Concept Lattices and Formal Concept Analysis

Radim Bělohlávek

Dept. Computer Science
Palacký University, Olomouc
radim.belohlavek@upol.cz

INFO

This is a preliminary version of a text on formal concept analysis and related methods.

FORMAL CONCEPT ANALYSIS

What is FCA?

– method of **analysis of object-attribute data**

– output 1: hierarchical structure of clusters (**concept lattice**)

– output 2: base of **attribute implications**

– existing **software** support

– documented **applications**

– nontrivial **open problems** (mathematical, algorithmic, methodological)

Origins of FCA

G. Birkhoff: **Lattice Theory.** AMS Col. Publ. 25, 1940.

M. Barbut: Note sur l'algèbre des techniques d'analyse hiérarchique. In: B. Matalon: *L'analyse hiérarchique.* Gauthier-Villars, Paris, 1965, pp. 125–146.

M. Barbut, B. Monjardet: **Ordre et Classification**, Vol. 2. Hachette, Paris, 1970.

R. Wille: **Restructuring lattice theory**: an approach based on hierarchies of concepts. In: Rival I.: *Ordered Sets.* Reidel, 1982, 445–470.

state of art (almost): B. Ganter, R. Wille: **Formal Concept Analysis: Mathematical Foundations.** Springer, 1999.

What is a concept?

- psychology (approaches: classical, prototype, exemplar, knowledge)
- logic (TIL)
- artificial intelligence (frames, learning of concepts)
- conceptual graphs (Sowa)
- “conceptual modeling”
- ...
- **traditional/Port-Royal logic**

Traditional/Port-Royal approach to concepts

- **concept :=** extent + intent
  - extent = objects covered by concept
  - intent = attributes covered by concept

Traditional/Port-Royal approach to concepts

- **concept := extent + intent**
  - extent = objects covered by concept
  - intent = attributes covered by concept
- **example: DOG**
  - extent of dog = collection of all dogs
  - intent of dog = collection of all dogs' attributes (barks, has four limbs, has tail, . . .)

Traditional/Port-Royal approach to concepts

- **concept := extent + intent**
  - **extent** = objects covered by concept
  - **intent** = attributes covered by concept
- **example: DOG**
  - extent of dog = collection of all dogs
  - intent of dog = collection of all dogs' attributes (barks, has four limbs, has tail, . . .)
- **concept hierarchy**
  - subconcept/superconcept relation
  - **concept1=(extent1,intent1)** $\leq$ **concept2=(extent2,intent2)** $\Leftrightarrow$ extent1 $\subseteq$ extent2 ($\Leftrightarrow$ intent1 $\supseteq$ intent2)
  - DOG $\leq$ MAMMAL $\leq$ ANIMAL

Basic notions of FCA

- **formal context** (input data table)
- **formal concept** (cluster in data)
- **concept lattice** (hierarchical system of clusters)
- **attribute implication** (dependency in data)

Formal context = input data

Def. **Formal context** is a triplet (X, Y, I) where
X . . . set of **objects**
Y . . . set of **attributes**
$I \subseteq X \times Y$ binary relation.

Interpretation: $(x, y) \in I$ **. . . object** x **has attribute** y

formal context ≈ **data table**

I	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×		×	×
x_3		×	×	×
x_4		×	×	×
x_5	×			

Formal concept = cluster in data

Def. **Induced operators** . . . mappings $^{\uparrow}: 2^X \to 2^Y$, $^{\downarrow}: 2^Y \to 2^X$ def. by:

$$A^{\uparrow} = \{y \in Y \mid \text{for each } x \in A : (x, y) \in I\}$$

$$B^{\downarrow} = \{x \in X \mid \text{for each } y \in B : (x, y) \in I\}$$

$A^{\uparrow}$ . . . attributes common to all objects from A

$B^{\downarrow}$ . . . objects sharing all attributes from B

Def. **Formal concept** in (X, Y, I) . . . (A, B), $A \subseteq X$, $B \subseteq Y$, s.t.

$$A^{\uparrow} = B \text{ and } B^{\downarrow} = A.$$

A . . . **extent** . . . objects covered by formal concept

B . . . **intent** . . . attributes covered by formal concept

Formal concepts as maximal rectangles

Thm. Formal concepts are exactly **maximal rectangles** in data table.

Example

I	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×		×	×
x_3		×	×	×
x_4		×	×	×
x_5	×			

formal concept $(A, B) = (\{x_1, x_2, x_3, x_4\}, \{y_3, y_4\})$

Further formal concepts

I	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×		×	×
x_3		×	×	×
x_4		×	×	×
x_5	×			

I	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×		×	×
x_3		×	×	×
x_4		×	×	×
x_5	×			

I	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×		×	×
x_3		×	×	×
x_4		×	×	×
x_5	×			

$$(A_1, B_1) = (\{x_1, x_3, x_4\}, \{y_2, y_3, y_4\})$$

$$(A_2, B_2) = (\{x_1, x_2\}, \{y_1, y_3, y_4\})$$

$$(A_3, B_3) = (\{x_1, x_2, x_5\}, \{y_1\})$$

Concept lattice

Def. **Subconcept-superconcept** ordering $\leq$ of formal concepts is defined by

$$(A_1, B_1) \leq (A_2, B_2) \quad \text{iff} \quad A_1 \subseteq A_2 \text{ (iff } B_2 \subseteq B_1).$$

Example DOG $\leq$ MAMMAL

Def. **Concept lattice** (**Galois lattice**) of (X, Y, I) is the set

$$\mathcal{B}(X, Y, I) = \{(A, B) \mid A^{\uparrow} = B, B^{\downarrow} = A\}$$

equipped with $\leq$.

Rem. $\mathcal{B}(X, Y, I)$ ... all concepts/clusters hidden in the data

Denote
$\text{Ext}(X, Y, I) = \{A \in 2^X \mid \langle A, B\rangle \in \mathcal{B}(X, Y, I) \text{ for some } B\}$ (extents of concepts)
$\text{Int}(X, Y, I) = \{B \in 2^Y \mid \langle A, B\rangle \in \mathcal{B}(X, Y, I) \text{ for some } A\}$ (intents of concepts)

Alternative notation

$\approx$ membership/characteristic function style

instead of $A \subseteq U$, consider corresponding $C_A \in 2^U$

that is: $C_A(u) = \begin{cases} 0 & \text{for } u \notin A \\ 1 & \text{for } u \in A \end{cases}$

we identify A with C_A, i.e. we write $A(u) = 0$, $A(u) = 1$,

i.e. $I(x, y) = 1$ if $\langle x, y \rangle \in I$

then

$$A^{\uparrow}(y) = \bigwedge_{x \in X} A(x) \to I(x, y)$$

$$B^{\downarrow}(x) = \bigwedge_{y \in Y} B(y) \to I(x, y)$$

where $\bigwedge$ denotes min and $\to$ is bivalent implication connective

Formal concepts as maximal rectangles

A rectangle in $\langle X, Y, I\rangle$ is a pair $\langle A, B\rangle$ such that for each $x \in A$ and $y \in B$ we have $\langle x, y\rangle \in I$ (that is: the rectangle corresponding to A and B is filled with 1's). For two rectangles $\langle A_1, B_1\rangle$ and $\langle A_2, B_2\rangle$ we put $\langle A_1, B_1\rangle \sqsubseteq \langle A_2, B_2\rangle$ iff $A_1 \subseteq A_2$ and $B_1 \subseteq B_2$.

Theorem Formal fuzzy concepts are exactly **maximal rectangles** (w.r.t. $\sqsubseteq$) in $\langle X, Y, I\rangle$.

Proof (by a simple reflection, viz prednasky) □

Related mathematical structures

Def. A **Galois connection** between sets X and Y is a pair $\langle f, g\rangle$ of mappings $f : 2^X \to 2^Y$ and $g : 2^Y \to 2^X$ satisfying for $A, A_1, A_2 \subseteq X$, $B, B_1, B_2 \subseteq Y$:

$$A_1 \subseteq A_2 \Rightarrow f(A_2) \subseteq f(A_1), \tag{1}$$
$$B_1 \subseteq B_2 \Rightarrow g(B_2) \subseteq g(B_1), \tag{2}$$
$$A \subseteq g(f(A)), \tag{3}$$
$$B \subseteq f(g(B)). \tag{4}$$

Lemma (**chaining of Galois connection**) For a Galois connection $\langle f, g\rangle$ between X and Y we have $f(A) = f(g(f(A)))$ and $g(B) = g(f(g(B)))$ for any $A \subseteq X$ and $B \subseteq Y$.

Proof We prove only $f(A) = f(g(f(A)))$ ($g(B) = g(f(g(B)))$ is dual): $f(A) \subseteq f(g(f(A)))$ follows from (4) by putting $B = f(A)$. Since $A \subseteq g(f(A))$ by (3), we get $f(A) \supseteq f(g(f(A)))$ by application of (1). $\square$

Remark For a Galois connection $\langle f, g\rangle$ between X and Y we put $\text{fix}(f, g) = \{\langle A, B\rangle \mid A \in 2^X, B \in 2^Y, A^\uparrow = B, B^\downarrow = A\}$. $\text{fix}(f, g)$ is called the set of all fixed points of $\langle f, g\rangle$.

Def. A **closure operator** on a set X is a mapping $C : 2^X \to 2^X$ satisfying for each $A, A_1, A_2 \subseteq X$

$$A \subseteq C(A), \tag{5}$$

$$A_1 \subseteq A_2 \Rightarrow C(A_1) \subseteq C(A_2), \tag{6}$$

$$A = C(C(A)). \tag{7}$$

Remark For a closure operator C on X we put fix(C) $= \{A \mid A \in 2^X A = C(A)\}$. fix($C$) is called the set of all fixed points of C.

Def. A **complete lattice** is a partially ordered set $\langle V, \leq \rangle$ such that for each $K \subseteq V$ there exists both the infimum inf(K) of K and the supremum sup(K) of K.

Recall: partial order, lower bound, upper bound, infimum, supremum, ...

Given a formal context $\langle X, Y, I \rangle$, the induced operators $^\uparrow$ and $^\downarrow$ will also be denoted by $^{\uparrow_I}$ and $^{\downarrow_I}$.

Theorem (fixpoints of closure operators) For a closure operator C on X, $\langle \text{fix}(C), \subseteq \rangle$ is a complete lattice with infima and suprema given by

$$\bigwedge_{j\in J} A_j = \bigcap_{j\in J} A_j, \tag{8}$$

$$\bigvee_{j\in J} A_j = C(\bigcup_{j\in J} A_j). \tag{9}$$

Proof Evidently, $\langle \text{fix}(C), \subseteq \rangle$ is a partially ordered set. First, we verify that for $A_j \in \text{fix}(C)$ we have $\bigcap_{j\in J} A_j \in \text{fix}(C)$, i.e. $\bigcap_{j\in J} A_j = C(\bigcap_{j\in J} A_j)$. $\bigcap_{j\in J} A_j \subseteq C(\bigcap_{j\in J} A_j)$ is obvious (a property of a closure operator). Conversely, we have $C(\bigcap_{j\in J} A_j) \subseteq \bigcap_{j\in J} A_j$ iff for each $j \in J$ we have $C(\bigcap_{j\in J} A_j) \subseteq A_j$ which is true. indeed, we have $\bigcap_{j\in J} A_j \subseteq A_j$ and so $C(\bigcap_{j\in J} A_j) \subseteq C(A_j) = A_j$. Now it is clear that $\bigcap_{j\in J} A_j$ is the infimum of A_j's (first, $\bigcap_{j\in J} A_j$ is less than each A_j; second, $\bigcap_{j\in J} A_j$ is above any $A \in \text{fix}(C)$ which is less than all A_j's).

Second, we verify $\bigvee_{j\in J} A_j = C(\bigcup_{j\in J} A_j)$. Since $\bigvee_{j\in J} A_j \supseteq A_j$ for any $j \in J$, we get $\bigvee_{j\in J} A_j \supseteq \bigcup_{j\in J} A_j$, and so $\bigvee_{j\in J} A_j = C(\bigvee_{j\in J} A_j) \supseteq C(\bigcup_{j\in J} A_j)$. On the other hand, $C(\bigcup_{j\in J} A_j)$ is a fixpoint which is above each A_j, and so it is above their supremum $\bigvee_{j\in J} A_j$, i.e. $C(\bigcup_{j\in J} A_j) \supseteq \bigvee_{j\in J} A_j$. To sum up, $\bigvee_{j\in J} A_j = C(\bigcup_{j\in J} A_j)$. □

Theorem (**binary relation induces Galois connection**) For each formal context $\langle X, Y, I\rangle$, the pair $\langle {}^{\uparrow_I}, {}^{\downarrow_I}\rangle$ forms a Galois connection between X and Y.

Proof Easy by direct verification (viz prednasky).

Remark Therefore, a concept lattice $\mathcal{B}(X, Y, I)$ is but a system of fixed points of the induced Galois connection $\langle {}^{\uparrow}, {}^{\downarrow}\rangle$, i.e. $\mathcal{B}(X, Y, I) = \text{fix}({}^{\uparrow}, {}^{\downarrow})$.

Conversely, a question arises as to whether each Galois connection $\langle f, g\rangle$ is induced by some binary relation I between X and Y.

Theorem (Galois connection is induced by binary relation) Let $\langle f, g\rangle$ be a Galois connection between X and Y. Then putting for each $x \in X$ and $y \in Y$

$$\langle x, y\rangle \in I \quad \text{iff} \quad y \in f(\{x\}) \qquad \text{or, equivalently, iff } x \in g(\{y\}), \tag{10}$$

I is a binary relation between X and Y such that the induced Galois connection $\langle {}^{\uparrow_I}, {}^{\downarrow_I}\rangle$ coincides with $\langle f, g\rangle$, i.e. $\langle {}^{\uparrow_I}, {}^{\downarrow_I}\rangle = \langle f, g\rangle$.

Proof First, let us show that $y \in f(\{x\})$ iff $x \in g(\{y\})$: From $y \in f(\{x\})$ we get $\{y\} \subseteq f(\{x\})$ from which we get $\{x\} \subseteq g(f(\{x\})) \subseteq g(\{y\})$, i.e. $x \in g(\{y\})$. In a similar manner, $x \in g(\{y\})$ implies $y \in f(\{x\})$. That is, we have $\langle x, y\rangle \in I$ iff $y \in f(\{x\})$ iff $x \in g(\{y\})$.

Now, for each $A \subseteq X$ we have $f(A) = f(\cup_{x\in A}\{x\}) = \cap_{x\in A} f(\{x\}) = \cap_{x\in A}\{y \in Y \mid y \in f(\{x\})\} = \cap_{x\in A}\{y \in Y \mid \langle x, y\rangle \in I\} = \{y \in Y \mid \text{ for each } x \in A : \langle x, y\rangle \in I\} = A^{\uparrow_I}$.

Dually, for $B \subseteq Y$ we get $g(B) = B^{\downarrow_I}$. □

Remark (1) The relation I induced from a Galois connection $\langle f, g\rangle$ by (10) will also be denetode by $I_{\langle f,g\rangle}$.

(2) It is easy to see that $I = I_{\langle {}^{\uparrow_I}\rangle, \langle {}^{\downarrow_I}\rangle}$ and $\langle {}^{\uparrow}, {}^{\downarrow}\rangle = \langle {}^{\uparrow_{I_{\langle\uparrow,\downarrow\rangle}}}, {}^{\downarrow_{I_{\langle\uparrow,\downarrow\rangle}}}\rangle$. That is,

there is a one-to-one correspondence between binary relations between I and Galois connections between X and Y.

Corollary (consequences of chaining) $\text{Ext}(X, Y, I) = \{B^{\downarrow} \mid B \in 2^Y\} = \{A^{\uparrow\downarrow} \mid A \in 2^X\}$; $\text{Int}(X, Y, I) = \{A^{\uparrow} \mid A \in 2^X\} = \{B^{\downarrow\uparrow} \mid B \in 2^Y\}$. Furthermore, $\mathcal{B}(X, Y, I) = \{\langle A, A^{\uparrow}\rangle \mid A \in \text{Ext}(X, Y, I)\} = \{\langle B^{\downarrow}, B\rangle \mid B \in \text{Int}(X, Y, I)\}$.

Proof (easy, viz prednasky)

Theorem (**from Galois connection to closure operator**) (1) If $\langle f, g\rangle$ is a Galois connection between X and Y then $C_X = f \circ g$ is a closure operator on X and $C_Y = g \circ f$ is a closure operator on Y.
(2) If $\langle f, g\rangle$ is induced by I, i.e. $\langle f, g\rangle = \langle {}^{\uparrow_I}, {}^{\downarrow_I}\rangle$, then $\mathcal{B}(X, Y, I)$ is isomorphic to $\langle \text{fix}(C_X), \subseteq\rangle$ and an isomorphism is given by sending $\langle A, B\rangle \in \mathcal{B}(X, Y, I)$ to $A \in \langle \text{fix}(C_X), \subseteq\rangle$. Moreover, $\mathcal{B}(X, Y, I)$ is dually isomorphic to $\langle \text{fix}(C_X), \subseteq\rangle$ and a dual isomorphism is given by sending $\langle A, B\rangle \in \mathcal{B}(X, Y, I)$ to $B \in \langle \text{fix}(C_Y), \subseteq\rangle$.

Proof (1) We show that $f \circ g : 2^X \to 2^X$ is a closure operator on X: (5) is $A \subseteq g(f(A))$ which is true by definition of a Galois connection.
(6): $A_1 \subseteq A_2$ impies $f(A_2) \subseteq f(A_1)$ which implies $g(f(A_1)) \subseteq g(f(A_2))$.
(7): Since $f(A) = f(g(f(A)))$, we get $g(f(A)) = g(f(g(f(A))))$.

(2) (viz prednasky) Follows immediately by definition of $\leq$ ($\langle A_1, B_2\rangle \leq \langle A_2, B_2\rangle$ iff $A_1 \subseteq A_2$ iff $B1 \supseteq B_2$) and by the above Corollary (in particular, by $\mathcal{B}(X, Y, I) = \{\langle A, A^{\uparrow}\rangle \mid A \in \text{Ext}(X, Y, I)\} = \{\langle B^{\downarrow}, B\rangle \mid B \in \text{Int}(X, Y, I)\}$). □

Remark We can see that for $\langle X, Y, I\rangle$, and $C_X = {}^{\uparrow\downarrow}$ and $C_Y = {}^{\downarrow\uparrow}$ we have $\text{Ext}(X, Y, I) = \text{fix}(C_X)$ and $\text{Int}(X, Y, I) = \text{fix}(C_Y)$.

Further issues in Galois connections etc.

(nebude pozadovano na zkousce)

Theorem (alternative definition of a Galois connection) $\langle f, g\rangle$ form a Galois connection between X and Y iff for each $A \subseteq X$ and $B \subseteq Y$ we have $A \subseteq B^{\downarrow}$ if and only if $B \subseteq A^{\uparrow}$.

Proof (easy, to be written)

Theorem (from closure operator to Galois connection) Let $C : 2^X \to 2^X$ be a closure operator in X. Define $I \subseteq X \times \text{fix}(C)$ by $\langle x, A\rangle \in I$ iff $x \in A$ for $x \in X$, $A \in \text{fix}(C)$. Then $\langle \text{fix}(C), \subseteq\rangle$ is isomorphic to $\mathcal{B}(X, \text{fix}(C), I)$.

Proof (to be written)

concept lattices in mathematics

- each complete lattice $(V, \leq)$ is isomorphic to some concept lattice, e.g. $(V, \leq) \cong \mathcal{B}(V, V, \leq)$;
- for partially ordered set $(V, \leq)$ ... $\mathcal{B}(V, V, \leq)$ is the MacNeille completion of $(V, \leq)$;

- V finitely dimensional vector space, V^* dual space, $a \perp \varphi$ means $\varphi(a) = 0$, then $\mathcal{B}(V, V^*, \perp)$ is the lattice of subspaces of V;
- ...

Main theorem of concept lattices

Theorem (Wille, 1982) (1) $\mathcal{B}(X,Y,I)$ is a complete lattice with infima and suprema given by

$$\bigwedge_{j\in J} \left\langle A_j, B_j \right\rangle = \langle \bigcap_{j\in J} A_j, (\bigcup_{j\in J} B_j)^{\downarrow\uparrow} \rangle \ , \bigvee_{j\in J} \left\langle A_j, B_j \right\rangle = \langle (\bigcup_{j\in J} A_j)^{\uparrow\downarrow}, \bigcap_{j\in J} B_j \rangle \ . \quad (11)$$

(2) Moreover, an arbitrary complete lattice $\mathbf{V} = (V, \leq)$ is isomorphic to $\mathcal{B}(X,Y,I)$ iff there are mappings $\gamma : X \to V$, $\mu : Y \to V$ such that

(i) $\gamma(X)$ is $\bigvee$-dense in V, $\mu(Y)$ is $\bigwedge$-dense in V;

(ii) $\gamma(x) \leq \mu(y)$ iff $(x,y) \in I$.

Proof (dukaz jen k casti (1); plyne z vyse uvedenych vysledku o Gal. konexich a uzav. operatorech, viz prednasky):

We check $\bigwedge_{j\in J} \left\langle A_j, B_j \right\rangle = \langle \bigcap_{j\in J} A_j, (\bigcup_{j\in J} B_j)^{\downarrow\uparrow} \rangle$: First, $\mathcal{B}(X,Y,I)$ is a complete lattice since it is isomorphic to a complete lattice $\text{Ext}(X,Y,I) = \text{fix}(^{\uparrow\downarrow})$ (and dually isomorphic to a complete lattice $\text{Int}(X,Y,I) = \text{fix}(^{\downarrow\uparrow})$). Moreover, infima in $\mathcal{B}(X,Y,I)$ correspond to infima in $\text{Ext}(X,Y,I)$ and to suprema in $\text{Int}(X,Y,I)$, from which we immediately get that the extent of $\bigwedge_{j\in J} \left\langle A_j, B_j \right\rangle$

is the infimum $\bigwedge_{j\in J} A_j$ of A_j's (taken in Ext(X,Y,I)) which is $\bigcap_{j\in J} A_j$, and that the intent of $\bigwedge_{j\in J} \left\langle A_j, B_j \right\rangle$ is the supremum $\bigvee_{j\in J} B_j$ of B_j's (taken in Int(X,Y,I)) which is $\bigcup_{j\in J}(B_j)^{\downarrow\uparrow}$.

Checking the formula for $\bigvee_{j\in J} \left\langle A_j, B_j \right\rangle$ is dual. □

Algorithms for concept lattices

Problem:

Input: (X, Y, I)

Output: $\mathcal{B}(X, Y, I)$ (possibly plus $\leq$)

Very good **survey and comparison** of algorithms:

Kuznetsov S. O., Obiedkov S. A.: Comparing performance of algorithms for generating concept lattices. *J. Experimental & Theoretical Artificial Intelligence* **14**(2003), 189–216.

- one of the first: Norris E. M.: An algorithm for computing the maximal rectangles of a binary relation. *J. ACM* **21**(1974), 356–366.
- often used 1: Ganter's NextClosure
- often used 2: Lindig's UpperNeighbor

NextClosure algorithm

suppose $X = \{1, \ldots, m\}$, $Y = \{1, \ldots, n\}$

for $A, B \subseteq Y$, $i \in \{1, \ldots, n\}$ put

$$A <_i B \quad \text{iff} \quad i \in B - A \text{ a } A \cap \{1, \ldots, i-1\} = B \cap \{1, \ldots, i-1\}.$$

and

$$A < B \quad \text{iff} \quad A <_i B \text{ for some } i.$$

$<$ ... **lexicographic ordering**

For $A \subseteq Y$, $i \in \{1, \ldots, n\}$, put

$$A \oplus i := ((A \cap \{1, \ldots, i-1\}) \cup \{i\})^{\downarrow\uparrow}.$$

Lemma The following assertions are true for any $B, D, D_1, D_2 \subseteq Y$:

(1) If $B <_i D_1$, $B <_j D_2$, and $i < j$ then $D_2 <_i D_1$;

(2) if $i \notin B$ then $B < B \oplus i$;

(3) if $B <_i D$ and $D = D^{\downarrow\uparrow}$ then $B \oplus i \subseteq D$;

(4) if $B <_i D$ and $D = D^{\downarrow\uparrow}$ then $B <_i B \oplus i$.

Proof (1) by easy inspection.

(2) is true because $B \cap \{1, \ldots, i-1\} \subseteq B \oplus i \cap \{1, \ldots, i-1\}$ and $i \in (B \oplus i) - B$.

(3) Putting $C_1 = B \cap \{1, \ldots, i-1\}$ and $C_2 = \{i\}$ we have $C_1 \cup C_2 \subseteq D$, and so $B \oplus i = (C_1 \cup C_2)^{\downarrow\uparrow} \subseteq D^{\downarrow\uparrow} = D$.

(4) By assumption, $B \cap \{1, \ldots, i-1\} = D \cap \{1, \ldots, i-1\}$. Furthermore, (3) yields $B \oplus i \subseteq D$ and so $B \cap \{1, \ldots, i-1\} \supseteq B \oplus i \cap \{1, \ldots, i-1\}$. On the other hand, $B \oplus i \cap \{1, \ldots, i-1\} \supseteq (B \cap \{1, \ldots, i-1\})^{\downarrow\uparrow} \cap \{1, \ldots, i-1\} \supseteq B \cap \{1, \ldots, i-1\}$. Therefore, $B \cap \{1, \ldots, i-1\} = B \oplus i \cap \{1, \ldots, i-1\}$. Finally, $i \in B \oplus i$ proving $B <_i B \oplus i$.

Theorem The least intent B^+ greater (w.r.t. $<$) than $B \subseteq Y$ is given by

$$B^+ = B \oplus i$$

where i is the greatest one with $B <_i B \oplus i$.

Proof Let B^+ be the least intent greater than B (w.r.t. to $<$). We have $B < B^+$ and thus $B <_i B^+$ for some i such that $i \in B^+$. By Lemma (4), $B <_i B \oplus i$, i.e. $B < B \oplus i$. Lemma (3) yields $B \oplus i \leq B^+$ which gives $B^+ = B \oplus i$ since B^+ is the least intent with $B < B^+$. It remains to show that i is the greatest one satisfying $B <_i B \oplus i$. Suppose $B <_k B \oplus k$ for $k > i$. By Lemma (1), $B \oplus k <_i B \oplus i$ which is a contradiction to $B \oplus i = B^+ < B \oplus k$ (B^+ is the least intent greater than B and so $B^+ < B \oplus k$). Therefore we have $k = i$.

```
NextClosure algorithm
 A:=leastIntent;
 store(A);
 while not(A=X) do
    A:=A+;
    store(A);
 endwhile.
```

complexity: time complexity of A^+ is $O(|X|^2 \cdot |Y|)$;
time complexity of NextClosure is $O(|X|^2 \cdot |Y| \cdot |\mathcal{B}(X, Y, I)|)$

⇒ **polynomial time delay complexity** (Johnson D. S., Yannakakis M., Papadimitrou C. H.: On generating all maximal independent sets. *Inf. Processing Letters* **27**(1988), 129–133.)

Note! Almost **no space requirements**. But: NextClosure does not directly give information about $\leq$.

UpperNeighbors algorithm

(nebude na zkousce pozadovan)

Idea:

start with the least formal concept $(\emptyset^{\uparrow\downarrow}, \emptyset^{\uparrow})$

for each (A, B) generate all its upper neighbors (and store the necessary information)

based on the following:

Thm. If $(A, B) \in \mathcal{B}(X, Y, I)$ is not the largest concept then $(A \cup \{x\})^{\uparrow\downarrow}$, with $x \in X - A$, is an extent of an upper neighbor of (A, B) iff for each $z \in (A \cup \{x\})^{\uparrow\downarrow} - A$ we have $(A \cup \{x\})^{\uparrow\downarrow} = (A \cup \{z\})^{\uparrow\downarrow}$.

UpperNeighbor procedure

```
min:=X − A;
neighbors:=∅;
for x ∈ X − A do
 B1 := (A ∪ {x})↑;  A1 := B1↓;
 if (min∩((A1 − A) − {x}) = ∅) then neighbors:=neighbors∪{(A1, B1)}
 else min:=min−{x};
enddo.
```

complexity polynomial time delay with delay $O(|X|^2 \cdot |Y|)$ (same as NextClosure)

Attribute implications

Def. **(Attribute) implication** (over attributes Y) is an expression $A \Rightarrow B$ where $A, B \subseteq Y$.

Why $A \Rightarrow B$? Primary reading: **"if object x has all attributes from A then x has all attributes from B"**

Denote $\mathrm{Imp} = \{A \Rightarrow B \mid A, B \subseteq Y\}$ (set of all attribute implications).

Def. $A \Rightarrow B$ **is true in C** $\subseteq Y$ if $A \subseteq C$ implies $B \subseteq C$;
denoted by $||A \Rightarrow B||_C = 1$ (or $C \models A \Rightarrow B$)

Def. **(Mod and Fml)** For a set $T \subseteq \mathrm{Imp}$ (set of attribute implications), $\mathcal{M} \subseteq 2^Y$ (set of sets of attributes), put

$$\begin{aligned} \mathrm{Mod}(T) &= \{C \in 2^Y \mid \text{for each } A \Rightarrow B \in T : ||A \Rightarrow B||_C = 1\}, \\ \mathrm{Fml}(\mathcal{M}) &= \{A \Rightarrow B \in \mathrm{Imp} \mid \text{for each } C \in \mathcal{M} : ||A \Rightarrow B||_C = 1\}. \end{aligned}$$

Rem. (1) $\mathrm{Mod}(T)$ ...models of T (all sets of attributes in which each implications from T are true); $\mathrm{Fml}(\mathcal{M})$ ...all implications true in (each set of attributes from) $\mathcal{M}$

(2) Put $\mathcal{X} = \text{Imp}$, $\mathcal{Y} = 2^Y$, define $\mathcal{I} \subseteq \mathcal{X} \times \mathcal{Y}$ by $\langle A \Rightarrow B, C \rangle \in \mathcal{I}$ iff $||A \Rightarrow B||_C = 1$. Then Mod and Fml form the Galois connection induced by $\langle \mathcal{X}, \mathcal{Y}, \mathcal{I} \rangle$. Therefore, we can use all properties of Galois connections for Mod and Fml.

(3) Mod and Fml . . . standard logical approach.

For $\mathcal{M} \subseteq 2^Y$ and $T = \{A_j \Rightarrow B_j \mid j \in J\}$:
$||T||_{\mathcal{M}} = 1$ (or $\mathcal{M} \models T$) iff for each $C \in \mathcal{M}$, $A \Rightarrow B \in T$: $||A \Rightarrow B||_C = 1$
(in words: T **is true in** $\mathcal{M}$)

Rem. Note: $||T||_{\mathcal{M}} = 1$ iff $\mathcal{M} \subseteq \text{Mod}(T)$ iff $T \subseteq \text{Fml}(\mathcal{M})$

Denote:

$$\begin{aligned} \text{Fml}(X, Y, I) \;&=\; \text{Fml}(\{\{x\}^{\uparrow} \mid x \in X\}) \ldots \textbf{implications true in data}, \\ &\quad\; (\{x\}^{\uparrow} \text{ is a row in table } \langle X, Y, I \rangle) \\ ||A \Rightarrow B||_{\langle X,Y,I \rangle} = 1 \;&\text{ iff }\; A \Rightarrow B \in \text{Fml}(X, Y, I) \end{aligned}$$

Sometimes: validity of $A \Rightarrow B$ in $\mathcal{B}(X, Y, I)$ means validity in $\text{Int}(X, Y, I)$.

Connection to predicate logic?

Rem. $\mathcal{M} \models A \Rightarrow B \ldots$ validity of a corresponding monadic formula $c(A \Rightarrow B)$ in a corresponding structure $c(\mathcal{M})$.

language given by unary relation symbols r_y $(y \in Y)$;
$A \Rightarrow B$ corresponds to **formula** $\varphi(A \Rightarrow B) = \&_{y \in A} r_y(x) \Rightarrow \&_{y \in B} r_y(x)$;
a set $\mathcal{M}$ of subsets of Y corresponds to **structure** $\mathbf{M}$ with support $M = \mathcal{M}$ in which
each r_y is interpreted by $r_y^{\mathbf{M}} = \{C \in \mathcal{M} \mid y \in C\}$.

Then:
$A \Rightarrow B$ is true in $\mathcal{M}$ (in the above sense) iff $\varphi(A \Rightarrow B)$ is true in $\mathbf{M}$ (in the standard sense of predicate logic).

Basic connection to FCA

Thm. $A \Rightarrow B$ is true in (X, Y, I) IFF $A \Rightarrow B$ is true in Int(X, Y, I) IFF $B \subseteq A^{\downarrow\uparrow}$ IFF $A^{\downarrow} \subseteq B^{\downarrow}$.

Proof nontrivial part is "if $A \Rightarrow B$ is true in (X, Y, I) then $A \Rightarrow B$ is true in $\mathcal{B}(X, Y, I)$": Let $A \Rightarrow B$ be true in (X, Y, I), i.e. $A^{\downarrow} \subseteq B^{\downarrow}$. Suppose $A \subseteq D$ for $\langle C, D\rangle \in \mathcal{B}(X, Y, I)$, i.e. $A \subseteq C^{\uparrow}$. This is equivalent to $C \subseteq A\downarrow$. Therefore $C \subseteq B^{\downarrow}$, which is equivalent to $B \subseteq C^{\uparrow} = D$, proving $A \Rightarrow B$ is true in $\mathcal{B}(X, Y, I)$.

Entailment, base

Def. $A \Rightarrow B$ **(semantically) follows** from a set T of implications ($T \models A \Rightarrow B$) if $A \Rightarrow B$ is true in each $C \subseteq Y$ which is a model of T, i.e.

$$T \models A \Rightarrow B \quad \textbf{iff} \quad A \Rightarrow B \in \text{Fml}(\text{Mod}(T)).$$

Meaning: $T \models A \Rightarrow B \ldots A \Rightarrow B$ is true whenever each $A_i \Rightarrow B_i \in T$ is true.

$T \subseteq \text{Imp}$ is called

- **closed** if it contains each implication which follows from T, i.e. $T = \text{FmlMod}(T)$,
- **non-redundant** if no implication from T follows from the rest (i.e. $T - \{A \Rightarrow B\} \not\models A \Rightarrow B$),
- **complete w.r.t.** $\langle X, Y, I\rangle$ if T is true in $\langle X, Y, I\rangle$ and each implication true in $\langle X, Y, I\rangle$ follows from T,
- **base w.r.t.** $\langle X, Y, I\rangle$ if it is complete w.r.t. $\langle X, Y, I\rangle$ and non-redundant.

Why base? To have less implications which carry the same information.

Lemma For $T \subseteq$ Imp:

1. T is true in $\langle X, Y, I\rangle$ IFF Mod$(T) \supseteq$ Int(X, Y, I),

2. each implication true in $\langle X, Y, I\rangle$ follows from T IFF Mod$(T) \subseteq$ Int(X, Y, I).

Proof "1.": T is true in $\langle X, Y, I\rangle$ IFF (by def.) $T \subseteq$ Fml(Int(X, Y, I)) IFF (by properties of Gal. conn.) Mod$(T) \supseteq$ Int(X, Y, I).

"2.": First, show Claim: ModFml(Int(X, Y, I)) = Int(X, Y, I).
Proof of Claim: "$\supseteq$" by properties of Gal. conn; "$\subseteq$": Let $A \in$ ModFml(Int$(X, Y, I$
Then $A \Rightarrow A^{\downarrow\uparrow} \in$ Fml(Int(X, Y, I)) (indeed: for $B \in$ Int(X, Y, I), we have: if $A \subseteq B$ then $A^{\downarrow\uparrow} \subseteq B^{\downarrow\uparrow} = B$, i.e. $||A \Rightarrow A^{\downarrow\uparrow}||_B = 1$). Thus, in particular, $||A \Rightarrow A^{\downarrow\uparrow}||_A = 1$ which means that if $A \subseteq A$ (which is true) then $A^{\downarrow\uparrow} \subseteq A$ which means $A \in$ Int(X, Y, I).

Second, each implication true in $\langle X, Y, I\rangle$ follows from T IFF (by def.) Fml(X, Y, I)
Fml(Mod(T)) IFF (by Fml(X, Y, I) = Fml(Int(X, Y, I))) Fml(Int(X, Y, I)) $\subseteq$ Fml(Mod(T)) IFF (by prop. of Gal. conn.) ModFml(Int(X, Y, I)) $\supseteq$ Mod(Fml(Mo
Mod(T) IFF (by Claim) Mod$(T) \subseteq$ Int(X, Y, I).

Corollary T is complete w.r.t. $\langle X, Y, I\rangle$ IFF Mod(T) = Int(X, Y, I).

Rules of entailment

Some rules of entailment (deduction):
$A \Rightarrow A$ is always true,
if $A \Rightarrow B$ and $B \Rightarrow C$ are true then $A \Rightarrow C$ is true (transitivity),
if $A \Rightarrow B$ is true and $B' \subseteq B$ then $A \Rightarrow B'$ is true (projectivity),
...

Is there a small set of simple rules for obtaining all consequences of a set T of attribute implications?

A consequence of theorem from relational databases (**caution!**, different notions, the same concept of entailment, Maier D.: The Theory of Relational Databases, Computer Science Press, 1983):

Thm. T is closed iff for each $A, B, C, D \subseteq Y$ we have

1. $A \Rightarrow A \in T$;
2. if $A \Rightarrow B \in T$ then $A \cup C \Rightarrow B \in T$;
3. if $A \Rightarrow B \in T$ and $B \cup C \Rightarrow D \in T$ then $A \cup C \Rightarrow D \in T$.

Proof (direct) "$\Rightarrow$" easy.
"$\Leftarrow$": Denote X^+ the largest X such that $X \Rightarrow X^+ \in T$ (this is correct: from $X \Rightarrow Y, X \Rightarrow Z \in T$ we get $X \Rightarrow Y \cup Z \in T$, SHOW using 1.–3.) Assume 1.–3. Let $T \vdash A \Rightarrow B$ mean that $A \Rightarrow B$ can be obtained from T using rules encoded in 1.–3. It is sufficient to show that if $T \models A \Rightarrow B$ then $T \vdash A \Rightarrow B$ (since then $A \Rightarrow B \in T$). By contradiction, assume $T \nvdash A \Rightarrow B$. We need $T \not\models A \Rightarrow B$, i.e. we need a set which is a model of T but not of $A \Rightarrow B$. We show that A^+ is such a set.
First, $A^+ \not\models A \Rightarrow B$: Clearly, $A \subseteq A^+$. We cannot have $B \subseteq A^+$ since then from $A \Rightarrow A^+ \in T$ we get (using 1.–3.) $A \Rightarrow B \in T$, a contradiction to $T \nvdash A \Rightarrow B$.
Second, we show that for each $C \Rightarrow D \in T$, $A^+ \models C \Rightarrow D$: Suppose $C \subseteq A^+$. We get $A^+ \Rightarrow C \in T$ (using $A^+ \Rightarrow A^+$ and projectivity which follows from 1.–3.). So we have $A \Rightarrow A^+, A^+ \Rightarrow C, C \Rightarrow D \in T$ and transitivity (follows from 1.–3.) gives $A \Rightarrow D \in T$, i.e. $D \subseteq A^+$.

Note (exercise): verify that using 1.–3. we have:
projectivity: $A \Rightarrow B \in T$, $C \subseteq B$ imply $A \Rightarrow C \in T$
transitivity: $A \Rightarrow B, B \Rightarrow C \in T$ imply $A \Rightarrow C \in T$

Pseudointents and Guigues-Duquenne base

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

Recall:
(1) A closure system is a sytem closed under arbitrary intersections.
(2) Closure systems vs. closure operators:

A **closure system** on a set X is a nonempty system $\mathcal{S} \subseteq 2^X$ which is closed under arbitrary intersections and contains X.

This means: the intersection of any members of $\mathcal{S}$ belongs to $\mathcal{S}$ (for any system $\{A_j \mid j \in J\} \subseteq \mathcal{S}$, $\cap_j A_j \in \mathcal{S}$); and $X \in \mathcal{S}$.

There is a one-to-one relationship between closure systems on X and closure operators on X. Given a closure operator C on X, $\mathcal{S}_C = \{A \in 2^X \mid A = C(A)\} = \text{fix}(C)$ is a closure system. Given a closure system on X, putting

$$C_{\mathcal{S}}(A) = \bigcap\{B \in \mathcal{S} \mid A \subseteq B\}$$

for any $A \in 2^X$, $C_{\mathcal{S}}$ is a closure operator on X. This is a one-to-one relationship, i.e. $C = C_{\mathcal{S}_C}$ and $\mathcal{S} = \mathcal{S}_{C_{\mathcal{S}}}$.

Lemma For a set T of attribute implications, $\text{Mod}(T) = \{A \subseteq Y \mid A \models T\}$ is a closure system.

Proof (1) $\text{Mod}(T) \neq \emptyset$ since $Y \in \text{Mod}(T)$.

(2) Let $C_j \in \text{Mod}(T)$ $(j \in J)$. For any $A \Rightarrow B \in T$, if $A \subseteq \bigcap_j C_j$ then for each $j \in J$: $A \subseteq C_j$, and so $B \subseteq C_j$ (since $C_j \in \text{Mod}(T)$, thus in particular $C_j \models A \Rightarrow B$), from which we have $B \subseteq \bigcap_j C_j$.

We showed that $\text{Mod}(T)$ is nonempty and is closed under intersections, i.e. $\text{Mod}(T)$ is a closure system.

Def. **Pseudointent** of (X, Y, I) is a subset $A \subseteq Y$ for which $A \neq A^{\downarrow\uparrow}$ and $B^{\downarrow\uparrow} \subseteq A$ for each pseudointent $B \subset A$.

Thm. (Guigues-Duquenne basis, stem basis)
The set $T = \{A \Rightarrow A^{\downarrow\uparrow} \mid A \text{ is a pseudointent of } (X, Y, I)\}$ of implications is a basis.

Proof We show that T is complete and non-redundant.

Complete: It suffices to show that $\text{Mod}(T) \subseteq \text{Int}(X, Y, I)$. Let $C \in \text{Mod}(T)$. Assume $C \neq C^{\downarrow\uparrow}$. Then C is a pseudointent (indeed, if $P \subset C$ is a pseudointent then since $||P \Rightarrow P^{\downarrow\uparrow}||_C = 1$, we get $P^{\downarrow\uparrow} \subseteq C$). But then $C \Rightarrow C^{\downarrow\uparrow} \in T$ and so $||C \Rightarrow C^{\downarrow\uparrow}||_C = 1$. But the last fact means that if $C \subseteq C$ (which is true) then $C^{\downarrow\uparrow} \subseteq C$ which would give $C^{\downarrow\uparrow} = C$, a contradiction with the assumption $C^{\downarrow\uparrow} \neq C$. Therefore, $C^{\downarrow\uparrow} = C$, i.e. $C \in \text{Int}(X, Y, I)$.

Non-redundant: Take any $P \Rightarrow P^{\downarrow\uparrow}$. We show that $T - \{P \Rightarrow P^{\downarrow\uparrow}\} \not\models P \Rightarrow P^{\downarrow\uparrow}$. Since $||P \Rightarrow P^{\downarrow\uparrow}||_P = 0$ (obvious, check), it suffices to show that $||T - \{P \Rightarrow P^{\downarrow\uparrow}\}||_P = 1$. That is, we need to show that for each $Q \Rightarrow Q^{\downarrow\uparrow} \in T - \{P \Rightarrow P^{\downarrow\uparrow}\}$ we have $||Q \Rightarrow Q^{\downarrow\uparrow}||_P = 1$, i.e. that if $Q \subseteq P$ then $Q^{\downarrow\uparrow} \subseteq P$. But this follows from the definition of a pseudointent (applt to P).

Lemma If P, Q are intents or pseudointents and $P \not\subseteq Q$, $Q \not\subseteq P$, then $P \cap Q$ is an intent.

Proof Let $T = \{R \Rightarrow R^{\downarrow\uparrow} \mid R \text{ a pseudointent}\}$ be the G.-D. basis. Since T is complete, it is sufficient to show that $P \cap Q \in \text{Mod}(T)$ (since then, $P \cap Q$ is a model of any implication which is true in $\langle X, Y, I \rangle$, and so $P \cap Q$ is an intent).

Obviously, P, Q are models of $T - \{P \Rightarrow P^{\downarrow\uparrow}, Q \Rightarrow Q^{\downarrow\uparrow}\}$, whence $P \cap Q$ is a model of $T - \{P \Rightarrow P^{\downarrow\uparrow}, Q \Rightarrow Q^{\downarrow\uparrow}\}$ (since the set of models is a closure system, i.e. closed under intersections).

Therefore, to show that $P \cap Q$ is a model of T, it is sufficient to show that $P \cap Q$ is a model of $\{P \Rightarrow P^{\downarrow\uparrow}, Q \Rightarrow Q^{\downarrow\uparrow}\}$. Due to symmetry, we only verify that $P \cap Q$ is a model of $\{P \Rightarrow P^{\downarrow\uparrow}$: But this is trivial: since $P \not\subseteq Q$, the condition "if $P \subseteq P \cap Q$ implies $P^{\downarrow\uparrow} \subseteq P \cap Q$" is satisfied for free. The proof is complete.

Lemma If T is complete, then for each pseudointent P, T contains $A \Rightarrow B$ with $A^{\downarrow\uparrow} = P^{\downarrow\uparrow}$

Proof For pseudointent P, $P \neq P^{\downarrow\uparrow}$, i.e. P is not an intent. Therefore, P cannot be a model of T (since models of a complete T are intents). Therefore, there is $A \Rightarrow B \in T$ such that $||A \Rightarrow B||_P = 0$, i.e. $A \subseteq P$ but $B \not\subseteq P$. As $||A \Rightarrow B||_{\langle X,Y,I\rangle} = 1$, we have $B \subseteq A^{\downarrow\uparrow}$ (Thm. on basic connections ...). Therefore, $A^{\downarrow\uparrow} \not\subseteq P$ (otherwise $B \subseteq P$, a contradiction). Therefore, $A^{\downarrow\uparrow} \cap P$ is not an intent (). By the foregoing Lemma, $P \subseteq A^{\downarrow\uparrow}$ which gives $P^{\downarrow\uparrow} \subseteq A^{\downarrow\uparrow}$. On the other hand, $A \subseteq P$ gives $A^{\downarrow\uparrow} \subseteq P^{\downarrow\uparrow}$. Altogether, $A^{\downarrow\uparrow} = P^{\downarrow\uparrow}$, proving the claim.

Thm. (Guigues-Duquenne base is smalest)
If T is the Guigues-Duquenne base and T' is complete then $|T| \leq |T'|$.

Proof Direct corollary of the above Lemma.

Computing Guigues-Duquenne base

$\mathcal{P}$ ... set of all pseudointents of $\langle X, Y, I \rangle$

THE base: $\{A \Rightarrow A^{\downarrow\uparrow} \mid A \in \mathcal{P}\}$

Q: What do we need? A: Compute all pseudointents.

Lemma The set of all P which are intents or pseudointents is a closure system.

Q: How to compute the fixed points (closed sets)?

For $Z \subseteq Y$, T a set of implications, put

$Z^T = Z \cup \bigcup\{B \mid A \Rightarrow B \in T, A \subset Z\}$

$Z^{T_0} = Z$
$Z^{T_n} = (Z^{T_{n-1}})^T \ (n \geq 1)$

define $C_T : 2^Y \rightarrow 2^Y$ by

$C_T(Z) = \bigcup_{n=0}^{\infty} Z^{T_n}$ (note: terminates, Y finite)

Thm. Let $T = \{A \Rightarrow A^{\downarrow\uparrow} \mid A \in \mathcal{P}\}$ (G.-D. base). Then
(1) C_T is a closure operator,
(2) P is a fixed point of C_T iff $P \in \mathcal{P}$ (pseudointent) or $P \in \text{Int}(X, Y, I)$ (intent).

Proof (1) easy

(2) $\mathcal{P} \cup \text{Int}(X, Y, I) \subseteq \text{fix}(C_T)$ easy. $\text{fix}(C_T) \subseteq \mathcal{P} \cup \text{Int}(X, Y, I)$: It suffices to show that if $P \in \text{fix}(C_T)$ is not an intent ($P \neq P^{\downarrow\uparrow}$) then P is an pseudointent. So take $P \in \text{fix}(C_T)$, i.e. $P = C_T(P)$, which is not an intent. Take any pseudointent $Q \subset P$. By definition (notice that $Q \Rightarrow Q^{\downarrow\uparrow} \in T$), $Q^{\downarrow\uparrow} \subseteq C_T(P) = P$ which means that P is a pseudointent. The proof is complete.

So: $\text{fix}(C_T) = \mathcal{P} \cup \text{Int}(X, Y, I)$

Intention: compute $\mathcal{P}$ by computing $\text{fix}(C_T)$ and excluding $\text{Int}(X, Y, I)$.

Computing $\text{fix}(C_T)$ by Ganter's next closure algorithm.

Caution! In order to compute C_T, we need T, i.e. we need $\mathcal{P}$, which **we do not know in advance.**

But we are not in *circulus vitiosus*: The part of T (or $\mathcal{P}$) which is needed is already available (computed).

Conceptual scaling

(na zkousce nebude pozadovano)

= way to deal with data tables with more general attributes (nominal, ordinal)

transformation (scaling) of general data table to a suitable formal context (only binary attributes)

For details see

B. Ganter, R. Wille: Formal Concept Analysis: Mathematical Foundations. Springer, 1999.

Selected applications

Software engineering

- G. Snelting: Reengineering of configurations based on mathematical concept analysis. *ACM Trans. Software Eng. Method.* 5(2):146–189, April 1996.
- G. Snelting, F. Tip: Understanding class hierarchies using concept analysis. *ACM Trans. Program. Lang. Syst.* 22(3):540–582, May 2000.
- U. Dekel, Y. Gill. Visualizing class interfaces with formal concept analysis. In *ACM OOPSLA'03 Conference*, pages 288–289, Anaheim, CA, October 2003.
- G. Ammons, D. Mandelin, R. Bodik, J. R. Larus. Debugging temporal specifications with concept analysis. In *Proc. ACM SIGPLAN'03 Conference on Programming Language Design and Implementation*, pages 182–195, San Diego, CA, June 2003.

Database views

- C. Carpineto, R. Romano: A lattice conceptual clustering system and its application to browsing retrieval. *Machine Learning* 24:95–122, 1996.
- Snášel *et al.*: Navigation through query result.

Analysis of texts (medical records, e-mails)

- R. Cole, P. Eklund: Scalability in formal context analysis: a case study using medical texts. *Computational Intelligence* 15:11–27, 1999.
- R. Cole: Analyzing e-mail collections using formal concept analysis (preprint).

Software support

- Toscana, Anaconda, . . .
- SW developed jointly by Dept. Comp. Sci., Palacký University, Olomouc and Dept. Comp. Sci., Technical University of Ostrava (public, to be released)

FCA of data with fuzzy attributes = fuzzy concept lattices

Motivation

- Fuzzy attributes ... expensive, small, etc.
- Concepts are fuzzy

Fuzzy sets and fuzzy logic

- scale of truth degrees (e.g. $[0, 1]$)
- logic: Hájek P.: Metamathematics of Fuzzy Logic. Kluwer, 1998.
- relational systems: Bělohlávek R.: Fuzzy Relational Systems: Foundations and Principles. Kluwer, 2002.

Pursued by Burusco, Fuentes-Gonzales, Pollandt, Bělohlávek *et al.*, ...

Basics from fuzzy logic

- structure of truth degrees: **complete residuated lattice**

  $\mathbf{L} = \langle L, \vee, \wedge, \otimes, \rightarrow, 0, 1\rangle$, where $\langle L, \vee, \wedge, 0, 1\rangle$ ... complete lattice,
  $\langle L, \otimes, 1\rangle$ ... commutative monoid,
  $\langle \otimes, \rightarrow\rangle$ ... **adjoint pair** (i.e. $x \leq y \rightarrow z$ iff $x \otimes y \leq z$)

  e.g. L is a finite subchain of $[0, 1]$, $\otimes$ ... left-continuous t-norm, (Gödel, Łukasiewicz) and

$$x \rightarrow y = \bigvee_{z \in L} \{z \mid x \otimes z \leq y\}$$

- example 1 (Łukasiewicz): $a \otimes b = \max(0, a + b - 1) + \rightarrow$

- example 2 (G¨odel): $a \otimes b = \min(a, b) + \rightarrow$

- **fuzzy set (L-set)** A in X ... $A\colon X \rightarrow L$
  $A(x)$ ... the truth degree of "x belongs to A"
  fuzzy relation I between X and Y: ... $I\colon X \times Y \rightarrow L$
  $I(x, y)$ ... the truth degree of "x is in relation to y"

- $\mathbf{A} \subseteq \mathbf{B}$ if $A(x) \leq B(x)$ for each $x \in X$
  more generally: **graded subsethood** between $\mathbf{L}$-sets

$$S(A, B) = \bigwedge_{x \in x} A(x) \rightarrow B(x)$$

Formal fuzzy context = input data

Def. **Formal fuzzy context** is a triplet (X, Y, I) where
X . . . set of **objects**
Y . . . set of **attributes**
$I : X \times Y \to L$ binary fuzzy relation.

Interpretation: $I(x, y)$ **. . . degree to which object** x **has attribute** y

formal fuzzy context $\approx$ **data table**

I	y_1	y_2	y_3	y_4
x_1	1	1	0	0.5
x_2	0.8	0.1	0	0.9
x_3	1	0.9	0.9	0
x_4	1	0.5	0.6	0.5
x_5	1	0	0	0.5

Formal fuzzy concept = fuzzy cluster in data

Def. **Induced operators** . . . mappings $^{\uparrow}: L^X \to L^Y$, $^{\uparrow}: L^X \to L^Y$ def. by:

$$A^{\uparrow}(y) = \bigwedge_{x \in X} A(x) \to I(x, y)$$

$$B^{\downarrow}(x) = \bigwedge_{y \in Y} B(y) \to I(x, y)$$

$A^{\uparrow}$ . . . fuzzy set of attributes common to all objects from A

$B^{\downarrow}$ . . . fuzzy set objects sharing all attributes from A

Def. **Formal fuzzy concept** in (X, Y, I) . . . (A, B), $A \in L^X$, $B \in L^Y$, s.t.

$$A^{\uparrow} = B \text{ and } B^{\downarrow} = A.$$

A . . . **extent** . . . objects covered by formal concept

B . . . **intent** . . . attributes covered by formal concept

- **(fuzzy) concept lattice** given by $\langle X, Y, I \rangle$

$$\mathcal{B}(X, Y, I) = \{\langle A, B \rangle \mid A^{\uparrow} = B,\ B^{\downarrow} = A\}$$

- **subconcept-superconcept hierarchy** $\leq$ in $\mathcal{B}(X, Y, I)$

$$\langle A_1, B_1 \rangle \leq \langle A_1, B_1 \rangle \text{ iff } A_1 \subseteq A_2 (\text{ iff } B_1 \supseteq B_2)$$

Further info:

Chapter 5 of R.B.: Fuzzy Relational Systems. Kluwer, New York, 2002.

Main theorem of fuzzy concept lattices

Several issues from bivalent case can be carried over to fuzzy setting. Examples: algorithms, the main theorem:

Theorem (1) $\mathcal{B}(X,Y,I)$ is a **completely lattice-type fuzzy ordered set** with infima and suprema given by

$$\bigwedge_{j\in J} \left\langle A_j, B_j \right\rangle = \langle \bigcap_{j\in J} A_j, (\bigcup_{j\in J} B_j)^{\downarrow\uparrow} \rangle \ , \bigvee_{j\in J} \left\langle A_j, B_j \right\rangle = \langle (\bigcup_{j\in J} A_j)^{\uparrow\downarrow}, \bigcap_{j\in J} B_j \rangle \ .$$

(2) Moreover, an arbitrary completely lattice-type fuzzy ordered set $\mathbf{V} = (V, \preceq)$ is isomorphic to $\mathcal{B}(X,Y,I)$ iff there are mappings $\gamma : X \times L \to V$, $\mu : Y \times L \to V$ such that

(i) $\gamma(X \times L)$ is $\bigvee$-dense in V, $\mu(Y \times L)$ is $\bigwedge$-dense in V;

(ii) $(\gamma(x,a) \leq \mu(y,b)) = (a \otimes b) \to I(x,y)$.

Non-standard issues

(ke zkousce jen prehledove)

In fuzzy setting, there arise new phenomena which are degenerate in bivalent setting. As an example, we present fatorization by similarity.

Similarity relation

Degree of similarity $\approx$ of $\langle A_1, B_1\rangle$ and $\langle A_2, B_2\rangle$ on $\mathcal{B}(X, Y, I)$

$$\langle A_1, B_1\rangle \approx \langle A_2, B_2\rangle = \bigwedge_{x\in X} A_1(x) \leftrightarrow A_2(x) \ (= \bigwedge_{y\in Y} B_1(y) \leftrightarrow B_2(y))$$

Given a truth degree $a \in L$ (a **threshold** specified by a user), the thresholded relation (a-cut) ${}^a{\approx}$ on $\mathcal{B}(X, Y, I)$ defined by

$$(\langle A_1, B_1\rangle, \langle A_2, B_2\rangle) \in {}^a{\approx} \text{ iff } (\langle A_1, B_1\rangle \approx \langle A_2, B_2\rangle) \geq a$$

denotes "being **similar in degree at least a**".
${}^a{\approx}$ is reflexive and symmetric, but need not be transitive.

A subset B of $\mathcal{B}(X, Y, I)$ is a **${}^a{\approx}$-block** if it is a maximal subset of $\mathcal{B}(X, Y, I)$ such that each two concepts from B are similar in degree at least a.

$\mathcal{B}(\mathbf{X}, \mathbf{Y}, \mathbf{I})/{}^{\mathbf{a}}{\approx}$ . . . the collection of all ${}^a{\approx}$-blocks.

Factorization by similarity

Put

$$\langle A, B\rangle_a := \bigwedge\{\langle A', B'\rangle \mid (\langle A, B\rangle, \langle A', B'\rangle) \in {}^a\!\approx\}$$
$$\langle A, B\rangle^a := \bigvee\{\langle A', B'\rangle \mid (\langle A, B\rangle, \langle A', B'\rangle) \in {}^a\!\approx\}.$$

Lemma ${}^a\!\approx$-blocks are exactly intervals of $\mathcal{B}(X, Y, I)$ of the form $[\langle A, B\rangle_a, (\langle A, B\rangle_a)^a]$, i.e.

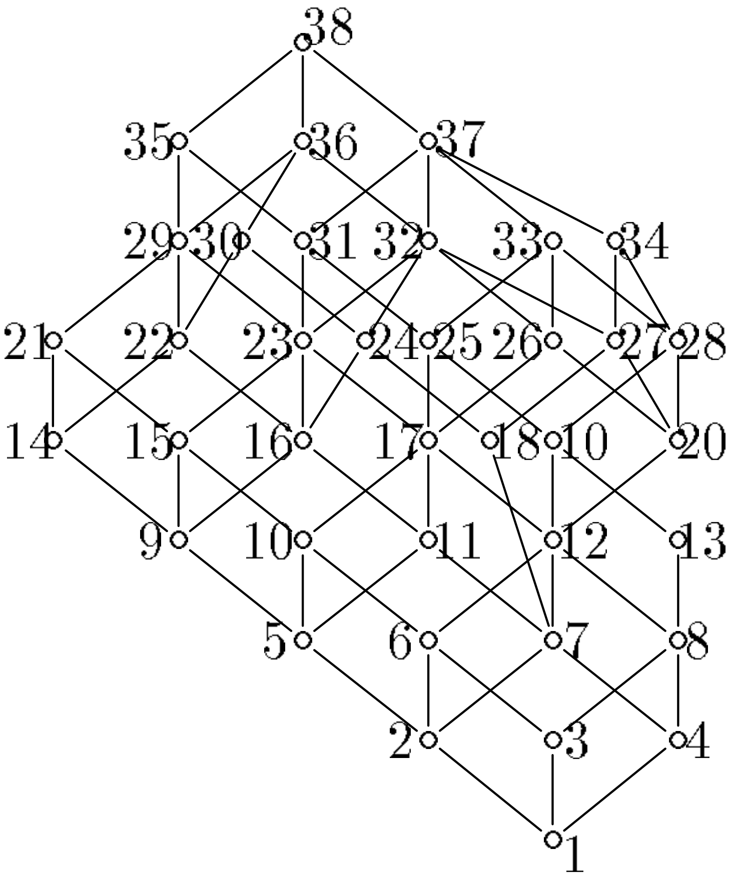
$$\mathcal{B}(X, Y, I)/{}^a\!\approx \;=\; \{[\langle A, B\rangle_a, (\langle A, B\rangle_a)^a] \mid \langle A, B\rangle \in \mathcal{B}(X, Y, I)\}.$$

Define a partial order $\preceq$ on blocks of $\mathcal{B}(X, Y, I)/{}^a\!\approx$ by $[c_1, c_2] \preceq [d_1, d_2]$ iff $c_1 \leq d_1$ (iff $c_2 \leq d_2$), where $[c_1, c_2], [d_1, d_2] \in \mathcal{B}(X, Y, I)/{}^a\!\approx$.

Theorem $\mathcal{B}(X, Y, I)/{}^a\!\approx$ equipped with $\preceq$ is a partially ordered set which is a complete lattice, the so-called
factor lattice of $\mathcal{B}(\mathrm{X}, \mathrm{Y}, \mathrm{I})$ by similarity $\approx$ and a threshold a.

Elements of $\mathcal{B}(X, Y, I)/{}^a\!\approx$ can be seen as **similarity-based granules** of formal concepts/clusters from $\mathcal{B}(X, Y, I)$.

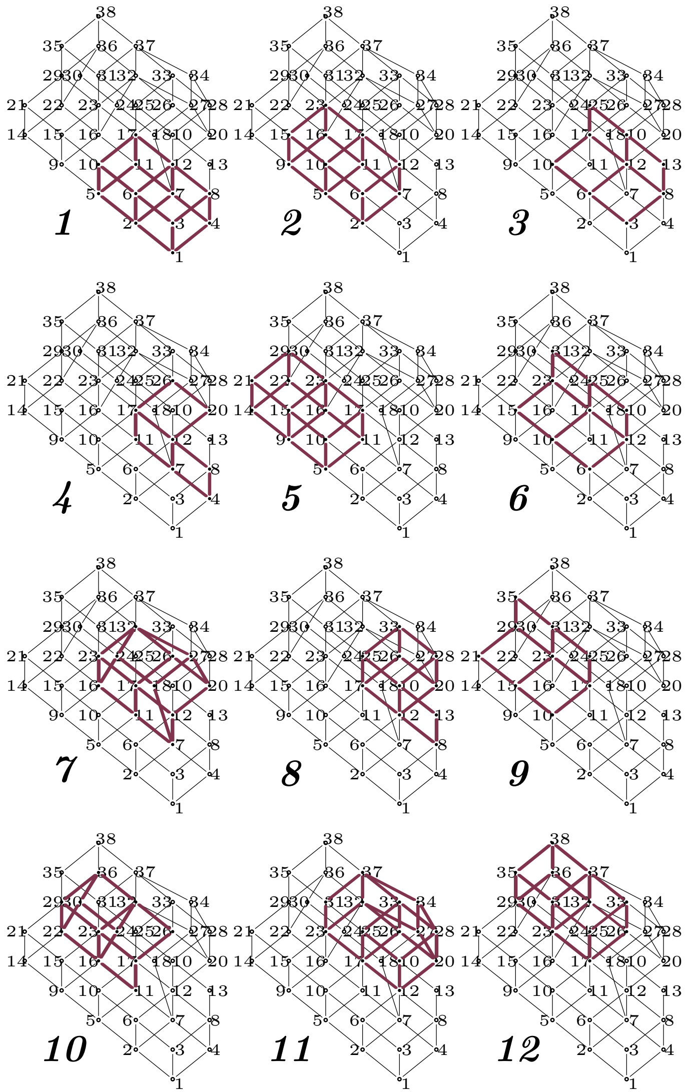
Factorization by similarity: example



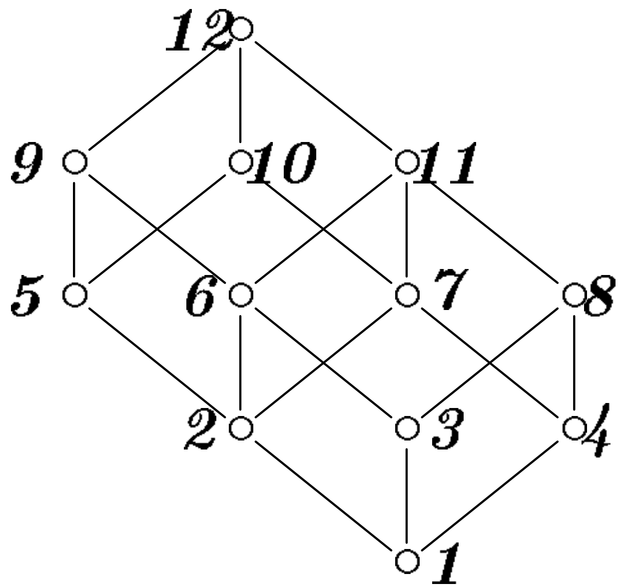
TOO LARGE!

Can we have cluters of 0.5-similar formal concepts instead?

Factorization by similarity: example



$^{a}\!\approx$-blocks



Factor lattice $\mathcal{B}(X,Y,I)/^{a}\!\approx$

Factorization directly from input data

Problem: Computation of $\mathcal{B}(X,Y,I)/{}^a\!\approx$ by definition is time demanding, can it be computed directly from input data?

Solution: It will turn out that our algorithm has a polynomial time delay and is much faster.

Some definitions: $(a \to C)(x) = a \to C(x) \qquad (a \otimes C)(x) = a \otimes C(x)$

Lemma If A is an extent then so is $a \to A$, similarly for intents.

FIRST, $\langle A,B\rangle_a$ **and** $\langle A,B\rangle^a$ **can be computed directly from** $\langle A,B\rangle$**:**

Lemma For $\langle A,B\rangle \in \mathcal{B}(X,Y,I)$, we have
(a) $\langle A,B\rangle_a = \langle (a \otimes A)^{\uparrow\downarrow}, a \to B\rangle$ (b) $\langle A,B\rangle^a = \langle (a \to A), (a \otimes B)^{\downarrow\uparrow}\rangle$.

Thus we have $(\langle A,B\rangle_a)^a = \langle a \to (a \otimes A)^{\uparrow\downarrow}, (a \otimes (a \to B))^{\downarrow\uparrow}\rangle$.

Lemma For $\langle A,B\rangle \in \mathcal{B}(X,Y,I)$ we have $\langle A,B\rangle_a = ((\langle A,B\rangle_a)^a)_a$.

SECOND, by Lemma **??** ${}^a\!\approx$**-blocks** $[c_1,c_2]$ **are uniquely given** by their suprema c_2, moreover, **by extents of suprema**, since each formal concept is uniquely given by its extent.

Factorization directly from input data: main result

Denote the set of all extents of suprema of ${}^a\!\approx$-blocks by $\mathbf{ESB(a)}$, i.e.

$$\mathrm{ESB}(a) = \{A \in L^X \mid \langle A, B\rangle \in \mathcal{B}(X,Y,I), [\langle A,B\rangle_a, \langle A,B\rangle] \in \mathcal{B}(X,Y,I)/{}^a\!\approx\}$$

Recall:

C is called a **fuzzy closure operator** in X if $A \subseteq C(A)$, $S(A_1, A_2) \leq S(C(A_1), C(A_2$
and $C(A) = C(C(A))$, for any $A, A_1, A_2 \in L^X$.
Fixed point of $C : L^X \to L^X$: fuzzy set A such that $A = C(A)$.
$\mathbf{fix(C)} = \{A \in L^X \mid A = C(A)\}$ . . . set of all fixed points of C.

Theorem Given input data $\langle X, Y, I\rangle$ and a threshold $a \in L$, a mapping

$$\mathbf{C_a} : \mathbf{A} \;\mapsto\; \mathbf{a} \to (\mathbf{a} \otimes \mathbf{A})^{\uparrow\downarrow}$$

is a fuzzy closure operator in X for which $\mathbf{fix(C_a) = ESB(a)}$.

Problem: How to generate $\mathrm{fix}(C_a) = \mathrm{ESB}(a)$?

Solution: **fuzzy adaptation of Ganter's algorithm** (R.B., 2002) for generating all formal concepts of a given fuzzy context, which is in fact an algorithm for generating the set of all fixed points of a fuzzy closure operator.

Factorization directly from input data: algorithm

Suppose $X = \{1, 2, \ldots, n\}$ and $L = \{0 = a_1 < a_2 < \cdots < a_k = 1\}$.
Put $(i, j) \leq (r, s)$ iff $i < r$ or $i = r, a_j \geq a_s$, for $i, r \in \{1, \ldots, n\}$, $j, s \in \{1, \ldots, k\}$.

In the following, we will freely refer to a_i just by i, i.e. we denote $(i, a_j) \in X \times L$ also simply by (i, j).

Put

$$\mathbf{A} \oplus (\mathbf{i}, \mathbf{j}) := C_a((A \cap \{1, 2, \ldots, i-1\}) \cup \{\, {}^{a_j}/i\})$$

and

$$\mathbf{A} <_{(\mathbf{i}, \mathbf{j})} \mathbf{C} \text{ iff } A \cap \{1, \ldots, i-1\} = C \cap \{1, \ldots, i-1\} \text{ and } A(i) < C(i) = a_j.$$

Finally, $\mathbf{A} < \mathbf{C}$ iff $A <_{(i,j)} C$ for some (i, j).

Lemma The least fixed point A^+ which is greater (w.r.t. $<$) than a given $A \in L^X$ is given by $A^+ = A \oplus (i, j)$ where (i, j) is the greatest one with $A <_{(i,j)} A \oplus (i, j)$.

Factorization directly from input data: algorithm

The algorithm for generating ${}^{a}\!\approx$-blocks:

INPUT: $\langle X, Y, I\rangle$ (data table with fuzzy attributes),

$a \in L$ (similarity threshold)

OUTPUT: $\mathcal{B}(X, Y, I)/{}^{a}\!\approx$ (${}^{a}\!\approx$-blocks $[c_1, c_2]$)

$A := \emptyset$

while $A \neq X$ **do**

$A := A^{+}$

store($[\langle(\mathbf{a} \otimes \mathbf{A})^{\uparrow\downarrow}, \mathbf{a} \rightarrow \mathbf{A}^{\uparrow}\rangle, \langle \mathbf{A}, \mathbf{A}^{\uparrow}\rangle]$)

Polynomial time delay complexity

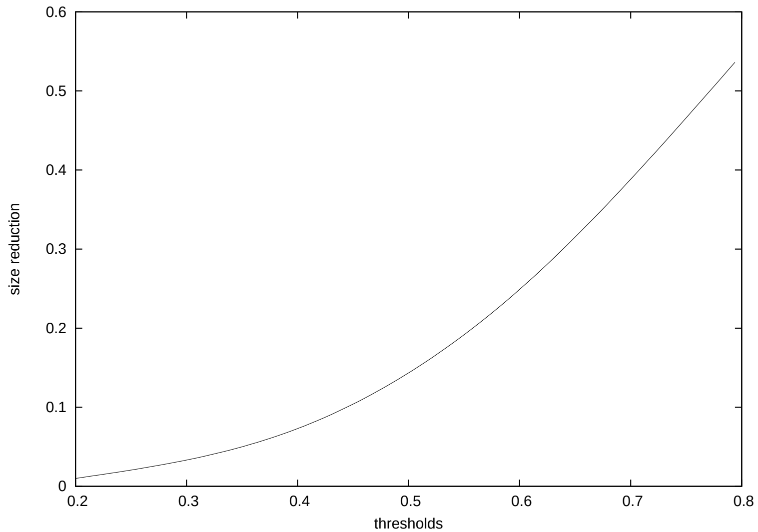
Ganter's algorithm, generating fix(C_a), has polynomial time delay complexity (in terms of size of the input $\langle X, Y, I\rangle$).

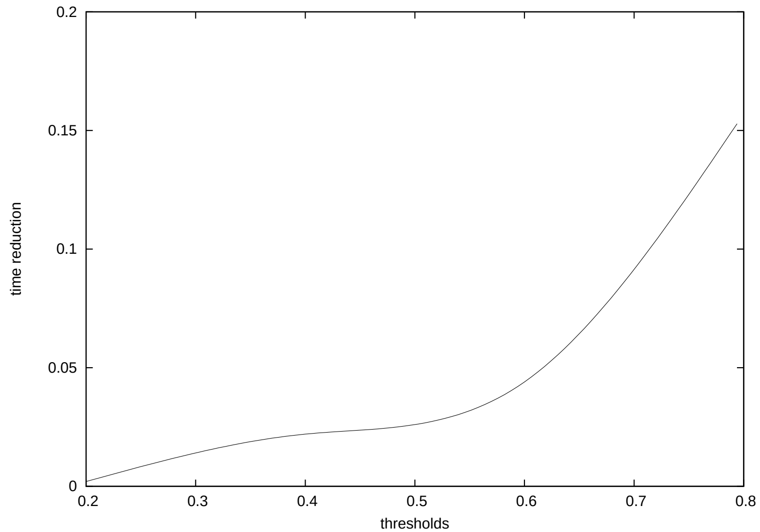
Since generating a ${}^{a}\!\approx$-block $[\langle(a \otimes A)^{\uparrow\downarrow}, a \rightarrow A^{\uparrow}\rangle, \langle A, A^{\uparrow}\rangle]$ from A takes a polynomial time, our algorithm is of polynomial time delay complexity as well.

Experiments

Łukasiewicz fuzzy logical connectives, $|\mathcal{B}(X,Y,I)| = 774$, time for computing $\mathcal{B}(X,Y,I) = 2292$ms

	0.2	0.4	0.6	0.8
size $\lvert\mathcal{B}(X,Y,I)/^a\approx\rvert$	8	57	193	423
naive algorithm (ms)	8995	9463	8573	9646
our algorithm (ms)	23	214	383	1517
reduction $\lvert\mathcal{B}(X,Y,I)/^a\approx\rvert/\lvert\mathcal{B}(X,Y,I)\rvert$	0.010	0.073	0.249	0.546
time reduction	0.002	0.022	0.044	0.157





Reduction $|\mathcal{B}(X,Y,I)/^a\approx|/|\mathcal{B}(X,Y,I)|$ and time reduction from Tab.

ASSOCIATION RULES

Association rules

- association rules = attribute implications + criteria of interestingness (support, confidence)

- introduced in 1993 (Agrawal R., Imielinski T., Swami A. N.: Mining association rules between sets of items in large databases. *Proc. ACM Int. Conf. of management of data*, pp. 207–216, 1993)

- but see GUHA method (in fact, association rules with statistics):
  - developed at 1960s by P. Hájek et al. (Academy of Sciences, Czech)
  - GUHA book available at `http://www.cs.cas.cz/ hajek/guhabook/`: Hájek P., Havránek T.: Mechanizing Hypothesis Formation. Mathematical Foundations for General Theory. Springer, 1978.

- one of main techniques in data mining

- good book: Adamo J.-M.: Data Mining for Association Rules and Sequential Patterns. Sequential and Parallel Algorithms. Springer, New York, 2001.

- good overview: Dunham M. H.: Data Mining. Introductory and Advanced Topics. Prentice Hall, Upper Saddle River, NJ, 2003.

Basic concepts

Association rule (over set Y of attributes) is an expression $A \Rightarrow B$ where $A, B \subseteq Y$ (sometimes we assume $A \cap B = \emptyset$).

Note: Association rules are just attribute implications in sense of FCA.

Data for mining (terminology in DM community): a set Y of **items**, a **database D of transactions**, $D = \{t_1, \dots, t_n\}$ where $t_i \subseteq Y$.

Note: one-to-one correspondence between databases D (over Y) and formal contexts (with attributes from Y): Given D, the corresponding $\langle X, Y, I\rangle_D$ is given by

$$\langle X, Y, I\rangle_D \dots X = D,\ \langle t_1, y\rangle \in I \Leftrightarrow y \in t_1;$$

given $\langle X, Y, I\rangle$, the corresponding $D_{\langle X,Y,I\rangle}$ is given by

$$D_{\langle X,Y,I\rangle} = \{\{x\}^{\uparrow} \mid x \in X\}.$$

(we will use both ways)

Why items and transactions?

original motivation:

item = product in a store

transaction = cash register transaction (set of items purchased)

association rule = says: when all items from A abre purchased then also all items from B are purchased

Example transactions $X = \{x_1, \ldots, x_5\}$, items $Y = \{be, br, je, mi, pb\}$ (beer, bread, jelly, milk, peanut butter)

I	be	br	je	mi	pb
x_1		×	×		×
x_2		×			×
x_3		×		×	×
x_4	×	×			
x_5	×			×	

For instance: a customer relaizing transaction x_3 bought bread, milk, and peanut butter.

Support and confidence

Def. Support of $A \Rightarrow B$ denoted by $\text{supp}(A \Rightarrow B)$ and defined by

$$\text{supp}(A \Rightarrow B) = \frac{|\{x \in X \mid \text{for each } y \in A \cup B : \langle x, y\rangle \in I\}|}{|X|},$$

i.e. $\text{supp}(A \Rightarrow B) \cdot 100\%$ of transactions contain $A \cup B$ (percentage of transactions where customers bought items from $A \cup B$).

Note that (in terms of FCA)

$$\text{supp}(A \Rightarrow B) = \frac{|(A \cup B)^{\downarrow}|}{|X|}.$$

Def. Confidence of $A \Rightarrow B$ denoted by $\text{conf}(A \Rightarrow B)$ and defined by

$$\text{conf}(A \Rightarrow B) = \frac{|\{x \in X \mid \text{for each } y \in A \cup B : \langle x, y\rangle \in I\}|}{|\{x \in X \mid \text{for each } y \in A : \langle x, y\rangle \in I\}|},$$

i.e. $\text{conf}(A \Rightarrow B) \cdot 100\%$ of transactions containing all items from A contain also all items from B (percentage of customers which by also (all from) B if they buy (all from) A.

Note that (in terms of FCA)

$$\text{conf}(A \Rightarrow B) = \frac{|(A \cup B)^{\downarrow}|}{A^{\downarrow}}.$$

We use both "support (confidence) is 0.3" and "support (confidence) is 30%".

Lemma supp($A \Rightarrow B$) $\leq$ conf($A \Rightarrow B$).

Lemma conf($A \Rightarrow B$) $= 1$ iff $||A \Rightarrow B||_{\langle X,Y,I \rangle} = 1$. That is, attribute implications which are true in $\langle X, Y, I \rangle$ are those which are fully confident.

More generally: for $B \subseteq Y$, put

$$\text{supp}(B) = \frac{|(A \cup B)^{\downarrow}|}{|X|}.$$

Example

What are association rules good for

main usage = marketing

usually, rules with large confidence (reliable) and smaller support are looked for

see also applications part (REFERATY)

Association rules problem

For prescribed values s and c, list all association rules with $\text{supp}(A \Rightarrow B) \geq s$ **and** $\text{conf}(A \Rightarrow B) \geq c$**.** (interesting rules)

most common technique: via frequent itemsets

1. find all frequent itemsets (see later)

2. generate rules from frequent itemsets

Def. For given s, an itemset (set of attributes) $B \subseteq Y$ is called **frequent** (large) **itemset** if supp$(B) \geq s$.

Example For $s = 0.3$ (30%),

$$L = \{\{be\}, \{br\}, \{mi\}, \{pb\}, \{br, pb\}, \}$$

How to generate interesting rules from large itemsets?

```
input
 <X,Y,I>, L (set of all frequent itemsets), s (support), c
(confidence)
output
 R (set of all asociation rules satisfying s and c)
algorithm (ARGen)
R:=O; //empty set
for each l in L do
 for each nonempty proper subset k of l do
  if  supp(l)/supp(k) >= c then
   add rule k=>(l-k) to R
```

Observe: supp(l)/supp(k) = conf($k \Rightarrow l - k$)

Example (previous cntd.) consider $c = 0.8$, take $l = \{br, pb\}$; there are two nonempty subsets k of l: $k = \{br\}$ and $k = \{pb\}$ then $br \Rightarrow pb$ IS NOT interesting since

$$\text{supp}(\{br, pb\})/\text{supp}(\{br\}) = 0.6/0.8 = 0.75 \not\geq c$$

while $pb \Rightarrow br$ IS interesting since

$$\text{supp}(\{pb, br\})/\text{supp}(\{pb\}) = 0.6/0.6 = 1.0 \geq c.$$

(efficient implementation later)

How to generate frequent itemsets (Apriori algorithm)

Lemma Any subset of a frequent itemset is frequent. If an itemset is not frequent then no of its supersets is frequent.
Proof Obvious.

basic idea of apriori algorithm: L_i . . . set of all frequent itemsets of size i (i.e. with i items), C_i . . . set of all itemsets of size i which are candidates for being frequent

1. in step i, C_i from L_{i-1} (if $i = 1$, put $C_1 = \{\{y\} \mid y \in Y\}$);
2. scanning $\langle X, Y, I\rangle$, generate L_i, the set of all those candidates from C_i which are frequent

How to get candidates C_i from frequent items L_{i-1}?

1. what means "a candidate": an itemset $B \subseteq Y$ is considered a candidate (for being frequent) if all of its subsets are frequent (in accordance with above Lemma)
2. getting C_i from L_{i-1}: find all $B_1, B_2 \in L_{i-1}$ such that $|B_1 - B_2| = 1$ and $|B_2 - B_1| = 1$ (i.e. $|B_1 \cap B_2| = i - 2$), and add $B_1 \cup B_2$ to C_i

Lemma If L_{i-1} is the set of all frequent itemsets of size $i-1$ then B is a candidate (i.e., all subsets of B are frequent) of size i iff $B = B_1 \cup B_2$ where $B_1, B_2 \in L_{i-1}$ are such that $|B_1 - B_2| = 1$ and $|B_2 - B_1| = 1$. Moreover, $|B_1 - B_2| = 1$ and $|B_2 - B_1| = 1$ iff $|B_1 \cap B_2| = i - 2$.

Example (previous cntd.) consider $s = 0.3$, $c = 0.5$

step 1:
$C_1 = \{\{br\}, \{br\}, \{je\}, \{mi\}, \{pb\}\}$
$L_1 = \{\{br\}, \{br\}, \{mi\}, \{pb\}\}$

step 2:
$C_2 = \{\{be, br\}, \{be, mi\}, \{be, pb\}, \{br, mi\}, \{br, pb\}, \{mi, pb\}\}$
$L_2 = \{\{br, pb\}\}$

stop (not itemset of size 3 can be frequent)

Algorithms

```
input
 L(i-1) //all frequent itemsets of size i-1
output
 C(i)   //candidates of size i
algorithm (Apriori-Gen)
C(i):=O; //empty set
for each B1 from L(i-1) do
 for each B2 from L(i-1) different from B1 do
  if intersection of B1 and B2 has just i-2 elements then
   add union of B1 and B2 to C(i)
```

down(B) means $B^{\downarrow}$

```
input
 <X,Y,I> //data table
 s       //prescribed support
output
 L   //set of all frequent itemsets
algorithm (Apriori)
k:=0;   //scan (step) number
L:=0;   //emptyset
C(0):={ {y} | y from Y}
repeat
 k:=k+1;
 L(k):=0;
 for each B from C(k) do
  if |down(B)| >= s x |X| do  // B is frequent
   add B to L(k)
 add all B from L(k) to L;
 C(k+1):=Apriori-Gen(L(k))
until C(k+1)=0;  \\empty set
```

ke zkousce z assoc. rules: to, co je na slajdech; dalsi veci (efektivni algoritmy, priklady) pristi semestr