Multi-Agent System for Shared Field of Apples

Gonçalo Nunes
ist199229

Inês Ji
ist199238

Ricardo Antunes
ist199316

ABSTRACT

The *tragedy of the commons* and *public goods games* are common problems of interest in game theory. Given the option to contribute a share of resources to a common cause, most agents will prefer to let others contribute instead, even if that might be detrimental to the group as a whole. This can arise in very simple and real problems. If we take a population exploiting an apple orchard or another such environment, each rational agent will want to maximize their share of apples, but if all agents harvest greedily, the field will deplete and there might not be any incentive for the agents with the least profit to keep maintaining the field. We model this problem in a game-theoretical framework and look at how various strategies and proportion of each strategy in the population affect the game, namely which strategies achieve a higher payoff and how long the game can last with a given population. We also add a mechanism to punish exploitative agents during the game.

1 INTRODUCTION

Public goods games show up naturally in a multitude of situations, in scenarios as diverse as farmers cooperating to maintain shared infrastructures [1] or the vaccination of populations [2, 3]. In such games, players must contribute a fraction of their resources (time, money, a portion of a harvest or something else) for a common objective. In the case presented here, the objective is to maximize the harvest and the game ends when there is nothing more to harvest. Ideally, agents would cooperate to achieve or maximize that common goal, but rational agents will always choose to defect, instead of cooperating, because it is in their interest to maximize their profit, even if failure to cooperate can prove worst for the population as a whole, giving rise to the *tragedy of the commons* [7].

Previous work has been done, from a game theoretical perspective, on the dynamics of systems with biological or self-renewing resources, where games might be played [4, 6]. A model for such a game has even been derived analytically through the replicator equation and analyzed, while looking at how to prevent the *tragedy of the commons* [5]. We propose a framework for a similar problem. We consider a shared field of apples, where players try to harvest as many apples for themselves and the apples can regrow. The game ends when there are no more apples on the field. It would be better for the community, if the field's sustainability was taken into account, to maximize the global harvest over time, but exploitative agents will prefer to keep picking apples and let others worry about the sustainability.

Additionally, a second game is added in association to the problem at hand. Harvesters may start a poll to ban an agent that has been deemed to be harvesting greedily and risking the sustainability of the harvest. The voting strategies considered vary by the behavior type of each agent.

Starting with a baseline with only agents that cooperate with each other, our goal is to analyze the changes in the game's end result, as we slowly introduce new and more adversarial types of agents.

2 METHODS

2.1 Environment

The environment consists of a $N \times M$ grid of cells where agents are spawn at random positions. Each cell may either be empty, or have an apple. The initial number of apples on the grid is random with a chance of 0.1 to spawn in each cell of the grid, this was to introduce variability into the environment. The apples have the ability to regrow and their regrowth rate is influenced by the presence of nearby apples in the neighbouring cells. When an agent eats an apple, the cell containing it becomes empty.

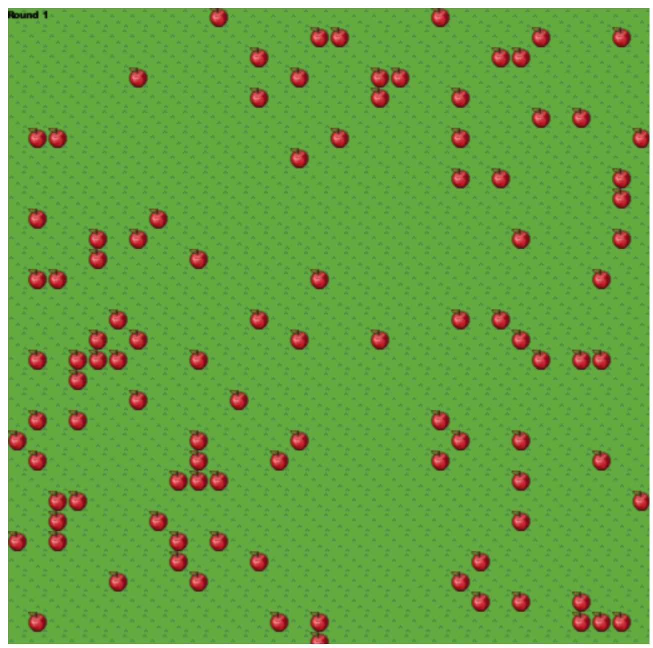


Figure 1: Environment without agents

The game is split into R rounds, each with T turns. At the end of a turn, agents choose between the following possible actions: eating the apple in the current cell (if there's one), staying in the same cell or moving to an adjacent cell.

At the end of each turn, each agent emits an event with type of action (Gather or Move) that they performed to other in-range agents. They also share with those agents the state of all cells it has observed until this turn, with each cell timestamped with its observation turn.

At the end of each round, a voting session occurs, where each agent may accuse another of gathering illegally. The most accused agent is then put on trial. All other agents share information on all illegal gathers they saw the accused agent performing. Then, other agents use this information to decide whether the agent should be banned or not, depending on their behavior. If a majority vote is 'yes', the accused agent will be expelled from the game, and thus unable to gather any more resources.

Formally, we can describe our environment as partially accessible, as agents can only perceive the state of cells and the actions taken in the last turn by other agents within a Manhattan distance V. Furthermore, it is non-deterministic, as there isn't a guarantee that the chosen action can be completed successfully (e.g. to agents may not move to the same cell). The environment is discrete, since it has a finite number of actions and perceptions. It is also static, as any changes happen between turns, and thus the environment doesn't change during action deliberation. In terms of memory, we can consider our system as non-episodic, as any action taken in a turn can influence others in the next turn.

2.2 Multi-Agent System

The agents are capable of autonomous actions in the environment in order to meet their objectives, in this case harvesting apples.

During votes, all non random agents behave the similarly. They propose a vote to ban another agent if they believe that agent to have harvested much more than they have or if they have done so in an unsustainable way. Other non-random agents will agree with the ban, if the accused has harvested more than they have. Random agents also vote at random.

2.2.1 *Agent Strategies.*

- **Cooperative**: Maximizes the growth of apples in the next turn, by eating only the apples which can no longer cause growth.
- **Greedy**: Continuously tries to maximize their individual harvest;
- **Adversarial**: Maximizes their share of the harvest, but behaves like a Cooperative agent in the presence of other agents;
- **Random-Walker**: Performs a random walk on the environment;

(a) Cooperative Agent **(b) Random-Walker** **(c) Greedy Agent** **(d) Adversarial Agent**

Figure 2: Types of agents

2.2.2 *Accusation Strategies.*

Every agent x has a list $A(x)$ of agents they know exist.

Random-Walker. A random agent x picks an agent y uniformly at random from A(x);

$$P_{Random-Walker}(Accuse = y)_x = \begin{cases} \frac{1}{length(A(x))} & \textbf{if } y \in \mathbf{A}(x) \\ 0 & \textbf{if } y \notin \mathbf{A}(x) \end{cases} \tag{1}$$

Adversarial, Cooperative, Greedy. Accuses the one agent seen to have eaten the most resources which were not surrounded (and thus could yet grow further). This is done with a function $SeenGathers(y, i)_x$ that returns the number of Gather actions for non-surrounded resources performed by agent y, seen by an agent x until round i;

$$GreedyAccuse(x, i) = \begin{cases} Nil & \textbf{if } (3) \\ argmax_{y_k}(SeenGathers(y_k, i)_x) & otherwise \end{cases} \tag{2}$$

$$\forall y_k \; SeenGathers(y_k, i)_x = 0 \tag{3}$$

2.2.3 *Voting Strategies.*

Random-Walker. Decides uniformly at random to ban the accused agent or not;

$$P_{Random-Walker}(Ban) = 0.5 \tag{4}$$

$$P_{Random-Walker}(\neg Ban) = 0.5 \tag{5}$$

Adversarial and Greedy. Given an agent x, an adversarial or greedy agent y votes to ban x as long as x isn't y;

$$PonderedBan(x)_y = \begin{cases} Ban & \text{if x} \neq \text{y} \\ \neg Ban & otherwise \end{cases} \tag{6}$$

Cooperative. A cooperative agent will vote to ban any agent that was seen performing a Gather action in a non socially acceptable (8) way. This is determined through the $KnownGathers(x, i)$ function, which returns the set of non socially acceptable gathers an agent x has performed (while in sight of another agent) until round i. This information is shared after agent x is accused, among all other agents.

$$ProtectiveBan(x, i)_y = \begin{cases} Ban & \text{if } \#KnownGathers(x, i) > 0 \\ \neg Ban & otherwise \end{cases} \tag{7}$$

2.2.4 *Action Strategies.*

There are only six types of actions: Up, Down, Left, Right, Stay and Gather. The first five represent the direction in which an agent moves. We consider Z to be a random variable over the discrete uniform distribution of the set of actions Move = {Up, Down, Left, Right}. Apples gathered when surrounded on all four sides by other apples are not counted towards the $SeenGathers(y, i)_x$ computation.

Random-Walker. If a random agent is in the same position as an apple, it will Gather it. Otherwise, it picks a direction from Z.

Greedy. Greedy agents also choose to Gather, whenever they are at an apple. When it comes to Move actions, they seek the apple $target_{i,j}$ closest to them (using the Manhattan distance as the distance metric) and pick the direction that minimizes the distance (ties can be broken arbitrarily)

$$argmin_{m \in Move} ManhattanDistance(move(x, m).position, target_{i,j})$$

If no such apple is known, a greedy agent will move (in the same manner) to the position with the oldest timestamp on its state, i.e., the cell it hasn't visited for the longest number of turns.

Cooperative. Cooperative agents will only Gather a resource when they believe it to be socially acceptable. We define socially acceptable as a resource that is surrounded on all four sides by another resource and none of those adjacent cells has another agent on it.

$$SociallyAcceptable(board_{k,l}) = \tag{8}$$

$$\begin{aligned} &= hasResource(board_{k+1,l}) \wedge \neg hasAgent(board_{k+1,l}) \wedge \\ &hasResource(board_{k-1,l}) \wedge \neg hasAgent(board_{k-1,l}) \wedge \\ &hasResource(board_{k,l+1}) \wedge \neg hasAgent(board_{k,l+1}) \wedge \\ &hasResource(board_{k,l-1}) \wedge \neg hasAgent(board_{k,l-1}) \end{aligned}$$

When this fails, a cooperative agent will do one of two things:

- If it's estimation (based on it's own knowledge of the board) of the resource count (9) is smaller than $\epsilon * length(A(x))$, where ϵ is a parameter of the game, then it will stay put to protect that resource. We considered $\epsilon = 5$;
- Otherwise, if there is an unoccupied socially acceptable resource, it will move there.
- Otherwise, it will move to the position with the oldest timestamp on its state, akin to the greedy agent.

$$estimateResourceCount = \sum_{i=1}^{N} \sum_{j=1}^{M} f(board_{i,j}) \tag{9}$$

$$f(board_{i,j}) = \begin{cases} 1 & \textbf{if hasResource(}board_{i,j}\textbf{)} \\ resourceFrequency & \textbf{otherwise} \end{cases} \tag{10}$$

Where resourceFrequency is the probably each cell holds a resource at the start of game (we considered resourceFrequency = 0.1).

Adversarial. An adversarial agent behaves like a greedy one, but it won't Gather if there are other agents in a sub-grid of side l around it to see it.

2.3 Framework Implementation

The Multi-Agent system described in the previous subsections and the proposed game were implemented in Python, while making use of the standard library and the Pygame module. The code developed has been made publicly available at https://github.com/Goncalo-Palma-Nunes/AASMA.

2.4 System Architecture

The agents (except Random-Walkers) are hybrid, as they share a similar reactive (decisions made based on their perception of the environment that may change at each turn $t \in T$) and proactive nature (actively plan and search for apples' positions to maximize their share of the harvest), as well as the same set of actions. What varies is how aggressively they harvest. Each agent has an internal state with information of what they believe the environment currently looks like and how much themselves have harvested.

2.4.1 Cooperative Agent Architecture. Cooperative agents have a very reactive and proactive nature, eating the closest acceptable apples that emerge. If they don't have knowledge of any such apple, then they start exploring the environment to look for more food, or, if there are few resources, stay on top of an apple, to prevent other agents from extinguishing them.

2.4.2 Adversarial Agent Architecture. The Adversarial Agent Architecture involves agents that prioritize searching for the nearest apple and eating it. They behave in a very similar way to to the Cooperative agents, with the exception that they ignore the long-term consequences of eating too many apples. They still avoid eating too many apples in front of other agents, to avoid being expelled. When they're alone, they eat as much as they can.

2.4.3 Greedy Agent Architecture. Greedy Agents are very simple agents which simply move towards and eat the apple they're closest to, disregarding any social consequences.

2.4.4 Random-Walker Agent Architecture. Random-Walkers have the simplest behaviour. They always pick a direction uniformly at random and take an action to move in that direction. Whenever they are in the same position $map_{i,j}$ as an apple, they harvest it.

3 EMPIRICAL EVALUATION

To empirically evaluate our multi-agent system, several different simulations were run with different population configurations and number of rounds. The number of turns per round was fixed at 100.

3.1 Configurations

We did two types of configuration to evaluate the system. In the first one, we ran 4 simulations of the game each with 20 rounds with 20 agents of the same type.

1st configuration.

- (A) 20 Greedy agents
- (B) 20 Random-Walker
- (C) 20 Cooperative agents
- (D) 20 Adversarial agents

For the second configuration, we ran 4 simulations with mixed population of agents, each with 50 rounds and 20 agents in total.

2nd configuration.

- (E) 10 Cooperative agents and 10 Random agents
- (F) 10 Cooperative agents and 10 Greedy agents
- (G) 10 Cooperative agents and 10 Adversarial agents
- (H) 19 Cooperative agents and 1 Adversarial agents

3.2 1st Configuration

In this configuration, we analyzed the behavior of four homogeneous groups of agents, each exhibiting a distinct behavioral strategy. The objective was to analyse the behavior of the group itself and compare their performance with the other groups over 20 rounds, with each round consisting of 100 turns, to showcase which type of agent has the better strategy.

After running the four simulations we analyse the following metrics:

- Average endowment per round
- Cumulative endowment
- Number of banned agents at the end of the game

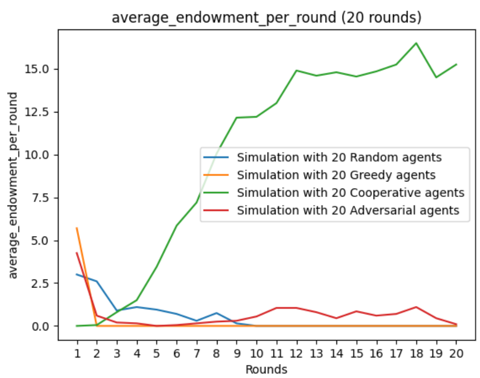


Figure 3: Average endowment per round

The average endowment per round highlights the efficiency and sustainability of each agent type's strategy. By examining the average resources accumulated by agents at the end of each round, we can identify trends and the overall effectiveness of each group's strategy.

We anticipated that the average endowment of cooperative agents would initially be lower than that of other agents. This is because cooperative agents focus on maximizing the growth of apples and consume only those apples that do not jeopardize the sustainability of the system. At the beginning, when there are fewer apples, cooperative agents consume less. They only start eating more once the apple population increases, ensuring long-term sustainability. In contrast, we expected greedy agents to have a high endowment at the beginning since they prioritize their individual harvest by consuming all available apples. However, this short-sighted behavior leads to their extinction after the first round, as shown in Figure 3. As for the adversarial agents, they exhibit behavior similar to greedy agents, but with the additional condition that they won't consume apples if there are other agents nearby. This behavior pattern is reflected in Figure 3, where the line corresponding to the adversarial agents initially shows high consumption in the first round when agents are scattered and there are many apples available. However, as they start exploring and encountering other agents, their consumption decreases rapidly since they refrain from eating when others are nearby. Another factor that leads to their lower average is that they always vote to ban one agent per round. By doing so, they ensure fewer agents compete for resources and to observe them eating, leading to a reduction in resource disputes as the rounds progress. In contrast to the deliberate behaviors of cooperative, greedy, and adversarial agents, random agents rely solely on chance to guide their actions, limiting their ability to make strategic choices to achieve objectives. As a result, their consumption of apples should fluctuates unpredictably. However, in the figure 3, we can see that their curve was initially high and slowly decreased until in the round 9 the apples became extinct, This observation suggests that despite the randomness in their actions, the number of agents on the board is sufficient to cover most apples, as the initial high average consumption by random agents can be attributed to their higher likelihood of consuming apples within the initial board. However, as the rounds progress, this probability gradually decreases, leading to a decline in consumption over time.

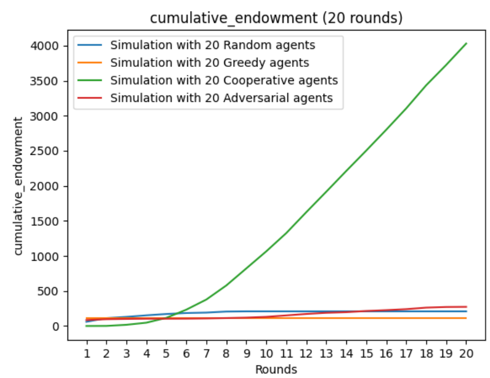


Figure 4: Cumulative endowment

The cumulative endowment metric functions similarly to the average endowment, measuring the total resources accumulated by the agents over all rounds. It provides insight into the overall performance and efficiency of each strategy. In Figure 4, we observe similar trends as in Figure 3. Notably, despite initial advantages in cumulative endowment for the other three strategies, cooperative agents demonstrate superior long-term performance in resource acquisition.

Through Figure 4, it becomes evident that cooperative agents outperform the other strategies significantly in the long run. While initially, the cumulative endowment of the other three strategies may appear higher, the cooperative agent strategy proves to be vastly superior over time. By the end of the game, cooperative agents have accumulated approximately 3700 more resources than the highest cumulative endowment achieved by any of the other three strategies. This observation highlights the remarkable effectiveness of the cooperative agent strategy in resource acquisition. Despite potential its initial setbacks, the collaborative efforts and resource-sharing mechanisms employed by cooperative agents ensure a steady and substantial accumulation of resources over the course of the game unlike the other three agents.

Agent Types	Random	Cooperative	Adversarial	Greedy
Value	4	0	18	0

Table 1: Number of banned agents at the end of the game

The number of banned agents indicates the quantity of agents excluded from further participation in the game due to being banned out. Analyzing the numbers provided in Table 1 offers valuable insights into the ban strategies of each type of agent when operating within homogeneous groups. These insights shed light on how ban strategies influenced the outcomes depicted in Figures 3 and 4.

Just like any behavior from the Random-Walker Agent, its voting process is completely random, resulting in the banning of four agents. It shows that even 16 agents were enough to cover the whole board for the game to end in the round 9. As anticipated, Cooperative Agents, adhering strictly to rules dictating bans only for agents violating resource-surrounding protocols, experienced no bans. Unsurprising, the Greedy Agents, known for their rapid resource accumulation, did not face bans since the game concluded swiftly in the first round, prior to the voting process. Contrastingly, Adversarial Agents, adopting a ban-all stance akin to Greedy Agents, endured the full 20 rounds. As predicted, this resulted in 18 agents being banned, leaving only two remaining. However, due to the requirement for a majority vote for a ban to pass, these two remaining agents were unable to ban each other. Consequently, the slowly dwindled population of agents led to low average in the Figure 3 as the few remaining agents were under less supervision, and thus behaved more greedily.

3.3 2nd Configuration

In this configuration, we analyze the outcome of inserting in a system with cooperative agents another different types of agents. By mixing agent types, we aim to explore how their interactions influence resource acquisition and overall game dynamics. Each simulation in this configuration comprises 50 rounds with a total of 20 agents. The metrics used evaluate the outcomes are the following:

- Average endowment per round
- Number of banned agents at the end of the game
- Total endowment at the end of the game

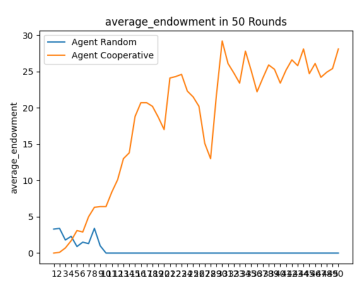
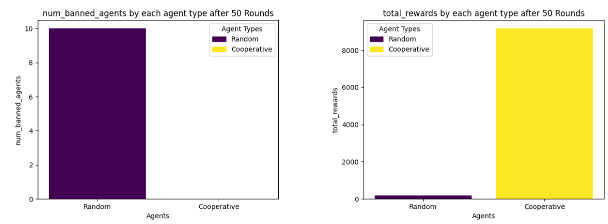


Figure 5: Average endowment for 50 rounds with 10 random agents and 10 cooperative agents

3.3.1 Setting E. By introducing 10 random agents into a system consisting of only 10 cooperative agents, we anticipate that the random behavior exhibited by the random agents will lead them to harvest apples indiscriminately, eventually violating the rule of cooperative agents regarding the consumption of non-surrounded resources. As expected, the behavior of random agents results in fluctuating average endowments, reflecting their unpredictable harvesting patterns. However, by observing Figure 5, we notice a significant decline in the average endowment of random agents after round 8, ultimately reaching 0. This abrupt decrease indicates that all random agents were banned by the cooperative agents, enforcing the rule against consuming non-surrounded resources.



(a) Number of banned agents **(b) Total endowment**

Figure 6: Setting E (10 cooperative agents and 10 random ones)

This outcome is further corroborated by Figure 6a, which illustrates that all 10 random agents were indeed banned by the cooperative agents. With the random agents eliminated from the system, the behavior of the cooperative agents thereafter mirrors that of a homogeneous cooperative population. Consequently, the endowment harvested by the cooperative agents steadily grows as the rounds progress, culminating in a substantially higher total endowment compared to the random agents (as shown in the Figure 6a).

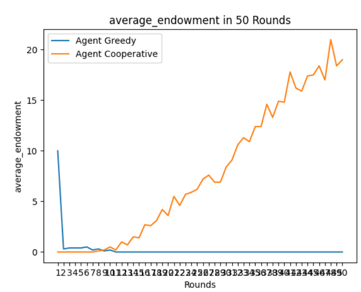
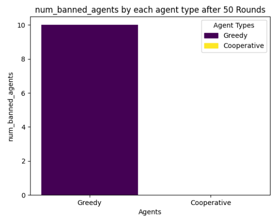
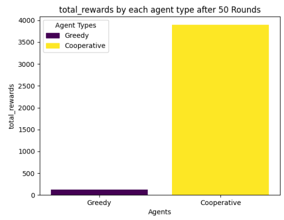


Figure 7: Average endowment for 50 rounds with 10 greedy agents and 10 cooperative agents

3.3.2 Setting F. The outcome of Setting F mirrors that of Setting E. As depicted in Figure 7, greedy agents initially strive to maximize their harvest, resulting in a high average endowment. However, cooperative agents systematically ban greedy agents each round, leading to the elimination of all greedy agents by round 11.



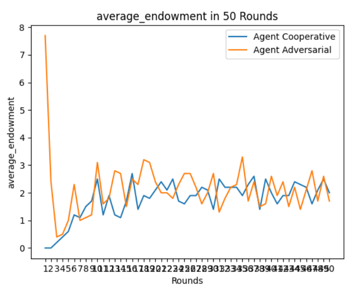
(a) **Number of banned agents**



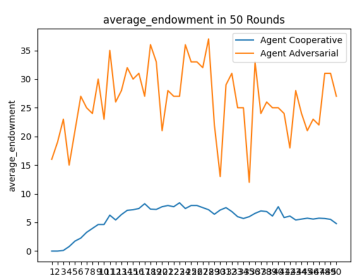
(b) **Total endowment**

Figure 8: Setting F (10 cooperative agents and 10 greedy ones)

In the figure 8b we can observe that the total endowment of greedy agents is lower than that of the random ones (Figure 7b) in a similar setup with cooperative agents due to their nature to greedily consume apples, and thus halting their growth faster. However, their outcome is the same in which all 10 agents are banned (Figure 8a), since they don't respect the rules of sustainability. This parallel between settings E and F highlights the consistent behavior of cooperative agents in enforcing rules and promoting sustainable resource management within the system. Despite initial gains by greedy agents, their non-compliance with sustainability rules leads to their eventual exclusion from the game.



(a) Average endowment for 50 rounds with 10 adversarial agents and 10 cooperative agents



(b) Average endowment for 50 rounds with 1 adversarial agents and 19 cooperative agents

Figure 9: Average endowment for settings (G) and (H) as described in subsection 3.1

3.3.3 *Setting G and H.*

When introducing adversarial agents into a system with cooperative agents, they consistently achieve a larger average endowment than their cooperative counterparts. This is attributed to their ability to evade expulsion while maximizing their harvest by avoiding detection from other agents. As result, we can see that in Figure 9a where there are half adversarial and half cooperative, its initial average endowment is high due to the presence of more scattered apples, allowing adversarial agents to consume without being seen. However, as rounds progress and the apples become confined in one place, their behavior becomes similar to cooperative agents which leads to their similar curves in the plot. As comparison, we can observe that in the Figure 9b, with only 1 adversarial agent introduced, the average endowment of cooperative agents is higher than in Figure 9a. Since there aren't as much adversarial agents, the apple growth ends up larger. This also reflects on the average of the adversarial agent.

4 CONCLUSION

As expected, in a public goods game, Greedy agents were the most destructive with regards to the total harvest, as they only acted to increase their payoff in the short term. For populations with a very low count of such agents and a majority of Cooperative or Adversarial agents, this effect is mitigated, as one of those greedy agents will always be banned at the end of a round.

Random-Walkers, although not as aggressive, can also pose a significant risk to the stability of a harvest, when present in the game, even leading to the extinction of all resources in some population configurations (namely those with a significant amount of Random-Walkers or Greedy agents, in comparison to Cooperative and Adversarial ones).

The presence of Cooperative agents always benefits other agents. Not only do they harvest sustainably, but they also prevent other agents from extinguishing resources, by protecting them when there are few. The best configurations we found (with maximum total reward) contained only Cooperative agents. When mixed with Random and Greedy agents, Cooperative agents initially have a low payoff, but eventually compensate after all non-Cooperative agents are expelled.

Adversarial agents perform better than Random and Greedy agents, as they still invest a bit in the resource's growth, while in sight of other agents. By themselves, they're not able to surpass Cooperative agents, as the resource growth never gets close to the ideal value. They excel when they're a small minority within a population of Cooperative agents.

5 FUTURE WORK

A possible direction of work is to introduce new actions that allow agents to spend a turn (or more) without moving nor eating, but that can increase the rate at which resources are replenished.

The accusation and voting system could also be extended to allow the ban of multiple agents per round. We expect this would greatly mitigate the destructive behavior of Greedy and Random-Walker agents.

Communication between agents could also be made more complex, for example, by allowing agents to lie about, or to filter, what they've observed.

The Cooperative behavior could also be improved further to make them more resilient against Adversarial agents. One possible direction could be to make them try to stick with other agents, and not letting any agent remain alone. With this, Adversarial agents would not be able to harvest resources in a non-sustainable manner as often.

Additionally, from a theoretical perspective, it would be interesting to derive an upper bound on the number of turns the game is expected to last, with a given probability, assuming a population made up entirely of Random-Walkers.

Finally, the work presented here considered fixed payoffs, that varied only according to each agent's strategy and current perception of the environment. It would be interesting to adapt the proposed framework to allow agents to adapt their strategies through reinforcement learning.

REFERENCES

[1] Abbie Turiansky, Collective action in games as in life: Experimental evidence from canal cleaning in Haiti, Journal of Development Economics, Volume 153, 2021, 102722, ISSN 0304-3878

[2] Soltanolkottabi M, Ben-Arieh D, Wu CH. Modeling Behavioral Response to Vaccination Using Public Goods Game. IEEE Trans Comput Soc Syst. 2019 Mar 7;6(2):268-276. doi: 10.1109/TCSS.2019.2896227. PMID: 32391406; PMCID: PMC7176036.

[3] Yong JC, Choy BKC. Noncompliance With Safety Guidelines as a Free-Riding Strategy: An Evolutionary Game-Theoretic Approach to Cooperation During the COVID-19 Pandemic. Front Psychol. 2021 Mar 16;12:646892. doi: 10.3389/fpsyg.2021.646892. PMID: 33796057; PMCID: PMC8008110.

[4] Abakumov, Alexander I. et al. "Game problems of harvesting in a biological community." Automation and Remote Control 77 (2016): 697-707.

[5] Bairagya, Joy Das, et al. "Eco-evolutionary games for harvesting self-renewing common resource: effect of growing harvester population." Journal of Physics: Complexity 4.2 (2023): 025002.

[6] Tilman AR, Plotkin JB, Akçay E. Evolutionary games with environmental feedbacks. Nat Commun. 2020 Feb 14;11(1):915. doi: 10.1038/s41467-020-14531-6. PMID: 32060275; PMCID: PMC7021758.

[7] Hardin G (1968) The tragedy of the commons. Science 162:1243–1248