CS 440: Introduction to Artificial Intelligence

Homework #3 Part 1 Probabilistic Reasoning

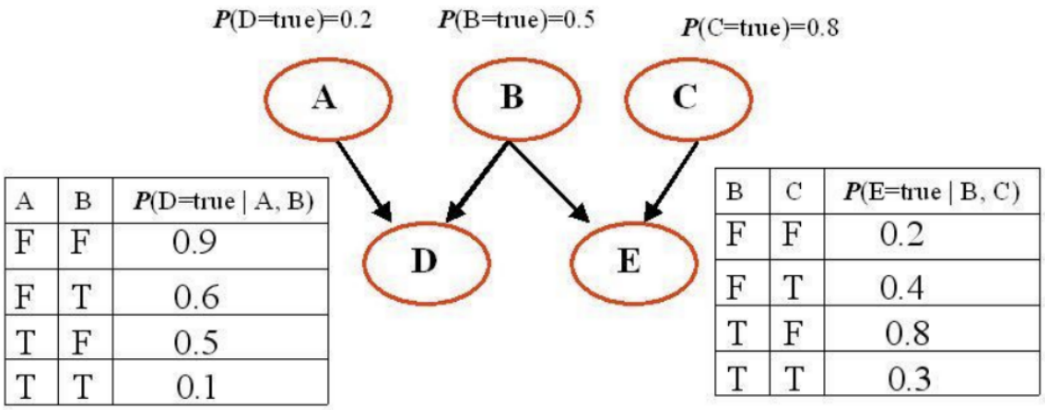
April 14, 2023

Professor Abdelsam Boularias

Jay Patwardhan 208001851
Alan Wu 208000574
Neel Shejwalkar 207004853

Problem 1

Consider the following Bayesian network, where variables A through E are all Boolean valued. Note: there is a typo in the image, it should be $P(A = true) = 0.2$ instead of $P(D = true) = 0.2$



A	B	P(D=true | A, B)
F	F	0.9
F	T	0.6
T	F	0.5
T	T	0.1

B	C	P(E=true | B, C)
F	F	0.2
F	T	0.4
T	F	0.8
T	T	0.3

Part A

What is the probability that all five of these Boolean variables are simultaneously true? [Hint: You have to compute the joint probability distribution. The structure of the Bayesian network suggests how the joint probability distribution is decomposed to the conditional probabilities available]

$P(A, B, C, D, E) = P(A) \cdot P(B) \cdot P(C) \cdot P(D|A = T, B = T) \cdot P(E|B = T, C = T)$
$= 0.2 \cdot 0.5 \cdot 0.8 \cdot 0.1 \cdot 0.3 = 0.0024$

Part B

What is the probability that all five of these Boolean variables are simultaneously false? [Hint: Answer similarly to above.]

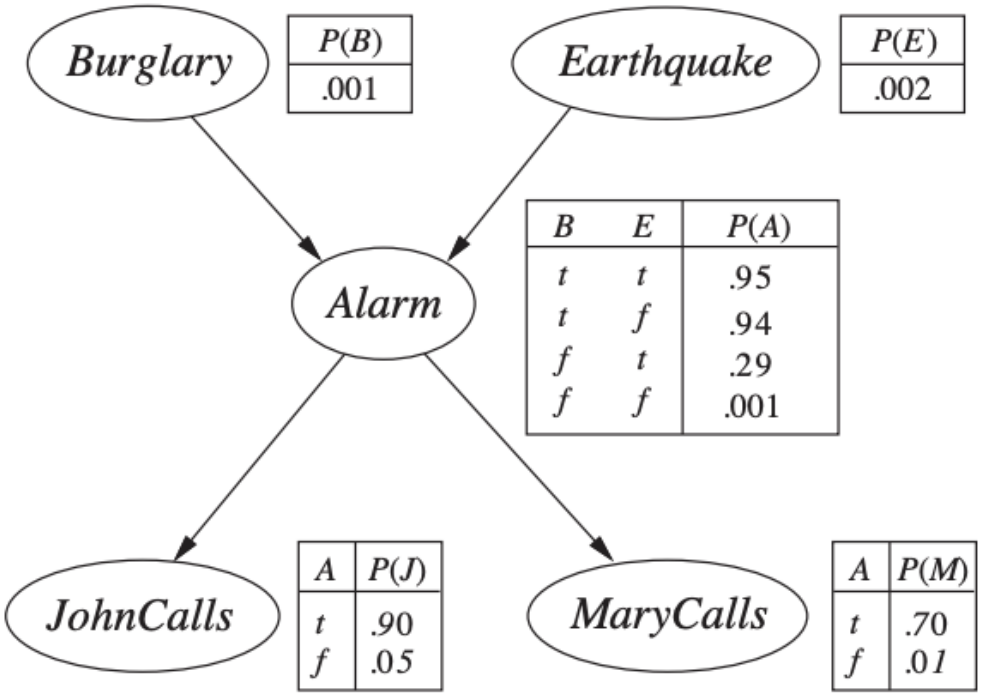
$P(\neg A, \neg B, \neg C, \neg D, \neg E) = P(\neg A) = P(\neg A) \cdot P(\neg B) \cdot P(\neg C) \cdot P(\neg D|A = F, B = F) \cdot P(\neg E|B = F, C = F)$
$= (1 - 0.2) \cdot (1 - 0.5) \cdot (1 - 0.8) \cdot (1 - 0.9) \cdot (1 - 0.2)$
$= 0.8 \cdot 0.5 \cdot 0.2 \cdot 0.1 \cdot 0.8 = 0.0064$

Part C

What is the probability that A is false given that the four other variables are all known to be true?

$P(\neg A|B, C, D, E) = \frac{P(\neg A, B, C, D, E)}{P(B, C, D, E)}$
$P(\neg A, B, C, D, E) = P(\neg A) \cdot P(B) \cdot P(C) \cdot P(D|A = F, B = T) \cdot P(E|B = T, C = T)$
$= (1 - 0.2) \cdot 0.5 \cdot 0.8 \cdot 0.6 \cdot 0.3 = 0.0576$
$P(B, C, D, E) = P(A, B, C, D, E) + P(\neg A, B, C, D, E) = 0.0024 + 0.0576 = 0.06$
$P(\neg A|B, C, D, E) = \frac{0.0576}{0.06} = 0.96$

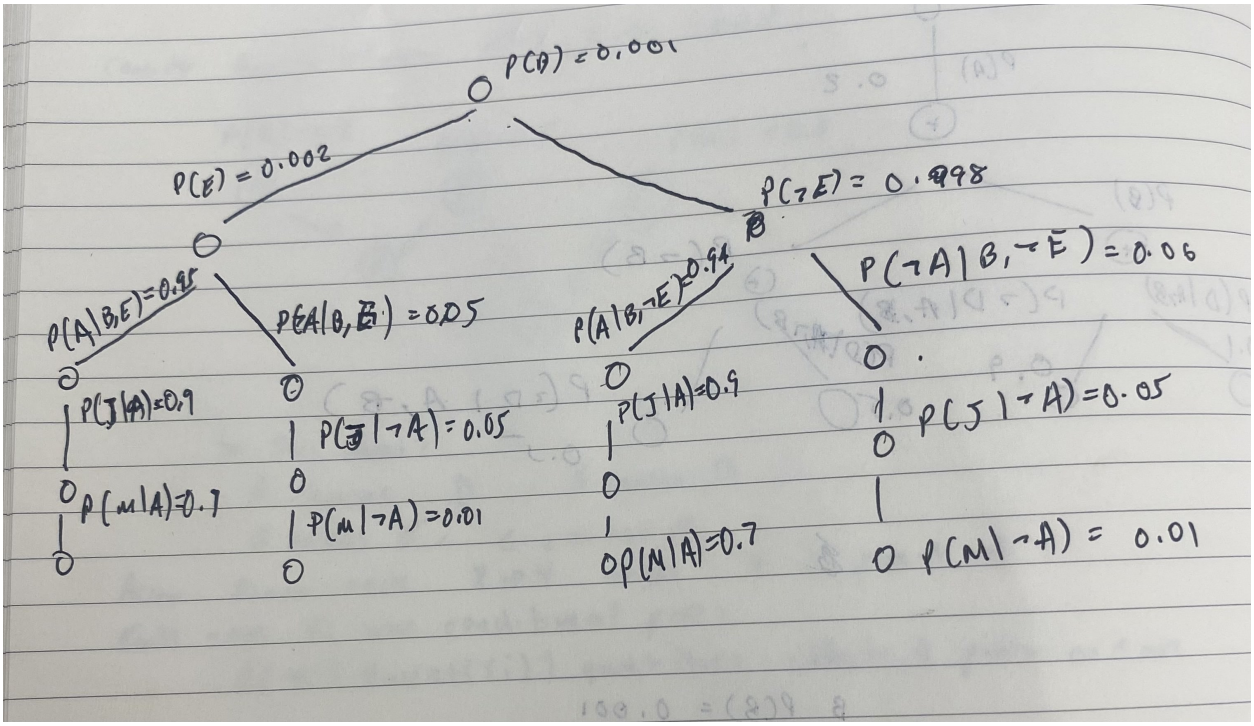
Problem 2



Part A

Calculate $P(Burglary|JohnsCalls = true, MaryCalls = true)$ and show in detail the calculations that take place. Use your book to confirm that your answer is correct.

The enumeration tree is shown:



$P(Burglary|JohnsCalls = true, MaryCalls = true) = P(B) \cdot \frac{P(J,M|B)}{P(J,M)}$
$= 0.001 \cdot \frac{(P(A|B)\cdot P(J,M|A))+(P(\neg A|B)\cdot P(J,M|\neg A))}{(P(J,M|A)\cdot P(A))+(P(J,M|\neg A)\cdot P(\neg A))}$

$P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{(P(B)\cdot P(E)\cdot P(A|B,E))+(P(B)\cdot P(\neg E)\cdot P(A|B,\neg E))}{P(B)}$
$= \frac{(0.001\cdot 0.002\cdot 0.95)+(0.001\cdot 0.998\cdot 0.94)}{0.001} = 0.94002$

$P(J, M|A) = P(J|A) \cdot P(M|A) = 0.9 \cdot 0.7 = 0.63$

$P(J, M|\neg A) = P(J|\neg A) \cdot P(M|\neg A) = 0.05 \cdot 0.01 = 0.0005$

$P(\neg A|B) = 1 - P(A|B) = 1 - 0.94002 = 0.05998$

$P(A) = (P(B) \cdot P(E)P(A|B, E)) + (P(B) \cdot P(\neg E)P(A|B, \neg E)) + (P(\neg B) \cdot P(E) \cdot P(A|\neg B, E)) + (P(\neg B) \cdot P(\neg E) \cdot P(A|\neg B, \neg E))$
$= (0.001 \cdot 0.002 \cdot 0.95) + (0.001 \cdot 0.998 \cdot 0.94) + (0.999 \cdot 0.002 \cdot 0.29) + (0.999 \cdot 0.998 \cdot 0.001)$
$= 0.0025$

$P(\neg A) = 1 - P(A) = 1 - 0.0025 = 0.99748$

Therefore:
$P(Burglary|JohnCalls = true, MaryCalls = true) = 0.001 \cdot \frac{(0.94002\cdot 0.63)+(0.05998\cdot 0.0005)}{(0.63\cdot 0.0025)+(0.0005\cdot 0.99748)}$
$= 0.2841718$

Part B

Suppose a Bayesian network has the from of a *chain* : a sequence of Boolean variables $X_1, ...X_n$ where $Parents(X_i) = X_{i-1}$ for $i = 2, ..., n$. What is the complexity of computing $P(X_1|Xn = true)$ using enumeration. What is the complexity with variable elimination?

The complexity for Enumeration:

Above is the binary tree since we only care about the true values for X_1 and X_N. We are looking for $P(X_1|X_n = true)$. Given the tree above, we know that the binary tree will have n levels that branch $n-2$ times. With the graph shown with 5 levels, it only branches 3 times. The above graph is missing the top level node. In this tree design there is a total of 2^{n-1} nodes. The space complexity will be $O(n)$ for dfs to traverse the tree and it will take $O(2^n)$ time complexity, with 2^{n-1} nodes absorbed into 2^n.

Now we will calculate the complexity for variable elimination:

Problem 3

Suppose you are working for a financial institution and you are asked to implement a fraud detection system. You plan to use the following information:

- When the card holder is travelling abroad, fraudulent transactions are more likely since tourists are prime targets for thieves. More precisely, 1of the transactions are fraudulent when she is not travelling. On average, 5card holder is travelling. If a transaction is fraudulent, then the likelihood of a foreign purchase increases, unless the card holder happens to be travelling. More precisely, when the card holder is not travelling, 10transactions are foreign purchases where as only 1other hand, when the card holder is travelling, then 90legitimacy of the transactions.
- Purchases made over the internet are more likely to be fraudulent. This is especially true for card holders who don't own any computer. Currently, 751transactions. For those who don't own any computer or smart phone, a mere 0.1is done over the internet, but that number increases to 1.1card company doesn't know whether a card holder owns a computer or smart phone, however it can usually guess by verifying whether any of the recent transactions involve the purchase of computer related accessories. In any given week, 10related item as opposed to just 0.1

Part A

Construct a Bayes Network to identify fraudulent transactions.

Solution.

Part B

What is the prior probability (i.e., before we search for previous computer related purchases and before we verify whether it is a foreign and/or an internet purchase) that the current transaction is a fraud? What is the probability that the current transaction is a fraud once we have verified that it is a foreign transaction, but not an internet purchase and that the card holder purchased computer related accessories in the past week?

Solution.