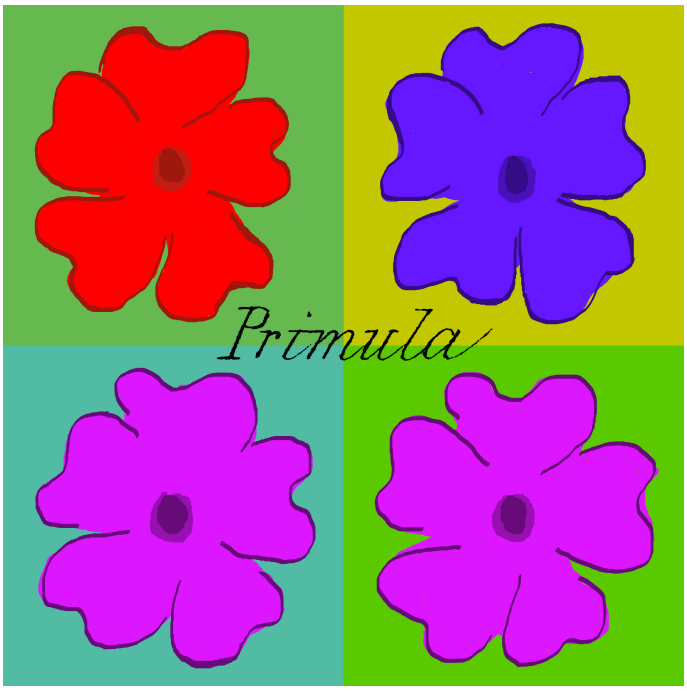
The *Primula* System: user's guide
Version 3.0
Example: Community detection

Manfred Jaeger
Institut for Datalogi, Aalborg Universitet, Selma Lagerløfs Vej 300, 9220 Aalborg Ø
jaeger@cs.aau.dk

Primula homepage: www.cs.aau.dk/∼jaeger/Primula

September 19, 2021



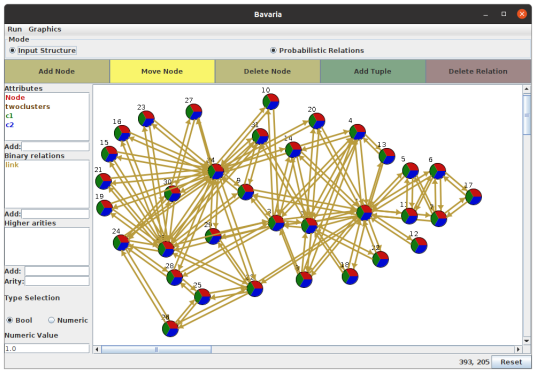
Community Detection

We use latent variable models for learning soft community membership degrees in single or multi-relational social networks. This example illustrates the methods described in

> Jiuchuan Jiang and Manfred Jaeger: Numeric Input Relations for Relational Learning with Applications to Community Structure Analysis, ArXiv 1506.05055, 2015.

0.1 Karate club network

We first consider the classic Karate Club network contained in `zachary.rdef`. We use the model `community_softclus_2c.rbn`. Opening the structure in Bavaria, and selecting both 'Input Structure' and 'Probabilistic Relations' for display shows:



Apart from the edge relation, the graph shows a Boolean node (type) attribute *node*, and two numeric node attributes *c1*,*c2* representing membership degrees in communities 1 and 2. The values for these membership degrees are initialized to a common default value. We now learn membership degrees by maximizing likelihood of the model contained in `community_softclus_2c.rbn`:

- Open the learn module; In Settings:Gradient Ascent select the 'Batch' top and 'LBFGS' sub-strategy.
- In the 'Relation Parameters' tab, select the three relations 'alpha','c1','c2'. *alpha* represents a single constant that is a bias parameter of the model.
- In the 'Learning' tab press 'Learn'
- Within a few restarts, a solution with a log-likelihood value (displayed at the bottom of the parameter value table) around -200. Stop the learning at this point.
- Press the 'Set' button to set the learned values in the input structure

The learned values for the *c1*,*c2* attributes are visualized by the saturation of the colors associated with the attributes: