Attribute Dependencies in a Fuzzy Setting

Cynthia Vera Glodeanu

Technische Universität Dresden,
01062 Dresden, Germany
Cynthia_Vera.Glodeanu@mailbox.tu-dresden.de

Abstract. We present a new framework for modelling users preferences in a fuzzy setting. Starting with a formal fuzzy context, the user enters so-called attribute dependency formulas based on his priorities. The method then yields the "interesting" formal concepts, that is, interesting from the point of view of the user. Our approach is designed for compounded attributes, i.e., attributes which include more than one trait. In this paper, after studying some properties of the formulas, we start investigating the computation of non-redundant bases for them. Such bases are wishful for a better overview of the preferences.

Keywords: Formal Concept Analysis, fuzzy data, data reduction

1 Introduction

Attribute dependency formulas were introduced in [1] and further studied in a series of papers, see for instance [2]. They were developed as a method of controlling the size of crisp concept lattices. The most appealing aspect of this method is that the reduction is done based on the user's preferences, namely he is allowed to define a sort of order on the attributes. In accordance with these preferences, the user receives just the "interesting" concepts, "interesting" from his point of view. In [1] such preferences were modelled in the language of Formal Concept Analysis as follows: An **attribute dependency formula (AD formula)** over a set M of attributes is $A \sqsubseteq B$, where $A, B \subseteq M$. The meaning of the formula is "the attributes from A are less important than the attributes from B". The AD formula $A \sqsubseteq B$ is **true** in $N \subseteq M$, written $N \models A \sqsubseteq B$, if

$$\text{if } A \cap N \neq \varnothing, \text{ then } B \cap N \neq \varnothing.$$

A formal concept $(C, D) \in \mathfrak{B}(G, M, I)$ satisfies $A \sqsubseteq B$ if $D \models A \sqsubseteq B$.

These formulas were the starting point of our work. However, we develop a different kind of AD formulas, namely some that are appropriate for compounded attributes, i.e., attributes which include more than one trait. For instance the notion "wealth" is a compounded attribute consisting of "investment" and "fluency". A person who is wealthy has to have high values on both investment and fluency. We develop such formulas for the fuzzy setting and automatically obtain the crisp case by choosing $\mathbf{L} = \{0, 1\}$ for the residuated lattice.

Some proofs are omitted due to lack of space but can be found in [3].

Local Proceedings in ISBN 978–84–695–5252–0,
Universidad de Málaga (Dept. Matemática Aplicada), Spain.

The paper is structured as follows: In Section 2 we give brief introductions to Fuzzy Sets, Fuzzy Logic and Formal Fuzzy Concept Analysis. Section 3 contains our new framework. Concluding remarks and further research topics are given in the last section.

2 Preliminaries

2.1 Fuzzy Sets and Fuzzy Logics

In this subsection we present some basics about fuzzy sets and fuzzy logic. The interested reader may find more details for instance in [4, 5].

A **complete residuated lattice with (truth-stressing) hedge** is an algebra $\mathbf{L} := (L, \wedge, \vee, \otimes, \rightarrow, ^*, 0, 1)$ such that: $(L, \wedge, \vee, 0, 1)$ is a complete lattice; $(L, \otimes, 1)$ is a commutative monoid; 0 is the least and 1 the greatest element; the adjointness property, i.e., $a \otimes b \leq c \Leftrightarrow a \leq b \rightarrow c$, holds for all $a, b, c \in L$. The hedge * is a unary operation on L satisfying the following conditions: i) $a^* \leq a$; ii) $(a \rightarrow b)^* \leq a^* \rightarrow b^*$; iii) $a^{**} = a^*$; iv) $\bigwedge_{i \in I} a_i^* = (\bigwedge_{i \in I} a_i)^*$; for every $a, b, a_i \in L$ $(i \in I)$. Elements of L are called **truth degrees**, $\otimes$ and $\rightarrow$ are (truth functions of) "fuzzy conjunction" and "fuzzy implication". The hedge * is a (truth function of) logical connective "very true", see [4, 6]. The properties (i)-(iv) have natural interpretations, i.e., (i) can be read as "if a is very true, then a is true", (ii) can be read as "if $a \rightarrow b$ is very true and if a is very true, then b is very true", etc. From the mathematical point of view, the hedge operator is a special kernel operator controlling the size of the fuzzy concept lattice.

A common choice of $\mathbf{L}$ is a structure with $L = [0, 1]$, $\wedge$ and $\vee$ being minimum and maximum, $\otimes$ being a left-continuous t-norm with the corresponding $\rightarrow$. The three most important pairs of adjoint operations on the unit interval are:

$$\text{Łukasiewicz: } a \otimes b := max(0, a + b - 1) \text{ with } a \rightarrow b := min(1,\ 1 - a + b),$$

$$\text{Gödel: } a \otimes b := min(a, b) \text{ with } a \rightarrow b := \begin{cases} 1, \ a \leq b \\ b, \ a \gneqq b \end{cases},$$

$$\text{Product: } a \otimes b := ab \text{ with } a \rightarrow b := \begin{cases} 1, & a \leq b \\ b/a, & a \gneqq b \end{cases}.$$

Typical examples for the hedge are the *identity*, i.e., $a^* := a$ for all $a \in L$, and the *globalisation*, i.e., $a^* := 0$ for all $a \in L \setminus \{1\}$ and $a^* := 1$ if and only if $a = 1$.

Let $\mathbf{L}$ be the structure of truth degrees. A **fuzzy set** (**L-set**) A in a universe U is a mapping $A : U \rightarrow L$, $A(u)$ being interpreted as "the degree to which u belongs to A". We denote by $u \in A$ the fact that $A(u) = 1$. If $U = \{u_1, \ldots, u_n\}$, then A can be denoted by $A = \{^{a_1}/u_1, \ldots, ^{a_n}/u_n\}$ meaning that $A(u_i)$ equals a_i for each $i \in \{1, \ldots, n\}$. Let $\mathbf{L}^U$ denote the collection of all **L**-sets in U. The operations with **L**-sets are defined component-wise. For instance, the intersection of **L**-sets $A, B \in \mathbf{L}^U$ is an **L**-set $A \cap B$ in U s. t. $(A \cap B)(u) = A(u) \wedge B(u)$ for each $u \in U$, etc. Binary fuzzy relations (**L**-relations) between X and Y can be thought of as **L**-sets in the universe $X \times Y$. For $A, B \in \mathbf{L}^U$, the **subsethood degree**, which generalises the classical subsethood relation $\subseteq$, is defined as $\mathrm{S}(A,B) := \bigwedge_{u\in U}(A(u) \to B(u))$. Therefore, $\mathrm{S}(A,B)$ represents a degree to which A is a subset of B. In particular, we write $A \subseteq B$ iff $\mathrm{S}(A,B) = 1$.

Fuzzy closure operators were introduced in [7] and studied further by Belohlavek *at al.*, see for instance [8, 9]. The definition given in [7] mirrors more a crisp thinking, representing a special case of the one given in [8]. Therefore, we will use the latter.

Definition 1. *Define on a set Y two mappings $\mathrm{C}, \kappa : \mathbf{L}^Y \to \mathbf{L}^Y$ satisfying*

$$A \subseteq \mathrm{C}(A), \qquad \kappa(A) \subseteq A, \tag{1}$$

$$\mathrm{S}(A_1, A_2)^* \leq \mathrm{S}(\mathrm{C}(A_1), \mathrm{C}(A_2)), \qquad \mathrm{S}(A_1, A_2)^* \leq \mathrm{S}(\kappa(A_1), \kappa(A_2)), \tag{2}$$

$$\mathrm{C}(A) = \mathrm{C}(\mathrm{C}(A)), \qquad \kappa(A) = \kappa(\kappa(A)), \tag{3}$$

for every $A, A_1, A_2 \in \mathbf{L}^Y$. Then, C *is called an* $\mathbf{L}^*$***-closure operator*** *and κ an* $\mathbf{L}^*$***-kernel operator*** *on Y. A system $\mathcal{S} := \{A_j \in \mathbf{L}^Y \mid j \in J\}$ is a* $\mathbf{L}^*$***-closure system*** *if for each $A \in \mathbf{L}^U$ it holds that*

$$\bigcap_{j\in J}(\mathrm{S}(A, A_j)^* \to A_j) \in \mathcal{S}. \tag{4}$$

The system $\mathcal{S}$ is called an $\mathbf{L}^*$***-kernel system*** *if for each $A \in \mathbf{L}^U$ it holds that*

$$\bigcup_{j\in J}(\mathrm{S}(A, A_j)^* \otimes A_j) \in \mathcal{S}. \tag{5}$$

For the globalisation, (2) becomes

$$A_1 \subseteq A_2 \Longrightarrow \mathrm{C}(A_1) \subseteq \mathrm{C}(A_2), \qquad A_1 \subseteq A_2 \Longrightarrow \kappa(A_1) \subseteq \kappa(A_2),$$

and (4) and (5) become $\bigcap_{j\in J, A\subseteq A_j} A_j$ and $\bigcup_{j\in J, A\subseteq A_j} A_j$, respectively.

Theorem 1. *([8, 9]) A system $\mathcal{S}$ which is closed under arbitrary intersections is an $\mathbf{L}^*$-closure system iff for each $a \in L$ and $A \in \mathcal{S}$ it holds that $a^* \to A \in \mathcal{S}$. A system $\mathcal{S}$ closed under arbitrary unions is an $\mathbf{L}^*$-kernel system iff for each $a \in L$ and $A \in \mathcal{S}$ it holds $a^* \otimes A \in \mathcal{S}$.*

2.2 Formal Fuzzy Concept Analysis

In the following we give brief introductions to Formal Fuzzy Concept Analysis [10, 5, 11].

A triple (G, M, I) is called a **formal fuzzy context** if $I : G \times M \to L$ is an $\mathbf{L}$-relation between the sets G and M and L is the support set of some residuated lattice. Elements from G and M are called **objects** and **attributes**, respectively. The $\mathbf{L}$-relation I assigns to each $g \in G$ and each $m \in M$ the truth degree $I(g,m) \in L$ to which the object g has the attribute m. For $\mathbf{L}$-sets $A \in \mathbf{L}^G$ and $B \in \mathbf{L}^M$, the **derivation operators** are defined by

$$A^{\uparrow}(m) := \bigwedge_{g\in G}(A(g)^* \to I(g,m)), \; B^{\downarrow}(g) := \bigwedge_{m\in M}(B(m)^* \to I(g,m)) \tag{6}$$

for $g \in G$ and $m \in M$. Then, $A^{\uparrow}(m)$ is the truth degree of the statement "m is shared by all objects from A", and $B^{\downarrow}(g)$ is the truth degree of "g has all attributes from B". The operators $^{\uparrow}, ^{\downarrow}$ form a so-called Galois connection with hedges ([11]). A **formal fuzzy concept (L-concept)** is a tuple (A, B) with $A \in \mathbf{L}^G, B \in \mathbf{L}^M$ such that $A^{\uparrow} = B$ and $B^{\downarrow} = A$. Then, A is called the **(fuzzy) extent** and B the **(fuzzy) intent** of (A, B). We denote the set of all **L**-concepts of a given context (G, M, I) by $\mathfrak{B}(G^*, M^*, I)$. Concepts serve for classification. Consequently, the super- and subconcept relation plays an important role. The concept (A_1, B_1) is a **subconcept** of (A_2, B_2), written $(A_1, B_1) \leq (A_2, B_2)$, iff $A_1 \subseteq A_2$ (or, equivalently, $B_1 \supseteq B_2$). Then, we call (A_2, B_2) the **superconcept** of (A_1, B_1). The set of all **L**-concepts ordered by this concept order forms a complete fuzzy lattice (with hedge), the so-called **fuzzy concept lattice** which is denoted by $\underline{\mathfrak{B}}(G^*, M^*, I) := (\mathfrak{B}(G^*, M^*, I), \leq)$, see [11].

3 Fuzzy Attribute Dependencies

Now we are ready to present our new framework. Given a fuzzy formal context, the user obtains the "interesting" concepts after entering a sort of order on the groups of attributes, and fix the truth values for their relevance. Recall, we designed this kind of AD formulas for compounded attributes. Thus, this notion is not the fuzzy equivalent one of the formulas presented in Section 1. For a straightforward fuzzification of those formulas see Remark 1.

Definition 2. *A* ***fuzzy attribute-dependency formula*** *(fAD) over a set M of attributes is an expression $A \sqsubseteq B$, where $A, B \in \mathbf{L}^M$ are* **L***-sets of attributes. $A \sqsubseteq B$ is* ***true*** *in an* **L***-set $N \in \mathbf{L}^M$ for $\alpha, \beta \in L \setminus \{0\}$ and $\alpha \leq \beta$, written $N \models_{\alpha,\beta} A \sqsubseteq B$, if the following condition is satisfied:*

$$\textit{if } \mathrm{S}(A, N) \geq \alpha, \ \textit{then } \mathrm{S}(B, N) \geq \beta. \tag{7}$$

For an fAD formula or a set T of fAD formulas, the values α and β are called the ***thresholds*** *of $A \sqsubseteq B$ and T. An* **L***-concept $(C, D) \in \mathfrak{B}(G, M, I)$ satisfies $A \sqsubseteq B$ if $D \models_{\alpha,\beta} A \sqsubseteq B$.*

For notational simplicity we will sometimes omit α and β from $\models_{\alpha,\beta}$ provided they are clear from the setting.

The set of all formal concepts from $\mathfrak{B}(G, M, I)$ that satisfy a given set T of fAD formulas is denoted by $\mathfrak{B}_T(G, M, I)$. We call $\mathfrak{B}_T(G, M, I)$ together with the restricted concept order the **fuzzy concept lattice of** (G, M, I) **constrained by** T and denote it by $\underline{\mathfrak{B}}_T(G, M, I)$.

The fAD formulas permit a two-sided modelling of the extracted **L**-concepts. On the one hand, α and β provide the thresholds to which an intent has to contain all elements of A and B. On the other hand, the truth degrees of the elements contained in A and B fix the thresholds to which we want the attributes to be contained in the intent of a concept satisfying the fAD formula. We will illustrate this fact in the forthcoming example.

In applications it is particularly useful to associate to the truth values of a residuated lattice $\mathbf{L}$ a Likert scale $\mathcal{L}$. This allows the user to have a better understanding of the truth values. For instance let $L = \{0, 0.25, 0.5, 0.75, 1\}$ be the support set of some residuated lattice with the associated Likert scale $\mathcal{L} = \{$not important, less important, important, very important, most important$\}$, i.e, 0 =not important, 0.25 =less important, etc.

Example 1. Consider the fuzzy context given in Figure 1. It represents the eval-

	good team player		good organisational skills			adaptive towards new			confidential		computer skills	
	a: collaborative	b: not discriminative	c: time management	d: problem solver	e: analytical thinking	f: environment	g: assignment	h: priority	i: judgement	j: discretion	k: word processing	l: database
1	0	0.5	0.5	1	1	0	0	0.5	0.5	0.5	1	1
2	1	1	1	1	1	1	1	1	0.5	0.5	0	0
3	0.5	0.5	0.5	1	1	0	1	1	1	1	0.5	0.5
4	1	0.5	1	1	1	0	1	1	0.5	0.5	1	1
5	0	0.5	0	0.5	0.5	0	0	0.5	0.5	0.5	0	0
6	1	1	0.5	1	1	1	1	1	0.5	0.5	0.5	0.5
7	0	0.5	0	0	0.5	0	0	0.5	0	0.5	0	0

Fig. 1. Fuzzy context about employees

uation of the employees of a small business regarding some qualities. Here each quality is compounded of two or more traits. For instance, an employee is a "good team player" if he/she is collaborative and not discriminative. The context has 44 $\mathbf{L}$-concepts with the Gödel logic which are far too many to be analysed by a busy manager. The manager however knows how good or bad the employees do their jobs and he is interested more in their collaboration than their organisational skills and more in their adaptivity than in their confidentiality. Therefore, he chooses the following two fAD formulas

$$\{{}^{0.5}/c, d, e\} \sqsubseteq \{a, b\} \text{ and } \{{}^{0.5}/i, {}^{0.5}/j\} \sqsubseteq \{f, g, h\}, \tag{8}$$

with $\alpha = 0.5$ and $\beta = 1$. Then, he obtains 11 $\mathbf{L}$-concepts which are given in Figure 2.

The manager realises that the company does neither send its employees to business trips nor to other companies and the employees should know their priorities. Therefore, he changes the second fAD formula into

$$\{{}^{0.5}/i, {}^{0.5}/j\} \sqsubseteq \{g, {}^{0.5}/h\}. \tag{9}$$

	Extent							Intent											
	1	2	3	4	5	6	7	a	b	c	d	e	f	g	h	i	j	k	l
1	0	0	0	0	0	0.5	0	1	1	1	1	1	1	1	1	1	1	1	1
2	0	0.5	0	0	0	0.5	0	1	1	1	1	1	1	1	1	1	1	0	0
3	0	1	0	0	0	0.5	0	1	1	1	1	1	1	1	1	0.5	0.5	0	0
4	0	0	0	0	0	1	0	1	1	0.5	1	1	1	1	1	0.5	0.5	0.5	0.5
5	0	1	0	0	0	1	0	1	1	0.5	1	1	1	1	1	0.5	0.5	0	0
6	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0	1	0	0	1	0	0	1	0	1	0	0
7	0.5	1	0.5	0.5	0.5	1	0.5	0	1	0	0	1	0	0	1	0	0.5	0	0
8	0.5	0.5	1	0.5	0.5	0.5	0.5	0	0.5	0	0	1	0	0	1	0	1	0	0
9	0.5	1	1	1	0.5	1	0.5	0	0.5	0	0	1	0	0	1	0	0.5	0	0
10	1	1	1	1	0.5	1	0.5	0	0.5	0	0	1	0	0	0.5	0	0.5	0	0
11	1	1	1	1	1	1	1	0	0.5	0	0	0.5	0	0	0.5	0	0.5	0	0

Fig. 2. Concepts satisfying (8)

Obviously the concepts from the first fAD formulas are a subset of the concepts from the second fAD formulas. The second couple of formulas yields 16 concepts, namely those from Figure 2 and 3.

	Extent							Intent											
	1	2	3	4	5	6	7	a	b	c	d	e	f	g	h	i	j	k	l
12	0	0	0.5	0.5	0	0.5	0	1	1	1	1	1	0	1	1	1	1	1	1
13	0	0.5	0.5	0.5	0	0.5	0	1	1	1	1	1	0	1	1	1	1	0	0
14	0	1	0.5	0.5	0	0.5	0	1	1	1	1	1	0	1	1	0.5	0.5	0	0
15	0	0	0.5	0.5	0	1	0	1	1	0.5	1	1	0	1	1	0.5	0.5	0.5	0.5
16	0	1	0.5	0.5	0	1	0	1	1	0.5	1	1	0	1	1	0.5	0.5	0	0

Fig. 3. Concepts satisfying (9) and the first fAD from (8)

We already have the framework of how to select interesting concepts based on the preferences of the user. Let us further investigate some properties of the fAD formulas.

Proposition 1. *Let T be a set of fAD formulas. Then, $\underline{\mathfrak{B}}_T(G,M,I)$ is a complete fuzzy lattice, which is a $\bigvee$-sublattice of $\underline{\mathfrak{B}}(G,M,I)$.*

Proof. Clearly, $\mathfrak{B}_T(G,M,I) \subseteq \mathfrak{B}(G,M,I)$ and $\underline{\mathfrak{B}}(G,M,I)$ with the restricted concept order, is a partially ordered subset of $\underline{\mathfrak{B}}(G,M,I)$. Further note that $\underline{\mathfrak{B}}_T(G,M,I)$ is bounded from below because the least **L**-concept of $\underline{\mathfrak{B}}(G,M,I)$ is $(M^{\downarrow}, M)$, concept which is compatible with every fAD formula. Now, we have to show that $\underline{\mathfrak{B}}_T(G,M,I)$ is closed under arbitrary suprema in $\underline{\mathfrak{B}}(G,M,I)$. To this end let $(A_j, B_j) \in \underline{\mathfrak{B}}_T(G,M,I)$ $(j \in J)$ be **L**-concepts. For any fAD formula $A \sqsubseteq B \in T$, we have $B_j \models A \sqsubseteq B$ for every $j \in J$. Now, if there is $j \in J$ such that $B_j(a) < \alpha$ for some $a \in A$, then $\cap_{j\in J} B_j(a) < \alpha$ and we are done because then $\cap_{j\in J}B_j \models A \sqsubseteq B$. Contrary, if for all $j \in J$ and all $a \in A$ we have $B_j(a) \geq \alpha$, then $\cap_{j\in J}B_j(a) \geq \alpha$ for all $a \in A$. Since $B_j \models A \sqsubseteq B$ holds for all $j \in J$, we also have that $B_j(b) \geq \beta$ for all $j \in J$ and $b \in B$. Due to the same argument as before, it holds $\cap_{j\in J}B_j(b) \geq \beta$ for all $b \in B$ and hence we have $\cap_{j\in J}B_j \models A \sqsubseteq B$, showing that $\underline{\mathfrak{B}}_T(G, M, I)$ is closed under arbitrary suprema. □

In general $\underline{\mathfrak{B}}_T(G, M, I)$ is not closed under arbitrary infima in $\underline{\mathfrak{B}}(G, M, I)$.

The fAD formulas are entered by the user. Thus, the chance that these formulas are redundant is very high. Wishful thinking suggests to have a set of non-redundant formulas because these are then easier to follow and to modify. Therefore, in the following we will develop methods for removing such redundancies.

Definition 3. *An* $\mathbf{L}$*-set* $N \in \mathbf{L}^M$ *is a* ***model*** *of a set* T *of fAD formulas if we have* $N \models A \sqsubseteq B$ *for each* $A \sqsubseteq B \in T$*. Let* $\mathrm{Mod}(T)$ *denote the set of all models of* T*, i.e.,*

$$\mathrm{Mod}(T) := \{N \in \mathbf{L}^Y \mid N \models A \sqsubseteq B \text{ for each } A \sqsubseteq B \in T\}.$$

An fAD formula $A \sqsubseteq B$ ***follows semantically*** *from* T*, written* $T \models A \sqsubseteq B$*, if for each* $N \in \mathrm{Mod}(T)$*, we have* $N \models A \sqsubseteq B$*.*

Lemma 1. *i)* $N \models A \sqsubseteq \{{}^{l_1}/y_1, \ldots, {}^{l_n}/y_n\}$ *iff* $N \models A \sqsubseteq \{{}^{l_i}/y_i\}$ *holds for each* $i \in \{1, \ldots, n\}$*.*
ii) For each set T *of fAD formulas and each fAD formula* φ*, we have* $T \models \varphi$ *iff* $\lfloor T \rfloor \models \varphi$*, where* $\lfloor T \rfloor := \{A \sqsubseteq \{{}^{l}/y\} \mid A \sqsubseteq B \in T \text{ and } B(y) = l\}$*.*

Proof. i) If $\mathrm{S}(A, N) < \alpha$, then we are done. Therefore, suppose $\mathrm{S}(A, N) \geq \alpha$. By the definition of S, we have $\mathrm{S}(\{{}^{l_1}/y_1, \ldots, {}^{l_n}/y_n\}, N) \geq \beta$ if and only if we have $N(\{{}^{l_i}/y_i\}) \geq \beta$ for all $i \in \{1, \ldots, n\}$ and thus $N \models A \sqsubseteq \{{}^{l_i}/y_i\}$ for all $i \in \{1, \ldots, n\}$.
ii) Follows by i), omitted due to lack of space. □

Thanks to Lemma 1 we may merge fAD formulas with the same left-hand side into a single fAD formula. The new formula is true in a model if and only if all its component fAD formulas are true in that model. This lemma allows us also to test semantic entailment in fAD formulas $A \sqsubseteq \{{}^{l}/y\}$ rather than on the whole $A \sqsubseteq B$.

In the following we will study the connection between the models of fAD formulas and $\mathbf{L}^*$-closure systems. It will turn out that any $\mathbf{L}^*$-closure system can be described by a set of fAD formulas.

Proposition 2. *Let* T *be a set of fAD formulas. Then,* $\mathrm{Mod}(T)$ *is an* $\mathbf{L}^*$*-closure system with* * *being the globalisation.*

Proof. Let $\{N_j \in \mathrm{Mod}(T) \mid j \in J\}$. We will be showing that $\mathrm{Mod}(T)$ is closed under arbitrary intersection, i.e., $\bigcap_{j\in J} N_j$ is a model of T. For any fAD formula

$A \sqsubseteq B \in T$ we have $N_j \models A \sqsubseteq B$ for every $j \in J$. Now, if there is $j \in J$ such that $N_j(a) < \alpha$ for some $a \in A$, then $\cap_{j \in J} N_j(a) < \alpha$ and we are done because then $\cap_{j \in J} N_j \models A \sqsubseteq B$. Contrary, if for all $j \in J$ and all $a \in A$ we have $N_j(a) \geq \alpha$, then $\cap_{j \in J} N_j(a) \geq \alpha$ for all $a \in A$. Since $N_j \models A \sqsubseteq B$ for all $j \in J$ holds, we also have that $N_j(b) \geq \beta$ for all $j \in J$ and $b \in B$. Due to the same argument as before, $\cap_{j \in J} N_j(b) \geq \beta$ for all $b \in B$ and hence we have $\cap_{j \in J} N_j \models A \sqsubseteq B$, showing that $\text{Mod}(T)$ is a closed under arbitrary intersection.

$\text{Mod}(T)$ is a closed under arbitrary intersection and due to Theorem 1 we just have to show that for any $N \in \text{Mod}(T)$ and any $a \in L$, $a^* \to N$ is a model of T. However, this condition only holds if * is the globalisation because, then we have

$$\text{S}(A, a^* \to N) = a^* \to \text{S}(A, N) = \begin{cases} 1, & a = 0, \\ \text{S}(A, N), & a = 1, \end{cases}$$

i.e., $a^* \to N$ trivially satisfies any fAD formula if $a = 0$ or we do not gain anything new to N in the case that $a = 1$. □

Remark 1. One may argue that due to Proposition 2 the fAD formulas are not strong enough. However, we consider that this is not the case. In the crisp setting, the AD formulas form a kernel system, and due to the connection between AD formulas and attribute implications (for details see [1]) one may efficiently compute a non-redundant base of formulas. For instance, if we generalise in a straight-forward way the AD formulas from [1] to the fuzzy setting, then we also obtain just crisp like kernel systems. This can be done as follows: Define a fAD formula $A \sqsubseteq B$, where $A, B \in \mathbf{L}^M$. Further, $A \sqsubseteq B$ is *true* in an $\mathbf{L}$-set $N \in \mathbf{L}^M$ for $\alpha, \beta \in L \setminus \{0\}$ and $\alpha \leq \beta$, written $N \models_{\alpha,\beta} A \sqsubseteq B$, if the following condition is satisfied:

$$\text{if } A \cap N \text{ } \alpha\text{-true, then } B \cap N \text{ } \beta\text{-true,}$$

where an $\mathbf{L}$-set $X \cap Y \in \mathbf{L}^M$ is α-true if there is at least one attribute $m \in M$ such that $(X \cap Y)(m) \geq \alpha$, where $\alpha \in L$. We need the thresholds in order to ensure that the obtained concepts are indeed relevant. Now, the models of such formulas form a crisp like kernel system which can be shown in an analogous way to Proposition 2.

Lemma 2. *For any $\mathbf{L}^*$-closure system $\mathcal{S}$ in M there is a set T of fAD formulas over M such that $\mathcal{S} = \text{Mod}(T)$.*

Proof. Define a set T of fAD formulas by $T := \{A \sqsubseteq C_{\mathcal{S}}(A) \mid A \in \mathbf{L}^M\}$, where $C_{\mathcal{S}}(A)$ is the closure of A given by the L^*-closure operator $C_{\mathcal{S}}$. Further, choose $\alpha = \beta = 1$. Let $N \in \mathcal{S}$, i.e., $N = C_{\mathcal{S}}(N)$. We have to show that N is a model of T, thus let $N \models A \sqsubseteq C_{\mathcal{S}}(A)$ for every $A \sqsubseteq C_{\mathcal{S}}(A) \in T$. If $\text{S}(A, N) < 1$, then $N \models A \sqsubseteq C_{\mathcal{S}}(A)$ and we are done. Now take $\text{S}(A, N) \geq 1$, i.e., $A \subseteq N$. Since $C_{\mathcal{S}}$ is a closure operator we have $C_{\mathcal{S}}(A) \subseteq C_{\mathcal{S}}(N) = N$, hence $\text{S}(C_{\mathcal{S}}(A), N) \geq 1$, i.e., $N \models A \sqsubseteq C_{\mathcal{S}}(A)$. Thus, N is a model of T and we have the first inclusion, namely $\mathcal{S} \subseteq \text{Mod}(T)$. For the converse inclusion let $N \in \text{Mod}(T)$. We have

$$\text{if } \text{S}(N, N) \geq 1, \text{ then } \text{S}(C_{\mathcal{S}}(N), N) \geq 1,$$

where $\mathrm{S}(N, N) \geq 1$ obviously holds and since N is a model of T we must also have $\mathrm{S}(C_{\mathcal{S}}(N), N) \geq 1$ yielding that $N = C_{\mathcal{S}}(N)$, i.e., $N \in \mathcal{S}$ and hence $\mathrm{Mod}(T) \subseteq \mathcal{S}$. □

According to Proposition 2, $\mathrm{Mod}(T)$ is an $\mathbf{L}^*$-closure system, and therefore there must exist an $\mathbf{L}^*$-closure operator $C_{\mathrm{Mod}(T)} : \mathbf{L}^M \to \mathbf{L}^M$ such that $N = C_{\mathrm{Mod}(T)}(N)$ if and only if $N \in \mathrm{Mod}(T)$. Hence, by definition $C_{\mathrm{Mod}(T)}(N)$ is the least model in $\mathrm{Mod}(T)$ which contains N. This definition of the $\mathbf{L}^*$-closure operator does not provide a useful method to compute the closure of a given N. First, because one has to iterate over all models in $\mathrm{Mod}(T)$ and second, such a iteration may be impossible if $\mathbf{L}$ is infinite because then $\mathrm{Mod}(T)$ is infinite.

Similarly to (fuzzy) attribute implications we proceed as follows: For any set T of fAD formulas and for any $\mathbf{L}$-set $N \in \mathbf{L}^M$ of attributes, we define an $\mathbf{L}$-set $N^T \in \mathbf{L}^M$ of attributes as follows:

$$N^T := N \cup \bigcup \{\beta \otimes B \mid A \sqsubseteq B \in T, \mathrm{S}(A, N) \geq \alpha\}. \tag{10}$$

Further, we define an $\mathbf{L}$-set $N^{T_n} \in \mathbf{L}^Y$ of attributes for each non-negative integer by

$$N^{T_n} := \begin{cases} N & \text{if } n = 0, \\ (N^{T_{n-1}})^T & \text{if } n \geq 1. \end{cases} \tag{11}$$

We define an operator $\mathrm{cl_T} : \mathbf{L}^M \to \mathbf{L}^M$ by

$$\mathrm{cl_T}(N) := \bigcup_{n=0}^{\infty} N^{T_n}. \tag{12}$$

Proposition 3. *For each $N \in \mathrm{Mod}(T)$ we have $\mathrm{cl_T}(N) = N$.*

Proof. Omitted due to lack of space. □

The next lemma shows that the $\mathbf{L}^*$-closure operator defined on the models of T coincides with the $\mathrm{cl_T}$-operator defined in (12).

Lemma 3. *Let T be a set of fAD formulas over M. Further let both M and $\mathbf{L}$ be finite. Then, $\mathrm{cl_T}$ is an $\mathbf{L}^*$-closure operator such that for each $N \in \mathbf{L}^M$, $C_{\mathrm{Mod}(T)}(N) = \mathrm{cl_T}(N)$.*

Proof. $C_{\mathrm{Mod}(T)}$ is an $\mathbf{L}^*$-closure operator, therefore it suffices to check that $C_{\mathrm{Mod}(T)}$ and $\mathrm{cl_T}$ coincide. To this end let $N \in \mathbf{L}^M$ be an $\mathbf{L}$-set of attributes. By the definition of $\mathrm{cl_T}$ we have $N \subseteq \mathrm{cl_T}(N)$. We still have to show that $\mathrm{cl_T}(N)$ belongs to $\mathrm{Mod}(T)$ and that $\mathrm{cl_T}(N)$ is the least model containing N. First of all note that the finiteness of $\mathbf{L}$ and M imply that $\mathbf{L}^M$ is finite and that there exists a non-negative integer k such that $\mathrm{cl_T}(N) = N^{T_k}$, where N^{T_k} is given by (11).

We still have to show that $\mathrm{cl_T}(N) \in \mathrm{Mod}(T)$, i.e., for any $A \sqsubseteq B \in T$ we have $\mathrm{cl_T}(N) \models A \sqsubseteq B$. Suppose that $\mathrm{S}(A, \mathrm{cl_T}(N)) \geq \alpha$. Then, $\mathrm{cl_T}(N) = N \cup \{\beta \otimes B\}$. Obviously, $\mathrm{S}(B, N \cup \{\beta \otimes B\}) \geq \beta$, proving that $\mathrm{cl_T}(N) \in \mathrm{Mod}(T)$ which contains N. For any $X \in \mathrm{Mod}(T)$ such that $N \subseteq X$ we have to show that $\mathrm{cl_T}(N) \subseteq X$. This easily follows by the properties of closure operators and by Proposition 3. In fact, we have that $\mathrm{cl_T}(N) \subseteq \mathrm{cl_T}(X) = X$. □

Based on the previous result we present an algorithm for the computation of the closure $C_{\text{Mod}(T)}(N)$ of a fuzzy attribute set $N \in \mathbf{L}^M$ provided $\mathbf{L}$ and M are finite.

Algorithm 1: Closure(N, T)

```
1 repeat
2     take A ⊑ B ∈ T such that S(A, N) ≥ α and S(B, N) < β;
3     set N to N ∪ {β ⊗ B};
4 until forall A ⊑ B ∈ T, (S(A, N) < α) or (S(A, N) ≥ α and S(B, N) ≥ β);
5 return N
```

If we choose more restrictive values for α and β, we have the following connection between fuzzy implications and fAD formulas:

Lemma 4. *Let* $\text{Imp}(T) := \{A \Rightarrow B \mid$ *for all* $A \sqsubseteq B \in T\}$ *and* T *be a set of fAD formulas. If we choose* $\alpha = \beta = 1$*, then the following holds*

$$\text{Mod}(\text{Imp}(T)) \subseteq \text{Mod}(T).$$

Proof. Omitted due to lack of space. □

Definition 4. *Two sets* T_1 *and* T_2 *of fAD formulas are called* ***equivalent****, written* $T_1 \equiv T_2$*, if for each* $\varphi_1 \in T_1$ *and* $\varphi_2 \in T_2$ *we have* $T_1 \models \varphi_2$ *and* $T_2 \models \varphi_1$*.*

Lemma 5. *Let* T_1 *and* T_2 *be sets of fAD formulas. Then, the following are equivalent:*
i) $\text{Mod}(T_1) = \text{Mod}(T_2)$*,*
ii) For any fAD formula φ *we have* $T_1 \models \varphi \Longleftrightarrow T_2 \models \varphi$*,*
iii) $T_1 \equiv T_2$*.*

Proof. Omitted due to lack of space. □

Now we are prepared to introduce non-redundant bases.

Definition 5. *A set* T_1 *of fAD formulas is called a* ***non-redundant base*** *of* T *if* $T \equiv T_1$ *and there is no* $T_2 \subset T_1$ *with* $T_2 \equiv T$*. A set* T_1 *of fAD formulas is called a* ***minimal base*** *of* T *if* $T \equiv T_1$ *and for each* T_2 *such that* $T \equiv T_2$*, we have* $|T_1| \leq |T_2|$*.*

Obviously, if T_1 is a minimal base of T, then T_1 is a non-redundant base of T. The converse implication is not true in general.

For a given set T of fAD formulas we may compute a non-redundant base as follows: First note that if $T_1 := T \setminus \{A \sqsubseteq B\}$ and $T_1 \models A \sqsubseteq B$, then $T \equiv T_1$. We may then remove fAD formulas $A \sqsubseteq B$ from T step-by-step until there is no $T_1 \subset T$ such that $T_1 \equiv T$. The computation of a non-redundant base with this method is quite laborious. In what follows, we present another connection between fuzzy attribute implications and fAD formulas which will considerably simplify this task.

Lemma 6. *Let T be a set of fAD formulas. We have* $\mathrm{Mod}(T) = \mathrm{Mod}(\mathrm{Imp}(T^*))$, *where*

$$\mathrm{Imp}(T^*) := \{\alpha \otimes A \Rightarrow \beta \otimes B \mid \forall A \sqsubseteq B \in T,\ \alpha, \beta \text{ thresholds of } T\}. \tag{13}$$

Proof. First note that an $\mathbf{L}$-set $N \in \mathbf{L}^M$ is a **model** of an attribute implication $A \Rightarrow B$ in a fuzzy setting if $||A \Rightarrow B||_N := \mathrm{S}(A, N)^* \to \mathrm{S}(B, N) = 1$.

Let $N \in \mathrm{Mod}(T)$ and $A \sqsubseteq B \in T$. We have two cases: i) $\mathrm{S}(A, N) \geq \alpha$ and $\mathrm{S}(B, N) \geq \beta$ both hold. Consider its first part. Then, for every attribute $m \in M$, we have $A(m) \to N(m) \geq \alpha$ which by the adjointness property gives us $\alpha \otimes A(m) \leq N(m)$ and therefore $\mathrm{S}(\alpha \otimes A, N) = 1$ and thus $\mathrm{S}(\beta \otimes B, N) = 1$. Hence, $\alpha \otimes A \Rightarrow \beta \otimes B||_N = \mathrm{S}(\alpha \otimes A, N)^* \to \mathrm{S}(\beta \otimes B, N) = 1^* \to 1 = 1$. ii) We have $\mathrm{S}(A, N) < \alpha$ which is equivalent to $\mathrm{S}(\alpha \otimes A, N) < 1$. Therefore, $||\alpha \otimes B \Rightarrow \beta \otimes B||_N = \mathrm{S}(\alpha \otimes A, N)^* \to \mathrm{S}(\beta \otimes B, N) = 0 \to \mathrm{S}(\beta \otimes B, N) = 1$. Cases i) and ii) show that N is a model of $\mathrm{Imp}(T^*)$.

For the converse inclusion let $N \in \mathrm{Mod}(\mathrm{Imp}(T^*))$. Then, we have

$$||\alpha \otimes A \Rightarrow \beta \otimes B||_N = \mathrm{S}(\alpha \otimes A, N)^* \to \mathrm{S}(\beta \otimes B, N) = 1 \tag{14}$$

for any fuzzy implication $A \Rightarrow B \in \mathrm{Imp}(T^*)$. Equation 14 holds if and only if one of the following situations appears: i) ($\mathrm{S}(\alpha \otimes A, N)^* = 1$ and $\mathrm{S}(\beta \otimes B, N) = 1$) $\iff$ ($\mathrm{S}(A, N) \geq 1$ and $\mathrm{S}(B, N) \geq 1$). ii) $\mathrm{S}(\alpha \otimes A, N)^* = 0$ is equivalent to $\mathrm{S}(\alpha \otimes A, N) < 1$ which further is equivalent to $\mathrm{S}(A, N) < \alpha$. The two cases prove $N \models_{\alpha,\beta} A \sqsubseteq B$. ☐

Thus, a fAD formula $A \sqsubseteq B$ with thresholds α, β is satisfied if and only if the corresponding implication from $\mathrm{Imp}(T^*)$, where * is the globalisation, holds with truth value 1. With this link between fAD formulas and fuzzy attribute implications we may easily compute a minimal base for any set T of fAD formulas. First, we build the set $\mathrm{Imp}(T^*)$ associated to T as given in (13). For this set we compute a minimal base of attribute implications $\mathcal{B}_{T^*}$. Finally, from $\mathcal{B}_{T^*}$ we obtain a minimal base of fAD formulas for T by

$$\mathcal{B}_T := \{A^{\blacklozenge} \sqsubseteq B^{\blacklozenge} \setminus A^{\blacklozenge} \mid A \Rightarrow B \in \mathcal{B}_{T^*}\},$$

where

$$A^{\blacklozenge} := \bigvee\{C \in \mathbf{L}^M \mid \alpha \otimes C = \alpha \otimes A\},\ B^{\blacklozenge} := \bigvee\{D \in \mathbf{L}^M \mid \alpha \otimes D = \alpha \otimes B\}.$$

Note that, unlike the crisp case, in the fuzzy setting a formal context does not have to have a unique stem base (see [12]). The uniqueness is ensured just in the case of the globalisation.

With the possibility of computing a non-redundant base, the user may review his choices and alter them conveniently.

4 Conclusion

We have presented two generalisations of crisp attribute dependency formulas into the fuzzy setting. Both variants allow the user to define a sort of order on the attributes. According to the entered constraints the user sees just a part of the concept lattice, namely the one containing the relevant concepts for him/her. Due to lack of space, the straightforward generalisation was just briefly presented in Remark 1.

The second approach was designed for compound attributes, such which incorporate more than one quality or specification. This time we required from the "interesting" concepts that if they contain all less important attributes with a threshold α, then they should also contain all more important attributes with a threshold β, were the thresholds are truth degrees such that $\alpha \leq \beta$. For such formulas, besides showing some of their properties, we focused mainly on the computation of their non-redundant bases.

Future work will focus on applying the method on real-world data and evaluating the outcomes by experts. Another research topic is the exploration of fAD formulas, where the user may alter the choices made without starting from scratch each time.

References

1. Belohlávek, R., Sklenar, V.: Formal concept analysis constrained by attribute-dependency formulas. In Ganter, B., Godin, R., eds.: ICFCA. Volume 3403 of Lecture Notes in Computer Science., Springer (2005) 176–191
2. Belohlávek, R., Vychodil, V.: Formal concept analysis with background knowledge: Attribute priorities. IEEE Transactions on Systems, Man, and Cybernetics, Part C **39**(4) (2009) 399–409
3. Glodeanu, C.: Attribute dependencies in a fuzzy setting. Technical Report MATH-AL-05-2012, TU Dresden (2012)
4. Hájek, P.: The Metamathematics of Fuzzy Logic. Kluwer (1998)
5. Belohlávek, R.: Fuzzy Relational Systems: Foundations and Principles. Volume 20 of IFSR Int. Series on Systems Science and Engineering. Kluwer Academic/Plenum Press (2002)
6. Hájek, P.: On very true. Fuzzy Sets and Systems **124**(3) (2001) 329–333
7. Biacino, L., Gerla, G.: An extension principle for closure operators. Journal of Mathematical Analysis and Appl. **198** (1996) 1–24
8. Belohlávek, R.: Fuzzy closure operators. Journal of Mathematical Analysis and Appl. **262** (2001) 473–489
9. Belohlávek, R., Funioková, T., Vychodil, V.: Fuzzy closure operators with truth stressers. Logic Journal of the IGPL **13**(5) (2005) 503–513
10. Pollandt, S.: Fuzzy Begriffe. Springer Verlag, Berlin Heidelberg New York (1997)
11. Belohlávek, R., Vychodil, V.: Fuzzy concept lattices constrained by hedges. JACIII **11**(6) (2007) 536–545
12. Belohlávek, R., Vychodil, V.: Attribute implications in a fuzzy setting. In: ICFCA. (2006) 45–60