Relational Calculator
a Tool for Analysing Social Networks

Vladimir Batagelj and Andrej Mrvar

University of Ljubljana

International Social Network Conference

London, July 6th -10th 1995

1. Relational networks, clusterings and permutations

Let $\mathbf{E} = \{X_i\}$ be a *set of units*. A *relational network* on $\mathbf{E}$ is called a structure

$$(\mathbf{E}, R_1, R_2, \ldots, R_k)$$

where $R_s \subseteq \mathbf{E} \times \mathbf{E}$, $s = 1, \ldots, k$.

For computational purposes relations are often represented by *binary matrices* $B(R) = [b_{ij}]$ where

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

A representation by *directed graphs* is usually used to visualize networks or to give intuitive meaning to some notions.

Besides networks, two additional types of objects are needed:

- *clusterings* $c \colon \mathbf{E} \to 0..n$.
  $c(i) = j$ iff unit i belongs to cluster j.
  $c(i) = 0$ – unit i is not selected. $c(i) \in 0, 1$ – sets.
- *permutations* $\pi \colon \mathbf{E} \to \mathbf{E}$.

2. Program `RelCalc`

Program `RelCalc` implements a computational system over these three data types.

In the current imlementation `RelCalc` implements the following tools for analysing relations:

2.1 Properties of relations

- reflexivity, transitivity, symmetricity, equivalence relation, acyclic relation, ...
- equal, unequal relations, subrelations.
- distances between relations: Hamming distance, ...

2.2 Basic operations on relations

- complement, transposition, union, intersection, symmetric sum, difference, multiplication, power, transitive closure.

2.3 Complex operations on relations

- *Factorisation*: returns as a result strongly connected components (clustering) and the factor relation.
- *Degree partition*: partition of units according to input, output or total degree.
- *Depth partition*; can be applied only if a relation is acyclic.
- *Core partition*: k-core is a maximal connected subgraph in which every unit has at least k neighbours from the same k-core.
- *Biconnected components*: maximal subgraphs without articulation points.
- *Transitive skeleton.*
- *Concatenation, substitution and domain alternation.*

2.4 Additional operations

- Creating standard and random relations, clusterings and permutations.
- Transforming clustering to permutation or equivalence relation.
- Rearranging relational matrix according to a given permutation.
- Transforming clustering to its canonical form (to compare two clusterings easier).
- Transforming clustering into binary clustering (to select 'interesting' units and their relations from network).

- Extracting subrelation, determined by a binary clustering, from a given relation.
- Shrinking of selected units to one unit (to make analysis of larger networks easier).
- Transforming permutation to clustering (orbital partition).
- Mirroring of permutation.
- Union and intersection of clusterings.
- Determining the direct neighbours of a given binary clustering with respect to a given relation.
- Comparing clusterings and permutations.

2.5 Log-file

RelCalc is usually used in the interactive mode. It writes the trace of all user's requests onto a log-file. This file can be later rerun to repeat the session and to restore the current state; or we can prepare on such files sequences of commands for standard procedures.

Example: dissimilarity between clusterings c_1 and c_2

$$d(c_1, c_2) = \mid \mathrm{eqv}(c_1 \sqcup c_2) \oplus \mathrm{eqv}(c_1 \sqcap c_2) \mid$$

```
% RelCalc 0.1, April 28, 1995
Msg Input first clustering
RDC ?                       %%%  C   1
Msg Input second clustering
RDC ?                       %%%  C   2
Msg Calculation
UnC(1,2)                    %%%  C   3
InC(1,2)                    %%%  C   4
ClR-3                       %%%  R   1
ClR-4                       %%%  R   2
Sum(1,2)                    %%%  R   3
Msg Result
```

3. Example: Set Operations and Dissimilarities

Using set operations on relations, we can often reveal interesting connections between these relations. In the case of the Wire room we have, for example:

$$\begin{aligned} \text{negative} \cap \text{positive} &= \emptyset \\ \text{negative} \cap \text{help} &= \{(W_6, W_7)\} \\ \text{positive} \setminus \text{games} &= \{(W_7, S_1), (S_1, W_7)\} \end{aligned}$$

The following partition into cliques was proposed in the literature

$$\begin{aligned} \mathcal{C} = \{ & \{I_1, W_1, W_2, W_3, W_4, S_1\},\ \{I_3\},\ \{W_5\}, \\ & \{W_6, W_7, W_8, W_9, S_4\},\ \{S_2\} \ \} \end{aligned}$$

Let $\text{eqv}(\mathcal{C})$ denote the corresponding equivalence relation. If we compute the relation

$$\text{games} \setminus \text{eqv}(\mathcal{C})$$

we can see that the only unit which is in relation with others is the unit W_5.

It is adjacent to $\{W_1, W_3, W_4, W_7, S_1\}$. All these units except W_7 are in first class of the partition $\mathcal{C}$. If we move the unit W_5 into the first class, we get a new partition

$$\begin{aligned} \mathcal{C}' &= \{ \{I_1, W_1, W_2, W_3, W_4, W_5, S_1\},\ \{I_3\}, \\ &\quad \{W_6, W_7, W_8, W_9, S_4\},\ \{S_2\} \} \end{aligned}$$

for which games $\setminus$ eqv$(\mathcal{C}') = \{(W_5, W_7), (W_7, W_5)\}$

We can measure the strength of connectedness between two relations by a *coefficient of determinacy of relation R by relation S*

$$K(R \Rightarrow S) = \frac{| R \cap S |}{| R |}$$

$K(R \Rightarrow S) = 1$, if $R = \emptyset$; or by a *coefficient of mutual determinacy of relations R and S*

$$K(R \Leftrightarrow S) = \frac{| R \cap S |}{| R \cup S |}$$

$K(R \Leftrightarrow S) = 1$, if $| R \cup S | = \emptyset$.

Since $K(\mathrm{eqv}(\mathcal{C}') \Rightarrow \mathrm{games}) = 0.96429$ the games relation can be almost completely explained by the property of belonging to the same clique in $\mathcal{C}'$.

We can also define dissimilarities between relations.

$$\begin{aligned} D_H(R,S) &= \mid R \oplus S \mid \\ D_U(R,S) &= \frac{\mid R \oplus S \mid}{\mid R \cup S \mid} \\ D_M(R,S) &= \frac{\max(\mid R \setminus S \mid, \mid S \setminus R \mid)}{\max(\mid R \mid, \mid S \mid)} \end{aligned}$$

where $D_U(R,S) = D_M(R,S) = 0$, if $R \cup S = \emptyset$. We computed the dissimilarity matrices for Wire Room relations and applied the hierarchical clustering to them. The resulting dendrograms show that they measure different aspects of the data.