Derivation digraphs for graded if-then rules

Lucie Urbanová, Vilém Vychodil

DAMOL
DATA ANALYSIS AND MODELING LAB
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

Overview

1. Preliminaries
2. Graded If-then formulas in FCA of data with fuzzy attributes and in similarity-based relational model
3. Graph-based method of reasoning with graded if-then formulas

Preliminaries: structures of truth degrees

Complete residuated lattice: $\mathbf{L} = \langle L, \wedge, \vee, \otimes, \rightarrow, {}^*, 0, 1 \rangle$:

Truth-stressing hedges (Takeuti+Titani, Baaz, Hájek,...)
idempotent truth-stressing **hedge**
mapping $^* \colon L \to L$ satisfying

$$1^* = 1, \qquad a^* \leq a, \qquad (a \to b)^* \leq a^* \to b^*, \qquad a^{**} = a^*,$$

meaning of *: truth function of logical connective "very true"

Two boundary hedges

1. **identity**, i.e. $a^* = a$ $(a \in L)$;
2. **globalization**: $a^* = \begin{cases} 1 & \text{if } a = 1, \\ 0 & \text{otherwise.} \end{cases}$ ... interp. "fully true"

Notes on extension of the Codd model

- Y: set of all attributes
- Relation Schemes = subsets $R \subseteq Y$
- For $R \subseteq Y$ and domains D_y: $\mathrm{Tupl}(R) = \prod_{y \in R} D_y$
- **Similarity relation**: each domain D_y equipped with binary $\mathbf{L}$-relation $\approx_y$ satisfying:

(Ref) for each $u \in D_y$: $u \approx_y u = 1$,
(Sym) for each $u, v \in D_y$: $u \approx_y v = v \approx_y u$.

- **Domain with similarity** = $\langle D_y, \approx_y \rangle$

Definition (Ranked Data Table (RDT))

Let $R \subseteq Y$ be a relation scheme and $\langle D_y, \approx_y \rangle$ be a domain with similarity (for each $y \in R$). A **ranked data table on R over $\{\langle D_y, \approx_y \rangle \mid y \in R\}$** is any finite $\mathbf{L}$-set in $\mathrm{Tupl}(R)$.

Notes on extension of the Codd model: continue

Example

Similarity-based query: "Show me all cars which are sold for (approximately) $7\,000$ € and are (approximately) 3 years old."

brand	*model*	*price*	*year*	*km*
Renault	Scénic	$6\,940$	2009	$115\,556$
Renault	Scénic	$7\,200$	2010	$101\,478$
Opel	Zafira	$7\,500$	2008	$130\,656$
Citroen	Picasso	$7\,925$	2009	$109\,015$
Volkswagen	Caddy	$8\,600$	2008	$122\,855$

Notes on extension of the Codd model: continue

Example

Similarity-based query: "Show me all cars which are sold for (approximately) $7\,000$ € and are (approximately) 3 years old."

$\mathcal{D}(t)$
0.97
0.8
0.75
0.54
0.1

brand	*model*	*price*	*year*	*km*
Renault	Scénic	$6\,940$	2009	$115\,556$
Renault	Scénic	$7\,200$	2010	$101\,478$
Opel	Zafira	$7\,500$	2008	$130\,656$
Citroen	Picasso	$7\,925$	2009	$109\,015$
Volkswagen	Caddy	$8\,600$	2008	$122\,855$

Fuzzy Attribute Implication (FAI)

Y - finite nonempty set of attributes (features)

Definition

Fuzzy attribute implication (graded if-then rule) on Y is any expression of the from $A \Rightarrow B$, where $A, B \in L^Y$. An $\mathbf{L}$-set T of FAIs on Y shall be called a **theory**.

Two basic interpretations:

1. attribute implications in FCA of data with fuzzy attributes
   - For every object: "If it is very true that an object has all attributes from A, then it has also all attributes from B."
2. similarity-based functional dependencies in an extension of the Codd model of data
   - For every pair of tuples: "If they have very similar values on attributes from A, then they have similar values on attributes from B."

Semantic entailment

1. Given a fuzzy attribute implication $A \Rightarrow B$:
   - $M \in L^Y$, $||A \Rightarrow B||_M = S(A, M)^* \rightarrow S(B, M)$
   - $||A \Rightarrow B||_{\mathcal{D}} = \bigwedge_{t_1, t_2 \in \mathrm{Tupl}(R)} \Big((t_1(A) \approx_{\mathcal{D}} t_2(A))^* \rightarrow (t_1(B) \approx_D t_2(B))\Big)$
     $t_1(C) \approx_{\mathcal{D}} t_2(C) = (\mathcal{D}(t_1) \otimes \mathcal{D}(t_2)) \rightarrow \bigwedge_{y \in R}(C(y) \rightarrow t_1(y) \approx_y t_2(y))$
2. **Model** of theory T (fuzzy set of FAIs):
   - $\mathrm{Mod}^{AI}(T) = \{M \in L^Y \,|$ for each $A, B \in L^Y : T(A \Rightarrow B) \leq ||A \Rightarrow B||_M\}$
   - $\mathrm{Mod}^{FD}(T) = \{\mathcal{D}|$ for each $A, B \in L^Y : T(A \Rightarrow B) \leq ||A \Rightarrow B||_D\}$
3. A degree $||A \Rightarrow B||_T$ to which $A \Rightarrow B$ **semantically follows** from T:
   - $||A \Rightarrow B||_T^{AI} = \bigwedge_{M \in \mathrm{Mod}^{AI}(T)} ||A \Rightarrow B||_M$
   - $||A \Rightarrow B||_T^{FD} = \bigwedge_{\mathcal{D} \in \mathrm{Mod}^{FD}(T)} ||A \Rightarrow B||_{\mathcal{D}}$

Theorem (Bělohlávek, Vychodil, 2005)

For any fuzzy set T of FAIs and any FAI $A \Rightarrow B$ we have: $||A \Rightarrow B||_T^{AI} = ||A \Rightarrow B||_T^{FD}$

Semantic entailment

1. Given a fuzzy attribute implication $A \Rightarrow B$:
   - $M \in L^Y$, $||A \Rightarrow B||_M = S(A, M)^* \rightarrow S(B, M)$
   - $||A \Rightarrow B||_{\mathcal{D}} = \bigwedge_{t_1, t_2 \in \mathrm{Tupl}(R)} \Big((t_1(A) \approx_{\mathcal{D}} t_2(A))^* \rightarrow (t_1(B) \approx_D t_2(B)) \Big)$
     $t_1(C) \approx_{\mathcal{D}} t_2(C) = (\mathcal{D}(t_1) \otimes \mathcal{D}(t_2)) \rightarrow \bigwedge_{y \in R} (C(y) \rightarrow t_1(y) \approx_y t_2(y))$
2. **Model** of theory T (fuzzy set of FAIs):
   - $\mathrm{Mod}^{AI}(T) = \{M \in L^Y \,|\, \text{for each } A, B \in L^Y : T(A \Rightarrow B) \leq ||A \Rightarrow B||_M\}$
   - $\mathrm{Mod}^{FD}(T) = \{\mathcal{D} \,|\, \text{for each } A, B \in L^Y : T(A \Rightarrow B) \leq ||A \Rightarrow B||_D\}$
3. A degree $||A \Rightarrow B||_T$ to which $A \Rightarrow B$ **semantically follows** from T:
   - $||A \Rightarrow B||_T^{AI} = \bigwedge_{M \in \mathrm{Mod}^{AI}(T)} ||A \Rightarrow B||_M$
   - $||A \Rightarrow B||_T^{FD} = \bigwedge_{\mathcal{D} \in \mathrm{Mod}^{FD}(T)} ||A \Rightarrow B||_{\mathcal{D}}$

Theorem (Bělohlávek, Vychodil, 2005)

For any fuzzy set T of FAIs and any FAI $A \Rightarrow B$ we have: $||A \Rightarrow B||_T^{AI} = ||A \Rightarrow B||_T^{FD}$

Armstrong-like axiomatic system

$\mathbf{L}$ finite, Y set of attributes, $A, B, C, D \in L^Y$:

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

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

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

Two kind of completeness can be shown for all finite residuated lattices:

- **Ordinary** completeness
  - T - crisp theory
  - $A \Rightarrow B$ is provable from T iff $A \Rightarrow B$ semantically follows from T (in degree 1)
- **Graded (Pavelka-style)** completeness
  - T - an $\mathbf{L}$-set
  - A degree to which $A \Rightarrow B$ semantically follows from T equals to degree to which $A \Rightarrow B$ is provable from T

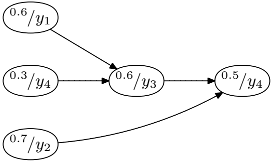
$$||A \Rightarrow B||_T = \bigvee\{c \in L \mid T \vdash A \Rightarrow c \otimes B\}$$

Bělohlávek, Vychodil, 2006: Completeness for general complete residuated lattices (+ infinitary deduction rule)

Construction of derivation digraphs for FAIs

David Maier (1980): graph-based approach - to normalize proof. Can this approach be extended for FAIs? **YES**.

Derivation digraphs:



- digraph - directed graph
- acyclic graph
- vertices are attributes from Y
- arcs correspond to FAIs used from an input theory

Construction of T-based $\mathbf{L}^*$-derivation DAG: T be a set of FAIs over Y.

- Any $\mathbf{D} = \langle V, \emptyset \rangle$ such that $\emptyset \neq V \subseteq Y \times L$ and for every $y \in Y$ there is at most one $a \in L$ such that $\langle y, a \rangle \in V$, is a T-based $\mathbf{L}^*$-derivation DAG;

Construction of derivation digraphs for FAIs: continue

- New vertex for attribute y

  If $\mathbf{D} = \langle V, A \rangle$ is a T-based $\mathbf{L}^*$-derivation DAG and there are $E \Rightarrow F \in T$, attribute $y \in Y$, and vertices $\langle y_1, a_1 \rangle \in V, \ldots, \langle y_k, a_k \rangle \in V$ such that for

$$s_0 = \bigwedge\{E(y) \to 0 \,|\, y \in Y \text{ and } y \notin \{y_1, \ldots, y_k\}\},$$
$$s_1 = \bigwedge\{E(y_i) \to a_i \,|\, i = 1, \ldots, k\},$$
$$m = \bigvee\{a \in L \,|\, \langle y, a \rangle \in V\},$$
$$d = \big((s_0 \wedge s_1)^* \otimes F(y)\big) \vee m,$$

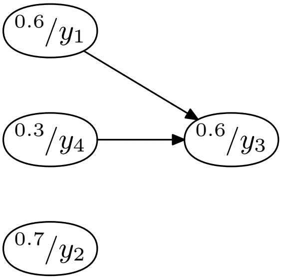
  we have $d > m$, then $\mathbf{D}' = \langle V', A' \rangle$, where

$$V' = V \cup \{\langle y, d \rangle\},$$
$$A' = A \cup \{\langle \langle y_i, a_i \rangle, \langle y, d \rangle \rangle \,|\, i = 1, \ldots, k\},$$

  is a T-based $\mathbf{L}^*$-derivation DAG.

One step in construction of DAG

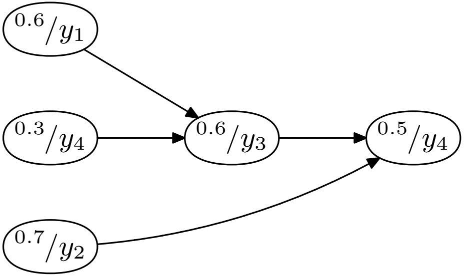
- $\mathbf{L}^*$: finite linearly ordered Łukasiewicz algebra with $L = \{0, 0.1, \ldots, 0.9, 1\} \subseteq [0, 1]$ $\wedge$ and $\vee$ usual minima and maxima, $\otimes, \rightarrow$ Łukasiewicz operations, $a^* = 1$ for $a = 1$, $a^* = 0.6$ for $0.6 \le a \le 0.9$, $a^* = 0.2$ for $0.2 \le a \le 0.5$, $a^* = 0$ for $0 \le a \le 0.1$
- set of initial vertices: $\{{}^{0.6}/y_1, {}^{0.7}/y_2, {}^{0.3}/y_4\}$, theory: $T = \{\ldots, \{{}^{0.7}/y_2, {}^{0.9}/y_3\} \Rightarrow \{{}^{0.9}/y_4\}, \ldots\}$
- adding vertex $\langle y_4, 0.5 \rangle$



$s_0 = 1$
$s_1 = 0.7$
$m = 0.3$
$d = (0.7^* \otimes 0.9) \vee 0.3$
$= (0.6 \otimes 0.9) \vee 0.3$
$= 0.5$

One step in construction of DAG

- $\mathbf{L}^*$: finite linearly ordered Łukasiewicz algebra with $L = \{0, 0.1, \ldots, 0.9, 1\} \subseteq [0, 1]$
  $\wedge$ and $\vee$ usual minima and maxima, $\otimes, \rightarrow$ Łukasiewicz operations,
  $a^* = 1$ for $a = 1$, $a^* = 0.6$ for $0.6 \leq a \leq 0.9$, $a^* = 0.2$ for $0.2 \leq a \leq 0.5$, $a^* = 0$ for $0 \leq a \leq 0.1$
- set of initial vertices: $\{{}^{0.6}/y_1, {}^{0.7}/y_2, {}^{0.3}/y_4\}$,
  theory: $T = \{\ldots, \{{}^{0.7}/y_2, {}^{0.9}/y_3\} \Rightarrow \{{}^{0.9}/y_4\}, \ldots\}$
- adding vertex $\langle y_4, 0.5\rangle$



$$
\begin{aligned}
s_0 &= 1 \\
s_1 &= 0.7 \\
m &= 0.3 \\
d &= (0.7^* \otimes 0.9) \vee 0.3 \\
&= (0.6 \otimes 0.9) \vee 0.3 \\
&= 0.5
\end{aligned}
$$

Derivation DAG: further notions

For T-based $\mathbf{L}^*$-derivation DAG $\mathbf{D}$ we have:

- **Yield of $\mathbf{D}$ on attribute y**: $\mathbf{D}(y) = \bigvee\{a \in L \mid \langle y, a\rangle \in V\}$
- $\mathbf{D}$ is called a **T-based $\mathbf{L}^*$-derivation DAG for $A \Rightarrow B$** if
  1. $\mathbf{D}(y) \geq B(y)$ for all $y \in Y$.
  2. If $A \neq \emptyset$, then the set of initial vertices of $\mathbf{D}$ is
  $$\{\langle y, A(y)\rangle \mid y \in Y \text{ and } A(y) > 0\}$$
  3. If $A = \emptyset$, then the set of initial vertices of $\mathbf{D}$ is $\{\langle y^{\#}, 0\rangle\}$, where $y^{\#}$ is a fixed attribute from Y.

Results: Completeness

Let T be a crisp theory

Theorem

If $T \vdash A \Rightarrow B$ (using (Ax),(Cut),(Mul)), then there is a T-based $\mathbf{L}^*$-derivation DAG for $A \Rightarrow B$.

Proof.

Assume that $T \vdash A \Rightarrow B$. Then, we can show (by induction) that there is a sequence $A_0, \ldots, A_n$ of $\mathbf{L}$-sets in Y such that $A_0 = A$, $B \subseteq A_n$ and for all $i = 0, \ldots, n-1$ there is some $E \Rightarrow F \in T$ and attribute $y \in Y$ such that

$$A_{i+1} = A_i \cup \left(S(E, A_i)^* \otimes \left\{{}^{F(y)}\!/y\right\}\right)$$

$A_i \subset A_{i+1}$, and $T \vdash A \Rightarrow A_i$ for all $i = 0, \ldots, n$. Based on this sequence we will construct T-based $\mathbf{L}^*$-derivation DAGs $\mathbf{D}_0, \ldots, \mathbf{D}_n$. $\mathbf{D}_0$ contains just the initial vertices corresponding to A and $\mathbf{D}_{i+1}$ results from $\mathbf{D}_i$ by adding the vertex $\langle y, A_{i+1}(y)\rangle$ where $A_{i+1}(y) = A_i(y) \vee (S(E, A_i)^* \otimes F(y))$. Obviously, for each $y \in Y$, we have $\mathbf{D}_i(y) = A_i(y)$. In a particular case, $\mathbf{D}_n$ is a T-based $\mathbf{L}^*$-derivation DAG for $A \Rightarrow B$. □

Theorem

If there is a T-based $\mathbf{L}^*$-derivation DAG for $A \Rightarrow B$, then $T \vdash A \Rightarrow B$.

Proof.

By induction on the complexity of T-based $\mathbf{L}^*$-derivation DAGs:
1) If $\mathbf{D}$ consists solely of the initial vertices, then $T \vdash A \Rightarrow B$ follows using (Ax).
2) Assume that $\mathbf{D}$ results from $\mathbf{D}' = \langle V', A' \rangle$ by adding vertex $\langle y, d \rangle$. Therefore, there are $E \Rightarrow F \in T$, attribute $y \in Y$, and vertices $\langle y_1, a_1 \rangle \in V', \ldots, \langle y_k, a_k \rangle \in V'$ such that $\mathbf{D}$ results by adding vertex $\langle y, d \rangle$, where $d = ((s_0 \wedge s_1)^* \otimes F(y)) \vee \mathbf{D}'(y)$. Let $B_{\mathbf{D}'} \in L^Y$ such that $B_{\mathbf{D}'}(y) = \mathbf{D}(y)$, $B \subseteq \{{}^{d}\!/y\} \cup B_{\mathbf{D}'}$. Consider $G \subseteq B_{\mathbf{D}'}$ such that $G(y_i) = a_i$ (for all $i = 1, \ldots, k$) and $G(y') = 0$ otherwise. It can be shown that from induction hypothesis $T \vdash A \Rightarrow G$ follow also $T \vdash A \Rightarrow \{{}^{S(E,G)^* \otimes F(y)}\!/y\} \cup B_{\mathbf{D}'}$. Since $s_0 \wedge s_1 = S(E, G)$, we have $T \vdash A \Rightarrow \{{}^{d}\!/y\} \cup B_{\mathbf{D}'}$ and by derived inference rule Projectivity $T \vdash A \Rightarrow B$. □

Theorem

If $\mathbf{L}$ is finite, then $A \Rightarrow B$ semantically follows from theory T in degree 1 iff there is a T-based $\mathbf{L}^*$-derivation DAG for $A \Rightarrow B$.

Proof.

Follows from previous theorems and completeness of fuzzy attribute logic. □

Theorem

If $\mathbf{L}$ is finite, then $||A \Rightarrow B||_T$ is the greatest degree $a \in L$ such that there is a T-based $\mathbf{L}^*$-derivation DAG for $A \Rightarrow a{\otimes}B$.

Proof.

Follow from the fact that

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

and from previous theorem. □

Results: Completeness

Let T be an $\mathbf{L}$-set
We can define crisp(T) by:

$$\text{crisp}(T) = \{A \Rightarrow T(A \Rightarrow B) \otimes B | A, B \in L^Y \text{ and } T(A \Rightarrow B) \otimes B \not\subseteq A\}$$

We know:

- $\text{Mod}(T) = \text{Mod}(\text{crisp}(T))$
- $||A \Rightarrow B||_T = ||A \Rightarrow B||_{\text{crisp}(T)}$.

Theorem

If $\mathbf{L}$ is finite and T is an $\mathbf{L}$-set of FAIs, then $||A \Rightarrow B||_T$ is the greatest degree $a \in L$ such that there is a T'-based $\mathbf{L}^*$-derivation DAG for $A \Rightarrow a \otimes B$, where $T' = \text{crisp}(T)$.

Results: determine degree $||A \Rightarrow B||_T$

Main ideas ($\mathbf{L}$, Y finite):

- A_T^+: largest $\mathbf{L}$-set such that $T \vdash A \Rightarrow A_T^+$
- Property (Bělohlávek, Vychodil, 2006): $||A \Rightarrow B||_T = S(B, A_T^+)$
- Informally: Final T-based $\mathbf{L}^*$-derivation DAG - cannot be enlarged

Theorem

Let $\mathbf{D}$ be a derivation DAG for $A \Rightarrow B$, then $\mathbf{D}$ is final iff $\mathbf{D}(y) = A_T^+(y)$ for all $y \in Y$

1. Construct a T-based $\mathbf{L}^*$-derivation DAG $\mathbf{D} = \langle V, \emptyset \rangle$ with
$$V = \{\langle y, A(y)\rangle \mid A(y) > 0\};$$
If $V = \emptyset$, put $V = \{\langle y^\sharp, 0\rangle\}$.
2. If $\mathbf{D}$ can be enlarged according to Definition (case 2.), then enlarge $\mathbf{D}$ and repeat step 2.; otherwise return $S(B, A_T^+)$, where $A_T^+(y) = \mathbf{D}(y)$ for all $y \in Y$.

Summary

- graph-based inference for FAIs
- degrees of semantic entailment of FAIs from a theory can be characterized by the existence of directed acyclic graphs
- characterization of syntactic closures via derivation DAGs
- algorithm for computing closures

Future work:

- further properties of derivation DAGs, complexity issues