June 23, 2023

I'm going over the Reinforcement learning book you've sent. I'm on chapter 2, which is about the multi armed bandit. I have the basic premise of the problem: trying to balance between exploration and exploitation when maximizing rewards.  I have a pretty good understanding of the terminology and a few methods including Upper Confidence Bound (identify which options has more benefit from exploration, usually with a wider graph) and Epsilon Greed (a method which separates the # of exploitations and # of explorations given an epsilon percentage). Though, I am still confused on the mathematical aspect and how to possibly apply on Webots. Chance wasn't available during lab hours, so couldn't ask. I was able to create an environment on Webots for the robots to run: created a robot variable from scratch with a camera and object detection, also created a new ball object using the same parameters as the robot bulls soccer ball. For now because I have more of an understanding of the basics, I was planning on looking back at the previous research papers to see if I can understand how they set up their reinforcement learning policies for their robots and possibly apply the same or improve the methods onto my own application. Additionally, I will still be going over the reinforcement learning textbook, doing research on each section. Once Chance is available I will be asking him about the mathematical aspect of RL. My goal is to possibly apply some sort of RL onto my Robot, since I finish setting up the platform. Not ready to kick the ball yet, but more just moving towards the ball.