Course: CS30A1570 Complex Systems

Assignment 8: Models of Cooperation in Social Systems

Completion Date: 27 April 2024

Submitted by: Turzo Ahsan Sami

Title: An Analysis of Strategies in Iterated Prisoner’s Dilemma Games

Introduction

The Prisoner’s Dilemma is a fundamental problem in game theory that demonstrates why two rational individuals might not cooperate, even if it appears that it is in their best interests to do so. This report investigates the effectiveness of various strategies in iterated Prisoner’s Dilemma games under different conditions and payoff matrices. The aim is to understand which strategies perform best and why, and to develop a new strategy that can perform well in these games.

Research Question

The key research question of this investigation is *“What are the most effective strategies in iterated Prisoner’s Dilemma games under different conditions and payoff matrices, and how can a new strategy be developed to perform well in these games?”* The tasks undertaken to complete this investigation are as follows:

- Task 1.1.1: The objective is to analyze the performance of different strategies in a two-person iterated Prisoner’s Dilemma game.
- Task 1.1.2: The aim is to modify the payoff matrix and observe its impact on the game’s outcomes.
- Task 1.1.3: The goal is to run a multi-agent iterated Prisoner’s Dilemma game and identify the strategy that yields the highest payoff.
- Task 1.1.4: The objective is to conduct a tournament among specific strategies and observe the winning strategy under different ratios of cooperators to defectors.
- Task 2.2.1: The aim is to implement a custom strategy in a two-person iterated Prisoner’s Dilemma game and evaluate its performance.
- Task 2.2.2: The goal is to implement a custom strategy in a multi-agent iterated Prisoner’s Dilemma game and assess its performance.
- Task 3: The objective is to develop and describe a unique strategy for a Prisoner’s Dilemma tournament among students.

Result

The following results were obtained by analyzing Prisoner's Dilemma Strategies in NetLogo.

Task 1.1.1: In the Human vs. Computer with random strategy setup of the experiment, ‘Defect’ showed to have the best average score for the human. The human score (920) is the highest here, and the computer score (145) is the lowest. This is because humans are always defecting, therefore benefiting from the computer's cooperation. It also ensures rewarding for the human when the computer also defects. However, this strategy isn't generally considered ideal as it creates distrust and might not be the best long-term approach. The results from the experiment are presented in Table 1.

Table 1: Human (varying strategy) vs Computer (random strategy)		
Entity	**Strategy**	**Average Score**
Computer	Random	680
Human	Random	670
Computer	Random	1168
Human	Cooperate	498
Computer	Random	145
Human	Defect	920
Computer	Random	670
Human	Tit-for-tat	665
Computer	Random	952
Human	Tit-for-two-tats	577
Computer	Random	162
Human	Unforgiving	867

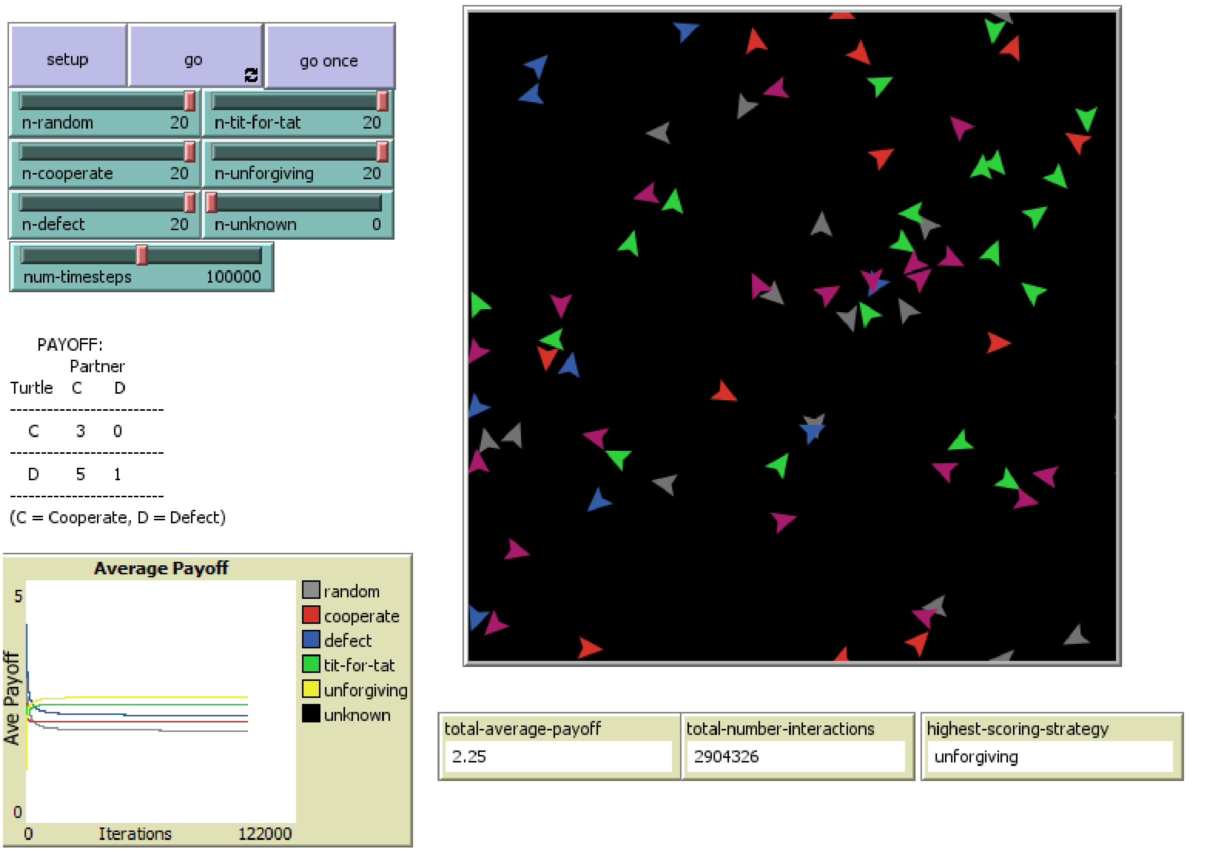
Task 1.1.2: In the modified version of the netlogo model with a new payoff matrix (Appendix: attachment 1) that rewards defecting, the outcome shows that the best strategies for the human are, Defect and Unforgiving. The defect and unforgiving strategies give the highest average scores for the human player against a computer using a random strategy with the modified payoff matrix and the score is much higher for the human following this strategy compared to the original in task 1.1.1. This is because the payoff matrix gives more reward to defecting rather than cooperating. The results of the experiment can be seen in table 2.

Table 2: Human (varying strategy) vs Computer (random strategy) with modified payoff matrix		
Entity	**Strategy**	**Average Score**
Computer	Random	862
Human	Random	904
Computer	Random	1508
Human	Cooperate	584
Computer	Random	300
Human	Defect	1200
Computer	Random	892
Human	Tit-for-tat	892
Computer	Random	1228
Human	Tit-for-two-tats	754
Computer	Random	322
Human	Unforgiving	1174

Task 1.1.3. In the PD-N-Person-Iterated-New.nlogo model where the number of agents with each strategy is set to 20 (except n-unknown to 0), the results show that ‘Unforgiving’ strategy always wins (Image 1). The “Unforgiving” strategy in the Prisoner’s Dilemma model is designed to cooperate initially, but if the partner defects, the “Unforgiving” player will defect indefinitely in all future interactions with that partner. This strategy is highly effective in a scenario where interactions are repeated with the same partners, as it encourages cooperation by punishing defection. Once a player defects against an “Unforgiving” player, the “Unforgiving” player will always defect in future rounds.

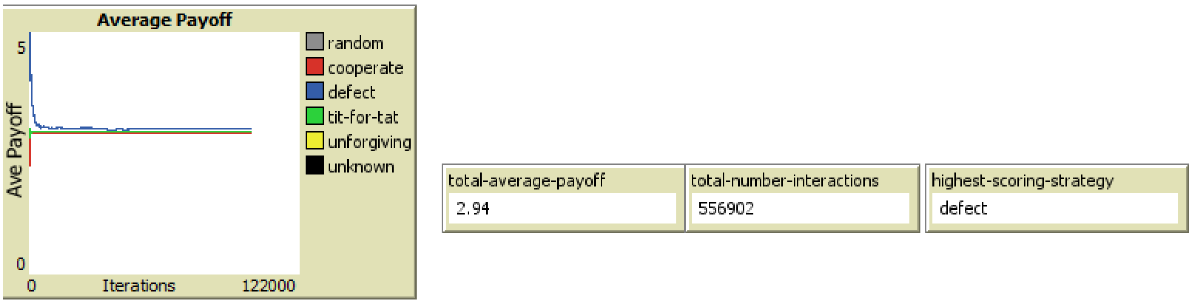
This effectively punishes the defecting player by denying them the higher payoff they would receive from mutual cooperation.

Image 1: PD-N-Person-Iterated netlogo model run (Winner: Unforgiving)



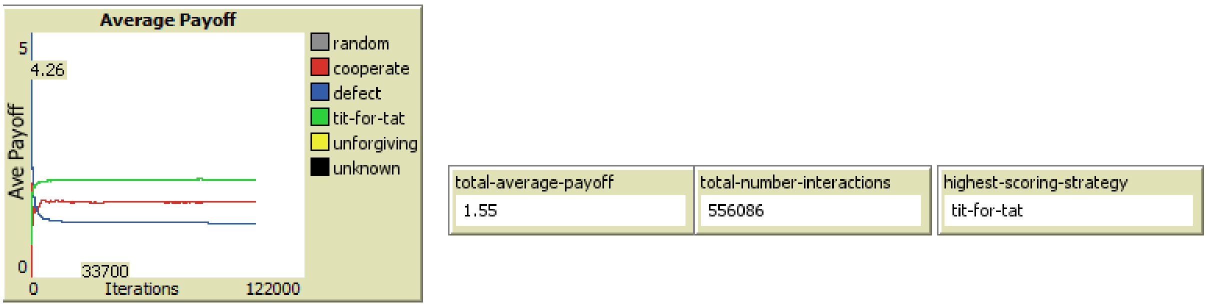
Task 1.1.4. In a tournament amongst Cooperate, Defect, and Tit-for-Tat strategy, the netlogo model run demonstrates that on each run, ‘Defect’ strategy always wins. However, ‘Tit-for-tat’ has very close score as well, with ‘Cooperate’ scoring much lower. The ‘Defect’ strategy comes out on top when there’s at least one defector (no-of-defect = 1) because it exploits the cooperators and tit-for-tat players in the first round. However, ‘Tit-for-tat’ players retaliate in subsequent rounds, leading to a lower overall score for everyone involved. When there are no defectors, ‘Tit-for-tat’ wins because it cooperates with other ‘Tit-for-tat’ and ‘Cooperate’ players, leading to high mutual payoffs. The ‘Cooperate’ strategy scores lower because it gets exploited by ‘Tit-for-tat’ players who are retaliating as a result of interactions with ‘Defect’ players (Image 2).

Image 2: PD-N-Person-Iterated netlogo model run (Winner: Defect)



Varying the ratio of cooperators to defectors changes the dynamics of the game. If there are more defector and very less number of cooperator (no-of-cooperate = 1), the system collapses into a low-trust state where everyone defects. This makes ‘Tit-for-tat’ a potential winner as they come out on top by forming cooperative clusters with other ‘Tit-for-tat’ players (Image 3).

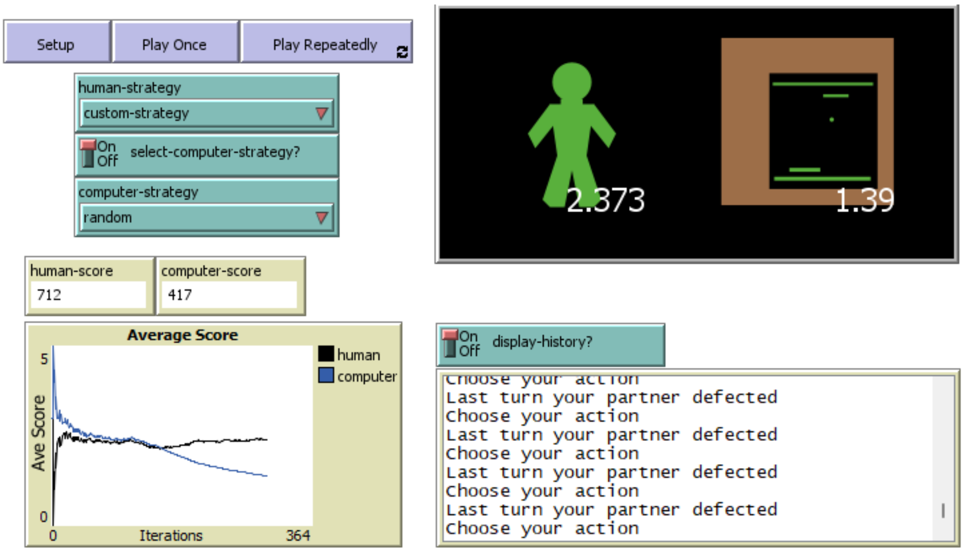
Image 3: PD-N-Person-Iterated netlogo model run (Winner: Tit-for-tat)



Task 2.1: In the Human vs. Computer with random strategy setup of the experiment, the PD-Two-Person-Iterated.nlogo has been modified so that a new strategy was implemented in the custom-strategy procedure (Appendix: attachment 2). Initially, the strategy is set to cooperate in the first two rounds. It calculates the frequency of cooperation of the opponent. If the opponent cooperates more than 50% of the time, the strategy is to assume that the opponent is cooperative and reciprocate accordingly. If the opponent has defected in the last two rounds, it is assumed that they are defecting and the stratyegy defects. Otherwise, the strategy is to play tit for tat. This strategy is shown to be clear winner after multiple simulation runs (Image 4).

The strategy involves cooperating initially, then calculating the frequency of cooperation of the opponent and adjusting behavior accordingly. This strategy wins every time against a random strategy due to its adaptability and responsiveness to the opponent’s actions. The strategy starts by cooperating in the first two rounds establishing trust if the opponent is using a reciprocal strategy. The strategy adapts based on the opponent’s behavior. If the opponent cooperates more than 50% of the time, the strategy assumes the opponent is generally cooperative and reciprocates accordingly. This can lead to a higher payoff when both players cooperate. If the opponent defects in the last two rounds, the strategy assumes they are defecting and responds by defecting as well. This helps to avoid being exploited by an opponent who defects frequently. If the opponent’s cooperation frequency is above 50% and they haven’t defected in the last two rounds, the strategy plays tit-for-tat, which is a highly effective strategy itself. This is why this strategy has much higher winning score over the random strategy.

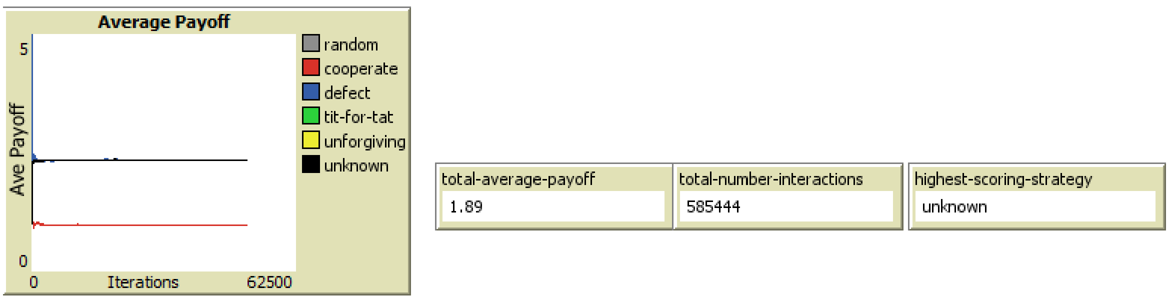
Image 4: modified PD-Two-Person-Iterated.nlogo (Winner: Human with Custom strategy)



Task 2.2: The PD-N-Person-Iterated.nlogo has been modified so that a new strategy was implemented in the unknown procedure (Appendix: attachment 3). This follows the same strategy of custom-strategy implemented in task 2.1. The newly implemented 'Unknown' strategy is shown to be the winner after multiple simulation runs (Image 4). However, 'defect' strategy is almost as close as the 'unknown' strategy.

The "unknown" strategy's success is attributed to its adaptability and unpredictability, which has given it an edge in the tournament. The "unknown" strategy was designed to learn and adapt based on the opponent's actions. Therefore, it effectively counteracts both the "cooperate" and "defect" strategies. Against "defectors", who always betray, the "unknown" strategy learns to always betray as well, ensuring it doesn't get exploited. Against "cooperators", who always cooperate, the "unknown" strategy alternates between cooperation and betrayal to maximize its own payoff.

Image 5: modified PD-N-Person-Iterated.nlogo (Winner: unknown)



Task 3: Three separate strategies have been implemented to the PD_Tournament.py game (Appendix: Attachment 4). The name of the strategies are Judas, Santa, and Sherlock. All these strategies are designed for the Prisoner's Dilemma game. Each strategy takes part in a scenario where cooperation can be beneficial, but betrayal can lead to higher immediate gains. The breakdown of each strategy is as follows:

- Judas: Cooperates in the first round, offering a chance to build trust. If Judas wins (higher score than opponent), it continues to cooperate to maintain the advantage. If Judas loses, it defects to punish the opponent and potentially even the score.

- Santa: Cooperates in the first round, promoting trust. This strategy uses a more complex approach based on its own cooperation history (my_history) and the total number of rounds played (N). It calculates a "magic number" based on the consistency of its cooperation and the overall number of cooperative actions taken. Every 25th round (or a multiple of 25, 25 because of December 25), Santa leans towards cooperation regardless of the situation. Based on the magic number and whether the round is even or odd, Santa might go on "mean streaks" of defection or "cooperative streaks". If none of the above conditions apply, Santa uses Tit-for-Tat, mimicking the opponent's previous action.
- Sherlock: Cooperates in the first two rounds, like Judas and Santa. Analyzes the opponent's cooperation history to assess their overall strategy. If the opponent cooperates more than half the time, Sherlock cooperates as well. If the opponent has defected in the last three rounds, Sherlock defects to punish the behavior. Otherwise, Sherlock uses Tit-for-Tat, mimicking the opponent's previous action.

These strategies offer different approaches to the Prisoner's Dilemma. Judas prioritize their own gain, while Santa and Sherlock try to balance cooperation with self-interest and punishment for bad behavior. Santa's strategy adds more complexity with its probabilistic approach and periodic cooperation. Sherlock focuses on analyzing the opponent's history to adapt its strategy accordingly.

The result of the game is as follows, here Santa is the winner with Judas and TitForTat closely following. In multiple iterations of the game, it was evident that in most of the cases either Santa, Judas or TitForTat wins.

Table 3: Result of the Prisoner’s Dilemma Tournament in python

	Altruist	Egoist	TitForTat	Grudger	Detective	Random	Judas	Santa	Sherlock	Average Score
Altruist	30	-10	30	30	2	2	30	30	30	144.7
Egoist	50	0	5	5	15	15	5	5	10	110.9
TitForTat	30	-1	30	30	28	16	30	30	30	193.3
Grudger	30	-1	30	30	12	14	30	30	30	175.6
Detective	44	-3	28	6	9	24	28	28	28	183.5
Random	44	-3	16	2	6	18	17	24	21	127.8
Judas	30	-1	30	30	28	17	30	30	30	194.2
Santa	30	-1	30	30	28	18	30	30	30	195.1
Sherlock	30	-2	30	30	28	9	30	30	30	185.4

Discussion

The results of the investigation reveal interesting dynamics in the performance of different strategies. In a two-person iterated Prisoner’s Dilemma game, the ‘Defect’ strategy yielded the highest average score for the human player when the computer used a random strategy. This is because the human player was always defecting, thereby benefiting from the computer’s cooperation and ensuring a reward when the computer also defected. However, this strategy is not generally considered ideal as it creates distrust and might not be the best long-term approach.

Modifying the payoff matrix to reward defecting more than cooperating led to the ‘Defect’ and ‘Unforgiving’ strategies yielding the highest average scores. This is because the modified payoff matrix incentivized defecting over cooperating.

In a multi-agent iterated Prisoner’s Dilemma game, the ‘Unforgiving’ strategy consistently yielded the highest payoff. This strategy, which cooperates initially but defects indefinitely if the partner defects, is highly effective in scenarios where interactions are repeated with the same partners, as it encourages cooperation by punishing defection.

In a tournament among the ‘Cooperate’, ‘Defect’, and ‘Tit-for-Tat’ strategies, the ‘Defect’ strategy consistently came out on top. However, varying the ratio of cooperators to defectors changed the dynamics of the game. If there were more defectors and fewer cooperators, the system collapsed into a low-trust state where everyone defected, allowing the ‘Tit-for-Tat’ strategy to come out on top by forming cooperative clusters with other ‘Tit-for-Tat’ players.

Implementing a custom strategy in both a two-person and a multi-agent iterated Prisoner’s Dilemma game showed that adaptability and responsiveness to the opponent’s actions can lead to success. The custom strategy, which cooperated initially, calculated the frequency of cooperation of the opponent, and adjusted behavior, accordingly, consistently won against a random strategy.

Conclusion

The investigation into the effectiveness of different strategies in iterated Prisoner’s Dilemma games revealed that adaptability, responsiveness to the opponent’s actions, and the ability to punish defection are key factors in a strategy’s success. The results also highlighted the impact of the payoff matrix and the ratio of cooperators to defectors on the dynamics of the game. The development of a new strategy that incorporates these elements could potentially perform well in these games. Further research could explore other factors that could influence the effectiveness of strategies in iterated Prisoner’s Dilemma games.

Appendices:

- Attachment 1: https://github.com/ahsan-sami-turzo/complex-system-code/blob/main/CS_Assignment_8/PD-Two-Person-Iterated_v_6.1.2.nlogo
- Attachment 2: https://github.com/ahsan-sami-turzo/complex-system-code/blob/main/CS_Assignment_8/PD-Two-Person-Iterated_v_6.1.3.nlogo
- Attachment 3: https://github.com/ahsan-sami-turzo/complex-system-code/blob/main/CS_Assignment_8/PD-N-Person-Iterated_v_6.1.2.nlogo
- Attachment 5: https://github.com/ahsan-sami-turzo/complex-system-code/blob/main/CS_Assignment_8/PD_Tournament.py