

CCOD

Kolesnikov Dmitriy

Lab 10(10.29.2016)

Clusterisation of English alphabet

Similarity table:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
A	1	0.2	0.1	0.1	0.2	0.2	0.1	0.3	0.1	0.1	0.2	0.1	0.7	0.8	0.1	0.1	0.1	0.4	0.2	0.3	0.1	0.4	0.3	0.65	0.65	0.3	A
B	0.2	1	0.1	0.85	0.7	0.6	0.3	0.3	0.1	0.1	0.55	0.1	0.1	0.1	0.4	0.7	0.1	0.85	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.55	B
C	0.1	0.1	1	0.4	0.1	0.1	0.6	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.9	0.55	0.9	0.1	0.6	0.1	0.1	0.1	0.1	0.1	0.1	0.1	C
D	0.1	0.85	0.4	1	0.4	0.3	0.6	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.8	0.65	0.75	0.3	0.55	0.1	0.35	0.1	0.1	0.1	0.1	0.1	D
E	0.2	0.7	0.1	0.4	1	0.9	0.1	0.1	0.2	0.2	0.52	0.48	0.1	0.1	0.1	0.1	0.1	0.9	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.6	E
F	0.2	0.6	0.1	0.3	0.9	1	0.1	0.1	0.1	0.1	0.55	0.4	0.1	0.1	0.1	0.1	0.95	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.55	F
G	0.1	0.3	0.6	0.6	0.1	0.1	1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.75	0.65	0.7	0.1	0.55	0.1	0.1	0.1	0.1	0.1	0.1	0.1	G
H	0.3	0.3	0.1	0.1	0.1	0.1	0.1	1	0.1	0.1	0.1	0.1	0.4	0.4	0.1	0.1	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.1	0.1	H
I	0.1	0.1	0.1	0.1	0.2	0.1	0.1	0.1	1	0.9	0.1	0.9	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.8	0.1	0.1	0.1	0.1	0.1	0.1	I
J	0.1	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.9	1	0.1	0.75	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.75	0.1	0.1	0.1	0.1	0.1	0.1	J
K	0.2	0.55	0.1	0.1	0.52	0.55	0.1	0.1	0.1	0.1	1	0.1	0.3	0.2	0.1	0.1	0.1	0.8	0.1	0.1	0.1	0.1	0.6	0.55	0.56	K	
L	0.1	0.1	0.1	0.1	0.48	0.4	0.1	0.1	0.9	0.75	0.1	1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.7	0.1	0.1	0.1	0.1	0.1	0.1	L
M	0.7	0.1	0.1	0.1	0.1	0.1	0.1	0.4	0.1	0.1	0.3	0.1	1	0.8	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.45	0.49	0.6	0.55	0.1	M
N	0.8	0.1	0.1	0.1	0.1	0.1	0.1	0.4	0.1	0.1	0.2	0.1	0.8	1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.3	0.45	0.55	0.55	0.1	N
O	0.1	0.4	0.9	0.8	0.1	0.1	0.75	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1	0.55	0.95	0.1	0.55	0.1	0.1	0.1	0.1	0.1	0.1	0.1	O
P	0.1	0.7	0.55	0.65	0.1	0.1	0.65	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.55	1	0.55	0.9	0.55	0.1	0.1	0.1	0.1	0.1	0.1	0.1	P
Q	0.1	0.1	0.9	0.75	0.1	0.1	0.7	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.95	0.55	1	0.1	0.51	0.1	0.1	0.2	0.1	0.1	0.1	0.1	Q
R	0.4	0.85	0.1	0.3	0.9	0.95	0.1	0.1	0.1	0.1	0.8	0.1	0.1	0.1	0.1	0.9	0.1	1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.55	R
S	0.2	0.1	0.6	0.55	0.1	0.1	0.55	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.55	0.55	0.51	0.1	1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	S
T	0.3	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.8	0.75	0.1	0.7	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1	0.1	0.1	0.1	0.1	0.1	0.1	T
U	0.1	0.1	0.1	0.35	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1	0.9	0.7	0.1	0.45	0.1	U
V	0.4	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.45	0.3	0.1	0.1	0.2	0.1	0.1	0.1	0.9	1	0.85	0.45	0.49	0.1	V
W	0.3	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.49	0.45	0.1	0.1	0.1	0.1	0.1	0.1	0.7	0.85	1	0.4	0.4	0.1	W
X	0.65	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.6	0.1	0.6	0.55	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.45	0.4	1	0.6	0.1	X
Y	0.65	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.55	0.1	0.55	0.55	0.1	0.1	0.1	0.1	0.1	0.1	0.45	0.49	0.4	0.6	1	0.1	Y
Z	0.3	0.55	0.1	0.1	0.6	0.55	0.1	0.1	0.1	0.1	0.56	0.1	0.1	0.1	0.1	0.1	0.1	0.55	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1	Z
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	

Final table:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	A	0	0.6	0.6	0.6	0.6	0.6	0	0	0	0.6	0	0.8	0.8	0.6	0.6	0.6	0.6	0.6	0	0	0	0	0.65	0.65	0.6
2	B	0.6	0	0.8	0.85	0.85	0.75	0	0	0	0.8	0	0.6	0.6	0.8	0.85	0.8	0.85	0.6	0	0	0	0	0.6	0.6	0.6
3	C	0.6	0.8	0	0.8	0.8	0.8	0.75	0	0	0	0.8	0	0.6	0.6	0.9	0.8	0.9	0.8	0.6	0	0	0	0.6	0.6	0.6
4	D	0.6	0.85	0.8	0	0.85	0.85	0.75	0	0	0	0.8	0	0.6	0.6	0.8	0.85	0.8	0.85	0.6	0	0	0	0.6	0.6	0.6
5	E	0.6	0.85	0.8	0.85	0	0.9	0.75	0	0	0	0.8	0	0.6	0.6	0.8	0.9	0.8	0.9	0.6	0	0	0	0.6	0.6	0.6
6	F	0.6	0.85	0.8	0.85	0.9	0	0.75	0	0	0	0.8	0	0.6	0.6	0.8	0.9	0.8	0.95	0.6	0	0	0	0.6	0.6	0.6
7	G	0.6	0.75	0.75	0.75	0.75	0.75	0	0	0	0	0.75	0	0.6	0.6	0.75	0.75	0.75	0.75	0.6	0	0	0	0.6	0.6	0.6
8	H	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	I	0	0	0	0	0	0	0	0	0.9	0	0.9	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0
10	J	0	0	0	0	0	0	0	0.9	0	0	0.9	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0
11	K	0.6	0.8	0.8	0.8	0.8	0.8	0.75	0	0	0	0	0.6	0.6	0.8	0.8	0.8	0.8	0.6	0	0	0	0.6	0.6	0.6	
12	L	0	0	0	0	0	0	0	0.9	0.9	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0
13	M	0.8	0.6	0.6	0.6	0.6	0.6	0.6	0	0	0	0.6	0	0	0.8	0.6	0.6	0.6	0.6	0.6	0	0	0	0.65	0.65	0.6
14	N	0.8	0.6	0.6	0.6	0.6	0.6	0.6	0	0	0	0.6	0	0.8	0	0.6	0.6	0.6	0.6	0.6	0	0	0	0.65	0.65	0.6
15	O	0.6	0.8	0.9	0.8	0.8	0.8	0.75	0	0	0	0.8	0	0.6	0.6	0	0.8	0.95	0.8	0.6	0	0	0	0.6	0.6	0.6
16	P	0.6	0.85	0.8	0.85	0.9	0.9	0.75	0	0	0	0.8	0	0.6	0.6	0.8	0	0.8	0.9	0.6	0	0	0	0.6	0.6	0.6
17	Q	0.6	0.8	0.9	0.8	0.8	0.8	0.75	0	0	0	0.8	0	0.6	0.6	0.95	0.8	0	0.8	0.6	0	0	0	0.6	0.6	0.6
18	R	0.6	0.85	0.8	0.85	0.9	0.95	0.75	0	0	0	0.8	0	0.6	0.6	0.8	0.9	0.8	0	0.6	0	0	0	0.6	0.6	0.6
19	S	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0	0	0	0.6	0	0.6	0.6	0.6	0.6	0.6	0.6	0	0	0	0	0.6	0.6	0.6
20	T	0	0	0	0	0	0	0	0.8	0.8	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	U	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.9	0.85	0	0	0
22	V	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.9	0	0.85	0	0	0
23	W	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.85	0.85	0	0	0	0
24	X	0.65	0.6	0.6	0.6	0.6	0.6	0.6	0	0	0	0.6	0	0.65	0.65	0.6	0.6	0.6	0.6	0.6	0	0	0	0	0.65	0.6
25	Y	0.65	0.6	0.6	0.6	0.6	0.6	0.6	0	0	0	0.6	0	0.65	0.65	0.6	0.6	0.6	0.6	0.6	0	0	0	0.65	0	0.6
26	Z	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0	0	0	0.6	0	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0	0	0	0.6	0.6	0.6

$$I = \{A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z\}$$

$$C_1 = \{\}$$

$$a_{OQ} = a_{QO} = 0.95$$

$$C_1 = \{O, Q\}$$

$$Max\ sum = C = 0.9 * 2 = 1.8 \Rightarrow C_1 = \{C, O, Q\}$$

Cluster №1 is created . Go in to the next step.

$$I = \{A, B, D, E, F, G, H, I, J, K, L, M, N, P, R, S, T, U, V, W, X, Y, Z\}$$

$$C_2 = \{\}$$

$$a_{UV} = a_{VU} = 0.9$$

$$C_2 = \{U, V\}$$

$$Max\ sum = W = 0.85 * 2 = 1.7 \Rightarrow C_2 = \{U, V, W\}$$

Cluster №2 is created . Go in to the next step.

$$I = \{A, B, D, E, F, G, H, I, J, K, L, M, N, P, R, S, T, X, Y, Z\}$$

$$C_3 = \{\}$$

$$a_{FR} = a_{RF} = 0.95$$

$$C_3 = \{F, R\}$$

$$Max\ sum = E, P = 0.9 * 2 = 1.8 \Rightarrow C_3 = \{E, F, P, R\}$$

Cluster №3 is created . Go in to the next step.

$$I = \{A, B, D, G, H, I, J, K, L, M, N, S, T, X, Y, Z\}$$

$$C_4 = \{\}$$

$$a_{BD} = a_{DB} = 0.85$$

$$C_4 = \{B, D\}$$

$$Max\ sum = K = 0.8 * 2 = 1.6 \Rightarrow C_4 = \{B, D, K\}$$

Cluster №4 is created . Go in to the next step.

$$I = \{A, G, H, I, J, L, M, N, S, T, X, Y, Z\}$$

$$C_5 = \{\}$$

$$a_{MN} = a_{NM} = 0.8$$

$$C_5 = \{M, N\}$$

$$Max\ sum = A = 0.8 * 2 = 1.6 \Rightarrow C_5 = \{A, M, N\}$$

Cluster №5 is created . Go in to the next step.

$$I = \{G, H, I, J, L, S, T, X, Y, Z\}$$

$$C_6 = \{\}$$

$$a_{IJ} = a_{JI} = 0.9$$

$$C_6 = \{I, J\}$$

$$Max\ sum = L = 0.9 * 2 = 1.8 \Rightarrow C_6 = \{I, J, L\}$$

Cluster №6 is created . Go in to the next step.

$$I = \{G, H, S, T, X, Y, Z\}$$

$$C_7 = \{\}$$

$$a_{XY} = a_{YX} = 0,65$$

$$C_7 = \{X, Y\}$$

$$Max\ sum = G, S, Z = 0.6 * 2 = 1.2 \Rightarrow C_7 = \{G, S, X, Y, Z\}$$

Cluster №7 is created . Go in to the next step.

Only zeros remain , so we divide other letters in separate clusters.

$$C_8 = \{H\}$$

$$C_9 = \{T\}$$

Final Cluster:

$$C = \{\{C, O, Q\}, \{U, V, W\}, \{E, F, P, R\}, \{B, D, K\}, \{A, M, N\}, \{I, J, L\}, \{G, S, X, Y, Z\}, \{H\}, \{T\}\}$$