edX edge

hw5_rl_q1_model_based_rl_grid
Question 1: Model-Based RL: Grid

4.0/4.0 points (graded)

Input Policy π
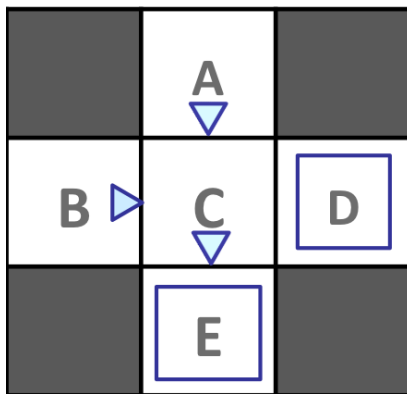
Assume: γ = 1

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, south, C) =

```
1
```
✔

T(B, east, C) =

```
1
```
✔

T(C, south, E) =

.75   ✔

T(C, south, D) =

.25   ✔

Submit