

CCOD

Alexander Bryuzgin(AS-36)

Lab 9(10.28.2016)

Clusterisation of 10 russian letters

For this practice I've chosen these 10 letters: **О, С, Ю, Ъ, Ы, Ь, И, Й, Р, У**. $\alpha = 0.5$

Here is a table of similarity percentage of letters

	О	С	Ю	Ъ	Ы	Ь	И	Й	Р	У	
О	1	0.8	0.7	0.4	0.3	0.4	0.1	0.1	0.15	0.1	R0
С	0.8	1	0.7	0.3	0.2	0.3	0.1	0.1	0.15	0.1	
Ю	0.7	0.7	1	0.35	0.25	0.35	0.1	0.1	0.1	0.1	
Ъ	0.4	0.3	0.35	1	0.85	0.95	0.15	0.15	0.4	0.1	
Ы	0.3	0.2	0.25	0.85	1	0.8	0.2	0.2	0.3	0.1	
Ь	0.4	0.3	0.35	0.95	0.8	1	0.1	0.1	0.4	0.1	
И	0.1	0.1	0.1	0.15	0.2	0.1	1	0.95	0.1	0.1	
Й	0.1	0.1	0.1	0.15	0.2	0.1	0.95	1	0.1	0.1	
Р	0.15	0.15	0.1	0.4	0.3	0.4	0.1	0.1	1	0.1	
У	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1	

There will be 3 steps of modifying this table

	О	С	Ю	Ъ	Ы	Ь	И	Й	Р	У	
О	1	0.8	0.7	0.4	0.4	0.4	0.2	0.2	0.4	0.1	R1
С	0.8	1	0.7	0.4	0.3	0.4	0.2	0.2	0.3	0.1	
Ю	0.7	0.7	1	0.4	0.35	0.4	0.2	0.2	0.35	0.1	
Ъ	0.4	0.4	0.4	1	0.85	0.95	0.2	0.2	0.4	0.1	
Ы	0.4	0.3	0.35	0.85	1	0.85	0.2	0.2	0.4	0.1	
Ь	0.4	0.4	0.4	0.95	0.85	1	0.2	0.2	0.4	0.1	
И	0.2	0.2	0.2	0.2	0.2	0.2	1	0.95	0.2	0.1	
Й	0.2	0.2	0.2	0.2	0.2	0.2	0.95	1	0.2	0.1	
Р	0.4	0.3	0.35	0.4	0.4	0.4	0.2	0.2	1	0.1	
У	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1	

	О	С	Ю	Ъ	Ы	Ь	И	Й	Р	У	
О	1	0.8	0.7	0.4	0.4	0.4	0.2	0.2	0.4	0.1	R2
С	0.8	1	0.7	0.4	0.4	0.4	0.2	0.2	0.4	0.1	
Ю	0.7	0.7	1	0.4	0.4	0.4	0.2	0.2	0.4	0.1	
Ъ	0.4	0.4	0.4	1	0.85	0.95	0.2	0.2	0.4	0.1	
Ы	0.4	0.4	0.4	0.85	1	0.85	0.2	0.2	0.4	0.1	
Ь	0.4	0.4	0.4	0.95	0.85	1	0.2	0.2	0.4	0.1	
И	0.2	0.2	0.2	0.2	0.2	0.2	1	0.95	0.2	0.1	
Й	0.2	0.2	0.2	0.2	0.2	0.2	0.95	1	0.2	0.1	
Р	0.4	0.4	0.4	0.4	0.4	0.4	0.2	0.2	1	0.1	
У	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1	

	О	С	Ю	Ъ	Ы	Ь	И	Й	Р	У	
О	1	0.8	0.7	0.4	0.4	0.4	0.2	0.2	0.4	0.1	R3
С	0.8	1	0.7	0.4	0.4	0.4	0.2	0.2	0.4	0.1	
Ю	0.7	0.7	1	0.4	0.4	0.4	0.2	0.2	0.4	0.1	
Ъ	0.4	0.4	0.4	1	0.85	0.95	0.2	0.2	0.4	0.1	
Ы	0.4	0.4	0.4	0.85	1	0.85	0.2	0.2	0.4	0.1	
Ь	0.4	0.4	0.4	0.95	0.85	1	0.2	0.2	0.4	0.1	
И	0.2	0.2	0.2	0.2	0.2	0.2	1	0.95	0.2	0.1	
Й	0.2	0.2	0.2	0.2	0.2	0.2	0.95	1	0.2	0.1	
Р	0.4	0.4	0.4	0.4	0.4	0.4	0.2	0.2	1	0.1	
У	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1	

Final table, where all values less than α and main diagonal will become zeros.

	О	С	Ю	Ъ	Ы	Ь	И	Й	Р	У	
О	0	0.8	0.7	0	0	0	0	0	0	0	Final
С	0.8	0	0.7	0	0	0	0	0	0	0	
Ю	0.7	0.7	0	0	0	0	0	0	0	0	
Ъ	0	0	0	0	0.85	0.95	0	0	0	0	
Ы	0	0	0	0.85	0	0.85	0	0	0	0	
Ь	0	0	0	0.95	0.85	0	0	0	0	0	
И	0	0	0	0	0	0	0	0.95	0	0	
Й	0	0	0	0	0	0	0.95	0	0	0	
Р	0	0	0	0	0	0	0	0	0	0	
У	0	0	0	0	0	0	0	0	0	0	

The iterations begin from maximum meaning

$$I = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

$$C_1 = \{\}$$

$$a_{46} = a_{78} = 0.95$$

$$C_1 = \{4, 6\}$$

$$\max \text{sum} = 5 \Rightarrow C_1 = \{4, 5, 6\}$$

$$I = \{1, 2, 3, 7, 8, 9, 10\}$$

$$C_2 = \{\}$$

$$a_{78} = 0.95$$

$$C_2 = \{7, 8\}$$

$$\max \text{sum} \emptyset \Rightarrow C_2 = \{7, 8\}$$

$$I = \{1, 2, 3, 9, 10\}$$

$$C_3 = \{\}$$

$$a_{12} = 0.8$$

$$C_3 = \{1,2\}$$

$$\max \text{sum} = 3 \Rightarrow C_3 = \{1,2,3\}$$

No other max sum

$$I = \{9,10\}$$

$$C_4 = \{9\}$$

$$C_5 = \{10\}$$

$$C = \{\{1,2,3\}, \{4,5,6\}, \{7,8\}, \{9\}, \{10\}\}$$

OR

$$C = \{\{O, C, Ю\}, \{Ъ, Ы, Ъ\}, \{И, Й\}, \{P\}, \{Y\}\}$$