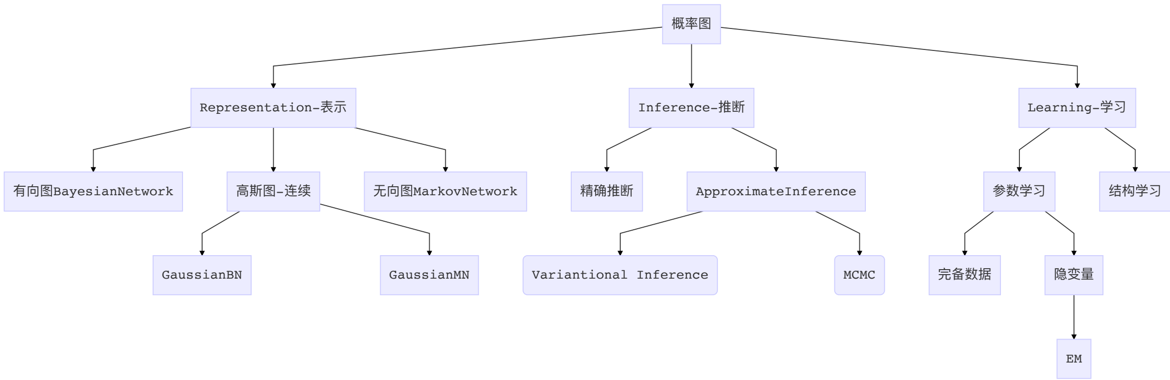
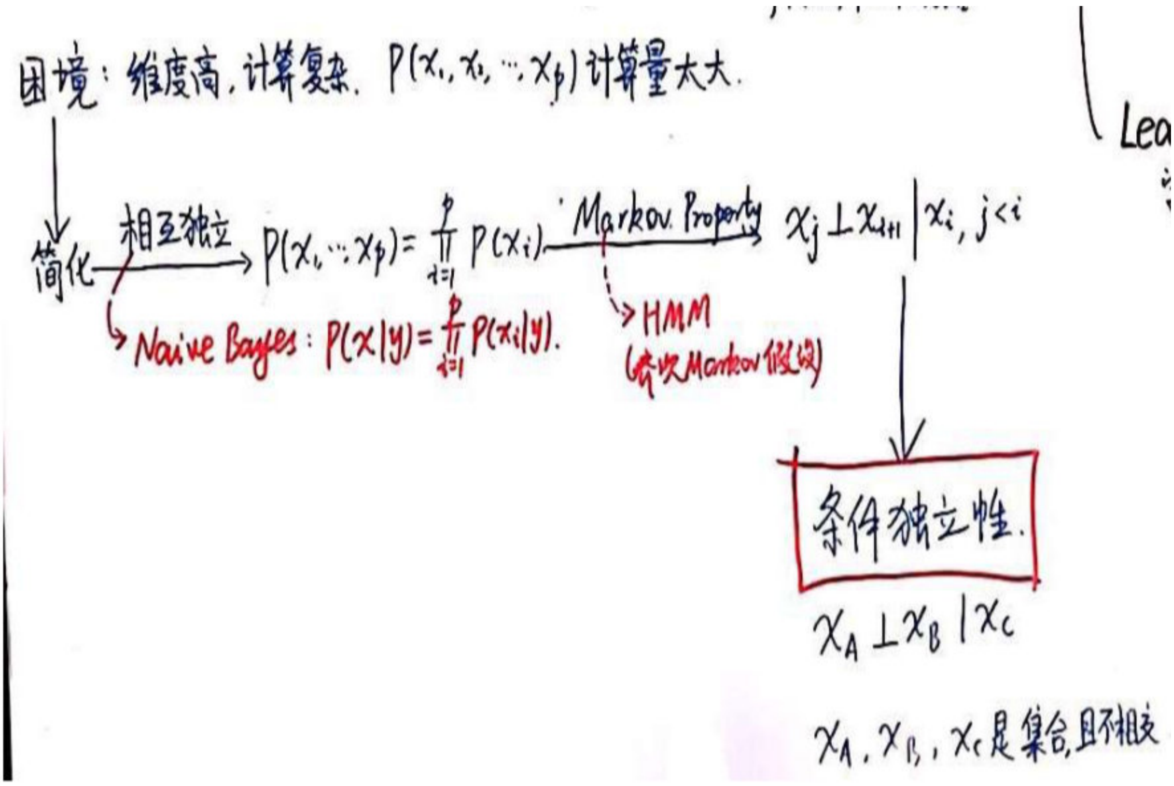


$$\text{Sum Rule: } P(x_1) = \int P(x_1, x_2)\, dx_2$$
$$\text{Product Rule: } P(x_1, x_2) = P(x_1) \cdot P(x_2|x_1) = P(x_2) \cdot P(x_1|x_2)$$

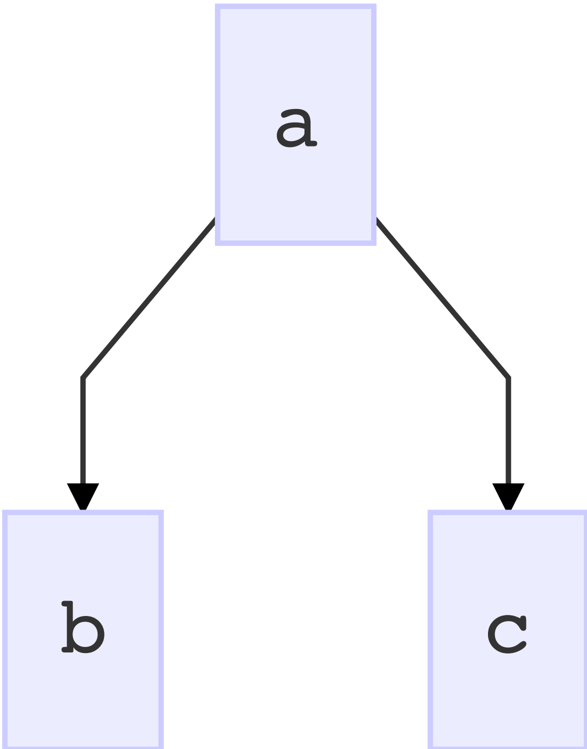
Chain Rule:

$$\begin{aligned} \mathrm{P}(X_4, X_3, X_2, X_1) &= \mathrm{P}(X_4 \mid X_3, X_2, X_1) \cdot \mathrm{P}(X_3, X_2, X_1) \\ &= \mathrm{P}(X_4 \mid X_3, X_2, X_1) \cdot \mathrm{P}(X_3 \mid X_2, X_1) \cdot \mathrm{P}(X_2, X_1) \\ &= \mathrm{P}(X_4 \mid X_3, X_2, X_1) \cdot \mathrm{P}(X_3 \mid X_2, X_1) \cdot \mathrm{P}(X_2 \mid X_1) \cdot \mathrm{P}(X_1) \end{aligned}$$



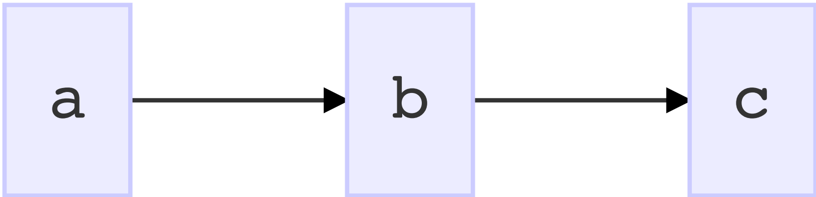
Bayesian Network

1. 拓扑排序构建图



$$
\begin{gathered}
ChainRule: P\left(x_1, x_2, \cdots, x_p\right) = P\left(x_1\right) \cdot \prod_{i=2}^{p} p\left(x_i \mid x_{1:i-1}\right) \\
P(a,b,c) = P(a)P(b|a)(c|a) \rightarrow \text{因子分解} \\
P(a,b,c) = P(a)P(b|a)(c|a,b) \rightarrow Chainrule \\
P(c|a) = P(c|a,b) \Rightarrow c \perp b|a \\
p(c|a) \cdot p(b|a) = p(c|a,b) \cdot p(b|a) = p(b,c|a) \\
p(c|a) \cdot p(b|a) = p(b,c|a)
\end{gathered}
$$

Tail to tail, 若a被观测，则路径被堵塞$tail \rightarrow head$

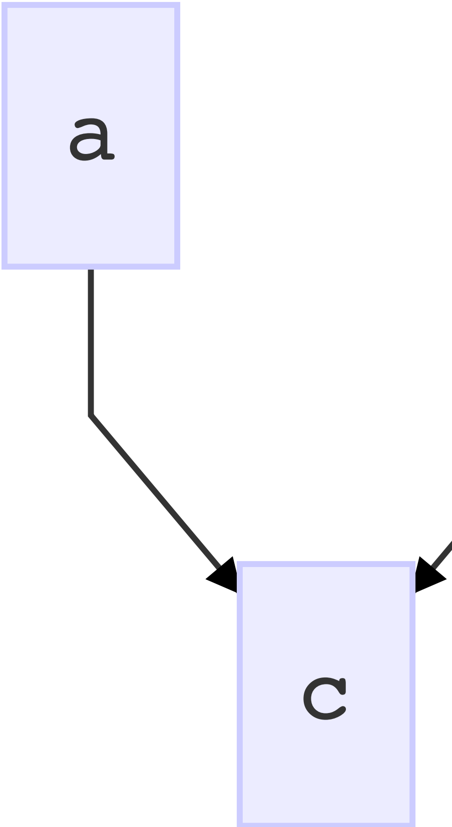


head to tail

$$
\begin{gathered}
P(a,b,c) = P(a)P(b|a)P(c|b) \\
P(a,b,c) = P(a)P(b|a)P(c|a,b) \\
P(c|b) = P(c|a,b)
\end{gathered}
$$

$$a \perp c|b$$

若b被观测，则路径被阻塞（independent）



head to head

默认情况下，$a \perp b$，路径阻塞的

若c被观测，则路径是通的

$$
\begin{gathered}
P(a,b,c) = P(a) \cdot P(b) \cdot P(c|a,b) \\
P(a,b,c) = P(a) \cdot P(b|a) \cdot (c|a,b) \\
P(b) = P(b|a)
\end{gathered}
$$

Inference

$$\text{sum rule } p(X)=\sum_{Y} p(X,Y)$$
$$\text{product rule } p(X,Y)=p(Y|X)p(X)$$

求概率： $P(x)=P\left(x_0, x_1, \cdots, x_p\right)$

边缘概率*marginal* probability：

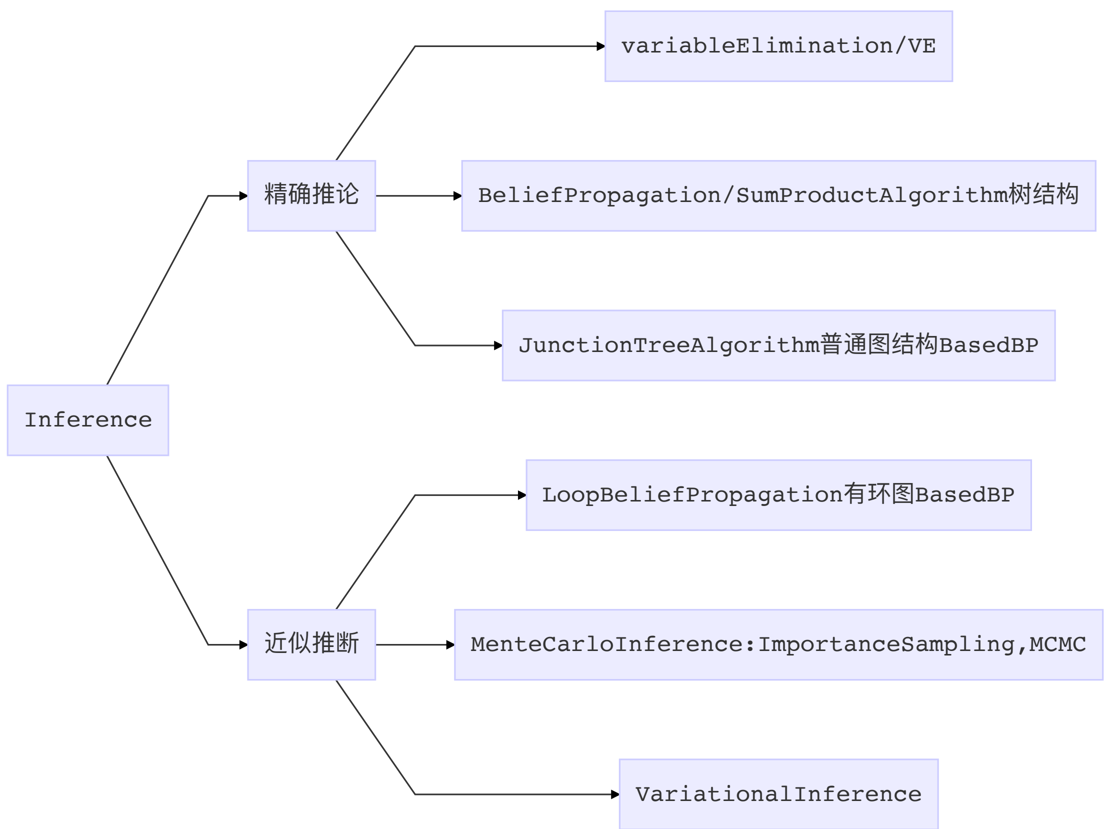
$$P\left(x_i\right)=\sum_{x_1} \cdot \sum_{x_{i-1}} \sum_{x_{i+1}} \cdots \sum_{x_p} p(x)$$

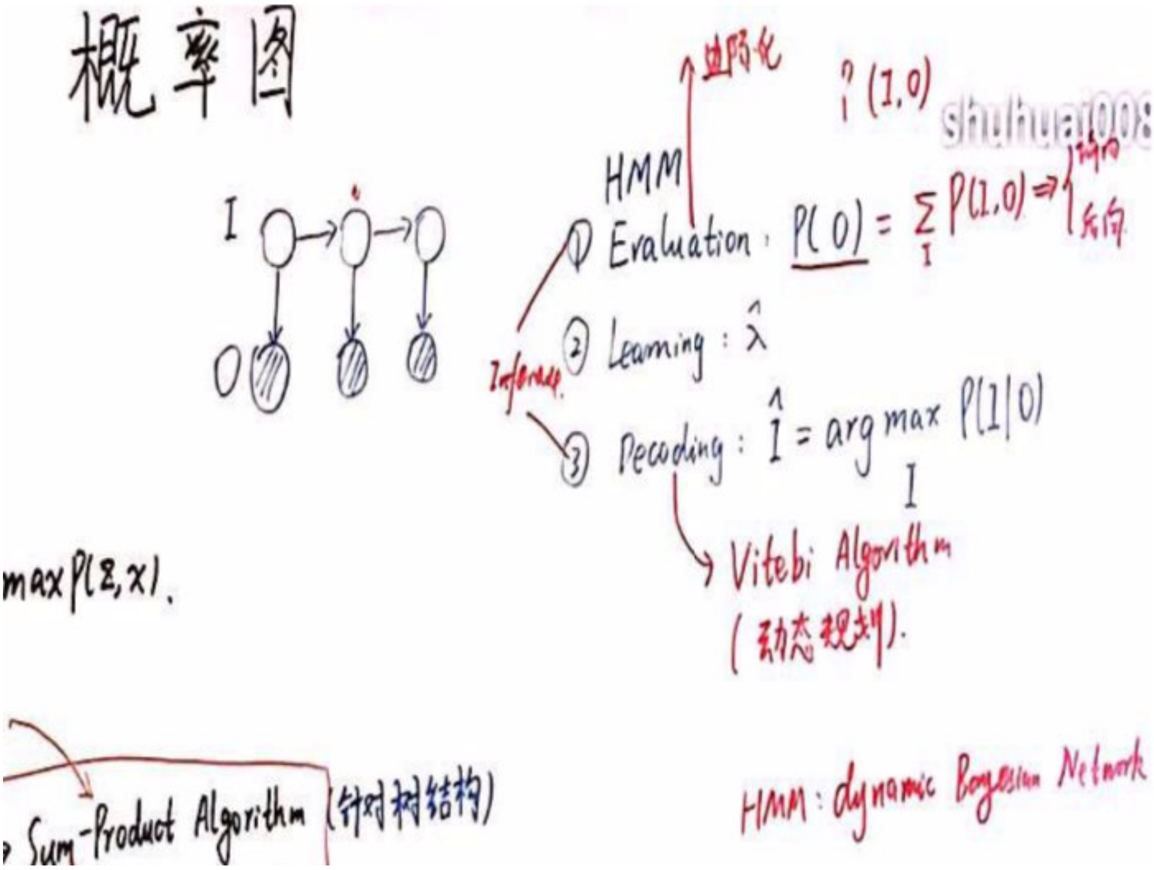
条件概率conditional probability：

$$P\left(x_A | x_B\right) \quad x=x_A \cup x_B$$

MAP Inference:

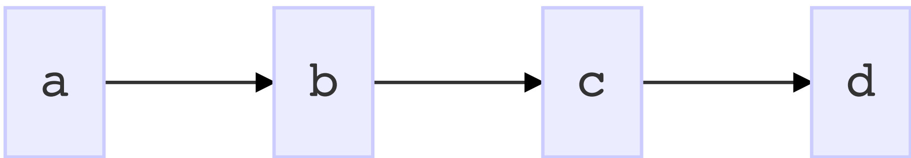
$$\hat{z}=\arg \max_{z} P(z|x) \propto \arg \max P(z, x)$$





Variable Elimination-乘法分配律

$$MAP \quad \tilde{X}_A = \arg\max_{X} P\left(x_A | x_B\right) = \arg\max P\left(x_A, x_B\right)$$



假设a,b,c,d均是离散的二值r,v {0,1}

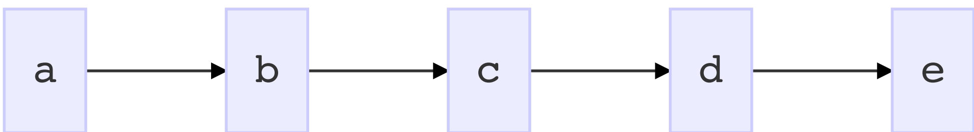
$$
\begin{gathered}
p(d) = \sum_{a,b,c} p(a,b,c,d) \\
= \sum_{a,b,c} p(a) \cdot p(b|a) \cdot p(c|b) \cdot p(d|c) \\
= p(a=0) \cdot p(b=0|a=0) \cdot p(c=0|b=0) \cdot p(d=0|c=0) \\
+ p(a=1) \cdot p(b=0|a=1) \cdot p(c=0|b=0) \cdot p(d=0|c=0) \\
+ \cdots \\
+ p(a=1) \cdot p(b=1|a=1) \cdot p(c=1|b=1) \cdot p(d=1|c=1) \\
= \sum_{b,c} p(c|b) \cdot p(d|c) \cdot \underbrace{\sum_{a} p(a) \cdot p(b|a)}_{\phi_a(b)} \\
= \sum_{c} p(d|c) \cdot \underbrace{\sum_{b} p(c|b) \cdot \phi_a(b)}_{\phi_b(c)} \\
= \phi_c(d)
\end{gathered}
$$

乘法对加法的分配律$ab + cb = b(a + c)$

Cons:

- Memoryless.重复计算
- Ordering NP-hard

Belief Propagation



Forward Algorithm

$$P(a,b,c,d,e) = P(a)P(b|a)\cdot P(c|b)\cdot P(d|c)\cdot P(e|d)$$

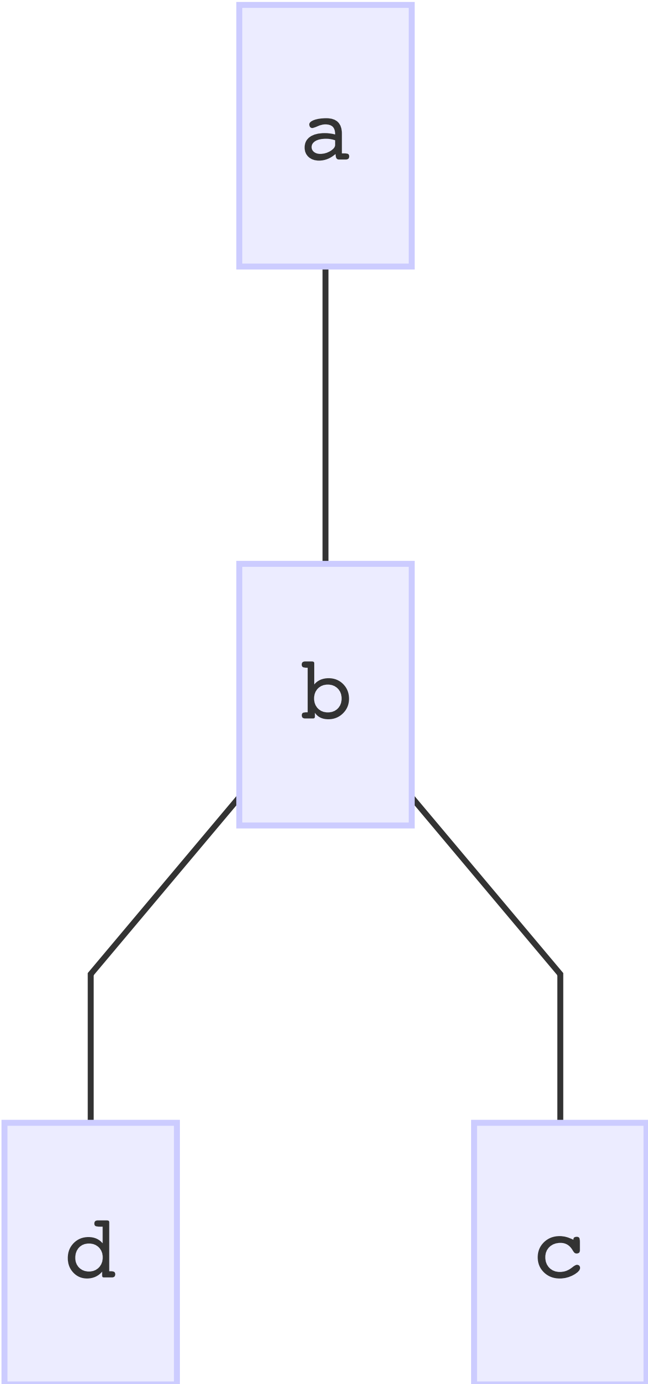
$$P(e) = \sum_{a,b,c,d} P(a,b,c,d,e)$$

$$= \sum_d p(e|d)\sum_c p(d|c)\underbrace{\sum_b p(c|b)\underbrace{\sum_a p(b|a)p(a)}_{m_{a\to b}(b)}}_{m_{b\to c}(c)}$$

$$p(c) = \sum_{a,b,d,e} p(a,b,c,d,e)$$

$$= \left(\sum_b p(c|b)\cdot\sum_a p(b|a)\cdot p(a)\right)\left(\sum_d p(d|c)\sum_e p(e|d)\right)$$

Forward-Backward Algorithm


a
b
d
c

$$
\begin{aligned}
& p(a,b,c,d) \\
=& \frac{1}{z}\psi_a(a)\psi_b(b)\cdot\psi_c(c)\cdot\varphi(d) \\
& \cdot\psi_{a,b}(a,b)\cdot\psi_{b,c}(b,c)\cdot\psi_{b,d}(b,d)
\end{aligned}
$$