CS 4700: Foundations of Artificial Intelligence

Spring 2020

Review questions for Quiz 2

For each of the following problems assume g(N)=$\sqrt{\log_{10} N}$.

1. Consider a multi-armed bandit problem with four arms, 1, 2, 3, and 4, each of which return a positive-valued reward. Imagine there have been 7 prior arm pulls – 2 pulls for each of arms 1, 2, and 3, and 1 pull for arm 4. Further, given the values of the rewards received up through that point, the UCB heuristic says to pull arm 4 as the 8th arm pull. After the N=8 arm pulls the relevant statistics are:
   - Sum_1=1 , N_1=2
   - Sum_2=2, N_2=2
   - Sum_3=3, N_3=2
   - Sum_4=4, N_4=2

   What is the smallest and largest values of the reward that arm 4 could ever have returned for its first pull?

2. Consider a multi-armed bandit with two arms. Arm 1 always gives a reward of 1 on each pull. Arm 2 always gives a reward of 0 on each pull. Other than during the initialization when each arm is pulled once, would arm 2 ever get pulled again? If no, explain why. If yes, give a value for the number of pulls.

3. True/False: Consider a multi-armed bandit with five arms. All five arms are identical and give a reward of 1 on each pull. After 10 arm pulls using the UCB algorithm each arm will have been pulled twice, regardless of how you break ties when arms have equal UCB values.

4. Consider a multi-armed bandit with two arms. After the first pull of each arm Arm 1 gave 0 and Arm 2 gave 1. How many additional arm pulls would it take to pull Arm 1 again, assuming in the meantime Arm 2 always gave a 0 for each further pull?