Evaluating the Effectiveness of the GRIM Strategy in Various Prisoner's Dilemma Scenarios

Kunal Gokhe
MT2309
Department of SCMS, SPPU
gokhe.kunal@scms.unipune.ac.in, github

Famit Dongarwar
MT2305
Department of SCMS, SPPU
dongarwar.famit@scms.unipune.ac.in, github

Abstract

The Prisoner's Dilemma is a fundamental model for studying cooperation and defection in game theory. This study evaluates the performance of the GRIM strategy in comparison to four other strategies: random, cooperate with probability, reactive, and zero-determinant (ZD). Through extensive simulation and analysis using histograms and cumulative distribution functions (CDFs), we demonstrate the long-term outcomes of each strategy. The results highlight that the GRIM strategy, while effective in maintaining cooperation when mutual trust is established, shows significant vulnerability to strategies that incorporate randomness or adaptive behaviours. Although GRIM can sustain cooperation, it is highly punitive, leading to mutual defection when trust is broken. This analysis underscores the importance of understanding the strengths and limitations of various strategies in repeated interactions, particularly in environments where cooperation is critical but cannot always be guaranteed.

Keywords: Prisoners Dilemma, strategy, statistics, grim trigger strategy. Etc.

1. Introduction

A classic game theory dilemma that highlights the difficulties of cooperative action in competitive settings is the Prisoner's Dilemma. Mathematicians Merrill Flood and Melvin Dresher first proposed it in the early 1950s, and Albert W [1]. Tucker later provided the traditional interpretation. The prisoner's dilemma is a game theory thought experiment that involves two rational agents, each of whom can cooperate for mutual benefit or betray their partner "defect" for individual reward. Two members of a criminal gang are arrested and imprisoned. Each prisoner is in different sell where no contact to each other, police don't have enough evidences to arrest both so they offer both one deal, if he testifies against his partner, he will go free while the partner will get three years in prison on the main charge. If both prisoners testify against each other, both will be sentenced to two years in jail.

Repetitive interactions have a significant impact on the dynamics of the circumstance. Both parties would have every incentive to defect if they knew there would only be one option available, as doing so would result in a larger payout than cooperating, regardless of the other player's decision. However, performers frequently maintain connections with both a significant past and a promising future. In a relationship like this, making wise decisions necessitates understanding the structural effects of strategic interaction. The payoff matrix for the prisoner’s dilemma is shown in figure 1 [3].

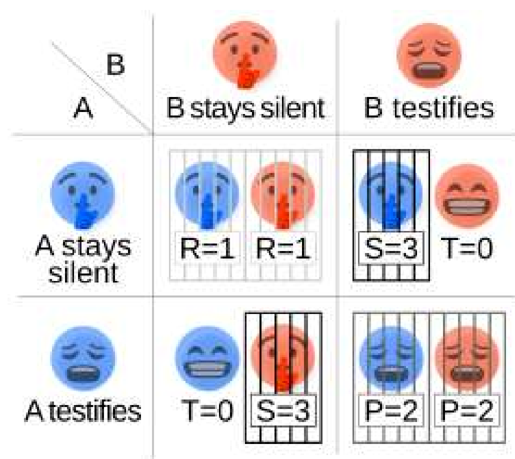

Figure 1. An example prisoner's dilemma payoff matrix [3]

This leads to four different possible outcomes for prisoners A and B:

- If A and B both remain silent, they will each serve one year in prison.
- If A testifies against B but B remains silent, A will be set free while B serves three years in prison.
- If A remains silent but B testifies against A, A will serve three years in prison and B will be set free.

- If A and B testify against each other, they will each serve two years.

The Dilemma: Without knowing what the other will do, each prisoner must decide whether to cooperate (keep silent) or defect (confess). Since this decision lessens the worst-case situation, rational self-interest advises that both should defect. If both of them defect, the consequence is not worse (1 year for each) than if they had both cooperated (2 years for each).

Implications: The Prisoner's Dilemma draws attention to the tension that exists between individual reason and the good of the group. It illustrates difficulties in circumstances like arms races, corporate competitiveness, and the provision of public goods, where individual motivations can result in poor outcomes for all parties involved. It has broad implications in economics, politics, and social sciences.

2. Prisoner's Dilemma Strategies

Two prisoners are segregated into separate rooms and are unable to speak with one another. Both prisoners are presumed to be aware of the rules of the game, to be unfaithful to one another, and to have no chance of reward or retaliation outside of the game. Each prisoner receives a larger reward for betraying the other (also known as "defecting"), regardless of what the other decides. Analysing the best answers from each player is part of the reasoning: B will either comply or disobey. A should defect if B cooperates, as being free is preferable to serving a year [2]. Because serving two years is preferable to serving three, if B defects, A should likewise defect. Therefore, regardless of B's approach, A should defect since that is the best course of action. Parallel reasoning will show that B should defect. For both players, defection is a strictly dominant strategy because it always yields a higher reward than cooperation. The only strong Nash equilibrium in the game is mutual defection. This Nash equilibrium is not Pareto efficient because the mutually ideal outcome of cooperation is irrational from a self-interested perspective.

In this section we have discussed few dilemma strategies and their known results based on the previously conducted research and simulation experiments.

Always Defect

Defect in every round, regardless of the opponent's actions

Results:

- Strengths: Easy to use and immune to cooperation-based tactics.
- Weaknesses: Usually results in mutual defection, which produces less than ideal outcomes for both players. Poor outcomes arise when faced with cooperative strategies.
- Outcome: Although it is strong in a one-shot game, it performs poorly in repeated encounters where cooperation could lead to longer-term benefits.

Always Cooperate

Cooperate in every round, no matter what the opponent does.

Results:

- Strengths: Promotes cooperation when the opposition cooperates as well.
- Weaknesses: Easily exploited by defectors; may lead to very unsatisfactory results when facing uncooperative opponents.
- Outcome: Generally, underperforms unless in an All-C player population or a highly cooperative environment.

Tit for Tat

Start by cooperating, and then mimic the opponent's previous move. If the opponent cooperates, you cooperate; if they defect, you defect in the next round.

Results:

- Strengths: Clear-cut, understanding, and forgiving. If both players begin to cooperate, it encourages cooperation between them.
- Weaknesses: Open to "noise" in the game, which can result in a vicious cycle of reprisals.
- Outcome: In IPD tournaments and simulations, Tit for Tat has proven to be very successful at promoting cooperation and resisting exploitation.

Random

Choose to cooperate or defect randomly in each round, with a fixed probability for each action.

Results:

- Strengths: Because its unpredictable, opponents find it challenging to predict and take advantage of.
- Weaknesses: Results are wildly inconsistent and lack a stable pattern of cooperation or defection.
- Outcome: Usually underperforms in terms of total rewards because it doesn't have a well-thought-out plan for encouraging cooperation or countering the opponent's moves.

Suspicious Tit for Tat

Similar to Tit for Tat, but starts by defecting in the first round. After the initial defect, it mimics the opponent's previous move

Results:

- Strengths: Protects against first-round exploitation and has the ability to penalize opponents who attempt to take advantage of a cooperative start.
- Weaknesses: May take longer to build mutual cooperation; less forgiving at first.
- Outcome: Good in situations where there is little initial trust, but compared to standard Tit for Tat, it might cause cooperation to return more slowly.

Cooperate with Probability

Cooperates with a fixed probability P P in each round and defects with probability $1 - P$. The probability P is chosen based on the desired level of cooperation

Results:

- Strengths: Offers flexibility in striking a balance between cooperation and defection; by modifying the probability, it can adjust to various environments.
- Weaknesses: Less predictable; if P is not tuned in, it may not consistently take advantage of opportunities for mutual cooperation.
- Outcome: Dependent on opponent behaviour and probability P selection, this strategy may work well in environments with different degrees of cooperation.

Reactive

Reacts to the opponent's last move but with a probabilistic approach. For instance, it might cooperate with a certain probability if the opponent cooperated and defect with a different probability if the opponent defected [4].

Results:

- Strengths: Able to respond to various opponent strategies and modify behaviour based on past performance.
- Weaknesses: May not always maintain cooperation and be less predictable, particularly if the probabilities are not calibrated properly.
- Outcome: Dependent on the particular probabilities used, performance can vary, but useful in settings where responses must be customized based on past interactions.

Zero Determinant

ZD strategies are a class of strategies in the iterated Prisoner's Dilemma that use a deterministic approach to achieve specific outcomes.

Results:

- Strengths: Enforces specific payoff relationships by allowing one player to set the terms of cooperation or defection.
- Weaknesses: Assumes rational, non-adaptive opponents and necessitates exact calibration of the strategy parameters.
- Outcome: Good at enforcing certain payoff structures, but because of its complexity and underlying assumptions, it must be applied carefully and is more frequently employed in theoretical analyses than in real-world situations.

3. Grim Trigger Strategy

In the Prisoner's Dilemma, the Grim Trigger strategy can be applied as follows:

Initially, both prisoners cooperate (stay silent). If one prisoner defects, the other prisoner will switch to a "grim" strategy and defect in all future interactions. Even if the defecting prisoner tries to cooperate again, the other prisoner will continue to defect, as the trust has been broken. The Grim Trigger strategy can help to maintain cooperation in the Prisoner's Dilemma by providing a deterrent against defection. If both prisoners know that defecting will lead to a permanent breakdown in cooperation, they may be more likely to cooperate in the first place. Working steps of strategy is as follows:

- **Initial Cooperation**: The GRIM strategy starts by cooperating with the other player.
- **Punishment Trigger**: If the opponent ever defects, the GRIM strategy retaliates by defecting in every subsequent round.
- **No Forgiveness**: Once the GRIM strategy starts defecting due to the opponent's defection, it will never return to cooperation, no matter what the opponent does afterward.

Rationale Behind the GRIM Strategy

- The GRIM strategy is simple to understand and implement, making it a powerful approach in many scenarios.
- In theory, if both players are rational and follow the GRIM strategy, they will both continue to cooperate indefinitely, leading to the best possible outcome for both (mutual cooperation).
- The threat of eternal punishment is meant to deter the opponent from defecting in the first place.

Advantages of the GRIM Strategy

The GRIM approach has many benefits, the head among them being its strong deterrent and ease of use. It ensures stability in recurring interactions by providing a strong incentive for the other party to continue cooperating—threatening eternal defection in response to a single act of betrayal. Because of the simplicity and clarity of this threat, both players can easily comprehend and anticipate the strategy, which may eventually lead to mutual cooperation.

Disadvantages of the GRIM Strategy

Even though it works well to prevent betrayal, the GRIM strategy has a number of serious drawbacks. Its main flaw is that it's rigid; if a player leaves the game, even unintentionally, the GRIM approach requires that they stay away forever. This can result in unfair consequences and the dissolution of potentially helpful relationships. This inability to forgive can lead to mutual destruction, wherein both parties ultimately suffer losses.

4. Results

In the context of the Prisoner's Dilemma, we have shown the comparisons between the "grim" strategy and other strategies (random, cooperate with probability, ZD, and reactive). Plots of the cumulative distribution function (PDF), histograms, and violin plots are used to display these comparisons. As demonstrated by the plots, the experiments' replication validates the Law of Large Numbers. Histograms and CDF plots show smooth curves and stable, consistent distributions as more experiments are carried out and the average results for each strategy approach their expected values. This convergence highlights the underlying performance characteristics of each strategy in the Prisoner's Dilemma.

Grim vs. Cooperate with probability ($p = 0.25$)

The distribution of average scores for the "grim" and "prob" (probability-based) strategies is displayed in the histogram. The "prob" strategy exhibits a slightly higher peak than the "grim" strategy, with a wider distribution. The "grim" strategy peaks at approximately 1.4. The "grim" CDF is mostly to the right of the "prob" CDF in the CDF plot, indicating that the "grim" strategy consistently scores marginally lower than the "prob" strategy. This implies that the "prob" approach might more frequently produce higher average scores. The results for grim versus cooperate with probability is shown below.

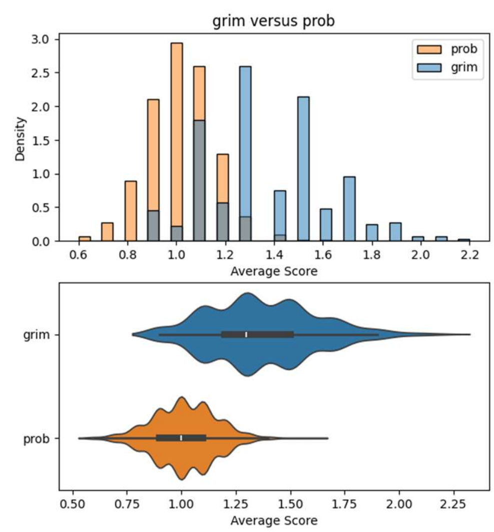

Figure 2. Histogram plot of grim and prob

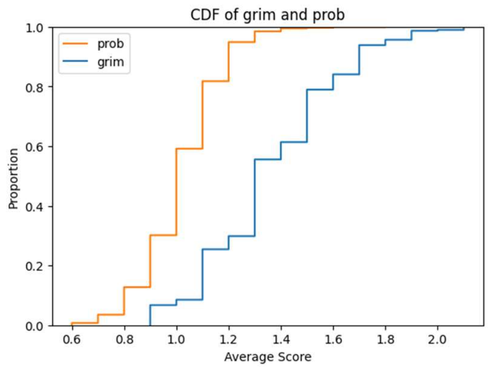

Figure 3. CDF plot of grim and prob

Note: In the above case the cooperation probability is 25% another 75% of the time it will defect.

Grim vs. Reactive

The distributions of the two strategies in this comparison are not similar, with peaks located around 1.4 for grim and 0.8 for reactive. There is a little bit more variance in the "reactive" strategy, with some examples scoring higher than 1.8. The "reactive" strategy has a larger percentage of lower scores early on, according to the CDF plot, but the "grim" strategy gradually catches up, showing that overall, the two strategies perform similarly, though the "grim" strategy might produce more consistent results. The results for grim with reactive strategy are shown below.

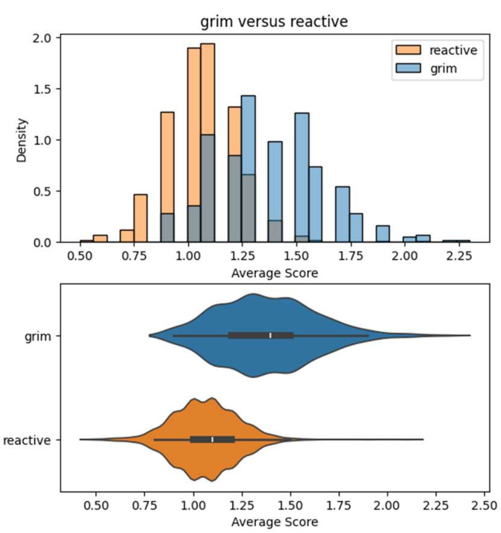

Figure 4. Histogram plot of grim and reactive

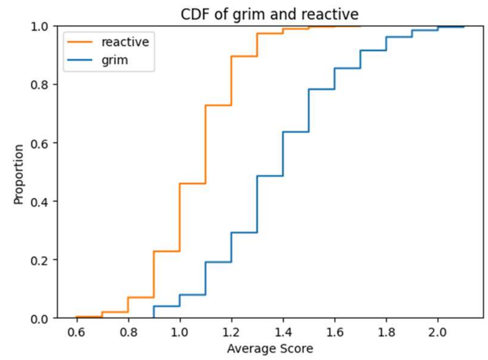

Figure 5. CDF plot of grim and reactive

Note: In the above case we have considered the reactive $y = 0.5, p = 0.25, q = 0.75$, y is probability of the cooperation in first round and with probabilities p or q after opponent cooperates or defects.

Grim vs. Random

In this case, the "grim" strategy has a single peak at 1.8, while the "random" strategy displays a bimodal distribution with peaks around 0.5 and 1.8. This suggests that while the "grim" strategy is more stable, the "random" strategy can result in very low or moderately high scores. Given that the CDF for "grim" is to the right of "random," the CDF plot unequivocally demonstrates that the "grim" strategy consistently performs better than the "random" strategy. The "random" strategy is generally less effective due to its high probability of low scores. The results for grim with random strategy are shown below.

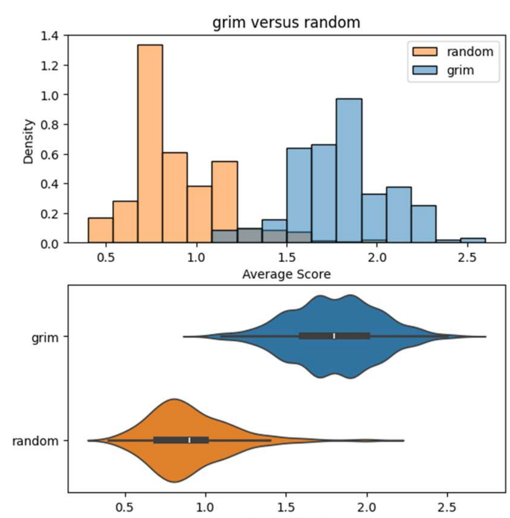

Figure 6. Histogram plot of grim and random

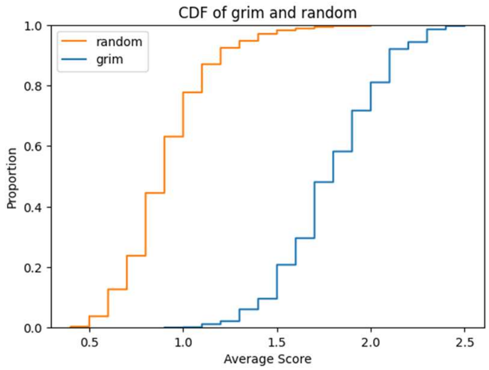

Figure 7. CDF plot of grim and random

Note: In the above case the cooperation probability is 50% another 50% of the time it will defect.

Grim vs. ZD

The distributions of the two strategies are with peaks located at 1.4 for grim and 0.7 for ZD. In comparison to the "grim" strategy, the "ZD" strategy exhibits a similar wider distribution, with more scores above 1.6. Given that the CDF for "ZD" is to the left of "grim" at lower score levels, the CDF plot suggests that the "grim" strategy can attain higher scores more frequently than the "ZD" strategy. The results for grim with ZD strategy are shown below.

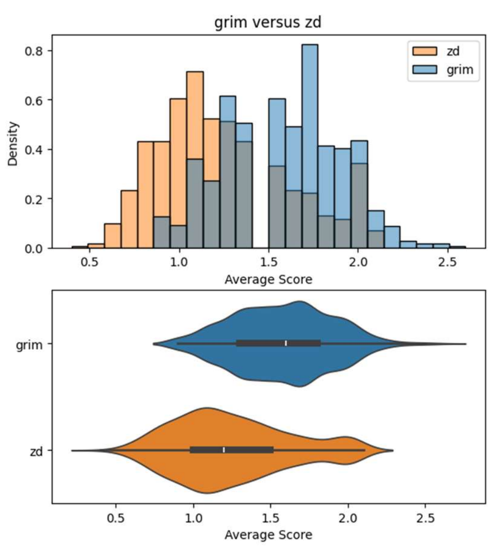

Figure 8. Histogram plot of grim and ZD

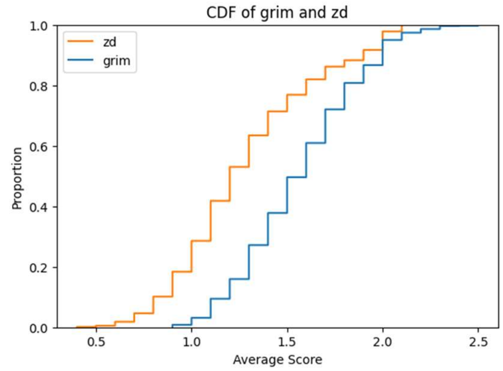

Figure 9. CDF plot of grim and ZD

Note: In this class of memory-one strategies that guarantee that a player's long-term average payoff in the infinitely repeated, two-player prisoner's dilemma (2IPD) will be related to his opponent's according to a fixed linear equation.

5. Application of the Prisoner's Dilemma

Economics: Businesses operating in a competitive market may encounter a scenario similar to the Prisoner's Dilemma. When two businesses are debating whether to reduce prices, for instance, it would be best for both of them to maintain high prices (cooperate). But if one lowers prices (defects), the other may feel compelled to follow as well, which would lower profits for both.

Politics: Nations debating whether to participate in an arms race encounter a similar circumstance. Both parties benefit from cooperation (disarmament); however, if one defected (built up arms), the other might feel pressured to follow the example of which would raise tensions and lower security.

Environmental Policies: When it comes to problems like climate change, nations are faced with a dilemma. Everyone gains when there is cooperation in reducing emissions, but there is a risk that global environmental degradation may result from the temptation to deviate and profit from others' reductions without making one's own.

Social Interaction: Situations like trust, teamwork, and rivalry in daily life frequently mirror the characteristics of the prisoner's dilemma. For example, two coworkers might have to choose between taking all the credit for a collaborative project (defect) or sharing it with one another (cooperate).

6. Conclusion

In this report we have conduct the tournament between various strategies particularly grim, random, prob, reactive and ZD strategy, we compare the grim with the other strategy and found out that grim performs well by giving the higher score. In conclusion, while the GRIM strategy excels in maintaining high average scores through consistent cooperation enforcement, it may be less adaptable in more complex environments where flexibility and forgiveness could yield better long-term outcomes. The analysis underscores the importance of strategy selection based on the specific dynamics of the game environment.

References

[1] Axelrod, R. (1980). Effective choice in the prisoner's dilemma. *Journal of conflict resolution*, *24*(1), 3-25.

[2] Kuhn, S. (1997). Prisoner's dilemma.

[3]https://en.wikipedia.org/wiki/Prisoner%27s_dilemma

[4] https://plato.stanford.edu/entries/prisoner-dilemma/strategy-table.html

[5] Dal Bó, P., & Fréchette, G. R. (2019). Strategy choice in the infinitely repeated prisoner's dilemma. *American Economic Review*, *109*(11), 3929-3952.

[6] Romero, J., & Rosokha, Y. (2018). Constructing strategies in the indefinitely repeated prisoner's dilemma game. *European Economic Review*, *104*, 185-219.

[7] Romero, J., & Rosokha, Y. (2023). Mixed Strategies in the Indefinitely Repeated Prisoner's Dilemma. *Econometrica*, *91*(6), 2295-2331.