Exercise 14.8

The Markov blanket of a variable is defined on the markov-blanket-page. Prove that a variable is independent of all other variables in the network, given its Markov blanket and derive the Equation (13.10).

The one remaining detail concerns the method of calculating the Markov blanket distribution $\mathbf{P}(X_i|mb(X_i))$, where $mb(X_i)$ denotes the values of the variables in X_i's Markov blanket, $MB(X_i)$. Fortunately, this does not involves any complex inference. As shown in Exercise 13.MARB, the distribution is given by

(13.10)

$$P(x_i|mb(X_i)) = \alpha\, P(x_i|parents(X_i)) \prod_{Y_j \in Children(X_i)} P(y_j|parents(Y_j)).$$

证明：

马尔可夫毯 mb(Xi)由三部分组成： Parents(Xi)：Xi 的父节点、

Children(Xi)：Xi 的子节点、

Spouse(Xi)：Xi 子结点的所有父结点

由 d-分离准则，只需证明 mb(Xi) d-分离了 {Xi} 和 除 Xi 和 mb(Xi)以外的其他结点即可：由于贝叶斯网络是有向无环图，这些其他结点，要么是 Xi 的祖先，要么是 Xi 的子孙，要么是 Xi 子孙的祖先，要么本来就与 Xi 不相通。对于前面三种情况，这些结点通往 Xi 的路径中，要么经过 Xi 的父节点，要么经过 Xi 的子节点，除此之外没有其他可行路径。而这些结点都被 mb(Xi)所覆盖，即 mb(Xi) d-分离了 其他结点和{Xi}。

综上，这一结构确保了马尔可夫毯覆盖了所有与 Xi 存在直接因果或依赖关系的变量，从而给定 mb(Xi)后可以使 Xi 相对于对网络中的其他变量独立

Exercise 14.10

Consider a simple Bayesian network with root variables *Cold*, *Flu*, and *Malaria* and child variable *Fever*, with a noisy-OR conditional distribution for *Fever* as described in Section 13.2.2. By adding appropriate auxiliary variables for inhibition events and fever-inducing events, construct an equivalent Bayesian network whose CPTs (except for root variables) are deterministic. Define the CPTs and prove equivalence.

（1）定义 CPT：（以下对随机变量的赋值中 1 表示 true，0 表示 false）

为了将噪声-或模型转换为确定性 CPT 的网络，需引入以下两种辅助变量：

抑制变量：I_{Cold}、I_{Flu}、$I_{Malaria}$ 分别表示感冒、流感、疟疾是否被抑制（即无法引发发烧）。

诱发变量：E_{Cold}、E_{Flu}、$E_{Malaria}$ 分别表示感冒、流感、疟疾是否实际导致发烧

定义上述变量的确定性 CPT：设 X∈{Cold,Flu,Malaria}

X	**$P(I_X$=True｜X)**
1	0
0	1
I_X	**$P(E_X$=True｜I_X)**
1	0
0	P_X （抑制关闭时，X 以 Px 的概率引起发烧）

则 Fever 的确定性 CPT 为：

P(Fever = 1| E_{Cold}, E_{Flu}, $E_{Malaria}$)=$\begin{cases}1 & 若\exists X \in \{\text{Cold, Flu, Malaria}\}，E_X = 1 \\ 0 & 其他情况\end{cases}$

（2）证明等价性：原噪声-或模型中，发烧的概率为：

P(Fever=1| Cold,Flu,Malaria)=1- $\prod_{X\in\{\text{Cold,Flu,Malaria}\}} (1-P_X)^X$

上标中的 X∈{0, 1}意为疾病 X 是否存在。

在（1）构造的模型中，发烧概率由诱发变量决定：

若 X=1，则 I_X = 0，那么 E_X= 1 的概率为 P_X

若 X= 0，则 I_X = 1，那么 E_X= 1 的概率为 0

因此 $P(E_X=1)=P_X\times X$，且由于各个 E_X 独立，则发烧概率为：

P(Fever = 1) = $1-\prod_{X\in\{\text{Cold,Flu,Malaria}\}} (1-E_X)$ =1- $\prod_{X\in\{\text{Cold,Flu,Malaria}\}} (1-P_X)^X$

与原模型完全一致。这表明新网络与原网络在发烧的条件概率分布上等价。

Exercise 14.17

Consider the Bayes net shown in Figure S13.12.

a. Which of the following are asserted by the network *structure*?

(i) $\mathbf{P}(B, I, M) = \mathbf{P}(B)\mathbf{P}(I)\mathbf{P}(M)$.

(ii) $\mathbf{P}(J \mid G) = \mathbf{P}(J \mid G, I)$.

(iii) $\mathbf{P}(M \mid G, B, I) = \mathbf{P}(M \mid G, B, I, J)$.

b. Calculate the value of $P(b, i, m, \neg g, j)$.

c. Calculate the value of $P(b, i, \neg m, g, j)$.

d. Calculate the probability that someone goes to jail given that they broke the law, have been indicted, and face a politically motivated prosecutor.

e. A **context-specific independence** (see page 420) allows a variable to be independent of some of its parents given certain values of others. In addition to the usual conditional independences given by the graph structure, what context-specific independences exist in the Bayes net in Figure S13.12?

f. Suppose we want to add the variable $P = PresidentialPardon$ to the network; draw the new network and briefly explain any links you add.

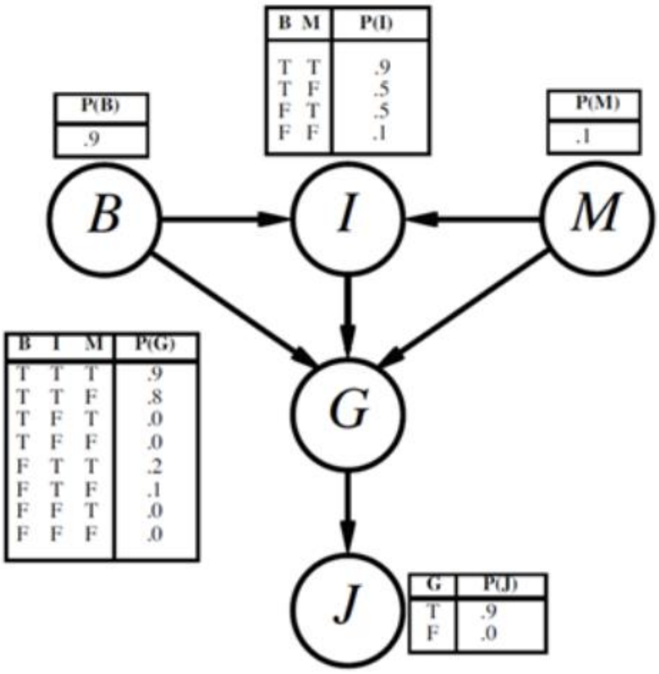


Figure S13.12 A Bayesian network for Exercise 13.BATF.

a. (ii),(iii)均可由该网络结构断言，(i)应该改为 P(B,I,M)=P(B) P(M) P(I | B=true,M=true)

b. P(b,i,m,¬g,j)=P(b)P(m)P(i|b,m) P(¬g |b,m,i) P(j |¬g)=0.9*0.1*0.9*(1-0.9)*0.0=0

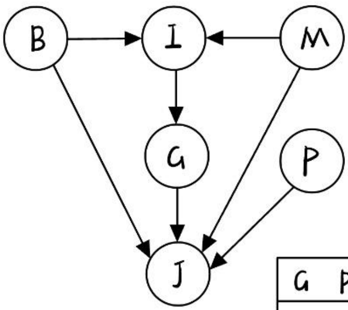
c. P(b,i,¬m,g,j)= P(b)P(¬m)P(i|b, ¬m) P(g |b,i, ¬m) P(j |g)

=0.9*(1-0.1)*0.5*0.8*0.9=0.2916

d. P(j | b,i,m)= $\sum_G \mathrm{P(j\,|\,G)P(G\,|b,m,i)}$ = $\mathrm{P(j\,|\neg g)P(\neg g|b,m,i)}$ + $\mathrm{P(j\,|\,g)P(g\,|b,m,i)}$ = 0.0*(1-0.9)+0.9*0.9=0.81

e. 当 I=False 时，G 与 B、M 无关（条件概率恒为 0），即及时违反了法律且面对有政治倾向的检察官，只要未被起诉就不会被抓捕。

f. 总统赦免可保证一定不入狱，即 P(j | p)=0。新的贝叶斯网络如下图：新节点 P 的概率分布未知，其他结点的 CPT 保持不变，故省略；改动后的 J 的 CPT 在图中画出。



G	P	P(J)
T	T	0
T	F	0.9
F	T	0
F	F	0

Exercise 14.21

Consider the query $\mathbf{P}(Rain \mid Sprinkler = true, WetGrass = true)$ in Figure 13.15(a) (page 435) and how Gibbs sampling can answer it.

a. How many states does the Markov chain have?

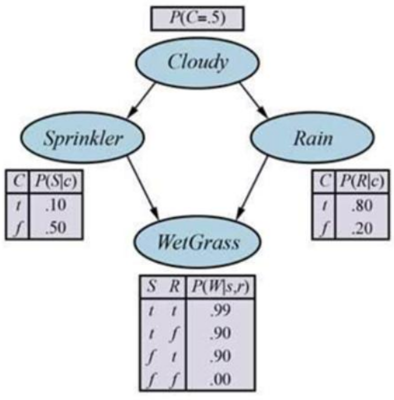
b. Calculate the **transition matrix Q** containing $k(\mathbf{y} \rightarrow \mathbf{y}')$ for all $\mathbf{y}$, $\mathbf{y}'$.

c. What does $\mathbf{Q}^2$, the square of the transition matrix, represent?

d. What about $\mathbf{Q}^n$ as $n \rightarrow \infty$?

e. Explain how to do probabilistic inference in Bayesian networks, assuming that $\mathbf{Q}^n$ is available. Is this a practical way to do inference?

Figure 13.15



(a)

a. 非证据变量为 Cloudy 和 Rain，因此状态空间的状态数为 $2^2=4$

b. 状态转移矩阵 Q 为：(给定 Sprinkler=true, WetGrass=true)

Q	Cloudy=true	Cloudy=false
Rain=true	0.5*0.8=0.4	0.5*0.2=0.1
Rain= false	0.6	0.9

c. Q^2 表示马尔可夫链在两步转移后的概率矩阵，即 $Q^2[i][j]$ 是从状态 i 经过两步转移到状态 j 的概率。

d. 当 $n\to\infty$ 时，Q^n 收敛到一个平稳分布矩阵，其中每一行均为平稳分布 π，即 π_i 是从任意初始状态经过无限步后最终到达状态 i 的概率，

e. 取一个足够大的 n，计算出 Q^n，直接将其作为 P(Rain | S=true, W=true)的分布。

可行性分析：在有多个非证据变量且只关注一个非证据变量的分布时，这样的方式得出的概率分布理论上是精确的。但实际上状态空间的数量会随变量数指数增长，即 Q 中的元素数量随变量数指数增长，计算 Q^n 所需的代价可能是庞大的。