# CSE 473 Final Exam – June 5, 2019

## Name:

This exam is take home and is due on **Thursday June 13th at 11:59 pm**. You can submit directly to Gradescope. **Please write your answers directly on the PDF and submit a scanned copy.** Don't forget to also write your name above.

This exam should not take significantly longer than 3 hours to complete if you have already carefully studied all of course material. Studying while taking the exam may take longer. :)

This exam is open book and open notes, but you must complete all of the work yourself with no help from others. Please feel free to post clarification questions to the course message board, but please do not discuss solutions.

If you show your work and *briefly* describe your approach to the longer questions, we will happily give partially credit, where possible.

## Question 1 – True/False – 30 points

Circle the correct answer for each True / False question. If you think a question is ambiguous, please add a very short explanation of the interpretation you are making, and we will do our best to grade accordingly.

1. True / False – Adding edges to a Bayes Net can restrict the set of distributions it can represent. (3 pt)

2. True / False – For answering conditional queries in Bayesian networks, rejection sampling has generally been observed to provide worse estimates that likelihood weighting (when given the same number of samples). (3 pt)

3. True / False – For Q-learning to converge, the agent must eventually start acting in the world according to the optimal policy. (3 pt)

4. True / False – Approximate Q-learning with feature vectors will always converge to the optimal policy, if the agent visits all of the states a sufficient number of times. (3 pt)

5. True / False – $P(A|B,C) * P(B|C) * P(C) = P(A,B,C)$ only if we assume $A$ is conditionally independent of $C$ given $B$. (3 pt)

6. True / False – Inference by enumeration can produce incorrect results if the Bayes network is dense (has many edges). (3 pt)

7. True / False – $P(B|A,C) * P(C|A,B)/P(B,C) = P(A|B,C)$, given no independence assumptions(3 pt)

8. True / False – All Markov models have a stationary distribution. (3 pt)

9. True / False – The forward algorithm for HMMs has linear time complexity in the number of states. (3 pt)

10. True / False – The number of parameters in a Bayesian network grows exponentially with the highest in-degree (number of parents) of a node in the network. (3 pt)

## Question 2 – Short Answer – 30 points

These short answer questions can be answered with a few sentences each. Please be brief, we will subtract points for very long responses (e.g. more than a sentence or two for each part of the question).

1. Short Answer – Briefly describe how you would decide which algorithm to use for answering queries to a Bayesian network. What is the key property of the network that, if known, would best help you make the appropriate decision. (5 pts)

2. Short Answer – Briefly describe one heuristic that can be used to order variables when doing variable elimination for a Bayesian network, and explain when it would work well in practice. (5 pts)

3. Short Answer – Briefly describe the conditions that are required for Q-learning to converge to the optimal policy. (5 pts)

4. Short Answer – HMMs can be seen as a special type of Bayesian network. Briefly describe one way in which they differ from the more general case. (5 pts)

5. Short Answer – Briefly describe the pros and cons of using the forward algorithm vs. a particle filter for HMMs. When would you use each and why? (5 its)

6. Short Answer – Briefly describe the difference between outcomes and events in joint probability models. (5 its)

## Question 3 – Markov Models – 15 points

One of the traffic lights on University Ave has gone faulty, and while it can still display one of three colors (red, yellow, or green) at a time, it no longer does so in the order you would expect. One UW student has figured out that the light transitions according to a Markov model given by the following CPT:

| $C_t$ | $C_{t+1}$ | $P(C_{t+1}\|C_t)$ |
|---|---|---|
| Red | Red | 0.8 |
| Red | Yellow | 0.1 |
| Red | Green | 0.1 |
| Yellow | Red | 0.2 |
| Yellow | Yellow | 0.3 |
| Yellow | Green | 0.5 |
| Green | Red | 0.0 |
| Green | Yellow | 0.7 |
| Green | Green | 0.3 |

When this fault occurred the light reset, and it followed the initial distribution of $P(Red) = 0.5$, $P(Green) = 0.5$, and $P(Yellow) = 0.0$ (consider this to be $t = 0$).

1. What is the probability that the light produces the sequence Green, Yellow, Green, Yellow, Red? (5 pts)

2. Fill in the values of the probabilities below, using the forward algorithm for Markov models. (10pts)

| | $t = 0$ | $t = 1$ | $t = 2$ |
|---|---|---|---|
| $P(C_t = Red)$ | | | |
| $P(C_t = Yellow)$ | | | |
| $P(C_t = Green)$ | | | |

## Question 4 – Hidden Markov Models – 10 points

Continuing from Q3, now imagine that a driver with very poor eyesight is driving up the street and cannot discern the actual color of the light. However, they can make some assumptions based on whether the brightness of the light is bright or dark. Consider the following emission distribution:

| $C_t$ | $B_t$ | $P(B_t \mid C_t)$ |
|---|---|---|
| Red | Bright | 0.25 |
| Red | Dark | 0.75 |
| Yellow | Bright | 0.8 |
| Yellow | Dark | 0.2 |
| Green | Bright | 0.5 |
| Green | Dark | 0.5 |

For this problem, assume that the driver initially has the belief at time t = 0 that $B(Red) = 0.3$, $B(Yellow) = 0.3$, and $B(Green) = 0.4$ and that the world behaves as an HMM with emissions and transition defined as above.

1. If at time t = 1 we observe $B_1$ = Bright, and at time t = 2 we observe $B_2$ = Dark (there is no observation at time t=0). What is the driver's belief at time t = 2 after considering the most recent observation? (15 pts)

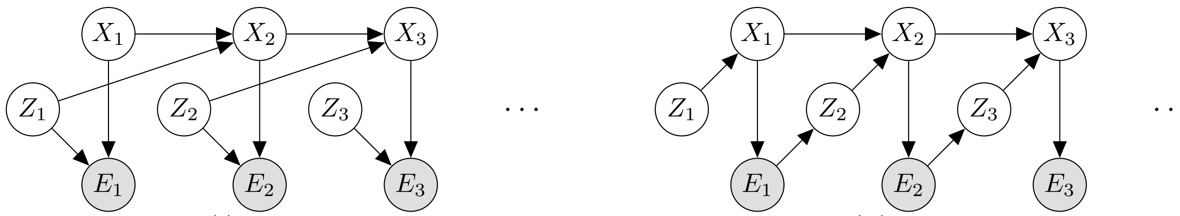
## Modified HMM Forward Algorithms – 20 pts

For a standard HMM the Elapse Time update and the Observation update are of the respective forms:

$$P(X_t \mid e_{1:t-1}) = \sum_{x_{t-1}} P(X_t \mid x_{t-1}) P(x_{t-1} \mid e_{1:t-1})$$

$$P(X_t \mid e_{1:t}) \propto P(X_t \mid e_{1:t-1}) P(e_t \mid x_t)$$

We now consider the following two HMM-like models:




(i) (ii)

For this problem, we want to do derive the forward algorithms (both elapse time and observation steps) for these Bayesian networks. More specifically, please show us the recursive updates for computing:

1. In model (i), $P(X_t, Z_t \mid e_{1:t-1}) = ?$ [5pts]

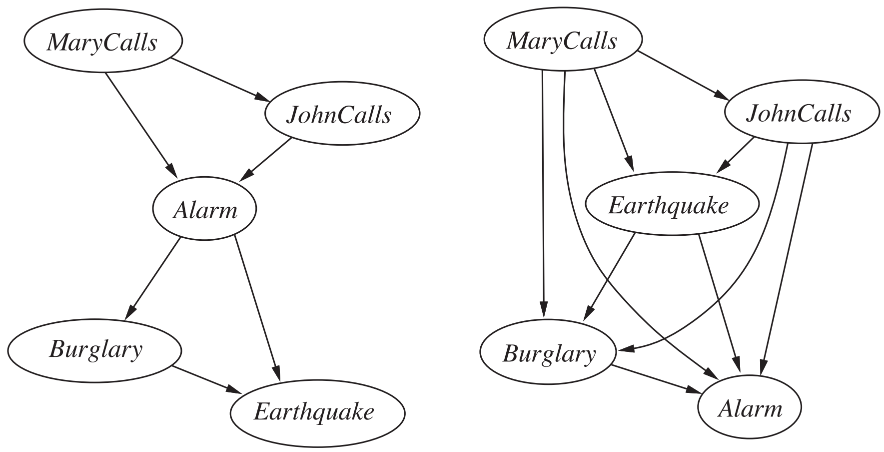
2. In model (i), $P(X_t, Z_t \mid e_{1:t}) \propto ?$ [5pts]

3. In model (ii), $P(X_t, Z_t \mid e_{1:t-1}) = ?$ [5pts]

4. In model (ii), $P(X_t, Z_t \mid e_{1:t}) \propto ?$ [5pts]

## Question 5 – Bayesian Networks – 35 points

Consider the following two Bayesian networks, which are variations on the alarm network we discussed in class:




(a) (b)

1. Based on the network structure alone, which network above makes the fewest independence assumptions? Briefly justify yoru answer. [3 pts]

2. Draw a new Bayesian network with the same set of random variables that makes as many independence assumptions as possible. [5 pts]

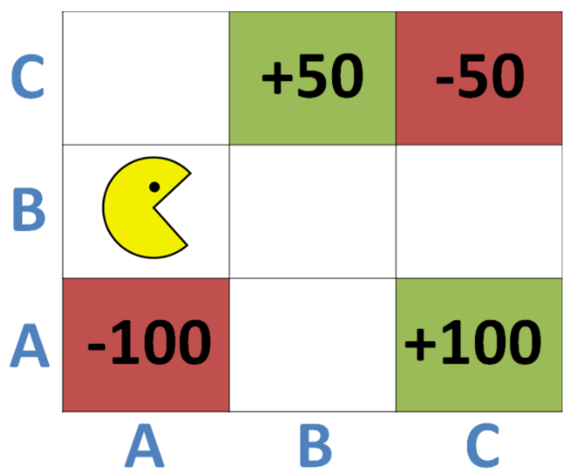
3. Write down two conditional independence assumptions encoded by the structure of network (a). If there are not two, write as many as possible. [6 pts]

4. Write down two conditional independence assumptions encoded by the structure of network (b). If there are not two, write as many as possible. [6 pts]

5. Simulate the execution of the variable elimination algorithm on network (a) to compute $P(Marycalls|Burglary = true)$. Since we have not given you the CPTs, you do not need to compute the entries. Instead, just list the tables that would be created and eliminated at each step of the computation. Use the most computationally efficient variable ordering. [15 pts]

## Question 6 – Reinforcement Learning – 15 points

Consider the grid-world given below and an agent who is trying to learn the optimal policy. States are named as (x-coordinate, y-coordinate) with horizontal axis x and vertical axis y, and the state after exiting is Done. Actions are North, South, East, West, and Exit denoted as N, S, E, W, and X for short. The Exit action can only be taken from shaded states, and Exit is the only action available in the shaded states. Rewards are only awarded for taking the Exit action from one of the shaded states. Taking this action moves the agent to the Done state, and the MDP terminates. Assume $\gamma = 1$ and $\alpha = 0.5$ for all calculations. In Q-Learning, all values are initialized to zero.




Now, assume the agent starts from (A, B) and observes the following sequence of episodes. Each step is a tuple containing $(s, a, s', r)$.

| | Episode 1 | Episode 2 | Episode 3 | Episode 4 | Episode 5 |
|---|---|---|---|---|---|
| **Step 1** | (A,B), N, (A,C), 0 | (A,B), E, (B,B), 0 | (A,B), E, (B,B), 0 | (A,B), E, (B,B), 0 | (A,B), E, (B,B), 0 |
| **Step 2** | (A,C), S, (A,B), 0 | (B,B), E, (C,B), 0 | (B,B), S, (B,A), 0 | (B,B), E, (C,B), 0 | (B,B), S, (B,A), 0 |
| **Step 3** | (A,B), N, (A,C), 0 | (C,B), N, (C,C), 0 | (B,A), E, (C,A), 0 | (C,B), W, (B,B), 0 | (B,A), N, (B,B), 0 |
| **Step 4** | (A,C), E, (B,C), 0 | (C,C), X, Done, -50 | (C,A), X, Done, +100 | (B,B), N, (B,C), 0 | (B,B), N, (B,C), 0 |
| **Step 5** | (B,C), X, Done, +50 | | | (B,C), X, Done, +50 | (B,C), X, Done, +50 |

1. Fill in the following Q-values obtained from direct evaluation from the samples? (10pt)

   (a) $Q((A, B), N) =$

   (b) $Q((B, B), E) =$

2. Which Q values are non-zero after running q-learning with the episodes above? (5pt)