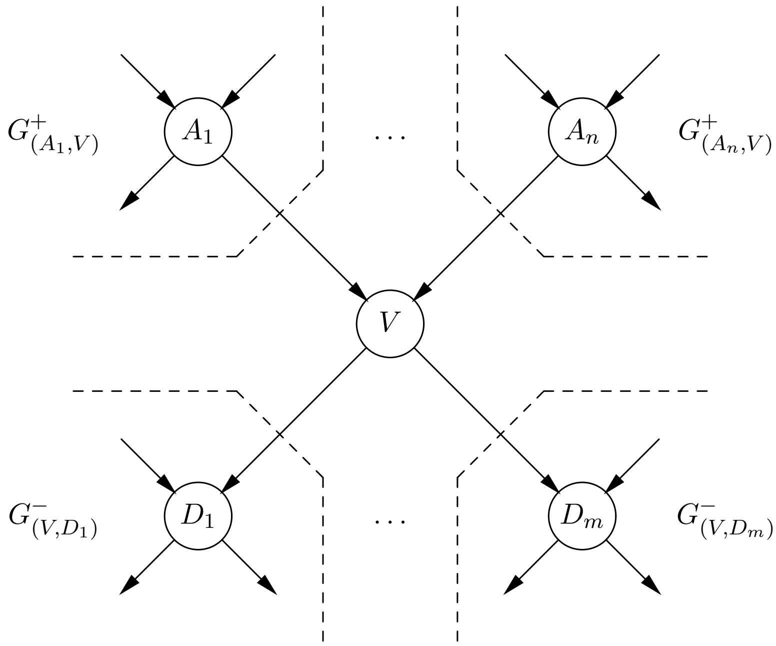
Formulas (Probabilistic Reasoning)

Pearl in a singly connected digraph



Consider a node V in a probabilistic network $B = (G, \Gamma)$. Let $\rho(V) = \{A_1, \ldots, A_n\}$ be the set of direct ancestors (parents) of V in G, and let $\sigma(V) = \{D_1, \ldots D_m\}$ be the set of its direct descendants (children). With Pearl's algorithm, node V computes the following parameters:

$$\begin{aligned}
\pi(V) &= \sum_{c_{\rho(V)}} \Big(\gamma\left(V \mid c_{\rho(V)}\right) \cdot \prod_{i=1,\ldots,n} \pi_V^{A_i}(c_{A_i}) \Big) \\
\lambda(V) &= \prod_{j=1,\ldots,m} \lambda_{D_j}^{V}(V) \\
\pi_{D_j}^{V}(V) &= \alpha \cdot \pi(V) \cdot \prod_{\substack{k = 1, \ldots, m \\ k \neq j}} \lambda_{D_k}^{V}(V) \\
\lambda_V^{A_i}(A_i) &= \alpha \cdot \sum_{c_V} \lambda(c_V) \cdot \sum_{c_{\rho(V)\setminus\{A_i\}}} \Big(\gamma\left(c_V \mid c_{\rho(V)\setminus\{A_i\}} \wedge A_i\right) \cdot \prod_{\substack{k = 1, \ldots, n \\ k \neq i}} \pi_V^{A_k}(c_{A_k}) \Big)
\end{aligned}$$

in order to perform *data fusion*: $\alpha \cdot \pi(V) \cdot \lambda(V)$, which results in V's (prior or posterior) probability distribution.

The MDL quality measure

Let $G = (V(G), A(G))$ be an acyclic digraph and let D be a dataset over N cases. Let $P(G)$ be a probability distribution over the set of acyclic graphs with node set V. Then, the MDL quality measure for graph G is given by

$$\begin{aligned} Q(G, D) &= \log P(G) - N \cdot H(G, D) - \frac{1}{2} \log N \cdot \sum_{V_i \in V} 2^{|\rho(V_i)|} \\ &= \log P(G) + \sum_{V_i \in V} q(V_i, \rho(V_i), D) \end{aligned}$$

where

$$-N \cdot H(G, D) = \sum_{V_i \in V} \sum_{c_{V_i}} \sum_{c_{\rho(V_i)}} N(c_{V_i} \wedge c_{\rho(V_i)}) \cdot \log\Big(\frac{N(c_{V_i} \wedge c_{\rho(V_i)})}{N(c_{\rho(V_i)})}\Big)$$

and $q(V_i, \rho(V_i), D)$ is the quality of node V_i.