Analysis of Genealogies
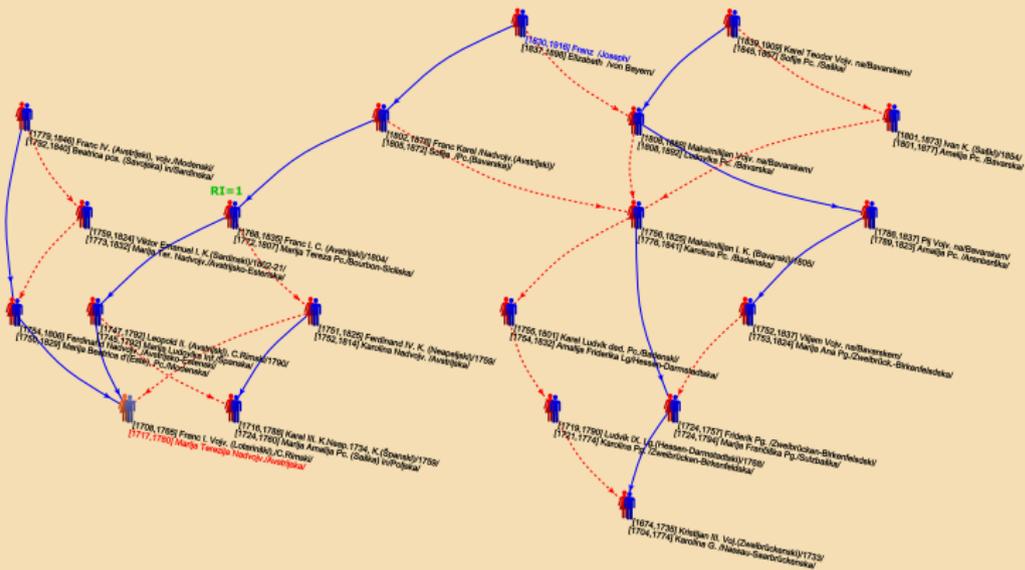
Andrej Mrvar

Sources of genealogies

People collect genealogical data for several different purposes:

- Research of different cultures in sociology, anthropology and history – kinship as fundamental social relation
- Genealogies of families and/or territorial units, e.g.,
  - genealogy of Ragusan (Dubrovnik) nobel families
  - Mormons genealogy
    `http://www.familytreemaker.com/`
  - genealogy of Škofja Loka district
  - genealogy of American presidents
- Special genealogies
  - Students and their PhD thesis advisors:
    - Theoretical Computer Science Genealogy:
      `http://sigact.acm.org/genealogy/`
    - Mathematics
      `https://genealogy.math.ndsu.nodak.edu/`

Pajek

Ibn Sina Abu ʿAli al-Husayn
Ibn al-Maraüübi Bahmanyär
al-Nïdäbüri Ghiyäth al-Din Abü al-Fath, 1068
al-Max Gdi al-Marwazï Saraf al-Din Muhammad
al-Tüxï Sharaf al-Din
al-Rïdï Fakhr al-Din Muhammad
Ibn Yünus Kamäl al-Din
Ibrähïm al-Misrï Qutb al-Din, 1222
al-Mukhirisï al-Asharï Afdr al-Din, 1264
al-Tüsï Nasïr al-Din
al-Bekhärï Shams al-Din
Chortlade Gregory, 1296
Bryennios Manuel, 1300
Metochites Theodore, 1315
Palamas Gregory, 1316
Kabasilas Nilos, 1363
Kydones Demetrios
Gemistos Georgios, 1380
Bessarion Basilios, 1436
Oresme Nicole, 1356
Leï Langenstein Heinrich , 1375 1363
von Gmunden Johannes , 1406
von Peùerbach Georg, 1440
Regiomontanus Johannes, 1457
Novara da Ferrara Domenico, 1483
von Dobachütz Leonhard, 1469
Copernicus Nicolaus, 1499
Rheticus Georg Joachim , 1535
Pascale Janue, 1472
Boüe Guillaume, 1406
Hussain Jacques, 1521
Argyropoulos Johannes, 1444
Rauchlin Johannes, 1477 1481
Pfauer Casper, 1545
Dietrich Sebastian, 1544
Andre Tycho, 1562
Otto Valentin, 1570
Schato Andreas, 1562
Meùzinger Theodor, 1559 1553
Stupanus Johannes, 1569
von Brunn Johann, 1615
Melanchthon Philipp, 1514 1511
Kaplor Johannes, 1601
Ursler Melchior, 1503
Lüeür Melchior, 1503
Caselius Johannes, 1560 1566
Metzger Georg, 1650 1646
Camerarius, Sr. Elias, 1563
Camerarius Rudolf, 1684
Calixï Georg, 1607
Guarewei Johann, 1643 1644
Rhodius Ambrosius, 1598 1600
Nottnagel-Christoph, 1630 1632 1634
Pasch Johann, 1683
Wolther-Michaël e. J., 1687 1689 1693
Planer Johann, 1696 1709
Hausen Christian, 1713
Kästner Abraham, 1739
Lichtenberg Georg, 1765
Thibaut Bernhard, 1796
Mayer Johann, 1773
Dirksen Enno, 1820
Jacobi Carl, 1825
Richelot Friedrich, 1831
Hesse Otto, 1840
Schelbear Wilhelm, 1853 1848
Neumann Carl, 1856
Mïày Wïllam, 1875
Lefschetz Solomon, 1911
Tucker Albert, 1932
Parsons Torrence, 1966
Pisanski Tomaž, 1981
Batageïj Vladimir, 1986
Divjak Saša
Mrvar Andrej, 1999

GEDCOM is standard for storing genealogical data, which is used to interchange and combine data from different programs. The following lines are extracted from the GEDCOM file of European Royal families.

```
0 HEAD                                          0 @I115@ INDI
1 FILE ROYALS.GED                               1 NAME William Arthur Philip/Windsor/
...                                             1 TITL Prince
0 @I58@ INDI                                     1 SEX M
1 NAME Charles Philip Arthur/Windsor/           1 BIRT
1 TITL Prince                                   2 DATE 21 JUN 1982
1 SEX M                                         2 PLAC St.Mary's Hospital, Paddington
1 BIRT                                          1 CHR
2 DATE 14 NOV 1948                              2 DATE 4 AUG 1982
2 PLAC Buckingham Palace, London                2 PLAC Music Room, Buckingham Palace
1 CHR                                           1 FAMC @F16@
2 DATE 15 DEC 1948                              ...
2 PLAC Buckingham Palace, Music Room            0 @I116@ INDI
1 FAMS @F16@                                     1 NAME Henry Charles Albert/Windsor/
1 FAMC @F14@                                     1 TITL Prince
...                                             1 SEX M
...                                             1 BIRT
0 @I65@ INDI                                     2 DATE 15 SEP 1984
1 NAME Diana Frances /Spencer/                   2 PLAC St.Mary's Hosp., Paddington
1 TITL Lady                                      1 FAMC @F16@
1 SEX F                                         ...
1 BIRT                                          0 @F16@ FAM
2 DATE 1 JUL 1961                                1 HUSB @I58@
2 PLAC Park House, Sandringham                   1 WIFE @I65@
1 CHR                                           1 CHIL @I115@
2 PLAC Sandringham, Church                       1 CHIL @I116@
1 FAMS @F16@                                     1 DIV N
1 FAMC @F78@                                     1 MARR
...                                             2 DATE 29 JUL 1981
...                                             2 PLAC St.Paul's Cathedral, London
```

Genealogies

Representation of genealogies using networks

Genealogies can be represented as networks in different ways:

- as Ore-graph,
- as p-graph,
- as bipartite p-graph.

Ore-graph:

In Ore-graph every person is represented by a vertex, marriages are represented with edges and relation *is a parent of* as arcs pointing from each of the parents to their children.
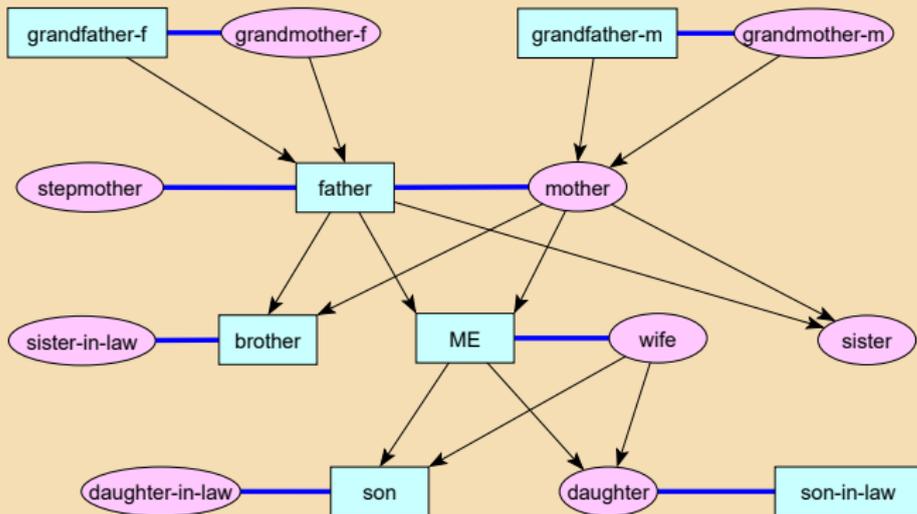
Read **family.ged** as *Ore-graph*.

p-graph:
In p-graph vertices represent individuals or couples. In the case that person is not married yet (s)he is represented by a vertex, otherwise person is represented with the partner in a common vertex. There are only arcs in p-graphs – they point from children to their parents.
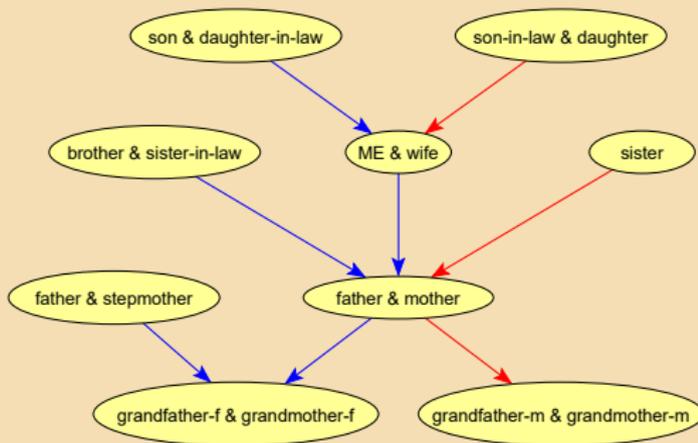
Read **family.ged** as *p-graph*.

Pajek

Genealogies

Sources
Representations
Sparse networks
Relinking
Comparison
Fragments
  Halfbrother-halfsister
Other analyses
  The highest number
  of marriages
  The highest number
  of children
  The shortest path
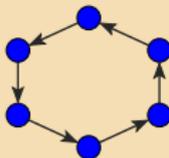 between two kings
Applying Pajek
 Ore graphs
  p-graphs
  bipartite p-graphs
Assignments

Bipartite p-graph:
Bipartite p-graph has two types of vertices – vertices representing couples (rectangles) and vertices representing individuals (circles for women and triangles for men). Arcs again point from children to their parents.
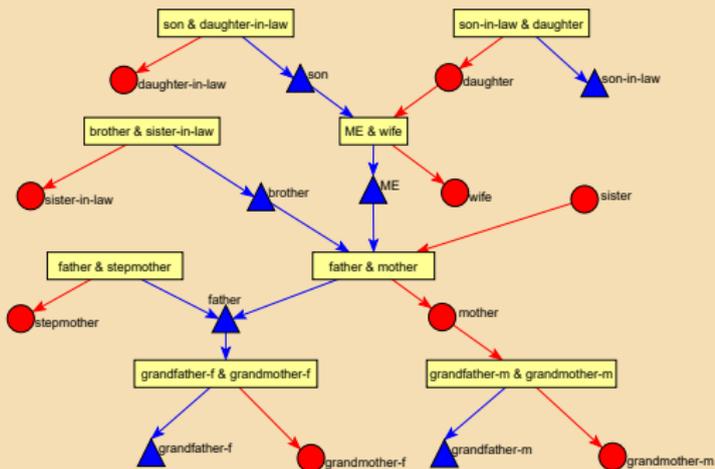
Read **family.ged** as *bipartite p-graph*.

Genealogies are sparse networks:

data	Ore-graph				p-graph			
	$|V|$	$|E|$	$|A|$	$\frac{|L|}{|V|}$	$|V_{ip}|$	$|V_{cp}|$	$|A_p|$	$\frac{|A_p|}{|V_p|}$
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

Advantages of p-graphs

- there are less vertices and lines in p-graphs;
- p-graphs are directed acyclic networks;

Cycle: Semicycle:
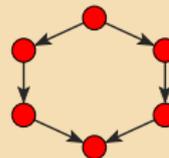
- in p-graphs every semi-cycle corresponds to a *relinking marriage*. There exist two types of relinking marriages:
  - blood marriage: e.g., marriage among brother and sister;
  - non-blood marriage: e.g., two brothers marry two sisters from another family;
- p-graphs are more suitable for analyses.

Advantages of bipartite p-graphs

Bipartite p-graphs have additional advantage: we can distinguish between *a married uncle and a remarriage of a father*. This property enables us, for example, to find marriages between stepbrothers and stepsisters.
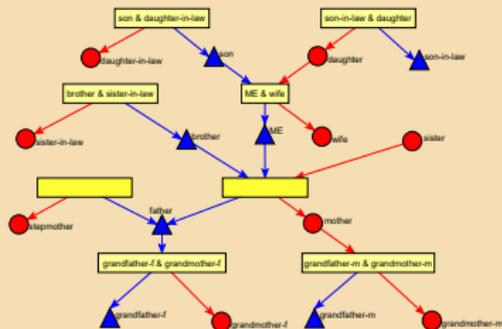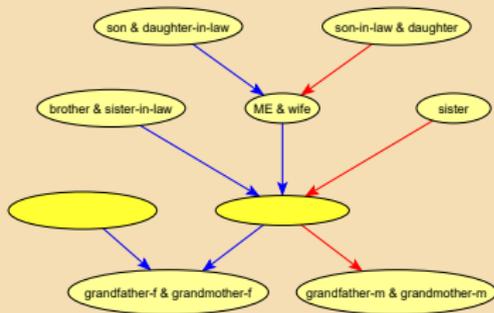
Relinking index is a measure of relinking by marriages among persons belonging to the same families. Special case of relinking is a blood-marriage.

Let n denotes number of vertices in p-graph, m number of arcs, and M number of maximal vertices (vertices having output degree 0, $M \geq 1$).

If we take a connected genealogy we get

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

For a trivial graph (having only one vertex) we define $RI = 0$.
RI has some interesting properties:

- $0 \leq RI \leq 1$
- If network is a forest/tree, then $RI = 0$ (no relinking).
- There exist genealogies having $RI = 1$ (the highest relinking).
- Relinking is usually computed for the largest biconnected component.

Network / **Acyclic Network** / **Info**

Patterns with Relinking Index = 1

Pajek

Genealogies

Sources

Representations

Sparse networks

Relinking

Comparison

Fragments
  Halfbrother-halfsister

Other analyses
  The highest number
  of marriages
  The highest number
  of children
  The shortest path
  between two kings

Applying Pajek
  Ore graphs
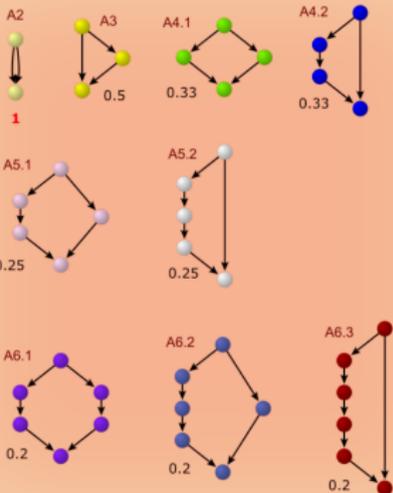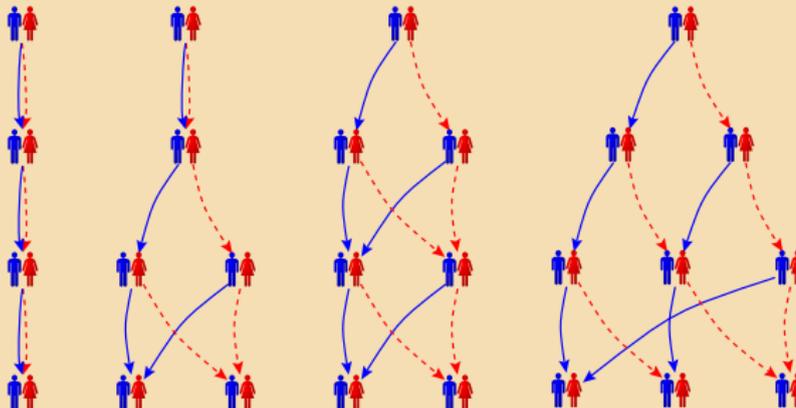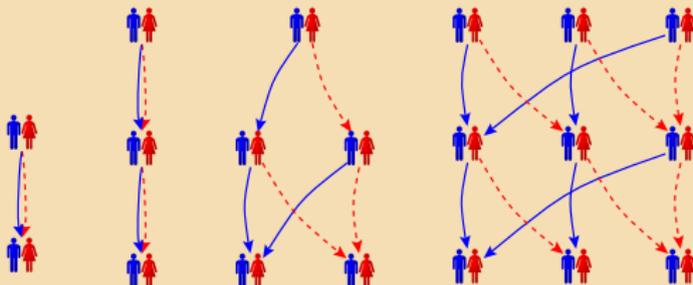 p-graphs
  bipartite p-graphs

Assignments
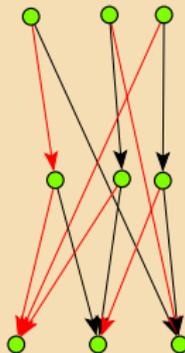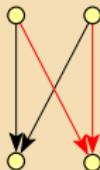
Relinking marriages (p-graphs with 2 up to 6 vertices)

Blood marriages

Relinking marriages
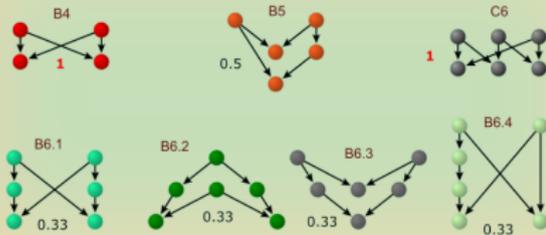
A2

A3   0.5

A4.1   0.33

A4.2   0.33

A5.1   0.25

A5.2   0.25

A6.1   0.2

A6.2   0.2

A6.3   0.2

B4   1

B5   0.5

C6   1

B6.1   0.33

B6.2   0.33

B6.3   0.33

B6.4   0.33

More blood marriages with RI=1

European nobility – MarijaTerezija.ged

Genealogy contains records about around 60,000 persons of noble origin (collected by Nenad Novaković). We can find there 333 blood marriages where two cousins married, 13 of them in the same component.
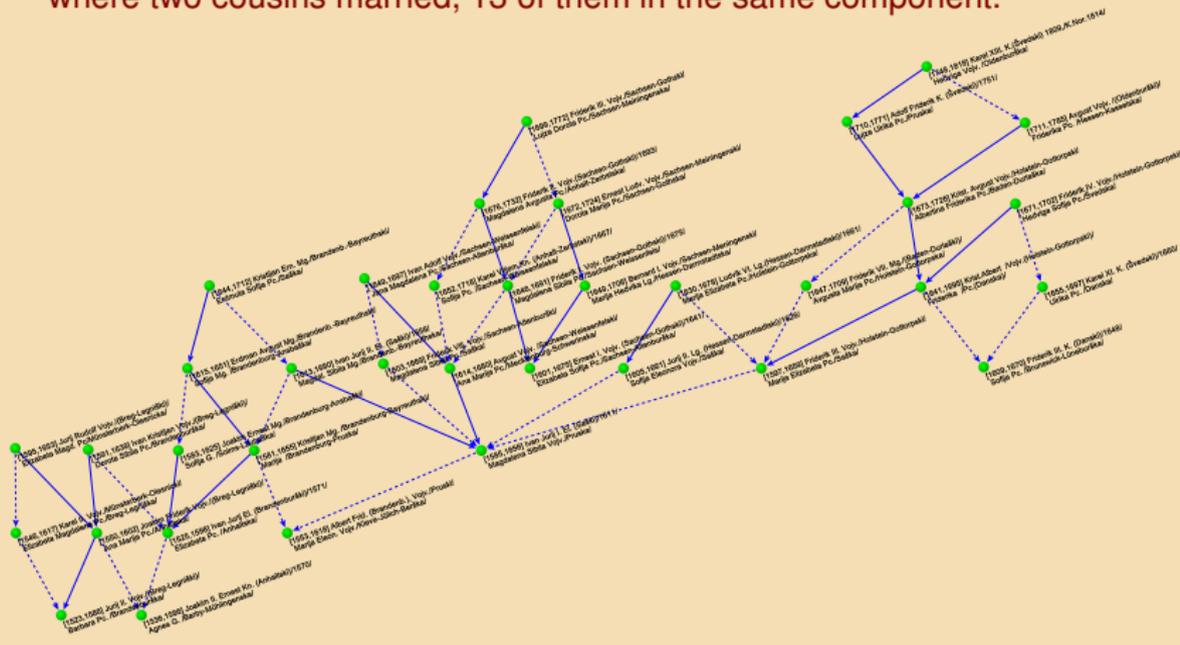
Genealogies

Noble familes are much more relinked than usual families.
Example: In 1854 **Franz Jozef (1830-1916)** married his sixteen year old cousin 'Sisi'.
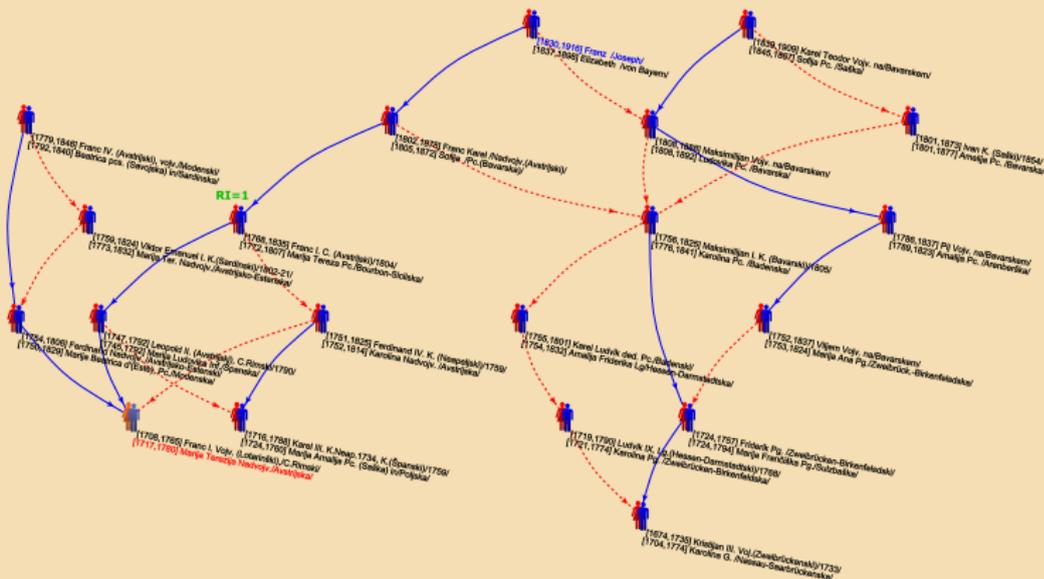
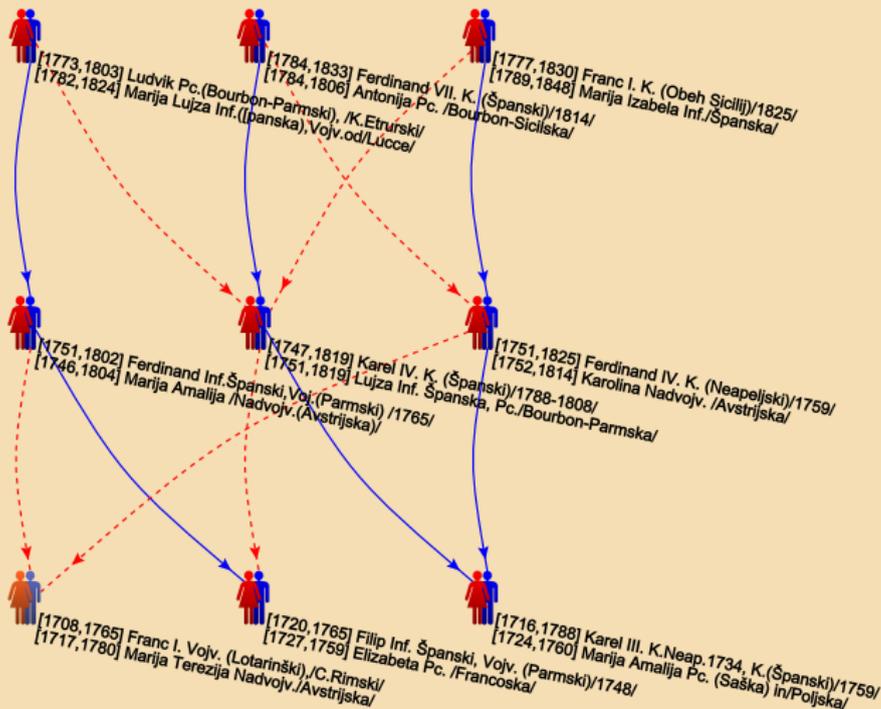**Question:** Can you find any other relinking marriages in this picture?

Brother-sister exchanges and marriages between cousins

[1773,1803] Ludvik Pc.(Bourbon-Parmski), /K.Etrurski/
[1782,1824] Marija Lujza Inf.((panska).Vojv.od/Lucce/

[1784,1833] Ferdinand VII. K.
[1784,1806] Antonija Pc. /Bourbon-Sicilska/

[1777,1830] Franc I. K. (Obeh Sicilij)/1825/
[1789,1848] Marija Izabela Inf./Spanska/

[1751,1802] Ferdinand Inf.Spanski,Vojv.(Parmski)
[1746,1804] Marija Amalija /Nadvojv.(Avstrijska)/

[1747,1819] Karel IV. K.
[1751,1819] Lujza Inf. Spanska,/1788-1808/

[1751,1825] Ferdinand IV. K. (Neapeljski)/1759/
[1752,1814] Karolina Nadvojv. /Avstrijska/

[1708,1765] Franc I. Vojv. (Lotarinški),
[1717,1780] Marija Terezija Nadvojv./Avstrijska/

[1720,1765] Filip Inf. Spanski, Vojv. (Parmski)/1748/
[1727,1759] Elizabeta Pc. /Francoska/

[1716,1788] Karel III. K.Neap.1734. K.(Španski)/1759/
[1724,1760] Marija Amalija Pc. (Saška) in/Poljska/

Comparing genealogies

For comparison, we took five genealogies:

- **Loka.ged** – genealogy of Škofja Loka district, Slovenia (P. Hawlina).
- **Silba.ged** – genealogy of the island Silba, Croatia (P. Hawlina).
  Special geographical position.
- **Ragusa.ged** – marriages among Ragusan (Dubrovnik) noble families between 12 and 16 century. Data collected by I. Mahnken (1960); entered to electronic form by P. Dremelj (1999).
  Very restricted marriage rules.
- **Tur.ged** – genealogy of Turkish nomads, Yörük.
  Data collected by Ulla C. Johansen and D.R. White (2001).
  A relinking marriage is a signal of commitment to stay within the nomad group.
- **Royal92.ged** – genealogy of European royal families – the Kings and Queens of Europe, including Great Britain.

Pajek

Frequency distribution of fragments

	fragment	Loka	Silba	Ragusa	Tur	Royal	$\sum$
	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
	No. Indi	47956	6427	5999	1269	3010	
	Largest bic.	4095	1340	1446	250	435	
	RI	0.55	0.78	0.74	0.75	0.37	

Pajek

Genealogies

Observations

- Generation jumps for more than one generation are very unlikely.

- There are many marriages B6.3 (spouses of two grandchildren are brother and sister) and B6.2 (two familes were relinked by a marriage between children and again (in the next generation) by a marriage between grandchildren)

- In Tur there are many marriages of types A4.1 and A6.1.

- For all genealogies number of relinking 'non-blood' marriages is much higher than number of blood marriages (this is especially true for Ragusa, exception is Royal92). There were economic reasons for non-blood relinking marriages: to keep the wealth and power within selected families.

type of marriage	Loka	Silba	Ragusa	Tur	Royal
blood-marriages	51	42	9	149	51
relinking-marriages	157	118	239	176	45

Number of individuals in genealogy Tur is much lower than in others, Silba and Ragusa are approximately of the same size, while Loka is much larger genealogy, what we must also take into account.

Pajek

Frequencies normalized with number of couples in p-graph $\times$ 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
	$\sum$	14.70	72.17	123.88	798.53	84.36

Bipartite p-graphs: Marriage between stepbrother and stepsister
Using p-graphs we cannot distinguish persons married several times. In this case we must use bipartite p-graphs. For example in bipartite p-graphs we can find marriages between stepbrothers and stepsisters.
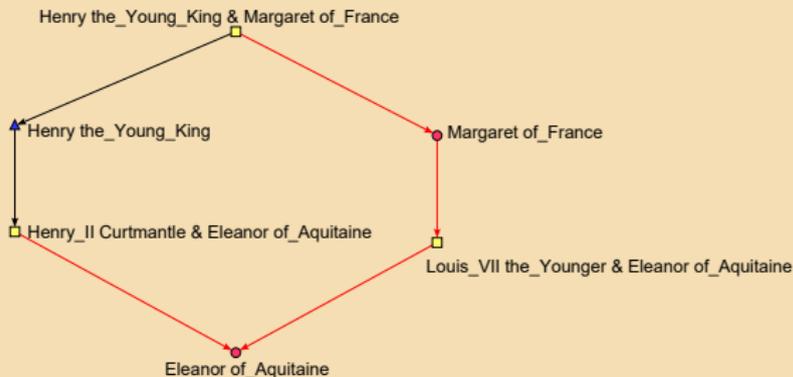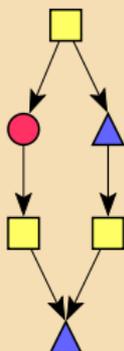
Bipartite p-graphs: Marriage among half-cousins
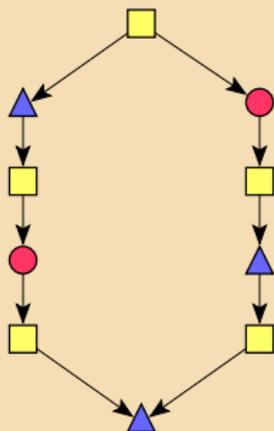
Benjamin Simoniti & Natalija Mavric
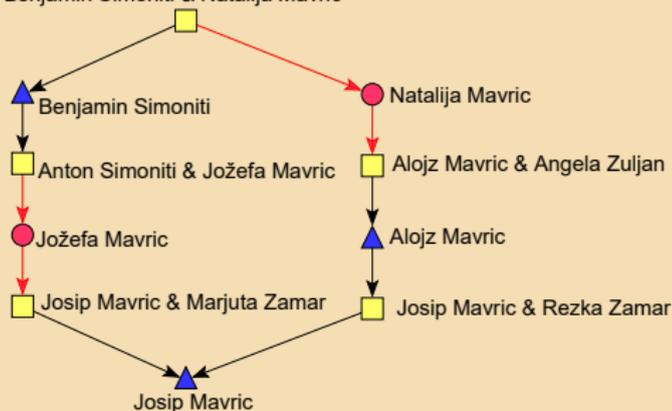
Benjamin Simoniti

Natalija Mavric

Anton Simoniti & Jožefa Mavric

Alojz Mavric & Angela Zuljan

Jožefa Mavric

Alojz Mavric

Josip Mavric & Marjuta Zamar

Josip Mavric & Rezka Zamar

Josip Mavric

Other analyses

People collecting data about their families are interested in several other 'standard' analyses:

- changes in relinking patterns over time;
- special situations: persons married several times, persons having the highest number of children;
- checking whether the two persons are relatives and searching for the shortest genealogical path between them;
- searching for all predecessors/successors of selected person and searching for person with the largest number of known predecessors or successors;
- the largest difference in age between husband and wife, the oldest/youngest person at the time of marriage, the oldest/youngest person at the time of child's birth;
- searching for the longest patrilineage and matrilineage;
- special situations $\rightarrow$ errors made in data entry (network consistency check).

The highest number of marriages

The highest number of children

English king Edward I (1237-1307) and his wife Eleanor (1241-1290) had 16 children who were born between 1255-1284 (in the picture a daughter without given name is missing). The youngest son (Edward) was the first among sons who survived a childhood. Eleanora had to try sixteen times to fulfill her most important duty as a queen: to give a birth to a men successor who later became a king. 10 out of 16 children died before age 10, only 3 of them lived longer than 40 years.

Searching for kinship relations

32nd American president F.D. Roosevelt (1882-1945).
43rd American president G.W. Bush (1946).
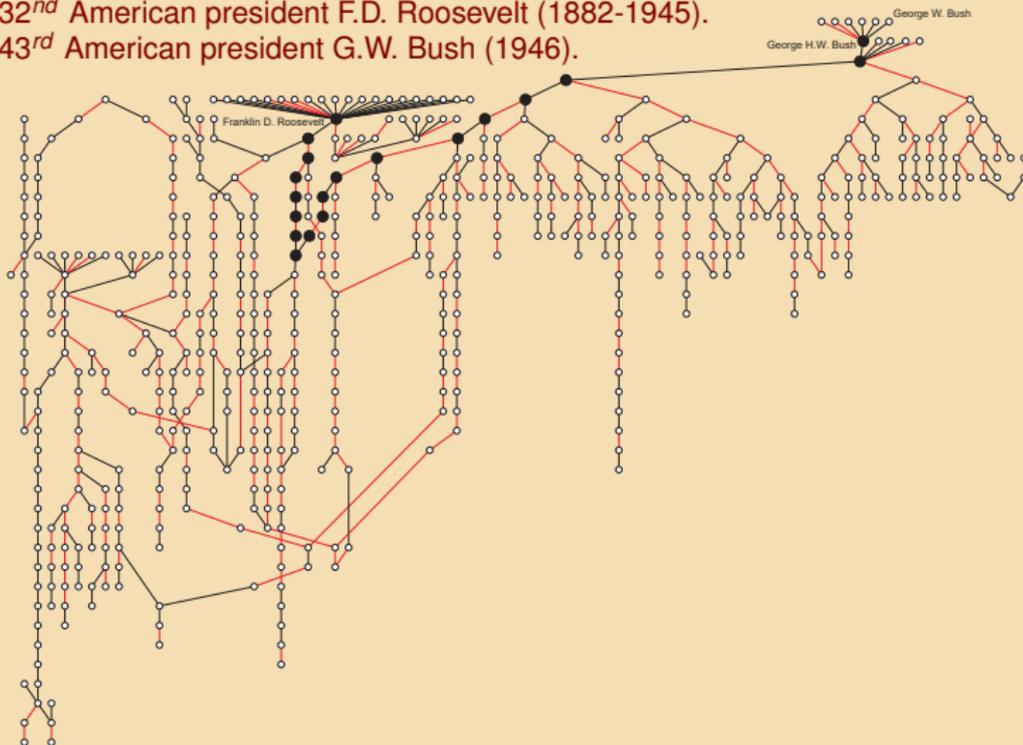
Pajek

The shortest genealogical paths between *Charles P. A. Windsor* (King of UK), *Juan Carlos* (ex King of Spain), and *Alexander* (Crown Prince of Yugoslavia) in Royal92.ged
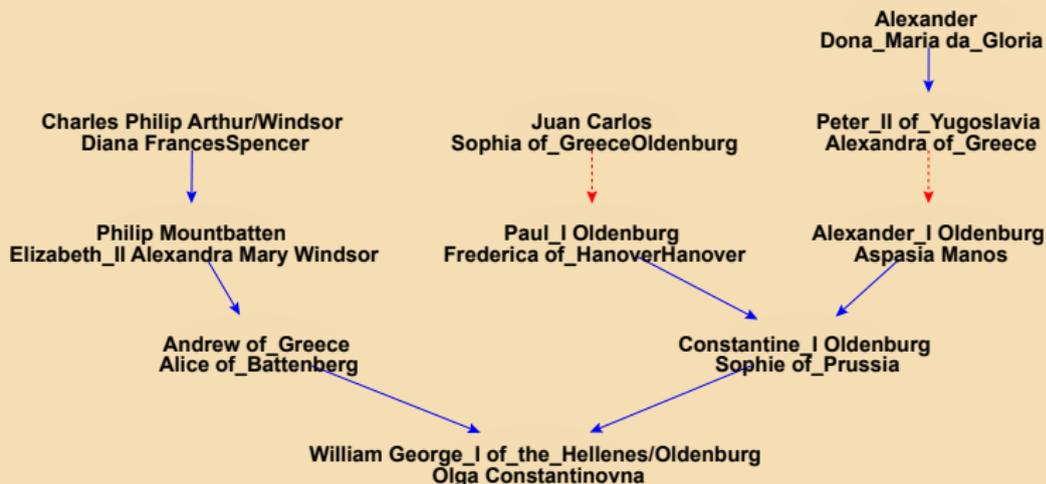
Alexander
Dona_Maria da_Gloria

Charles Philip Arthur/Windsor
Diana FrancesSpencer

Juan Carlos
Sophia of_GreeceOldenburg

Peter_II of_Yugoslavia
Alexandra of_Greece

Philip Mountbatten
Elizabeth_II Alexandra Mary Windsor

Paul_I Oldenburg
Frederica of_HanoverHanover

Alexander_I Oldenburg
Aspasia Manos

Andrew of_Greece
Alice of_Battenberg

Constantine_I Oldenburg
Sophie of_Prussia

William George_I of_the_Hellenes/Oldenburg
Olga Constantinovna

General instructions:

1. For drawing in layers we compute *genealogical depth partition* (works for Ore graphs, p-graphs and bipartite p-graphs):
   Network / **Acyclic Network** / **Create Partition** / **Depth Partition** / **Genealogical**

2. After we find some interesting person or couple (e.g. with high number of children or high number of marriages), we extract and draw his/her/its 1-neighbourhood to check the results

3. In case that genealogy is read as p-graph (recall: p-graphs are acyclic) we get *Relinking Index* with:
   Network / **Acyclic Network** / **Info**

4. Many interesting results can be found with Pajek with a single command or combining few commands. In case that several commands are needed it is better to use *a macro*.

Analyses using *Ore graph*   *Example: royal92.ged*

Before reading genealogy as Ore graph be sure that **Different relations for male and female links** is the only option checked in **Options / Read-Write** (when reading genealogy arcs from father to children are assigned relation number (and line value) 1; arcs from mother to children are assigned relation number (and line value) 2; and edges, which represent marriages, are assigned relation number (and value) 3.

1 Number of children: *Output degree* (number of outgoing arcs), but first we have to remove marriages (edges): **Network / Create New Network / Transform / Remove / all Edges**
Vertex 1261: Edward_I (Longshanks): 18 children (with 2 viwes).

2 If we are not interested in number of children only but in number of all descendants, we again need network without edges and compute *Output Domain*: **Network / Create Vector / Centrality / Proximity Prestige / Output**
2018: Sceaf has 1157 descendants.

3 Similarly, for number of ancestors, compute *Input Domain*: **Network / Create Vector / Centrality / Proximity Prestige / Input**
Prince Henry (116) and Prince William (115) have 598 ancestors.

4 Number of marriages: *Degree* (number of edges attached to vertex), but first we have to remove the parent relation (arcs): **Network** / **Create New Network** / **Transform** / **Remove** / **all Arcs**
828: Henry_VIII Tudor had 6 viwes.

5 Shortest genealogical path among any two individuals:

   1 Transform all arcs to edges: **Network** / **Create New Network** / **Arcs** → **Edges** / **All**

   2 Find the shortest path(s) as usually.

   3 Extract subnetwork induced by vertices on the shortest path (reported in the obtained partition) from original network (the one with arcs and edges).

   4 Compute genealogical layers.

   5 Draw in layers and optimize the layout.

   Find the shortest genealogical path between ex Spanish King Juan Carlos (435) and Prince William (115).

When some operations need several commands to be executed it is more convenient to use a macro (e.g. `ShortestPath.mcr`).

Genealogies

Sources
Representations
Sparse networks
Relinking
Comparison
Fragments
Halfbrother-halfsister
Other analyses
The highest number of marriages
The highest number of children
The shortest path between two kings
Applying Pajek
Ore graphs
p-graphs
bipartite p-graphs
Assignments

6 **Longest patrilineage and matrilineage in Ore graph**
Patrilineage: Lineal descent traced through the male line - in each generation at least one male descendant is born (important for keeping the surname and other privileges).
Matrilineage: Lineal descent traced through the female line - in each generation at least one female descendant is born.

As *Network* select genealogy, as *Partition* select Gender partition, as *Vector* select Year of Birth.
For the longest patrilineage run (**Macro** / **Play**) LongestPatrilineage.mcr, for the longest matrilineage run LongestMatrilineage.mcr.
In Draw window select **Options** / **Mark Vertices Using** / **Vector Values** and set number of decimals to 0 **(Options** / **Layout** / **Decimal Places 0**).
To find the longest path not taking gender into account run LongestChain.mcr.

Searching for relinking marriages - fragments - in p-graphs:

- Read genealogy as p-graph - before reading check only:
  Options / **ReadWrite GEDCOM – Pgraph**.

- Read *Pajek project file frag16.paj* with all 16 fragments:
  File / **Project File** / **Read** (or **F1**).

- Select first fragment (*Marriage among children*) as the first
  network and genealogy as the second network.

- Run **Networks** / **Fragment (First in Second)**.
  Before starting search (**Find**) make sure that only *"Extract
  Subnetwork"* and *"Same vertices determine one fragment at
  most"* are checked.

- As result we get all fragments found. Instead of running this
  command by hand 15 more times we can also run **Macro** /
  Repeat Last Command. Since the second network is always
  the same, we must fix it (**Fix (Second) Network**) and answer
  15 – run the command for the 15 remaining fragments.

...Genealogies...

...Applying Pajek to p-graphs...

Pajek

Genealogies

Sources
Representations
Sparse networks
Relinking
Comparison
Fragments
 Halfbrother-halfsister
Other analyses
 The highest number
 of marriages
 The highest number
 of children
 The shortest path
 between two kings
Applying Pajek
 Ore graphs
 p-graphs
 bipartite p-graphs
Assignments

As result we get:

1. Vector with 16 cells ("*Number of Fragments*"), explaining how many times each of 16 fragments occurred.
   royal92.ged: 0, 3, 22, 7, 0, 13, 0, 8, 16, 1, 0, 5, 0, 13, 16, 0

2. 16 new networks which contain fragments found ("*Subnetwork induced by...*").

3. Corresponding 16 new partitions, telling which vertices are involved in fragments found.

Visualizing results:

Select one of the resulting networks. Network can be drawn using macro `DrawFragments.mcr`. Before running select network with fragments found (*"Subnetwork induced by Sub fragments like N2 in N1"*), compatible partition (*"Sub fragments like N2 in N1"*) and vector that we would like to use in visualization (e. g. *"Year of Birth of Men"*).

We can optimize picture several times using **Layers** / **Optimize layers in x direction**. Do not forget to select **Options** / **Mark Vertices Using** / **Vector Values**.

We can use two vectors in Pajek visualizations. For example, we can show year of birth of women as well.

Select vector (*"Year of Birth of Women"*) and partition obtained as result of corresponding fragment searching (e. g. *"Sub fragments like N2 in N1"*).

Run **Operations** / **Vector + Partition** / **Extract Subvector** and leave the answer as suggested (*1-**). Then select the obtained vector as the second vector and draw the network again.

Sources
Representations
Sparse networks
Relinking
Comparison
Fragments
Halfbrother-halfsister
Other analyses
The highest number
of marriages
The highest number
of children
The shortest path
between two kings
Applying Pajek
Ore graphs
p-graphs
bipartite p-graphs
Assignments

Searching for relinking marriages in bipartite p-graphs:

Marriages among half brother and half sister (or marriages among half cousins) can be found only in *bipartite pgraphs*:

- In **Options** / **Read-Write** check the following two items: **GEDCOM - Pgraph** and **Bipartite Pgraph**.
- Read the fragment stored in file *HalfBrotherSister*.net* or (*HalfCausins*.net*) and set it as the first network.
- Select genealogy (which must be loaded as bipartite p-graph) as the second network.
- **Networks** / **Fragment (First in Second)** / **Find**.
- Draw obtained fragments as explained earlier.
- But looking at the results some fragments look strange...

Providing more details in fragment searching:

Note that only fragments with *triangle* or *circle* at the bottom really represent marriages between half brother and half sister (or marriages between half cousins). Other fragments are probably errors produced by data entry or represent some other situation (e.g. person belonging to two or more families as a child).

The problem can be solved by providing more details in fragment searching: If we select **Check values of lines** or **Check relation numbers** only fragments that match also in *values of lines* (*relation numbers*) are reported as a result.
But things get more complicated here: We have two fragments for searching for marriages among half brother and half sister:

- father married twice:
  `HalfBrotherSister-Father2TimesMarried.net`
- mother married twice:
  `HalfBrotherSister-Mother2TimesMarried.net`

and 8 fragments for marriages among halfcousins: `HalfCausins*.net`
In genealogy *royal92.ged* we find one marriage halfbrother-halfsister and one marriage among halfcousins.

royale.ged

Ore graph:

1 *2758: Roberto/di Borbone/: 24 children (with 2 viwes).*

2 *Two persons with 11000 descendants, including 6023: P,pin/Carolingien/.*

3 *Five persons with 7487 ancestors, including 7487: Gabrielle/de Habsbourg-Lorraine/.*

4 *Two persons with 7 viwes, one of them 16329: Charles/Stuart/.*

5 *Find the shortest genealogical path between two persons with the highest number of marriages: i.e. 16329: Charles/Stuart/, and 14812: Ivan/Rurikide/.*

p-graph:

6 *frequencies of 16 fragments in p-graph:*
   1, 28, 218, 114, 0, 181, 0, 159, 453, 28, 0, 76, 9, 340, 319, 29

Bipartite p-graph:

7 In genealogy *royale.ged* we cannot find any marriages between halfbrothers and halfsisters.
   We can find 18 marriages among halfcousins, all of them are correct.

Marija Terezija.ged

Ore graph:

1 *4034: Algirdas Vel. Kn. (Litovski)/1345/: 24 children (with 2 viwes).*

2 *Two persons with 21194 descendants, including 37051: Achaemenes.*

3 *Four persons with 10147 ancestors, including 13105: Jožef Emanuel Pc./Liechtensteinski/.*

4 *4314: Maksimilijan I. (Avstrijski),/C.Rimski/: 8 viwes.*

5 *Find shortest genealogical path between "Karolina G. /Nassau-Saarbrückenska/ (1704-1735)" and "Marija Terezija Nadvojv./Avstrijska/ (1717-1780)" (vertices 13360 and 8500).*

p-graph:

6 *frequencies of 16 fragments in p-graph:*
   18, 43, 333, 185, 2, 220, 0, 194, 538, 34, 0, 64, 4, 361, 367, 24

Bipartite p-graph:

7 In *MarijaTerezija.ged* we can find 12 marriages between halfbrothers and halfsisters, but if we look into details we find out that only four of them are correct (three of them where father is married twice, and one of them where mother is married twice).

We can find 30 marriages among halfcousins, 28 of them are the right ones, only two of them are wrong.

Pajek

1. Read and draw genealogy stored in gedcom file `family.ged` as *Ore graph*, *p-graph* and *bipartite p-graph*. For depth partition use: **Network** / **Acyclic Network** / **Create Partition** / **Depth Partition** / **Genealogical**

2. `pres2020.ged` is a collection of the presidents of the US, with their ancestors and descendants updated to 2020. Find some interesting genealogical paths between american presidents, e.g. US President No. 26 – Theodore /Roosevelt/ (1463) and US President No. 43 – George Walker /Bush/ (1191).

3. The GEDCOM file `Isle_of_Man.ged` contains the combined genealogies of approximately twenty families from the British Isle of Man. Describe the overall structure of this network and the sections with structural relinking. Which types of relinking do occur?

4. The GEDCOM file `Silba.ged` contains genealogy of families from the Croatian island Silba. Search for different relinking marriages.

5. The GEDCOM file `Ragusa.ged` contains genealogy of noble families between 12th and 16th century from Dubrovnik (Ragusa is old name for Dubrovnik, Croatia). Which types of relinking do occur?