Yinhong Qin
yq2021@nyu.edu

Homework 5

Problem 1

The chance to win the car is greater after changing the chosen door.

Proof. Let A represents the event "Randomly choose a door first", B represents the event "The door opened by the host" and C denotes the event "The car is behind the door".
Thus

$$P(A = 1) = P(A = 2) = P(A = 3) = \frac{1}{3}$$

And we also should notice that A and C is independent. Thus

$$P(C = 1) = P(C = 2) = P(C = 3) = \frac{1}{3}$$

Without loss of generality, let us assume the first randomly picked door is door 1, and the host open door 3. Thus, we can have

$$\begin{aligned}
P(B = 3|A = 1, C = 1) &= \frac{1}{2} \\
P(B = 3|A = 1, C = 2) &= 1 \\
P(B = 3|A = 1, C = 3) &= 0
\end{aligned}$$

With the Bayes Rule, we can compute the probability that "The car is behind door 1 when first randomly pick door 1 and the host open the door 3"

$$\begin{aligned}
P(C = 1|A = 1, B = 3) &= \frac{P(B = 3|A = 1, C = 1)P(A = 1, C = 1)}{\sum_{i=1}^{3} P(B = 3|A = 1, C = i)P(A = 1, C = i)} \\
&= \frac{P(B = 3|A = 1, C = 1)P(C = 1)}{\sum_{i=1}^{3} P(B = 3|A = 1, C = i)P(C = i)} \\
&= \frac{P(B = 3|A = 1, C = 1)}{\sum_{i=1}^{3} P(B = 3|A = 1, C = i)} \\
&= \frac{\frac{1}{2}}{\frac{1}{2} + 1 + 0} \\
&= \frac{1}{3}
\end{aligned}$$

Thus, the probability of event "The car is behind door 2 when first randomly pick door 1 and the host open the door 3" is

$$P(C = 2|A = 1, B = 3) = 1 - P(C = 1|A = 1, B = 3) = \frac{2}{3}$$

□

Proof above indicates that after first randomly picking and the host opened a empty door, switching to another door will have the greater chance to win the car.

Yinhong Qin
yq2021@nyu.edu

Homework 5

Problem 2

The factorization of the probability distribution $p(x_1, \cdots, x_5)$ implied by this given directed graph is:

$$p(x_1, \cdots, x_5) = \prod_{i=1}^{5} p(x_i|parents(i)) = p(x_1)p(x_3)p(x_2|x_1)p(x_4|x_1, x_3)p(x_5|x_2, x_4)$$

With Bayes ball algorithm, the answer to questions below are:

1. x_2 and x_4 are independent.
   False. A valid path is $x_2 \to x_1 \to x_4$.

2. x_2 and x_4 are conditionally independent given x_1, x_3 and x_5.
   False. A valid path is $x_2 \to x_5 \to x_4$.

3. x_2 and x_4 are conditionally independent given x_1 and x_3.
   True.

4. x_5 and x_3 are conditionally independent given x_4.
   False. A valid path is $x_5 \to x_2 \to x_1 \to x_4 \to x_3$.

5. x_5 and x_3 are conditionally independent given x_1, x_2 and x_4.
   True.

6. x_1 and x_3 are conditionally independent given x_5.
   False. A valid path is $x_1 \to x_2 \to x_5 \to x_4 \to x_3$.

7. x_1 and x_3 are conditionally independent given x_2.
   True.

8. x_2 and x_3 are independent.
   True.

9. x_2 and x_3 are conditionally independent given x_5.
   False. A valid path is $x_2 \to x_5 \to x_4 \to x_3$.

10. x_2 and x_3 are conditionally independent given x_4 and x_5.
    False. A valid path is $x_2 \to x_1 \to x_4 \to x_3$.

Problem 3

The graph after moralization is listed below:

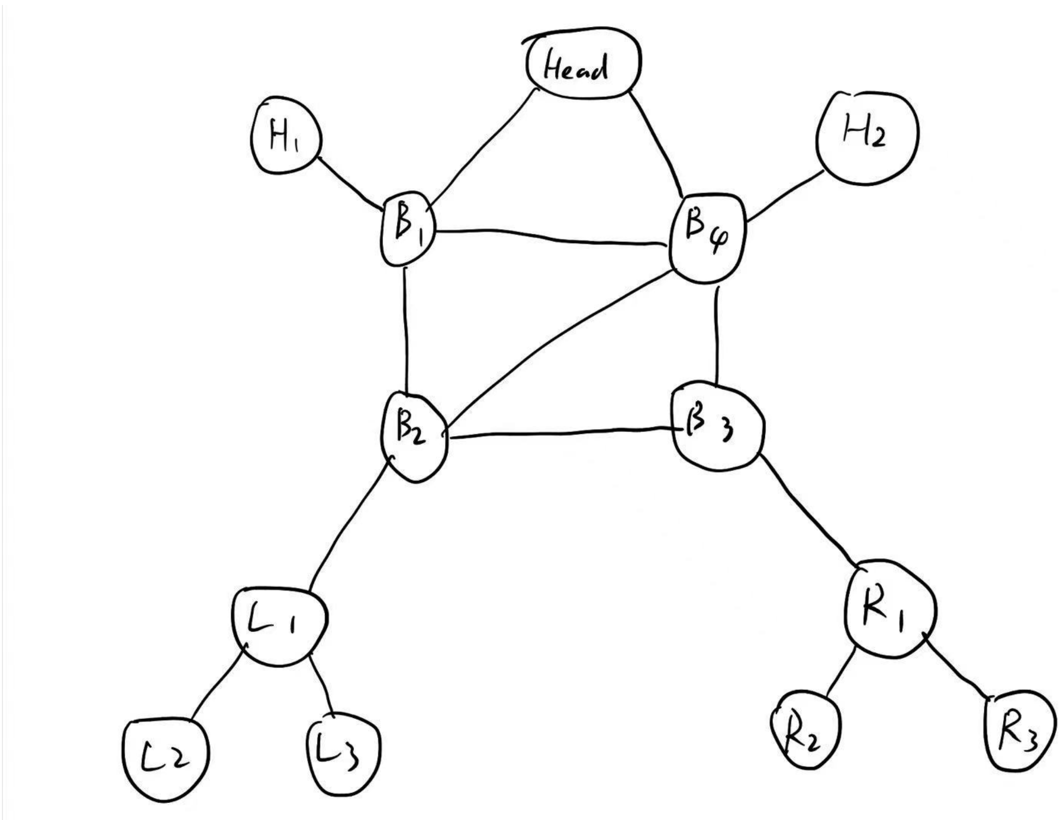

Figure 1: Graph after moralization

The junction tree constructed on this moralization graph is listed below:

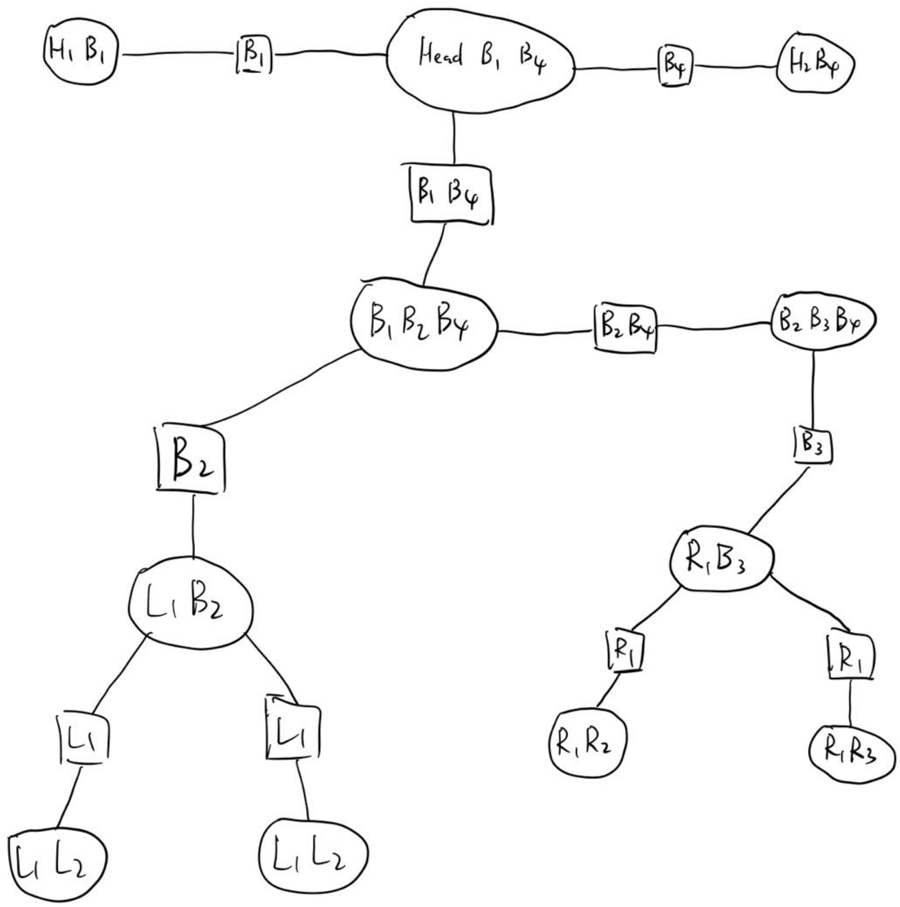

Figure 2: Junction tree built on moralization graph

Problem 4

The junction tree of given undirected graph is listed below:

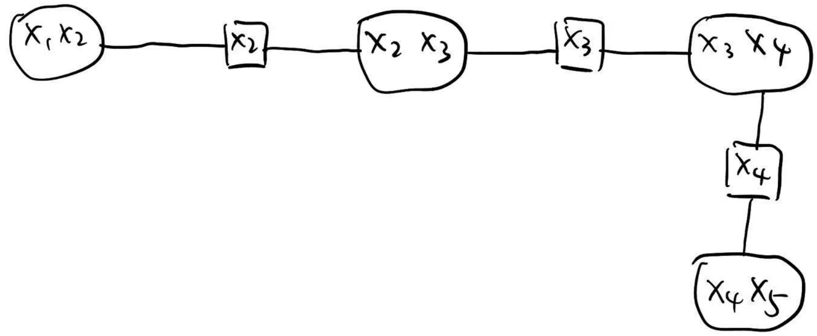

Figure 3: Junction tree built given undirected graph

For this junction tree. we can firstly perform forward message passing from node x_1x_2 to node x_4x_5. After that, we can perform the backward message passing from node x_4x_5 to node x_1x_2.

The computed pairwise marginals and single variable marginals are listed below. We can notice that the single variable marginal from different pairwise marginal are the same.

$p(x_1, x_2)$	$x_2 = 0$	$x_2 = 1$	$p(x_1)$
$x_1 = 0$	0.04046	0.44509	0.48555
$x_1 = 1$	0.32370	0.19075	0.51445
$p(x_2)$	0.36416	0.63584	

$p(x_2, x_3)$	$x_3 = 0$	$x_3 = 1$	$p(x_2)$
$x_2 = 0$	0.26012	0.10405	0.36416
$x_2 = 1$	0.05780	0.57803	0.63584
$p(x_3)$	0.31792	0.68208	

$p(x_3, x_4)$	$x_4 = 0$	$x_4 = 1$	$p(x_3)$
$x_3 = 0$	0.11922	0.19870	0.31792
$x_3 = 1$	0.63945	0.04263	0.68208
$p(x_4)$	0.75867	0.24133	

$p(x_4, x_5)$	$x_5 = 0$	$x_5 = 1$	$p(x_4)$
$x_4 = 0$	0.56900	0.18967	0.75867
$x_4 = 1$	0.06033	0.18100	0.24133
$p(x_5)$	0.62933	0.37067	

Problem 5

The most likely sequence of emotional states is:

Day 1	Day 2	Day 3	Day 4	Day 5
Happy	Angry	Angry	Angry	Angry

We can use the **Viterbi Algorithm** to get this answer. For this question, the state transition probability matrix is (Let "Happy" = 1 and "Angry = 2"):

$$A = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$$

The initial probalbility matrx is:

$$\pi = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

The emission matrix is (Let "smile = 1", "frown = 2" and so on):

$$B = \begin{bmatrix} 0.4 & 0.1 & 0.3 & 0.2 \\ 0.1 & 0.4 & 0.2 & 0.3 \end{bmatrix}$$

The observation sequence is:

$$O = \begin{bmatrix} smile & yell & frown & frown & laugh \end{bmatrix} = \begin{bmatrix} 1 & 4 & 2 & 2 & 3 \end{bmatrix}$$

Let $\delta_t(i)$ denotes the maximum probabilty when we reach state i at time t and our current observation is $o_1, o_2, \cdots, o_t$. The initial value of $\delta_t(i)$ is:

$$\delta_{t=1}(i) = \pi_i * B(i, o_1), \quad i = 1, 2\ (Happy\ or\ Angry)$$

For this question,

$$\begin{aligned} \delta_1(Happy) &= \delta_1(1) = \pi_1 * B(Happy, smile) \\ &= \pi_1 * B(1, 1) \\ &= 1 * 0.4 = 0.4 \end{aligned}$$

$$\begin{aligned} \delta_1(Angry) &= \delta_1(2) = \pi_1 * B(Angry, smile) \\ &= \pi_2 * B2, 1) \\ &= 0 * 0.1 = 0 \end{aligned}$$

The iteration of $\delta_t(i)$ is:

$$\delta_t(i) = \max_{1\leqslant j\leqslant N}[\delta_{t-1}(j)A_{ji}] \times B(i, o_t)$$

N is the number of states. For this question, $N = 2$.

Let $\Psi_t(i)$ denotes the $(t-1)$th node which has the max probability of the optimal path. In other words, it denotes the previous node of our current state.

$$\Psi_t(i) = \arg\max_{1\leqslant j\leqslant N}[\delta_{t-1}(j)A_{ji}]$$

Its initial value is set to 0.

$$\Psi_1(i) = 0$$

It will be updated at the same time when δ_i is updated.

The computation of this question is listed below:

t = 1

$$\delta_1(1) = 0.4, \quad \delta_1(2) = 0$$
$$\Psi_1(1) = 0, \quad \Psi_1(2) = 0$$

t = 2

$$\begin{aligned}
\delta_2(1) &= \max_{1\leqslant j\leqslant N}[\delta_1(j)A_{j1}] \times B(1, o_2) \\
&= \max_{1\leqslant j\leqslant N}[\delta_1(1)A_{11}, \delta_1(2)A_{21}] \times B(1, yell) \\
&= \max_{1\leqslant j\leqslant N}[0.4 \times 0.8, 0 \times 0.2] \times 0.2 = 0.064 \\
\Psi_2(1) &= 1
\end{aligned}$$

$$\begin{aligned}
\delta_2(2) &= \max_{1\leqslant j\leqslant N}[\delta_1(j)A_{j2}]B(2, o_2) \\
&= \max_{1\leqslant j\leqslant N}[\delta_1(1)A_{12}, \delta_1(2)A_{22}] \times B(2, yell) \\
&= \max_{1\leqslant j\leqslant N}[0.4 \times 0.2, 0 \times 0.8] \times 0.3 = 0.024 \\
\Psi_2(2) &= 1
\end{aligned}$$

t = 3

$$
\begin{aligned}
\delta_3(1) &= \max_{1\leqslant j\leqslant N}[\delta_2(j)A_{j1}] \times B(1, o_3) \\
&= \max_{1\leqslant j\leqslant N}[\delta_2(1)A_{11}, \delta_2(2)A_{21}] \times B(1, frown) \\
&= \max_{1\leqslant j\leqslant N}[0.064 \times 0.8, 0.024 \times 0.2] \times 0.1 = 0.00512 \\
\Psi_3(1) &= 1
\end{aligned}
$$

$$
\begin{aligned}
\delta_3(2) &= \max_{1\leqslant j\leqslant N}[\delta_2(j)A_{j2}]B(2, o_3) \\
&= \max_{1\leqslant j\leqslant N}[\delta_2(1)A_{12}, \delta_2(2)A_{22}] \times B(2, frown) \\
&= \max_{1\leqslant j\leqslant N}[0.064 \times 0.2, 0.024 \times 0.8] \times 0.4 = 0.00768 \\
\Psi_3(2) &= 2
\end{aligned}
$$

t = 4

$$
\begin{aligned}
\delta_4(1) &= \max_{1\leqslant j\leqslant N}[\delta_3(j)A_{j1}] \times B(1, o_3) \\
&= \max_{1\leqslant j\leqslant N}[\delta_3(1)A_{11}, \delta_3(2)A_{21}] \times B(1, frown) \\
&= \max_{1\leqslant j\leqslant N}[0.00512 \times 0.8, 0.00768 \times 0.2] \times 0.1 = 0.0004096 \\
\Psi_4(1) &= 1
\end{aligned}
$$

$$
\begin{aligned}
\delta_4(2) &= \max_{1\leqslant j\leqslant N}[\delta_3(j)A_{j2}]B(2, o_3) \\
&= \max_{1\leqslant j\leqslant N}[\delta_3(1)A_{12}, \delta_3(2)A_{22}] \times B(2, frown) \\
&= \max_{1\leqslant j\leqslant N}[0.00512 \times 0.2, 0.00768 \times 0.8] \times 0.4 = 0.0024576 \\
\Psi_4(2) &= 2
\end{aligned}
$$

t = 5

$$\begin{aligned}
\delta_5(1) &= \max_{1\leqslant j\leqslant N}[\delta_4(j)A_{j1}] \times B(1, o_2) \\
&= \max_{1\leqslant j\leqslant N}[\delta_4(1)A_{11}, \delta_4(2)A_{21}] \times B(1, laugh) \\
&= \max_{1\leqslant j\leqslant N}[0.0004096 \times 0.8, 0.0024576 \times 0.2] \times 0.3 = 0.000147456 \\
\Psi_5(1) &= 2
\end{aligned}$$

$$\begin{aligned}
\delta_5(2) &= \max_{1\leqslant j\leqslant N}[\delta_4(j)A_{j2}]B(2, o_2) \\
&= \max_{1\leqslant j\leqslant N}[\delta_4(1)A_{12}, \delta_4(2)A_{22}] \times B(2, laugh) \\
&= \max_{1\leqslant j\leqslant N}[0.0004096 \times 0.2, 0.0024576 \times 0.8] \times 0.2 = 0.000393216 \\
\Psi_5(2) &= 2
\end{aligned}$$

Since $\delta_5(1) < \delta_5(2)$, the final emotional state is "Angry". Then, using the $\Psi_t(i)$ sequence to trace back the optimal path, we can have:

$$\begin{aligned}
emotional\ state_4 = \Psi_5(emotional\ state_5) = \Psi_5(2) = 2 = Angry \\
emotional\ state_3 = \Psi_4(emotional\ state_4) = \Psi_4(2) = 2 = Angry \\
emotional\ state_2 = \Psi_3(emotional\ state_3) = \Psi_3(2) = 2 = Angry \\
emotional\ state_1 = \Psi_2(emotional\ state_2) = \Psi_2(2) = 1 = Happy
\end{aligned}$$

Thus, the most likely sequence of emotional states is:

Day 1	Day 2	Day 3	Day 4	Day 5
Happy	Angry	Angry	Angry	Angry