

A Type-2 Fuzzy Relational Database Model

Nguyen Hoa

Faculty of Information Technology
Saigon University
nguyenhhoa@sgu.edu.vn

Abstract: This paper introduces a type-2 fuzzy relational database model (T-2FRDB) as an extension of the type-1 fuzzy relational database models with a full set of basic fuzzy relational algebraic operations that can represent and query uncertain and imprecise information in the real world applications. In this model, the membership degree of tuples in a fuzzy relation is represented by fuzzy numbers on $[0, 1]$, the fuzzy relational algebraic operations are defined by using the extension principle for computing the minimum and maximum values of such fuzzy numbers. Some properties of the type-2 fuzzy relational algebraic operations in T-2FRDB are also formulated and proven as extensions of those of fuzzy relational algebraic operations in the type-1 fuzzy relational database models.

Keywords: Fuzzy set; fuzzy relation; type-2 fuzzy relational database; type-2 fuzzy relational algebraic operation.

I. INTRODUCTION

As we know, the classical relational database model is very useful for modeling, designing and implementing large-scale systems. However, it is restricted for representing and handling uncertain and imprecise information of objects in practice ([2], [4]). For example, applications of the classical relational database model cannot deal with the query “find all patients who are *young* and have lung cancer and a *high* treatment cost”, where *young* and *high* are the vague notion and imprecise value ([7], [18]).

So far, there have been many relational database models studied and built ([1], [3], [6], [8], [13], [15], [17]), based on the fuzzy theory for modeling objects about which information may be vague and imprecise to overcome the limitations of the classical relational database model. Such models are called *fuzzy relational database models*.

There are two main approaches to represent fuzzy relations in fuzzy relational database models: (1) representing each fuzzy relation as a set of tuples whose each attribute may take a fuzzy set (or a possibility distribution is inferred from a fuzzy

set) ([1], [3], [6], [9], [14], [16]), whereby the membership degree of tuples for the relation is hidden in that of their attribute values; (2) representing each fuzzy relation as a fuzzy set of tuples whose each attribute only takes a single and precise value ([5], [7], [8], [10]–[13], [15], [17]), whereby the membership degree of tuples for the relation also is that for the fuzzy set.

Fuzzy relational database models that are built based on one of above approaches are extensions of classical relational database model with fuzzy sets. They have different capabilities for expressing and dealing with uncertain and imprecise information.

There are two types of the models based on the second approach, called the *type-1 fuzzy relational database model* (T-1FRDB), or the (ordinary) fuzzy relational database model, whereby the membership degree of tuples is assigned to a number in $[0, 1]$, and the *type-2 fuzzy relational database model* (T-2FRDB), whereby the membership degree of tuples is expressed as a fuzzy number on $[0, 1]$. Many type-1 fuzzy relational database models have been proposed ([7], [8], [11]–[13], [15]). However, it is due to the membership degree of tuples is expressed as a number in $[0, 1]$, these models were restricted in representing the associated imprecise degree of attribute values. In the real world relational databases, since attribute values of tuples may be imprecise, there are many situations in which we do not know exactly the membership degree of tuples as a number in $[0, 1]$ but only can estimate it as an approximate number (or a fuzzy number) on $[0, 1]$. Some type-2 fuzzy relational database models have been introduced to overcome the shortcoming of type-1 fuzzy relational database models ([5], [10], [17]). However, in [10], the data representative notions were not defined really formally and some fuzzy relational algebraic operations were missing. Also, in ([5], [17]), the set of basic fuzzy relational algebraic operations was not complete. Thus, the abilities of expressing and dealing with imprecise information of those models were limited.

In this paper, we propose a type-2 fuzzy relational database

model (T-2FRDB) as an extension of type-1 fuzzy relational database models to overcome the shortcomings of the models in ([5], [10], [17]). In our T-2FRDB, the notions of the data representative model are completely defined formally, the full set of basic type-2 fuzzy relational algebraic operations is built based on the extension principle for computing the minimum and maximum values of fuzzy numbers. Some properties of these algebraic operations are also formulated as theorems and proven coherently. T-2FRDB allows for representing soft queries associated with fuzzy sets to handle imprecise information in practice.

The basis of mathematics to develop T-2FRDB is presented in Section II. Schemas and relations of T-2FRDB are introduced in Section III. Section IV and V present type-2 fuzzy relational algebraic operations and their properties in T-2FRDB. Finally, Section VI concludes the paper and outlines further research directions in the future.

II. FUZZY SETS AND FUZZY RELATIONS

In this section, we present some notions about fuzzy sets and fuzzy relations as the mathematical basis for developing T-2FRDB model. Fuzzy sets are used to represent and execute soft queries meanwhile fuzzy relations to define relations in T-2FRDB.

A. Fuzzy Sets

For a classical set, an element is to be or not to be in the set or, in other words, the membership degree of an element in the set is binary. For a fuzzy set, the membership degree of an element in the set is expressed by a real number in the interval $[0, 1]$. The fuzzy set is extended from the classical set as in [18] and is defined as follows.

Definition 1. A fuzzy set A on a domain X is defined by a membership function μ_A from X to the closed interval $[0, 1]$. For each $x \in X$, $\mu_A(x)$ is the membership degree of x for A .

We note that a classical set A on X also is a fuzzy set [7] with the membership function $\mu_A(x) = 1, \forall x \in A$ and $\mu_A(x) = 0, \forall x \notin A$ or even a element e in X is also considered as a special fuzzy set on X with the membership function $\mu_e(e) = 1$ and $\mu_e(x) = 0, \forall x \in X$ and $x \neq e$. The fuzzy set A on X as in the above definition is called the *ordinary fuzzy set* and can be denoted by $A = \{x: \mu_A(x) | x \in X\}$. In addition, the notation $A(x)$ can be used to replace $\mu_A(x)$.

The *support* of a fuzzy set A on X is the classical set that contains all the elements of X that have nonzero membership degrees in A . The *height* $h(A)$ of a fuzzy set A on X is the largest membership degree obtained by any element in that set. It means $h(A) = \sup_{x \in X} A(x)$. A fuzzy set A is called *normal* when $h(A) = 1$ and it is called *subnormal* when $h(A) < 1$. An α -cut of

a fuzzy set A is the classical set ${}^{\alpha}A = \{x | A(x) \geq \alpha\}$. A *strong α -cut* of a fuzzy set A is the classical set ${}^{\alpha+}A = \{x | A(x) > \alpha\}$.

The operations on fuzzy sets are defined in general based on functions from the Cartesian product of the closed intervals $[0, 1]$ to the closed interval $[0, 1]$. However, the section only presents the standard operations in [18] which are applied in computing the relations of T-2FRDB.

Definition 2. Let A, B be two fuzzy sets on X and have the membership functions μ_A, μ_B , respectively. The *complement of A , union, intersection and difference of A and B* are defined by their membership functions as follows.

1. $\mu_{A^c}(x) = 1 - \mu_A(x)$,
 2. $\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x))$,
 3. $\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x))$,
 4. $\mu_{A - B}(x) = \min(\mu_A(x), 1 - \mu_B(x))$,
- $\forall x \in X$.

The fuzzy numbers are special fuzzy sets that are used to represent the fuzzy relations in T-2FRDB. The fuzzy numbers are defined in [7] as follows.

Definition 3. A fuzzy number A is a fuzzy set on the real number set R such that:

1. A is a normal fuzzy set.
2. ${}^{\alpha}A$ is a closed interval for every $\alpha \in (0, 1]$.
3. The support of A is bounded.

Example 1. The fuzzy set *about_0.5* that is given by the membership function and its graph as Figure 1 below is a fuzzy number.

$$\text{about_0.5}(x) = \begin{cases} 2x & \forall x \in [0, 0.5] \\ 2(1-x) & \forall x \in (0.5, 1] \\ 0 & \forall x \notin [0, 1] \end{cases}$$

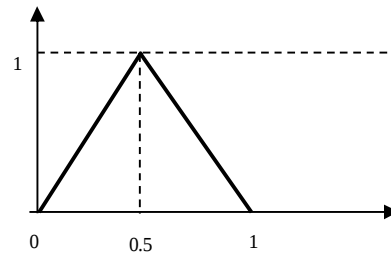


Fig. 1. Fuzzy number *about_0.5*

For computing and combining the membership degrees of tuples in the type-2 fuzzy relational algebraic operations, we use the operations MIN and MAX to determine the minimum and maximum values of fuzzy numbers. The operations MIN and MAX are defined by the extension principle in [7] as below.

Definition 4. Let A and B be two fuzzy numbers. The *minimum*

and *maximum values* of A and B are fuzzy numbers that are defined by:

1. $\text{MIN}(A, B)(z) = \sup_{z=\min(x,y)} \min[A(x), B(y)]$,
 2. $\text{MAX}(A, B)(z) = \sup_{z=\max(x,y)} \min[A(x), B(y)]$
- $\forall z \in \mathbb{R}$.

Example 2. Let $A = \{1: 1, 0.9: 0.8, 0.8: 0.3\}$ and $B = \{0.6: 0.3, 0.5: 1, 0.4: 0.4\}$ be two fuzzy sets, then $\text{MIN}(A, B) = \{0.6: 0.3, 0.5: 1, 0.4: 0.4\}$.

The fuzzy set of type 2 is extended from the ordinary fuzzy set in [7] as below.

Definition 5. Let $\mathfrak{F}([0, 1])$ be the set of all ordinary fuzzy sets on $[0, 1]$. A *type-2 fuzzy set* A on a domain X is defined by a membership function μ_A from X to $\mathfrak{F}([0, 1])$. For each $x \in X$, $\mu_A(x)$ is the membership degree of x for A .

We note that, it is due to each number in $[0, 1]$ to be considered as a special ordinary fuzzy set, so each ordinary fuzzy set is also considered as a special type-2 fuzzy set.

B. Fuzzy Relations

The ordinary fuzzy relation and type-2 fuzzy relation are the foundation for fuzzy relational database models. The fuzzy relation and type-2 fuzzy relation are defined in [7] by extending the notion of the classical relation based on fuzzy sets as follows.

Definition 6. Let $A_1, A_2, \dots, A_k$ be non-empty sets, an k -ary *ordinary fuzzy relation* R on k sets $A_1, A_2, \dots, A_k$ is an ordinary fuzzy set on the Cartesian product $A_1 \times A_2 \times \dots \times A_k$.

Definition 7. Let $A_1, A_2, \dots, A_k$ be non-empty sets, an k -ary *type-2 fuzzy relation* R on k sets $A_1, A_2, \dots, A_k$ is a type-2 fuzzy set on the Cartesian product $A_1 \times A_2 \times \dots \times A_k$.

We note that the ordinary fuzzy relation is also called *type-1 fuzzy relation*. The membership functions of the ordinary fuzzy relation and type-2 fuzzy relation are $\mu_R: A_1 \times A_2 \times \dots \times A_k \rightarrow [0, 1]$ and $\mu_R: A_1 \times A_2 \times \dots \times A_k \rightarrow \mathfrak{F}([0, 1])$, respectively.

III. T-2FRDB SCHEMAS AND RELATIONS

A. Type-2 Fuzzy Relational Schemas

A type-2 fuzzy relational schema in T-2FRDB consists of a set of attributes associated with a membership function of a type-2 fuzzy set, as the basis for determining type-2 fuzzy relations, and is defined as follows.

Definition 8. A *type-2 fuzzy relational schema* is a pair $R = (U, \mu)$, where

1. $U = \{A_1, A_2, \dots, A_k\}$ is a set of pairwise different attributes.

2. μ is a function that maps each $(v_1, v_2, \dots, v_k) \in D_1 \times D_2 \times \dots \times D_k$ to a fuzzy number on $[0, 1]$, where D_i is the domain of the attribute A_i ($i = 1, \dots, k$).

As in the classical relational database, the notations $R(U, \mu)$ and R can be used to replace $R = (U, \mu)$. In addition, each $t = (v_1, v_2, \dots, v_k)$ is called a tuple on the set of attributes $\{A_1, A_2, \dots, A_k\}$.

Example 3. A type-2 fuzzy relational schema **PATIENT** in T-2FRDB describing patients can be as follows.

PATIENT(P_NAME, AGE, DISEASE, D_COST, μ)
where $\mu: \text{string} \times \text{integer} \times \text{string} \times \text{real} \rightarrow \mathfrak{F}([0, 1])$, $\mathfrak{F}([0, 1])$ is the set of all fuzzy numbers on $[0, 1]$, **string**, **integer** and **real** are the domains of the attributes P_NAME, DISEASE, AGE and D_COST, respectively.

B. Type-2 Fuzzy Relations

The notion of the type-2 fuzzy relation in T-2FRDB is extended from that of the ordinary fuzzy relation in T-1FRDB as the following definition.

Definition 9. Let $U = \{A_1, A_2, \dots, A_k\}$ be a set of k pairwise different attributes. A *type-2 fuzzy relation* r over the type-2 fuzzy relational schema $R(U, \mu)$, is a finite set of tuples $\{t_1, t_2, \dots, t_n\}$ on the set of $\{A_1, A_2, \dots, A_k\}$ in which each tuple t_i is associated with the fuzzy number $\mu(t_i)$ representing the membership degree of t_i in r , for every $i = 1, 2, \dots, k$. The notation $t.A$ or $t[A]$ is used to denote the value of attribute A of tuple t in r . The membership degree of t_i in r is denoted by $\mu_r(t_i)$.

For each set of attributes $X \subseteq \{A_1, A_2, \dots, A_k\}$, the notation $t[X]$ is used to denote the rest of t after eliminating the value of attributes not belonging to X .

Note that, as in the classical database model, if we only care about a unique relation over a schema then we can unify its symbol name with its schema's name.

Example 4. A type-2 fuzzy relation over the schema **PATIENT** in Example 3 can be given as Table 1.

Table 1. Relation PATIENT

P_NAME	AGE	DISEASE	D_COST	μ
L.V. Tam	53	Lung cancer	350	0.9
L.T. Huong	65	Cirrhosis	40	about_0.5
N.T. Trang	29	Bronchitis	70	1.0
T.T. Tu	21	Hepatitis	30	high

In the relation, the attributes P_NAME, AGE, DISEASE, and D_COST describe the information about the name, age, disease and daily treatment cost of each patient, respectively. In reality, while diagnosing, the disease of each patient is not always determined certainly by the physicians. Similarly, the daily treatment cost for patients is also not known definitely even as the patients know about their diseases. Here, the

conventional unit for the treatment cost is 1000 (VND).

We note that, $\mu(t)$ represents the membership degree of each tuple t in the relation (by Definition 9), it means the precise degree of the information of the attribute values that is expressed by t . For example, consider the tuple t_1 (the first tuple) in the relation PATIENT, assuming that the information of the patient's name (*L.V. Tam*) expressed by t_1 is precise, then $\mu(t_1) = 0.9$ of t_1 represents the aggregated precise degree of the information about the age (53), diagnosed disease (*Lung cancer*) and daily treatment cost (350 thousands VND) of the patient. With t_4 , we do not know precisely both the information of the attribute values that it represents and its membership degree in the relation PATIENT. We are able to just estimate $\mu(t_4)$ is *high* where *high* = {0.5:0, 0.6:0.5, 0.7:0.8, 0.8:0.9, 0.9:1.0, 1:1.0} is a fuzzy number on [0, 1].

In the real world applications, fuzzy sets that represent the membership degrees of tuples in a fuzzy relation, such as *high* and *about_0.5* above, will be defined compatibly and consistently based on the meaning and precise degree of the information that these tuples express. The fuzzy sets *high* and *about_0.5* that are given as in this example only simply illustrates for Definition 9.

Now, the notion of a type-2 fuzzy relational database is defined as follows.

Definition 10. A type-2 fuzzy relational database over a set of attributes is a set of type-2 fuzzy relations corresponding with the set of their type-2 fuzzy relational schemas.

We note that when $\mu(t) \in [0, 1]$ for every tuple t in a type-2 fuzzy relation, this relation becomes a type-1 fuzzy relation. In other words, a type-1 fuzzy relational database is a particular case of an T-2 FRDB by Definition 10.

IV. SELECTION OPERATION ON T-2FRDB

A. Syntax of Selection Conditions

The syntax of selection conditions in T-2FRDB is extended from those in [11] with type-2 fuzzy relations as the following definition.

Definition 11. Let R be a schema in T-2FRDB and X be a set of relational tuple variables and θ be a binary relation from $\{=, \neq, \leq, <, >, \geq\}$. Then *selection conditions* are inductively defined and have one of the following forms:

1. $x.A \theta v$, where $x \in X$, A is an attribute in R and v is a precise value.
2. $x.A \rightarrow v$, where $x \in X$, A is an attribute in R , $\rightarrow$ a binary fuzzy relation and v fuzzy set value.
3. $x.A_1 \theta x.A_2$, where $x \in X$, A_1 and A_2 are two different attributes in R .
4. $\neg E$ if E a selection condition.

5. $E_1 \wedge E_2$ if E_1 and E_2 are selection conditions on the same relational tuple variable.
6. $E_1 \vee E_2$ if E_1 and E_2 are selection conditions on the same relational tuple variable.

Example 5. Consider the schema PATIENT in Example 4, the selection of “all patients who are *young* and get hepatitis” can be expressed by the selection condition $x.AGE \rightarrow \text{young} \wedge x.DISEASE = \text{hepatitis}$.

B. Semantics of Selection Conditions

The semantics of selection conditions is the satisfied degree measure of them for tuples in a fuzzy relation and is defined as below.

Definition 12. Let $R(U, \mu)$ be a fuzzy relational schema in T-2FRDB, r be a relation over R , x be a tuple variable and t be a tuple in r . The *interpretation* of selection conditions with respect to R , r and t , denoted by $Int_{R,r,t}$, is the partial mapping from the set of all selection conditions to the set of all fuzzy numbers on [0, 1] that is inductively defined as follows:

1. $Int_{R,r,t}(x.A \theta v) = \mu_r(t)$ if $t.A \theta v$ and $Int_{R,r,t}(x.A \theta v) = 0$ otherwise.
2. $Int_{R,r,t}(x.A \rightarrow v) = \text{MIN}(\mu_r(t), \mu_\varphi(t))$, with $\varphi = x.A \rightarrow v$.
3. $Int_{R,r,t}(x.A_1 \theta x.A_2) = \mu_r(t)$ if $t.A_1 \theta t.A_2$ and $Int_{R,r,t}(x.A_1 \theta x.A_2) = 0$ otherwise.
4. $Int_{R,r,t}(\neg E) = 1 - Int_{R,r,t}(E)$.
5. $Int_{R,r,t}(E_1 \wedge E_2) = \text{MIN}(Int_{R,r,t}(E_1), Int_{R,r,t}(E_2))$.
6. $Int_{R,r,t}(E_1 \vee E_2) = \text{MAX}(Int_{R,r,t}(E_1), Int_{R,r,t}(E_2))$.

We note that, v is a fuzzy set in $x.A \rightarrow v$, so $\varphi = x.A \rightarrow v$ is a binary fuzzy relation. Consequently, φ also is a fuzzy set. In particular, φ is the fuzzy set whose elements are tuples in r . For each $t \in r$, $\mu_\varphi(t) = v(t.A)$.

Intuitively, $Int_{R,r,t}(x.A \theta v)$ and $Int_{R,r,t}(x.A \rightarrow v)$ are respectively the satisfied degrees of the conditions $t.A \theta v$ and $t.A \rightarrow v$ for the tuple t in r while $Int_{R,r,t}(x.A_1 \theta x.A_2)$ is the satisfied degree of the condition $t.A_1 \theta t.A_2$ for the tuple t in r .

In the classical relational database model, for each tuple t and a relation r , $\mu_r(t) \in \{0, 1\}$, so the interpretation of selection conditions with respect to r and t always takes one of two values 0 or 1. This also means that the concept of the interpretation of selection conditions in the classical relational database model is a particular case of that of the interpretation of selection conditions in type-1 fuzzy relational database models and T-2FRDB.

Example 6. Let the fuzzy set *young* represent the young age of a patient whose membership function is defined by

$$\mu_{\text{young}}(x) = \begin{cases} 1, & \forall x \in [0, 20] \\ (35 - x)/15, & \forall x \in (20, 35), \\ 0, & \forall x \geq 35 \end{cases}$$

then the interpretation of the selection conditions $E_1 = \text{"x.AGE} \rightarrow \text{young}"$, $E_2 = \text{"x.DISEASE} = \text{hepatitis}"$, and $E = \text{"x.AGE} \rightarrow \text{young} \wedge \text{x.DISEASE} = \text{hepatitis}"$ with respect to $r = \text{PATIENT}$ and t_4 (the fourth tuple) in Example 4 are

$$\text{Int}_{R,r,t_4}(E_1) = \text{MIN}(\mu_r(t_4), \text{young}(21)) = \text{MIN}(\text{high}, 0.93) = \{0.5:0, 0.6:0.5, 0.7:0.8, 0.8:0.9, 0.9:1.0, 0.93:1.0\},$$

$$\text{Int}_{R,r,t_4}(E_2) = \mu_r(t_4) = \text{high}, \text{ and}$$

$$\begin{aligned} \text{Int}_{R,r,t_4}(E) &= \text{Int}_{R,r,t_4}(\text{x.AGE} \rightarrow \text{young} \wedge \text{x.DISEASE} = \text{hepatitis}) \\ &= \text{MIN}(\text{Int}_{R,r,t_4}(E_1), \text{Int}_{R,r,t_4}(E_2)) = \text{MIN}(\{0.5:0, 0.6:0.5, 0.7:0.8, 0.8:0.9, 0.9:1.0, 0.93:1.0\}, \text{high}) = \{0.5:0, 0.6:0.5, 0.7:0.8, 0.8:0.9, 0.9:1.0, 0.93:1.0\}. \end{aligned}$$

Let $\text{approx_0.9} = \{0.5:0, 0.6:0.5, 0.7:0.8, 0.8:0.9, 0.9:1.0, 0.93:1.0\}$, we have $\text{Int}_{R,r,t_4}(E) = \text{approx_0.9}$.

Now, the selection operation in T-2FRDB is extended from the selection operation in [11] as follows.

Definition 13. Let $R(U, \mu)$ be a relational schema in T-2FRDB, r be a type-2 fuzzy relation over R and ϕ be a selection condition. The *selection* on r with respect to ϕ , denoted by $\sigma_\phi(r)$, is the type-2 fuzzy relation r' over R , including all tuples defined by

$$r' = \{t \in r \mid \text{Int}_{R,r,t}(\phi) \neq 0 \wedge \mu_r(t) = \text{Int}_{R,r,t}(\phi)\}.$$

We note that the number 0 in $\text{Int}_{R,r,t}(\phi) \neq 0$ also is the fuzzy number 0 on $[0, 1]$.

Example 7. Consider the relation $r = \text{PATIENT}$ in Example 4, the query “Find all patients who are young and have hepatitis” can be done by the selection operation $r' = \sigma_\phi(\text{PATIENT})$ with $\phi = \text{"x.AGE} \rightarrow \text{young} \wedge \text{x.DISEASE} = \text{hepatitis}"$.

The selection is implemented by checking the satisfaction of all tuples in PATIENT for the selection condition ϕ . From the result computed in Example 6 above, we can see only the tuple t_4 satisfies ϕ with the value of membership function being $\text{approx_0.9} = \{0.5:0, 0.6:0.5, 0.7:0.8, 0.8:0.9, 0.9:1.0, 0.93:1.0\}$. Therefore, the result of the selection is the relation $r' = \sigma_\phi(\text{PATIENT})$ as in Table 2.

Table 2. Relation $\sigma_\phi(\text{PATIENT})$

P_NAME	AGE	DISEASE	D_COST	μ
T. T. Tu	21	Hepatitis	30	approx_0.9

V. OTHER OPERATIONS ON T-2FRDB

As for the classical relational database and T-1FRDB, other basic relational algebraic operations on T-2FRDB are the projection, Cartesian product, join, intersection, union, and difference. We now extend those operations of T-1FRDB for T-2FRDB taking into account the fuzzy membership degree of tuples in relations.

A. Projection

A projection of a type-2 fuzzy relation on a set of attributes is a new type-2 fuzzy relation in which only the attributes in that set are considered for each tuple of the new relation as the following definition.

Definition 14. Let $R = (U, \mu)$ be a type-2 fuzzy relational schema, r be a relation over R and $L = \{A_1, A_2, \dots, A_k\}$ be a subset of U . The *projection* of r on L , denoted by $\Pi_L(r)$, is the type-2 fuzzy relation r' over the schema R' , determined by:

1. $R' = (L, \mu')$, where μ' is the mapping from $D_1 \times D_2 \times \dots \times D_k$ to the set of all fuzzy numbers on $[0, 1]$, D_i is the value domain of A_i ($i = 1, \dots, k$).
2. $r' = \{t' = t[L] \mid t \in r \text{ and } \mu'_{r'}(t') = \text{MAX}_{t \in r} \{\mu_r(t) \mid t' = t[L]\}\}$.

Example 8. The projection of the relation PATIENT in Table 1 on $L = \{P_NAME, DISEASE\}$ is the relation $\Pi_L(\text{PATIENT})$ as in Table 3.

Table 3. Relation $\Pi_L(\text{PATIENT})$

P_NAME	DISEASE	μ'
L. V. Tam	Lung cancer	0.9
L. T. Huong	Cirrhosis	about_0.5
N. T. Trang	Bronchitis	1.0
T. T. Tu	Hepatitis	high

B. Cartesian Product

For the Cartesian product of two type-2 fuzzy relations, as in the classical relational database and T-1FRDB, we assume the sets of attributes of their schemas are disjoint as the definition below.

Definition 15. Let U_1, U_2 be two sets of attributes that have not any common element and r_1, r_2 be two fuzzy relations over two type-2 fuzzy relational schemas $R_1 = (U_1, \mu_1)$ and $R_2 = (U_2, \mu_2)$, respectively. The *Cartesian product* of r_1 and r_2 , denoted by $r_1 \times r_2$, is the type-2 fuzzy relation r over R , determined by:

1. $R = (U, \mu)$, where $U = U_1 \cup U_2$, μ is the mapping from $D_1 \times D_2 \times \dots \times D_{k+m}$ to the set of all fuzzy numbers on $[0, 1]$, $k = |U_1|$, $m = |U_2|$, D_i is the value domain of $A_i \in U_1 \cup U_2$.
2. $r = \{t = (v_1, v_2, \dots, v_k, v_{k+1}, v_{k+2}, \dots, v_{k+m}) \mid t_1 = (v_1, v_2, \dots, v_k), t_2 = (v_{k+1}, v_{k+2}, \dots, v_{k+m}), t_1 \in r_1, t_2 \in r_2 \text{ and } \mu_r(t) = \text{MIN}(\mu_{1r_1}(t_1), \mu_{2r_2}(t_2))\}$.

C. Join

The join of two fuzzy relations in T-2FRDB is extended from the natural join of two relations in the classical relational database and the join in T-1FRDB. The join in T-2FRDB is defined as follows.

Definition 16. Let U_1 and U_2 be two sets of attributes such that the value domains of two attributes of the same name A in U_1

and U_2 , respectively are identical. Let r_1, r_2 be two fuzzy relations over the type-2 fuzzy relational schemas $R_1 = (U_1, \mu_1)$ and $R_2 = (U_2, \mu_2)$, respectively and $\{A_k, \dots, A_l\} = U_1 \cap U_2$. The natural join of r_1 and r_2 , denoted by $r_1 \bowtie r_2$, is the type-2 fuzzy relation r over the schema R , determined by:

1. $R = (U, \mu)$, where $U = U_1 \cup U_2$, μ is the mapping from $D_1 \times D_2 \times \dots \times D_n$ to the set of all fuzzy numbers on $[0, 1]$, $n = |U|$, D_i is the value domain of $A_i \in U_1 \cup U_2$.
2. $r = \{t = (v_1, \dots, v_j, v_{k_1}, \dots, v_{k_l}, v_{m_1}, \dots, v_n) \mid t_1 = (v_1, \dots, v_j, v_{k_1}, \dots, v_{k_l}), t_2 = (v_{k_1}, \dots, v_{k_l}, v_{m_1}, \dots, v_n), t_1 \in r_1, t_2 \in r_2 \text{ such that } v_k = t_1[A_k] = t_2[A_k], \dots, v_l = t_1[A_l] = t_2[A_l] \text{ and } \mu_r(t) = \min(\mu_{r_1}(t_1), \mu_{r_2}(t_2))\}$.

Example 9. Let $U_1 = \{P_ID, DISEASE\}$, $U_2 = \{P_NAME, DISEASE\}$ be two sets of attributes and $PATIENT_1, PATIENT_2$ be two type-2 fuzzy relations over two schemas $R_1 = (U_1, \mu_1)$ and $R_2 = (U_2, \mu_2)$ respectively as Tables 4 and 5. It is easy to see that $\min(\{1: 1, 0.9: 0.8, 0.8: 0.3\}, \{0.6: 0.3, 0.5: 1, 0.4: 0.4\}) = \{0.6: 0.3, 0.5: 1, 0.4: 0.4\}$. So, the result of the join of $PATIENT_1$ and $PATIENT_2$ is the relation $PATIENT$ over the schema $R = (U_1 \cup U_2, \mu)$ computed as in Table 6.

 Table 4. Relation $PATIENT_1$

P_ID	DISEASE	μ_1
PT005	Bronchitis	0.8
PT006	Gall-stone	$\{1: 1, 0.9: 0.8, 0.8: 0.3\}$

 Table 5. Relation $PATIENT_2$

P_NAME	DISEASE	μ_2
L. V. Lan	Bronchitis	0.9
N.T.Thu	Gall-stone	$\{0.6: 0.3, 0.5: 1, 0.4: 0.4\}$

 Table 6. Relation $PATIENT = PATIENT_1 \bowtie PATIENT_2$

P_ID	P_NAME	DISEASE	μ
PT005	L. V. Lan	Bronchitis	0.8
PT006	N.T.Thu	Gall-stone	$\{0.6: 0.3, 0.5: 1, 0.4: 0.4\}$

D. Intersection, Union, and Difference

By extending the operations of ordinary fuzzy sets in Definition 2, the set operations on the type-2 fuzzy relations in T-2FRDB are defined in turn as below.

Definition 17. Let r_1 and r_2 be two type-2 fuzzy relations over the same schema $R(U, \mu)$. The intersection of r_1 and r_2 , denoted by $r_1 \cap r_2$, is the type-2 fuzzy relation r over R including tuples t 's defined by

$$r \cap s = \{t \mid \mu_{r \cap s}(t) = \min(\mu_r(t), \mu_s(t))\}.$$

Definition 18. Let r_1 and r_2 be two type-2 fuzzy relations over the same schema $R(U, \mu)$. The union of r_1 and r_2 , denoted by $r \cup s$, is the type-2 fuzzy relation r over R including tuples t 's

defined by

$$r \cup s = \{t \mid \mu_{r \cup s}(t) = \max(\mu_r(t), \mu_s(t))\}.$$

Definition 19. Let r_1 and r_2 be two type-2 fuzzy relations over the same schema $R(U, \mu)$. The difference of r_1 and r_2 , denoted by $r - s$, is the type-2 fuzzy relation r over R including tuples t 's defined by

$$r - s = \{t \mid \mu_{r - s}(t) = \min(\mu_r(t), 1 - \mu_s(t))\}.$$

E. Property of Algebraic Operations

In this section, we propose some properties of the fuzzy relational algebraic operations in T-2FRDB as an extension from those in the classical relational database and T-1FRDB. Clearly, these properties say that T-2FRDB model is coherent and consistent.

Theorem 1. Let r be a fuzzy relation over the schema R in T-2FRDB, ϕ_1 and ϕ_2 be two selection conditions. Then

$$\sigma_{\phi_1}(\sigma_{\phi_2}(r)) = \sigma_{\phi_2}(\sigma_{\phi_1}(r)) = \sigma_{\phi_1 \wedge \phi_2}(r) \quad (1)$$

where, the expressions ϕ_1 and ϕ_2 in $\phi_1 \wedge \phi_2$ are assumed to have the same tuple variable.

Proof: Let $s = \sigma_{\phi_2}(r)$, we have

$$\begin{aligned} \sigma_{\phi_1}(\sigma_{\phi_2}(r)) &= \{t \in s \mid \text{Int}_{R,s,t}(\phi_1) \neq 0\} \quad (\text{Definition 13}) \\ &= \{t \in r \mid \text{Int}_{R,r,t}(\phi_2) \neq 0 \wedge \text{Int}_{R,s,t}(\phi_1) \neq 0\} \\ &= \{t \in r \mid \text{Int}_{R,r,t}(\phi_2) \neq 0 \wedge \text{Int}_{R,r,t}(\phi_1) \neq 0\} \quad (\text{Because } s \subseteq r) \\ &= \{t \in r \mid \min(\text{Int}_{R,r,t}(\phi_2), \text{Int}_{R,r,t}(\phi_1)) \neq 0\} \\ &\quad (\text{Definition 12}) \\ &= \{t \in r \mid \text{Int}_{R,r,t}(\phi_2 \wedge \phi_1) \neq 0\} = \sigma_{\phi_1 \wedge \phi_2}(r). \end{aligned}$$

So, $\sigma_{\phi_1}(\sigma_{\phi_2}(r)) = \sigma_{\phi_1 \wedge \phi_2}(r)$ is proven. The equation $\sigma_{\phi_2}(\sigma_{\phi_1}(r))$

$= \sigma_{\phi_2 \wedge \phi_1}(r)$ is proven similarly. Since $\phi_1 \wedge \phi_2 \Leftrightarrow \phi_2 \wedge \phi_1$ (the logical conjunction of selection conditions is commutative), hence $\sigma_{\phi_1 \wedge \phi_2}(r) = \sigma_{\phi_2 \wedge \phi_1}(r)$. Therefore, it results in $\sigma_{\phi_1}(\sigma_{\phi_2}(r)) =$

$\sigma_{\phi_2}(\sigma_{\phi_1}(r))$ and so $\sigma_{\phi_1}(\sigma_{\phi_2}(r)) = \sigma_{\phi_2}(\sigma_{\phi_1}(r)) = \sigma_{\phi_1 \wedge \phi_2}(r)$. Thus,

Theorem 1 is proven.

Theorem 2. Let r be a fuzzy relation over the schema R in T-2FRDB, A and B be two subsets of attributes of R and $A \subseteq B$. Then

$$\Pi_A(\Pi_B(r)) = \Pi_A(r) \quad (2)$$

Proof: Because $A \subseteq B$, so $A \cap B = A$. From Definition 14, it is easy to see that the sides of (2) are the relations over the same schema with the set of attributes $A \cap B = A$. By the property of the projection of the classical relations, Definition 9 and Definition 14, it follows that two classical sets of tuples which are collected respectively from two relations $\Pi_A(\Pi_B(r))$ and $\Pi_A(r)$ are the same. Moreover, by Definition 14, the operation MAX in two sides of (2) is executed on the same value set of the membership degrees of tuples of r . From that $\Pi_A(\Pi_B(r)) =$

$\Pi_{A \cap B}(r) = \Pi_A(r)$. Thus, the equation (2) is proven.

Theorem 3. Let R_1 , R_2 and R_3 be the schemas in T-2FRDB such that if they have the attributes of the same name then such attributes have the same value domain. Let r_1 , r_2 and r_3 be the fuzzy relations over R_1 , R_2 and R_3 respectively. Then

$$r_1 \bowtie r_2 = r_2 \bowtie r_1 \quad (3)$$

$$(r_1 \bowtie r_2) \bowtie r_3 = r_1 \bowtie (r_2 \bowtie r_3) \quad (4)$$

Proof: Clearly, $r_1 \bowtie r_2$ and $r_2 \bowtie r_1$ are two relations over the same schema. By the property of the join of the classical relations, Definition 9 and Definition 16, it follows that two classical sets of tuples which are collected respectively from two relations $r_1 \bowtie r_2$ and $r_2 \bowtie r_1$ are the same. In addition, the operation MIN of two fuzzy numbers (two membership degrees of two tuples in r_1 and r_2 , respectively) has commutativity. From that leading the join of tuples has commutativity. So, by Definition 16 we have $r_1 \bowtie r_2 = r_2 \bowtie r_1$.

By Definition 16, clearly $(r_1 \bowtie r_2) \bowtie r_3$ and $r_1 \bowtie (r_2 \bowtie r_3)$ are two relations over the same schema. By the property of the join of the classical relations, Definition 9 and Definition 16, it follows that two classical sets of tuples which are collected respectively from two relations $(r_1 \bowtie r_2) \bowtie r_3$ and $r_1 \bowtie (r_2 \bowtie r_3)$ are the same. Let A be a common attribute in U_1 , U_2 and U_3 of R_1 , R_2 and R_3 , because the operation MIN of two fuzzy numbers and the identical operation of attribute values have associativity, from that the join of tuples has associativity. Thus, by Definition 16, it results in $(r_1 \bowtie r_2) \bowtie r_3 = r_1 \bowtie (r_2 \bowtie r_3)$.

Because the Cartesian product (Definition 15) is a particular case of the join we have the straight corollary of Theorem 3 below.

Corollary 1. Let R_1 , R_2 and R_3 be schemas in T-2FRDB such that each pair of them has not any common attribute, r_1 , r_2 and r_3 be fuzzy relations over R_1 , R_2 and R_3 respectively. Then

$$r_1 \times r_2 = r_2 \times r_1 \quad (5)$$

$$(r_1 \times r_2) \times r_3 = r_1 \times (r_2 \times r_3) \quad (6)$$

Theorem 4. Let r_1 , r_2 and r_3 be fuzzy relations over the same schema R in T-2FRDB. Then

$$r_1 \cap r_2 = r_2 \cap r_1 \quad (7)$$

$$(r_1 \cap r_2) \cap r_3 = r_1 \cap (r_2 \cap r_3) \quad (8)$$

$$r_1 \cup r_2 = r_2 \cup r_1 \quad (9)$$

$$(r_1 \cup r_2) \cup r_3 = r_1 \cup (r_2 \cup r_3) \quad (10)$$

Proof: Because the intersection and union operations of sets and the MIN and MAX operations of fuzzy numbers have commutativity and associativity. So, by Definitions 17 and 18, it follows that the equations (7), (8), (9) and (10).

VI. CONCLUSION

In this paper, we have proposed a type-2 fuzzy relational database model, abbreviated to T-2FRDB, as an extension of the type-1 fuzzy relational database models. In T-2FRDB, the

membership degrees of tuples in a relation are represented by the fuzzy numbers on the interval $[0, 1]$. The data model and fuzzy relational algebraic operations in T-2FRDB have been defined formally and consistently. Computing and associating the membership degrees of tuples in manipulating of the algebraic operations are implemented by the operations MAX and MIN using the extension principle. F-2FRDB allows expressing and executing the soft queries that are associated with fuzzy sets for dealing with imprecise information in real databases. A set of basic properties of the algebraic operations in F-2FRDB has also been proposed as theorems and proven completely.

In the next steps, we will extend the notions of the key and fuzzy functional dependency in T-1FRDB for T-2FRDB and build a type-2 fuzzy relational database management system based on T-2FRDB with the familiar querying and manipulating language like SQL for representing and handling the imprecise information in the real world applications.

REFERENCES

- [1] S. Chakraborty, "Codd's relational data model and fuzzy logic: a practical approach to find the computer solution", *International Journal of Advanced Technology & Engineering Research*, vol.2, 2012, pp. 21-27.
- [2] E.F. Codd, "A relational model of data for large shared data banks", *Communications of the Association for Computing Machinery*, vol. 13, 1970, pp. 377-387.
- [3] J.C. Cubero, J.M. Medina, O. Pons, and M.A. Vila, "Data summarization in relational databases through fuzzy dependencies", *International Journal of Information Sciences*, vol. 121, 1999, pp. 22-43.
- [4] C.J. Date, "An introduction to database systems", *Addison-Wesley*, 8th Edition, 2008.
- [5] S.M. Darwish, T.F. Mabrouk, and Y.F. Mokhtar, "Enriching vague queries by type-2 fuzzy orderings", *Lecture Notes on Information Theory*, vol. 2, 2014, pp.177- 185.
- [6] D. Dubois, and H. Prade, "Using fuzzy sets in flexible querying: why and how?", *Proceedings of the Workshop on Flexible Query-Answering Systems*, Denmark, 1996, pp. 89-103.
- [7] G.J. Klir, and B. Yuan, "Fuzzy sets and fuzzy logic--Theory and applications", *Prentice Hall PTR*, 1994.
- [8] J. Mishra, and S. Ghosh, "Uncertain query processing using vague set or fuzzy set: which one is better?", *International Journal of Computers Communications & Control*, vol. 9, 2014, pp. 730-740.
- [9] X. Meng, Z.M. Ma, and X. Zhu, "A Knowledge-based fuzzy query and results ranking approach for relational databases", *Journal of Computational Information Systems*, vol. 6, 2010, pp. 2037-2044.

- [10] Niewiadomski, "A type-2 fuzzy approach to linguistic summarization of data", *IEEE Transactions on Fuzzy Systems*, vol. 16, 2008, pp. 198–212.
- [11] H. Nguyen, "A fuzzy relational database model", *Journal of Information and Communications Technology*, vol. 5(1), 2015, pp. 37-45.
- [12] S. Parsons, "Current approaches to handling imperfect information in data and knowledge bases," *IEEE Transactions on Knowledge and Data Engineering*, vol. 8, 1996, pp. 353-372.
- [13] F.E. Petry, "Fuzzy databases: Principles and applications", *Kluwer Academic Publishers*, 1996.
- [14] O. Pivert, and H. Prade, "Dealing with aggregate queries in an uncertain database model based on possibilistic certainty", *Proceedings of the 15th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU)*, Montpellier, France, 2014, pp. 150-159.
- [15] K. V. S. V. N. Raju, and A.K. Majumdar, "Fuzzy functional dependencies and lossless join decomposition of fuzzy relational database systems", *ACM Transactions on Database Systems*, vol. 13, 1988, pp. 129-166.
- [16] A.A. Sabour, A.M. Gadallah, and H.A. Hefny, "Flexible querying of relational databases: fuzzy set based approach", *Advanced Machine Learning Technologies and Applications Communications in Computer and Information Science*, vol. 488, 2014, pp. 446-455.
- [17] P. C. Saxena, and D. K. Tayal, "Normalization in type-2 fuzzy relational data model based on fuzzy functional dependency using fuzzy functions", *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 20, 2012, pp. 99-138.
- [18] L.A. Zadeh, "Fuzzy sets", *Information and Control*, vol. 8, 1965, pp. 338-353.

AUTHOR'S BIOGRAPHY



Nguyen Hoa is a lecturer of Saigon University. He received B.S degree (1982) in Mathematics from Vinh Pedagogical University, M.E degree (2003) in Computer Science from Ho Chi Minh City University of Technology and Ph.D degree (2008) in Computer Science from Vietnam National University, Ho Chi Minh City.

His research interest includes imprecise and uncertain knowledge representation, fuzzy databases and probabilistic databases.