

Fuzzy Subset Dependencies

Jyothi Singaraju

Assoc. Professor & Head

Dept. of Computer Science

Sri Padmavathi Mahila Visvavidyalayam (Women's University), Tirupati, India

ABSTRACT

The fuzzy relational databases which use the special types of constraints, called dependencies, as a semantic tool for expressing properties of the data like a classical relational databases. Fuzzy functional dependencies (ffd) and fuzzy multivalued dependencies (fmvd) are the most common types of dependencies. Similar to classical relational database, a full utilization of fuzzy multivalued dependencies in the design of fuzzy relational databases requires that fuzzy embedded multivalued dependencies (fEmvd) and fuzzy subset dependencies (fsd). These multivalued dependencies hold in a projection of a relation but not necessarily in the relation itself. The main object of this paper is to extend the concept of Embedded Multivalued Dependency (EMVD) and Subset Dependency (SD) to fuzzy relational database.

Key Words: Fuzzy relational databases, Fuzzy functional dependencies, Fuzzy multivalued dependencies, Fuzzy embedded multivalued dependencies, Fuzzy subset dependencies.

1. INTRODUCTION

The fuzzy relational databases [7,8,13,14,16] which use the special types of constraints, called dependencies, as a semantic tool for expressing properties of the data like a classical relational databases [2,3,15]. Fuzzy functional dependencies [13,14] and fuzzy multivalued dependencies [7,8] are the most common types of dependencies. Similar to classical relational database [15], a full utilization of fuzzy multivalued dependencies in the design of fuzzy relational databases requires that fuzzy embedded multivalued dependencies (fEmvds) and fuzzy subset dependencies. These multivalued dependencies hold in a projection of a relation but not necessarily in the relation itself.

The main object of this paper is to extend the concept of Embedded Multivalued Dependency (EMVD) and Subset Dependency (SD) to fuzzy relational database.

This paper is organized in the following manner. In section 2, the basic definitions and concepts related to the classical relational databases and fuzzy relational databases are presented. Fuzzy integrity constraints are discussed in section 3. An Embedded fuzzy multivalued dependency is defined in section 4. Section 5 defines fuzzy subset dependencies and proves the soundness and completeness of associated inference rules. Concluding remarks are given in section 6.

2. BASIC DEFINITIONS

In this section the basic terminology, notations, definitions and concepts related to the classical relational data model [2,3,15] are given and a few definitions and concepts from fuzzy relations are reviewed.

Classical Relational Data Model

The universe of a relational database denoted by R , is a finite set of elements A_1, A_2, \dots, A_n called attributes. The

domain of an attribute A_j , for $j = 1, 2, \dots, n$, is written as $\text{dom}(A_j)$. In this paper, the letters A, B, \dots are used for single attributes and the letters X, Y, \dots are used for set of attributes. The union of the two sets X and Y is written as XY .

A relation r on the set of attributes (A_1, A_2, \dots, A_n) is a subset of the Cartesian product $\text{dom}(A_1) \times \dots \times \text{dom}(A_n)$. The elements of the relation are called tuples or rows. Normally, r is used an instance of a relation scheme R . If t is a tuple and A is an attribute in relation R , then $t[A]$ is a tuple which corresponds to an attribute A .

Data dependencies are constraints imposed on data in a database. In addition to a set of attributes, a set of data dependencies is also an essential part of a relation scheme. The class of functional dependencies (fds) was the first type of data dependencies. Let r be an instance of a relation scheme R . Relation r satisfies fd: $X \rightarrow Y$, if for all t_1 and t_2 in r , $t_1[X] = t_2[X]$ implies $t_1[Y] = t_2[Y]$.

A multivalued dependency (mvd) is a constraint on the set R of attributes and is of the form $X \twoheadrightarrow Y$, where X and Y are subsets of R . Relation r satisfies mvd $X \twoheadrightarrow Y$, if for every pair of tuples in r , say t_1 and t_2 , there is another tuple t_3 in r such that $t_1[X] = t_2[X] = t_3[X]$ implies $t_1[Y] = t_3[Y]$ and $t_2[Z] = t_3[Z]$, where $Z = R - XY$.

An Embedded multivalued dependency (Emvd) is a constraint on R and is written as $X \twoheadrightarrow Y / Z$ where XYZ is a proper subset of R . Relation r satisfies Emvd $X \twoheadrightarrow Y / Z$, if mvd $X \twoheadrightarrow Y$ holds in the projection $r[XYZ]$.

Note that the notation $X \twoheadrightarrow Y / Z$ means $X \twoheadrightarrow Y$ and $X \twoheadrightarrow Z$ for XYZ in R . The logical equivalence for a mvd cannot be extended to any Emvd.

Inference axiom for Emvd:

EMD1 (Projection) : If $X \twoheadrightarrow Y / Z$, $Y' \subseteq Y$ and $Z' \subseteq Z$, then $X \twoheadrightarrow Y' / Z'$.

Subset dependencies were first introduced by Sagiv and Walecka [15]. Subset dependencies are generalizations of embedded multivalued dependencies.

A subset dependency (SD) is a constraint on the set R of attributes that contains XYZ and is of the form $Z(X) \rightarrow Z(Y)$ for all tuples t_1 and t_2 in r with $t_1[X] = t_2[X]$, there exists a tuple t_3 in r such that $t_1[Y] = t_3[Y]$ and $t_2[Z] = t_3[Z]$.

Generally the notation $Z(X) \subset Z(Y)$ is used for the subset dependency. For our convenience we denote the subset dependency as $Z(X) \rightarrow Z(Y)$.

Fuzzy Set Theory

A fuzzy set A in $U = \{ u_1, \dots, u_n \}$, $A \subseteq U$ will be written as

$$A = \{ \mu(u_1)/u_1, \mu(u_2)/u_2, \dots, \mu(u_n)/u_n \}$$

where

$$\mu_A(u) = U \rightarrow [0, 1].$$

The membership function $\mu_A(u) \in [0, 1]$ is the membership grade of u in A , with the grades 1 and 0 representing full membership and non-membership, respectively.

Let A and B be fuzzy subset of U . The basic operations performed on fuzzy sets are [1,4,5,6,9,10,17]

$$\mu_{A \cup B}(u) = \max(\mu_A(u), \mu_B(u)),$$

$$\mu_{A \cap B}(u) = \min(\mu_A(u), \mu_B(u)),$$

$$\mu_{\bar{A}}(u) = 1 - \mu_A(u).$$

The Cartesian product $A \times B$ is defined to be fuzzy subset of $U_1 \times U_2 = U$, where

$$\mu_{A \times B}(u) = \min(\mu_A(u), \mu_B(u)).$$

If U is the Cartesian product of n universes of discourse $U_1 \times U_2 \times \dots \times U_n$, then an n -ary fuzzy relation R in U is a fuzzy subset of $U_1 \times U_2 \times \dots \times U_n$ and is characterized by the n -variate membership function

$$\mu_r: U_1 \times U_2 \times \dots \times U_n \rightarrow [0, 1].$$

FUZZY RELATIONS

A fuzzy relation scheme R and a fuzzy relation r on a relation scheme $R(A_1 A_2 \dots A_n)$ are same as classical relational data model. As in classical relations, r will be represented as a table with an additional column for $\mu_r(t)$ denoting the membership value of the tuple t in r . For any tuple not present in the table we assume $\mu_r(t) = 0$. The tuples for which $\mu_r(t) > 0$ are only considered in this paper.

Example 2.1: Consider the relation scheme $COMPANY(N, E, S, P, Y)$, where N stands for Company name, E for Number of employees, S for Sales, P for Profit, and Y for Year. An instance r of a relation R is shown in Fig.2.1. In this case some of the domain values are crisp, where as $\text{dom}(E)$, $\text{dom}(S)$, and $\text{dom}(P)$ are fuzzy subsets.

Fig 2.1: An instance r of $COMPANY(N, E, S, P, Y)$

N	E	S	P	Y	μ
A	10	11000	1100	1985	0.37
A	11	10000	1600	1986	0.67
A	14	10000	1150	1987	0.50
A	12	13000	1800	1988	0.67
A	9	8000	1200	1989	0.50

Suppose that the universe of discourse U_e for the number of employees is in the range 0-30. U_p , the universe of discourse of profit is in the range of Rs. 500-3,000 and U_s , the universe of salary is in the range of Rs.5,000-30,000. The membership function μ_{SE} , μ_{HS} , and μ_{HP} of the fuzzy sets, small number of employees, High-sales and High-profit, are as given below:

$$\mu_{SE}(E) = \begin{cases} (1 + |e - 10| / 4)^{-1} & \text{for } e \geq 10 \\ 1 & \text{for } e < 10, \end{cases}$$

$$\mu_{HS}(S) = \begin{cases} (1 + |s - 12000| / 4000)^{-1} & \text{for } s \leq 12000 \\ 1 & \text{for } s > 12000, \end{cases}$$

$$\mu_{HP}(P) = \begin{cases} (1 + |p - 1600| / 400)^{-1} & \text{for } p \leq 1600 \\ 1 & \text{for } p > 1600. \end{cases}$$

In this example, $\mu_r(t)$ can be represented as the truth value for the fuzzy proposition, “ y has high-sales and high-profit”, for the tuple t . Thus the truth value of the fuzzy proposition “Company A has high-sales and high-profit” is 0.67.

3. FUZZY DATA CONSTRAINTS

To extend the classical relational data model to deal with fuzzy information it would be necessary to consider integrity constraints that may also involve compound integrity constraints.

The integrity constraints in relational database systems can be classified into two major types. One is domain dependency and another is data dependency. As an example of fuzzy domain dependency, consider a relation $STUDENTS(\text{Name}, \text{Age}, \text{Height}, \text{Course}, \text{Marks})$. An integrity constraint may be stated as “most of the DBMS students are young” or “most of

the A.I students are intelligent”. As an example of a fuzzy data dependency, consider a relation scheme $EMPLOYEE(\text{Name}, \text{Department}, \text{Job}, \text{Experience}, \text{salary})$, where an integrity constraint may be stated as “in any department employees having similar jobs and experience must have almost equal salaries”. This type of data dependency is called fuzzy functional dependency [13, 14]. As another example of fuzzy data dependency consider a relation scheme $COMPANY(\text{Name}, \text{Employees}, \text{Sales}, \text{Profit})$ for which an integrity constraint is “if any company has almost equal sales then one set of profits are almost equal and

another set of employees are almost equal". This type of compound data constraint can be viewed as fuzzy multivalued dependency[7, 8].

To deal with these data constraints the concept of particularization has been introduced by Zadeh[17, 18]. In order to evaluate the particularization of a fuzzy relation due to a compound fuzzy proposition, it is necessary to examine how the possibility distribution of a compound fuzzy proposition can be obtained from the possibility distributions of constituent atomic propositions. For this purpose, Zadeh[17,18] developed the translation rules. In this paper, implication rule of the translation rules are discussed. Suppose F and G are fuzzy subsets of the universe U and V , respectively, we associate possibility distribution Π_X and Π_Y

$$\Pi (X \rightarrow Y) = R_s , \quad (3.1)$$

where R_s is a fuzzy subset of $U \times V$ with membership function

$$\mu_{R_s}(u, v) = \begin{cases} 1 & \text{if } \mu_F(u) \leq \mu_G(v), \\ 0 & \text{otherwise.} \end{cases} \quad (3.2)$$

We consider the translation rule R_s with some modification[7, 8] to determine the possibility distribution associated with a compound conditional fuzzy proposition, "If X is F then (Y is G_1 and Z is H_2) or (Y is G_2 and Z is H_1) or (Y is G) or (Z is H)" as

$$\Pi (X \rightarrow Y) = R_s^* , \quad (3.3)$$

where R_s^* is a fuzzy subset of $U \times V \times W$ with membership function

$$\mu_{R_s^*}(u, v, w) = \begin{cases} 1 & \text{if } \mu_F(u) \leq \mu_{(G_1 \cap H_2) \cup (G_2 \cap H_1) \cup G \cup H}(v, w), \\ 0 & \text{otherwise.} \end{cases} \quad (3.4)$$

where \cap and \cup are "and" and "or" operators, respectively.

The other modified translation rule is defined as follows:

Let $F = F_1 \cup F_2 \cup F_3$, $G = G_1 \cup G_2 \cup G_3$, $H = H_1 \cup H_2 \cup H_3$ be fuzzy subsets in U , V , and W respectively. The possibility distribution $\Pi (Z (X) \rightarrow Z (Y))$ associated with compound fuzzy conditional proposition " if X is F_1 then Y is G_2 and Z is H_3 " is given by

$$\Pi (Z (X) \rightarrow Z (Y)) = R_s^{**} \quad (3.5)$$

where R_s^{**} is a fuzzy subset of $U \times V \times W$ with membership function

$$\mu_{R_s^{**}}(u, v, w) = \begin{cases} 1 & \text{if } \mu_{F_1}(u) \leq \mu_{G_2 \cap H_3}(v, w), \\ 0 & \text{otherwise.} \end{cases} \quad (3.6)$$

Next, the fuzzy measure EQUAL(EQ) is defined to compare two elements of a given domain[13, 14] in the fuzzy relations with membership function as

$$\mu_{EQ} : U \times U \rightarrow [0, 1], \quad (3.7)$$

where μ_{EQ} satisfies the following conditions. For all $a, b \in U$

$$\mu_{EQ}(a, a) = 1 \quad (\text{reflexivity}), \quad (3.8)$$

$$\mu_{EQ}(a, b) = \mu_{EQ}(b, a) \quad (\text{symmetry}). \quad (3.9)$$

That is EQUAL is a resemblance relation over U . It may be noted that EQUAL is not assumed to be transitive. The resemblance of EQUAL over a composite domain is defined by

$$\mu_{EQ}(t_1, t_2) = \min\{ \mu_{EQ}(t_1[A_1], t_2[A_1]), \dots, \mu_{EQ}(t_1[A_n], t_2[A_n]) \}, \quad (3.10)$$

where t_1, t_2 be two tuples in $D = \text{dom}(A_1) \times \dots \times \text{dom}(A_n)$.

In this paper, the extended fuzzy resemblance relation EQUAL [8] is used to compare a set of elements of a given domain. A fuzzy relation EQUAL(EQ) over a universe of discourse U is defined as a fuzzy subset $U \times U \times U$, having the membership function

$$\mu_{EQ}(a, b, c) = \min\{ \mu_{EQ}(a, b), \mu_{EQ}(b, c), \mu_{EQ}(c, a) \}, \quad (3.11)$$

where $a, b, c \in U$.

The membership function μ_{EQ} satisfies the reflexivity and symmetry properties and does not satisfy the transitivity. In terms of possibility theory, $\mu_{EQ}(a, b, c)$ can be represented as the possibility of a, b , and c to be 'equal'. Similar to (3.1), the fuzzy measure EQUAL of a set of elements over a composite domain is defined as

$$\mu_{EQ}(t_1, t_2) = \min\{ \mu_{EQ}(t_1[A_1], t_2[A_1], t_3[A_1]), \dots, \mu_{EQ}(t_1[A_n], t_2[A_n], t_3[A_n]) \}. \quad (3.12)$$

Example 3.1: Consider the relation scheme COMPANY(N, E, S, P) and its relation r discussed in example 2.1. The equality is defined over different domains as follows.

$$\begin{aligned} \mu_{EQ}(a, b) &= 0 & \text{for } a \neq b, a, b \in \{\text{dom}(N), \text{dom}(Y)\} \\ \mu_{EQ}(a, b) &= 1/(1 + \beta |a - b|) \end{aligned}$$

where

$$\begin{aligned} \beta &= 1 & \text{for } a, b \in \text{dom}(E), \\ &= 1/1000 & \text{for } a, b \in \text{dom}(S), \\ &= 1/100 & \text{for } a, b \in \text{dom}(P), \end{aligned}$$

All membership functions defined above are reflexive and symmetric. The fuzzy measure EQUAL can be used to represent the approximate equality of the domain values. The EQUAL relation is used to represent the fuzzy data dependencies.

with atomic propositions " X is F " and " Y is G ", respectively, where $\Pi_X = F$ and $\Pi_Y = G$. Mamdani[9] suggested methods for a conditional fuzzy propositions, "if X is F then Y is G ". This method of translation rule is called the implication rule of inference. Mizumoto et al.[10] suggested several improved methods for implication rules. The translation rules for conditional fuzzy propositions are based on the implication in standard sequence logic S_{α} (called R_s), or the Godelian implication rule using G_{α} logic(called R_g).

In this paper, the translation rule R_s is used to determine the possibility distribution associated with a conditional fuzzy proposition " If X is F then Y is G " given by

Definition 3.1: A fuzzy dependency (ffd) $f : X \rightsquigarrow Y$ with $X, Y \subseteq R$, holds in a fuzzy relation r on R , if for all tuples t_1 and t_2 of r (ie., $\mu_r(t_i) > 0$ for $i = 1, 2$), we have

$$\mu_{EQ}(t_1[X], t_2[X]) \leq \mu_{EQ}(t_1[Y], t_2[Y]). \quad (3.13)$$

Definition 3.2: A fuzzy multivalued dependency (fmvd) $f : X \rightsquigarrow\rightsquigarrow Y$ holds in a fuzzy relation r on R , if for all tuples t_1 and t_2 of r there exists another tuple t_3 in r such that

$$\mu_{EQ}(t_1[X], t_2[X], t_3[X]) \leq \max \{ \min(\mu_{EQ}(t_1[Y], t_3[Y]), \mu_{EQ}(t_2[Z], t_3[Z])), \min(\mu_{EQ}(t_2[Y], t_3[Y]), \mu_{EQ}(t_1[Z], t_3[Z])), \mu_{EQ}(t_1[Y], t_2[Y], t_3[Y]), \mu_{EQ}(t_1[Z], t_2[Z], t_3[Z]) \} \quad (3.14)$$

where $Z = R - XY$.

4. EMBEDDED FUZZY MULTIVALUED DEPENDENCY

In this section, the extension of the concept of embedded multivalued dependency(emvd) to fuzzy relational database.

Definition 4.1: An Embedded fuzzy multivalued dependency (Efmvd) $m : X \rightsquigarrow\rightsquigarrow Y$ with $X, Y, Z \subseteq R$, holds in a fuzzy relation r over a set of attributes R if the fmvd $X \rightsquigarrow\rightsquigarrow Y/Z$ holds in $r(XYZ)$ (ie., in the projection of r onto XYZ).

As in the case of classical database, the fmvd $X \rightsquigarrow\rightsquigarrow Y/Z$ and the Efmvd $X \rightsquigarrow\rightsquigarrow Y/Z$ are syntactically the same but semantically different. The fmvd $X \rightsquigarrow\rightsquigarrow Y/Z$ is defined only on relations over XYZ while the Efmvd $X \rightsquigarrow\rightsquigarrow Y/Z$ is defined on relation over any set of attribute that contains XYZ . However, the fmvd $X \rightsquigarrow\rightsquigarrow Y/Z$ and the Efmvd $X \rightsquigarrow\rightsquigarrow Y/Z$ express the same constraints on relations over XYZ .

Example 4.1: Consider the relation scheme COMPANY(N, E, S, P, Y), where N = Company name, E = Number of employees, S = Sales, P = Profit, Y = Year. An instance r of a relation R is shown in Fig.4.1. The resemblance relations of this relation is discussed in example 3.1. The fuzzy relation r of COMPANY satisfies the following Embedded fuzzy multivalued dependencies.

$$\begin{aligned} NS &\rightsquigarrow\rightsquigarrow P/E \\ NP &\rightsquigarrow\rightsquigarrow E/S \\ E &\rightsquigarrow\rightsquigarrow S/P \end{aligned}$$

Fig 4.1: An instance r of COMPANY(N, E, S, P, Y).

N	E	S	P	Y	μ
A	10	11000	1100	1985	0.37
A	11	10000	1600	1986	0.67
A	14	10000	1150	1987	0.50
A	12	13000	1800	1988	0.67
A	9	8000	1200	1989	0.50

The inference axiom of Efmvd is,

EFM1 (Projection): If $X \rightsquigarrow\rightsquigarrow Y/Z$, $Y' \subseteq Y$ and $Z' \subseteq Z$, then $X \rightsquigarrow\rightsquigarrow Y'/Z'$.

To prove the soundness of the axiom, consider an instance r of R , then $X \rightsquigarrow\rightsquigarrow Y/Z$ means

$$\mu_{EQ}(t_1[X], t_2[X], t_3[X]) \leq \max \{ \min(\mu_{EQ}(t_1[Y], t_3[Y]), \mu_{EQ}(t_2[Z], t_3[Z])), \min(\mu_{EQ}(t_2[Y], t_3[Y]), \mu_{EQ}(t_1[Z], t_3[Z])), \mu_{EQ}(t_1[Y], t_2[Y], t_3[Y]), \mu_{EQ}(t_1[Z], t_2[Z], t_3[Z]) \} \quad \text{---(1)}$$

Case 1: Let 1st part of the right hand side be the maximum. That is,

$$\mu_{EQ}(t_1[X], t_2[X], t_3[X]) \leq \min(\mu_{EQ}(t_1[Y], t_3[Y]), \mu_{EQ}(t_2[Z], t_3[Z]))$$

Since $Y' \subseteq Y$ and $Z' \subseteq Z$, by FF1, $Y \rightsquigarrow Y'$ and $Z \rightsquigarrow Z'$.

By definition 3.1, we have

$$\begin{aligned} \mu_{EQ}(t_1[Y], t_3[Y]) &\leq \mu_{EQ}(t_1[Y'], t_3[Y']) \\ \mu_{EQ}(t_1[Z], t_3[Z]) &\leq \mu_{EQ}(t_1[Z'], t_3[Z']) \end{aligned}$$

therefore,

$$\begin{aligned} \min(\mu_{EQ}(t_1[Y], t_3[Y]), \mu_{EQ}(t_2[Z], t_3[Z])) \\ \leq \min \{ \mu_{EQ}(t_1[Y'], t_3[Y']), \mu_{EQ}(t_1[Z'], t_3[Z']) \} \quad \text{--- (2)} \end{aligned}$$

From (1) and (2), we have

$$\mu_{EQ}(t_1[X], t_2[X], t_3[X]) \leq \min \{ \mu_{EQ}(t_1[Y'], t_3[Y']), \mu_{EQ}(t_2[Z'], t_3[Z']) \}$$

That is, $X \rightsquigarrow\rightsquigarrow Y'/Z'$ holds in r . Proof of other three cases is similar.

5. FUZZY SUBSET DEPENDENCY

An instance r of a relation scheme $R(A_1 A_2 \dots A_n)$ satisfies a classical Sd: $Z(X) \rightarrow Z(Y)$, for all tuples t_1 and t_2 in r , if $t_1[X] = t_2[X]$, then there exists another tuple t_3 in r such that $t_3[Y] = t_1[Y]$ and $t_3[Z] = t_2[Z]$, where X, Y , and Z are sets of attributes and both X and Y are disjoint from Z , and $t[X]$ is the projection of a tuple t over the attribute X .

The above statement prescribes that if the first set of tuples agree on X-column, then the second set of tuples must agree on Y-column and the third set of tuples must agree on Z-column. This is too restricted. In a fuzzy database, we define the fuzzy subset dependency(fsd) as “ if X is approximately equal on first part then Y is approximately equal on second part and Z is approximately equal on third part”.

Let $R(A_1 A_2 \dots A_n)$ be a relation scheme, and let $X = A_{i1}, A_{i2}, \dots A_{il}; Y = A_{j1}, A_{j2}, \dots A_{jp}$ and $Z = A_{k1}, A_{k2}, \dots A_{km}$ be subsets of $A_1, A_2, \dots A_n$. A fuzzy proposition ‘X is Equal’ defines a fuzzy subset of $\text{dom}(A_{i1}) \times \text{dom}(A_{i2}) \times \dots \text{dom}(A_{il})$ based on our interpretation of ‘equality’, with the membership function determined by (3.13). A generalization of a Sd: $Z(X) \rightarrow Z(Y)$ in R is called $\text{fsd}: Z(X) \rightsquigarrow Z(Y)$. The compound fuzzy conditional proposition, ‘if X is equal on first part then Y is equal on second part and Z is equal on third part’, can be translated into fuzzy relation using the translation rule R_s^{**} . The possibility distribution determined by (3.5) and (3.6) define the fuzzy subset dependency in the following way:

Definition 5.2: A fuzzy subset dependency (fsd): $Z(X) \rightsquigarrow Z(Y)$, where $X, Y, Z \subseteq R$ and both X and Y are disjoint from Z, holds in a fuzzy relation r, if for all tuples t_1 and t_2 of r there exists another tuple t_3 in r, such that

$$\mu_{EQ}(t_1[X], t_2[X]) \leq \min \{ \mu_{EQ}(t_1[Y], t_3[Y]), (\mu_{EQ}(t_2[Z], t_3[Z])) \} \quad (5.1)$$

The fuzzy subset dependency introduced here reduces to classical subset dependency, when the resemblance relation takes the binary values $\{0,1\}$ and $\mu_{EQ}(X_1, X_2) = 0$ for $X_1 \neq X_2$ where $X_j \in U_j$ for $j = 1, 2$. Therefore, if $\mu_{EQ}(t_1[X], t_2[X]) = 1$ (ie., $t_1[X] = t_2[X]$) by definition 5.2, $\min \{ \mu_{EQ}(t_1[Y], t_3[Y]), (\mu_{EQ}(t_2[Z], t_3[Z])) \}$ must be equal to 1, then $\mu_{EQ}(t_1[Y], t_3[Y]) = 1$ and $\mu_{EQ}(t_2[Z], t_3[Z]) = 1$. Hence $t_1[Y] = t_3[Y]$ and $t_2[Z] = t_3[Z]$.

Example 5.1: Let us consider the relation scheme COMPANY(N, D, E, S, P), where N = Company Name, D= Department, E = Number of Employees, S = Sales, P = Profit, and the resemblance relations discussed in example 3.1. The fuzzy relation r in Fig.5.1 satisfies the following fuzzy subset dependency.

$E(DS) \rightsquigarrow E(P)$
Fig 5.2: An instance r of COMPANY relation

N	D	E	S	P	μ
A	x	10	11000	1100	0.37
A	x	8	10000	1600	0.67
A	y	9	8000	1200	0.50

Note that the $\text{fsd}: E(DS) \rightsquigarrow E(P)$ does not permit the tuple $t = (A \ y \ 11 \ 7500 \ 1300)$ to be inserted in the database because r already contains the tuples $t_1 = (A \ y \ 9 \ 8000 \ 1200)$, $t_2 = (A \ x \ 10 \ 11000 \ 1100)$, and

$$\mu_{EQ}(t[DS], t_1[DS]) > \min \{ \mu_{EQ}(t[P], t_2[P]), (\mu_{EQ}(t_1[E], t_2[E])) \}$$

However, insertion of thesis tuple would not violate a classical Sd: $E(DS) \rightarrow E(P)$. The integrity constraint “ if in any department, sales are approximately equal on first part then profit must be approximately equal on second part and the number of employees must be approximately equal on third part”. By suitably selecting μ_{EQ} , the $\text{fsd}: E(DS) \rightsquigarrow E(P)$ provides a more acceptable model for such integrity constraints.

5.1 Inference Rules for Fuzzy Subset Dependencies

A set of sound and complete inference axioms are presented for fsds, which are similar to classical Sds []. Let $R(A_1 A_2 \dots A_n)$ be a relation scheme and S be a set of fsds. An instance r of R satisfies all fsds in S. In the following axioms, X, Y, Z, V and W are subsets of the relation scheme R.

FS1 (Reflexivity) : If $W \subseteq V$ and $Z \cap V = \emptyset$, then $Z(V) \rightsquigarrow Z(W)$, holds in r.

FS2 (Augmentation) : If $W \subseteq V$ and $Z \cap V = \emptyset$ and $Z(X) \rightsquigarrow Z(Y)$ holds in r, then
 $Z(XV) \rightsquigarrow Z(YW)$ holds in r.

FS3 (Transitivity): If $Z(X) \rightsquigarrow Z(Y)$ and $Z(Y) \rightsquigarrow Z(W)$ hold in r, then $Z(X) \rightsquigarrow Z(W)$ holds in r.

FS4 (Union): If $Z(X) \rightsquigarrow Z(Y)$ and $Z(X) \rightsquigarrow Z(W)$ hold in r, then $Z(X) \rightsquigarrow Z(YW)$ holds in r.

FS5 (Decomposition): If $Z(X) \rightsquigarrow Z(YW)$ holds in r, then $Z(X) \rightsquigarrow Z(Y)$ and $Z(X) \rightsquigarrow Z(W)$ hold in r.

FS6 (Pseudo-transitivity) : If $Z(X) \rightsquigarrow Z(Y)$ and $Z(YW) \rightsquigarrow Z(V)$ hold in r, then
 $Z(XW) \rightsquigarrow Z(V)$ holds in r.

To prove the soundness of these axioms, consider an instance r of R and let t_1, t_2, t_3 be three tuples in r.

FS1 (Reflexivity): Since $W \subseteq V$ and $Z \cap V = \emptyset$, we have $Z \cap W = \emptyset$. So both V and W are disjoint from Z . By the definition 5.1 $Z(V) \rightsquigarrow Z(W)$ trivially holds in r .

FS2 (Augmentation): Since $Z(X) \rightsquigarrow Z(Y)$ holds in r ,

$$\mu_{EQ}(t_1[X], t_2[X]) \leq \min \{ \mu_{EQ}(t_1[Y], t_3[Y]), \mu_{EQ}(t_2[Z], t_3[Z]) \} \quad \text{--- (1)}$$

and $W \subseteq V$ and $Z \cap V = \emptyset$, by reflexivity axiom FS1 we have $Z(V) \rightsquigarrow Z(W)$. That is,

$$\mu_{EQ}(t_1[V], t_2[V]) \leq \min \{ \mu_{EQ}(t_1[W], t_3[W]), \mu_{EQ}(t_2[Z], t_3[Z]) \} \quad \text{--- (2)}$$

from (1) and (2)

$$\begin{aligned} \min \{ \mu_{EQ}(t_1[X], t_2[X]), \mu_{EQ}(t_1[V], t_2[V]) \} &\leq \\ \min [\min \{ \mu_{EQ}(t_1[Y], t_3[Y]), \mu_{EQ}(t_2[Z], t_3[Z]) \}, & \\ \min \{ \mu_{EQ}(t_1[W], t_3[W]), \mu_{EQ}(t_2[Z], t_3[Z]) \}] & \end{aligned}$$

Hence,

$$\mu_{EQ}(t_1[XV], t_2[XV]) \leq \min \{ \mu_{EQ}(t_1[YW], t_3[YW]), \mu_{EQ}(t_2[Z], t_3[Z]) \}$$

That is, $Z(XV) \rightsquigarrow Z(YW)$ holds in r .

FS3 (Transitivity): Since $Z(X) \rightsquigarrow Z(Y)$ holds in the fuzzy relation r ,

$$\mu_{EQ}(t_1[X], t_2[X]) \leq \min \{ \mu_{EQ}(t_1[Y], t_3[Y]), \mu_{EQ}(t_2[Z], t_3[Z]) \} \quad \text{--- (1)}$$

We consider another tuple t_4 in r and $Z(X) \rightsquigarrow Z(Y)$ and $Z(Y) \rightsquigarrow Z(W)$ hold in r (given), we get

$$\mu_{EQ}(t_1[X], t_2[X]) \leq \min \{ \mu_{EQ}(t_1[Y], t_4[Y]), \mu_{EQ}(t_2[Z], t_4[Z]) \} \quad \text{--- (2)}$$

and

$$\mu_{EQ}(t_1[Y], t_3[Y]) \leq \min \{ \mu_{EQ}(t_1[W], t_4[W]), \mu_{EQ}(t_3[Z], t_4[Z]) \} \quad \text{--- (3)}$$

These three conditions imply that

$$\mu_{EQ}(t_1[X], t_2[X]) \leq \min \{ \mu_{EQ}(t_1[W], t_4[W]), \mu_{EQ}(t_2[Z], t_4[Z]) \}$$

Thus $Z(X) \rightsquigarrow Z(W)$ also holds in r .

The remaining axioms follow from these three axioms.

FS4 (Union) : $Z(X) \rightsquigarrow Z(Y)$

(given)

$$Z(X) \rightsquigarrow Z(XY) \quad \text{(by FS2)} \quad \text{--- (1)}$$

$$Z(X) \rightsquigarrow Z(V) \quad \text{(given)}$$

$$Z(XY) \rightsquigarrow Z(YV) \quad \text{(by FS2)} \quad \text{--- (2)}$$

$$Z(X) \rightsquigarrow Z(YV) \quad \text{(from (1) and (2) and FS3)}$$

FS5 (Decomposition) : $Z(X) \rightsquigarrow Z(YV)$

(given)

$$Z(YV) \rightsquigarrow Z(Y) \quad \text{(by FS1)} \quad \text{--- (1)}$$

$$Z(X) \rightsquigarrow Z(Y) \quad \text{(by FS3 and from (1))}$$

$$Z(YV) \rightsquigarrow Z(V) \quad \text{(by FS1)} \quad \text{--- (2)}$$

$$Z(X) \rightsquigarrow Z(V) \quad \text{(by FS3 and from (2))}$$

FS6 (Pseudo-transitivity) : $Z(X) \rightsquigarrow Z(Y)$

(given)

$$Z(XW) \rightsquigarrow Z(XW) \quad \text{(by FS2)} \quad \text{--- (1)}$$

$$Z(YW) \rightsquigarrow Z(V) \quad \text{(given)} \quad \text{--- (2)}$$

$$Z(XW) \rightsquigarrow Z(V) \quad \text{(from (1) and (2) and FS3)}$$

The following definition describes the closure of a set of attributes with respect to a set of fsds.[54]

Definition 5.2: Suppose S is a set of fsds of relation scheme R and let $W \subseteq R$. Then W^+ , the closure of W with respect to S , is the set of attributes $A \in R$, such that $Z(W) \rightsquigarrow Z(A)$ can be obtained from S using fsd inference axioms (FS1-FS6).

Lemma 5.1: $Z(W) \rightsquigarrow Z(V)$ follows from the inference axioms of fsds, iff $V \subseteq W^+$.

Proof: LET $V = \{ A_1, A_2, \dots, A_k \}$ and suppose that $Z(W) \rightsquigarrow Z(V)$ follows from the fsd inference axioms. Then for each i , $Z(W) \rightsquigarrow Z(A_i)$ holds by the decomposition axiom. So $V \subseteq W^+$.

Suppose $V \subseteq W^+$. By definition 5.3, for each i , $Z(W) \rightsquigarrow Z(A_i)$ is implied by the fsd axioms. Hence according to the union rule, $Z(W) \rightsquigarrow Z(V)$ follows.

Lemma 5.2: Natural join of fuzzy relations preserves fuzzy subset dependencies.

Proof: Let S_1 and S_2 be two sets of fsds satisfied by two fuzzy relations r_1 and r_2 respectively.

Let $r = r_1 \bowtie r_2$, we show that r satisfies all fsds in $S_1 \cup S_2$. Now we assume that r does not satisfy the fsds: $Z(X) \rightsquigarrow Z(Y)$ in $S_1 \cup S_2$. That is, for any two tuples t_1 and t_2 of r , there exists another tuple t_3 of r such that

$$\mu_{EQ}(t_1[X], t_2[X]) > \min \{ \mu_{EQ}(t_1[Y], t_3[Y]), \mu_{EQ}(t_2[Z], t_3[Z]) \} \quad \text{--- (1)}$$

We know that s is in $S_1 \cup S_2$. When $s \in S_j$ ($j \in \{1, 2\}$), then X, Y, Z are attributes of r_j , where X and Y are disjoint from Z . Therefore (1) leads to the contradiction that r_j violates $s \in S_j$, for $j \in \{1, 2\}$. Hence r satisfies $S_1 \cup S_2$.

It can be similarly shown that if r_i satisfies S_i , for $i = 1, 2, \dots, n$ then $r = \bigcap_{i=1}^n r_i$ satisfies $\bigcup_{i=1}^n S_i$.

Next, the completeness of the fsd axioms is examined. The following example shows that the fsd inference axioms FS1-FS6 are not always complete.

Example 5.2: We select an additional condition to the resemblance relations over $\text{dom}(A_i)$, $i = 1, 2, 3$, that is, for all $a_{li}, a_{lj} \in \text{dom}(A_i)$ and for all $a_{rk}, a_{rp}, a_{rm} \in \text{dom}(A_r)$, $r = 2, 3$, if $a_{rk}, a_{rp} \neq a_{rm}$, then

$$\mu_{EQ}(a_{li}, a_{lj}) > \min\{\mu_{EQ}(a_{2k}, a_{2m}), \mu_{EQ}(a_{3p}, a_{3m})\} \quad \text{--- (1)}$$

Consider the fsd $Z(A_1 A_2) \rightsquigarrow Z(A_3)$ over $R(A_1 A_2 A_3 A_4)$ (here $Z = A_4$). Let r be an instance of R . Since r satisfies $Z(A_1 A_2) \rightsquigarrow Z(A_3)$, for any three tuples t_1, t_2 , and t_3 of r such that

$$\mu_{EQ}(t_1[A_1A_2], t_2[A_1A_2]) \leq \min\{\mu_{EQ}(t_1[A_3], t_3[A_3]), \mu_{EQ}(t_2[A_4], t_3[A_4])\} \quad \text{--- (2)}$$

Let ϕ be a proposition defined by

$$\phi = \{ (t_1[A_1] \neq t_2[A_1]) \text{ and } (t_1[A_2] = t_2[A_2]) \}.$$

We now show that r also satisfies $Z(A_2) \rightsquigarrow Z(A_3)$. Assume that the proposition ϕ is true. From (2.24) $\mu_{EQ}(t_1[A_1A_2], t_2[A_1A_2]) = \mu_{EQ}(t_1[A_1], t_2[A_1])$. By (1) and (2), we have $t_1[A_3] = t_3[A_3]$ and $t_2[A_4] = t_3[A_4]$. Hence by reflexivity of resemblance relations, r satisfies $Z(A_2) \rightsquigarrow Z(A_3)$.

Now we assume that ϕ is false (i.e., either $t_1[A_1] = t_2[A_1]$ or $t_1[A_2] \neq t_2[A_2]$). From (2.24) $\mu_{EQ}(t_1[A_1A_2], t_2[A_1A_2]) = \mu_{EQ}(t_1[A_1], t_2[A_2])$. So, from (2), r satisfies $Z(A_2) \rightsquigarrow Z(A_3)$.

Example 5.2 specifies that depending upon the type of the resemblance relations used for defining the fsds, it is possible to imply new fsds that cannot be inferred using FS1-FS6. To infer such fsds, we have to consider additional inference axioms that depend on the resemblance relations used for comparing the domain values. A complete set of fsds where additional restrictions are imposed on EQUAL. In this connection, it will be useful to find a class of fsds for which the inference axioms FS1-FS6 constitute a complete set. The following theorem establishes completeness of these fsd axioms.

Theorem 5.1: The inference axioms form a complete set of inference axioms for fuzzy subset dependencies of a relation scheme $R(A_1 A_2 \dots A_n)$ 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 $\mu_{EQ}(a_i, b_i) = 0$. (5.2)

Proof: Let fsd: $Z(X) \rightsquigarrow Z(Y)$ be in fsd $(R) - \Sigma^+$. We need to show that r satisfies all fsds in Σ^+ but does not satisfy the fsd $Z(X) \rightsquigarrow Z(Y)$. Consider r to be an instance of R . Let t_1, t_2 and t_3 be three tuples in r . Let tuples t_2 and t_3 be a_1, a_2, \dots, a_n . The tuple t_1 is defined by

$$t_1[A_i] = \begin{cases} a_i & \text{if } A_i \in X^+ \\ b_i & \text{otherwise,} \end{cases}$$

where $\mu_{EQ}(a_j, b_j) = 0$ for $a_j, b_j \in \text{dom}(A_j)$, $j = 1, 2, \dots, n$.

First we show that r satisfies all the fsds in S . Let $Z(W) \rightsquigarrow Z(V)$ be an fsd in S . If $W \not\subseteq X^+$, then for $A_j \in (W - X^+)$, $\mu_{EQ}(t_1[A_j], t_2[A_j]) = \mu_{EQ}(b_j, a_j) = \mu_{EQ}(a_j, b_j) = 0$. So, $\mu_{EQ}(t_1[W], t_2[W]) = 0$. Hence by (5.1) r trivially satisfies $Z(W) \rightsquigarrow Z(V)$. When $W \subseteq X^+$, by lemma 5.1, $Z(X) \rightsquigarrow Z(W)$ and by transitivity, $Z(X) \rightsquigarrow Z(V)$. Applying lemma 5.1 again, $V \subseteq X^+$. Since $WV \subseteq X^+$, by construction $\mu_{EQ}(t_1[W], t_2[W]) = \mu_{EQ}(t_1[V], t_3[V]) = 1.0$ and $\mu_{EQ}(t_2[Z], t_3[Z]) = 1.0$ (given). Therefore r satisfies the fsd $Z(W) \rightsquigarrow Z(V)$.

Since s cannot be inferred from S using the inference axioms, we now show that r does not satisfy s . from the definition of r , $t_1[X] = t_2[X]$. In order that r satisfies s , we must have

$$\mu_{EQ}(t_1[Y], t_3[Y]) = \mu_{EQ}(t_2[Z], t_3[Z]) = 1.0$$

But this requires that $Y \subseteq X^+$, thereby implying that s follows from S by the inference axioms.

Since this conclusion violates our original assumption, the fuzzy relation r cannot satisfy the fsd $Z(X) \rightsquigarrow Z(Y)$. Hence the inference rules are complete.

The condition (5.2) defines a class of fsds for which the inference axioms FS1-FS3 constitute a complete set.

6. CONCLUSIONS

This paper deals with fuzzy embedded multivalued dependency and fuzzy subset dependency for treating precise as well as imprecise data. The definition of fuzzy subset dependency as given in definition 5.1 is not unique way of generalizing sd in a fuzzy database. For example, on introducing another translation rule for conditional fuzzy proposition such as Godelian implication rule[5, 12, 19] a different set of relation is to have binary truth values. Thus a relation may either satisfy the fsd: $Z(X) \rightsquigarrow Z(Y)$ or not. Condition (5.2) is only a sufficient condition for the axioms FS1-FS3 to be complete. It may be possible to find a restriction on EQUAL that is both necessary and sufficient for the completeness of the axioms.

REFERENCES

- [1] Baldwin. J.F, “Fuzzy Logic and its Application to Fuzzy Reasoning”, in *Advances in Fuzzy Set Theory and Applications*, M.M.Gupta et al., eds., North-Holland, Amsterdam, 1979.
- [2] Beeri. C, Fagin. R, and Howard. J. H, “A Complete Axiomatization for Functional and Multivalued Dependencies in Database Relations”, *ACM SIGMOD Conf.*, 1977, 47-61.
- [3] Codd. E.F, “Further Normalization of the Database Relational Model” in *Database Systems*, R. Rustin, ed., Prentice-Hall, Englewood Cliffs, NJ, 1972.
- [4] Dubois. D and Prade. H, “Fuzzy Sets and Systems: Theory and Applications”, Academic Press, New York, 1980.
- [5] Kandle. A, *Fuzzy Mathematical Techniques with Applications*, Addison-Wesley, California, 1986.
- [6] Kaufmann. A, “An Introduction to the Theory of Fuzzy Subsets”, Vol. 1, Academic Press, New York, 1975.
- [7] Jyothi. S, “Design Theory for Fuzzy Relational Databases”, Ph.D Thesis, S.V.University, Tirupati, 1992.
- [8] Jyothi. S and Syam Babu. M, “Multivalued Dependencies in Fuzzy Relational Databases and Lossless Join Decomposition”, *Fuzzy Sets and Systems*, 88, 1997.
- [9] Mamdani. E.H, “Applications of Fuzzy Logic to Approximate Reasoning using Linguistic Systems”, *IEEE Trans, Comput*, C-26, 1977.
- [10] Mizumoto. M, Fukumi. S and Tanaka. K, “Some Methods of Fuzzy Reasoning” in *Advances in Fuzzy Set Theory and Applications*, M.M.Gupta et al., eds., North-Holland, Amsterdam, 1979.
- [11] Prade. H and Testemale. C, “Generalizing Database Relational Algebra for the Treatment of Incomplete or Uncertain Information and Vague Queries”, *Information Sciences*, 3, 1984.
- [12] Prade. H, “Lipski’s Approach to Incomplete Information Database Restated and Generalized in the Setting of Zadeh’s Possibility Theory”, *Information Systems*, 9, 1984.
- [13] Raju. K.V. S. V. N and Majumdar. A. K, “Fuzzy Relational Databases”, Ph.D Thesis, Andhra university, Waltair, 1988.
- [14] Raju. K.V. S. V. N and Majumdar. A. K, “Fuzzy Functional Dependencies and Lossless Join Decomposition of Fuzzy Relational Database Systems” *ACM Trans., Database Systems*, 13, 1988.
- [15] Sagiv. Y, and Walecka. S. F, “Subset Dependencies and a Completeness Result for a Subclass of Embedded Multivalued Dependencies”, *JACM*, 29, 2, 373-393, 1982.
- [16] Senoi. S and Melton. A, “proximity Relations in the Fuzzy Relational Database Model”, North-Holland, Elsevier, 1989.
- [17] Zadeh. L. A, “A Theory of Approximate Reasoning” in *Machine Intelligence*, J. Hayes et al. eds., 9, Ellis Horwood Ltd., Sussex, UK, 1979.
- [18] Zadeh. L. A, “PRUF-aA Meaning Representation Language for Natural Language” in *Fuzzy reasoning and Applications*, E. H. Mamdani and B.R. Gaines, eds., Academic Press, New York, 1981.
- [19] Zemankova-Leech. M and Kandel. A, “Fuzzy Relational Databases-A key to Expert Systems”, Verlag TUV, Rhineland, 1984.