CS726 Homework - 1

Name: Sandarbh Yadav

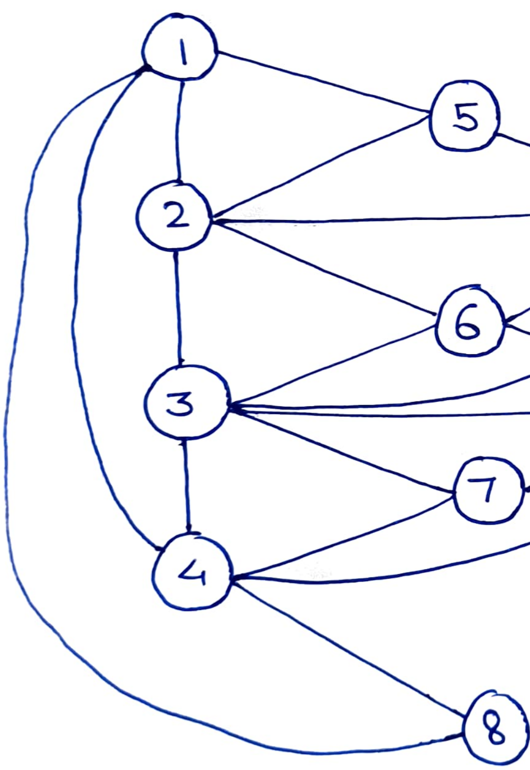
Roll no: 22D0374

Q1 (a) ϕ empty set.

x_5 is non descendant of x_4 & x_4 has no parent

so $x_5 \perp\!\!\!\perp x_4$ and vice versa

(b) we add an edge for every directed edge and also add an edge between parents of each node. Moralized structure is.



Extra edges added are: (1,2), (2,3), (3,4), (1,4), (2,9), (3,9), (3,10), (4,10)

1 (c) No the undirected graphical model is not triangulated because there exists a cycle of length 4. between nodes 1, 2, 3 & 4.

1 (d) No, the markov blanket of each node is unique. So any other correct and minimal undirected graph is not possible as there is only one unique way of moralizing.

1 (e) CIs which hold in original BN but not in undirected graph are:

$x_1 \perp\!\!\!\perp x_2$

$x_2 \perp\!\!\!\perp x_3$

$x_3 \perp\!\!\!\perp x_4$

$x_1 \perp\!\!\!\perp x_4$

$x_9 \perp\!\!\!\perp x_2 \mid x_5$

$x_{10} \perp\!\!\!\perp x_3 \mid x_6$

$x_3 \perp\!\!\!\perp x_9$

$x_4 \perp\!\!\!\perp x_{10}$

Q2 (a) Given: Out of 1000 people, 20 have cold, 2 have TB & 1 have corona.

$\alpha_{11} = 0.8$ $\alpha_{12} = 0.3$

$\alpha_{21} = \alpha_{22} = \alpha_{23} = 0.2$

$\alpha_{32} = 0.7$

2a (i) $P(D_1)$

2 out of 1000 people have TB so potential values are:

$D_1 = 0$	0.998
$D_1 = 1$	0.002

2a(ii) $P(S_3 \mid Pa(S_3)) = P(S_3 \mid D_2)$

$P_r(S_j = 0 \mid Pa(S_j)) = \Pi_{i \in Pa(S_j)} (1 - \alpha_{ij})^{D_i}$

$P(S_3 = 0 \mid D_2 = 0) = (1 - \alpha_{23})^0 = 1$

$P(S_3 = 1 \mid D_2 = 0) = 1 - P(S_3 = 0 \mid D_2 = 0) = 1 - 1 = 0$

$P(S_3 = 0 \mid D_2 = 1) = (1 - \alpha_{23})^1 = 1 - 0.2 = 0.8$

$P(S_3 = 1 \mid D_2 = 1) = 1 - P(S_3 = 0 \mid D_2 = 1) = 1 - 0.8 = 0.2$

	$S_3 = 0$	$S_3 = 1$
$D_2 = 0$	1	0
$D_2 = 1$	0.8	0.2

Values denote $P(S_3 \mid D_2)$

2a(iii) $P(S_1 \mid Pa(S_1)) = P(S_1 \mid D_1, D_2)$

$$P(S_1 = 0 \mid D_1, D_2) = (1-\alpha_{11})^{D_1}(1-\alpha_{21})^{D_2}$$
$$= (1-0.8)^{D_1}(1-0.2)^{D_2}$$
$$= (0.2)^{D_1}(0.8)^{D_2}$$

$P(S_1 = 0 \mid D_1 = 0, D_2 = 0) = (0.2)^0(0.8)^0 = 1$

$P(S_1 = 1 \mid D_1 = 0, D_2 = 0) = 1 - P(S_1 = 0 \mid D_1 = 0, D_2 = 0) = 1 - 1 = 0$

$P(S_1 = 0 \mid D_1 = 0, D_2 = 1) = (0.2)^0(0.8)^1 = 0.8$

$P(S_1 = 1 \mid D_1 = 0, D_2 = 1) = 1 - P(S_1 = 0 \mid D_1 = 0, D_2 = 1) = 1 - 0.8 = 0.2$

$P(S_1 = 0 \mid D_1 = 1, D_2 = 0) = (0.2)^1(0.8)^0 = 0.2$

$P(S_1 = 1 \mid D_1 = 1, D_2 = 0) = 1 - P(S_1 = 0 \mid D_1 = 1, D_2 = 0) = 1 - 0.2 = 0.8$

$P(S_1 = 0 \mid D_1 = 1, D_2 = 1) = (0.2)^1(0.8)^1 = 0.16$

$P(S_1 = 1 \mid D_1 = 1, D_2 = 1) = 1 - P(S_1 = 0 \mid D_1 = 1, D_2 = 1) = 1 - 0.16 = 0.84$

	$S_1 = 0$	$S_1 = 1$
$D_1 = 0, D_2 = 0$	1	0
$D_1 = 0, D_2 = 1$	0.8	0.2
$D_1 = 1, D_2 = 0$	0.2	0.8
$D_1 = 1, D_2 = 1$	0.16	0.84

Values denote $P(S_1 \mid D_1, D_2)$

2 (b) Marginal probability $P(S_1 = 1)$ can be expressed in terms of parents of S_1 ie D_1 and D_2

$$P(S_1) = \sum_{D_1, D_2} P(S_1 \mid D_1, D_2)\, P(D_1)\, P(D_2)$$

This factorization is according to graph.

Also since we have equation of $P(S_1 = 0 \mid D_1, D_2)$ ~~we~~ can calculate it and then

~~$P(S_1) = 1 - P($~~

$$P(S_1 = 1) = 1 - P(S_1 = 0)$$

$$= 1 - \sum_{D_1, D_2} P(S_1 = 0 \mid D_1, D_2)\, P(D_1)\, P(D_2)$$

$$= 1 - \Big[(1 - B_1)(1 - B_2) + (1 - B_1) B_2 (1 - \alpha_{21}) + B_1 (1 - B_2)(1 - \alpha_{11}) + B_1 B_2 (1 - \alpha_{11})(1 - \alpha_{21}) \Big]$$

$$= 1 - \Big[1 - \cancel{B_1} - \cancel{B_2} + \cancel{B_1 B_2} + \cancel{B_2} - \cancel{B_1 B_2} - B_2 \alpha_{21} + \cancel{B_1 B_2 \alpha_{21}} + \cancel{B_1} - \cancel{B_1 B_2} - B_1 \alpha_{11} + \cancel{B_1 B_2 \alpha_{11}} + \cancel{B_1 B_2} - \cancel{B_1 B_2 \alpha_{11}} - \cancel{B_1 B_2 \alpha_{21}} + B_1 B_2 \alpha_{11} \alpha_{21} \Big]$$

$$= 1 - 1 + B_2 \alpha_{21} + B_1 \alpha_{11} - B_1 B_2 \alpha_{11} \alpha_{21}$$

$$= \boxed{B_1 \alpha_{11} + B_2 \alpha_{21} - B_1 B_2 \alpha_{11} \alpha_{21}}$$

This is expression of $P(S_1 = 1)$

2 (c) $S_1 = 1$ Fever

$S_2 = 0$ No cough — At most one disease

$S_3 = 1$ No smell

We can calculate probabilities $P(D_1=1, D_2=0, D_3=0 \mid S_1=1, S_2=0, S_3=1)$, $P(D_1=0, D_2=1, D_3=0 \mid S_1=1, S_2=0, S_3=1)$ & $P(D_1=0, D_2=0, D_3=1 \mid S_1=1, S_2=0, S_3=1)$ and whichever is maximum will be the disease.

Using ~~Bayes rule~~. conditional probability definition

$$P(D_1, D_2, D_3 \mid S_1=1, S_2=0, S_3=1) = \frac{P(D_1, D_2, D_3, S_1=1, S_2=0, S_3=1)}{P(S_1=1, S_2=0, S_3=1)}$$

Denominator will be same for all 3 diseases so we consider only numerator which can be factorized using Bayesian network as

$$P(S_1=1 \mid D_1, D_2)\, P(S_2=0 \mid D_1, D_2, D_3)\, P(S_3=1 \mid D_2)\, P(D_1)\, P(D_2)\, P(D_3)$$

1. $D_1 = 1, D_2 = 0, D_3 = 0$

   $P(S_3=1 \mid D_2=0)$ term will be zero so prob is zero

2. $D_1 = 0, D_2 = 1, D_3 = 0$

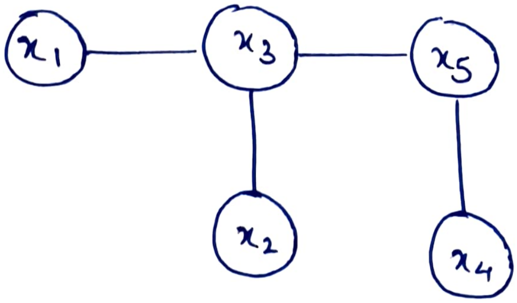
   All terms are non zero so there will be some finite value of probability.

3. $D_1 = 0, D_2 = 0, D_3 = 1$

   Again $P(S_3=1 \mid D_2=0)$ term will be zero so prob is zero

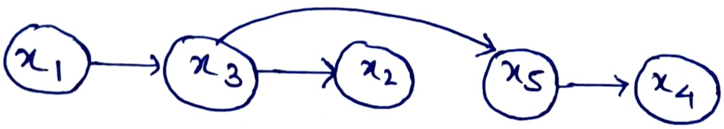
Hence, the most likely disease is Corona.

Q3 (a)



undirected graphical model corresponding to above potentials.

3 (b) Using x_1 x_3 x_2 x_5 x_4, the Bayesian network obtained is :

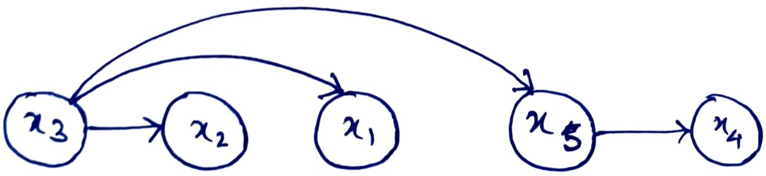


$x_3 \perp x_1 \mid \phi$ × $x_2 \perp x_1, x_3 \mid \phi$ × $x_5 \perp x_1, x_2, x_3 \mid \phi$ × $x_4 \perp x_1, x_2 x_3 x_5 \mid \phi$ ×

$x_3 \perp \phi \mid x_1$ ✓ $x_2 \perp x_1 \mid x_3$ ✓ $x_5 \perp x_1, x_2 \mid x_3$ $x_4 \perp x_1 x_2 x_3 \mid x_5$

Since the graph is chordal, the Bayesian network is perfect

3 (c) Using x_3 x_2 x_1 x_5 x_4, the Bayesian n/w obtained is :



$x_2 \perp x_3 \mid \phi$ × $x_1 \perp x_2 x_3 \mid \phi$ × $x_5 \perp x_1 x_2 x_3 \mid \phi$ × $x_4 \perp x_1 x_2 x_3 x_5 \mid \phi$ ×

$x_2 \perp \phi \mid x_3$ ✓ $x_1 \perp x_2 \mid x_3$ ✓ $x_5 \perp x_1, x_2 \mid x_3$ $x_4 \perp x_1, x_2 x_3 \mid x_5$

Since the graph is chordal, the Bayesian network is perfect.