BCS - Neuroeconomics

Takneek-2023

Pool - Aryans

Rishav Bikarwar - 200792
Aditya Bangar - 210069
Shubham Devang Patel - 210709
Mridul Gupta - 220672
Sagar Arora - 220933

Final Documentation

April 2023

Introduction

The world defined in the problem is a grid consisting creatures named as Macpan. They are allowed four types of movements i.e up, down, right, left. Canteens exists at random places in the world to give out food necessary for Macpen to reproduce(1 Macpan $\rightarrow$ 2 Macpen) with food equally distributed among them.

There exists a *ghost gang* which takes a certain amount of food from everyone and if the food level of Macpan falls below 0 they die.

There exists 3 types of Macpen:

1. Helpful - Which gives some amount of food to other if present in excess.
2. Ungrateful - Doesn't share food at all.
3. Tit-for-tat - Will share food with the others based on their history with the grid community.

In this problem, we are required to analyze the type-wise population development of the grid world by simulating a virtual environment in python.

Environment

Here we'll present all about the world and its parameters.

General

Below mentioned are all controlled parameters

- Height & Width of the grid.
- Number of canteens in the world.
- Starting number of Macpen.
- Initial number of helpful, ungrateful & tit-for-tat Macpen.
- Number of days of simulation and hours per day.
- Reproduction threshold.
- Food given by canteen to a Macpan in a day.
- Food taken by the ghost gang each time they visit.
- Hunger: food lost while taking each step
- Vision Radius is the radius when an unvisited canteen becomes visible to a macpan

Macpen

- Each Macpan is given an ID.
- Initially, all Macpen spawn randomly all over the world. Just after spawning, each Macpan has some food which is a random value between MIN_START_FOOD and $FOOD\_THRESHOLD\_FOR\_REPRODUCTION$ and they scourge over the world in search of canteens while maintaining a history and thus developing an understanding about the world.
- All Macpen continuously move by 1 step in every iteration (every 1 hour) and never rest. When they die, they are despawned off the world.
- Each Macpan has a list named history_list which stores the grids coordinate visited and food acquired at that particular grid.
- Every Macpan is given a donate attribute which which classifies them into helpful if positive and ungrateful if negative.
- After reproducing each Macpan is divided into 2 Macpen with equal amount of food distributed from the amount that parent had at the time of reproduction. Each child Macpan is spawned into the same cell where the parent reproduced.
- Each child Macpen retains no history of their parent and their behavior (helpful, unhelpful or tit-for-tat) is the same as that of their parent.

Canteen

- Initially, canteens spawn randomly across the entire game board.
- The position of the canteens change every 5 days to remove saturation from the simulation
- Canteen provides a certain amount of food to a Macpan at most once a day when a Macpan visits the canteen. Each canteen thus maintains a list of people it has fed in that day as well as the available food. At the end of each day, its records are reset.
- Canteen operates all day.
- Canteens have a limit to the total food they can provide to all the Macpen combined per day.

Donation Rate

- The donation rate for helpful and ungrateful is fixed 1 & -1 respectively.
- Tit-for-tat case:

- For tit-for-tat case initial value is set at 0 and they act as helpful at the very first case.
- After first case, their interactions with others defines their history.
- Donation rate changes dynamically i.e either positive or negative depending upon their previous interactions with the grid community and thus they can donate or request accordingly.

Interactions

We have implemented an interaction policy using which the Macpen interact which each other and depending upon their nature and current state choose whether to share their food with a given Macpan.

Excess food: For a Macpan, excess food is defined as the difference between its current food possession and the amount of food taken away by the ghost gang. Thus, when a Macpan decides to share their food with another, it ensures its survival.

As defined in the problem statement, the helpful will always share its excess food. The decision for tit-for-tat Macpen will depend on the donate parameter. Also, the amount of food donated by a Macpan is the minimum amongst:

- The requirement of the other Macpan to survive, should the Ghost Gang visit.
- The excess currently possessed by the donating Macpan.

When more than 2 Macpen (say n) arrive at a cell, a random combination out of $\binom{n}{2}$ possible combinations take place.

Movement

We have implemented a reward based movement policy along with vision for its movement. We have implemented QLearning algorithm for Macpan to decide its next move through reinforced learning.

QLearning

In the Q-Learning algorithm, the goal is to learn iteratively the optimal Q-value function using the Bellman Optimality Equation. To do so, we store all the Q-values in a table that we will update at each time step using the Q-Learning iteration:

$$q^{new}(s,a) = (1-\alpha)\underbrace{q(s,a)}_{\text{old value}} + \alpha\overbrace{\left(R_{t+1} + \gamma \max_{a'} q(s',a')\right)}^{\text{learned value}}$$

where α is the learning rate, an important hyper parameter that we need to tune since it controls the convergence.

Reinforcement learning involves an agent, a set of states S S, and a set A A of actions per state. By performing an action a , the agent transitions from state a to new state. Executing an action in a specific state provides the agent with a reward (a numerical score).

Rewards

The goal of the agent is to maximize its total reward. It does this by adding the maximum reward attainable from future states to the reward for achieving its current state, effectively influencing the current action by the potential future reward. This potential reward is a weighted sum of expected values of the rewards of future steps starting from the current state.
In our algorithm the rewards are mapped to food available in a cell Macpan enters. There is also a reproduction reward for the Macpan when the Macpan reproduces. The Q-value is basically the quality of the action and is updated using Bellman equation.

Working

Initially, the Macpan has no history so he sets out to explore the world and later on he uses his memory to take actions. To accomplish this we set the initial exploration value to 1 and learning rate as 0.1. As a Macpan gathers food, he learns about the grid and starts moving by his learning to ensure to get food to reproduce. This is done by decaying the exploration probability gradually. The decision made by Macpan also accounts for its future, so it tries to maximise the food he will get in the subsequent moves. This is done by using the discount with next q value. The Macpan will perform one of the valid actions with the maximum Q value.

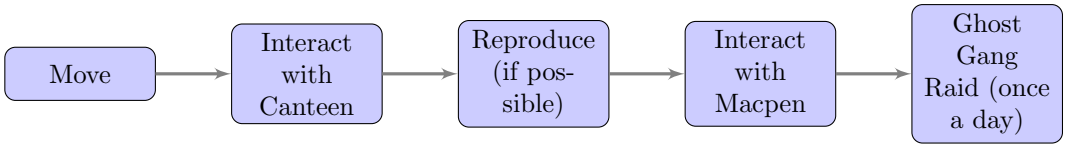
Vision

We have implemented vision along with QLearning to control the movement of Macpan. Whenever a Macpan comes within vision radius from an unvisited

canteen, it becomes visible to Macpan and the Macpan starts moving towards the canteen with a probability of 0.2

Simulation Process

The operations in each iteration are performed in the following order:



In every iteration, the Macpen will move into a new position. If there is a canteen in that cell, it will first take food from the canteen. After this, it will interact with other Macpen present in the same cell. Now, if a Macpan has sufficient food to reproduce, it will split into 2. Next, if the ghost gang is to arrive at the end of iteration, it will do so and sweep off food from every player currently in the grid world. Talk about movement (list, probabilities, movement is random with heuristic)

Other environments- different definition of excess food (helpful are more likely to prosper), Canteen have unlimited food Canteen can share food with a Macpan multiple times in a day. When each Macpan reproduces, its children are of different types.

Analysis

We performed simulations under different initial conditions , keeping initial count of Macpen 300. Macpen of different types formed their own groups in each simulation.
Overall, the tit-tat population continued to thrive in almost all cases where the helpful and unhelpful population would perish. Also, on studying the trends of clustering we observed that helpful and unhelpful population coudn't co-exist. Same is for titat population with positive attribute couldn't exist with unhelpful. There is also an interesting observation that most of the population is found in clusters istead of being spred over the grid. This could be due to the fact that different generations grow together and when ghost gangs come, an entire colony vanishes.
In the end, the following trend is observed in populations over many simulations: titat > unhelpful > helpful. This could be because of the unbaised nature of titat as it learns to adapt to live in conditions that are favourable to either of the populations.

QLEARNING:
LEARNING_RATE=0.1 START_EXPLORATION_PROB=1
MIN_EXPL_PROB=0.4
DISCOUNT=0.9
EXPL_RATE_DECAY=0.001
REPRODUCTION_REWARD=40
VISION_RADIUS=10 VISION_MOVEMENT_PROBABILITY=0.2
ENVIRONMENT:
INITIAL_HELPFUL=100 INITIAL_TITTAT=100 INITIAL_UNHELPFUL=100
CANTEEN_COUNT=50 CANTEEN_FOOD_PER_PERSON=30 CANTEEN_FOOD_DAILY_LIMIT=2400
GHOST_VAL=8
REPRODUCTION_THRESHOLD=40
FOOD_LOST_PER_MOVE=0.041666666666666664

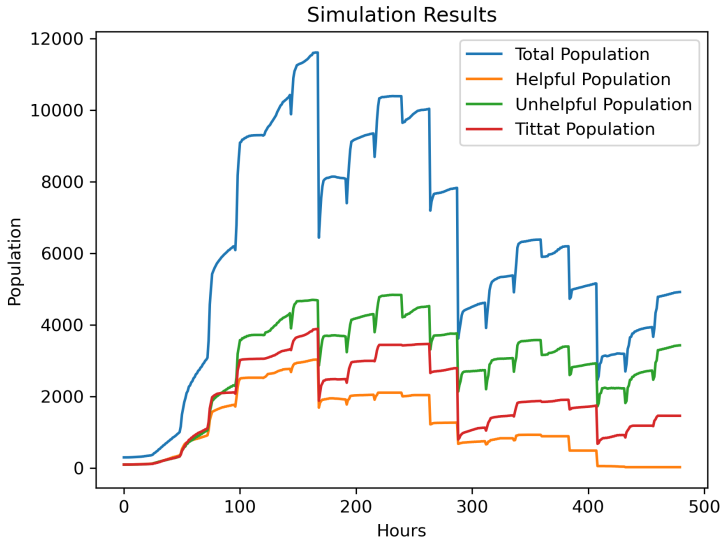


Figure 1: Simulation 1

In the first simulation, we found that the unhelpful population has the highest survival rate.

QLEARNING:
LEARNING_RATE=0.1 START_EXPLORATION_PROB=1
MIN_EXPL_PROB=0.4
DISCOUNT=0.9
EXPL_RATE_DECAY=0
REPRODUCTION_REWARD=50
VISION_RADIUS=10 VISION_MOVEMENT_PROBABILITY=0.2
ENVIRONMENT:
INITIAL_HELPFUL=100 INITIAL_TITTAT=100 INITIAL_UNHELPFUL=100
CANTEEN_COUNT=100 CANTEEN_FOOD_PER_PERSON=50 CANTEEN_FOOD_DAILY_LIMIT=4000
GHOST_VAL=25
REPRODUCTION_THRESHOLD=50
FOOD_LOST_PER_MOVE=0.041666666666666664

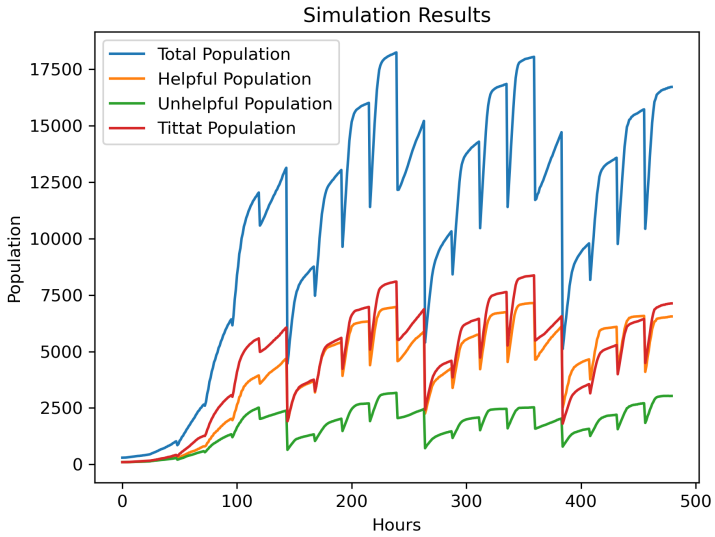


Figure 2: Simulation 2

In this simulation, the ghost value and the canteen count were increased. As a result, the population of unhelpful fell drastically with respect to helpful and titat population. The titat and helpful had almost similar populations at the end.

QLEARNING:
LEARNING_RATE=0.1 START_EXPLORATION_PROB=1
MIN_EXPL_PROB=0.4
DISCOUNT=0.9
EXPL_RATE_DECAY=0
REPRODUCTION_REWARD=50
VISION_RADIUS=10 VISION_MOVEMENT_PROBABILITY=0.2
ENVIRONMENT:
INITIAL_HELPFUL=100 INITIAL_TITTAT=100 INITIAL_UNHELPFUL=100
CANTEEN_COUNT=50 CANTEEN_FOOD_PER_PERSON=70 CANTEEN_FOOD_DAILY_LIMIT=5600
GHOST_VAL=15
REPRODUCTION_THRESHOLD=50
FOOD_LOST_PER_MOVE=0.041666666666666664

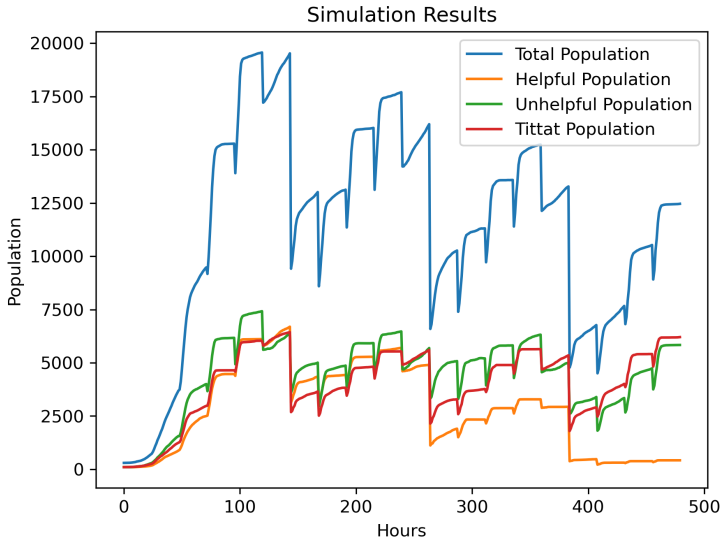


Figure 3: Simulation 3

In third simulation, all the parameters were same but the canteen count was decreased to 50. The helpful population fell drastically wheras the titat and unhelpful were almost equal.

QLEARNING:
LEARNING_RATE=0.1 START_EXPLORATION_PROB=1
MIN_EXPL_PROB=0.4
DISCOUNT=0.9
EXPL_RATE_DECAY=0.1
REPRODUCTION_REWARD=60
VISION_RADIUS=10 VISION_MOVEMENT_PROBABILITY=0.2
ENVIRONMENT:
INITIAL_HELPFUL=100 INITIAL_TITTAT=100 INITIAL_UNHELPFUL=100
CANTEEN_COUNT=50 CANTEEN_FOOD_PER_PERSON=50 CANTEEN_FOOD_DAILY_LIMIT=10000
GHOST_VAL=25
REPRODUCTION_THRESHOLD=60
FOOD_LOST_PER_MOVE=0.041666666666666664

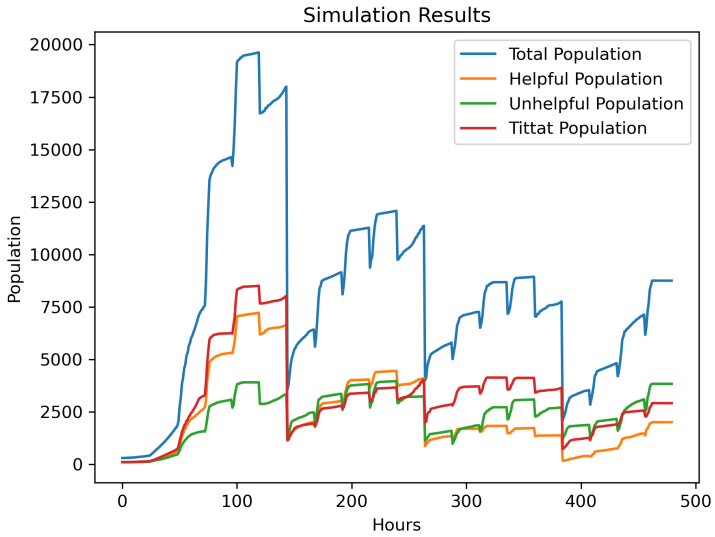


Figure 4: Simulation 4

Reproduction threshold was slightly increased and the distinction between three populations became more clearer. With the unhelpful population forming the major part of the population.

Similar Environment

- Alternative definition of excess food: Instead of defining excess food in terms of the limit to be safe from the Daku, if we define another EXCESS THRESHOLD, which would be greater than the threshold required to be safe from the Daku. In that case, the skew between the populations would decrease and helpful people would also more likely be prosperous. We can think of this situation as the helpful people helping the people rationally/selfishly as compared to the selfless help they used to provided before.

- Canteen share their food multiple times in a day with the Macpen: In this case the population boom is likely to be bigger however we do not expect the distribution/movement to differ since our movement is not governed by the award received at a particular spot.

- Canteen has unlimited food supply: We expect more density around canteens in this case. It isn't due to our Macpen revolving more around the canteen, it is simply the fact that since the canteens can feed an infinite number of people, our randomly moving Macpen reproduce A LOT around the canteens thus the numbers around the canteen are higher.

- When each Macpan reproduces, its children are now of a random type, not necessarily the same type as that of the parent: In this case, we expect a more or less equitable distribution amongst the 3 types since each offspring can now be of any type.