Dr. Christian Săcărea

"Babeş–Bolyai" University of Cluj-Napoca
Department of Computer Science

Exercise sheet Nr. 1

Knowledge Discovery in Wide Area Networks

Organization
hands-on exercise means

- *autonomous work* on the practice sheet in small teams of 3-4 students, under supervision
- *no general repetition* of lecture material
- *no demonstration* of the sample solution (will be provided later)

necessary for that is

- making notes during the lecture
- performing autonomous follow-up course work before the exercise
- bringing material and your notes to the exercise
- developing own activity

Why this exercise concept?

- active development of the lecture material is more effective
- discovering relationships in the material
- learning structured thinking and autonomous working
- learning team work
- learning to explain things
- exercise for the exams ;-)
- *You have finished your study of . . . Your personal strengths include pro-activity and team work, you are communicative and willing to cooperate.* (typical job advertisement)

GROUPWORK:

(G 1)

a) Consider the formal context of a patient suffering from Anorexia Nervosa (see Lectures). Use `ConExp` and `FCA Tools Bundle` to determine the set of concepts and to draw the concept lattices. Use `ConExp` to determine an implicational base. Find out all features of these software tools.

b) Consider the concept lattices of white wines and red wines, respectively. Rebuild the original data set as a formal context and afterwards redraw the concept lattice using `ConExp`, `FCA Tools Bundle`, and `Elba` from the ToscanaJ suite.

c) What is according to you a nice concept lattice? How can we improve the graphical appearance of a concept lattice such that it is considered as nice by humans?

d) Give an example of your choice for a data set in form of a formal context and build its lattice. If yo would program it, how would you proceed in order to help the user to have an overview of the concept hierarchy even for larger lattices?

(G 2)

Regard the following formal context K, given as a cross table:

	needs water to live	lives in water	lives on land	needs chlorophyll to produce food	two seed leaves	one seed leaf	can move around	has limbs	suckles its offspring
Leech	x	x					x		
Bream	x	x					x	x	
Frog	x	x	x				x	x	x
Spike-Weed	x	x		x		x			
Reed	x	x	x	x		x			
Bean	x		x	x	x				
Maize	x		x	x		x			

a) Specify the following sets:

1. $\{Bean\}'$
2. $\{lives\ on\ land\}'$
3. $\{two\ seed\ leaves\}''$
4. $\{Frog, Maize\}'$
5. $\{needs\ chlorophyll\ to\ produce food, can\ move\ around\}'$
6. $\{lives\ in\ water, lives\ on\ land\}'$
7. $\{needs\ chlorophyll\ to\ produce food, can\ move\ around\}'$

b) Extend $\mathbb{K}$ with both an object and an attribute.

c) Generate all concepts using the naive method, i.e., by finding out all maximal rectangles of crosses. Compare the results with the concept list generated by `FCA Tools Bundle`. Try to draw by hand the concept lattice, then compare your results with those of one of the available software tools. Compare the results. Figure out improvements.

(G 3)

a) Recall: how is the derivation operator $(\cdot)'$ defined?

b) Let $\mathbb{K} = (G, M, I)$ be a formal context and let $A, B \subseteq G$. Prove the following statements:

1. $A \subseteq B$ implies $B' \subseteq A'$
2. $A \subseteq A''$
3. $A' = A'''$
4. For $C \subseteq G$ and $D \subseteq M$ holds: (C, D) is a formal concept if and only if there is some $E \subseteq G$such that $C = E''$ and $D = E'$.