Soundness and Completeness of Inference Rules for New Vague Functional Dependencies

Dženan Gušić\textsuperscript{1,*}

\textsuperscript{1}University of Sarajevo, Faculty of Sciences and Mathematics, Department of Mathematics, Zmaja od Bosne 33-35, 71000 Sarajevo, Bosnia and Herzegovina

Abstract. In this paper we introduce a new definition of vague functional dependency based on application of appropriately chosen similarity measures. The definition is adjusted in order to be applicable to both, the imprecise and precise vague functional dependencies. Ultimately, the set of inference rules for new vague functional dependencies is given, and is proven to be sound and complete.

1 Introduction

Let $U$ be the universe of discourse.

Suppose that $V$ is a vague set in $U$.

Now, there exist functions $t_V : U \to [0, 1]$, $f_V : U \to [0, 1]$, such that $t_V (u) + f_V (u) \leq 1$ for $u \in U$.

We shall write

$$V = \{ [t_V (u), f_V (u)] : u \in U \},$$

where $[t_V (u), 1 - f_V (u)] \subseteq [0, 1]$ is the vague value joined to $u \in U$.

Recall that the vague value $[t_V (u), 1 - f_V (u)]$ reduces to the fuzzy value $t_V (u) = 1 - f_V (u) \in [0, 1]$ if it happens that $t_V (u) = 1 - f_V (u) \in (0, 1)$.

If it happens that $t_V (u) = 1 - f_V (u) = 1$, then the vague value $[t_V (u), 1 - f_V (u)]$ reduces to the ordinary value $t_V (u) = 1 - f_V (u) = 1 \in [0, 1]$.

The ordinary case $t_V (u) = 1 - f_V (u) = 0$, we read as: the element $u$ does not belong to the vague set $V$. In such scenario, we write

$$V = \{ [u_1, 0.2, 0.7], 0.2, [1, 1] \}$$

instead of

$$V = \{ [u_1, 0.2, 0.7], [u_2, 0.2, 0.7], [u_3, 0.2, 0.7] \},$$

where $V = \{ u_1, u_2, u_3 \}$ is some universe of discourse, and $V$ is some vague set in $U$.

Let $R (A_1, A_2, \ldots, A_n)$ be a relation scheme on domains $U_1, U_2, \ldots, U_n$, where $A_i$ is an attribute on the universe of discourse $U_i, i \in [1, 2, \ldots, n] = I$.

Suppose that $V (U_i)$ is the family of all vague sets in $U_i, i \in I$.

A vague relation $r$ on $R (A_1, A_2, \ldots, A_n)$ is a subset of the cross product $V (U_1) \times V (U_2) \times \cdots \times V (U_n)$.

A tuple $t$ of $r$ is then of the form

$$(t, A_1, \ldots, t, A_n),$$

where $t, A_i$ is a vague set in $U_i, i \in I$.

Note that we may (more freely speaking) consider $t, A_i$ the value of the attribute $A_i$ on $t$.

A vague relation $r$ on $R (A_1, A_2, \ldots, A_n)$ can be visually represented as a two-dimensional table with $n$ columns and the table headings $A_1, A_2, \ldots, A_n$, where each horizontal row of the table is a tuple of $r$, and each column of the table contains the attribute values under the corresponding heading.

Let $R (Name, Int, Succ)$ be a relation scheme on domains $U_1 = \{ Em, Ted, Jim, Katie, Sara, Tina, Joe, John \}$, $U_2 = \{ 115, 130, 145 \}$, $U_3 = \{ 5, 10 \}$, where $Int$ (as intelligence) and $Succ$ (as success) are vague attributes on universes $U_2$ and $U_3$, respectively, and $Name$ as ordinary attribute on the universe of discourse $U_1$.

Let $r$ be the vague relation instance on $R (Name, Int, Succ)$ given by Table 1.

\begin{table}[h]
\centering
\begin{tabular}{ccc}
\hline
Name & Int & Succ \\
\hline
1 & \{ Ted \} & \{ [115, [1, 1]], [10, [1, 1]] \} \\
2 & \{ Sara \} & \{ [115, [0.7, 0.9]], [5, [0.6, 0.9]], [130, [0.9, 0.95]], [10, [0.8, 0.95]] \} \\
3 & \{ Jim \} & \{ [130, [0.8, 0.9]], [10, [1, 1]], [145, [0.85, 0.95]] \} \\
4 & \{ Katie \} & \{ [145, [1, 1]], [5, [0.9, 0.95]] \} \\
\hline
\end{tabular}
\end{table}

The vague sets $\{ [115, [1, 1]], [10, [1, 1]] \}$, given in the first row of the Table 1, mean that the knowledge about Ted’s intelligence and success are very accurate. More precisely, one knows that his intelligence and success are exactly 115 and 10, respectively. Having in mind that the ranges of person’s intelligence and success are determined
by the sets \{115, 130, 145\} and \{5, 10\}, we may say that Ted is very successful person with regard to his intelligence. Sara’s intelligence is determined by the vague set \{(115, [0.7, 0.9]), (130, [0.9, 0.95])\}. Since the truth value 0.7 is quite high, the false value 0.1 = 1 - 0.9 is pretty small, and the difference 0.9 - 0.7 = 0.2 is also very small, we conclude that Sara’s intelligence must be close to 115. However, 0.9 > 0.7, 0.05 = 1 - 0.95 < 0.1, and 0.95 - 0.9 = 0.05 < 0.2 = 0.9 - 0.7, so Sara’s intelligence is definitively closer to 130 (from bellow) than to 115 (note that 115 \not\in \{(115, [0.7, 0.9]), (130, [0.9, 0.95])\}). Reasoning in the same way, we conclude that Sara’s success is between 5 and 10, and it is closer to 10 than to 5. The data about Katie are quite precise. As opposed to Ted, however, she is a very intelligent person who is not so successful. Compared to Ted and Katie, Sara is a relatively intelligent person who is relatively successful. Finally, Jim is a pretty intelligent person who is also very successful.

For the basic relational concepts, see, e.g., [10].

Let \( r_1 \) be the fuzzy relation instance on \( R(\text{Name}, \text{Int}, \text{Succ}) \) given by Table 2 (now, we assume that \( \text{Int} \) and \( \text{Succ} \) are fuzzy attributes on \( U_2 \) and \( U_3 \), respectively).

Table 2.

| Name | Int   | Succ |
|------|-------|------|
| \( t_1 \) (Emily) | \{(130, 0.5), (145, 0.5), (5, 0.5)\} | \{(5, 1)\} |
| \( t_2 \) (John) | \{(115, 0.5), (130, 0.5)\} | \{(5, 1)\} |

In Table 2, \( a_1 \in (0, 1) \) denotes the membership value of the element 130 \in U_2 to the fuzzy set \( \{(130, a_1), (145, a_2)\} \), etc., \( a_5 \in (0, 1) \) denotes the membership value of the element 115 \in U_2 to the fuzzy set \( \{(115, a_5)\} \).

The authors in [8] and [1], for example, apply fuzzy membership values to incorporate fuzzy data into relational database theory.

Note that the fuzzy relation instance \( r_1 \) may be represented as the vague relation instance given by Table 3.

Table 3.

| Name | Int   | Succ |
|------|-------|------|
| \( t_1 \) (Emily) | \{(130, 0.5), (145, 0.5), (5, 0.5)\} | \{(5, 1)\} |
| \( t_2 \) (John) | \{(115, 0.5), (130, 0.5)\} | \{(5, 1)\} |

Similarly, the relation instance \( r_2 \) on \( R(\text{Name}, \text{Int}, \text{Succ}) \) given by Table 4 (now, we assume that the attributes \( \text{Int} \) and \( \text{Succ} \) are ordinary attributes on \( U_2 \) and \( U_3 \), respectively), may be represented as the vague relation instance given by Table 5.

Table 4.

| Name | Int | Succ |
|------|-----|------|
| \( t_1 \) (Joe) | \{130\} | \{5\} |
| \( t_2 \) (Tina) | \{145\} | \{10\} |

For the ordinary relational database theory, see [16].

The aforementioned examples show clearly that the vague relation concept represents a natural generalization of the ordinary relation concept and the fuzzy relation concept. While the relation theory is not able to handle imprecise data almost at all, and the knowledge about fuzzy data has its own limitations, the quality of the information about vague data is obviously much more refined.

Let \( a_1 = \{t_{v_i}(u_1), 1 - t_{v_i}(u_1)\} \subseteq [0, 1] \) and \( a_2 = \{t_{v_2}(u_2), 1 - t_{v_2}(u_2)\} \subseteq [0, 1] \) be the vague values joined to \( u_1 \in U_1 \) and \( u_2 \in U_2 \), respectively, where \( V_i = \{u_i, t_{v_i}(u_i), 1 - t_{v_i}(u_i)\} : u_i \in U_i \) is a vague set in the universe of discourse \( U_i, i \in \{1, 2\} \).

We define the similarity measure \( SE(a_1, a_2) \) between the vague values \( a_1 \) and \( a_2 \) following Lu-Ng [12].

\[
SE(a_1, a_2) = \frac{1}{2} \sqrt{1 - \frac{\{t_{v_2}(u_2) - t_{v_1}(u_1)\} - \{f_{v_1}(u_1) - f_{v_2}(u_2)\}}{2}}.
\]

Note that \( SE(a_1, a_2) \in [0, 1] \).

Moreover, \( SE(a_1, a_2) = SE(a_2, a_1) \), \( SE(a_1, a_2) = 1 \) if and only if \( a_1 = a_2 \), and \( SE(a_1, a_2) = 0 \) if and only if \( a_1 = [0, 0], a_2 = [1, 1] \) or \( a_1 = [0, 1], a_2 = [a, a], a \in [0, 1] \) \( a_1 = [1, 1], a_2 = [0, 0] \) or \( a_1 = [a, a], a_2 = [0, 1], a \in [0, 1] \).

Note that several authors, including Chen [5], [6], Hong-Kim [7], Li-Xu [11], Szmidt-Kaczprzyk [15], Grzegorzewski [9], proposed various definitions of similarity measures between vague sets and distances between intuitionistic fuzzy sets. According to Lu-Ng [12], however, the similarity measure given above, reflects reality in a more appropriate manner when it comes to more general cases.

Let

\[
A = \{u, [t_A(u), 1 - t_A(u)] : u \in U\}
\]

and

\[
B = \{u, [t_B(u), 1 - t_B(u)] : u \in U\}
\]

be two vague sets in some universe of discourse \( U \).

We define the similarity measure \( SE(A, B) \) between the vague sets \( A \) and \( B \) as follows:

\[
SE(A, B) = \frac{1}{|U|} \sum_{u \in U} SE([t_A(u), 1 - t_A(u)], [t_B(u), 1 - t_B(u)])
\]

\[
= \frac{1}{|U|} \sum_{u \in U} \sqrt{1 - \frac{\{t_A(u) - t_B(u)\} - \{f_A(u) - f_B(u)\}}{2}}
\]

\[
\sqrt{1 - \{t_A(u) - t_B(u)\} + \{f_A(u) - f_B(u)\}}.
\]
Lemma 1 Let \( R(A_1, A_2, \ldots, A_n) \) be a relation scheme on domains \( U_1, U_2, \ldots, U_m \), where \( A_i \) is an attribute on the universe of discourse \( U_i \), \( i \in I \). Let \( r \) be a vague relation instance on \( R(A_1, A_2, \ldots, A_n) \). Suppose that \( X \subseteq [A_1, A_2, \ldots, A_n] \), then

\[
SE_X(t_1, t_2) \geq SE_X(t_1, t_2)
\]

for any \( t_1 \) and \( t_2 \) in \( r \).

Proof. Let \( t_1, t_2 \in r \). Since,

\[
\{SE(t_1[A], t_2[A]) : A \in Y\} \subseteq \{SE(t_1[A], t_2[A]) : A \in X\},
\]

it immediately follows that

\[
SE_X(t_1, t_2) = \min_{A \in X} \{SE(t_1[A], t_2[A])\} \geq \min_{A \in X} \{SE(t_1[A], t_2[A])\} = SE_X(t_1, t_2).
\]

\[\square\]

Lemma 2 Let \( R(A_1, A_2, \ldots, A_n) \) be a relation scheme on domains \( U_1, U_2, \ldots, U_m \), where \( A_i \) is an attribute on the universe of discourse \( U_i \), \( i \in I \). Let \( r \) be a vague relation instance on \( R(A_1, A_2, \ldots, A_n) \). Suppose that \( X = \{A_1, A_2, \ldots, A_k\} \), where \( X \) is a subset of \([A_1, A_2, \ldots, A_n] \). If \( SE_X(t_1, t_2) \geq \theta \) for all \( j \in \{1, 2, \ldots, k\} \), where \( t_1 \) and \( t_2 \) are some tuples in \( r \), then \( SE_X(t_1, t_2) \geq \theta \).

Proof. Since,

\[
SE_X(t_1, t_2) = \min_{A \in X} \{SE(t_1[A], t_2[A])\},
\]

the assertion follows.

\[\square\]

Lemma 3 Let \( R(A_1, A_2, \ldots, A_n) \) be a relation scheme on domains \( U_1, U_2, \ldots, U_m \), where \( A_i \) is an attribute on the universe of discourse \( U_i \), \( i \in I \). Let \( r \) be a vague relation instance on \( R(A_1, A_2, \ldots, A_n) \). Suppose that \( X = \{A_1, A_2, \ldots, A_n\} \). If \( SE_X(t_1, t_2) \geq \theta \) and \( SE_X(t_1, t_2) \geq \theta \), where \( t_1 \), \( t_2 \), and \( t_3 \) are some three, mutually distinct tuples in \( r \), and \( X \) is a subset of \([A_1, A_2, \ldots, A_n] \), then the inequality \( SE_X(t_1, t_2) \geq \theta \) does not necessarily hold true.

Proof. Let \( r \) be the vague relation instance on \( R(\text{Name}, \text{Int}, S\text{ucc}) \) given by Table 1. Suppose that \( X = \{\text{Int}, S\text{ucc}\} \). We have (see matrices \( I \) and \( S \)),

\[
SE_X(t_1, t_2) = \min_{A \in X} \{SE(t_1[A], t_2[A])\}.
\]

where, for example, 0.78 means that

\[
0.78 = SE(t_1[\text{Int}], t_2[\text{Int}]) = SE(t_2[\text{Int}], t_3[\text{Int}]) = \frac{1}{3} SE([0.7, 0.9], [0, 0]) + \frac{1}{3} SE([0.9, 0.95], [0.8, 0.9]) + \frac{1}{3} SE([0, 0.8], [0.95, 0.95])
\]
Now, \( SE_X(t_1, t_2) > \theta \) and \( SE_X(t_2, t_3) > \theta \), where \( \theta = 0.5 \) for example.

However, \( SE_X(t_1, t_3) < \theta \). This completes the proof. \( \square \)

Note that the fact that \( SE(t_1 [A], t_2 [A]) \in [0, 1] \) for all \( A \in X \), yields that \( SE_X(t_1, t_2) \in [0, 1] \).

Moreover,

\[
SE_X(t_1, t_2) = \min_{A \in X} \{ SE(t_1 [A], t_2 [A]) \} \\
= \min_{A \in X} \{ SE(t_2 [A], t_1 [A]) \} \\
= SE_X(t_2, t_1).
\]

Furthermore, \( SE_X(t_1, t_2) = 1 \) if and only if \( SE(t_1 [A], t_2 [A]) = 1 \) for all \( A \in X \).

Finally, \( SE_X(t_1, t_2) = 0 \) if and only if \( SE(t_1 [A], t_2 [A]) = 0 \) for all \( A \in X \).

Let \( R(A_1, A_2, ..., A_n) \) be a relation scheme on domains \( U_1, U_2, ..., U_n \), where \( A_i \) is an attribute on the universe of discourse \( U_i \), \( i \in I \). A fuzzy relation instance \( r \) on \( R(A_1, A_2, ..., A_n) \) is a subset of the cross product \( U_i \times U_j \times ... \times U_k \) of the power sets of the domains of the attributes.

A tuple \( t \) of \( r \) is then of the form \( (t_1 [A_1], t_2 [A_2], ..., t_n [A_n]) \), where \( t_i [A_i] \subseteq U_i \) for \( i \in I \), and where we assume that \( t_i [A_i] \neq \emptyset \) for \( i \in I \).

As we already noted, any fuzzy attribute value is described by some set of crisp values, where each of the crisp values is similar to the fuzzy value. More precisely, each attribute domain \( U_i \), \( j \in I \) is equipped with some similarity relation \( s_j : U_j \times U_j \rightarrow [0, 1] \), where \( s_j : U_j \times U_j \rightarrow [0, 1] \) is said to be a similarity relation on \( U_j \), if for every \( x, y, z \in U_j \), the conditions: \( s_j(x, x) = 1 \), \( s_j(x, y) = s_j(y, x) \), \( s_j(x, z) \geq \max_i \{ s_j(x, y_i), s_j(y, z_i) \} \) hold true.

Thus, while in the case of relational database model we were able to check if \( t_i [A_i] = t_k [A_k] \), now, in the case of fuzzy relational database model, we are able to define how conformant \( t_i [A_i] \) and \( t_k [A_k] \) are. In particular (see, [14]), if \( s_j : U_j \times U_j \rightarrow [0, 1] \) is a similarity relation on \( U_j \), and \( t_i [A_i] = d_j, t_k [A_k] = d_k \), then the conformance \( \varphi(A_i, t_i) \) of the attribute \( A_i \) on tuples \( t_j \) and \( t_k \) is defined by

\[
\varphi(A_i, t_j, t_k) = \min_{x \in d_j} \left\{ \max_{y \in d_k} \{ s_j(x, y) \} \right\},
\]

Hence, we calculate \( \varphi(A_i, t_j, t_k) \) instead of calculating \( \varphi(A_i, t_j[A_i], t_k[A_i]) \), i.e., instead of checking whether or not \( t_j[A_i] = t_k[A_i] \). Consequently, we calculate \( \varphi(X, t_j, t_k) \) instead of checking whether or not \( t_j[X] = t_k[X] \), which is the conformance of the attribute set \( X \) on tuples \( t_j \) and \( t_k \), defined by

\[
\varphi(X, t_j, t_k) = \min_{A \in X} \{ \varphi(A, t_j, t_k) \}.
\]

For the similarity-based fuzzy relational database approach, we refer to [2]-[4].

Now, the condition: \( t_j[X] = t_k[X] \) implies that \( t_j[Y] = t_k[Y] \), could be read as: if \( \varphi(X, t_j, t_k) \) resp. \( \varphi(Y, t_j, t_k) \) is the conformance of the attribute set \( X \) resp. \( Y \) on tuples \( t_j \) and \( t_k \), then \( \varphi(Y, t_j, t_k) \geq \varphi(X, t_j, t_k) \) for every pair of tuples \( t_j \) and \( t_k \) in \( r \), \( \varphi(Y, t_j, t_k) \geq \varphi(X, t_j, t_k) \).

Consider the following example.

Let \( R(Tea, Exp, Sal) \) be a relation scheme on domains \( U_1 = \{ Grace, Harry, Oscar \}, U_2 = \{ low, high \}, U_3 = \{ 3800USD, 4500USD \} \), where \( Exp \) (as experience) and \( Sal \) (as salary) are fuzzy attributes on universes \( U_2 \) and \( U_3 \), respectively, and \( Tea \) (as teachers) is ordinary attribute on the universe of discourse \( U_1 \).
Let $s_2 : U_2 \times U_2 \rightarrow [0, 1]$ be the similarity relation on $U_2$ defined by $s_2$(low, high) = 0.3, and $s_3 : U_3 \times U_3 \rightarrow [0, 1]$ be the similarity relation on $U_3$ defined by $s_3(3800USD, 4500USD) = 0.5$. Let $r_3$ be the fuzzy relation instance on $R(\text{Tea}, \text{Exp}, \text{Sal})$ given by Table 6.

| Tea   | Exp          | Sal         |
|-------|--------------|-------------|
| $t_1$ | [Grace]      | [low, high] | (3800USD, 4500USD) |
| $t_2$ | [Oscar]      | [low]       | (3800USD)          |

Consider the dependency: teachers with similar experiences should have similar salaries. Note that the values of the attributes: experience and salary may be imprecise. This fact, as well as the fact that the word similar is applied within dependency, imply that this dependency can be taken as an example of fuzzy functional dependency. It can be written in the form $X \rightarrow_{F} Y$, where $X = \{\text{Exp}\}$ and $Y = \{\text{Sal}\}$.

Let’s check if the fuzzy relation instance $r_3$ satisfies $X \rightarrow_{F} Y$. We obtain,

$$\varphi(X[t_1, t_2]) = \min_{A \in X} \{\varphi(A[t_1, t_2])\} = \varphi(\text{Exp}[t_1, t_2])$$

$$= \min \left\{ \min \left( \max \{s_2(\text{low}, \text{low})\}, \max \{s_2(\text{high}, \text{low})\} \right), \right.$$

$$\left. \min \left( \max \{s_2(\text{low}, \text{low}), s_2(\text{low}, \text{high})\} \right) \right\}$$

$$= \min \left\{ \min \{\max \{1, \max \{0.3\}\}, \min \{\max \{1, 0.3\}\}\} \right.$$

$$\left. = \min \{\max \{1, 0.3\}, \max \{0.3, 1\}\} \right\} = \min \{0.3, 1\} = 0.3,$$

$$\varphi(Y[t_1, t_2]) = \min_{A \in Y} \{\varphi(A[t_1, t_2])\} = \varphi(\text{Sal}[t_1, t_2])$$

$$= \min \left\{ \min \left( \max \{s_3(3800USD, 3800USD)\}, \right.$$

$$\left. \max \{s_3(3800USD, 4500USD)\} \right) \right\}$$

$$= \min \left\{ \min \{\max \{1, \max \{0.5\}\}, \min \{\max \{1, 0.5\}\}\} \right.$$

$$\left. = \min \{\max \{1, 0.5\}, \min \{1\}\} \right\} = \min \{0.5, 1\} = 0.5.$$

The condition $\varphi(Y[t_1, t_2]) \geq \varphi(X[t_1, t_2])$ is satisfied. This means that $r_3$ satisfies $X \rightarrow_{F} Y$.

Assume for a moment that $t_2[\text{Exp}] = \{\text{low}, \text{high}\}$. Now,

$$\varphi(X[t_1, t_2]) = \varphi(\text{Exp}[t_1, t_2])$$

$$= \min \left\{ \min \{\max \{s_2(\text{low}, \text{low})\}, s_2(\text{low}, \text{high})\}, \right.$$

$$\left. \min \left( \max \{s_2(\text{low}, \text{low}), s_2(\text{low}, \text{high})\} \right) \right\}$$

$$= \min \left\{ \min \{\max \{1, \max \{0.3\}\}, \min \{\max \{0.3, 1\}\}\} \right.$$

$$\left. = \min \{\max \{1, 1\}, \min \{1, 1\}\} \right\} = \min \{1, 1\} = 1.$$

In this case, $\varphi(Y[t_1, t_2]) = 0.5 < 1 = \varphi(X[t_1, t_2])$, i.e., $r_3$ violates $X \rightarrow_{F} Y$.

Note that the dependency: teachers with similar experiences should have similar salaries, tells the truth about the real-world. Obviously, both scenarios $t_2[\text{Exp}] = \{\text{low}\}$ and $t_2[\text{Exp}] = \{\text{low}, \text{high}\}$ are possible in reality. In the first case, Grace and Oscar have similar experiences and similar salaries, where the salaries are more similar than the experiences are. In the second case, their salaries are similar, but not identical, although their experiences are identical. This discussion shows that the dependency: teachers with similar experiences should have similar salaries, makes sense by itself, and that the instance $r_3$ should satisfy this dependency in both cases, $t_2[\text{Exp}] = \{\text{low}\}$ and $t_2[\text{Exp}] = \{\text{low}, \text{high}\}$. The inequalities $\varphi(Y[t_1, t_2]) \geq \varphi(X[t_1, t_2])$, $\varphi(Y[t_1, t_2]) < \varphi(X[t_1, t_2])$, where $t_2[\text{Exp}] = \{\text{low}\}$, $t_2[\text{Exp}] = \{\text{low}, \text{high}\}$, respectively, tell us, however, that the condition $\varphi(Y[t_1, t_2]) \geq \varphi(X[t_1, t_2])$ is not adequate for determining whether or not the fuzzy relation instance $r$ satisfies the fuzzy functional dependency $X \rightarrow_{F} Y$. If this condition is satisfied, the instance $r$ satisfies the dependency $X \rightarrow_{F} Y$ for sure. Otherwise, if the condition fails, the instance $r$ may or may not satisfy $X \rightarrow_{F} Y$.

In order to overcome these difficulties and correct the irregularities, Sozat and Yazici [14] introduced the following definition.

Let $R(A_1, A_2, \ldots, A_n)$ be a relation scheme on domains $U_1, U_2, \ldots, U_n$, where $A_i$ is an attribute on the universe of discourse $U_i$, $i \in I$. Suppose that $r$ is a fuzzy relation instance on $R(A_1, A_2, \ldots, A_n)$. Furthermore, let $X$ and $Y$ be subsets of $\{A_1, A_2, \ldots, A_n\}$, and $\theta \in [0, 1]$. Fuzzy relation instance $r$ is said to satisfy the fuzzy functional dependency $X \rightarrow_{F} Y$, if for every pair of tuples $t_1$ and $t_2$ in $r$, $\varphi(Y[t_1, t_2]) \geq \min_{\theta}(\varphi(X[t_1, t_2]))$.
Thus, if it happens that $\varphi (Y [t_1, t_2]) \geq \varphi (X [t_1, t_2])$ for $t_1, t_2 \in r$, then

$$\varphi (Y [t_1, t_2]) \geq \varphi (X [t_1, t_2])$$

$$\geq \min \{ \theta, \varphi (X [t_1, t_2]) \}$$

for every $t_1, t_2 \in r$, and $\theta \in [0, 1]$, i.e., $r$ satisfies $X \rightarrow^\theta Y$ for every $\theta \in [0, 1]$.

More generally, if it happens that for every $t_1, t_2 \in r$, either $\varphi (Y [t_1, t_2]) \geq \varphi (X [t_1, t_2])$ or $\varphi (Y [t_1, t_2]) \geq \varphi (X [t_1, t_2])$, then

$$\varphi (Y [t_1, t_2]) \geq \min \{ \theta, \varphi (X [t_1, t_2]) \}$$

for every $t_1, t_2 \in r$, i.e., $r$ satisfies $X \rightarrow^\theta Y$.

Consequently, if it happens that for some $t_1, t_2 \in r$, $\varphi (Y [t_1, t_2]) < \varphi (X [t_1, t_2])$, and $\varphi (Y [t_1, t_2]) < \varphi (X [t_1, t_2])$, then the instance $r$ violates the dependency $X \rightarrow^\theta Y$.

In particular, the instance $r_3$ satisfies the dependency $X \rightarrow^\theta_Y$ for every $\theta \in [0, 1]$ (see, Table 6). Furthermore, if $t_2 [Exp] = [low, high]$, then the instance $r_3$ satisfies resp. the conditions $S E Y (t_1, t_2) \geq \min \{ \theta, S E X (t_1, t_2) \}$.

The value $\theta \in [0, 1]$ that appears in the notation $X \rightarrow^\theta Y$ is called the linguistic strength of the fuzzy functional dependency. If $\theta = 1$, the fuzzy functional dependency $X \rightarrow^\theta Y$ becomes $X \rightarrow Y$.

Now, one could try to say that some vague relation instance $r$ on $R (A_1, A_2, \ldots, A_n)$ satisfies the vague functional dependency $X \rightarrow^\theta_Y$, if for every pair of tuples $t_1$ and $t_2$ in $r$, $S E Y (t_1, t_2) \geq \min \{ \theta, S E X (t_1, t_2) \}$ (see, e.g., [12, 17]).

Recall the vague relation instance $r$ given by Table 1.

Consider the dependency: $\text{the intelligence level of a person more or less determines the degree of success.}$

Since the values of the attributes $\text{intelligence}$ and $\text{success}$ may be imprecise, we may consider this dependency as a vague functional dependency. We can write it in the form $X \rightarrow_{V Y} Y$, where $X = \{ \text{Int} \}$ and $Y = \{ \text{Suc} \}$ (see, Table 1).

Let’s check if the vague relation instance $r$ satisfies $X \rightarrow_{V Y} Y$.

Since (see, matrices $I$ and $S$),

$$S E Y (t_3, t_4) = S E (t_3 [S ucc], t_4 [S ucc]) = 0.13$$

and

$$S E X (t_3, t_4) = S E (t_3 [I n t], t_4 [I n t]) = 0.46,$$

it follows that the instance $r$ violates $X \rightarrow_{V Y} Y$. The vague functional dependency: $\text{the intelligence level of a person more or less determines the degree of success}$, however, tells the truth about the real-world in the same way the fuzzy functional dependency: $\text{teachers with similar experiences should have similar salaries}$ does. Moreover, the scenario presented by the vague relation instance $r$ (Table 1), makes sense in reality. This actually means that the instance $r$ should somehow satisfy the dependency $X \rightarrow_{V Y} Y$.

Reasoning as in the case of fuzzy functional dependencies, we conclude that the condition $S E Y (t_1, t_2) \geq S E X (t_1, t_2), t_1, t_2 \in r$, must be adapted. We introduce the following definition.

Let $R (A_1, A_2, \ldots, A_n)$ be a relation scheme on domains $U_1, U_2, \ldots, U_n$, where $A_i$ is an attribute on the universe of discourse $U_i$, $i \in I$. Suppose that $r$ is a vague relation instance on $R (A_1, A_2, \ldots, A_n)$. Furthermore, let $X$ and $Y$ be subsets of $\{ A_1, A_2, \ldots, A_n \}$, and $\theta \in [0, 1]$. Vague relation instance $r$ is said to satisfy the vague functional dependency $X \rightarrow_{V Y} Y$, if for every pair of tuples $t_1$ and $t_2$ in $r$, $S E Y (t_1, t_2) \geq \min \{ \theta, S E X (t_1, t_2) \}$.

Thus, if $S E Y (t_1, t_2) \geq S E X (t_1, t_2)$ for $t_1, t_2 \in r$, the instance $r$ satisfies $X \rightarrow_{V Y} Y$ for $\theta \in [0, 1]$. If for every pair of tuples $t_1$ and $t_2$ in $r$, $S E Y (t_1, t_2) \geq S E X (t_1, t_2)$ or $S E Y (t_1, t_2) \geq S E X (t_1, t_2)$, then the instance $r$ satisfies $X \rightarrow_{V Y} Y$. Finally, if for some $t_1, t_2 \in r$, the conditions $S E Y (t_1, t_2) \leq S E X (t_1, t_2)$ and $S E Y (t_1, t_2) \leq \theta$ hold true, the instance $r$ violates $X \rightarrow_{V Y} Y$.

Now, the vague relation instance $r$ given by Table 1, satisfies resp. violates the vague functional dependency $X \rightarrow_{V Y} Y$ (\{Int\} $\rightarrow_{V Y} \{ \text{Suc} \}$), if $\theta \in [0, 0.13]$ resp. $\theta \in [0.13, 1]$.

If $\theta = 1$, the vague functional dependency $X \rightarrow_{V Y} Y$ becomes $X \rightarrow Y$.

For yet another definition of vague functional dependency, called $\alpha$-vague functional dependency, see [13].

### 3 Soundness of inference rules for vague functional dependencies

The following rules are the inference rules for vague functional dependencies (VFDs).

**VF1** Inclusive rule for VFDs: If $X \rightarrow_{V Y} Y$ holds, and $\theta_1 \geq \theta_2$, then $X \rightarrow_{V Y} Y$ holds.

**VF2** Reflexive rule for VFDs: If $X \geq Y$, then $X \rightarrow_{V Y} Y$ holds.

**VF3** Augmentation rule for VFDs: If $X \rightarrow_{V Y} Y$ holds, then $X Z \rightarrow_{V Y} Y Z$ holds.

**VF4** Transitivity rule for VFDs: If $X \rightarrow_{V Y} Y$ holds and $Y \rightarrow_{V Y} Z$ hold true, then $X \rightarrow_{\min(\theta_1, \theta_2)} Y Z$ holds true.

Here, $XZ$ means $X \cup Z$.

**Theorem 4** The inference rules: VF1, VF2, VF3 and VF4 are sound.

**Proof.** (proof for VF1) We must prove that $r$ satisfies $X \rightarrow_{V Y} Y$, if $r$ is any vague relation instance on $R (A_1, A_2, \ldots, A_n)$ which satisfies the vague functional dependency $X \rightarrow_{V Y} Y$. Theorem 4
Suppose that $r$ satisfies $X \overset{\theta_1}{\rightarrow} Y$, but violates $X \overset{\theta_2}{\rightarrow} Y$.

Since $r$ violates $X \overset{\theta_2}{\rightarrow} Y$, we know that there are tuples $t_1$ and $t_2$ in $r$, such that

$$SE_Y(t_1, t_2) < \min \{ \theta_2, SE_X(t_1, t_2) \}.$$ 

Therefore, $SE_Y(t_1, t_2) < \theta_2$ and $SE_Y(t_1, t_2) < SE_X(t_1, t_2)$.

The instance $r$ satisfies the dependency $X \overset{\theta_2}{\rightarrow} Y$.

Hence,

$$SE_Y(t_1, t_2) \geq \min \{ \theta_1, SE_X(t_1, t_2) \}.$$ 

If $\min \{ \theta_1, SE_X(t_1, t_2) \} = \theta_1$, then

$$SE_Y(t_1, t_2) \geq \theta_1 \geq \theta_2.$$ 

This contradicts the fact that $SE_Y(t_1, t_2) < \theta_2$.

If $\min \{ \theta_1, SE_X(t_1, t_2) \} = SE_X(t_1, t_2)$, then

$$SE_Y(t_1, t_2) \geq SE_X(t_1, t_2).$$ 

This contradicts the fact that $SE_Y(t_1, t_2) < SE_X(t_1, t_2)$.

Consequently, $r$ satisfies $X \overset{\theta_2}{\rightarrow} Y$.

The cases VF2-VF4 are discussed similarly. This completes the proof. \hfill $\Box$

### 4 Soundness of additional inference rules for vague functional dependencies

The following inference rules are additional inference rules for vague functional dependencies. We shall prove that these rules follow from the rules: VF1, VF2, VF3 and VF4. This will mean that the vague functional dependencies obtained by the additional rules can certainly be obtained by successive application of the rules: VF1, VF2, VF3 and VF4. The additional inference rules, however, can make such an effort much shorter and easier.

**VF5** Union rule for VFDs: If $X \overset{\theta_1}{\rightarrow} Y$ and $X \overset{\theta_2}{\rightarrow} Z$ hold true, then $X \overset{\min(\theta_1, \theta_2)}{\rightarrow} YZ$ holds also true.

**VF6** Pseudo-transitivity rule for VFDs: If $X \overset{\theta_1}{\rightarrow} Y$ and $WY \overset{\theta_2}{\rightarrow} Z$ hold true, then $WX \overset{\min(\theta_1, \theta_2)}{\rightarrow}YZ$ holds true.

**VF7** Decomposition rule for VFDs: If $X \overset{\theta}{\rightarrow} Y$ holds, and $Z \subseteq Y$, then $X \overset{\theta}{\rightarrow} Z$ also holds.

**Theorem 5** The inference rules: VF5, VF6 and VF7 are sound.

**Proof.** (proof I for VF5) Let $r$ be any vague relation instance (on $R(A_1, A_2, ..., A_n))$ which satisfies $X \overset{\theta_1}{\rightarrow} Y$ and $X \overset{\theta_2}{\rightarrow} Z$. We shall prove that $r$ satisfies $X \overset{\min(\theta_1, \theta_2)}{\rightarrow} YZ$.

Suppose that $r$ violates $X \overset{\min(\theta_1, \theta_2)}{\rightarrow} YZ$.

Now, there exist tuples $t_1$ and $t_2$ in $r$, such that

$$SE_{YZ}(t_1, t_2) < \min (\min(\theta_1, \theta_2), SE_X(t_1, t_2)).$$

It follows that $SE_{YZ}(t_1, t_2) < \theta_1$, $SE_{YZ}(t_1, t_2) < \theta_2$, $SE_{YZ}(t_1, t_2) < SE_X(t_1, t_2)$.

Since $r$ satisfies $X \overset{\theta_1}{\rightarrow} Y$ and $X \overset{\theta_2}{\rightarrow} Z$, we have that

$$SE_Y(t_1, t_2) \geq \min (\theta_1, SE_X(t_1, t_2)),$$

$$SE_{Z}(t_1, t_2) \geq \min (\theta_2, SE_X(t_1, t_2)).$$

We deduce,

$$SE_{YZ}(t_1, t_2) = \min_{\theta \in YZ} \{ SE(t_1[A], t_2[A]) \},$$

$$\min_{\theta \in YZ} \{ SE(t_1[A], t_2[A]) \} = \min (\min (\theta_1, SE_X(t_1, t_2)), \min (\theta_2, SE_X(t_1, t_2))) = \min (\theta_1, \theta_2, SE_X(t_1, t_2)) = \min (\theta_1, \theta_2, SE_X(t_1, t_2)).$$

The obtained inequality contradicts the fact that

$$SE_{YZ}(t_1, t_2) < \min (\theta_1, \theta_2), SE_X(t_1, t_2)).$$

Therefore, $r$ satisfies $X \overset{\min(\theta_1, \theta_2)}{\rightarrow} YZ$.

(proof II for VF5) We have:

1) $X \overset{\theta_1}{\rightarrow} Y$ (input),
2) $X \overset{\theta_2}{\rightarrow} Z$ (input),
3) $X \overset{\theta_1}{\rightarrow} XY$ (VF3, 1) augment with $X$),
4) $XY \overset{\theta_2}{\rightarrow} YZ$ (VF3, 2) augment with $Y$),
5) $X \overset{\min(\theta_1, \theta_2)}{\rightarrow} YZ$ (VF4, 3, 4)).

The cases VF6 and VF7 are discussed similarly. This completes the proof. \hfill $\Box$

### 5 Completeness of inference rules for vague functional dependencies

Let $R(A_1, A_2, ..., A_n)$ be a relation scheme on domains $U_1, U_2, ..., U_n$, where $A_i$ is an attribute on the universe of discourse $U_i$, $i \in I$.

Suppose that $V$ is some set of vague functional dependencies on $[A_1, A_2, ..., A_n]$. The closure $V^*$ of $V$ is the set of all vague functional dependencies that can be derived from $V$ by repeated applications of the inference rules: VF1, VF2, VF3 and VF4.

Note that the set $V^*$ is infinite one regardless of whether the set $V$ is finite or not. Namely, if $X \overset{\theta}{\rightarrow} Y$
Belongs to $V$, then, by VF1, $X \xrightarrow{\theta} Y$ belongs to $V^*$ for all $\theta \in \{0, \theta\}$.

Let $X \xrightarrow{\theta} Y$ be some vague functional dependency on $\{A_1, A_2, \ldots, A_n\}$. The dependency $X \xrightarrow{\theta} Y$ may or may not belong to $V^*$. The limit strength of $X \xrightarrow{\theta} Y$ (with respect to $V^*$) is the number $\theta(X) \in [0, 1]$, such that $X \xrightarrow{\theta(X)} Y$ belongs to $V^*$, and $\theta \leq \theta(X)$ for each $X \xrightarrow{\theta} Y$ that belongs to $V^*$.

If $X \xrightarrow{\theta} Y$ belongs to $V^*$, then the limit strength $\theta(X)$ obviously exists. Otherwise, if $X \xrightarrow{\theta} Y$ does not belong to $V^*$, the limit strength $\theta(X)$ does not necessarily exist.

Let $X$ be a subset of $\{A_1, A_2, \ldots, A_n\}$, and $\theta$ be a number in $[0, 1]$. The closure $X^+ (\theta, V)$ of $X$ (with respect to $V^*$) is the set of attributes $A \in \{A_1, A_2, \ldots, A_n\}$, such that $X \xrightarrow{\theta} A$ belongs to $V^*$.

Now, if $A \in X$, then, by VF2, $X \xrightarrow{\theta} A$ belongs to $V^*$. Hence, by VF1, $X \xrightarrow{\theta} A$ belongs to $V^*$. Therefore, $A \in X^+ (\theta, V)$. Since $A \in X$, we conclude that $X \subseteq X^+ (\theta, V)$.

Theorem 6 Let $R (A_1, A_2, \ldots, A_n)$ be a relation scheme on domains $U_1, U_2, \ldots, U_n$, where $A_i$ is an attribute on the universe of discourse $U_i$, $i \in I$. Let $V^*$ be the closure of $V$, where $V^*$ is some set of vague functional dependencies on $\{A_1, A_2, \ldots, A_n\}$. Suppose that $X \xrightarrow{\theta} Y$ is some vague functional dependency on $\{A_1, A_2, \ldots, A_n\}$. Then, $X \xrightarrow{\theta} Y$ belongs to $V^*$ if and only if $Y \subseteq X^+ (\theta, V)$.

Proof. ($\Rightarrow$) Suppose that $X \xrightarrow{\theta} Y$ belongs to $V^*$.

Now, by VF7, $X \xrightarrow{\theta} A$ belongs to $V^*$ for every $A \in Y$.

Hence, $A \in X^+ (\theta, V)$ for every $A \in Y$, i.e., we have that $Y \subseteq X^+ (\theta, V)$.

($\Leftarrow$) Let $Y \subseteq X^+ (\theta, V)$.

Now, $A \in X^+ (\theta, V)$ for every $A \in Y$.

Therefore, $X \xrightarrow{\theta} A$ belongs to $V^*$ for every $A \in Y$.

Hence, by V5, $X \xrightarrow{\theta} Y$ belongs to $V^*$.

This completes the proof. \qed

Theorem 7 The set $\{VF1, VF2, VF3, VF4\}$ is complete set.

Proof. Let $R (A_1, A_2, \ldots, A_n)$ be a relation scheme on domains $U_1, U_2, \ldots, U_n$, where $A_i$ is an attribute on the universe of discourse $U_i$, $i \in I$.

Let $V^*$ be the closure of $V$, where $V^*$ is some set of vague functional dependencies on $\{A_1, A_2, \ldots, A_n\}$.

Suppose that $X \xrightarrow{\theta} Y$ is some vague functional dependency on $\{A_1, A_2, \ldots, A_n\}$ which is not a member of $V^*$.

In order to prove the theorem, it is enough to prove that there exists a vague relation instance $r$ on $R (A_1, A_2, \ldots, A_n)$ which satisfies $A \xrightarrow{\theta} B$ if $A \xrightarrow{\theta} B$ belongs to $V^*$, and violates $X \xrightarrow{\theta} Y$.

$r$ can be constructed in the following way.

Fix some $\theta' \in \{\theta^*, \theta\}$, where

$$\theta^* = \max \{\theta_i (V)\}.$$ \hspace{1cm} (5)

Now, if $A \xrightarrow{\theta} B$ is a vague functional dependency whose limit strength $\theta_i (V)$ is less than $\theta$, then $1/\theta_i (V) < \theta$. Otherwise, if $\theta_i (V) \geq \theta$, then $\theta' < \theta$.

Suppose that $U_1 = U_2 = \ldots = U_n = \{u\}$. Let

$$V_1 = \{(u, [t_1, (u, 1 - f_1 (u)]) : u \in U\} = \{(u, [t_1, (u, 1 - f_1 (u)]) : u \in U\}$$

and

$$V_2 = \{(u, [t_2, (u, 1 - f_2 (u)]) : u \in U\} = \{(u, [t_2, (u, 1 - f_2 (u)]) : u \in U\}$$

be two vague sets in $U$, such that

$$SE (a, b) = \theta'.$$

Hence,

$$SE (V_1, V_2) = SE (a, b) = \theta'.$$

Let $X^+ (\theta, V)$ be the closure of $X$.

Now, let $r = \{t_1, t_2\}$ be the vague relation instance on $R (A_1, A_2, \ldots, A_n)$ given by Table 7.

| Table 7. | attributes of $X^+ (\theta, V)$ | other attributes |
| --- | --- | --- |
| $t_1$ | $V_1, V_1, \ldots, V_1$ | $V_1, V_1, \ldots, V_1$ |
| $t_2$ | $V_1, V_1, \ldots, V_1$ | $V_2, V_2, \ldots, V_2$ |

It is not so hard to prove that $r$ satisfies $A \xrightarrow{\theta} B$ if $A \xrightarrow{\theta} B$ belongs to $V^*$, and violates $X \xrightarrow{\theta} Y$.

This completes the proof. \qed

References

[1] J. F. Baldwin, Knowledge engineering using a fuzzy relational inference language, IFAC Proceedings Volumes 16, 14-20 (1983)

[2] B. P. Buckles, F. E. Petry, A fuzzy representation of data for relational databases, Fuzzy Sets and Systems 7, 213-216 (1982)

[3] B. P. Buckles, F. E. Petry, Fuzzy databases and their applications, Fuzzy Inform. Decision Processes 2, 361-371 (1982)

[4] B. P. Buckles, F. E. Petry, Uncertainty models in information and database systems, J. Inform. Sci. 11, 77-87 (1985)

[5] S. M. Chen, Similarity Measures Between Vague Sets and Between Elements, IEEE Transactions on Systems, Man and Cybernetics 27, 153-159 (1997)
[6] S. M. Chen, Measures of Similarity Between Vague Sets, Fuzzy Sets and Systems 74, 217-223 (1995)
[7] D. H. Hong, C. Kim, A note on Similarity Measures Between Vague Sets and Between Elements, Information Sciences 115, 83-96 (1999)
[8] C. Giardina, I. Sack, D. Sinha, Fuzzy Field Relational Database Tech. Report 8332 (Elect. Engng. and Computer Science Dept., Stevens Institute of Technology, Hoboken, NJ, 1983)
[9] P. Grzegorzewski, Distances Between Intuitionistic Fuzzy Sets and/or Interval-valued Fuzzy Sets Based on the Hausdorff Metric, Fuzzy Sets and Systems 148, 319-328 (2004)
[10] M. Levene, G. Loizou, A Guided Tour of Relational Databases and Beyond (Springer-Verlag, London, 1999)
[11] F. Li, Z. Xu, Measures of Similarity Between Vague Sets, Journal of Software 12, 922-927 (2001)
[12] A. Lu, W. Ng, Managing Merged Data by Vague Functional Dependencies, In: Atzeni, P., Chu, W., Lu, H., Zhou, S., Ling, T.-W. (eds) ER 2004 LNCS, vol. 3288, pp. 259-272, Springer-Verlag, Berlin Heidelberg, (2004)
[13] J. Mishra, S. Ghosh, A New Functional Dependency in a Vague Relational Database Model, International Journal of Computer Applications 39, 29-36 (2012)
[14] M. Sozat, A. Yazici, A complete axiomatization for fuzzy functional and multivalued dependencies in fuzzy database relations, Fuzzy Sets and Systems 117, 161-181 (2001)
[15] E. Szmidt, J. Kacprzyk, Distances Between Intuitionistic Fuzzy Sets, Fuzzy Sets and Systems 114, 505-518 (2000)
[16] J. D. Ullman, Principles of Database Systems (Computer Science Press, Rockville, MD, 1982)
[17] F. Zhao, Z. M. Ma, Functional Dependencies in Vague Relational Databases, In: IEEE International Conference on Systems, Man, and Cybernetics, vol. 5, pp. 4006-4010, Taipei, Taiwan, (2006)