WHY WOULD YOU DO THAT?

BCS-IIT Kanpur : Semester Project 2021

Akanksha Singh, Ayushi Chaudhary, Siddhant Singh, Subiksha Shree, Suyash Mallik

INTRODUCTION

Neuroeconomics seeks to explain human decision-making, the ability to process multiple alternatives and to follow a course of action. In general, a population of people thrives when they depict certain social behaviors. Small individual choices can have huge impacts on the population. In this project we'll be taking a look at how multi-agent systems interact and produce macro-effects as a result of micro-choices, and how those results in turn affect the future decisions of the agents.

OBJECTIVE

Learn about evolutionary game theory and use deep learning methods to implement game theoric models in a virtual environment. Develop a Q-learning model to choose the optimal strategy for survival from a given set. The project was focused on developing a model that is close to human decision-making. The objective of the simulations designed for this project is to find the traits that would survive with time in an environment constrained by certain conditions and probabilities. The project is based on multi-agent interactions and reinforcement learning. A range of factors has been considered to implement the simulations based on OOP concepts, deep learning, and game theory models.

THE STRATEGIES

- Always cooperative (AC) : Agents assigned the maximum probability to share food and get food in return.
- Tit-for-tat (TFT) : Agents decide whether they will share their food or not based on the history of the agent-at-mercy (needy agent).
- Alternatively cooperate (ALT) : Sharing alternates between cooperativeness and competitiveness for every iteration of gathering food.
- Always defective (AD) : Agents are assigned the least probability to share the food they have acquired.
- "Intelligent" : Actions taken by the s-agent in the QL model

ABSTRACT

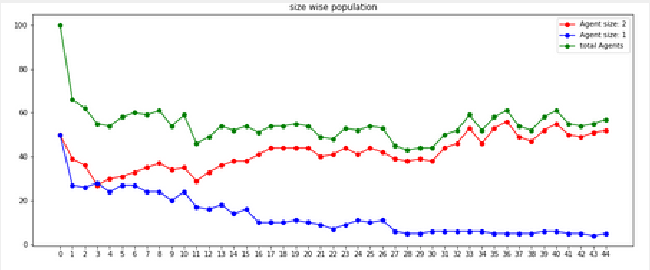
Cooperativeness/Competitiveness and Altruism/Selfishness are two significant features adopted by natural algorithms that observe substantial variations and improve the quality of the existing solutions. The agents are introduced in an environment where they learn to choose an optimal strategy for their survival.
When with excess food, the agents learn an optimal strategy to survive their way through the environment by estimating which immediate action is available to it at night during the time of sharing to achieve evolution

FRAMEWORK

A Q-table is initialized with (states, actions). The Q-table is global in our case to ensure a larger collection of data to refer to. Here,
states= the strategy of the needy agent that our agent is interacting with.
actions= strategy chosen by the s-agent to deal with the needy agent.
The agents get a positive reward if either the population of other agents lowers down as compared to when compared to the previous iteration of its own population increases and vice-versa: thereby making the environment more competitive. A larger number of agents are employed to obtain an optimal policy with the accuracy.

CONCLUSIONS

Lack of a punishing mechanism gives rise to cheaters and decline of cooperators within the environment.TFT strategy is assumed a good punishing mechanism for the cheaters, namely the AD agents and sometimes the ALT agents. Factors that may not be directly relevant to the act of cooperation (here, size), can affect the evolutionary pattern of the environment. In a general multi-agent environment, seemingly irrelevant factors may prove to be disadvantageous for the agents that lack them.1. The ALT strategy became vestigial to the agents.The agents learn a complex sharing-behaviour that is a mix of mostly AC and TFT.



Size-wise development of the environment when 4 staunch agents present
Size 2>Size 1

Strategy-wise development of the environment when 4 staunch agents present

Strategy-wise development of the environment when 4 staunch agents plus 1 learning agent present

Literature

1. C. WICKRAMAGE AND D. N. RANASINGHE, "MODELLING ALTRUISTIC AND SELFISH BEHAVIOURAL PROPERTIES OF ANT COLONY OPTIMISATION," 2014 14TH INTERNATIONAL CONFERENCE ON ADVANCES IN ICT FOR EMERGING REGIONS (ICTER), 2014, PP. 85-90, DOI: 10.1109/ICTER.2014.7083884.
2. STEPHENS, C. (1996). MODELLING RECIPROCAL ALTRUISM. THE BRITISH JOURNAL FOR THE PHILOSOPHY OF SCIENCE, 47(4), 533-551. RETRIEVED JUNE 18, 2021, FROM HTTP://WWW.JSTOR.ORG/STABLE/687923
3. AXELROD, R. (1984). THE EVOLUTION OF COOPERATION. NEW YORK: BASIC BOOKS.
4. LI, YUXI. (2018). DEEP REINFORCEMENT LEARNING. RETRIEVED JULY 18, 2021, FROM HTTPS://ARXIV.ORG/ABS/1810.06339V1

MENTOR: SHIVANSHU TYAGI

https://github.com/ayucd/why-would-you-do-that
https://github.com/s-akanksha/Why_Would_You_Do_That