CS 181 Artificial Intelligence (Fall 2018), Final Exam

Name (in Chinese): ______________________

ID#: ______________________

Instructions

- Time: 9–10:40am (100 minutes)
- This exam is closed-book, but you may bring an A4-size cheat sheet. Put all the study materials and electronic devices (with the exception of a calculator) into your bag and put your bag in the front, back, or sides of the classroom.
- You can write your answers in either English or Chinese.
- Two blank pieces of paper are attached, which you can use as scratch paper. Raise your hand if you need more paper.

1 Multiple choice (10 pt)

1.1 Each question has only one correct answer

Please write your answers in the table below.

1	2	3	4	5

1. Consider an HMM with state variables $\{X_T\}$ and evidence variables $\{E_T\}$. Which of the following equations is correct?

   A. $P(e_{t+1}, e_{1:t}|X_{t+1}) = P(e_{t+1}|X_{t+1})P(X_{t+1}|e_{1:t})$

   B. $P(X_{t+1}|e_{1:t+1}) = P(e_{t+1}|X_{t+1}, e_{1:t})P(X_{t+1}|e_{1:t})$

   C. $P(X_{t+1}|e_{1:t}) = \sum_{X_t} P(X_t|e_{1:t})P(X_{t+1}|X_t, e_{1:t})$

   D. $P(X_{t+1}, e_{t+1}|X_t) = P(X_{t+1}|X_t)P(e_{t+1}|X_t)$

   E. None of the above.

2. Consider an MDP with the set of states $S = \{s_i \mid 1 \leq i \leq N,\ i \in \mathbf{Z}^+\}$ and the set of actions $A = \{a_i \mid 1 \leq i \leq N,\ i \in \mathbf{Z}^+\}$. $N \in \mathbf{Z}^+$ is a finite constant. Which of the following transition function is valid? A transition tuple (s_i, a_j, s_k) means state s_i transits to state s_k after taking action a_j.

   A. $T(s_i, a_j, s_k) = \begin{cases} 0.5, & \text{if } j = k, \\ 0.5, & \text{otherwise.} \end{cases}$

   B. $T(s_i, a_j, s_k) = \begin{cases} 0, & \text{if } j = k, \\ 1, & \text{otherwise.} \end{cases}$

   C. $T(s_i, a_j, s_k) = \begin{cases} 1, & \text{if } i = k, \\ 0, & \text{otherwise.} \end{cases}$

   D. $T(s_i, a_j, s_k) = \begin{cases} \frac{1}{N}, & \text{if } i = k, \\ 0, & \text{otherwise} \end{cases}$

   E. None of the above

3. Which of the following statements is correct?

   A. In TD-Learning, when the learning rate α is 0.5, it is equivalent to taking the average of all the previous samples.

   B. Q-learning can converge to the optimal policy even if you are acting sub-optimally when you explore.

   C. When applying ϵ-greedy learning strategy, discounting ϵ faster results in lower overall regret

   D. When applying ϵ-greedy learning strategy, discounting ϵ slower results in lower overall regret

4. Which of the following statements about EM (Expectation Maximization) is WRONG?

   A. We maximize the marginal likelihood when we use EM to learn a GMM (Gaussian mixture model).

   B. EM can be used to learn models with hidden variables.

   C. When using EM to learn a GMM, we assign each data point to one Gaussian distribution during the E-step.

   D. K-means is a special case of EM for GMM.

5. How many components are there in a context-free grammar?

A. 2
B. 3
C. 4
D. 5

Solution:

1. C
2. C
3. B
4. C
5. C

1.2 Each question has one or more correct answers

Please write your answers in the table below.

1	2	3	4	5

1. Consider an HMM with state variables $\{X_T\}$ and evidence variables $\{E_T\}$. Which of the following statements is/are correct?

   A. An HMM is a special case of a Dynamic Bayes Net.
   B. Given the evidence variables, the state variables are conditionally independent.
   C. The time complexity and space complexity of computing $P(X_t|e_{1:t})$ by using the forward algorithm are both $O(t|X|)$.
   D. The time complexities of the forward algorithm and the Viterbi algorithm are the same.
   E. None of the above.

2. Which of the following statements of particle filtering is/are correct?

   A. Each iteration of particle filtering consists of *propagate forward*, *observe* and *resample*.
   B. Particle filtering is equivalent to *likelihood weighting*.
   C. In *propagate forward*, we sample the next state of each particle based on the transition model.
   D. In particle filtering, the evidence variables have no influence on the set of sampled particles.
   E. None of the above.

3. Which of the following statements is/are correct? Recall that the Bellman Equation of MDP is: $V(s) = \max_a \sum_{s'} T(s,a,s')[R(s,a,s') + \gamma V(s')]$, and S, A denote the sizes of the state space and action space.

   A. $V(s)$ is the undiscounted expected utility of state s when $\gamma = 0$.
   B. Setting $\gamma < 1$ helps Value Iteration converge.
   C. Value Iteration is guaranteed to converge to the optimal values in $O(S^2A)$.
   D. Policy Iteration converges to the optimal policy much faster than Value Iteration under some conditions.
   E. None of the above.

4. Which of the following statements about regression is/are correct?

   A. The output is discrete in regression.

   B. The output is real-valued in regression.

   C. We aim to minimize the error over the training set (possibly with regularization) during training.

   D. We aim to minimize the error over the testing set (possibly with regularization) during testing.

   E. None of the above.

5. Which of the following statements about k-means is/are correct?

   A. The time complexity is $O(tkn)$, where n is the number of data points, k is the number of clusters, and t is the number of iterations.

   B. K-means may perform poorly when there are noisy data points very far away from the other data points.

   C. The results of k-means can be different given different initializations.

   D. K-means minimizes the total Euclidean distances between data points and their cluster centers.

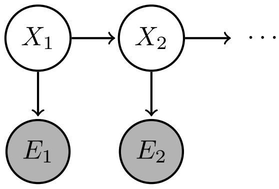
   E. None of the above.

Solution:

1. AD
2. AC
3. BD
4. BC
5. ABC

2 Most likely explanation of HMM (10 pt)

Consider an HMM with state variables $\{X_T\}$ and evidence variables $\{E_T\}$. For each time step, the state X_t can take one of the three values $\{A, B, C\}$.



2.1 Most likely explanation (4 pt)

The transition probabilities, emission probabilities and initial distribution are given in the following tables.

X_t	X_{t+1}	$P(X_{t+1} \mid X_t)$
A	A	0
A	B	0.5
A	C	0.5
B	A	0.8
B	B	0
B	C	0.2
C	A	0.4
C	B	0.4
C	C	0.2

X_t	E_t	$P(E_t \mid X_t)$
A	α	0.2
A	β	0.8
B	α	0.5
B	β	0.5
C	α	0.25
C	β	0.75

X_1	$P(X_1)$
A	0.2
B	0.4
C	0.4

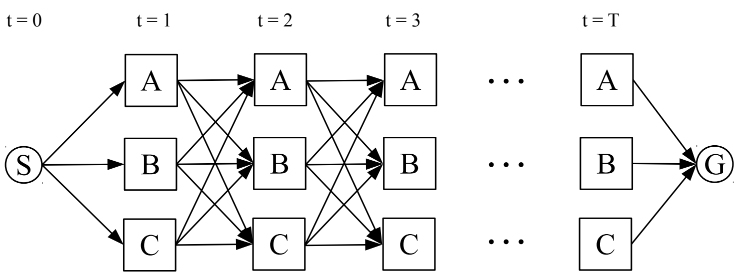
Please compute the most likely explanation of $E_1 = \alpha, E_2 = \alpha, E_3 = \beta$.

$\arg\max_{X_1,X_2,X_3} P(X_1, X_2, X_3 | E_1 = \alpha, E_2 = \alpha, E_3 = \beta)$:____________________

$\max_{X_1,X_2,X_3} P(X_1, X_2, X_3, E_1 = \alpha, E_2 = \alpha, E_3 = \beta)$:____________________

2.2 Shortest path and most likely explanation (6 pt)

We can represent the state transitions through the following directed graph, also called a *Trellis Diagram*.



We wish to formulate the most likely explanation problem as finding the shortest path between S and G. Assume that the length of a path is the sum of the length of all the edges along the path. You need to specify the length of each edge in the diagram in terms of the transition probabilities $P(X_{t+1}|X_t)$, the emission probabilities $P(E_t|X_t)$ and the initial probabilities $P(X_1)$. The values of the evidence variables are $e_1, ..., e_T$.

Definition: Let $w^t_{Y \to Z}$ be the length of the edge for the transition from state Y at time t to state Z at time $t+1$.

Please specify the following lengths. If the length can be any constant number, you should write **c**.

$w^0_{S \to Y}$: ____________________

$w^t_{Y \to Z}$ $(1 \leq t < T)$: ____________________

$w^T_{Y \to G}$: ____________________

Solution:

(1) $arg\max_{X_1,X_2,X_3} P(X_1, X_2, X_3|E_1 = \alpha, E_2 = \alpha, E_3 = \beta) = C, B, A$
$\max_{X_1,X_2,X_3} P(X_1, X_2, X_3, E_1 = \alpha, E_2 = \alpha, E_3 = \beta) = 0.0128$

(2) $-\log P(X_1 = Y)P(E_1 = e_1|X_1 = Y)$,

$-\log P(X_{t+1} = Z|X_t = Y)P(E_{t+1} = e_{t+1}|X_{t+1} = Z)$,

c

(Maybe some other answers.)

Such as

$-\log P(X_1 = Y)$,

$-\log P(X_{t+1} = Z|X_t = Y)P(E_t = e_t|X_t = Y)$,

$-\log P(E_T = e_T|X_T = Y)$

3 Reinforcement Learning (10 pt)

3.1 MDP and value iteration (5 pt)

Consider an MDP with the set of states $S = \{s_i \mid 1 \le i \le N,\ i \in \mathbf{Z}^+\}$ and the set of actions $A = \{a_i \mid 1 \le i \le N,\ i \in \mathbf{Z}^+\}$. $N \in \mathbf{Z}^+$ is a finite constant. Suppose the transition function is

$$T(s_i, a_j, s_k) = \begin{cases} 1, & \text{if } j = k, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Suppose the reward function is

$$R(s_i, a_j, s_k) = \begin{cases} i, & \text{if } j = k, \\ 0, & \text{otherwise} \end{cases}$$

Please use a simple expression to represent the optimal policy π_i for each state.

$\pi_i =$ ________________________

Solution:

$\pi_i = a_N$

(b) Let $N = 5$. Suppose the reward function is

$$R(s_i, a_j, s_k) = \begin{cases} i, & \text{if } i > j, j = k, \\ -i, & \text{if } i < j, j = k, \\ 0, & \text{otherwise} \end{cases}$$

Please apply the Value Iteration Algorithm described in the lecture to fill the first row of the following table.

V_1					
V_0	0	0	0	0	0
	s_1	s_2	s_3	s_4	s_5

Solution:

V_1	0	2	3	4	5
V_0	0	0	0	0	0
	s_1	s_2	s_3	s_4	s_5

(c) Under the same setting of (b), suppose we continue iterating and get the values of V_2, as shown below.

V_2	$\underline{V_2(s_1)}$	$\underline{V_2(s_2)}$	$\underline{V_2(s_3)}$	$\underline{V_2(s_4)}$	$\underline{V_2(s_5)}$
V_1	...	...	...	...	...
V_0	0	0	0	0	0
	s_1	s_2	s_3	s_4	s_5

Please answer the following questions:

If $V_2(s_1) = 0$, the range of discounting rate γ is: ____________________

If $V_2(s_2) \neq 2$, the range of discounting rate γ is: ____________________

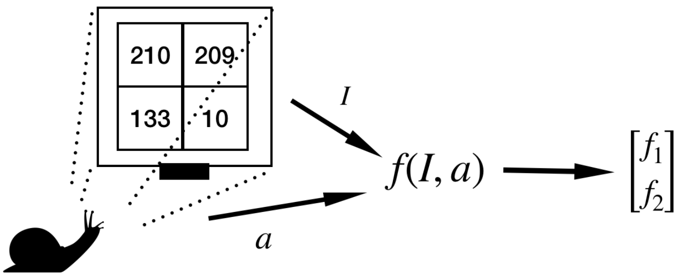
Solution:

$\gamma \leq 0.2$
$1 \geq \gamma > 0.8$

3.2 Mr. Snail's Video Game (5 pt)

Mr. Snail bought a tiny video game but he is not good at playing it. So he wants to train an agent to play it automatically. He does not know the MDP underlying the game, and the state space is too large for traditional methods, so he decides to use approximate Q-learning with linear Q-functions. The video game's resolution is $\mathbf{2} \times \mathbf{2}$, which means a screen-shot (frame) of this video game is a 2×2 image, containing only 4 pixels (big enough for Mr. Snail!). Each pixel is an integer ranging from 0 to 255. The agent takes the image input of the video game as state s (denote this 2×2 image as I) and it can perform one of three actions $\{a_1, a_2, a_3\}$ at each time step.

To extract features, Mr. Snail uses a pretrained fixed feature extractor $f(I, a)$ which takes image I and action a as input and outputs a feature vector containing two features $[f_1, f_2]$. After feature extraction, the approximate Q-value can be calculated as $Q(s, a) = w_1 f_1 + w_2 f_2$



(a) If we use traditional Q-leaning, the number of Q-table entries ($Q(s, a)$) is:

____________________ (you can use power and product in your answer)

Solution:

3×256^4

(b) If we use the following feature extractor $f(I, a)$:

$f(I, a_1)$	$I \cdot [1, 0]^T$
$f(I, a_2)$	$I \cdot [1, -1]^T$
$f(I, a_3)$	$I \cdot [0.5, 0.5]^T$

Suppose the current weights are $w_1 = 1.0$ and $w_2 = -3.0$. The agent is at state (image) I, takes action a_2, and gets to the next state I' with reward $r = -120$.

$$I = \begin{bmatrix} 100 & 250 \\ 30 & 80 \end{bmatrix} \qquad I' = \begin{bmatrix} 150 & 200 \\ 30 & 10 \end{bmatrix}$$

Assume the discount factor $\gamma = 1$ (no discount) and learning rate $\alpha = 0.1$. What are the updated values of w_1 and w_2?

w_1 ______________________________

w_2 ______________________________

Solution:

sample $= r + \gamma \cdot max_{a'} Q(I', a')$

$f(I', a_1) = [150, 30]^T, \quad Q(I', a_1) = 60$
$f(I', a_2) = [-50, 20]^T, \quad Q(I', a_2) = -110$
$f(I', a_3) = [175, 20]^T \quad Q(I', a_3) = 115$

sample $= -120 + 115 = -5$
$f(I, a_2) = [-150, -50]^T, Q(I, a_2) = 0$
difference = sample - $Q(I, a_2) = -5$
$f_1(I, a_2) = -150, \quad f_2(I, a_2) = -50,$
$w_1' = w_1 + 0.1 * -5 * -150 = 1 + 75 = 76$
$w_2' = w_2 + 0.1 * -5 * -50 = -3 + 25 = 22$

4 Naïve Bayes (10 pt)

Bob wants to borrow money from a bank. The table below shows the historical data of borrower information and whether each borrower paid back money in time. Learn a naïve Bayes classifier (without smoothing) from the table and use it to judge whether Bob (*Home owner: No, Marital status: Married, Job experience: 3*) can pay back the money in time or not. You must show how you get your prediction based on the probabilities in the naïve Bayes that you learn from the data.

Home owner	Marital Status	Job experience	Pay back in time
Yes	Single	3	Yes
No	Married	4	Yes
No	Single	5	Yes
Yes	Married	4	Yes
No	Divorced	2	No
No	Married	4	Yes
Yes	Divorced	2	Yes
No	Married	3	No
No	Married	3	Yes
Yes	Single	2	No

Solution:

$$P(\mathrm{Y} = \mathrm{Yes}) = \frac{7}{10}$$
$$P(\text{Home owner} = \text{No}|\mathrm{Y} = \text{Yes}) = \frac{4}{7}$$
$$P(\text{Marital status} = \text{Married}|\mathrm{Y} = \text{Yes}) = \frac{4}{7}$$
$$P(\text{Job experience} = 3|\mathrm{Y} = \text{Yes}) = \frac{2}{7}$$
$$P(\text{Bob can pay back in time}) = \frac{7}{10} \times \frac{4}{7} \times \frac{4}{7} \times \frac{2}{7} = 0.065$$

$$P(\mathrm{Y} = \mathrm{No}) = \frac{3}{10}$$
$$P(\text{Home owner} = \text{No}|\mathrm{Y} = \text{No}) = \frac{2}{3}$$
$$P(\text{Marital status} = \text{Married}|\mathrm{Y} = \text{No}) = \frac{1}{3}$$
$$P(\text{Job experience} = 3|\mathrm{Y} = \text{No}) = \frac{1}{3}$$
$$P(\text{Bob can NOT pay back in time}) = \frac{3}{10} \times \frac{2}{3} \times \frac{1}{3} \times \frac{1}{3} = 0.022$$

Bob can pay back in time.

5 Parsing (10 pt)

Consider the following probabilistic context-free grammar:

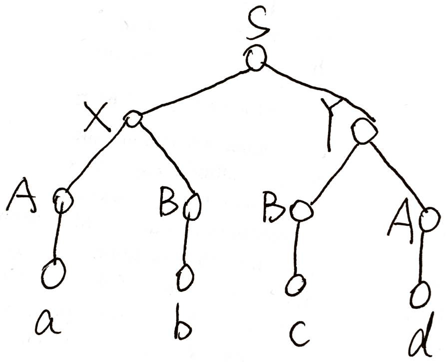
S → X Y [1.0]	
X → A B [0.6]	X → B C [0.4]
Y → B A [0.5]	Y → B B [0.5]
A → a [0.4]	A → d [0.6]
B → a [0.2]	B → b [0.6]
B → c [0.2]	C → b [1.0]

Given the sentence "*a b c d*", fill the parse table of probabilistic CYK and draw the best parse tree.

	1	2	3	4
0				
1				
2				
3				

Solution:

	1	2	3	4
0	A[0.4] B[0.2]	X[0.144]		S[0.00864]
1		B[0.6] C[1.0]	Y[0.06]	
2			B[0.2]	Y[0.06]
3				A[0.6]


S
X
Y
A
B
B
A
a
b
c
d