh L.T., Generalized Positive Boolean Dependencies. *J.EIK*, **28** (1992), 363-370.

[12] Maier D. *The Theorem of Relational Databases.* Computer Science Press , (1983).

[13] Sagiv Y., Delobel C., Parker D.S., and Fagin R. An Equivalence Between Relational Database Dependencies and a Fragment of Propositional Logic. *J.ACM*, **28** (1981), 435-453.

[14] Thi V.D. Logical dependencies and irredundant relations. *Computers and Artificial Intelligence*, **7** (1988), 165-184.

[15] Thi V.D. Strong dependencies and s-semilatties. *Acta Cybernetica*, **8** (1987), 195-202.

[16] Ullman J.D. *Principles of Database Systems.* (Second Edition.) Computer Science Press, Potomac, Md., 1982.

Received July 30, 1993