Strong dependencies and *s*-semilattices

VU DUC THI

1. Introduction

The full family of functional dependencies was first axiomatized by W. W. Armstrong [1]. Different kinds of functional dependencies have also been investigated in relational data base theory. The full family of strong dependencies has been introduced and axiomatized [2], [3], [4].

In this paper *s*-semilattices and strong operations are defined. We investigate connections between full families of strong dependencies, *s*-semilattices and strong operations. We prove that there are one-to-one correspondences between them, and *s*-semilattices completely determine both full families of strong dependencies and strong operations. We give a necessary and sufficient condition for an arbitrary family of sets to be a full family of strong dependencies. A necessary and sufficient condition for a relation to represent a given full family of strong dependencies is also given. Finally, we show that for a given *s*-semilattice I, we can construct a concrete relation R, the full family of strong dependencies of which is determined by I.

We start with some necessary definitions formulated in [3].

Definition 1.1. Let $R=\{h_1, \ldots, h_m\}$ be a relation over the finite set of attributes Ω, and $A, B \subseteq \Omega$. We say that B strongly depends on A in R (denoted $A \xrightarrow[R]{s} B$) iff

$$(\forall h_i, h_j \in R)((\exists a \in A)(h_i(a) = h_j(a)) \rightarrow (\forall \in B)(h_i(b) = h_j(b))).$$

Let $S_R = \{(A, B)\colon A \xrightarrow[R]{s} B\}$. S_R is called the full family of strong dependencies of R.

Definition 1.2. Let Ω be a finite set, and denote by $P(\Omega)$ its power set. Let $Y \subseteq P(\Omega) \times P(\Omega)$. We say that Y is a full family of strong dependencies over Ω if for all $A, B, C, D \subseteq \Omega$, $a \in \Omega$, S1 $(\{a\}, \{a\}) \in Y$;

S2 $(A, B) \in Y, \quad (B, C) \in Y, \quad B \neq \emptyset \rightarrow (A, C) \in Y$;

S3 $(A, B) \in Y, \quad C \subseteq A, \quad D \subseteq B \rightarrow (C, D) \in Y$;

S4 $(A, B) \in Y, \quad (C, D) \in Y \rightarrow (A \cup C, B \cap D) \in Y$;

S5 $(A, B) \in Y, \quad (C, D) \in Y \rightarrow (A \cap C, B \cup D) \in Y$.

Definition 1.3. Let $I \subseteq P(\Omega)$. We say that I is a $\cap$-semilattice over Ω if $\Omega \in I$, and $A, B \in I \rightarrow A \cap B \in I$. Let $M \subseteq P(\Omega)$. Denote by M^+ the set $\{\cap M' \colon M' \subseteq M\}$. Then we say M generates I if $M^+ = I$.

J. Demetrovics in [3] showed that for a given $\cap$-semilattice I, there is exactly one family N which generates I and has minimal cardinality.

Lemma 1.4. ([3]). Let $I \subseteq P(\Omega)$ be a $\cap$-semilattice over Ω. Let

$$N = \{A \in I \colon \forall B, C \in I \colon A = B \cap C \rightarrow A = B \text{ or } A = C\}.$$

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 $\Omega \in N$.)

It can be seen that if $N_1(N_2)$ is the minimal generator of $I_1(I_2)$ and $I_1 \neq I_2$, then $N_1 \neq N_2$ holds.

2. The results

Definition 2.1. Let $I \subseteq P(\Omega)$. We say that I is an s-semilattice over Ω if I satisfies

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

(2) for all $A \in N \setminus \Omega$

$$(\exists a \in A)((\forall B \in N \setminus \Omega)(A \not\subset B) \rightarrow a \notin B),$$

where N is the minimal generator of I.

Definition 2.2. The mapping $F \colon P(\Omega) \rightarrow P(\Omega)$ is called a strong operation over Ω if for every $a, b \in \Omega$ and $A \in P(\Omega)$, the following properties hold:

(1) $F(\emptyset) = \Omega$,

(2) $a \in F(\{a\})$,

(3) $b \in F(\{a\}) \rightarrow F(\{b\}) \subseteq F(\{a\})$,

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

It is easy to see that the set $\{F(\{a\}) \colon a \in \Omega\}$ determines the set $\{F(A) \colon A \in P(\Omega)\}$.

The following theorem shows that there is an one-to-one correspondence between s-semilattices and strong operations.

Theorem 2.3. Let F be a strong operation over Ω. Let $I_F=\{F(A)\colon A\in P(\Omega)\}$. Then I_F is an s-semilattice over Ω. Conversely, if I is an s-semilattice over Ω, then there is exactly one strong operation F so that $I_F=I$, where $F(\emptyset)=\Omega$, and for all $a\in\Omega$,

$$F(\{a\})=\begin{cases}\bigcap\limits_{\substack{a\in A_i\\ A_i\in N\setminus\Omega}} A_i & \text{if } \exists A_i\colon a\in A_i (N \text{ is the minimal generator of } I),\\ \Omega & \text{otherwise.}\end{cases}$$

Proof. It is clear that for arbitrary strong operation F

$$\forall A, B\in P(\Omega)\colon F(A\cup B) = F(A)\cap F(B), F(\emptyset) = \Omega$$

and $A\subseteq B\rightarrow F(B)\subseteq F(A)$. Consequently, $I_F=\{F(A)\colon A\in P(\Omega)\}$ is a $\cap$-semilattice over Ω. Denote by N_F the minimal generator of I_F. For all $A\in N_F\setminus\Omega$ if there is no attribute a such that $F(\{a\})=A$, then if $A=F(B)(|B|\geqq 2)$ holds, then, according to the definition of strong operation, $A=\bigcap\limits_{b_i\in B} F(\{b_i\})$. This contradicts the definition of minimal generator. Consequently, there is an attribute $a\in\Omega$ so that $F(\{a\})=A$. It is obvious that $a\in A$. It is clear that $A, B\in N_F$ implies $A\neq B$, and by (3) in the definition of strong operation, for all $A\in N_F\setminus\Omega$:

$$(\exists a\in A)\big((\forall B\in N_F\setminus\Omega)(A \not\subseteq B) \rightarrow a\notin B\big).$$

Consequently, I_F is an s-semilattice over Ω.

Conversely, we now suppose that I is an s-semilattice over Ω. Denote by N the minimal generator of I. We define the following operation F:

$$F(\emptyset) = \Omega,$$

and for all b,

$$F(\{b\})=\begin{cases}\bigcap\limits_{\substack{b\in A_i\\ A_i\in N\setminus\Omega}} A_i & \text{if } \exists A_i\colon b\in A_i,\\ \Omega & \text{otherwise.}\end{cases}$$

It can be seen that for all $A\in N\setminus\Omega$, where $\exists a\in A\colon A_i\in N\setminus\Omega$ and $A\not\subseteq A_i\rightarrow a\notin A_i$, we have $F(\{a\})=A$. For all different elements $A (A\in N\setminus\Omega)$ it is easy to see that there is an $a\in A$ so that $F(\{a\})=A$. Consequently, $\forall A\in N\setminus\Omega\colon \exists a\in\Omega\colon F(\{a\})=A$. We now show that F is a strong operation over Ω. It can be seen that $b\in F(\{b\})$, and if there is an $A_i\in N\setminus\Omega$ such that $b\in A_i$, then $F(\{b\})\in N^+$. If $a\in F(\{b\})$ holds, then

$$F(\{a\})=\bigcap\limits_{\substack{a\in A_i\\ A_i\in N\setminus\Omega}} A_i\subseteq \bigcap\limits_{\substack{b\in A_i\\ A_i\in N\setminus\Omega}} A_i = F(\{b\}).$$

On the other hand, it can be seen that the set $\{F(\{b\})\colon b\in\Omega\}$ determines the set $\{F(A)\colon A\in P(\Omega)\}$. Consequently, F is a strong operation over Ω. It is easy to see that $I=\{F(A)\colon A\in P(\Omega)\}$. If we suppose that there is a strong operation F' such that $I_{F'}=I$ then for all $a\in\Omega$ $\exists b\in\Omega\colon F(\{a\})=F'(\{b\})$. It obvious that $a\in F'(\{b\})$. Consequently, $F'(\{a\})\subseteq F(\{a\})$. On the other hand, there is an attribute c so that $F'(\{a\})=F(\{c\})$. Clearly, $F(\{a\})\subseteq F'(\{a\})$ by $a\in F(\{c\})$. Consequently, for all $A\in P(\Omega)$, $F'(A)=F(A)$. The proof is complete.

Based on Theorem 2.3, it is easy to see that s-semilattices determine the strong operations, and for arbitrary s-semilattice I over Ω, $|N|$ in not greater than $|\Omega|+1$. Clearly, there is an algorithm to decide for a given family of sets $N \subseteqq P(\Omega)$ whether N is the minimal generator of some s-semilattice or not. The following theorem gives necessary and sufficient conditions for an arbitrary family of sets to be a full family of strong dependencies over Ω.

Theorem 2.4. Let $Y \subseteqq P(\Omega) \times P(\Omega)$. Y is a full family of strong dependencies over Ω if and only if there is a family $\{E_i\colon i=1, \ldots, l;\ \bigcup_{i=1}^{l} E_i = \Omega\}$ of subsets of Ω such that

(i) for all $A \subseteqq \Omega$, $(\emptyset, A) \in Y$,

(ii) for any $A, B \subseteqq \bigcup_{E_i \cap A \neq \emptyset} E_i \rightarrow (A, B) \in Y$,

(iii) $\big((C, D) \in Y,\ C \cap E_i \neq \emptyset\big) \rightarrow D \subseteqq E_i$.

Proof. First we suppose that Y is a full family of strong dependencies over Ω. Then by (S1), (S3), (S5) for each $a \in \Omega$ we can construct an E_i $(E_i \subseteqq \Omega)$ so that $(\{a\}, E_i) \in Y$, and $\forall E'\colon E \subset E'$ implies $(\{a\}, E') \notin Y$. It is obvious that $a \in E_i$, and we obtain n such E_i's, where $n = |\Omega|$. Thus, we have the set $E = \{E_i\colon i=1, \ldots, n;\ \bigcup_{i=1}^{n} E_i = \Omega\}$. It is easy to see that for all $A \subseteqq \Omega$ we have $(\emptyset, A) \in Y$. We now assume that $A = \{a_1, \ldots, a_k\colon a_j \in \Omega, j=1, \ldots, k\} \neq \emptyset$ and B_1 is a set such that $(A, B_1) \in Y$, $\forall B_2\colon B_1 \subset B_2$ implies $(A, B_2) \notin Y$. According to the construction of E, it is clear that for each a_j there is an $E_{i_j} \in E$ so that $(\{a_j\}, E_{i_j}) \in Y$. By (S4) we have $\big(\bigcup_{j=1}^{k} a_j, \bigcap_{j=1}^{k} E_{i_j}\big) = \big(A, \bigcap_{j=1}^{k} E_{i_j}\big) \in Y$. By the definition of B_1 we obtain $\bigcap_{j=1}^{k} E_{i_j} \subseteqq B_1$. On the other hand, by $(A, B_1) \in Y$ and by (S3), we have $(\{a_j\}, B_1) \in Y$ for all j $(j=i, \ldots, k)$. Consequently, $B_1 \subseteqq \bigcap_{j=1}^{k} E_{i_j}$ holds, i.e. $B_1 = \bigcap_{j=1}^{k} E_{i_j}$. It is obvious that $\bigcap_{E_i \cap A \neq \emptyset} E_i \subseteqq \bigcap_{j=1}^{k} E_{i_j}$. Thus, for all $B \big(B \subseteqq \bigcap_{E_i \cap A \neq \emptyset} E_i\big)$: $B \subseteqq B_1$. Hence $(A, B) \in Y$ holds. If $(C, D) \in Y$, $C \cap E_i \neq \emptyset$, then we assume that $a_1 \in C \cap E_i$. On the other hand, suppose that a is an attribute such that $(\{a\}, E_i) \in Y$, and $\forall E'\colon E_i \subset E'$ implies $(\{a\}, E') \notin Y$. By $a_1 \in E_i$ and by (S3), $(\{a\}, \{a_1\}) \in Y$ holds. By (S3) and $a_1 \in C$ we obtain $(\{a_1\}, D) \in Y$. Consequently, by (S2), $(\{a\}, D) \in Y$ holds. According to the definition of E we have $D \subseteqq E_i$.

The proof of the reverse direction is easy, and so it is omitted. The proof is complete.

Based on Theorem 2.4 we prove the following result, which shows that between full families of strong dependencies and strong operations there exists a one-to-one correspondence.

Theorem 2.5. Let Y be a full family of strong dependencies over Ω. We define the mapping $F_Y\colon P(\Omega) \rightarrow P(\Omega)$ as follows:

$F_Y(A)=\{a\in\Omega\colon (A, \{a\})\in Y\}.$

Then F_Y is a strong operation over Ω. Conversely, if F is an arbitrary strong operation over Ω, then there is exactly one full family of strong dependencies Y so that $F_Y=F$, where

$$Y = \{(A, B)\colon A, B\in P(\Omega)\colon B \subseteq F(A)\}.$$

Proof. Suppose that Y is a full family of strong dependencies over Ω. It is obvious that $\forall a\in\Omega\colon a\in F_Y(\{a\})$. By Theorem 2.4 we have $(C, D)\in Y$, $C\cap E_i\neq\emptyset$ imply $D\subseteq E_i$. It can be seen that in Theorem 2.4, for any $a\in\Omega$,

$$F_Y(\{a\})\in\{E_i\colon i = 1, \ldots, n;\ |\Omega| = n;\ \bigcup_{i=1}^{n} E_i = \Omega\}.$$

Consequently, $(\{b\}, F_Y(\{b\}))\in Y, b\in F_Y(\{a\})$, i.e. $b\cap F_Y(\{a\})\neq\emptyset$ implies $F_Y(\{b\})\subseteq F_Y(\{a\})$. By (iii) in Theorem 2.4 we obtain $(A, F_Y(A))\in Y$, $\forall a\in A\colon A\cap F_Y(\{a\})\neq\emptyset$ imply $F_Y(A)\subseteq F_Y(\{a\})$. Thus, $F_Y(A)\subseteq \bigcap_{a\in A} F_Y(\{a\})$. On the other hand, by (S5) in the definition of full family of strong dependencies we have $\forall a\in A\colon (\{a\}, F_Y(\{a\}))\in Y$ implies $(A, \bigcap_{a\in A} F_Y(\{a\}))\in Y$, i.e. $\bigcap_{a\in A} F_Y(\{a\})\subseteq F_Y(A)$. Consequently, $F_Y(A)= \bigcap_{a\in A} F_Y(\{a\})$ holds. Conversely, assume that F is a strong operation over Ω, and $Y=\{(A, B)\colon B\subseteq F(A)\}$. We have to show that Y is a full family of strong dependencies. By Theorem 2.4 we set $E=\{F(\{a\})\colon a\in\Omega,\ |\Omega|= \quad \}$. By the definition of Y and by $\bigcap_{F(\{a\})\cap A\neq\emptyset} F(\{a\}\subseteq F(A)$, it is obvious that $B\subseteq \bigcap_{F(\{a\})\cap A\neq\emptyset} F(\{a\})$ implies $(A, B)\in Y$. On the other hand, if $(C, D)\in Y$ and $C\cap F(\{a\})\neq\emptyset$, then we assume that $b\in C\cap F(\{a\})$, hence by (iii) in the definition of strong operation $b\in F(\{a\})$ implies $F(\{b\})\subseteq F(\{a\})$. It is obvious that $D\subseteq F(C)=\bigcap_{d\in c} F(\{d\})$. By $b\in C$, and $\bigcap_{d\in C} F(\{d\})\subseteq F(\{b\})$ we obtain $D\subseteq F(\{a\})$. It is clear that $\forall A\subseteq\Omega\colon (\emptyset, A), (A, \emptyset)\in Y$. It can be seen that $F=F_Y$. Now, we suppose that there is a full family of strong dependencies Y' so that $F_{Y'}=F$. By the definition of Y and F we obtain $Y'\subseteq Y$. If $(A, B)\in Y$ holds, then $B\subseteq F(A)= F_{Y'}(A)$ holds. By the definition of F_Y we have $(A, B)\in Y'$. Consequently, $Y'=Y$ holds. The proof is complete.

Remark 2.6. Clearly, if F_1 and F_2 are strong operations over Ω $(F_1\neq F_2)$, then $Y_1\neq Y_2$, where $Y_i=\{(A, B)\colon B\subseteq F_i(A)\}, i=1, 2.$

Based on Theorem 2.3 and Theorem 2.5 the next corollary is obvious.

Corollary 2.7. Let Y be a full family of strong dependencies over Ω. We define the mapping $F\colon P(\Omega)\to P(\Omega)$ as follows:

$$F_Y(A)\colon \{a\in\Omega\colon (A, \{a\})\in Y\}.$$

Let $I_Y=\{F_Y(A)\colon A\in P(\Omega)\}$. Then I_Y is an s-semilattice over Ω. Conversely, if I is an arbitrary s-semilattice over Ω, then there is exactly one full family of strong dependencies Y such that $I_Y=I$, where

$$Y = \{(A,B):\ A,B\in P(\Omega), A\neq\emptyset, \exists A_i\in N\setminus\Omega:\ A_i\cap A\neq\emptyset, B\subseteq \bigcap_{\substack{A_i\cap A\neq\emptyset\\ A_i\in N\setminus\Omega}},\ N \text{ is the minimalgenerator of } I\}\cup$$

$$\cup\{(A,B):\ A=\emptyset \text{ or } \overline{\exists} A_i\in N\setminus\Omega:\ A_i\cap A\neq\emptyset,\ B\in P(\Omega)\}.$$

Corollary 2.7 shows that between full families of strong dependencies and s-semilattices there is a one-to-one correspondence and the s-semilattices determine the full families of strong dependencies.

It is proved (see [2], [3], [4]) that if Y is a full family of strong dependencies over Ω, then there exists a relation R over Ω so that $S_R=Y$.

With the aid of the concept of s-semilattice we can construct for a given full family of strong dependencies Y a simple concrete relation R such that $S_R=Y$.

The equality sets of the relation are defined in [4] as follows:

Definition 2.8. Let $R=\{h_1, \dots, h_m\}$ be a relation over Ω. For $1\leq i<j\leq m$ denote by E_{ij} the set $\{a\in\Omega:\ h_i(a)=h_j(a)\}$.

Definition 2.9. Let Y be a full family of strong dependencies over Ω. We say that a relation R represents Y iff $S_R=Y$.

We now prove the following theorem which gives a necessary and sufficient condition for a relation to represent a given full family of strong dependencies.

Theorem 2.10. Let Y be a full family of strong dependencies, and $R=\{h_1, \dots, h_m\}$ be a relation over Ω. Then R represents Y iff for each $a\in\Omega$,

$$F_Y(\{a\}) = \begin{cases} \bigcap_{a\in E_{ij}} E_{ij} & \text{if } \exists E_{ij}:\ a\in E_{ij}, \\ \Omega & \text{otherwise,} \end{cases} \tag{1}$$

where $F_Y(A)=\{a\in\Omega:\ (A,\{a\})\in Y\}$, and E_{ij} is the equality set of R.

Proof. By Theorem 2.5, $S_R=Y$ if and only if $F_{S_R}=F$ holds. Consequently, first we show that $F_{S_R}(\{a\})=\bigcap_{a\in E_{ij}} E_{ij}$ if $\exists E_{ij};\ a\in E_{ij}$, and in other case $F_{S_R}(\{a\})=\Omega$ holds. Clearly, $F_{S_R}(\{a\})=\{b\in\Omega:\ \{a\}\xrightarrow[R]{s}\{b\}\}$. According to the definition of strong dependency we know that for any $a\in\Omega$,

$$\{a\}\xrightarrow[R]{s} B \leftrightarrow \{a\}\xrightarrow[R]{f} B,$$

where $a\neq\emptyset$, and $\{a\}\xrightarrow[R]{f} B$ denotes that B functionally depends on $\{a\}$ in R, i.e.

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

(see [4]). Let us denote by T the set $\{E_{ij}:\ a\in E_{ij}\}$. It is obvious that if $T=\emptyset$, then $\{a\}\xrightarrow[R]{f}\Omega$, i.e. $F_{S_R}(\{a\})=\Omega$ holds. If $T\neq\emptyset$ holds, then we set $A=\bigcap_{a\in E_{ij}} E_{ij}$.

If $T=E$ (E is the set of all equality sets of R), then it is obvious that $\{a\}\xrightarrow[R]{f} A$. If $T\subset E$ then for E_{ij}: $E_{ij}\notin T$ we obtain $h_i(a)\neq h_j(a)$. Consequently, we have also $\{a\}\xrightarrow[R]{f} A$. Denote by A' the set with the following properties:

(i) $\{a\}\xrightarrow[R]{f} A'$,

(ii) $A'\subset A''$ implies $\{a\}\xrightarrow[R]{f} A''$, i.e. A'' does not functionally depend on $\{a\}$.

It can be seen that $A'=A$. According to the definition of F_{S_R}, we obtain $F_{S_R}(\{a\})=$ $=\bigcap_{a\in E_{ij}} E_{ij}$. Thus, if $S_R=Y$ then we have (1). Conversely, if F_Y satisfies (1), then according to the above considerations, for any $a\in\Omega$ we have $F_Y(\{a\})=F_{S_R}(\{a\})$. Because F_Y and F_{S_R} are strong operations over Ω, and by Theorem 2.5 we obtain $\forall A\subseteq\Omega$: $F_{S_R}(A)=F_{S_Y}(A)$. Consequently, $F_Y=F_{S_R}$ holds. The proof is complete.

Definition 2.11. Let R be a relation, and F a strong operation over Ω. We say that the relation R represents F iff $F_{S_R}=F$.

By Theorem 2.10 the next corollary is obvious.

Corollary 2.12. Let F be a strong operation and R a relation over Ω. Then R represents F iff for all $a\in\Omega$,

$$F(\{a\})=\begin{cases}\bigcap_{a\in E_{ij}} E_{ij} & \text{if } \exists E_{ij}\colon a\in E_{ij},\\ \Omega & \text{otherwise.}\end{cases}$$

Clearly, from a relation R we can construct the set of all equality sets of R. Consequently, the following corollary is also obvious.

Corollary 2.13. Let R be a relation and F a strong operation over Ω. Then there is an algorithm which decides whether R represents F or not. This algorithm requires time polynomial in the number of rows and columns of R.

Based on Theorem 2.10 the next proposition is straightforward and so its proof will be omitted.

Proposition 2.14. Let Y be a full family of strong dependencies over Ω. Denote N the minimal generator of s-semilattice I_Y.

Suppose that $N-\Omega=\{B_1, \dots, B_l\}$. We set $R=\{h_0, h_1, \dots, h_l\}$ as follows:

$$\text{for all } a\in\Omega\colon h_0(a)=0,$$

$$\text{for each } i\ (i=1, \dots, l),\ h_i(a)=\begin{cases}0 & \text{if } a\in B_i,\\ i & \text{otherwise.}\end{cases}$$

Then $S_R=Y$ holds.

Clearly, Proposition 2.14 shows that from a given s-semilattice I we can construct a simple concrete relation R such that $I=I_{S_R}$. Because between $\cap$-semilattices and minimal generators there is a one-to-one correspondence, it can be seen that (based on Theorem 2.3, Corollary 2.7, and Proposition 2.14) from the minimal generators of s-semilattices we can construct suitable full families of strong dependencies, strong operations, and relations.

Acknowledgements

The author would like to take this opportunity to express deep gratitude to Professor J. Demetrovics for his help, valuable comments and suggestions.

COMPUTER AND AUTOMATION INSTITUTE
HUNGARIAN ACADEMY OF SCIENCES
KENDE U. 13—17
BUDAPEST, HUNGARY
H—1502

References

[1] ARMSTRONG, W. W., Dependency structures of data base relationships, Information Processing 74, North-Holland (1974) 580—583.
[2] CZÉDLI, G., Függőségek relációs adatbázis modellben, Alkalmazott Matematikai Lapok 6 (1980) 131—143.
[3] DEMETROVICS, J., Relációs adatmodell logikai és strukturális vizsgálata, MTA—SZTAKI Tanulmányok, Budapest, 114 (1980) 1—97.
[4] DEMETROVICS J. and GY. GYEPESI, On the functional dependency and some generalizations of it, Acta Cybernetica, Tom. 5, Fasc. 3 (1981) 295—305.

(Received Okt. 27, 1986)