CS47100 Assignment 4

Due date: Wednesday December 7, 2022 (11:59pm)

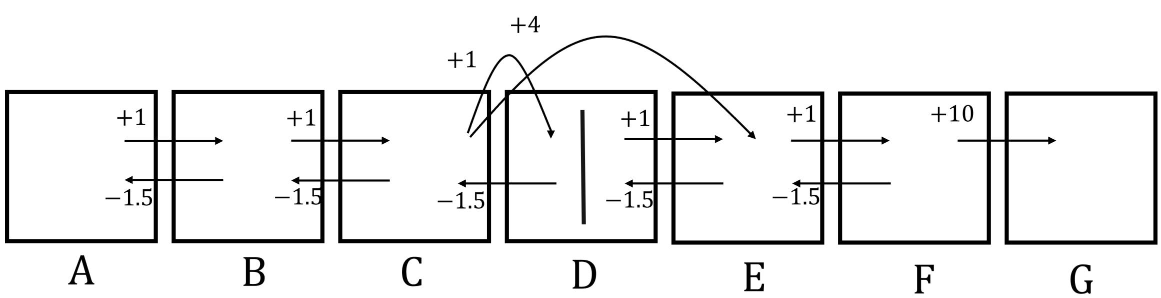
This assignment will involve written and programming exercises.

Part 1: Written Assignment (40 pts + 10 pts Bonus)

MDP (24 pts)

1. Considering an MDP problem modeling race track shown as below. There is a single hurdle in square D. The terminal state is G (i.e., $V(G) = 0$). The agent can either move to *left* (if there is a square on its left) or *right* (if there is a square on its right), except for when the agent is in square C—When the agent is in square C, it can *not* move *right* into the hurdle square D. Instead, the agent can choose to *jump* in square C, which may result in either a fall to the hurdle square D, or a successful hurdle jump to square E. Rewards are shown below in the graph (for example, moving "*right*" from Square A to Square B results in a reward of +1). Assume the discount factor $\gamma = 1$.

Rewards:



Actions:

- *right*: Deterministically move to the right
- *left*: Deterministically move to the left
- *jump*: Stochastically jump to the right. This action is available for square C only.
  $T(C, jump, E) = 0.5$ (jump succeeds)
  $T(C, jump, D) = 0.5$ (jump fails)

(a) (8 pts) For the policy π of always taking actions *right* (when in Squares A, B, D–F) or *jump* (when in Square C), please compute $V^{\pi}(C)$.

(b) (8 pts) Is the previous policy π optimal? If yes, please show why it is optimal. If no, please conduct one step of policy improvement to improve it.

(c) (8 pts) Perform two iterations of value iteration and compute $V_2(B), Q_2(B, right), Q_2(B, left)$. Iteration 0 corresponds to the initialization of $V_0(s) = 0$ for all states s.

Naive Bayes (16 Pts)

2. Consider training a Naive Bayes classifier $Y = N(\boldsymbol{X})$ to predict whether an online comment is biased or not. X_1, X_2, X_3 are binary features extracted from the online comment, and Y is the binary label to predict (i.e., biased=1 and non-biased=0). The training dataset is given in the table as below:

X_1	0	0	1	0	1	1	1	1
X_2	0	1	0	1	1	0	1	1
X_3	1	1	1	0	0	1	1	0
Y	0	0	0	1	1	0	0	1

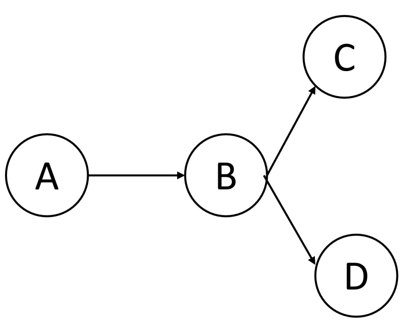
(a) (2 pts) Given this dataset, what is the maximum likelihood estimation for the prior probability P(Y) in the Naive Bayes classifier?

(b) (10 pts) What is the maximum likelihood estimation for $P(X_2|Y)$? When Laplace smoothing with $k = 3$ is used, what is the estimate for $P(X_2 = 0|Y = 1)$?

(c) (4 pts) Assume all parameters of your Naive Bayes classifier are estimated through maximum likelihood estimation with Laplace smoothing ($k = 3$). Given a new online comment characterized by $(X_1 = 0, X_2 = 1, X_3 = 1)$, what is the prediction of your Naive Bayes classifier regarding whether it is biased?

Bonus Question: Approximate Inference (10 Pts)

3. Consider the following Bayesian network:



A	$P(A)$
$+a$	1/3
-a	2/3

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

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

B	D	$P(D\|B)$
$+b$	$+d$	1/3
$+b$	$-d$	2/3
-b	$+d$	1/4
-b	$-d$	3/4

(a) (2 pts) We generated the following set of random samples of this Bayesian network using simple sampling:

(1) $< +a, +b, +c, -d >$

(2) $< +a, -b, -c, -d >$

(3) $< -a, -b, +c, +d >$

(4) $< -a, +b, +c, +d >$

(5) $< -a, +b, -c, +d >$

Use these samples to estimate $P(B = +b)$.

(b) (3 pts) Given the samples in the previous question, show how to estimate $P(B = +b|A = -a, C = +c, D = +d)$ using rejection sampling.

(c) (5 pts) You are now interested in estimating $P(B = +b|A = -a, C = +c, D = +d)$ using likelihood weighting sampling, and the samples you generated are:

(1) $< -a, +b, +c, +d >$

(2) $< -a, -b, +c, +d >$

(3) $< -a, -b, +c, +d >$

What is your estimate of $P(B = +b|A = -a, C = +c, D = +d)$ in this case?

Submission

Upload your answers to the **written** questions as a pdf format file in Gradescope:

- For your pdf file, use the naming convention `username_hw#.pdf`. For example, your TA with username *mmostafi* would name his pdf file for HW4 as `mmostafi_hw4.pdf`.
- To make grading easier, please start a new page in your pdf file for each subquestion. Hint: use a `\newpage` command in LaTeX after every question ends. For example, use a `\newpage` command after each of part (a)-(c) of Question 1.
- After uploading to Gradescope, mark each page to identify which question is answered on the page. (Gradescope will facilitate this.)

Part 2: Programming Assignment (60 pts)

For the programming assignments we will use the Pacman project 3 designed for the course CS188 at UC Berkeley: `https://inst.eecs.berkeley.edu/~cs188/fa20/project3/`

In project 3, you will complete the reinforcement learning questions for Pacman and a crawler robot. Similar to the previous homework, this project include an autograder (which is available at the Berleley site) for you to grade your answers on your machine.

Please remember that solutions to any assignment should be your own. Using other people solutions, within or outside Purdue goes against the course academic honesty policy. The TAs will use code similarity measures to detect plagiarism cases with projects on Github when grading the assignment.

TODO:

1. Complete Project 3, Questions 1-7 described on the Berkeley site. Submit your modified versions of `qlearningAgents.py`, `analysis.py`, `valueIterationAgents.py` for grading. We will multiply your original score returned by `autograder.py` (20 pts in total) by 3.

Submission

Please upload the following files: `qlearningAgents.py`, `analysis.py`, and `valueIterationAgents.py` to Gradescope.