Connect 4 Bot Using Reinforcement Learning

Randy Deng
Joshua Kassab

End Goal:

Our goal for this project is to implement a near-perfect Connect 4 Bot using reinforcement learning. We plan on using Q-Learning to train the bot and OpenAI Gym to develop the Connect 4 programming environment. If time permits, we would also like to explore other algorithms such as Deterministic Policy Gradient (DPG), Cross-Entropy Method (CEM), and Deep SARSA. Our stretch goal would be to build a framework that will allow us to compare an ensemble of reinforcement learning algorithms with their single counterparts.

Motivation:

Reinforcement learning has recently been at the forefront of machine learning and applied to a wide variety of situations ranging from playing classical video games and beating professionals in board games [1, 2] to improving elevator performance [3]. Though not within the scope of this class, we believe reinforcement learning to be a good topic to explore and learn about in addition to the classical machine learning methods covered in class.

Methodology:

There are three main phases to our project:

1. Phase one is designing the Connect 4 learning environment in OpenAI Gym that will allow us to easily interface with it. There are currently no available environments online for Connect 4, which is why we will create our own.
2. Phase two is training our bot using Q-Learning. This will require learning about how each of the parameters may affect our bot and tuning them accordingly.
3. Phase three is our stretch goal, which is training multiple bots using different algorithms and creating an ensemble to make decisions.

Some of the metrics we plan on collecting include training time, games played, and win rate.

References:

[1] Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., ... & Chen, Y. (2017). Mastering the game of go without human knowledge. *Nature*, *550*(7676), 354.

[2] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*.

[3] Crites, R. H., & Barto, A. G. (1996). Improving elevator performance using reinforcement learning. In *Advances in neural information processing systems* (pp. 1017-1023).