

Properties of models of fuzzy attribute implications

Radim Bělohlávek and Vilém Vychodil

Dept. Computer Science, Palacký University, Tomkova 40, CZ-779 00, Olomouc, Czech Republic

email: {radim.belohlavek, vilem.vychodil}@upol.cz

Abstract—This paper deals with closure properties of models of fuzzy attribute implications. Fuzzy attribute implications are particular IF-THEN rules which can be interpreted in data tables with fuzzy attributes and/or ranked data tables over domains with similarity relations. We show that models of any set of fuzzy attribute implications form a fuzzy closure system with hedge. Conversely, we show that each fuzzy closure system with hedge can be seen as a system of models of some set of fuzzy attribute implications. Furthermore, we show applications of the closure properties: we describe semantic entailment from fuzzy attribute implications using least models and show a method for generating of non-redundant sets of dependencies which are true in a given data table with fuzzy attributes.

Keywords—completeness in data, fuzzy attribute implication, fuzzy closure operator, hedge, model, non-redundant basis

I. INTRODUCTION

The study of closure properties of model classes of various families of formulas proved to be important from both the theoretical and applicational points of view. Model-theoretical properties of theories are the subject of study of model theory [10] which is now a well-developed part of mathematical logic. In this paper we deal with closure properties of models of fuzzy attribute implications and show their applications in analysis of data with fuzzy attributes. Fuzzy attribute implications (FAIs) were thoroughly studied in [2], [4], [5], [6], [8], see also [16] for a related approach. FAIs are formulas of the form $A \Rightarrow B$, where A and B are fuzzy sets of attributes, meaning “if A , then B ”. In our previous papers we introduced two basic interpretations of FAIs: (i) interpretation in data tables with fuzzy attributes, (ii) interpretation in ranked data tables over domains with similarities, see [6] and [7] for surveys. In this paper we focus on closure properties of models of FAIs. For technical reasons, we focus only on the first type of interpretation of FAIs. On the other hand, the presented results can also be extended for the other interpretation. We show that there is a one-to-one correspondence between collections of all models of sets of FAIs and particular fuzzy closure systems. This result will be further used to obtain a characterization of semantic entailment from sets of FAIs using the concept of a least model. Finally, we show an application of the results for data analysis—a method for computing of non-redundant bases of data tables with fuzzy attributes.

II. PRELIMINARIES

In this section we survey basic notions of fuzzy logic, fuzzy sets, and fuzzy closure operators with hedges which will be used in further sections. A complete residuated lattice with hedge, which is our basic structure of truth degrees, 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 $a \leq b \rightarrow c$ ($a, b, c \in L$); hedge $*$ satisfies, for each $a, b \in L$,

$$1^* = 1, \quad (1)$$

$$a^* \leq a, \quad (2)$$

$$(a \rightarrow b)^* \leq a^* \rightarrow b^*, \quad (3)$$

$$a^{**} = a^*. \quad (4)$$

Each $a \in L$ are called a truth degree. \otimes and \rightarrow are (truth functions of) “fuzzy conjunction” and “fuzzy implication”. Hedge $*$ is a (truth function of) logical connective “very true”, see [13], [14]. Properties (1)–(4) have natural interpretations, e.g. (2) can be read: “if a is very true, then a is true”, (3) can be read: “if $a \rightarrow b$ 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 \rightarrow . Three most important pairs of adjoint operations on the unit interval are: Łukasiewicz ($a \otimes b = \max(a + b - 1, 0)$, $a \rightarrow b = \min(1 - a + b, 1)$), Gödel ($a \otimes b = \min(a, b)$, $a \rightarrow b = 1$ if $a \leq b$, $a \rightarrow b = b$ else), Goguen (product): ($a \otimes b = a \cdot b$, $a \rightarrow b = 1$ if $a \leq b$, $a \rightarrow b = \frac{b}{a}$ else). Complete residuated lattices include also finite structures of truth degrees. For instance, one can put $L = \{a_0 = 0, a_1, \dots, a_n = 1\} \subseteq [0, 1]$ ($a_0 < \dots < a_n$) with \otimes given by $a_k \otimes a_l = a_{\max(k+l-n, 0)}$ and the corresponding \rightarrow given by $a_k \rightarrow a_l = a_{\min(n-k+l, n)}$. Such an \mathbf{L} is called a finite Łukasiewicz chain. Another possibility is a finite Gödel chain which consists of L and restrictions of Gödel operations on $[0, 1]$ to L . A special case of both of these chains is the Boolean algebra with $L = \{0, 1\}$ (structure of truth degrees of classical logic).

Note that two boundary cases of hedges are (i) identity, i.e. $a^* = a$ ($a \in L$); (ii) globalization [17]:

$$a^* = \begin{cases} 1 & \text{if } a = 1, \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

Given \mathbf{L} which serves as a structure of truth degrees, we define usual notions: an \mathbf{L} -set (fuzzy set) A in universe U is a mapping $A : U \rightarrow L$, $A(u)$ being interpreted as “the degree to which u belongs to A ”. If U is a finite universe $U = \{u_1, \dots, u_n\}$ then an \mathbf{L} -set A in U can be denoted by $A = \{a_1/u_1, \dots, a_n/u_n\}$ meaning that $A(u_i)$ equals a_i ($i = 1, \dots, n$). 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 $\{1/u, 0.5/v, 0/w\}$, etc. 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, intersection of \mathbf{L} -sets $A, B \in \mathbf{L}^U$ is an \mathbf{L} -set $A \cap B$ in U such that $(A \cap B)(u) = A(u) \wedge B(u)$ ($u \in U$); $A^*(u) = A(u)^*$ ($u \in U$). For $a \in L$ and $A \in \mathbf{L}^U$, we define \mathbf{L} -sets $a \otimes A$ (a -multiple of A) and $a \rightarrow A$ (a -shift of A) by $(a \otimes A)(u) = a \otimes A(u)$, $(a \rightarrow A)(u) = a \rightarrow A(u)$ ($u \in U$). Binary \mathbf{L} -relations (binary fuzzy relations) between U and V can be thought of as \mathbf{L} -sets in $U \times V$. Given $A, B \in \mathbf{L}^U$, we define a subsethood degree

$$S(A, B) = \bigwedge_{u \in U} (A(u) \rightarrow B(u)), \quad (6)$$

which generalizes the classical subsethood relation \subseteq in a fuzzy setting. Described verbally, $S(A, B)$ 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, we have $A \subseteq B$ iff $A(u) \leq B(u)$ for each $u \in U$. In the following we use well-known properties of residuated lattices and fuzzy structures which can be found in monographs [1], [13].

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 $B \in \mathbf{L}^U$,

$$\bigcap_{i \in I} (S(B, A_i)^* \rightarrow A_i) \in \mathcal{S}. \quad (7)$$

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 $C: \mathbf{L}^U \rightarrow \mathbf{L}^U$ satisfying, for each $A, B_1, B_2 \in \mathbf{L}^U$,

$$A \subseteq C(A), \quad (8)$$

$$S(B_1, B_2)^* \leq S(C(B_1), C(B_2)), \quad (9)$$

$$C(A) = C(C(A)). \quad (10)$$

Theorem 1 (see [3]): $C: \mathbf{L}^U \rightarrow \mathbf{L}^U$ is an \mathbf{L}^* -closure operator on U iff it satisfies (8) and the following condition:

$$S(B_1, C(B_2))^* \leq S(C(B_1), C(B_2)). \quad (11)$$

A system $\mathcal{S} \subseteq \mathbf{L}^U$ which is closed under arbitrary intersections is an \mathbf{L}^* -closure system iff, for each $a \in L$ and $A \in \mathcal{S}$, we have $a^* \rightarrow A \in \mathcal{S}$. ■

Theorem 2 (see [3]): Let $\mathcal{S} = \{A_i \in \mathbf{L}^U \mid i \in I\}$ be an \mathbf{L}^* -closure system. Then $C_{\mathcal{S}}: \mathbf{L}^U \rightarrow \mathbf{L}^U$ defined by

$$C_{\mathcal{S}}(B) = \bigcap_{i \in I} (S(B, A_i)^* \rightarrow A_i) \quad (12)$$

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 $C: \mathbf{L}^U \rightarrow \mathbf{L}^U$ 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 2, there is a one-to-one correspondence between \mathbf{L}^* -closure operators on U and \mathbf{L}^* -closure systems on U [3]. In detail, let C be an \mathbf{L}^* -closure operator on U , \mathcal{S} be an \mathbf{L}^* -closure system on U . Then \mathcal{S}_C is an \mathbf{L}^* -closure system on U , $C_{\mathcal{S}}$ is an \mathbf{L}^* -closure operator on U , and 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 [3] for more details.

III. FUZZY ATTRIBUTE IMPLICATIONS

We now recall basic notions of fuzzy attribute logic (FAL). More details can be found in [2], [4], [6].

Let Y denote a *finite set of attributes*, each $y \in Y$ will be called an *attribute*. A *fuzzy attribute implication (over attributes Y)* is an expression $A \Rightarrow B$, where $A, B \in \mathbf{L}^Y$ are fuzzy sets of attributes. Fuzzy attribute implications (FAIs) are the formulas of fuzzy attribute logic. In order to consider semantic validity (truth) of FAIs, we introduce a semantic component in which we evaluate FAIs and their formal interpretation. The intuitive meaning we wish to give to $A \Rightarrow B$ is: “if it is (very) true that an object has all attributes from A , then it has also all attributes from B ”. Formally, for an \mathbf{L} -set $M \in \mathbf{L}^Y$ of attributes, we define a *truth degree* $\|A \Rightarrow B\|_M \in L$ to which $A \Rightarrow B$ is true in M by

$$\|A \Rightarrow B\|_M = S(A, M)^* \rightarrow S(B, M), \quad (13)$$

where $S(\cdot \cdot \cdot)$ denote subsethood degrees, see (6). The degree $\|A \Rightarrow B\|_M$ can be understood as follows: if M (semantic component) represents presence of attributes of some object, i.e. $M(y)$ is truth degree to which “the object has the attribute $y \in Y$ ”, then $\|A \Rightarrow B\|_M$ is the truth degree to which “if the object has all attributes from A , then it has all attributes from B ”, which corresponds to the desired interpretation of $A \Rightarrow B$. Note also that the hedge $*$ servers as a modifier of interpretation of $A \Rightarrow B$, see [2], [4], [6] for details.

Let T be a set of fuzzy attribute implications. An \mathbf{L} -set $M \in \mathbf{L}^Y$ is called a *model of T* if, for each $A \Rightarrow B \in T$, $\|A \Rightarrow B\|_M = 1$. The set of all models of T will be denoted by $\text{Mod}(T)$, i.e.

$$\text{Mod}(T) = \{M \in \mathbf{L}^Y \mid \text{for each } A \Rightarrow B \in T: \|A \Rightarrow B\|_M = 1\}. \quad (14)$$

A degree $\|A \Rightarrow B\|_T \in L$ to which $A \Rightarrow B$ semantically follows from T is defined by

$$\|A \Rightarrow B\|_T = \bigwedge_{M \in \text{Mod}(T)} \|A \Rightarrow B\|_M. \quad (15)$$

Described verbally, $\|A \Rightarrow B\|_T$ is defined to be the degree to which “ $A \Rightarrow B$ is true in each model of T ”. Hence, degrees $\|\cdot \cdot \cdot\|_T$ defined by (15) represent degrees of semantic entailment from T . Let us note that degrees $\|\cdot \cdot \cdot\|_T$ can also be fully described via the (syntactic) concept of a *provability degree*, see [6] for a survey.

IV. CLOSURE PROPERTIES OF MODELS OF FAIS

In this section we focus on problems related with closure properties of models of sets of fuzzy attribute implications. We show that all models of a set of fuzzy attribute implications form an \mathbf{L}^* -closure system. Furthermore, we show several important properties of the closure operator (with hedge) induced by the system of all models.

Until otherwise mentioned, we let \mathbf{L} be a complete residuated lattice with hedge $*$, Y be a set of attributes, T be a set of FAIs over Y . Notice that $\text{Mod}(T)$ is a collection of fuzzy sets in universe Y , i.e. $\text{Mod}(T) \subseteq \mathbf{L}^Y$. Also, note that each degree of truth $\|\cdot \cdot \cdot\|_M$ in M and thus all the derived notions like model and semantic entailment depend on $*$. That

is, for two complete residuated lattices with hedges \mathbf{L}_1 and \mathbf{L}_2 which differ only in definitions of hedges, a set of FAIs T has possibly different models. This has particular relevance for the closure properties of models which depend on the hedge as it is shown below. In classical setting, the collection of all models of a set of attribute implications form a closure system. In our setting, we have the following

Theorem 3: $\text{Mod}(T)$ is an \mathbf{L}^* -closure system in Y .

Proof: Due to Theorem 1, it suffices to check that (i) $\text{Mod}(T)$ is closed under arbitrary intersections; and that (ii) all a^* -shifts of models from $\text{Mod}(T)$ belong to $\text{Mod}(T)$.

“(i)”: Take an I -indexed system $\{M_i \in \text{Mod}(T) \mid i \in I\}$ of models of T . We show that $\bigcap_{i \in I} M_i$ is a model of T . Thus, we check that, for each $A \Rightarrow B \in T$, $\|A \Rightarrow B\|_{\bigcap_{i \in I} M_i} = 1$. Since each M_i is a model of T , for any $A \Rightarrow B \in T$, we have $\|A \Rightarrow B\|_{M_i} = 1$, i.e. $S(A, M_i)^* \leq S(B, M_i)$. Now, using the fact that $(\bigwedge_{j \in J} a_j)^* \leq \bigwedge_{j \in J} a_j^*$, we get

$$\begin{aligned} S(A, \bigcap_{i \in I} M_i)^* &= (\bigwedge_{i \in I} S(A, M_i))^* \leq \bigwedge_{i \in I} S(A, M_i)^* \leq \\ &\leq \bigwedge_{i \in I} S(B, M_i) = S(B, \bigcap_{i \in I} M_i), \end{aligned}$$

proving $\|A \Rightarrow B\|_{\bigcap_{i \in I} M_i} = 1$. Hence, $\bigcap_{i \in I} M_i \in \text{Mod}(T)$.

“(ii)”: Take $M \in \text{Mod}(T)$ and any truth degree $a \in L$. It suffices to show that $a^* \rightarrow M$ belongs to $\text{Mod}(T)$. Recall that $a^* \rightarrow M$ is an \mathbf{L} -set of attributes such that, for each attribute $y \in Y$, $(a^* \rightarrow M)(y) = a^* \rightarrow M(y)$. Since M is a model of T , for each $A \Rightarrow B \in T$, we have $\|A \Rightarrow B\|_M = 1$, i.e., $S(A, M)^* \leq S(B, M)$. Using (3), (4), and monotony of \rightarrow in the second argument, we further get that

$$\begin{aligned} S(A, a^* \rightarrow M)^* &= (a^* \rightarrow S(A, M))^* \leq \\ &\leq a^{**} \rightarrow S(A, M)^* = a^* \rightarrow S(A, M)^* \leq \\ &\leq a^* \rightarrow S(B, M) = S(B, a^* \rightarrow M). \end{aligned}$$

The latter inequality yields $\|A \Rightarrow B\|_{a^* \rightarrow M} = 1$ for each FAI $A \Rightarrow B \in T$, i.e. $a^* \rightarrow M \in \text{Mod}(T)$. ■

Remark 1: Degrees $\|\cdot\|_T$ were introduced as degrees of entailment from ordinary sets of FAIs. In fuzzy setting it is natural and often desirable to consider entailment from fuzzy collections T of formulas, where each $T(\varphi)$ (degree to which formula φ belongs to T) is interpreted as the degree to which “ φ is prescribed by T ”. In case of FAIs, one can proceed as follows. For a fuzzy set T of FAIs, $M \in \mathbf{L}^Y$ is called a model of T if $\|A \Rightarrow B\|_M \geq T(A \Rightarrow B)$. The collection of all models of a fuzzy set T of FAIs will be denoted by $\text{Mod}(T)$. Observe that if T is a crisp set, i.e. if for each $A \Rightarrow B$ we have $T(A \Rightarrow B) \in \{0, 1\}$, then due to the correspondence between crisp sets and ordinary sets, $\text{Mod}(T)$ coincides with (14). Now, for a fuzzy set T of FAIs, one can define degrees of entailment by $\|A \Rightarrow B\|_T = \bigwedge_{M \in \text{Mod}(T)} \|A \Rightarrow B\|_M$, see [6]. Hence, entailment from ordinary sets of FAIs (15) can be seen as a particular case of more general a concept, namely, the concept of entailment from fuzzy sets of FAIs. From this point of view it may be surprising that each collection $\text{Mod}(T)$ of models of a fuzzy set T of FAIs is also an \mathbf{L}^* -closure system (we postpone proof of this claim to a full version of the paper).

The following assertion shows that each \mathbf{L}^* -closure system is a system of all models of some set of FAIs.

Theorem 4: Let S be an \mathbf{L}^* -closure system in Y . Then there is a set T of FAIs over Y such that $S = \text{Mod}(T)$.

Proof: Put $T = \{A \Rightarrow C_S(A) \mid A \in \mathbf{L}^Y\}$. Let $M \in S$, i.e. we have $M = C_S(M)$. Then $S(A, M)^* \leq S(C_S(A), C_S(M)) = S(C_S(A), M)$ by (9), which gives $\|A \Rightarrow C_S(A)\|_M = 1$, i.e. M is a model of T . This proves $S \subseteq \text{Mod}(T)$. Conversely let $M \notin S$, i.e. $M \neq C_S(M)$. This is, we have $M \subset C_S(M)$ by (8), which further gives $S(C_S(M), M) \neq 1$. Hence, $\|M \Rightarrow C_S(M)\|_M = S(M, M)^* \rightarrow S(C_S(M), M) = 1^* \rightarrow S(C_S(M), M) = S(C_S(M), M) \neq 1$, i.e. $M \notin \text{Mod}(T)$, showing $\text{Mod}(T) \subseteq S$. ■

Remark 2: Theorem 3 and Theorem 4 now yield that \mathbf{L}^* -closure systems and models of sets of FAIs coincide. This has important consequences. For instance, any \mathbf{L}^* -closure system can be described by a set of FAIs.

From Theorem 2 and Theorem 3 we derive that $\text{Mod}(T)$ induces an \mathbf{L}^* -closure operator $C_{\text{Mod}(T)} : \mathbf{L}^Y \rightarrow \mathbf{L}^Y$ such that $M = C_{\text{Mod}(T)}(M)$ iff $M \in \text{Mod}(T)$. Directly from its definition, $C_{\text{Mod}(T)}(M)$ is the least model in $\text{Mod}(T)$ that contains M (in sense of the \mathbf{L} -set inclusion “ \subseteq ”). Thus, if $M' \in \text{Mod}(T)$ such that $M \subseteq M'$, then $C_{\text{Mod}(T)}(M) \subseteq M'$.

The definition of $C_{\text{Mod}(T)}$ provides no feasible way to compute the closure of a given M because its direct application suggests iteration over all models in $\text{Mod}(T)$. Also, $\text{Mod}(T)$ may be infinite provided \mathbf{L} is infinite. We now give a constructive description of $C_{\text{Mod}(T)}$ which can be used to compute the closures.

First, for any set T of FAIs and any fuzzy set $M \in \mathbf{L}^Y$ of attributes define a fuzzy set $M^T \in \mathbf{L}^Y$ of attributes by

$$M^T = M \cup \bigcup \{B \otimes S(A, M)^* \mid A \Rightarrow B \in T\}. \quad (16)$$

Using (16), for each nonnegative integer n we define a fuzzy set $M^{Tn} \in \mathbf{L}^Y$ of attributes by

$$M^{Tn} = \begin{cases} M & \text{for } n = 0 \\ (M^{T(n-1)})^T & \text{for } n \geq 1. \end{cases} \quad (17)$$

Finally, we define an operator $cl_T : \mathbf{L}^Y \rightarrow \mathbf{L}^Y$ by

$$cl_T(M) = \bigcup_{n=0}^{\infty} M^{Tn}. \quad (18)$$

Lemma 1: For each $M \in \text{Mod}(T)$, $cl_T(M) = M$.

Proof: Take $A \Rightarrow B \in T$. Since M is a model of T , we have $\|A \Rightarrow B\|_M = 1$, i.e. $S(A, M)^* \leq S(B, M)$. Hence, by adjointness, $S(A, M)^* \otimes B \subseteq M$. Thus, (16) yields that $M^T \subseteq M$. The converse inclusion is trivial. That is, (17) and (18) give that $M = M^{T0} = M^{T1} = M^{T2} = \dots$ and consequently $M = \bigcup_{n=0}^{\infty} M^{Tn} = cl_T(M)$. ■

Theorem 5: Let \mathbf{L} and Y be both finite, T be a set of FAIs over Y . Then cl_T is an \mathbf{L}^* -closure operator such that, for each $M \in \mathbf{L}^Y$, $C_{\text{Mod}(T)}(M) = cl_T(M)$.

Proof: Since $C_{\text{Mod}(T)}$ is an \mathbf{L}^* -closure operator, it suffices to check that $C_{\text{Mod}(T)}$ coincides with cl_T . This can be proved by showing that, for each \mathbf{L} -set $M \in \mathbf{L}^Y$ of attributes, $cl_T(M)$ is the least model in $\text{Mod}(T)$ which is greater than or equal to M . The claim that $M \subseteq cl_T(M)$ follows directly from the definition of cl_T , i.e. it remains to show that $cl_T(M)$ belongs to $\text{Mod}(T)$ and that $cl_T(M)$ is the least model containing M .

Observe that since both \mathbf{L} (structure of truth degrees) and Y (set of attributes) are finite, there exists a nonnegative integer k such that $cl_T(M) = M^{T_k}$, where M^{T_k} is defined by (17). Indeed, the finiteness of \mathbf{L} and Y gives that \mathbf{L}^Y is finite, i.e. in the chain $M^{T_0} \subseteq M^{T_1} \subseteq \dots \subseteq M^{T_k} \subseteq \dots$ formed of fuzzy sets of attributes defined by (17), there can be only finitely many proper inclusions “ \subset ”. Thus, k satisfying $cl_T(M) = M^{T_k}$ always exists and $M^{T_k} = M^{T_{k+1}} = M^{T_{k+2}} = \dots$.

We now prove that $cl_T(M) \in \text{Mod}(T)$. Take any $A \Rightarrow B \in T$. Since $cl_T(M) = M^{T_k} = M^{T_{k+1}}$, using (16), we get that

$$\begin{aligned} S(A, cl_T(M))^* \otimes B &= S(A, M^{T_k})^* \otimes B \subseteq (M^{T_k})^T = \\ &= M^{T_{k+1}} = M^{T_k} = cl_T(M). \end{aligned}$$

Thus, by adjointness, $S(A, cl_T(M))^* \otimes B \subseteq cl_T(M)$ yields $S(A, cl_T(M))^* \leq B(y) \rightarrow cl_T(M)(y)$ ($y \in Y$). Hence, it follows that $S(A, cl_T(M))^* \leq S(B, cl_T(M))$, i.e. we have that $\|A \Rightarrow B\|_{cl_T(M)} = 1$, showing $cl_T(M) \in \text{Mod}(T)$.

We have shown that $cl_T(M)$ is a model of T which contains M . Take any $M' \in \text{Mod}(T)$ such that $M \subseteq M'$. It remains to show that $cl_T(M) \subseteq M'$. But this is immediate because from $M \subseteq M'$ it follows that $cl_T(M) \subseteq cl_T(M')$ (operator cl_T is obviously monotone) and Lemma 1 yields $cl_T(M') = M'$. ■

Remark 3: Theorem 5 shows that for finite \mathbf{L} and Y , the operator $C_{\text{Mod}(T)}$ induced by the \mathbf{L}^* -closure system of models $\text{Mod}(T)$ coincides with the operator cl_T given by (18). This enables us to compute the closure $C_{\text{Mod}(T)}(M)$ of a given $M \in \mathbf{L}^Y$ (provided \mathbf{L} is finite) using the following algorithm.

INPUT: set of FAIs T , $M \in \mathbf{L}^Y$

OUTPUT: $C_{\text{Mod}(T)}(M)$

- 1) If, for each $A \Rightarrow B \in T$, $B \otimes S(A, M)^* \subseteq M$, then go to step 3); else go to step 2);
- 2) take any $A \Rightarrow B \in T$ such that $B \otimes S(A, M)^* \not\subseteq M$; set M to $M \cup (B \otimes S(A, M)^*)$; goto step 1);
- 3) return M .

Let us note that for certain hedges, the algorithm can be improved. For instance, if $*$ is globalization, see (5), then we can introduce an extended variant of the LinClosure algorithm which is in its basic form known from database systems, see [15]. It is then possible to compute the closure of M in a linear time depending on the size of T . We postpone details due to the full version of the paper.

The following assertion shows that least models can be used to characterize degrees of semantic entailment from sets of FAIs. This is an important property which generalizes the well-known property from classical setting. An interesting thing to stress is that even if we work with arbitrary degrees $\|\cdot\|_T$ of entailment (i.e., not only with 0 and 1), we are still able to characterize $\|\cdot\|_T$ using subsethood degrees (see (6) in Section II).

Theorem 6: For each set T of FAIs and $A \Rightarrow B$,

$$\|A \Rightarrow B\|_T = \|A \Rightarrow B\|_{C_{\text{Mod}(T)}(A)} = S(B, C_{\text{Mod}(T)}(A)).$$

Proof: Clearly, we have $\|A \Rightarrow B\|_T \leq \|A \Rightarrow B\|_{C_{\text{Mod}(T)}(A)}$ because $C_{\text{Mod}(T)}(A) \in \text{Mod}(T)$. Moreover, $C_{\text{Mod}(T)}$ satisfies

condition (8), i.e., we have

$$\begin{aligned} \|A \Rightarrow B\|_{C_{\text{Mod}(T)}(A)} &= \\ &= S(A, C_{\text{Mod}(T)}(A))^* \rightarrow S(B, C_{\text{Mod}(T)}(A)) = \\ &= 1 \rightarrow S(B, C_{\text{Mod}(T)}(A)) = S(B, C_{\text{Mod}(T)}(A)). \end{aligned}$$

Take any $M \in \text{Mod}(T)$. Applying (9), we get

$$\begin{aligned} S(B, C_{\text{Mod}(T)}(A)) \otimes S(A, M)^* &\leq \\ &\leq S(B, C_{\text{Mod}(T)}(A)) \otimes S(C_{\text{Mod}(T)}(A), C_{\text{Mod}(T)}(M)) \leq \\ &\leq S(B, C_{\text{Mod}(T)}(M)) = S(B, M). \end{aligned}$$

By adjointness, the following inequality

$$S(B, C_{\text{Mod}(T)}(A)) \leq S(A, M)^* \rightarrow S(B, M) = \|A \Rightarrow B\|_M$$

is true for each model $M \in \text{Mod}(T)$. Hence, it follows that $S(B, C_{\text{Mod}(T)}(A)) \leq \|A \Rightarrow B\|_T$. ■

Remark 4: Theorem 6 and the algorithm in Remark 3 allow us to determine $\|A \Rightarrow B\|_T$ using a single model of T which can be computed by the algorithm in Remark 3. Thus, for a finite \mathbf{L} , a set T of FAIs, and a FAI $A \Rightarrow B$, we first compute $C_{\text{Mod}(T)}(A)$ and then compute the subsethood degree $S(B, C_{\text{Mod}(T)}(A))$ which equals, by Theorem 6, to the degree $\|A \Rightarrow B\|_T$ to which $A \Rightarrow B$ is entailed by T .

V. NON-REDUNDANT BASES OF DATA TABLES WITH FUZZY ATTRIBUTES

In classical setting, by a non-redundancy of a set of formulas T is meant a property which says that no formula in T is entailed by the other ones from T . In our setting, we define non-redundancy as follows. A set T of FAIs is *non-redundant* if there is no $A \Rightarrow B \in T$ such that $\|A \Rightarrow B\|_{T - \{A \Rightarrow B\}} = 1$. In other words, T is non-redundant iff, for each $A \Rightarrow B \in T$, we have $\|A \Rightarrow B\|_{T - \{A \Rightarrow B\}} \neq 1$. As one can see, our concept of non-redundancy naturally extends the ordinary one in a fuzzy setting—if \mathbf{L} (our structure of truth degrees) is the two-element Boolean algebra (structure of truth degrees of the classical logic), we get exactly the usual concept.

Sets T_1 and T_2 of FAIs are *equivalent*, written $T_1 \equiv T_2$ if, for each $A \Rightarrow B$, we have $\|A \Rightarrow B\|_{T_1} = \|A \Rightarrow B\|_{T_2}$. Described verbally, two sets of FAIs are equivalent if they fully agree on degrees of semantic entailment, i.e. the sets are indistinguishable in sense of their semantic consequences. Hence, one can substitute two equivalent sets of FAIs and still have the same degrees of entailment. From this point of view it might be desirable to find, for a given T , a non-redundant set T' which is equivalent to T because T' would be a least set which encompasses the same information about semantic entailment as T . The following assertion will be used to remove redundant FAIs from a set of FAIs.

Lemma 2: Let T be a set of FAIs, $T' = T - \{A \Rightarrow B\}$. If $\|A \Rightarrow B\|_{T'} = 1$, then $T \equiv T'$.

Proof: Clearly, it suffices to show that $\text{Mod}(T) = \text{Mod}(T')$. By definition of T' , we get $\text{Mod}(T) \subseteq \text{Mod}(T')$. Now, take any $M \in \text{Mod}(T')$. Take $C \Rightarrow D \in T$. If $C \Rightarrow D \in T'$, then $\|C \Rightarrow D\|_M = 1$ because M is a model of T' . If $C \Rightarrow D \notin T'$, then $C \Rightarrow D$ is $A \Rightarrow B$. In that case, $\|A \Rightarrow B\|_{T'} = 1$,

i.e. $A \Rightarrow B$ is true in degree 1 in each model of T' and that includes M . Hence, $M \in \text{Mod}(T)$. ■

Remark 5: Using the observation in Lemma 2, one can find a non-redundant set of FAIs which is equivalent to the original one by removing of FAIs. In more detail, if \mathbf{L} is finite, one can check that $\|A \Rightarrow B\|_{T'} = 1$ using the procedure in Remark 4. Since we are interested in entailment in degree 1 (full entailment), the criterion for $\|A \Rightarrow B\|_{T'} = 1$ further simplifies. In fact, we have $\|A \Rightarrow B\|_{T'} = 1$ iff $B \subseteq C_{\text{Mod}(T')}(A)$, where $C_{\text{Mod}(T')}(A)$ can be computed by the procedure described in Remark 3. To sum up, we have the following algorithm:

INPUT: set of FAIs T

OUTPUT: non-redundant set of FAIs equivalent to T

- 1) If, for each $A \Rightarrow B \in T$, $B \not\subseteq C_{\text{Mod}(T - \{A \Rightarrow B\})}(A)$, then go to step 3); else go to step 2);
- 2) take $A \Rightarrow B \in T$ such that $B \subseteq C_{\text{Mod}(T - \{A \Rightarrow B\})}(A)$; remove $A \Rightarrow B$ from T ; go to step 1);
- 3) return T .

Note that a non-redundant set obtained by the algorithm is not given uniquely. The resulting non-redundant sets can vary depending in which order we remove FAIs from the original set. It can happen that we can arrive even to non-redundant bases with different sizes.

A non-redundant set T is interesting if T describes all dependencies which are true in a given data set. In that case, T is a least description of the dependencies in the data set, i.e. no proper subset of T can do this job. In the rest of this section we show a method of generating non-redundant sets of FAIs describing dependencies which are true in particular data sets. Our data sets will be so-called data tables with fuzzy attributes. A definition follows.

Let X and Y be finite sets of objects and attributes (Y plays the same role as before), respectively, $I: X \times Y \rightarrow L$ be a binary \mathbf{L} -relation between X and Y . $\mathcal{T} = \langle X, Y, I \rangle$ is called a *data table with fuzzy attributes*. \mathcal{T} represents a table which assigns to each object $x \in X$ (table row) and each attribute $y \in Y$ (table column) a truth degree $I(x, y) \in L$ (table entry) to which object x has attribute y . A degree $\|A \Rightarrow B\|_{\mathcal{T}}$ to which a fuzzy attribute implication $A \Rightarrow B$ (over Y) is *true* in $\mathcal{T} = \langle X, Y, I \rangle$ is defined by

$$\|A \Rightarrow B\|_{\mathcal{T}} = \bigwedge_{x \in X} \|A \Rightarrow B\|_{I_x}, \quad (19)$$

where each I_x ($x \in X$) denotes the row of \mathcal{T} corresponding to object x , i.e. $I_x \in \mathbf{L}^Y$ is an \mathbf{L} -set of attributes defined by $I_x(y) = I(x, y)$ ($y \in Y$). A set T of attribute implications is said to be *complete in data table* $\mathcal{T} = \langle X, Y, I \rangle$ if, for each $A \Rightarrow B$, $\|A \Rightarrow B\|_{\mathcal{T}} = \|A \Rightarrow B\|_{\mathcal{T}}$ (degree to which $A \Rightarrow B$ semantically follows from T is just the degree to which $A \Rightarrow B$ is true in $\mathcal{T} = \langle X, Y, I \rangle$). A set T of attribute implications is called a *non-redundant basis of* $\mathcal{T} = \langle X, Y, I \rangle$ if T is complete in \mathcal{T} and T is non-redundant.

Non-redundant bases of data tables can be computed either directly or indirectly. In [2], [5], [8] we have presented several algorithms for direct computation of non-redundant bases. That is, the algorithms take a data table with fuzzy attributes

as their input and produce its non-redundant basis directly. The algorithms from [2], [5], [8] have several limitations, we will comment on this later on. Another approach to generating of non-redundant bases of data tables is to start with a set which is complete (in a given \mathcal{T}) yet reasonably small and find a non-redundant basis by its reduction. This can be seen as an indirect procedure because we first need a complete set of FAIs which is further processed by the algorithm in Remark 5. We now describe particular complete sets which can be computed and further reduced. Because of the computational issues, in what follows we assume that \mathbf{L} is finite.

Given a data table with fuzzy attributes $\mathcal{T} = \langle X, Y, I \rangle$, for $A \in \mathbf{L}^X$ (fuzzy set of objects) and $B \in \mathbf{L}^Y$ (fuzzy set of attributes) we define $A^\uparrow \in \mathbf{L}^Y$ (fuzzy set of attributes) and $B^\downarrow \in \mathbf{L}^X$ (fuzzy set of objects) by $A^\uparrow(y) = \bigwedge_{x \in X} (A(x) \rightarrow I(x, y))$, $B^\downarrow(x) = \bigwedge_{y \in Y} (B(y) \rightarrow I(x, y))$. By definition, A^\uparrow is the collection of all attributes shared by all objects from A ; B^\downarrow is the collection of all objects sharing all attributes from B . Moreover, for each $B \in \mathbf{L}^Y$ we can consider $B^{\downarrow*\uparrow} \in \mathbf{L}^Y$ defined by $((B^\downarrow)^*)^\uparrow$.

In order to describe a set of FAIs which is complete in a given $\mathcal{T} = \langle X, Y, I \rangle$, we introduce a particular collection $\mathcal{P} \subseteq \mathbf{L}^Y$ of \mathbf{L} -set of attributes defined by:

- (i) $P \neq P^{\downarrow*\uparrow}$, and
- (ii) for each $Q \in \mathcal{P}$ such that $Q \subset P$: $Q^{\downarrow*\uparrow} \subseteq P$.

Since both \mathbf{L} and Y , the system $\mathcal{P} \subseteq \mathbf{L}^Y$ is uniquely given (one can check by induction, see also [12]). Now, the following assertion says that \mathcal{P} determines a set of FAIs which is complete in \mathcal{T} .

Theorem 7: *Let \mathcal{T} be a data table with fuzzy attributes, \mathcal{P} be defined as above. Then $T = \{P \Rightarrow P^{\downarrow*\uparrow} \mid P \in \mathcal{P}\}$ is complete in \mathcal{T} .*

Proof: We first show that $\text{Mod}(T)$ coincides with $\{M \in \mathbf{L}^Y \mid M = M^{\downarrow*\uparrow}\}$. Clearly, for each $M \in \mathbf{L}^Y$ such that $M = M^{\downarrow*\uparrow}$ we have $S(P, M)^* \leq S(P^{\downarrow*\uparrow}, M^{\downarrow*\uparrow})$ because $\downarrow*\uparrow$ satisfies (9). Thus, $S(P, M)^* \leq S(P^{\downarrow*\uparrow}, M)$, showing $\|P \Rightarrow P^{\downarrow*\uparrow}\|_M = 1$, i.e., we can see that M is a model of T . Conversely, let M be a model of T . For each $Q \in \mathcal{P}$, we have $\|Q \Rightarrow Q^{\downarrow*\uparrow}\|_M = 1$, i.e. if $Q \subset M$, then $Q^{\downarrow*\uparrow} \subseteq M$. Thus, if $M \neq M^{\downarrow*\uparrow}$, then $M \in \mathcal{P}$. On the other hand, $\|M \Rightarrow M^{\downarrow*\uparrow}\|_M \neq 1$, contradicting the fact that M is a model of T . Thus, we must have $M = M^{\downarrow*\uparrow}$. We have shown that models of T are exactly the fixed points of $\downarrow*\uparrow$. The latter observation is equivalent to that T is complete in \mathcal{T} , see [2] and [4, Theorem 4]. ■

Remark 6: If $*$ (the hedge) is globalization, see (5), then \mathcal{P} defined above is a so-called system of pseudo-intents, see [2], [6] for details. For the other hedges, \mathcal{P} need not be a system of pseudo-intents. If \mathcal{P} is a system of pseudo-intents, then T defined as in Theorem 7 is a non-redundant basis [2], [6]. In [2], [5], [8] we have shown several algorithms to compute systems of pseudo-intents (and thus non-redundant bases of data tables with fuzzy attributes, see Remark 6). The algorithm from [2] works with polynomial time delay but is suitable only for $*$ being globalization. Another approach presented in [5], [8] is based on graph-theoretical representation of systems of

hdg.	0*	0.25*	0.5*	0.75*	1*
* ₁	0	0	0	0	1
* ₂	0	0	0.5	0.5	1
* ₃	0	0.25	0.25	0.25	1
* ₄	0	0.25	0.5	0.5	1
* ₅	0	0.25	0.5	0.75	1

\mathcal{T}	y	z
x_1	0.5	0.5
x_2	1	0
x_3	0.25	0.75

Fig. 1. Hedges (left) and input data table (right)

$$\begin{aligned}
T_1 &= \{ \{y, 0.5/z\} \Rightarrow \{y, z\}, \{0.75/y\} \Rightarrow \{y\}, \\
&\quad \{0.5/y, 0.75/z\} \Rightarrow \{y, z\}, \{0.25/y, z\} \Rightarrow \{y, z\}, \\
&\quad \{0.25/y, 0.25/z\} \Rightarrow \{0.25/y, 0.5/z\}, \{ \} \Rightarrow \{0.25/y\} \}, \\
T_2 &= \{ \{0.5/y, 0.75/z\} \Rightarrow \{0.75/y, z\}, \\
&\quad \{0.25/y, z\} \Rightarrow \{0.75/y, z\}, \{ \} \Rightarrow \{0.25/y\} \}, \\
T_3 &= \{ \{y, 0.5/z\} \Rightarrow \{y, 0.75/z\}, \{0.75/y\} \Rightarrow \{y\}, \\
&\quad \{0.5/y, 0.75/z\} \Rightarrow \{y, 0.75/z\}, \{0.25/y, z\} \Rightarrow \{y, z\}, \\
&\quad \{0.25/y, 0.25/z\} \Rightarrow \{0.25/y, 0.5/z\} \}, \\
T_4 &= \{ \{0.5/y, 0.75/z\} \Rightarrow \{0.75/y, 0.75/z\}, \{0.25/y, z\} \Rightarrow \{0.75/y, z\}, \\
&\quad \{0.25/y, 0.25/z\} \Rightarrow \{0.25/y, 0.5/z\}, \{ \} \Rightarrow \{0.25/y\} \}, \\
T_5 &= \{ \{0.25/y, z\} \Rightarrow \{0.5/y, z\}, \{ \} \Rightarrow \{0.25/y\} \}.
\end{aligned}$$

Fig. 2. Non-redundant bases of data tables with fuzzy attributes

pseudo-intents. This procedure works with any hedge, however it can be applied only to small data tables.

Remark 7: For a data table \mathcal{T} with fuzzy attributes we can compute its non-redundant basis as follows. We first determine the system \mathcal{P} which can be done with a polynomial time delay (the idea is analogous to that presented in [2] see also [9], [11]). Then, due to Theorem 7, we have a set $T = \{P \Rightarrow P \downarrow^{*\uparrow} \mid P \in \mathcal{P}\}$ which is complete in \mathcal{T} . We then use the algorithm described in Remark 5 to get a subset $T' \subseteq T$ which is non-redundant and equivalent to T . Hence, T' is a non-redundant basis of the input data table \mathcal{T} .

Example 1: Let \mathbf{L} be the five element Łukasiewicz algebra with $L = \{0, 0.25, 0.5, 0.75, 1\} \subseteq [0, 1]$, \wedge and \vee being minimum and maximum, respectively. \mathbf{L} can be endowed with five hedges which are displayed in Fig. 1 (left). Each hedge is represented by a row in the table Fig. 1 (left). Hedge $*_1$ is the globalization on L , see (5); $*_5$ is the identity on L ; and $*_2, *_3, *_4$ represent three intermediate hedges. Consider a data table with fuzzy attributes $\mathcal{T} = \langle X, Y, I \rangle$ given by Fig. 1 (right). That is, $X = \{x_1, x_2, x_3\}$ (set of objects), $Y = \{y, z\}$ (set of attributes), and the table entries in Fig. 1 (right) define \mathbf{L} -relation $I: X \times Y \rightarrow L$. For each hedge $*_1, \dots, *_5$ we can compute a non-redundant basis of \mathcal{T} using the algorithm described in Remark 7. Some of the non-redundant bases we can get this way are depicted in Fig. 2 (T_i is a non-redundant basis of \mathcal{T} which was computed using hedge $*_i$). This example demonstrates that sizes of bases may vary and are dependent on the chosen hedge.

VI. CONCLUSIONS AND FURTHER RESEARCH

This paper shows that fuzzy closure systems with hedges can be seen as collections of models of fuzzy attributes

implications (FAIs) and *vice versa*. Thus, we have established a characterization of sets of models (of FAIs) by their closure properties. Furthermore, we showed applications of the closure properties. First, we used the concept of a least model to characterize degrees of entailment. Second, we showed that this essential property can be used to find non-redundant bases of data tables with fuzzy attributes in an efficient way.

Future research will focus on further issues related to model-theoretical properties of FAIs and non-redundant bases including the following problems:

- Given two hedges $*_1$ and $*_2$ on a complete residuated lattice, is there any relationship between non-redundant bases of data tables generated using $*_1$ and $*_2$? For instance, can we say anything about the sizes of bases?
- A non-redundant basis need not be minimal in terms of its size. Also, a non-redundant basis can contain extraneous attributes (attributes that can be removed without losing the completeness). It is thus interesting to explore the possibility to find bases with minimal size and without any extraneous attributes.

VII. ACKNOWLEDGMENT

Supported by grant No. 1ET101370417 of GA AV ČR, by grant No. 201/05/0079 of the Czech Science Foundation, and by institutional support, research plan MSM 6198959214.

REFERENCES

- [1] Bělohlávek R.: *Fuzzy Relational Systems: Foundations and Principles*. Kluwer, Academic/Plenum Publishers, New York, 2002.
- [2] Bělohlávek R., Chlupová M., Vychodil V.: Implications from data with fuzzy attributes. In: Proc. AISTA 2004, 5 pages. ISBN 2-9599776-8-8.
- [3] Bělohlávek R., Funioková T., Vychodil V.: Fuzzy closure operators with truth stressers. *Logic Journal of IGPL* **13**(5)(2005), 503–513.
- [4] Bělohlávek R., Vychodil V.: Fuzzy attribute logic: attribute implications, their validity, entailment, and non-redundant basis. In: Proc. IFSA 2005, vol. I, pp. 622–627, ISBN 7-302-11377-7.
- [5] Bělohlávek R., Vychodil V.: Fuzzy attribute implications: computing non-redundant bases using maximal independent sets. In: S. Zhang and R. Jarvis (Eds.): *AI 2005*, LNAI **3809**, pp. 1126–1129, 2005.
- [6] Bělohlávek R., Vychodil V.: Attribute implications in a fuzzy setting. In: Missaoui R., Schmid J. (Eds.): *ICFCA 2006*, LNAI **3874**, pp. 45–60, 2006.
- [7] R. Bělohlávek, V. Vychodil. Data tables with similarity relations: functional dependencies, complete rules and non-redundant bases. In: Lee M. L., Tan K. L., Wuwongse V. (Eds.): *DASFAA 2006*, LNCS **3882**, pp. 644–658, 2006.
- [8] Bělohlávek R., Vychodil V.: Computing non-redundant bases of if-then rules from data tables with graded attributes (to appear in Proc. IEEE-GrC 2006).
- [9] Carpineto C., Romano G.: *Concept Data Analysis. Theory and Applications*. J. Wiley, 2004.
- [10] Chang C. C., Keisler H. J.: *Model Theory*. North-Holland, Amsterdam, 1973.
- [11] Ganter B., Wille R.: *Formal Concept Analysis. Mathematical Foundations*. Springer, Berlin, 1999.
- [12] Guigues J.-L., Duquenne V.: Familles minimales d'implications informatives résultant d'un tableau de données binaires. *Math. Sci. Humaines* **95**(1986), 5–18.
- [13] Hájek P.: *Metamathematics of Fuzzy Logic*. Kluwer, Dordrecht, 1998.
- [14] Hájek P.: On very true. *Fuzzy Sets and Systems* **124**(2001), 329–333.
- [15] Maier D.: *The Theory of Relational Databases*. Computer Science Press, Rockville, 1983.
- [16] Pollandt S.: *Fuzzy Begriffe*. Springer-Verlag, Berlin/Heidelberg, 1997.
- [17] Takeuti G., Titani S.: Globalization of intuitionistic set theory. *Annals of Pure and Applied Logic* **33**(1987), 195–211.