Karan Sheth
KS6208

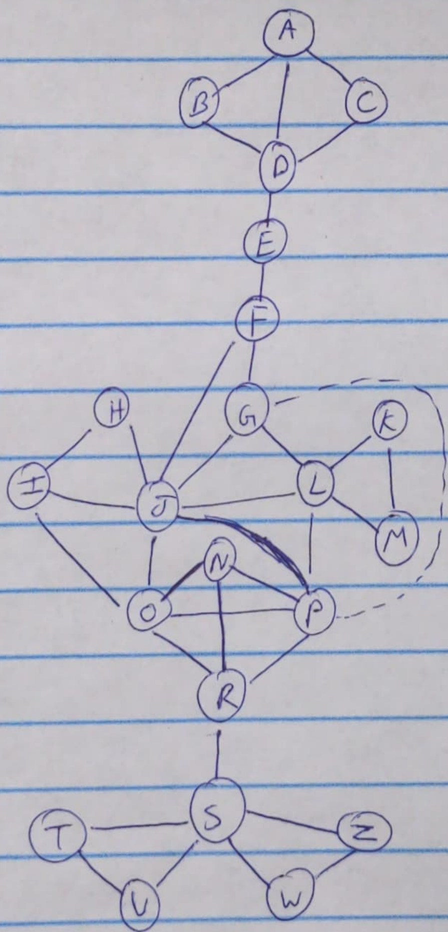
Problem 1

To convert the given ~~a~~ graph into a junction tree we perform the following steps:-

Moralization

→ Convert directed graph to undirected graph

→ Connect nodes that have a common children.



Here, we introduced ~~two~~ few new edges

→ J - P

→ O - N

→ A - D

→ F - J

→ I - O

→ G - P

→ N - R

After this we perform triangulation to eliminate any loop with more than or equal to 4 nodes.

~~This gives us the following graph.~~

Since there are no such loops in above graph, we can directly find out the cliques and apply kruskal's algorithm.

Cliques:

→ ABD → ACD → DE → EF → FGJ

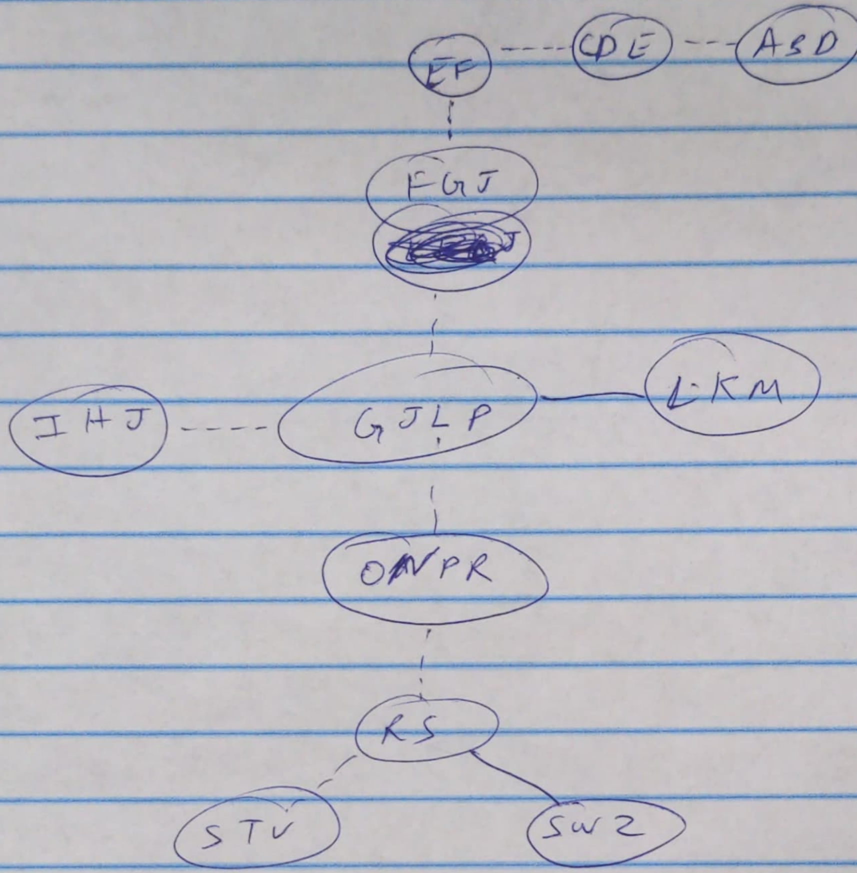
→ GJLP → LKM → HIJ → JPO → ONPR

→ JIO → RS → STV → SWZ

Using this we build Kruskal Table as follows:

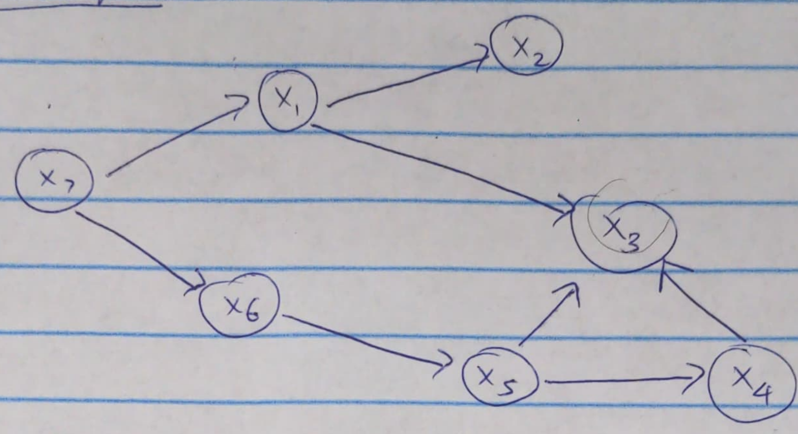
	DE	EF	RS	ABD	ACD	FGJ	LKM	HIJ	JPO	JIO	STV	SWZ	G	O
DE	-	1	0											
EF		-	0											
RS			-											
ABD				-										
ACD					-									
FGJ						-								
LKM							-							
HIJ								-						
JPO									-					
JIO										-				
STV											-			
SWZ												-		
GJLP													-	
ONPR														-

And the using the above we can convert it into JTA as below.


EF
CDE
ASD
FGJ
IHJ
GJLP
LKM
ONPR
RS
STV
SWZ

Problem 2

Given Graph :-



$$p(x_1, \dots x_7) = \prod_{i=1}^{7} p(x_i | \text{parents}_i)$$

$$= \cancel{p(x_7)(x}$$

$$= p(x_7)\, p(x_1|x_7)\, p(x_2|x_1)\, p(x_6|x_7)\, p(x_5|x_6)$$
$$\cancel{p(x_3|x_5, x_1)\,p(} \; p(x_3|x_1, x_4, x_5)\, p(x_4|x_5)$$

Rearanging for simplicity we get :-

$$p(x_1 \dots x_7) = p(x_1|x_7)\, p(x_2|x_1)\, p(x_3|x_1, x_4, x_5)\, p(x_4|x_5)\, p(x_5|x_6)$$
$$p(x_6|x_7)\, p(x_7)$$

Now using bayes ball theorem we find the following :-

(a) x_2 and x_6 are independent : $\boxed{\text{FALSE}}$

$x_2 - x_1 - x_7 - x_6$

$\rightarrow x_2 - x_1 - x_7$ Markov chain, Hence ball will go through

$\rightarrow x_1 - x_7 - x_6$ Two effects rule, ball goes through

Hence False x_2 and x_6 are dependent

(b) x_2 and x_6 are independent, given x_1, x_3, x_5 : $\boxed{\text{TRUE}}$

→ $x_2 - x_1 - x_7 - x_6$: Ball stops at x_1 (Markov ~~chat~~ chain)

→ $x_2 - x_1 - x_3 - x_5 - x_6$: Ball stops at x_1

(c) x_1 and x_4 are condinally independen given x_5 : $\boxed{\text{TRUE}}$

→ $x_1 \rightarrow x_3 \leftarrow x_4$: Two causes, Ball ~~goes through~~ Stops since x_3 is not given

→ $x_1 - x_3 - x_5 - x_4$: Same as above

→ $x_1 - x_7 - x_6 - x_5 - x_4$: Ball stops at $(x_6 - x_5 - x_4)$ since x_5 is given (Markov chain)

(d) $x_5 \perp\!\!\!\perp x_2 \mid x_1, x_3$: $\boxed{\text{TRUE}}$

→ $x_5 - x_3 - x_1 - x_2$: Ball stops at x_1 (Two effects)

→ $x_5 - x_6 - x_7 - x_1 - x_2$: Ball stops at x_1 (Markov chain)

(e) $x_5 \perp\!\!\!\perp x_1 \mid x_3, x_2, x_4$: $\boxed{\text{FALSE}}$

$x_5 - x_6 - x_7 - x_1$: $x_5 - x_6 - x_7$ (Markov chain)

$x_6 - x_7 - x_1$ (Two affects) Ball goes through

(f) $x_4 \perp\!\!\!\perp x_7 \mid x_6$: $\boxed{\text{TRUE}}$

$x_4 - x_5 - x_6 - x_7$: Ball stops at x_6 (Markov chain)

$x_4 - x_3 - x_1 - x_7$: Ball stops at x_3 (Two causes)

(g) $x_4 \perp\!\!\!\perp x_7 \mid x_5$: $\boxed{\text{TRUE}}$

$x_4 - x_5 - x_6 - x_7$: Ball stops at x_5 (Markov chain)

$x_4 - x_3 - x_1 - x_7$: Ball stops at x_3 (Two causes)

(h) x_1 ~~and~~ $\perp\!\!\!\perp$ $x_5 \mid x_6, x_7$: $\boxed{\text{TRUE}}$

$x_1 - x_7 - x_6 - x_5$: Ball stops at x_7 (Two effects)

$x_1 - x_3 - x_5$: Ball stops at x_3 (Two causes)

(i) $x_5 \perp\!\!\!\perp x_1$: $\boxed{\text{FALSE}}$

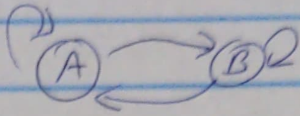
$x_1 - x_7 - x_6 - x_5$: Ball goes Through (Two effects + Markov chain)

(j) $x_2 \perp\!\!\!\perp x_4 \mid x_1$: $\boxed{\text{TRUE}}$

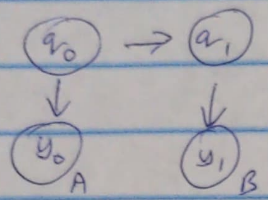
$x_2 - x_1 - x_3 - x_4$: Ball stops at x_1 (Two effects)

$x_2 - x_1 - x_7 - x_6 - x_5 - x_4$: Ball stops at x_1 (Markov chain)

Problem 3

Given finite state machine is :- 

and we construct JTA as :-



we initialize the junction tree using the given values as :-

(q_0, y_0) — $[q_0]$ — (q_0, q_1) — $[q_1]$ — (q_1, y_1)

Init :-

	y_0 = A	B
$q_0=1$	1/32	3/32
$q_0=2$	7/16	7/16

1	1

q_0 \ q_1	1	2
1	1/3	2/3
2	1/4	3/4

1	1

	y_1 = A	B
$q_1=1$	1/4	3/4
$q_1=2$	1/2	1/2

Now we apply the message passing iteration step by step :-

$\longrightarrow$ $\longleftarrow$

q_0 \ y_0	A	B
1	1/32	
2	7/16	

q_0	1	2
	1/32	7/16

q_0 \ q_1	1	2
1	1/3	2/3
2	1/4	3/4

q_1	1	2
	3/4	1/2

q_1 \ y_1	A	B
1		3/4
2		1/2

1/32	
7/16	

q_0 1 2 $\longrightarrow$ q_1 1 2 $\longleftarrow$ q_1 1 2

q_0	1	2
	1/32	7/16

q_0 \ q_1	1	2
1	1/128	1/96
2	21/256	21/128

q_1	1	2
	3/4	1/2

	3/4
	1/2

1/32	
7/16	

q_0 1 2 $\longleftarrow$ q_1 1 2 $\longrightarrow$ 1 2

q_0	1	2
	7/384	63/256

q_0 \ q_1	1	2
1	1/128	1/96
2	21/256	21/128

	1	2

$\frac{23}{256}$ $\frac{67}{384}$

	3/4
	1/2

So the likelihood,

$$p(y) = \frac{7}{384} + \frac{63}{256} = \frac{1}{128} + \frac{1}{96} + \frac{21}{256} + \frac{21}{128}$$

$$= \frac{23}{256} + \frac{67}{384}$$

$$= \frac{203}{768} = \boxed{0.264}$$

$$p(q_0 = 1 | y) = \frac{7/384}{7/384 + 63/256} = \frac{2}{29}$$

$$p(q_0 = 2 | y) = \frac{63/256}{7/384 + 63/256} = \frac{27}{29}$$

$$p(q_1 = 1 | y) = \frac{23/256}{23/256 + 67/384} = \frac{69}{203}$$

$$p(q_1 = 2 | y) = \frac{67/384}{23/256 + 67/384} = \frac{134}{203}$$

Problem 4

Initial centers = $\mu_1 = (-4,2)$ $\mu_2 = (4,-5)$

Data-points = $(-5,3)$, $(-3,2)$, $(-4,5)$, $(-3,4)$, $(3,-4)$, $(4,-2)$, $(6,-6)$, $(8,-3)$

0, 1, 2, 3, 4, 5, 6, 7

So lets say we have points $x_0 \dots x_7$ as above we initially calculate distance of each point to the clusters centers as

Points	$\mu_1(-4,2)$	$\mu_2(4,-5)$	min ($\mu_1(-4,2)$, $\mu_2(4,-5)$ which clusters?
-5,3	1.414	12.04	Cluster 1
-3,2	1	9.89	Cluster 1
-4,5	3	12.80	Cluster 1
-3,4	2.23	11.40	Cluster 1
3,-4	9.21	1.4	Cluster 2
4,-2	8.94	3	Cluster 2
6,-6	12.80	2.2	Cluster 2
8,-3	13	4.4	Cluster 2

Now based on above we calculate new centers :-

$$\mu_1 = \left(\frac{-5+-3+-4+-3}{4}, \frac{3+2,+5,+4}{4}\right) = \left(\frac{-15}{4}, \frac{15}{4}\right)$$

$$\mu_2 = \left(\frac{3+4+6+8}{4}, \frac{-4+-2+-6+-3}{4}\right) = \left(\frac{21}{4}, \frac{-15}{4}\right)$$

P.T.O →

with these new centers we do ~~as the~~ the same as above and calculate distances.

Points	$\mu_1(-15/4, 15/4)$	$\mu_2(21/4, -15/4)$	Which Cluster ?
-5, 3	1.45	12.27	1
-3, 2	1.90	10.05	1
-4, 5	1.27	12.73	1
-3, 4	0.79	11.31	1
3, -4	10.27	2.26	2
4, -2	9.65	2.15	2
6, -6	13.78	2.37	2
8, -3	13.55	2.85	2

New cluster centers :- They are same as before since cluster is same

$\mu_1 = \left(\frac{-15}{4}, \frac{15}{4}\right) \quad \mu_2 = \left(\frac{21}{4}, \frac{-15}{4}\right)$

Cluster 1 : [(-5, 3), (-3, 2), (-4, 5), (-3, 4)]

Cluster 2 :- [(3, -4), (4, -2), (6, -6), (8, -3)]

Problem 5

KL Divergence :- $KL(p||q) = \sum_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)}$

where both $p(x)$ and $q(x)$ are probability distrubution
we have to use Jensen's inequality to prove ~~that~~ that given KL Divergence is non-negative.

Jensen's inequality for distrubution is as follows:-

$$f(E[x]) \leq E[f(x)]$$

and for continuos densities it is :-

$$f\left(\int x p(x)\, dx\right) \leq \int f(x) p(x)\, dx$$

Now, applying ~~Jen~~ above jensen's inequality to the given KL-divergence we get :-

$$KL(p||q) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)}\, dx$$

$$= - \int_{-\infty}^{\infty} p(x) \log \frac{q(x)}{p(x)}\, dx$$

$$\geq - \log \int_{-\infty}^{\infty} \frac{\cancel{p(x)}\, q(x)}{\cancel{p(x)}}\, dx \quad \text{... Applying Jensen's inequality}$$

$$\geq 0 \quad \text{.... since } \boxed{\log 1 = 0}$$

$\boxed{\text{Hence Non-Negative}}$

We first flip the log to cancel out $p(x)$ when jensen's inequality is applied

Problem 6

Given the CNN architechture we have to find size of input image (x, y)

so lets assume size is x, y and calculate the output size (which is already known)

Conv layer 1 :- Input size $= (x, y)$

output channels $= 5$, filter size $= (3, 3)$

stride $= (2, 3)$

Output $= \left[\frac{x-3}{2} + 1, \frac{y-3}{3} + 1\right]$

Relu has no effect on size.

Maxpool layer 1 :- Region size $= (2, 2)$, stride $= (2, 2)$

Output size $= /2$

$$= \left[\frac{x-3}{4} + \frac{1}{2}, \frac{y-3}{6} + \frac{1}{2}\right]$$

Now for next lets assume input as x, y again P.T.O →

Conv laye 2 :- Input size $= x, y$

output channels $= 7$ Inputs channels $= 5$

filter $= 3 \times 3$ stride $= 2 \times 2$

Output size $= \left[\frac{x-2}{4} + \frac{1}{2}, \frac{y-9}{6} + \frac{1}{2}\right]$

After Maxpool with region : 2×2 and stride 3×3

we have output size as

$$= \left[\frac{x-7}{12} + \frac{1}{3}, \frac{y-9}{18} + \frac{1}{3}\right]$$

Conv layer, 2 $\quad$ Input size $= \left[\frac{x-1}{4} , \frac{y}{6} \right]$

filter $= 3\times3$ $\quad$ stride $= 2\times2$

output size $= \left[\frac{\frac{x-1}{4} - 3}{2} + 1 , \frac{y/6 - 3}{2} + 1 \right]$

$= \left[\frac{x-13}{8} + 1 , \frac{y-18}{12} + 1 \right]$

Maxpool : stride $= 3\times3$ $\quad$ region : 2×2

output size $= \left[\frac{x-13}{24} + \frac{1}{3} , \frac{y-18}{36} + \frac{1}{3} \right]$

Now actual output size given is $(7, 14)$

$\therefore \frac{x-13}{24} + \frac{1}{3} = 7$

$\therefore x = \left(7 - \frac{1}{3}\right) \times 24 + 13$

$= 173$

$\therefore \frac{y-18}{36} + \frac{1}{3} = 14$

$\therefore y = \left(14 - \frac{1}{3}\right) \times 36 + 18$

$\therefore \boxed{y = 510}$

Image size $= (173, 510)$