Analysis of Game Parameters in Prisoner's Dilemma

Shreyas Marwah

2310110604

Shiv Nadar Institute of Eminence, Delhi NCR

sm451@snu.edu.in

Abstract

This project aims to investigate how the initial parameters:

1) Defector-reward multiplier (α)
2) Initial fraction of cooperators in the total population (ϕ_0)

influence the **composition of population at equilibrium state (ϕ*)** and the **risk of total collapse into full defection (ψ)** in the spatial Prisoners Dilemma. To this end, we employ a Monte Carlo simulation approach, averaging outcomes over 100 independent runs for each parameter configuration.

To automate this simulation, we use the pynetlogo python library.

Introduction

In Game Theory, one of the most studied games is the Prisoner's Dilemma. The dilemma demonstrates a conflict between rational individual behaviour and the benefits of cooperation in certain situations. The setup of the game is as follows:

Two suspects are apprehended by the police. The police do have enough evidence to convict these two suspects. As a result, they separate the two, visit each of them, and offer both the same deal: "If you confess, and your accomplice remains silent, he goes to jail for 10 years, and you can go free. If you both remain silent, only minor charges can be brought upon both of you and you guys get 6 months each. If you both confess, then each of you two gets 5 years."

Even the layman can clearly see that cooperation gets them both the least amount of cumulative punishment. Morality also pushes people to do no harm to others. Thus, one would expect the two prisoners to cooperate and suffer equally. However, objective mathematical analysis would reveal that the equilibrium lies in mutual defection, not cooperation.

Background

Payoff Matrix

In games where the players have a set of choices that they can make, each with corresponding awards or punishments, a payoff matrix is commonly used for illustration and analysis. The payoff matrix for this specific game is as given below:

	Cooperate	**Defect**
Cooperate	(0.5, 0.5)	(0, 10)
Defect	(10, 0)	(5, 5)

Prisoner's Dilemma Payoff Matrix

In general, the payoff matrix for a simple prisoner's dilemma is described by the payoff matrix given below, where b is the benefit of cooperation, c is the cost of cooperation and b>c>0.

	Cooperate	**Defect**
Cooperate	(b-c, b-c)	(-c, b)
Defect	(b, -c)	(0, 0)

General Prisoner's Dilemma Payoff Matrix

This payoff matrix represents the possible choices made by the row player (who selects a row) and the column player (who selects a column). Each cell shows the outcome based on their combined choices. The numbers in brackets represent the payoffs, where the first number is the reward for the row player and the second number is for the column player. Traditionally, you play as the row player, against an opponent who is the column player.

Let us determine the best course of action on a case-by-case basis.

Case 1: Column player cooperates.

- If you cooperate you both get 6 months in jail.

- If you defect, you get to go free, while the other player is incarcerated for 10 years.
- Thus, for you, the better strategy is to defect

Case 2: Column player defects.

- If you cooperate, you are incarcerated for 10 years
- If you defect, you get a 5-year sentence
- Thus, for you, the better strategy is to defect

We see that our best course of action is independent of what the other player chooses. A rational player would, thus, always defect in such a scenario. We typically consider ourselves to be rational creatures, hence, we would expect everyone to defect all the time. But that is rarely the case, we see often, people choose to cooperate. Let us explore the reason of this contradiction.

Reciprocity

Reciprocity, in simple terms, is the concept of doing what others do to you. The most basic form of reciprocity is **direct reciprocity**, where players keep track of the previous interaction with other players and simply copy their opponent's choice for the next interaction. A player that follows such a strategy is called a tit-for-tat player. The payoff matrix in such a case is:

	TFT	**Defect**
TFT	$(\frac{b-c}{1-\omega}, \frac{b-c}{1-\omega})$	(-c, b)
Defect	(b, -c)	(0, 0)

Direct reciprocity payoff matrix

Where ω is the probability of meeting the same player again. This can also be interpreted as a measure of the viscosity of the population.

Another form of reciprocity is **indirect reciprocity** where all individuals keep track of the "reputation" of all other individuals. If an individual has cooperated with others in the past then others are likely to cooperate with them, and vice versa. Players that follow such a policy are called discriminators. The payoff matrix for such a case, if q is the probability that the discriminator can discriminate correctly, is:

	Discr.	**Defect**
Discr.	(b-c, b-c)	(-c(1-q), b(1-q))
Defect	(b(1-q), -c(1-q))	(0, 0)

Indirect reciprocity payoff matrix

The setup for this project, however, demonstrates **spatial reciprocity**, where players only interact with their Moore neighbourhood. In every iteration, for each player, a score is calculated according to the following scheme:

1) If the player cooperates, the score is equal to the number of immediate neighbours that cooperated.
2) If the player defects, then the score is equal to the number of neighbouring cooperators multiplied by a user-defined defection-award value.

For an interaction between a single player and a neighbour, we can describe the playoff matrix as such:

	Cooperate	**Defect**
Cooperate	(1, 1)	(0, α)
Defect	(α, 0)	(0, 0)

Where α is the defector-award multiplier.

Based on this score, the player decides their move for the next iteration. The chooses to follow the same strategy as the neighbour that had the highest score in the previous iteration. Over multiple iterations with a large enough number of players, interesting patterns emerge. Varying the initial parameters like α and ϕ_0 also greatly affects how the game plays out.

Methodology

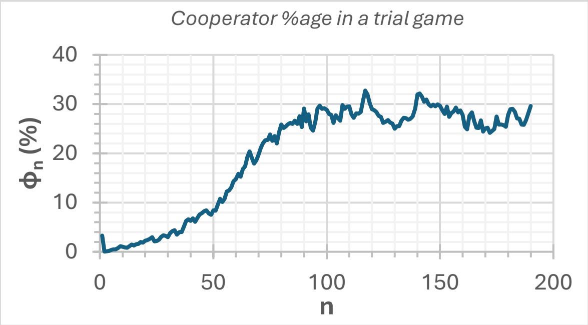
In a game where players have only two choices—**to cooperate or defect**—the state of the population at any (n^{th}) iteration can be fully described by the **fraction of cooperators**, $\boldsymbol{\phi_n}$, at that time step. Since there are only two possible strategies, the fraction of defectors is simply $\mathbf{1 - \phi_n}$. Thus $\boldsymbol{\phi_n}$, alone, can fully describe the state of the game at any given time.

To simulate the spatial prisoner's dilemma, we use the "Prisoner's Dilemma Basic Evolutionary" model from the NetLogo Models library. The internal representation and code of this model is largely irrelevant for us except the two initial parameters which must be modified for each run.

The simulation cannot be allowed to run indefinitely; hence an end state must be defined.

Through trial and error, it is determined that 100 ticks are enough for a game to reach a stable state (where the ratio of defectors to cooperators does not change much), regardless of the initial parameters.

Let ϕ_n be the fraction of cooperators at the **n**th **tick.** The graph below shows the evolution of ϕ_n with respect to n in a trial game.



The initial parameters for this game were chosen to demonstrate a worst-case scenario, where the game takes the highest possible number of ticks to reach equilibrium. Even in this case, we see that the proportion of cooperators in the population becomes hovers around a constant mean after 100 ticks. Hence, we can safely say that the proportion of cooperators at 100 ticks is close to equilibrium.

We can see, however, that this value is quite unstable, even after reaching equilibrium. This necessitates that an average be taken over several runs to give an accurate measure of the cooperator proportion at the equilibrium state. For this purpose, we shall run each unique combination of the two parameters 100 times and take the mean of their end states.

Thus, if ϕ_{100} is the proportion of cooperators at the 100th tick, ϕ^* is simply ϕ_{100} averaged over 100 independent runs.

$$\phi^*(\alpha, \phi_0) = \text{avg}(\phi_{100})$$

Simulation Setup

We define two input parameters:

- ϕ_0 (%) = *[5,10,30,50,70,90,95]*
- α = *[1.25, 1.33, 1.5, 1.67, 2]*

We generate all unique combinations of these values and run the simulation 100 times, each for 100 ticks. The end-state (cooperator fraction at 100th tick) is stored in a CSV file for each of the 100 runs for each unique combination of starting parameters.

Through pure chance, the simulation sometimes evolves in such a way that the proportion of cooperators becomes zero. This is an irrecoverable state as all the neighbours of every player are now defectors. Hence, they will always choose to defect as the neighbour with the highest score is guaranteed to be a defector. For the calculation of equilibrium cooperator fraction, we discard such runs.

For measuring the risk of total collapse of cooperation, we simply count the number of simulations where the cooperation percentage at the 100th tick is zero, i.e. ϕ_{100} **= 0.** This can be interpreted as the percentage chance that a simulation with some starting parameters can collapse into complete defection, where no cooperators are left.

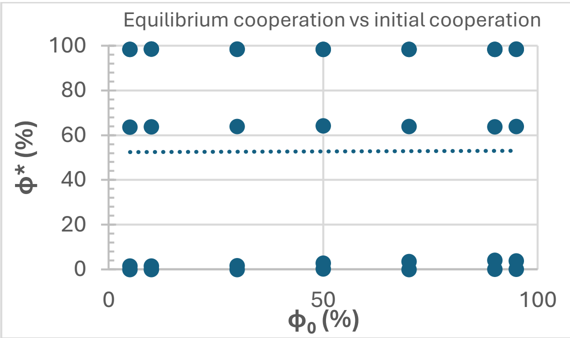
Thus, ψ is simply given by counting the number of runs where total collapse of cooperation was observed and dividing this quantity by the total number of runs, i.e. 100.

$$\psi(\alpha, \phi_0) = \text{count}(\phi_{100} = 0) / 100$$

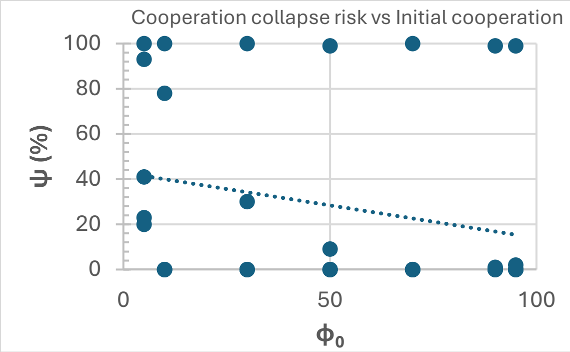
Observation

Graphs for correlation analysis

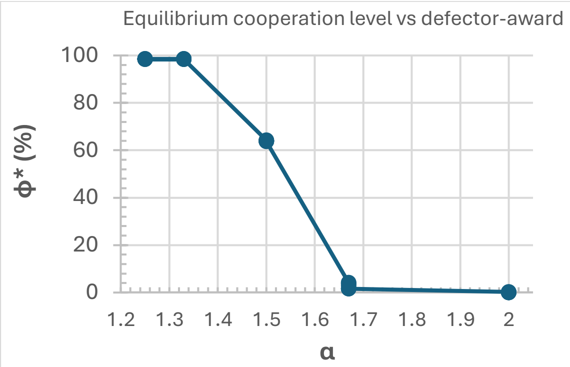
To analyse how the two starting parameters affect our risk of collapse into defection and the equilibrium cooperation fraction, we may simply plot graphs between each of the two inputs and each of the two outputs. If a discernible pattern emerges in any of the 4 graphs, then we know that there is some relation between the input and output of that graph.



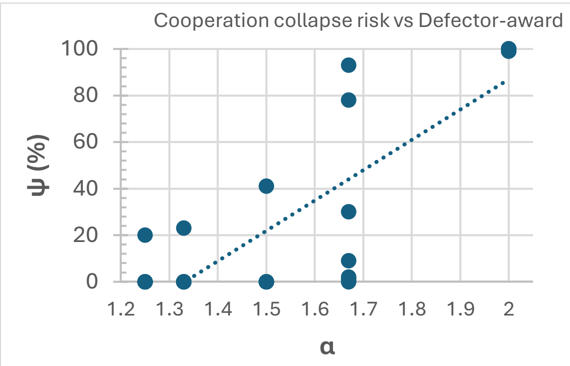
We see that this graph has values which are spread randomly, with no correlation. This likely means that the equilibrium state is not affected by the initial fraction of cooperators in the total population.



This graph has shown some correlations between the risk of total collapse of cooperation and the initial cooperation level. In general, the risk decreases as the initial cooperation level increases.



In this graph, we see that the equilibrium cooperation fraction depends strongly on the defector award multiplier. As the defector award multiplier increases, the equilibrium cooperation level decreases sharply.



Here we see that the risk of complete collapse of cooperation is somewhat correlated with the defector award multiplier. In general, the risk of collapse of cooperation increases with the defector award multiplier.

Analysis and Conclusion

The defector award multiplier can be seen as a measure of the advantage that a defector has over a cooperator. Naturally, if a defector wins by a greater margin over cooperator, then they are likely to be more numerous in the total population. This sheds light on how **ϕ^*** is affected by **α**. Over the course of the game, the fraction of cooperators gradually converges to the equilibrium level. Therefore, altering the initial cooperation level only changes the starting point, not the underlying dynamics of the game, and thus does not affect the final equilibrium. Hence, we saw earlier that **ϕ^*** is independent of **ϕ_0**.

We know that due to the nature of the simulation itself, **ϕ** fluctuates a lot. Hence, if **ϕ_n** is very low, then the chance of **ϕ_{n+1}** reaching 0 is quite high. Thus, the risk of collapse of cooperation (**ψ**) would naturally be high if:

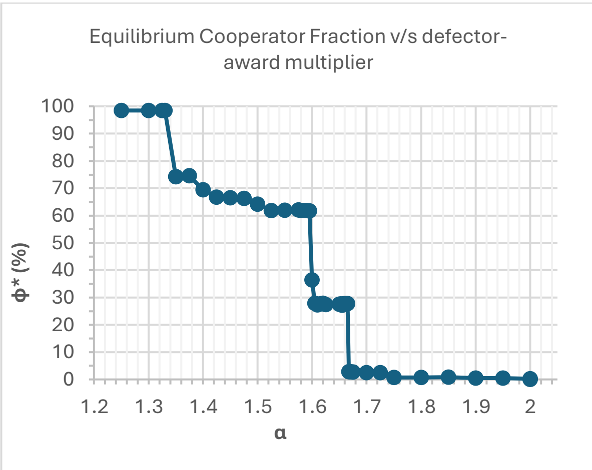
1) we start the simulation off with a very low **ϕ_0** or,

2) **α** is such that the resulting **ϕ^*** is close to 0, causing **ϕ_n** to approach 0 as n increases

This would corroborate our observation of the correlation between **ψ** and **ϕ_0** and, **ψ** and **α**.

Relation between ϕ^* and α

The relation between **ϕ^*** and **α** is of particular interest, because it appears that **ϕ^*** is entirely dependent on **α**. If we re-plot this graph with increased resolution, we obtain:



We see that this graph exhibits sharp transitions at **$\alpha \in \{1.33, 1.6, 1.66\}$**. Nearly the entire range of **$\phi^* \in [0,1]$** is spanned by varying **α** in the range **[1.25, 2]**.

Summary: The equilibrium state, described by the fraction of cooperators in the total population, is entirely dependent on the defector award multiplier and independent of the initial level of cooperation. The risk of collapse of cooperation is dependent on both

the initial fraction of cooperators in the total population and the defector award multiplier.

Limitations and Future Work

The reason for the stepped nature of the relationship between equilibrium cooperator fraction and the defector award multiple is still elusive. Perhaps a comprehensive mathematical analysis of the equilibrium state in spatial prisoner's dilemma could shed more light on the same.

The values and results obtained in this project only apply to a 101x101 grid of players. Whether or not these can be generalised to a larger setup is unknown. Perhaps with a larger number of players, some of the fluctuations observed in the value of **ϕ** could be eliminated.

Bibliography

- Wilensky, U. (2002). NetLogo Prisoner's Dilemma Basic Evolutionary model. http://ccl.northwestern.edu/netlogo/models/Prisoner'sDilemmaBasicEvolutionary. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.
- Wilensky, U. (1999). NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.
- https://pynetlogo.readthedocs.io/en/latest/