Attribute Dependencies for Data with Grades I

Radim Belohlavek, Vilem Vychodil
radim.belohlavek@acm.org, vychodil@acm.org

DATA ANALYSIS AND MODELING LAB

Dept. Computer Science
Palacky University, Olomouc, Czech Republic

european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness

INVESTMENTS IN EDUCATION DEVELOPMENT

Aims

– present content of our paper summarizing results on attribute implications in data with grades
– overview of main results
– structure of our paper:
  - Introduction
  - Scales of grades and basic principles of fuzzy logic
  - Graded attribute implications and their semantics (*)
  - Logic of graded attribute implications (*)
  - Bases of graded attribute implications (part II)
  - Algorithms
  - Reducing graded attribute implications to ordinary ones via thresholding (*)
  - Relationship to functional dependencies over domains with similarities (*)
  - Conclusions
– preliminary forms of most results appeared in conference proceedings, since 2004

We assume . . .

- basic familiarity with classical attribute implications
- basic familiarity with mathematical fuzzy logic
  - residuated lattices
  - truth-stressing hedges
  - fuzzy sets and relations
  - notation: $||\varphi||_{\mathbf{M}}$ . . . truth degree of φ in **M**, etc.
  - truth-functionality:

$$||\varphi \Rightarrow \psi|| = ||\varphi|| \rightarrow ||\psi||, \quad ||(\forall x)\varphi|| = \bigwedge_e ||\varphi||_e, \quad \text{etc.}$$

Graded attribute implications and their semantics

From ordinary attribute implications . . .

i.e. expressions

$$A \Rightarrow B \qquad \text{where } A, B \subseteq \text{finite set } Y \text{ of attributes}$$

such as

(a) $\{\text{prime}, > 2\} \Rightarrow \{\text{odd}\}$

(b) $\{\text{flight No.}\} \Rightarrow \{\text{depart. time, arriv. time}\}$

which are known as

(a) attribute implications (association rules):
presence of attributes in A implies presence of attributes in B

(b) functional dependencies:
match on attributes in A implies match on attributes in B

. . . we want to go to graded attribute implications

i.e. expressions

$$A \Rightarrow B \qquad \text{where } A, B \text{ are } L\text{-sets in } Y \text{ of attributes}$$

i.e. expressions of the form

$$\{{}^{a_1}\!/y_1, \dots, {}^{a_p}\!/y_p\} \Rightarrow \{{}^{b_1}\!/z_1, \dots, {}^{b_q}\!/z_q\},$$

with basic intended meaning:

presence of attributes y_is to degree at least a_is
implies
presence of attributes z_js to degree at least b_js

Functional dependencies? Very interesting in this view (later).

Example:

$$\{{}^{0.5}\!/\text{unhealthy food}, {}^{0.9}\!/\text{little activity}\} \Rightarrow \{{}^{0.7}\!/\text{high cholesterol}\}.$$

roughly says:

somewhat unhealthy food and very little activity implies rather high cholesterol

hence
an object x satisfies $A \Rightarrow B$ means: if x satisfies A then x satisfies B

with degrees involved, things are not so simple, care needed

x satisfies $A \Rightarrow B$ means: if x satisfies A then x satisfies B

consider

$$\{{}^{1}/y_1, {}^{0.5}/y_3\} \Rightarrow \{{}^{0.8}/y_2, {}^{1}/y_4\}$$

and three objects x_i (object x_i = L-set M_i =row i):

I	y_1	y_2	y_3	y_4
x_1	1.0	0.9	0.8	1.0
x_2	1.0	0.7	0.8	1.0
x_3	0.9	0.5	0.8	1.0

Does x_i satisfy A, B? Does x_i satisfy $A \Rightarrow B$? ...matters of degree.

When does object M fully satisfiy $A \Rightarrow B$ ($||A \Rightarrow B||_M = 1$)?

Two options we want to capture:

(1) if $A \subseteq M$ then $B \subseteq M$
(2) $S(A, M) \leq S(B, M)$

. . . which brings us to . . .

Definition

Let $\mathbf{L}$ be a complete residuated lattice $\mathbf{L}$ with a truth-stressing hedge *. The *degree* $||A \Rightarrow B||_M \in L$ *to which* $A \Rightarrow B$ *is valid in an* $\mathbf{L}$*-set* M of attributes is defined by

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

- option (1) results when * is globalization
- option (2) results when * is identity
- generalizes ordinary semantics ($L = \{0, 1\}$)

Theories, models, semantic entailment

Definition

- theory = (fuzzy) set T of implications
- models of T:

$$\mathrm{Mod}(T) = \{M \in \mathbf{L}^Y \mid \text{for each } A, B \in \mathbf{L}^Y\text{: } T(A \Rightarrow B) \leq ||A \Rightarrow B||_M\}$$

- degree to which $A \Rightarrow B$ semantically follows from theory T:

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

Some basic results

Theories may be reduced to equivalent crisp theories:

Theorem

Let T be a theory, $A \Rightarrow B$ be a graded attribute implication. For the crisp theory $\mathrm{c}(T)$ defined by

$$\mathrm{c}(T) = \{A \Rightarrow T(A \Rightarrow B) \otimes B \mid A, B \in \mathbf{L}^Y \text{ and } T(A \Rightarrow B) \otimes B \neq \emptyset\} \quad (1)$$

we have

$$\begin{aligned} \mathrm{Mod}(T) &= \mathrm{Mod}(\mathrm{c}(T)), & (2) \\ ||A \Rightarrow B||_T &= ||A \Rightarrow B||_{\mathrm{c}(T)}. & (3) \end{aligned}$$

Entailment may be reduced to entailment to degree 1:

Theorem

For a graded theory T and an implication $A \Rightarrow B$ we have

$$||A \Rightarrow B||_T = \bigvee\{c \in L \mid ||A \Rightarrow c \otimes B||_T = 1\}.$$

... behind is:

Lemma

$c \leq ||A \Rightarrow B||_M$ *iff* $||A \Rightarrow c \otimes B||_M = 1$.

Conceptually similar to Hájek's treatment of Pavelka's logic within BL-style Łukasiewicz logic.

A system $\mathcal{S} \subseteq \mathbf{L}^Y$ of $\mathbf{L}$-sets in Y is called an $\mathbf{L}^*$-closure system if it is closed under intersections and a^*-shifts, i.e. satisfies the following conditions:

$$\text{if } A_j \in \mathcal{S} \text{ for } j \in J \text{ then } \bigcap_{j \in J} A_j \in \mathcal{S},$$
$$\text{if } a \in L \text{ and } A \in \mathcal{S} \text{ then } a^* \rightarrow A \in \mathcal{S}.$$

Now, $\mathrm{Mod}(T)$ are just $\mathbf{L}^*$-closure systems:

Theorem

$\mathrm{Mod}(T)$ is an $\mathbf{L}^*$-closure system in Y for any graded theory T of implications over Y.
Let $\mathcal{S}$ be an $\mathbf{L}^*$-closure system in Y. Then there exists a theory T of graded attribute implications over Y such that $\mathcal{S} = \mathrm{Mod}(T)$.

Denote by $C_{\mathrm{Mod}(T)}(A)$ the least model of T containing A, where $C_{\mathrm{Mod}(T)}$ is the $\mathbf{L}^*$-closure operator corresponding to T.

Theorem

For every theory T and a graded attribute implication $A \Rightarrow B$,

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

Theorem

For every $\langle X, Y, I\rangle$,

$$||A \Rightarrow B||_{\langle X,Y,I\rangle} = ||A \Rightarrow B||_{\mathrm{Int}(X^*,Y,I)} = S(B, A^{\downarrow\uparrow}).$$

Logic of graded attribute implications

Ordinary-style logic

Armstrong-like rules:

(Ax) infer $A \cup B \Rightarrow A$,

(Cut) from $A \Rightarrow B$ and $B \cup C \Rightarrow D$ infer $A \cup C \Rightarrow D$,

(Mul) from $A \Rightarrow B$ infer $c^* \otimes A \Rightarrow c^* \otimes B$,

for each $A, B, C, D \in \mathbf{L}^Y$, and $c \in L$.

With ordinary notions of proof we have:

Theorem

Let L and Y be finite. For a set T of graded attribute implications and a graded attribute implication $A \Rightarrow B$,

$$T \vdash A \Rightarrow B \quad \textit{iff} \quad ||A \Rightarrow B||_T = 1.$$

More in paper (derived rules, other rules regarding similarity, etc.).

Pavelka-style logic

Aim: Capture syntactically degree of entailment.

Direct option using the framework of Pavelka's abstract logic:
RB+VV: A Logic of graded attributes (submitted).

Involves deduction rules operating with truth degrees, the notion of degree of provability, etc.

Continues our previous research in Pavelka-style logic:

- RB: Fuzzy equational logic. Archive for Math. Logic 41(2002), 83-90.
- RB: Birkhoff variety theorem and fuzzy logic. Archive for Math. Logic 42(8)(2003), 781-790.
- RB+VV: Fuzzy Horn logic I: proof theory. Archive for Math. Logic 45(1)(2006), 3-51.
- RB+VV: Fuzzy Horn logic II: implicationally defined classes. Archive for Math. Logic 45(2)(2006), 149-177.
- RB+VV: Fuzzy Equational Logic, Springer 2005, 283 pp.

Indirect option:
simulate degree of provability using bivalent notion of provability

Define the degree $|A \Rightarrow B|_T \in L$ to which $A \Rightarrow B$ is provable from T by

$$|A \Rightarrow B|_T = \bigvee\{c \in L \,|\, \mathrm{c}(T) \vdash A \Rightarrow c \otimes B\},$$

Theorem

Let $\mathbf{L}$ and Y be finite. Then for every fuzzy set T of fuzzy attribute implications and $A \Rightarrow B$ we have

$$|A \Rightarrow B|_T = ||A \Rightarrow B||_T.$$

Reducing graded attribute implications to ordinary ones via thresholding

Builds on the idea from

RB: Reduction and a simple proof of characterization of fuzzy concept lattices. Fundamenta Informaticae 46(4)(2001), 277-285

to represent fuzzy attributes by crisp attributes:

Given table $\mathcal{T} = \langle X, Y, I \rangle$ with graded attributes, denote by $\mathcal{T}^{\times}$ the table with crisp attributes defined by:

$$\mathcal{T}^{\times} = \langle X \times L, Y \times L, I^{\times} \rangle, \quad \text{where}$$
$$\langle \langle x, a \rangle, \langle y, b \rangle \rangle \in I^{\times} \text{ iff } a \otimes b \leq I(x, y).$$

Theorem

$\mathcal{B}(X, Y, I)$ (fuzzy concept lattice) is isomorphic to $\mathcal{B}(X \times L, Y \times L, I^{\times})$ (ordinary concept lattice).

In the present context involving a hedge, the same role is played by:

$$\mathcal{T}^{\times} = \langle X \times *(L), Y \times L, I^{\times} \rangle, \quad \text{where}$$

$$*(L) = \{a^* \mid a \in L\}, \text{ and}$$

$$\langle \langle x, a \rangle, \langle y, b \rangle \rangle \in I^{\times} \text{ iff } a \otimes b \leq I(x, y).$$

That is $\mathcal{B}(X^*, Y, I)$ is isomorphic to $\mathcal{B}(X \times *(L), Y \times L, I^{\times})$.

Do fuzzy attribute implications of $\mathcal{T} = \langle X, Y, I \rangle$ correspond to ordinary attribute implications of $\mathcal{T}^{\times} = \langle X \times *(L), Y \times L, I^{\times} \rangle$?

For an $\mathbf{L}$-set $B \in \mathbf{L}^Y$ we define the ordinary subset $\lfloor B \rfloor$ of $Y \times L$ by

$$\lfloor B \rfloor = \{\langle y, a\rangle \in Y \times L \,|\, a \leq B(y)\}.$$

For a subset $D \subseteq Y \times L$ we define the $\mathbf{L}$-set $\lceil D \rceil$ in Y by

$$\lceil D \rceil(y) = \bigvee \{a \,|\, \langle y, a\rangle \in D\}.$$

We are "translating" between fuzzy and crisp (data, implications).
Then:

Theorem

For a data table $\mathcal{T} = \langle X, Y, I\rangle$ with graded attributes, the corresponding $\mathcal{T}^{\times} = \langle X \times *(L), Y \times L, I^{\times}\rangle$, and arbitrary $A \in \mathbf{L}^Y$, $B \in \mathbf{L}^Y$ and $C, D \subseteq Y \times L$, we have

$$||A \Rightarrow B||_{\mathcal{T}} = 1 \quad \text{if and only if} \quad ||\lfloor A \rfloor \Rightarrow \lfloor B \rfloor||_{\mathcal{T}^{\times}} = 1;$$
$$||C \Rightarrow D||_{\mathcal{T}^{\times}} = 1 \quad \text{if and only if} \quad ||\lceil C \rceil \Rightarrow \lceil D \rceil||_{\mathcal{T}} = 1.$$

Does the theorem mean that problems regarding graded implications can be reduced to the corresponding problems of ordinary implications?

No:
Important example:

Bases(non-redundant fully informative implications) of $\mathcal{T} = \langle X, Y, I\rangle$ may be strictly smaller than bases of $\mathcal{T}^{\times} = \langle X \times *(L), Y \times L, I^{\times}\rangle$.

Why?: The dependencies reflecting the algebraic structure **L** of the set of grades are implicitly taken into account in the definition of entailment of graded implications over Y and need not be present in T. Their counterparts, however, are "not known" to the definition of (bivalent) semantic entailment of ordinary implications over $Y \times L$, and need thus be explicitly present in $T^{\times}$.

More in paper.

Relationship to functional dependencies over domains with similarities

Just a brief invitation . . .

– Ordinary attribute implications are naturally interpreted in relational databases.
– Graded attribute implications are naturally interpreted in principle in the same way in an interesting extension of relational databases in which domains are equipped with similarity relations.
– There are what is behind similarity-based queries such as "show houses with price approximately $200K and location near Vestal, NY".
– Hot topic in DB research.
– Graded implications play the role of functional dependencies, but: exact match (of ordinary model) is replaced by approximate match (w.r.t. similarity).
– New kind of dependencies, data mining appeal.

. . . in the paper we show that

such functional dependencies have the same logic as graded implication with the semantics described above, in particular:

Theorem

For every fuzzy set T of graded attribute implications and every graded attribute implication $A \Rightarrow B$ we have

$$||A \Rightarrow B||_T^{\mathrm{FD}} = ||A \Rightarrow B||_T^{\mathrm{AI}}.$$

Problems and solutions regarding entailment of functional dependencies may be translated to the equivalent ones regarding graded implications.

This extends the results from:

Fagin R.: Functional dependencies in a relational database and propositional logic. IBM J. Research and Development, 21(6)(1977), 543–544.

mentioned in Chap. 14 of Maier's book.