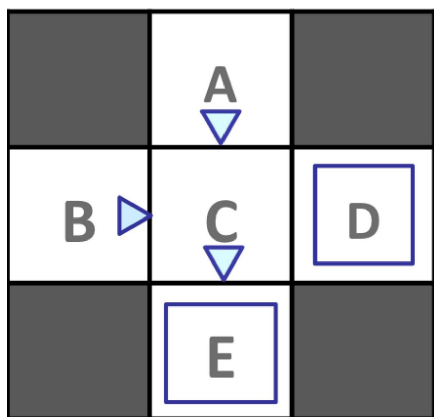


Quiz 3: Passive Reinforcement Learning

Quiz 3: Passive Reinforcement Learning

5/5 points (ungraded)

Input Policy π



Assume: $\gamma = 1$

Observed Episodes (Training)

Episode 1

A, south, C, -1
C, south, E, -1
E, exit, x, +10

Episode 2

B, east, C, -1
C, south, D, -1
D, exit, x, -10

Episode 3

B, east, C, -1
C, south, E, -1
E, exit, x, +10

Episode 4

A, south, C, -1
C, south, E, -1
E, exit, x, +10

What are the estimates for the following quantities as obtained by direct evaluation:

$$\hat{V}^{\pi}(A) =$$



$$\hat{V}^{\pi}(B) =$$



$$\hat{V}^{\pi}(C) =$$



$\hat{V}^{\pi}(D) =$



$\hat{V}^{\pi}(E) =$



Submit

✓ Correct (5/5 points)