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

Analysis and Visualization of Temporal Networks using Pajek

Andrej Mrvar
Vladimir Batagelj

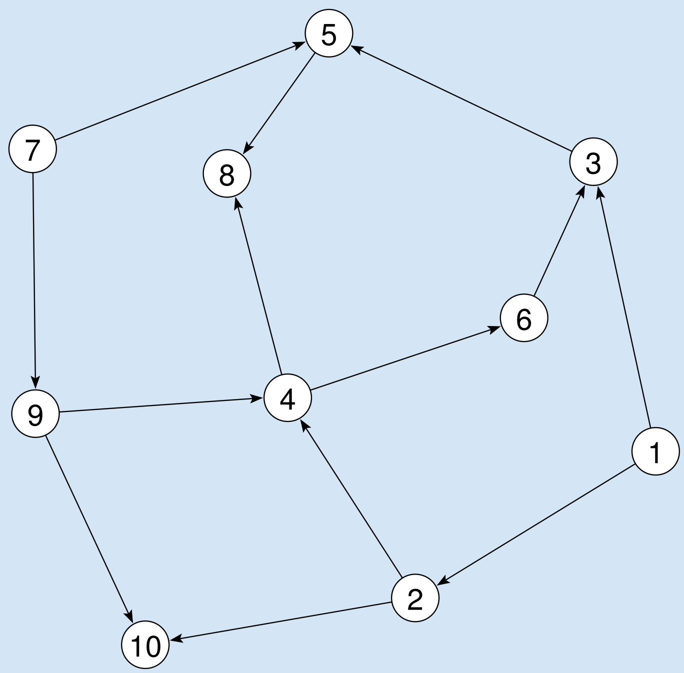
University of Ljubljana
Slovenia

SFI/SAS Joint Workshop
The Dynamics of Groups and Institutions
Brdo pri Kranju, June 7 – 11, 2004

Graph

Graph $G = (V, L)$:

- V – set of *vertices*
- L – set of *lines*: directed lines or *arcs* and undirected lines or *edges*.



Network

$N = (V, L, F_V, F_L)$.

- graph $G = (V, L)$
- function $f : V \to X$ from F_V, assigns a meaning to vertices:
  vertex label;
  class number to which the vertex belongs (partition);
  number associated with vertex (vector);
  type of vertex: source, intermediate, sink; ...
- function $g : L \to Y$ from F_L assigns a meaning to lines:
  name of the street;
  type of relation: friendship, business relation, sport activity;
  distance;
  capacity of link; ...

Temporal network

Temporal network

$$N_T = (V, L, F_V, F_L, T)$$

is obtained if the *time* T is attached to an ordinary network. T is a set of linearly ordered *time points* $t \in T$.

In temporal network vertices $v \in V$ and lines $l \in L$ are not necessarily present or active in all time points. Activity of vertices and lines can be described by a *functional description* – as logical functions:
$v(t) = true \Leftrightarrow$ vertex $v \in V$ is active in time point t;
$l(t) = true \Leftrightarrow$ line $l \in L$ is active in time point t.

The following *consistency condition* have to be satisfied:

$$\forall t \in T \forall l \in L : (l(u,v)(t) \Rightarrow u(t) \wedge v(t))$$

If a line $l(u,v)$ is active in time point t then also its endpoints u and v should be active in time t.

. . . Temporal network

We will denote the network consisting of lines and vertices active in time $t \in T$ by $N(t)$ and call it the *time slice* in time point t.

Properties of vertices and lines from F_V and F_L can also be time dependent.

Examples of temporal networks

- network of friendship in the schoolclass, institution or any other human group over several years (Newcomb fraternity);
- changes in signed graphs over time (Sampson monastery data (SVG));
- network of phone calls inside selected set of phone numbers (used by the police in the investigation of organized crime);
- citation or collaboration networks from a selected scientific area;
- network of transitions of a ball in a football game;
- changes in HIV networks;
- relations among actors in different episodes of movies;
- births, marriages and deaths in genealogies; ...

Activity intervals

One of realizations of functional description are *activity intervals* – for each vertex and line list of intervals when vertex/line is present is given. Intervals are separated by commas and have one of the following forms:

t_i – in time point t_i,
t_i-t_j – from time point t_i to time point t_j,
t_i-$*$ – from time point t_i on.

```
*Vertices 3
1 "a" [5-10,12-14]
2 "b" [1-3,7]
3 "e" [4-*]
*Edges
1 2 1 [7]
1 3 1 [6-8]
```

Time events

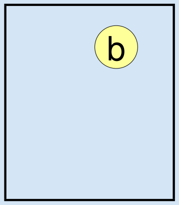
Development of network in time can be also described by the sequence of *time events*:

command	description
TI t	initial event: following events happen at start of time segment t
TE t	end event: following events happen at end of time segment t
AV vns	add vertex v with label n and properties s
HV v	hide vertex v
SV v	show vertex v
DV v	delete vertex v
AA uvs	add arc (u, v) with properties s
HA uv	hide arc (u, v)
SA uv	show arc (u, v)
DA uv	delete arc (u, v)
AE uvs	add edge $(u: v)$ with properties s
HE uv	hide edge $(u: v)$
SE uv	show edge $(u: v)$

DE uv	delete edge $(u: v)$
CV vs	change properties of vertex v to s
CA uvs	change property of arc (u, v) to s
CE uvs	change property of edge $(u: v)$ to s
CT uv	change type (arc/edge) of line (u, v)
CD uv	change direction of arc (u, v)
PE uvs	replace pair of bidirected arcs by a single edge
AP uvs	add pair of arcs with properties s
DP uv	delete pair of arcs
EP uvs	replace edge by pair of opposite arcs

Example of network described by time events 1

TI 1
AV 2 ”b”

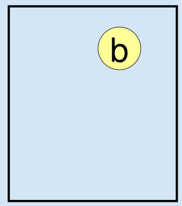


◀ ▶ ⏮ ● ⏭ ✱ ☛ X

Example of network described by time events 4

TI 1

AV 2 ”b”

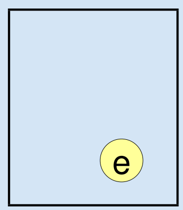


TE 3

HV 2

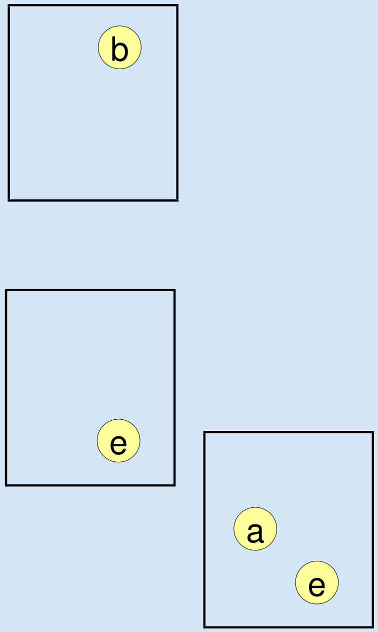
TI 4

AV 3 ”e”



Example of network described by time events 5

TI 1
AV 2 ”b”
TE 3
HV 2
TI 4
AV 3 ”e”
TI 5
AV 1 ”a”



Example of network described by time events 6

TI 1

AV 2 ”b”

TE 3

HV 2

TI 4

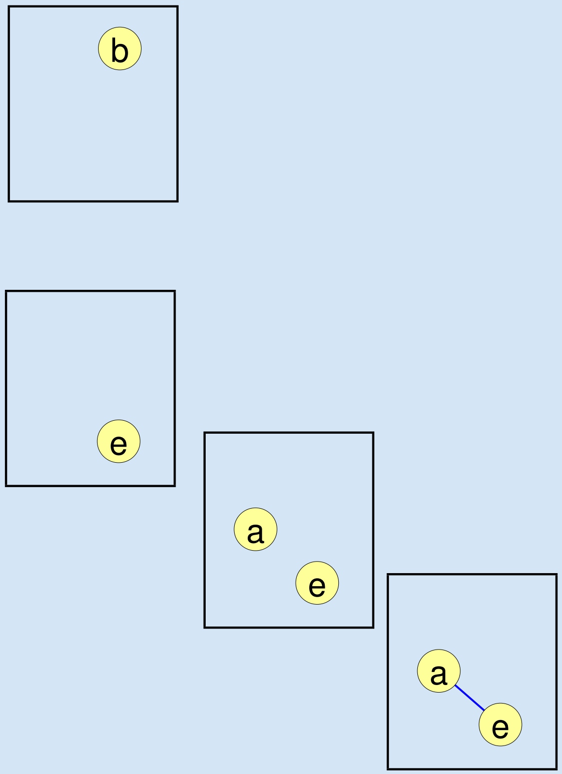
AV 3 ”e”

TI 5

AV 1 ”a”

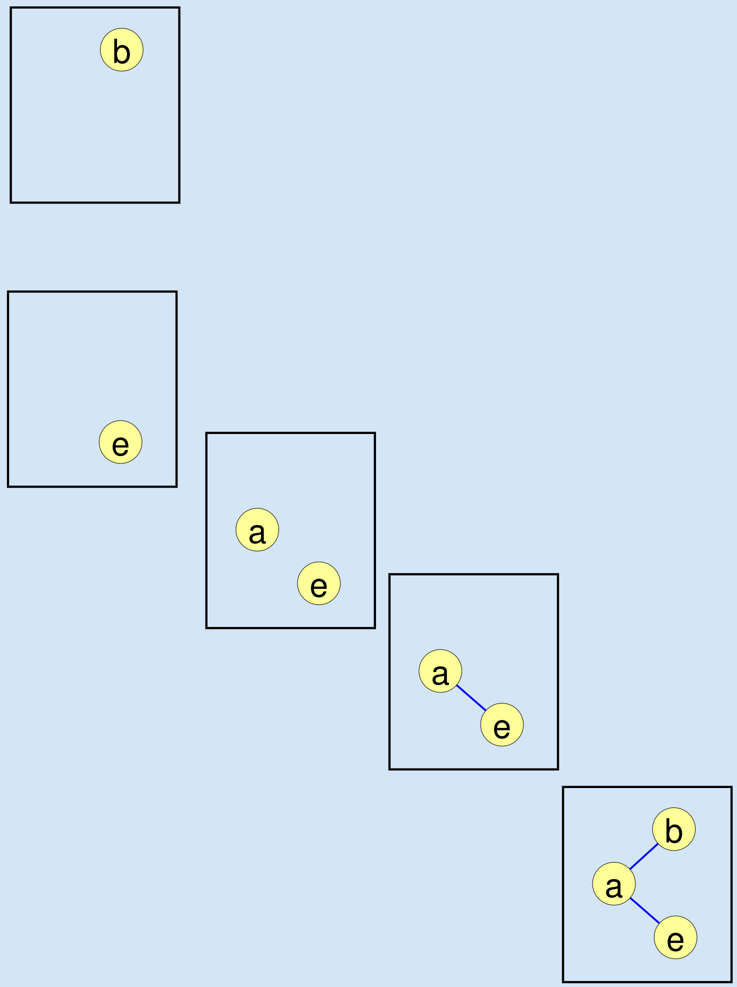
TI 6

AE 1 3 1



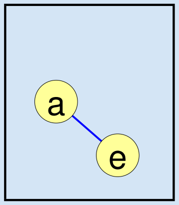
Example of network described by time events 7

TI 1
AV 2 ”b”
TE 3
HV 2
TI 4
AV 3 ”e”
TI 5
AV 1 ”a”
TI 6
AE 1 3 1
TI 7
SV 2
AE 1 2 1



Example of network described by time events 8

TE 7
DE 1 2
DV 2



Example of network described by time events 9

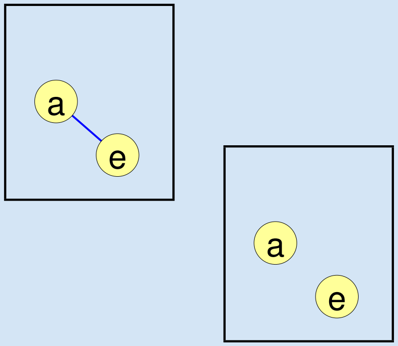
TE 7

DE 1 2

DV 2

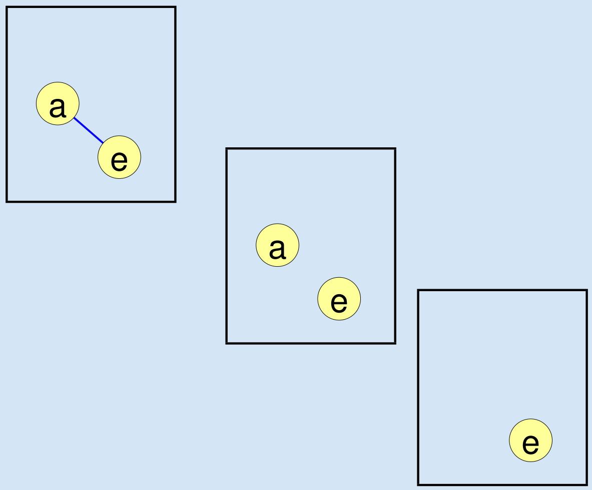
TE 8

DE 1 3



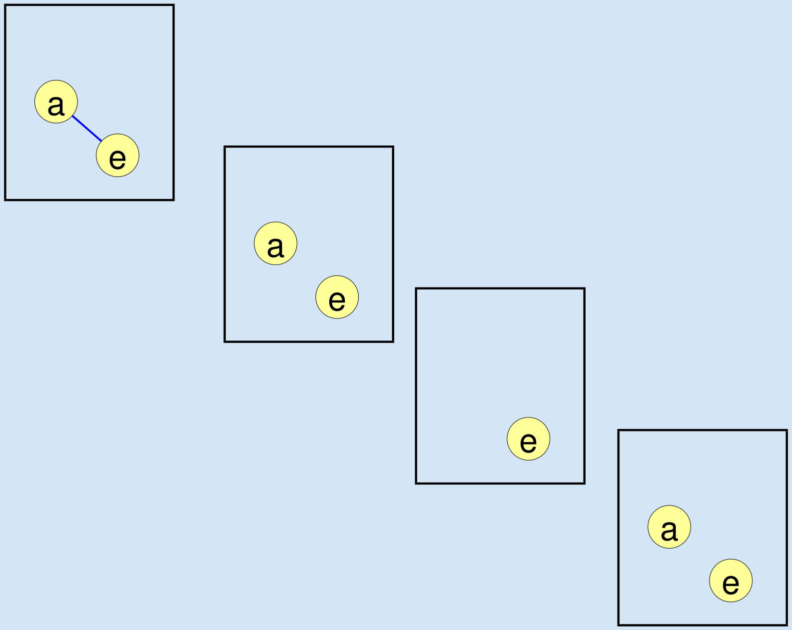
Example of network described by time events 11

TE 7
DE 1 2
DV 2
TE 8
DE 1 3
TE 10
HV 1



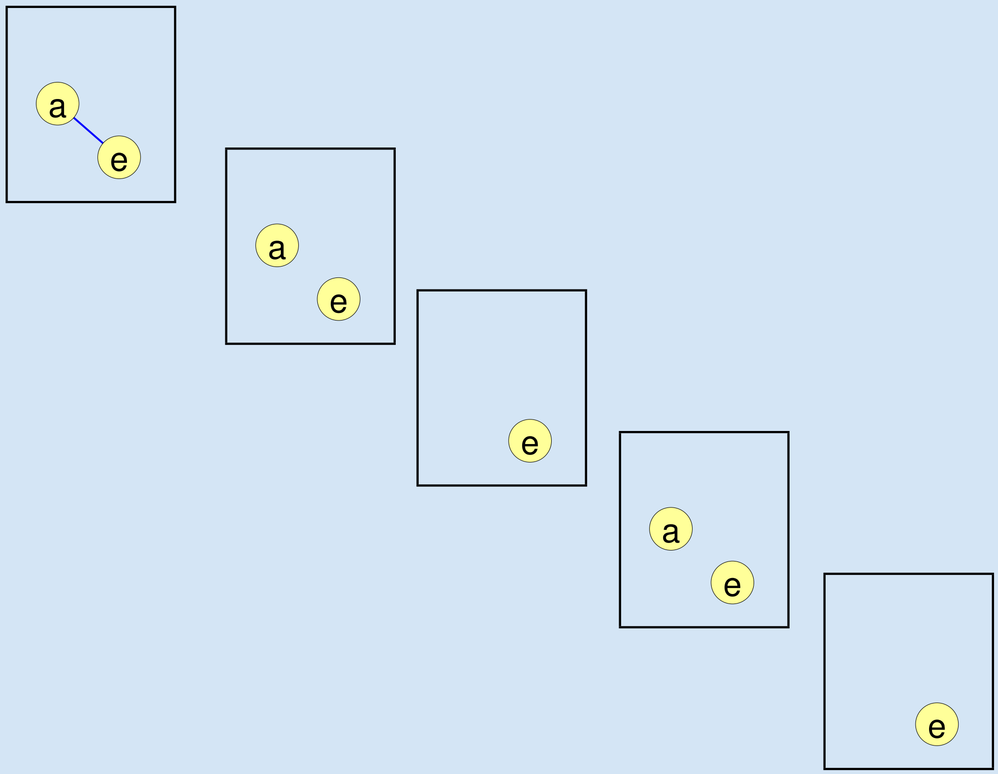
Example of network described by time events 12

TE 7
DE 1 2
DV 2
TE 8
DE 1 3
TE 10
HV 1
TI 12
SV 1



Example of network described by time events 10

TE 7
DE 1 2
DV 2
TE 8
DE 1 3
TE 10
HV 1
TI 12
SV 1
TE 14
DV 1



Example

The data about actors and relations among them in the long-running German soap opera called ‘Lindenstrasse’ were collected for Graph Drawing 1999 competition.

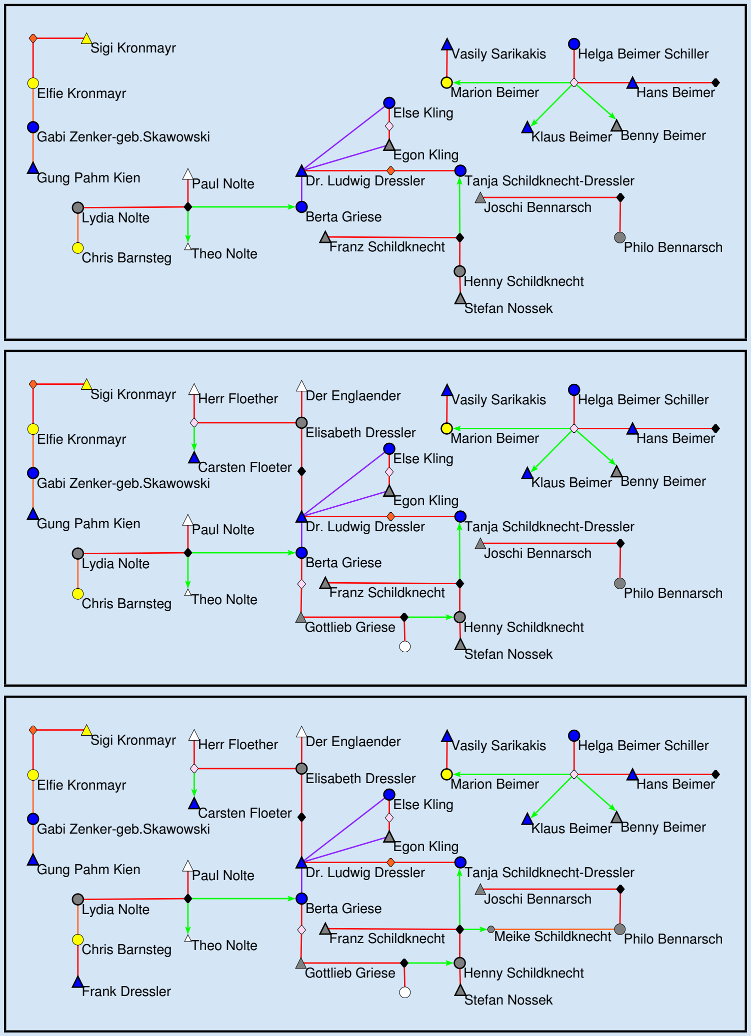
For each actor her/his name, gender, birthdate, and several other records are available. Additionally for each actor episode numbers in which the actor played actively are given.

For each line in the network its meaning is given: family relation, business relation, unfriendly relationships, . . .

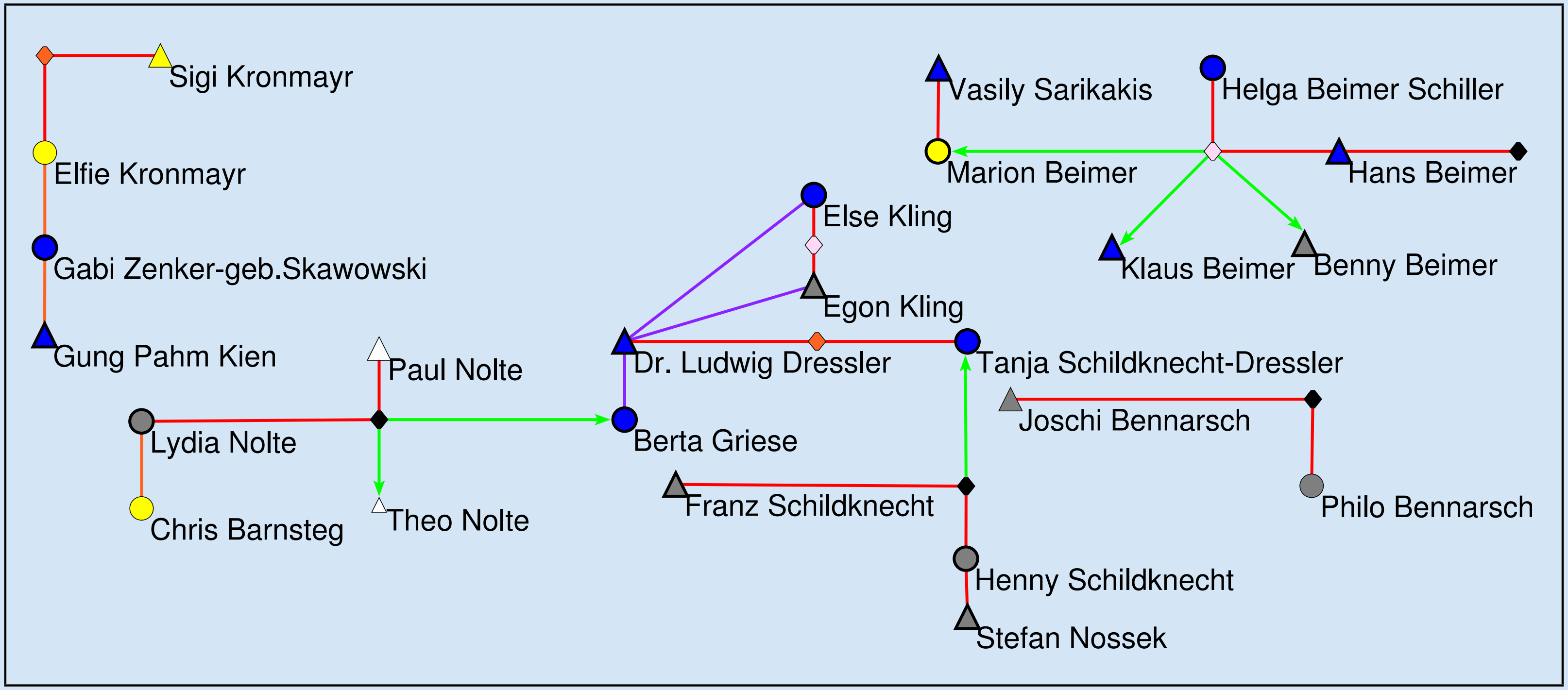
Properties of vertices are represented by different shapes, colors, sizes of vertices: e.g. triangles correspond to men, circles to women; and properties of lines are represented by colors: green line stands for family relation, blue for business relations, ...

In the printed version dynamics can be represented by the sequence of layouts only.

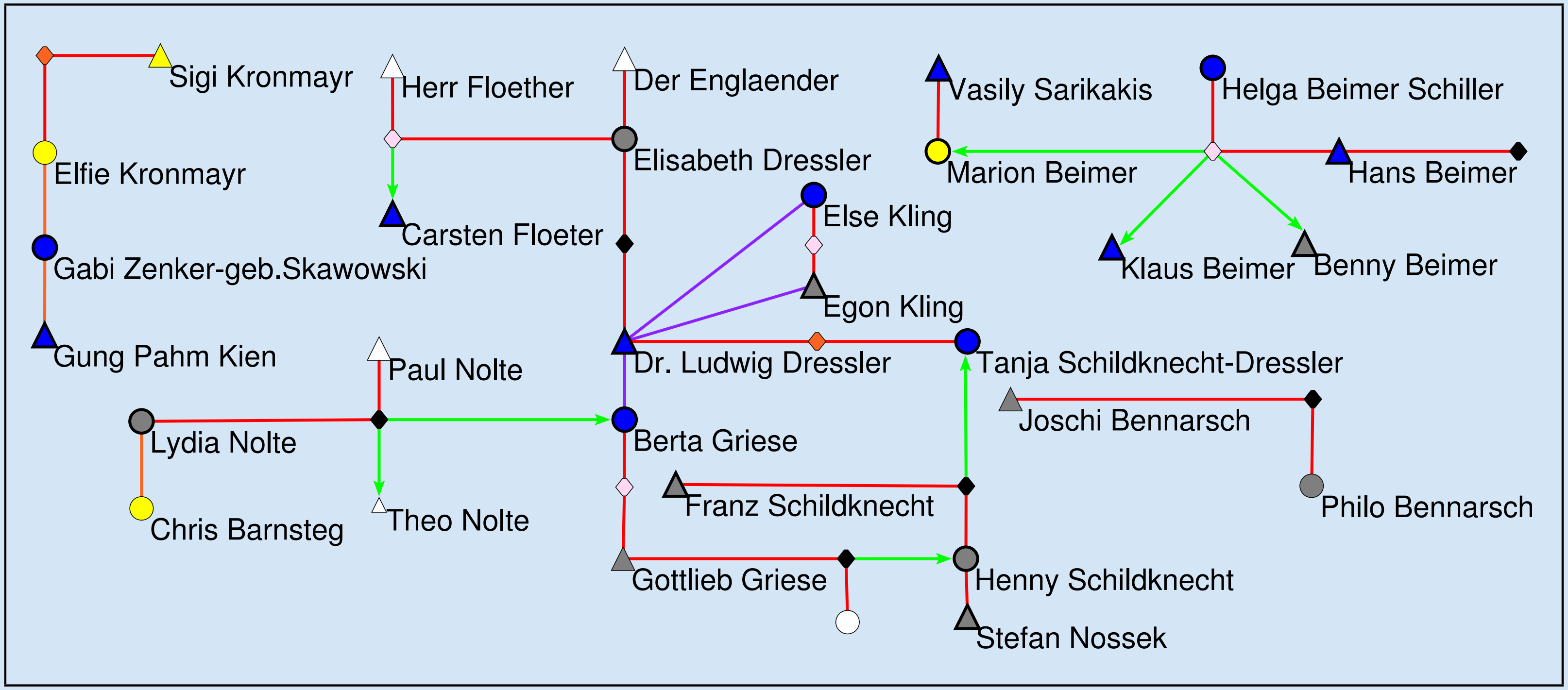
Relations among actors in 5^{th}, 6^{th} and 7^{th} episode



Episode 5

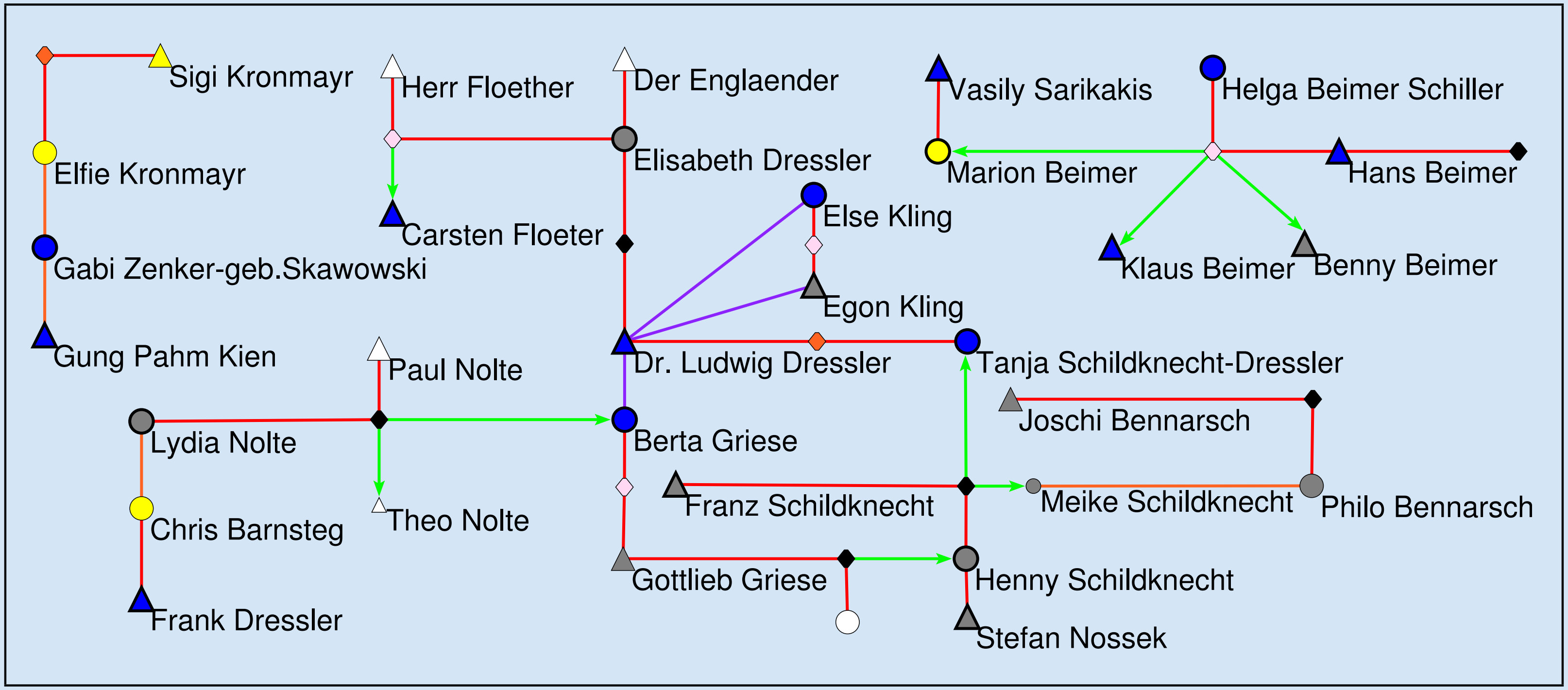

Sigi Kronmayr
Elfie Kronmayr
Gabi Zenker-geb.Skawowski
Gung Pahm Kien
Paul Nolte
Lydia Nolte
Chris Barnsteg
Theo Nolte
Vasily Sarikakis
Helga Beimer Schiller
Marion Beimer
Hans Beimer
Else Kling
Klaus Beimer
Benny Beimer
Egon Kling
Dr. Ludwig Dressler
Tanja Schildknecht-Dressler
Berta Griese
Joschi Bennarsch
Franz Schildknecht
Philo Bennarsch
Henny Schildknecht
Stefan Nossek

Episode 6



Sigi Kronmayr
Herr Floether
Der Englaender
Vasily Sarikakis
Helga Beimer Schiller
Elfie Kronmayr
Elisabeth Dressler
Marion Beimer
Hans Beimer
Else Kling
Carsten Floeter
Gabi Zenker-geb.Skawowski
Klaus Beimer
Benny Beimer
Egon Kling
Gung Pahm Kien
Paul Nolte
Dr. Ludwig Dressler
Tanja Schildknecht-Dressler
Joschi Bennarsch
Lydia Nolte
Berta Griese
Franz Schildknecht
Philo Bennarsch
Chris Barnsteg
Theo Nolte
Gottlieb Griese
Henny Schildknecht
Stefan Nossek

Episode 7



Sigi Kronmayr
Herr Floether
Der Englaender
Vasily Sarikakis
Helga Beimer Schiller
Elfie Kronmayr
Elisabeth Dressler
Marion Beimer
Hans Beimer
Carsten Floeter
Else Kling
Gabi Zenker-geb.Skawowski
Klaus Beimer
Benny Beimer
Egon Kling
Gung Pahm Kien
Paul Nolte
Dr. Ludwig Dressler
Tanja Schildknecht-Dressler
Joschi Bennarsch
Lydia Nolte
Berta Griese
Franz Schildknecht
Meike Schildknecht
Philo Bennarsch
Chris Barnsteg
Theo Nolte
Gottlieb Griese
Henny Schildknecht
Frank Dressler
Stefan Nossek

Temporal Networks in Pajek

Some options for analysis and visualizaton of sequences of networks are available in Pajek. Macro submenu enables to run the last command executed by Pajek several times applied to different successive objects.

Example: after loading several networks in Pajek, execute degree partition on the first network and all other networks.

If the result of the command is also a constant, all constants are stored in a new vector. Some of these constants can be used for comparison of temporal networks in different time points: number of vertices, arcs, edges, network densities, average degree, centralization indices, diameter, number of interesting fragments, main core number, number of components.

For graphical representation of changing properties of networks over time statistical packages **R** or **SPSS** can be called. We can also compute correlations among properties: e.g., it is known that density is lower if the number of vertices is higher but average degree is a expected to be constant.

Sequences of networks can be visualized by applying spring embedders.

Random Networks

Several models:

Model by Erdős and Rényi (1960): each pair of vertices is connected with equal probability p.

But in several networks (e.g., connections among WWW pages, citation networks, Internet topology, actors networks) a new vertex would preferably connect to such vertex already in the system that has higher degree. One of models describing such network formation is *extended model* by:
R. Albert in A.-L. Barabási:
Topology of evolving networks: local events and universality.

Networks grow and evolve by local events, such as

- the addition of new vertices,
- the addition of new lines,
- rewiring of lines from one vertex to another.

Extended Model Network Generation Algorithm

We start with m_0 isolated vertices, and at each time point we perform one of the following three operations:

1. *With probability p we add m ($m \leq m_0$) new lines.* For this we randomly select a vertex as the starting point of the new line. The other end of the line, however, is selected with probability
$$P(i) = \frac{deg_i + 1}{\sum_j (deg_j + 1)}, \qquad (1)$$
incorporating the fact that new lines preferentially point to popular vertices, with a high number of connections. This process is repeated m times.

2. *With probability q we rewire m lines.* For this we randomly select a vertex i and a line l_{ij} connected to it. Next we remove this line and replace it with a new line $l_{ij'}$ that connects i with vertex j' chosen with probability $P(j')$ given by (1). This process is repeated m times.

3. *With probability $1 - p - q$ we add a new vertex.* The new vertex has m new lines that are connected to vertices i already present in the network with probability $P(i)$.