

Project 2

Nicolas Cofre
nicolas.cofre@gatech.edu

Date: March 15, 2020

Abstract

This project consists in the solution of the Lunar Lander task from OpenAI Gym environment. The objective is to safely land the spacecraft between two flags. The use of the engine and missing the landing target are costly. Using the Deep Q Network reinforcement learning methodology I look for a policy that maximizes the reward.

1 Introduction

Deep Q Network (DQN) is closely related to the well-known tabular Q-learning algorithm. The DQN implementation used here can be briefly described as follows:

- The Q function is approximated using a neural network, that maps the state space S to the action space A . $Q(\theta) : S \rightarrow A$
- The $Q(\theta)$ has a lagged copy or target network $Q(\theta^-)$
- There is a memory buffer which is used draw samples and minimize $L(\theta) = Q(\theta) - r - \gamma Q(\theta^-)$ for the observed rewards r

2 Parameter tuning

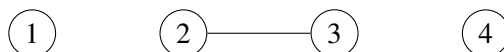
Hovering was observed if gamma was set too high, for instance $\gamma = 0.99$. Intuitively, this means that the spacecraft is very patient, meaning that it prefer to delay the reward of the landing at the expense of more fuel consumption but avoiding to crash. If we decrease the value of gamma, the spacecraft will be more in a hurry to get the landing reward before the time decreases the payoff due to discounting.

The size of the buffer

The exploration rate

The lag of the target network update

The discount rate (maybe it will prefer to fly)





This template is based on the standard L^AT_EX article class, hence the arguments of article class are acceptable (a4paper, 10pt and etc.). Alternative engines are pdfL^AT_EX and X_YL^AT_EX.

References

- [1] V. Quadrini, “Financial Frictions in Macroeconomic Fluctuations,” *FRB Richmond Economic Quarterly*, vol. 97, no. 3, pp. 209–254, 2011.
- [2] C. T. Carlstrom and T. S. Fuerst, “Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis,” *The American Economic Review*, pp. 893–910, 1997.
- [3] Q. Li, L. Chen, and Y. Zeng, “The Mechanism and Effectiveness of Credit Scoring of P2P Lending Platform: Evidence from Renrendai.com,” *China Finance Review International*, vol. 8, no. 3, pp. 256–274, 2018.