ECE421: Introduction to Machine Learning — Fall 2024

Worksheet 6: Markov Decision Process

Q0 (Probability Review) This question checks your knowledge of basic properties such as independence and conditional independence. Note that the TAs will not review these questions during the tutorial session.

0.a Suppose the variables A and B are Boolean variables (*i.e.*, A can have the value a or $\neg a$, and B can have the value b or $\neg b$). Suppose that A is independent of B. Determine the missing values in the joint probability mass function $P(A, B)$ below.

$P(\neg a, \neg b)$	0.1
$P(\neg a, b)$	0.3
$P(a, \neg b)$	
$P(a, b)$	

0.b Suppose A, B and C are Boolean variables and that B is independent of C given A. Determine the missing values in the joint probability mass function $P(A, B, C)$ below.

$P(\neg a, \neg b, \neg c)$	0.01
$P(\neg a, \neg b, c)$	0.02
$P(\neg a, b, \neg c)$	0.03
$P(\neg a, b, c)$	
$P(a, \neg b, \neg c)$	0.01
$P(a, \neg b, c)$	0.1
$P(a, b, \neg c)$	
$P(a, b, c)$	

Q1 **(Discounting)** Consider the Mars Rover example that we discussed in class. The figure below denote illustrates the environment for this problem, containing five states, namely a, b, c, d, and e.

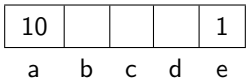

Figure 1: Mars Rover Environment

Assume that the transition dynamics is deterministic. Mars Rover has actions `Left` and `Right` in states b, c, and d, and the only action available in states a and e is `exit`, which ends its mission (*i.e.*, non more transitions after `exit` action). The values in the figure show the reward for the exit action in the two terminal states. Assume a reward of zero for `Left` and `Right` actions.

1.a Find the optimal policy when the discount factor is $\gamma = 1$.

1.b Find the optimal policy when the discount factor is $\gamma = 0.1$.

1.c For which γ are `Left` and `Right` equally good when in state d?

Q2 **(Multiple Choice)** For each question, choose only one option.

2.a Convolutional neural networks have fewer parameters than fully connected networks because they use shared weights. ☐ True ☐ False

2.b It is best to keep the learning rate of a neural network constant as learning progresses. ☐ True ☐ False

2.c In a multi-layered neural network, if the activation of a hidden unit is zero, then the gradients of the weights of all of its incoming connections are zero. ☐ True ☐ False