MTA SZTAKI Közlemények 37/1987 pp. 223-248

RETRIEVAL FROM FUZZY DATABASE BY FUZZY RELATIONAL ALGEBRA

LE TIEN VUONG

Institute of Informatics and Cybernetics,
Hanoi, Viet Nam

HO THUAN

Computer and Automation Institute,
Hungarian Academy of Sciences

1. INTRODUCTION

In the real world, there exists many things and events which can not or need not be precisely defined. This imprecision may be classified into two types: randomness and fuzziness. The concept of randomness is associated with objective things and has been studied by probability theory. The fuzziness is related to subjective phenomena, and has been studied by fuzzy sets theory. The concept of fuzzy sets has been introduced by L. Zadeh/75/, and applied to various fields.

By fuzzy data, we mean data which have the properties of fuzziness, that is, related to human beings or a result of their subjective observations.

In order to represent and to process fuzzy data in a digital computer we must first choose a suitable mathematical model. There are some models for representation of fuzzy data, for example, fuzzy sets and linguistic variable /L. Zadeh 75/, fuzzy graphs/M. Umano 79/ and fuzzy data classification /W. Belke, et al. 78/. From the view-point of easy implemen-

This work was partially supported by the Hungarian National Science Foundation, Grant No. 1066 / 1985

tation and wide application for most database systems, we extend the Cood's model to a relational fuzzy database. The concept of a relational fuzzy database and a somewhat different point of view about the normalization in this model will be discussed detaily in /LE and HO/. In this paper we only introduce a concept of extended relational model with corresponding relational operations for fuzzy relations. These relational operations have been written in PL/1 and experimentalized on an EC1022 computer.

The physical and logical representation of fuzzy data of this system can be seen in /Le Thien Vuong 83,85/.

2. THE CONCEPT OF FUZZY SETS

In this section we shall present briefly the concept of fuzzy sets which are required in the sequel. Here we only want to repeat some definitions of fuzzy sets theory. More details of the discussion may be seen in /L.A. Zadeh, 75/.

We introduce the following definitions.

Definition 2.1

Let $U = \{u\}$ be a universe of discourse /i.e., a collection of objects/, denoted generically by u; then a fuzzy set $\underset{\sim}{A}$ of U is a set of ordered pairs $\{(\mu_{\underset{\sim}{A}}(u)/u)\}$, $u \in U$, where $\mu_{\underset{\sim}{A}}(u)$ is the grade of membership of u in $\underset{\sim}{A}$ and $\mu_{\underset{\sim}{A}}: U \to [0,1]$ is the membership function.

In order to capture more information about the entities from a mini-world we choose fuzzy sets to apply to our data base.

Let us consider a complete and distributive lattice V which is the set of all fuzzy sets of the universe of discourse U

$$V(U) = \{\mu / \mu : U \to [0,1] \}$$

/see Negoita et al. 75, pp. 15-16/.

In this lattice are defined the operations union, intersection and complementation as follows:

Definition 2.2

Let $\underset{\sim}{A}$ and $\underset{\sim}{B}$ be two fuzzy sets of U.

a/ UNION /denoted by $\underset{\sim}{A} \cup \underset{\sim}{B}$ or by $\underset{\sim}{A}\underset{\sim}{B}$/

The union of two fuzzy sets $\underset{\sim}{A}$ and $\underset{\sim}{B}$ of U is defined

$$\mu_{\underset{\sim}{A} \cup \underset{\sim}{B}}(u) = \mathrm{Max}(\mu_{\underset{\sim}{A}}(u), \mu_{\underset{\sim}{B}}(u)) \quad \text{for} \quad u \in U$$

or

$$\mu_{\underset{\sim}{A} \cup \underset{\sim}{B}}(u) = \mu_{\underset{\sim}{A}}(u) \vee \mu_{\underset{\sim}{B}}(u) \quad \text{for} \quad u \in U;$$

b/ INTERSECTION /denoted by $\underset{\sim}{A} \cap \underset{\sim}{B}$/

The intersection of two fuzzy sets $\underset{\sim}{A}$ and $\underset{\sim}{B}$ of U is defined by

$$\mu_{\underset{\sim}{A} \cap \underset{\sim}{B}}(u) = \mathrm{Min}(\mu_{\underset{\sim}{A}}(u), \mu_{\underset{\sim}{B}}(u)) \quad \text{for} \quad u \in U$$

or

$$\mu_{\underset{\sim}{A} \cap \underset{\sim}{B}}(u) = \mu_{\underset{\sim}{A}}(u) \wedge \mu_{\underset{\sim}{B}}(u) \quad \text{for} \quad u \in U$$

c/ COMPLEMENTATION

$\neg \underset{\sim}{A}$ is the complementation of fuzzy set $\underset{\sim}{A}$ of U defined by

$$\mu_{\neg \underset{\sim}{A}}(u) = 1 - \mu_{\underset{\sim}{A}}(u) \qquad \text{for} \quad u \in U;$$

d/ EQUALITY

$$\underset{\sim}{A} = \underset{\sim}{B} \Leftrightarrow \mu_{\underset{\sim}{A}}(u) = \mu_{\underset{\sim}{B}}(u) \qquad \text{for} \quad u \in U;$$

e/ CONTAINMENT

$$\underset{\sim}{A} \subseteq \underset{\sim}{B} \Leftrightarrow \mu_{\underset{\sim}{A}}(u) \leq \mu_{\underset{\sim}{B}}(u) \qquad \text{for} \quad u \in U.$$

The symbols $\wedge$, $\vee$ and $\neg$ stand for minimum, maximum and complement.

Remark

The symbols $\wedge$, $\vee$ and $\neg$ stand for AND, OR and NOT in ordinary logic.

Definition 2.3

Let $U = \{u\}$ and $V = \{v\}$ be two arbitrary universes of discourse. A fuzzy /binary/ relation R from U to V is a fuzzy subset of the Cartesian product $U \times V = \{(u,v)\}$, characterized by the membership function $\mu_R: U \times V \to [0,1]$.
A fuzzy relation R in $U \times V$ is expressed by

$$R = \{(\mu(u_1,v_1) / (u_1,v_1), \ldots, \mu(u_n,v_m)/(u_n,v_m) : u_i \in U, v_j \in V\}$$

Remark

Generally, an n-ary fuzzy relation R in $U_1 \times \ldots \times U_n$ is characterized by a membership function over $U_1 \times \ldots \times U_n$ /see L.A. Zadeh 75/.

Definition 2.4

A linguistic variable is characterized by a quintuple (A,T,U,G,M) in which A is the name of the variable; T denotes the term set of A, that is, the set of names of linguistic values of A; G is the syntactic rule for generating the names in T and M is the semantic rule for associating with each t in T its meaning M(t), which is a fuzzy set of U.

The meaning of a linguistic value t can be defined as

$$M(t) = \{(\mu_t(u) / u) : u \in U\} .$$

3. RELATIONAL MODEL AND EXTENDED RELATION

3.1 The data model

The relational operations developed in this paper are applicable to most database systems, specially to fuzzy database systems. The concept of relational model may be seen in /E.F. Cood, 70/.

To simplify later discussions, we shall choose the following sample data model of a system consisting of two n-ary normalized relations /fig. 1 and fig. 2/.

Figure 1 represents some entries of the relation

EMP(MNO,MNAME,AGE,DNO,SAL),

where every row of the table corresponds to the data of an object /an employee/. The elements of EMP consists of an employee's person ordinal number, name, age, department ordinal number and his salary. Figure 2 represents some entries of the relation

DEPT(DNO,DNAME)

where a row of the DEPT-relation consists of the department ordinal number and the name of a department.

EMP (MNO	NAME	AGE	DNO	SAL)
100	Fischer	25	10	1000
101	Neuman	20	11	1500
102	King	20	11	2000
103	Shmid	30	12	2000
104	John	50	13	3000

Fig.1: EMP-relation

DEPT (DNO	DNAME)
10	A
11	B
12	C
13	D

Fig.2: DEPT-relation

Such a systems should be limited to consist of only precise data. There exists yet another possibility for describing a relation in which some linguistic values /fuzzy term/ have been used as fuzzy data /see Fig. 3/.

F-EMP (MNO	NAME	AGE	DNO	SAL)
100	Fischer	25	10	1000
101	Neuman	young	11	1500
102	King	young	11	high
103	Shmid	30	12	2000
104	John	OLD	13	very high

Fig.3: Extended relation F-EMP

In /V. Tahani, 77/ the author had given a method for fuzzy queries processing based on a data model with precise situations. In /M. Umano, 79/ the author had introduced a logical and physical representation for the fuzzy sets as the fuzzy data. These fuzzy sets have been processed by 52 operations.

In this paper we consider not only precise /certain/ data but also fuzzy data. This data can be processed by relational operations.

3.2 Extended relational model

Following customary notation, we use the letters A,B,... to denote single attributes, X,Y,... to denote sets of attributes. ER to denote an extended relation and ATTR(ER) to denote the attribute set of the relation ER. Let $U(A_i)$ or U_i be the usual regular domain of an attribute A_i. The set U_i must be added by the corresponding set of fuzzy term, i.e. the linguistic values of A_i /see Def. 2.4/ in order to enable to process the data represented in Fig. 3.

Every $t \in T$ can be evaluated by semantic rules /see Def. 2.4/ into a fuzzy set. Then the extended domain of attribute A_i is $D(A_i) = U(A_i) \cup T(A_i)$ or briefly $D_i = U_i \cup T_i$.
In order to describe easily the nonprocedural sublanguage as SEQUEL, we assume that the data of the system have the logical and physical representation as in /Le Tien Vuong, 83,85/.

We can formulate an extended relational model as follows.

Definition 3.1

An extended database D_e is defined as a collection of extended relations

$$D_e = \{ER_1, \ldots, ER_n\},$$

where every extended relation is a subset of Cartesian product of the extended domains

$$ER_i \subseteq \{U(A_{i_1}) \cup T(A_{i_1})\} x \ldots x \{U(A_{i_{m_i}}) \cup T(A_{i_{m_i}})\}$$

in which $U(A_{i_j})$ or U_{i_j}, $j = 1, \ldots, m_i$ are basic domains and $T(A_{i_j})$ or T_{i_j} are sets of fuzzy terms of the corresponding attributes A_{i_j}.

To illustrate the Definition 3.1 we consider the following example /see Fig. 3/. The extended relation F-EMP consists of the following domains:

U_1(MNO) : all 6-space digits,

U_2(NAME) : all possible individual names,

U_3(AGE) : all integers 1-150,

T_3(AGE) : the set of all fuzzy terms as ols, young,...

U_4(DNO) : all possible 2-space digits,

U_5(SAL) : all 4-space digits,

T_5(SAL) : the set of all fuzzy term as high, middle,...

We have

$$\text{F-EMP} \subseteq U_1 \times U_2 \times \{U_3 \cup T_3\} \times U_4 \times \{U_5 \cup T_5\}.$$

Here the sets T_1, T_2 and T_4 are empty.

3.3 The relational operations on an extended relation

The purpose of this section is to extend the operations of relational algebra to extended relations in order to process a class of fuzzy queries. These fuzzy queries should be processed not only by operations of fuzzy sets but also by relational operations.

Let r denote a tuple of ER; $A_i \in \text{ATTR(ER)}$; $r[A_i]$ is the value of attribute A_i in a tuple $r \in \text{ER}$ with $r[A_i] = a_{ij} \in D(A_i)$. If $X = \{A_1,\ldots,A_k\}$ is a subset of ATTR(ER) and r is a tuple over ATTR(ER) then the restriction of r on X will be called the X-value of r denoted by $r[X]$ and is expressed by

$$r[X] = r[A_1,\ldots,A_k] = (r[A_1],\ldots,r[A_k]).$$

Based on Definition 2.2 and Definition 2.4, we can introduce the following concepts.

If $r_1[A] = u_1$, $r_2[A] = u_2$, $r_1, r_2 \in ER$ then the equality of $r_1[A]$ and $r_2[A]$ is defined by

$$r_1[A] = r_2[A] \Leftrightarrow u_1 = u_2 \quad \text{if} \quad u_1, u_2 \in U(A) \quad \text{or}$$

$$M(u_1) = M(u_2) \quad \text{if} \quad u_1, u_2 \in T(A).$$

In general case, where $X \subseteq ATTR(ER)$ the equality of $r_1[X]$ and $r_2[X]$ can be defined analogously:

Let be $r_1[X] = (r_1[A_1], \ldots, r_1[A_k])$,

$$r_2[X] = (r_2[A_1], \ldots, r_2[A_k]),$$

$$r_i[A_j] \in D_j, \quad i = 1,2; \quad j = 1, \ldots, k.$$

The equality of $r_1[X]$ and $r_2[X]$ is defined by

$$r_1[X] = r_2[X] \Leftrightarrow r_1[A_i] = r_2[A_i] \text{ for } i = 1, \ldots, k;$$

$$A_i \in ATTR(ER).$$

In the following, we introduce some definitions of extended relational operations.

a. Projection

Based on the above concepts the projection of an extended relation should be defined as follows:

Definition 3.2

Let ER be an extended relation over a set of attributes $ATTR(ER)$ and $X \subseteq ATTR(ER)$ with $X = \{A_1, \ldots, A_k\}$.

The projection of ER over X is defined by

$ER[X] = \{r[X]: r \in ER\}$

$= \{(r[A_1],\ldots,r[A_k]): r \in ER \wedge A_i \in X \wedge X \subseteq ATTR(ER) \wedge i=1,\ldots,k.\}$

In more intuitive terms, ER[X] is obtained from ER by deleting all columns of ER not corresponding to attributes in X and identifying duplicate tuples in what remains.

Remark

The Projection introduced here is analogous to Codd's projection /see E.F. Codd, 70/. The difference consists only in that two tuples are equal if all corresponding components of tuples are equal in the sense of fuzzy sets /see Definition 2.2/.

We have now a simple example /see Figure 3/.

Example

Let X = {AGE, DNO}

F-EMP[X] =

(AGE	DNO)
25	10
young	11
30	12
old	13

b. Join

The equijoin of two extended relations is defined as follows:

Definition 3.3

Let ER and ES be two extended relations over the attributes sets ATTR(ER) and ATTR(ES), respectively. $A \in ATTR(ER)$ and $B \in ATTR(ES)$ are two attributes. The equijoin of ER and ES on A and B is defined by

$ER[A=B]\,ES = \{r = (w,s): w \in ER \wedge s \in ES \wedge w[A] = s[B]\}$.

In the case A and B are the same attribute names, one of these two columns is deleted in the result. The join is called natural join (denoted by ⋈). By analogy, the natural join of two relations ER and ES on ATTR(ER)∩ATTR(ES) can be formulated as follows:

$$ER \bowtie ES = \{r \mid (\exists w \in ER: w = r[ER] \wedge \exists s \in ES: s = r[ES])\ .$$

Example /see Figure 2 and 3/.

F-EMP ⋈ DEPT =

(MNO	NAME	AGE	DNO	SAL	DNAME)
100	Fischer	25	10	1000	A
101	Neuman	young	11	1500	B
102	King	young	11	high	B
103	Shmid	30	12	2000	C
104	John	old	13	very high	D

c. Section

Definition 3.4

Let ER be an extended relation over ATTR(ER). A is an attribute and c is a value from $D(A) = U(A) \cup T(A)$. The selection A=c from ER (denoted by ER [A = c]) is defined by

$$ER[A=c] = \{r \in ER: (r[A]=c' \wedge c' \in U(A) \wedge c \in U(A) \wedge c=c')$$
$$\vee (r[A]=\underset{\sim}{u} \wedge \underset{\sim}{u} \in T(A) \wedge c \in U(A) \wedge \mu_{\underset{\sim}{u}}(c)=1)$$
$$\vee (r[A]=u_i \wedge u_i \in U(A) \wedge c \in T(A) \wedge \mu_c(u_i) \geq \tau)$$
$$\vee (r[A]=\underset{\sim}{u} \wedge \underset{\sim}{u} \in T(A) \wedge c \in T(A) \wedge (\exists u_i \in U(A):$$
$$\mu_{\underset{\sim}{u} \cap c}(u_i) \geq \tau \wedge \tau \in [0,1]))\}.$$

The value $\tau \in [0,1]$ is given by the user. If the user doesn't specify then $\tau = 0.5$ is assumed as default value.

We shall use F-EMP relation in Fig. 3 on an example.

Example

If we define $\text{young}(x) = \begin{cases} 1 & \text{for } x \leq 24 \\ 0.5 & \text{for } 25 \leq x \leq 30, \\ 0 & \text{otherwise} \end{cases}$ then

F-EMP [AGE = young] =

(MNO	NAME	AGE	DNO	SAL)
100	Fischer	25	10	1000
101	Neuman	young	11	1500
102	King	young	11	high
103	Shmid	30	12	2000

Remark

The remaining operations as UNION, INTERSECTION, DIFFERENCE for the extended relations are computed as in the theory of ordinary sets.

3.4 Fuzzy database and fuzzy relational algebra

Wirh the extended relational operations in the subsection 3.3 we have a possibility to process the data of a system more easily and finely. Therefore user's requirements can be satisfies more efficiently. In order to obtain an efficient database management system and an easy data processing system at high level the fuzzy data must be considered as the fuzzy sets. The stored data in Fig. 3 are neither "certain" nor "fuzzy term" data, but they are expressed in form of fuzzy sets. The information about every entity of a mini- world can be expressed simply by fuzzy sets and then more effectively processed.

In this case, any non-fuzzy value of a basic set /universe of discourse/ is assigned by a grade of membership and has a form $(1/u_i)$. Every value or fuzzy term of Fig. 3 is considered as a unique name of a corresponding fuzzy set. In /M. Umano, 79/ the author had given a method for data storage and processing in fuzzy database. The users requests are processed by 52 fuzzy sets operations. In this subsection we introduce a new method for processing of fuzzy data and fuzzy queries by fuzzy relation algebra. When no confusion occurs, $\underset{\sim}{U}$ is used to denote set of all fuzzy sets over a basic set U of an attribute. A fuzzy relation over an attribute set $\{A_1,\ldots,A_n\}$ /denoted by $FR(A_1,\ldots,A_n)$ with $U_i = U(A_i))$ and a fuzzy database are defined as follows.

Definition 3.5

A fuzzy database D_F is a collection of fuzzy relations

$$D_F = \{FR_1,\ldots,FR_n\},$$

where every fuzzy relation of fuzzy sets $\underset{\sim}{U}_1,\ldots,\underset{\sim}{U}_{n_i}$ in $U_1,\ldots,U_{n_i}$ is defined by a membership function

$$\mu_{FR}: \underset{\sim}{U}_1 \times\ldots\times \underset{\sim}{U}_{n_i} \rightarrow [0,1].$$

In other words, a fuzzy relation of the sets $\underset{\sim}{U}_1,\ldots,\underset{\sim}{U}_{n_i}$ is a subset of Cartesian product

$$FR \subseteq \underset{\sim}{U}_1 \times\ldots\times \underset{\sim}{U}_{n_i}$$

where U_i is determined on the universe of discourse of attribute A_i.

Remark

The Cartesian product was defined by /L.A. Zadeh, 75/

$$\mu: \underset{\sim}{U}_1 \times\ldots\times \underset{\sim}{U}_{n_i} \rightarrow [0,1]$$

where $\mu_{U_1 \times \ldots \times U_{n_i}}(r) = \bigwedge_{j=1}^{n_i} \mu_{U_j}(r_j)$, $r = (r_1, \ldots, r_{n_i}) \in \underset{\sim}{U}_1 \times \ldots \times \underset{\sim}{U}_{n_i}$,

$r_j \in \underset{\sim}{U}_j$.

To illustrate a fuzzy relation of a fuzzy database in table form we add a special attribute name (in the sense of M. Umano, 79/. Below we introduce some fuzzy relational operations.

a. <u>Fuzzy projection</u>

Definition 3.6

Let FR be a fuzzy relation in ATTR(FR). A fuzzy projection of a fuzzy relation FR over a subset $X = \{A_1, \ldots, A_k\} \subseteq$ ATTR(FR) /denoted by π_X(FR)) is defined by

$$\pi_X(FR) = FR[X] = \{ (\bigvee_{\bar{X}} \mu_{FR}(r)/r(A_1, \ldots, A_k)) = \bigvee_{\bar{X}} \mu_{FR}(r)/(r[A_1], \ldots, r[A_k])$$

$$: \bar{X} = ATTR(FR) \setminus X \wedge A_i \in X \wedge \mu_{FR}(r)/r \in FR\}.$$

The fuzzy projection of a fuzzy relation is defined similarly as a projection of an extended relation. The duplicate tuples are identified and assigned the maximum μ- value.

We illustrate this Definition 3.6 with the following simple example.

Example

We consider two fuzzy relations

FR1 (μ_1	A_1	A_2	A_3)
0.5	a	A	1
0.7	a	A	2
1.0	b	B	1
0.6	b	A	2

Fig. 4: FR1-fuzzy relation

FR2 (μ_2	A_1	A_2	A_4)
0.6	a	A	x
0.8	a	A	y
0.9	b	A	y
0.9	b	B	x

Fig. 5: FR2-fuzzy relation

Let $X = \{A_1, A_2\} \subseteq ATTR(FR_1)$

$\pi_X(FR1) =$

(μ_1	A_1	A_2)
0.7	a	A
0.6	b	A
1.0	b	B

b. <u>Fuzzy join</u>

The definition of fuzzy join of the fuzzy relations is represented as follows.

Definition 3.7

Let FR and FS be two fuzzy relations in attribute set ATTR(FR) and ATTR(FS), respectively. $A \in ATTR(FR)$ and $B \in ATTR(FS)$ are two comparable attributes. A fuzzy equi-join of FR and FS on A and B is defined by

$$FR[A=B]FS = \{\mu/(w,s):\ w[A] = s[B] \wedge \mu_{FR}(w)/w \in FR \wedge \mu_{FS}(s)/s \in FS \wedge \mu = \mu_{FR}(w) \wedge \mu_{FS}(s)\}.$$

The fuzzy sets w[A] and s[B] are compared by the rules in Definition 2.2. If A and B are the some attribute names then either A or B is deleted from result. This join is called natural and denoted by * . The Definition 3.7 can be generalized with the same way for a fuzzy join on a common attribute set $ATTR(FR) \cap ATTR(FS)$.

Example

/See Fig. 4 and 5/.

FR1 * FR2 =

(μ	A_1	A_2	A_3	A_4)
0.5	a	A	1	x
0.5	a	A	1	y
0.6	a	A	2	x
0.7	a	A	2	y
0.9	b	B	1	x
0.6	b	A	2	y

c. Fuzzy selection

Based on Definition 2.2, a fuzzy selection from a fuzzy relation is defined as follows.

Definition 3.8

Let FR be a fuzzy relation, A be an attribute of ATTR(FR) and c be an element of $\underset{\sim}{U}(A)$. A fuzzy selection A = c (denoted by $\sigma_{A=c}(FR)$) of relation FR is defined by

$$\sigma_{A=c}(FR) = \{\mu_{FR}(r)/r: \mu_{FR}(r)/r \in FR \wedge c \in \underset{\sim}{U}(A) \wedge$$

$$(\exists u_i \in U(A): \mu_{c \cap r[A]}(u_i) \geq \tau \wedge \tau \in [0,1])\}.$$

Example (see Fig. 4)

$\sigma_{A_1=a}(FR1) =$

(μ_1	A_1	A_2	A_3)
0.5	a	A	1
0.7	a	A	2

There are two tuples satisfy the condition A_1=a.

Remark

Other set-operations of the fuzzy relations can be defined based on the Definition 2.2. In this case it must be assumed that the fuzzy relations must have the same attribute set and every fuzzy relation is considered as a fuzzy set (level-2 fuzzy set).

Let FR1 and FR2 be two fuzzy relations on $\underset{\sim}{U} = \underset{\sim}{U}_1 \times \ldots \times \underset{\sim}{U}_n$ where

$$FRj = \{\mu_{FRj}(\underset{\sim}{u}_1, \ldots, \underset{\sim}{u}_n)/(\underset{\sim}{u}_1, \ldots, \underset{\sim}{u}_n): \underset{\sim}{u}_i \in \underset{\sim}{U}_i, \quad i=1,\ldots,n\}$$

$$= \{\mu_{FRj}(\underset{\sim}{u})/\underset{\sim}{u}: \underset{\sim}{u} \in \underset{\sim}{U}\}, \quad j=1,2.$$

Briefly we represent this set operation in the following Fig. 6.

Name	Notation	Representation
fuzzy union	$FR1 \cup FR2$	$FR1 \cup FR2 = \{\mu_{FR1}(\underset{\sim}{u}) \vee \mu_{FR2}(\underset{\sim}{u}) / \underset{\sim}{u}: \underset{\sim}{u} \in \underset{\sim}{U}\}$
fuzzy intersection	$FR1 \cap FR2$	$FR1 \cap FR2 = \{\mu_{FR1}(\underset{\sim}{u}) \wedge \mu_{FR2}(\underset{\sim}{u}) / \underset{\sim}{u}: \underset{\sim}{u} \in \underset{\sim}{U}\}$
difference (minus)	$FR1 \ominus FR2$	$FR1 \ominus FR2 = \{(\mu_{FR1}(\underset{\sim}{u}) - \mu_{FR2}(\underset{\sim}{u})) \vee 0 / \underset{\sim}{u}: \underset{\sim}{u} \in \underset{\sim}{U}\}$
Cartesian product	$FR1 \otimes FR2$	$FR1 \otimes FR2 = \{\mu_{FR1}(\underset{\sim}{u}) \wedge \mu_{FR2}(\underset{\sim}{v}) / (\underset{\sim}{u}\ \underset{\sim}{v}): \underset{\sim}{u} \in \underset{\sim}{U},\ \underset{\sim}{v} \in \underset{\sim}{V}\}$
		$= \{\mu_{FR1}(\underset{\sim}{u}_1, \ldots, \underset{\sim}{u}_n) \wedge \ _{FR2}(\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_n) / (\underset{\sim}{u}_1, \ldots, \underset{\sim}{v}_n)$
		$: \underset{\sim}{u}_i \in \underset{\sim}{U}_i \ \ \underset{\sim}{v}_j \in \underset{\sim}{V}_j\}$

Fig. 6: Set operations on fuzzy relation

4. APPLICATIONS OF FUZZY RELATIONAL EXPRESSIONS TO FUZZY RELATIONS BY EXAMPLES OF SEQUEL-2

In this section we shall formulate some applications of fuzzy relational expressions to fuzzy relations which is an extended version of Codd's model.

The retrieval from fuzzy database is performed on the basis of relational algebraic operations. In the general case, the representation of some user's request in the form of a relational expression is difficult. Here we would like to limit our observation only to a queries class which can be formulated nonprocedurally in terms of fuzzy relational expression, namely in terms of the SEQUEL-language /see D.D. Chamberlin et al. 78/.

Retrieval method

The user's request is formulated in the form of a fuzzy relational expression. (i.e. of the form {t| (t)}, see Ullman/.

Any predicate /atom/ of a fuzzy relational expression contains fuzzy sets as constants /with their unique names in the representation/ and fuzzy equality as predicate operator.

The predicates are evaluated with respect to a tuple of a fuzzy relation by a two-valued logic. If a predicate satisfies the conditions, then it obtains a truth-value 1. If it doesn't, there it obtains a truth-value 0.

The predicates are connected by the Boolean operations OR, AND and NOT.

As for evaluation of fuzzy relational expression, the predicates are evaluated ranging over all tuples of the relations and if the predicates are true then the tuple constucted from the components corresponding to target list of the expression.

The grade of membership of any tuple in a fuzzy relation is computed and processed basing on Definition 3.6-Definition 3.8.

Remark

The predicates introduced here are evaluated by a two-valued logic. In the general case, they could be considered as fuzzy predicates which are evaluated not only by two-valued logic but also by fuzzy logic /see Le Tien Vuong and Ho/.

Example

Here we shall consider as example a fuzzy database which consists of two fuzzy relations F-EMP and F-DEPT /see fig. 7 and 8/.

F-EMP(μ	MNO	NAME	AGE	DNO	SAL)
0.8	101	A	20	10	very high
0.9	102	B	25	11	high
0.8	103	A	more or less 20	10	1000
0.9	104	D	young	12	2000
0.8	105	D	old	12	1500
1.0	106	B	50	11	500

Fig. 7: F-EMP relation

F-DEPT (μ	DNO	DNAME	LOC
1.0	10	x	London
1.0	11	y	London
1.0	12	z	Manchester
1.0	13	w	Liverpool

Fig. 8: F-DEPT relation

Relation F-EMP describes a set of employees, giving the employee person ordinal number, name, age, department ordinal number and salary for each employee. The F-DEPT relation consists of department ordinal number, name and location of each department. The attribute name shows the grade of membership of each relation tuple.

Some fuzzy sets in Fig. 7 and 8 can be defined basing upon the techniques developed by /L.A. Zadeh, 75/ as follows.

more or less 20 := {0.5/19, 1.0/20, 0.6/21},

$$\text{young }(x) = \begin{cases} 1 \text{ for } x \leq 24 \\ 0.5 \text{ for } 25 \leq x \leq 30 \\ 0 \text{ otherwise} \end{cases}$$

$$\text{old }(x) = \begin{cases} 1 \text{ for } x \geq 60 \\ 0.5 \text{ for } 55 \leq x < 60 \\ 0 \text{ otherwise} \end{cases}$$

$$\text{high }(x) = \begin{cases} 1 \text{ for } x \geq 1800 \\ 0.6 \text{ for } 1500 \leq x < 1800 \\ 0 \text{ otherwise} \end{cases}$$

$$\text{very high }(x) = [\text{high }(x)]^2 .$$

The following examples will be represented in terms of SEQUEL 2.

Q1. List of the names of employees who are 20 yr old.

```
SELECT UNIQUE NAME
FROM F-EMP
WHERE AGE = 20
```

The selection AGE = 20 is first executed from F-EMP relation. The following tuples satisfy this condition

(0.8	101	A	20		10	very high)
(0.8	103	A	more or less 20		10	1000)
(0.9	104	D	young		12	2000).

Then the projection on the NAME attribute is executed with eliminating the duplicate tuples. The last result is

{0.8/A , 0.9/D}.

Q2. List of names and employee person ordinal number who young and have a high salary:

```
SELECT MNO, NAME
FROM F-EMP
WHERE AGE = young AND SAL = high
```

The first, second and fourth tuple of F-EMP relation are satisfy the selection conditions. .
After the projection is executed on NAME and MNO we have the result:

{0.8/(101,A), 0.9/(102,B), 0.9/(104,D)}

Q3. List the departments which have no employees.

```
SELECT DNO
FROM F-DEPT
MINUS
SELECT DNO
FROM F-EMP
```

The first projection is performed on attribute DNO from F-DEPT relation. The result is {1.0/10, 1.0/11, 1.0/12, 1.0/13}. The second projection is performed on DNO from F-EMP relation an we obtain following tuples {0.8/10, 1.0/11, 0.9/12}. Now the minus operations /see Fig. 6/ are executed and we have the following result /for $\tau = 0.5$/

{1.0/13}.

Q4. List the names of all employees and the locations where they work.

```
SELECT UNIQUE F-EMP.NAME,F-DEPT.LOC
FROM F-EMP,F-DEPT
WHERE F-EMP.DNO = F-DEPT.DNO
```

The join operation is first executed on attribute DNO of F-EMP relation and F-DEPT relation. We have the following result

(0.8	101	A	20	10	very hight	x	London)
(0.8	103	A	more or less 20	10	1000	x	London)
(0.9	102	B	25	11	high	y	London)
(1.0	106	B	50	11	500	y	London)
(0.9	104	D	young	12	2000	z	Manchester)
(0.8	105	D	old	12	1500	z	Manchester)

The projection (with eliminating of duplicate tuples) is executed on F-EMP.NAME and F-DEPT.LOC. We have the following result

{0.8/(A,London), 1.0/(B,London), 0.9/ (D,Manchester)}

5. CONCLUSION

In this paper the Codd's relational model of data is extended using fuzzy sets theory. We propose fuzzy relational model in which fuzzy data are represented by fuzzy relations.

We have introduced some fuzzy relational operations, namely, projection, join, selection, union, intersection, difference... . The retrieval method for such a fuzzy database is presented in detail by several examples. This method is based on fuzzy relational algebra. By generalizing appropriately comparison operations between fuzzy sets, we hope that the processing capacities of a relational fuzzy database can be more powerful. That is the aim of our subsequent paper.

ACKNOWLEDGEMENT

The authors wish to thank Prof. N.D. Ngoc for his valuable remarks and suggestions.

REFERENCES

BELKE 78 W. BELKE, et...: Zur Nutzung unscharfer Mengen im formal System der Klassifizierung. TU-Infor. 08-02-78, Dresden.

CODD 70 E.F. Codd: A relational model of data for large shared data banks, C.ACM 13,6(1970), 377-387.

CHAMBERLIN et al. 78: D.D. Chamberlin et al.: SEQUEL-2: A unified approach to data definition, manipulation and control, IBM-Jour., Res. and Dec, 20,6(1978) 560-575.

LE 83 Le Tien Vuong: Untersuchung zur ternaeren Dekomposition einer Relation und zur Anwendung der unscharfen Menge in CRM, Diss. TU-Dresden (1983).

LE 85 Le Tien Vuong: On the applications of fuzzy sets theory to relational database, J. Computer Sc., and Cyb. Vol. 1, N.4 (1985) (in vietnamese).

LE and HO in preparation

NEGOITA et al. 75: C.V. Negoita, et al.: Applications of fuzzy sets to systems analysis, Birkhhauser Verlag, Basel und Stuttgart, 1975.

TAHANI 77 V. Tahani: A conceptual framework for fuzzy queries processing- a step toward very intelligent database system, Inf.Proc. Manag. 13(1977) 289-303.

UMANO 79 M. Umano: Representation and manipulation of fuzzy data, Diss., Osaka Univ. Japan 2. 1979.

ZADEH 75 L.A. Zadeh: The concept of a linguistic variable and its application to approximate reasoning, Inf. Sc. Vol. 8, 199-248, 301-357 (1975).

ULLMAN Principles of database systems, second edition, Computer Science Press (1982).

Retrieval from fuzzy database by fuzzy relational algebra

Le Tien Vuong, Ho Thuan

Summary

In this paper a fuzzy relational model for fuzzy data - considered as an extended version of Codd's model-is developed.

A method for fuzzy data and queries processing is introduced, based on the fuzzy relational operations, namely, projection, join, selection, union, intersection.

These are illustrated by several examples and their applications to fuzzy database are presented in detail by the sublanguage SEQUEL-2.

Fuzzi adathalmazból való visszakeresés fuzzi relációs algebra segitségével

Le Tien Vuong, Ho Thuan

Összefoglaló

A cikkben a szerzők a Codd-féle relációs modellnek egy kiterjesztését ismertetik, amelyben "fazzi" /"fuzzy"/ adatok is szerepelhetnek. A kiterjesztés az un. fazzi /"fuzzy"/ relációs adatmodell. Bevezetnek több "fazzi" relációs müveletet /vetités, összekapcsolás, kiválasztás, unió, metszet/. A mondottakat több egyszerü példán szemléltetik, valamint az un. SEQUEL-2 nyelv segitségével is leirják ezeket a példákat.