Calvin Kwei Hoi Tan

Pre-Training to Speed Up Reinforcement Learning

Traditional Reinforcement Learning (RL) methods require an Artificial Intelligent (AI) agent to perform an action given the current state and learn both the risk/reward (values/features) and a strategy of actions (policy) to perform that maximizes reward. These methods often are time consuming when training. They may not converge to a solution, especially if the environment contains a large set of actions with sparse extrinsic reward signals, thus requiring research in existing or innovative techniques to potentially address these issues.

A technique that found success in accelerating training in other areas of Machine Learning is pre-training. Pre-training, specifically in Deep Learning, allows the network to be initialized in a particular region of parameter space often unreachable through random initialization and has proven to have better results compared to traditional methods of training [1]. A recent success like the Atari 2600 games [2] capitalizes on this strategy by training an initial policy through supervised learning on human moves and further trained using RL, ultimately demonstrating significant improvements on training time. For these games, however, there exists a dataset of examples that can be used to pre-train an agent, which may not be true for other environments.

Another technique that is used to improve training on sparse reward signal environments is Auxiliary Tasks [3]. Auxiliary Tasks introduce pseudo-rewards functions during the training process that help the agent select and learn aspects of the environment (not possible solely with extrinsic reward) and has shown speedup in learning [4]. The Auxiliary Task technique simultaneously updates the RL agent as well as the pseudo-reward function during training, which differs from the proposed pre-training method of initializing parameters for the agent.

The goal in this research is to explore the principles of pre-training RL agents using a Deep Learning approach for the card game Gin Rummy, a high-dimensional environment containing sparse reward signals between each state [5] that does not converge using conventional RL techniques. The Deep Learning model will take an Auxiliary Task approach: it will be focused on learning an aspect of the game to learn features similar to the reward prediction task [4], however will not trained during the RL algorithm like a typical Auxiliary Task.

To achieve this goal, the following objectives must be met:

- Apply pre-training to learn features of the Gin Rummy environment
- Synthesize the pre-trained network as initialization weights for a RL agent
- Train the RL agent using a Deep RL algorithm
- Outperform the baseline agent in a series of games [6]

The approach is to pre-train a Deep Neural Network (DNN) through supervised learning using hundreds of thousands of self-play examples generated using a baseline agent [6]. The DNN will be tuned to learn possible features of the environment such as the ability to predict when to knock [6], which card the player will discard, and which card the opponent will discard. The initialized weights of the DNN will be used as the initial parameters for the Deep RL algorithm. The agent will be trained using a Deep Q-Learning Network (DQN), however other algorithms like Asynchronous Advantage Actor-Critic (A3C) [2] will be attempted. With the combination of the pre-trained NN and Deep RL algorithm, the agent will be evaluated against a baseline agent following simple heuristics [6].

Reinforcement Learning has many real-world applications that have strict time constraints, thus it is important to explore different methodologies to address the issue of training time and convergence of the model solution given sparse reward signals.

[1] D. Erhan, A. Courville, Y. Bengio, and P. Vincent, *Why Does Pre-training Help Deep Learning?* The Journal of Machine Learning Research, 2010. [Online]. Available: http://proceedings.mlr.press/v9/erhan10a/erhan10a.pdf

[2] G. Cruz Jr, Y. Du, and M. Taylor, *Pre-training Neural Networks with Human Demonstrations for Deep Reinforcement Learning*. 2nd ed, 2019. [Online]. Available: https://arxiv.org/pdf/1709.04083.pdf

[3] J. Hare, *Dealing with Sparse Rewards in Reinforcement Learning*. 2nd ed, 2019. [Online]. Available: https://arxiv.org/pdf/1910.09281.pdf

[4] M. Jaderberg, V. Mnih, W.M. Czarnecki, et al, *Reinforcement Learning with Unsupervised Auxiliary Tasks*. 1st ed, 2016. [Online]. Available: https://arxiv.org/pdf/1611.05397.pdf

[5] "Gin Rummy", *Games in RLCard*. DATA Lab at Texas A&M University, 2020. [Online]. Available: https://rlcard.org/games.html#gin-rummy

[6] T. Neller, *Gin Rummy EAAI Undergraduate Research Challenge*. Gettysburg College: Department of Computer Science, 2019. [Online]. Available: http://cs.gettysburg.edu/~tneller/games/ginrummy/eaai/