Reducing the size of if-then rules generated from data tables with graded attributes*

Radim Belohlavek and Vilem Vychodil
Department of Computer Science, Palacky University, Olomouc
Tomkova 40, CZ-779 00 Olomouc, Czech Republic
{radim.belohlavek, vilem.vychodil}@upol.cz

Abstract

Presented is a method for reducing the size of if-then rules generated from data tables with graded attributes. Data tables with graded attributes represent data structures describing graded properties of objects. We show a method for generating of particular reduced and minimal set of if-then rules which describe all dependencies which are valid in a given data table with graded attributes. We present theoretical insight, algorithms, and illustrative examples.

1. Introduction

This paper contributes to the filed of extraction of interesting patterns from input data. In our situation, we are dealing with input data which take form of a data table with rows denoting objects, columns denoting attributes, and table entries indicating to what degrees objects (given by rows) have corresponding attributes (given by columns). Data tables of this form are called *data tables with graded attributes* because they express information about objects (table rows) and their graded attributes (attributes are given by columns and the grades are given by the corresponding table entries). In this paper we are interested in extraction of rules describing dependencies between attributes in such data tables. In particular, we will be interested in rules $A \Rightarrow B$, saying "if each object (from the table) has all attributes from A, then it has all attributes from B". Rules of this form are called attribute implications in formal concept analysis [10, 11], association rules in data mining [20], and functional dependencies [17] in database terminology. In all the fields just mentioned, there have been efforts to find descriptions of all if-then rules which are valid in a particular data table. The set of all if-then rules valid in a given data table is usually large. The problem of finding a minimal description of all valid if-then rules is thus an appealing one. In [2, 4] and [6, 7], we have described so-called non-redundant bases of data tables with graded attributes and showed various methods for their computation. Non-redundant bases of data tables graded attributes are particular sets of if-then rules which fully describe all if-then rules which are valid in the data table. Furthermore, in some important cases we have proved that the computed bases have minimal size in terms of the number of if-then rules contained in the bases [4]. In this paper, which is a continuation of [2, 4], we show that size of if-then rules contained in non-redundant bases can be reduced by removing extraneous attributes. In addition to that, we present an efficient algorithm which allows us to compute, given a data table with graded attributes, a minimal non-redundant basis with reduced if-then rules. In Section 2 and Section 3, we present preliminaries of fuzzy logic (our formal framework) and fuzzy attribute implications (if-then rules we focus on). Section 4 contains the main results which are further illustrated by examples.

2. Preliminaries

In this section we introduce basic notions of fuzzy logic and fuzzy set theory, for details see [1, 12, 14, 16]. Our basic structures of truth degrees will be so-called complete residuated lattices with hedges. A complete residuated lattice with hedge 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 $a, b \in L$, (i) $1^* = 1$, (ii) $a^* \leq a$, (iii) $(a \rightarrow b)^* \leq a^* \rightarrow b^*$, and (iv) $a^{**} = a^*$. Each $a \in L$ is 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" and properties of hedges have natural interpretations, see [14, 15]. A common choice of $\mathbf{L}$ is a structure with $L = [0, 1]$ (unit interval), $\wedge$ and $\vee$ being minimum and maximum, $\otimes$ being

*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.

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 (e.g., finite Łukasiewicz and Gödel chains). Two boundary cases of hedges are (i) identity, i.e. $a^* = a$ ($a \in L$); (ii) so-called globalization [19]: $1^* = 1$, $a^* = 0$ ($a < 1$). A special case of a complete residuated lattice with hedge is the two-element Boolean algebra $\langle \{0,1\}, \wedge, \vee, \otimes, \rightarrow, ^*, 0, 1 \rangle$, denoted by $\mathbf{2}$ (structure of truth degrees of classical logic). For each $\mathbf{L}$ we consider a derived truth function $\ominus$ defined by

$$a \ominus b = a \otimes ((a \rightarrow b)^* \rightarrow 0). \tag{1}$$

If * is globalization, we get

$$a \ominus b = \begin{cases} 0 & \text{if } a \leq b, \\ a & \text{else.} \end{cases} \tag{2}$$

Note that the derived truth function $\ominus$ can be seen as a particular subtraction of truth degrees. This is apparent especially in case of globalization, see (2).

Until otherwise mentioned, we assume that $\mathbf{L}$ denotes a complete residuated lattice (with hedge *) which serves as a structure of truth degrees. Using $\mathbf{L}$, we define the following notions. An $\mathbf{L}$-set (a fuzzy set) A in universe U is a mapping $A\colon U \rightarrow L$, $A(u)$ being interpreted as "the degree to which u belongs to A". If U is a finite universe $U = \{u_1, \ldots, u_n\}$ then an $\mathbf{L}$-set A in U can be denoted by $A = \{{}^{a_1}/u_1, \ldots, {}^{a_n}/u_n\}$ meaning that $A(u_i)$ equals a_i ($i = 1, \ldots, n$). For brevity, we introduce the following convention: we write $\{\ldots, u, \ldots\}$ instead of $\{\ldots, {}^1/u, \ldots\}$, 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. Denote by $|A|$ the cardinality of the support set of A, i.e. $|A| = |\{u \in U \,|\, A(u) > 0\}|$. The operations with $\mathbf{L}$-sets are defined componentwise. For instance, union of $\mathbf{L}$-sets $A, B \in \mathbf{L}^U$ is an $\mathbf{L}$-set $A \cup B$ in U such that $(A \cup B)(u) = A(u) \vee B(u)$ ($u \in U$); for * being globalization, we get

$$(A \ominus B)(u) = \begin{cases} 0 & \text{if } A(u) \leq B(u), \\ A(u) & \text{else.} \end{cases} \tag{3}$$

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 $S(A,B) \in L$ by $S(A,B) = \bigwedge_{u \in U} \big(A(u) \rightarrow B(u)\big)$. $S(A,B)$ is called a subsethood degree of A in B and it generalizes the classical subsethood relation $\subseteq$ in a fuzzy setting. In particular, we write $A \subseteq B$ iff $S(A,B) = 1$; and $A \subset B$ iff $S(A,B) = 1$ and $A \neq B$. As a consequence, we have $A \subseteq B$ iff $A(u) \leq B(u)$ for each $u \in U$.

3. Fuzzy attribute implications

Let Y denote a *finite set of attributes.* Each $\mathbf{L}$-set $M \in \mathbf{L}^Y$ of attributes can be seen as a set of graded attributes because M prescribes, for each attribute $y \in Y$, a degree $M(y) \in L$. 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 [4]. The intuitive meaning we wish to give to $A \Rightarrow B$ is: "if it is (very) true that an object has all (graded) attributes from A, then it has also all (graded) 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

$$||A \Rightarrow B||_M = S(A,M)^* \rightarrow S(B,M). \tag{4}$$

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 * present in (4) serves as a modifier of interpretation of $A \Rightarrow B$ and plays an important technical role, see [2, 4, 5] for details. See also [18] for a related approach.

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 $\mathrm{Mod}(T)$. The set $\mathrm{Mod}(T)$ of all models of T form a particular fuzzy closure system in Y, see [8] for details. Thus, for each $\mathbf{L}$-set $M \in \mathbf{L}^Y$ we can consider its closure $[M]_T$ in $\mathrm{Mod}(T)$ which is then the least model of T containing M. A *degree* $||A \Rightarrow B||_T$ *to which* $A \Rightarrow B$ *semantically follows from* T is defined by

$$||A \Rightarrow B||_T = \textstyle\bigwedge_{M \in \mathrm{Mod}(T)} ||A \Rightarrow B||_M. \tag{5}$$

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 $||\cdots||_T$ defined by (5) represent degrees of semantic entailment from T. Sets T and T' of FAIs are called *semantically equivalent,* written $T \equiv T'$, if $||A \Rightarrow B||_T = ||A \Rightarrow B||_{T'}$ is true for each $A \Rightarrow B$. Let us note that degrees $||\cdots||_T$ can also be fully described via the (syntactic) concept of a *provability degree*, see [5] for a survey.

Fuzzy attribute implications are meant to be interpreted in data tables with fuzzy attributes. A *data table with fuzzy attributes* 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), 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 "object x has attribute y". $\langle X, Y, I \rangle$ can be thought of 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 seen as a set I_x of graded attributes (a fuzzy set of attributes) to which an attribute $y \in Y$ belongs to a degree $I_x(y) = I(x,y)$. Furthermore, a *degree* $||A \Rightarrow B||_{\langle X,Y,I \rangle}$ *to which* $A \Rightarrow B$ *is true in data table* $\langle X,Y,I \rangle$ is defined by

$$||A \Rightarrow B||_{\langle X,Y,I \rangle} = \bigwedge_{x \in X} ||A \Rightarrow B||_{I_x}. \tag{6}$$

By definition, $||A \Rightarrow B||_{\langle X,Y,I \rangle}$ is a degree to which "$A \Rightarrow B$ is true in each row of table $\langle X,Y,I \rangle$", 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". T is called *complete in* $\langle X,Y,I \rangle$ if $||A \Rightarrow B||_T = ||A \Rightarrow B||_{\langle X,Y,I \rangle}$, i.e. if, for each $A \Rightarrow B$, a degree to which T entails $A \Rightarrow B$ coincides with a degree to which $A \Rightarrow B$ is true in $\langle X,Y,I \rangle$. If T is complete in $\langle X,Y,I \rangle$ and no proper subset of T is complete in $\langle X,Y,I \rangle$, then T is called a *non-redundant basis of* $\langle X,Y,I \rangle$. If T is complete in $\langle X,Y,I \rangle$ and for each set T' which is complete in $\langle X,Y,I \rangle$ we have $|T| \leq |T'|$, then T is called a *minimal basis of* $\langle X,Y,I \rangle$. It is easily seen that each minimal basis of $\langle X,Y,I \rangle$ is also a non-redundant basis of $\langle X,Y,I \rangle$ (but not *vice versa* in general).

We need the following notions of formal concept analysis of data tables with fuzzy attributes [3, 5]. Given a data table $\langle X,Y,I \rangle$, for $A \in \mathbf{L}^X$, $B \in \mathbf{L}^Y$ we define $A^{\uparrow} \in \mathbf{L}^Y$ (fuzzy set of attributes) and $B^{\downarrow} \in \mathbf{L}^X$ (fuzzy set of objects) by

$$A^{\uparrow}(y) = \bigwedge_{x \in X}(A(x)^* \rightarrow I(x,y)), \tag{7}$$

$$B^{\downarrow}(x) = \bigwedge_{y \in Y}(B(y) \rightarrow I(x,y)). \tag{8}$$

Operators ${}^{\downarrow}, {}^{\uparrow}$ form so-called Galois connection with hedge, see [3]. The set of all fixed points of ${}^{\downarrow}, {}^{\uparrow}$ (so-called fuzzy concepts, i.e. particular conceptual clusters present in the input data) hierarchically ordered by the subconcept-superconcept relation is called a *fuzzy concept lattice with hedge*, see [3, 5]. Moreover, for each $M \in \mathbf{L}^Y$, put

$$M^{T^*} = M \cup \bigcup\{B \,|\, A \Rightarrow B \in T \text{ and } A \subset M\}, \tag{9}$$

$$M^{T^*_n} = \begin{cases} M & \text{for } n = 0 \\ (M^{T^*_{n-1}})^{T^*} & \text{for } n \geq 1, \end{cases} \tag{10}$$

and define an operator $cl_{T^*} \colon \mathbf{L}^Y \rightarrow \mathbf{L}^Y$ by

$$cl_{T^*}(M) = \bigcup_{n=0}^{\infty} M^{T^*_n}. \tag{11}$$

Let $\langle X,Y,I \rangle$ be a data table with fuzzy attributes. A system of fuzzy sets of attributes $\mathcal{P} \subseteq \mathbf{L}^Y$ is called a *system of pseudo-intents of* $\langle X,Y,I \rangle$ if, for each $P \in \mathbf{L}^Y$, we have:

$$P \in \mathcal{P} \text{ iff } P \neq P^{\downarrow\uparrow} \text{ and for each } Q \in \mathcal{P} \text{ such that } Q \neq P \colon ||Q \Rightarrow Q^{\downarrow\uparrow}||_P = 1. \tag{12}$$

If $\mathcal{P}$ is a system of pseudo-intents of $\langle X,Y,I \rangle$, then each $P \in \mathcal{P}$ is called a *pseudo-intent of* $\langle X,Y,I \rangle$. The following assertion (see [2, 4, 5]) shows important properties of systems of pseudo-intents.

Theorem 1. *Let* $\mathbf{L}$ *be a finite residuated lattice with globalization,* $\langle X,Y,I \rangle$ *be a data table with fuzzy attributes. Then*

(i) *there is a unique* $\mathcal{P}$ *satisfying* (12)*;*

(ii) $T = \{P \Rightarrow P^{\downarrow\uparrow} \,|\, P \in \mathcal{P}\}$ *is a minimal basis of* $\langle X,Y,I \rangle$*;*

(iii) *for* $T = \{P \Rightarrow P^{\downarrow\uparrow} \,|\, P \in \mathcal{P}\}$*, the operator* cl_{T^*} *defined by* (11) *is a closure operator and* $\mathcal{P} \cup \{M^{\downarrow\uparrow} \,|\, M \in \mathbf{L}^Y\}$ *is a set of all its fixed points.* □

Remark 2. From Theorem 1 it follows that for * being globalization a minimal basis (i.e., a particular non-redundant basis) of $\langle X,Y,I \rangle$ can be computed using fixed points of the closure operator cl_{T^*}. Namely, the minimal basis is determined by fuzzy sets $P \in \mathbf{L}^Y$ of attributes such that $P = cl_{T^*}(P)$ and $P \neq P^{\downarrow\uparrow}$. In [2] we have presented an algorithm which enables us to compute the minimal basis with polynomial time delay. The algorithm uses graded version of Ganter's NEXTCLOSURE algorithm [4, 5], cf. also [10, 11, 13]. In the following section we present an extension of this algorithm which allows us to compute minimal bases with reduced FAIs.

4. Reduction of attributes in FAIs

Throughout this section we assume that $\mathbf{L}$ is a finite residuated lattice with globalization. Recall that due to Theorem 1, a non-redundant basis of $\langle X,Y,I \rangle$ given by the theorem is a minimal one, i.e. one cannot find a set of fuzzy attribute implications which is complete in $\langle X,Y,I \rangle$ and consists of a strictly smaller number of FAIs. Still, it can happen that FAIs in the basis contain extraneous attributes, i.e. attributes that can be removed from particular FAIs without losing completeness of the basis. We may try to reduce the left-hand (right-hand) sides of FAIs in the basis, thus making the basis more readable for users.

In what follows we introduce a tractable concept of a Supp-reduced set of FAIs which will be used to compute reduced bases of data tables. Let us first demonstrate the problem by an example. Consider a set T of FAIs:

$$T = \{\{{}^{0.4\!}/b, {}^{0.8\!}/c, {}^{0.3\!}/d\} \Rightarrow B, \{{}^{0.2\!}/b\} \Rightarrow \{{}^{0.4\!}/b, {}^{0.8\!}/c\}\},$$

where B is some fuzzy set of attributes. Looking at set T, we see that the right-hand side (consequent) of the second fuzzy attribute implication is contained in the left-hand side (antecedent) of the first one, i.e. we have $\{{}^{0.4\!}/b, {}^{0.8\!}/c\} \subseteq \{{}^{0.4\!}/b, {}^{0.8\!}/c, {}^{0.3\!}/d\}$. Now, if we remove $\{{}^{0.4\!}/b, {}^{0.8\!}/c\}$ from the antecedent of the first FAI and we replace it by antecedent of the second FAI, we get

$$T' = \{\{{}^{0.2\!}/b, {}^{0.3\!}/d\} \Rightarrow B, \{{}^{0.2\!}/b\} \Rightarrow \{{}^{0.6\!}/b, {}^{0.8\!}/c\}\},$$

which is equivalent to the original set T but the first FAI in

T' has shorter antecedent than the corresponding FAI in T. Thus, we have constructed a new set T' of FAIs from T by reducing of one FAI in T by another FAI. This kind of reducibility of FAIs which leads to sets of FAIs with shorter antecedents is defined by the following

Definition 3. Let T be a set of FAIs, $A \Rightarrow B$ and $C \Rightarrow D$ be two distinct FAIs from T. If

(i) $D \subseteq A$,

(ii) $C \subseteq [A]_{T-\{A \Rightarrow B\}}$, and

(iii) $|(A \ominus D) \cup C| < |A|$,

then $A \Rightarrow B$ *is called* Supp-*reducible by* $C \Rightarrow D$ *in* T. ■

Remark 4. Note that Definition 3 (i) says that consequent of one FAI must be contained in antecedent of another; (ii) is a technical condition saying that C is a subset of the least model of $T - \{A \Rightarrow B\}$ containing A; this property ensures semantic equivalence of sets of FAIs (we will see later); (iii) ensures that the reduction will lead to a shorter FAI.

We introduce reduced sets of FAIs based on the previous concept of a Supp-reducibility as follows:

Definition 5. Let T be a set of FAIs. If there are no FAIs $A \Rightarrow B, C \Rightarrow D \in T$ such that $A \Rightarrow B$ is Supp-reducible by $C \Rightarrow D$ in T, then T is called a Supp-*reduced set of FAIs.* ■

The following assertion says that each FAI which is Supp-reducible by another FAI in T can be equivalently replaced by a FAI with shorter left-hand side.

Lemma 6. *Let T be a set of FAIs. Let $A \Rightarrow B, C \Rightarrow D \in T$ such that $A \Rightarrow B$ is* Supp-*reducible by $C \Rightarrow D$ in T. Then we have* $\mathrm{Mod}(T) = \mathrm{Mod}(T')$*, where T' is a set of FAIs defined by $T' = T - \{A \Rightarrow B\} \cup \{(A \ominus D) \cup C \Rightarrow B\}$.*

Proof. Let $M \in \mathrm{Mod}(T - \{A \Rightarrow B\})$. It suffices to check that $||A \Rightarrow B||_M = 1$ iff $||(A \ominus D) \cup C \Rightarrow B||_M = 1$.

"$\Rightarrow$": Let $||A \Rightarrow B||_M = 1$. Assume that $(A \ominus D) \cup C \subseteq M$. Since $C \Rightarrow D \in T - \{A \Rightarrow B\}$ and $C \subseteq M$, we get $D \subseteq M$. That is, $A \subseteq (A \ominus D) \cup D \subseteq M$. From the latter inclusion and $||A \Rightarrow B||_M = 1$ it follows that $B \subseteq M$. Hence, we have shown $B \subseteq M$ assuming $(A \ominus D) \cup C \subseteq M$ which proves that for M we have $||(A \ominus D) \cup C \Rightarrow B||_M = 1$.

"$\Leftarrow$": Let $||(A \ominus D) \cup C \Rightarrow B||_M = 1$, $A \subseteq M$. We clearly have $A \ominus D \subseteq M$. Since M is a model of $T - \{A \Rightarrow B\}$ satisfying $A \subseteq M$, Definition 3 (ii) yields $C \subseteq [A]_{T-\{A \Rightarrow B\}} \subseteq M$. Altogether, we get that $(A \ominus D) \cup C \subseteq M$. Thus, the assumption of $||(A \ominus D) \cup C \Rightarrow B||_M = 1$ yields $B \subseteq M$, which proves $||A \Rightarrow B||_M = 1$. □

Furthermore, we can use Lemma 6 repeatedly to find, for any set T of FAIs, a Supp-reduced set of FAIs as it is shown by the following

Theorem 7. *For each set T of FAIs there is a* Supp-*reduced set T' of FAIs such that $T \equiv T'$ and $|T'| \leq |T|$.*

Algorithm 9 (Generate Minimal Supp-reduced basis).

Input: data table $\langle X, Y, I \rangle$ with fuzzy attributes

Output: Supp-reduced minimal basis of $\langle X, Y, I \rangle$

```
if ∅↓↑ ≠ ∅:
   T := {∅ ⇒ ∅↓↑}
else:
   T := ∅
P := ∅, T_P := T
while P ≠ Y:
   P := NextClosure(P, T_P)
   P_c := P↓↑
   if P ≠ P_c:
      add P ⇒ P_c to T_P
      repeat:
         REDUCED := true
         for each Q ⇒ R ∈ T:
            if R ⊆ P and |(P ⊖ R) ∪ Q| < |P|:
               P := (P ⊖ R) ∪ Q
               REDUCED := false
      until REDUCED = true
      add P ⇒ P_c to T
return T
```

Proof. Put $T' = T$. If there are $A \Rightarrow B, C \Rightarrow D \in T'$ such that $A \Rightarrow B$ is Supp-reducible by $C \Rightarrow D$ in T', remove $A \Rightarrow B$ from T' and add $(A \ominus D) \cup C \Rightarrow B$ to T'. Repeat the process until T' is Supp-reduced. Since Y and $\mathbf{L}$ are finite, the process halts after finitely many steps. Obviously, for T and T' we have $|T'| \leq |T|$. In order to prove $T \equiv T'$ is suffices to show $\mathrm{Mod}(T) = \mathrm{Mod}(T')$. Claim $\mathrm{Mod}(T) = \mathrm{Mod}(T')$ follows by induction and repeated use of Lemma 6. □

Remark 8. Theorem 7 yields that given $\langle X, Y, I \rangle$, we can get a minimal non-redundant Supp-reduced basis of $\langle X, Y, I \rangle$: we first determine the minimal basis of $\langle X, Y, I \rangle$ using the algorithm from [2] and then we use procedure shown in Theorem 7 to get semantically equivalent Supp-reduced set T' of FAIs. Since T is a minimal basis and T and T' are semantically equivalent, we get that $|T| = |T'|$. In words, T' is complete in $\langle X, Y, I \rangle$ and has the same size as a minimal basis of $\langle X, Y, I \rangle$. Thus, T' is also a minimal basis of $\langle X, Y, I \rangle$ which is in addition Supp-reduced. Note, however, that this direct application of procedure given by proof of Theorem 7 is inefficient.

Interestingly, there is a way to obtain a minimal Supp-reduced basis of $\langle X, Y, I \rangle$ which more efficient than that described in Remark 8. The algorithm for computing of minimal bases presented in [2] uses graded extension of Ganter's NextClosure algorithm, see [10, 11]. Using NextClosure, one can list in lexical order all pseudo-intents which determine the basis. The listing of pseudo-intents in lexical order can be used to speed up the process of reduction. In what follows we present an extension of the algorithm from [2] in which the reduction of sizes of antecedents of if-then rules will be performed during the computation of the basis itself.

The procedure is described in Algorithm 9. In the algorithm we denote by NEXTCLOSURE$(P, T_{\mathcal{P}})$ the lexically smallest fixed point of $cl_{T^*_{\mathcal{P}}}$ which is a successor of P.

Proof of the correctness of Algorithm 9. Due to the limited scope of this paper, we present only a sketch of the proof. The detailed proof is postponed to the full version of this paper. Correctness of the algorithm can be proved by induction on the number of iterations of the while-loop (see lines 6–20). Observe that FAIs are added to T only at line 20. To avoid confusion, we let $P_{\mathcal{P}}$ denote P computed at line 7 (which may differ from P at line 20 because P may be modified at line 17). In particular, if we consider $Q \Rightarrow R \in T$ then by $Q_{\mathcal{P}}$ we mean the value of P which has been computed at line 7 during the loop which resulted in adding $Q \Rightarrow R$ to T. In the beginning of each iteration, which starts by checking the condition at line 6, we can inductively assume that (i) T is Supp-reduced; (ii) for each $Q \Rightarrow R \in T$, $Q_{\mathcal{P}}$ is a pseudo-intent such that

$$Q \subseteq Q_{\mathcal{P}} \subset Q_{\mathcal{P}}^{\downarrow\uparrow} = R;$$

and prove that if $P_{\mathcal{P}} \neq P_c$ then (a) $P_{\mathcal{P}}$ is a pseudo-intent and that (b) by adding $P \Rightarrow P_c$ to T (see line 20), we do not violate (i) and (ii). After that, we get that T is Supp-reduced and due to repeated use of Lemma 6, T has the same models as $T_{\mathcal{P}} = \{P \Rightarrow P^{\downarrow\uparrow} \mid P \text{ is a pseudo-intent}\}$. This proves that T and $T_{\mathcal{P}}$ are semantically equivalent. Thus, $|T| \leq |T_{\mathcal{P}}|$ and T is complete in $\langle X, Y, I\rangle$. Since $T_{\mathcal{P}}$ is a minimal basis of $\langle X, Y, I\rangle$, we get $|T| = |T_{\mathcal{P}}|$, i.e. T is also a minimal basis of the input data table $\langle X, Y, I\rangle$. □

Remark 10. (1) The efficiency of Algorithm 9 partly depends on our ability to compute fixed points of closure operator cl_{T^*} given by (11). This can be done with a graded extension of LINCLOSURE algorithm, see [9], that has linear time complexity with respect to the size of T.

(2) Notice that during the computation of Algorithm 9 we do not check the time-consuming condition Definition 3 (ii). Validity of the condition follows from the fact that we process all pseudo-intents in lexical order. This will be explained in more detail in a full version of this paper.

(3) Other notions of reduction (i.e., other than Supp-reducibility) might be possible, see e.g. [17], and this still needs to be explored. The benefit of the presented approach is that the algorithm determines the reduced basis in a single step. The usual methods (for binary data) proceed in two steps (generating of bases and then reduction).

(4) Algorithm 9 can also be used for ordinary (binary) data tables, i.e. for data table whose entries are not arbitrary degrees but only 0 (an object does not have an attribute) and 1 (an object has an attribute). In detail, for $\mathbf{L} = \mathbf{2}$, Algorithm 9 produces a Supp-reduced basis of an ordinary data table. Also note that due to relationship of data tables with graded attributes and data table over domains with similarity relations, we can use Algorithm 9 for determining Supp-reduced minimal bases of functional dependencies in data tables over domains with similarity relations which naturally appear in an extension of Codd's relational model which takes into account similarities on domains, see [6, 7].

So far, we focused ourselves on reduction of antecedents (left-hand sides) of FAIs. It is also possible to reduce consequents (right-hand sides) of FAIs. Note, however, that this type of reduction may not always be desirable because this may hide interesting attributes (consequences of left-hand sides) from users. The following theorem describes a reduction of right-hand sides of FAIs obtained by Algorithm 9.

Theorem 11. *Let $\langle X, Y, I\rangle$ be a data table with fuzzy attributes, T be the output of Algorithm 9 for $\langle X, Y, I\rangle$. Then we have $T \equiv T'$ for $T' = \{Q \Rightarrow R \ominus Q_{\mathcal{P}} \mid Q \Rightarrow R \in T\}$ where each $Q_{\mathcal{P}}$ denotes the same fuzzy set of attributes as in proof of Algorithm 9.*

Proof. Put $T'' = \{Q_{\mathcal{P}} \Rightarrow R \mid Q \Rightarrow R \in T\}$. Since $T \equiv T''$, it suffices to prove $\mathrm{Mod}(T') = \mathrm{Mod}(T'')$.

"$\subseteq$": Take $M \in \mathrm{Mod}(T')$ and $Q_{\mathcal{P}} \Rightarrow R \in T''$. We prove that $||Q_{\mathcal{P}} \Rightarrow R||_M = 1$. Assume $Q_{\mathcal{P}} \subseteq M$. Since $Q \subseteq Q_{\mathcal{P}}$, we get $Q \subseteq M$, i.e. assumption $||Q \Rightarrow R \ominus Q_{\mathcal{P}}||_M = 1$ gives $R \ominus Q_{\mathcal{P}} \subseteq M$ from which we get $R \subseteq (R \ominus Q_{\mathcal{P}}) \cup Q_{\mathcal{P}} \subseteq M$, proving $||Q_{\mathcal{P}} \Rightarrow R||_M = 1$. Hence, $M \in \mathrm{Mod}(T'')$.

"$\supseteq$": Let $M \in \mathrm{Mod}(T'')$, i.e. we have $||Q_{\mathcal{P}} \Rightarrow R||_M = 1$ for each $Q_{\mathcal{P}} \Rightarrow R \in T''$. For any $Q \Rightarrow R \ominus Q_{\mathcal{P}} \in T'$, we can prove the following claim: if, for each $Q' \Rightarrow R' \ominus Q'_{\mathcal{P}} \in T'$ such that $Q'_{\mathcal{P}}$ is lexically smaller than $Q_{\mathcal{P}}$, we have $||Q' \Rightarrow R'||_M = 1$, then $||Q \Rightarrow R||_M = 1$. We omit the proof of the claim due to the limited scope of this paper. Using the claim together with induction, we get that $||Q \Rightarrow R||_M = 1$ (and thus $||Q \Rightarrow R \ominus Q_{\mathcal{P}}||_M = 1$) for each $Q \Rightarrow R \ominus Q_{\mathcal{P}} \in T'$, i.e. this way we prove $M \in \mathrm{Mod}(T')$. □

Example 12. Let $\mathbf{L}$ be a five-element Łukasiewicz chain with globalization, where $L = \{0, 0.25, 0.5, 0.75, 1\}$ with its genuine ordering $0 < 0.25 < 0.5 < 0.75 < 1$. Consider the data table from Fig. 1. The data table describes birth/death rates and infant mortality of selected countries. In detail, the set X of objects contains countries "Germany", "Japan", ...; the set Y consists of six attributes: "low birth rate" (denoted "bl"), "high birth rate" (denoted "bh"), "low death rate" (denoted "dl"), "low infant mortality" (denoted "ml"), "high infant mortality" (denoted "mh"). The minimal basis of this table given by algorithm described in [2] is depicted in Fig. 2. The Supp-reduced minimal basis computed by Algorithm 9 is depicted in Fig. 3. As one can see, FAIs from the reduced basis are more readable. Compare, for instance, $\{{}^{0.75\!}/bl, dh, {}^{0.5\!}/ml, {}^{0.5\!}/mh\} \Rightarrow Y$ and the corresponding FAI $\{dh, {}^{0.5\!}/ml\} \Rightarrow Y$ from the reduced basis.

	birth rate low(bl)	birth rate high(bh)	death rate low(dl)	death rate high(dh)	infant mortality low(ml)	infant mortality high(mh)
Germany	1	0	0	0.75	1	0
Japan	0.75	0	0	0.5	1	0
Poland	0.75	0	0	0.75	0.5	0
Russia	0.75	0	0	1	0	0.5
USA	0	0.25	0	0.5	0.75	0

Figure 1. Illustrative data table

$\{bl, {}^{0.5}/dh\} \Rightarrow \{bl, {}^{0.75}/dh, ml\}$
$\{{}^{0.75}/bl, {}^{0.25}/bh, {}^{0.5}/dh, ml\} \Rightarrow Y$
$\{{}^{0.75}/bl, dh, {}^{0.5}/ml, {}^{0.5}/mh\} \Rightarrow Y$
$\{{}^{0.75}/bl, dh, {}^{0.75}/mh\} \Rightarrow Y$
$\{{}^{0.75}/bl, dh\} \Rightarrow \{{}^{0.75}/bl, dh, {}^{0.5}/mh\}$
$\{{}^{0.75}/bl, {}^{0.75}/dh, ml\} \Rightarrow \{bl, {}^{0.75}/dh, ml\}$
$\{{}^{0.75}/bl, {}^{0.5}/dh, {}^{0.75}/ml\} \Rightarrow \{{}^{0.75}/bl, {}^{0.5}/dh, ml\}$
$\{{}^{0.25}/bl, {}^{0.5}/dh\} \Rightarrow \{{}^{0.75}/bl, {}^{0.5}/dh\}$
$\{{}^{0.5}/bh, {}^{0.5}/dh, {}^{0.75}/ml\} \Rightarrow Y$
$\{{}^{0.25}/bh, {}^{0.5}/dh\} \Rightarrow \{{}^{0.25}/bh, {}^{0.5}/dh, {}^{0.75}/ml\}$
$\{{}^{0.25}/dl, {}^{0.5}/dh\} \Rightarrow Y$
$\{{}^{0.75}/dh\} \Rightarrow \{{}^{0.75}/bl, {}^{0.75}/dh\}$
$\{{}^{0.5}/dh, ml\} \Rightarrow \{{}^{0.75}/bl, {}^{0.5}/dh, ml\}$
$\{{}^{0.5}/dh, {}^{0.25}/ml\} \Rightarrow \{{}^{0.5}/dh, {}^{0.5}/ml\}$
$\{{}^{0.5}/dh, {}^{0.25}/mh\} \Rightarrow \{{}^{0.75}/bl, dh, {}^{0.5}/mh\}$
$\{\} \Rightarrow \{{}^{0.5}/dh\}$

Figure 2. Minimal basis.

$\{bl\} \Rightarrow \{{}^{0.75}/dh, ml\}$
$\{{}^{0.25}/bl, {}^{0.25}/bh, {}^{0.75}/ml\} \Rightarrow Y$
$\{dh, {}^{0.5}/ml\} \Rightarrow Y$
$\{dh, {}^{0.75}/mh\} \Rightarrow Y$
$\{dh\} \Rightarrow \{{}^{0.5}/mh\}$
$\{{}^{0.75}/dh, ml\} \Rightarrow \{bl\}$
$\{{}^{0.25}/bl, {}^{0.75}/ml\} \Rightarrow \{ml\}$
$\{{}^{0.25}/bl\} \Rightarrow \{{}^{0.75}/bl\}$
$\{{}^{0.5}/bh\} \Rightarrow Y$
$\{{}^{0.25}/bh\} \Rightarrow \{{}^{0.75}/ml\}$
$\{{}^{0.25}/dl\} \Rightarrow Y$
$\{{}^{0.75}/dh\} \Rightarrow \{{}^{0.75}/bl\}$
$\{ml\} \Rightarrow \{{}^{0.75}/bl\}$
$\{{}^{0.25}/ml\} \Rightarrow \{{}^{0.5}/ml\}$
$\{{}^{0.25}/mh\} \Rightarrow \{{}^{0.75}/bl, dh, {}^{0.5}/mh\}$
$\{\} \Rightarrow \{{}^{0.5}/dh\}$

Figure 3. Supp-reduced minimal basis.

5. Conclusions

This paper presents a method for generating of reduced minimal sets of if-then rules from data tables with graded attributes which characterize all dependencies valid in the data tables. The benefit of the presented approach is that, unlike the usual methods, our algorithm determines the reduced basis in a single step. Future research will focus on further types of reductions and their algorithmic tractability.

References

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

[2] Belohlavek R., Chlupová M., Vychodil V.: Implications from data with fuzzy attributes. In: Proc. AISTA 2004, 5 pages, ISBN 2–9599776–8–8.

[3] Belohlavek R., Vychodil V.: Reducing the size of fuzzy concept lattices by hedges. In: Proc. FUZZ-IEEE 2005, pp. 663–668, ISBN 0–7803–9158–6.

[4] Belohlavek R., Vychodil V.: Fuzzy attribute logic: attribute implications, their validity, entailment, and non-redundant basis. In: Proc. IFSA 2005, pp. 622–627. Tsinghua University Press and Springer, ISBN 7–302–11377–7.

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

[6] Belohlavek R., Vychodil V.: Functional dependencies of data tables over domains with similarity relations. In: *Proc. IICAI 2005*, pp. 2486–2504, ISBN 0–9727412–1–6.

[7] Belohlavek R., Vychodil V.: 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] Belohlavek R., Vychodil V.: Properties of models of fuzzy attribute implications (to appear in Proc. SCIS ISIS 2006).

[9] Belohlavek R., Vychodil V.: Graded LinClosure (submitted).

[10] Ganter B.: Algorithmen zur formalen Begriffsanalyse. In: Ganter B., Wille R., Wolff K. E. (Hrsg.): *Beiträge zur Begriffsanalyse*. B. I. Wissenschaftsverlag, Mannheim, 1987, 241–254.

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

[12] Goguen J. A.: The logic of inexact concepts. *Synthese* **18**(1968-9), 325–373.

[13] Guigues J.-L., Duquenne V.: Familles minimales d'implications informatives resultant d'un tableau de données binaires. *Math. Sci. Humaines* **95**(1986), 5–18.

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

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

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

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

[18] Pollandt S.: *Fuzzy Begriffe.* Springer-Verlag, Berlin/Heidelberg, 1997.

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

[20] Zhang C., Zhang S.: *Association Rule Mining. Models and Algorithms.* Springer, Berlin, 2002.