MACHINE LEARNING COMS 4771
HOMEWORK 5 SOLUTIONS

1 Probability (10 points)

Let $T \in \{1,2,3\}$ indicate the door that the car is hidden behind, $I \in \{1,2,3\}$ denote the initially chosen door, and $H \in \{1,2,3\}$ be the door the host opened. Suppose the fact is that I initially chose door 1, and the host opened door 3. We have:

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

$$P(H=3|T=1, I=1) = \frac{1}{2},$$

$$P(H=3|T=2, I=1) = 1,$$

$$P(H=3|T=3, I=1) = 0,$$

$P(T=i|I=j) = P(T=i), i, j \in \{1,2,3\}$, since T and I are independent.

Applying Bayesian rule:

$$\begin{aligned}
P(T=1|H=3, I=1) &= \frac{P(H=3|T=1, I=1)P(T=1|I=1)}{\sum\limits_{k=1}^{3} P(H=3|T=k, I=1)P(T=k|I=1)} \\
&= \frac{P(H=3|T=1, I=1)P(T=1)}{\sum\limits_{k=1}^{3} P(H=3|T=k, I=1)P(T=k)} \\
&= \frac{P(H=3|T=1, I=1)}{\sum\limits_{k=1}^{3} P(H=3|T=k, I=1)} \\
&= \frac{\frac{1}{2}}{\frac{1}{2}+1+0} \\
&= \frac{1}{3}
\end{aligned}$$

$$P(T=2|H=3, I=1) = 1 - P(T=1|H=3, I=1) = \frac{2}{3}.$$

$P(T=1|H=3, I=1) < P(T=2|H=3, I=1)$. Therefore, I will get greater chance to win the car if I switches to choose door 2.

2 Bayesian Network Conditional Independence (10 points)

From the definition of a Bayesian Network,

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

We use Bayes Ball rules for the questions:
1. False (go through 1)
2. False (go through 5)
3. True
4. False (3-4-1-2-5)
5. True
6. False (1-2-5-4-3)
7. True
8. True
9. False (3-4-5-2)
10. False (3-4-1-2)

3 Junction Tree Construction (5 points)

After Moralization and Triangulation, the graph is as in **Figure 1**. The constructed junction tree is as in **Figure 2**.

4 Junction Tree Algorithm (15 points)

The constructed junction tree is as in **Figure 3.**
In implementation we pick clique x_{n-1}, x_n as root. So we start with sending messages from x_1, x_2 to x_2, x_3.(First collect step). After we reached the root we start distribute operation by sending information from x_{n-1}, x_n to x_{n-2}, x_{n-1}. After all information is sent, we normalize the tables.

The joint probability distributions over clique's are given in the below tables. Also marginal distribution calculated from table over each variable is given. Note that common marginals match between tables. Also our seperator values ends up being equal to those marginals as expected.

$p(x_1, x_2)$	$x_2 = 0$	$x_2 = 1$	$p(x_1)$
x_1=0	0.040462	0.445087	0.485549
x_1=1	0.323699	0.190751	0.514451
p(x_2)	0.364162	0.635838	

$p(x_2, x_3)$	$x_3 = 0$	$x_3 = 1$	$p(x_2)$
$x_2 = 0$	0.260116	0.104046	0.364162
$x_2 = 1$	0.057803	0.578035	0.635838
p(x_3)	0.317919	0.682081	

$p(x_3, x_4)$	$x_4 = 0$	$x_4 = 1$	$p(x_3)$
$x_3 = 0$	0.119220	0.198699	0.317919
$x_3 = 1$	0.639451	0.042630	0.682081
p(x_4)	0.758671	0.241329	

$p(x_4, x_5)$	$x_5 = 0$	$x_5 = 1$	$p(x_4)$
$x_4 = 0$	0.569003	0.189668	0.758671
$x_4 = 1$	0.060332	0.180997	0.241329
p(x_5)	0.629335	0.370665	

Matlab Code

```
%Main function of hw5 problem 4
n = 5;
psis = cell(n-1, 1);

for i = 1:(n-1)
psis{i} = rand(2,2);
end
[ marginals ] = JCT4MarkovChain( psis );

p_test = cell(4,1);
p_test{1} = [0.1, 0.7; 0.8, 0.3];
p_test{2} = [0.5, 0.1; 0.1, 0.5];
p_test{3} = [0.1, 0.5; 0.5, 0.1];
p_test{4} = [0.9, 0.3; 0.1, 0.3];
[ m_test] = JCT4MarkovChain(p_test);
```

```
%Junction Tree Algorithm
function [ marginals ] = JCT4MarkovChain( potentials )
% potentials = cell of potentials
% marginals = output marginals

marginals = potentials;
n = size(marginals,1);
separators = ones(n-1,2);
%Forward
for i=1:n-1
separators(i,:) = sum(marginals{i});
marginals{i+1} = marginals{i+1}.*(separators(i,:)'*[1,1]);
end
%Backward
for i=1:n-1
s_old = separators(n-i,:);
separators(n-i,:) = sum(marginals{n-i+1},2)';
marginals{n-i} = marginals{n-i}.*([1;1]*(separators(n-i,:)./s_old));
end
%Normalize
for i=1:n
marginals{i} = marginals{i}/sum(sum(marginals{i}));
end

end
```

5 Hidden Markov Model (10 points)

We use the ArgMax Junction Tree Aalgorithm here. Run JTA but replace sums with max, then find biggest entry in separators. The most likely sequence of Mario's emotional states for the first five days is:

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

Matlab Code

```
function [ H ] = argMaxInfer( T, E, O, I )
% Input:
%   T = transition probabilities
%   E = emission probabilities
%   O = Observed states
%   I = initial probabilities
% Output:
%   H = the most likely hidden states

t = size(T, 1);
n = size(O, 2);
psi = zeros(t, t, n);
phi = zeros(t, n);
phi(:, 1) = I;
% forward
for i = 2 : n
```

```
k = O(1, i);
psi(:, :, i) = diag(phi(:, i - 1)) *T * diag(E(:,k));
phi(:, i) = max(psi(:, :, i));
end
% backward
for i = n - 1 : -1 : 1
phi_new = max(psi(:, :, i + 1), [], 2);
psi(:, :, i) = psi(:, :, i) * diag(phi_new ./ phi(:, i));
phi(:, i) = phi_new;
end
[¬, H] = max(phi);
end
```

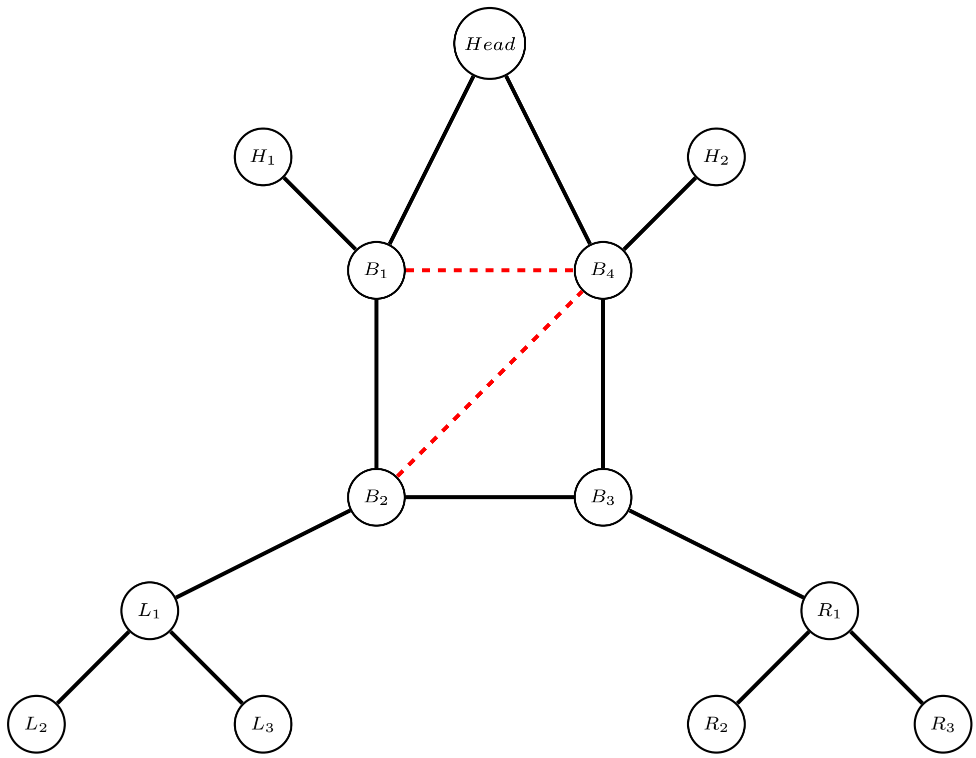

Figure 1: The graph after Moralization and Triangulation.

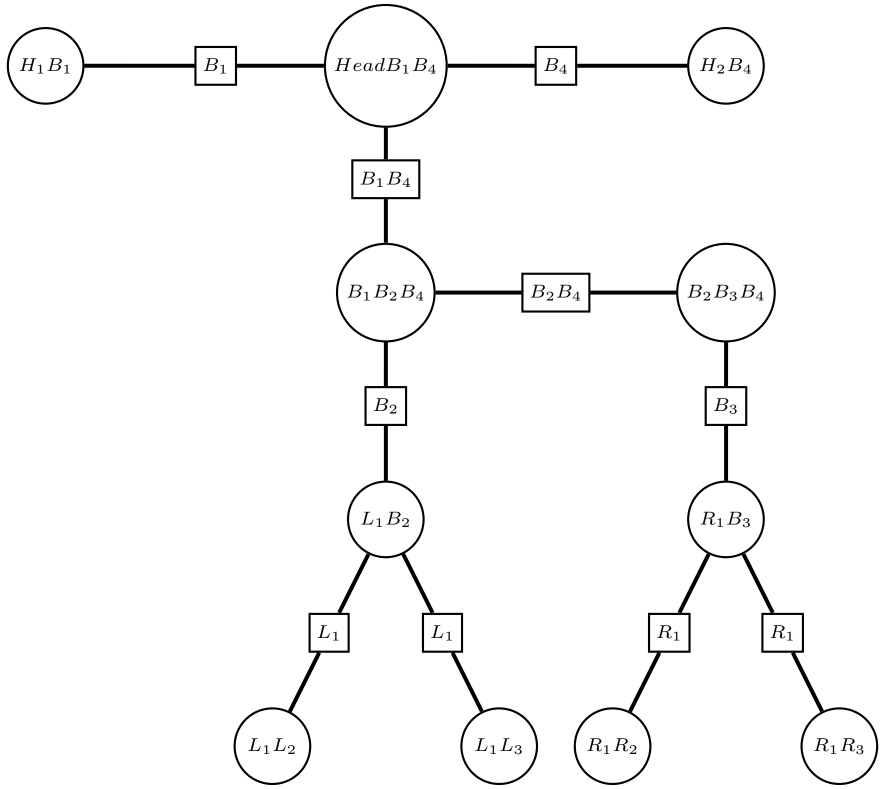

Figure 2: The constructed junction tree of graph with separators.

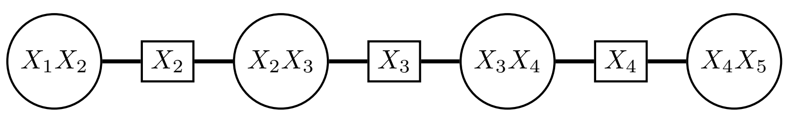

Figure 3: The constructed junction tree of graph with separators.