On Blockmodeling

Anuška Ferligoj and **Vladimir Batagelj**

University of Ljubljana

Patrick Doreian

University of Pittsburgh

Fifth International Conference on Logic and Methodology

October 3-6, 2000 Cologne, Germany

Basic Notions

Network

Let $E = \{X_1, X_2, \ldots, X_n\}$ be a finite set of *units or units*. The units are related by binary *relations*

$$R_t \subseteq E \times E, \quad t = 1, \ldots, r$$

which determine a *network*

$$\mathcal{N} = (E, R_1, R_2, \ldots, R_r)$$

In the following we restrict our discussion to a single relation R described by a corresponding binary matrix $\mathbf{R} = [r_{ij}]_{n \times n}$ where

$$r_{ij} = \begin{cases} 1 & X_i R X_j \\ 0 & \text{otherwise} \end{cases}$$

In some applications r_{ij} can be a nonnegative real number expressing the strength of the relation R between units X_i and X_j.

Example: Student Government

The Student Government network consists of communication interactions among twelve members and advisors of the Student Government at the University in Ljubljana (Tina Hlebec, 1993). The results of the measurement are not real interactions among actors but cognition about communication interactions. Data were collected with face to face interviews. Interviews were conducted in May 1992.

Communication flow among actors was identified by the following question:

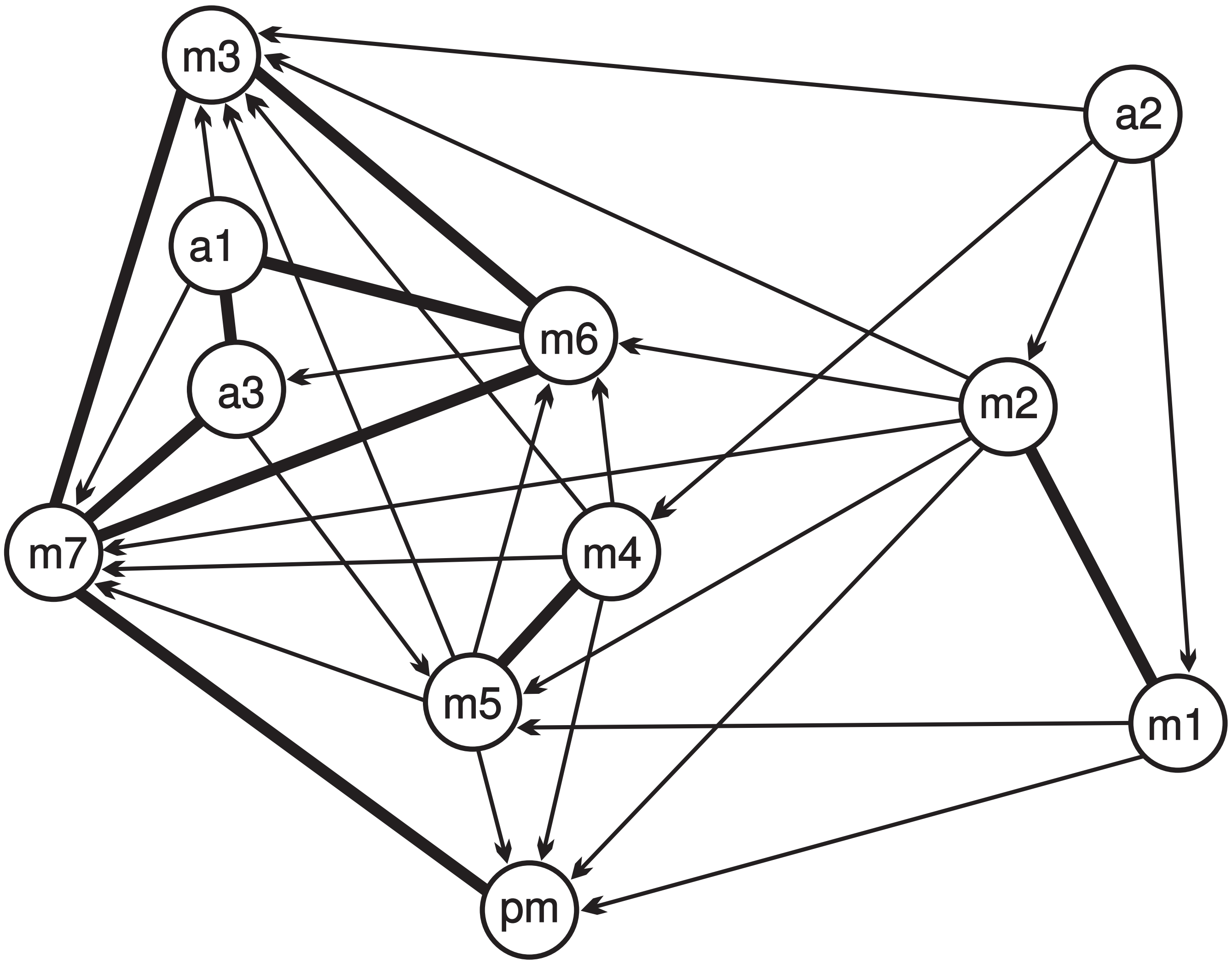
> Of the members and advisors of the Student Government, whom do you (most often) talk with?

The content of the communication flow was limited to the matters of the Student Government. The time frame was also defined: the question was referred to the six months period. One respondent refused to cooperate in the experiment. As he was not considered in the analysis, the network consists of eleven actors.

Student Government Matrix

		m 1	p 2	m 3	m 4	m 5	m 6	m 7	m 8	a 9	a 10	a 11
minister 1	1	0	1	1	0	0	1	0	0	0	0	0
p.minister	2	0	0	0	0	0	0	0	1	0	0	0
minister 2	3	1	1	0	1	0	1	1	1	0	0	0
minister 3	4	0	0	0	0	0	0	1	1	0	0	0
minister 4	5	0	1	0	1	0	1	1	1	0	0	0
minister 5	6	0	1	0	1	1	0	1	1	0	0	0
minister 6	7	0	0	0	1	0	0	0	1	1	0	1
minister 7	8	0	1	0	1	0	0	1	0	0	0	1
adviser 1	9	0	0	0	1	0	0	1	1	0	0	1
adviser 2	10	1	0	1	1	1	0	0	0	0	0	0
adviser 3	11	0	0	0	0	0	1	0	1	1	0	0

Student Government Network



Cluster, clustering

One of the main procedural goals of blockmodeling is to identify, in a given network, *clusters* (classes) of units that share structural characteristics defined in terms of R. The units within a cluster have the same or similar connection patterns to other units. They form a *clustering*

$$\mathcal{C} = \{C_1, C_2, \ldots, C_k\}$$

which is a partition of the set E:

$$\bigcup_i C_i = E$$

$$i \neq j \Rightarrow C_i \cap C_j = \emptyset$$

Each partition determines an equivalence relation (and vice versa). Let us denote by $\sim$ the relation determined by partition $\mathcal{C}$.

Block

A clustering $\mathcal{C}$ partitions also the relation R into *blocks*

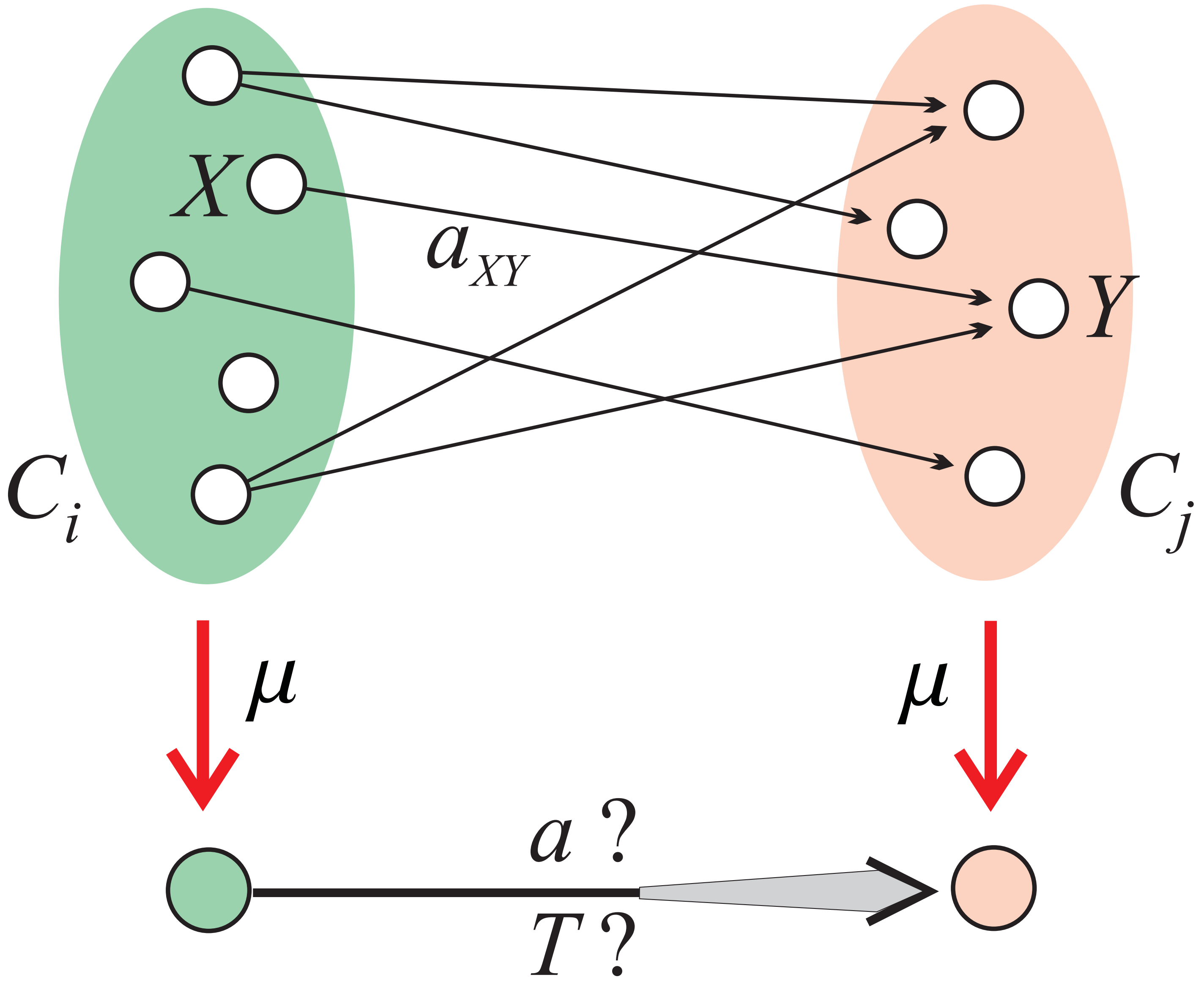
$$R(C_i, C_j) = R \cap C_i \times C_j$$

Each such block consists of units belonging to clusters C_i and C_j and all arcs leading from cluster C_i to cluster C_j. If $i = j$, a block $R(C_i, C_i)$ is called a *diagonal* block.

Blockmodel and Blockmodeling

The goal of blockmodeling is to reduce a large, potentially incoherent network to a smaller comprehensible structure that can be interpreted more readily. Blockmodeling, as an empirical procedure, is based on the idea that units in a network can be grouped according to the extent to which they are equivalent, according to some *meaningful* definition of equivalence.

Blockmodeling Scheme



Generalized Blockmodeling

A *blockmodel* consists of structures obtained by identifying all units from the same cluster of the clustering $\mathcal{C}$. For an exact definition of a blockmodel we have to be precise also about which blocks produce an arc in the *reduced graph* and which do not, and of what *type*. Some types of connections are presented in the figure on the next slide.

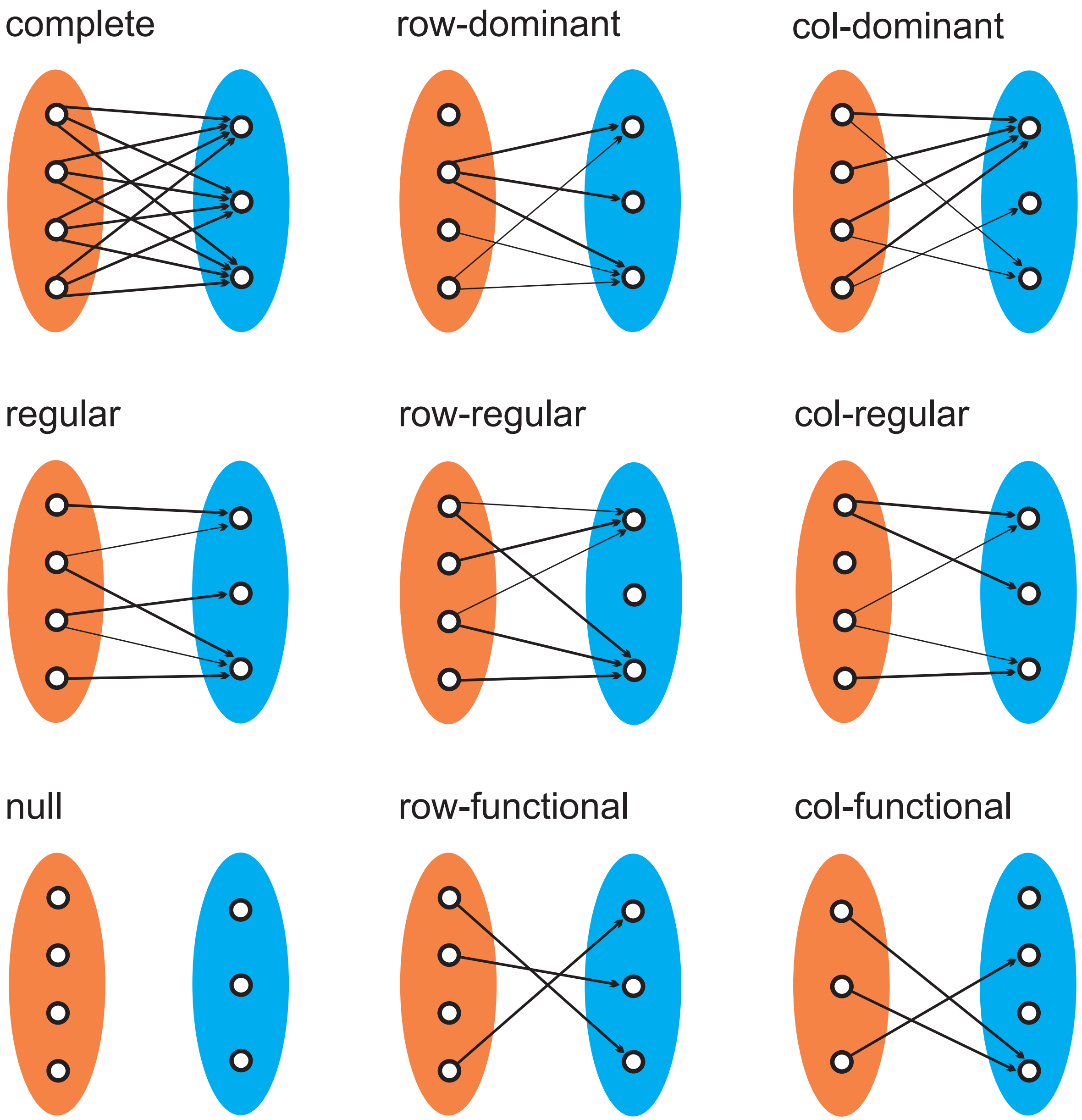
A block is *symmetric* if

$$\forall x, y \in C_i \times C_j : (xRy \Leftrightarrow yRx)$$

Note that for nondiagonal blocks this condition involves a pair of blocks $R(C_i, C_j)$ and $R(C_j, C_i)$.

The reduced graph can be presented by relational matrix, called also *image matrix*.

Block Types

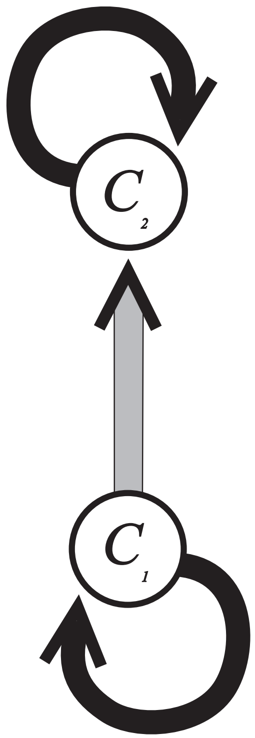

complete
row-dominant
col-dominant
regular
row-regular
col-regular
null
row-functional
col-functional

Characterizations of Types of Blocks

null	nul	all 0 (except may be diagonal)
complete	com	all 1 (except may be diagonal)
row-regular	rre	each row is 1-covered
col-regular	cre	each column is 1 -covered
row-dominant	rdo	$\exists$ all 1 row (except may be diagonal)
col-dominant	cdo	$\exists$ all 1 column (except may be diagonal)
regular	reg	1-covered rows and 1-covered columns
non-null	one	$\exists$ at least one 1

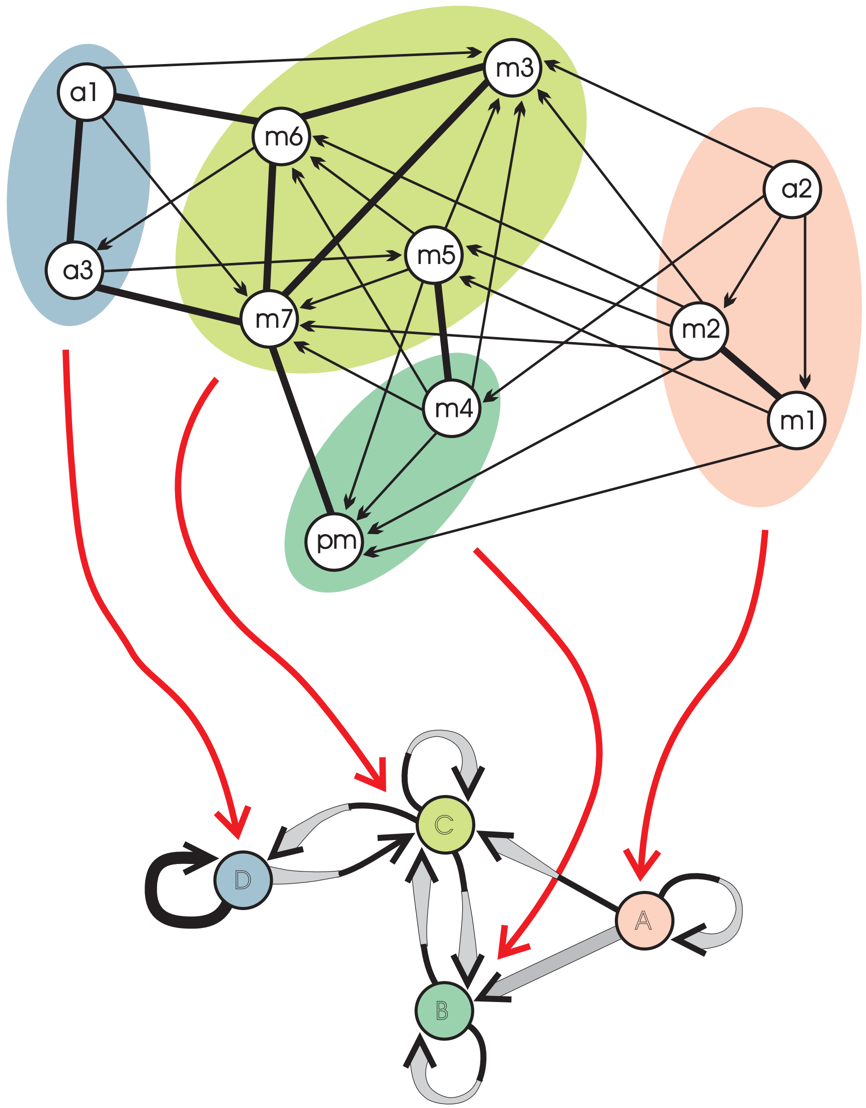
Block Types and Matrices

$$
\begin{array}{cccc|cccc}
1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\
1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 0
\end{array}
$$



	C_1	C_2
C_1	complete	regular
C_2	null	complete

Blockmodeling Example



Blockmodeling and Optimization

A criterion function

The problem of establishing a partition of units in a network in terms of a selected type of equivalence is a special case of **clustering problem** that can be formulated as an optimization problem: determine the clustering $\mathcal{C}^*$ for which

$$P(\mathcal{C}^*) = \min_{\mathcal{C} \in \Phi} P(\mathcal{C})$$

where $\mathcal{C}$ is a clustering of a given *set of units* E, Φ is the set of all feasible clusterings and $P : \Phi \to \mathbb{R}$ the *criterion function*.

One of the possible ways of constructing a criterion function that directly reflects the considered equivalence is to measure the fit of a clustering to an ideal one with perfect relations within each cluster and between clusters according to the considered equivalence.

Local Optimization

For solving the blockmodeling problem we use a local optimization procedure (relocation algorithm):

Determine the initial clustering $\mathcal{C}$;
repeat:
  if in the neighborhood of the current clustering $\mathcal{C}$
  there exists a clustering $\mathcal{C}'$ such that $P(\mathcal{C}') < P(\mathcal{C})$
  then move to clustering $\mathcal{C}'$.

The neighborhood in this local optimization procedure is determined by the following two transformations:

- *moving* a unit X_k from cluster C_p to cluster C_q (*transition*);
- *interchanging* units X_u and X_v from different clusters C_p and C_q (*transposition*).

Benefits from Optimization Approach

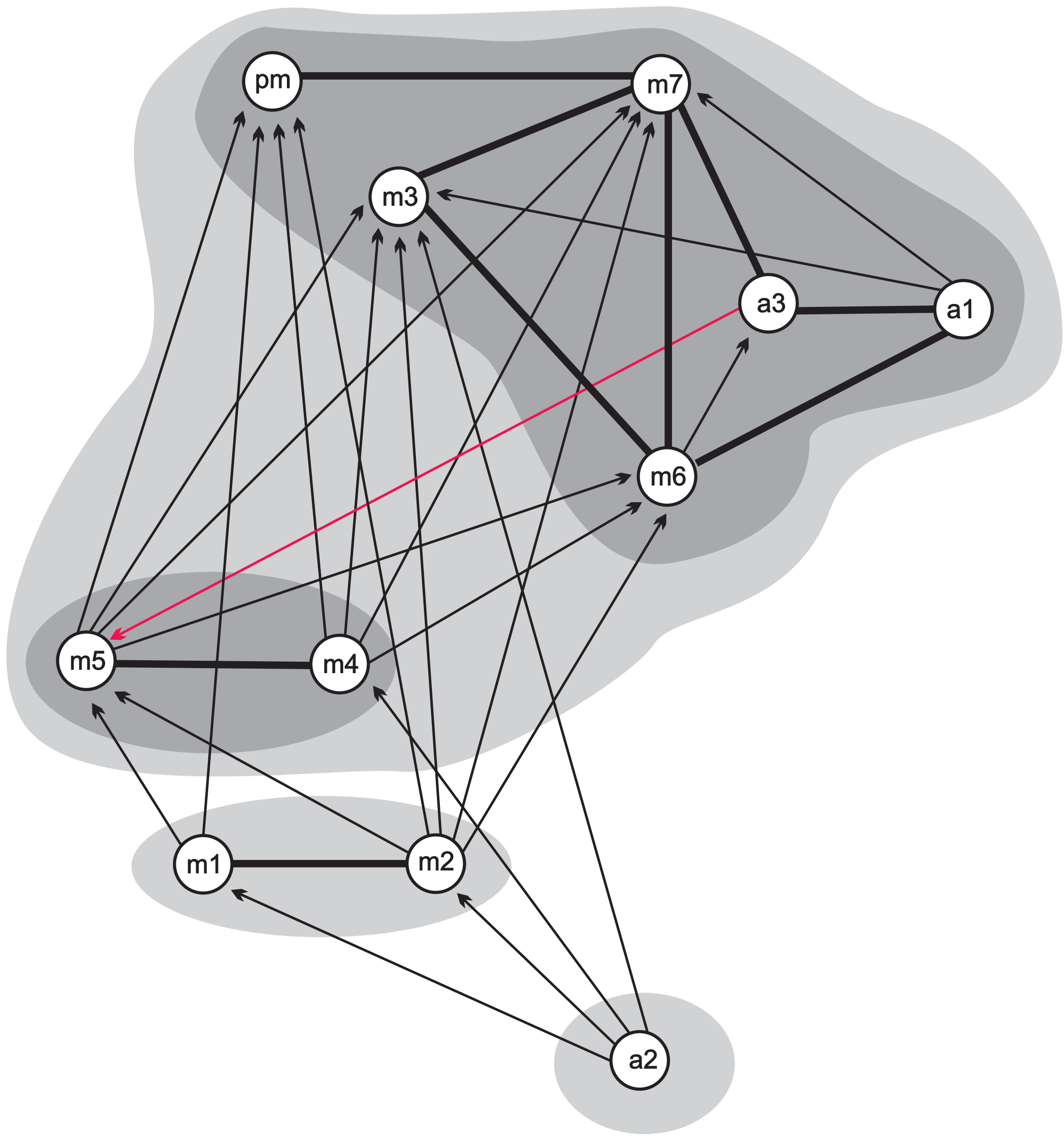
- *ordinary / inductive blockmodeling*: Given a network $\mathcal{N}$ and set of types of connection $\mathcal{T}$, determine the model $\mathcal{M}$;
- *evaluation of the quality of a model, comparing different models, analyzing the evolution of a network* (Sampson data, Doreian and Mrvar 1996): Given a network $\mathcal{N}$, a model $\mathcal{M}$, and blockmodeling μ, compute the corresponding criterion function;
- *model fitting / deductive blockmodeling*: Given a network $\mathcal{N}$, set of types $\mathcal{T}$, and a model $\mathcal{M}$, determine μ which minimizes the criterion function (Batagelj, Ferligoj, Doreian, 1998).
- we can fit the network to a partial model and analyze the residual afterward;
- we can also introduce different constraints on the model, for example: units x and y are of the same type; or, types of units x and y are not connected; …

Pre-Specified Blockmodels

The pre-specified blockmodeling starts with a blockmodel specified, in terms of substance, *prior to an analysis*. Given a network, a set of ideal blocks is selected, a reduced model is formulated, and partitions are established by minimizing the criterion function. The pre-specified blockmodeling is supported by the program MODEL 2 (Batagelj, 1996).

As an example of pre-specified blockmodel we present a symmetric acyclic blockmodel of Student Government. The obtained clustering in 4 clusters is almost exact – acyclic model with symmetric clusters. The only error is produced by the arc $(a3, m5)$.

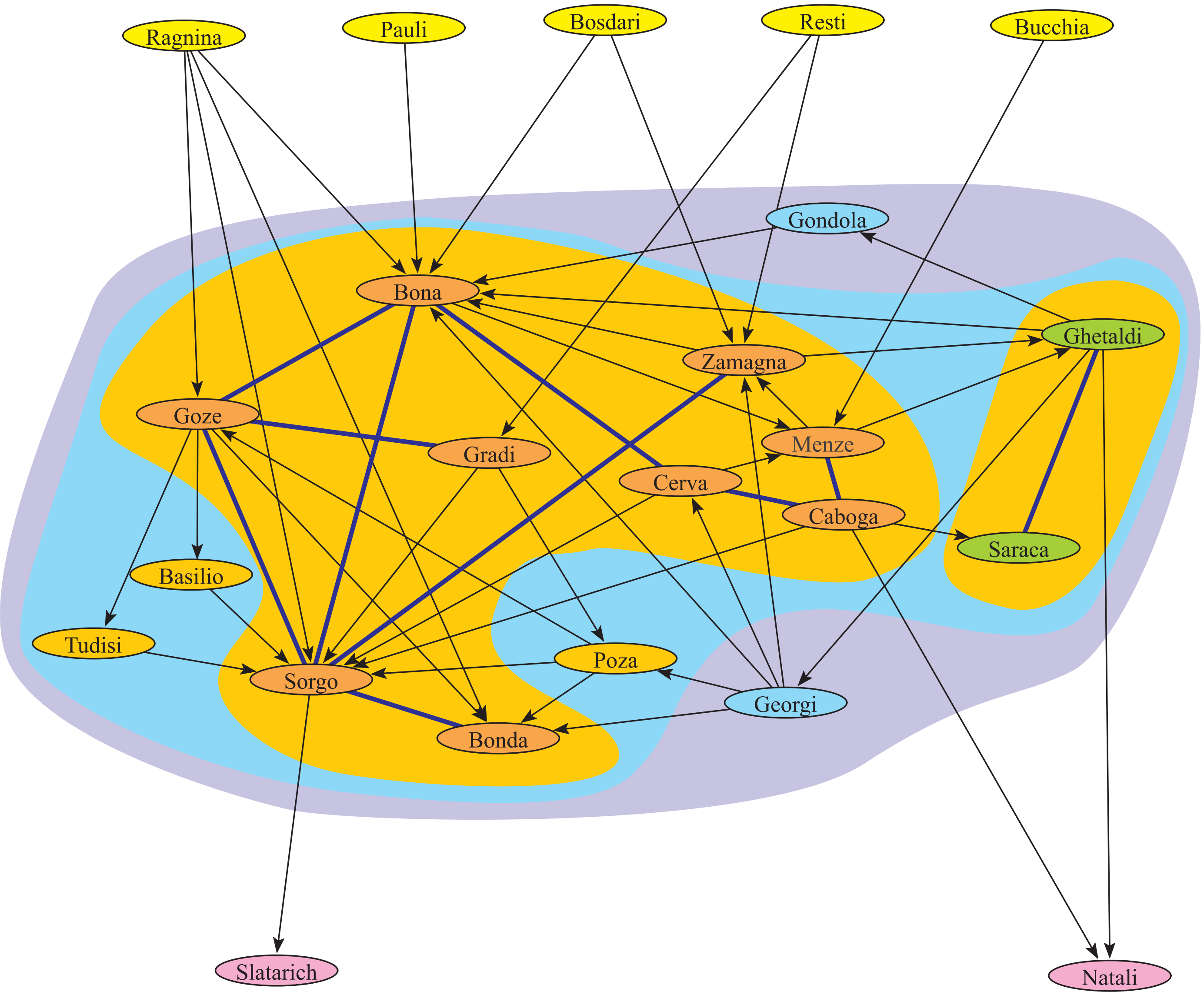
A Symmetric Acyclic Blockmodel of Student Government



Ragusan Noble Families Marriage Network, 18th and 19th Century

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Basilio	1	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	1	·	·
Bona	2	·	·	·	·	·	·	1	·	·	·	2	·	2	·	·	·	·	·	·	·	1	·	·
Bonda	3	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	2	·	·
Bosdari	4	·	1	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	1
Bucchia	5	·	·	·	·	·	·	·	·	·	·	·	·	1	·	·	·	·	·	·	·	·	·	·
Caboga	6	·	·	·	·	·	·	1	·	·	·	·	·	1	1	·	·	·	·	1	·	1	·	·
Cerva	7	·	1	·	·	·	1	·	·	·	·	·	·	1	·	·	·	·	·	·	·	1	·	·
Georgi	8	·	1	2	·	·	·	1	·	·	·	·	·	·	·	·	4	·	·	·	·	·	·	1
Ghetaldi	9	·	1	·	·	·	·	·	1	·	1	·	·	·	1	·	·	·	·	1	·	·	·	·
Gondola	10	·	1	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
Goze	11	1	2	1	·	·	·	·	·	·	·	2	2	·	·	·	·	·	·	·	·	2	1	·
Gradi	12	·	·	·	·	·	·	·	·	·	·	1	·	·	·	·	1	·	·	·	·	3	·	·
Menze	13	·	·	·	·	·	1	·	·	1	·	·	·	·	·	·	·	·	·	·	·	·	·	1
Natali	14	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
Pauli	15	·	1	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
Poza	16	·	·	2	·	·	·	·	·	·	·	1	·	·	·	·	1	·	·	·	·	1	·	·
Ragnina	17	·	1	1	·	·	·	·	·	·	·	1	·	·	·	·	·	·	·	·	·	1	·	·
Resti	18	·	·	·	·	·	·	·	·	·	·	·	1	·	·	·	·	·	·	·	·	·	·	1
Saraca	19	·	·	·	·	·	·	·	·	1	·	·	·	·	·	·	·	·	·	·	·	·	·	·
Slatarich	20	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
Sorgo	21	·	2	1	·	·	·	·	·	·	·	1	·	·	·	·	·	·	·	·	1	1	·	1
Tudisi	22	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	1	·	·
Zamagna	23	·	1	·	·	·	·	·	·	2	·	·	·	·	·	·	·	·	·	·	·	1	·	·

A Symmetric-Acyclic Decomposition of the Ragusan Families Network



Final Remarks

The current, local optimization based, programs for generalized blockmodeling can deal only with networks with at most some hundreds of units. What to do with larger networks is an open question. For some specialized problems also procedures for (very) large networks can be developed (Doreian, Batagelj, Ferligoj, 1998).

Another interesting problem is the development of blockmodeling of valued networks.

`http://vlado.fmf.uni-lj.si/pub/networks/`