"Prisoner's Dilemma:Reactive Strategies R(y,p,q)"

Shubham Godase MT2308, Gaurav Lute MT2313

August 2024

1 Introduction

The Prisoner's Dilemma is a classic problem in game theory, which has been extensively studied in various fields, including mathematics, economics, and psychology. This problem is often used to illustrate the concept of strategic decision-making and the potential for cooperation or defection in social dilemmas.

In the traditional formulation of the Prisoner's Dilemma, two suspects are arrested and interrogated separately. Each suspect has to decide whether to defect (D) or remain coperate (C). If both suspects confess, they each receive a moderate sentence. If one confesses and the other remains silent, the confessor receives a light sentence, while the silent one receives a heavy sentence. If both remain silent, they each receive a minimal sentence.

The Prisoner's Dilemma is a well-known scenario in game theory that explores how individuals might end up with a worse outcome by acting in their own self-interest rather than cooperating. In this situation, two suspects are arrested for a crime they committed together and are interrogated separately. Each suspect must decide whether to cooperate with the other by staying silent or to defect by confessing. If both stay silent, they receive a light sentence. If one defects while the other stays silent, the defector goes free and the silent one gets a heavy sentence. If both defect, they both receive a moderate sentence. The dilemma arises because defecting is the dominant strategy for both suspects, leading them to defect each other and end up with a worse outcome than if they had both cooperated. This paradox highlights the challenge of achieving mutual benefit when individual incentives drive decisions.

2 Representative Strategies

Various strategies have been developed for the Presoner's Delima, ranging from simple fixed strategies to complex ones that adapt based on the opponent's previous actions:

Payoff Matrix

In a two person zero or constant sum game, the resulting gain can be represented in the form of a matrix which is called Pay-off Matrix or Gain matrix. In other words payoff matrix is a table in which strategies of one player are listed in rows and those of the other player in columns and the cells show payoffs to each player such that the payoff of the row player is listed first.

	Cooperate	Defect
Cooperate	Reward(R)	Sucker's Payoff(S)
Defect	Temptation (T)	Punishment (P)

Reward (R) :

The benefit received when both individuals cooperate.

Sucker's Payoff (S) :

The cost incurred when one cooperates while the other defects.

Temptation (T) :

The greater benefit obtained when one defect while the other cooperates.

Punishment (P) : The shared cost when both individuals defects.

2.1 Tit-for-Tat (TFT):

In our exploration of programmed strategies for the Prisoner's Dilemma, we introduce the 'Tit-for-Tat' strategy, a classic and widely studied approach in game theory. Tit-for-Tat is known for its simplicity and its adaptive nature,which makes it a compelling subject of analysis.

The Tit-for-Tat strategy follows a straight forward principle: it starts by cooperating ('C') in the first round. In subsequent rounds, it mirrors its opponent's previous move.If the opponent cooperated in the last round, 'Tit-for-Tat' responds with cooperation in the current round. Conversely,if the opponent defect ('D') in the previous round, 'Tit-for-Tat' retaliates with defect. This strategy is reactive and forgiving, as it responds to the opponent's actions while giving them the opportunity to revert to cooperation.

2.2 Always Cooperate (All-C):

Always cooperate, no matter what the other player does. A Cooperate unless the other player defects, then punish them to some degree. Try to figure out what someone's strategy is, then play what's best against that.

In some cases, the Always Cooperate strategy can lead to better outcomes in the long run, especially in repeated games or when there is a high level of trust between players. For instance, if both players always cooperate, they can achieve a higher payoff than if one or both players defect.

While the Always Cooperate strategy can be a valuable approach in certain situations, it's important to consider the potential risks and limitations. In many real-world scenarios, more nuanced strategies, like Tit-for-Tat or conditional cooperation, may be more effective in achieving long-term cooperation and avoiding exploitation.

2.3 Always Defect (All-D):

Always Defect strategy is a simple and deterministic approach where a player consistently chooses to defect in every round of the game, regardless of the opponent's actions. This strategy leads to a sub optimal outcome if both players defect, resulting in harsher penalties compared to mutual cooperation. However, it can be rational in single-play scenarios or when there's significant uncertainty about the opponent's intentions.

In repeated games or when interacting with multiple players, "Always Defect" may offer short-term gains and can be effective in avoiding exploitation. Yet, it often fasters instability and conflict within groups. The choice to use this strategy depends on the game's structure, the potential benefits of co-operation versus defection, and the frequency of interactions.

2.4 Random Strategy:

In our exploration of strategic approaches for the Prisoner's Dilemma, we introduce the 'Random Select' strategy, which brings a refreshing twist to decision-making. The 'Random Select' strategy is distinct in its reliance on chance: it randomly chooses between cooperation ('C') and defect ('D') in each round of the game. This method stands apart from more deterministic strategies, as it does not consider the opponent's previous actions or follow any specific decision-making rules. Instead, it leverages randomness to guide its choices.

The 'Random Select' strategy introduces an element of unpredictability into the Prisoner's Dilemma. Unlike deterministic strategies that follow predefined patterns, 'Random Select' relies on chance.

This randomness can lead to a variety of outcomes, making it difficult for opponents to anticipate the player's choices.

2.5 Grim Trigger:

In this strategy, a player initially cooperates with the other player and continues to cooperate as long as the other player does too. However, if the other player defects even once, the Grim Trigger strategy dictates that the player will defect in all future rounds as punishment, essentially never returning to cooperation. It's a form of retaliation that ensures cooperation only if both players remain mutually cooperative throughout the game.

2.6 Pavlov Strategy:

In this strategy, a player will repeat their previous move (cooperate or defect) if it resulted in a good outcome (win). However, if the outcome was bad (loss), the player switches their move in the next round. For example, if both players cooperated and received a good payoff, they will continue to cooperate. But if one defects and gets a poor result, they will change their strategy in the next round. This strategy encourages cooperation but also adapts based on outcomes.

2.7 Discriminating Altruist:

The Discriminating Altruist strategy in the Iterated Prisoner's Dilemma (IPD) refers to a player who cooperates with others as long as they have never defected against them. Once another player defects, the Discriminating Altruist no longer cooperates with that player, refusing to engage in future cooperation. This strategy rewards trust and cooperation but punishes any instance of betrayal by permanently withholding cooperation. It fosters a conditional form of altruism based on the behavior of others.

2.8 ZD Strategy:

In this strategy, a player's long-term average payoff is related to their opponent's average payoff through a fixed linear equation. Essentially, the ZD strategy ensures that the player's payoff can be expressed as a linear function of the opponent's payoff over the long run. This approach aims to maintain a predictable and stable relationship between the payoffs of the two players throughout the game.

3 Reactive Strategy R(y,p,q)

Cooperates with probability y in first round and with probabilities p or q after opponent cooperates or defects.

The Reactive Strategy is characterized by probabilistic decisions based on the opponent's previous actions. It is parameterized by y, p, and q.

1. y: Probability of cooperating in the first round.
2. p: Probability of cooperating after the opponent cooperated in the previous round.
3. q: Probability of cooperating after the opponent defected in the previous round.

3.1 Why choose this stratergy?

I chose this strategy due to its flexibility in adapting to the opponent's behavior, allowing for a mix of cooperation and defection based on the opponent's actions. The strategy balances between Tit-for-Tat's retaliatory nature and the potential to exploit consistent cooperators.including the nature of the game, the characteristics of the opponent, and the long-term goals of the player. By carefully considering these factors, players can choose a reactive strategy that is likely to lead to favorable outcomes.

Adaptability: Reactive strategies can adapt quickly to changing game circumstances, making them suitable for dynamic environments.

Unpredictability: The random element in reactive strategies makes them difficult to anticipate, which can be an advantage in competitive games.

Fairness: Reactive strategies can ensure fairness and unbiased outcomes, which is essential in certain game scenarios.

Simplicity: Reactive strategies are often easy to implement and understand, making them a convenient choice for game developers and researchers.

Flexibility: Reactive strategies can be applied to various game contexts and opponents, making them a versatile approach.

Mitigating Uncertainty: Reactive strategies can help mitigate the impact of unknown factors in games with high levels of uncertainty or randomness.

Competitive Edge: The unpredictability and adaptability of reactive strategies can provide a competitive edge in games where opponents are using more deterministic approaches.

4 Simulation Setup and Results:

The simulation results are calculated by running the simulation 1000 times and averaging the payoffs for each strategy.Here's a breakdown of how the payoffs are calculated for each strategy.

4.1 Reactive Strategy

The simulations were conducted over 100 rounds of the iterated Prisoners Delima, comparing the Reactive Strategy (with varying y, p, q values) against other standard strategies like Tit-for-Tat, All-C, and All-D.

4.1.1 Cumulative Payoff Analysis

The cumulative payoffs for each strategy pair were plotted over time. The Reactive Strategy with parameters $y = 0.7, y = 0.7, p = 0.9, p = 0.9, q = 0.2, q = 0.2$ performed comparably to Tit-for-Tat when facing a Tit-for-Tat opponent, showing a steady accumulation of payoffs. Against All-D, the Reactive Strategy adjusted by defecting more frequently, preventing exploitation.(See figure 1)

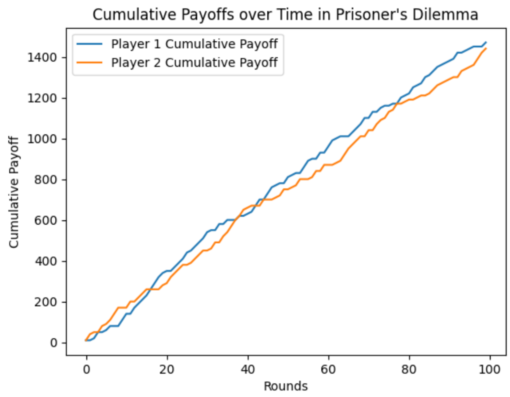

Figure 1: Cummulative Payoffs over time PD

The reactive strategy's cumulative payoff for Player 1 is expected to surpass Player 2's random strategy over 1000 rounds. By adapting to the opponent's previous action, the reactive strategy exploits mistakes and increases its payoff over time. In contrast, the random strategy's lack of adaptation leads to a flat or decreasing payoff, as it fails to account for the opponent's behavior and makes more mistakes.

4.1.2 Payoff Distribution

Barplot of payoffs over 1000 rounds were created for each strategy comparison. The Reactive Strategy exhibited a bimodal distribution, reflecting its adaptability cooperating frequently against cooperators and defecting against defectors.(See figure 2)

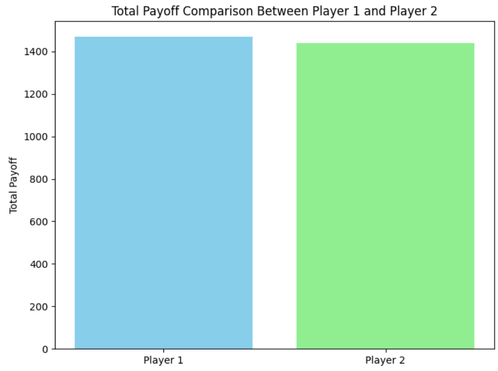

Figure 2: Cummulative Total payoff

4.1.3 "Reactive strategy vs Tit-for-Tat"

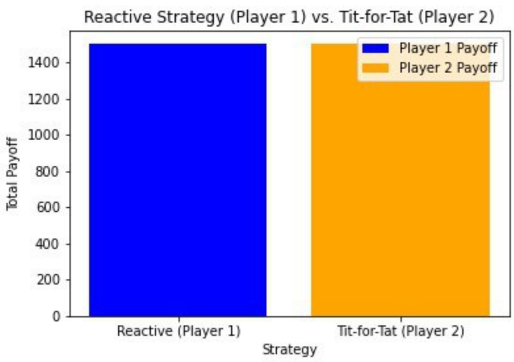

Figure 3: Reactive vs Tit-for-tat strategy

- Reactive Strategy vs. Tit for Tat Strategy:

  Reactive Strategy wins: 245 times

  Tit for Tat Strategy wins: 242 times

  Ties: 513times

4.1.4 " Reactive strategy vs All Cs"

- Reactive Strategy vs. All Cs Strategy:

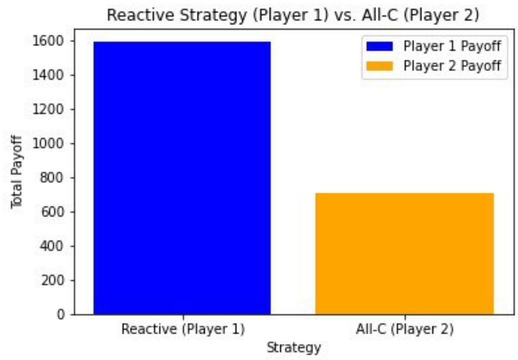

Figure 4: Reactive vs All Cs strategy

  Reactive Strategy wins: 1000 times

  All Cs Strategy wins: 0 times

  Ties: 0 times

4.1.5 "Reactive strategy va All Ds "

Figure 5: Reactive vs All Ds

- Reactive Strategy vs. All Ds Strategy:

  Reactive Strategy wins: 0 times

  All Ds Strategy wins: 1000 times

  Ties: 0 times

Figure 6: Reactive vs Random

4.1.6 "Reactive strategy vs Random"

- Reactive Strategy vs. Random Strategy:

  Player 1 wins: 447 times

  Player 2 wins: 500 times

  Ties: 53 times

4.1.7 "Reactive vs Grim Trigger Strategy"

Figure 7: Reactive vs Grim Trigger

- Reactive Strategy vs. Grim Trigger:

  Reactive Strategy wins: 503 times

  Grim Trigger wins: 497 times

  Ties: 0 times

4.1.8 "Reactive vs Pavlov Strategy"

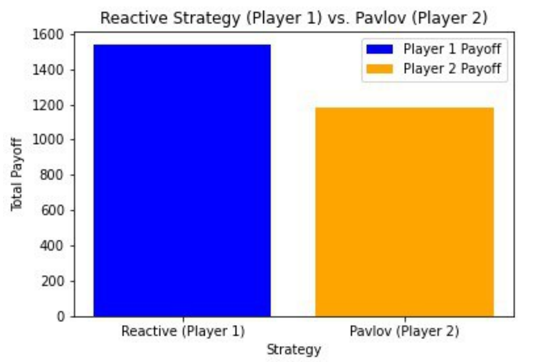

Figure 8: Reactive vs Pavlov strategy

- Reactive Strategy vs. Pavlov:

  Reactive Strategy wins: 993 times

  Pavlov wins: 3 times

  Ties: 4 times

4.1.9 "Reactive vs Discriminating Altruist"

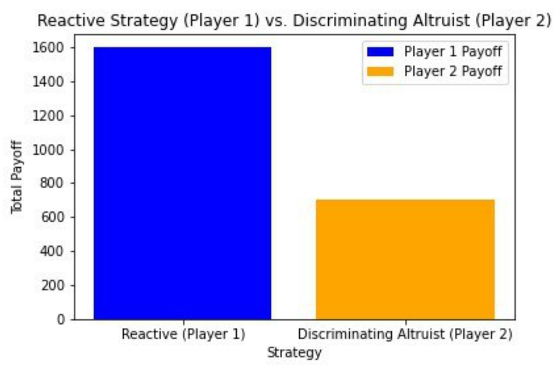

Figure 9: Reactive vs Discriminating Altruist

- Reactive Strategy vs. Discriminating Altruist:

  Reactive Strategy wins: 1000 times

  Discriminating Altruist wins: 0 times

  Ties: 0 times

4.1.10 "Reactive vs ZD Strategy"

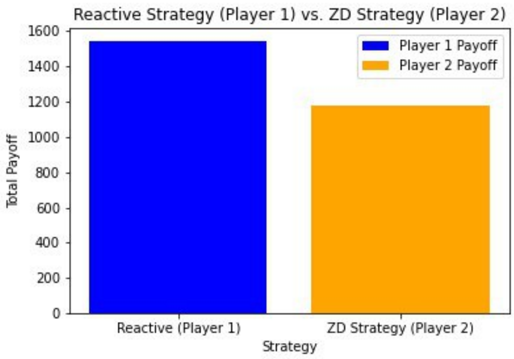

Figure 10: Reactive vs ZD strategy

- Reactive Strategy vs. ZD Strategy:

  Reactive Strategy wins: 989 times

  ZD Strategy wins: 6 times

  Ties:5 times

5 Conclusions

We conclude that simulates and compares the performance of five different strategies in the Prisoner's Dilemma game: Reactive, Random, All Cs, All Ds, and Tit for Tat etc. The Reactive strategy is a probabilistic strategy that adapts to the opponent's previous action, while the other strategies are more straightforward.

The code first defines the payoff matrix for the Prisoner's Dilemma game, which specifies the rewards for each possible combination of actions (Cooperate or Defect) taken by the two players. It then implements the "Reactive strategy", which is a probabilistic strategy that takes into account the opponent's previous action. The code runs multiple iterations of the game for each strategy, calculates the total payoffs for each player, and plots the results. The plots show the cumulative payoffs over time for the Reactive strategy and the total payoffs for each player under each strategy. The "Reactive Strategy" effectively balances between cooperation and defection based on the opponent's behavior, making it versatile in different Prisoners Delimma scenarios. It outperforms simple strategies like All-C or All-D by avoiding exploitation and maintaining a competitive payoff.

6 Future Scope

1. Artificial Intelligence and Machine Learning:

   Researchers are exploring AI strategies inspired by the Prisoner's Dilemma to improve decision-making in autonomous systems. It has potential applications in robotics,autonomous vehicles, and algorithmic trading.

2. Blockchain and Cryptoeconomics:

   Cryptoeconomic systems often involve multiple participants with conflicting interests. Game theory, including the Prisoner's Dilemma,is used to design incentive mechanisms and ensure network security.

3. Social Media and Online Platforms:

   Understanding user interactions, content moderation, and misinformation spread can benefit from game theory insights. The Prisoner's Dilemma may inform platform policies and strategies.

4. Healthcare and Epidemiology:

   The dilemma can be applied to understand vaccination decisions, antibiotic use,and cooperation in healthcare settings. It has relevance in modeling the spread of infectious diseases and designing interventions.

7 References

1. Under the guidance of Dr.Bhalachandra Pujari SCMS

2. `https://github.com/htpusa/reactive-PD`

3. The Continuous Prisoner's Dilemma: I. Linear Reactive Strategies Author links open overlay panel LINDI M WAHL, MARTIN A NOWAK *

4. `https://github.com/shubhamgodase2002/prisoners_dilemma_reactive_startergy-y-p-q-`

5. `https://github.com/gauravlute01/prisoner-s-dilemma-reactive-strategy`