

ON FUNCTIONAL DEPENDENCIES AND MULTIVALUED DEPENDENCIES IN FUZZY RELATIONAL DATABASES

HO THUAN, TRAN THIEN THANH

Abstract. In this paper, we present a new definition of fuzzy functional dependency and fuzzy multivalued dependency based on similarity in fuzzy relational database, for which thresholds are defined for each attributes of relation scheme. The soundness and completeness of inference rules, similar to Armstrong's axioms are proved.

Tóm tắt. Trong bài báo này chúng tôi trình bày định nghĩa cho phụ thuộc hàm và phụ thuộc đa trị mờ trên mô hình cơ sở dữ liệu mờ dựa trên quan hệ tương tự với ngưỡng tương tự xác định riêng cho mỗi thuộc tính. Tính xác đáng và đầy đủ của hệ tiên đề tương tự hệ tiên đề Armstrong cũng được chứng minh.

1. INTRODUCTION

Relational databases have been studied since Codd's. Such databases can only deal with well-defined and unambiguous data. But in the real world there exist data which cannot be well-defined in a certain clear sense and under a certain crisp form (often called fuzzy data). The databases for the above mentioned data have been investigated by different authors (see [7]). The fuzzy database models are an extension of the classical relational model. It is based on the fuzzy set theory invented by Zadeh to capture the imprecise parts of the real world.

In general, two approaches have been proposed for the introduction of fuzziness in the relational model. The first one uses the principle of replacing the ordinary equivalence among domain values by measures of nearness such as *similarity relationships*, *proximity relationship*, and *distinguishability function* (see [8]). The second major effort has involved a variety of approaches that directly use *possibility distributions* for attribute value (see [5]). There have also been some mixed models combining these approaches [12].

This paper takes the similarity-based fuzzy relational databases as the reference model in our study presented here.

The data dependencies are the most important topics in theory of relational databases. Several authors have proposed extended dependencies in fuzzy relational database models. In [1, 2, 4, 6, 10, 12] have been given definitions of fuzzy functional dependencies and fuzzy multivalued dependencies in fuzzy relational data models. These dependencies are extension of dependencies of classical relational model. In this article, we give the definitions of fuzzy functional dependency (abbr. (α, β) -ffd) and fuzzy multivalued dependency (abbr. (α, β) -fmvd). These dependencies are extension of dependencies in classical model and more general than definitions of Raju, Mazumdar, etc. We also show that the inference rules of (α, β) -ffd, (α, β) -fmvd, which are similar to Armstrong's axioms for classical relational databases, are sound and complete.

This paper is organized as follows. Section 2 present some basic definitions of the similarity-based relational databases. In Section 3 and Section 4, we introduce an extension of functional and multivalued dependencies; Armstrong's axioms for fuzzy functional and multivalued dependencies are presented; the soundness and completeness are proved. Section 5 concludes this paper with some perspectives of the present work.

2. SIMILARITY-BASED FUZZY RELATIONAL DATABASES

The similarity-based fuzzy relational database model is a generalization of the original relational model. It is allowed an attribute value to be a subset of the associated domain. Domains for this model are either discrete scalars or discrete numbers drawn from either a finite or infinite set. The

equivalence relation over the domain is replaced by a fuzzy similarity relation to identify similar tuples exceeding a given threshold of similarity.

Definition 2.1. A similarity relation is a mapping $s : D \times D \rightarrow [0, 1]$ such that for $x, y, z \in D$,

$$\begin{aligned} s(x, x) &= 1 \text{ (reflexivity),} \\ s(x, y) &= s(y, x) \text{ (symmetry),} \\ s(x, z) &\geq \max_{y \in D} \{ \min[s(x, y), s(y, z)] \} \text{ (max-min transitivity).} \end{aligned}$$

Definition 2.2. A fuzzy relation scheme is a triple $\mathcal{S} = (R, \vec{s}, \vec{\alpha})$, where $R = \{A_1, A_2, \dots, A_n\}$ is a set of attributes, $\vec{s} = (s_1, s_2, \dots, s_n)$ is a set of associated similarity relations, $\vec{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_n)$ is a set of associated thresholds ($\alpha_i \in [0, 1], 1 \leq i \leq n$).

Definition 2.3. A fuzzy relation instance r on scheme $\mathcal{S} = (R, \vec{s}, \vec{\alpha})$ is a subset of the cross product $P(D_1) \times P(D_2) \times \dots \times P(D_n)$, where $D_i = \text{dom}(A_i)$, and $P(D_i) = 2^{D_i} - \emptyset$.

Let X, Y be sets of attributes in R , $X = (A_h)_{h \in I}, I \subseteq \{1, \dots, n\}$.

α_X denotes the vector of thresholds for a set of attributes X , i.e. $\alpha_X = (\alpha_h)_{h \in I}$.

α_{XY} denotes the vector of thresholds for a set of attributes $X \cup Y$ (XY for short).

In order to approximate equality between tuples of fuzzy relation, a fuzzy measure, a similarity relation τ is defined as follows.

Definition 2.4. Let r be a fuzzy relation instance on scheme $\mathcal{S} = (R, \vec{s}, \vec{\alpha})$, t_1 and t_2 are two tuples in r . The similarity measure of two tuples t_1 and t_2 on attribute A_k in R denoted by $\tau(t_1[A_k], t_2[A_k])$ is given as

$$\tau(t_1[A_k], t_2[A_k]) = \min_{x \in d_k^1, y \in d_k^2} \{s_k(x, y)\},$$

where $t_1 = (d_1^1, d_2^1, \dots, d_n^1)$, $t_2 = (d_1^2, d_2^2, \dots, d_n^2)$.

If $\tau(t_1[A_k], t_2[A_k]) \geq \alpha_k$ then tuples t_1 and t_2 are said to be similar on A_k with threshold α_k .

Definition 2.5. Let r be a fuzzy relation on scheme $\mathcal{S} = (R, \vec{s}, \vec{\alpha})$, X be a subset of R , t_1 and t_2 are two tuples in r . The similarity measure of two tuples t_1 and t_2 on a set of attributes X denoted by $\tau(t_1[X], t_2[X])$ is given as

$$\tau(t_1[X], t_2[X]) = \left(\tau(t_1[A_{i_1}], t_2[A_{i_1}]), \tau(t_1[A_{i_2}], t_2[A_{i_2}]), \dots, \tau(t_1[A_{i_k}], t_2[A_{i_k}]) \right),$$

where $X = A_{i_1} A_{i_2} \dots A_{i_k}$.

If $\tau(t_1[X], t_2[X]) \geq \alpha_X$ then two tuples t_1 and t_2 are said to be similar on X with thresholds α_X .

3. FUZZY FUNCTIONAL DEPENDENCY (α, β) -FFD

Definition 3.1. Let r be any fuzzy relation instance on scheme $\mathcal{S} = (R, \vec{s}, \vec{\alpha})$, X and Y are subsets of R with associated thresholds α_X, α_Y , respectively. Fuzzy relation instance r is said to satisfy the fuzzy functional dependency (α, β) -ffd, denoted by $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ if, for every pair of tuples t_1 and t_2 in

r , $\tau(t_1[X], t_2[X]) \geq \alpha_X$ then $\tau(t_1[Y], t_2[Y]) \geq \alpha_Y$.

Definition 3.2. A scheme $\mathcal{S} = (R, \vec{s}, \vec{\alpha})$ is said to satisfy the (α, β) -ffd $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$, if every relation instance r on \mathcal{S} satisfies $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$.

Remark 1. The definition ffd of Raju *et al.* is a special case of Definition 3.1. (i.e., if any instance r that satisfies ffd $X \overset{\alpha_0}{\rightsquigarrow} Y$ then r also satisfies (α, β) -ffd $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$), where $\alpha_X = \underbrace{(\alpha_0, \alpha_0, \dots, \alpha_0)}_{|X| \text{ times}}$ and $\alpha_Y = \underbrace{(\alpha_0, \alpha_0, \dots, \alpha_0)}_{|Y| \text{ times}}$ with α_0 is a constant in $[0, 1]$.

The inference rules for (α, β) -ffds

FFD1 (*Reflexivity*): If $Y \subseteq X$ then $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$

FFD2 (*Augmentation*): If $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ then $XW \overset{(\alpha_{XW}, \alpha_{YW})}{\rightsquigarrow} YW$

FFD3 (*Transitivity*): If $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$, and $Y \overset{(\alpha_Y, \alpha_Z)}{\rightsquigarrow} Z$ then $X \overset{(\alpha_X, \alpha_Z)}{\rightsquigarrow} Z$

Theorem 3.1. *Rules FFD1-FFD3 are sound.*

Proof. Let r be a relation instance on scheme $\mathcal{S} = (R, \vec{s}, \vec{\alpha})$.

Reflexivity: Suppose that $Y \subseteq X \subseteq R$.

$\forall t_1, t_2 \in r, \tau(t_1[X], t_2[X]) \geq \alpha_X$. Since $X \subseteq Y$, then $\tau(t_1[Y], t_2[Y]) \geq \alpha_Y$. Thus $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$

holds in r .

Augmentation: Suppose that $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ holds in $r, Z \subseteq R$.

$\forall t_1, t_2 \in r, \tau(t_1[XZ], t_2[XZ]) \geq \alpha_{XZ}$.

We have $\tau(t_1[X], t_2[X]) \geq \alpha_X$ and $\tau(t_1[Z], t_2[Z]) \geq \alpha_Z$. (1)

Since $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ holds in r then $\tau(t_1[Y], t_2[Y]) \geq \alpha_Y$. (2)

Combining (1) with (2), we obtain $\tau(t_1[YZ], t_2[YZ]) \geq \alpha_{YZ}$.

Hence $XZ \overset{(\alpha_{XZ}, \alpha_{YZ})}{\rightsquigarrow} YZ$ holds in r .

Transitivity: Suppose that $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$, and $Y \overset{(\alpha_Y, \alpha_Z)}{\rightsquigarrow} Z$ hold in r .

$\forall t_1, t_2 \in r, \tau(t_1[X], t_2[X]) \geq \alpha_X$.

Since $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ holds in r then $\tau(t_1[Y], t_2[Y]) \geq \alpha_Y$.

Since $Y \overset{(\alpha_Y, \alpha_Z)}{\rightsquigarrow} Z$ holds in r then $\tau(t_1[Z], t_2[Z]) \geq \alpha_Z$.

Therefore, $Y \overset{(\alpha_Y, \alpha_Z)}{\rightsquigarrow} Z$ holds in r .

The following rules are easily obtained from FFD1-FFD3.

FFD4 (*Union*): If $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ and $X \overset{(\alpha_X, \alpha_Z)}{\rightsquigarrow} Z$ then $X \overset{(\alpha_X, \alpha_{YZ})}{\rightsquigarrow} YZ$

FFD5 (*Pseudo-transitivity*): If $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ and $YW \overset{(\alpha_{YW}, \alpha_Z)}{\rightsquigarrow} Z$ then $XW \overset{(\alpha_{XW}, \alpha_Z)}{\rightsquigarrow} Z$

FFD6 (*Decomposition*): If $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ and $Z \subseteq Y$ then $X \overset{(\alpha_X, \alpha_Z)}{\rightsquigarrow} Z$

Theorem 3.2. *Rules FFD1-FFD6 are complete on scheme $\mathcal{S} = (R, \vec{s}, \vec{\alpha})$ when the following condition holds:*

For each $A_i \in R$, there exists at least one pair of elements $a_i, b_i \in \text{dom}(A_i)$ such that $s_i(a_i, b_i) \leq \alpha_i$.

Proof. Let F be a set of (α, β) -ffds on scheme $\mathcal{S} = (R, \vec{s}, \vec{\alpha})$, and suppose that $f = X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ does not follow from F by the rules FFD1-FFD6.

Consider the relation instance r on scheme \mathcal{S} with two tuples as follows

Attributes of X^+				Other attributes			
a_1	a_2	\dots	a_k	a_{k+1}	\dots	a_n	
a_1	a_2	\dots	a_k	b_{k+1}	\dots	b_n	

It is easily shown that all (α, β) -ffds in F are satisfied by r , and f is not satisfied by r .

We conclude that whenever $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ does not follow from F by the rules FFD1-FFD6 then F does not logically imply $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$. That is, the rules FFD1-FFD6 are complete.

4. FUZZY MULTIVALUED DEPENDENCY (α, β) -FMVD

Definition 4.1: Let r be any fuzzy relation instance on scheme $\mathcal{S} = (R, \vec{s}, \vec{\alpha})$, X and Y are subsets of R , with associated thresholds α_X, α_Y , respectively. Relation r is said to satisfy the fuzzy multivalued dependency (α, β) -fmvd, denoted by $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ if, for every two tuples t_1, t_2 in $r, \tau(t_1[X], t_2[X]) \geq \alpha_X$ then there exists a tuple t_3 in r such that $\tau(t_1[X], t_3[X]) \geq \alpha_X, \tau(t_1[Y], t_3[Y]) \geq \alpha_Y$, and $\tau(t_2[Z], t_3[Z]) \geq \alpha_Z$, where $Z = R - XY$.

Definition 4.2. A scheme $\mathcal{S} = (R, \vec{s}, \vec{\alpha})$ is said to satisfy the (α, β) -fmvd $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ if every relation instance r on \mathcal{S} satisfies $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$.

Remark 2. The definition fmvd of Mazumdar et al. [1] is a special case of Definition 4.1 (i.e. if relation r holds fmvd $X \rightsquigarrow_{\alpha_0} Y$ then r also holds (α, β) -fmvd $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$, where $\alpha_X = \underbrace{(\alpha_0, \alpha_0, \dots, \alpha_0)}_{|X| \text{ times}}$ and $\alpha_Y = \underbrace{(\alpha_0, \alpha_0, \dots, \alpha_0)}_{|Y| \text{ times}}$).

By Definition 4.1 it is easy to show following remarks.

Remark 3. Relation r satisfies $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ if, for every two tuples t_1, t_2 in r , $\tau(t_1[X], t_2[X]) \geq \alpha_X$ then there exists a tuple t_3 in r such that $\tau(t_2[X], t_3[X]) \geq \alpha_X$, $\tau(t_2[Y], t_3[Y]) \geq \alpha_Y$, and $\tau(t_1[Z], t_3[Z]) \geq \alpha_Z$, where $Z = R - XY$.

Remark 4. Relation r satisfies $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ if, for every two tuples t_1, t_2 in r , $\tau(t_1[X], t_2[X]) \geq \alpha_X$ then there exists a tuple t_3 in r such that $\tau(t_1[XY], t_3[XY]) \geq \alpha_{XY}$, and $\tau(t_2[X(R - Y)], t_3[X(R - Y)]) \geq \alpha_{X(R - Y)}$.

The inference rules for (α, β) -fmvds.

FMVD1 (*Complementation*): If $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ then $X \overset{(\alpha_X, \alpha_Z)}{\rightsquigarrow} Z$, where $Z = R - XY$

FMVD2 (*Augmentation*): If $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$, $V \subseteq W$ then $XW \overset{(\alpha_{XW}, \alpha_{YV})}{\rightsquigarrow} YV$

FMVD3 (*Transitivity*): If $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ and $Y \overset{(\alpha_Y, \alpha_Z)}{\rightsquigarrow} Z$ then $X \overset{(\alpha_X, \alpha_{(Z - Y)})}{\rightsquigarrow} (Z - Y)$

Theorem 4.1. Rules FMVD1-FMVD3 are sound.

Proof. Let r be a relation instance on scheme $\mathcal{S} = (R, \vec{s}, \vec{\alpha})$.

Complementation: Suppose $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ holds in r .

$$\forall t_1, t_2 \in r, \tau(t_1[X], t_2[X]) \geq \alpha_X. \quad (1)$$

Since $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ then $\exists t_3 \in r$, such that

$$\tau(t_2[X], t_3[X]) \geq \alpha_X, \quad (2)$$

$$\tau(t_2[Y], t_3[Y]) \geq \alpha_Y, \quad (3)$$

$$\tau(t_1[Z], t_3[Z]) \geq \alpha_Z, \text{ where } Z = R - XY. \quad (4)$$

$$\text{Combining (1) with (2), we have } \tau(t_1[X], t_3[X]) \geq \alpha_X. \quad (5)$$

$$\text{Since } W = R - XZ \subseteq Y, \text{ and by (3) then } \tau(t_2[W], t_3[W]) \geq \alpha_W. \quad (6)$$

From (5) and (6), it follows that $X \overset{(\alpha_X, \alpha_Z)}{\rightsquigarrow} Z$ holds in r .

Augmentation: Suppose $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ holds in r and $V \subseteq W$.

$$\forall t_1, t_2 \in r, \tau(t_1[XW], t_2[XW]) \geq \alpha_{XW}.$$

$$\text{Clearly, } \tau(t_1[X], t_2[X]) \geq \alpha_X \quad (7)$$

$$\text{and } \tau(t_1[W], t_2[W]) \geq \alpha_W. \quad (8)$$

Since $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ holds in r , and by (7), we have $\exists t_3 \in r$, such that

$$\tau(t_1[X], t_3[X]) \geq \alpha_X, \quad (9)$$

$$\tau(t_1[Y], t_3[Y]) \geq \alpha_Y, \quad (10)$$

$$\tau(t_2[Z], t_3[Z]) \geq \alpha_Z. \quad (11)$$

From (8), (9), (10) and (11) we have

$$\tau(t_1[W], t_3[W]) \geq \alpha_W, \text{ hence } \tau(t_1[XW], t_3[XW]) \geq \alpha_{XW}. \quad (12)$$

Since $\tau(t_1[W], t_3[W]) \geq \alpha_W$ and $V \subseteq W$, then $\tau(t_1[V], t_3[V]) \geq \alpha_V$.

$$\text{By (10), we obtain } \tau(t_1[YV], t_3[YV]) \geq \alpha_{YV}. \quad (13)$$

$$\text{Since } U = R - XYW \subseteq Z \text{ then } \tau(t_2[U], t_3[U]) \geq \alpha_U. \quad (14)$$

From (12), (13) and (14), it follows that $XW \overset{(\alpha_{XW}, \alpha_{YV})}{\rightsquigarrow} YV$ holds in r .

Transitivity: Suppose $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$, and $Y \overset{(\alpha_Y, \alpha_Z)}{\rightsquigarrow} Z$ hold in r .

$$\forall t_1, t_2 \in r, \tau(t_1[X], t_2[X]) \geq \alpha_X. \quad (15)$$

Since $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ holds in r then $\exists t_3 \in r$ such that

$$\tau(t_1[XY], t_3[XY]) \geq \alpha_{XY}, \quad (16)$$

$$\tau(t_2[X(R-Y)], t_3[X(R-Y)]) \geq \alpha_{X(R-Y)}. \quad (17)$$

Since $Y \overset{(\alpha_Y, \alpha_Z)}{\rightsquigarrow} Z$ holds in r , then $\exists t_4 \in r$, such that

$$\tau(t_1[YZ], t_4[YZ]) \geq \alpha_{YZ}, \quad (18)$$

$$\tau(t_3[Y(R-Z)], t_4[Y(R-Z)]) \geq \alpha_{Y(R-Z)} \quad (19)$$

First, we show that $\tau(t_1[X], t_4[X]) \geq \alpha_X$.

By (18), we have $\tau(t_1[X \cap YZ], t_4[X \cap YZ]) \geq \alpha_{X \cap YZ}$. (20)

Since $X - YZ \subseteq R - Z$, and by (5), it follows that $\tau(t_3[X - YZ], t_4[X - YZ]) \geq \alpha_{X - YZ}$.

By (17) and $X - YZ \subseteq R - Z$ then $\tau(t_2[X - YZ], t_4[X - YZ]) \geq \alpha_{X - YZ}$.

Therefore, $\tau(t_1[X - YZ], t_4[X - YZ]) \geq \alpha_{X - YZ}$. (21)

Combining (20) with (21), we obtain $\tau(t_1[X], t_4[X]) \geq \alpha_X$.

Next, we show that $\tau(t_1[Z - Y], t_4[Z - Y]) \geq \alpha_{Z - Y}$.

By (18), it is easy to see that $\tau(t_1[Z - Y], t_4[Z - Y]) \geq \alpha_{Z - Y}$.

Final, we show that $\tau(t_2[W], t_4[W]) \geq \alpha_W$, where $W = R - X(Z - Y)$.

Since $R - XYZ \subseteq R - Z$, and $R - XYZ \subseteq R - Y$, by (18) and (19),

we obtain $\tau(t_2[R - XYZ], t_4[R - XYZ]) \geq \alpha_{R - XYZ}$. (22)

From (16), (18) and (19), we have $\tau(t_2[Y], t_4[Y]) \geq \alpha_Y$. (23)

Since $W \subseteq Y(R - XY)$, by (22) and (23), it follows that $\tau(t_2[W], t_4[W]) \geq \alpha_W$.

Consequently, $X \overset{(\alpha_X, \alpha_{Z-Y})}{\rightsquigarrow} Z - Y$ holds in r .

The following rules are easily to obtained

FMVD4 (*Reflexivity*): If $Y \subseteq X$ then $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$

FMVD5 (*Union*): If $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ and $X \overset{(\alpha_X, \alpha_Z)}{\rightsquigarrow} Z$ then $X \overset{(\alpha_X, \alpha_{YZ})}{\rightsquigarrow} YZ$.

FMVD6 (*Decomposition*): If $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ and $X \overset{(\alpha_X, \alpha_Z)}{\rightsquigarrow} Z$ then $X \overset{(\alpha_X, \alpha_{(Z-Y)})}{\rightsquigarrow} (Z - Y)$

and $X \overset{(\alpha_X, \alpha_{(Y-Z)})}{\rightsquigarrow} (Y - Z)$

FMVD7 (*Pseudo-transitivity*): If $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ and $YW \overset{(\alpha_{YW}, \alpha_Z)}{\rightsquigarrow} Z$ then

$XW \overset{(\alpha_{XW}, \alpha_{(Z-YW)})}{\rightsquigarrow} (Z - YW)$

Rules relate (α, β) -ffds and (α, β) -fmvds

FFD-FMVD1 (*Replication*): If $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$ then $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$

FFD-FMVD2 (*Coalescence*): If $X \overset{(\alpha_X, \alpha_Z)}{\rightsquigarrow} Z$, $Y \overset{(\alpha_Y, \alpha_{Z'})}{\rightsquigarrow} Z'$, where $Z' \subseteq Z$, $Y \cap Z = \emptyset$

then $X \overset{(\alpha_X, \alpha_{Z'})}{\rightsquigarrow} Z'$

FFD-FMVD3 (*Mixed pseudotransitivity*): If $X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$, $XY \overset{(\alpha_{XY}, \alpha_Z)}{\rightsquigarrow} Z$ then $X \overset{(\alpha_X, \alpha_{(Z-Y)})}{\rightsquigarrow} (Z - Y)$

Definition 4.3. Let F and G be sets of (α, β) -ffd and (α, β) -fmvd on relation scheme $S = (R, \vec{s}, \vec{\alpha})$. The closure of $F \cup G$, denoted by $(F, G)^+$, is the set of all (α, β) -ffds and (α, β) -fmvds that can be derived from $F \cup G$ by repeated application of rules in the set $\{\text{FFD1-FFD6, MVD1-FMVD7, FFD-FMVD1-FFD-FMVD3}\}$.

Theorem 4.2. Let $S = (R, \vec{s}, \vec{\alpha})$ be a scheme relation, X be a subset of R , then we can partition $R - X$ into sets of attributes Y_1, Y_2, \dots, Y_k , such that if $Z \subseteq R - X$, then $X \overset{(\alpha_X, \alpha_Z)}{\rightsquigarrow} Z$ holds if and only if Z is the union of some of the Y_i 's.

Proof. Similar to classical case, see the proof in [11].

We call the above sets Y_1, Y_2, \dots, Y_k constructed for X from a set of (α, β) -ffds and (α, β) -fmvds D the *dependency basis* for X (with respect to D).

Theorem 4.3. *Rules FFD1-FFD6, FMVD1-FMVD7, FFD-FMVD1-FFD-FNVD3 are complete on scheme $\mathcal{S} = (R, \vec{s}, \vec{\alpha})$ when the following condition holds:*

For each $A_i \in R$, there exists at least one pair of elements $a_i, b_i \in \text{dom}(A_i)$ such that $s_i(a_i, b_i) \leq \alpha_i$.

Proof. Suppose F, G are sets of (α, β) -ffds and (α, β) -fmvds on scheme \mathcal{S} , d is a (α, β) -ffd or (α, β) -fmvd with left side is X and d does not follow from F, G by axioms.

Let $\{Y_1, Y_2, \dots, Y_k\}$ be a dependency basics for X respect to $F \cup G$.

We set $X^* = \{A \in R \mid X \overset{(\alpha_X, \alpha_A)}{\rightsquigarrow} A \in (F, G)^+\}$.

Since $X \overset{(\alpha_X, \alpha_{X^*})}{\rightsquigarrow} X^* \in (F, G)^+$ then $X \overset{(\alpha_X, \alpha_{X^*})}{\rightsquigarrow} X^* \in (F, G)^+$. Hence $X \overset{(\alpha_X, \alpha_{R-X^*})}{\rightsquigarrow} R - X^* \in$

$(F, G)^+$. By Theorem 4.2, we have $R - X^* = \bigcup_{i=1}^m W_i$, where $W_i \in \{Y_1, Y_2, \dots, Y_k\}$.

We now construct the following relation r on scheme \mathcal{S}

$$\begin{array}{ccccccc} X^* & W_1 & W_2 & \dots & W_m & & \\ \hline (a_i)_{i \in I_0} & (a_i)_{i \in I_1} & (a_i)_{i \in I_2} & \dots & (a_i)_{i \in I_m} & & \\ (a_i)_{i \in I_0} & (b_i)_{i \in I_1} & (a_i)_{i \in I_2} & \dots & (a_i)_{i \in I_m} & & \\ (a_i)_{i \in I_0} & (a_i)_{i \in I_1} & (b_i)_{i \in I_2} & \dots & (a_i)_{i \in I_m} & & \\ (a_i)_{i \in I_0} & (b_i)_{i \in I_1} & (b_i)_{i \in I_2} & \dots & (a_i)_{i \in I_m} & & \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \\ (a_i)_{i \in I_0} & (b_i)_{i \in I_1} & (b_i)_{i \in I_2} & \dots & (b_i)_{i \in I_m} & & \end{array}$$

where $X^* = (A_i)_{i \in I_0}$, $W_j = (A_i)_{i \in I_j}$, $I_j \subseteq \{1, 2, \dots, m\}$, $j = 0, \dots, m$.

First we show that r satisfies F and G .

Suppose $U \overset{(\alpha_U, \alpha_V)}{\rightsquigarrow} V \in F$. We set $W = \bigcup_{i \in I} W_i$, where $I = \{i : W_i \cap U \neq \emptyset, 1 \leq i \leq n\}$.

Clearly $U \subseteq X^*W$ then $X^*W \overset{(\alpha_{X^*W}, \alpha_V)}{\rightsquigarrow} V \in (F, G)^+$.

By Theorem 4.2, we have $X \overset{(\alpha_X, \alpha_W)}{\rightsquigarrow} W \in (F, G)^+$. Since $X \overset{(\alpha_X, \alpha_{X^*})}{\rightsquigarrow} X^* \in (F, G)^+$, by union rule then $X \overset{(\alpha_X, \alpha_{X^*W})}{\rightsquigarrow} X^*W \in (F, G)^+$. From FFD-FMVD3, $X \overset{(\alpha_X, \alpha_{(V-X^*W)})}{\rightsquigarrow} (V - X^*W) \in (F, G)^+$, implies $V - X^*W \subseteq X^*$.

Let t_1 and t_2 be tuples of r such that $\tau(t_1[U], t_2[U]) \geq \alpha_U$. By construction of r , it follows that $t_1[U] = t_2[U]$, $t_1[W] = t_2[W]$, $t_1[X^*] = t_2[X^*]$ so $\tau(t_1[X^*W], t_2[X^*W]) \geq \alpha_{X^*W}$. Hence $\tau(t_1[V], t_2[V]) \geq \alpha_V$. Therefore, r satisfies (α, β) -ffd $U \overset{(\alpha_U, \alpha_V)}{\rightsquigarrow} V$.

Now assume that $U \overset{(\alpha_U, \alpha_V)}{\rightsquigarrow} V \in G$. Since $U \subseteq X^*W$, by FMVD2 then $X^*W \overset{(\alpha_{X^*W}, \alpha_V)}{\rightsquigarrow} V \in (F, G)^+$.

We have $X \overset{(\alpha_X, \alpha_{X^*W})}{\rightsquigarrow} X^*W \in (F, G)^+$, by FMVD3 then $X \overset{(\alpha_X, \alpha_{(V-X^*W)})}{\rightsquigarrow} (V - X^*W) \in (F, G)^+$.

By Theorem 4.2, $V - X^*W = \bigcup_{i \in I_1} W_i$, $I_1 \subseteq \{1, \dots, m\}$. For any pair tuples $t_1, t_2 \in r$,

$\tau(t_1[U], t_2[U]) \geq \alpha_U$, by construction of r , $t_1[U] = t_2[U]$. There exists a tuple t_3 which is defined by $t_3[V - X^*W] = t_1[V - X^*W]$, and $t_3[R - (V - X^*W)] = t_2[R - (V - X^*W)]$.

It is easy to see that t_3 is a tuple in r and $t_3[U] = t_1[U]$, $t_3[V] = t_1[V]$ (because $t_1[W] = t_3[W]$), $t_3[R - UV] = t_2[R - UV]$.

Hence $\tau(t_1[U], t_3[U]) \geq \alpha_U$, $\tau(t_1[V], t_3[V]) \geq \alpha_V$, $\tau(t_2[R - UV], t_3[R - UV]) \geq \alpha_{R-UV}$.

Therefore, r satisfies (α, β) -fmvd $U \overset{(\alpha_U, \alpha_V)}{\rightsquigarrow} V$.

We now show that d does not hold by r .

Suppose that $d = X \overset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$. Since $d \notin (F, G)^+$ then $Y \not\subseteq X^*$.

By construction of r , there exist two tuples $t_1, t_2 \in r$, such that $t_1[Y] \neq t_2[Y]$. Furthermore, we have $\tau(t_1[Y], t_2[Y]) \not\geq \alpha_Y$. But $t_1[X] = t_2[X]$ then $\tau(t_1[X], t_2[X]) \geq \alpha_X$. Hence, r does not hold $X \underset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y$.

Now assume that $d = X \underset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y \notin (F, G)^+$, and d holds on r .

By construction of r , it is easy to show that $Y = X' \cup W'$, where $X' \subseteq X^*$, $W' = \bigcup_{j \in J} W_j$,

$J \subseteq \{1, \dots, m\}$.

We have $X \underset{(\alpha_X, \alpha_{X'})}{\rightsquigarrow} X' \in (F, G)^+$, and by Theorem 4.2, $X \underset{(\alpha_X, \alpha_{W'})}{\rightsquigarrow} W' \in (F, G)^+$. From union rule,

we have $X \underset{(\alpha_X, \alpha_Y)}{\rightsquigarrow} Y \in (F, G)^+$, contradiction. Thus d does not hold on r . The proof is complete.

5. CONCLUSION

This paper deals with fuzzy data dependencies in fuzzy relational databases. We give the definitions of fuzzy functional and multivalued dependencies. Furthermore, we discuss the inference rules of these dependencies. The soundness and completeness are proved. A further study involving the definitions of fuzzy join dependency, normal forms for the fuzzy relational databases has been on going.

REFERENCES

- [1] Bhattacharjee T. K., Mazumdar A. K., Axiomatisation of fuzzy multivalued dependencies in a fuzzy relational data model, *Fuzzy Sets and Systems* **96** (1998) 343–352.
- [2] Buckles B. P., Petry F. E., Information-theoretic characterization of fuzzy relational databases, *IEEE Trans. Syst. Man Cybernet* **13** (1983) 74–77.
- [3] Chen G., Kerre E. E., Vandenbulcke J., A computational algorithm for the FFD transitive closure and a complete axiomatization of fuzzy functional dependence, *International Journal of Intelligent Systems* **9** (1994) 421–439.
- [4] Cubero J. C., Vila M. A., A new definition of fuzzy functional dependency in fuzzy relational databases, *International Journal of Intelligent Systems* **9** (1994) 441–448.
- [5] Dubois D., Prade H., *Possibility Theory : An Approach to Computerized Processing of Uncertainty*, Plenum Press, New York, 1988.
- [6] Jyothi S., Babu M. S., Multivalued dependencies in fuzzy relational databases and lossless join decomposition, *Fuzzy Sets and Systems* **88** (1997) 315–332.
- [7] Kraft D. H., Petry F. E., Fuzzy information systems: managing uncertainty in databases and information retrieval systems, *Fuzzy Sets and Systems* **90** (1997) 183–191.
- [8] Petry F., Bosc P., *Fuzzy Databases : Principles and Applications*, Kluwer, Norwell, MA, 1996.
- [9] S. Sheno, A. Melton, An extended version of the fuzzy relational database model, *Information Sciences* **52** (1990) 35–52.
- [10] S. Sheno, A. Melton, and L. T. Fan, Functional dependencies and normal forms in the fuzzy relational database model, *Information Sciences* **60** (1992) 1–28.
- [11] Ullman J. D., *Principles of Database system*, Comp. Science Press, 1980.
- [12] L. T. Vuong, H. Thuan, A relational databases extended by application of fuzzy set theory and linguistic variables, *Computers and Artificial Intelligence* **8** (2) (1989) 153–168.
- [13] Yazici M. et al., The integrity constrains for similarity - based fuzzy relational databases, *International Journal of Intelligent Systems* **13** (1998) 641–659.

Received August 10, 2000

Ho Thuan - Institute of Information Technology.

Tran Thien Thanh - Pedagogical Institute of Quy Nhon.