Siyue Wu
Zhecheng Wang
Zeyu Jin

Exploring Cooperation and Competition Mechanisms in Multi-Agent DQN

Background/Motivation

Multi-agent system has long been a hot topic in reinforcement learning, where agents are interacting with each other through cooperation and competition. On the other hand, deep learning has seen a rapid resurgent in recent years, making reinforcement learning successful in a much more complex and stochastic environment with extremely higher-dimension states. In this work, we are aimed to leverage **sharing experience replay** and **evolutionary elimination and update** to boost multi-agent deep reinforcement learning. Our target is to explore effective mechanisms of cooperation and competition for (1) shortening the training time compared to independent learning and (2) reaching a higher overall performance in multi-agent system than learning independently.

Games

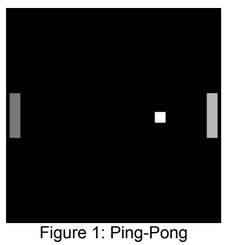


Figure 1: Ping-Pong

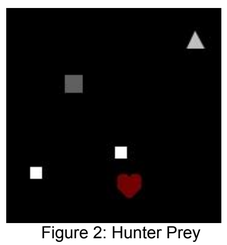
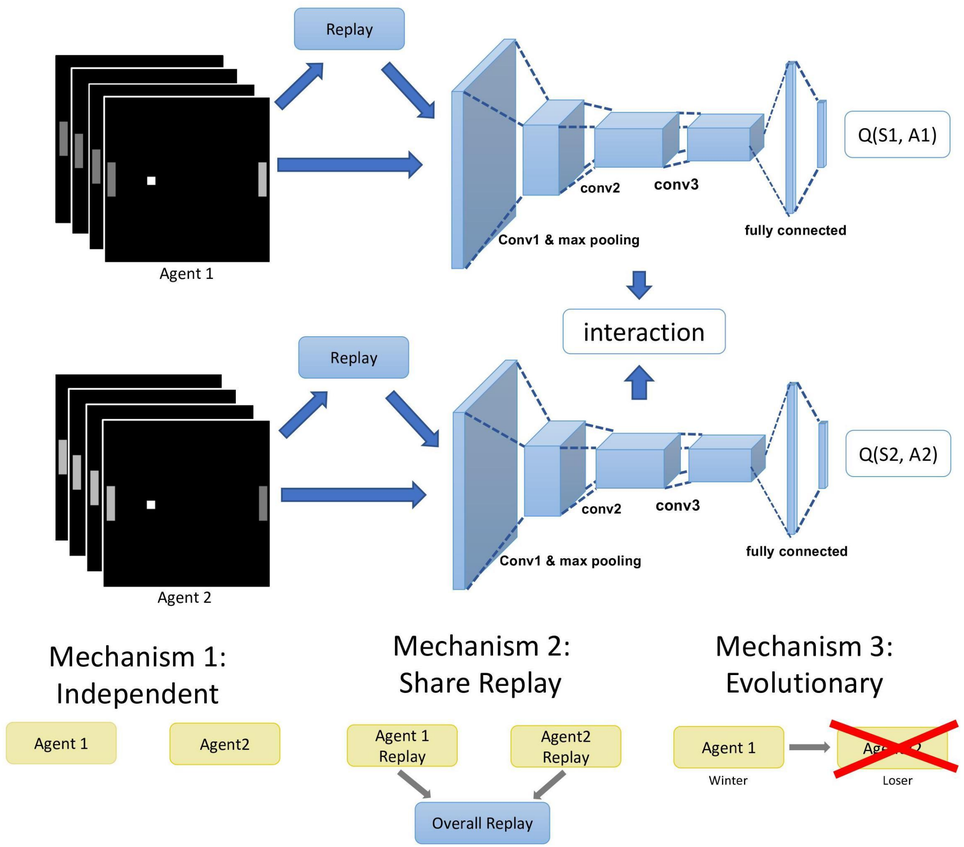


Figure 2: Hunter Prey

Ping Pong: two paddles (two agents) and one ball. The reward for failing catching the ball is -10 and the other agent gets reward +10 in the meantime.

Hunter Prey: two hunters (dark grey square and light grey triangle), two preys (white squares) and some bonus (red heart). Reward for keeping alive is +1, reward for catching the bonus is +10 and reward for being killed by hunter is -10.

Models/Algorithms



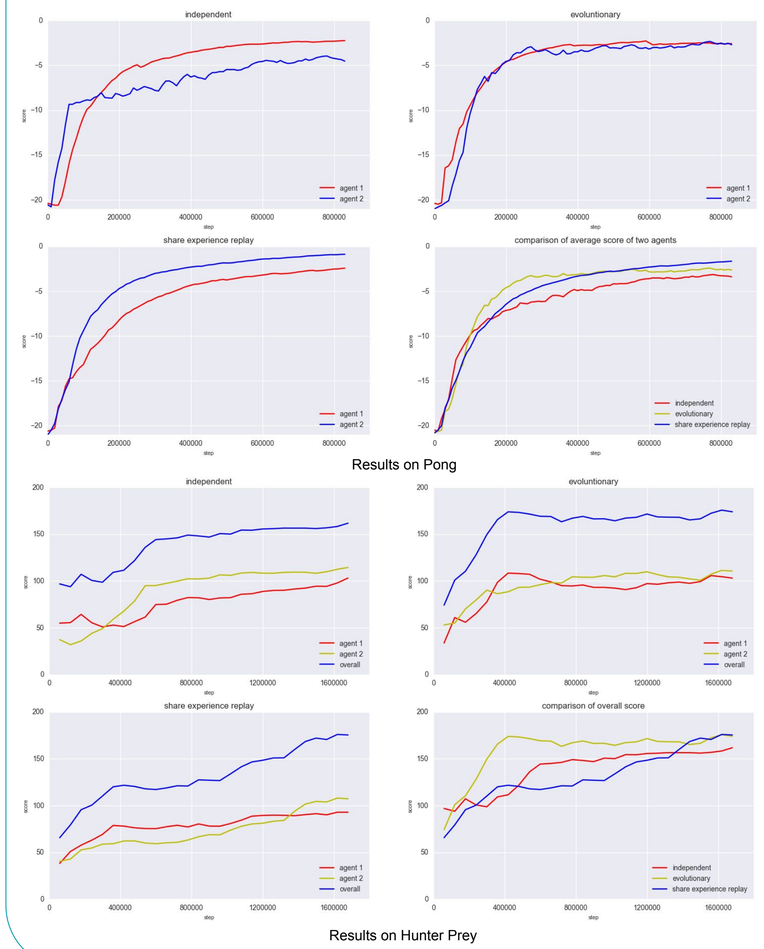
Our model is based on Deep Q-Learning. An agent regards other agents as a part of environment. We developed two different mechanisms for competition and cooperation of agents and compared them with independent learning.

Independent: two agent are assumed to act independently, competitively and do not share any information.

Share replay (cooperative): there is only one replay deque for both agents, both model would get the batches from same replay memory.

Evolutionary (competitive): inspired by natural selection, we evaluate the performances of two agents periodically, then replace the model parameters of loser with that of winner, and continue learning.

Results



Results on Pong

Results on Hunter Prey

Discussion & Future Work

- For both Pong and Hunter-prey, adding evolutionary mechanism can speed up training. Moreover, it can keep all agents to learn synchronously, while independent and replay sharing will result in gap between agents.
- Both replay sharing and evolutionary mechanisms can reach higher scores than independent learning.
- Future work can be focused on designing more elaborate evolutionary and replay sharing mechanisms.