COOPERATION AND COMPETITION IN SEQUENTIAL SOCIAL DILEMMAS

MULTI-AGENT REINFORCEMENT LEARNING

ISH HANDA 3035238565

Under the supervision of Professor Francis C.M. Lau

University of Hong Kong

Computer Science

Introduction

Artificial Intelligence has been a hot research topic to solve problems and accomplish tasks in the STEM, robotics, economics, and many other fields. My project explores the application of AI in economics and psychology. It focuses on understanding the dynamics of cooperation and competition in multi-agent systems, using multi-agent deep reinforcement learning.

Reinforcement Learning

A Reinforcement learning agent is a type of machine learning inspired by behaviorist psychology, where a software agent learns by interacting with an environment to maximize its cumulative reward along the course of its interactions. It is often referred to as the science of making optimal decisions. There are state of the art provably convergent algorithms for solving the single-agent RL problems and tasks. These algorithms used by the software agent to determine its actions is called its policy. These policies can be deterministic, stoichiometric or even neural networks with observations as input and actions as outputs.

RL vs other ML architectures

1. Unsupervised - In RL, there is no supervisor, the agent gets a reward (positive or negative) for its action either after each time step or at the end of the task.
2. Performance-Based - in RL, there is no accurate input or output, the algorithm focuses on finding the balance between exploration and exploitation i.e. investigating the environment or staying in the confines of the given information.
3. Sequential – in RL, data is seen as sequential time steps and the previous state's actions affect the next time step.

Social Dilemmas, Game Theory and Matrix Games

A social dilemma is a social situation where an individual can profit from being selfish unless everyone decides to do so, then everyone loses. Game theory is the term used by economists to understand such social situations. An important characteristic of game theory is that one person's reward is determined by not just his/her action but the actions of everyone involved. The notion of matrix games is a simple framework in game theory that deals with scenarios with multiple agents but just one state with an associated reward. They can be used to represent each player's reward structure as an n-dimensional matrix [5]. Two well-known two-person matrix games are Prisoner's Dilemma (figure 1) and Rock-Paper-Scissors (figure 2).

		Player A: Cooperate	Player A: Betray
Player B	Cooperate	1, 1	0, 3
	Betray	3, 0	2, 2

Figure 1: Matrix form of Prisoners Dilemma

		Player 2: Rock	Player 2: Paper	Player 2: Scissors
Player 1	Rock	0, 0	1, -1	-1, 1
	Paper	-1, 1	0, 0	1, -1
	Scissors	1, -1	-1, 1	0, 0

Figure 2: Matrix form Rock Paper Scissors

Multi-Agent Reinforcement Learning

Combined with the simplicity and generality of the setting of RL as well as the philosophy of game theory opens a new field of research, known as multi-agent reinforcement learning (MARL). MARL can have potential benefits, being that they can make parallel computation faster as multiple agents can learn new information from sharing experiences through communication, teaching, or imitation. However, it also faces challenges those of single-agent reinforcement learning of exploration vs exploitation and curse of dimensionality. New problems also arise in multi-agent systems such as identifying specific goals and the need for coordination among agents and policies [4].

Original Paper

Matrix-based games tend to assume the choice whether to cooperate or not as binary actions, limiting the application of a social dilemma like approach in multi-agent systems. Google DeepMind introduced the concept of sequential social dilemmas [1] that constitute of these matrices based social dilemmas along with the requirement for agents to learn policies that incorporate strategic planning. Using Deep Q-network algorithm as the learning framework for multiple self-interested independent agents, they analyzed the dynamics of the policies learned in two Markov game 1. Fruit gathering game and 2. Wolfpack hunting game. The analysis reflected on the dependency of varying environmental factors on the learned policies, elaborating on the emergence of confliction from competition and the effect of the sequential nature of real-world social dilemmas on cooperation.

Paper Idea

This idea is an extension to the above, to further our understanding of the dynamics of policies learned by multiple agents. The above paper assumed that agents are self-interested independent agents, however, this case is not always the reality. This project would consider agents with different degrees of awareness (independent, fully aware, and in-directly tracking) built upon on the same Deep Q-network framework. A comparative analysis would be performed on a combination of these awareness levels. This project will also consider another scenario wherein multiple agents could actually be based on different degrees of awareness (for example, one agent-independent and other one is fully aware), tackling a more real life like scenario, wherein different people would have different mindsets.

Objective

The aim of this final year project would be to write a research paper on the mentioned Idea above. This is will help in providing a different analysis in multi-agent system theory to share with the world, helping push this industry forward.

Methodology

Since this research area is a new area of interest, there are not many multi-agent reinforcement learning toolkits or packages, but there are a few single agent reinforcement learning toolkits which allow us to make custom environments and algorithms which can incorporate multi-agent systems as well.

OpenAI Gym

The gym toolkit for python [2] would serve as the base for developing the Deep Q-network based reinforcement learning algorithm. The gym library provides an easy-to-use platform for the development of reinforcement learning algorithms as compared to other platforms such as google dopamine which are either very new and often lack the functionality that gym can offer.

Unity and ML-agents

Unity and ML-agents, an asset package by unity for developing reinforcement learning algorithms [3], will be used in developing the environment for the two Markov games mentioned above, fruit gathering game and wolfpack hunting game for our comparison. The same environment could be developed using gym, but Unity provides a visual and easy drag-drop with less code process for building a custom environment. It also integrates 2D/3D models of multiple formats as well as other multimedia tools, allowing the environment to be more aesthetically pleasing as well as easy to understand.

TensorFlow/Keras

The Deep Q-Network algorithm is the reinforcement learning algorithm of choice of this paper, that integrates neural networks in the decision-making aspect of the re-enforcement algorithm. Keras, wrapper of TensorFlow or TensorFlow would be used to create the neural network. These libraries are official machine learning libraries developed by Google which are efficient and well reputed as compared to other options available.

Project Schedule

The project will be treated as a software engineering life cycle, with the schedule as follows

Life Cycle		Time
Inception	Research Study	Sept 2018 - Oct 2018
Elaboration 1	a) Create an Outline of Research b) Write a Thesis Statement	Oct 2018
Elaboration 2	Code Development a) Develop both the Environments b) Develop the RL algorithms	Nov 2018 - Jan 2019
Construction	a) Conduct Research and Analysis b) Draft Paper	Feb 2019 - March 2019
Transition	Revise and Fill Gaps	April 2019

References

[1] Joel Z. Leibo, Vinicius Zambaldi, Marc Lanctot, Janusz Marecki, Thore Graepal. Multi-agent Reinforcement Learning in Sequential Social Dilemmas. Google DeepMind. arXiv:1702.03037.

[2] Greg Brockman and Vicki Cheung and Ludwig Pettersson and Jonas Schneider and John Schulman and Jie Tang and Wojciech Zaremba. OpenAI Gym. OpenAI. arXiv:1606.01540

[3] Juliani, A., Berges, V., Vckay, E., Gao, Y., Henry, H., Mattar, M., Lange, D. (2018). Unity: A General Platform for Intelligent Agents. arXiv preprint arXiv:1809.02627. https://github.com/Unity-Technologies/ml-agents.

[4] L. Bus¸oniu, R. Babuska, and B. De Schutter. Multi-agent reinforcement learning: An overview, Chapter 7 in Innovations in Multi-Agent Systems and Applications. (D. Srinivasan and L.C. Jain, eds.), vol. 310 of Studies in Computational Intelligence, Berlin, Germany: Springer, pp. 183–221, 2010.

[5] Gon¸calo Neto. From Single-Agent to Multi-Agent Reinforcement Learning: Foundational Concepts and Methods, 2005.