Soheil Shirvani
96246592
sshiro30@ucr.edu

Assignment 1

Question 1: states $(1,3), (2,3)$ are terminal, thus $V_\pi(S=\text{terminal}) = 0$

~~States $(1,3), (2,3)$ are terminal states and do not have any outgoing transitions. Therfore, the state values for terminal states are thier immediate reward.~~

~~$V_\pi((1,3)) = 5$~~ ~~$V_\pi((2,3)) = 5$~~

Question 2:

$$V_\pi(S) = \sum_{a \in A} \pi(a|S) \sum_{S',r} P(S',r|S,a)\left[r + \gamma V_\pi(S')\right]$$

$\pi(a|S) = 0.25$, $\gamma = 0.9$

$$V((1,1)) = 0.25 \times \left[0.8 \times (0 + V_\pi((2,1))) + 0.1 \times (0 + V_\pi((1,2))) + 0.1 \times (0 + V_\pi((1,1)))\right] \times 0.9$$

$$+ 0.25 \times \left[0.8\, V_\pi((1,1)) + 0.1 V_\pi((1,2)) + 0.1 V_\pi((2,1))\right] \times 0.9$$

$$+ 0.25 \times \left[0.8\, V_\pi((1,2)) + 0.1 V_\pi((2,1)) + 0.1 V_\pi((1,1))\right] \times 0.9$$

$$+ 0.25 \times \left[0.8\, V_\pi((1,1)) + 0.1 V_\pi((2,1)) + 0.1 V_\pi((1,1))\right] \times 0.9$$

$$V_\pi(1,2) = 0.25 \times \Big[0.8\, V_\pi((2,2))^{\times 0.9} + 0.1\big(-5 + V_\pi((1,3))^{0 \times 0.9}\big) + 0.1\, V_\pi((1,1))^{\times 0.9} \Big]$$
$$+ 0.25 \times \Big[0.8\, V_\pi((1,2))^{\times 0.9} + 0.1(-5) + 0.1\, V_\pi((1,1))^{\times 0.9} \Big]$$
$$+ 0.25 \times \Big[0.8(-5) + 0.1\, V_\pi((2,2))^{\times 0.9} + 0.1\, V_\pi((1,2))^{\times 0.9} \Big]$$
$$+ 0.25 \Big[0.8\, V_\pi((1,1)) + 0.1\, V_\pi((2,2)) + 0.1\, V_\pi(1,2) \Big] \times 0.9$$

$$V_\pi(2,1) = 0.25 \times \Big[0.8\, V_\pi((2,1)) + 0.1\, V_\pi((2,2)) + 0.1\, V_\pi((2,1)) \Big] \times 0.9$$
$$+ 0.25 \times \Big[0.8\, V_\pi((1,1)) + 0.1\, V_\pi((2,2)) + 0.1\, V_\pi((2,1)) \Big] \times 0.9$$
$$+ 0.25 \times \Big[0.8\, V_\pi((2,2)) + 0.1\, V_\pi((2,1)) + 0.1\, V_\pi((1,1)) \Big] \times 0.9$$
$$+ 0.25 \times \Big[0.8\, V_\pi((2,1)) + 0.1\, V_\pi((2,1)) + 0.1\, V_\pi((1,1)) \Big] \times 0.9$$

$$V_\pi(2,2) = 0.25 \times \Big[0.8\, V_\pi((2,2))^{\times 0.9} + 0.1\big(5 + V_\pi((1,3))^{0 \times 0.9}\big) + 0.1\, V_\pi((2,1))^{\times 0.9} \Big]$$
$$+ 0.25 \times \Big[0.8\, V_\pi((1,2))^{\times 0.9} + 0.1(5) + 0.1\, V_\pi((2,1))^{\times 0.9} \Big]$$
$$+ 0.25 \times \Big[0.8(5) + 0.1\, V_\pi(2,2)^{\times 0.9} + 0.1\, V_\pi((1,2))^{\times 0.9} \Big]$$
$$+ 0.25 \times \Big[0.8\, V_\pi(2,1) + 0.1\, V_\pi(2,2) + 0.1\, V_\pi(1,2) \Big]$$

Question 3:

$$V_\pi(1,1) = 0.25\left[V_\pi(2,1)^{\times 0.9} + V_\pi(1,2)^{\times 0.9} + 2V_\pi(1,1)^{\times 0.9}\right]$$

$$V_\pi(1,2) = 0.25\left[V_\pi(2,2)^{\times 0.9} + V_\pi(1,1)^{\times 0.9} - 5 + V_\pi(1,2)^{\times 0.2}\right]$$

$$V_\pi(2,1) = 0.25\left[2V_\pi(2,1)^{\times 0.9} + V_\pi(1,1)^{\times 0.9} + V_\pi(2,2)^{\times 0.9}\right]$$

$$V_\pi(2,2) = 0.25\left[V_\pi(2,2)^{\times 0.9} + V_\pi(1,2)^{\times 0.9} + 5 + V_\pi(2,1)^{\times 0.9}\right]$$

$$\begin{cases} V_\pi(1,1) = -0.388 \\ V_\pi(1,2) = -1.337 \\ V_\pi(2,1) = 0.388 \\ V_\pi(2,2) = 1.337 \end{cases}$$

Question 4:

$$V_\pi'(1,1) = [0.8V_\pi'(2,1) + 0.1V_\pi'(1,1) + 0.1V_\pi'(1,2)] \times 0.9$$

$$V_\pi'(1,2) = [0.8V_\pi'(1,1) + 0.1V_\pi'(2,2) + 0.1V_\pi'(1,2)] \times 0.9$$

$$V_\pi'(2,1) = [0.8V_\pi'(2,2) + 0.1V_\pi'(2,1) + 0.1V_\pi(1,1)] \times 0.9$$

$$V_\pi'(2,2) = 0.8 \times 5 + 0.1 \times 0.9V_\pi'(2,2) + 0.1 \times 0.9V_\pi'(1,2)$$

$$\begin{cases} V_\pi'(1,1) = 3.328 \\ V_\pi'(1,2) = 3.098 \\ V_\pi'(2,1) = 3.819 \\ V_\pi'(2,2) = 4.702 \end{cases}$$

Question 5:

$$V^*(s) = \max_{a \in A} \sum_{s',r} p(s',r|s,a)[r + \gamma V^*(s')]$$

$$V^*(1,1) = [0.8\, V_\pi'(2,1) + 0.1\, V_\pi'(1,1) + 0.1\, V_\pi'(1,2)] \times 0.9 = 3.328$$

$$V^*(1,2) = [0.8\, V_\pi(1,1) + 0.1\, V_\pi'(2,2) + 0.1\, V_\pi(1,2)] \times 0.9 = 3.098$$

$$V^*(2,1) = [0.8\, V_\pi'(2,2) + 0.1\, V_\pi'(2,1) + 0.1\, V_\pi'(1,1)] \times 0.9 = 4.028$$

$$V^*(2,2) = [4 + 0.09\, V_\pi'(2,2) + 0.09\, V_\pi'(1,2)] = 4.702$$

Since $V_\pi^*(2,1) \neq V_\pi'(2,1)$ it doesn't satify bellman optimality

Question 6:

Since bellman optimallity is not satisfied the $\pi(s)$ is <u>not optmal</u>

counter example is state $(2,1)$. $\pi''(2,1) =$ up instead of left

$$\Rightarrow V^*(2,1) = 0.9[0.8\, V_\pi'(2,2) + 0.1\, V_\pi'(1,1) - 0.1 \times 5 = 3.184$$

since $V_\pi''(2,1) > V_\pi(2,1) \Rightarrow \pi_\pi'$ is <u>not optimal</u>