Machine Learning Project: part 1

Ignace Bleukx

March 17, 2020

1 Literature Review

In this first assignment we try to combine basic principles from game theory with the work concerning multi-agent reinforcement learning. Most literature included in this literature review will therefore more of less fall into one of these categories. First we give an overview of the relevant literature. Afterwards, we give a detailed list of the contributions for each paper.

Shoham and Leyton-Brown [5] introduces elementary concepts from game theory. Bloembergen [1] introduces basic concepts from multi-agent systems, and explains how reinforcement learning algorithms can be used to reach equilibriums in simple games. Replicator dynamics are introduced to model evolutionary concepts in multi-agent systems. Bloembergen [1] also introduces lenient reinforcement learning to overcome difficulties when bad initial exploration leads to convergence to wrong equilibria.

We use the game-theoretic reinforcement learning framework OpenSpiel for all experiments. The practical details are outlined in Lanctot et al. [3]. Details about solving the Prisoner's Dilemma using reinforcement learning algorithms are found in Harper et al. [2].

Article	Contribution
Multi-agent systems: Algorithmic, Game-Theoretic,and Logical Foundations, Shoham and Leyton-Brown [5]	Game theory (utility, payoff functions, strategies, zero-sum games, Pareto optimality, Nash equilibria, existence of Nash equilibria), Finding Nash equilibria (minmax and maxmin algorithms)
Multi-agent learning dynamics, Bloembergen [1]	Multi-agent systems as a way to solve many problems using sensor data as input and rewards as output, evolutionary modelling (replicator dynamics as selection strategy)
OpenSpiel: A Framework for Reinforcement Learning in Games, Lanctot et al. [3]	The OpenSpiel framework: installation, design, implemented games and algorithms, visualization
Reinforcement learning produces dominant strategies for the Iterated Prisoner's Dilemma, Harper et al. [2]	Definition of Prisoner's dilemma, example values for parameter tuning (learning rate, discount factor)
The replicator equation on graphs, Ohtsuki and Nowak [4]	Phase plots as a visual representation of graphing evolutionary policies by using replicator dynamics.

References

[1] Daan Bloembergen. *Multi-agent learning dynamics.* PhD thesis, 05 2015.

[2] Marc Harper, Vincent Knight, Martin Jones, Georgios Koutsovoulos, Nikoleta E. Glynatsi, and Owen Campbell. Reinforcement learning produces dominant strategies for the iterated prisoner's dilemma. *PLOS ONE*, 12(12):1–33, 12 2017. doi: 10.1371/journal.pone.0188046. URL `https://doi.org/10.1371/journal.pone.0188046`.

[3] Marc Lanctot, Edward Lockhart, Jean-Baptiste Lespiau, Vinicius Zambaldi, Satyaki Upadhyay, Julien Pérolat, Sriram Srinivasan, Finbarr Timbers, Karl Tuyls, Shayegan Omidshafiei, Daniel Hennes, Dustin Morrill, Paul Muller, Timo Ewalds, Ryan Faulkner, János Kramár, Bart De Vylder, Brennan Saeta, James Bradbury, David Ding, Sebastian Borgeaud, Matthew Lai, Julian Schrittwieser, Thomas Anthony, Edward Hughes, Ivo Danihelka, and Jonah Ryan-Davis. Openspiel: A framework for reinforcement learning in games, 2019.

[4] Hisashi Ohtsuki and Martin A. Nowak. The replicator equation on graphs. *Journal of theoretical biology*, 243 1:86–97, 2006.

[5] Yoav Shoham and Kevin Leyton-Brown. *Multi-agent systems: Algorithmic, Game-Theoretic,and Logical Foundations.* 2009.