Ministry of Science and Higher Education of the Russian Federation

National Research University Higher School of Economics

Faculty of Computer Science

School of Data Analysis and Artificial Intelligence

HOMEWORK REPORT

Practical homework №3

Subject: *Ordered Sets for Data Analysis*

Student	*M.D. Kirdin*
Teaching Assistant	*A. Tomat*
Teaching Assistant	*M. Zueva*
Professor	*S.O. Kuznetsov*

Moscow, 2024

QUESTION 1

Task. For the context given in the following table:

1. find all formal concepts using the CbO algorithm (show the tree);
2. draw the concept lattice.

	a	b	c	d	e
1		1	1		
2	1	1	1		1
3	1	1		1	
4	1			1	1

Solution. Let's draw the concept lattice.

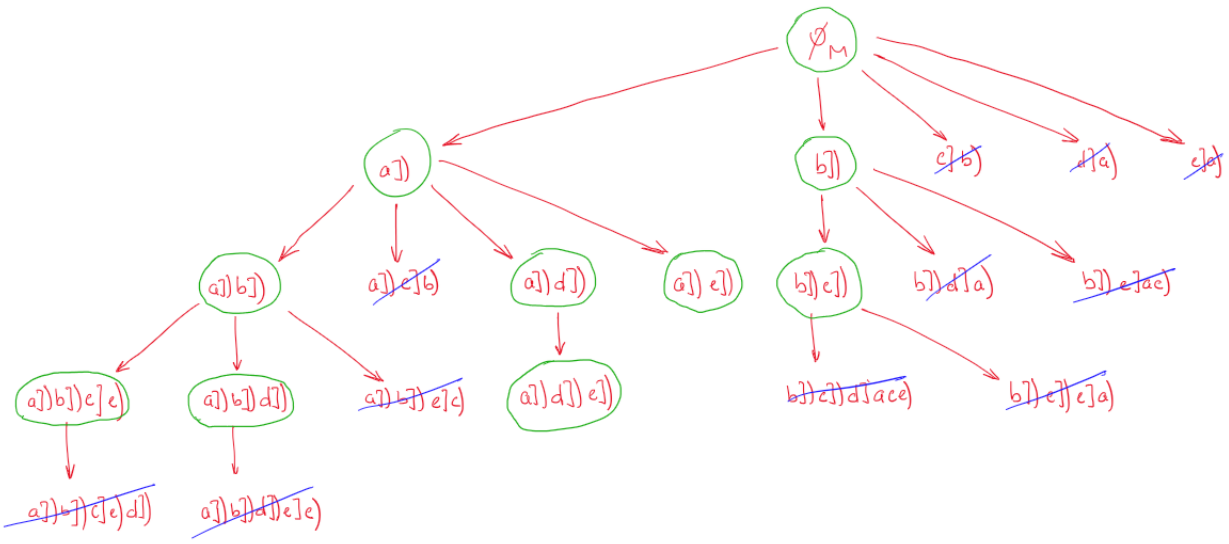

Figure 1. Concept lattice produced by CbO algorithm

On fig. 1 green circles indicate formal concepts and blue strikethroughs indicate non-canonical intent generations. From it we can infer that there are following formal concepts: $(G, \varnothing_M)$, $(\{2, 3, 4\}, \{a\})$, $(2, \{a, b, c, e\})$, $(3, \{a, b, d\})$, $(\{3, 4\}, \{a, d\})$, $(4, \{a, d, e\})$, $(\{2, 4\}, \{a, e\})$, $(\{1, 2, 3\}, b)$, $(\{1, 2\}, \{b, c\})$.

QUESTION 2

Task. For the following many-valued context given in the following table:

1. Binarize the data given in the table. Use nominal scales for the features (Brand, Color). For the feature (RAM), use the ordinal scale (≥ 8, ≥ 16, ≥ 32) given in the table on the right. For the feature (is_touch), use the dichotomic scale (is_touch, not is_touch).

2. Find the minimal positive and minimal negative hypotheses of the binarized context. (Show the concept lattices of both the positive and negative contexts).

3. Classify the objects <Razer, Black, 32, Yes>, <Toshiba, Red, 18, Yes>, <Mac, Pink, 8, No> using the found hypotheses.

Brand	Colour	RAM	is_touch	class
Lenovo	Black	16	No	+
HP	Black	16	Yes	+
Lenovo	Black	16	Yes	+
Razer	Silver	32	No	+
Razer	Gold	32	No	+
Toshiba	Pink	4	Yes	-
Toshiba	White	16	Yes	-
HP	Silver	8	No	-
Mac	Gold	16	No	-

	≥ 8	≥ 16	≥ 32
4			
8	1		
16	1	1	
32	1	1	1

Solution. Binarized data can be seen in table below:

	len	hp	raz	tsb	mac	blk	sil	gol	pnk	wht	≥ 8	≥ 16	≥ 32	tch_y	tch_n	class
1	+	-	-	-	-	+	-	-	-	-	+	+	-	-	+	+
2	-	+	-	-	-	+	-	-	-	-	+	+	-	+	-	+
3	+	-	-	-	-	+	-	-	-	-	+	+	-	+	-	+
4	-	-	+	-	-	-	-	-	-	-	+	+	+	-	+	+
5	-	-	+	-	-	-	-	+	-	-	+	+	+	-	+	+
6	-	-	-	+	-	-	-	-	+	-	-	-	-	+	-	-
7	-	-	-	+	-	-	-	-	-	+	+	+	-	+	-	-
8	-	+	-	-	-	-	+	-	-	-	+	-	-	-	+	-
9	-	-	-	-	+	-	-	+	-	-	+	+	-	-	+	-
10	-	-	+	-	-	+	-	-	-	-	+	+	+	+	-	τ
11	-	-	-	+	-	-	-	-	-	-	+	+	-	+	-	τ
12	-	-	-	-	+	-	-	-	-	-	+	-	-	-	+	τ

Using this data we can construct positive and negative concept lattices.

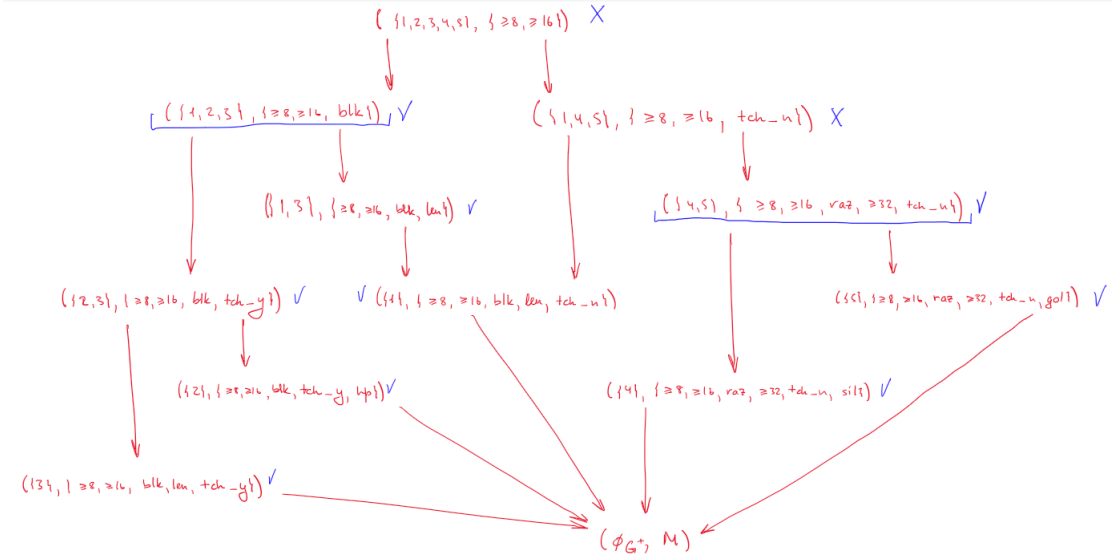

Figure 2. Positive concept lattice

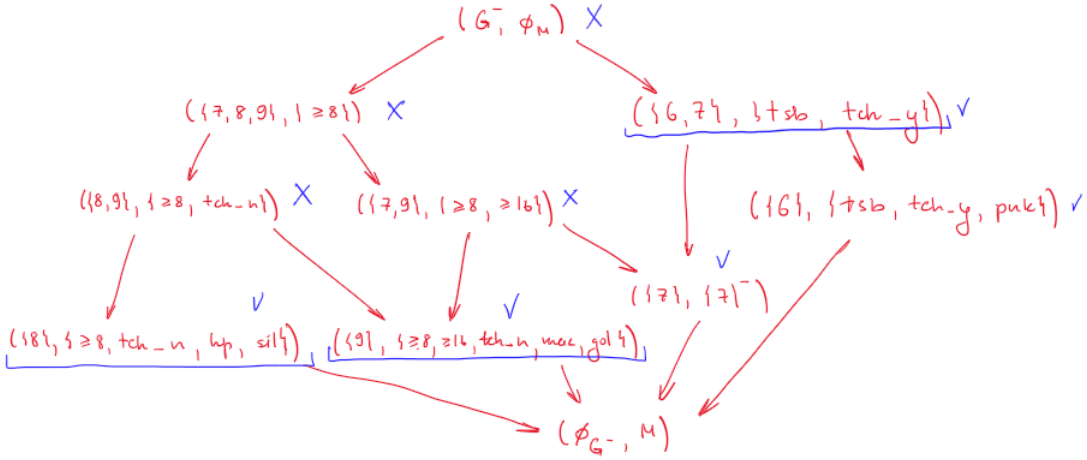

Figure 3. Negative concept lattice

According to fig. 2, $\{\geq 8, \geq 16, \text{blk}\}$ and $\{\geq 8, \geq 16, \geq 32, \text{raz}, \text{tch\_n}\}$ are minimal positive hypotheses. $\{\geq 8, \geq 16\}$ is not a positive hypothesis since it is part of a negative hypothesis. Those, according to fig. 3 are: $\{\text{tsb}, \text{tch\_y}\}$, $\{\geq 8, \text{tch\_n}, \text{hp}, \text{sil}\}$, $\{\geq 8, \geq 16, \text{tch\_n}, \text{mac}, \text{gol}\}$. Therefore, since:

- $\{10\}^\tau$ contains $\{\geq 8, \geq 16, \text{blk}\}$ and no negative hypotheses, 10 is a positive observation;
- $\{11\}^\tau$ contains $\{\text{tsb}, \text{tch\_y}\}$ and no positive hypotheses, 11 is a negative observation;

- $\{12\}^{\tau}$ does not contain any positive or negative observations, 12 should be labelled as undetermined.