CS 594: Modern Reinforcement Learning Homework 3 Due Sunday of Week 9 11:59 PM

For this assignment you will experiment with a number of RL algorithms and write a report with the results. This assignment must be done individually. That it, each student must run their own experiments and write their own report. But you may (and in fact are encouraged to) give and receive unlimited help with your setup, code, debugging and other aspects. The end goal is that everyone taking the course feels comfortable running modern RL algorithms on their own but help getting to that point is perfectly fine.

I have updated the resources on the course website with more code examples and other resources that may help with this. You are free to choose any reasonable way of implementing each of the experiments based on your experiences, resources, and eventual project plans. As you find interesting / good ways to perform the experiments, please share them on Piazza so others can benefit from your experiences. You also shouldn't feel obligated to perform exactly the experiments described if you want to do something more interesting or related to your eventual project. I will be satisfied as far as the course goes as long as your report describes a reasonable set of experiments.

Please submit your report via gradescope. Please make sure to tag which page(s) each experiment is on at the appropriate step in the submission process. For full credit the description of each experiment should include (a) what specific question you are answering, (b) your experiment design (including what algorithm and environment you selected) and a discussion of how the design answers your question, and (c) appropriate data (e.g. plots or summary statistics) summarizing the results of your experiment as well as discussion and interpretation of your results (which may include reflecting on your experiment design and whether the design itself was successful).

- 1. **Tuning:** Choose an RL algorithm with at least one parameter. Keeping all but one of them fixed, explore tuning the remaining parameter. You are encouraged to draw inspiration from the discussion of parameter tuning in the papers we have read so far this semester.
- 2. Value Learning: Choose any algorithm which works primarily by learning values (e.g. DQN). Examine the effects of adding or substracting one of the ideas we have read about from it. This could be done by adapting an implementation (e.g. adding double q-learning to an algorithm that does not already use it), comparing two existing implementations (e.g. DQN vs Dueling DQN), or ablating a feature from a more complex algorithm (e.g. as Rainbow did). You may choose your metric when examining this and there are many options: performance of final model, performance of intermediate models, wall-clock time to learn an acceptable model, consistency of learned-model performance.
- 3. **Policy Gradients:** Choose any policy gradient or actor-critic method and examine the effects of adding or subtracting a feature. E.g. you could compare REINFORCE-type estimates to those you get when adding a baseline, or baseline methods to actor-critic. Beyond the metrics from the previous question you could explore the variance of the estimates from the two versions of the algorithm.
- 4. Free Choice: Design an experiment of your own choice. This could be a richer experiment with value learning or policy gradient methods, a comparison between them or other algorithms (e.g. TRPO), or something else entirely that relates to your project or interests.