On the importance of fuzzy attribute implications

Vilem Vychodil, *Member, IEEE*

Abstract— **Our paper deals with the expressive power of fuzzy attribute implications which are if-then rules describing dependencies among graded attributes. In our previous work, we have shown that fuzzy attribute implications are important in data mining because they can be used as a concise description of all if-then dependencies which are hidden in object-attribute fuzzy relational data. Fuzzy attribute implications can be seen as formulas (in the narrow sense) of the form $A \Rightarrow B$ where both A and B are conjunctions of subformulas containing constants for truth degrees acting as (constants for) threshold truth degrees. This paper investigates possibility to replace sets of fuzzy attribute implications with fuzzy sets of ordinary if-then formulas in the sense of Pavelka's abstract logic. We reveal the impossibility to replace fuzzy attribute implications by the ordinary formulas without losing their expressive power. From the technical point of view, we present counterexamples demonstrating that fuzzy sets of the ordinary attribute implications cannot be used to describe sets of fixed points of arbitrary fuzzy closure operators.**

I. INTRODUCTION

If-then rules describing dependencies among attributes are the subject of study of various disciplines. The rules appear as association rules in data mining, attribute implications in formal concept analysis [9], and functional dependencies in relational database systems [14]. In their basic setting, the rules describe dependencies among Boolean yes/no attributes. That is, validity of such rules is considered with respect to the fact that each of the attributes either applies or does not apply to a specified object or a set of objects. It is often the case, however, that attributes are graded rather than just bivalent yes/no attributes. For instance, "average household income", "low fuel consumption", and "very high price" are all examples of attributes for which it makes more sense to consider degrees to which the attributes apply to objects rather than just applies/does not apply. Therefore, there is a need to develop a theory of if-then rules that are graded counterparts of the classical association rules, attribute implications, and functional dependencies.

In the past few decades, there have been several approaches to if-then rules in graded setting with varying technical clarity and scientific quality. Most of the results did not go far beyond definitions. This is partly because a lot of the approaches have been devised long before the fuzzy logic in narrow sense [10], [11] and the theory of fuzzy relational systems [1] were developed. In a series of papers, we have introduced fuzzy attribute implications which seem to be "the right" counterparts of the classical attribute implications, see [4] and [5] for a survey of results. The nontrivial results on nonredundant bases and axiomatization [7] of the entailment of fuzzy attribute implications support the claim that fuzzy attribute implications are among the various if-then rules used in data mining of fuzzy relational data the most appropriate ones.

One of the benefits of using fuzzy attribute implications is their great expressive power. For example, we have shown [6] that fuzzy attribute implications can be used to describe fixpoints of arbitrary fuzzy closure operators. In addition to that, there is a tractable procedure to compute a nonredundant base (certain minimal description) of such fixpoints using fuzzy attribute implications. In this paper, we show that these results cannot be obtained with weaker formulas. Namely, we focus on ordinary if-then rules whose semantics and entailment will be defined as it is usual in Pavelka's abstract fuzzy logic [10], [15]. Intuitively, such rules may seem (nearly) as expressive as fuzzy attribute implications but we show, by means of counterexamples, that they are not. We further argue that the weaker formulas are not suitable for basic data mining tasks in case of object-attribute data with graded attributes.

This paper is organized as follows, Section II presents preliminaries. In Section III, we recall notions of fuzzy attribute implications, their models, entailment, and related notions. In Section IV, we formulate the problem in a more detail and, finally, in Section V, we present the counterexamples.

II. PRELIMINARIES

In this section we present an overview of notions of fuzzy logic and fuzzy set theory we will be using in the sequel. Details can be found e.g. in [1], [10], [11], a good introduction to fuzzy logic and fuzzy sets is presented in [13]. We will use complete residuated lattices with hedges as basic structures of truth degrees. A complete residuated lattice with a truth-stressing hedge (shortly, a hedge) [11], [12] is an algebra $\mathbf{L} = \langle L, \wedge, \vee, \otimes, \rightarrow, {}^*, 0, 1\rangle$ such that $\langle L, \wedge, \vee, 0, 1\rangle$ is a complete lattice with 0 and 1 being the least and greatest element of L, respectively; $\langle L, \otimes, 1\rangle$ is a commutative monoid (i.e. $\otimes$ is commutative, associative, and $a \otimes 1 = 1 \otimes a = a$ for each $a \in L$); $\otimes$ and $\rightarrow$ satisfy so-called adjointness property: $a \otimes b \leq c$ iff $a \leq b \rightarrow c$ for each $a, b, c \in L$; hedge * satisfies (i) $1^* = 1$, (ii) $a^* \leq a$, (iii) $(a \rightarrow b)^* \leq a^* \rightarrow b^*$, and (iv) $a^{**} = a^*$, for each $a, b \in L$. Elements a of L are called truth degrees. $\otimes$ and $\rightarrow$ are (truth functions of) "fuzzy conjunction" and "fuzzy implication". Hedge * is a (truth function of) logical

Vilem Vychodil is with the Dept. Systems Science and Industrial Engineering, T. J. Watson School of Engineering and Applied Science, Binghamton University–SUNY, PO Box 6000, Binghamton, NY 13902–6000, USA (phone: +1 (607) 777 5690; fax: +1 (607) 777 4094; email: vychodil@binghamton.edu). He is also with the Dept. Computer Science, Palacky University, Olomouc; Tomkova 40, CZ-779 00 Olomouc, Czech Republic. Support by grant no. 1ET101370417 Grant Agency of the Academy of Sciences of Czech Republic, grant No. 201/05/0079 of the Czech Science Foundation, and by institutional support, research plan MSM 6198959214, is gratefully acknowledged.

connective “very true”, see [11], [12]. Properties of hedges have natural interpretations, e.g. (ii) can be read: “if a is very true, then a is true”, (iii) can be read: “if $a \to b$ is very true and if a is very true, then b is very true”, etc.

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

$$\text{Łukasiewicz:} \quad \begin{aligned} a \otimes b &= \max(a+b-1,0), \\ a \to b &= \min(1-a+b,1), \end{aligned} \tag{1}$$

$$\text{Gödel:} \quad \begin{aligned} a \otimes b &= \min(a,b), \\ a \to b &= \begin{cases} 1, & \text{if } a \le b, \\ b, & \text{otherwise,} \end{cases} \end{aligned} \tag{2}$$

$$\text{Goguen (product):} \quad \begin{aligned} a \otimes b &= a \cdot b, \\ a \to b &= \begin{cases} 1, & \text{if } a \le b, \\ \frac{b}{a}, & \text{otherwise.} \end{cases} \end{aligned} \tag{3}$$

Two boundary cases of (truth-stressing) hedges are (i) identity, i.e. $a^* = a$ ($a \in L$); (ii) globalization [16]:

$$a^* = \begin{cases} 1, & \text{if } a = 1, \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

A special case of a complete residuated lattice with hedge is the two-element Boolean algebra $\langle\{0,1\},\wedge,\vee,\otimes,\to,{}^*,0,1\rangle$, denoted by $\mathbf{2}$, which is the structure of truth degrees of the classical logic. That is, the operations $\wedge,\vee,\otimes,\to$ of $\mathbf{2}$ are the truth functions (interpretations) of the corresponding logical connectives of the classical logic and $0^* = 0$, $1^* = 1$.

Having $\mathbf{L}$, we define usual notions: an $\mathbf{L}$-set (fuzzy set) A in universe U is a mapping $A : U \to L$, $A(u)$ being interpreted as “the degree to which u belongs to A”. If $U = \{u_1,\dots,u_n\}$ then A can be denoted by $A = \{{}^{a_1}\!/u_1,\dots,{}^{a_n}\!/u_n\}$ meaning that $A(u_i)$ equals a_i for each $i = 1,\dots,n$. For brevity, we introduce the following convention: we write $\{\dots,u,\dots\}$ instead of $\{\dots,{}^{1}\!/u,\dots\}$, and we also omit elements of U whose membership degree is zero. For example, we write $\{u,{}^{0.5}\!/v\}$ instead of $\{{}^{1}\!/u,{}^{0.5}\!/v,{}^{0}\!/w\}$ and so on. Let $\mathbf{L}^U$ denote the collection of all $\mathbf{L}$-sets in U. The operations with $\mathbf{L}$-sets are defined componentwise. For instance, the intersection of $\mathbf{L}$-sets $A,B \in \mathbf{L}^U$ is an $\mathbf{L}$-set $A \cap B$ in U such that $(A \cap B)(u) = A(u) \wedge B(u)$ for each $u \in U$, etc. Binary $\mathbf{L}$-relations (binary fuzzy relations) between X and Y can be thought of as $\mathbf{L}$-sets in the universe $X \times Y$. That is, a binary $\mathbf{L}$-relation $I \in \mathbf{L}^{X\times Y}$ between a set X and a set Y is a mapping assigning to each $x \in X$ and each $y \in Y$ a truth degree $I(x,y) \in L$ (a degree to which x and y are related by I).

Given $A,B \in \mathbf{L}^U$, we define a subsethood degree

$$S(A,B) = \textstyle\bigwedge_{u\in U}\bigl(A(u) \to B(u)\bigr), \tag{5}$$

which generalizes the classical subsethood relation $\subseteq$. $S(A,B)$ represents a degree to which A is a subset of B. In particular, we write $A \subseteq B$ iff $S(A,B) = 1$. As a consequence, $A \subseteq B$ iff $A(u) \le B(u)$ for each $u \in U$.

A fuzzy set $A \in \mathbf{L}^U$ is called crisp if, for each $u \in U$, $A(u) \in \{0,1\}$. Hence, by a slight abuse of notation, we can identify crisp fuzzy sets in U with the ordinary subsets of U.

In the following we use well-known properties of residuated lattices and fuzzy structures which can be found in monographs [1], [10], [11], [13]. Throughout the rest of the paper, $\mathbf{L}$ denotes an arbitrary complete residuated lattice with a hedge.

A system $\mathcal{S} = \{A_i \in \mathbf{L}^U \mid i \in I\}$ is said to be closed under S^*-intersections iff, for each $B \in \mathbf{L}^U$,

$$\textstyle\bigcap_{i\in I}(S(B,A_i)^* \to A_i) \in \mathcal{S}. \tag{6}$$

A system $\mathcal{S} \subseteq \mathbf{L}^U$ closed under S^*-intersections is called an $\mathbf{L}^*$-closure system in U. An $\mathbf{L}^*$-closure operator (or, a fuzzy closure operator with hedge *) on a set U is a mapping $C\colon \mathbf{L}^U \to \mathbf{L}^U$ satisfying, for each $A,B_1,B_2 \in \mathbf{L}^U$,

$$A \subseteq C(A), \tag{7}$$

$$S(B_1,B_2)^* \le S(C(B_1),C(B_2)), \tag{8}$$

$$C(A) = C(C(A)). \tag{9}$$

Note that if * is globalization, the notions of an $\mathbf{L}^*$-closure system and an $\mathbf{L}^*$-closure operator will become the usual notions of a fuzzy closure system and a fuzzy closure operator [1].

Theorem 1 (see [2]): Let $\mathcal{S} = \{A_i \in \mathbf{L}^U \mid i \in I\}$ be an $\mathbf{L}^*$-closure system. Then $C_{\mathcal{S}}\colon \mathbf{L}^U \to \mathbf{L}^U$ defined by

$$C_{\mathcal{S}}(B) = \textstyle\bigcap_{i\in I}(S(B,A_i)^* \to A_i) \tag{10}$$

is an $\mathbf{L}^*$-closure operator. Moreover, for each $B \in \mathbf{L}^U$, we have $B \in \mathcal{S}$ iff $B = C_{\mathcal{S}}(B)$.

Let $C\colon \mathbf{L}^U \to \mathbf{L}^U$ be an $\mathbf{L}^*$-closure operator. Then $\mathcal{S}_C = \{A \in \mathbf{L}^U \mid A = C(A)\}$ is an $\mathbf{L}^*$-closure system. ■

In addition to Theorem 1, there is a one-to-one correspondence between $\mathbf{L}^*$-closure operators on U and $\mathbf{L}^*$-closure systems on U [2]. In detail, let C be an $\mathbf{L}^*$-closure operator on U, $\mathcal{S}$ be an $\mathbf{L}^*$-closure system on U. Then we have $C = C_{\mathcal{S}_C}$ and $\mathcal{S} = \mathcal{S}_{C_{\mathcal{S}}}$, i.e. the mappings $C \mapsto \mathcal{S}_C$ and $\mathcal{S} \mapsto C_{\mathcal{S}}$ are mutually inverse. See [2] for more details.

III. Fuzzy Attribute Implications

A. Fuzzy attribute implications and their interpretation

In this section we introduce fuzzy attribute implications [4] and define their interpretation in fuzzy sets of attributes. Suppose Y is a finite nonempty set of attributes. A (*fuzzy*) *attribute implication* (*over attributes* Y) is an expression $A \Rightarrow B$, where $A,B \in \mathbf{L}^Y$ (A and B are fuzzy sets of attributes).

The intended meaning of $A \Rightarrow B$ is: “if it is (very) true that an object has all attributes from A, then it has also all attributes from B” with the logical connectives being given by $\mathbf{L}$. Fuzzy attribute implications are meant to be interpreted in data tables with fuzzy attributes [4]. A data table with fuzzy attributes can be seen is a triplet $\langle X,Y,I\rangle$ where X is a set of objects, Y is a finite set of attributes (the same as above in the definition of a fuzzy attribute implication), and $I \in \mathbf{L}^{X\times Y}$ is a binary $\mathbf{L}$-relation between X and Y assigning to each object $x \in X$ and each attribute $y \in Y$

a degree $I(x, y)$ to which x has y. $\langle X, Y, I\rangle$ can be seen as a table with rows and columns corresponding to objects $x \in X$ and attributes $y \in Y$, respectively, and table entries containing degrees $I(x, y)$.

A row of a table $\langle X, Y, I\rangle$ corresponding to an object $x \in X$ can be identified with a fuzzy set I_x of attributes to which an attribute $y \in Y$ belongs to a degree $I_x(y) = I(x, y)$. For fuzzy set $M \in \mathbf{L}^Y$ of attributes, we define a *degree* $||A \Rightarrow B||_M \in L$ *to which* $A \Rightarrow B$ *is true in* M by

$$||A \Rightarrow B||_M = S(A, M)^* \rightarrow S(B, M). \qquad (11)$$

It is easily seen that if M is a fuzzy set of attributes of some object x, i.e. $M = I_x$, then $||A \Rightarrow B||_M$ is the degree to which "if it is (very) true that x has all attributes from A then x has all attributes from B". More generally, for a system $\mathcal{M}$ of $\mathbf{L}$-sets in Y, define a degree $||A \Rightarrow B||_{\mathcal{M}}$ to which $A \Rightarrow B$ is true in (each M from) $\mathcal{M}$ by

$$||A \Rightarrow B||_{\mathcal{M}} = \textstyle\bigwedge_{M \in \mathcal{M}} ||A \Rightarrow B||_M. \qquad (12)$$

Finally, given a data table $\langle X, Y, I\rangle$ and putting $\mathcal{M} = \{I_x \,|\, x \in X\}$, $||A \Rightarrow B||_{\mathcal{M}}$ is a *degree to which* $A \Rightarrow B$ *is true in each row of table* $\mathcal{T}$, i.e. a degree to which "for each object $x \in X$: if it is (very) true that x has all attributes from A, then x has all attributes from B". This degree is denoted by $||A \Rightarrow B||_{\langle X,Y,I\rangle}$ and is called a *degree to which* $A \Rightarrow B$ *is true in* $\langle X, Y, I\rangle$.

B. Models, semantic entailment, and bases

We now focus on models of sets of fuzzy attribute implications and related notions as models will play a key role in our considerations on the expressive power of FAIs. Let T be a set of fuzzy attribute implications. $M \in \mathbf{L}^Y$ is called a *model of* T if $||A \Rightarrow B||_M = 1$ for each $A \Rightarrow B \in T$. The set of all models of T is denoted by $\mathrm{Mod}(T)$, i.e.

$$\mathrm{Mod}(T) = \{M \in \mathbf{L}^Y \,|\, \text{for all } A \Rightarrow B \in T\text{:} \quad ||A \Rightarrow B||_M = 1\}. \qquad (13)$$

Using results from [2], the set $\mathrm{Mod}(T)$ of all models of a set T of FAIs over Y is an $\mathbf{L}^*$-closure system on Y. The associated closure operator $C_{\mathrm{Mod}(T)}$ sends each $M \in \mathbf{L}^Y$ to the least model of T containing M. Using (12) with $\mathcal{M}$ being $\mathrm{Mod}(T)$, we define a degree $||A \Rightarrow B||_T$ to which $A \Rightarrow B$ *semantically follows from* T by

$$||A \Rightarrow B||_T = ||A \Rightarrow B||_{\mathrm{Mod}(T)}, \qquad (14)$$

i.e. the semantic entailment from T is defined as a degree to which a given FAIs is true in each model of T.

Let $\langle X, Y, I\rangle$ be a data table with fuzzy attributes. A set T of fuzzy attribute implications is called a *base of* $\langle X, Y, I\rangle$ if $||A \Rightarrow B||_T = ||A \Rightarrow B||_{\langle X,Y,I\rangle}$. Thus, T is a base of $\langle X, Y, I\rangle$ iff, for each $A \Rightarrow B$, the degree to which T entails $A \Rightarrow B$ coincides with the degree to which $A \Rightarrow B$ is true in $\langle X, Y, I\rangle$. If T is a base of $\langle X, Y, I\rangle$ and no proper subset of T is a base of $\langle X, Y, I\rangle$, then T is called a *non-redundant base of* $\langle X, Y, I\rangle$. Our previous results [4] have shown a characterization of (nonredundant) bases of data tables with fuzzy attributes: each table has at least one base and in many important cases, we are able to describe a nonredundant base which is also a minimal one (in terms of the number of FAIs contained).

C. Fuzzy concept lattices with hedge

Fuzzy concept lattices are particular structures of conceptual clusters which are induced by object-attribute data table with fuzzy attributes. In the sequel, we will use such structures to find counterexamples.

Given a data table with fuzzy attributes $\langle X, Y, I\rangle$, for $A \in \mathbf{L}^X$, $B \in \mathbf{L}^Y$ (i.e. A is a fuzzy set of objects, B is a fuzzy set of attributes), we define fuzzy sets $A^{\uparrow} \in \mathbf{L}^Y$ (fuzzy set of attributes), $B^{\downarrow} \in \mathbf{L}^X$ (fuzzy set of objects) by

$$A^{\uparrow}(y) = \textstyle\bigwedge_{x \in X} \big(A(x)^* \rightarrow I(x, y)\big), \qquad (15)$$

$$B^{\downarrow}(x) = \textstyle\bigwedge_{y \in Y} \big(B(y) \rightarrow I(x, y)\big). \qquad (16)$$

Operators ${}^{\downarrow}, {}^{\uparrow}$ form so-called Galois connection with hedge [2], [8]. Pairs $\langle A, B\rangle$, where $A \in \mathbf{L}^X$, $B \in \mathbf{L}^Y$, $A^{\uparrow} = B$, and $B^{\downarrow} = A$ are naturally interpreted as concepts (clusters) hidden in the input data represented by I. Namely, $A^{\uparrow} = B$ and $B^{\downarrow} = A$ say that B (so called intent) is the collection of all attributes shared by all objects from A, and A (so called extent) is the collection of all objects sharing all attributes from B. The structure of all concepts equipped with the subconcept-superconcept hierarchy is called a fuzzy concept lattice induced by $\langle X, Y, I\rangle$.

The following assertions show model-theoretical characterizations of bases of data tables with fuzzy attributes.

Theorem 2 (see [2]): Let $\langle X, Y, I\rangle$ *be a data table with fuzzy attributes. Then*

$$\mathrm{Int}(X, Y, I) = \{B \in \mathbf{L}^Y \,|\, B = B^{\downarrow\uparrow}\} \qquad (17)$$

is an $\mathbf{L}^*$*-closure system such that* ${}^{\downarrow\uparrow} : \mathbf{L}^Y \rightarrow \mathbf{L}^Y$ *is the corresponding* $\mathbf{L}^*$*-closure operator. As a consequence, we get* $\mathrm{Int}(X, Y, I) = \{B^{\downarrow\uparrow} \,|\, B \in \mathbf{L}^Y\}$. ■

Theorem 3 (see [4]): Let $\langle X, Y, I\rangle$ *be a data table with fuzzy attributes. Then a set* T *of fuzzy attribute implications is a base of* $\langle X, Y, I\rangle$ *iff* $\mathrm{Mod}(T) = \mathrm{Int}(X, Y, I)$. ■

IV. Expressive Power of FAIs

Using FAIs, we can express if-then dependencies among graded attributes, e.g., FAI $\{{}^{1}\!/u, {}^{0.6}\!/v\} \Rightarrow \{{}^{0.8}\!/w\}$ says that if an object has u (to degree 1) and v to degree 0.6, then it has w to degree 0.8. The degrees 1, 0.6, and 0.8 in the FAI can be seen as threshold truth degrees. Intuitively, FAIs with threshold degrees have greater "expressive power" than ordinary if-then dependency formulas of the form $A \Rightarrow B$, where A and B are ordinary subsets of Y. In this paper we prove this claim and more.

Obviously, the notion of "expressive power" needs to be clarified. We will take the model-theoretical approach. For each family of formulas $\mathfrak{F}$, we can consider all systems of models which may be delimited by all possible theories (collections of formulas) from $\mathfrak{F}$. Denote this collection of all possible systems of models by $\mathrm{Mod}(\mathfrak{F})$. Hence, for each theory T consisting of formulas from $\mathfrak{F}$, $\mathrm{Mod}(\mathfrak{F})$ contains the system of all models of T. Clearly, if $\mathfrak{F}_1$ and $\mathfrak{F}_2$ are two families of formulas, $\mathfrak{F}_1$ has greater expressive power than $\mathfrak{F}_2$ if $\text{Mod}(\mathfrak{F}_2) \subseteq \text{Mod}(\mathfrak{F}_1)$. In other words, $\mathfrak{F}_1$ has greater expressive power than $\mathfrak{F}_2$ if each model system which can be delimited by formulas from $\mathfrak{F}_2$ can also be delimited by formulas from $\mathfrak{F}_1$, i.e. the collection of all systems of models that can be described by $\mathfrak{F}_1$ is greater.

Let us take a look at the expressive power of FAIs and related formulas. Due to observations from [2], [6], $\mathbf{L}^*$-closure systems are exactly the systems of models of FAIs. In a more detail, for each set T of FAIs over Y, $\text{Mod}(T)$ is an $\mathbf{L}^*$-closure system [2] in Y and, conversely, for each $\mathbf{L}^*$-closure system $\mathcal{S}$ in Y, there is a set T of FAIs such that $\mathcal{S} = \text{Mod}(T)$. Namely, we can take

$$T = \{A \Rightarrow C_{\mathcal{S}}(A) \mid A \in \mathbf{L}^Y\},$$

where $C_{\mathcal{S}}$ is the $\mathbf{L}^*$-closure operator induced by $\mathcal{S}$, see [6]. Thus, if $\mathfrak{F}$ denotes the family of all FAIs over Y, then

$$\text{Mod}(\mathfrak{F}) = \{\text{Mod}(T) \mid T \subseteq \mathfrak{F}\} \tag{18}$$

is the collection of all $\mathbf{L}^*$-closure systems in Y. Thus, (18) gives us a clear description of the expressive power of FAIs.

We will now compare the expressive power of FAIs with other types of formulas. Namely, we focus on if-then formulas which do not have threshold truth degrees. Such formulas take form of implications between two groups of (symbols of) attributes connected by conjunctions:

$$(y_1 \,\&\cdots\&\, y_m) \Rightarrow (y'_1 \,\&\cdots\&\, y'_n), \tag{19}$$

with the usual truth-functional interpretation. For simplicity, observe that formulas (19) can be seen as FAIs which allow for threshold truth degrees 0 and 1, only. Indeed, fuzzy attribute implication $A \Rightarrow B$ (over Y) will be called *crisp* if both A and B are crisp fuzzy sets. We can see that (19) can be seen as a FAI

$$\{y_1, \ldots, y_m\} \Rightarrow \{y'_1, \ldots, y'_n\}, \tag{20}$$

and the interpretation of (19) coincides with the interpretation of the corresponding FAI (with $\&$ interpreted as the min-conjunction). Thus, the ordinary if-then rules (classical attribute implications [9]) with their usual interpretation can be seen as particular FAIs.

We now show that crisp FAIs have strictly smaller expressive power than the general FAIs. That is, we are going to show that there are $\mathbf{L}^*$-closure systems which cannot be described by crisp FAIs. We prove even more, because we focus on Pavelka-style entailment [15] from crisp FAIs and related model classes. Even in this more general setting, crisp FAIs are not as expressive as general FAIs.

Recall that the abstract fuzzy logic (also known as Pavelka-style fuzzy logic) considers entailment from fuzzy sets of formulas. In our setting, we are interested in fuzzy sets of crisp FAIs. Let T be a fuzzy set of crisp FAIs. Then, $M \in \mathbf{L}^Y$ is called a model of T if, for each crisp $A \Rightarrow B$,

$$T(A \Rightarrow B) \leq ||A \Rightarrow B||_M. \tag{21}$$

Let $\text{Mod}(T)$ denote the set of all models of a fuzzy set T of crisp FAIs. The degree $||A \Rightarrow B||_T$ to which $A \Rightarrow B$ semantically follows from a fuzzy set T of crisp FAIs is defined by $||A \Rightarrow B||_T = ||A \Rightarrow B||_{\text{Mod}(T)}$.

For technical reasons, we now show that semantic entailment from fuzzy sets of crisp FAIs is reducible to semantic entailment from ordinary sets of certain general FAIs. Here we use a procedure which is a particularization of a more general procedure which we have introduced in [4]. Let T be a fuzzy set of crisp FAIs. For each crisp FAI of the from (20) and a truth degree $b \in L$, we can consider a FAI defined by

$$\{y_1, \ldots, y_m\} \Rightarrow \{{}^{b}/y'_1, \ldots, {}^{b}/y'_n\}, \tag{22}$$

i.e. the antecedent of the FAI remains the same and the consequent contains threshold degrees b instead of the original threshold degrees 1. Denote the resulting fuzzy attribute implication by $c(A \Rightarrow B, b)$, where $A \Rightarrow B$ abbreviates (20).

Now, given a fuzzy set T of crisp FAIs, we can construct a set $c(T)$ of FAIs as follows

$$c(T) = \{c(A \Rightarrow B, b) \mid T(A \Rightarrow B) = b > 0\}.$$

One can check [4] that

$$\text{Mod}(T) = \text{Mod}(c(T)).$$

In other words, fuzzy sets of crisp FAIs have the same expressive power as ordinary sets composed of FAIs of the form $c(A \Rightarrow B, b)$. Thus, in order to show that the expressive power of general FAIs is strictly greater than the expressive power of crisp FAIs with Pavelka-style entailment it suffices to find an $\mathbf{L}^*$-closure system which is not a system of models of any set of FAIs of the form $c(A \Rightarrow B, b)$.

In the next section we show particular $\mathbf{L}^*$-closure systems which serve as counterexamples. We will induce these $\mathbf{L}^*$-closure systems by data tables with fuzzy attributes. This will show that the $\mathbf{L}^*$-closure systems, which cannot be delimited by these "weaker formulas", naturally appear in the analysis of object-attribute data with graded attributes. As a consequence, the weaker formulas cannot be used to describe systems of intents of formal concepts, i.e. the formulas are not sufficient for particular data mining tasks studied earlier related to data with graded attributes, see [4] for a survey.

V. Analysis and the Counterexamples

In this section, we present counterexamples to the claim that fuzzy sets of crisp fuzzy attribute implications have the same expressive power as sets of (general) fuzzy attribute implications. We will present minimal examples which the least truth degrees possible. We focus on founding counterexamples which use identity and globalization on Łukasiewicz and Gödel chains because these structures of truth degrees are perhaps the most widely used structures of truth degrees in fuzzy logic.

We will see that it is sufficient to take just three truth degrees: 0 (falsity), 0.5 (mid falsity/truth), and 1 (full truth). Thus, we put $L = \{0, 0.5, 1\}$ with the natural ordering $0 < 0.5 < 1$. In addition to that, we will consider operations on L given by (1) and (2), and two hedges: identity and globalization (4). For each of these four residuated lattices with hedges, we show a counterexample. That is, considering each of the structures of truth degrees, we are going to show a particular $\mathbf{L}^*$-closure system which cannot be delimited as a system of models of any fuzzy set of crisp FAIs. All the $\mathbf{L}^*$-closure systems used as counterexamples will be induced by data tables with graded attributes because we want to show that the counterexamples quite naturally appear in data analysis, i.e., that they are not just some artificially constructed structures.

A. Counterexample for globalization and/or Gödel chain

We begin with the simpler case. Consider the following data table with fuzzy attributes:

	y
x	0

That is, we have $X = \{x\}$, $Y = \{y\}$, and $I(x, y) = 0$. If

- the hedge * is globalization (and $\otimes$ arbitrary), or
- the hedge * is identity and $\otimes$ is the Gödel multiplication,

then in both cases we get $\mathrm{Int}(X, Y, I) = \{\{\}, \{y\}\}$. Due to Theorem 2, $\mathrm{Int}(X, Y, I)$ is an $\mathbf{L}^*$-closure system. We now claim that this particular closure system cannot be described by a fuzzy set of crisp fuzzy attribute implications. That is, we prove that there is no fuzzy set T of crisp FAIs such that $\mathrm{Mod}(T) = \mathrm{Int}(X, Y, I) = \{\{\}, \{y\}\}$.

Since we have a single attribute $Y = \{y\}$, we have four pairwise different crisp FAIs:

$$\{\} \Rightarrow \{\}, \quad \{\} \Rightarrow \{y\}, \quad \{y\} \Rightarrow \{\}, \quad \{y\} \Rightarrow \{y\}.$$

Clearly, $\{\} \Rightarrow \{\}$, $\{y\} \Rightarrow \{\}$, and $\{y\} \Rightarrow \{y\}$ are true in each $M \in \mathbf{L}^Y$ to degree 1, i.e. such FAIs cannot be used to "disqualify" any model. In a more detail, if we have two fuzzy sets T and T' of crisp FAIs such that

(i) for each $M \in \mathbf{L}^Y$: $||A \Rightarrow B||_M = 1$, and
(ii) for each $C \Rightarrow D$ different from $A \Rightarrow B$: $T(C \Rightarrow D) = T'(C \Rightarrow D)$,

then $\mathrm{Mod}(T) = \mathrm{Mod}(T')$. From this point of view, the degrees to which $\{\} \Rightarrow \{\}$, $\{y\} \Rightarrow \{\}$, and $\{y\} \Rightarrow \{y\}$ belong to a theory are not important because these formulas do not have an influence on the set of models of such theory.

Hence, the only nontrivial crisp fuzzy attribute implication is $\{\} \Rightarrow \{y\}$. Since we are interested in theories which consist of fuzzy sets of FAIs, we have just two nontrivial theories:

$$T_1(\{\} \Rightarrow \{y\}) = 0.5,$$
$$T_2(\{\} \Rightarrow \{y\}) = 1.$$

Again, we do not consider $T(\{\} \Rightarrow \{y\}) = 0$ because each fuzzy set of attributes is a model of such T. The latter fuzzy sets T_1 and T_2 of crisp fuzzy attribute implications correspond with the following ordinary sets of (general) fuzzy attribute implications:

$$c(T_1) = \{\{\} \Rightarrow \{{}^{0.5}/y\}\},$$
$$c(T_2) = \{\{\} \Rightarrow \{y\}\}.$$

From previous observations,

$$\mathrm{Mod}(c(T_1)) = \mathrm{Mod}(T_1),$$
$$\mathrm{Mod}(c(T_2)) = \mathrm{Mod}(T_2).$$

Clearly, if * is globalization (4) or $\otimes$ is the Gödel multiplication (2), we have

$$\mathrm{Mod}(c(T_1)) = \{\{{}^{0.5}/y\}, \{y\}\}, \quad (23)$$
$$\mathrm{Mod}(c(T_2)) = \{\{y\}\}. \quad (24)$$

Therefore, the $\mathbf{L}^*$-closure system $\mathrm{Int}(X, Y, I) = \{\{\}, \{y\}\}$ induced by a data table with $I(x, y) = 0$ cannot be described by a fuzzy set of crisp fuzzy attribute implications because the only $\mathbf{L}^*$-closure systems which can be described by fuzzy set of crisp FAIs are (23) and (24), respectively. That concludes our first counterexample.

B. Counterexample for a Łukasiewicz chain with identity

The previous counterexample did not include the situation where * is identity and $\otimes$ is the Łukasiewicz multiplication, see (1). In this case, we can also find an $\mathbf{L}^*$-closure system which is not a model class of any fuzzy set of crisp FAIs. In this case, however, we cannot induce the $\mathbf{L}^*$-closure system by a data table with just one attribute. Indeed, if we take

	y
x	0

,

we get $\mathrm{Int}(X, Y, I) = \{\{\}, \{{}^{0.5}/y\}, \{y\}\}$. Hence, if we take $T = \emptyset$, i.e., for each $A \Rightarrow B$, $T(A \Rightarrow B) = 0$, we have $\mathrm{Int}(X, Y, I) = \{\{\}, \{{}^{0.5}/y\}, \{y\}\} = \mathrm{Mod}(T)$. In other words, such an $\mathbf{L}^*$-closure system equals to $\mathbf{L}^Y$, which is a system that can trivially be described by the empty theory. If we take

	y
x	0.5

,

we obtain $\mathrm{Int}(X, Y, I) = \{\{{}^{0.5}/y\}, \{y\}\}$. In this case, we can take a fuzzy set T of crisp FAIs such that $T(\{\} \Rightarrow \{y\}) = 0.5$. One can check that for such T, $\mathrm{Int}(X, Y, I) = \mathrm{Mod}(T)$. Finally, for

	y
x	1

,

we have $\mathrm{Int}(X, Y, I) = \{\{y\}\}$, which can be described by a fuzzy set T of crisp FAIs such that $T(\{\} \Rightarrow \{y\}) = 1$.

Therefore, we must consider $\mathrm{Int}(X, Y, I)$ induced by a larger data table. We will see that it is sufficient to keep the only object x and two attributes y and z. That is, we consider $X = \{x\}$, $Y = \{y, z\}$, and I given by the following table:

	y	z
x	0	0.5

One can check that we have

$$\mathrm{Int}(X, Y, I) = \{\{{}^{0.5}/z\}, \{{}^{0.5}/y, z\}, \{y, z\}\}. \quad (25)$$

Again, we will show that there is no fuzzy set of crisp fuzzy attribute implications over $Y = \{y, z\}$ such that its models are exactly the fuzzy sets of attributes from $\mathrm{Int}(X, Y, I)$. In this case, we have $(2^2)^2 = 16$ pairwise different crisp FAIs over Y, namely:

$$\begin{array}{llll} \{\}\Rightarrow\{\}, & \{\}\Rightarrow\{y\}, & \{\}\Rightarrow\{z\}, & \{\}\Rightarrow\{y,z\}, \\ \{y\}\Rightarrow\{\}, & \{y\}\Rightarrow\{y\}, & \{y\}\Rightarrow\{z\}, & \{y\}\Rightarrow\{y,z\}, \\ \{z\}\Rightarrow\{\}, & \{z\}\Rightarrow\{y\}, & \{z\}\Rightarrow\{z\}, & \{z\}\Rightarrow\{y,z\}, \\ \{y,z\}\Rightarrow\{\}, & \{y,z\}\Rightarrow\{y\}, & \{y,z\}\Rightarrow\{z\}, & \{y,z\}\Rightarrow\{y,z\}. \end{array}$$

Only a few of the above-listed FAIs can be used to disqualify models. First, we can omit all FAIs $A \Rightarrow B$ such that $B \subseteq A$ because such FAIs are true in each fuzzy sets of attributes to degree 1. Hence, following the arguments from the previous counterexample, their appearance in fuzzy sets of crisp FAIs in not interesting from the model-theoretical point of view. Therefore, our set of crisp FAIs reduces to:

$$\begin{array}{llll} \{\}\Rightarrow\{y\}, & \{\}\Rightarrow\{z\}, & \{\}\Rightarrow\{y,z\}, & \{y\}\Rightarrow\{z\}, \\ \{y\}\Rightarrow\{y,z\}, & \{z\}\Rightarrow\{y\}, & \{z\}\Rightarrow\{y,z\}. & \end{array}$$

Furthermore, we can skip all crisp fuzzy attribute implications $A \Rightarrow B$ such that $B \cap A \neq \emptyset$. Indeed, this is due to the fact that for each FAI $A \Rightarrow B$ over Y and $B' \in \mathbf{L}^Y$ such that

$$B'(y) = \begin{cases} 0, & \text{if } B(y) \leq A(y), \\ B(y), & \text{otherwise,} \end{cases}$$

we have $||A \Rightarrow B||_M = ||A \Rightarrow B'||_M$ for each $M \in \mathbf{L}^Y$. Hence, the original set of 16 crisp fuzzy attribute implications reduces to

$$\{\}\Rightarrow\{y\}, \ \{\}\Rightarrow\{z\}, \ \{\}\Rightarrow\{y,z\}, \ \{y\}\Rightarrow\{z\}, \ \{z\}\Rightarrow\{y\}.$$

Suppose now, by contradiction, that $\mathrm{Int}(X, Y, I)$ given by (25) can be described by a fuzzy set of the previous five crisp FAIs. That is, we assume that there is a fuzzy set T of crisp FAIs, containing at most the previous five FAIs to a nonzero degree, such that $\mathrm{Mod}(T) = \mathrm{Int}(X, Y, I)$. Let us analyze how T would look like. Since $\mathrm{Mod}(T) \neq \mathbf{L}^Y$, T must be nonempty, i.e. there must be a crisp FAI $A \Rightarrow B$ such that $T(A \Rightarrow B) \neq 0$. Observe that for $\{{}^{0.5}\!/z\} \in \mathrm{Int}(X, Y, I)$,

$$\begin{aligned} ||\{\}\Rightarrow\{y\}||_{\{0.5/z\}} &= 1 \wedge 1 \rightarrow 0 \wedge 1 = 1 \rightarrow 0 = 0, \\ ||\{\}\Rightarrow\{z\}||_{\{0.5/z\}} &= 1 \wedge 1 \rightarrow 1 \wedge 0.5 = 1 \rightarrow 0.5 = 0.5, \\ ||\{\}\Rightarrow\{y,z\}||_{\{0.5/z\}} &= 1 \wedge 1 \rightarrow 0 \wedge 0.5 = 1 \rightarrow 0 = 0, \\ ||\{y\}\Rightarrow\{z\}||_{\{0.5/z\}} &= 0 \wedge 1 \rightarrow 1 \wedge 0.5 = 0 \rightarrow 0.5 = 1, \\ ||\{z\}\Rightarrow\{y\}||_{\{0.5/z\}} &= 1 \wedge 0.5 \rightarrow 0 \wedge 1 = 0.5 \rightarrow 0 = 0.5. \end{aligned}$$

Since we assume that $\{{}^{0.5}\!/z\} \in \mathrm{Mod}(T)$, the truth degrees $||\cdots||_{\{0.5/z\}}$ computed above yield that T must be a union of some of the following theories:

$$\begin{array}{ll} T_1(\{\}\Rightarrow\{z\}) = 0.5, & T_1(\cdots) = 0 \text{ otherwise}, \\ T_2(\{y\}\Rightarrow\{z\}) = 0.5, & T_2(\cdots) = 0 \text{ otherwise}, \\ T_3(\{y\}\Rightarrow\{z\}) = 1, & T_3(\cdots) = 0 \text{ otherwise}, \\ T_4(\{z\}\Rightarrow\{y\}) = 0.5, & T_4(\cdots) = 0 \text{ otherwise}. \end{array}$$

That is, $T = T_{i_1} \cup \cdots \cup T_{i_k}$, where $\{i_1, \ldots, i_k\} \subseteq \{1, \ldots, 4\}$. Note that T must be a union of at least one of $T_1, \ldots, T_4$ because $\mathrm{Mod}(T) \neq \mathbf{L}^Y$. Now, put $M = \{{}^{0.5}\!/y, {}^{0.5}\!/z\}$ and observe that

$$\begin{aligned} ||\{\}\Rightarrow\{z\}||_M &= 1 \wedge 1 \rightarrow 1 \wedge 0.5 = 1 \rightarrow 0.5 = 0.5, \\ ||\{y\}\Rightarrow\{z\}||_M &= 0.5 \wedge 1 \rightarrow 1 \wedge 0.5 = 0.5 \rightarrow 0.5 = 1, \\ ||\{z\}\Rightarrow\{y\}||_M &= 1 \wedge 0.5 \rightarrow 0.5 \wedge 1 = 0.5 \rightarrow 0.5 = 1. \end{aligned}$$

That is, $M = \{{}^{0.5}\!/y, {}^{0.5}\!/z\}$ is a model of each $T_1, \ldots, T_4$. Therefore, M is also a model of any union of $T_1, \ldots, T_4$. As a consequence, M must be a model of T and, consequently, it must belong to $\mathrm{Int}(X, Y, I)$ because we have assumed $\mathrm{Mod}(T) = \mathrm{Int}(X, Y, I)$—a contradiction to the fact that $M \notin \mathrm{Int}(X, Y, I)$.

Remark: Let us stress that general FAIs can be used as an alternative means for describing concept lattices. Each concept lattice has its nonredundant base T, consisting of FAIs, such that the models of T are exactly the intents of conceptual clusters [4]. By the counterexamples presented in this section we have demonstrated that weaker formulas are not powerful enough to do this job.

Conclusions: We have shown that sets of fuzzy attribute implications have strictly greater expressive power than fuzzy sets of the ordinary attribute implications by providing several closure structures that can be described by sets of fuzzy attribute implications but cannot be described by fuzzy sets of ordinary attribute implications. This has several practical consequences for mining nonredundant bases of object-attribute data tables with fuzzy attributes. Our future research will focus on the expressive power of formulas with semantics related to FAIs, e.g., fuzzy Horn clauses, see [3].

REFERENCES

[1] Belohlavek R.: *Fuzzy Relational Systems: Foundations and Principles*. Kluwer Academic/Plenum Publishers, New York, 2002.

[2] Belohlavek R., Funioková T., Vychodil V.: Fuzzy closure operators with truth stressers. *Logic J. of IGPL* **13**(5)(2005), 503–513.

[3] Belohlavek R., Vychodil V.: *Fuzzy Equational Logic*. Springer, Berlin, 2005.

[4] Belohlavek R., Vychodil V.: Attribute implications in a fuzzy setting. In: Missaoui R., Schmid J. (Eds.): *Proc. ICFCA 2006*, LNAI **3874**, pp. 45–60, 2006.

[5] Belohlavek R., Vychodil V.: Data tables with similarity relations: functional dependencies, complete rules and non-redundant bases. In: Lee M. L., Tan K. L., and Wuwongse V. (Eds.): *Proc. DASFAA 2006*, LNCS **3882**, pp. 644–658, 2006.

[6] Belohlavek R., Vychodil V.: Properties of models of fuzzy attribute implications. *Proc. SCIS & ISIS 2006,* pp. 291–296, Tokyo Institute of Technology, 2006, ISSN 1880–3741.

[7] Belohlavek R., Vychodil V.: Fuzzy attribute logic over complete residuated lattices. *J. Exp. Theor. Artif. Int.* **18**(4)(2006), 471–480.

[8] Belohlavek R., Vychodil V.: Fuzzy concept lattices constrained by hedges. *JACIII* **11**(6)(2007), 536–545.

[9] Ganter B., Wille R.: *Formal Concept Analysis. Mathematical Foundations.* Springer, 1999.

[10] Gerla G.: *Fuzzy Logic. Mathematical Tools for Approximate Reasoning.* Kluwer, Dordrecht, 2001.

[11] Hájek P.: *Metamathematics of Fuzzy Logic.* Kluwer, Dordrecht, 1998.

[12] Hájek P.: On very true. *Fuzzy Sets and Systems* **124**(2001), 329–333.

[13] Klir G. J., Yuan B.: *Fuzzy Sets and Fuzzy Logic. Theory and Applications.* Prentice Hall, 1995.

[14] Maier D.: *The Theory of Relational Databases.* Computer Science Press, Rockville, 1983.

[15] Pavelka J.: On fuzzy logic I, II, III. *Z. Math. Logik Grundlagen Math.* **25**(1979), 45–52, 119–134, 447–464.

[16] Takeuti G., Titani S.: Globalization of intuitionistic set theory. *Annals of Pure and Applied Logic* **33**(1987), 195–211.