VNU. JOURNAL OF SCIENCE, NAT. SCI., t.XI, n°4 - 1995

ON ARMSTRONG RELATIONS IN THE CLASS O MULTI-VALUED POSITIVE BOOLEAN DEPENDENCIES

Vu Ngoc Loan
College of Natural Sciences - VNU

Abstract. A class of multi-valued positive Boolean dependencies (MVPBD) is introdced. Some results about the present of sets of multivalued positive Boolean dependencies and Armstrag relations for a given set of MVPBDs are presented.

1. INTRODUCTION

In the paper, a type of multivalued logic and some of its properties are presented. On the basis of this logic, a class of multivalued Boolean dependencies that is a generalization of soae kinds of dependencies such as the equational dependencies, the positive Boolean dependencies is ntroduced. The main purpose of the paper is to develop some results about Armstrong relations, which have been obtained from the generalized positive Boolean dependencies in [6]. Some result about the presentation of sets of multivalued positive Boolean dependencies are given. The pape also show that in some cases a m-Armstrong relation for a set of multivalued positive Boolean dpendencies does not exist. With the aid of Equivalence Theorem, a necessary and sufficient condition for a relation is an Armstrong relation for a set of multi-valued positive Boolean depedencies is formulated.

2. BASIS DEFINITIONS

Definition 2.1. Let $U = \{A_1, A_2, ..., A_n\}$ be a nonempty finite set. Say that, a multi-valued logic is defined over U if for each attribute $A_i \in U$, $1 \le i \le n$ there exist a finite set B called the valuation domain of the attribute A_i, which satisfies the following:

1. $B_i \subset [0, 1]$,
2. If $s \in B_i$ then $1 - s \in B_i$ and
3. $1 \in B_i$.

Let $K = \bigcup_{i=1}^{n} B_i$. Each element s in K is called a logical constant. For $s_1, s_2 \in K$ we define logical connectives $\vee$, $\wedge$, $\rightarrow$, $\neg$, $\approx$ on K as follows: $s_1 \vee s_2 = \max\{s_1, s_2\}$, $s_1 \wedge s_2 = \min\{s_1, s_2\}$, $s_1 \rightarrow s_2 = \max\{1 - s_1, s_2\}$, $\neg s_1 = 1 - s_1$, $s_1 \approx s_2 = 1$ if $s_1 = s_2$, and $s_1 \approx s_2 = 0$ if $s_1 \neq s_2$. These connectives are also called disjunction, conjunction, implication, negation and comparison correspondingly.

Let $B = B_1 \times B_2 \times ... \times B_n$. A mapping $x : U \rightarrow K$ such that $x(A_i) \in B_i$ with $1 \le i \le n$ is said to be a valuation over U. If $x(A_i) = x_i$, $1 \le i \le n$, then x is denoted by $(x_1, x_2, .., x_n) \in B$. It is clear that the set of all valuations over U is finite.

Definition 2.2. Elements of U are also called logical variable or elementary variabes. Logical istants in K and logical variables in U are said to be formulas.

Let g, h be formulas, then we can construct new formulas by using logical connectives $\vee$, $\wedge$, $\rightarrow$ ·, $\approx$ and the given formulas. It means that $(g \wedge h)$, $(g \vee h)$, $\neg(g)$, $(g \rightarrow h)$, $(h \approx g)$ are new mulas. By F we denote the set of all formulas constructed from U using logical connectives ntioned above. Each $f \in F$ is said to be a multi-valued Boolean dependency (MVBD).

Assume that $f \in F$, $x = (x_1, x_2, ..., x_n) \in B$. We now denote by $f(x)$ the truth valued of f the valuation x. This value is defined as follows: If f is a variable A_i then $f(x) = x_i$. When f :onstructed by formulas g, h and logical connectives $\vee$, $\wedge$, $\rightarrow$, $\neg$, $\approx$ we define $f(x)$ as follows: $f = (g * h)$ then $f(x) = g(x) * h(x)$, where $* \in \{\vee, \wedge, \rightarrow, \neg, \approx\}$.

ieorem 2.1. *For any $\{x_1, x_2, ..., x_h\} \subseteq B$, $y_1, y_2, ..., y_h \in K$, there always exists a formula f :h that $f(x_i) = y_i$, $1 \leq i \leq h$.*

oof. For each $x_i \in B$ with $1 \leq i \leq h$, suppose $x_i = (x_{i_1}, x_{i_2}, ..., x_{i_n})$. By g_i, $1 \leq i \leq h$ we iote $(A_1 \approx x_{i_1}) \wedge (A_2 \approx x_{i_2}) \wedge ... \wedge (A_n \approx x_{i_n}) \wedge y_i$. Set $f = g_1 \vee g_2 \vee ... \vee g_h$. It is not difficult verify that the formula f satisfies the theorem.

It is clear that, for any $f \in F$, $x \in B$ then $f(x) \in K$. For brevity, instead of $(g \wedge h)$, $(g \vee h)$, $\neg(g)$, $(g \rightarrow h)$, $(h \approx g)$ we write $g \wedge h$, $g \vee h$, $\neg g$, $g \rightarrow h$, $h \approx g$ correspondingly.

finition 2.3. Let T be a set of valuations over U and let g and h be formulas. We say that ind h equivalent over T, written by $g \stackrel{T}{=} h$ iff for any $x \in T$, the equality $g(x) = h(x)$ holds. early, this is and equivalence relation. When $T = B$ we say that g and h equivalent and write $= h$.

Let $U = \{A_1, A_2, ..., A_n\}$ be a nonempty finite set of symbols called attributes. Corresponding each attribute A_i there is a set d_i, $1 \leq i \leq n$ called the domain of A_i. Assume that every d_i itains at least two elements. Suppose that there is a multi-valued logic defined over U. Let iote each domain of A_i, $1 \leq i \leq n$ by B_i, assume that B_i satisfies conditions in Definition 2.1.

A subset R of $d_1 \times d_2 \times ... \times d_n$ is called a relation over U. Elements of R are called tuples and usually denote tuples by u, v or t, etc. The class of all relations over U is denoted by $REL(U)$. r $R \in REL(U)$, $t \in R$, $A \in U$, we denote by $t.A$ the value of t for the attribute A, and by $t.X$ set $\{t.A | A \in X\}$.

finition 2.4. For each set d_i, $1 \leq i \leq n$, we consider a mapping $\alpha_i : d_i \times d_i \rightarrow B_i$, satisfying following:

1. $(\forall a \in d_i)(\alpha_i(a, a) = 1)$,
2. $(\forall a, b \in d_i)(\alpha_i(a, b) = \alpha_i(b, a))$,
3. $(\forall s \in B_i, \exists a, b \in d_i)(\alpha_i(a, b) = s)$.

It is not hard to verify that, the mappings α_i in the above example satisfy the demands of finition 2.4.

For each $m \in [0, 1]$, $f \in F$, $\Sigma \in F$, we set $T^m{}_f = \{x \in B | f(x) \geq m\}$ and $T^m{}_\Sigma = \{x \in$ $\forall f \in \Sigma,\ f(x) \geq m\}$.

finition 2.5. Let f and g be formulas and $m \in [0, 1]$. We say that g m-implies f or f is implied from g, written $g|\stackrel{m}{-}f$, iff for any $x \in B$, satisfying $g(x) \geq m$, then $f(x) \geq m$. Two mulas f and g are said to be m-equivalent if $f|\stackrel{m}{-}g$ and $g|\stackrel{m}{-}f$. For $m \in [0, 1]$, $\Sigma \subseteq F$, $f \in F$, that, Σ m-implies f or f is m-implied from Σ denoted by $\Sigma|\stackrel{m}{-}f$ iff $T^m{}_\Sigma \subseteq T^m{}_f$. Let $\Gamma \subseteq F$. The set Γ is said to be m-implied from Σ denoted by $\Sigma|\stackrel{m}{-}\Gamma$ if $\Sigma|\stackrel{m}{-}f$ for any $f \in \Gamma$. say that, Σ and Γ are m-equivalent, written $\Sigma \stackrel{m}{=} \Gamma$, if $\Sigma|\stackrel{m}{-}\Gamma$ and $\Gamma|\stackrel{m}{-}\Sigma$.

Each formula $f \in F$ is said to be a multi-valued positive Boolean dependency (MVPBD) if $f(e) = 1$ for $e \in B$ and $e = (1, 1, ..., 1)$. By F_p we denote the set of all MVPBDs oer U. Let $R \in REL(U)$ and $u, v \in R$. Then $(\alpha_1(u.A_1, v.A_1), ..., \alpha(u.A_n, v.A_n))$ is said to be valuation over U and denoted by $\alpha(u, v)$. We set $T_R = \{\alpha(u, v)\ u, v \in R\}$. Note that, for any $R \in REL(U)$, we always have $e \in T_R$.

Definition 2.6. Let R be a relation over U and $f \in F_p$. We say that the relation R m-satisfies MVPBD f, denoted by $R^m(f)$, if $T_R \subseteq T^m{}_f$. For $\Sigma \subseteq F_p$, the relation R is said to m-satisfy the set Σ of MVPBDs, written $R^m(\Sigma)$ if $R^m(f)$ for any $f \in \Sigma$. This is equivalent to $T_R \subseteq T^m{}_f$. If $R \in REL(U)$ and R does not m-satisfy f, then we written $\neg R(f)$. If R does not m-satisfy Σ, write $\neg R^m(\Sigma)$.

Definition 2.7. For $m \in [0, 1]$, $\Sigma \subseteq F_p$ and $f \in F_p$, $\Sigma| \overset{m}{=} f$ means that, for all $R \in REL(U)$ if $R^m(\Sigma)$ then $R^m(f)$.$\Sigma| \overset{m}{=}_2 f$, means that, for all $R \in REL(U)$ and R has only tw tuples. if $R^m(\Sigma)$ then also $R^m(f)$.

3. m-REPRESENTATION AND ARMSTRONG RELATION

Let $\Sigma \subseteq F_p$. We set $SAT^m(\Sigma) = \{R|R \in REL(U)$ and $R^m(\Sigma)\}$. In case $\Sigma = \{f\}$ instead of $SAT^m(\{f\})$ we write $SAT^m(f)$. For $R \in REL(U)$, we define $LD^m(R) = \{f|f \in F_p, R^*(f), R \in REL(U)\}$. When $R = \{R\}$, we use $LD^m(R)$ instead of $LD^m(R)$. Thus $LD^m(R) = \{f|f \in F_p, R^m(f)\}$. It is obvious that $LD^m(R) = \bigcap_{R \in R} LD^m(R)$.

Definition 3.1. Let $\Sigma \subseteq F_p$, $f \in F_p$ and R be a relation over U. We define $\Sigma_m{}^+ = \{f|\Sigma| \overset{m}{=} f\}$. We say that the relation R m-represents Σ if $LD^m(R) \supseteq \Sigma_m{}^+$, R exactly m-reprsents Σ if $LD^m(R) = \Sigma_m{}^+$. If R exactly m-represents Σ then R is said to be an m-Armstrong elation for Σ.

Theorem 3.1. (Equivalence Theorem) [8]. *Let $m \in [0, 1]$, $\Sigma \subseteq F_p$, $f \in F_p$, then the folowing are equivalent*

1. $\Sigma| \overset{m}{—} f$ 2. $\Sigma| \overset{m}{=} f$ 3. $\Sigma| \overset{m}{=}_2 f$

From Definition 3.1 and Equivalence Theorem, it is not hard for us to get the tw following corollaries.

Corollary 3.1. *Let Σ, $\Gamma \subseteq F_p$, and let $m \in [0, 1]$. The following are equivalent*

1. $\Sigma \overset{m}{=} \Gamma$
2. $\Sigma| \overset{m}{—} \Gamma$ and $\Gamma| \overset{m}{—} \Sigma$
3. $\Sigma| \overset{m}{=} \Gamma$ and $\Gamma| \overset{m}{=} \Sigma$
4. $\Sigma_m{}^+ = \Gamma_m{}^+$
5. $T^m{}_\Sigma = T^m{}_\Gamma$

Corollary 3.2. *Let $m \in [0, 1]$, $\Sigma \subseteq F_p$. The following hold*

1. $\Sigma_m{}^+ \overset{m}{=} \Sigma$
2. $T^m{}_{\Sigma_m{}^+} = T^m{}_\Sigma$

Lemma 3.1. *Let $m \in [0, 1]$ and R be a nonempty relation over U. Then there awass exists a MVPBD f such that $T^m{}_f = T_R$.*

Proof. Suppose k is the smallest integer in K such that $k \geq m$. It is clear that $e \in T_R$. Taking Theorem 2.1 into account we can construct a formula f, which satisfies $f(e) = 1$ aid for any

T_R with $x \neq e$ then $f(x) = k$, in case $x \in B \setminus T_R$ then $f(x) = 0$. It is not hard to verify that atisfies the demands of the lemma. The proof is complete.

rollary 3.3. *Let $\Sigma, \Gamma \subseteq F_p$, $\boldsymbol{R}, \boldsymbol{S} \subseteq REL(U)$ and let $f \in F_p$. The following are valid:*

1. $SAT^m(\Sigma) = SAT^m(\Gamma)$ *iff* $T^m{}_\Sigma = T^m{}_\Gamma$
2. *If* $\Sigma \subseteq \Gamma$ *then* $SAT^m(\Sigma) \supseteq SAT^m(\Gamma)$
3. $SAT^m(\Sigma \cup \Gamma) = SAT^m(\Sigma) \cap SAT^m(\Gamma)$
4. $SAT^m(\Sigma \cap \Gamma) \supseteq SAT^m(\Sigma) \cup SAT^m(\Gamma)$
5. $\Sigma \subseteq LD^m(SAT^m(\Sigma))$
6. *If* $\boldsymbol{R} \subseteq \boldsymbol{S}$ *then* $LD^m(\boldsymbol{R}) \supseteq LD^m(\boldsymbol{S})$
7. $\boldsymbol{R} \subseteq SAT^m(LD^m(\boldsymbol{R}))$
8. $LD^m(\boldsymbol{R} \cup \boldsymbol{S}) = LD^m(\boldsymbol{R}) \cap LD^m(\boldsymbol{S})$
9. $LD^m(\boldsymbol{R} \cap \boldsymbol{S}) \supseteq LD^m(\boldsymbol{R}) \cup LD^m(\boldsymbol{S})$
10. $SAT^m(LD^m(SAT^m(\Sigma))) = SAT^m(\Sigma)$
11. $LD^m(SAT^m(LD^m(\boldsymbol{R}))) = LD^m(\boldsymbol{R})$

oof. It is not hard to verify properties 1, 2, 3, 4.

5. To prove this property, it is enough for us to show that if $f \in LD^m(SAT^m(\Sigma))$ then $f \notin \Sigma$. t $\boldsymbol{R} = SAT^m(\Sigma)$. Because if $f \notin LD^m(SAT^m(\Sigma))$ then $f \notin LD^m(\boldsymbol{R})$ so there is $R \in \boldsymbol{R}$ such at $\neg R^m(f)$, t.e. $f \notin \Sigma$.

6. Let $f \in LD^m(S)$, then for any $S \in \boldsymbol{S}$ we have $S^m(f)$, it mean that $T_s \subseteq T^m{}_f$. Since $\boldsymbol{R} \subseteq \boldsymbol{S}$ en for any $R \in \boldsymbol{R}$, we have $T_R \subseteq T^m{}_f$, it means $R^m(f)$.

7. Set $\Sigma = LD^m(\boldsymbol{R})$, suppose $R \notin SAT^m(LD^m(R))$ then $R \notin SAT^m(\Sigma)$ therefore there exist $\in \Sigma = LD^m(\boldsymbol{R})$ such that $\neg R^m(f)$. This proves that $R \in \boldsymbol{R}$. Property 7. has been proved.

The other parts of the corollary are also easy to verify. The proof is completed.

heorem 3.2. *Let Σ be the set of MVPBDs over U, and let R be a nonempty relation over U. r $m \in [0, 1]$, then a necessary and sufficient condition for a relation R be an m-Armstrong lation for the set Σ of MVPBDs is $T_R = T^m{}_\Sigma$.*

roof.

1. Sufficient condition. Suppose we have $T_R = T^m{}_\Sigma$, then $T_R \subseteq T^m{}_\Sigma$ therefore $R^m(\Sigma)$. ecause of $T_R \subseteq T^m{}_\Sigma$ and taking Corollary 3.2 into account we obtain $R^m(\Sigma_m{}^+)$. Hence $\Sigma_m{}^+ \subseteq$ $D^m(R)$ (1). On the other hand, for any $f \in LD^m(R)$ then $R^m(f)$, which means that $T_R \subseteq T^m{}_f$. he hypothesis $T_R = T^m{}_\Sigma$ implies $T^m{}_\Sigma \subseteq T^m{}_f$, so $\Sigma \overset{m}{\models} f$. Based on Equivalence Theorem 3.1. e can get $f \in \Sigma_m{}^+$ and so $LD^m(R) \subseteq \Sigma_m{}^+$ (2). The assertions (1) and (2) complete the proof r the sufficient condition of the theorem.

2. Necessary condition. Let R be a m-Armstrong relation for the set Σ of MVPBDs. This eans that $LD^m(R) = \Sigma_m{}^+$. We let $\Gamma = LD^m(R)$. It is easy to see that $\Gamma_m{}^+ = \Gamma = \Sigma_m{}^+$. y Corollary 3.2 we have $T^m{}_\Sigma = T^m{}_\Gamma$. From $R^m(LD^m(R))$ we obtain $T_R \subseteq T^m{}_\Gamma$, therefore $_R \subseteq T^m{}_\Sigma$ (3). Now we have to show that $T^m{}_\Sigma \subseteq T_R$. Indeed, using Lemma 3.1, we can onstruct a MVPBD f such that $T^m{}_f = T_R$. From that equality we get $T_R \subseteq T^m{}_f$ and therefore $\in LD^m(R) = \Sigma_m{}^+$. Thus we have proved $\Sigma \overset{m}{\models} f$. Using Equivalence Theorem 3.1 we infer $\overset{m}{\vdash} f$ (4). For any $x \in T^m{}_\Sigma$ from (4) we obtain $x \in T^m{}_f = T_R$. Hence $T^m{}_\Sigma \subseteq T_R$ (5). From (3) nd (5) we complete the proof of the necessary condition of the theorem.

heorem 3.3. *For $m \in [0, 1]$, $\Sigma \subseteq F_p$, then m-Armstrong relation for Σ does not always exist.*

roof. We shall show that by an anti-example. Suppose that U is a set of two attributes A and

B having correspondingly their domains to be $dom(A) = \{a_1, a_2, a_3\}$, $dom(A) = \{b_1, b_2, b_3\}$ Mappings α_j, $1 \le j \le 2$ are defined as follows:

$$\alpha_1(a_i, a_i) = \alpha_2(b_i, b_i) = 1, \ 1 \le i \le 3,$$
$$\alpha_1(a_1, a_2) = \alpha_1(a_2, a_1) = \frac{2}{5}, \ \alpha_1(a_1, a_3) = \alpha_1(a_3, a_1) = 0,$$
$$\alpha_1(a_2, a_3) = \alpha_1(a_3, a_2) = \frac{3}{5}, \ \alpha_2(b_1, b_2) = \alpha_2(b_2, b_1) = 0,$$
$$\alpha_2(b_1, b_3) = \alpha_2(b_3, b_1) = \frac{1}{4}, \ \alpha_2(b_2, b_3) = \alpha_2(b_3, b_2) = \frac{3}{4}.$$

Remark 1:

1. If $a_i \neq a_j$ then $\alpha_1(a_i, a_j) \neq 1$.
2. $\alpha_2(b_i, b_j) = \frac{3}{4}$ iff $(b_i, b_j) = (b_2, b_3)$ or $(b_i, b_j) = (b_3, b_2)$.

Suppose $T \subseteq B$, $T = \{(1, 1), (\frac{2}{5}, \frac{3}{4}), (\frac{3}{5}, \frac{3}{4})\}$. By applying Lemma 2.2 there exista formula such that $T^m{}_f = T$. Setting $\Sigma = \{f\}$ we get $T^m{}_\Sigma = T$.

Suppose P is a relation containing all tuples over $\{A, B\}$. Let Q be subrelation of$P$ such tha $T_Q \subseteq T^m{}_\Sigma$.

Remark 2:

1. There are not two tuples u, v in the relation Q such that $u.B = v.B$.
2. From Remark 1 we see that for any $u \in Q$, then $u.B = b_2$ or $u.B = b_3$.
3. Paying attention to Remark 2.1, 2.2 we conclude that the number of tuplesin Q is no greater than two therefore we can not find any subrelation of P such that $T_Q = T^m{}_\Sigma$. The proo is complete.

REFERENCES

1. Armstrong W. W. Dependency Structure of Database Relationships. Proc. IFP 74, Nortl Holland, Amsterdam, (1974), 580-583.
2. Beeri C., Dowd M., Faging R., and Statman R. On the structure of Armstrong lelations foi Functional Dependencies. J. ACM, 31, 1 (Jan. 1984), 30-46.
3. Berman J. and Blok W. J. Generalized Boolean dependencies. Abstract of AMS, 6 (1985), 163.
4. Berman J. and Blok W. Y. Positive Boolean dependencies. Inf. Processing Letter, 27 (1988), 147-150.
5. Maier D. The theory of Relational Databases. Computer Sciences Press, Rockvill Md., 1983.
6. Nguyen Xuan Huy and Le Thi Thanh. Generalized Positive Boolean Dependencie J. Inform. Process. Cybernet. EIK, 28, 6 (1992), 363-370.
7. Sagiv Y., Delobel C., Parker D. S., and Faging R. An Equivalence Between Relationl Database Dependencies and a Fragment of Propositional Logic. J. ACM, 28, 3 (1981), 435-453. Also a correction to this paper in J. ACM, 34, 4 (1987), 1016-1018.
8. Vu Ngoc Loan and Nguyen Xuan Huy. On the Equivalence Theorem for a class ofmultivalued positive Boolean dependencies. Journal of Computer Science and Cybernetics Vol. 10, 4 (1994), 11-15 (in Vietnamese).

CHÍ KHOA HỌC ĐHQGHN, KHTN, t.XI, n°4, 1995

VỀ QUAN HỆ ARMSTRONG TRONG LỚP CÁC PHỤ THUỘC LOGIC DƯƠNG ĐA TRỊ

Vũ Ngọc Loãn
Đại học Khoa học tự nhiên - ĐHQGHN

Trên cơ sở trình bày về một kiểu phụ thuộc Boole dương đa trị mà nó là tổng quát hóa của số các lớp phụ thuộc lôgic như phụ thuộc cân bằng, phụ thuộc Boole dương tổng quát ... bài đã phát triển một số kết quả về các quan hệ Armstrong mà chúng đã được đề cập đến trong phụ thuộc Boole dương tổng quát [6]. Ngoài việc đưa ra một số kết quả về việc biểu diễn các phụ thuộc trong các lớp phụ thuộc Boole dương đa trị, kết quả chính trong bài báo là đưa ra tiêu chuẩn cần và đủ để cho một quan hệ R là quan hệ m-Armstrong cho một tập các phụ thuộc Boole dương đa trị cho trước. Trên cơ sở đó bài báo cũng khẳng định rằng với một tập Σ phụ thuộc Boole dương đa trị và với một tham số m thì nói chung quan hệ Armstrong cho nó không phải lúc nào cũng tồn tại. Một số khẳng định tương đương đề cập trong các hệ quả 3.1, cũng là có lợi khi xem xét đến bài toán viên cũng như các quan hệ m-Armstrong trong lớp phụ thuộc Boole dương đa trị.