A New Functional Dependency in a Vague Relational Database Model

Jaydev Mishra
College of Engineering and Management, Kolaghat
West Bengal, India.

Sharmistha Ghosh
Vellore Institute of Technology University,
Vellore, Tamilnadu, India.

ABSTRACT
In order to model the real world with imprecise and uncertain information, various extensions of the classical relational data model have been studied in literature using fuzzy set theory. However, vague set, as a generalized fuzzy set, has more powerful ability to process fuzzy information than fuzzy set. In this paper, we have proposed a vague relational database model and have defined a new kind of vague functional dependency (called α **-vfd)** based on the notion of α -equality of tuples and the idea of similarity measure of vague sets. Next, we present a set of sound vague inference rules which are similar to Armstrong's axioms for the classical case. Finally, partial α -**vfd** and vague key have been studied with the new notion of α -**vfd** and also tested with examples.

General Terms
Vague Database Design

Keywords
Vague set, similarity measure of vague sets, α **-vfd, partial** α **-vfd, vague key**.

1. INTRODUCTION
Information in the real world is very often imprecise or uncertain in nature. Fuzzy set theory, introduced by Zadeh in 1965 [1] has been widely used in literature [2, 3, 4, 5, 6, 7, 8, 9] to incorporate such imprecise data into classical relational databases. Extensive research has been carried out in this direction and several fuzzy relational database models have been proposed to model vague information in relational databases. Also, based on such fuzzy relational models, there have been many studies on different data integrity constraints [2, 3, 4, 8, 9], fuzzy relational algebra [5], fuzzy query languages [10] and so on.

However, vague set theory was put forward by Gau and Buehrer [11] in 1993 as a more efficient tool to deal with imperfect or ambiguous data. A vague set, conceived as a generalization of the concept of fuzzy set, is a set of objects each of which has a grade of membership whose value is a continuous sub-interval of [0,1]. On the contrary, it is well known that a traditional fuzzy set *F* in the universe of discourse *U* is characterized by a single membership function μ_F that assigns to each object $u \in U$ a single membership value $\mu_F(u)$ which is a real number lying between 0 and 1. $\mu_F(u)$ is called the grade of membership of the element *u* in the set *F*. However, in real life it is hard to make sure of the precision degree that an element belongs to a fuzzy set. Further, it was pointed out by Gau *et al*. in [11] that the single membership value in the fuzzy set theory combines the evidence for $u \in U$ and the evidence against $u \in U$ without indicating how much there is of each. To resolve this problem, they had introduced the concept of vague set *V* which is characterized by a truth membership function t_V and a false membership function f_V. Thus, a vague set separates the positive and negative evidence for membership of an element in the set and provides lower and upper bounds on the grade of membership of an element in the set. These lower and upper bounds are used to create a sub-interval on [0, 1], namely, $[t_V(u), 1 - f_V(u)]$, to generalize the membership function $\mu_V(u)$ of fuzzy sets, where $t_V(u) \le \mu_V(u) \le 1 - f_V(u)$.

Since vague sets have been introduced to deal with imprecise information in a more efficient manner than traditional fuzzy sets, classical relational databases may also be extended to represent and deal with uncertain data with the concept of vague set theory. The extended database model is then called a vague relational database model. However, compared to fuzzy relational databases, much less research has been reported so far in the area of vague relational database. This is, in particular, true for the study of vague functional dependency. But, it is well known that data dependencies play an important role in any database design and implementation of functional dependency of one set of attributes upon another is one of the most vital concepts in relational databases. Thus, similar to the theory of classical relational databases, vague functional dependencies can also be used as a guideline for the design of a vague relational schema.

Zhao et al. [12] have proposed a vague relational model based on vague set theory and a new similarity measure (SE) between vague sets. They have also extended and studied the vague relational algebra in ref. [13]. In particular, Zhao et al. have focused on the issues of vague functional dependency and vague Armstrong's axioms in [12]. According to Zhao et al., a vague relation *r* on a relational schema *R* satisfies the vague functional dependency **vfd**: $X \xrightarrow{vfd} Y$, where $XY \subseteq R$ if $(\forall t \in r)(\forall s \in r)(SE(t[X], s[X]) \le (SE(t[Y], s[Y]))$. It may be clearly observed that the above definition of **vfd** fails if the $SE(t[Y], s[Y])$ is slightly less than $SE(t[X], s[X])$, which should not be the case in reality. However, such situations have been taken care by Zhao et al. [12] in the satisfaction degree of **vfd**s. Such vague functional dependencies have been utilized by Lu & Ng [14] to maintain the consistency of a vague database.

In this paper, we propose a new kind of vague functional dependency (called α -**vfd**) based on the concept of α -equality of tuples and the notion of similarity measure between vague sets which can resolve the problem in such situations at definition level. Next, we present a set of sound vague inference rules which are similar to Armstrong's axioms for the classical case. Finally, **partial** α **-vfd** and **vague key** have also been studied with the new notion of α -**vfd** and tested with examples.

The organization of the paper is as follows:

The basics of vague set theory have been reviewed in **section 2**. The vague relational model has been proposed in **section 3**. **Section 4** discusses the similarity measure of vague sets used in this study. We present the newly proposed vague functional dependency (α -**vfd**) in **section 5**. The vague inference rules have also been investigated and the definitions of partial α -vfd and vague key have been introduced in the same section. The final conclusions are reported in **section 6**.

2. BASICS OF VAGUE SET

In this section, we review some basic definitions on the theory of vague sets [11].

Let *U* be the universe of discourse where an element of *U* is denoted by *u*.

Definition 2.1

A **vague set** *V* in the universe of discourse *U* is characterized by two membership functions given by:

(i) a truth membership function $t_V : U \to [0,1]$,

and

(ii) a false membership function $f_V : U \to [0,1]$,

where $t_V(u)$ is a lower bound of the grade of membership of *u* derived from the 'evidence for *u*', and $f_V(u)$ is a lower bound on the negation of *u* derived from the 'evidence against *u*', and $t_V(u) + f_V(u) \le 1$.

Thus, the grade of membership $\mu_V(u)$ of *u* in the vague set *V* is bounded by a subinterval $[t_V(u), 1 - f_V(u)]$ of [0,1] i.e., $t_V(u) \le \mu_V(u) \le 1 - f_V(u)$.

Then, the **vague set *V*** is written as $V = \{\langle u, [t_V(u), 1 - f_V(u)]\rangle : u \in U\}$.

Here, the interval $[t_V(u), 1 - f_V(u)]$ is said to be the vague value to the object *u* and is denoted by $V_V(u)$. For example, in a voting process, the vague value [0.5, 0.8] can be interpreted as "the vote for a resolution is 5 in favour, 2 against and 3 neutral (abstentious)."

The precision of knowledge about *u* is clearly characterized by the difference $(1 - f_V(u) - t_V(u))$. If this is small, the knowledge about *u* is relatively precise. However, if it is large, we know correspondingly little. If $t_V(u)$ is equal to $(1 - f_V(u))$, the knowledge about *u* is precise, and vague set theory reverts back to fuzzy set theory. If $t_V(u)$ and $(1 - f_V(u))$ are both equal to 1 or 0, depending on whether *u* does or does not belong to *V*, the knowledge about *u* is very exact and the theory goes back to that of ordinary sets. Thus any crisp or fuzzy set may be considered as a special case of vague set.

However, it may be noted that interval-valued fuzzy sets (*i-v* fuzzy sets) [15] are not vague sets. In *i-v* fuzzy sets, an interval based membership value is assigned to each element of the universe considering only the "evidence for *u*", without considering the "evidence against *u*". In vague sets both are independently proposed by the decision maker. This makes a major difference in judgment about the grade of membership.

Next, we present several special vague sets and various operations on vague sets that are obvious extensions of the corresponding definitions for ordinary sets and fuzzy sets.

Definition 2.2

A vague set *V* is an **empty** vague set, denoted by ϕ , if and only if its truth-membership function $t_V(u) = 0$ and false-membership function $f_V(u) = 1$ for all *u* on *U*.

Definition 2.3

The **complement** of a vague set *V* is denoted by V^c and is defined by the truth-membership and false-membership functions $t_{V^c}(u)$ and $f_{V^c}(u)$ as follows:

$$t_{V^c}(u) = f_V(u) \text{ and } f_{V^c}(u) = t_V(u).$$

Definition 2.4

A vague set *A* is **contained** in another vague set *B*, written as $A \subseteq B$, if and only if, $t_A \le t_B$ and $f_A \ge f_B$.

Definition 2.5

Two vague sets *A* and *B* are **equal**, written as *A* = *B*, iff, $A \subseteq B$ and $B \subseteq A$, that is, $t_A = t_B$ and $f_A = f_B$.

Definition 2.6

The **union** of two vague sets *A* and *B* is a vague set *C*, denoted as $C = A \bigcup B$, whose truth-membership and false-membership functions are related to those of *A* and *B* by $t_C = \max(t_A, t_B)$ and $f_C = \min(f_A, f_B)$.

Definition 2.7

The **intersection** of two vague sets *A* and *B* is a vague set *C*, written as $C = A \bigcap B$, such that $t_C = \min(t_A, t_B)$ and $f_C = \max(f_A, f_B)$.

Definition 2.8

Let $U = U_1 \times U_2 \times \cdots \times U_n$ be the Cartesian product of n universes, and $V_i, i = 1,2,\ldots,n$ be vague sets in their corresponding universe of discourses $U_i, i = 1,2,\ldots,n$

respectively. Also, let $u_i \in U_i, i = 1,2,\ldots,n$. Then the Cartesian product $V = V_1 \times V_2 \times \cdots \times V_n$ is defined to be a vague set of $U = U_1 \times U_2 \times \cdots \times U_n$ with the truth and false membership functions defined as follows:

$$t_V(u_1,u_2,\ldots,u_n) = \min\left(t_{V_1}(u_1), t_{V_2}(u_2),\ldots,t_{V_n}(u_n)\right)$$

$$f_V(u_1,u_2,\ldots,u_n) = \max\left(f_{V_1}(u_1), f_{V_2}(u_2),\ldots, f_{V_n}(u_n)\right)$$

.

3. VAGUE RELATIONAL DATABASE MODEL (VRDB)

In this section, we make an attempt to extend the classical relational database model to incorporate imprecise or vague data by means of vague set theory, which results in the Vague Relational Database model (VRDB).

Table-I : A vague relational instance *r* of "EMP_SAL" relation

NAME	EXP	SAL
Smith	{<12, [1, 1]>, <15, [.9, .95]>, <20, [0, 0]>}	{<45000, [.8, .95]>}
David	{<12, [.9, .95]>, <15, [.95, 95]>, <20, [.2, .3]>}	{<45000, [1, 1]>, <75000, [.1, .2]>}
Sachin	{<12, [.1, .15]>, <15, [.2, 3]>, <20, [1, 1]>}	{<75000, [1, 1]>}
John	{ <20, [.9, .95]>}	{<45000, [.2, .3]>, <75000, [.9, .95]>}

Definition 3.1

Let $A_i, i = 1,2,\ldots,n$ be n attributes defined on the universes of discourse sets U_i, respectively. Then, a **vague relation** *r* on the relation schema $R(A_1, A_2,\ldots,A_n)$ is defined as a subset of the Cartesian product of a collection of vague subsets:

$r \subseteq V(U_1) \times V(U_2) \times \ldots \times V(U_n)$,

where $V(U_i)$ denotes the collection of all vague subsets on a universe of discourse U_i.

Each tuple *t* of *r* consists of a Cartesian product of vague subsets on the respective U_i's, i.e., $t[A_i] = \pi(A_i)$, where $\pi(A_i)$ is a vague subset of the attribute A_i defined on U_i for all *i*. The relation *r* can thus be represented by a table with *n* columns.

It may be observed that the vague relation can be considered as an extension of classical relations (all vague values are [1, 1]) and fuzzy relations (all vague values are [*a*, *a*], $0 \le a \le 1$). It is clear that the vague relation can capture more information about vagueness.

Example 3.1.1: Consider the vague relational instance *r* over EMP_SAL (NAME, EXP, SAL) given above in **Table-I**. In *r*, EXP(Experience) and SAL(Salary) are vague attributes. The first tuple in *r* means the employee with NAME = "Smith" has the experience of {<12, [1, 1]>, <15, [.9, .95]>, <20, [0, 0]>} and the salary of {<45000, [.8, .95]>}, which are vague sets. Here the vague data <12, [1, 1]> means the evidence in favour of "The experience is 12" is 1 and the evidence against it is 0. Similarly the vague data <45000, [.8, .95]> indicates the evidence in favour of "The salary is Rs. 45000" is 0.8 while the evidence against it is 0.05 and so on.

4. SIMILARITY MEASURE OF VAGUE DATA

There have been some studies in literature which discuss the topic concerning how to measure the degree of similarity between vague sets [16, 17, 18]. In [18], it was pointed out by Lu et al. that the similarity measure in [16, 17] did not fit well in some cases. They have proposed a new similarity measure between vague sets which turned out to be more reasonable in more general cases. The same has been used in the present work which is defined as follows:

Definition 4.1: ***Similarity Measure between two vague values***

Let *x* and *y* be two vague values such that $x = [t_x, 1\text{-}f_x]$, $y = [t_y, 1\text{-}f_y]$, where $0 \le t_x \le 1\text{-}f_x \le 1$, and $0 \le t_y \le 1\text{-}f_y \le 1$. Let $SE(x, y)$ denote the similarity measure between *x* and *y*. Then,

$$SE(x,y) = \sqrt{\left(1 - \frac{\left|(t_x - t_y) - (f_x - f_y)\right|}{2}\right)\left(1 - \left|(t_x - t_y) + (f_x - f_y)\right|\right)}$$

Definition 4.2: ***Similarity Measure between two vague sets***

Let $U = \{u_1, u_2, u_3, \ldots, u_n\}$ be the universe of discourse. Let *A* and *B* be two vague sets on *U*, where:

$$A = \{< u_i, [t_A(u_i), 1 - f_A(u_i)] >, \forall u_i \in U\},$$

where $t_A(u_i) \le \mu_A(u_i) \le 1\text{-} f_A(u_i)$ and $1 \le i \le n$.

$$B = \{< u_i, [t_B(u_i), 1 - f_B(u_i)] >, \forall u_i \in U\},$$

where $t_B(u_i) \le \mu_B(u_i) \le 1\text{-} f_B(u_i)$ and $1 \le i \le n$.

Now, the similarity measure between *A* and *B*, denoted by *SE (A, B)* is defined as:

$$SE(A,B) = \frac{1}{n}\sum_{i=1}^{n} SE([t_A(u_i), 1 - f_A(u_i)], [t_B(u_i), 1 - f_B(u_i)])$$

$$= \frac{1}{n}\sum_{i=1}^{n}\sqrt{\left(1 - \frac{\left|(t_A(u_i) - t_B(u_i)) - (f_A(u_i) - f_B(u_i))\right|}{2}\right)\left(1 - \left|(t_A(u_i) - t_B(u_i)) + (f_A(u_i) - f_B(u_i))\right|\right)}$$

5. DESIGN OF VAGUE RELATIONAL DATABASE MODEL

Functional Dependency (**fd**) is one of the most important data dependencies which play a crucial role in design of any logical database. Here we proceed to extend the concept of functional dependency in the light of vague set theory. This is termed as Vague Functional Dependency (**vfd**). In the present work, we have proposed a new **vfd**, called α **-vfd**, which is based on the idea of α - equality of two vague tuples. Here, $\alpha \in [0, 1]$ is a choice parameter and the concept of α - equality has been introduced using the notion of similarity measure of vague data as follows:

5.1 Vague Functional Dependency (vfd)

Definition 5.1.1: α ***- equality of two vague tuples***

Let $r(R)$ be a vague relation on the relational schema $R(A_1, A_2, \ldots, A_n)$. Let t_1 and t_2 be any two vague tuples in r. Let $\alpha \in [0, 1]$ be a threshold or choice parameter, predefined by the database designer, and $X = \{ X_1, X_2, \ldots, X_k \} \subseteq R$. Then the vague tuples t_1 and t_2 are said to be α-equal on X if $SE(t_1[X_i], t_2[X_i]) \geq \alpha \ \forall \ i=1,2,3\ldots k$. We denote this equality by the notation $t_1[X](VE)_\alpha t_2[X]$.

Then, the following proposition is straightforward from the above definition.

Proposition 5.1.1: If $0 \leq \alpha_2 \leq \alpha_1 \leq 1$, then $t_1[X](VE)_{\alpha_1} t_2[X] \Rightarrow t_1[X](VE)_{\alpha_2} t_2[X]$.

Next, we define our α **-vfd** as follows:

Definition 5.1.2: ***Vague Functional Dependency (*** α **-vfd*)***

Let $X, Y \subset R = \{A_1, A_2 \ldots A_n\}$. Choose a threshold value $\alpha \in [0, 1]$. Then a vague functional dependency (**vfd**), denoted by $X \underset{\alpha}{\overset{vfd}{\rightarrow}} Y$, is said to exist if, whenever $t_1[X](VE)_\alpha t_2[X]$, it is also the case that $t_1[Y](VE)_\alpha t_2[Y]$.

We may read this **vfd** as follows: the set of attributes X vague functionally determines the set of attributes Y at α-level of choice. In another terminology, the set of attributes Y is vague functionally determined by the set of attributes X at α-level of choice. During the course of analysis, different α values may be set by the database designer so, that the above **vfd** may be termed as α **-vfd**. Also, we have the following straightforward proposition for α **-vfd**.

Proposition 5.1.2 : If $0 \leq \alpha_2 \leq \alpha_1 \leq 1$, then $X \underset{\alpha_1}{\overset{vfd}{\rightarrow}} Y \Rightarrow X \underset{\alpha_2}{\overset{vfd}{\rightarrow}} Y$.

Example 5.1.1: Consider the vague relational instance r presented in **Table-I**. Let us now check whether $EXP \underset{\alpha}{\overset{vfd}{\rightarrow}} SAL$ holds to a certain α-level of choice or not.

First, we obtain SE(t_p[EXP], t_q[EXP]) and SE(t_p[SAL], t_q[SAL]) for every pair of tuples t_p & t_q in r, respectively, using ***Definition 4.2***. The results are shown below by the matrices E and S respectively in **Table-II**.

Table-II : Matrices E & S showing Similarity Measure

E	t_1	t_2	t_3	t_4	S	t_1	t_2	t_3	t_4
t_1	1	.91	.3	.18	t_1	1	.87	.16	.3
t_2	.91	1	.48	.35	t_2	.87	1	.18	.47
t_3	.3	.48	1	.89	t_3	.16	.18	1	.88
t_4	.18	.35	.89	1	t_4	.3	.47	.88	1

For α = **0.8** (given by decision maker), we can see from the above two matrices, that for any pair of tuples t_p & t_q

if SE(t_p[EXP], t_q[EXP]) $\geq$ α, then it is also the case that SE(t_p[SAL], t_q[SAL]) $\geq$ α.

So, we can say that the **vfd** $EXP \underset{.8}{\overset{vfd}{\rightarrow}} SAL$ holds for the vague relational instance r.

Also, from above matrices E and S, we can say that the **vfd** $EXP \underset{.85}{\overset{vfd}{\rightarrow}} SAL$ holds true. However the **vfd** $EXP \underset{.9}{\overset{vfd}{\rightarrow}} SAL$ does not hold because for tuples t_1 and t_2, SE(t_1[EXP], t_2[EXP]) = 0.91 ≥ 0.9, but SE(t_p[SAL], t_q[SAL]) = 0.87 ≤ 0.9.

5.2 Inference rules for α -vfd

It is well known that in classical relational databases, functional dependencies satisfy a set of inference rules called Armstrong's axioms. In this section, we have derived a set of inference rules for our proposed α **-vfd**. These vague inference rules are similar to Armstrong's axioms for **fd**. We call them **vague Armstrong's axioms** and are given as follows:

(A1) α **-vfd reflexive rule**: If $Y \subseteq X \subseteq R$, then $X \underset{\alpha}{\overset{vfd}{\rightarrow}} Y$.

(A2) α **-vfd augmentation rule**: If $X \underset{\alpha}{\overset{vfd}{\rightarrow}} Y$ and $Z \subseteq R$, then $XZ \underset{\alpha}{\overset{vfd}{\rightarrow}} YZ$.

(A3) α **-vfd transitive rule**: If $X \xrightarrow[\alpha_1]{vfd} Y$ and $Y \xrightarrow[\alpha_2]{vfd} Z$, then $X \xrightarrow[\min(\alpha_1,\alpha_2)]{vfd} Z$.

Theorem 5.2.1 : Vague Armstrong's axioms (A1)- (A3) are sound.

Proof:

(A1) α **-vfd reflexive rule**:

Let $t_1[X](VE)_\alpha t_2[X]$ is true, i.e., $SE(t_1[X_i], t_2[X_i]) \geq \alpha \ \forall X_i \in X$.

Then, $SE(t_1[X_i], t_2[X_i]) \geq \alpha \ \forall X_i \in Y$ holds ($\because Y \subseteq X$) i.e., $t_1[Y](VE)_\alpha t_2[Y]$ is also true.

This implies $X \xrightarrow[\alpha]{vfd} Y$ holds. Hence **proved**.

(A2) α **-vfd augmentation rule**:

Let $X \xrightarrow[\alpha]{vfd} Y$. Now, from ***Definition 5.1.2***, for any two tuples t_1 and t_2 if $t_1[X](VE)_\alpha t_2[X]$ **……….(i)** is true, then $t_1[Y](VE)_\alpha t_2[Y]$ **……….(ii)** is also true. Next, suppose $t_1[XZ](VE)_\alpha t_2[XZ]$ **……….(iii)** is true. This implies,

$SE(t_1[X_i], t_2[X_i]) \geq \alpha \ \forall X_i \in XZ$

$\Rightarrow SE(t_1[X_i], t_2[X_i]) \geq \alpha \ \forall X_i \in Z$

$\Rightarrow t_1[Z](VE)_\alpha t_2[Z]$ **……….(iv)**

Then from **(ii)** and **(iv)**, we get $t_1[YZ](VE)_\alpha t_2[YZ]$ **……….(v)**

Thus, for any two tuples t_1 and t_2 if $t_1[XZ](VE)_\alpha t_2[XZ]$, then it is also the case that $t_1[YZ](VE)_\alpha t_2[YZ]$ which implies $XZ \xrightarrow[\alpha]{vfd} YZ$. Hence **proved**.

(A3) α **-vfd transitive rule**:

Let us assume that both the **vfd**s $X \xrightarrow[\alpha_1]{vfd} Y$ and $Y \xrightarrow[\alpha_2]{vfd} Z$ hold in the relation $r(R)$.

<u>Case I:</u> $\alpha_1 \geq \alpha_2$ so that min (α_1, α_2) = α_2.

Given that $X \xrightarrow[\alpha_1]{vfd} Y$ and $0 \leq \alpha_2 \leq \alpha_1 \leq 1$.

So, using ***Proposition 5.1.2*** we get $X \xrightarrow[\alpha_2]{vfd} Y$. ---------(i)

Then, from (i) we can write

$t_1[X](VE)_{\alpha_2} t_2[X] \Rightarrow t_1[Y](VE)_{\alpha_2} t_2[Y]$ ----(ii)

Again, since $Y \xrightarrow[\alpha_2]{vfd} Z$ holds, so we have

$t_1[Y](VE)_{\alpha_2} t_2[Y] \Rightarrow t_1[Z](VE)_{\alpha_2} t_2[Z]$ -----(iii)

Combining (ii) and (iii), we get

$t_1[X](VE)_{\alpha_2} t_2[X] \Rightarrow t_1[Z](VE)_{\alpha_2} t_2[Z]$

which implies $X \xrightarrow[\alpha_2]{vfd} Z$.

Hence for α_2 = min (α_1, α_2), if $X \xrightarrow[\alpha_1]{vfd} Y$ and $Y \xrightarrow[\alpha_2]{vfd} Z$, then $X \xrightarrow[\alpha_2]{vfd} Z$.

<u>Case II:</u> $\alpha_2 \geq \alpha_1$ follows similarly. Hence **proved**.

Using the above vague Armstrong's axioms, the following results are also derived for α **-vfd**.

(A4) α **-vfd decomposition rule**: If $X \xrightarrow[\alpha]{vfd} YZ$, then $X \xrightarrow[\alpha]{vfd} Y$ and $X \xrightarrow[\alpha]{vfd} Z$.

Proof: Given that $X \xrightarrow[\alpha]{vfd} YZ$ --------- (i)

Since Y$\subset$YZ, by α **-vfd reflexive rule**, we have $YZ \xrightarrow[\alpha]{vfd} Y$ --------- (ii)

From (i) and (ii) using α **-vfd transitive rule**, we get $X \xrightarrow[\alpha]{vfd} Y$.

$X \xrightarrow[\alpha]{vfd} Z$ also follows similarly.

Thus, if $X \xrightarrow[\alpha]{vfd} YZ$, then $X \xrightarrow[\alpha]{vfd} Y$ and $X \xrightarrow[\alpha]{vfd} Z$. Hence **proved**.

(A5) α **-vfd union rule**: If $X \xrightarrow[\alpha_1]{vfd} Y$ and $X \xrightarrow[\alpha_2]{vfd} Z$, then $X \xrightarrow[\min(\alpha_1,\alpha_2)]{vfd} YZ$.

Proof: Given that $X \xrightarrow[\alpha_1]{vfd} Y$ --------- (i) and $X \xrightarrow[\alpha_2]{vfd} Z$ ---------(ii)

From (i) we may write $X \xrightarrow[\alpha_1]{vfd} XY$ --------- (iii) (using α **-vfd augmentation rule**).

Similarly, from (ii) we can write $XY \xrightarrow[\alpha_2]{vfd} YZ$ --------- (iv) .

Thus from (iii) and (iv) using α **-vfd transitive rule**, we get $X \xrightarrow[\min(\alpha_1,\alpha_2)]{vfd} YZ$. Hence **proved**.

(A6) α **-vfd pseudo transitive rule**: If $X \xrightarrow[\alpha_1]{vfd} Y$ and $WY \xrightarrow[\alpha_2]{vfd} Z$, then $WX \xrightarrow[\min(\alpha_1,\alpha_2)]{vfd} Z$.

Proof: Given that $X \xrightarrow[\alpha_1]{vfd} Y$ --------- (i)

and $WY \xrightarrow[\alpha_2]{vfd} Z$ --------- (ii)

From (i), using α **-vfd augmentation rule** we can write $WX \xrightarrow[\alpha_1]{vfd} WY$. --------- (iii)

From (iii) and (ii) using α **-vfd transitive rule**, we get $WX \xrightarrow[\min(\alpha_1,\alpha_2)]{vfd} Z$. Hence **proved**.

5.3 Partial vague functional dependency (partial α -vfd)

After validation of Armstrong's axioms in the vague environment with our present notion of α **-vfd**, let us define partial vague functional dependency (**partial** α **-vfd**) as follows:

Definition 5.3.1: ***Partial vague functional dependency*** **(partial** α **-vfd**)

Y is called partially vague functionally dependent on *X* at α -level of choice, i.e., $X \xrightarrow[\alpha]{vfd} Y$ partially, if $X \xrightarrow[\alpha]{vfd} Y$ hold and also there exists a non empty set $X' \subset X$, such that, $X' \xrightarrow[\alpha]{vfd} Y$.

The concept of **partial** α **-vfd** then expresses the fact that after removal of an attribute X_i from X, the dependency still holds i.e., for an attribute $X_i \in X$, X-{ X_i} still vague functionally determines Y at α -level of choice. The notion of **partial** α **-vfd** is needed to define vague key.

Example 5.3.1

Let the relational schema be R={A, B, C, D, E} and the set of **vfds** *V* on R be given by

V={ $ABC \xrightarrow[0.75]{vfd} D$ and $AC \xrightarrow[0.8]{vfd} D$ }. Then it may be easily observed that the **vfd** $ABC \xrightarrow[0.75]{vfd} D$ is a **partial** α **-vfd**.

5.4 Vague Key

In classical relational database, key is known to be a special case of functional dependency. Let us now extend the idea of classical key in the vague environment to define vague key with α -level of choice where $\alpha \in [0, 1]$ is a choice parameter defined by the database designer. A formal definition of **vague key** is as follows:

Definition 5.4.1: ***Vague Key***

Let $K \subseteq R$ and V be a set of **vfd**s for R. Then, K is called a **vague key** of R at α -level of choice where $\alpha \in [0, 1]$ iff $K \xrightarrow[\alpha]{vfd} R \in V$ and $K \xrightarrow[\alpha]{vfd} R$ is not a **partial** α **-vfd**.

Example 5.4.1: Let us assume a relation schema R= (A, B, C, D) and a set of **vfd**s

V = { $A \xrightarrow[0.75]{vfd} B$, $A \xrightarrow[0.8]{vfd} C$, $A \xrightarrow[0.7]{vfd} D$ } of R. Find a **vague key** of R.

Solution: Given $A \xrightarrow[0.75]{vfd} B$ --------(i), $A \xrightarrow[0.8]{vfd} C$ -------- (ii) and $A \xrightarrow[0.7]{vfd} D$ -------- (iii)

Applying α **-vfd union rule** on (i) and (ii), we get $A \xrightarrow[0.75]{vfd} BC$ -------- (iv)

Again, applying α **-vfd union rule** on (iii) and (iv), we get $A \xrightarrow[0.7]{vfd} BCD$ -------- (v) Also, $A \xrightarrow[1]{vfd} A$ is trivial ---- (vi)

Thus from (v) and (vi) using α **-vfd union rule**, we get $A \xrightarrow[0.7]{vfd} ABCD$ i.e., $A \xrightarrow[0.7]{vfd} R$

which implies that *A* is a **vague key** of *R* at 0.7-level of choice.

6. CONCLUSION

In this paper, we present an extension of the classical relational database model with the concepts of vague set theory, a generalized version of fuzzy sets. The paper mainly concentrated on the study of functional dependency in vague relational database. For this purpose, we have introduced a new kind of vague functional dependency (called α **-vfd**) based on the idea of α -equality of tuples and similarity measure of vague sets. We have also derived the vague inference rules and defined partial α -vfd and vague key in the paper.

The work may be extended to study Multivalued Dependency and Normalization using α **-vfd** which constitute an important part of a relational database design.

7. REFERENCES

[1] Zadeh L. A., 1965, “Fuzzy Sets”, Information and Control, Vol. 8, No.3, pp. 338-353.

[2] Al-Hamouz S., Biswas R., 2006, “Fuzzy Functional Dependencies in Relational Databases”, International Journal of Computational Cognition, Vol. 4, No.1, pp. 39-43.

[3] Liao S. Y., Wang H. Q. and Liu W.Y., 1993, “Functional Dependencies with Null values, Fuzzy values, and Crisp values”, IEEE Transaction on Fuzzy Systems, Vol. 7, No. 1, pp. 97-103.

[4] Liu, W. Y., 1997, “Fuzzy Data Dependencies and Implication of Fuzzy Data Dependencies”, Fuzzy Sets and Systems, Vol. 92, No. 3, pp. 341-348.

[5] Ma Z. M., Zhang W.J., Ma W. Y., and Mili F., 2002, “Handling Fuzzy Information in Extended Possibility-Based Fuzzy Relational Databases”, International Journal of Intelligent Systems, Vol. 17, No. 10, pp. 925-942.

[6] Mishra J. and Ghosh S., 2008, “A Study of Fuzzy Relational Database”, International Journal of Computational Cognition, Vol. 6, No. 4, pp. 45-50.

[7] Yazici A. and Sozat M. I., 1998, “The Integrity Constraints for Similarity-Based Fuzzy Relational Databases”, International Journal of Intelligent Systems, Vol. 13, No. 7, pp. 641-659.

[8] Yazici A. and Sozat M. I., 2001, "A Complete Axiomatization for Fuzzy Functional and Multivalued Dependencies in Fuzzy Database Relations", Fuzzy Sets and Systems, Vol. 117, No. 2, pp. 161-181.

[9] Singaraju J., 2010, "Fuzzy Subset Dependencies", International Journal of Computer Applications, Vol. 8, No. 13, pp. 29-36.

[10] Bosc P. and Pivert O., 1995, "SQLf: A Relational Database Language for Fuzzy Querying", IEEE Transactions on Fuzzy Systems, Vol. 3, No. 1, pp. 1-17.

[11] Gau W. L. and Buehrer D. J., 1993, "Vague Sets", IEEE Trans. Syst. Man, Cybernetics, Vol. 23, No. 2, pp. 610-614.

[12] Zhao F. and Ma Z. M., 2006, "Functional Dependencies in Vague Relational Databases", IEEE International Conference on Systems, Man and Cybernetics.

[13] Zhao F., Ma Z. M., and Yan L., 2007, "A Vague Relational Model and Algebra", Fourth International Conference on Fuzzy Systems and Knowledge Discovery, (FSKD).

[14] Lu A. and Ng W., 2009, "Maintaining consistency of vague databases using data dependencies", Data and Knowledge Engineering, Vol. 68, No. 7, pp. 622-641.

[15] Zadeh L. A., 1975, "The Concept of a Linguistic Variable and its Application to Approximate Reasoning - I", Information Sciences, Vol. 8, pp. 199-249.

[16] Chen S. M., 1997, "Similarity Measure between Vague Sets and between Elements", IEEE Trans. Syst. Man, Cybern., Vol. 27, No. 1, pp.153-158.

[17] Hong D. H. and Kim C., 1999, "A Note on Similarity Measures between Vague Sets and between Elements", Information Sciences, Vol. 115, pp. 83-96.

[18] Lu A., and Ng W., 2004, "Managing Merged Data by Vague Functional Dependencies", LNCS 3288, pp. 259-272, Springer-Verlag Berlin Heidelberg.