BITS Pilani

Pilani | Dubai | Goa | Hyderabad

PROBABILISTIC GRAPHICAL MODEL

SESSION # 8 : PROBLEMS

The instructor is gratefully acknowledging the authors who made their course materials freely available online.

Problem 1 Statement

We have 3 biased coins A, B, and C such that the probability of getting heads on a single toss of each of them are respectively 0.2, 0.6 and 0.8. One of the coins is selected uniformly at random and then flipped 3 times to get outcomes

Problem Statement

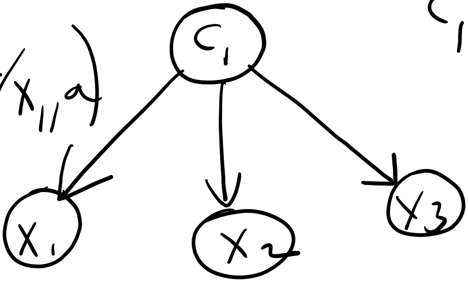
1) Draw a Bayesian network corresponding to this setup.

2) Which coin is most likely to have been drawn out of the bag if the observed values of X_1, X_2 and X_3 are Heads, Heads and Tails?

Problem 1 Solution

The Bayesian network is as below:

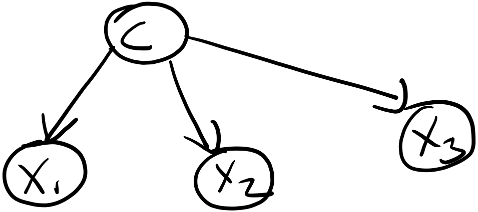
$P\left(\frac{X_2}{a}\right) = P(X_2 / X_1, a)$



$C_1 \rightarrow$ random variable representing which coin is selected

$X_1, X_2, X_3 \rightarrow$ flips on the selected coin

Problem 1 Solution



$X_1/a \rightarrow$
X_1/b
X_1/c

C	H	T
a	0.2	0.8
b	0.6	0.4
c	0.8	0.2

C	H	T
a	0.2	0.8
b	0.6	0.4
c	0.8	0.2

C	H	T
a	0.2	0.8
b	0.6	0.4
c	0.8	0.2

$P(X_1 = T/a)$ $\rightarrow P(H/a)$

$P(T/a)$

Problem | Solution

To solve the 2nd part of the question we calculate $P(HHT/a)$ $P(HHT/b)$ and $P(HHT/c)$ and take the coin that gives us the largest probability

$$P(HHT/a) = P(X_1 = H, X_2 = H, X_3 = T/a)$$

$$= P(X_1 = H/a)\,P(X_2 = H/a)\,P(X_3 = T/a)$$

$$= (0.2)^2(0.8) = 0.0032$$

innovate achieve lead

Problem 1 Solution

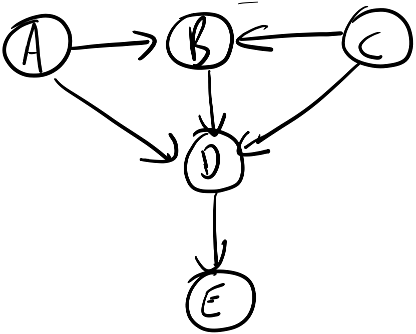
$$P(HHT/b) = (0.6)^2(0.4) = 0.144$$

$$P(HHT/c) = (0.8)^2(0.2) = 0.128$$

$\therefore$ b is the coin that was most likely to have produced HHT

Problem 2 Statement

Consider the Bayesian network below:



1) $P(A, B, C) \stackrel{?}{=} P(A)\,P(B)\,P(C)$

2) $P(E/D) \stackrel{?}{=} P(E/D, B)$

3) $P(C/A, B, D) \stackrel{?}{=} P(C/A, B, D, E)$

innovate achieve lead

Problem 2 Solution

(i) is false since there is an active path between A and B, and between B and C

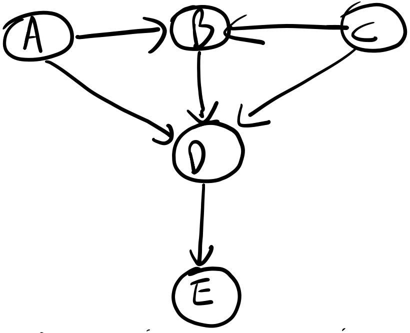
(ii) The paths between B and E becomes inactive when D is specified, so (ii) is true

(iii) C is conditionally independent of E given A, B, D since all paths between C and E become inactive

∴ (iii) is true

Problem 3 Statement

For the graph below, calculate $P(A=0, B=0, C=1, D=0, E=0)$



with CPDs given as in the next slide.

Problem 3 Statement

$P(A=0) = 0.9$, $P(C=0) = 0.1$

$P(B=0/A, C)$

A	C	P(B)
0	0	0.9
0	1	0.5
1	1	0.5
1	0	0.1

D	P(E)
0	0.9
1	0.0

$P(E=0/D)$

A	B	C	P(D)
0	0	0	0.9
0	0	1	0.8
0	1	0	0.0
0	1	1	0.0
1	0	0	0.2
1	0	1	0.1
1	1	0	0.0
1	1	1	0.0

$P(D=0/A, B, C)$

Problem 3 Statement

1) Calculate $P(A=0, B=0, C=1, D=0, E=0)$

2) Calculate $P(E=0 / A=0, B=0, C=0)$

Problem 3 Solution

① The joint probability can be expressed as a factorization

$$P(A=0)\,P\left(B=0 / A=0, C=1\right)P(C=1)\,P\left(D=0 / A=0, B=0, C=1\right)$$

$$P\left(E=0 / D=0\right)$$ ← where did we get this from?

$$= 0.9 \times 0.5 \times 0.9 \times 0.8 \times 0.9$$

$$= 0.2916$$

innovate achieve lead

Problem 3 Solution

$$P(E=0 / A=0, B=0, C=0) =$$

$$= \frac{\left[P(E=0, A=0, B=0, C=0, D=0) + P(E=0, A=0, B=0, C=0, D=1) \right]}{P(A=0, B=0, C=0)}$$

$$P(E / A, B, C) = \frac{P(E, A, B, C)}{P(A, B, C)}$$

Problem 3 Solution

$$P(E=0, A=0, B=0, C=0, D=0)$$

$$= P(E=0 / D=0)\, P(A=0)\, P(B=0 / A=0, C=0)\, P(C=0)\, P(D=0 / A=0, B=0, C=0)$$

$$P(E=0, A=0, B=0, C=0, D=1)$$

$$= P(E=0 / D=1)\, P(A=0)\, P(B=0 / A=0, C=0)\, P(C=0)\, P(D=1 / A=0, B=0, C=0)$$

Problem 3 Solution

$$P\left(E=0 \mid A=0, B=0, C=0\right) = \frac{P(A=0)\,P\left(\frac{B=0}{A=0,\,C=0}\right)P(C=0)\left[P\left(\frac{E=0}{D=1}\right)P\left(\frac{D=1}{A=0,\,B=0,\,C=0}\right) + P\left(\frac{E=0}{D=0}\right)P\left(\frac{D=0}{A=0,\,B=0,\,C=0}\right)\right]}{P(A=0,\,B=0,\,C=0)}$$

Problem 3 Solution

$$P(A=0, B=0, C=0) = P(A=0)\, P\left(\frac{B=0}{A=0, C=0}\right) P(C=0) \sum_{D,E} P\left(\frac{D}{A=0, B=0, C=1/0}\right) P\left(E/D\right)$$

After substitution we get

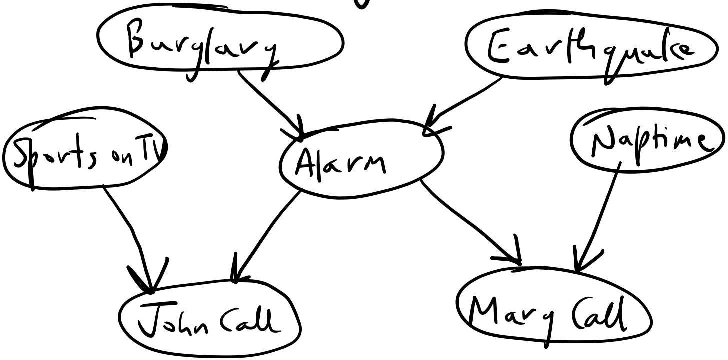
$$\frac{\sum_{D} P\left(D/A=0, B=0, C=0\right) P\left(\frac{E=0}{D}\right)}{\sum_{D'} P\left(D'/A=0, B=0, C=0\right) \sum_{E} P\left(E'/D'\right)}$$

Problem 3 Solution

$$= \frac{0.9 \times 0.9 + 0.1 \times 0}{0.9 \times 0.1 + 0.1 \times 0 + 0.9 \times 0.9 + 0.1 \times 1} \approx 0.81$$

Problem 4 Statement

Consider the following alarm network shown below:



Problem 4 Statement

Construct a Bayesian network structure over the nodes Burglary, Earthquake, Sports on TV, Naptime, John Call, Mary Call which is a minimal I-map for the marginal distribution over those variables defined by the network on the previous slide. Be sure to get all dependencies that remain from the original network.

Problem 4 Solution

The key idea is to regard Alarm as unobserved

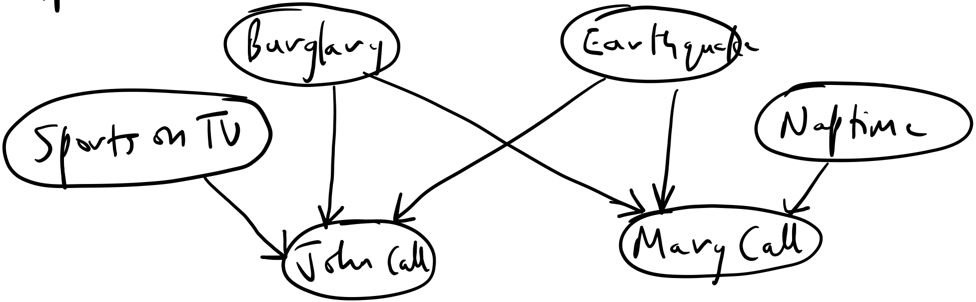
Since we marginalize over Alarm.

We must preserve all the old independencies and add new edges to account for the fact that Alarm is removed

If Alarm is unobserved, there exist active paths between Alarm's parents and Alarm's children

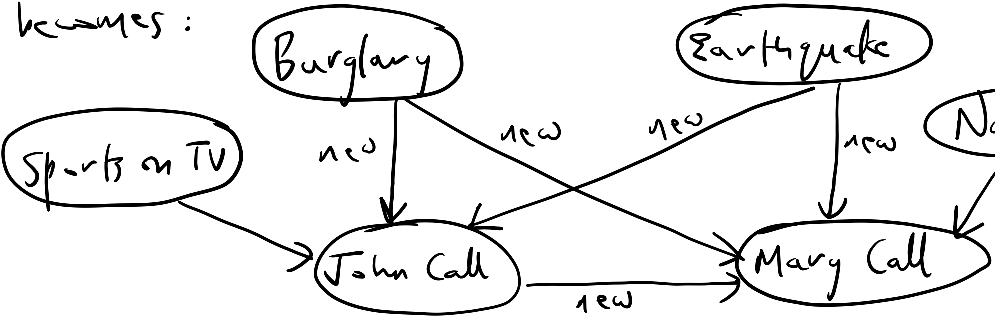
Problem 4 Solution

Thus as a first step our graph needs to look like this:



Problem & Solution

Now we observe that any two children of Alarm also have an active path between them, so the graph becomes :



Problem 4 Solution

Now it appears that there is an active path between Sports on TV and Mary Call through John Call such that

$$\text{Sports on TV} \perp \text{Mary Call} \mid \text{John Call}$$

However in the original graph we had

Sports on TV $\rightarrow$ John Call $\leftarrow$ Alarm $\rightarrow$ Mary Call

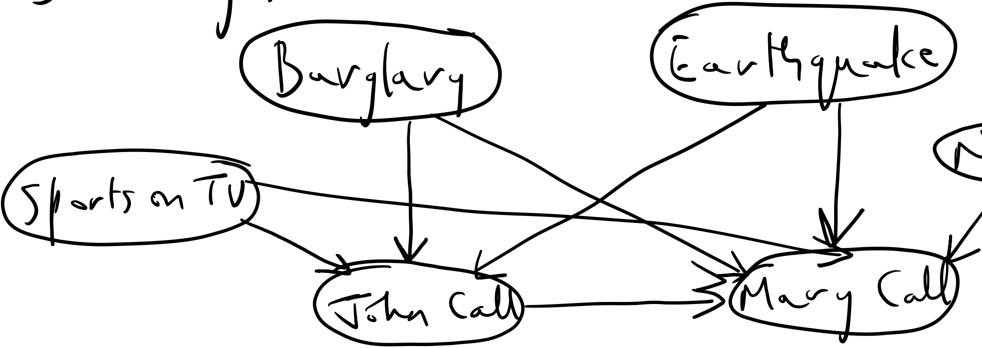
Problem & Solution

John Call is a converging node, specifying it leads to an active path from Sports on TV to Mary Call, which is the opposite of what we have in our new graph

We compensate for this by adding a direct edge between Sports on TV and Mary Call

Problem 4 Solution

So we get



Problem & Solution

Any more edges? The original graph looked Symmetric, should we not add an edge between <u>Naptime</u> and <u>John Call</u> to preserve symmetry?

In the <u>old graph</u>:

Naptime → MaryCall ← Alarm → John Call -

(MaryCall converging)

Problem 4 Solution

Now, we have

$$\text{Naptime} \rightarrow \text{MaryCall} \leftarrow \text{JohnCall}$$

Naptime and JohnCall related in the same way as before (since Alarm is unobserved)

So no edge needed between Naptime and John Call

Problem 5

This problem will construct a practical example of a non-positive distribution where local independencies do not imply global ones.

Let P be any distribution over $\mathcal{X} = \{X_1, X_2, \ldots X_n\}$ and let $\mathcal{X}' = \{X_1', X_2' \ldots X_n'\}$

We construct a distribution $P'(\mathcal{X}, \mathcal{X}')$ whose marginal over $X_1, X_2, \ldots X_n$ is the same as P and where each X_i is deterministically equal to X_i'.

Problem 5

Let H be a Markov network containing no edges other than $X_i - X_j$. Question: Is H an I-map for P?

By which we mean, does every global independency asserted by H also exist in P?

innovate achieve lead

Problem 5 Solution

X_1 —— X_1'

X_2 —— X_2'

$\vdots$

$\Big\}\ H$

In graph H we see local independencies of the form $(X_i \perp \text{all variables other than } X_i \,\&\, X_i' \mid X_i')$

These local independencies are also satisfied in P' since once X_i' is known, X_i has to be equal to X_i' and has no dependence on any other variable.

Problem 5 Solution

What about global independencies asserted by H? We see that X_i is D-separated from every other node X_j given the empty set, so X_i should be independent of X_j.

But do these independencies hold in P'?

No because P was any distribution on $X_1, X_2, \ldots X_n$ and may not support these independencies.

Problem 6

We shall show now that for a non-positive distribution, the pairwise independencies do not imply local independencies.

Let P be any distribution over $\mathcal{X} = \{X_1, X_2, \dots X_n\}$ and consider two auxiliary sets of variables $\mathcal{X}'$ and $\mathcal{X}''$ and define $\mathcal{X}^* = \mathcal{X} \cup \mathcal{X}' \cup \mathcal{X}''$

We now construct a distribution P' over $\mathcal{X}^*$ such that its marginal over $X_1, X_2, \dots X_n$ is the same as that of P, and such that $X_i = X_i' = X_i''$ deterministically.

Problem 6

Let H be the empty Markov network over $\mathcal{X}^*$. Does H satisfy pairwise independence assertions in P'?

Are all local independencies asserted by H also found in P'?

Problem 6 Solution

We note that $X_i \perp X_i' \mid \mathcal{X}^* - \{X_i, X_i'\}$ since $\mathcal{X}^* - \{X_i, X_i'\}$ contains X_i'' and $X_i = X_i'' = X_i'$ so

$$P'(X_i / X_i'', X_i') = P'(X_i / X_i'')$$

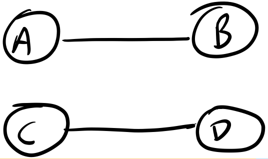
Similarly X_i and X_j are independent given all other nodes

Thus H satisfies all pairwise independencies but not local or global independencies

Problem 7

Let a, b, c, d be binary variables taking the values 0 and 1. Let $P(a, b, c, d)$ be the joint distribution such that $P(a^0, b^0, c^0, d^0) = 0.5$ and $P(a^1, b^1, c^1, d^1) = 0.5$

Can the Markov network constructed according to local independencies be the following graph?



Problem 7 Solution

We see that if we know the value of B (either 0 or 1), the value of A is known → we do not care about the value of C or D.

$\therefore P(A/B,C,D) = P(A/B)$. Thus $(A \perp C,D \mid B)$

In the given graph A is H-separated from C and D once B is specified → does this mean that the given graph is a I-map?

Problem 7 Solution

We see similarly $B \perp C, D \mid A$, $C \perp A, B \mid D$ and so on. All these independencies are supported by the distribution and the graph

However the graph is not an I-map for P because it asserts false independencies like $A \perp C$ given the empty set