Fuzzy functional dependencies and direct bases

J.M. Rodríguez-Jiménez, E. Rodríguez-Lorenzo, Karell Bertet[2], P. Cordero, M. Enciso, and A. Mora

Universidad de Málaga, Andalucía Tech, Spain,
e-mail: {jmrodriguez,estrellarodlor,amora}@ctima.uma.es, {pcordero,enciso}@uma.es

Abstract. This work is devoted to a fuzzy extension of the notion of functional dependency that is defined over datatables with fuzzy values and based on similarity relations. Specifically, we propose the first method to calculate direct-optimal bases for fuzzy functional dependencies over fuzzy attribute tables and domains with similarity relations.

1 Introduction

Functional dependency (FD) is a well known concept in the database area. The importance of this notion must be found in the normalization theory, the core of the relational model [5]. In [8], we introduced a logic for the management of functional dependencies, named Simplification Logic ($\mathbf{SL}_{\mathrm{FD}}$), which constitutes a solid alternative to Armstrong's Axioms. The main novelty of $\mathbf{SL}_{\mathrm{FD}}$ is that it is strongly based on the Simplification Rule, which allows us to narrow the implications by removing attributes. This new axiomatic system has provided the definition of true automated deduction methods to solve functional dependency problems [14].

Fuzzy Set Theory has provided a generalization of the relational model in order to incorporate to the data some degree of uncertainty, vagueness, truthlikeness, incompleteness and imprecision. In [2, 17–19], the authors have presented some extensions of the relational model, several definitions of FFDs and complete axiomatic systems to reason with them. Nevertheless, these axiomatic systems are not designed to build automated methods to make inference with fuzzy functional dependencies. Besides that, there are two dimensions to categorize [11] the wide range of FFD definitions: the way in which vagueness is incorporated to the data and level of fuzzification of the functional dependency.

In [6, 7] we have introduced two definitions of FFD together with the corresponding sound and complete axiomatic systems and their automated reasoning methods. Both definitions incorporate different levels of fuzzification of the dependency while data remain crisp. In [11], we establish this classification and introduce a general FFD definition. In the relational model, the atomic element is the attribute value. Our approach associates a degree to each value of the attribute, providing the maximum level of uncertainty in tables. A sound and complete axiomatic system for these dependencies is also introduced in [11] and an automated reasoning method for this approach is introduced in [16]. Once the automated reasoning method has been provided, the notion of normal forms becomes to be the spotlight. It would be interesting a definition of a normal form for FFD that ensures an smaller cost for the reasoning methods considering some minimality criteria. For the classical case (crisp functional dependencies or attribute implications) different answers have been provided, depending of the notion of minimality and also depending of the environment where the implication notion is used [12, 13]. K. Bertet and B. Monjardet, in [4], established the equivalence of five definitions presented by different authors in several areas which correspond with the same notion of basis. Other well-known property used to define another kind of bases is *directness*, i.e., a single traversal of the implicational system is enough to compute the closure of a given set of attributes. A basis fulfilling this property is named *direct basis*. This property is usually accompanied by some minimality criteria. We are particularly interested in those ones with minimum size (number of attributes). In [4] a method to calculate the direct-optimal basis for classical attribute implications is introduced.

In this work, we show the notion of fuzzy attribute table and the fuzzy extension of the notion of functional dependency based on similarities in Section 3. Section 4 focusses on the Fuzzy Simplification Logic ($\mathbb{FSL}$) introduced in [11]. Then, in Section 5, the necessity of a kind of normal form (named direct optimal basis) is justified in order to improve de efficiency of the automated reasoning method. Finally, in Section 6, a method to compute direct optimal basis.

2 Preliminaries

We assume that basic notions related to Fuzzy Logic are known. In this framework, it is usual to replace the truthfulness value set $\{0, 1\}$ (false and true) by the interval $[0, 1]$ (*truth degrees*) enriched with some operations. Our approach uses the unit interval $[0, 1]$, the infimum (denoted by $\wedge$) as the universal quantifier, the supremun (denoted by $\vee$) as the existential quantifier, an arbitrary left-continuous t-norm (denoted by $\otimes$) as the conjunction and the residuum defined by $a \to b = \sup\{x \in [0, 1] \mid x \otimes a \leq b\}$. That is, the system of truth values is the residuated complete lattice $([0, 1], \vee, \wedge, 0, 1, \otimes, \to)$ where $([0, 1], \otimes, 1)$ is a commutative monoid and $(\otimes, \to)$ is an adjoin pair ($a \otimes b \leq c$ iff $a \leq b \to c$). We also use fuzzy sets in the standard way. Thus, for instance, the union of two fuzzy sets $A, B \in [0, 1]^{\Omega}$ is the fuzzy set such that $(A \cup B)(x) = A(x) \vee B(x)$. Since we work with fuzzy sets with finite support, we denote each fuzzy set A by its graph $\{(x, A(x)) \mid x \in \Omega, A(x) > 0\}$.

On the other hand, we show the basic concepts of the Relational Model, with emphasis on functional dependencies. Given a family of sets $\{D_a \mid a \in \Omega\}$, named *domains*, indexed in a finite non-empty set Ω of elements, named *attributes*, a relation is a subset of the cartesian product of the domains $R \subseteq \mathbb{D} = \prod_{a \in \Omega} D_a$. The elements in this product $t = (t_a)_{a \in \Omega} \in \mathbb{D}$ will be named *tuples*.

Now, some issues concerning the database notation are summarized: Given $A, B \subseteq \Omega$, AB denotes $A \cup B$ and D_A denotes $\prod_{a \in A} D_a$. Let $t \in R$ be a tuple, then $t_{/A}$ denotes the *projection* of t to D_A: if $t = (t_a)_{a \in \Omega}$ then $t_{/A} = (t_a)_{a \in A}$.

Definition 1. *A formula* $A \mapsto B$*, where* $A, B \subseteq \Omega$*, is named a* functional dependency *(FD). A relation* $R \subseteq \mathbb{D}$ *is said to satisfy* $A \mapsto B$ *if, for all tuples* $t_1, t_2 \in R$*,* $t_{1/A} = t_{2/A}$ *implies that* $t_{1/B} = t_{2/B}$*.*

3 Fuzzy Attributes Tables and FFDs

In the literature, some very similar definitions of fuzzy functional dependency have been proposed [1, 15, 17–19]. The first step in order to fuzzify the model is considering similarity relations instead of the equality. Thus, each domain D_a is provided with a similarity relation $\rho_a : D_a \times D_a \to [0, 1]$, that is, a reflexive ($\rho_a(x, x) = 1$ for all $x \in D_a$) and symmetric ($\rho_a(x, y) = \rho_a(y, x)$ for all $x, y \in D_a$) fuzzy relation. Given $A \subseteq \Omega$, the extension to the set $\mathbb{D}$ is the following: for all $t, t' \in \mathbb{D}$, $\rho_A(t, t') = \bigwedge_{a \in A} \rho_a(t_a, t'_a)$. A suitable definition of fuzzy functional dependency based on these similarities is the following [10]: A *fuzzy functional dependency* (FFD) is an expression $A \xrightarrow{\vartheta} B$ where $A, B \subseteq \Omega$ and $\vartheta \in [0, 1]$ and we say that the FFD hols in a relation $R \subseteq \mathbb{D}$ if,

$$\vartheta \leq \bigwedge_{t,t' \in R} \rho_A(t, t') \to \rho_B(t, t') \tag{1}$$

However, although similarities are used in the definition of functional dependency, the table definition remains classical. The same occurs in most of the fuzzy extensions of the functional dependency in the literature [15, 18, 19].

Since the value of each attribute is the atomic element in the classical relational model, if we would like to introduce uncertainty at the ground level, the values assigned to each attribute in each tuple have to be capable to be fuzzified. Thus, we propose to introduce a rank associated to each value which indicates the truthfulness degree of the value of this attribute in this tuple. This extension of the classical relational table is named *Fuzzy Attributes Tables* and constitutes a generalization of other fuzzy data tables appeared in the literature [15, 18, 19]. That is, for each tuple $t = (t_a)_{a \in \Omega} \in \mathbb{D}$, we consider a map $R \colon \mathbb{D} \to [0, 1]^{\Omega}$ or, equivalently $R \colon \mathbb{D} \times \Omega \to [0, 1]$, which renders a tuple of truth values $R(t) = (r_a)_{a \in \mathcal{A}}$.

For each tuple t, t_a denotes the value of the attribute a in the tuple t and $R(t)(a)$ is the truthfulness of the value t_a. We would like to remark that, it is possible that $R(t)(a) = 0$ for all attribute $a \in \Omega$.

When we work with Fuzzy Attributes Tables, we also consider similarity relations in domains in the same way as previous works [6,9,10]. Fuzzy Attributes Table is an extension of the original table in the classical relational model by adding the degree of certainty to the values of each attribute.

Example 1. We consider a table to store some patients with a mark from infection in their skin. The table is built with the set of attributes $\mathcal{A} = \{n, a, p, c, l\}$ where n represent the name, a the age, p the percent of extension in the skin, c the mark color, l the localization in the skin.

The domain of the attributes are $D_n = \{\text{John}, \text{Peter}, \text{Ann}, \text{Dave}, \text{Peter}\}$, $D_a = \{n \in \mathbb{N} \mid 0 \leq n \leq 120\}$, $D_p = \{n \in \mathbb{N} \mid 0 \leq n \leq 100\}$, $D_c = \{$ Black, Brown, Purple, Red, Yellow $\}$ and $D_l = \{$Arm, Face, Foot, Hand, Leg$\}$. We build the similarity relations in each domain D_a as follows:

ρ_c	b	w	p	r	y
b	1	0.7	0.4	0.3	0.1
w	0.7	1	0.5	0.4	0.2
p	0.4	0.5	1	0.8	0.2
r	0.3	0.4	0.8	1	0.1
y	0.1	0.2	0.2	0.1	1

ρ_l	a	f	o	h	l
a	1	0.2	0.4	0.9	0.4
f	0.2	1	0.2	0.2	0.2
o	0.4	0.2	1	0.6	0.8
h	0.9	0.2	0.6	1	0.3
l	0.4	0.2	0.8	0.3	1

$$\rho_n(t_n, t'_n) = \begin{cases} 1 & \text{if } t_n = t'_n \\ 0 & \text{if } t_n \neq t'_n \end{cases} \qquad \rho_a(t_a, t'_a) = 1 - \frac{|t_a - t'_a|}{max_a - min_a} \qquad \rho_p(t_p, t'_p) = 1 - \frac{|t_p - t'_p|}{100}$$

And finally, we consider the following Fuzzy Attributes Table.

name	age	percent	colour	localization
Ann/1	26/0.9	2/0.8	Brown/0.8	Hand/0.8
Albert/1	33/0.7	4/0.7	Black/0.6	Leg/0.9
Mary/1	21/0.6	7/0.6	Purple/0.9	Arm/0.9
Dave/1	43/0.4	4/0.9	Yellow/0.8	Foot/0.8
Peter/1	24/0.1	3/0.7	Brown/0.7	Arm/0.6

The information represented corresponds to patient's names and the description of their infection. Note that the name attribute is crisp, as a particular case.

The next step to provide a notion of Fuzzy attribute table over domains with similarity relations is to relate the similarity degree between two values of an attribute (in two tuples) with their truthfulness degree.

Definition 2. *Thus, let R be a fuzzy attributes table and let $t, t' \in R$ two tuples, for all $a \in \Omega$, the relative similarity degree is defined as*

$$\rho_a^R(t, t') = (R(t)(a) \otimes R(t')(a)) \to \rho_a(t_a, t'_a)$$

The above definition may be generalized to subsets of attributes $A \subseteq \Omega$ as usual, that is, $\rho_A^R(t, t') = \bigwedge_{a \in A} \rho_a^R(t, t')$

The definition of the relative similarity relation presented in Definition 2 may be used in Equation (1) so that the definition of fuzzy functional dependency remains with no change.

Definition 3. *A* fuzzy functional dependency *is an expression $A \xrightarrow{\vartheta} B$ where $A, B \subseteq \Omega$, $A \neq \varnothing$ and $\vartheta \in [0, 1]$. A Fuzzy Attributes Table R is said to satisfy $A \xrightarrow{\vartheta} B$ if the following condition holds:*

$$\vartheta \leq \bigwedge_{t, t' \in \mathbb{D}} \rho_A^R(t, t') \to \rho_B^R(t, t'))$$

Example 2. The Fuzzy Attributes Table given in Example 1 is a model of the fuzzy functional dependency $colour, percent \xrightarrow{0.6} localization$ for the Łukasiewicz t-norm ($a \otimes b = \max\{0, a + b - 1\}$) and its residuum ($a \to b = \min\{1 - a + b, 1\}$).

4 Fuzzy Simplification Logic

Fuzzy Simplification Logic ($\mathbb{FSL}$) is a Pavelka style fuzzy logic for reasoning about FFDs defined over fuzzy attribute tables. Some complete axiomatic system over several kind of FFDs have been defined [2,18,19]. However, like in the case of classical FDs and Armstrong's Axioms, these fuzzy inference systems are not oriented to develop automated methods to manipulate FFDs. A logic for the management of FFDs over fuzzy attribute tables, named $\mathbb{FSL}$, was introduced in [11]. In the axiomatic system of this logic, the transitivity role is played by a novel rule, named *simplification rule*, which leads to define automated reasoning methods. In this section, $\mathbb{FSL}$ is introduced. Its language is the following:

Definition 4. *Given a finite set of attribute symbols Ω, we define the language*

$$\mathbb{L} = \{A \xrightarrow{\vartheta} B \mid \vartheta \in [0,1] \text{ and } A, B \in 2^{\Omega}\}$$

Concerning the semantic, the models are given by a fuzzy attribute table $R\colon \mathbb{D} \to [0,1]^{\Omega}$ over a family of domains $\{(D_a, \rho_a) \mid a \in \Omega\}$. We say that $R \models A \xrightarrow{\vartheta} B$ if R satisfies $A \xrightarrow{\vartheta} B$, $R \models \Gamma$ means that R satisfies every FFD in the set Γ and $\Gamma \models A \xrightarrow{\vartheta} B$ denotes that $R \models \Gamma$ implies $R \models A \xrightarrow{\vartheta} B$.

Definition 5. *The axiomatic system for $\mathbb{FSL}$ has one axiom scheme and three inference rules:*

[`Ax`] Reflexive Axioms: $\vdash A \xrightarrow{1} A$

[`InR`] Inclusion Rule: $A \xrightarrow{\vartheta_1} B \vdash A \xrightarrow{\vartheta_2} B'$ *when* $\vartheta_2 \le \vartheta_1$ *and* $B' \subseteq B$.

[`CoR`] Composition Rule: $A \xrightarrow{\vartheta_1} B,\ C \xrightarrow{\vartheta_2} D \vdash AC \xrightarrow{\vartheta_1 \wedge \vartheta_2} BD$

[`SiR`] Simplification Rule: $A \xrightarrow{\vartheta_1} B,\ C \xrightarrow{\vartheta_2} D \vdash C\text{-}B \xrightarrow{\vartheta_1 \otimes \vartheta_2} D\text{-}B$ *when* $A \subseteq C$ *and* $A \cap B = \varnothing$.

The next definition presents the well known notions of *syntactic inference* ($\vdash$) and *equivalence* ($\equiv$).

Definition 6. *Let $\Gamma, \Gamma' \subseteq \mathbb{L}$ and $\varphi \in \mathbb{L}$. We say that φ is (syntactically) inferred from Γ, denoted $\Gamma \vdash \varphi$, if there exist $\varphi_1 \dots \varphi_n \in \mathbb{L}$ such that $\varphi_n = \varphi$ and, for all $1 \le i \le n$, we have that φ_i belongs to Γ, is an axiom or is obtained by applying the inference rules to formulas in $\{\varphi_j \mid 1 \le j < i\}$.*
Γ and Γ' are said to be equivalent*, denoted $\Gamma \equiv \Gamma'$, if $\Gamma \vdash \varphi'$, for all $\varphi' \in \Gamma'$, and $\Gamma' \vdash \varphi$, for all $\varphi \in \Gamma$.*

Theorem 1 ([11]). *The axiomatic system of $\mathbb{FSL}$ is sound and complete.*

5 Closures and direct bases of FFDs

In [16], we propose an automated reasoning method to decide if a formula $A \xrightarrow{\vartheta} B$ can be derived from a theory Γ (a set of fuzzy functional dependencies). That is, an automated algorithm to compute the membership function for the closure of Γ defined as follows:

$$\Gamma^+ = \{A \xrightarrow{\vartheta} B \mid \Gamma \vdash A \xrightarrow{\vartheta} B\} \tag{2}$$

Notice that, as a consequence of [`InR`], Γ^+ assigns an infinite set of pairs (B, ϑ) to every set A. If the set B is also fixed then Γ^+ gives an interval (consequence of [`InR`]) whose supremum will be denoted as $\vartheta^+_{A,B}$

$$\vartheta^+_{A,B} = \sup\{\vartheta \in [0,1] \mid A \xrightarrow{\vartheta} B \in \Gamma^+\} \tag{3}$$

On the other hand, if we fix the value of ϑ then a subset of 2^Ω is obtained. This set is finite and, by [`InR`] and [`CoR`], is an ideal of $(2^\Omega, \subseteq)$. The maximum element of this ideal will be denoted by A^+_ϑ.

$$A^+_\vartheta = \max\{B \subseteq \Omega \mid A \xrightarrow{\vartheta} B \in \Gamma^+\} \tag{4}$$

And finally, for each attribute set A, the closure of A is defined as the fuzzy set

$$A^+ \in [0,1]^\Omega \text{ with } A(x) = \vartheta^+_{A,\{x\}} \tag{5}$$

Note that the closure of a (crisp) set of attributes is a fuzzy set in $[0,1]^\Omega$ and $A^+_\vartheta = \mathrm{Cut}_\vartheta(A^+) = \{x \in \Omega \mid A^+(x) \geq \vartheta\}$. The following proposition is straightforward from definition and relates these sets.

Proposition 1. *Let Γ be a set of fuzzy functional dependencies, $A, B \subseteq \Omega$ and $\vartheta \in (0,1]$. Then*

$$\Gamma \vdash A \xrightarrow{\vartheta} B \quad \textit{if and only if} \quad \vartheta \leq \vartheta^+_{A,B} \quad \textit{if and only if} \quad B \subseteq A^+_\vartheta$$

Thus, the method for solving the implication problem (i.e. for checking if $\Gamma \vdash A \xrightarrow{\vartheta} B$) is strongly based on the computation of A^+. Algorithm 1 computes these closures.

Algorithm 1: Closure Algorithm

Data: Γ, A
Result: A^+
$\mathrm{X} := \{(x,1) \mid x \in A\}$;
`/*` X `will be the closure of` A`, which is a fuzzy set. */`
repeat
  $\mathrm{Xold} := \mathrm{X}$; $\Sigma := \varnothing$;
  foreach $\mathrm{B} \xrightarrow{\vartheta} \mathrm{C} \in \Gamma$ **do**
    if *there exists* $b \in \mathrm{B}$ *with* $b \neq a$ *for all* $(a, \kappa) \in \mathrm{X}$ **then** $\eta := 0$;
    else $\eta := \min\{\kappa \mid (b, \kappa) \in \mathrm{X} \text{ with } b \in \mathrm{B}\}$;
    if $\eta \otimes \vartheta \neq 0$ **then** $\mathrm{X} := \mathrm{X} \cup \{(c, \eta \otimes \vartheta) \mid c \in \mathrm{C}\}$;
    if $\eta \neq 1$ *and* $\mathrm{C} \not\subseteq \mathrm{Cut}_\vartheta(\mathrm{X})$ **then**
      $\Sigma := \Sigma \cup \{\mathrm{B} \smallsetminus \mathrm{Cut}_1(\mathrm{X}) \xrightarrow{\vartheta} \mathrm{C} \smallsetminus \mathrm{Cut}_\vartheta(\mathrm{X})\}$
  $\Gamma := \Sigma$;
until $\mathrm{X} = \mathrm{Xold}$;
return "A^+ is " X

Example 3. In this example abc^+ are going to be computed from the set

$$\Gamma = \{cd \xrightarrow{0.6} e, ac \xrightarrow{0.7} def, f \xrightarrow{0.5} dg, de \xrightarrow{0.9} ch, dh \xrightarrow{0.4} a\}$$

by considering the Łukasiweicz product. The initial set X is $\{(a,1),(b,1),(c,1)\}$ and the sketch of the trace of Algorithm 1 is depicted in Table 1. The output is

$$abc^+ = \{(a,1),(b,1),(c,1),(d,0.7),(e,0.7),(f,0.7),(g,0.2),(h,0.6)\}$$

$B \xrightarrow{\vartheta} C \in \Gamma$	Σ	X
$\Gamma = \{cd \xrightarrow{0.6} e, ac \xrightarrow{0.7} def, f \xrightarrow{0.5} dg, de \xrightarrow{0.9} ch, dh \xrightarrow{0.4} a\}$		
	$\varnothing$	$\{(a,1),(b,1),(c,1)\}$
$cd \xrightarrow{0.6} e$	$\{d \xrightarrow{0.6} e\}$	$\{(a,1),(b,1),(c,1)\}$
$ac \xrightarrow{0.7} def$	$\{d \xrightarrow{0.6} e\}$	$\{(a,1),(b,1),(c,1),(d,0.7),(e,0.7),(f,0.7)\}$
$f \xrightarrow{0.5} dg$	$\{d \xrightarrow{0.6} e, f \xrightarrow{0.5} g\}$	$\{(a,1),(b,1),(c,1),(d,0.7),(e,0.7),(f,0.7),(g,0.2)\}$
$de \xrightarrow{0.9} ch$	$\{d \xrightarrow{0.6} e, f \xrightarrow{0.5} g\}$	$\{(a,1),(b,1),(c,1),(d,0.7),(e,0.7),(f,0.7),(g,0.2),(h,0.6)\}$
$dh \xrightarrow{0.4} a$	$\{d \xrightarrow{0.6} e, f \xrightarrow{0.5} g\}$	$\{(a,1),(b,1),(c,1),(d,0.7),(e,0.7),(f,0.7),(g,0.2),(h,0.6)\}$
$\Gamma = \{d \xrightarrow{0.6} e, f \xrightarrow{0.5} g\}$		
	$\varnothing$	$\{(a,1),(b,1),(c,1),(d,0.7),(e,0.7),(f,0.7),(g,0.2),(h,0.6)\}$
$d \xrightarrow{0.6} e$	$\varnothing$	$\{(a,1),(b,1),(c,1),(d,0.7),(e,0.7),(f,0.7),(g,0.2),(h,0.6)\}$
$f \xrightarrow{0.5} g$	$\{f \xrightarrow{0.5} g\}$	$\{(a,1),(b,1),(c,1),(d,0.7),(e,0.7),(f,0.7),(g,0.2),(h,0.6)\}$

Table 1. Algorithm's schema

Moreover, to decide, for instance, if $\Gamma \vdash abc \xrightarrow{0.5} dh$ holds, we need to check $\{d,h\} \subseteq \mathrm{Cut}_{0.5}(abc^+) = \{a,b,c,d,e,f,h\}$. In this case the answer is affirmative.

As we have mentioned in the introduction, the aim of this work is to establish good properties to be demanded to the set of fuzzy functional dependencies in order to ensure the best behavior of the closure algorithm. Following the idea introduced in [4] for classical implications, we introduce a desirable property named directness.

Definition 7. *Let Γ be a set of fuzzy functional dependencies in Ω. We say that Γ is* direct *if, for each subset $X \subseteq \Omega$,*

$$X^+ = X \cup \bigcup_{\substack{A \xrightarrow{\vartheta} B \in \Gamma \\ X \subseteq A}} \{(b,\vartheta) \mid b \in B\}$$

And Γ is said to be a direct optimal basis if, for any direct base Γ', we have that $\Gamma \equiv \Gamma'$ implies $\|\Gamma\| \leq \|\Gamma'\|$ where $\|\Gamma\|$ denotes the size of Γ[1].

[1] That is, $\|\Gamma\| = \sum_{A \xrightarrow{\vartheta} B \in \Gamma} (|A| + |B|)$ where $|A|$ denotes the cardinality of A.

The following theorem ensures the existence and unicity of a direct optimal basis equivalent to each one.

Theorem 2. *For any set of fuzzy functional dependencies Γ in Ω there exists a unique direct optimal basis Γ_d such that $\Gamma_d \equiv \Gamma$.*

The proof of the above theorem follows the same scheme as the equivalent one provided in [4] for classical implications (crisp functional dependencies).

6 Computing direct-optimal basis

In this section, we propose the first method that calculates direct-optimal basis for fuzzy functional dependencies over fuzzy attribute tables and domains with similarity relations. We stress there is not exist in the literature, as far as we know, any work for any fuzzy extension of FD considering to compute direct optimal basis of FFDs. The method proposed here is directly based on $\mathbf{SL}_{\text{FD}}$ with the following main operations that use the operations of reduction, the rules of simplifications, and the strong simplification rule based on the logic.

In some areas, the management of formulas is limited to unitary ones. Thus, the use of Horn Clauses in Logic Programming is widely accepted. Such a language restriction allows an improvement in the performance of the methods, which are more direct and lighter.

Definition 8. *Let Γ be a set of fuzzy functional dependencies in Ω. We say that Γ is a* proper unit theory *if, for all $A \xrightarrow{\vartheta} B \in \Gamma$, the set B is a singleton not included in A and $\vartheta > 0$.*

It is not difficult to conclude that there is a proper unit theory equivalent to any set of fuzzy functional dependencies Γ:

$$\Gamma_u = \{A \xrightarrow{\vartheta} a \mid \vartheta > 0, A \xrightarrow{\vartheta} B \in \Gamma, a \in B \smallsetminus A\}$$

The algorithm for computing direct optimal basis, that we present here, has four stages: First, it transform the set of FFDs in a proper unit theory; second, it computes a direct basis by applying the following derived rule, named Strong Simplification:

$$[\texttt{sSiR}]\ A \xrightarrow{\vartheta_1} a,\ aB \xrightarrow{\vartheta_2} b \vdash AB \xrightarrow{\vartheta_1 \otimes \vartheta_2} b \qquad \text{when } a, b \notin A \cup B.$$

The third step is to narrow the set of FFDs applying the following equivalence

$$[\texttt{NarrEq}]\ \text{If } A \subseteq C \text{ and } \vartheta_1 \geq \vartheta_2, \qquad \{A \xrightarrow{\vartheta_1} b, C \xrightarrow{\vartheta_2} b\} \equiv \{A \xrightarrow{\vartheta_1} b\}.$$

Finally, the method applies, when it is possible, the Composition Equivalence: $\{A \xrightarrow{\vartheta} B, A \xrightarrow{\vartheta} C\} \equiv \{A \xrightarrow{\vartheta} BC\}$.

Theorem 3. *Let Γ be a set of fuzzy functional dependencies. Algorithm 2 renders the unique direct-optimal base equivalent to Γ.*

Algorithm 2: DirectOptimal

input : A set of fuzzy functional dependencies Γ in Ω
output: The direct-optimal basis Γ_{do} equivalent to Γ
begin
  $\Gamma_u := \{A \xrightarrow{\vartheta} a \mid \vartheta > 0, A \xrightarrow{\vartheta} B \in \Gamma, a \in B \smallsetminus A\}$
  foreach $A \xrightarrow{\vartheta_1} a \in \Gamma_u$ **do**
    foreach $Ca \xrightarrow{\vartheta_2} b \in \Gamma_u$ **do**
      if $a \neq b$ *and* $b \notin A$ **then** add $AC \xrightarrow{\vartheta_1 \otimes \vartheta_2} b$ to Γ_u;
  foreach $A \xrightarrow{\vartheta_1} b \in \Gamma_u$ **do**
    foreach $C \xrightarrow{\vartheta_2} b \in \Gamma_u$ **do**
      if $A \subseteq C$ *and* $\vartheta_1 \geq \vartheta_2$ **then** delete $C \xrightarrow{\vartheta_2} b$ from Γ_u;
  $\Gamma_{do} := \Gamma_u$
  foreach $A \xrightarrow{\vartheta_1} B \in \Gamma_{do}$ **do**
    foreach $C \xrightarrow{\vartheta_2} D \in \Gamma_{do}$ **do**
      if $A = C$ *and* $\vartheta_1 = \vartheta_2$ **then** replace $A \xrightarrow{\vartheta_1} B$ and $C \xrightarrow{\vartheta_2} D$ by $A \xrightarrow{\vartheta_1} BD$ in Γ_{do};
  return Γ_{do}

7 Conclusions and further works

In this work, we propose the first method to calculate direct-optimal basis for fuzzy functional dependencies over fuzzy attribute tables and domains with similarity relations. A discussion about the cost of the algorithm and possible improvements for it is now under consideration. As future work, we are also going to extend these results to a more expressive logic that we introduced in [3].

Acknowledgment

Supported by Grants TIN2011-28084 and TIN2014-59471-P of the Science and Innovation Ministry of Spain.

References

1. Belohlavek, R., Vychodil, V.: Relational model of data over domains with similarities: An extension for similarity queries and knowledge extraction. In: Information Reuse and Integration, 2006 IEEE International Conference on. pp. 207–213 (2006)
2. Belohlavek, R., Vychodil, V.: Tables with similarity relations: Functional dependencies, complete rules and non-redundant bases in database systems for advanced applications. LNCS 3882, 644–658 (2006)

3. Belohlávek, R., Cordero, P., Enciso, M., Mora, A., Vychodil, V.: An efficient reasoning method for dependencies over similarity and ordinal data. LNCS 7647, 408–419 (2012)
4. Bertet, K., Monjardet, B.: The multiple facets of the canonical direct unit implicational basis. Theor. Comput. Sci. 411(22-24), 2155–2166 (2010)
5. Codd, E.F.: The relational model for database management: version 2. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA (1990)
6. Cordero, P., Enciso, M., Mora, A., Guzmán, I.P.: A complete logic for fuzzy functional dependencies over domains with similarity relations. LNCS 6692, 261–269 (2009)
7. Cordero, P., Enciso, M., Mora, A., de Guzmán, I., Rodríguez-Jiménez, J.: An efficient algorithm for reasoning about fuzzy functional dependencies. Lecture Notes in Computer Science, vol. 6692, pp. 412–420 (2011)
8. Cordero, P., Enciso, M., Mora, A., de Guzmán, I.P.: SLFD logic: Elimination of data redundancy in knowledge representation. Lecture Notes in Computer Science 2527, 141–150 (2002)
9. Cordero, P., Enciso, M., Mora, A., de Guzmán, I.P.: Reasoning about fuzzy functional dependencies. In: In XIV Spanish Conference on Fuzzy Logic and Technology. pp. 121–126 (2008)
10. Cordero, P., Enciso, M., Mora, A., de Guzmán, I.P.: A complete logic for fuzzy functional dependencies over t-norms. In: In XV Spanish Conference on Fuzzy Logic and Technology. pp. 205–210 (2010)
11. Cordero, P., Enciso, M., Mora, A., de Guzmán, I.P., Rodríguez-Jiménez, J.M.: Specification and inference of fuzzy attributes. In: FOCI. pp. 107–114 (2011)
12. Guigues, J.L., Duquenne, V.: Famille minimale d'implications informatives résultant d'un tableau de données binaires. Mathématiques et Sciences Humaines 24(95), 5–18 (1986)
13. Maier, D., Sagiv, Y., Yannakakis, M.: On the complexity of testing implications of functional and join dependencies. J. ACM 28(4), 680–695 (1981)
14. Mora, A., Cordero, P., Enciso, M., Fortes, I., Aguilera, G.: Closure via functional dependence simplification. Int. J. Comput. Math. 89(4), 510–526 (2012)
15. Raju, K.V.S.V.N., Majumdar, A.K.: Fuzzy functional dependencies and lossless join decomposition of fuzzy relational database systems. ACM Trans. Database Syst. 13, 129–166 (1988)
16. Rodríguez-Jiménez, J.M., Cordero, P., Enciso, M., Mora, A.: Automated inference with fuzzy functional dependencies over graded data. LNCS 7903, 254–265 (2013)
17. Sözat, M.I., Yazici, A.: A complete axiomatization for fuzzy functional and multivalued dependencies in fuzzy database relations. Fuzzy Sets and Systems 117(2), 161–181 (2001)
18. Tyagi, B., Sharfuddin, A., Dutta, R., Tayal, D.K.: A complete axiomatization of fuzzy functional dependencies using fuzzy function. Fuzzy Sets and Systems 151(2), 363–379 (2005)
19. Yahia, S.B., Ounalli, H., Jaoua, A.: An extension of classical functional dependency: dynamic fuzzy functional dependency. Information Sciences 119(3-4), 219–234 (1999)