Homework 1 (due: 9/27/2023)

1. For each of the following statements, either prove it is true or give a counterexample.

   (a) If $P(a|b, c) = P(b|a, c)$, then $P(a|c) = P(b|c)$

   (b) If $P(a|b, c) = P(a)$, then $P(b|c) = P(b)$

   (c) If $P(a|b) = P(a)$, then $P(a|b, c) = P(a|c)$

2. Given the full joint distribution shown below (Figure 13.3), calculate the following

   (a) $\mathbf{P}(\textit{toothache})$.

   (b) $\mathbf{P}(\textit{Catch})$.

   (c) $\mathbf{P}(\textit{Cavity}|\textit{catch})$.

   (d) $\mathbf{P}(\textit{Cavity}|\textit{toothache} \vee \textit{catch})$.

	toothache		*¬toothache*	
	catch	*¬catch*	*catch*	*¬catch*
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

Figure 13.3 A full joint distribution for the Toothache, Cavity, *Catch* world.

3. We wish to transmit an n-bit message to a receiving agent. The bits in the message are independently corrupted (flipped) during transmission with ϵ probability each. With an extra parity bit sent along with the original information, a message can be corrected by the receiver if at most one bit in the entire message (including the parity bit) has been corrupted. Suppose we want to ensure that the correct message is received with probability at least $1 - \delta$. What is the maximum feasible value of n? Calculate this value for the case $\epsilon = 0.002$, $\delta = 0.01$.

$$P(x_1, \cdots, x_n) = \prod_{i=1}^{n} \theta(x_i|\textit{parents}(X_i)) \tag{14.1}$$

$$P(a|b) = \frac{P(a \wedge b)}{P(b)} \tag{13.3}$$

$$\mathbf{P}(Y) = \sum_{z \in Z} P(Y, z) \tag{13.6}$$

4. Equation (13.1) on page 433 defines the joint distribution represented by a Bayesian network in terms of the parameter $\theta(X_i|\textit{Parents}(X_i))$. This exercise asks you to derive the equivalence between the parameters and the conditional probabilities $\mathbf{P}(X_i|\textit{Parents}(X_i))$ from this definition. (For part **c**, show the resulting expression reduces to $\theta(z|y)$ to be consistent with other parts of the problem.)

   (a) Consider a simple network $X \to Y \to Z$ with three Boolean variables. Use Equations (13.3) and (13.6) (pages 434 and 455) to express the conditional probability $P(z|y)$ as the ratio of two sums, each over entries in the joint distribution $\mathbf{P}(X, Y, Z)$.

(b) Now use Equation (13.1) to write this expression in terms of the network parameters $\theta(X)$, $\theta(Y|X)$, and $\theta(Z|Y)$.

(c) Next, expand out the summations in your expression from part (b), writing out explicitly the terms for the true and false values of each summed variable. Assuming that all network parameters satisfy the constraint $\sum_{x_i} \theta(x_i|\textit{Parents}(X_i)) = 1$, show that the resulting expression reduces to $\theta(z|y)$.

(d) Generalize this derivation to show that $\theta(X_i|\textit{Parents}(X_i)) = \mathbf{P}(X_i|Parents(X_i))$ for any Bayesian network.

5. Consider the Bayesian network in Figure 14.2.

 (a) If no evidence is observed, are *Burglary* and *Earthquake* independent? Prove this from the numerical semantics and from the topological semantics.

 (b) If we observe *Alarm=true*, are *Burglary* and *Earthquake* independent? Justify your answer by calculating whether the probabilities involved satisfy the definition of conditional independence.

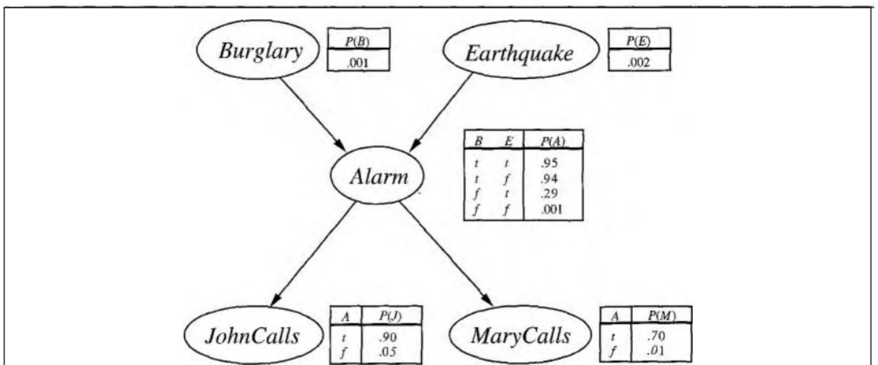

Figure 14.2 A typical Bayesian network, showing both the topology and the conditional probability tables (CPTs). In the CPTs, the letters B, E, A, J, and M stand for *Burglary, Earthquake, Alarm, JohnCalls*, and *MaryCalls*, respectively.

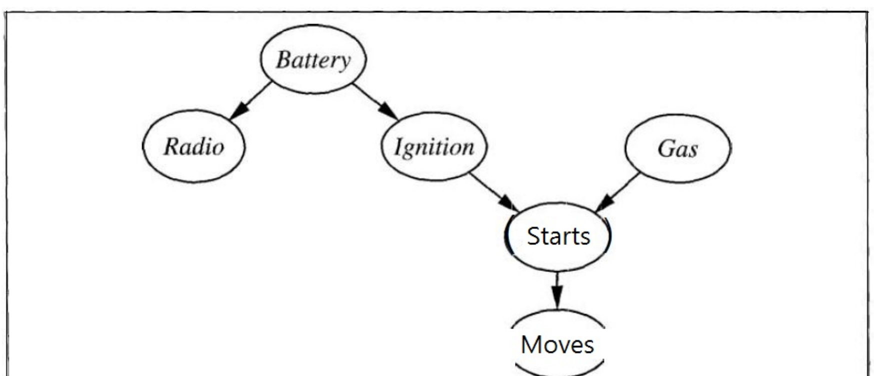

Figure 14.21 A Bayesian network describing some features of a car's electrical system and engine. Each variable is Boolean, and the true value indicates that the corresponding aspect of the vehicle is in working order.

6. Consider the network for car diagnosis shown in Figure 14.21.
   (a) Extend the network with the Boolean variables *IcyWeather* and *StarterMotor*.
   (b) Give reasonable conditional probability tables for all the nodes.
   (c) How many independent values are contained in the joint probability distribution for eight Boolean nodes, assuming that no conditional independence relations are known to hold among them?
   (d) How many independent probability values do your network tables contain?
   (e) The conditional distribution for *Starts* could be described as a **noisy-AND** distribution. Define this family in general and relate it to the noisy-OR distribution.

7. Consider the modified burglary network shown in Figure 1.
   (a) Implement the variable elimination algorithm for computing $P(B|j,m)$ using MATLAB. (Turn in a printout of your MATLAB code.)
   (b) Find $P(B|J,M)$ and $P(E|J,M)$ using the program implemented in (a). There are a total of 16 probabilities. (Turn in a printout of the outputs from your MATLAB code. You may use the command `diary` in MATLAB.)

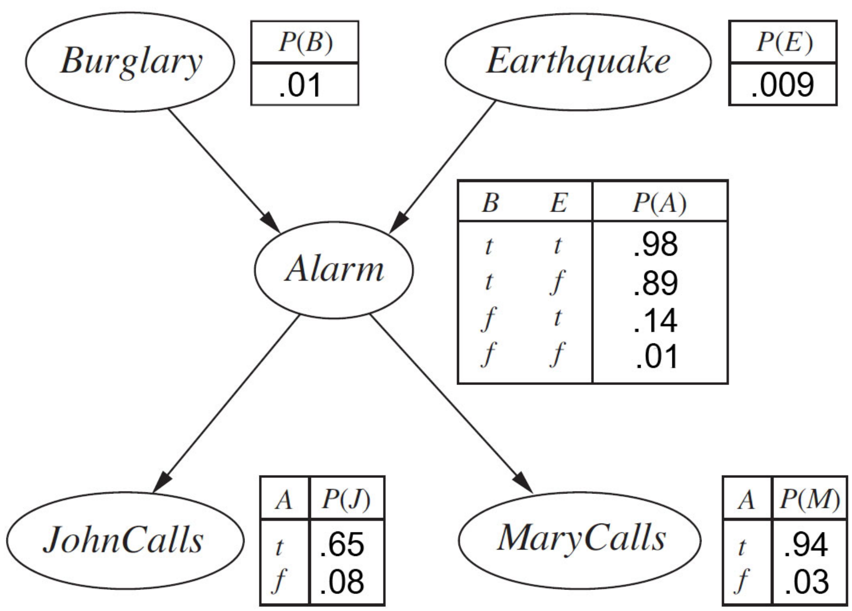

Figure 1: Modified burglary network.

8. Consider the modified weather network shown in Figure 2.
   (a) Is S and R independent? Is S and R conditionally independent given C? Prove your answers from the numerical semantics.
   (b) Find $P(S,R|W)$ exactly. Explain the relationship between S and R when the grass is wet. (Show all your work.)

(c) Implement the direct sampling algorithm to estimate $P(S, R|W)$ using MATLAB. You need to use an enough number of samples for good estimates for the exact values. How many samples does it need to achieve an accuracy of ± 0.01? (Turn in printouts of your MATLAB code and outputs.)

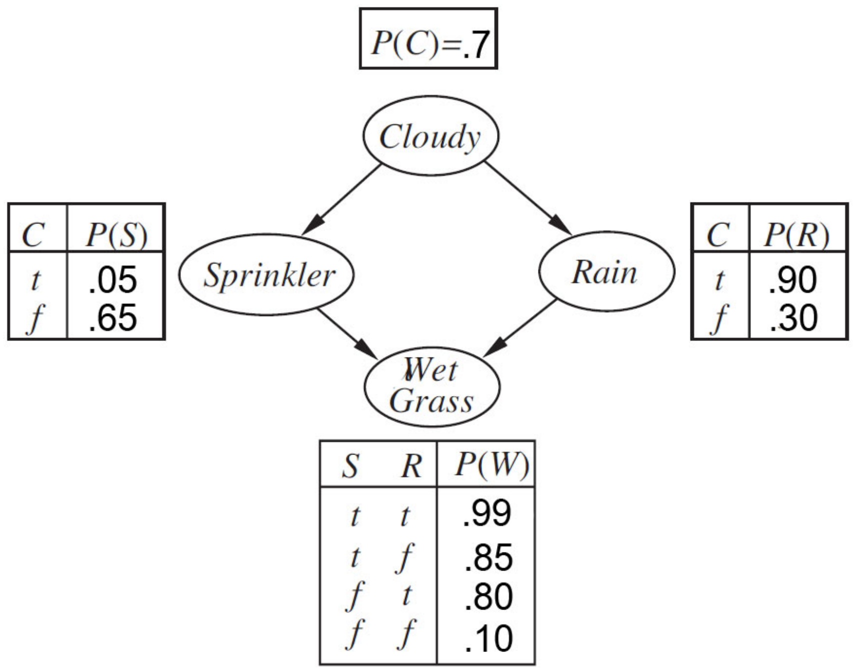

Figure 2: Modified weather network.