

Homework Assignment 1B

Zhaoyang Xu

Study group: Yi Liu, Jingzheng Li, Zhaoyang Xu

September 15, 2020

1 Grid World (25 points)

1.1

-76.79	-73.18	-74.53
-81.45	-68.08	-76.19
-76.97	-63.8	-74.53
-86.3	-39.17	-13.25

(a) $\gamma = 1.0$

-16.72	-15.58	-13.63
-20.84	-14.46	-15.24
-18.66	-15.31	-13.63
-23.69	-9.11	-2.5

(b) $\gamma = 0.9$

Figure 1: $12V_{rand}$ with two γ

1.2

-10.0	-9.0	-15.0
-11.0	-3.0	-15.0
-3.0	-2.0	-15.0
-14.0	-1.0	0.0

(a) $\gamma = 1.0$

-8.6	-8.44	-10.0
-8.74	-2.71	-10.0
-2.71	-1.9	-10.0
-10.0	-1.0	0.0

(b) $\gamma = 0.9$

Figure 2: $12V^*$ with two γ

By policy iteration, algorithm on page 37 from slide. We calculated the Bellman equation first, used the Bellman optimality equation to change policy.

With different discount γ , the threshold is different. For $\gamma = 1$, the threshold is bigger than when $\gamma = 0.9$.