CS 4100: Foundations of Artificial Intelligence

Midterm 2

4/9/2020

- Exam start time: 11:45 AM. Exam finish time: 1:45 PM.
- Please upload your answers as a PDF file to your git folder, no later than 1:45 PM.
- The exam is closed book, closed notes except a one-page crib sheet.
- Please use non-programmable calculators only.
- If you are not sure of your answer you may wish to provide a brief explanation. All short answer sections can be successfully answered in a few sentences *at most*.

First Name	Zhiyuan
Last Name	Cao
NEU ID (optional)	001693668

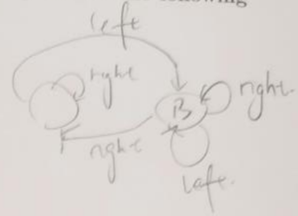
For staff use only

Q. 1	Q. 2	Q. 3	Q. 4	Q. 5	Q. 6	Q. 7	Total
/14	/10	/14	/ 16	/ 20	/14	/12	/100

1 Reinforcement Learning (14 points)

Consider the following game, which has only two states A, B and two actions the agent can choose from: Right, Left. Suppose an agent plays one episode in the game, and generates the following sequence of actions and rewards:

t	s_t	a_t	s_{t+1}	r_t
0	A	Left	B	2
1	B	Left	B	-4
2	B	Right	B	4
3	B	Right	A	3
4	A	Right	A	-1



Assume a discount factor of $\gamma = 0.5$ and a learning rate $\alpha = 0.5$, and initialize all values to 0.

1. **(4 pt)** What are the following Q-values learned by running Q-learning with the above experience sequence?

   2 + 0.5 · 0 $Q(A, Left) =$ <u>1</u>, $Q(B, Right) =$ <u>2.75</u>

   2   -1   2+0.5[3+0.5(2)-2] =

2. **(4 pt)** In model-based reinforcement learning, we first estimate the transition function $T(s, a, s')$ and the reward function $R(s, a, s')$. Fill in the following estimates of T and R, estimated from the experience above. Write 'N/A' if not applicable or undefined.

   $\hat{T}(A, Right, A) =$ <u>1</u>, $\hat{R}(A, Right, A) =$ <u>-1</u>

   $\hat{T}(A, Right, B) =$ <u>N/A</u>, $\hat{R}(A, Right, B) =$ <u>N/A</u>

   $\hat{T}(B, Right, A) =$ <u>0.5</u>, $\hat{R}(B, Right, A) =$ <u>3</u>

   $\hat{T}(B, Right, B) =$ <u>0.5</u>, $\hat{R}(B, Right, B) =$ <u>4</u>

3. Assume we had a **different experience** that ended up with the following estimates of the transition and reward functions:

s	a	s'	$\hat{T}(s,a,s')$	$\hat{R}(s,a,s')$
A	Right	A	1	12
A	Left	A	0.5	1
A	Left	B	0.5	1
B	Right	A	1	-5
B	Left	B	1	10

   (a) **(4 pt)** Give the optimal policy $\hat{\pi}^*(s)$ and $\hat{V}^*(s)$ for the MDP with transition function $\hat{T}$ and reward function $\hat{R}$.
   Hint: find the optimal policy first, and then use Bellman equation to calculate the value function.

   $\hat{\pi}^*(A) =$ <u>Right</u>, $\hat{\pi}^*(B) =$ <u>Left</u>, $\hat{V}^*(A) =$ <u>12</u>, $\hat{V}^*(B) =$ <u>10</u>

(b) **(2 pt)** If we repeatedly feed this experience sequence into a TD-value learning algorithm, what values will it converge to? Assume that the learning rate α is properly chosen so that convergence is guaranteed.

i. the values found above, $\hat{V}^*$

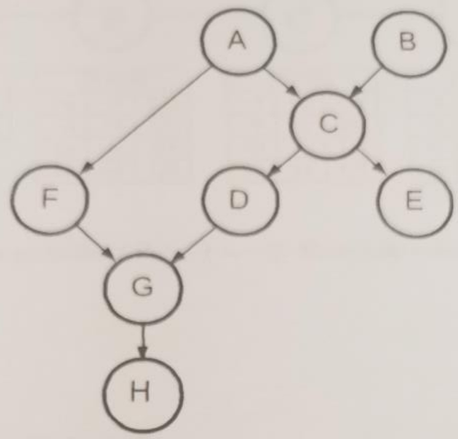
ii. the optimal values, V^*

iii. neither $\hat{V}^*$ nor V^*

iv. not enough information to determine

2 Bayesian Networks and D-separation (10 points)

Consider the following Bayes' net.

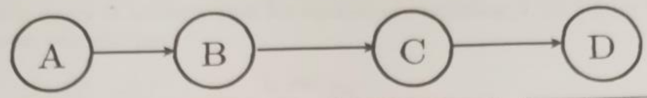


For each expression, circle to indicate whether it is True or False. If False, provide one active path.

- ○ True ○ False It is guaranteed that $A \perp\!\!\!\perp B$
- ○ True ○ False It is guaranteed that $A \perp\!\!\!\perp D | C$
- ○ True ○ False It is guaranteed that $F \perp\!\!\!\perp C | H$ F → G → H , C → D → H
- ○ True ○ False It is guaranteed that $F \perp\!\!\!\perp C | A$
- ○ True ○ False It is guaranteed that $F \perp\!\!\!\perp E | A, G$ F → G A E → E → D → G

3 Inference in Bayesian Networks (14 points)

Assume the following Bayes' net, and the corresponding distributions over its variables:



$P(A)$	
$-a$	3/4
$+a$	1/4

$P(B \mid A)$		
$-a$	$-b$	2/3
$-a$	$+b$	1/3
$+a$	$-b$	4/5
$+a$	$+b$	1/5

$P(C \mid B)$		
$-b$	$-c$	1/4
$-b$	$+c$	3/4
$+b$	$-c$	1/2
$+b$	$+c$	1/2

$P(D \mid C)$		
$-c$	$-d$	1/8
$-c$	$+d$	7/8
$+c$	$-d$	5/6
$+c$	$+d$	1/6

1. **(4 pt)** Compute the probability $P(+c \mid +a, -d)$. Show your work.

2. You are given the following samples:

$+a$	$+b$	$-c$	$-d$
$+a$	$-b$	$+c$	$-d$
$-a$	$+b$	$+c$	$-d$
$-a$	$-b$	$+c$	$-d$

$+a$	$-b$	$-c$	$+d$
$+a$	$+b$	$+c$	$-d$
$-a$	$+b$	$-c$	$+d$
$-a$	$-b$	$+c$	$-d$

(a) **(2 pt)** Assume that these samples came from performing Prior Sampling, and calculate the sample estimate of $P(+a, +c)$.

2/8

(b) **(2 pt)** Now we will estimate $P(+c \mid +a, -d)$. Above, clearly cross out the samples that would **not** be used when doing Rejection Sampling for this task, and write down the sample estimate of $P(+c \mid +a, -d)$ below.

2/3

3. **(2 pt)** Using Likelihood Weighting Sampling to estimate $P(-a \mid +b, -d)$, the following samples were obtained. Fill in the weight of each sample in the corresponding row.

Sample				**Weight**
$-a$	$+b$	$+c$	$-d$	P(b|-a) P(-d|c) = 0.277
$+a$	$+b$	$+c$	$-d$	0.17
$+a$	$+b$	$-c$	$-d$	0.025
$-a$	$+b$	$-c$	$-d$	0.042

4. **(2 pt)** From the weighted samples in the previous question, estimate $P(-a \mid +b, -d)$.

$(5/18 + 1/24) / (5/18 + 5/30 + 1/40 + 1/24) = 0.625$

5. **(2 pt)** Which query is better suited for likelihood weighting, $P(D \mid A)$ or $P(A \mid D)$? Justify your answer in one sentence.

P(D | A) is better.

Since the order is $A \to D$

0.5
$0.57 + 0.28 = 0.35$

4 HMM (16 points)

X_1	$\Pr(X_1)$
0	0.3
1	0.7

X_t	X_{t+1}	$\Pr(X_{t+1}\|X_t)$
0	0	0.4
0	1	0.6
1	0	0.8
1	1	0.2

X_t	O_t	$\Pr(O_t\|X_t)$
0	A	0.9
0	B	0.1
1	A	0.5
1	B	0.5

1. **(8 pt)** Use HMM filtering to compute the probability distribution $P(X_2, \underbrace{O_1 = A}_{0.8}, \underbrace{O_2 = B}_{0.5})$. Show your work.

$P(X_2, O_1=A, O_2=B)$
$= P(O_2=B \mid X_2) \sum_{X_1} P(X_2 \mid X_1) P(X_1, O_1=A)$
$= \langle 0.5 \;\; 0.8 \rangle ((\langle 0.2, 0.8 \rangle \cdot 0.5 \cdot 0.7) + (\langle 0.6, 0.4 \rangle \cdot 0.9 \cdot 0.25))$
$= \langle 0.5 \cdot ((0.2 \cdot 0.5 \cdot 0.7) + 0.8 \cdot 0.9 \cdot 0.25), \;\; 0.9 \cdot (0.8 \cdot 0.5 \cdot 0.7 + 0.4 + 0.4 \cdot 0.9)$
$= \langle 0.125, 0.75 \rangle$

2. **(8 pt)** Compute the probability $P(X_1 = 1 | O_1 = A, O_2 = B)$. Show your work.

$P(X_1 \mid O_1 = A, O_2 = B)$

$= \dfrac{P(X_1, O_1=A, O_2=B)}{P(O_1=A, O_2=B)}$

$= \dfrac{\sum P(X_1=1) P(X_2 \mid X_1) P(O=A \mid X_1=1) \cdot P(O_2 \mid B=X_2)}{P(X=1, O=A, O_2=B) + \sum_{X_2} P(X_1=0) P(X_2 \mid X_1=0)}$

$= \dfrac{0.028 + 0.035}{0.028 + 0.035 + 0.0108 + 0.189}$

$= \dfrac{63}{2628} = \dfrac{21}{876} = \dfrac{7}{292}$

5 Classification (20 points)

Sample	IsHeavy	IsSmelly	IsSpotted	IsSmooth	IsPoisonous
A	0	0	0	0	0
B	0	0	1	0	0
C	1	1	0	1	0
D	1	0	0	1	1
E	0	1	1	0	1
F	0	0	1	1	1
G	0	0	0	1	1
H	1	1	0	0	1
X	1	1	1	1	?
Y	0	1	0	1	?
Z	1	1	0	0	?

Assume you and your friends are on an island. To survive you must eat mushrooms indigenous to the island. Three people try mushrooms A, B, and C, and they are fine. Five of your friends try mushrooms D, E, F, G, and H, and they are ill. Now you have the training data of A-H, where you know whether they are poisonous, but you do not know about X, Y, and Z.

1. **(2 pt)** What is the entropy of IsPoisonous? $= -5/8 \log 5/8 - 3/8 \log 3/8$

   $H(\text{IsPoisonous}) = 0.9544$

2. **(5 pt)** Which attribute will you choose as the root of a decision tree? What is the resulting information gain from choosing this attribute?

   isSmooth

   $IG(\text{IsPoisonous} \mid \text{IsSmooth})$

   $= 0.9544 - (0.8113 \times 0.5 + 1.0 \times 0.5)$

   $= 0.0487$

3. **(5 pt)** Build a decision tree classifier from the training data, and classify X, Y, and Z. No need to show the computation process. You can just draw the tree. Hint: you should be able to find out the construction by eyeballing the data points.

   IsSmooth

   1 / \ 0

   IsSmelly   IsSmelly

   1 / \ 0   1 / \ 0

   0   1   1   0

ISP=

4. **(8 pt)** Build a Naïve Bayes classifier from the training data, and classify X, Y, and Z. Do not use Laplace smoothing. Show your work.

X: $P(\text{poisonous} \mid 1, 1, 1, 1) = \frac{2}{5} \cdot \frac{2}{5} \cdot \frac{2}{5} \cdot \frac{3}{5} \cdot \frac{5}{8} = 0.024$

$P(-\text{poisonous} \mid 1\ 1\ 1\ 1) = 0.004$

Y: $P(\text{poisonous} \mid 1, 1, 0, 0) = \frac{3}{5} \cdot \frac{3}{5} \cdot \frac{6}{8} = 0.5$

$P(-\text{poi} \mid \ ) = 0.012$

Z: $P(p \mid 1\ 1\ 0\ 0) = \frac{2}{5} \cdot \frac{3}{5} \cdot \frac{5}{8} = 0.1$

$P(-p \mid \ ) = 0.012$

Naive Bayes classifier

$P(\text{IsHairy} \mid \text{IsPoisonous}) = \frac{2}{5}$

$P(\text{IsHairy} \mid -\text{IsP}) = \frac{1}{3}$

$P(\text{IsSpotted} \mid \text{IsP}) = \frac{2}{5}$

$P(\text{IsSpotted} \mid -\text{IsP}) = \frac{1}{3}$

$P(\text{IsSmelly} \mid +\text{IsP}) = \frac{2}{5}$

$P(-\text{IsSmelly} \mid -\text{IsP}) = \frac{1}{3}$

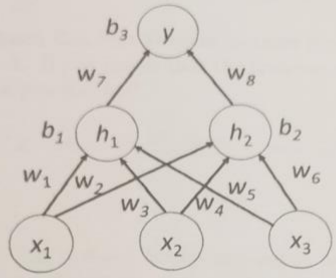
$P(\text{IsSmooth} \mid +\text{IsP}) = \frac{3}{5}$

$P(\text{IsSmooth} \mid -\text{IsP}) = \frac{1}{3}$

6 Neural Networks (14 points)

Your task is to design a multi-layer perceptron which receives three binary valued (i.e., 0 or 1) inputs x_1, x_2, x_3 and outputs 1 if *exactly* two of the inputs are 1, and outputs 0 otherwise. All of the units use a hard threshold activation function:

$$f(z) = \begin{cases} 1 & \text{if } z \geq 0 \\ 0 & \text{if } z < 0 \end{cases}$$



1. **(11 pt)** Specify the weights and biases which correctly implement this function.

   $w_1 =$ 1, $w_2 =$ 1, $w_3 =$ 1, $w_4 =$ 1, $w_5 =$ 1, $w_6 =$ 1

   $w_7 =$ -1, $w_8 =$ 1, $b_1 =$ -2.5 $b_2 =$ -1.5 $b_3 =$ -0.5

2. **(3 pt)** Briefly explain how this MLP works.

From ~~input~~ to Hidden Layer, we set all weights to 1. input and h_1 and h_2 are ~~two~~ sums of inputs.

But we should let sum=2 become the label.

Other value are 0, 1, 3. so b_1 and b_2 are threshold to control it. In the correct way $h_1 = -0.5$ and $h_2 = 0.5$ then $h_1 + h_2 = 0.5 + bias = 0$ other will get small than zero

7 Short Answers (12 points)

1. **(4 pt)** Briefly explain the problem of overfitting. Name two specific ways that try to deal with this problem.

   Overfitting is too many features in the model and the model just apply one only data set

   Salution: ① Regalarzation ② Dropout layers

2. **(2 pt)** Suppose you use Batch Gradient Descent to train a neural network and you plot the training error at every epoch. If you notice that the training error consistently goes up, what is likely going on? How can you fix this?

   Lerny rate has problem, we should adjust it

3. **(2 pt)** ○ True ● False In general, for Q-learning to converge to the optimal Q-values, it is necessary that actions get chosen according to $\text{argmax}_a Q(s, a)$.

4. **(2 pt)** ○ True ● False The number of parameters in a Bayesian network is exponential in the number of arcs in the graph.

5. **(2 pt)** Suppose you have a multi-class classification problem with three classes, and a small labeled data set (about 500 examples). Which algorithm you should choose for the classification, provided that you care more about the prediction time than the accuracy of the model?

   (a) Neural network

   (b) Decision tree

   (c) Logistic regression

   (d) KNN

   (e) Random forest