Artificial Intelligence
Lab – 6
Yash Ubale – 2014CSB1040

Q-1) This is the output after running value iteration for 5 iterations on the default BookGrid

Grading: You … We will check your value … values, and … convergence (e.g. after … iterations).

Question 2 (…

`BridgeGrid` … minal state and a high-re… terminal state … of which is a chasm of … negative rew… With the default discount … and the defau… the bridge. Change only ON… the discount … causes the agent to attem… cross the brid… (Noise refers to how ofte… agent ends u… perform an action.) The de… corresponds to:

```
python gridworld.py -a value -i 100 -g BridgeGrid --discount 0.9
```

Q-2) After putting the values answerDiscount = 0.9 answerNoise = 0 the following output is obtained

… ed from a … vector is … ed on the … mputation … eady been

Note: A policy synthesized from values of depth k (which reflect the next k rewards) will actually

Q-3)

a) Prefer the close exit (+1), risking the cliff (-10)

answerDiscount = 0.7

answerNoise = 0.1

answerLivingReward = -5

We need to get to the nearer goal state with shorter and risky path so living reward will have low value and noise should be less so as to take the shorter path

b) Prefer the close exit (+1), but avoiding the cliff (-10)

answerDiscount = 0.5

answerNoise = 0.4

answerLivingReward = -2

We need to get to the nearer goal state but with longer and safe path, so living reward is increased and noise is added so that it takes longer path

c) Prefer the distant exit (+10), risking the cliff (-10)

answerDiscount = 0.9

answerNoise = 0.1

answerLivingReward = -2

We need to get to the farther goal state but with shorter and risky path so, we have reduced the noise and increased to discount so that the thinks for ‘long term’ goals. Noise is reduced so that it takes shorter and risky path

d) Prefer the distant exit (+10), avoiding the cliff (-10)

answerDiscount = 0.9

answerNoise = 0.4

answerLivingReward = -1

We need to get to the farther goal state but with longer and safe path, so living reward is increased and noise is added so that it takes longer path. Discount is high so that it thinks for ‘long term’ goals

e) Avoid both exits and the cliff (so an episode should never terminate)

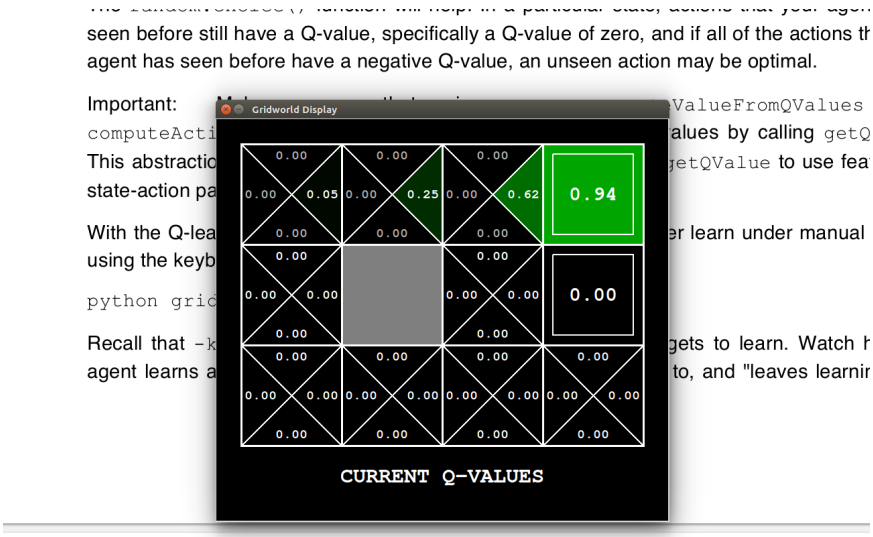
answerDiscount = 0

answerNoise = 0

answerLivingReward = 1

We need our agent to avoid exits and not to use shorter path so living reward is made possitive so that it keeps on getting reward just by not going to the terminal state. Discount is made zero so that it doesn’t think about any goal

Q-4) This is the output after manual movements and setting the noise to 0.0 after 5 iterations by moving only North and East



Q-5) After running for 100 iterations python gridworld.py -a q -k 100