Assignment 1

Deep Reinforcement Learning

Spring 2022

Sutton - Exercise 3.1

Different examples of Markov Decision Processes can be found in Sutton. Credit is awarded for any answers that specify the state, actions and rewards.

Sutton - Exercise 3.3

Although this question is subjective, in general, the line should be drawn for properties of the agent that do not impact the environment or rewards. For example, in driving, the state can be determined by the cars speed and acceleration. In this case, it's not necessary to track the angle at which an individual hits the accelerator, as the velocity and acceleration may be used to estimate the individual's arrival time to their destination. In general, it is useful to choose the minimum amount of parameters to sufficiently achieve one's optimization objective. Sometimes, parameters that have a very small impact on the system's optimization objective may also be not included, as introducing complexity to a model increases the difficulty in finding solutions.

Sutton - Exercise 3.4

Give a table analogous to that in Example 3.3, but for $p(s_0, r|s, a)$. It should have columns for s, a, s_0, r, and $p(s_0, r|s, a)$, and a row for every 4-tuple for which $p(s_0, r|s, a) > 0$

Investing in Stocks

Let's say you start off at day one, with two stocks S_1 and S_2 to choose to invest in. The return of stock one is uniformly distributed (as a percentage) between [-1,1]. The return of stock two is uniformly distributed (as a percentage) between [-2,8]. Assume you start off on day one with 100 dollars.

Part 1

Write a simulation that runs for one hundred iterations, and randomly invests in stock one with a fifty percent chance, and in stock two with a fifty percent chance. What is your final return on day 100, assuming you start day 1 with 100 dollars? Please attach your code to this assignment.

Part 2

In this problem, what is the state variable(s)? What is the action variable(s)? What are the possible states and actions? What are the possible rewards each day?

Part 3

If you wanted to maximize your mean return, which stock is better to invest in? **Hint:** The solution is very short, you could write it in a sentence or two.