# CS221 Fall 2018 Homework [blackjack]

SUNet ID: prabhjot
Name: Prabhjot Singh Rai

By turning in this assignment, I agree by the Stanford honor code and declare that all of this is my own work.

# Problem 1

(a) The equation $V_{opt}$ is given by:

$$V_{opt} = \begin{cases} 0 & \text{if } endState \\ max_{\text{a } \epsilon \text{ actions}} Q_{opt}(s,a) & \text{otherwise} \end{cases}$$

where,

$$Q_{opt}(s,a) = \sum_{s'} T(s,a,s')[\text{Reward}(s,a,s') + \gamma V_{opt}(s')]$$

**Iteration 1**
As per the question, $V$ is assigned values of 0. Therefore,

$$V_{opt} = \{-2:0, -1:0, 0:0, 1:0, 2:0\}$$

**Iteration 2**

(a) State 0

i. Action +1

$$\begin{aligned} Q_{opt} &= 0.3 * [-5+0] + 0.7 * [-5+0] \\ &= -5 \end{aligned}$$

ii. Action −1

$$\begin{aligned} Q_{opt} &= 0.8[-5+0] + 0.2 * [-5+0] \\ &= -5 \end{aligned}$$

Therefore, $V^1_{opt}(0) = -5$

(b) State 1

i. Action $+1$

$$\begin{aligned} Q_{opt} &= 0.3 * [100 + 0] + 0.7 * [-5 + 0] \\ &= 26.5 \end{aligned}$$

ii. Action $-1$

$$\begin{aligned} Q_{opt} &= 0.8 * [-5 + 0] + 0.2 * [100 + 0] \\ &= 16 \end{aligned}$$

Therefore, $V^1_{opt}(1) = 26.5$

(c) State $-1$

i. Action $+1$

$$\begin{aligned} Q_{opt} &= 0.3 * [-5 + 0] + 0.7 * [20 + 0] \\ &= 12.5 \end{aligned}$$

ii. Action $-1$

$$\begin{aligned} Q_{opt} &= 0.8 * [20] + 0.2 * [-5] \\ &= 15 \end{aligned}$$

Therefore, $V^1_{opt}(-1) = 15$

(d) State 2, since it's an end state $V^1_{opt}(2) = 0$

(e) State $-2$, since it's an end state $V^1_{opt}(-2) = 0$

Therefore,

$$V_{opt} = \{-2 : 0, -1 : 15, 0 : -5, 1 : 26.5, 2 : 0\}$$

(b) (your solution)

# Problem 2

(a) (your solution)

(b) (your solution)