Can Artificial Agents Replicate Biological Cooperation? An Evolutionary Prisoner’s Dilemma Study

Author: Odiljonov Shohrukhbek, Iqboljon og’li

Abstract

The Prisoner’s Dilemma (PD) is a unique dilemma in the sense that it raises the question of whether to pursue unsustainable short-term gain or maintain long-term cooperation. The original study in the book The Evolution of Cooperation by Robert Axelrod showed that strategies that valued sustained long-term cooperation, like Tit-for-Tat, outperformed their counterparts; however, the strategies were developed by humans and refined over several tournaments. In this study, I am curious whether AI can replicate the same result when given a blank sheet, without any human strategies. To test this, I developed a digital simulation with 100 starting agents scattered across a region. Each agent plays only the agents within a certain radius, and is given a partial memory of their opponent’s previous moves. A game lasts a random number of rounds to limit any form of cheating that may occur. The best agents are then given the option to reproduce at the cost of fitness points, where the offspring inherit their parent’s data with slight mutation.

The results proved that even without human intervention, cooperation can be sustained. Cooperative agents formed small altruistic clusters, which eventually led to defectors going extinct and cooperators thriving. The cooperation rate stagnated around 0.75, which also aligns with Robert Axelrod’s conclusion: “Cooperation can evolve, if players are friendly but firm - willing to cooperate, but ready to retaliate if betrayed.”

Introduction

This study examines whether cooperation can emerge on its own when artificial agents start from scratch, without any built-in human strategies. To test this, I created a digital simulation where 100 agents, each with a simple neural network and some memory, interact with nearby agents in repeated Prisoner’s Dilemma games. Agents that do well can reproduce and pass on their strategies, while those that do not are removed. This approach lets us see whether cooperation can develop naturally through evolution rather than being programmed.

The results show that cooperation can develop and remain steady without human involvement. Agents naturally formed small groups that helped one another, allowing them to outperform defectors and eventually leading defectors to disappear. The cooperation rate settled at about 0.75, supporting Robert Axelrod’s idea that cooperation works best when individuals are friendly but firm.

Hypothesis:

I predict that the rate of cooperation amongst agents will grow steadily initially. However, as the population becomes increasingly cooperative, it becomes vulnerable to defecting agents. As the defecting agent’s population increases, the cooperation rate will collapse, and might even lead to total extinction. If not, the cooperation rate should stabilize and reset the cycle.

Methods:

Simulation Overview: I developed an evolutionary simulation in which agents are randomly spawned across a region and interact in a repeated Prisoner's Dilemma (PD) environment. Each agent is equipped with a small neural network that determines its probability of cooperating in every interaction. It also possesses a rolling memory of each agent's previous actions. An agent can produce an offspring if certain conditions are met and if the agent is willing to pay the cost.

Agent Architecture

- **Neural Network:**
    - Input: 3 neurons (bias, normalized fitness (HP), opponent's last action average)
    - Hidden Layer: 5 neurons, tanh activation
    - Output: 1 neuron, probability of cooperating
- **Memory:** Each agent tracks its opponent's last 10 moves, supporting conditional strategies similar to Tit-for-Tat. (Emergent cooperation from mutual acknowledgment exchange in multi-agent reinforcement learning, 2024)
- **Position:** Agents possess spatial coordinates (x, y) to simulate biological spatial structure.

Agent Interaction

- **Spatial Interactions:** Agents interact only with those within a set radius (interaction_radius = 20) (Buesser et al., 2013)
- **Repeated Games:** Each agent pair plays 40-70 rounds of PD per generation.
- **Payoffs:**
    - Mutual Cooperation: +3 HP each
    - Defection: cooperator -1, defector + 5 (High reward & punishment for being cooperative)
    - Mutual Defection: -4 HP each (Punishment for being too greedy)
- **HP Decay:** Agents lose HP every generation (adaptive_decay) proportional to population density to simulate metabolic costs and resource limitation.
- **Movement:** Each agent moves slightly (+-2 units in x and y) each generation for dynamic clustering.

Reproduction

- **Decision-making:** The neural network outputs a reproduction probability if the agent has sufficient HP.
- **Inheritance:** Clones inherit their parents' neural network, with slight mutations to their biases and partial memory.
- **Placement:** Each new clone is placed near the parent to ensure safe clusters.

Termination and Data Collection

- The simulation stops either after MAX_GENERATIONS or when the agent population goes extinct.
- Metrics recorded per generation:
    - Average HP of alive agents
    - Cooperation rate (rate of agents cooperating)
    - Number of alive agents
    - Cooperation distribution
    - The Neural Network weighs and exports the best agent.

Experiment:

Baseline Cooperation Evolution

- **Objective:** Observe whether AI agents evolve stable cooperative behaviours over generations.
- **Setup:** 100 agents randomly spread across a field with 100 vertices, optional reproduction, and partial memory.
- **Outcomes:** Cooperation rate, average HP, agent population, agent with the most HP by the 50th Generation.

Results

- **Experiment 1: Baseline Cooperation Evolution.** The experiment began with 100 agents randomly distributed across a field. Initially, each agent showed varying levels of cooperation, resulting in a mean of around 0.5 in the early generations (Figure 1). Over successive generations, however, the cooperation rate grew at a near-constant rate, as agents began forming cooperative clusters. It stabilized at ~0.75 by Generation 30

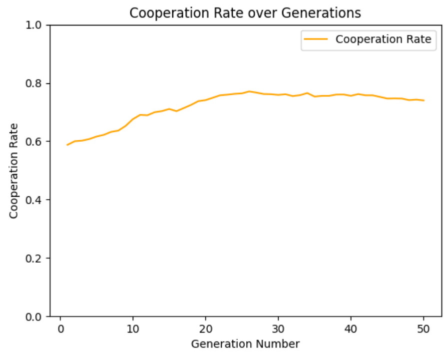

Figure 1

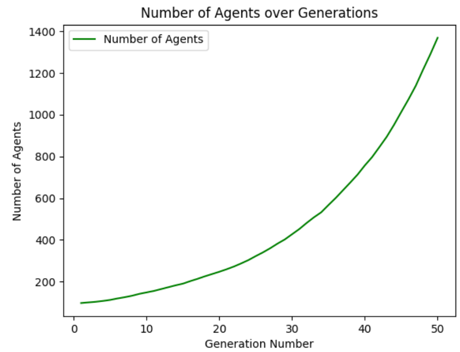

Figure 2

The total number of agents alongside the average fitness (HP) grew exponentially (Figures 2 and 3, respectively). The average HP increased from 1541 in Generation 1 to 2000793 by Generation 50. Similarly, the agent population grew from 97 to 1369 in the same time span. This suggests that cooperative agents not only survived but also thrived within their regional clusters, which created a positive feedback loop between cooperation and overall fitness.

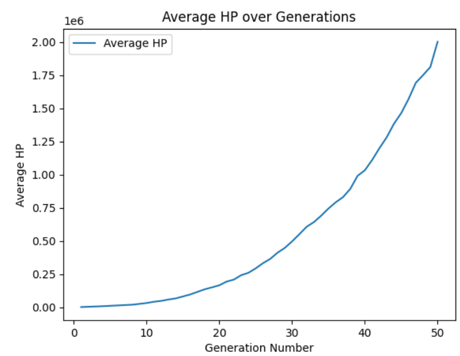

Figure 3

Interestingly, despite the potential advantage of short-term defection, defectors failed to dominate the agent population. The emergence of cooperative strategies was further reinforced by inherited memory alongside spatial clustering, which allowed cooperative agents to identify and preferentially interact with other cooperators. This localized reinforcement is similar to the formation of cooperative niches observed in natural ecosystems (Nowak, May 1992).

The stagnation of the cooperation rate at around 0.75 indicates a quasi-stable evolutionary equilibrium. This dynamic reflects what evolutionary theorists describe as a mixed equilibrium state, where cooperation persists despite evolutionary pressure toward selfishness. This rate of cooperation also mirrors Tit-for-Tat, in which cooperation is encouraged but not to the point of exploitation.

Conclusion

Overall, these results demonstrate that artificial agents with partial memory, spatial interaction, and evolutionary selection can evolve to cooperate, even when defection offers a higher short-term reward.

References

(2024). Emergent cooperation from mutual acknowledgment exchange in multi-agent reinforcement learning. Autonomous Agents and Multi-Agent Systems 38. https://link.springer.com/article/10.1007/s10458-024-09666-5

Buesser, Tomassini, P., Antonioni, M. & Alberto. (2013). Opportunistic Migration in Spatial Evolutionary Games. https://arxiv.org/abs/1309.5229