*Tạp chí Tin học và Điều khiển học, T. 13, S. 3 (1997) (89–93)*

# ARMSTRONG RELATIONS AND STRONG DEPENDENCIES

VU DUC THI

**Abstract.** In this paper, the concept of strong scheme is introduced. We prove that the membership problem for strong dependencies is solved by an algorithm in polynomial time.

We give a necessary and sufficient condition for a relation to be Armstrong relation of a given strong scheme.

**Keyword and Phrases:** Relation, relation datamodel, strong scheme, membership problem, SD-relation implication problem, SD-relation equivalence problem, closure, closed set.

## RESULTS

**Definition 1.** Let $U$ be a nonempty finite set, $R = \{h_1, ..., h_m\}$ relation over $U$ and $A$, $U \subseteq U$. We say that $B$ strongly depend on $A$ in $R$ (denote $A \xrightarrow[R]{s} B$) if $\forall h_i$, $h_j \in r$ have $h_i(a) = h_j(a)$ for some $a \in A$ implies $h_i(b) = h_j(b)$ for all $b \in B$. Let $S_R = \{(A, B) : A \xrightarrow[R]{s} B\}$.

$R_R$ is called a full family of strong dependencies of $R$. Where we write $(A, B)$ or $A \to B$ when $R$, $s$ are clear from the context.

**Definition 2.** Let $U = \{a_1, ..., a_n\}$ be a nonempty finite set attributes. A strong dependency $(SD)$ is a statement of the form $A \to B$, where $A, B \subseteq U$. The $SD$ $A \to B$ hold in a relation $R = \{h_1, ..., h_m\}$ over $U$ if $A \xrightarrow[R]{s} B$. We also say that $r$ satisfies the $SD$ $A \to B$.

**Definition 3.** Let $U = \{a_1, ..., a_n\}$ be a nonempty finite set attributes and $P(U)$ its power set. Let $Y \subseteq P(U) \times P(U)$. We say $Y$ is a $s$-family if all $A$, $B$, $C$, $D \subseteq U$ and $a \in U$

$S1:$ $(\{a\}, \{a\}) \in Y$

$S2:$ $(A, B) \in Y$, $(B, C) \in Y$, $B$ is no empty set $\Longrightarrow (A, C) \in Y$,

$S3:$ $(A, B) \in Y$, $C \subseteq A$, $D \subseteq B \Longrightarrow (C, D) \in Y$,

$S4:$ $(A, B) \in Y$, $(C, D) \in Y \Longrightarrow (A \cup C, B \cap D) \in Y$,

$S5: (A, B) \in Y, (C, D) \in Y \Longrightarrow (A \cap C, B \cup D) \in Y.$

Clearly, $S_R$ is a $s$-family over $U$. It is know [6] that if $Y$ is an arbitrary $s$-family, then there is a relation $R$ over $U$ such that $S_R = Y$.

**Definition 4.** A strong scheme $G$ is a pair $\langle U, S \rangle$, where $U$ is a finite set of attributes, and $S$ a set of $SDs$ over $U$.

Let $S^+$ be a set of all $SDs$ that can be derived from $S$ by rules Definition 3.

It can be seen [6] that if $G = \langle U, S \rangle$ is a strong scheme then there is a relation $R$ over $U$ such that $S_R = S^+$. Such a relation called Armstrong relation of $G$.

**Definition 5.** The mapping $F: P(U) \to P(U)$ is called a strong operation over $U$ if every $a, b \in U$ and $A \in P(U)$ the following properties hold:
(1) $F(0) = U$,
(2) $a \in F(\{a\})$,
(3) $b \in F(\{b\})$ implies $F(\{b\}) \subseteq F(\{a\})$,
(4) $F(A) = \bigcap_{a \in A} F(\{a\})$.
Clearly, if $A \subseteq B$ then $F(B) \subseteq F(A)$, and $F(A) \cup F(B) = F(A \cup B)$. It can be seen that the set $\{F(\{a\}) : a \in U\}$ determines the set $\{F(A) : A \in P(U)\}$.

**Definition 6.** *Let $R = \{h_i, ..., h_m\}$ be a relation over $U$ and $A, B \subseteq U$. Then we say $B$ functionally depends on $A$ in $R$ (denote $A \to B$) if*

$$\forall h_i, h_j \in R) \ (\forall a \in A) \ (h_i(a) = h_j(a)) \Longrightarrow (\forall b \in B) \ (h_i(b) = h_j(b))$$

Let $F_R = \{(A, B) : A, B \subseteq U, A \xrightarrow[R]{s} B\}$

**Theorem 1.** [11] *Let $S$ a $s$-family over $U$. We define the mapping $F_s$ as follows: $F_s(A) = \{a \in U : (A, \{a\}) \in S\}$. Then $F_s$ is a strong operation over $U$. Conversely, if $F$ is a strong operation over $U$ then there is exactly one $s$-family over $U$ such that $F_s = F$, where $S = \{(A, B) : B \subseteq F(A)\}$.*

This theorem shows that between $s$-familys and strong operation there exists an one-to-one corresponding.

**Lemma 1.** *Let $G = \langle U, S \rangle$ be a strong sheme, and $A \in P(U)$. We set $A^+ = \{a \in U : A \to a \in S^+\}$. Then $A \to B \in S^+$ holds iff $B \subseteq A^+$ hold.*

*Proof.* If $B \in A^+$ then by definition of $A^+$ we have $A \to \{a\}$ $\forall a \in B$. By S5 $A \to B$ holds. The converse case is clear.

**Algorithm 1.** (Finding $\{a\}^+$)

*In put:* Given a strong scheme $G = \langle U, S = \{A_i \rightarrow B_i : i = 1, ..., m\}\rangle$, $a \in U$.

*Output:* Compute $\{a\}^+$.

We compute $\{a\}^+$ by induction

Step 1: $a \in U$ we set $X^0 = \{a\}$

Step $i+1$: If there is an $A_j \rightarrow B_j \in S$ so that $A_j \cap X^{(i)} \neq 0$ and $B \not\subset X^{(i)}$ then $X^{(i+1)} = X^{(i)} \cup (\bigcup_{A_j \cap X^{(i)} \neq 0 Z} B_j)$.

In the converse case we set $\{a\}^+ = X^1$. It is easy to see that there is a $t$ such that

$\{a\} =\subseteq X^{(1)} \subseteq \cdots \subseteq X^{(t)} = X^{(t+1)} = \cdots$ and we set $\{a\}^+ = X^{(t)}$.

**Proposition 1.** *For each $a \in U$, Algorithm 1 computes $\{a\}^+$.*

*Proof.* We have to prove that an attribute $a' \in \{a\}^+$ iff $a' \in X^{(t)}$ holds.

$\Leftarrow$: we have prove by the induction. It is obvious that $X^0 = a \in \{a\}^+$. We assume that $X(1) \subseteq \{a\}^+$, and $a' \in X^{(i+1)} - X^{(i)}$. Then there is $A_j \rightarrow B_j \in S$ so that $acapX^{(i)} \neq 0$, $a' \in B_j - X^{(i)}$. By the inductive hypothesis $\{a\} \rightarrow X^{(i)}$ holds. According to $S3$ we have $X^{(i)} \rightarrow T = A_j \cap X^{(i)}$. hence, $\{a\} \rightarrow T$ holds. By $S3$ we have $T \rightarrow B_j$. By S2 $\{a\} \rightarrow B_j$ holds. Consequently, $\{a\} \rightarrow T$ holds, i.e. $a' \in \{a\}^+$.

$\Rightarrow$: From rules $S4$ and $S5$ we can see that $SD$ $\{a_{i1}, ..., a_{is}\} \rightarrow \{b_{j1}, ..., b_{jp}\}$ is equivalent to the set of $DSs$ $\{\{a_{i1}\} \rightarrow \{b_{j1}\}, ..., \{a_{is}\} \rightarrow \{b_{jp}\}\}$. Thus, we can assume that the set $S$ only contains $SDs$ form $\{b\} \rightarrow \{c\}$ we call a sequence $SDs$ $(f_1, ..., f_m)$ is a derivation of a $SD$ $\{a\} \rightarrow B$ iff $f_m = \{a\} \rightarrow B$ and for each $i : 1 \leq i \leq m$ one of the following holds:

(1) $f_i \in S$ or $f_i = \{a\} \rightarrow \{a\}$,

(2) $f_i$ is the rusult of applying the $S2$ to two of $SDs$ $f_1, ..., f_{i-1}$,

(3) $f_i$ is the rusult of applying the $S3$ to one of $SDs$ $f_1, ..., f_{i-1}$,

(4) $f_i$ is the rusult of applying the $S4$ or $S5$ to two of $SDs$ $f_1, ..., f_{i-1}$,

By induction on the leghth of the (shortest) derivation of $\{a\} \rightarrow B$ we show the converse. Because the set $S$ only contains the $Sds$ form $\{b\} \rightarrow \{c\}$ and by rules $S4$ we assume (without loss of generality) that in the derivation of a $SDs\{a\} \rightarrow B$ $f_1 = \{a\} \rightarrow \{a_j\}$, where $a_j \in U$ and each $f_i$ has the form $\{a\} \rightarrow C$ or $f_i \in S$. We assume that $a' \in \{a\}^+$ and $a' \in B$. If the sequence $\{f_1, ..., f_m\}$ of $\{a\} \rightarrow B$ has $m = 1$ then it is obvios that $B \subseteq X^{(t)}$. Now we consider the devation $\{f_1, ..., f_{i+1}\}$ of $\{a\} \rightarrow B$. If $f_{i+1} = \{a\} \rightarrow B = \{a'\}$ then by the construction of $X$ we obtain $a' \in X^{(t)}$. It is easy to see that if $f_{i+1}$ is the result of applying $S2$

to $f_p$ and $f_p$ $(1 \leq p \leq i,\ 1 \leq q \leq i)$ or is result of applying $S3$ to $f_s$ $(1 \leq s \leq i)$ then by the induction hypothesis and the construction of $X^{(t)}$ $B \subseteq X^{(t)}$ holds. If $f_{i+1}$ is the result of applying $S5$ ($S4$) two $SDs$ $f_p = \{a\} \rightarrow C$ and $f_q = \{a\} \rightarrow D$ $(p,\ q \leq i)$ then by the induction hypothesis there are $X^{(l)}$ and $X^{(h)}$ such that $C \subseteq X^{(l)}$, $D \subseteq X^{(h)}$. We set $s = \max(l,\ h)$, then $B = C \cup D \subseteq X^{(s)}$ and $B = C \cap D \subseteq X^{(l)}$. Thus, $B \subseteq X^{(t)}$, i.e. $a' \in X^{(t)}$. The proposition is proved.

It can be seen that Algorithm 1 is polynomial time in the $|U|$, $|G|$, and $A^+ = \bigcap_{a \in A} \{a\}^+$. Thus, the following proposition is clear.

**Proposition 2.** (The membership problem)

*Let $G = \langle U,\ S\rangle$ be a strong scheme over $U$, and $A \rightarrow B$ is a strong dependency, then there is a pollynomial time algrithm deciding whether $A \rightarrow B \in S^+$.*

**Definition 7.** Let $R = \{h_1, ..., h_m\}$ be a relation over $U$. $R_E$ is the equality set of $R$, i.e. $E_{ij} = \{a \in U : h_i(a) = h_j(a)\}$ and $E_R = \{E_{ij} : 1 \leq \subset \leq j \leq m\}$. We set $E(a) = \bigcap_{a \in E_{ij}} E_{ij}$ if there is a such $E_{ij}$ in the converse case set $E(a) = U$, where $a \in U$.

Denote $E_R^* = \{E(a) : a \in U\}$. $E_R^*$ is called the attribute-equality set of $R$.

Clearly, we can compute $E_R^*$ in polynmial time in the size of $R$.

**Theorem 2.** *Let $G = \langle U,\ S\rangle$ be a strong scheme, $R = \{h_1, ..., h_m\}$ a relation over $U$ and $E_R^*$ in the attribute equality set of $R$.*

*Then necessary and sufficient condition for $R$ to be Armtrong relation of strong sheme $G$ is for each $a \in U : \{a\}^+ = E(a)$, where $E(a) \in E_R^*$.*

*Proof.* We set $A_R^+ = \{a \in U\ :\ A \xrightarrow[R]{s} \{a\}\}$.

Now we show that $\{a\}_R^+ = E(a)$ for each $a \in U$. By definition of strong dependency we know that for any $a \in U : \{a\} \xrightarrow[R]{s} B$ iff $\{a\} \xrightarrow[R]{f} B$.

Denote $H = \{E_{ij} : a \in E_{ij}\}$ It can be seen that if $H = \emptyset$ then $\{a\} \xrightarrow[R]{s} U$.

We assume that $H \neq \emptyset$. It is easy to see that if $H = E_R$ holds, then by the definition of $E(a)$ and definiton of strong dependency $\{a\} \xrightarrow[R]{s} E(a)$ holds.

If $H \subseteq E_R$ holds then for $E_{ij} \notin H$ we obatain $h_i(a) \neq h_j(a)$. Consequently, $\{a\} \xrightarrow[R]{s} E(a)$.

It can seen that $\forall E : E \supset E(a)$ we obtain $\{a\} \xrightarrow[R]{s} E(a)$.

According to the definition of $\{a\}_R^+$ we have $\{a\}_R^+ = E(a)$.

Clearly, by Theorem 1 we can see that $S_R = S^+$ iff for each $a \in U : \{a\}^+ = \{a\}_R^+$ holds. Thus, if $S_R = S^+$ holds then $\{a\}^+ = E(a)$ for all $a \in U$.

Conversely, according to Theorem 1 and base on $\{a\}_R^+ = E(a)$ for all $a \in U$ we obain $S_R = S^+$. The theorem is proved.

## REFERENCES

1. Arstrong W. W. *Dependency Structutes of Database Relationships.* Information Processing 74, Holland publ. Co., (1974) 580-583.
2. Beeri C., Dowd M, Fagin R., Staman R., *On the Structure of Armstrong relations for Functional Dependencies.* J. ACM, **31** (1) (1984) 30-46.
3. Beeri C., Bernstein P. A., *Compuiational problems related to the design of normal form relational schemas.* ACM Trans. on Database syst., **4** (1) (1979) 30-59.
4. Burosch G., Demetrovics J., Katona G. O. H., *The poset of closures as a model of changing databases,* Order 4 (1987) 127-142.
5. Czedli G. *On the dependencies in the relational model of data.* J. EIK, **17** (2-3) (1981) 103-112.
6. Demetrovics J., *Relacios adatmodell logikai es strukturalis vizsgalata.* MTA-SZTAKI Tanulmanyok, Budapest, 114 (1980) 1-97.
7. Demetrovics J., Gyepesi G., *On the functional dependency and generalizations of it.* Acta Cybernetica Hungary, Vol. 3 (1983) 295-305.
8. Demetrovics J., Thi. V. D., *Relations and minimal keys.* Acta Cybernetica Hungary VIII/3 (1988) 279-285.
10. Thi V. D., *Minimal keys and Antikeys.* Acta Cybernetica Hungary, Tom. 7, Fasc. 4 (1986) 361-371.
11. Thi. V. D. Strong dependencien and s-semilattices. Acta Cybernatica, **VIII** (2) (1987) 195-202.
12. Thi V. V., *Logical dependencies and irredundant relations.* Computers and Artificial Intelligence, **7** (1988) 165-184.
13. Gottlob G., Libkin L., *Investigations on Armtrong ralations, dependency inference, and excluded functional dependencies.* Acta Cybernetica Hungary, Tom. 9 Fasc. 4 (1990) 385-402.
14. Mannila H., Raiha K. J., *Algorithms for inferring functional dependencies.* Proc. of the thirteenth international conferebce on very large data bases, Brighton, September, 1987.
15. Lucchesi C. L., Osborn S. L., *Candidate keys for relations.* J. Comput. Syst. Scien., **17** (2) (1978) 270-279.
16. Demetrovics J., Thi V. D., *On the algorithms for generating Armstrong relations and for inferring functional depencies in the relation datamodel.* Computer and Mathematics with Applications, Gred Britain, **26** (4) (1993) 43-55.
17. Demetrovich J., Thi V. D., *Some relults about normal form for functional dependency in the relational.* Discrete Applied Mathematics, **69** (1966) 61-74.

*Institute of Information Technology*
*National Centre for Sciences and Technology of Vietnam.*

*Received: April 30, 1996*