Project Proposal

Dongqiangzi Ye	Haozhen Wu	Jennifer Cao
dye25@wisc.edu	hwu84@wisc.edu	qcao44@wisc.edu
9080082994	9068149831	907999836

Although deep neural networks themselves have made various breakthrough in different problems such as machine translation, computer vision and time series prediction, they can achieve even more amazing result when combined with Reinforcement learning. For example the Alpha Go from Google Deepmind.

Reinforcement learning refers to goal-oriented algorithms, which learn to represent a complex objective (goal) or maximize along a particular dimension over many steps. Compared to common supervised learning, where we can immediately see the outcome of a particular prediction, Reinforcement learning may reward a prediction after several steps. We think about the world for supervised learning is static, whereas for Reinforcement learning it is dynamic and the prediction will be based on previous experience and current information of the environment. For example, maximize the points won in a game over many moves. They can start from a blank slate, and under the right conditions they achieve superhuman performance. Like a child incentivized by spankings and candy, these algorithms are penalized when they make the wrong decisions and rewarded when they make the right ones – this is reinforcement.

Reinforcement algorithms that incorporate deep learning can beat world champions at the game of GO as well as human experts playing numerous Atari video games. While that may sound trivial, it's a vast improvement over their previous accomplishments, and the state of the art is progressing rapidly.

We expect Reinforcement learning to perform better even in more ambiguous environment as long as we are able to run almost infinitely number of iteration and we have a correct environment function that takes current state and action as input and output the reward.

Open AI Gym is a toolkit for Reinforcement Learning research, which contains a collection of RL environments with a common interface. Besides, it also provides researchers with a community through a website(gym.openai.com), where people can share results and compare different algorithms.

Reinforcement learning algorithms are trying to maximize the total reward as an agent interacts with the environment. To achieve this goal, Gym breaks an agent's experience into different episodes, randomly initializes agent's state from a distribution and maximize the expectation of total reward per episode, thus achieving a higher level of performance in as few episodes as possible.

There are multiple environments available in Gym, ranging from easy to hard and contains a lot of classic benchmarks. For example, the Atari Environment contains classic Atari games, 2D and 3D robots can control a robot in simulation, classic control and toy context includes a bunch of small-scale tasks from RL literature. By sharing a common interface Gym allows people to develop general RL algorithms. Among these environments, we can see there exists a lot of interesting applications of reinforcement learning in computer vision.

Existing reinforcement learning methods can divide into two approaches, value-base and policy-based. They typically use convolutional neural networks as part of a policy network or Q-network. In this work, we will analyze different methods in OpenAi Gym Atari, such as DQN, double DQN, policy gradient, actor critic and so all, from classical algorithms to state-of-the-art algorithms. We plan to reimplement these algorithms and compare their properties and performance.

Time	Goal
Feb 15th - Mar 1st	Read related research papers and create the webpage
Mar 1st - Mar 15th	Set up environment and try some toy examples
Mar 15th - Mar 24th	Try classical algorithms (e.g. DQN, policy gradient)
Mar 24th - Mar 27th	Write mid-term report
Mar 27th - Apr 12th	Try state-of-the-art algorithms
Apr 12th - Apr 19th	Analizy results
Apr 19th - May 6th	Prepare the presentation, meanwhile finish the webpage