CMPUT 366 Programming Assignment 1 Report
Greg Knoblauch & Bryce Cartman

Question 2:

Alpha = **0.001**
Emu = **0.01**
Epi = **0.01**
numEpisodes = **1,000,000**

The average sum returned is: **-0.043464**
The learned policy is:
Usable Ace:

S H H S S S S S H S 20
H S S H H S H S S S 19
H H H H H H S S S H 18
H H H H H S H H H H 17
H H H H H H H H H H 16
H H H H H H S H H H 15
H H H H H H H H H H 14
H H H H H H H H H H 13
H H H H H H H H H S 12
1 2 3 4 5 6 7 8 9 10

No Usable Ace:

S S S S S S S S S S 20
S S S S S S S S S S 19
S S S S S S S S S S 18
S S S S S S S S S S 17
H S S S S S H H S H 16
H H S S S H H H H H 15
S H H S H H H H H H 14
H H H S H H H H H H 13
S H H H S H H H H H 12
1 2 3 4 5 6 7 8 9 10

Question 3:

For this question we built a script to iterate though possible Emu and Epi values between 0.01 and 0.1, along with a corresponding alpha and number of episodes. While we were unable to complete our script using 10,000,000 episodes, due to time constraints for all values, we did get a new maximum. Are findings are below which gave us the best average return.

CMPUT 366 Programming Assignment 1 Report
Greg Knoblauch & Bryce Cartman

Alpha = **0.01**
Emu = **0.1** with 1mil episodes
Epi = **0.01** with 1mil episodes
numEpisodes = We used 1,000,000 episodes for learning and for our deterministic policy which produced our bests results as shown below. Note: anything that was less than 1,000,000 episodes produced results which were random, sometimes better and sometimes worse.

The average sum returned: **-0.026875**

The learned policy is:
Usable Ace:
S S S S S S S S S 20
S S S S S S S S S 19
H S S S S S S H H 18
H H S H H S H H H H 17
H H H H H H H H H H 16
H H H H H H H H H H 15
H H H H H H H H H H 14
H H H H H H H H H H 13
H H H H H H H H H H 12
1 2 3 4 5 6 7 8 9 10

No Usable Ace:
S S S S S S S S S 20
S S S S S S S S S 19
S S S S S S S S S 18
S S S S S S S S S 17
H S S S S H H H H 16
H H H S S H H H H H 15
H H H H H H H H H H 14
H H H H H H H H H H 13
H H H H H H H H H H 12
1 2 3 4 5 6 7 8 9 10