Autonomous Agents - Assignment 3 Report

Nicolò Girardi, Jorge Sáez Gómez, Johan Sundin

October 17, 2013

1 Introduction

In this assignment we worked on a predator versus prey grid world scenario, in which both the predator and the prey are learning agents [1]. [2] There are small changes made to the grid world scenario compared to previous assignments.

2 Theoretical description of the algorithms

For this assignment we have implemented the following algorithms: Independent Q-Learning [2], Minimax Q-Learning [4], Generalized Infinitesimal Gradient Ascent Win or Learn Fast(GIGA-WoLF) [5]. Each algorithm will be described concisely below.

2.1 Independent Q-Learning

In Independent Q-Learning each agent in the environment treats all other agents as being part of the environment. An agent employing Independent Q-Learning does not try to model other agents, or try to learn their behavior [2]. The agent then uses Q-Learning to learn a state-actionvalue function, mapping actions in states to values. A problem with this algorithm is that the Q-Learning algorithm assumes a static environment, in which state transition probabilities do not change. This assumption is not valid for independent Q-Learning, as other learning agents might change their behavior over time. Since the agent models other agents as part of the environment, the environment thus becomes dynamic.

2.2 Minimax Q-Learning

2.3 GIGA-WoLF

3 Method

Something here...

3.1 Scoring system

$$score = reward \cdot 0.9^{T-1}$$

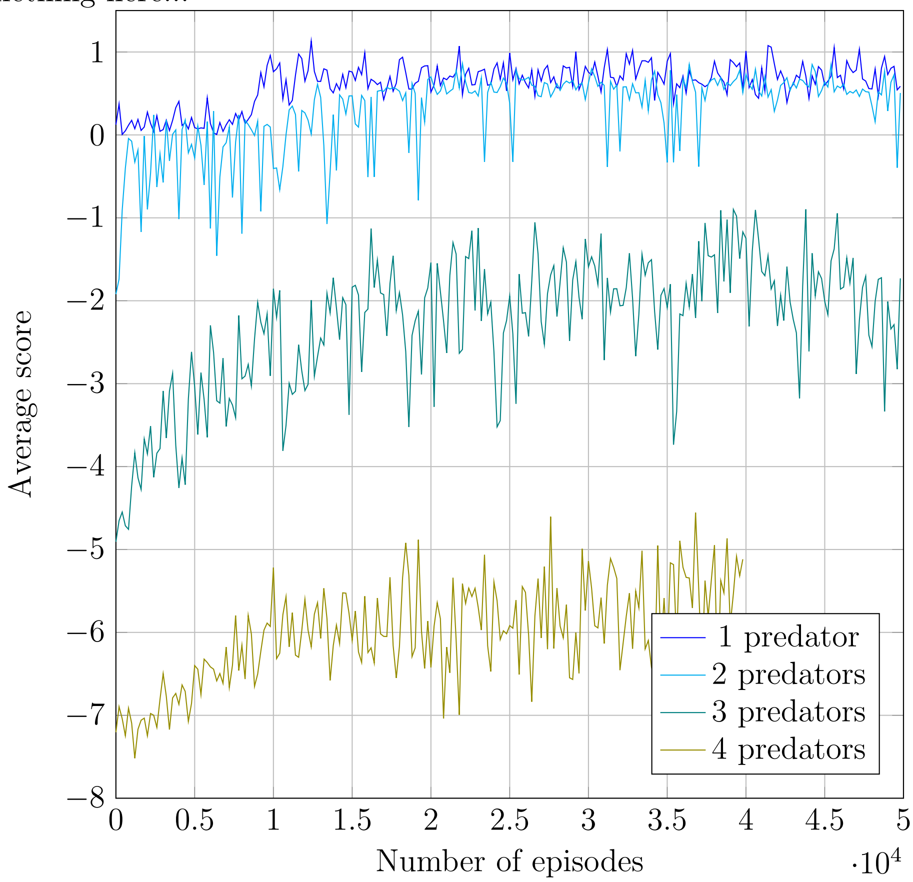
3.2 State-space encoding

In order to make the presented algorithms converge as fast as possible, we made use of the reduced state-space we devised for the first assignment. However, we had to adapt it to the multi-predator scenario in order to make it work with all cases we will be considering for this assignment. To this end, several ideas needed to be introduced.

First of all, we will extend our original idea for a reduced state-space encoding and use the distances from a certain agant to the rest of them, instead of encoding the positions of every agent on the world. This allows us to use two less variables for the state-space encoding. Furthermore, it is now essential to know which agent "we are", and this needs to be encoded in the state representation somehow. If this information was not encoded, then it would be impossible to distinguish between predators in a multi-predator setting, and the algorithms would not be able to know which predator should be applied a certain action from the usual set {North, South, East, West, Stay}.

4 Results

Something here...



5 Discussion

Something here...

6 Conclusion

Something here...

References

[1] *Assignment.*
Assignments Autonomous Agents, D.M Roijers, September 2, 2013

[2] *Vlassis Nicos*
A Concise Introduction to Multiagent Systems and Distributed Artificial Intelligence
Morgan and Claypool Publishers, 2007

[3] *Sutton and Barto.*
Reinforcement Learning: An Introduction, Richard S. Sutton, Andrew G Barto.
`http://webdocs.cs.ualberta.ca/~sutton/book/ebook/`

[4] Michael L. Littman
Markov games as a framework for multi-agent reinforcement learning
Brown University / Bellcore
`http://www.cs.rutgers.edu/~mlittman/papers/refer.html`

[5] Michael Bowling
Convergence and no-regret in Multiagent Learning
`http://webdocs.cs.ualberta.ca/~bowling/papers/04nips.pdf`