

1. *We have not explicitly considered or given pseudocode for any Monte Carlo methods in this chapter. What would they be like? Why is it reasonable not to give pseudocode for them? How would they perform on the Mountain Car task?*

The Monte Carlo is basically  $n$ -step Sarsa with  $n = T$ . I expect it to perform poorly on the Mountain Car example: until it doesn't reach the goal, it doesn't learn anything, so the first episode might be really long. As opposed to this,  $n$ -step Sarsa tries new actions after it sees that the previous actions didn't lead to the end of the episode, so it will get to the goal line eventually.