HW8

Seth Jaksik

1001541359

December 8, 2020

Task 1

environment2.txt, -0.04, 1, 20

```
    utilities    :
 0.812   0.868   0.918   1.000
 0.762   0.000   0.660  −1.000
 0.705   0.655   0.611   0.387

policy:
>       >       >       o
^       x       ^       o
^       <       <       <
```

environment2.txt, -0.04, 0.9, 20

```
    utilities    :
 0.509   0.650   0.795   1.000
 0.399   0.000   0.486  −1.000
 0.296   0.254   0.345   0.130

policy:
>       >       >       o
^       x       ^       o
^       >       ^       <
```

Task 2

For the non-terminal states, I would assign a reward of -0.01. This is due to the fact that most chess games do not end quickly, and this would incentivize the algorithm to not take a large penalty for taking more intermediate steps.

For the discount factor γ, I would choose a higher gamma, such as 1.0 to incentivize the algorithm to focus on the end goal of winning the match rather than achieving a reward as fast as possible.

Task 3

Part a)

Up Action assumed optimal:

$$U(2,2) = 0.8 * U_h((2,2),(2,3)) + 0.1 * U(2,2) + 0.1 * U(2,2) = 0.8 * U_h((2,2),(2,3)) + 0.2 * X$$
$$= 0.8 * (\text{-}0.04 + 0.9) + 0.2 * X$$
$$= 0.688 + 0.2 * X$$
$$\therefore X = 0.688 + 0.2 * X$$
$$\therefore 0.8 * X = 0.688$$
$$\therefore X = 0.86$$

$U(2,2)$ for the Up action is 0.86

Part b)

(-0.05,0.05)