Proposal

Chuqiao Yuan and Jay Li

November 23, 2018

1 Discussion of the Problem and Algorithms

We will investigate **multi-agent reinforcement learning algorithms for partially observed Markov decision processes** with the test environment of **Pacman game**.

Specifically, we will develop algorithms for multiple ghosts with partial observation of the game state, with the goal of capturing the Pacman. The algorithms we will develop are predictive multi-agent RL, Bayesian RL, and multi-agent RL with knowledge transfer. The specifics of these algorithms are discussed in the Algorithms section.

To test and measure the performance of developed algorithms, we will modify the Pacman Project developed at Berkeley to model the environment we will test against.

2 Related Course Topics

Related course topics include Bayesian networks, Markov decision process, hidden Markov models, and reinforcement learning.

3 Research Question

We want to focus on realizing and comparing different multi-agent reinforcement learning algorithms for POMDPs. We will implement the algorithms in the following section for ghost agents and try to compare the time efficiency, space efficiency, winning rates and total values of these algorithms under the scenarios of different ghost numbers, different incomplete information retrieval methods and different layouts.

4 Algorithms

Our project will cover 3 multi-agent reinforcement learning algorithms. The ideas and intentions behind the algorithms are quite different, so it is interesting to learn performance of the 3 algorithms under distinct conditions. We will further discuss about the architecture of the algorithms below.

1. Predictive Multi-Agent Reinforcement Learning

Marinescu, Dusparic, Taylor, Cahill and Clarke (2014) [2] proposed predictive-MARL to solve multi-agent games under dynamic and uncertain environments. The key idea is to set up a neural network to infer about the world based on historical observations and changes in the world. With predictions provided by the neural network, each agent is developed by Q-learning algorithm individually. In general, it is a decentralized multi-agent reinforcement learning algorithm with a built-in neural network.

2. Bayesian Reinforcement Learning

Amato and Oliehoek (2013) [1] tried to apply Bayesian adaptive techniques to solve POMDPs. They used specific prior distribution to model the prior distribution of the initial states and updated their knowledge about the world using posterior distributions of the observations. They provided two frameworks. One is for a group of agents with full communication, another is for a group of decentralized agents developing their best-response models. To model the realistic environment, we will adopt the model with decentralized agents.

3. Multi-agent Reinforcement Learning with Knowledge Transfer

To examine centralized multi-agent reinforcement learning algorithm, we will discover about the negotiation-based algorithm proposed by Zhou, Yang, Chen and Gao (2016)[3]. We will implement communication between multiple agents based on this algorithm and see how communication between agents will impact the final performance of the agents.

5 Expectation

We will implement the above 3 algorithms for the ghost agents in the Pacman framework. The ghosts should be able to collaboratively capture the Pacman with fair success rates depending on the algorithms and environments. We expect to compare the performance of different algorithms under diverse environment parameters and explain why specific approaches perform the best under some environments.

6 Group Members

Jay Li and Chuqiao Yuan. We will develop algorithms, tests as well as reports together. The amount of work will be about equal.

References

[1] C. Amato and F. A. Oliehoek. Bayesian reinforce- ment learning for multiagent systems with state uncertainty. 2013.

[2] Taylor A et al. Marinescu A, Dusparic I. Decentralised multi-agent reinforcement learning for dynamic and uncertain environments. 2014.

[3] Chen C et al. Zhou L, Yang P. Multiagent reinforcement learning with sparse interactions by negotiation and knowledge transfer. *IEEE transactions on cybernetics*, 2016.