

Homework 3

Due date: 18th Jun., 2020

- (100 points) **(Matlab Exercise for the random walk)** Consider the Markov reward process (MRP) shown below, where all episodes start in the center state C, then proceed wither left or right by one state on each step, with equal probability. Episodes terminate either on the extreme left or the extreme right. When an episode terminates on the right, a reward of +1 occurs; all other rewards are zero. This task is undiscounted.



Measure the performance in terms of the root-mean-squared (RMS) error between the learned value function and the true value function, averaged over the five states, then averaged over 100 runs, using TD(0) and Monte Carlo. For TD(0), the performance is measured when $\alpha=0.05, 0.1, 0.15$, while for Monte Carlo, the performance is measured when $\alpha=0.1, 0.2, 0.3, 0.4$. In all cases the approximated value functions was initialized to the value $V(s)=0.5$ for all s . Plot a figure as shown in the following.

