

Input: a data set of n variables, i.e., genomic alterations or patterns, and m samples.

Result: a graphical model representing all the relations of “selective advantage.”

Pruning based on the Suppes' criteria

- 1 Let $G \leftarrow$ a directed graph over the vertices n
- 2 **forall** arcs $(a, b) \in G$ **do**
- 3 Compute a score $S(\cdot)$ for the nodes a and b in terms of Suppes' criteria.
- 4 Remove the arc (a, b) if Suppes' criteria are not met.
- 5 **end**

Likelihood fit on the prima facie directed acyclic graph

- 6 Let $\mathcal{M} \leftarrow$ the subset of the remaining arcs $\in G$, that maximize the log-likelihood of the model, computed as:
 $LL(D \mid \mathcal{M}) - ((\log m)/2) \dim(\mathcal{M})$, where D denotes the input data, m denotes the number of samples, and $\dim(\mathcal{M})$ denotes the number of parameters in \mathcal{M} (see Koller and Friedman(2009)).
- 7 **return** *The resulting graphical model \mathcal{M} .*