

## Cvičení 6

Pojmy potřebné pro zvládnutí tohoto cvičení:

formální kontext, Galoisova konexe, operátory  $\uparrow$  a  $\downarrow$ , formální koncept (pojem) - rozsah (extent), obsah (intent), uspořádání konceptů, supremum a infimum v konceptuálním svazu, částečné uspořádání v konceptuálním svazu (pojmy - koncepty), transakční data, množina položek (itemset  $K \subseteq I$ ), asociační pravidlo, podpora (support),  $\mu$ -frequent itemset ( $\mu$ FIS),  $\mu$  = minimální podpora, spolehlivost (confidence),  $c$  = minimální spolehlivost.

**Příklad 1:** Pro daný kontext  $(X, Y, I)$  a dané podmnožiny  $(A_i \subseteq X, B_i \subseteq Y)$ :

- nalezněte  $A_i^\uparrow$  (intent množiny  $A_i$ ) a  $B_i^\downarrow$  (extent množiny  $B_i$ ),
- určete, v jakých množinových relacích jsou zadané množiny a jaké množinové relace platí pro získané intenty a extenty,
- ověřte, které ze zadaných podmnožin jsou uzavřené (tj.  $A_i = (A_i^\uparrow)^\downarrow$ ,  $B_i = (B_i^\downarrow)^\uparrow$ ).

a) Pro podmnožiny  $A_1 = \{x_4\}$ ,  $A_2 = \{x_3, x_4\}$ ,  $A_3 = \{x_3, x_4, x_5\}$  určete  $A_i^\uparrow$  a pro podmnožiny  $B_1 = \{y_6\}$ ,  $B_2 = \{y_3, y_6\}$ ,  $B_3 = \{y_2, y_3, y_6\}$  určete  $B_i^\downarrow$ .

	y <sub>1</sub>	y <sub>2</sub>	y <sub>3</sub>	y <sub>4</sub>	y <sub>5</sub>	y <sub>6</sub>
x <sub>1</sub>		1		1		1
x <sub>2</sub>		1			1	
x <sub>3</sub>			1		1	1
x <sub>4</sub>	1	1	1			1
x <sub>5</sub>				1		1

b) Pro podmnožiny  $A_1 = \{T_4, T_6\}$ ,  $A_2 = \{T_1, T_2, T_6\}$  určete  $A_i^\uparrow$  a pro podmnožiny  $B_1 = \{a, c\}$ ,  $B_2 = \{b, e\}$  určete  $B_i^\downarrow$ .

	a	b	c	d	e
T <sub>1</sub>		1		1	
T <sub>2</sub>		1			1
T <sub>3</sub>			1		
T <sub>4</sub>	1	1	1		
T <sub>5</sub>				1	
T <sub>6</sub>		1	1		
T <sub>7</sub>					1

c) Pro podmnožiny  $A_1 = \{x_1, x_3\}$ ,  $A_2 = \{x_3, x_5\}$  určete  $A_i^\uparrow$  a pro podmnožiny  $B_1 = A_1^\uparrow$ ,  $B_2 = A_2^\uparrow$  určete  $B_i^\downarrow$ .

	y <sub>1</sub>	y <sub>2</sub>	y <sub>3</sub>	y <sub>4</sub>	y <sub>5</sub>	y <sub>6</sub>
x <sub>1</sub>	1	0	1	1	0	1
x <sub>2</sub>	0	1	1	0	1	1
x <sub>3</sub>	1	0	1	1	0	0
x <sub>4</sub>	0	0	0	1	1	1
x <sub>5</sub>	1	0	1	1	0	0

**Příklad 2:** Pro kontexty z příkladu 1 určete množiny všech konceptů a vytvořte konceptuální svazy - pomocí jedinečných průniků. Pro vytvořené konceptuální svazy určete, zda

jsou distributivní, modulární, komplementární či dokonce booleovské.

- a) jeden konceptuální svaz na daném kontextu,
- b) a druhý konceptuální svaz na nějakých  $T_{kach}$ ,
- c) a další konceptuální svaz na daném kontextu.

**Příklad 3:** Pro níže uvedená data v podobě tabulky transakcí v lékárně:

- a) napište 0/1 reprezentaci, konceptu?
- b) vytvořte konceptuální svaz, h) najděte menší množinu položek se stejnou podporou, jako má větší množina položek (tj. generator) - co jim odpovídá ve formálním konceptu?
- c) nakreslete příslušný Rymon tree (ve vztahu k algoritmu Apriori),
- d) vypište frekventované množiny položek (frequent itemsets) z Rymon tree, pro které platí  $\mu \geq 0.1$ ,
- e) vytvořte všechna asociační pravidla nad FIS (frequent item sets),
- f) určete jejich spolehlivost,
- g) najděte množiny položek, které nelze rozšířit beze změny podpory (tj. closed itemset) - co jim odpovídá ve formálním

Transaction	Content
T1	Paralen, Vitamín C
T2	Paralen, Stodal
T3	Fastum gel
T4	Paralen, Vitamín C, Stodal
T5	Fastum gel, Ibalgin
T6	Paralen, Ibalgin
T7	Paralen, Vitamín C

**Příklad 4:**

Mějme množinu položek  $I = \{1, 2, 3, 4, 5\}$  a k ní příslušných  $m$  transakcí. Určete:

- a) Co platí pro podpory podmnožin položek (itemsets)  $K, K' \subseteq I$ , jestliže  $K' \subseteq K$ ?  
[sup( $K'$ ) ... sup( $K$ )]
- b) Co platí pro spolehlivost pravidel  $A \Rightarrow K$  a  $A \Rightarrow K'$ , kde  $A, K, K' \subseteq I$ , pokud  $A \cap K' = \emptyset = A \cap K$  a  $K' \subseteq K$ ?  
[conf( $A \Rightarrow K'$ ) ... conf( $A \Rightarrow K$ )]
- c) Co platí pro spolehlivost pravidel  $A \Rightarrow K$  a  $A' \Rightarrow K$ , kde  $A, A', K \subseteq I$ , pokud  $A \cap K = \emptyset = A' \cap K$  a  $A' \subseteq A$ ?  
[conf( $A' \Rightarrow K$ ) ... conf( $A \Rightarrow K$ )]

Můžeme vybrat následující podmnožiny množiny položek:  $A' = \{1\}$ ,  $A = \{1, 2\}$ ,  $K' = \{3, 4\}$ ,  $K = \{3, 4, 5\}$ .

**Příklad 5:** Je slovně popsána binární heterogenní relace  $I \subseteq L \times A$ :

- Adam hraje šachy a ragby.
- Béd'a hraje na ukulele a šachy.
- Cyril hraje ragby, vodní polo a tenis.

- Dan hraje na ukulele a vodní polo a tenis.
  - Ernesto hraje na ukulele a tenis.
  - Filip hraje šachy a tenis.
- 
- a) Danou relaci zapište jako formální kontext (Lidi, Aktivity, I).
  - b) Pomocí algoritmu jedinečných průniků nalezněte všechny koncepty.
  - c) Pomocí uspořádání konceptů nakreslete výsledný konceptuální svaz.
  - d) Určete vlastnosti vytvořeného svazu (úplnost, ohraničenost - tj. existenci svazové nuly a svazové jednotky, distributivnost, modularitu, komplementaritu, zda je Booleovský).
  - e) Uvažujte danou binární heterogenní relaci jako transakční data a vytvořte Rymon tree.
  - f) Nalezněte všechna asociační pravidla s podporou  $\mu \geq 0.3$  a se spolehlivostí  $c \geq 0.5$ .
  - g) Zamyslete se nad vztahem mezi nalezenými koncepty a možnostmi dostatečně podporovaných položek.
  - h) Lze využít nalezený konceptuální svaz pro detekci asociačních pravidel?

**Příklad 6:** Pro níže uvedená data v podobě homogenního neorientovaného grafu bez smyček  $G = (V, E)$  určete všechny koncepty nad formálním kontextem  $(V, V, I + A(G))$ , kde  $A(G)$  je matice sousednosti grafu  $G$  a matice  $I$  je jednotková. Uveďte, co jednotlivé koncepty reprezentují a co reprezentují koncepty ve tvaru  $\text{ext}_k = \text{int}_k$ .

- a)  $V = \{1, 2, 3, 4, 5, 6\}$ ,  $E = \{(1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (2, 5), (3, 6), (4, 5), (5, 6)\}$
- b)  $V = \{1, 2, 3, 4, 5, 6, 7, 8\}$ ,  $E = \{(1, 2), (1, 4), (1, 5), (2, 3), (2, 4), (2, 5), (3, 6), (4, 5), (4, 7), (5, 6), (5, 7), (6, 8), (7, 8)\}$