School of Computing
National University of Singapore
CS5340: Uncertainty Modeling in AI
Semester 1, AY 2018/19

Exercise 1

Question 1

a)

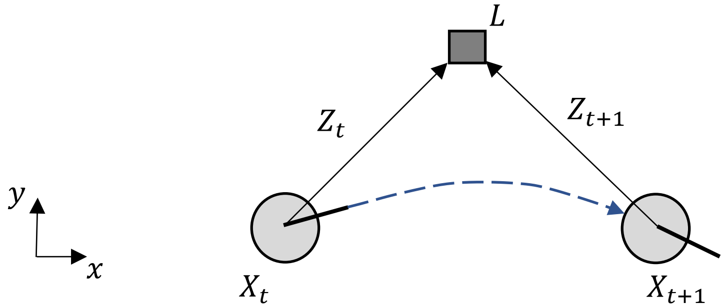


Fig. 1.1

Fig. 1.1 shows a mobile robot that traverses from pose X_t to X_{t+1} over time t to $t+1$. The robot is equipped with an 1-dimensional range sensor that returns the distances Z_t and Z_{t+1} of a landmark structure L in the environment from the poses X_t and X_{t+1} respectively. Let U_t denotes the control command given by the user to move the robot from X_t to X_{t+1}.

(i) Taking $\{U_t, L, X_t, X_{t+1}, Z_t, Z_{t+1}\}$ as random variables, state whether each of these random variables is an observed or latent/hidden random variable. Explain your answers.

Answer:

observed variables: $\{u_t, z_t, z_{t+1}\}$ (Inputs from user and observations from sensor);
latent variables: $\{x_t, x_{t+1}, l\}$

(ii) Given the following conditional independencies:

$$L \perp U_t \mid \emptyset, \quad X_t \perp L \mid U_t, \quad X_{t+1} \perp \{L, U_t\} \mid X_t,$$

$$Z_t \perp \{U_t, X_{t+1}\} \mid \{X_t, L\}, \quad Z_{t+1} \perp \{U_t, X_t, Z_t\} \mid \{L, X_{t+1}\}.$$

Write the factorized probability and draw the Bayesian network that represents the joint distribution $p(u_t, l, x_t, x_{t+1}, z_t, z_{t+1})$ assuming the following topological ordering of the random variables:

$$\{U_t, L, X_t, X_{t+1}, Z_t, Z_{t+1}\}.$$

Show all your workings clearly.

Answer:

From chain rule:

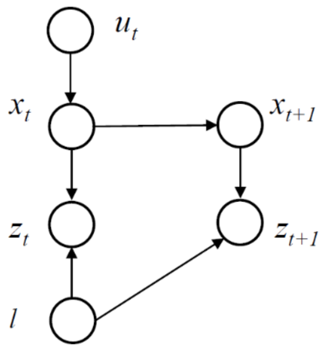
$$p(u_t, l, x_t, x_{t+1}, z_t, z_{t+1}) = p(u_t)p(l|u_t)p(x_t|u_t, l)p(x_{t+1}|u_t, l, x_t)$$
$$p(z_t|u_t, l, x_t, x_{t+1})p(z_{t+1}|u_t, l, x_t, x_{t+1}, z_t)$$

$$p(l|u_t) = p(l) \quad since \quad l \perp u_t|\emptyset$$
$$p(x_t|u_t, l) = p(x_t|u_t) \quad since \quad x_t \perp l|u_t$$
$$p(x_{t+1}|u_t, l, x_t) = p(x_{t+1}|x_t) \quad since \quad x_{t+1} \perp \{l, u_t\}|x_t$$

$$p(z_t|u_t, l, x_t, x_{t+1}) = p(z_t|x_t, l) \quad since \quad z_t \perp \{u_t, x_{t+1}\}|\{x_t, l\}$$
$$p(z_{t+1}|u_t, l, x_t, x_{t+1}, z_t) = p(z_{t+1}|l, x_{t+1}) \quad since \quad z_{t+1} \perp \{u_t, x_t, z_t\}|\{l, x_{t+1}\}$$

$$\Rightarrow p(u_t, l, x_t, x_{t+1}, z_t, z_{t+1}) = p(u_t)p(l)p(x_t|u_t)p(x_{t+1}|x_t)p(z_t|x_t, l)p(z_{t+1}|l, x_{t+1})$$

Graphic model:



(iii) Write the following probability distribution $p(z_t, z_{t+1} \mid l)$ in terms of the factorized probability obtained in (ii). Simplify your answer.

Answer:

$$p(z_t, z_{t+1}|l) = \frac{p(z_t, z_{t+1}, l)}{p(l)} = \frac{p(z_t, z_{t+1}, l)}{\sum_{z_t}\sum_{z_{t+1}} p(z_t, z_{t+1}, l)},$$

where

$$p(z_t, z_{t+1}, l) = \sum_{u_t}\sum_{x_t}\sum_{x_{t+1}} p(u_t)p(l)p(x_t|u_t)p(x_{t+1}|x_t)p(z_t|x_t, l)p(z_{t+1}|x_{t+1}, l),$$

$$= p(l)\sum_{u_t}\sum_{x_t}\sum_{x_{t+1}} p(\cancel{u_t})\frac{p(x_t,u_t)}{p(\cancel{u_t})}p(x_{t+1}|x_t)p(z_t|x_t,l)p(z_{t+1}|x_{t+1},l)$$

$$= p(l)\sum_{x_t}\sum_{x_{t+1}} p(x_{t+1}|x_t)p(z_t|x_t,l)p(z_{t+1}|x_{t+1},l)\sum_{u_t} p(x_t,u_t)$$

$$= p(l)\sum_{x_t}\sum_{x_{t+1}} p(x_{t+1}|x_t)p(z_t|x_t,l)p(z_{t+1}|x_{t+1},l)p(x_t)$$

and

$$p(l) = \sum_{u_t}\sum_{x_t}\sum_{x_{t+1}} p(u_t)p(l)p(x_t|u_t)p(x_{t+1}|x_t)\sum_{z_t} p(z_t|x_t, l)\sum_{z_{t+1}} p(z_t|x_{t+1}, l),$$

$$since \quad \sum_{z_t} p(z_t|x_t, l) = 1 \quad and \quad \sum_{z_{t+1}} p(z_t|x_{t+1}, l) = 1,$$

$$p(l) = \sum_{u_t}\sum_{x_t}\sum_{x_{t+1}} p(u_t)p(l)p(x_t|u_t)p(x_{t+1}|x_t)$$

$$= p(l)\sum_{u_t}\sum_{x_t}\sum_{x_{t+1}} p(\cancel{u_t})\frac{p(x_t,u_t)}{p(\cancel{u_t})}\frac{p(x_{t+1},x_t)}{p(x_t)}$$

$$= p(l)\sum_{x_t}\sum_{x_{t+1}}\frac{p(x_{t+1},x_t)}{p(x_t)}\sum_{u_t} p(x_t,u_t)$$

$$= p(l)\sum_{x_t}\sum_{x_{t+1}}\frac{p(x_{t+1},x_t)}{p(\cancel{x_t})}p(\cancel{x_t}) = p(l)$$

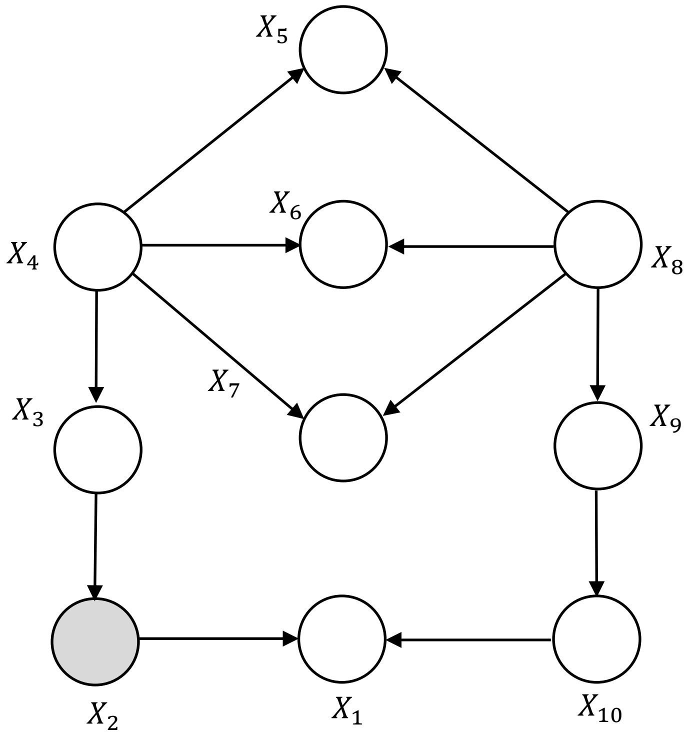
We get:

$$p(z_t, z_{t+1}|l) = \frac{p(z_t, z_{t+1}, l)}{p(l)} = \frac{p(z_t, z_{t+1}, l)}{\sum_{z_t}\sum_{z_{t+1}} p(z_t, z_{t+1}, l)}$$

$$= \frac{\cancel{p(l)}\sum_{x_t}\sum_{x_{t+1}} p(x_{t+1}|x_t)p(z_t|x_t,l)p(z_{t+1}|x_{t+1},l)p(x_t)}{\cancel{p(l)}}$$

$$= \sum_{x_t}\sum_{x_{t+1}} p(x_{t+1}|x_t)p(z_t|x_t,l)p(z_{t+1}|x_{t+1},l)p(x_t)$$

b) For each of the Bayesian networks shown in Fig. 1.2, determine the largest set of nodes X_B such that $X_1 \perp X_B \mid X_2$. Explain your answers.



Answer:

(i) $x_1 \perp \{x_3, x_4\} | x_2$

Bayes ball can move to $x_{10}, x_9, x_8, x_7, x_6, x_5$, but blocked from $\{x_4, x_5\}$ because of head-to-head structures in $\{x_5, x_6, x_7\}$ (d-separation).

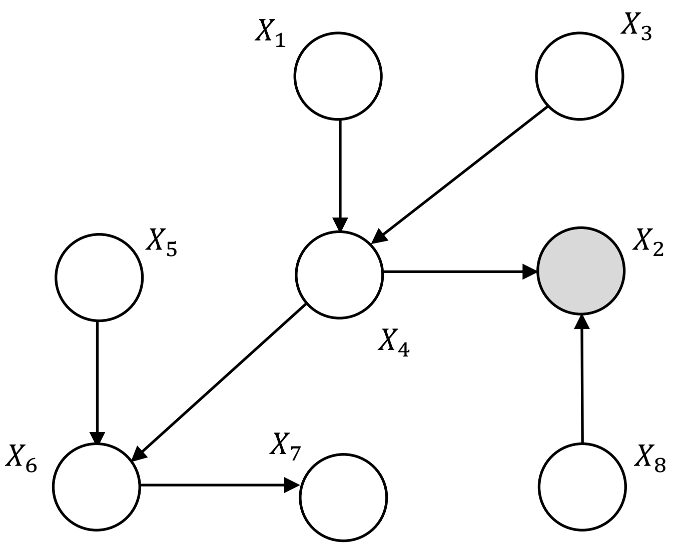


Fig. 1.2

Answer:

(ii) $x_1 \perp x_5 | x_2$

This is because of the head-to-head structure at $\{x_5 \rightarrow x_6 \leftarrow x_4\}$. Since x_6 is not observed, path is blocked from x_1 to x_5. Since x_2 is observed, path from x_1 to x_3 is not blocked because x_2 is a descendant of x_4. Similarly, the path from x_1 to x_8 is not blocked since x_2 is observed.

Question 2

Consider the graph shown in Fig. 2.1:

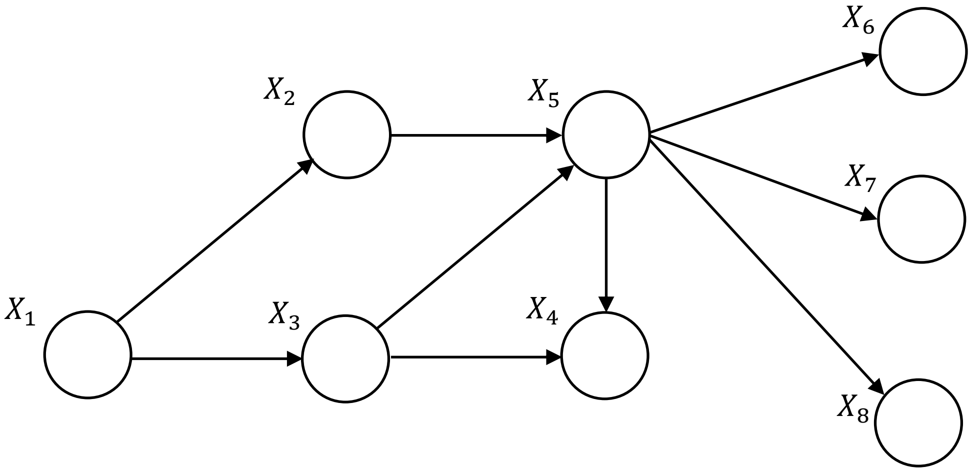
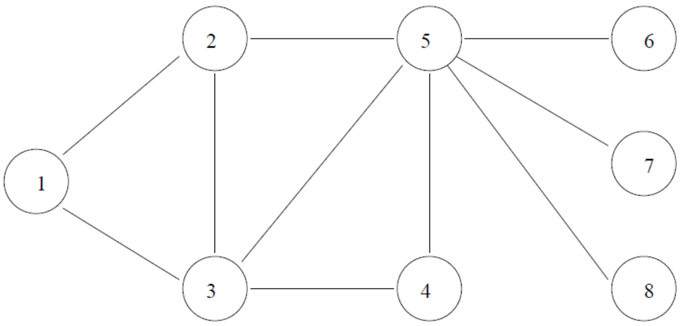


Fig 2.1

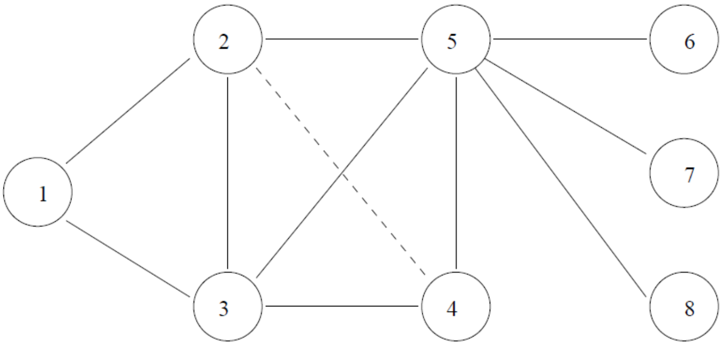
a) What is the corresponding moral graph?

Answer:



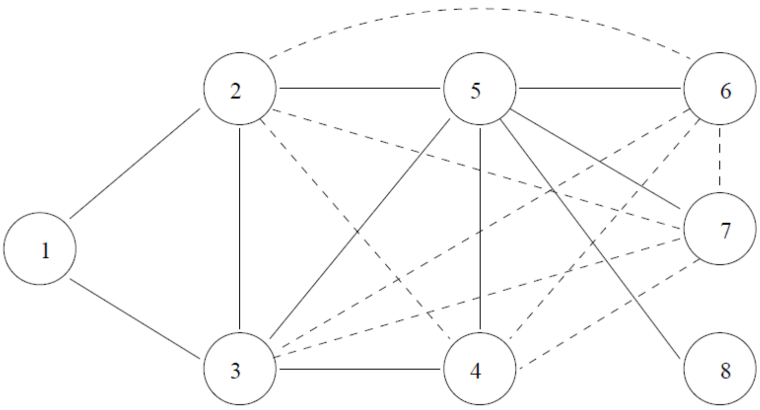
b) What is the reconstituted graph from the UNDIRECTEDGRAPHELIMINATE algorithm on the moral graph with the ordering {8,7,6,5,4,3,2,1}?

Answer:



c) What is the reconstituted graph from the UNDIRECTEDGRAPHELIMINATE algorithm on the moral graph with the ordering {8,5,6,7,4,3,2,1}?

Answer:



d) Suppose you wish to calculate $p(x_1|x_8)$. Which ordering is preferable? Why?

Answer:

The ordering in (b) is preferable because the maximum clique is of size 4 whereas the ordering in (c) results in a maximum clique of size 6.

<u>Question 3</u>

What is the treewidth of the graph below?

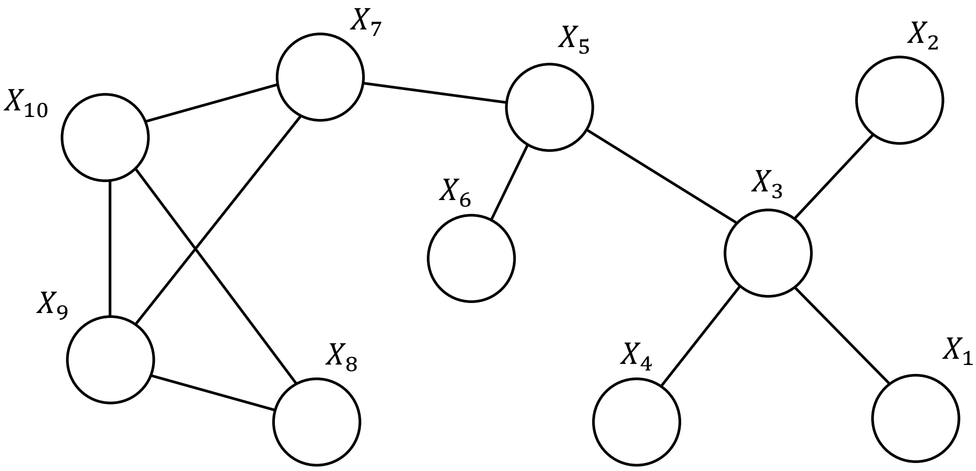


Fig 3.1

Answer:

- Elimination process adds new edges between (remaining) neighbors of the node, and this creates new “elimination cliques” in the graph.

- Treewidth: one less than the smallest achievable cardinality of the largest elimination clique over all possible elimination orderings.

We choose the elimination order that introduces the least “elimination cliques” into the graph, i.e. $I = \{X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}\}$.

I introduces no new “elimination cliques” into the graph. We can see that the maximum clique is 3. Hence, the treewidth is 3-1 = 2.

Question 4

Consider the following random variables. X_1 and X_2 represent the outcomes of two independent fair coin tosses. X_3 is the indicator function of the event that the outcomes are identical.

a) Specify a directed graphical model that describes the joint probability distribution (i.e. specify the graph and the conditional distributions).

Answer:

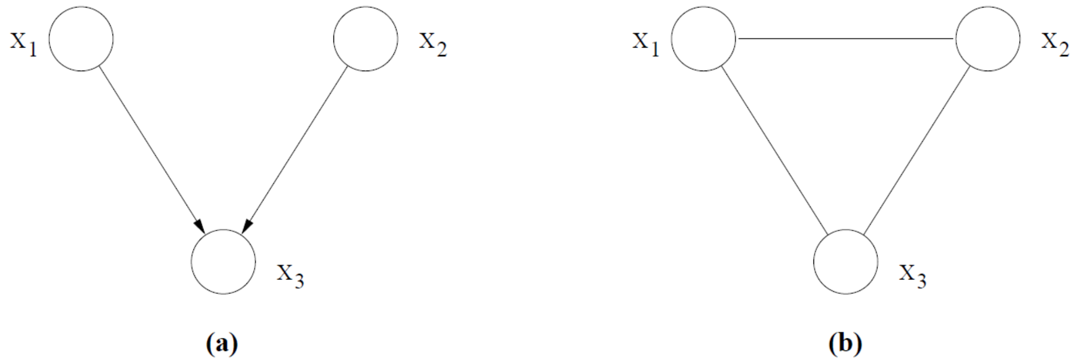


Figure 1: (a) The directed graphical model representing X_1, X_2, X_3 (b) An undirected graphical model representation for the same problem.

See figure (a) for the model. Let 1 denote heads and 0 denote tails. Then

$$P(X_1, X_2, X_3) = P(X_1)P(X_2)P(X_3|X_1, X_2).$$

where

$$P(X_1 = 1) = P(X_1 = 0) = \frac{1}{2},$$
$$P(X_2 = 1) = P(X_2 = 0) = \frac{1}{2},$$

$$P(X_3 = 0|X_1 = 0, X_2 = 0) = 0,$$
$$P(X_3 = 0|X_1 = 0, X_2 = 1) = 1,$$
$$P(X_3 = 0|X_1 = 1, X_2 = 0) = 1,$$
$$P(X_3 = 0|X_1 = 1, X_2 = 1) = 0,$$
$$P(X_3 = 1|X_1 = 0, X_2 = 0) = 1,$$
$$P(X_3 = 1|X_1 = 0, X_2 = 1) = 0,$$
$$P(X_3 = 1|X_1 = 1, X_2 = 0) = 0,$$
$$P(X_3 = 1|X_1 = 1, X_2 = 1) = 1,$$

b) Specify an undirected graphical model that describes the joint probability distribution (i.e. give the graph and specify the clique potentials).

Answer:

See figure (b) for the model. From part (a), there is a factor that includes all three

variables, hence there must be a corresponding clique with all three variables X_1, X_2 and X_3. Since there is only one clique, we have

$$P(X_1, X_2, X_3) = \frac{1}{Z}\Phi_1(X_1, X_2, X_3)$$

where $\Phi_1(X_1, X_2, X_3) = P(X_1, X_2, X_3) = P(X_1)P(X_2)P(X_3|X_1, X_2)$ and $Z = 1$.

c) In both cases, list all conditional independencies that are implied by the graph.

Answer:

From part (a), we can only state $X_1 \perp X_2|\varnothing$. For part (b), there is no conditional independence statements that is implied by the graph.

d) In both cases, list any additional conditional dependencies that are displayed by this probability distribution but are not implied by the graph.

Answer:

Other than $X_1 \perp X_2|\varnothing$, we also have $X_1 \perp X_3|\varnothing$ and $X_2 \perp X_3|\varnothing$. To see this, note that $P(X_1 = 0) = P(X_1 = 1) = P(X_2 = 0) = P(X_2 = 1) = P(X_3 = 0) = P(X_3 = 1) = 1/2$. We also have $P(x_1, x_2) = 1/4 = P(x_1)P(x_2)$ for all values of x_1, x_2. Similarly $P(x_2, x_3) = 1/4 = P(x_2)P(x_3)$ for all values of x_2, x_3.
This is a case where all three random variables are pairwise independent but are not mutually independent (i.e. $P(x_1, x_2, x_3) \neq P(x_1)P(x_2)P(x_3)$). Note also, that two of the conditional independence are numerical instead of structural and are easily destroyed by slight change in the parameter values i.e. if the coins are slightly unfair.

Question 5

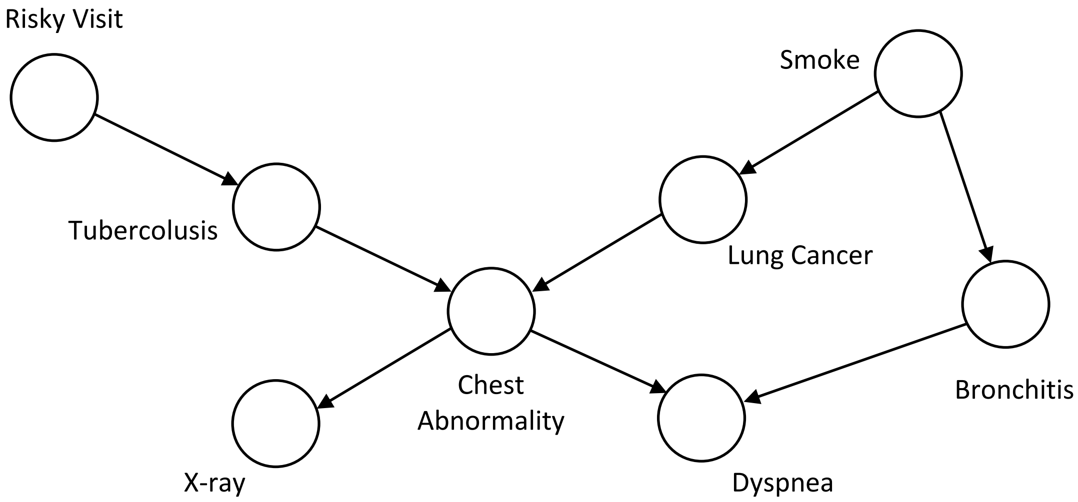


Fig. 5.1

The graphical model shown above describes some relationships among variables associated with chest abnormality. Answer the following questions based on the graphical model.

a) True or False. Justify your choice. $Smoke \perp Dyspnea \mid Bronchitis$.

Answer:
False. Although *Bronchitis* blocks one path from *Smoke* to *Dyspnea*, there is another path through *Lung Cancer* and *Chest Abnormality*.

b) True or False. Justify your choice. $Bronchitis \perp X\text{-}ray \mid Cancer$.

Answer:
True. One path is blocked by *Cancer* while another path is blocked by *Dyspnea*.

c) True or False. Justify your choice. $Smoke \perp Risky\,Visit \mid Dyspnea$.

Answer:
False. *Dyspnea* is a descendent of *Chest Abnormality* hence the path through *Tubercolusis, Chest Abnormality, Cancer* is no longer blocked by *Chest Abnormality*.

d) True of False. Justify your choice. $X\text{-}ray \perp Smoke \mid \{Cancer, Bronchitis\}$.

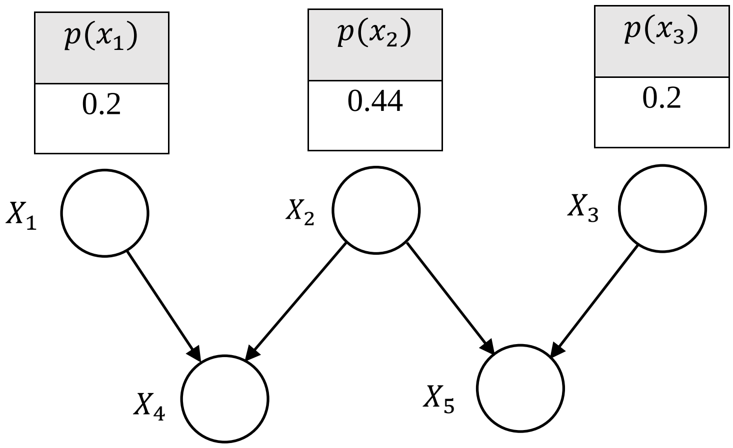
Answer:
True. One path is blocked by *Lung Cancer* while another is blocked by *Bronchitis*.

Question 6

Evaluate (give the distribution tables) the following probabilities:

$$p(x_1\,|x_5),\quad p(x_2\,|x_4),\quad p(x_3\,|x_2),\quad p(x_4\,|x_3),\quad p(x_5)$$

for the Bayesian network shown in Fig. 6.1, where each random variable takes a binary state, i.e. $x_i \in \{T, F\}$. Show all your workings clearly.



X_1	X_2	$p(x_4\|x_1,x_2)$
T	T	0.35
T	F	0.6
F	T	0.01
F	F	0.95

X_2	X_3	$p(x_5\|x_2,x_3)$
T	T	0.35
T	F	0.6
F	T	0.01
F	F	0.95

Fig. 6.1

Answer:

For $p(x_2 \mid x_4)$:

$$p(x_2|x_4) = \frac{p(x_2, x_4)}{\sum_{x_2} p(x_2, x_4)} = \frac{p(x_2, x_4)}{p(x_4)}$$

$$\begin{aligned} p(x_2, x_4) &= \sum_{x_1}\sum_{x_3}\sum_{x_5} p(x_1)p(x_2)p(x_3)p(x_4|x_1, x_2)p(x_5|x_2, x_3) \\ &= \sum_{x_1} p(x_1)p(x_2)p(x_4|x_1, x_2) \\ &= p(x_2)\sum_{x_1} p(x_1)p(x_4|x_1, x_2) \end{aligned}$$

$$p(x_2, x_4) = p(x_2)\{p(x_1 = 0)p(x_4|x_1 = 0, x_2) + p(x_1 = 1)p(x_4|x_1 = 1, x_2)\}$$

x_1	x_2	$p(x_4 = 1\|x_1, x_2)$	$p(x_4 = 0\|x_1, x_2)$
T	T	0.35	0.65
T	F	0.6	0.4
F	T	0.01	0.99
F	F	0.95	0.05

Given the following distribution tables:
We can compute:
1) $x_2 = 0, x_4 = 0$:

$$\begin{aligned} p(x_2, x_4) &= p(x_2 = 0)\{p(x_1 = 0)p(x_4 = 0|x_1 = 0, x_2 = 0) + p(x_1 = 1)p(x_4 = 0|x_1 = 1, x_2 = 0)\} \\ &= (0.56)\{(0.8)(0.05) + (0.2)(0.4)\} \\ &= 0.0672 \end{aligned}$$

2) $x_2 = 0, x_4 = 1$:

$$\begin{aligned} p(x_2, x_4) &= p(x_2 = 0)\{p(x_1 = 0)p(x_4 = 1|x_1 = 0, x_2 = 0) + p(x_1 = 1)p(x_4 = 1|x_1 = 1, x_2 = 0)\} \\ &= (0.56)\{(0.8)(0.95) + (0.2)(0.6)\} \\ &= 0.4928 \end{aligned}$$

3) $x_2 = 1, x_4 = 0$:

$$\begin{aligned} p(x_2, x_4) &= p(x_2 = 1)\{p(x_1 = 0)p(x_4 = 0|x_1 = 0, x_2 = 1) + p(x_1 = 1)p(x_4 = 0|x_1 = 1, x_2 = 1)\} \\ &= (0.44)\{(0.8)(0.99) + (0.2)(0.65)\} \\ &= 0.4057 \end{aligned}$$

4) $x_2 = 1, x_4 = 1$:

$$\begin{aligned} p(x_2, x_4) &= p(x_2 = 1)\{p(x_1 = 0)p(x_4 = 1|x_1 = 0, x_2 = 1) + p(x_1 = 1)p(x_4 = 1|x_1 = 1, x_2 = 1)\} \\ &= (0.44)\{(0.8)(0.01) + (0.2)(0.35)\} \\ &= 0.0343 \end{aligned}$$

Hence, the probability distribution of $p(x_2, x_4)$ is summarized as:

x_2	x_4	$p(x_2, x_4)$
0	0	0.0672
0	1	0.4928
1	0	0.4057
1	1	0.0343

Since

$$p(x_4) = \sum_{x_2} p(x_2) \sum_{x_1} p(x_1) p(x_4|x_1, x_2),$$

$p(x_4) = \sum_{x_2} p(x_2, x_4) = p(x_2 = 0, x_4) + p(x_2 = 1, x_4)$

$p(x_4 = 0) = p(x_2 = 0, x_4 = 0) + p(x_2 = 1, x_4 = 0) = 0.0672 + 0.4057 = 0.4729$
$p(x_4 = 1) = p(x_2 = 0, x_4 = 1) + p(x_2 = 1, x_4 = 1) = 0.4928 + 0.0343 = 0.5271$

Since $\quad p(x_2|x_4) = \frac{p(x_2,x_4)}{p(x_4)}$,
if $x2 = 0, x_4 = 0, \quad p(x2|x4) = \frac{p(x_2=0,x_4=0)}{p(x_4=0)} = \frac{0.0672}{0.4729} = 0.1421;$
if $x2 = 0, x_4 = 1, \quad p(x2|x4) = \frac{p(x_2=0,x_4=1)}{p(x_4=1)} = \frac{0.4928}{0.5271} = 0.9349.$

For $\boldsymbol{p(x_1\,|x_5), p(x_3\,|x_2),\; p(x_4\,|x_3)}$:

$p(x_1|x_5) = p(x_1)$ since x_4 d-separates x_1 and x_5.
$p(x_3|x_2) = p(x_3)$ since x_5 d-separates x_2 and x_3.
$p(x_4|x_3) = p(x_4)$ since x_5 d-separates x_3 and x_4.

	$p(x_1)$	$p(x_2)$	$p(x_3)$
=1	0.2	0.44	0.2
=0	0.8	0.56	0.8

For $\boldsymbol{p(x_5)}$:

$$\begin{aligned} p(x_5) &= \sum_{x_1} \sum_{x_2} \sum_{x_3} \sum_{x_4} p(x_1)p(x_2)p(x_3)p(x_4|x_1, x_2)p(x_5|x_2, x_3) \\ &= \sum_{x_2} \sum_{x_3} p(x_2)p(x_3)p(x_5|x_2, x_3) \end{aligned}$$

x_2	x_3	$p(x_5 = 1 \mid x_2, x_3)$	$p(x_5 = 0 \mid x_2, x_3)$
T	T	0.35	0.65
T	F	0.6	0.4
F	T	0.01	0.99
F	F	0.95	0.05

$\boldsymbol{x_2}$	$\boldsymbol{x_3}$	$\boldsymbol{x_5 = 0}$
0	0	$p(x_2 = 0)p(x_3 = 0)p(x_5 = 0 \mid x_2 = 0, x_3 = 0) =$ (0.56)(0.8)(0.05) = 0.0224
0	1	$p(x_2 = 0)p(x_3 = 1)p(x_5 = 0 \mid x_2 = 0, x_3 = 1) =$ (0.56)(0.2)(0.99) = 0.11088
1	0	$p(x_2 = 1)p(x_3 = 0)p(x_5 = 0 \mid x_2 = 1, x_3 = 0) =$ (0.44)(0.8)(0.40) = 0.1408
1	1	$p(x_2 = 1)p(x_3 = 1)p(x_5 = 0 \mid x_2 = 1, x_3 = 1) =$ (0.44)(0.2)(0.65) = 0.0572

$\boldsymbol{x_2}$	$\boldsymbol{x_3}$	$\boldsymbol{x_5 = 1}$
0	0	$p(x_2 = 0)p(x_3 = 0)p(x_5 = 1 \mid x_2 = 0, x_3 = 0) =$ (0.56)(0.8)(0.95) = 0.4256
0	1	$p(x_2 = 0)p(x_3 = 1)p(x_5 = 1 \mid x_2 = 0, x_3 = 1) =$ (0.56)(0.2)(0.01) = 0.00112
1	0	$p(x_2 = 1)p(x_3 = 0)p(x_5 = 1 \mid x_2 = 1, x_3 = 0) =$ (0.44)(0.8)(0.60) = 0.2112
1	1	$p(x_2 = 1)p(x_3 = 1)p(x_5 = 1 \mid x_2 = 1, x_3 = 1) =$ (0.44)(0.2)(0.35) = 0.0308

$\boldsymbol{x_5}$	$\boldsymbol{p(x_5)}$
0	0.0224+0.11088+0.1408+0.0572=0.33128
1	0.4256+0.00112+0.2112+0.0308=0.66872

Question 7

Give the junction tree of the Bayesian network shown in Fig. 7.1 using the following elimination order: $\{X_7, X_6, X_5, X_4, X_3, X_2, X_1\}$. Show all your workings clearly.

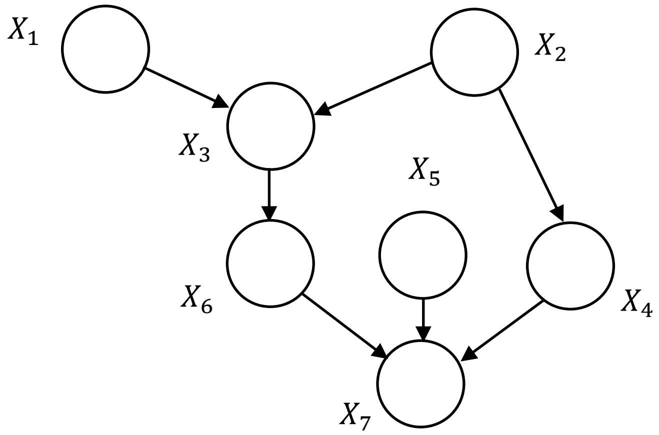
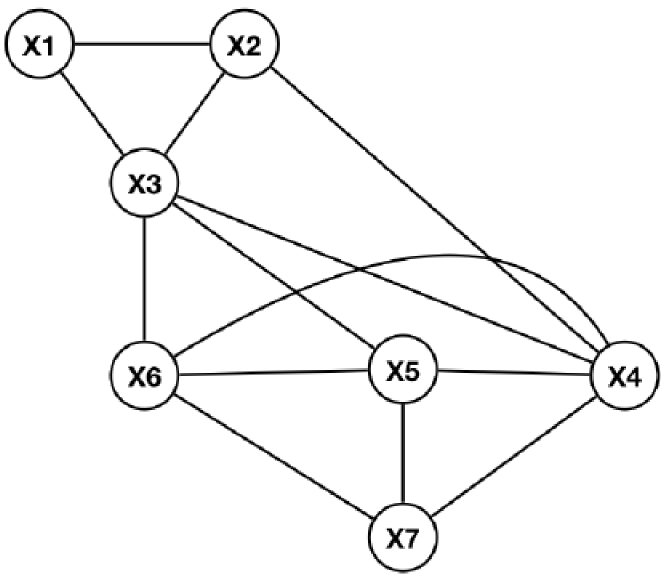


Fig. 7.1

Answer:

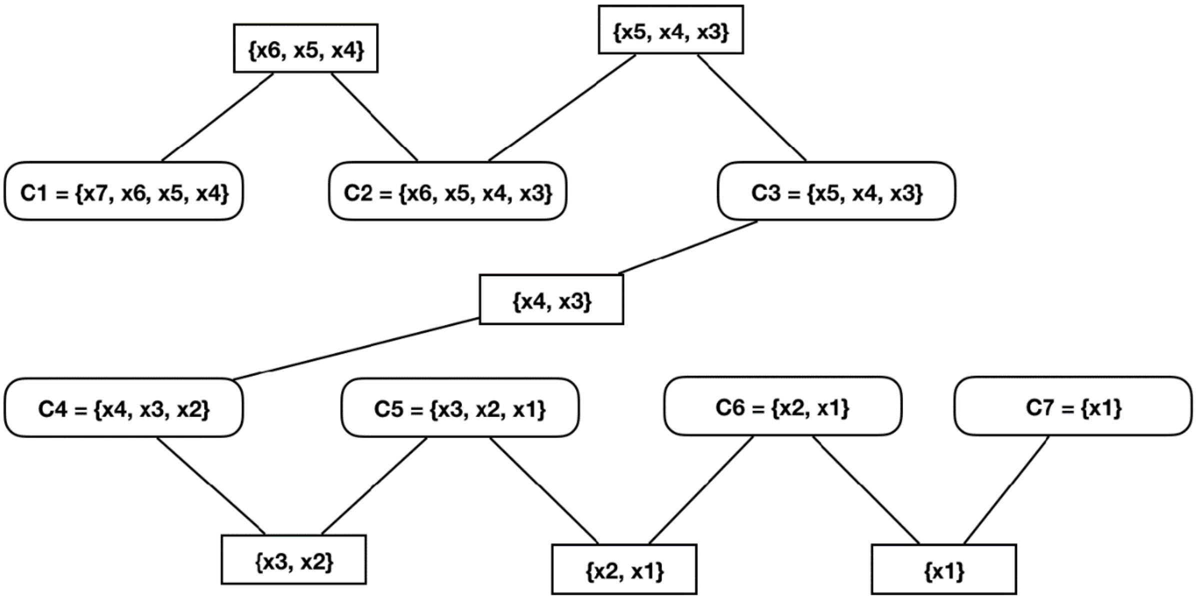
1. Moralization
2. Triangulation



3. Form clusters from elimination clusters:

$$C_1 : \{x_7, x_6, x_5, x_4\}$$
$$C_2 : \{x_6, x_5, x_4, x_3\}$$
$$C_3 : \{x_5, x_4, x_3\}$$
$$C_4 : \{x_4, x_3, x_2\}$$
$$C_5 : \{x_3, x_2, x_1\}$$
$$C_6 : \{x_2, x_1\}$$
$$C_7 : \{x_1\}$$

4. Get minimum spanning tree



--End--