

STROJNO UČENJE

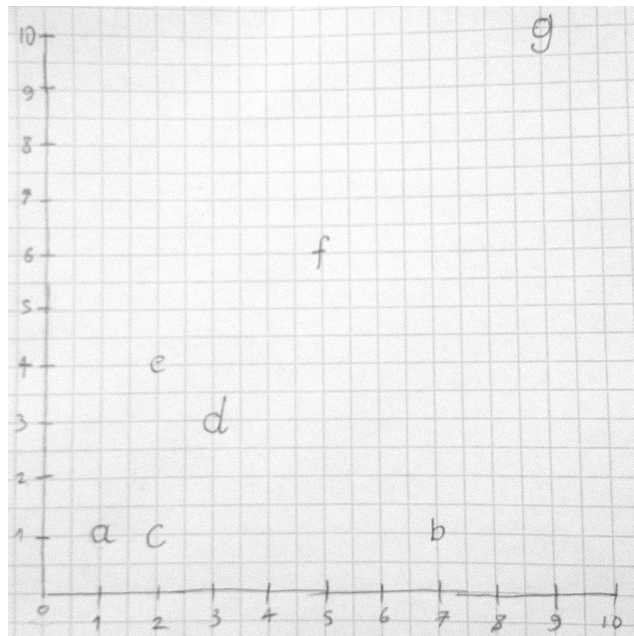
5. Domaća Zadaća

Krešimir Špes

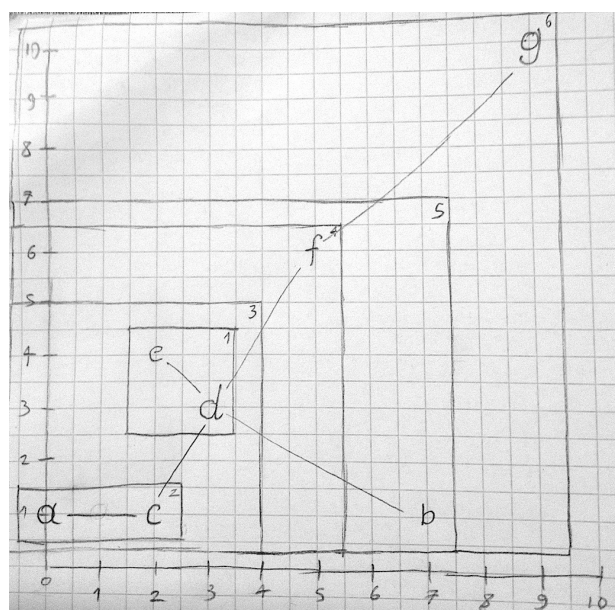
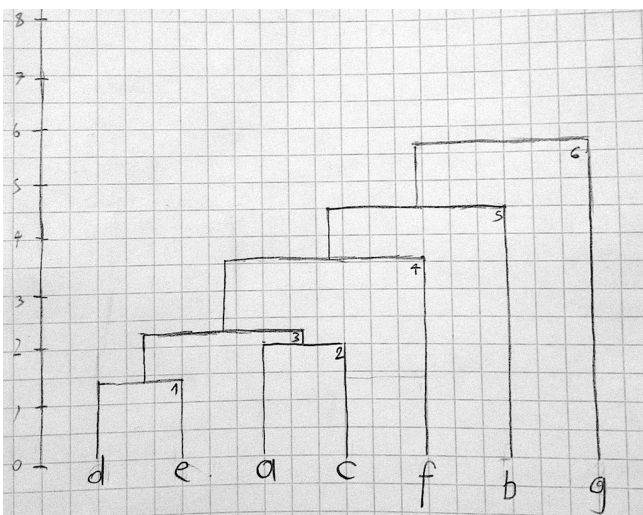
0036419866

ak. god. 2011. / 2012.

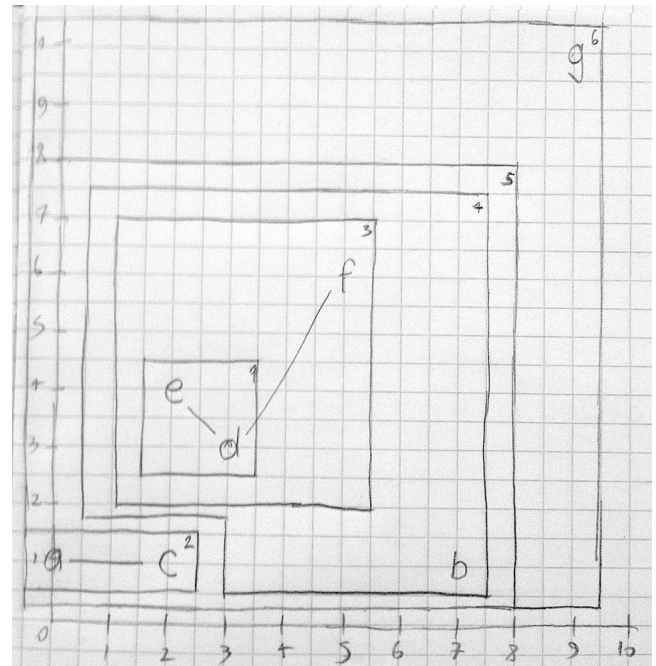
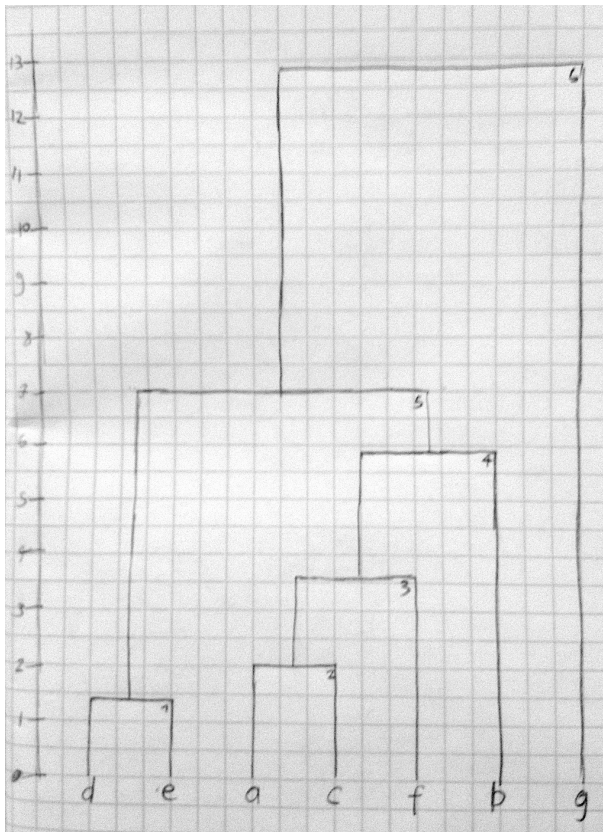
a)



Lijevo je dan dendrogram a desno grafička ilustracija grupa. U gornjem desnom kutu svake grupe je dana brojčana oznaka (vidljiva i u dendrogramu) radi lakše vizualizacije hijerarhije. Crte u desnom crtežu odznacavaju udaljenosti grupa koje su u svakom koraku bile najmanje.



b)



c) Gledajući vizualno, najbolje je odrezati dendrogram tamo gdje su veliki skokovi u udaljenostima. U prethodnom primjeru (b podzadatak) to bi najbolje bilo na udaljenosti 10 ili možda 6.5.

u prvom slučaju dobili i grupe $\{a,b,c,d,e,f\}$ i $\{g\}$ a u drugom $\{e,d\}$, $\{a,b,c,f\}$ i $\{g\}$

d)

Matrica udaljenosti između primjera:

PRIMJER	a	b	c	d	e	f	g
a	0	7.0	2.0	3.6	3.6	7.07	12.72
b	7.0	0	5.0	4.47	5.83	5.38	9.21
c	2.0	5.0	0	2.23	3.0	5.83	11.4
d	3.6	4.47	2.23	0	1.41	3.6	9.21
e	3.6	5.83	3.0	1.41	0	3.6	9.21
f	7.07	5.38	5.83	3.6	3.6	0	5.65
g	12.72	9.21	11.4	9.21	9.21	5.65	0

Matrica sličnosti između primjera:

$$s(\mathbf{x}^a, \mathbf{x}^b) = \frac{1}{1 + d(\mathbf{x}^a, \mathbf{x}^b)} \cdot \quad ==> \quad \text{korištena formula za izračun mjere sličnosti iz udaljenosti}$$

PRIMJER	a	b	c	d	e	f	g
a	1.0	0.12	0.33	0.22	0.22	0.12	0.07
b	0.12	1.0	0.17	0.18	0.15	0.16	0.1
c	0.33	0.17	1.0	0.31	0.25	0.15	0.08
d	0.22	0.18	0.31	1.0	0.41	0.22	0.1
e	0.22	0.15	0.25	0.41	1.0	0.22	0.1
f	0.12	0.16	0.15	0.22	0.22	1.0	0.15
g	0.07	0.10	0.08	0.10	0.10	0.15	1.0

Nad matricom sličnosti, od algoritama grupiranja koje smo upoznali kroz ovaj kolegij mogu raditi: k-medioda, maksimizacija očekivanja, algoritam hijerarhijskog grupiranja uz metode jednostruke ili potpune povezanosti