CHAPTER III
THEORETICAL BACKGROUND

Fuzzy Sets

Let U be a classical set of objects, called the universe of discourse. An element of U is denoted by u.

Definition 3.1. A fuzzy set F in a universe of discourse U is characterized by a membership function

$$\mu_F : U \to [0,1], \tag{3.1}$$

where for each u, $\mu_F(u)$ denotes the grade of membership of u in the fuzzy set F.

The notation which is often used in the literature for defining fuzzy set is

$$F = \{\mu(u_1)/u_1,\ \mu(u_2)/u_2,\ \ldots,\ \mu(u_n)/u_n\}, \tag{3.2}$$

where $u_i \in U$ and $1 \leq i \leq n$. The membership functions are found following a user's individual idea or from statistical data. For more details, refer to Dubois and Prade(1988). Note that a classical subset A of U can be viewed as a fuzzy subset with the membership function μ_A taking binary values, i.e.,

$$\mu_A = 1, \quad \text{if} \quad u \in A,$$
$$= 0 \quad \text{otherwise.}$$

Thus, the function maps elements of a universal set on to the set containing 0 and 1. This can be indicated by

$$\mu_A : U \to \{0, 1\}.$$

<u>Example 3.1.</u> Let the universe of discourse U of ages selected be

$$U = \{5, 10, 20, 30, 40, 50, 60, 70, 80\}$$

and the fuzzy set labels be infant, adult, young, and old. These fuzzy sets can be expressed as

young = {1/5, 1/10, .8/20, .5/30, .2/40, .1/50},
adult = {.8/20, 1/30, 1/40, 1/50, 1/60, 1/70, 1/80},
old = {.1/20, .2/30, .4/40, .6/50, .8/60, 1/70, 1/80},
infant = { }.

Operations on fuzzy sets have been developed from classical set theory operations. Let A and B be two fuzzy subsets in a universal set with membership functions μ_A and μ_B, respectively. Then the membership functions of A ∪ B, A ∩ B, and $\bar{A}$ are given by

$$\mu_{A \cup B} = \max(\mu_A(u), \mu_B(u)), \quad (3.3)$$

$$\mu_{A \cap B} = \min(\mu_A(u), \mu_B(u)), \quad (3.4)$$

$$\mu_{\bar{A}} = 1 - \mu_A(u). \quad (3.5)$$

Example 3.1 yields

young ∪ old = {1/5, 1/10, .8/20, .5/30, .4/40, .6/50, .8/60, 1/70, 1/80},
young ∩ old = {.1/20, .2/30, .2/40, .1/50},
not old = {1/5, 1/10, .9/20, .8/30, .6/40, .4/50, .2/60}.

Given two fuzzy subsets A and B in U, B is fuzzy subset of A, denoted by B ⊆ A, if

$$\mu_B(u) \leq \mu_A(u) \quad \text{for all } u \in U \quad (3.6)$$

From Example 3.1, $\mu_{old}(u) \leq \mu_{adult}(u)$ for all u ∈ U, then old ⊆ adult. Two fuzzy subsets are said to be equal if A ⊆ B and B ⊆ A.

Possibility Distributions

Let X be a variable taking values in U and $\mu_F(u)$ be the grade of membership of u in F for all $u \in U$. The proposition "X is F" induces a possibility distribution Π_x that is equal to F, i.e.,

$$\Pi_x = F \tag{3.7}$$

This distribution is characterized by a possibility distribution function, $\pi_x : U \rightarrow [0,1]$, that is equal to μ_F and associates with each $u \in U$. The possibility that X can take u as it value is denoted by

$$\pi_x(u) = Poss(X = u) = \mu_F(u) \text{ for } u \in U. \tag{3.8}$$

For example, in the proposition "Somchai is young", let young be fuzzy set in Example 3.1, and the possibility distribution Π_x = young. We can then conclude that the possibility that Somchai have the age = 30 is 0.5 or

$$\pi_x(30) = Poss(X = 30) = 0.5.$$

More details about possibility as well as about the differences between possibility and probability are given in Zadeh(1987b), Zimmermann(1986), and Dubois and Prade(1988).

Fuzzy Relations

Let $U = U_1 \times U_2 \times \ldots \times U_n$ be a Cartesian product of n universes and A_1, A_2, ..., A_n be fuzzy sets in U_1, U_2, ..., U_n, respectively. The Cartesian product $A_1 \times A_2 \times \ldots \times A_n$ is defined to be a fuzzy subset of $U_1 \times U_2 \times \ldots \times U_n$, where

$$\mu_{A_1 \times A_2 \times \ldots \times A_n}(u_1 \ldots u_n) = \min(\mu_{A_1}(u_1), \ldots, \mu_{A_n}(u_n)), \tag{3.9}$$

where $u_i \in U_i$, $i = 1,2,\ldots,n$.

Example 3.2. Let $U_1 = U_2 = \{2, 4, 6\}$, $A_1 = \{.5/2, 1/4, .6/6\}$, and $A_2 = \{1/2, .6/4\}$, then $A_1 x A_2 = \{.5/(2,2), 1/(4,2), .6/(6,2) .5/(2,4), .6/(4,4), .6/(6,4)\}$.

An n-ary fuzzy relation r is a fuzzy subset of a Cartesian product of some universes. Let U_1, U_2, ...,U_n be n universes of discourse, then a fuzzy relation r is a fuzzy subset of $U_1 \times U_2 \times \ldots \times U_n$ and is characterized by the n-variate membership function

$$\mu_r : U_1 \times U_2 \times \ldots \times U_n \text{ --> } [0,1]. \quad (3.10)$$

Note that a fuzzy relation is a fuzzy set defined on the Cartesian product of ordinary sets $U_1 \times U_2 \times \ldots \times U_n$, where tuples $(u_1, u_2, \ldots, u_n)$ may have varying degrees of membership within the relation.

Example 3.3. Let r be a fuzzy relation between the two sets X = {Chiangmai, Bangkok} and Y = {Songkla, Chiangmai, Cholburi} that represents the concept "far." This relation can be written as

r(X,Y) = {1/(Chiangmai,Songkla), 0/(Chiangmai,Chiangmai), .6/(Chiangmai,Cholburi), .75/(Bangkok,Songkla), .7/(Bangkok,Chiangmai), .3/(Bangkok,Cholburi)}.

For a comparison of elements in a single set, a fuzzy relation must be defined to be a fuzzy measure as given by

Definition 3.2. A fuzzy relation EQUAL(EQ) over a universe of discourse U is defined to be a fuzzy subset of U x U, where μ_{EQ} satisfies the conditions: For all a, b $\in$ U,

$$\mu_{EQ}(a, a) = 1 \quad \text{(reflexivity)},$$
$$\mu_{EQ}(a, b) = \mu_{EQ}(b, a) \quad \text{(Symmetry)}. \quad (3.11)$$

The relations in Definition 3.2, satisfying reflexive and symmetric properties, are called compatibility relations or tolerance relations or proximity relations(Krir and Folger, 1988) or EQUAL relations(Raju and Majumdar, 1988).

The possibilistic interpretation of $\mu_{EQ}(a,b)$ is the possibility of treating a and b as EQUAL.

Fuzzy Relational Data Models

Since data and relations can be fuzzy or non-fuzzy, classical relational data models can be extended into the four difference types shown in Table 3.1. A relational data model with fuzzy data is a model in which data are fuzzy, but relations are ordinary.

Before dealing with the other types of models in Table 3.1, we consider first the following notations used in fuzzy relational models. Let the relation scheme R be a finite set of attribute name $\{A_1, A_2, \ldots, A_n\}$ and denoted by $R(A_1A_2\ldots A_n)$ or simply by R. Corresponding to each attribute name A_i, $1 \leq i \leq n$, is a set $dom(A_i)$, called the domain of A_i. However, unlike a classical relations, in fuzzy relational models, $dom(A_i)$ may be a fuzzy set or even a set of fuzzy sets. Hence, along with each attribute A_i, we associate a set U_i, called the universe of discourse for domain value of A_i.

Definition 3.3 A fuzzy relation r on the relation scheme $R(A_1A_2\ldots A_n)$ is a fuzzy subset of $dom(A_1) x dom(A_2) x \ldots x dom(A_n)$.

By this definition, Type I fuzzy relations, $dom(A_i)$ can only be a fuzzy set or a classical set. In Type II fuzzy relations, $dom(A_i)$ can be a set of fuzzy sets.

Table 3.1. Relational data models and their extensions to fuzzy fuzzy data value and fuzzy relation.

	Non-fuzzy data value.	Fuzzy data value.
Non-fuzzy relations.	Classical relational data models.	Relational data models with fuzzy data.
Fuzzy relations.	Type I fuzzy relational data models.	Type II fuzzy relational data models.

The representation of fuzzy relations is shown by table as classical relations, but has an addition column of membership value of tuple t in r denoted by $\mu_r(t)$. This table contains only tuples for which $\mu_r(t) > 0$.

<u>Type I Fuzzy Relational Data Model.</u> In this type of data model, $dom(A_i)$ may be a crisp subset or fuzzy subset of U_i. Let μ_{A_i} denote the membership function of $dom(A_i)$ for i=1,2,...,n. Then, by (3.9), $dom(A_1) x dom(A_2) x \ldots x dom(A_n)$ is a fuzzy subset of $U = U_1 \times U_2 \times \ldots \times U_n$. Hence, this type of relation r is also a fuzzy subset of U with membership function μ_r. Also, it follows from (3.6) and (3.9) that for all $(u_1 u_2 \ldots u_n) \in U$, μ_r must satisfy

$$\mu_r(u_1 u_2 \ldots u_n) \leq \min[\mu_{A_1}(u_1), \mu_{A_2}(u_2), \ldots, \mu_{A_n}(u_n)]. \quad (3.12)$$

In possibilistic interpretation, μ_r can be treat as a possibility distribution function in U. Thus $\mu_r(u_1 u_2 \ldots u_n)$ determines the possibility that the tuple $t \in U$ has $t[A_i]=u_i$, for i = 1,2,...,n.

Example 3.4. Let LIKES(Student,Course) be a relation scheme, where dom(Student) and dom(Course) are ordinary sets. In a fuzzy relation r shown in Table 3.2., μ_r is a fuzzy measure of association between Student and Course or a possibility distribution function. Thus the possibility that Somchai liking EDP is 0.80.

Example 3.5. Let R(N,J,X,S) be a relation scheme of highly experienced and highly salaried employees, where N = Employee's name, J = Job, X = Experience, S = Salary, dom(N) and dom(J) are ordinary sets, and dom(X) and dom(S) are fuzzy sets High-Experience and High-Salary in appropriate universes. Let U_x and U_s be sets of positive integer 0-30 and 10,000-50,000, respectively. The membership function μ_{HX} and μ_{HS} of the fuzzy sets Hign-Experience and High-Salary are

$$\mu_{HX}(x) = (1+|x-10|/4)^{-1}, \quad \text{for } x \leq 10,$$
$$= 1 \quad , \quad \text{for } x > 10,$$

$$\mu_{HS}(x) = (1+|s-30{,}000|/10{,}000)^{-1}, \quad \text{for } s \leq 30{,}000,$$
$$= 1 \quad , \quad \text{for } s > 30{,}000.$$

In this example, $\mu_r(t)$ can be interpreted as the truth value of the fuzzy proposition "Y has high experience and high salary" for the tuple t. Thus the truth value of the fuzzy proposition "Jukkrit has high experience and hign salary" is 0.67.

Table 3.2. An instance r of LIKES

Student	Course	μ
Somchai	OS	0.75
Eak	OS	0.90
Somchai	EDP	0.80
Veera	EDP	0.95

Table 3.3. An instance r of Highly Experienced and Highly Salaried Employees

Name	Job	Experience	Salary	μ
Jukkrit	Engineer	8	30,000	0.67
Waruth	Manager	9	35,000	0.80
Jarunee	Secretary	8	20,000	0.50
Satid	Engineer	12	40,000	1.00
Paiboon	Engineer	9	30,000	0.80

<u>Type II Fuzzy Relational Data Model.</u> This type considers a further generalization of the fuzzy relationl data model, where for any attribute A_i, dom(A_i) may be a set of fuzzy sets in U_i. This generalization, a tuple $t=(a_1a_2...a_n)$ in D = dom(A_1)xdom(A_2) x...xdom(A_n), becomes a fuzzy subset of $U = U_1xU_2x...xU_n$ with

$$\mu_t(u_1u_2...u_n) = \min[\mu_{a_1}(u_1),\mu_{a_2}(u_2),...,u_{a_n}(u_n)], \quad (3.13)$$

where $u_i \in U_i$, for i = 1,2...,n. Since (3.13) holds for all $u_i \in U_i$, i = 1,2,...,n and according to Definition 3.3 a Type II fuzzy relation r is a fuzzy subset of D, it follows from (3.6) that the membership function

$$\mu_r \;: D \dashrightarrow [0,1] \qquad (3.14)$$

must satisfy the condition

$$\mu_r(t) \leq \max_{(u_1 u_2 \ldots u_n) \in U} [\min\{\mu_{a_1}(u_1), \mu_{a_2}(u_2), \ldots, \mu_{a_n}(u_n)\}], \qquad (3.15)$$

where $t = (a_1 a_2 \ldots a_n) \in D$.

In this type, μ_r may be interpreted either as a possibility measure of association among the data values or as a truth value of a fuzzy predicate associated with r. We can treat a_i as a possibility distribution on U_i. In other words, for a tuple $t = (a_1 a_2 \ldots a_n) \in D$, the possibility of $t[A_i] = u_i$ is equal to $\mu_a(u_i)$. For example, suppose that an instance of the relation Employee(Name, Salary) contains tuples (Somsri, S), where S={0.3/10,000, 0.6/20,000, 0.8/30,000}, where S represents the possibility distribution for the salary of Somsri, i.e., Poss(Salary of Somsri = 20,000) = 0.6.

Based on the possibilistic interpretation, for a tuple t of r, we obtain

$$\begin{aligned} &\text{Poss}(t[A_1] = u_1,\ t[A_2] = u_2, \ldots, t[A_n] = u_n) \\ &\qquad = \min\{\mu_r(t), \mu_t(u_1 u_2 \ldots u_n)\}, \end{aligned} \qquad (3.16)$$

where $u_i \in U_i$, $i = 1,2,\ldots,n$ and μ_t is given by (3.14).

<u>Example 3.6.</u> Let EMPLOYEE(N,D,J,X,S,I) be a relation scheme, where N = Employee's Name, D = Department, j = job, X = Experience, S = salary , I = Income Tax. The dom(N), dom(D), and dom(J) are ordinary sets. But dom(X), dom(S), and dom(I) are a set of fuzzy sets in the universes U_x, U_s, and U_I, respectively. Let U_x, U_s, and U_I be sets of positive integers in

Fuzzy Relational Operations

Since a fuzzy relation is a fuzzy subset of the Cartesian product of its attribute domains, the concepts of union, intersection, and cross-product of fuzzy sets can be extended to fuzzy relations such as join and projection operation in the following:

Projection. Let r be a fuzzy relation of the relation scheme $R(A_1A_2...A_n)$, $R_i(A_{i_1}...A_{i_k})$ be a subset of R, and $t[R_i]$ be the restriction of t on attribute R_i, i.e., for $t = (a_1a_2...a_n)$, $t[R_i] = (a_{i_1}...a_{i_k})$. The projection $r_i = P_{R_i}(r)$ is a k-ary fuzzy relation in $dom(A_{i_1})xdom(A_{i_2})x...xdom(A_{i_k})$. The membership function μ_r is given by

$$\mu_{r_i}(t) = \max_{t_r} \{\mu_r(t_r) \mid t_r[R_i] = t\}, \qquad (3.17)$$

where t_r is a tuple of r and $t \in dom(A_{i_1})xdom(A_{i_2})x...xdom(A_{i_k})$.

Example 3.7. The projection of the fuzzy relation r in Table 3.3 over R_{JS} = {JOB, SALARY} is shown in Table 3.5.

Another operation that can be viewed to be the inverse of projection is called a cylindrical extension. The cylindrical extension of r_i on R is denoted by $\hat{r}_i = C_R(r_i)$. This operation is an n-ary fuzzy relation in $D = dom(A_1)xdom(A_2)x...xdom(A_n)$. The membership μ_r is given by

$$\mu_{\hat{r}_i}(t) = \mu_{r_i}(t[R_i]) \quad \text{for } t \in D, \qquad (3.18)$$

From the definition of projection and cylindrical extension, we can see that for any instance r of a relation scheme R and $R_i \subseteq R$

$$\mu_{\hat{r}_i}(t) \geq \mu_r(t), \qquad (3.19)$$

where $\hat{r}_i = C_R(P_{R_i}(r))$ and $t \in D$; in other words,

$$r \subseteq C_R(P_{R_i}(r)). \qquad (3.20)$$

<u>Join.</u> Let $\rho = \{R_1, R_2, \ldots, R_s\}$ be a set of relation schemes $R(A_1A_2\ldots A_n) = R_1R_2\ldots R_s$, and let $\{r_1, r_2, \ldots, r_s\}$ be a set of fuzzy relations, where r_i is an instance of R_i, $i = 1,2,\ldots,s$. The natural join of these fuzzy relations is a fuzzy relation of the relation scheme R written as

$$r = r_1 x r_2 x \ldots x r_s \qquad (3.21)$$

or denoted by $x_{i=1}^{s} r_i$. The membership of r is given by

$$\mu_r(a_1a_2\ldots a_n) = \min(\mu_{\hat{r}_1}(a_1a_2\ldots a_n), \mu_{\hat{r}_2}(a_1a_2\ldots a_n), \ldots, \mu_{\hat{r}_s}(a_1a_2\ldots a_n)], \qquad (3.22)$$

where $a_i \in dom(A_i)$ and $\hat{r}_j$ is the cylindrical extension of r_j on R, for $i = 1,2,\ldots,n$ and $j = 1,2,\ldots,s$.

Table 3.5. Projection of an instance of Highly Experienced and Highly Salaried Employees

Job	Salary	μ
Engineer	30,000	0.8
Engineer	40,000	1.0
Manager	35,000	0.8
Secretary	20,000	0.5

Example 3.8. The join of instance r in Table 3.2 with an instance r_1 in Table 3.6 is shown in Table 3.7.

Consider the fuzzy relations r_i, $i = 1,2,\ldots,s$, obtained from the projection of a fuzzy relation r on $R_i, i = 1,2,\ldots,s$ and by (3.21) and (3.23), one has

$$r \subseteq P_{R_1(r)} x P_{R_2(r)} x \ldots x P_{R_s(r)}, \quad (3.23)$$

i.e.,

$$r \subseteq m_r(r), \quad (3.24)$$

where $m_r(r) = P_{R_1(r)} x P_{R_2(r)} x \ldots x P_{R_s(r)}$.

Fuzzy Integrity Constraints

Integrity constraints in relational database systems are constraints imposed on attributes, domains, tuples, and relations. Such constraints can be classified into two groups (Date, 1981; Raju and Majumdar, 1988; Yang, 1986):

1. Domain Dependency - which specifies the characteristics of an attribute (restricts admissible domain values).

2. Data Dependency - which specifies a relationship among several attributes in a database. This type of constraint is related to the structure of a database and has greater impact on the design of the database systems. There are several types of data dependencies, such as functional dependency, multivalued depandency, join dependency, etc.

Table 3.6. An instance r_1 of TEACHING

Teacher	Course	μ
Montri	OS	0.80
Montri	EDP	0.60
Somsak	OS	0.60
Somsak	EDP	0.90

Table 3.7. Natural Join of Fuzzy Relations in Tables 3.2 and 3.6

Student	Teacher	Course	μ
Somchai	Montri	OS	0.75
Somchai	Somsak	OS	0.60
Somchai	Montri	EDP	0.60
Somchai	Somsak	EDP	0.80
Eak	Montri	OS	0.80
Eak	Somsak	OS	0.60
Veera	Montri	EDP	0.60
Veera	Somsak	EDP	0.90

Fuzzy relational database systems, have also these two types of fuzzy integrity constraints. In order to deal with these constraints, the particularization concept(Zadeh, 1981) was introduced, as well as the translation rules of fuzzy calculus to evaluate the particularization of a fuzzy relation due to a compound fuzzy proposition. This study uses the translation rule R_s to determine the possibility distribution of fuzzy conditional propositions given below.

<u>Definition 3.4.</u> Let F and G be fuzzy subset of U and V, respectively. The possibility distribution $\Pi(X \sim> Y)$ associated with the conditional fuzzy proposition "If X is F then Y is G" is given by

$$\Pi(X \sim> Y) = R_s, \tag{3.25}$$

where R_s is a fuzzy subset of UxV with the membership function

$$\mu_R(u,v) = 1, \quad \text{if } \mu_F(u) \leq \mu_G(v),$$
$$= 0 \quad \text{otherwise.} \tag{3.26}$$

From this definition, we can see that μ_R defines a hard partition of U x V, i.e., R_s is an ordinary subset of U x V.

From Definition 3.2, EQUAL can be extended over the composite domain for the purpose of studying fuzzy data dependency in the following. Let $D = dom(A_1) x dom(A_2) x ... x dom(A_n)$, and t_1, t_2 be two tuples in D. The extension over D of the fuzzy relation EQUAL defines a fuzzy subset of D x D, with the membership function

$$\mu_{EQ}(t_1,t_2) = \min \{\mu^1_{EQ}(t_1[A_1], t_2[A_1]),$$
$$\mu^2_{EQ}(t_1[A_2], t_2[A_2]), ...$$
$$\mu^n_{EQ}(t_1[A_n], t_2[A_n])\}. \tag{3.27}$$

There are various ways to define μ_{EQ} over a domain of Type I and Type II relations. In practice, however, it should be selected or defined by the database designer during the database creation.

Fuzzy Functional Dependencies

In relational database systems, a functional dependency $f: X \sim\to Y$ that is satisfied in an instance r of a relation scheme, if for each pair of tuple t_1 and t_2 of r that $t_1[X] = t_2[X]$, is $t_1[Y] = t_2[Y]$. In fuzzy relational database systems, a fuzzy functional dependency should define from a given fuzzy proposition "If X is equal then Y is equal."

Definition 3.5. A fuzzy functional dependency(ffd) $X \sim\to Y$ with $X, Y \subseteq R$ holds in the fuzzy relation r on R, if for all t_1 and t_2 of r, one has

$$\mu_{EQ}(t_1[X], t_2[X]) \leq \mu_{EQ}(t_1[Y], t_2[Y]). \tag{3.28}$$

By this definition, the following set of sound and complete inference rules for ffds hold:

FFD1(Reflexivity): If $Y \subseteq X$, then $X \sim\to Y$.

FFD2(Augmentation): If $X \sim\to Y$, then $XZ \sim\to YZ$.

FFD3(Transitivity): If $X \sim\to Y$ and $Y \sim\to Z$, then $X \sim\to Z$.

FFD4(Union): If $X \sim\to Y$ and $X \sim\to Z$, then $X \sim\to YZ$.

FFD5(Decomposition): If $X \sim\to YZ$, then $X \sim\to Y$ and $X \sim\to Z$.

FFD6(Pseudotransitivity): If $X \sim\to Y$ and $YW \sim\to Z$, then $XW \sim\to Z$.