CMPS 4660/6660 Reinforcement Learning – Fall 2020

Homework 3 (15 points)

Due 10/13/20 before class

Note: Please clearly justify your answer to each of the following questions. **Questions marked with *** are required for graduate students only.**

1. **TD(0) (5 points)**

   Consider the random walk example we discussed in class (Example 6.2 in Sutton and Barto's book). From the results shown in the left graph of the random walk example it appears that the first episode results in a change in only $V(A)$. What does this tell you about what happened on the first episode? Why was only the estimate for this one state changed? By exactly how much was it changed?

2. **n-step TD (5 points)**

   Consider applying n-step TD method to the random walk example (Example 7.1 in Sutton and Barto's book). Why do you think a larger random walk task (19 states instead of 5) was used in this example? Would a smaller walk have shifted the advantage to a different value of n? How about the change in left-side outcome from 0 to -1 made in the larger walk? Do you think that made any difference in the best value of n?

3. ***** Equivalence of Forwards and Backwards TD(λ) (5 points)**

   Consider an episode where a state s may be visited multiple times. Prove that the sum of *offline* updates is identical for forward-view and backward-view TD(λ).