MSBA –Optimization 2
Project 3 – Reinforcement Learning

Deliverables

One python (.ipynb or .py) file and one HTML file, submitted to Canvas. Your report should go into some detail about how you solved the problem, include some graphs that explain your results, and include relevant code chunks in the final output. 66% of your grade will be based on whether you create a good/thorough RL or not, the remaining 34% will be based on the quality the presentation of your analysis. We will re-run your code. If we don't score the same as you report or your python file doesn't run, we will go through your code and give partial credit accordingly. The easier it is to read your code the easier it is for us to understand what you're doing, so use a lot of comments in your code!

Problem Overview

You work for a video game company. Your company is considering using RL to automate opponents' actions when playing against the person playing the game. Your company is very new to RL, so your boss wants you to explore RL on some simple Atari games first. You will do this for 2 Atari games. First you will develop an RL strategy to play pong, and then you will choose 1 other game to automate.

Specifics

1) Write a reinforcement learning player that plays pong. We have built 2 simple engines in class that did not perform very well. Your job is to write one that does perform well! You may use any of the strategies we talked about in class to improve the performance. You can try actor/critic, you can use a memory buffer, you can remove frames without a ball, you can choose an action after every 4 frames and play the same action for 4 frames in a row, or you can use a linear annealed strategy. Additionally, you could try to modify the reward after each frame; for example, by giving yourself a reward when your paddle hits the ball and a penalty when your opponent's paddle hits the ball. If you try this, be careful not to put yourself in a situation where the RL wants to play long points which give you lots of rewards for hitting balls lots of times! Train this RL for as long as you want, but you need to get it to win. After you train, play 200 games with the trained RL; the average score of these 200 games needs to be a winning score.
2) Now fit another RL for no more than 5,000 games. After training is done, play 200 games with your trained NN. What is your average score over those 200 games? The group with the highest average score gets 5 points extra credit. I would not be surprised if after 5,000 games of training your RL still loses many points…It's ok if the network structure is the same as step 1, but you must start training over from the beginning for this part of the project.
3) Pick another Atari game to play. There is a signup sheet on canvas, which is editable by you, where you can sign up for a game to play. No game can be played by more than 2 groups. Some of these games may be hard, and some may be easy. I have not tried most of them. Create a RL that can play this game. You can try the same things as you tried in pong. Modifying the reward won't be the same, but you are free to modify it as you see fit; just be careful you aren't creating competing objectives – you want to win! That's not to say you have to build an RL that wins at this game, just give it a shot, do the best you can.
4) Pretend you are a developer at the gaming company. Your boss is interested in potentially using RL to automate players. Your team has been asked to write a report about the effectiveness of RL. Write this project as if this is what you're going to deliver to your boss. Your boss is pretty technical and understands RL, so don't be afraid to include quantitative material. Your boss is also busy, so be sure to include some visualizations to get the important points across. Would it be beneficial to hire an RL expert to work at your company?