Contemporary Data Processing Technology (CCOD)

Laboratory Work (4.11.2016)

Igor Rapinchuck(AS-37)

Clusterisation of Greek Alphabet

{ Α, Β, Γ, Δ, Ε, Ζ, Η, Θ, Ι, Κ, Λ, Μ, Ν, Ξ, Ο, Π, Ρ, Σ, Τ, Φ, Χ, Ψ, Ω}

Process of clusterisation:

Table of similarity percentage of letters:

	Α	Β	Γ	Δ	Ε	Ζ	Η	Θ	Ι	Κ	Λ	Μ	Ν	Ξ	Ο	Π	Ρ	Σ	Τ	Υ	Φ	Χ	Ψ	Ω
Α	1	0.1	0.1	0.8	0.1	0.1	0.1	0.1	0.1	0.1	0.8	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Β	0.1	1	0.2	0.1	0.3	0.1	0.2	0.1	0.1	0.2	0.1	0.1	0.1	0.2	0.2	0.2	0.7	0.1	0.2	0.1	0.1	0.2	0.1	0.1
Γ	0.1	0.2	1	0.1	0.2	0.1	0.1	0.1	0.8	0.1	0.1	0.1	0.1	0.1	0.1	0.3	0.3	0.2	0.8	0.1	0.1	0.1	0.1	0.1
Δ	0.8	0.1	0.1	1	0.1	0.1	0.1	0.2	0.1	0.1	0.8	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2
Ε	0.1	0.3	0.2	0.1	1	0.1	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.7	0.1	0.1	0.2	0.7	0.2	0.1	0.1	0.1	0.1	0.1
Ζ	0.1	0.1	0.1	0.1	0.1	1	0.1	0.1	0.1	0.1	0.1	0.1	0.8	0.2	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.2	0.1	0.1
Η	0.1	0.2	0.1	0.1	0.2	0.1	1	0.1	0.1	0.1	0.1	0.1	0.3	0.3	0.1	0.6	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Θ	0.1	0.1	0.1	0.2	0.1	0.1	0.1	1	0.1	0.1	0.2	0.1	0.1	0.2	0.8	0.2	0.2	0.1	0.1	0.1	0.6	0.1	0.1	0.7
Ι	0.1	0.1	0.8	0.1	0.1	0.1	0.1	0.1	1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.1	0.8	0.2	0.2	0.1	0.2	0.1
Κ	0.1	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1	0.1	0.1	0.2	0.1	0.1	0.1	0.2	0.2	0.1	0.1	0.1	0.2	0.1	0.1
Λ	0.8	0.1	0.1	0.8	0.1	0.1	0.1	0.2	0.1	0.1	1	0.2	0.1	0.1	0.1	0.3	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.3
Μ	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	1	0.2	0.1	0.1	0.3	0.1	0.2	0.1	0.1	0.1	0.2	0.1	0.1
Ν	0.1	0.1	0.1	0.1	0.1	0.8	0.3	0.1	0.1	0.2	0.1	0.2	1	0.1	0.1	0.3	0.1	0.1	0.1	0.1	0.1	0.2	0.1	0.1
Ξ	0.1	0.2	0.1	0.1	0.7	0.2	0.3	0.2	0.1	0.1	0.1	0.1	0.1	1	0.1	0.2	0.2	0.7	0.1	0.1	0.1	0.1	0.3	0.1
Ο	0.1	0.2	0.1	0.2	0.1	0.1	0.1	0.8	0.1	0.1	0.1	0.1	0.1	0.1	1	0.2	0.2	0.1	0.1	0.1	0.7	0.1	0.1	0.8
Π	0.1	0.2	0.3	0.1	0.1	0.1	0.6	0.2	0.1	0.1	0.3	0.3	0.3	0.2	0.2	1	0.3	0.2	0.2	0.1	0.1	0.1	0.1	0.4
Ρ	0.1	0.7	0.3	0.1	0.2	0.1	0.2	0.2	0.2	0.2	0.1	0.1	0.1	0.2	0.2	0.3	1	0.1	0.2	0.1	0.3	0.1	0.3	0.1
Σ	0.1	0.1	0.2	0.1	0.7	0.2	0.1	0.1	0.1	0.2	0.1	0.2	0.1	0.7	0.1	0.2	0.1	1	0.1	0.1	0.1	0.2	0.1	0.1
Τ	0.1	0.2	0.8	0.1	0.2	0.1	0.1	0.1	0.8	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.1	1	0.3	0.3	0.1	0.3	0.1
Υ	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.1	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.3	1	0.2	0.3	0.8	0.1
Φ	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.6	0.2	0.1	0.1	0.1	0.1	0.1	0.7	0.1	0.3	0.1	0.3	0.2	1	0.1	0.3	0.3
Χ	0.1	0.2	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.2	0.3	0.2	0.2	0.1	0.1	0.1	0.1	0.2	0.1	0.3	0.1	1	0.1	0.1
Ψ	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.3	0.1	0.1	0.3	0.1	0.3	0.8	0.3	0.1	1	0.2
Ω	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.7	0.1	0.1	0.3	0.1	0.1	0.1	0.8	0.4	0.1	0.1	0.1	0.1	0.3	0.1	0.2	1

1 Iteration:

In the final table, where all values less than α and main diagonal will become ***zeros***.

		А	В	Г	Δ	Е	Z	Н	Θ	I	К	Λ	М	N	Ξ	О	П	Р	Σ	Т	Υ	Φ	Χ	Ψ	Ω
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
А	1	0	0	0	0.8	0	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0
В	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.7	0	0	0	0	0	0	0
Г	3	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0
Δ	4	0.8	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0
Е	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0.7	0	0	0	0.7	0	0	0	0	0	0
Z	6	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0
Н	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.6	0	0	0	0	0	0	0	0	0
Θ	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0.7	0	0	0.8
I	9	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0
К	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Λ	11	0.8	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
М	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N	13	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ξ	14	0	0	0	0	0.7	0	0	0	0	0	0	0	0	0	0	0	0	0.7	0	0	0	0	0	0
О	15	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0.7	0	0	0.8
П	16	0	0	0	0	0	0	0.6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Р	17	0	0.7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Σ	18	0	0	0	0	0.7	0	0	0	0	0	0	0	0	0.7	0	0	0	0	0	0	0	0	0	0
Т	19	0	0	0.8	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Υ	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0
Φ	21	0	0	0	0	0	0	0	0.7	0	0	0	0	0	0	0.7	0	0	0	0	0	0	0	0	0.7
Χ	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ψ	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0
Ω	24	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0	0.8	0	0	0	0	0	0.7	0	0	0

First set I = {1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24} and C_1 = { }.

$a_{4\,11}$ = 0.8 are maximum. Then C_1= {4, 11}.

$a_{4\,1} + a_{11\,1}$ = 1.6 are maximum. Then C_1= {4, 11, 1}

There are no j such that $a_{4\,j} + a_{11\,j} + a_{1\,j}$ is maximum. Then final C_1= {4, 11, 1} = **{ Δ, Λ, А }**

After deleting 4, 11, 1 rows and columns the table is:

		В	Г	Е	Ζ	Η	Θ	Ι	Κ	Μ	Ν	Ξ	Ο	Π	Ρ	Σ	Τ	Υ	Φ	Χ	Ψ	Ω
		2	3	5	6	7	8	9	10	12	13	14	15	16	17	18	19	20	21	22	23	24
В	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0.7	0	0	0	0	0	0	0
Г	3	0	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0
Е	5	0	0	0	0	0	0	0	0	0	0	0.7	0	0	0	0.7	0	0	0	0	0	0
Ζ	6	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0
Η	7	0	0	0	0	0	0	0	0	0	0	0	0	0.6	0	0	0	0	0	0	0	0
Θ	8	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0.7	0	0	0.8
Ι	9	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0
Κ	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Μ	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ν	13	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ξ	14	0	0	0.7	0	0	0	0	0	0	0	0	0	0	0	0.7	0	0	0	0	0	0
Ο	15	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0.7	0	0	0.8
Π	16	0	0	0	0	0.6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ρ	17	0.7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Σ	18	0	0	0.7	0	0	0	0	0	0	0	0.7	0	0	0	0	0	0	0	0	0	0
Τ	19	0	0.8	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Υ	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0
Φ	21	0	0	0	0	0	0.7	0	0	0	0	0	0.7	0	0	0	0	0	0	0	0	0.7
Χ	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ψ	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0
Ω	24	0	0	0	0	0	0.8	0	0	0	0	0	0.8	0	0	0	0	0	0.7	0	0	0

2 Iteration:

I = {2, 3, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24}, C_2 = { }.

$a_{8\,15}$ = 0.8 are maximum. Then C_2= {8, 15}.

$a_{8\,24} + a_{15\,24}$ = 1.6 are maximum. Then C_2 = {8, 15, 24}.

$a_{8\,21} + a_{15\,21} + a_{24\,21}$ = 2.1 are maximum. Then C_2 = {8, 15, 24, 21}.

There are no j such that $a_{8\,j} + a_{15\,j} + a_{24\,j} + a_{1\,j}$ is maximum. Then final C_2= {8, 15, 24, 21} = **{ Θ, Ο, Ω, Φ }**

After deleting 8, 15, 24, 21 rows and columns the table is:

		В	Г	Е	Z	Н	I	К	М	N	Ξ	П	Р	Σ	Т	Y	X	Ψ
		2	3	5	6	7	9	10	12	13	14	16	17	18	19	20	22	23
В	2	0	0	0	0	0	0	0	0	0	0	0	0.7	0	0	0	0	0
Г	3	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0.8	0	0	0
Е	5	0	0	0	0	0	0	0	0	0	0.7	0	0	0.7	0	0	0	0
Z	6	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0
Н	7	0	0	0	0	0	0	0	0	0	0	0.6	0	0	0	0	0	0
I	9	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0
К	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
М	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N	13	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0
Ξ	14	0	0	0.7	0	0	0	0	0	0	0	0	0	0.7	0	0	0	0
П	16	0	0	0	0	0.6	0	0	0	0	0	0	0	0	0	0	0	0
Р	17	0.7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Σ	18	0	0	0.7	0	0	0	0	0	0	0.7	0	0	0	0	0	0	0
Т	19	0	0.8	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0
Y	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8
X	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ψ	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0

3 Iteration:

I = {2, 3, 5, 6, 7, 9, 10, 12, 13, 14, 16, 17, 18, 19, 20, 22, 23}, C_3 = { }.

$a_{19\,9}$ = 0.8 are maximum. Then C_3= {19, 9}.

$a_{19\,3} + a_{9\,3}$ = 1.6 are maximum. Then C_3= {19, 9, 3}.

There are no j such that $a_{19\,j} + a_{9\,j} + a_{3\,j}$ is maximum. Then final C_3= {19, 9, 3} = **{ T, I, Г }**

After deleting 19, 9, 3 rows and columns the table is:

		В	Е	Z	Н	К	М	N	Ξ	П	Р	Σ	Υ	Х	Ψ
		2	5	6	7	10	12	13	14	16	17	18	20	22	23
В	2	0	0	0	0	0	0	0	0	0	0.7	0	0	0	0
Е	5	0	0	0	0	0	0	0	0.7	0	0	0.7	0	0	0
Z	6	0	0	0	0	0	0	0.8	0	0	0	0	0	0	0
Н	7	0	0	0	0	0	0	0	0	0.6	0	0	0	0	0
К	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0
М	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N	13	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0
Ξ	14	0	0.7	0	0	0	0	0	0	0	0	0.7	0	0	0
П	16	0	0	0	0.6	0	0	0	0	0	0	0	0	0	0
Р	17	0.7	0	0	0	0	0	0	0	0	0	0	0	0	0
Σ	18	0	0.7	0	0	0	0	0	0.7	0	0	0	0	0	0
Υ	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8
Х	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ψ	23	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0

4 Iteration:

I = {2, 5, 6, 7, 10, 12, 13, 14, 16, 17, 18, 20, 22, 23}, C_4 = { }.

$a_{6\,13}$ = 0.8 are maximum. Then C_4= {6, 13}.

There are no j such that $a_{6\,j} + a_{13\,j}$ is maximum. Then final C_4= {6, 13} = **{ Z, N }**

After deleting 6, 13 rows and columns the table is:

		В	Е	Н	К	М	Ξ	П	Р	Σ	Y	Х	Ψ
		2	5	7	10	12	14	16	17	18	20	22	23
В	2	0	0	0	0	0	0	0	0.7	0	0	0	0
Е	5	0	0	0	0	0	0.7	0	0	0.7	0	0	0
Н	7	0	0	0	0	0	0	0.6	0	0	0	0	0
К	10	0	0	0	0	0	0	0	0	0	0	0	0
М	12	0	0	0	0	0	0	0	0	0	0	0	0
Ξ	14	0	0.7	0	0	0	0	0	0	0.7	0	0	0
П	16	0	0	0.6	0	0	0	0	0	0	0	0	0
Р	17	0.7	0	0	0	0	0	0	0	0	0	0	0
Σ	18	0	0.7	0	0	0	0.7	0	0	0	0	0	0
Y	20	0	0	0	0	0	0	0	0	0	0	0	0.8
Х	22	0	0	0	0	0	0	0	0	0	0	0	0
Ψ	23	0	0	0	0	0	0	0	0	0	0.8	0	0

5 Iteration:

I = {2, 5, 7, 10, 12, 14, 16, 17, 18, 20, 22, 23}, C_5 = { }.

$a_{23\,20}$ = 0.8 are maximum. Then C_5= {23, 20}.

There are no j such that $a_{23\,j} + a_{20\,j}$ is maximum. Then final C_5= {23, 20} = **{ Ψ, Y }**

After deleting 23, 20 rows and columns the table is:

		В	Е	Н	К	М	Ξ	П	Р	Σ	Х
		2	5	7	10	12	14	16	17	18	22
В	2	0	0	0	0	0	0	0	0.7	0	0
Е	5	0	0	0	0	0	0.7	0	0	0.7	0
Н	7	0	0	0	0	0	0	0.6	0	0	0
К	10	0	0	0	0	0	0	0	0	0	0
М	12	0	0	0	0	0	0	0	0	0	0
Ξ	14	0	0.7	0	0	0	0	0	0	0.7	0
П	16	0	0	0.6	0	0	0	0	0	0	0
Р	17	0.7	0	0	0	0	0	0	0	0	0
Σ	18	0	0.7	0	0	0	0.7	0	0	0	0
Х	22	0	0	0	0	0	0	0	0	0	0

6 Iteration:

I = {2, 5, 7, 10, 12, 14, 16, 17, 18, 22}, C_6 = { }.

$a_{14\,5}$ = 0.7 are maximum. Then C_6= {14, 5}.

$a_{14\,18}$ + $a_{5\,18}$ = 1.4 are maximum. Then C_6= {14, 5, 18}.

There are no j such that $a_{14\,j}$ + $a_{5\,j}$ + $a_{18\,j}$ is maximum. Then final C_6= {14, 5, 18} = **{ Ξ, E, Σ }**

After deleting 14, 5, 18 rows and columns the table is:

		B	H	K	M	П	P	X
		2	7	10	12	16	17	22
B	2	0	0	0	0	0	0.7	0
H	7	0	0	0	0	0.6	0	0
K	10	0	0	0	0	0	0	0
M	12	0	0	0	0	0	0	0
П	16	0	0.6	0	0	0	0	0
P	17	0.7	0	0	0	0	0	0
X	22	0	0	0	0	0	0	0

7 Iteration:

I = {2, 7, 10, 12, 16, 17, 22}, C_7 = { }.

$a_{17\,2}$ = 0.7 are maximum. Then C_7= {17, 2}.

There are no j such that $a_{17\,j}$ + $a_{2\,j}$ is maximum. Then final C_7= {17, 2} = **{ P, B }**

After deleting 17, 2 rows and columns the table is:

		Н	К	М	П	Х
		7	10	12	16	22
Н	7	0	0	0	0.6	0
К	10	0	0	0	0	0
М	12	0	0	0	0	0
П	16	0.6	0	0	0	0
Х	22	0	0	0	0	0

8 Iteration:

I = {7, 10, 12, 16, 22}, C_7 = { }.

$a_{16\,7}$ = 0.6 are maximum. Then C_7= {16, 7}.

There are no j such that $a_{16\,j} + a_{7\,j}$ is maximum. Then final C_7= {16, 7} = **{ П, Н }**

After deleting 16, 7 rows and columns the table is:

		К	М	Х
		10	12	22
К	10	0	0	0
М	12	0	0	0
Х	22	0	0	0

Now $a_{10\,12} = a_{10\,22} = a_{12\,22} = 0$. Then {10}, {12}, {22} are three separated clusters. **({K}, {M}, {X})**

In this way, when α = 0.55, we have **11 clusters:**

{ Δ, Λ, А }, { Θ, О, Ω, Φ }, { Т, І, Г }, { Z, N }, { Ψ, Y }, { Ξ, Е, Σ }, { Р, В }, { П, Н }, {К}, {М}, {Х}