Machine Learning Project: part 1

Ignace Bleukx AND Quinten Bruynseraede

March 19, 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, along with their contributions.

Article	Contribution
Multi-agent systems: Algorithmic, Game-Theoretic,and Logical Foundations, Shoham and Leyton-Brown [7]	This paper provides a thorough explaination of the different aspect of game theory, including different types of equilibria. These concepts are of importance to us since we will investigate whether or not our learning algorithms converge to one of these equilibria. Furthermore, the paper provides a detailed description of different types of games, such as cooperative games and non-cooperative games, as well as the notion of games in normal form.
Multi-agent learning dynamics, Bloembergen [2]	This thesis on multi agent learning dynamics provides essential information about different game theory aspects. Not all sections are relevant for our initial research on matrix games. Mainly section 2.3 on evolutionary game theory and chapter 3 are relevant. In this last chapter, the replicator dynamics of many matrix games are investigated and explained very clearly. In this chapter we find an example of the learning pattern we would like to observe with our application of different learning algorithms.
OpenSpiel: A Framework for Reinforcement Learning in Games, Lanctot et al. [5]	The paper provides the documentation of the OpenSpiel framework. All aspects of the library are explained, from installation to implemented algorithms and games. Many design choices of the framework are clarified which helps to understand the philosophy behind the framework. In the paper, the game theory aspects are briefly touched upon, as well as important concepts of the implemented learning algorithms. This paper is of very much importance to us as we will use (and potetially extend) the OpenSpiel framework for this assignement.
Reinforcement learning produces dominant strategies for the Iterated Prisoner's Dilemma, Harper et al. [4]	This document contains a detailed description of the prisoners dilemma. Since this is one of the matrix games we will examine in the first part of the assignement, this belongs to the relevant lecture on this list. Furhermore, some examples of parameters for the training algorithms are given, which will help to produce meaningfull results when training the learning algorithms of choice.

The replicator equation on graphs, Ohtsuki and Nowak [6]	The paper provides an insight on the visualization of the replicator dynamics using phase, as well as some examples relevant to our research. These examples include the prisoners dilemma and biased rock-paper-scissors.
Analyzing Reinforcement Learning algorithmsusing Evolutionary Game Theory, Bloembergen [1]	This thesis provides a rich source of information on the reinforcement learning branch for evolutionary game theory. Many algorithms are examined, some of which are available in OpenSpiel. The paper also contains the exact parameter settings used to achieve the presented results. These paramters can be used by our agents to reproduce favorable results of the paper.
Evolutionary Dynamics of Multi-Agent Learning:A Survey, Bloembergen et al. [3]	Like other papers, this document provides a basic knowledge of game theory, as well as reinforcement learning. For our research, mainly the part about lenient FAQ-learning as a way to increase the robustness of Q-learning, is important. FAQ-learning is able to recover from bad exploration in the start of the run, while normal Q-learning is sometimes not.
Extended Replicator Dynamics as a Key to Reinforcement Learning in Multi-agent Systems, Tuyls et al. [8]	To model stochastic policies, populations of players are used. These populations can be described using evolutionary concepts, such as selection and mutation. This paper explains the transition from regular to evolutionary game theory. We received insight on the dynamics of a population through the central notion of replicator dynamics. These selection mechanisms can be extended with mutation, based on the Boltzmann mechanism. To overcome converge to suboptimal equilibria, lenience towards mistakes is introduced in this paper.

2 Independent learning

3 Dynamics of learning

3.1 Lenient Boltzmann Q-Learning dynamics

Based on Bloembergen et al. [3], we implemented Lenient Boltzmann Q-Learning dynamics. This extension introduces two parameters. The first parameter, κ, is the degree of leniency: the number of rewards that are examined in each step before updating the policy. We expect to see improved robustness when introducing leniency. Intuitively, this is done by increasing the area of attraction for optimal equilibria. Secondly, τ introduces entropy into the population: a learner will favour exploration over exploitation when τ increases. We expect fixed points of the dynamics to stray away from any Nash equilibria as τ is increased. However, introducing a s small amount of entropy may prevent attraction from unwanted equilibria early on.

Figures 1 through 3 show the influence of the parameter κ on the dynamics. For games such as the Prisoner's Dilemma with one equilibrium (in this case located at 0,0), the difference isn't very interesting. However, in Figure 2, we clearly see how there is no longer attraction to the suboptimal equilibrium at $(\frac{2}{5}, \frac{3}{5})$. The influence of τ is very clear, as shown in Figure 4. As τ increases, exploration of new solutions is greatly preferred. Therefore, an agent will not further capitalize on improvements in his strategy. As a result, policies seem to converge to random guessing.

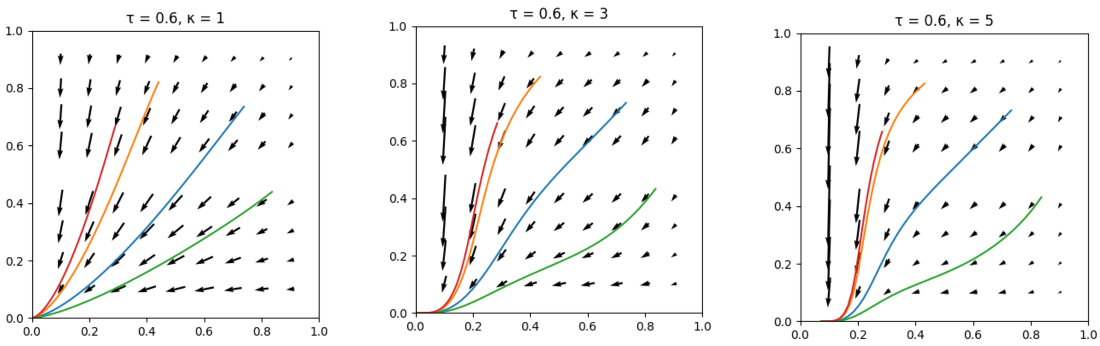

Figure 1: LFAQ for the Prisoner's dilemma

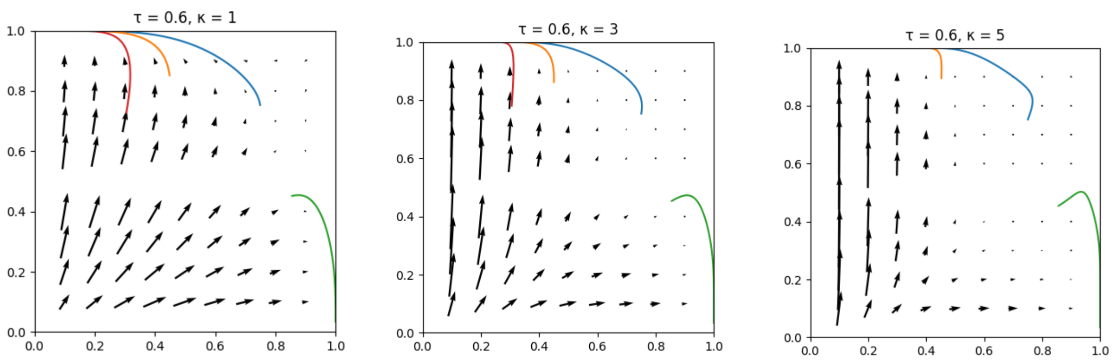

Figure 2: LFAQ for the Battle of the Sexes

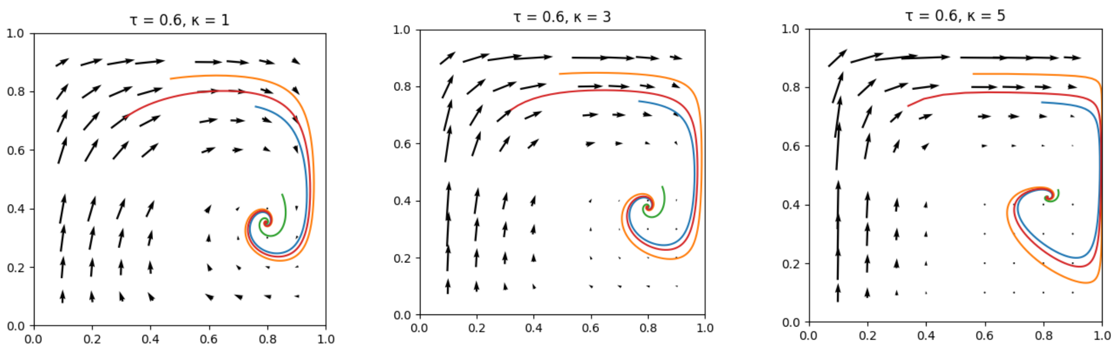

Figure 3: LFAQ for the Matching Pennies game

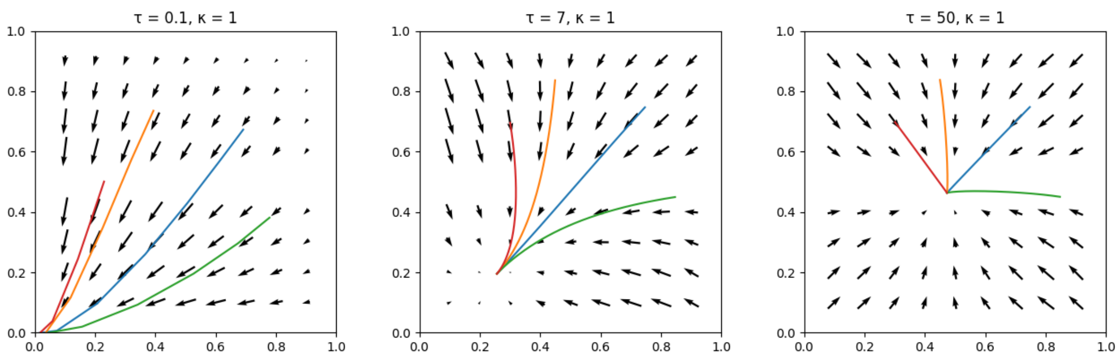

Figure 4: Influence of increasing τ on the Prisoner's Dilemma

References

[1] Daan Bloembergen. Analyzing reinforcement learning algorithmsusing evolutionary game theory. *Journal of theoretical biology*, 2010.

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

[3] Daan Bloembergen, Karl Tuyls, Daniel Hennes, and Michael Kaisers. Evolutionary dynamics of multi-agent learning:a survey. *Journal of Artificial Intelligence Research*, 2015.

[4] 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`.

[5] 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.

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

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

[8] Karl Tuyls, Dries Heytens, Ann Nowe, and Bernard Manderick. Extended replicator dynamics as a key to reinforcement learning in multi-agent systems. In Nada Lavrač, Dragan Gamberger, Hendrik Blockeel, and Ljupčo Todorovski, editors, *Machine Learning: ECML 2003*, pages 421–431, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg.