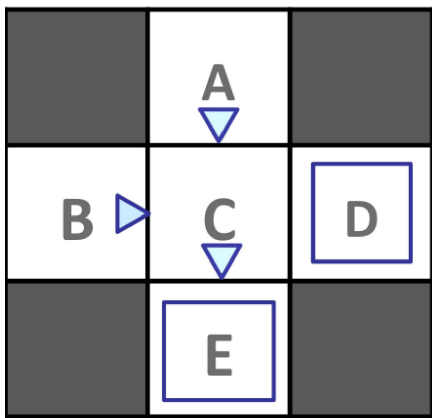


Quiz 2: Model-Based Learning

Quiz 2: Model-Based Learning

4/4 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 model would be learned from the above observed episodes?

$T(A, \text{south}, C) =$



$T(B, \text{east}, C) =$



$T(C, \text{south}, E) =$

0.75



T(C, south, D) =

0.25



Submit

✓ Correct (4/4 points)