**CPS841 Reinforcement Learning Assignment 3 Question 1 Report  
Tony Misic 500759917 --- Daniel Bellissimo 500749419**

**Question 5.5 \_**

**First Visit:**

By looking at the algorithm we can see that at the beginning we would go the last state, which would be a non-terminal state. Next, we set G = 0 and continue looping. During each iteration except the last, we continuously add 1 to the sum of rewards, by the last state we are at a total of 10. This is because we end on a non-terminal state. Therefore if we took the average of the Returns we would get 10 / 11, which equal 10, which is the value.

**Every Visit:**

Ever visit follows the same principle expect you take the average for each state iteration not just the first case of the state. This means that we get