



Foto: Stefan Ernst, *Gartenkreuzspinne* / *Araneus diadematus*

Analiza rodovnikov s programom Pajek

Andrej Mrvar

Univerza v Ljubljani
Slovenija

Viri rodoslovnih podatkov

Ljudje zbirajo rodoslovne podatke za različne namene:

- Raziskave različnih kultur v sociologiji, antropologiji in zgodovini – sorodstvo kot osnovna socialna relacija
- Rodovniki družin in / ali teritorialnih enot, npr.,
 - rodovniki Mormonov: <http://www.familytreemaker.com/>
 - rodovnik ljudi živečih na področju Škofje Loke:
<http://genealogy.ijp.si>
 - rodovnik ameriških predsednikov:
<ftp://www.dcs.hull.ac.uk/public/genealogy/>
- Posebni rodovniki
 - Študenti in njihovi mentorji pri izdelavi doktoratov:
 - * Theoretical Computer Science Genealogy:
<http://sigact.acm.org/genealogy/>
 - * Mathematics

Standard GEDCOM

GEDCOM je standard za shranjevanje rodoslovnih podatkov, ki se uporablja tudi za izmenjavo podatkov pridobljenih z različnimi programi za vnos. Naslednje vrstice so izločene iz datoteke GEDCOM Evropskih kraljevih družin.

```
0 HEAD
1 FILE ROYALS.GED
...
0 @I58@ INDI
1 NAME Charles Philip Arthur/Windsor/
1 TITL Prince
1 SEX M
1 BIRT
2 DATE 14 NOV 1948
2 PLAC Buckingham,Palace,London,England
1 CHR
2 DATE 15 DEC 1948
2 PLAC Buckingham,Palace,Music Room,England
1 FAMS @F16@
1 FAMC @F14@
...
```

...

0 @I65@ INDI
1 NAME Diana Frances /Spencer/
1 TITL Lady
1 SEX F
1 BIRT
2 DATE 1 JUL 1961
2 PLAC Park House, Sandringham, Norfolk, England
1 CHR
2 PLAC Sandringham, Church, Norfolk, England
1 FAMS @F16@
1 FAMC @F78@

...

0 @I115@ INDI
1 NAME William Arthur Philip/Windsor/
1 TITL Prince
1 SEX M
1 BIRT
2 DATE 21 JUN 1982
2 PLAC St. Mary's Hosp., Paddington, London, England
1 CHR
2 DATE 4 AUG 1982
2 PLAC Music Room, Buckingham, Palace, England
1 FAMC @F16@

...



...

0 @I116@ INDI

1 NAME Henry Charles Albert/Windsor/

1 TITL Prince

1 SEX M

1 BIRT

2 DATE 15 SEP 1984

2 PLAC St. Mary's Hosp., Paddington, London, England

1 FAMC @F16@

...

...

0 @F16@ FAM

1 HUSB @I58@

1 WIFE @I65@

1 CHIL @I115@

1 CHIL @I116@

1 DIV N

1 MARR

2 DATE 29 JUL 1981

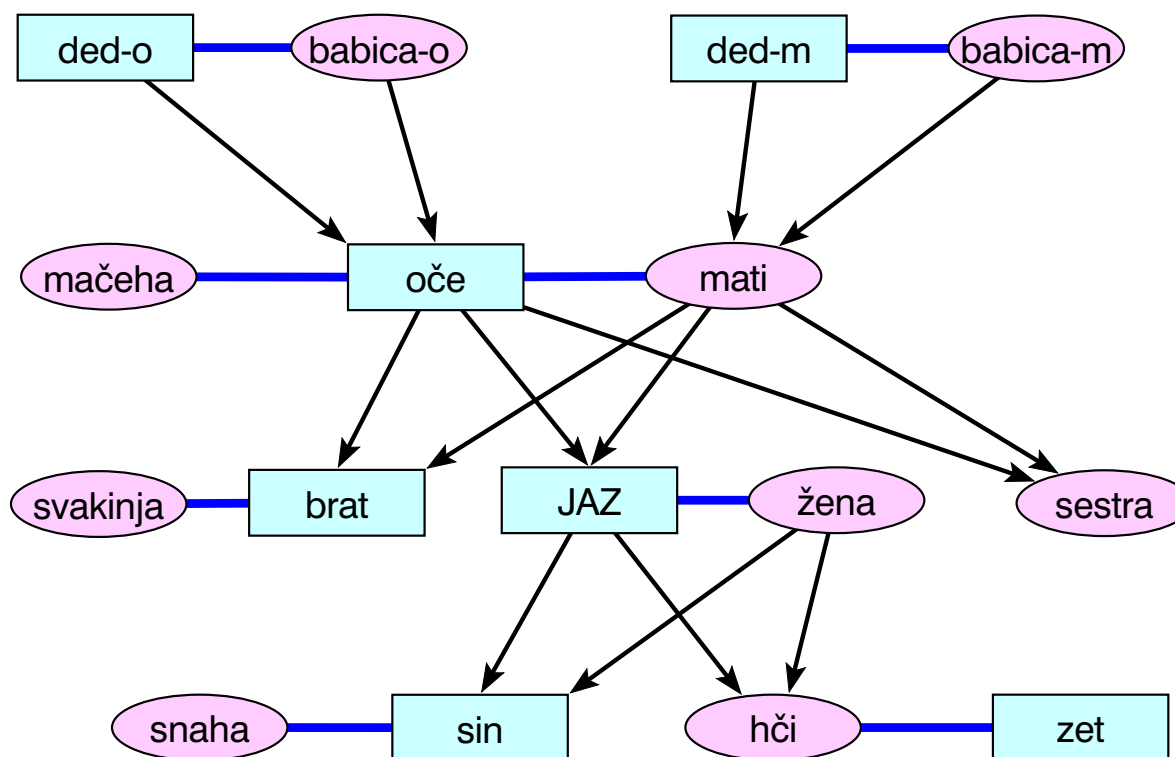
2 PLAC St. Paul's, Cathedral, London, England

Predstavitev rodovnikov z omrežji

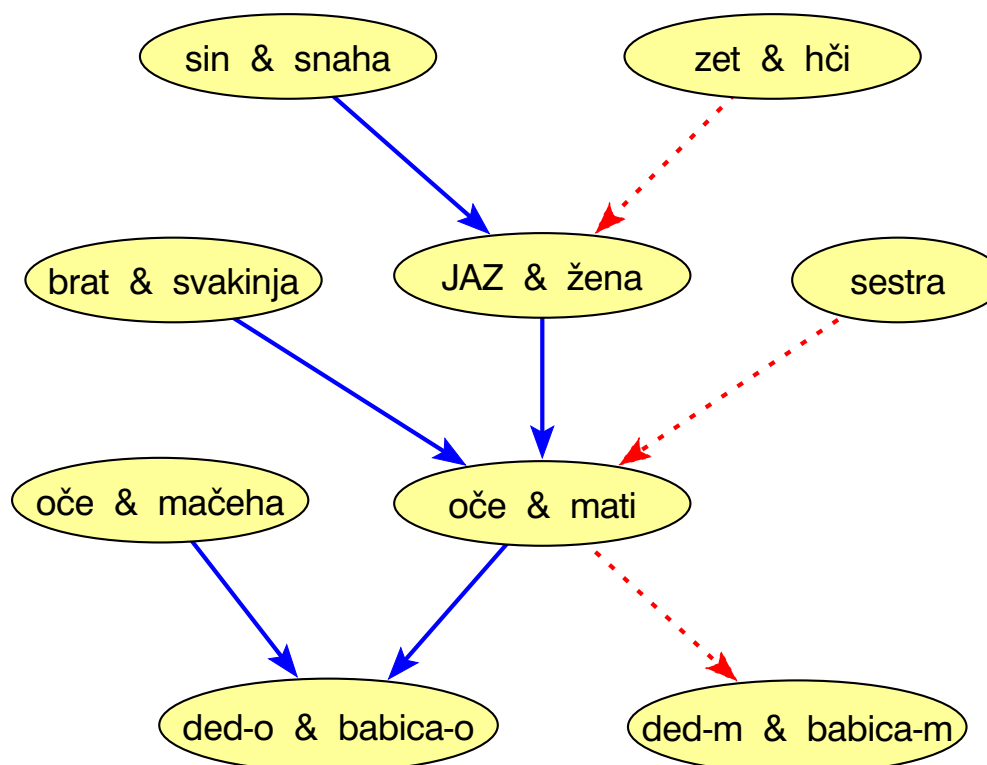
Rodovnike lahko predstavimo z omrežji na več različnih načinov:

- kot navadne rodovnike (Ore-graph),
- kot parne rodovnike (p-graph),
- kot dvodelne parne rodovnike (bipartite p-graph).

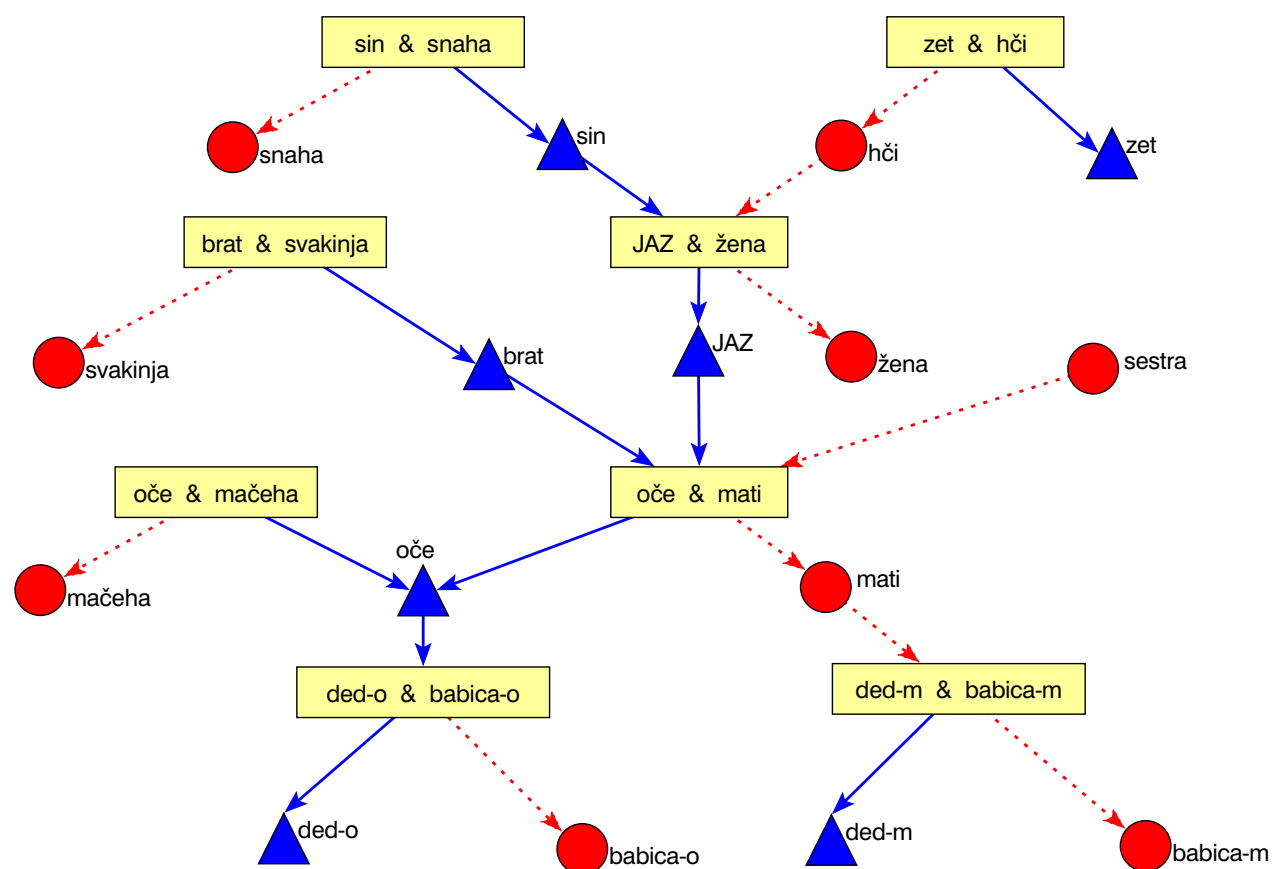
Navadni rodovnik: V navadnem rodovniku je vsaka oseba predstavljena s svojo točko, poroke so predstavljene z neusmerjenimi povezavami, relacija *je eden od staršev od* pa z usmerjenimi povezavami, ki kažejo od vsakega od staršev do njihovih otrok.



Parni rodovnik: V parnem rodovniku predstavljajo točke posameznike ali pare. Če neka oseba še ni poročena, je predstavljena s svojo točko, sicer je predstavljena s svojim partnerjem v skupni točki. V parnem rodovniku obstajajo samo usmerjene povezave, ki kažejo od otrok na njihove starše.



Dvodelni parni rodovnik: ima dve vrsti točk – točke, ki predstavljajo poročene pare (pravokotniki) in točke, ki predstavljajo posameznike (okrogle točke za ženske in trikotniki za moške). Usmerjene povezave tudi tu kažejo od otrok na njihove starše.

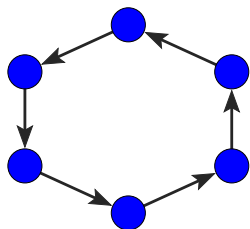


Rodovniki so redka omrežja

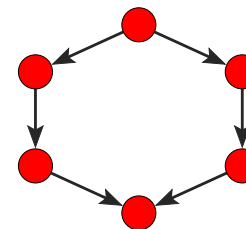
	navadni rodovnik				parni rodovnik			
podatki	$ V $	$ E $	$ A $	$\frac{ L }{ V }$	$ V_{ip} $	$ V_{cp} $	$ A_p $	$\frac{ A_p }{ V_p }$
Bruno	15512	4841	18664	1.52	6000	5289	10053	0.89
Combo	20350	7248	26199	1.64	6931	7945	14845	1.00
Dodderer	16761	5650	22425	1.68	6029	5652	11765	1.01
Drame	29606	8256	41814	1.69	13254	8939	21862	0.99
Little	25968	8778	34640	1.67	9212	8850	18233	1.01
President	2145	978	2223	1.49	218	1042	1222	0.97
Tillotsn	42559	12796	54043	1.57	15177	15959	31234	1.00
Loka	47956	14154	68052	1.71	19189	16039	36192	1.03
Silba	6427	2217	9627	1.84	2001	2479	5281	1.18
Ragusa	5999	2002	9315	1.88	2066	2310	5336	1.22
Tur	1269	407	1987	1.89	0	956	1114	1.17
Royal	3010	1138	3724	1.62	719	1422	2259	1.06

Prednosti parnih rodovnikov

- v parnem rodovniku je manj točk in povezav;
- parni rodovniki so usmerjena aciklična omrežja



Sklenjena veriga:



- v parnem rodovniku pomeni vsaka sklenjena veriga (semi-cycle) *prepletenost porok*. Obstajata dve vrsti prepletenosti:
 - krvne poroke: npr., poroka med bratom in sestro.
 - ne-krvne poroke: npr., dva brata se poročita z dvema sestrama iz druge družine.

Pri dvodelnih parnih rodovnikih lahko dodatno razlikujemo še med *poročenim stricem in ponovno poročenim očetom* ter tako poiščemo npr. *poroke med polbrati in polsestrami*.

Indeks prepletenosti

Indeks prepletenosti (*relinking index*) meri, kako pogosto so se pripadniki istih družin poročali med sabo. Poseben primer prepletenosti so krvne poroke.

Naj bo n število točk v parnem rodovniku, m število povezav in M število končnih točk (točk z izhodno stopnjo 0, $M \geq 1$).

V povezanem rodovniku velja

$$RI = \frac{m - n + 1}{n - 2M + 1}$$

Za omrežje z eno samo točko postavimo $RI = 0$.

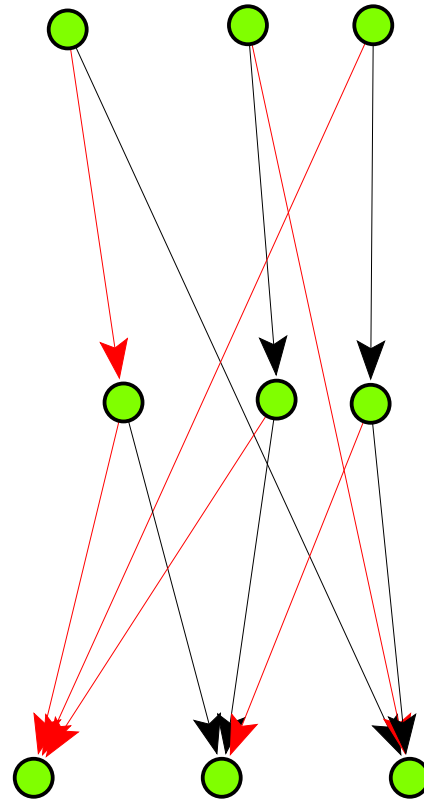
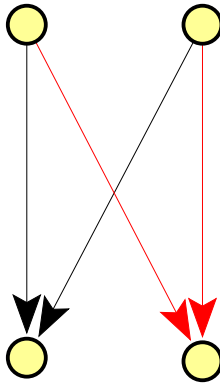
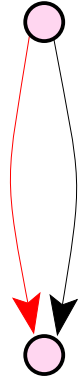
* $0 \leq RI \leq 1$

* Če je rodovnik gozd/drevo, potem je $RI = 0$ (ni prepletenosti).

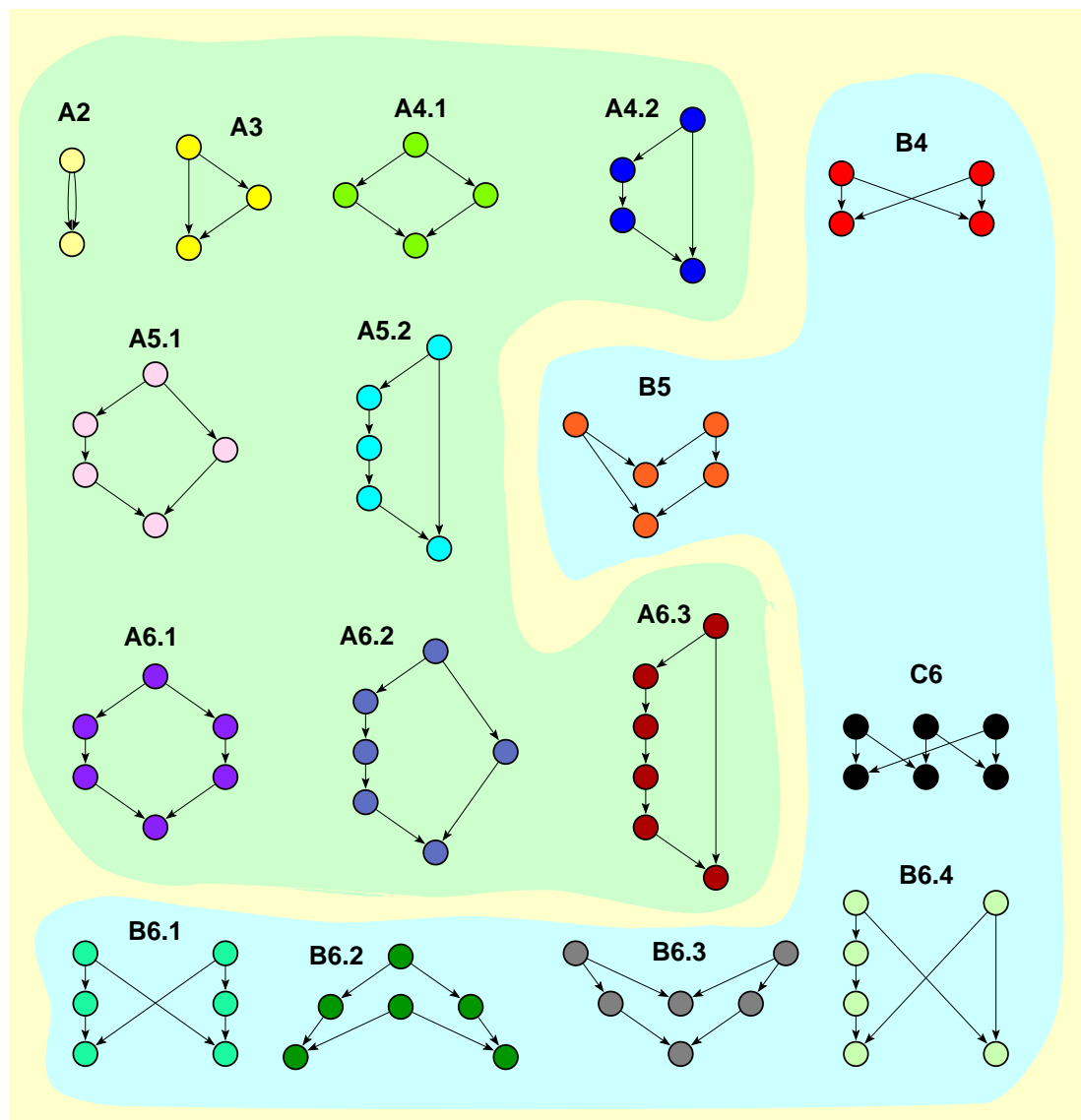
* Obstajajo rodovniki z $RI = 1$ (največja možna prepletenost).

* Prepletenost ponavadi računamo za največjo dvopovezano komponento.

Vzorci z indeksom prepletenosti enakim 1



Prepletene poroke (parni rodovniki z 2 do 6 točkami)



Primerjave rodovnikov

Rodovnike lahko primerjamo po tem kolikokrat se v njih pojavljajo karakteristični vzorci. Kot primer smo vzeli pet rodovnikov:

- Loka.ged – rodovnik s področja Škofje Loke (P. Hawlina).
- Silba.ged – rodovnik hrvaškega otoka Silba (P. Hawlina).
Poseben geografski položaj.
- Ragusa.ged – poroke med dubrovniškimi plemiškimi družinami med 12 in 16 stoletjem. Podatke je zbrala I. Mahnken (1960); v elektronsko obliko pa jih je vnesla P. Dremelj (1999).
Zelo omejena pravila dovoljenih porok – pripadnik plemiške družine naj bi se poročil s pripadnikom druge plemiške družine.
- Tur.ged – rodovnik turškega nomadskega plemena Yörük. Podatke sta zbrala Ulla C. Johansen in D.R. White (2001)
Prepletena poroka je znak pripravljenosti ostati v nomadski skupini.
- Royal.ged – rodovnik evropskih kraljevih družin.

















Silba, Dubrovnik



Aydin, jugozahodna Turčija



Frekvenčna porazdelitev vzorcev

vzorec	Loka	Silba	Ragusa	Tur	Royal	Σ
 A2	1	0	0	0	0	1
 A3	1	0	0	0	3	4
 A4.1	12	5	3	65	21	106
 B4	54	25	21	40	7	147
 A4.2	0	0	0	0	0	0
 A5.1	9	7	4	15	13	48
 A5.2	0	0	0	0	0	0
 B5	19	11	47	19	8	104
 A6.1	28	28	2	65	13	140
 A6.2	0	2	0	0	1	3
 A6.3	0	0	0	0	0	0
 C6	10	12	19	15	5	61
 B6.1	0	1	2	0	0	3
 B6.2	27	39	63	54	12	194
 B6.3	47	30	82	46	13	218
 B6.4	0	0	5	3	0	8
Št. pos.	47956	6427	5999	1269	3010	
Naj. dv. komp.	4095	1340	1446	250	435	
RI	0.55	0.78	0.74	0.75	0.37	

















Opazanja

- Generacijski preskoki za več kot eno generacijo so zelo malo verjetni.
- Obstaja veliko porok tipa B6.3 (dva vnuka se poročita v isto družino) in B6.2 (dve družini sta povezani s poroko med otrokoma in kasneje še enkrat s poroko med vnukoma)
- V rodovniku Tur obstaja veliko porok A4.1 in A6.1.
- V vseh rodovnikih je število 'ne-krvnih' porok veliko večje od števila krvnih porok (kar še posebej velja za rodovnik Ragusa, izjema je Royal). Za tako vrsto porok obstajajo ekonomski razlogi: ohraniti bogastvo in moč znotraj izbranih družin.

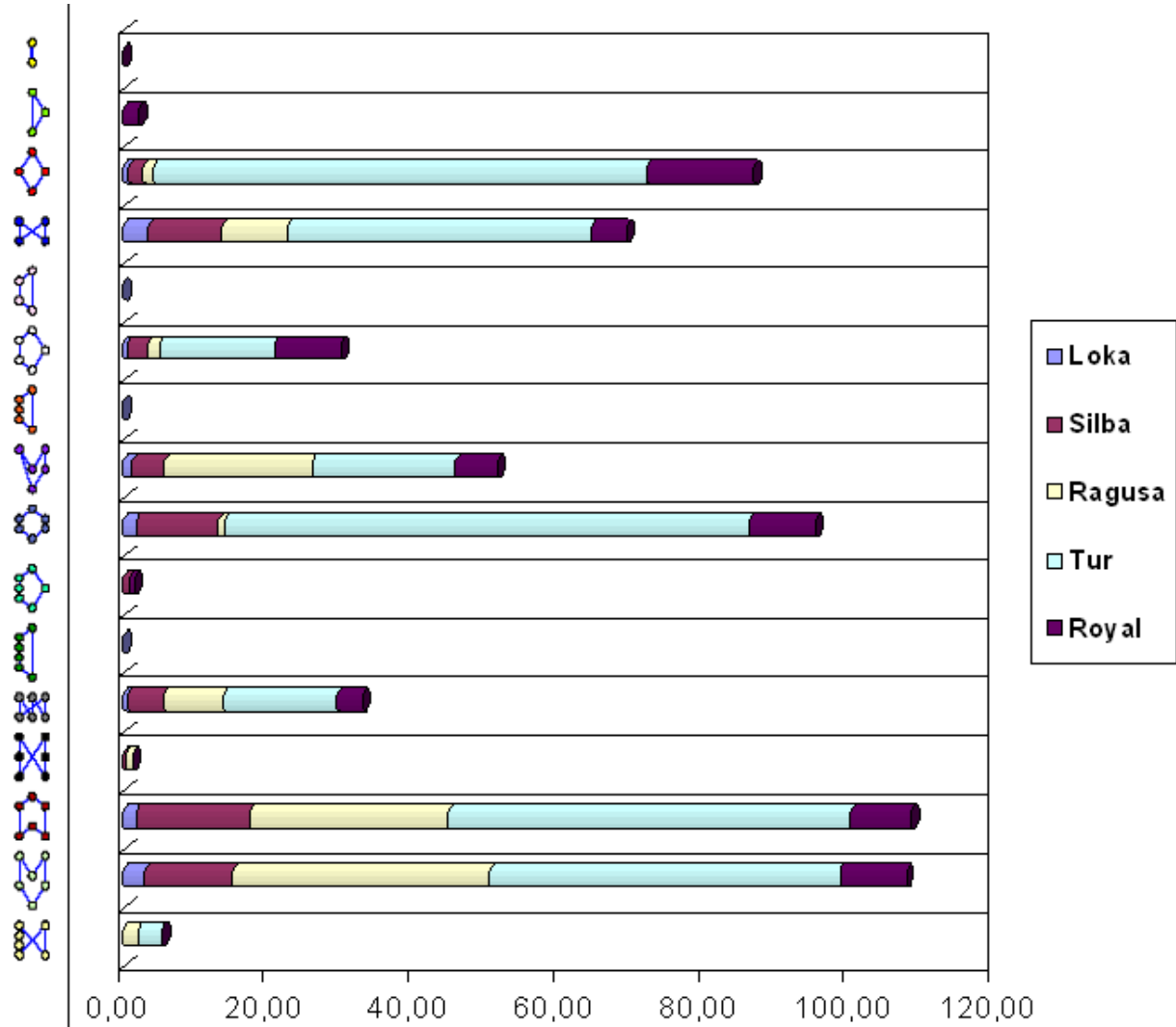
vrsta poroke	Loka	Silba	Ragusa	Tur	Royal
krvne poroke	51	42	9	149	51
ostale prepletene poroke	157	118	239	176	45

Število posameznikov v rodovniku Tur je veliko manjše kot v ostalih rodovnikih, Silba in Ragusa sta približno enako velika, medtem je Loka veliko večji rodovnik, kar moramo pri primerjavah pravtako upoštevati.

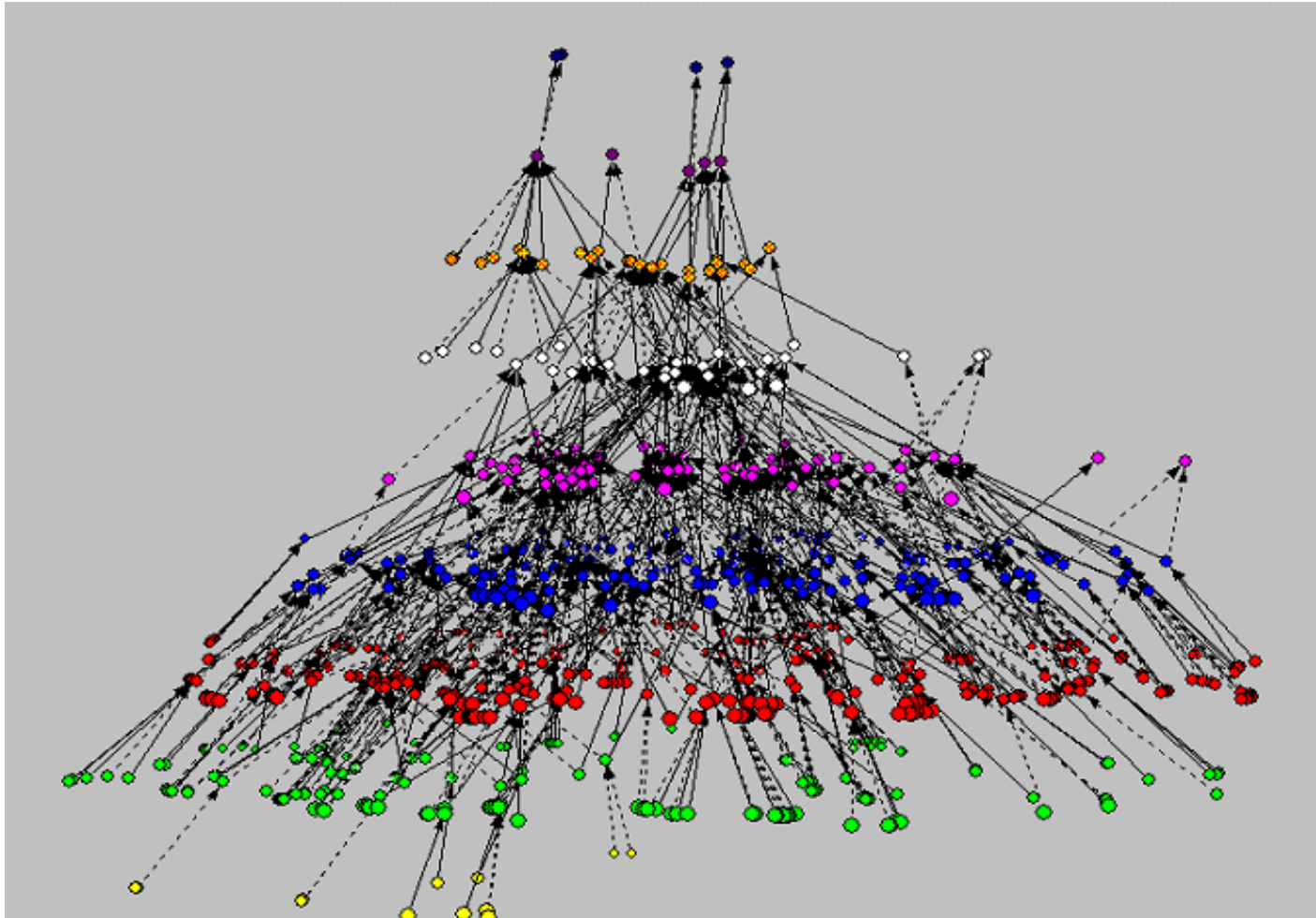
Frekvence normalizirane s številom parov v parnem rodovniku $\times 1000$

	pattern	Loka	Silba	Ragusa	Tur	Royal
	A2	0.07	0.00	0.00	0.00	0.00
	A3	0.07	0.00	0.00	0.00	2.64
	A4.1	0.85	2.26	1.50	159.71	18.45
	B4	3.82	11.28	10.49	98.28	6.15
	A4.2	0.00	0.00	0.00	0.00	0.00
	A5.1	0.64	3.16	2.00	36.86	11.42
	A5.2	0.00	0.00	0.00	0.00	0.00
	B5	1.34	4.96	23.48	46.68	7.03
	A6.1	1.98	12.63	1.00	169.53	11.42
	A6.2	0.00	0.90	0.00	0.00	0.88
	A6.3	0.00	0.00	0.00	0.00	0.00
	C6	0.71	5.41	9.49	36.86	4.39
	B6.1	0.00	0.45	1.00	0.00	0.00
	B6.2	1.91	17.59	31.47	130.22	10.54
	B6.3	3.32	13.53	40.96	113.02	11.42
	B6.4	0.00	0.00	2.50	7.37	0.00
	Σ	14.70	72.17	123.88	798.53	84.36

Normalizirane frekvence

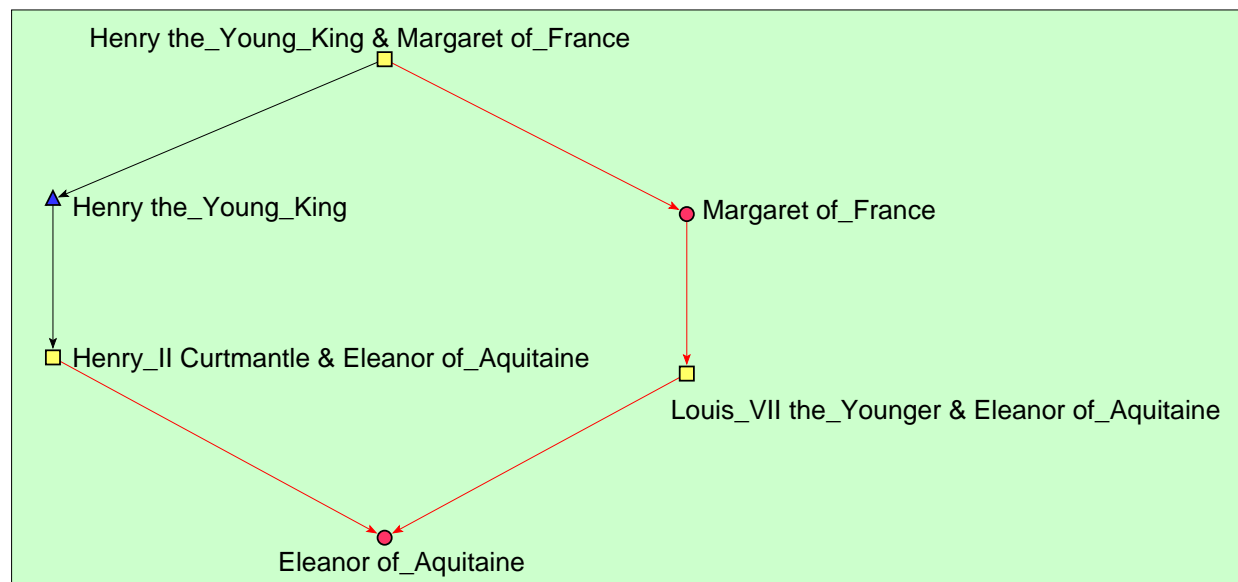
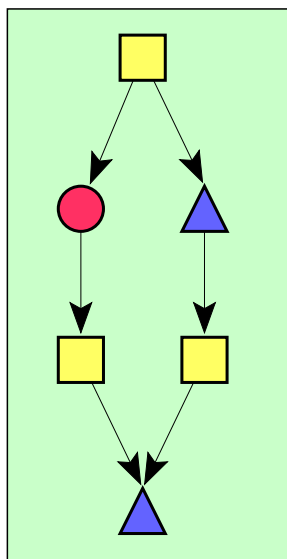


Prepletenost v rodovniku nomadskega plemena



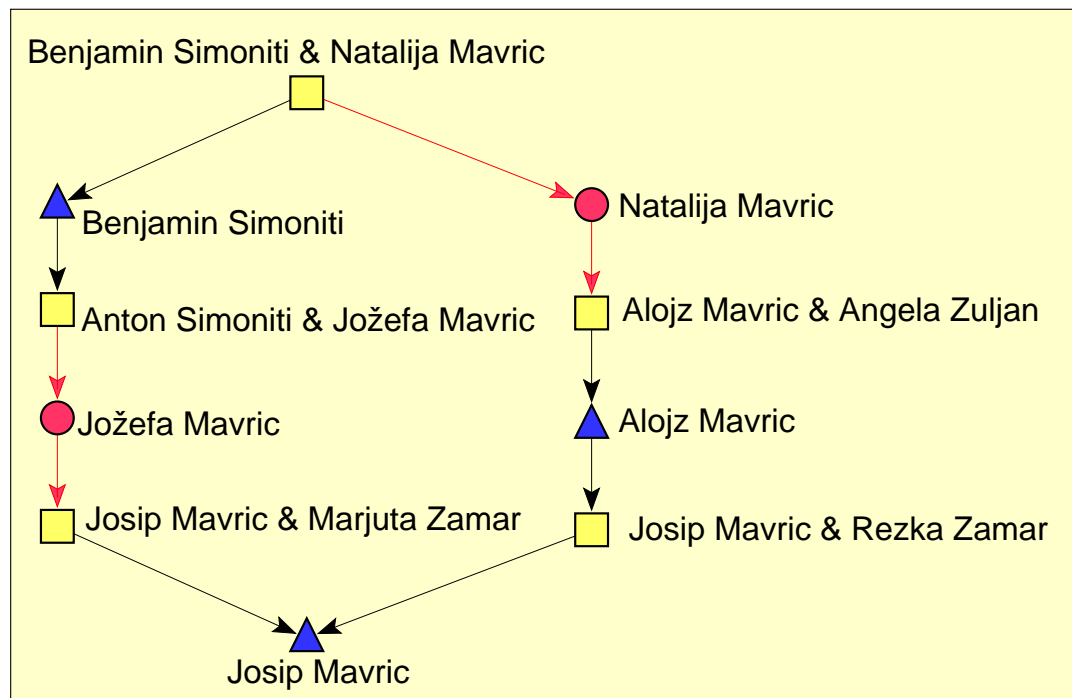
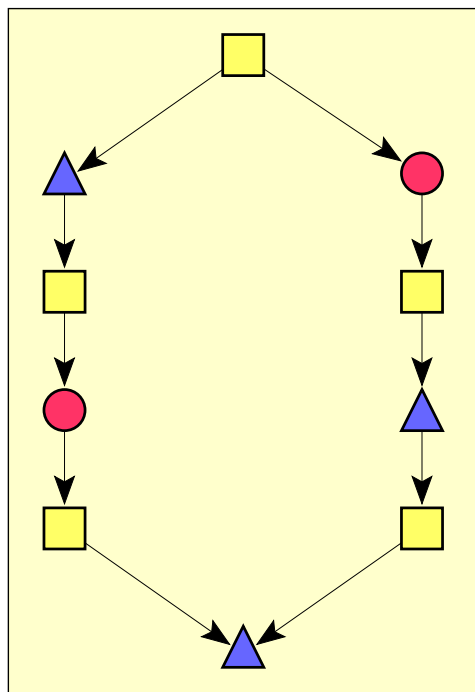
Dvodelni parni rodovniki: Poroka med polbratom in polsestro

Pri parnih rodovnikih iz same strukture omrežja ne moremo prepoznati večkrat poročenih oseb. V tem primeru moramo uporabiti dvodelne parne rodovnike. V njih lahko najdemo poroke med polbrati in polsestrami. V omenjenih petih rodovnikih smo našli le en tak primer v rodovniku Royal.ged.



Dvodelni parni rodovniki: Poroka med polbratranci

Pravtako ne obstaja veliko porok med polbratranci / polsestričnami. En tak primer smo našli v rodovniku Loka in štiri v rodovniku turških nomadskih plemen.

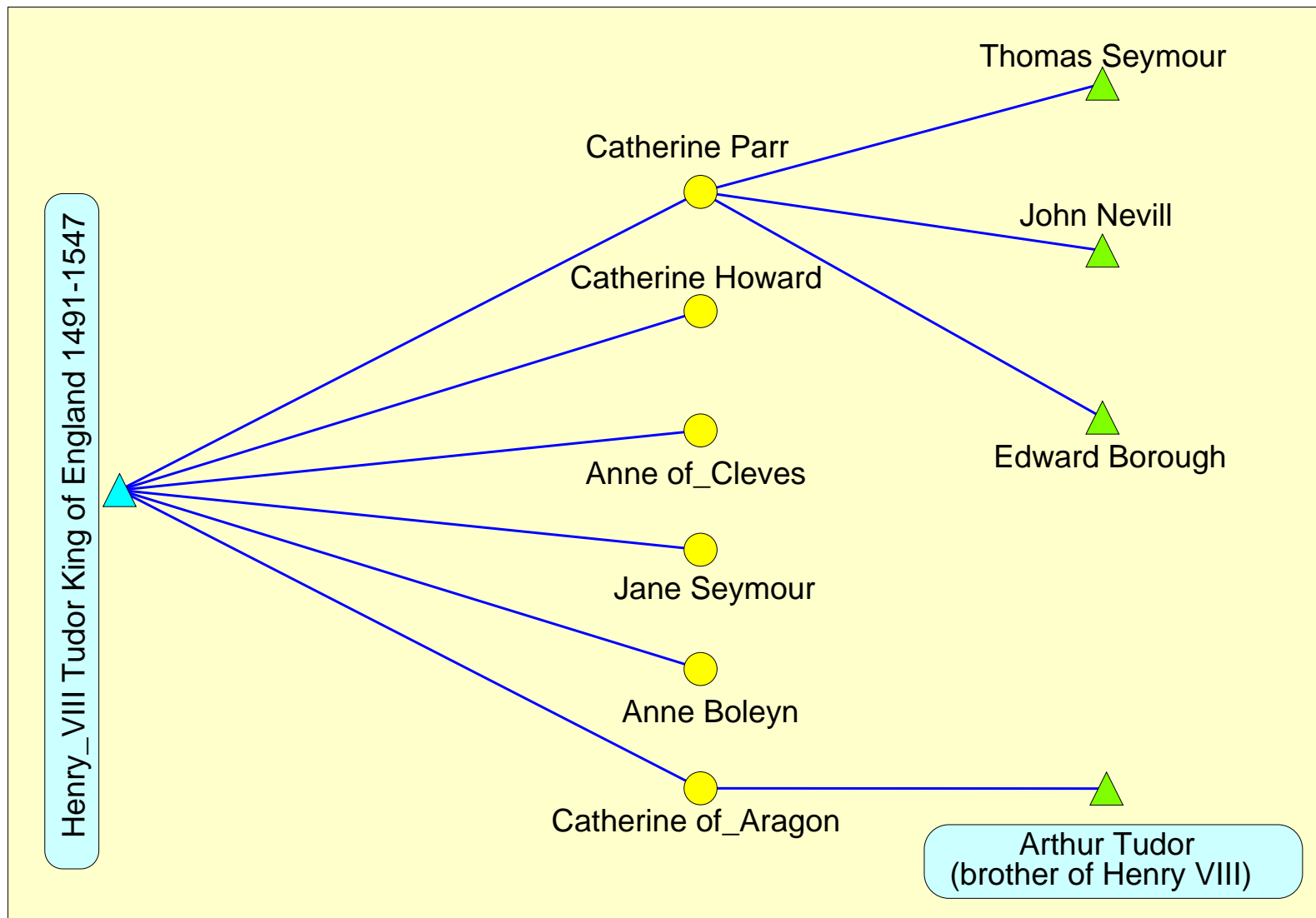


Druge analize

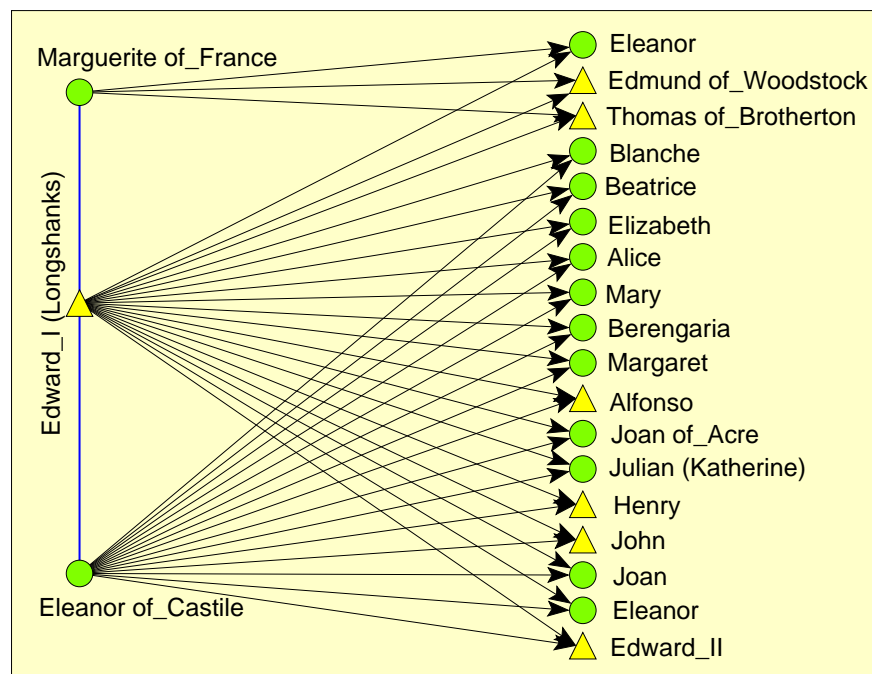
Posameznike, ki zbirajo podatke o svojih družinah, zanimajo še številne druge podrobnosti:

- posebnosti: velikokrat poročene osebe, posamezniki z velikim številom otrok;
- preverjanje ali sta dve osebi v sorodstveni zvezi in iskanje najkrajše sorodstvene vezi med njima;
- iskanje vseh prednikov/potomcev izbrane osebe in iskanje osebe z največjim znanim številom prednikov ali potomcev;
- največja razlika v starosti med možem in ženo, najstarejša/najmlajša oseba ob času poroke, najstarejša/najmlajša oseba ob rojstvu otroka;
- spreminjanje vzorcev porok skozi čas;
- iskanje najdaljše moške/ženske verige;
- posebni primeri → napake storjene pri vnosu podatkov (konsistentnost).

Največkrat poročena oseba...

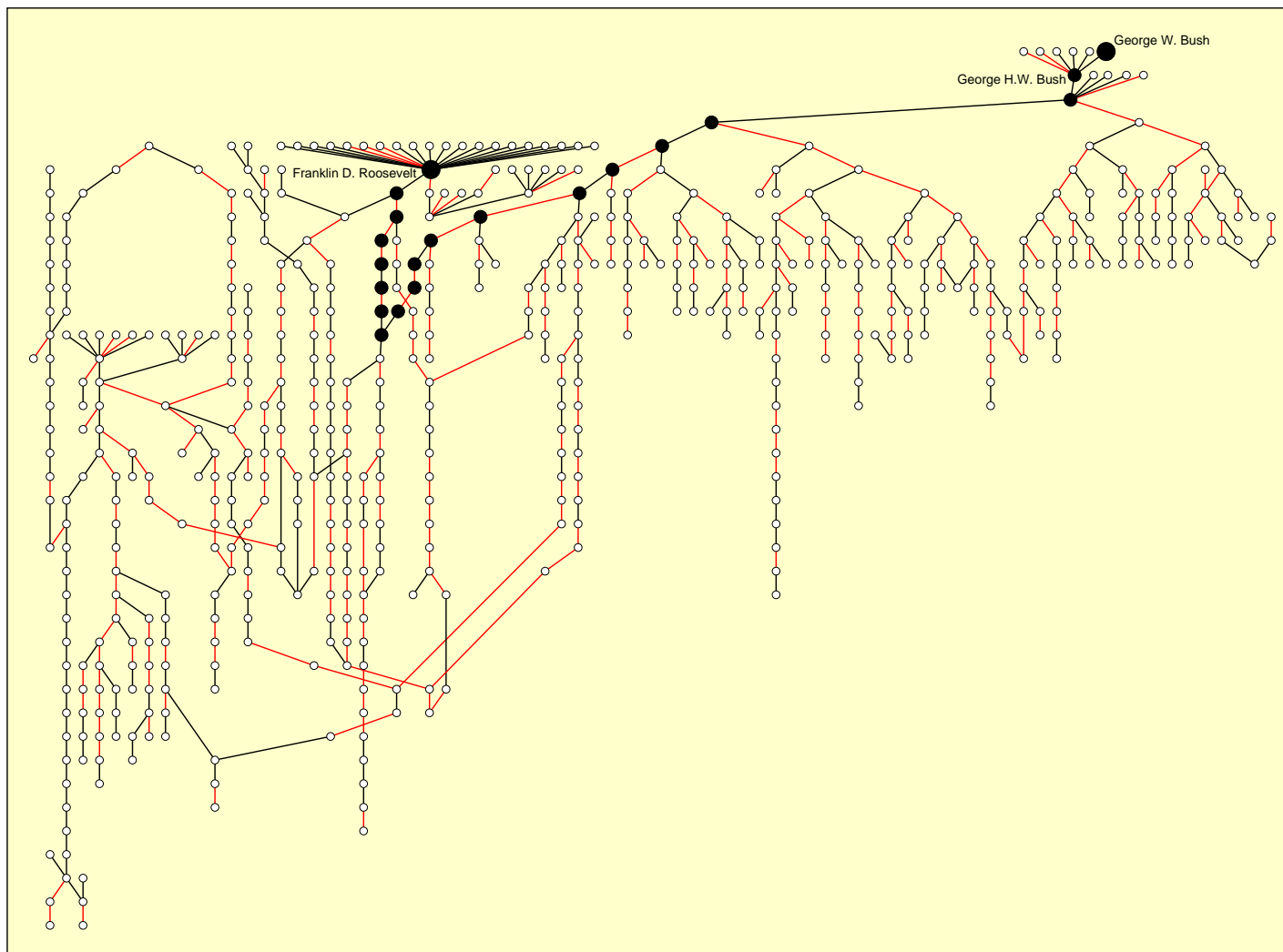


Največje število otrok...

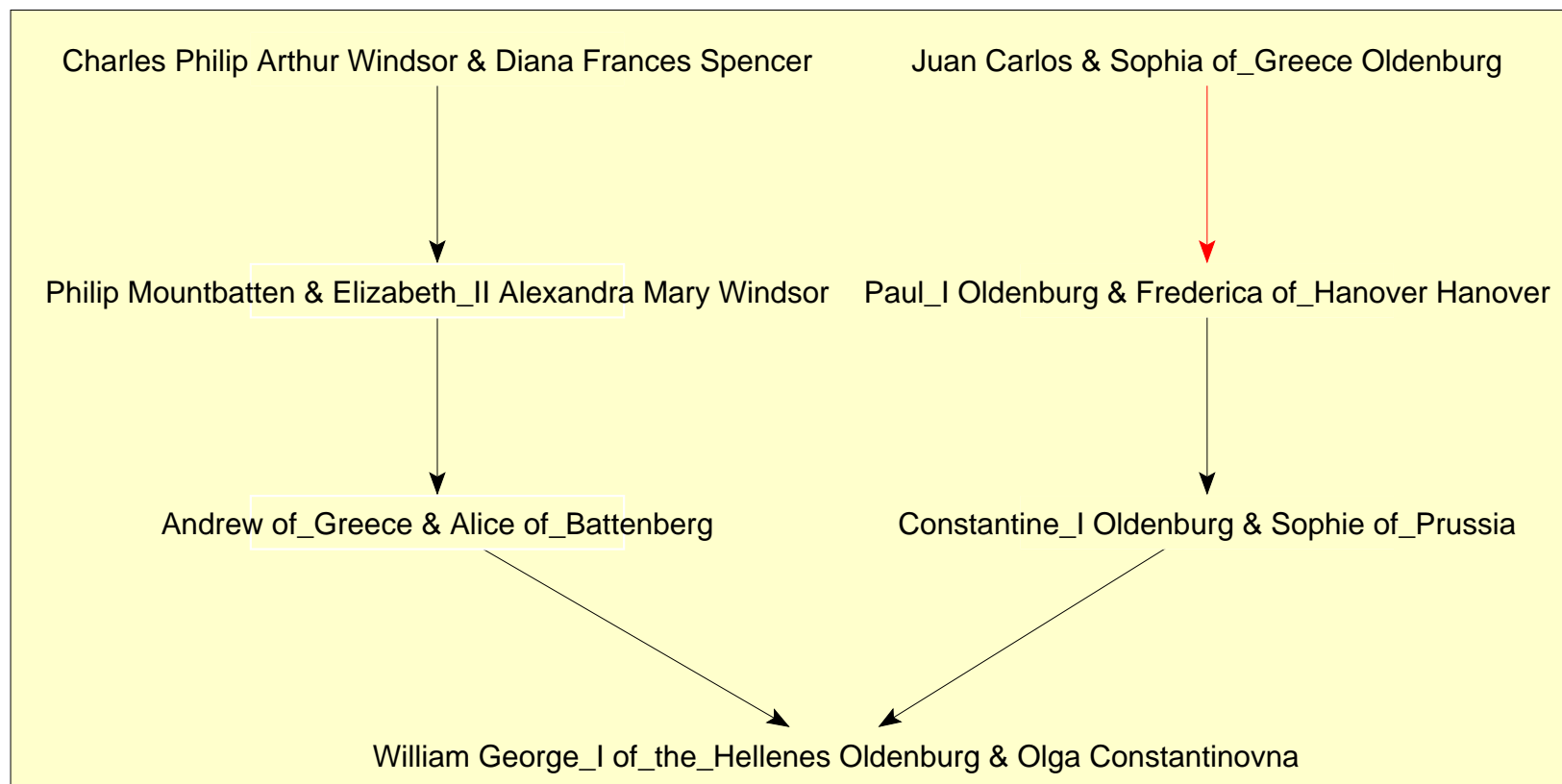


Angleški kralj Edward I (1237-1307) in njegova žena kraljica Eleanor (1241-1290) sta imela 16 otrok rojenih med letoma 1255 in 1284 (na sliki manjka hčerka brez imena). Najmlajši sin (Edward) je bil prvi od sinov, ki je preživel nevarna otroška leta. Eleanor je morala poskusiti šestnajstkrat, da je izpolnila najpomembnejšo nalogo angleške kraljice: svojemu možu zagotoviti naslednika moškega spola. Do desetega leta je umrlo deset od šestnajstih otrok, več kot 40 let so dočakali le trije.

Največja povezana komponenta v rodovniku ameriških predsednikov



Najkrajša sorodstvena pot med angleškim princem *Charlesom* in španskim kraljem *Juanom Carlosom* v Royal.ged



Osnovne sorodstvene relacije

V antropologiji se za osnovne sorodstvene relacije štejejo

(http://www.era.anthropology.ac.uk/Era_Resources/Era/Kinship/prologTerm2.html):

Tip	Angleško ime	Slovensko ime
P	Parent	Starši
F	Father	Oče
M	Mother	Mati
C	Child	Otrok
S	Son	Sin
D	Daughter	Hči
G	Sibling	Brat-sestra
B	Brother	Brat
Z	Sister	Sestra
E	Spouse	Zakonec
H	Husband	Mož
W	Wife	Žena

Računanje sorodstvenih relacij

Ko Pajek prebere datoteko GEDCOM kot navaden rodovnik, ustvari naslednje tri relacije:

F: *_je oče od_*

M: *_je mati od_*

E: *_je poročen z_*

Dodatno moramo ustvariti še dve dvojiški diagonalni matriki za razločevanje oseb po spolu:

L: *_moški_* / 1-moški, 0-ženska

J: *_ženska_* / 1-ženska, 0-moški

Ostale *osnovne* relacije lahko izračunamo z množenjem omrežij (matrik) in drugimi operacijami nad omrežji, ki so na voljo v Pajku.

<i>_je eden od staršev</i> _ (parent_of.mcr)	$P = F \cup M$
<i>_je otrok od</i> _ (child_of.mcr)	$C = P^T$
<i>_je sin od</i> _ (son_of.mcr)	$S = L * C$
<i>_je hči od</i> _ (daughter_of.mcr)	$D = J * C$
<i>_je mož od</i> _ (husband_of.mcr)	$H = L * E$
<i>_je žena od</i> _ (wife_of.mcr)	$W = J * E$
<i>_je brat-sestra od</i> _ (sibling_of.mcr)	$G = ((F^T * F) \cap (M^T * M)) \setminus I$
<i>_je brat od</i> _ (brother_of.mcr)	$B = L * G$
<i>_je sestra od</i> _ (sister_of.mcr)	$Z = J * G$

Izračunamo lahko še številne *izpeljane* relacije npr.:

_je stric od _ (uncle_of.mcr)

$$U = B * P$$

_je teta od _ (aunt_of.mcr)

$$A = Z * P$$

_je polbrat-polsestra od _ (semisibling_of.mcr)

$$G_e = (P^T * P) \setminus I$$

_je eden od starih staršev od _

$$GP = P^2$$

_je dedek od _

$$GF = F * P = L * GP$$

_je nečakinja od _

$$N_i = D * G$$

Lastnosti relacij

Tip	Ime	Sime.	Tran.	Acikl.
P	Parent			x
F	Father			x
M	Mother			x
C	Child			x
D	Daughter			x
S	Son			x
G	Sibling	x	x	
Z	Sister		x	
B	Brother		x	
E	Spouse	x		
H	Husband			x
W	Wife			x
U	Uncle			x
A	Aunt			x
Ge	Semi-sibling	x		

Velikosti sorodstvenih relacij

Tip	Turks	Ragusa	Loka	Silba	Royal
P-Starši	1987	9315	68052	9627	3724
F-Oče	1022	4956	34330	4998	2010
M-Mati	965	4359	33722	4629	1714
C-Otrok	1987	9315	68052	9627	3724
D-Hči	857	3577	32647	4518	1589
S-Sin	1130	5738	35405	5109	2135
G-Brat-sestra	2485	8782	69347	7803	2858
Z-Sestra	2256	6949	66874	7314	2634
B-Brat	2714	10615	71820	8292	3082
E-Zakonec	407	2002	14154	2217	1138
H-Mož	407	2002	14154	2217	1138
W-Žena	407	2002	14154	2217	1138
U-Stric	3816	16665	81695	11372	3453
A-Teta	3477	10644	80995	10564	2973
Ge-Polbrat/sestra	2926	10763	76746	8972	3372
# Posameznikov	1269	5999	47956	6427	3010

Relativne velikosti sorodstvenih relacij

Kin Type	Turks	Ragusa	Loka	Silba	Royal
P-Starši	1.000	1.000	1.000	1.000	1.000
F-Oče	0.514	0.532	0.504	0.519	0.540
M-Mati	0.486	0.468	0.496	0.481	0.460
C-Otrok	1.000	1.000	1.000	1.000	1.000
D-Hči	0.431	0.384	0.480	0.469	0.427
S-Sin	0.569	0.616	0.520	0.531	0.573
G-Brat-sestra	1.250	0.943	1.019	0.811	0.767
Z-Sestra	1.135	0.746	0.983	0.760	0.707
B-Brat	1.366	1.140	1.055	0.861	0.828
E-Zakonec	0.205	0.215	0.208	0.230	0.306
H-Mož	0.205	0.215	0.208	0.230	0.306
W-Žena	0.205	0.215	0.208	0.230	0.306
U-Stric	1.920	1.789	1.200	1.181	0.927
A-Teta	1.750	1.143	1.190	1.097	0.798
Ge-Polbrat/sestra	1.473	1.155	1.128	0.932	0.905