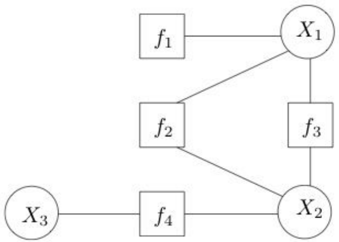
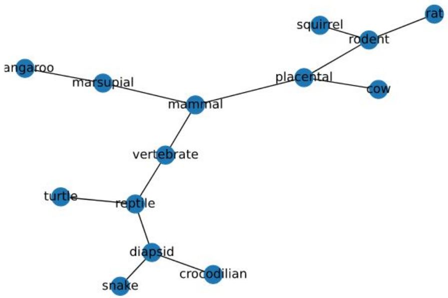
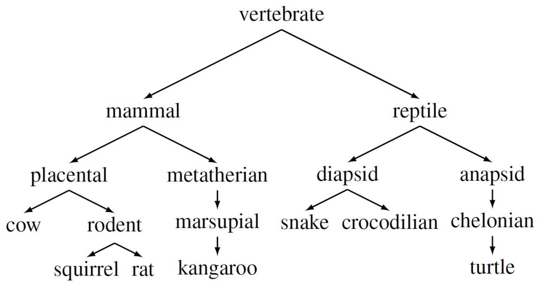




Graphical Models
Belief Propagation Algorithm

Clément Bonnet, Charbel-Raphaël Segerie



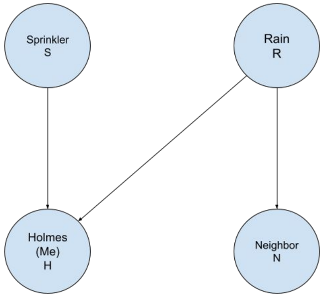
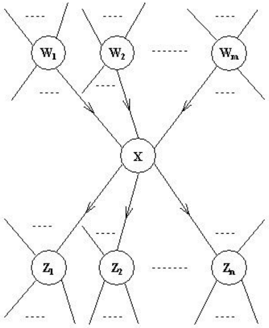
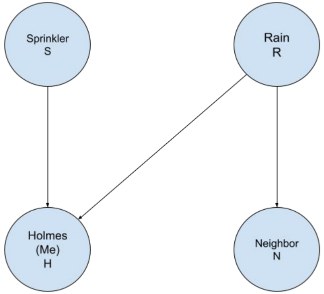
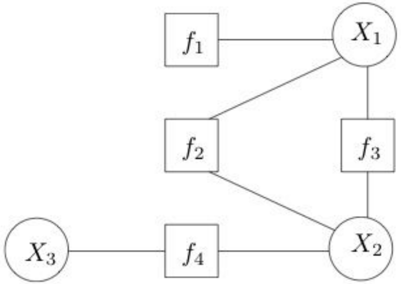
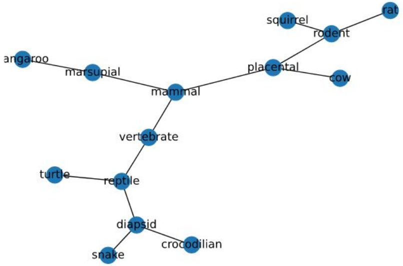


Table of Contents

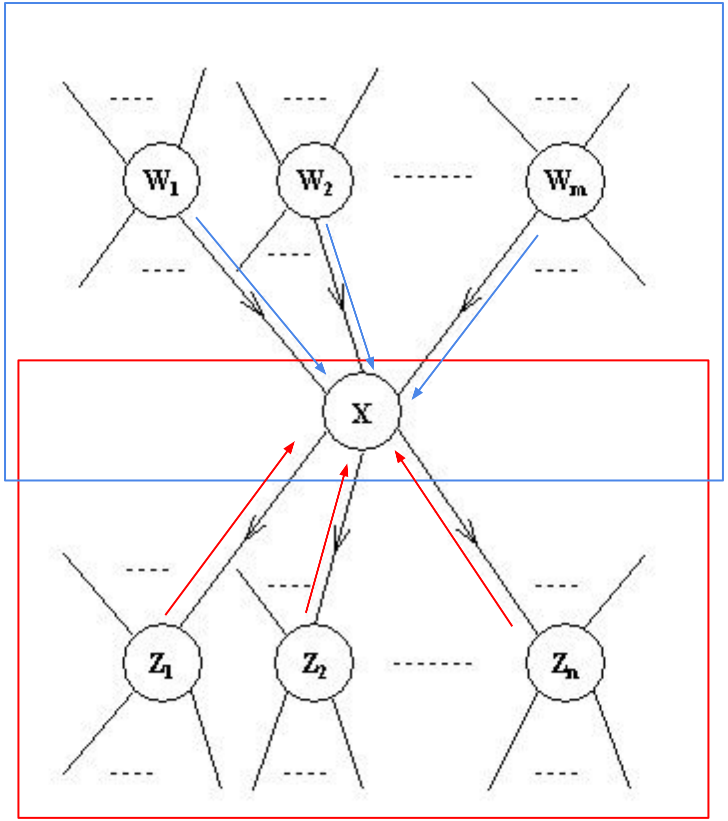








Bayesian network: Propagation Rules



Acyclic graph, but messages flow in both direction.

message Child -> Parent

Likelihood $\lambda(x) = \mathbb{P}(Z_1, \ldots, Z_n | X=x) = \prod_{j=1}^{n} \lambda_{Z_j}(x)$

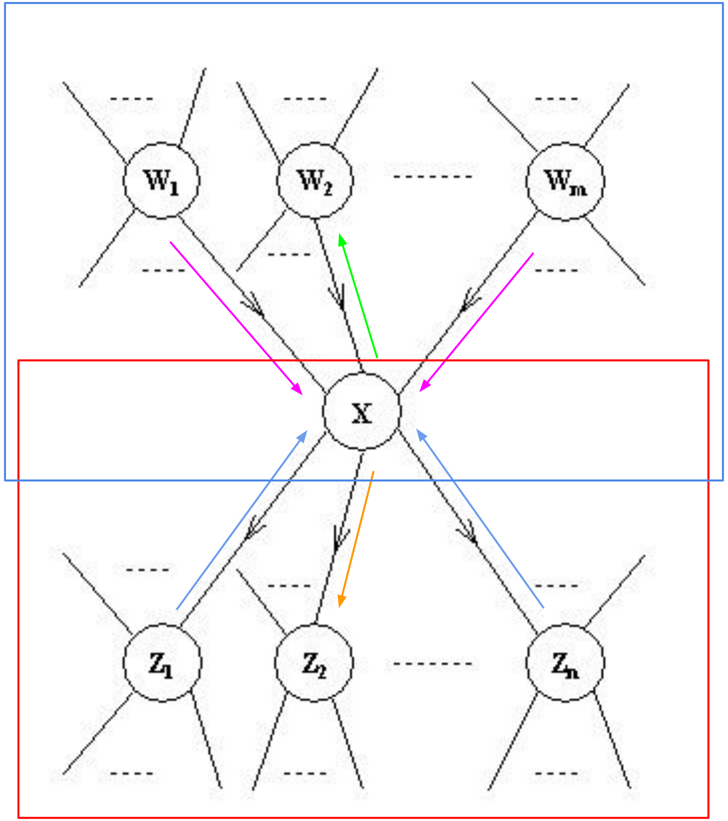
message Parent-> Child

Prior $\pi(x) = \sum_{\mathbf{w} \in \{0,1\}^m} \mathbb{P}(X=x|\mathbf{w}) \prod_{k=1}^{m} \pi_X(w_k)$

Belief

$$
\begin{aligned}
BEL(x) &\equiv \alpha\lambda(x)\pi(x) \\
&= \alpha \boxed{\left[\prod_{j=1}^{n} \lambda_{Z_j}(x)\right]} \boxed{\left[\sum_{\mathbf{w}} P(x \mid \mathbf{w}) \prod_{k=1}^{n} \pi_X(w_k)\right]}
\end{aligned}
$$

Bayesian network: Updates



Model of the world specified by the user

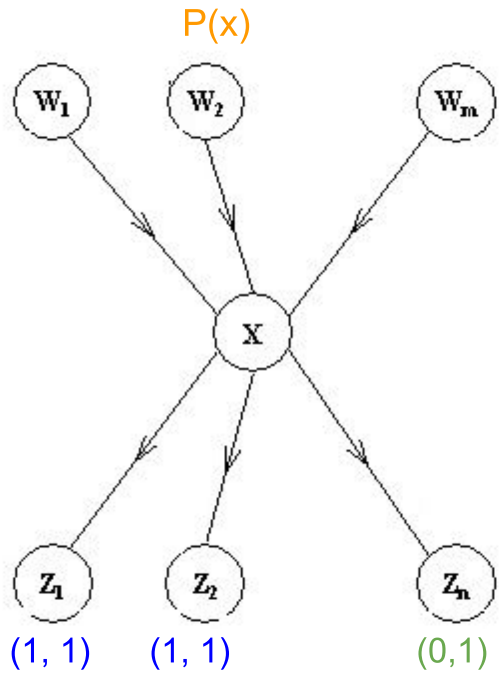
Likelihood:

$$\lambda_X(w_i) \propto \sum_x \boxed{\lambda(x)} \sum_{\substack{\mathbf{w}' \in \{0,1\}^m \\ w'_i = w_i}} \mathbb{P}(X = x|\mathbf{w}') \prod_{k\neq i} \pi_X(w'_k)$$

Prior:

$$\pi_{Z_j}(x) \propto \boxed{\pi(x)} \prod_{k\neq j} \lambda_{Z_k}(x)$$

Bayesian network: Boundary Conditions



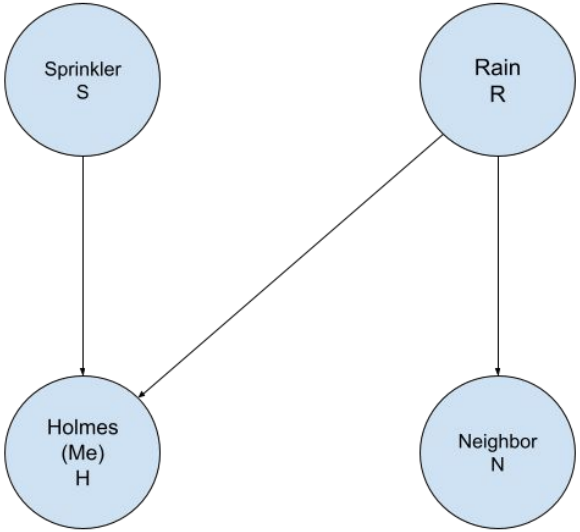
Example for binary states

Root nodes: If w is a node with no parents, we set the prior value π value equal to the prior probabilities of P(x)

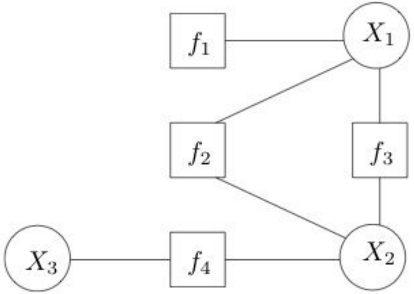
Anticipatory nodes: If X is a childless node that has not been instantiated, we set the λ value as a vector of all 1's

Evidence nodes: If evidence $X=x_i$ is obtained, we set the λ value to (0,...,0,1,0,...0) with 1 at the i_{th} position.

Bayesian network: Case study

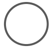


Factor Graphs: Theory



Factor Graph = Bipartite Graph

 Factor node

 Variable node

Total likelihood:

$$p(\mathbf{x}) = \prod_{f \in F} \mathbf{f}_f(\mathbf{x}_f)$$

Propagation:

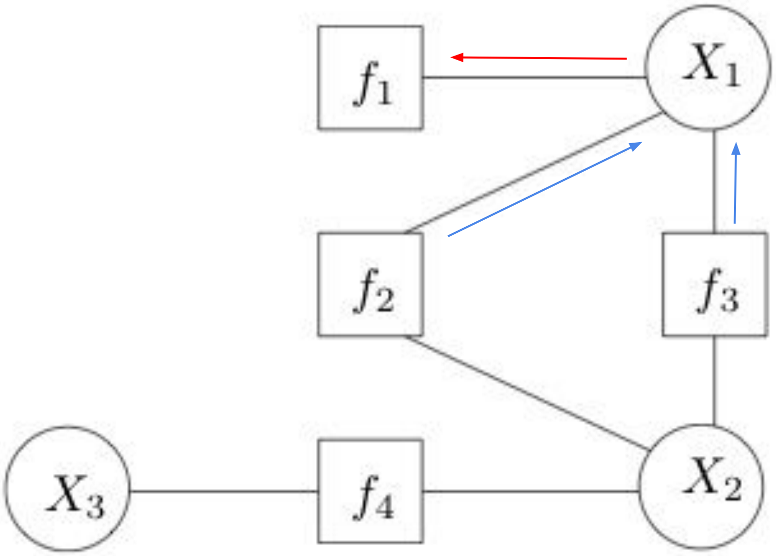
$$\mu_{v \to f}(x_v) = \prod_{f^\star \in N(v) \setminus \{f\}} \mu_{f^\star \to v}(x_v)$$

$$\mu_{f \to v}(x_v) = \sum_{\mathbf{x}'_f : x'_v = x_v} \mathbf{f}_f(\mathbf{x}'_f) \prod_{v^* \in N(f) \setminus \{v\}} \mu_{v^* \to f}(x'_{v^*})$$

Marginals:

$$p_{X_v}(x_v) \propto \prod_{f \in N(v)} \mu_{f \to v}(x_v)$$

Factor Graphs: example of propagation: variable to factor node



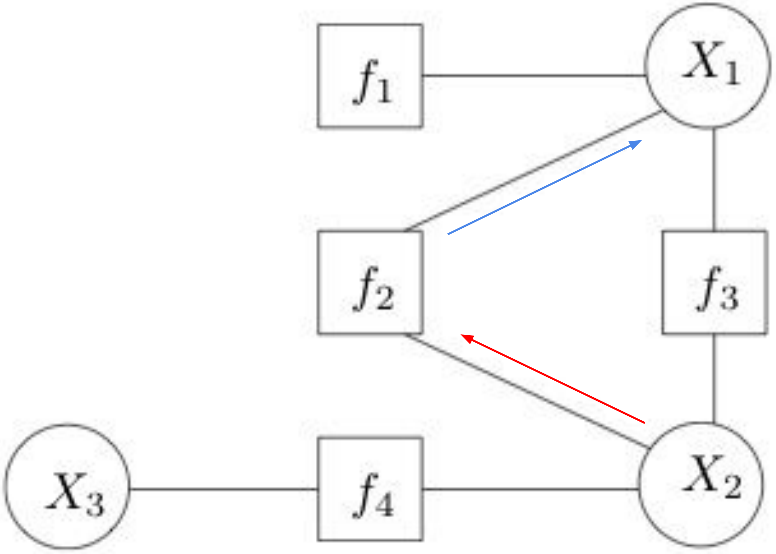
$$\overrightarrow{\mu_{v \to f}}(x_v) = \prod_{f^\star \in N(v) \setminus \{f\}} \overrightarrow{\mu_{f^\star \to v}}(x_v)$$

Messages from variable X to factor nodes f are vectors of size NbStates(X)

Messages from factor f to variable X are also vectors of size NbStates(X)

We just multiply coordinate by coordinates the two incoming messages

Factor Graphs: example of propagation: factor to variable node



$$\overrightarrow{\mu_{f\to v}}(x_v) = \sum_{\mathbf{x}'_f : x'_v = x_v} \mathrm{f}_f(\mathbf{x}'_f) \prod_{v^* \in N(f) \setminus \{v\}} \overrightarrow{\mu_{v^* \to f}}(x'_{v^*})$$

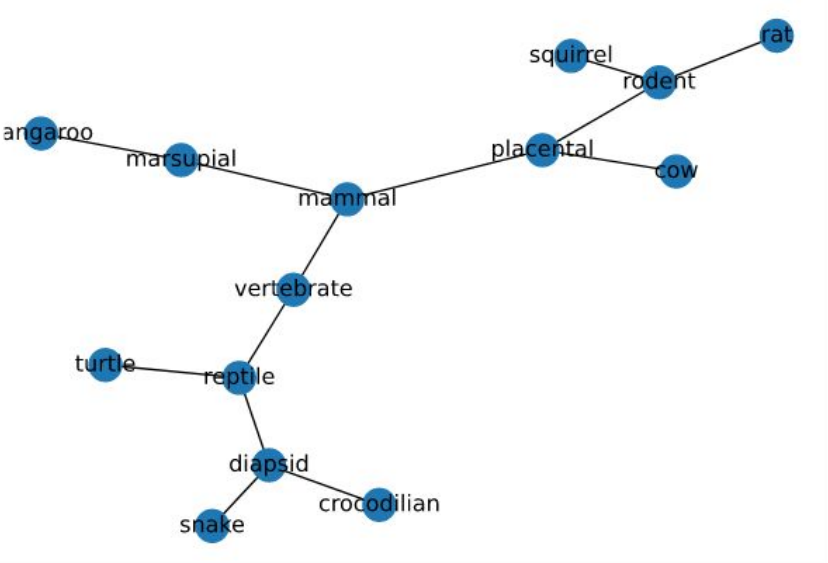
Messages from variable X to factor nodes f are vectors of size NbStates(X)

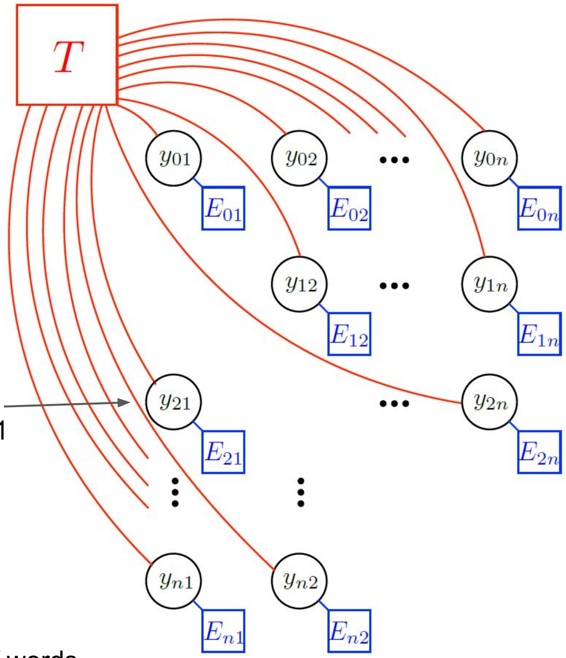
Messages from factor f to variable X are also vectors of size NbStates(X)

We just multiply coordinate by coordinates the incoming messages

Factor Graphs: Taxonomy Project

From a corpus of abstracts on Wikipedia, and a selection of words (x1, …, xn), the objective is to discover the taxonomic structure.





1 if word x2 is the parent of x1

$$P(y|x) \propto \prod_F \phi_F(y)$$

Embedding of the pair of words

$$\boxed{\phi_{E_{i,j}}(y_{i,j})} = \begin{cases} \exp(\mho(x_i, x_j) \cdot w) & \text{if } y_{i,j} = 1 \\ \exp(0) = 1 & \text{if } y_{i,j} = 0 \end{cases}$$

$$\boxed{\phi_T(y_{i,j})} = \begin{cases} 1 & \text{if } y \text{ forms a legal taxonomy tree} \\ 0 & \text{otherwise} \end{cases}$$