1 RL Coding Exercises

1. **Playing with Cartpole, Randomly**

   - Open a Python 3 Notebook named "HW7 - First Name, Last Name" at colab.research.google.com. NOTE: You're welcome to write and run the code in your own computer instead to visualize the environment. Just paste the code in Colaboratory when you're done so it can be shared with us.
   - Recreate the cartpole code we wrote in class
   - Comment the code such that any beginner could understand it
   - Take any policy that exceeds 200 reward once during our random policy search. Write down that policy as a comment, and run that exact policy 5 more times. Comment (or more preferably, plot with Matplotlib.pyplot) the reward it attains each time.

2. **Playing with Cartpole II, Noisily**

   - Use the same Python 3 Notebook and algorithm from the above section
   - During our algorithm, if a random policy has reward > 100, but still < 200, save the policy's parameters. This is our "almost-good" random policy.
   - Instead of trying a whole new random policy, we'll build off of this "almost-good" one
   - Create a new random vector, and divide this new vector by some constant (ex: 10)
   - Add this new random vector to our "almost-good" policy's parameters, to perturb it in some random direction. This is our new policy.
   - Run this new algorithm a few times, and record your thoughts/observations as a comment. (Feel free to tweak it).
   - Share the file with cmsc389f@gmail.com