Game Theory & The Prisoner's Dilemma: Explorations in Time, Space & Evolution

Joe Cooke, Kaiyi Qian, Ashe Raymond-Barker, Rosenwyn Silver, Callum Stewart

Year 4 Project
School of Mathematics
University of Edinburgh
March 6, 2025

Abstract

In this paper, we discuss the field of Game Theory, where real-world interactions are modelled by mathematical structures. We start by introducing the basic components of one of these models, called a game. We define and discuss Nash's concept of an equilibrium: a steady-state of a game. We then explore topological results like Brouwer's and Kakutani's Fixed Point Theorems, and explain their relevance to the study of Game Theory. The remainder of the paper focuses on the Prisoner's Dilemma, applying it in various different contexts. We look at continuous and repeated formulations of the Prisoner's Dilemma; in the latter we examine Axelrod's tournament. We also consider a variation of the continuous Prisoner's Dilemma where the strategy space takes the form of a torus. Finally we look at repeated discrete games on the torus, discussing how strategies evolve over time, with applications to evolutionary biology.

Declaration

I declare that this thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text.

(*Joe Cooke, Kaiyi Qian, Ashe Raymond-Barker, Rosenwyn Silver, Callum Stewart*)

Acknowledgements

We would like to extend our deepest gratitude to our supervisor, Prof. Minhyong Kim, for his invaluable support, and encouragement throughout this project. His expertise and insightful feedback have not only shaped our understanding but also refined our approach to this project. This project would not have been possible without his continuous support, wisdom, and generosity in sharing his knowledge.

Contents

Chapter 1

Introduction

According to Camus, life is the sum of our choices. While this is commonly taken to affirm the weight and meaning of our own decisions, it could be argued that a more accurate interpretation is that life is the sum of *everyone's* choices. If you drive for hours to see a world-famous musician live only to get stuck in traffic with hundreds of others going to the same concert, or if you arrive at a party to find all the food already eaten, this sentiment will undoubtedly reveal itself to you in the form of crushing despair. In a more professional context, imagine that you choose to start a solo business venture. While the decision to build something from the ground up was yours and yours alone, the success of your business will ultimately depend on who chooses to invest in your idea, and beyond that the decisions occurring in the market itself. Even a hermit who has never interacted with another soul shares the same planet as everyone else, so the effects of how larger societies choose to treat the environment will ripple out, eventually reaching them in one way or another. While our own decisions tend to impact ourselves the most, it is the sum of *everyone's* choices that truly forms the paths we take.

This all being said, an astute reader might have noticed that this is not a paper in philosophy; such a reader would be correct and should not be expecting any grander exploration of decision theory, free will or the philosophical underpinnings of Utility theory. While these are important topics to consider while reading this paper, many aspects of these fields will be largely taken for granted. Rather, this preamble has served as an introduction to the *mathematical field of Game Theory*. While examples of its applications date back as far back as that of the beginnings of human history (for reasons that will become apparent), its establishment as a mathematical field was largely thanks to the work of Von Neumann in the early 20th century. Since then, it has been expanded upon by many, the most notable being John Nash (whose name will be mentioned frequently throughout this paper).

But what exactly *is* Game Theory? Put succinctly, Game Theory is the formal study of strategic interactions between **players** in various settings, and the phenomena that arise from these interactions. Here, the players refer to the participants of the given setting, and the setting itself is called a **game**. In a game, each player has some number of choices, known as **strategies**. The strategies chosen by all players lead to an overall outcome of the game, and each player has *preferences* on the set of possible outcomes. These preferences can then be given numerical values which can be compared against each other, called **payoffs** or **utilities**. When studying games, we assume that every player is *rational*, meaning each are aware of their possible choices and from that, seek to maximise their own payoffs. There is some further discussion on how to interpret this assumption later in the paper (see 2.1), but the theory of what makes a player rational or whether it is representative of all of humanity at any given time is best left to psychologists and philosophers. Sections 1.3 and 1.4 of Ken Binmore's *Playing For Real* [1] give an overview of rationality in relation to Game Theory, for those interested.

The reason for introducing Game Theory in a way that some might have deemed unnecessary, however, is to highlight the following: a player's preferences are defined not just from their strategy, but on *everyone's* strategies. Imagine that Alice, after getting stuck in traffic trying to attend a large concert, learns her lesson and invites three of her friends Bob, Charlie and Daniela to a smaller concert. Alice would enjoy herself the most if all her friends came with her, therefore *preferring* this to none of her friends joining. However, there could be additional scenarios where some of her friends decide to join but others don't, where each have their *own* preferences. For example, Bob would enjoy himself the most if Charlie didn't go and vice versa (they have history), while Daniela would have the most fun with just Alice; Daniela would also enjoy herself if only one of Bob or Charlie joined them, but not if both were there. This example serves to emphasise that when dealing with interactions in games, every player can affect the overall satisfaction of every other player. As a whole, the strategies and preferences of players embedded within a game can produce very interesting results depending on how we define them, but as it turns out these games have *universal properties* that explain a wide range of natural and social phenomena. The most interesting of these properties are called **equilibria**, which we will explore in much greater depth in the following chapter.

Now imagine a new scenario: two criminals have been arrested on multiple charges and are each being interrogated in separate rooms, with no way of communicating. The police candidly tell them that there is currently a lack of evidence for their greater crimes, so each is being sentenced to one year in prison on a lesser charge. However, both are given an ultimatum. If they agree to testify against their partner and *confess*, they will be set free and their partner will serve five years on the primary charge. But this is only true if their partner stays quiet: if they *both* confess, they will each serve three years in prison. What should either do if they wish to serve as little time as possible? This problem is famous within Game Theory and is known as the Prisoner's Dilemma: it is an example of a *toy game* in that it is simple but creates interesting discussion when analysed. When looking at toy games it is often helpful to visualise each player's strategies and payoffs using a *payoff table*:

		Player 2	
		Silent	Confess
Player 1	Silent	(3,3)	(0,5)
	Confess	(5,0)	(1,1)

Note that we can consider the Prisoner's Dilemma either in terms of minimising cost (i.e. serving the shortest sentence) or maximising gain (in that serving less time results in a greater satisfaction or payoff): the table above uses the latter and will be the standard used throughout this paper. In any case, we can see from the table that the outcome that has the highest total payoff across both players is when each stay quiet. So, wouldn't they just do that? It's true that neither player confessing is a good outcome for both, but there is also an incentive to defect and rat on their partner to *get an even better outcome for themselves*. If we were also to fix the strategy of Player 2 i.e. Player 1 somehow knows Player 2 will stay quiet, it is in Player 1's best interest to confess if they wish to maximise their payoff, as they would receive a payoff of 5 instead of 3. But if Player 1 knows Player 2 will confess, it is still in Player 1's best interest to confess, as they will receive a payoff of 1 instead of 0. This is also true if Player 2 were to know Player 1's strategy. We say then that staying quiet in the Prisoner's Dilemma is *strictly dominated* by confessing, as if the opponent's strategy is known beforehand confessing is always the *best* response.

We can also consider played strategies not individually but as a collection. Looking at the above table we can see four possible outcomes: both players stay quiet, both players confess, Player 1 stays quiet but Player 2 confesses and finally, Player 2 stays quiet but Player 1 confesses. We call any single 'joint' outcome a **strategy profile**. Now, consider the strategy profile of both players confessing, but each knows the other will do the same. As this outcome stands, both players will receive a payoff of 1. But if one were to change their strategy at the last second, would doing so be beneficial? The answer is no, as their payoff would decrease from 1 to 0. The fact that this is true for *both* players makes the outcome of both players confessing special, in that such an outcome is called a **Nash Equilibrium**. This concept will appear throughout our paper, just as it appears throughout all of Game Theory: it is a fundamental solution concept i.e. a possible way to 'solve' a game.

Solution concepts attempt to answer a fundamental question about the Prisoner's Dilemma (and games in general): what strategy is 'rational', given that one player cannot predict the behaviour of the other? A common error (known as the *twins* fallacy) is to assume that two players operating 'rationally' will arrive at the same conclusion and therefore play the same strategy. This is incorrect for a few reasons, but most share the common link that such an assumption assumes a structure on the Prisoner's Dilemma that does not accurately represent it i.e. they are analysing a different game. If both players were to *always* apply the same logic and from that play the same strategy there is effectively only one choice being made, so it is no longer a two-player game! The prisoner's dilemma hinges on the fact that the two players act *independently* of each other. In general, rationality only extends as far as aiming to maximise one's own payoff; it cannot be used to predict other players' behaviour, at least when only considering one-shot games (we shall explore repeated instances later).

People often draw the wrong conclusions from the Prisoner's Dilemma. The explorations referring to strictly dominated strategies and Nash Equilibria imply that, given we cannot predict the other player's behaviour, a rational player should confess and thereby betray their partner. From there, some conclude that

Game Theory denies co-operation in *all* cases. What Game Theory attempts to explain does not refer to the nature of humanity or other thinking beings, but of their *conditions*: even if players have a natural tendency towards co-operation, the Prisoner's Dilemma itself does not foster an environment encouraging it. What can be said though, of *alterations* to the Prisoner's Dilemma? This is what our paper aims to explore, through a mixture of conjectures and computational approaches. Up until this point, the discussion has been mostly theoretical: we must now delve into Game Theory with some mathematical rigour.

Chapter 2

Strategic Games

2.1 Strategic Games and Nash Equilibria

We begin by introducing the fundamental ideas we will need for our discussion, starting with the concept of a game. We use many of the basic definitions found in Martin J. Osborne and Ariel Rubinstein's "A Course in Game Theory" [2], and adopt much of their notation.

Definition 2.1. *A **strategic game** is a structure*

$$\langle N, (A_i)_{i \in N}, (\succsim_i)_{i \in N} \rangle$$

where:

- $N = \{1, \ldots, n\}$ *is the set of* n *players.*
- *For* $i \in N$, A_i *is the set of actions available to Player* i.
- $\succsim_i$ *on* $A = \times_{j \in N} A_j$ *is the preference relation of Player* i.

The product A can be thought of as the set of states or **strategy profiles** of the game. If the game has n players, then elements of A are n-tuples $(a_1, ..., a_n)$ where $a_j \in A_j$ represents Player j's strategy. The preference relation $\succsim_i$ represents how Player i **weighs** the states of the game. For example, if Player i prefers the strategy profile $a := (a_1, ..., a_n)$ over the strategy profile $b := (b_1, ..., b_n)$, then we write $a \succsim_i b$. If each player has only finitely many actions available to them (A_i is a finite set for each $i \in N$) then the game is **finite**.

Sometimes it is more convenient to work with concrete functions than relations. When this is the case, we instead use the **Utility function**. A Utility function is a function for each player $u_i : A \to \mathbb{R}$ assigning to each state of the game a numerical value, representing how highly they value this outcome. For states $a, b \in A$, if Player i has $a \succsim_i b$, then we must have $u_i(a) \geq u_i(b)$.

We can think of $u_i(a)$ as the value of the reward Player i receives from the outcome a. This is often more useful than the preference relation, as $a \succsim_i b$ just tells us a is better than b, but the values $u_i(a), u_i(b)$ tell us *how much* better a is than b. In other words a Utility function is a *quantitative* measure, as opposed to a preference relation which is a *qualitative* one. When we think of a game in this alternative way we will instead write it as $G = \langle N, (A_i)_{i\in N}, (u_i)_{i\in N}\rangle$.

Game Theory attempts to model real world situations using this simple game structure. Perhaps the most famous example is the Prisoner's Dilemma, which we will explore in depth throughout this paper. We provide the following example to help illustrate this definition.

Consider a boy and girl eating at an Italian restaurant. The girl orders pizza and the boy orders pasta. Halfway through the meal, the boy decides he doesn't like his pizza. He can choose whether to ask the girl to swap, or keep quiet and finish his pasta. If he asks to swap, the girl can accept or decline the swap. The girl would prefer to eat the pizza she originally ordered, but may still swap to accommodate her friend. We can model this as a game $G = \langle N, (A_i)_{i\in N}, (\succsim)_{i\in N}\rangle$ with the following components:

- The set of players $N = \{1,\ 2\}$ where Player 1 is the boy and Player 2 the girl.
- The set of actions $A_1 = \{\text{ask, don't ask}\}$ and $A_2 = \{\text{accept, reject}\}$.
- The preference relations $\succsim_1$ and $\succsim_2$.

We can define the preference relations as follows. Player 1's ideal situation is swapping with Player 2. But Player 1 would rather suffer in silence with his dish than ask for a swap and get rejected. This leads to the preference relations for Player 1:

$$\text{(ask, accept)} \succsim_1 \text{(don't ask, -)} \succsim_1 \text{(ask, reject)}.$$

In the middle action profile we have left Player 2's action blank, since if Player 1 does not ask then Player 2 has no choice to make. For Player 2's preferences, she would rather keep her dish, but if Player 1 asks to swap she would rather swap to be considerate to him. This leads to the following preference relations for Player 2:

$$\text{(don't ask, -)} \succsim_2 \text{(ask, accept)} \succsim_2 \text{(ask, reject)}.$$

We have now formulated this as a strategic game, and can analyse possible best strategies for each player. If we want to instead talk about Utility functions, we can assign values to each of the states of the game above so that the values satisfy these preference relations.

In order to analyse strategies for players, it would be helpful to have a notion of when a game is in standstill; a state where no player wants to move. We now rigorously define the previously explored concept of Nash equilibria. For this, we need some notation. We denote by $A_{-i} = \times_{\substack{j\in N\\ j\neq i}} A_j$ the set of actions of all players except Player i. Each $a_{-i} \in A_{-i}$ is an action profile with Player i's action removed. Then $(a_{-i}, b) \in A$ is the action profile defined by Player i playing action b, and the other players' actions described by a_{-i}. Given $a_{-i} \in A_{-i}$, we define $B_i(a_{-i})$ to be the set of Player i's best responses to a_{-i}:

$$B_i(a_{-i}) = \{a^* \in A_i : (a_{-i}, a^*) \succsim_i (a_{-i}, a_i) \text{ for all } a_i \in A_i\}.$$

We call the (set-valued) function B_i the **best response function** of Player i. In other words, given an action profile of all other players' actions, the best response function returns the set of best actions for Player i to respond with.

Definition 2.2. *An action profile $a^* \in A$ is a* ***Nash Equilibrium*** *if $\forall i \in N, a_i \in A_i$ we have: $u_i(a_i^*, a_{-i}^*) \geq u_i(a_i, a_{-i}^*)$.*

Thus a Nash Equilibrium is a state of the game where every player is playing their best possible action in response to the actions of the other players. To better understand this definition, we look at an example.

The Election Problem [2, Exercise 19.1]

Consider a game where n people choose whether to enter a political race. If they choose yes, they must choose a position in the interval $[0, 1]$. This is an example of a game with a *continuous* strategy space, which we will look at more in 4.2. There is a distribution of citizens' political positions given by a density function f on $[0, 1]$; each political candidate attracts the citizens that are closer to them than to any other candidate. If k candidates are in the same position, they win $1/k$ of the nearest votes each. The preference of each candidate is as follows:

$$\text{unique winner} \succsim \text{joint winner} \succsim \text{don't compete} \succsim \text{lose.}$$

Let us consider the possible Nash Equilibria of this game in the case $n = 2$, when there are only two candidates. We start by noting that there cannot be a Nash Equilibrium where both players choose to compete and one loses, because then the losing player would choose to not compete instead, as not competing is preferable to losing. There also clearly cannot be an equilibrium where both do not compete, as changing your position to compete would win you all of the votes. Thus, the equilibria must be found in a scenario where the two players draw for first, or where one does not compete. Note that drawing for first can happen in two cases; one where the candidates choose the same position, and one where they choose different positions but are awarded the same amount of votes.

First consider if only one competes, resulting in the competing player getting 100% of the vote. Then the non-competing player will always be better off by copying the competing player's position and drawing with $\frac{1}{2}$ of the vote, as drawing for first is preferable to not competing. So there cannot be a Nash Equilibrium where one player does not compete.

Now consider the scenario where the candidates draw for first. If they choose different positions, then one can always improve their results by moving closer to the other player to 'steal' some of their votes. Suppose they draw for first by choosing the same position. We claim the only way this is a Nash Equilibrium is if this position is the median of the density function. First, to see that it is a Nash Equilibrium, if both candidates choose the median they each receive $\frac{1}{2}$ of the vote. If one candidate moves away from the median they decrease their votes and lose, as the other candidate will now cover more than half of the voters. Thus it is a Nash Equilibrium. Consider now if the candidates choose the same position which is not the median. Then either player can improve their position by moving to the median, as this will grant them more than half of the vote. So the only Nash Equilibrium in the case $n = 2$ is when both candidates choose the median position, with equal numbers of citizens on each side.

Suppose instead that $n = 3$. We claim that there are now no Nash Equilibria. Let us start by checking if the Nash Equilibrium from the $n = 2$ case is still an equilibrium. Suppose all three players choose the median. Then they each get $\frac{1}{3}$ of the total vote and draw for first. Suppose one candidate, which we call p_1, moves to a position $a \in [0, 1]$ left of the median, such that the proportion of voters with a position to the left of a is $\frac{1}{2} - \varepsilon$ with $\varepsilon < \frac{1}{2}$. Now p_1 gets $\frac{1}{2} - \varepsilon$ of the total vote from the voters to the left of a, and half of the share of the votes between p_1 and the median. Since there is ε of the total vote between p_1 and the median, p_1 now gets a proportion of the vote given by

$$\frac{1}{2} - \varepsilon + \frac{\varepsilon}{2} = \frac{1 - \varepsilon}{2}.$$

The votes that each of p_2 and p_3 get by staying at the median is the remainder of the vote divided by 2, since they must share the rest of the votes between them. Thus they each get a proportion of $\frac{1}{2}(1 - \frac{1-\varepsilon}{2}) = \frac{1+\varepsilon}{4}$ of the vote. It follows that p_1 wins as long as $\frac{1-\varepsilon}{2} > \frac{1+\varepsilon}{4}$ which occurs when $\varepsilon < \frac{1}{3}$. We can conclude that if all three players draw for first at the median, one player can become the unique winner by moving to a position less than a third of the vote away. Thus the scenario where all players choose the median is no longer a Nash Equilibrium.

It can be shown in the same way as the $n = 2$ case why the other scenarios are still not Nash Equilibria.

Of course, in the real world this model has a lot of limitations. Political parties might favour integrity over winning popularity, sticking to their morals even when they know it might lose them votes. It would be nice to think that political parties prefer to represent their own values, rather than simply conforming to the most popular position. This could be incorporated into the game; we could assign to each candidate a value in $[0, 1]$ that represents the candidate's **true belief**, and introduce a new payoff that decays the further the candidate moves from this belief. But again, in reality, it seems impossible to place a nuanced political stance on a scale from 0 to 1!

In Game Theory, when we look at the best response for a player, we typically assume selfishness. We expect every player to care about maximising their individual payoff, with no regard for what the outcome is for anyone else. Perhaps there are real-world situations where this is true: in a first-price auction, players only care about beating the bid of every other player. Since the only options are win or lose, in order for one to win, every other participant must lose. However, there are many real-world situations where we do care about the outcomes of others and where we value *compromise*. This can be driven simply by human empathy or in the interest of maintaining positive relationships, such as in political negotiations where countries may want to reach a mutual agreement to preserve peace.

If we wish to model this type of situation with Game Theory, we have to define an alternative motive for players. So far we have assumed that a player's goal is to maximise their own payoff. What if instead players wish to maximise the game's **total payoff**?

Definition 2.3. *For a game* $\langle N, (A_i)_{i\in N}, (u_i)_{i\in N}\rangle$ *with* $|N| = n$*, we define the* ***total payoff*** *of an action profile* $a \in A$ *to be* $U(a) = \sum_{j=1}^{n} u_j(a)$*.*

This is the sum of the payoffs of all players, and suggests a possible way of defining a fair strategy. We could say that a fair strategy $a \in A$ is one satisfying $U(a) \geq U(b)$ for all $b \in A$. Note that it is entirely possible for there to be more than one fairest strategy by this definition.

If we assume that the players are altruistic rather than selfish, then every player could want to maximise this total payoff, leading to all players agreeing to play a fair strategy. This seems like a nicer outcome, but it is flawed. Consider the following payoff table.

		Player 2	
		Y_1	Y_2
Player 1	X_1	(5,5)	(10,0)
	X_2	(2,3)	(3,2)

Looking at the payoffs, intuitively it would seem that the action profile (X_1, Y_1) with payoffs $(5, 5)$ is the fairest strategy. While this *is* a fair strategy by the above definition since it sums to 10, so is the action profile (X_1, Y_2) with payoffs $(10, 0)$ which also sum to 10. But this one seems to be the most unfair strategy! Player 1 gets the entire payoff, whilst Player 2 gets nothing.

If the number of players n is large, we can instead maximise the **mean payoff** $\frac{1}{n}U(a)$ to solve this problem. However, we see in the $n = 2$ case above that we still get (X_1, Y_2) as a fair strategy. How can we eliminate this fake 'fair strategy'? A possible solution is to involve the distance between the payoffs so that payoffs which are closer together are deemed fairer. We adapt the definition of fair strategy as follows. Given a strategy profile $a \in A$, define the function

$$f(a) = \frac{1}{n}U(a) - \sqrt{\frac{1}{n}\sum_{i=1}^{n}(u_i(a) - \frac{1}{n}U(a))^2}.$$

Note that the negative term is the standard deviation, so this function punishes payoff profiles that deviate from the mean, like (X_1, Y_2) above. Now we redefine the notion of fair strategy as follows.

Definition 2.4. *For a game* $\langle N, (A_i)_{i\in N}, (u_i)_{i\in N}\rangle$ *with* $|N| = n$*, define a* ***fair strategy*** *to be an action profile* $a^* \in A$ *such that* $f(a^*) = \max_{a\in A} f(a)$*.*

Going back to the example above, we now have that $f(X_1, Y_1) = 5$ and $f(X_1, Y_2) = 0$, and so (X_1, Y_1) is now the unique fair strategy. We have eliminated the 'fake' fair strategy, since the distance between the two payoffs is large. We revisit this idea in Chapter 4 for the Prisoner's Dilemma.

For now, we define another type of strategy.

Definition 2.5. *Let A, B be strategies in a game G. If playing A always gives a better outcome than playing B, then we say A **strictly dominates** B. If playing A is always at least as a good as playing B, we say A **weakly dominates** B.*

If A strictly (weakly) dominates all other strategies, we say A is a **strictly (weakly) dominant strategy**. Possibly the most famous example of a strictly dominant strategy in a game is Defection in the Prisoner's Dilemma, which we explore further in chapter 4. The notion of dominance of a strategy can be very useful at eliminating dominated strategies to find the Nash Equilibria. We consider an example of this below.

The Traveller's Dilemma

Proposed by Kaushik Basu in [3], the Traveller's Dilemma presents a situation where two passengers on an aeroplane have lost their luggage. Somewhat surprisingly, both travellers were carrying an identical antique of value between £2 and £100. The airline wishes to compensate the passengers, but does not know the value of the antique. They separately ask each player to choose a value of the antique in the range $[2, 100]$. There are two possible outcomes:

1. The players choose the same value $x \in [2, 100]$.
2. The players choose different values $x, y \in [2, 100]$ with $x < y$.

In case 1, both players receive the same compensation of x pounds. In case 2, the airline assumes that the lower value x is the correct one, and that the player who chose y is trying to extort more money from them. The player who chose y gets $x - 2$ pounds in compensation, with a £2 fine for lying, whereas the player who chose x receives $x + 2$ pounds, with a £2 bonus for being honest. We can use the idea of dominating strategies to reason what the best choice is for the players.

An intuitive solution is for both players to choose 100, so that they are both awarded a large prize. However the strategy of choosing 99 weakly dominates the strategy of choosing 100. To see this, we show that choosing 99 is better than or the same as choosing 100 in every scenario. Consider the three possibilities:

(i) Your opponent chooses 100. Then choosing 99 gives you the £2 reward, and you get £101. Choosing 100 gives you £100, so 99 is the better strategy.

(ii) Your opponent chooses 99. Then by choosing 99, you both get £99. Choosing £100 you will be fined for dishonesty and get only £97, so choosing 99 is still the better strategy.

(iii) Your opponent chooses $a < 99$. Again choosing 99, you are given the fine and receive £$(a - 2)$. Choosing 100 has no effect as you still picked the higher number and receive £$(a - 2)$. So choosing 99 has the same result as choosing 100.

From the above cases, we conclude that the strategy of choosing 99 weakly dominates the strategy of choosing 100. Thus we may eliminate the strategy of choosing 100, deducing that there is no point in ever playing it. However, we may apply the same logic to see that the strategy of choosing 98 weakly dominates the strategy of choosing 99. In fact, choosing $i-1$ always weakly dominates choosing i. Recursively, this leads to the best strategy for both players being to choose 2! This is the unique Nash Equilibrium; assuming your opponent picks 2, picking anything other than 2 will yield you £0.

Here we used strategic dominance to eliminate strategies with induction and find the Nash Equilibrium. The result seems a little counter-intuitive; the players receive a very low reward relative to what they could have received by cooperating. It turns out that this is a theme that crops up a lot in Game Theory!

2.2 Existence of Nash Equilibria

There are various results about when we can guarantee that a Nash Equilibrium exists. We investigate a major one, Nash's Theorem, in Chapter 3. For now we look at a less general equilibria existence condition. In order to prove it, we first need to state Kakutani's Fixed Point Theorem.

Theorem 2.6. *[2, Lemma 20.1] [Kakutani's Fixed Point Theorem] Let X be a compact convex subset of $\mathbb{R}^n$ and let $f : X \to X$ be a set-valued function for which*

1. *for all $x \in X$ the set $f(x)$ is nonempty and convex,*
2. *the graph of f is closed.*

Then there exists $x^* \in X$ such that $x^* \in f(x^*)$.

This is a generalisation of Brouwer's Fixed Point Theorem, and we outline the proof of both in Chapter 3. Note that since f is set-valued, a fixed point corresponds to inclusion in the set, rather than equality. To better understand Kakutani's Fixed Point Theorem, we explore the definitions of the terms used.

Definition 2.7. *A set valued function $f : X \to X$ has a* ***closed graph*** *if for all sequences $\{x_n\}, \{y_n\} \in X$ such that $y_n \in f(x_n)$ for all n and $x_n \to x$, $y_n \to y$, we have $y \in f(x)$.*

If we assume Kakutani's Fixed Point Theorem for now, we can prove the following proposition which outlines conditions for a Nash Equilibrium.

Proposition 2.8. *[2, Proposition 20.3] The strategic game $\langle N, (A_i)_{i\in N}, (\succsim_i)_{i\in N}\rangle$ has a Nash Equilibrium if for all $i \in N$*

- *the set A_i of actions of Player i is a non-empty compact convex subset of a Euclidean space*

and the preference relation $\succsim_i$ is

- *continuous*
- *quasi-concave on A_i.*

We start by familiarising ourselves with the definitions in the proposition.

Theorem 2.9. *[Heine-Borel] A subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded.*

Here a closed set is one that contains all of its limit points.

Definition 2.10. *A subset X of $\mathbb{R}^n$ is **convex** if for any $x, y \in X$ and $\lambda \in [0,1]$, we have that $\lambda x + (1-\lambda)y \in X$.*

Essentially this means that for any two points in the set, the line between them also lies in the set.

Definition 2.11. *A preference relation $\succsim_i$ on a set A is **continuous** if for all i, if there exist sequences $(a^k)_k, (b^k)_k \in A$ such that $a^k \succsim_i b^k$ for all k and $a^k \to a$, $b^k \to b$, then $a \succsim_i b$.*

This means that if there are two sequences of action profiles, where every action in the first sequence is preferred to every action in the second, then the limit of the first sequence is preferred to the limit of the second.

Definition 2.12. *A preference relation $\succsim_i$ over a set A is **quasi-concave** on A_i if for every $a^* \in A$ the set $\{a_i \in A_i : (a^*_{-i}, a_i) \succsim_i a^*\}$ is convex.*

Suppose that we are given an action profile $a^* \in A$, which includes the action of Player i, denoted a_i. The set $\{a_i \in A_i : (a^*_{-i}, a_i) \succsim_i a^*\}$ is the set of actions that Player i would be better off changing to, given the rest of the player's actions are fixed. So, if a^* is a Nash Equilibrium, this set will contain a_i, meaning that Player i cannot improve their outcome given this action profile.

Now that we understand the terms, we are ready to prove this result.

Proof of Proposition 2.8. Consider the set-valued function $B : A \to A$ defined by $B(a) = \times_{i \in N} B_i(a_{-i})$, where B_i is the best response function of Player i and a_{-i} is the action profile a with Player i's action removed. For each i, the relation $\succsim_i$ is continuous, therefore the corresponding Utility function u_i is also continuous. Also, A_i is compact, so by the Extreme Value Theorem, u_i will attain a maximum on A_i for every i. Thus $B_i(a_{-i})$ is non-empty for all $i \in N$, as it contains this maximum. Since $\succsim_i$ is quasi-concave on A_i, each $B_i(a_{-i})$ is convex by definition. Next, we show that B has a closed graph. Consider sequences $(a^n), (b^n) \in A$, with $a^n \to a$ and $b^n \to b$. Assume that $b^n \in B(a^n)$ for all n. Then for all $i \in N$, it follows that $b^n_i \in B_i(a^n_{-i})$, so $(a^n_{-i}, b^n_i) \succsim_i (a^n_{-i}, c_i)$ for all $c_i \in A_i$. Then by continuity of $\succsim_i$, it follows that $(a_{-i}, b_i) \succsim_i (a_{-i}, c_i)$ for all $c_i \in A_i$, so $b_i \in B(a_{-i})$ for all i, meaning $b \in B(a)$. Thus, B has a closed graph.

We have shown that the graph of B is closed and that $B(a)$ is non-empty and convex for all $a \in A$, so by Kakutani's Fixed Point Theorem, B has a fixed point. This fixed point corresponds to a Nash Equilibrium of the game since if $a \in B(a)$, we have that a_i is a best response for Player i for all $i \in N$, so every player is playing (one of) their best responses, and cannot improve. □

This proposition turns out to be very useful in the proof of Nash's Theorem, a much more powerful result that we prove in Chapter 3. For now, we see how this result can be useful in its own right in the next section.

2.3 Symmetric Games

We now look at different types of games, each with their own properties and results. The first is the notion of a symmetric game.

Definition 2.13. *A 2-player game is* ***symmetric*** *if* $A_1 = A_2$ *and* $(a_1, a_2) \succsim_1 (b_1, b_2)$ *if and only if* $(a_2, a_1) \succsim_2 (b_2, b_1)$ *for all* $a \in A$ *and* $b \in B$*.*

In other words, both players have the same options available to them, and if one player receives a certain payoff by responding to the other's choice, the other player would receive the same payoff if the actions were swapped.

An example you have probably played yourself is rock-paper-scissors, where Player 1's payoff for playing rock against Player 2's scissors is the same as Player 2's payoff for playing rock against Player 1's scissors (assuming payoff of 1 for a win,0 for a draw, and -1 for a loss). The same is true for rock against paper, scissors against scissors, scissors against paper and **all possible strategy profiles**, hence the game is symmetric. In Section 2.4 we will look at equilibria properties of rock-paper-scissors.

For a symmetric game, we call a Nash Equilibrium a **symmetric equilibrium** if it is of the form (a, a) for some $a \in A$. We come to the following result, which illustrates the power of results like Proposition 2.8 in the previous section.

Proposition 2.14. *A two-person symmetric game satisfying the conditions of Proposition 2.8 has a symmetric Nash Equilibrium.*

Proof. Since $A_1 = A_2$, define A^* to be the shared set of actions, i.e. $A^* = A_1 = A_2$ and let $a_1, a_2 \in A^*$ be any action in A^* of Player 1 or Player 2 respectively. Then, let $B_1 : A^* \to A^*$ be Player 1's best response function, and $B_2 : A^* \to A^*$ be Player 2's best response function. If $b_1 \in B_1(a_2)$ for some $a_2 \in A^*$ then

$$(b_1, a_2) \succsim_1 (a_1, a_2)$$

for all $a_1 \in A^*$. But since our game is symmetric, we also have

$$(a_2, b_1) \succsim_2 (a_2, a_1)$$

So if Player 1 plays a_2, Player 2 prefers b_1 over all of its other possible actions. We therefore have $b_1 \in B_2(a_2)$, and as such $B_1(a_2) \subseteq B_2(a_2)$ for all $a_2 \in A^*$. Applying the same logic but given $b_1 \in B_2(a_2)$, we get that $B_2(a_2) \subseteq B_1(a_2)$, and as such $B_1(a_2) = B_2(a_2)$ for all $a_2 \in A^*$. Thus $B_1 = B_2$.

If from here we can show that either B_1 or B_2 contains a fixed point a^*, then the other will also contain a^*; both (identical) actions from both players will be a best response to the other and so we will have found a symmetric Nash Equilibrium. Without loss of generality, we examine the properties of Player 1; we begin by showing that $B_1(a_2)$ is non-empty for all $a_2 \in A^*$. We have that $A^* = A_1 = A_2$ is compact and therefore is both closed and bounded; we are also given that it is non-empty. Looking at Player 1's preference relation $\succsim_1$, we can define for any of Player 2's possible actions $a_2 \in A^*$ Player 1's respective Utility function $u_1^{a_2} : A^* \to \mathbb{R}$ such that for $a_1 \in A^*$, $u_1^{a_2}(a_1) = u_1(a_1, a_2)$ where u_1 is Player 1's general Utility function. Since $\succsim_1$ is a continuous relation, so must every function $u_1^{a_2}$ be continuous; since A^* is compact we have by the Extreme Value Theorem that each $u_1^{a_2}$ must attain a maximum for some $a_1 \in A^*$. That is, for each of Player 2's actions there exists an action for Player 1 that maximises their payoff; we therefore have that $B_1(a_2)$ is non-empty for all $a_2 \in A^*$.

We have that $\succsim_1$ is quasi-concave, so by definition B_1 is convex; we also have that $\succsim_1$ is continuous and it follows from this that the graph of B_1 is closed. Hence B_1 satisfies the conditions of Kakutani's Theorem (2.6), so there exists $a \in A^*$ such that $a \in B_1(a)$. Since we deduced that $B_1 = B_2$, the same is true for B_2 for the same a. So there exists an a that is a best response to itself for both, and so (a, a) is a symmetric Nash Equilibrium. □

Symmetric games are often simpler to analyse because we can consider the best responses for one player only and apply the same findings to the other player. We see in Chapter 4 that the Prisoner's Dilemma is a symmetric game, which simplifies our investigation somewhat.

2.4 Mixed Strategy Nash Equilibria

Let us consider rock-paper-scissors as a strategic game with two players. Say we have players $N = \{1, 2\}$, a set of possible actions for Player i given by $A_i = \{R, P, S\}$, and a natural Utility function for Player i, given by $u_i : A_1 \times A_2 \to \mathbb{R}$ defined according to the traditional rules of rock-paper-scissors, and giving a payoff of 1 for a win, 0 for a draw, and -1 for a loss. For example, we have that $u_1(R, S) = 1$, $u_1(R, R) = 0$ and $u_1(R, P) = -1$, while $u_2(X, Y) = u_1(Y, X)$ for all $X, Y \in \{R, P, S\}$.

It is not difficult to see that, under Definition 2.2, this game has no Nash Equilibria; given any action profile $(a_1, a_2) \in A_1 \times A_2$ at least one player will always be able to increase their payoff by changing their strategy, assuming that the other player's strategy remains fixed.

This is perhaps unsurprising; in a game of rock paper scissors, if Player 1 knows exactly what Player 2 is going to play, then they are guaranteed to be able to win, but Player 2 is unlikely to be happy with this status quo! In our rock-paper-scissors game, no equilibrium is possible if both players know exactly which action the other is going to take. What if the players were able to choose a strategy which selects rock, paper and scissors each with a given probability? This is the notion of a **mixed strategy**, which we introduce in this section. In Chapter 3 we will show Nash's famous theorem that, if we allow mixed strategies, every finite strategic game has a **mixed strategy Nash Equilibrium**.

As with a Nash Equilibrium, a mixed strategy Nash Equilibrium models a situation where a game is at steady state; given that the choices of the other players remain fixed, no player can improve their standing by changing their own strategy. Where the idea of a mixed strategy Nash Equilibrium differs, is that we allow player's choices to be nondeterministic; player's strategies now take the form of probabilistic distributions.

We previously defined a strategic game to be the triple $\langle N, (A_i), (\succsim_i)\rangle$, where $\succsim_i$ is the preference relation of Player i over the set of action profiles $A = \times_{i\in N} A_i$. In a mixed-strategy context, where we allow player's choice to be nondeterministic, we instead consider a strategic game G to be of the form $G = \langle N, (A_i), (u_i)\rangle$. Here the *expected value* of the Utility function $u_i : A \to \mathbb{R}$ is used to represent the preference relation of Player i over the set of **lotteries on** A.
We shall denote the set of probability distributions on A_i by (ΔA_i), and refer to a distribution $\alpha \in (\Delta A_i)$ as a **mixed strategy** of player i. In a mixed-strategy context, we call an element $a \in A_i$ a **pure strategy** of player i. If the set A_i is finite, then for $\alpha \in \Delta(A_i)$ we let $\alpha(a)$ denote the probability that α assigns to $a \in A_i$, and we define the **support** of α to be the set of elements $a \in A_i$ for which $\alpha(a) > 0$. This is the set of pure strategies that the player assigns a non-zero probability to.

Definition 2.15. *[2, Definition 32.1] Let $G = \langle N, (A_i), (u_i))\rangle$ be a strategic game. The* ***mixed extension*** *of G is the strategic game $\langle N, (\Delta(A_i)), (U_i)\rangle$, where $\Delta(A_i)$ is the set of probability distributions over A_i, and $U_i : \times_{j\in N}\Delta(A_j) \to \mathbb{R}$ maps $\alpha \in \times_{j\in N}\Delta(A_j)$ to the expected value under u_i of the lottery over A that is induced by α. That is, assuming A is finite,*

$$U_i(\alpha) = \sum_{a\in A}\left(\prod_{j\in N}\alpha_j(a_j)\right)u_i(a),$$

where a_j is the (pure) action of Player j in the action profile a.

Definition 2.16. *[2, Definition 32.3] A* ***mixed strategy Nash Equilibrium*** *of a strategic game is a Nash Equilibrium of its mixed extension.*

Lemma 2.17. *[2, Lemma 33.2] Let $G = \langle N, (A_i), (u_i)\rangle$ be a finite strategic game. Then $\alpha^* \in \times_{i\in N}\Delta(A_i)$ is a mixed strategy Nash Equilibrium of G if and only if for every player $i \in N$, every pure strategy in the support of α_i^* is a best response to α_{-i}^*.*

Proof. The game G has mixed extension $\langle N, (\Delta(A_i)), U_i\rangle$. Suppose $\alpha^* \in \times_{i\in N}\Delta(A_i)$ is a mixed strategy Nash Equilibrium of G, but there exists an action a_i in the support of α_i^* such that a_i is not a best response to α_{-i}^*. We note that the function U_i is multilinear, in the sense that

$$U_i(\alpha_{-j}, \lambda\beta_j + (1-\lambda)\gamma_j) = \lambda U_i(\alpha_{-j}, \beta_j) + (1-\lambda)U_i(\alpha_{-j}, \gamma_j),$$

for any mixed strategy profile α, any mixed strategies β_j, γ_j of Player j and any $\lambda \in [0, 1]$. We prove this in Section 3.4. By linearity of U_i, we then have that Player i would be better off transferring the probability assigned to a_i to an action which is a best response to α_{-i}^*. So α^* is not a Nash Equilibrium, since Player i can improve their situation.

Now suppose that every pure strategy in the support of α_i^* is a best response to α_{-i}^* for every player $i \in N$, but α^* is not a Nash Equilibrium. Then there exists $j \in N$ such that there is a mixed strategy α'_j that yields a higher payoff for Player j than α^*_j does in response to α^*_{-j}. Again by linearity of U_i, at least one action in the support of α'_j gives a higher payoff than some action in the support of α^*_j. Then this action in α^*_j is not a best response to α^*_{-j} as claimed. $\square$

Lemma 2.17 is invaluable when trying to find mixed strategy equilibria of finite games. Consider the following strategic game, in which two players each guess a number in the set $\{1, \ldots, K\}$:

A Guessing Game [2, Exercise 36.1]: *Players 1 and 2 each choose a member of the set $\{1, \ldots, K\}$. If the players choose the same number then Player 2 pays \$1 to Player 1; otherwise no payment is made. Each player maximises his expected monetary payoff. Find the mixed strategy Nash Equilibria of this game.*

Let $\alpha^* = (\alpha^*_1, \alpha^*_2)$ be a mixed strategy equilibrium of this game. Let $\alpha^*_i(j)$, for $j \in \{1, \ldots, K\}$, be the probability that α^*_i assigns to the number j. It follows from Lemma 2.17, that if $\alpha^*_1(j) > 0$, then the pure strategy of picking j must be a best response by Player 1 to α^*_2.

Therefore, by considering the game's payoffs, we must have that $\alpha^*_2(j)$ is equal to the maximum probability assigned to any number in the set $\{1, \ldots, K\}$ by α^*_2. In particular, this implies that $\alpha^*_2(j) > 0$, as otherwise α^*_2 assigns probability of 0 to every number in $\{1, \ldots, K\}$, which is impossible.

By Lemma 2.17 again, we must then have that the pure strategy of picking j is a best response by Player 2 to α^*_1, and therefore that $\alpha^*_1(j)$ is equal to the *minimum* probability assigned to any number in the set $\{1, \ldots, K\}$ by α^*_1. Observe that if the minimum probability assigned to a number in the set $\{1, \ldots, K\}$ by α^*_1 was greater than $\frac{1}{K}$, then we would have that $\sum_{i=1}^{K} \alpha^*_1(i) > 1$, so we must have that $\alpha^*_1(j) \leq \frac{1}{K}$. But we require that $\sum_{i=1}^{K} \alpha^*_1(i) = 1$, so the condition that $\alpha^*_1(j) \leq \frac{1}{K}$ whenever $\alpha^*_1(j) > 0$ forces that $\alpha^*_1(j) = \frac{1}{K}$ for all $j \in \{1, \ldots, K\}$.

Now let $j \in \{1, \ldots, K\}$. We know that $\alpha^*_1(j) = \frac{1}{K} > 0$, so by our earlier analysis we have that $\alpha^*_2(j) > 0$, and that $\alpha^*_2(j)$ is equal to the maximum probability assigned to any number in $\{1, \ldots, K\}$ by α^*_2. It follows that all of the probabilities $\alpha^*_2(i)$, for $i \in \{1, \ldots, K\}$ must be equal, and as they must also sum to 1, we have that $\alpha^*_2(i) = \frac{1}{K}$ for all $i \in \{1, \ldots, K\}$.

Hence we have shown that this game has a unique mixed strategy Nash Equilibrium, where both players assign equal probability to every number in the set $\{1, \ldots, K\}$. Further, the expected payoff for Player 1 at this equilibrium is given by

$$\sum_{i=1}^{K} \alpha^*_1(i)\alpha^*_2(i) = \sum_{i=1}^{K} \frac{1}{K^2} = \frac{1}{K}.$$

The observant reader may have noticed a gap in this argument; we started by assuming that a mixed strategy equilibrium of this game exists! We could show directly that both players assigning equal probability to each number in $\{1, \ldots K\}$ *is* an equilibrium, but the next chapter introduces a key result that makes this completely unnecessary. In Chapter 3 we present a proof of Nash's Theorem (Theorem 3.1), which guarantees the existence of a mixed strategy Nash Equilibrium for any finite game.

For now, let us return to the rock-paper-scissors game, that we introduced at the beginning of this section. It is clear that there are no Nash Equilibria where either player uses a pure strategy; if Player 1 knows that Player 2 will play rock with probability 1, then Player 1's unique best response is to play paper with probability 1. This guarantees Player 1 the maximum possible expected payoff of 1. But this is not an equilibrium, as Player 2, knowing that Player 1 will play paper, can increase their payoff from -1 to 1, by instead playing the pure strategy that picks scissors. The same clearly holds with Player 1 in place of Player 2, and paper or scissors in place of rock. But what about mixed strategy equilibria?

Assume that $\alpha^* = (\alpha_1^*, \alpha_2^*)$ is a mixed strategy Nash Equilibrium of this game, and that $\alpha_i^*(X) = 0$ for some $i \in \{1, 2\}$ and $X \in \{R, P, S\}$. Due to the symmetry of the game, we may assume without loss of generality that $i = 1$ and $X = R$, so that $\alpha_1^*(R) = 0$. Consider the expected payoffs for Player 2's pure strategies: rock has expected payoff given by

$$U_2(\alpha_1^*, R) = (1) \cdot \alpha_1^*(S) + (0) \cdot \alpha_1^*(R) + (-1) \cdot \alpha_1^*(P) = \alpha_1^*(S) - \alpha_1^*(P),$$

where R here denotes Player 2's degenerate mixed strategy that assigns probability 1 to R and probability 0 to S, P. Similarly, as $\alpha_1^*(R) = 0$, we have $U_2(\alpha_1^*, P) = -\alpha_1^*(S)$, and $U_2(\alpha_1^*, S) = \alpha_1^*(P)$. By Lemma 2.17 we know that any pure strategy in the support of α_2^* must be a best response to α_1^*. We note that P cannot be a best response to α_1^* as it has payoff $-\alpha_1^*(S) \leq 0$, and S has payoff $\alpha_1^*(P) \geq 0$. So for P to be a best response we must have that $\alpha_1^*(P) = \alpha_1^*(S) = 0$, but then we have that α_1^* assigns zero probability to every pure action, which is impossible. By Lemma 2.17 we then know that $\alpha_2^*(P) = 0$.
If only one of R, S was a pure-strategy best response by Player 2 to α_1^*, Lemma 2.17 would then tell us that α_2^*, Player 2's equilibrium strategy, is pure, as it only assigns positive probability to one action in $\{R, P, S\}$. But we have already discussed above that no equilibrium is possible where either Player's strategy is pure. It follows that R and S must both be best responses to α_1^* and therefore $\alpha_1^*(S) - \alpha_1^*(P) = \alpha_1^*(P)$, so that $\alpha_1^*(S) = 2\alpha_1^*(P)$ As we already have that $\alpha_1^*(R) = 0$, this tells us that $\alpha_1^*(S) = 2/3$ and $\alpha_1^*(P) = 1/3$.
As α_1^* assigns positive probability to both S and P, it follows from Lemma 2.17 again that both S and P must be pure best responses by Player 1 to α_2^*. Considering the payoffs to Player 1's pure strategies, this tells us that $U_1(S, \alpha_2^*) = U_1(P, \alpha_2^*)$, and so

$$\alpha_2^*(P) - \alpha_2^*(R) = \alpha_2^*(R) - \alpha_2^*(S),$$

but we know that $\alpha_2^*(P) = 0$, so we calculate that $\alpha_2^*(S) = 2\alpha_2^*(R)$, and therefore we have $\alpha_2^*(S) = 2/3$, $\alpha_2^*(R) = 1/3$.
We have now completely determined the mixed strategy profile $\alpha^* = (\alpha_1^*, \alpha_2^*)$, as

being given by $\alpha_1^* = (0, 1/3, 2/3)$, and $\alpha_2^* = (1/3, 0, 2/3)$, where the first entry of each vector represents the probability assigned to R, the second to P, and the third to S. We can now directly calculate that the payoff of α^* to Player 1 is given by

$$\begin{aligned} U_1(\alpha_1^*, \alpha_2^*) &= \frac{1}{9}u_1(P,R) + \frac{2}{9}u_1(P,S) + \frac{2}{9}u_1(S,R) + \frac{4}{9}u_1(S,S) \\ &= \frac{1}{9}(1) + \frac{2}{9}(-1) + \frac{2}{9}(-1) + \frac{4}{9}(0) \\ &= -\frac{1}{3}. \end{aligned}$$

But we can immediately see that, given Player 2's equlibrium strategy $\alpha_2^* = (1/3, 0, 2/3)$, Player 1's pure strategy R has an expected payoff of $2/3 > -1/3$, which contradicts that α^* is an equilibrium. It follows that in any mixed strategy Nash Equilibrium of this game, all three pure strategies R, P, S must be in the support of each Player's equilibrium strategy, and therefore by Lemma 2.17, that the pure strategies R, P, S must all have equal payoff, for each player. Calculating the pure-strategy payoffs for each Player and setting them to be equal then gives a system of three equations in three unknowns, which can be easily solved to yield that the only possible mixed strategy Nash Equilibrium of this game is where each Player assigns equal probability (1/3) to each of the pure actions R, P, S. Assuming the result of Nash's Theorem (Theorem 3.1), which we prove in Chapter 3, we therefore have that the unique mixed strategy Nash Equilibrium of rock-paper-scissors is for each player to play rock, paper, and scissors, each with equal probability.

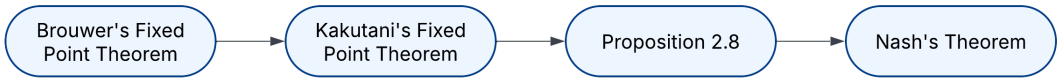
Compare this with our discussion of the pure version of rock-paper-scissors at the beginning of this section. If Player 1 knows exactly which pure strategy Player 2 is going to use, then they can take advantage of this information and guarantee a win, which makes any equilibrium impossible. In the mixed extension of rock-paper-scissors, the unique Nash Equilibrium is given by each player playing the only mixed strategy that gives their opponent no information they can take advantage of.

Chapter 3

Nash's Theorem

3.1 Outlining the Proof

Nash's Theorem is a central theorem in the study of equilibria in games, giving us conditions for the existence of Nash Equilibria. It is a seemingly simple result, that relies on advanced mathematics for its proof. In this chapter, we work towards the proof of Nash's Theorem by following a path of proofs. To help us keep track, this diagram shows the implication order of the results.



Note that we have already proved Proposition 2.8 in Section 2.2, but assumed Kakutani's Fixed Point Theorem in the proof. In this section we will revisit Kakutani's Theorem and discuss the proof, which itself requires Brouwer's Fixed Point Theorem. Once we have understood these results, we will be ready to prove Nash's Theorem, which we state here to see what we are working towards. The original formulation is due to John Nash [4]. We use the following statement of Nash's Theorem, given by Osborne and Rubinstein [2]:

Theorem 3.1. *[2, Proposition 33.1][Nash's Theorem] Every finite strategic game has a mixed strategy Nash Equilibrium.*

There are many famous fixed point theorems in mathematics; they establish conditions for a function f to have a point x such that $x = f(x)$ (or in the case of a set-valued function, $x \in f(x)$). You may be surprised at the application of topological results to Game Theory, as their uses in the area are not immediately obvious, but Kakutani's and Brouwer's Theorems will be vital in the proof of Nash's Theorem in Section 3.4.

3.2 Brouwer's Fixed Point Theorem

This section explains the ideas behind the proof of Brouwer's Fixed Point Theorem [5], drawing on ideas from Tony Carbery's General Topology Lecture Notes

[6]. In order to state Brouwer's Fixed Point Theorem, we need to understand the closed unit ball. It is given by $\overline{B}_1^n := \{x \in \mathbb{R}^n : |x| \leq 1\}$. As the name suggests, the closed unit ball is a closed set as it contains its boundary $|x| = 1$. It is also bounded by 1, so by the Heine-Borel Theorem, the closed unit ball is compact in $\mathbb{R}^n$.

Theorem 3.2. *[6, Theorem 6.1] [Brouwer's Fixed Point Theorem] Let $f : \overline{B}_1^n \to \overline{B}_1^n$ be a continuous function. Then there exists $x \in \overline{B}_1^n$ such that $f(x) = x$.*

Brouwer's Fixed Point Theorem says that a continuous function on the closed unit ball must have a fixed point. We show that this is just a special case of Kakutani's Fixed Point Theorem. We have already remarked that the closed unit ball $\overline{B}_1^n$ is compact. Now we show that it is convex.

Consider $x, y \in \overline{B}_1^n$. By definition $|x| \leq 1$ and $|y| \leq 1$. Let $\lambda \in [0, 1]$ be a constant, so that $|1-\lambda| = 1-\lambda$ since $0 \leq \lambda \leq 1$. Then by the triangle inequality,

$$\begin{aligned} |\lambda x + (1-\lambda)y| &\leq |\lambda x| + |(1-\lambda)y| \\ &\leq \lambda|x| + (1-\lambda)|y| \\ &\leq \lambda + (1-\lambda) = 1. \end{aligned}$$

Thus $\lambda x + (1-\lambda)y \in \overline{B}_1^n$, so $\overline{B}_1^n$ is convex. We have shown that $\overline{B}_1^n$ is a compact convex subset of $\mathbb{R}^n$, which is what we require for Kakutani's Fixed Point Theorem. We now show that a continuous function $f : \overline{B}_1^n \to \overline{B}_1^n$ has a closed graph. Note that $f(x)$ is trivially convex for all $x \in \overline{B}_1^n$ because f is not set-valued, so each set $f(x)$ contains just one value.

Now we take $\{x_n\}, \{y_n\} \in \overline{B}_1^n$ such that $y_n = f(x_n)$ for all n. Suppose $x_n \to x$ and $y_n \to y$. Since f is continuous, we have that $\lim_{n\to\infty} f(x_n) = f(x)$. So,

$$f(x) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} y_n = y.$$

Thus $y = f(x)$, and f has a closed graph.

The above discussion shows that the continuous function $f : \overline{B}_1^n \to \overline{B}_1^n$ satisfies the hypotheses of Kakutani's Fixed Point Theorem, but is not set-valued, so Kakutani's Theorem is a generalisation of Brouwer's to set-valued functions.

We now seek to prove Brouwer's Fixed Point Theorem. We will give an outline omitting some of the more complex details, which can be found in [6]. Note that when $n = 1$, the ball $\overline{B}_1^n$ is just the closed interval $[-1, 1]$, so we have $f : [-1, 1] \to [-1, 1]$. Define $g(x) := f(x) - x$, which is also continuous since it is the sum of continuous functions. The Intermediate Value Theorem tells us that since g is continuous, it attains every value on the interval $[-1, 1]$. In particular, it attains 0, so $g(x) - x = 0$ for some x. Thus this x has $g(x) = x$, so x is a fixed point.

When $n \geq 2$, the proof becomes more complicated. The proof relies on several results in topology and one possibly unexpected theorem from calculus. We state three results from topology that will help us with the proof. To begin, we denote by C^2 the set of functions with continuous first and second derivative. The three results are as follows:

1. For any continuous function $f : \overline{B}_1^n \to \overline{B}_1^n$, there exists a sequence of C^2 functions $f_k : \overline{B}_1^n \to \overline{B}_1^n$ that converge uniformly to f on $\overline{B}_1^n$.

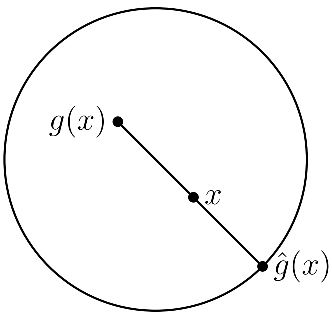
2. If every C^2 function $f : \overline{B}_1^n \to \overline{B}_1^n$ has a fixed point, then every continuous function $g : \overline{B}_1^n \to \overline{B}_1^n$ also has a fixed point.

3. There is no C^2 function $g : \overline{B}_1^n \to \partial\overline{B}_1^n$ such that $g(x) = x$ for all $x \in \partial\overline{B}_1^n$.

The first two statements are reasonably intuitive to understand, but the third requires some explanation. The set $\partial\overline{B}_1^n$ is the boundary of $\overline{B}_1^n$, given by $\partial\overline{B}_1^n = \{x \in \mathbb{R}^n : |x| = 1\}$. Thus statement 3 says that there is no twice differentiable function from the closed ball to its boundary such that the function fixes every point on the boundary.

Why are these results useful? If we assume the three statements, Brouwer's Fixed Point Theorem follows relatively easily, as we show here:

Proof of Theorem 3.2. Assume statements 1-3 above are true. Suppose for a contradiction that Brouwer's Fixed Point Theorem is false. Then there exists a continuous function $f : \overline{B}_1^n \to \overline{B}_1^n$ with no fixed point. It follows from statement 2 that there must exist a C^2 function $g : \overline{B}_1^n \to \overline{B}_1^n$ with no fixed point. For any point $x \in \overline{B}_1^n$, we may draw a line segment L_x starting from $g(x)$ passing through x, since $g(x) \neq x$. Then define by $\hat{g} : \overline{B}_1^n \to \partial\overline{B}_1^n$ the map sending $x \in \overline{B}_1^n$ to the unique intersection of L_x with the boundary $\partial\overline{B}_1^n$. Note that if $g(x)$ is on the boundary, then there are 2 intersections, and we take the one that is not $g(x)$ to be $\hat{g}(x)$. To help visualise the function $\hat{g}(x)$, we present this picture of the $n = 2$ case:



This helps us to see why intersection is unique as long as $g(x)$ is not on the boundary: since the line segment starts at $g(x)$ and passes through x, there is only one intersection because the line only leaves the circle at one point. Now if $x \in \partial\overline{B}_1^n$, the line segment L_x intersects the boundary at x. It follows that $\hat{g}(x) = x$ for all $x \in \partial\overline{B}_1^n$. By writing down an explicit formula for $\hat{g}$, which is a little tedious and we leave as an exercise, the function $\hat{g}$ is C^2. Additionally, it acts as the identity on the boundary, which contradicts statement 3. We have reached a contradiction, so Brouwer's Fixed Point Theorem must be true. □

We have shown that statements 1-3 are sufficient to prove Brouwer's Fixed Point Theorem, but we have not proved the statements themselves. Statements 1 and 2 are fairly simple to prove using the Weierstrass approximation Theorem. Statement 3 is a little more complex and (perhaps surprisingly) relies on the Divergence Theorem from calculus. The full proofs are contained in [6].

We have now proven that any continuous function on the unit closed ball has a fixed point. We also showed that this is a weaker version of Kakutani's Fixed Point Theorem, which we need in the proof of Nash's Theorem. We omit the full proof here, but explain how it works and its relation to Brouwer's Theorem in the following section.

3.3 Kakutani's Fixed Point Theorem

Kakutani's Fixed Point Theorem has a very complex proof containing ideas from topology about simplices and simplical subdivisions. Here we discuss the ideas behind Kakutani's proof [7], where Brouwer's Fixed Point Theorem plays a vital role.

We take X and $f : X \to X$ as in Kakutani's Theorem (2.6). Note that in our formulation of Brouwer's Fixed Point Theorem, we used the closed unit ball, but the theorem applies more generally to any compact, convex subset of $\mathbb{R}^n$. We wish to split X into subsets that we can then apply Brouwer's Theorem to.

For the purpose of the proof, we wish to switch from thinking about compact convex subsets of $\mathbb{R}^n$ to **simplices**. A n-dimensional simplex has $n+1$ vertices $u_0, ..., u_n$ and has the form

$$\{\theta_0 u_0 + \theta_1 u_1 + ... + \theta_n u_n : \sum_{i=0}^{n} \theta_i = 1 \text{ and } \forall i, \theta_i \geq 0\}.$$

We present the correspondence between the subset X of $\mathbb{R}^n$ and an n-dimensional simplex as follows. We have the subset $X \subseteq \mathbb{R}^n$ and the set-valued function $f : X \to X$. We are looking for a fixed point: a point $x \in X$ such that $x \in f(x)$. We can embed X into an n-dimensional simplex S with injection $i : X \to S$. An important result in mathematics says that there exists a **retraction** function $g : S \to X$ such that $g \circ i = \mathrm{id}_X$. This is a function from S that fixes elements of X. The diagram below helps to illustrate this, with the retraction shown in red.

$$\begin{tikzcd}
X \arrow[r, hook, "i"'] \arrow[d, "f"'] & S \arrow[l, bend right, "g"', red] \arrow[dl, "h"] \\
X &
\end{tikzcd}$$

On the diagram we also have the function $h : S \to X$ which we define by $h := f \circ g$. Since $X \subseteq S$, we can alternatively think of h as going from $S \to S$ instead. Since g is the retraction, it fixes every element of X, so h with domain restricted to X is f.

Suppose we know there is a fixed point $a \in S$ of $h : S \to S$. Then $a \in h(a)$. But $h(a) \subseteq X$ since $h = f \circ g$ and f maps from X to X. So $a \in X$ and $a \in h(a) = f(a)$ since h acts as f on X. Thus a fixed point of h on the simplex S is **the same as** a fixed point of f on the subset X of $\mathbb{R}^n$, so we can instead investigate fixed points on the simplex.

From the above reasoning, we now consider a simplex S of dimension n. The proof works by dividing the simplex S into smaller simplices of the same dimension n. Denote by S_i the simplex S divided into i simplices of dimension n. So we have simplices $S_0, S_1, S_2, ...$ Note that $S_0 = S$, and this sequence of simplices can continue infinitely.

For each i, we can construct a map $\phi_i : S_i \to S$. Consider a point v_j which is the vertex of a subsimplex in S_i. The map ϕ_i sends this vertex to $y_j \in f(v_j)$, which we can choose since the set $f(v_j)$ is nonempty by hypothesis. This map is continuous and by Brouwer's Theorem has a fixed point $x_i \in S$ such that $\phi_i(x_i) = x_i$.

If we collect a sequence of the fixed points for $i = 0, 1, ...$, we get a sequence $x_0, x_1, ...$ such that $x_m = \phi_m(x_m)$ for all $m = 0, 1, ...$ By doing the subdivision infinitely many times, this becomes an infinite sequence. The simplex S is closed and bounded, so has a convergent subsequence by the Bolzano-Weierstrass Theorem, which converges to some $x^* \in S$. It turns out this x^* satisfies $x^* \in f(x^*)$, so it is the fixed point we are looking for.

The proof that this is indeed the fixed point makes use of the properties of the continuity and convexity of f. The details can be found in Kakutani's original paper [7].

3.4 Nash's Theorem

Now that we have proved or outlined the proof of the three preliminary results, we can prove the crowning jewel of this section: Nash's Theorem. Because of how important it is, we state it again here.

Theorem 3.1. *[2, Proposition 33.1][Nash's Theorem] Every finite strategic game has a mixed strategy Nash Equilibrium.*

Proof. Let $G = \langle N, (A_i), (u_i) \rangle$ be a finite strategic game. We wish to show that the mixed extension of G satisfies the assumptions of Proposition 2.8. In the mixed extension $\langle N, (\Delta(A_i)), (U_i) \rangle$ of G, the set of actions of player i, is given by the set $\Delta(A_i)$, of probability distributions over A_i. Let $|A_i| = m_i \in \mathbb{N}$, then the set $\Delta(A_i)$, of actions of Player i in the mixed extension of G corresponds to the simplex

$$S_i := \left\{ (p_1, \dots, p_{m_i}) \in \mathbb{R}^{m_i} \;\middle|\; p_k \geq 0, \forall k \text{ and } \sum_{k=1}^{m_i} p_k = 1 \right\},$$

where p_k is the probability that Player i assigns to their $k-$th pure strategy. Clearly we have that S_i is nonempty; by setting $p_k = 1$ and $p_j = 0$ for $j \neq k$, we obtain a vector in S_i that corresponds to Player i's $k-$th pure strategy.

The simplex S_i is bounded, as it is contained in the unit hypercube given by the Cartesian product $[0, 1]^{m_i}$. Note that S_i is the intersection of the sets $V = \{(x_1, \dots, x_{m_i}) \in \mathbb{R}^{m_i} \mid x_k \geq 0, \forall k\}$, and $W = \{(x_1, \dots, x_{m_i}) \in \mathbb{R}^{m_i} \mid \sum_{k=1}^{m_i} x_k = 1\}$.

The set V is clearly closed. To see that W is closed, note that the mapping $f : \mathbb{R}^{m_i} \to \mathbb{R}$, given by $f(x_1, \ldots, x_{m_i}) = x_1 + \cdots + x_{m_i}$ is continuous. We therefore have that $W = f^{-1}(\{1\})$ is closed, as the set $\{1\}$ is closed, and the preimage of a closed set is closed for a continuous function f. It follows that S_i is a closed and bounded subset of $\mathbb{R}^{m_i}$, and is therefore compact by Theorem 2.9.

Now let $a, b \in S_i$, so that $a = (p_1, \ldots, p_{m_i})$ and $b = (q_1, \ldots, q_{m_i})$, where $p_k, q_k \geq 0$ for all $1 \leq k \leq m_i$ and $\sum_{k=1}^{m_i} p_k = \sum_{k=1}^{m_i} q_k = 1$. Let $\lambda \in [0, 1]$. Then we have that

$$\begin{aligned} \lambda a + (1-\lambda)b &= \lambda(p_1, \ldots, p_{m_i}) + (1-\lambda)(q_1, \ldots, q_{m_i}) \\ &= (\lambda p_1, \ldots, \lambda p_{m_i}) + ((1-\lambda)q_1, \ldots, (1-\lambda)q_{m_i}) \\ &= (q_1 + \lambda(p_1 - q_1), \ldots, q_{m_i} + \lambda(p_{m_i} - q_{m_i}))\,. \end{aligned}$$

Firstly, for any $1 \leq k \leq m_i$ we have that

$$q_k + \lambda(p_k - q_k) = \lambda p_k + (1-\lambda)q_k \geq \lambda p_k + (0)q_k = \lambda p_k \geq 0,$$

where we have used that $\lambda \in [0, 1]$ and $p_k, q_k \geq 0$. Next we have that

$$\begin{aligned} \sum_{k=1}^{m_i} (q_k + \lambda(p_k - q_k)) &= \sum_{k=1}^{m_i} ((1-\lambda)q_k + \lambda p_k) \\ &= (1-\lambda)\sum_{k=1}^{m_i} q_k + \lambda \sum_{k=1}^{m_i} p_k \\ &= (1-\lambda)(1) + \lambda(1) \\ &= 1, \end{aligned}$$

where we have used that $\sum_{k=1}^{m_i} p_k = \sum_{k=1}^{m_i} q_k = 1$. We have therefore shown that, for any $\lambda \in [0, 1]$, and $a, b \in S_i$, we have $\lambda a + (1-\lambda)b \in S_i$, and thus the set S_i is convex.
In sum then, we have shown that the set of actions of Player i in the mixed extension of G, is a nonempty, compact, convex subset of $\mathbb{R}^{m_i}$.

We next turn our attention to Player i's preference relation in the mixed extension of G, given here by U_i. Let $\alpha \in \times_{j \in N} \Delta(A_j)$ be a mixed strategy profile, β_j, γ_j be mixed strategies of Player j and $\lambda \in [0, 1]$. We have shown that the set $\Delta(A_j)$ is convex, so $\lambda\beta_j + (1-\lambda)\gamma_j$ is a mixed strategy of Player j, given by $(\lambda\beta_j + (1-\lambda)\gamma_j)(a_j) = \lambda\beta_j(a_j) + (1-\lambda)\gamma_j(a_j)$ for $a_j \in A_j$. We calculate that

$$U_i(\alpha_{-j}, \lambda\beta_j + (1-\lambda)\gamma_j) = \sum_{a \in A} \left(\prod_{\substack{k \in N \\ k \neq j}} \alpha_k(a_k) \right) (\lambda\beta_j(a_j) + (1-\lambda)\gamma_j(a_j))\, u_i(a)$$

$$= \sum_{a \in A} \left(\prod_{\substack{k \in N \\ k \neq j}} \alpha_k(a_k) \right) \lambda\beta_j(a_j) u_i(a) + \sum_{a \in A} \left(\prod_{\substack{k \in N \\ k \neq j}} \alpha_k(a_k) \right) (1-\lambda)\gamma_j(a_j) u_i(a)$$

$$= \lambda \sum_{a \in A} \left(\prod_{\substack{k \in N \\ k \neq j}} \alpha_k(a_k)\beta_j(a_j) \right) u_i(a) + (1-\lambda) \sum_{a \in A} \left(\prod_{\substack{k \in N \\ k \neq j}} \alpha_k(a_k)\gamma_j(a_j) \right) u_i(a)$$
$$= \lambda U_i(\alpha_{-j}, \beta_j) + (1-\lambda) U_i(\alpha_{-j}, \gamma_j),$$

where we note that the sums and products given here are finite, as G is a finite game. So we have that the function U_i is multilinear, in the sense that

$$U_i(\alpha_{-j}, \lambda\beta_j + (1-\lambda)\gamma_j) = \lambda U_i(\alpha_{-j}, \beta_j) + (1-\lambda) U_i(\alpha_{-j}, \gamma_j), \tag{3.4.1}$$

for any $\lambda \in [0,1]$. Let $\alpha \in \times_{j \in N} \Delta(A_j)$, and let S_α be the set

$$S_\alpha := \{\alpha_i \in \Delta(A_i) | U_i(\alpha_{-i}, \alpha_i) \geq U_i(\alpha)\}.$$

Let $\beta_i, \gamma_i \in S_\alpha$ and let $\lambda \in [0,1]$. We wish to show that $\lambda\beta_i + (1-\lambda)\gamma_i \in S_\alpha$, so that S_α is a convex set, and therefore U_i is quasi-concave in $\Delta(A_i)$. Using (3.4.1), and that $\beta_i, \gamma_i \in S_\alpha$ we have that

$$\begin{aligned} U_i(\alpha_{-i}, \lambda\beta_i + (1-\lambda)\gamma_i) &= \lambda U_i(\alpha_{-i}, \beta_i) + (1-\lambda) U_i(\alpha_{-j}, \gamma_i) \\ &\geq \lambda U_i(\alpha) + (1-\lambda) U_i(\alpha) \\ &= U_i(\alpha). \end{aligned}$$

So U_i is quasi-concave in $\Delta(A_i)$, as required. Finally, (3.4.1) tells us that the expected payoff given by U_i is linear in the probabilities assigned by each player's mixed strategy. Any multilinear function $\psi : \times_{i \in N} \mathbb{R}^{m_i} \to \mathbb{R}$ is continuous; indeed this is a special case of the property that multivariate polynomials are continuous. The restriction of such a continuous function to the strategy spaces of the players (each $\Delta(A_i)$ can be identified with a subset of $\mathbb{R}^{mi}$, as we discussed earlier) then also gives a continuous function. It therefore follows that the Utility function U_i is continuous, and so we have shown that the mixed extension of G satisfies all of the assumptions of Proposition 2.8, and thus has a Nash Equilibrium. □

So we have proved Nash's Theorem, and can now fully justify our implicit use of it in showing that the number guessing and rock-paper-scissors games of Section 2.4 have unique Nash Equilibria!

Chapter 4

The Prisoner's Dilemma

4.1 Discrete Prisoner's Dilemma

Now that we have covered the foundational aspects of Game Theory, we will return to the game that motivated our original discussion: the Prisoner's Dilemma. We begin by re-introducing the classical Prisoner's Dilemma in a more general sense, and go on to consider variations.

The Prisoner's Dilemma
We have two players A and B. In a round of the Prisoner's Dilemma, each player chooses to either stay quiet and *Cooperate* (C) with their partner, or *Defect* (D) against their partner and confess what they know. Writing (x, y) where x is Player A's payoff and y is Player B's, the payoffs are as follows:

		Player B C	Player B D
Player A	C	(R,R)	(S,T)
Player A	D	(T,S)	(P,P)

where

$$\mathbf{T} = \text{Temptation}, \quad \mathbf{R} = \text{Reward}, \quad \mathbf{P} = \text{Punishment}, \quad \mathbf{S} = \text{Sucker}.$$

These payoffs satisfy the following relation:

$$T > R > P > S.$$

If we have different conditions on T, R, P and S, for example if $R > T$, the dilemma does not arise. This is because there is no reason to ever change from the state (C, C). We need temptation to be the highest value to provide incentive for one player to betray another; betrayal and trust are the crux of the Prisoner's Dilemma.

The Prisoner's Dilemma is named after the context it was first given in, but when boiling it down to its essence we can interpret it in many ways. Consider two siblings given by their mother a plate of cookies to share. They can split them evenly (Cooperate) or choose to steal some of their sibling's share (Defect). If one sibling steals and the other shares, the stealing sibling receives more cookies. However if they both choose to steal, their squabble leads to some of the cookies falling on the floor, and they share the remaining between them. Of course, both sharing results in both siblings getting cookies, but by choosing to share you are at risk of having no cookies at all!

Proposition 4.1. *The Prisoner's Dilemma is a symmetric game.*

First we note that both players have the same set of actions $S = \{C, D\}$. If Player A chooses $x \in S$ and Player B chooses $y \in S$, we denote their respective payoffs by $u_A(x, y)$ and $u_B(x, y)$. We can see from the payoff matrix that the game is symmetric since $u_A(x, y) = u_B(y, x)$ for all $x, y \in S$. It follows that an effective strategy for Player A will be an effective strategy for Player B, and vice-versa.

But what makes an effective strategy in the Prisoner's Dilemma? Looking at the payoffs, we can deduce that Defection strictly dominates cooperation. Consider the game from Player A's perspective. If B chooses to Cooperate, A's best response is to Defect. If B chooses to Defect, A's best response is still to Defect. Since the game is symmetric, the same is true for B. Thus Defect dominates Cooperate. The result below follows.

Theorem 4.2. *(D,D) is the unique Nash Equilibrium of the Prisoner's Dilemma.*

This follows from the argument above; both players are always better off playing Defect than Cooperate, so neither player will move from the state (D,D). It is unique because if any player is not defecting, they cannot be at a Nash Equilibrium as they are better off by moving to Defect. In fact, this is a symmetric equilibrium as introduced in Section 2.3.

In our example, the siblings will both go to steal the cookies and some will fall on the floor in the commotion. Overall, this appears to be a worse situation: less cookies are being eaten! If the siblings trusted each other and cooperated, they could increase the total number of cookies eaten. But the risk that their sibling could steal makes also stealing safer.

Note that all of this discussion assumes selfishness of the players; they prioritise maximising their own payoff rather than creating the fairest outcome for everyone. In many real world applications of the Prisoner's Dilemma, players are looking out for their own interest. But what if we want to meet in the middle? Consider negotiations between strikers from a workers' union and their employer. Although the two parties have opposing demands, it is in both of their best interest to reach a compromise as soon as possible: the employer so that work can resume, and the employees so that they can be paid. We can apply our discussion of fair strategies from Chapter 2 to the Prisoner's Dilemma.

Setting $n = 2$ for the Prisoner's Dilemma, we have that a fair strategy is a strategy profile $(x, y) \in S \times S$ maximising the function f given by

$$f((x,y)) = \frac{1}{2}U((x,y)) - \sqrt{\frac{1}{2}\sum_{i=1}^{n}[u_i((x,y)) - \frac{1}{2}U((x,y))]^2}.$$

We can calculate f for each strategy profile of the Prisoner's Dilemma. We get that

$$f((C,C)) = \frac{1}{2}(2R) - \sqrt{\frac{1}{2}(0)} = R,$$

$$f((D,D)) = \frac{1}{2}(2P) - \sqrt{\frac{1}{2}(0)} = P,$$

$$f((C,D)) = f((D,C)) = \frac{1}{2}(S+T) - \sqrt{\frac{1}{2}[(S - \frac{1}{2}(S+T))^2 + (T - \frac{1}{2}(S+T))^2]}$$

$$= \frac{1}{2}(S+T) - \sqrt{\frac{(T-S)^2}{4}} = \frac{1}{2}(S+T) - \frac{1}{2}(T-S) = S.$$

From the payoff relations above, we have that the maximum value of f is $f((C,C)) = R$, so (C,C) is the fairest strategy. In fact $f((C,C)) > f((D,D)) > f((C,D)) = f((D,C))$, so the least fair strategy is when one player Cooperates and the other Defects. This follows our intuitive idea of fairness, as this state of play corresponds to one Player betraying another.

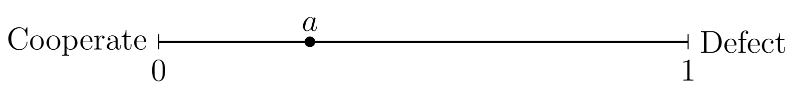
4.2 Continuous Prisoner's Dilemma

So far we have studied the Prisoner's Dilemma where players only have two choices: Cooperate or Defect. The discrete Prisoner's Dilemma assumes a binary world where the only choices are to be 100% cooperative or 100% defective. It is more realistic to view this as a spectrum, with most people lying somewhere in between. This is what motivates our exploration into a continuous Prisoner's Dilemma.

As a motivating example, think back to the example of the siblings stealing or sharing cookies, but now consider that they are arguing over a carton of their favourite milkshake. While one sibling could still drink their sibling's entire share or cooperate and steal none, they could also choose to drink some *fraction* of their sibling's share of the milkshake. Similarly to before, imagine that if they *both* try to steal, their rowdiness causes them to spill some of the milkshake, but now the *amount* they spill depends on how much they were each trying to steal. In this case, one sibling entirely cooperating but the other only stealing a small amount of milkshake results in a lesser reward compared to if they decided to steal a large amount. But if their sibling also steals to some extent, less milkshake is lost from rowdiness than if one had tried to steal everything.

If we want our model to better capture the nuance of human interaction, we need to introduce continuity. We can modify the Prisoner's Dilemma by allowing players to choose an extent to which they will Defect.

Suppose each player can choose a number $a \in [0,1]$ where $a = 0$ corresponds to full cooperation, and $a = 1$ corresponds to fully defecting. An example is shown below.



Note that here we are taking Cooperate and Defect to be *opposite* positions, so that increasing the extent you Cooperate corresponds to a direct decrease in the extent you Defect.

This continuous Prisoner's Dilemma allows for infinitely many states of the game, a huge leap from the four we had in the discrete case! A state of the game can now be seen as a pair (x, y) with $x, y \in [0, 1]$. This is reminiscent of a coordinate system, and we take the natural next step to put the Prisoner's Dilemma on a square.

Now we consider both player's choices simultaneously. We can think of the strategy space as a square with Player A's choice in the horizontal direction and Player B's in the vertical. Then, given that the choice between cooperation and defection is represented as a continuous function, we can now also consider the payoffs of Player A and Player B on this strategy space:

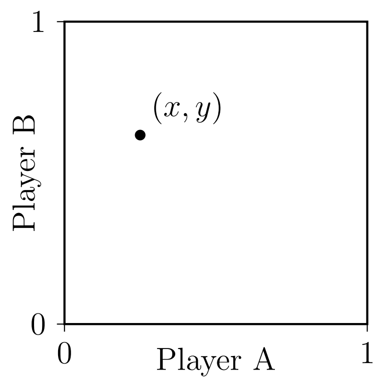


Figure 4.1: Strategy Space

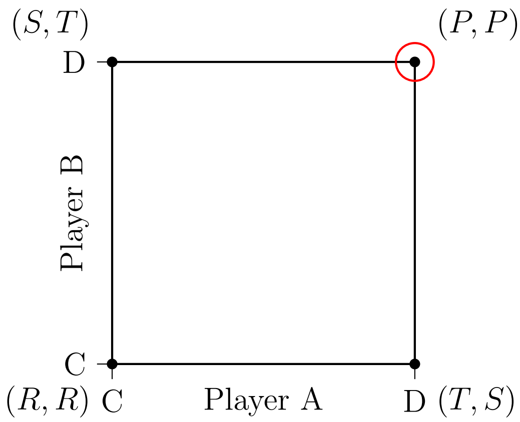


Figure 4.2: Payoff Space

The red circle in Figure 4.2 represents the Nash Equilibria in the discrete case. However, now that our game is continuous, this is not necessarily still an equilibrium. From the discrete case, we know the payoffs for the four corners of the square, but we have not defined payoffs for the values in between. We wish to know if (D, D) is still a Nash Equilibrium, and if any others arise in the continuous case.

To do this, we need to define a new Utility function that assigns a payoff to each player given $x, y \in [0, 1]$. We would like this Utility function to preserve our payoffs from the discrete case, when both players play either 0 or 1.

A Mixed Strategy Interpretation

We start the discussion with one possible interpretation: viewing continuity as a mixed strategy game. Previously we thought of a point $a \in [0, 1]$ as a position in between cooperation and defection, with the size of a representing the level of defection. Alternatively we can think of the point a as the *likelihood* of a player defecting, and $1 - a$ the likelihood of cooperating, turning this into a probabilistic problem. Now we refer to a point on the strategy space as (p_a, p_b) which represents the likelihood of Player A and Player B defecting respectively.

So $(0.5, 0.3)$ would correspond to Player A defecting with the probability of 50% and Player B defecting with the probability of 30%.

Can we find a function that will map from the space of likelihood of defection to the space of possible payoffs? YES! Now that we have defined our game in terms of probabilities, we can use expected values.

Let p_a and p_b be the probabilities of players A and B defecting, respectively. Then the four possible outcomes are:

1. Both players Cooperate with a probability of $(1-p_a)*(1-p_b)$ and payoff is (R, R).

2. Player A Cooperates and Player B defects with a probability of $(1-p_a)*(p_b)$ and payoff is (S, T).

3. Player A defects and Player B Cooperates with a probability of $(p_a)*(1-p_b)$ and payoff is (T, S).

4. Both players Defect with a probability of $(p_a)*(p_b)$ and payoff is (P, P).

Therefore, we look to find a Utility function $U : (p_a, p_b) \mapsto (\alpha, \beta)$ that takes in the probability of Defection for players A and B and outputs their expected scores. Recall that $p_a, p_b \in [0, 1]$ and $\alpha, \beta \in [S, T]$. We define our function by

$$U(p_a, p_b) = \begin{bmatrix} g(p_a, p_b) \\ h(p_a, p_b) \end{bmatrix}$$

Note that our values of α and β should be equal to the expected payoffs on the continuous payoff space given the strategies. Then

$$\alpha = g(p_a, p_b) = (1-p_a)(1-p_b)R + (1-p_a)p_bS + p_a(1-p_b)T + p_ap_bP$$

Because we want the continuous Prisoner's Dilemma to remain a symmetric game, we define β, Player B's expected payoff to be:

$$\beta = h(p_a, p_b) = (1-p_b)(1-p_a)R + (1-p_b)p_aS + p_b(1-p_a)T + p_bp_aP$$

So manipulating these gives the probabilistic Utility function for the continuous Prisoner's Dilemma:

$$U(p_a, p_b) = \begin{bmatrix} p_ap_b(P+R-T-S) + p_a(T-R) + p_b(S-R) + R \\ p_bp_a(P+R-T-S) + p_b(T-R) + p_a(S-R) + R) \end{bmatrix}, \tag{4.2.1}$$

where the payoff ordering satisfies $T > R > P > S$. We can use this Utility function to prove the following theorem.

Theorem 4.3. *With the Utility function 4.2.1, (D,D) is the unique Nash Equilibrium of the continuous Prisoner's Dilemma.*

Let us differentiate our Utility function with respect to p_a, the likelihood of Player A defecting and with respect to p_b, the likelihood of Player B defecting.

$$\frac{dU}{dp_a} = \begin{bmatrix} p_b(P+R-T-S)+T-R \\ p_b(P+R-T-S)+S-R \end{bmatrix} = \begin{bmatrix} (1-p_b)(T-R)+p_b(P-S) \\ (1-p_b)(S-R)+p_b(P-T) \end{bmatrix}$$

Then by symmetry or explicit calculation:

$$\frac{dU}{dp_b} = \begin{bmatrix} p_a(P+R-T-S)+S-R \\ p_a(P+R-T-S)+T-R \end{bmatrix} = \begin{bmatrix} (1-p_a)(S-R)+p_a(P-T) \\ (1-p_a)(T-R)+p_a(P-S) \end{bmatrix}$$

Now recall we have defined the payoff ordering as $T > R > P > S$ and that $x, y \in [0, 1]$. So we have:

$$(1-p_a), (1-p_b) \geq 0, \tag{4.2.2}$$
$$(T-R), (P-S) > 0, \tag{4.2.3}$$
$$(S-R), (P-T) < 0. \tag{4.2.4}$$

Hence we have

$$\frac{dU}{dp_a} = \begin{bmatrix} \frac{dg}{dp_a} > 0 \\ \frac{dh}{dp_a} < 0 \end{bmatrix}, \qquad \frac{dU}{dp_b} = \begin{bmatrix} \frac{dg}{dp_b} < 0 \\ \frac{dh}{dp_b} > 0 \end{bmatrix}.$$

Note that $\frac{dU}{dp_a}$ measures the change Player A and B's expected payoffs as the probability of Player A defecting increases and $\frac{dU}{dp_b}$ measures the same but as the probability of Player B defecting increases.

In both cases of the differentiated Utility function, the change in expected payoff for the player whose probability of defecting is increasing is strictly positive, whilst the other player's is strictly negative. This illustrates that for both players in this game, no matter what their opponent plays, their Utility score will always be higher if they choose to increase their probability of defection. At the limit, both players will eventually reach a stalemate at (D,D) and hence this is the unique Nash Equilibrium.

Remark. *Notice also that because the change in expected payoff for the player's opponent is strictly negative, no matter what the opponent plays, their Utility score will always be lower if the player in question chooses to increase their probability of defection.*

This is a nice result, but we stress the fact that this is *one* possible Utility function for the continuous Prisoner's Dilemma, and there are many, many more. What if we could find a **general** form for a Utility function?

The General Case

Suppose we want to investigate the class of polynomials in x and y that can be considered Utility functions for the continuous Prisoner's Dilemma. We return to thinking of $x, y \in [0, 1]$ as extents of defection, and define the set $\mathbf{n} = \{(i, j) :$

$i, j \in \mathbb{N}$ and $i + j = n\}$. We define an n-degree polynomial for Player A's Utility function as follows.

$$P_A^n(x,y) = a_1x + a_2x^2 + ... + a_nx^n + b_1y + b_2y^2 + ... + b_ny^n + \sum_{(i,j)\in\mathbf{n}} c_{i,j}x^iy^i + d.$$

To ensure the game is symmetric, we set $P_B^n(x,y) := P_A^n(y,x)$. Using the payoffs from the discrete case, we can deduce that for this to be a Utility function for the continuous Prisoner's Dilemma, the following conditions must be imposed:

1. $P_A^n(0,0) = R \implies d = R,$
2. $P_A^n(0,1) = S \implies b_1 + b_2 + ... + b_n = S - R < 0,$
3. $P_A^n(1,0) = T \implies a_1 + a_2 + ... + a_n = T - R > 0,$
4. $P_A^n(1,1) = P \implies a_1 + \cdots + a_n + b_1 + \cdots + b_n + \sum_{(i,j)\in\mathbf{n}} c_{i,j} + d = P$
$$\implies \sum_{(i,j)\in\mathbf{n}} c_{i,j} = P + R - T - S.$$

Note that the mixed strategy Utility function 4.2.1 was the case $n = 2$ with $a_2 = b_2 = 0$. We saw in this case that the derivatives of the Utility functions implied that (D,D) was the unique equilibrium. Can we deduce a more general version of this result?

We start by noticing that there can be no general polynomial for the $n = 1$ case, as the conditions can only be fulfilled if $P + R - T - S = 0$, which may be true but is not always.

An interesting investigation is to try to impose conditions on the coefficients such that (D,D) is still the unique Nash Equilibrium with this Utility function. From our mixed strategy Utility function, we saw that if the derivative of Player A's Utility function with respect to x is positive and with respect to y is negative, the result follows.

We consider when this is true in the $n = 2$ case.

Theorem 4.4. *If $n = 2$, and*

$$\max\{S - P, R - T\} < a_2 < \min\{P - S, T - R\},$$

$$\max\{S - R, P - T\} < b_2 < \min\{R - S, T - P\},$$

then (D, D) is the unique Nash Equilibrium of the continuous Prisoner's Dilemma.

Proof. From the above discussion, we can write

$$P_A(x,y) = a_1x + a_2x^2 + b_1y + b_2y^2 + (P + R - T - S)xy + R,$$

$$P_B(x,y) = P_A(y,x).$$

Differentiating this we get

$$\frac{dP_A}{dx} = a_1 + 2a_2x + b_1 + 2b_2y + (P + R - T - S)y$$

First we show that $a_1 + 2a_2x > 0$. We consider two cases.
Case one: If $a_2 < 0$, then $a_1 + 2a_2x > a_1 + 2a_2$ since $x \in [0,1]$. Then

$$a_1 + 2a_2x > a_1 + 2a_2 = (a_1 + a_2) + a_2 = T - R + a_2 > 0,$$

since $a_2 > R - T$.
Case two: If $a_2 \geq 0$, then $a_1 + 2a_2x > a_1$ since $x \in [0,1]$. Then

$$a_1 + 2a_2x > a_1 = T - R - a_2 > 0,$$

since $a_2 < T - R$.

We have covered both cases so $a_1 + 2a_2x > 0$. If $P+R-T-S > 0$ then $\frac{dP_A}{dx} > 0$. Suppose instead that $P+R-T-S < 0$. Then $(P+R-T-S)y > P+R-T-S$ since $y \in [0,1]$. We consider the same cases.
Case one: If $a_2 < 0$ then

$$a_1 + 2a_2x + (P + R - T - S)y > a_1 + 2a_2 + P + R - T - S = a_2 + P - S > 0,$$

since $a_2 > S - P$.
Case two: If $a_2 \geq 0$ then

$$a_1 + 2a_2x + (P + R - T - S)y > a_1 + P + R - T - S = (T - R - a_2) + P + R - T - S$$

$$= P - S - a_2 > 0,$$

since $a_2 < P - S$.

In all cases $\frac{dP_A}{dx} > 0$ and by symmetry also $\frac{dP_B}{dy} > 0$. We can also prove $\frac{dP_A}{dy}, \frac{dP_B}{dx} < 0$ using the same method and the conditions for b_2. We omit this calculation here, as it is very similar. Since $\frac{dP_A}{dx} > 0$, $\frac{dP_B}{dy} > 0$ and $\frac{dP_A}{dy}, \frac{dP_B}{dx} < 0$, the unique equilibrium is (D, D). $\square$

We see that this is not always the case when $n = 2$. Consider the common payoff values $(T, R, P, S) = (5, 3, 1, 0)$ and the polynomial Utility function given by

$$p_A(x, y) = 2x^2 + y + y^2 - xy + 3.$$

Note that this satisfies the necessary conditions to be a Utility function for the continuous Prisoner's Dilemma. We differentiate to get

$$\frac{dP_A}{dx} = 4x - y.$$

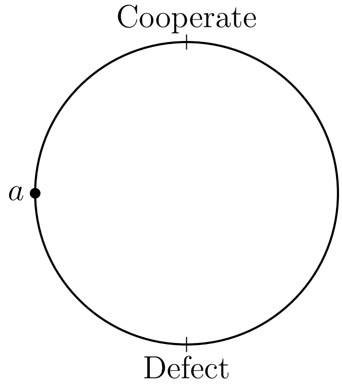
At the point (0,1), we have that $\frac{dP_A}{dx} = -1 < 0$, so the conditions on the derivative to force (D, D) as an equilibrium are not satisfied.

This is an interesting possible direction to explore after the project; are there conditions we can impose on the Utility function in general to force (D, D) to be a Nash Equilibrium? And vice versa, are there conditions we can impose to create new, possibly 'nicer' Nash Equilibria?

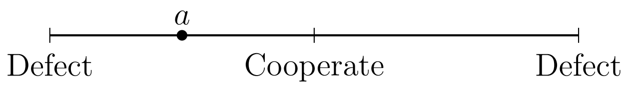
4.3 The Prisoner's Dilemma on a Torus

In the previous section, our strategy space terminated at either end of the Cooperate-Defect spectrum. Now imagine a continuous space that loops from Cooperate to Defect and then back to Cooperate again. Instead of the probability of defection being represented as an interval, this space represents the entirety of possible strategies, distributed in a cyclical manner around a continuous loop, with Cooperate and Defect the two extreme points.

Although we can picture it like this:

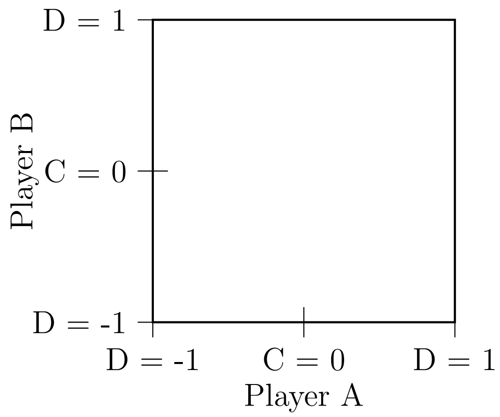


For the sake of our investigation, we can also imagine it as follows:



Here, the points at each end of the line are the same point.

We can again create a strategy space for two players. However, let us now use the interval $x, y \in [-1, 1]$. Here $x, y = -1$ and $x, y = 1$ are equivalent to complete defection and $x, y = 0$ is equivalent to complete cooperation.



Remember here that on the X-axis $D = -1$ and $D = 1$ are the same point and this is similarly the case for the Y-axis. We have combined the two continuous loops of Player A and Player B in order to create a torus to map the continuous Prisoner's Dilemma onto. Line $X = -1$ is connected to line $X = 1$ and line $Y = -1$ is connected to line $Y = 1$.

We use the following pictures to explain how the square maps to the torus. In the resulting torus, the four points circled in green become the same point, corresponding to the pure strategy (D, D).

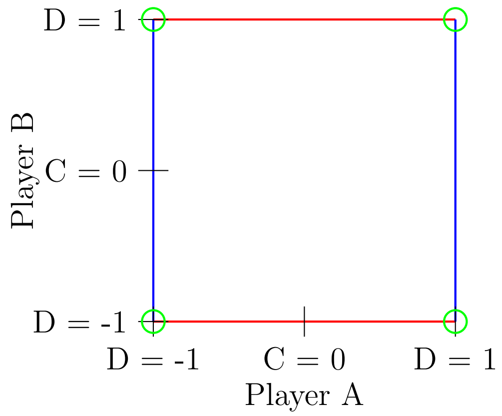


Figure 4.3

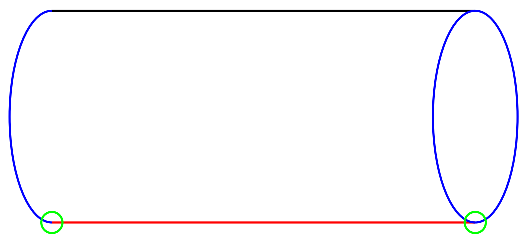
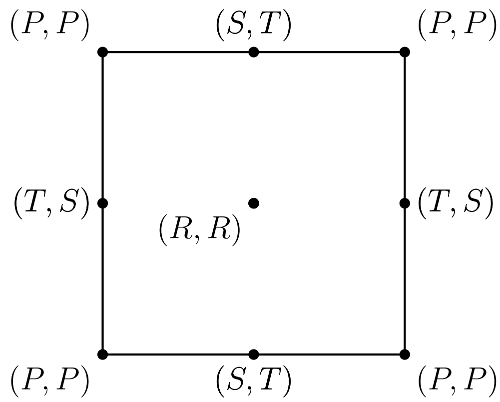


Figure 4.4

We start by joining the red sides of Figure 4.3 to get Figure 4.4. Note that the points $(-1, 1)$ and $(-1, -1)$ have merged into a single point, as have $(1, 1)$ and $(1, -1)$. Now to make the torus, the blue faces are connected to form a loop, and the final two green points merge to one.

Now the point (x, y) still corresponds to the level of defection, although now we say the level of defection by Player A and Player B is equal to $|x|$ and $|y|$ respectively.

We can now draw the associated payoff torus for Player A and Player B:



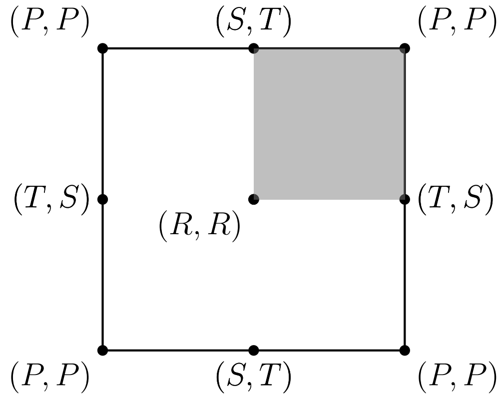
Let us consider the mixed strategy interpretation of the continuous game. Notice that the mixed strategy Utility function of our square

$$U(p_a, p_b) = \begin{bmatrix} p_a p_b(P + R - T - S) + p_a(T - R) + p_b(S - R) + R \\ p_b p_a(P + R - T - S) + p_b(T - R) + p_a(S - R) + R) \end{bmatrix}, \quad (4.3.1)$$

where:

$$T > R > P > S, \quad p_a, p_b \in [0, 1]$$

now only defines the top right corner of our torus:



Is there a way to alter this in order to find a Utility function that defines the entire space over the torus? Notice that on our graph, for any given value of p_b, the value of p_a and $-p_a$ should give the same expected payoff. Similarly, for any value of p_a, the values p_b and $-p_b$ should also give the same expected payoff. Hence, let us make use of $|p_a|$ and $|p_b|$ in our function.

This gives a Utility function for the probabilistic interpretation of the continuous Prisoner's Dilemma on a torus:

$$U(p_a, p_b) = \begin{bmatrix} |p_a||p_b|(P+R-T-S) + |p_a|(T-R) + |p_b|(S-R) + R \\ |p_b||p_a|(P+R-T-S) + |p_b|(T-R) + |p_a|(S-R) + R) \end{bmatrix}, \tag{4.3.2}$$

where:

$$T > R > P > S, \quad |p_a|, |p_b| \in [0,1]$$

However, notice that now we are using $|p_a|$ and $|p_b|$ on a torus such that -1 and 1 refer to the same point, $|p_a|$ and $|p_b|$ and by extension the Utility function are no longer differentiable at $|p_a|, |p_b| = 0, 1$. We denote the Left-Hand Derivative as **LHD** and the Right-Hand Derivative as **RHD**. Then:

$$|\mathbf{p_a}| = \mathbf{0} : \begin{cases} LHD : (-p_a)' = -1 \\ RHD : (p_a)' = 1 \end{cases} \qquad |\mathbf{p_a}| = \mathbf{1} : \begin{cases} LHD : (p_a)' = 1 \\ RHD : (-p_a)' = -1 \end{cases}$$

This will clearly prove problematic when attempting to find the Nash Equilibrium for this case. To solve this issue, we replace $|p_a|$ and $|p_b|$ with smooth approximation functions.

Definition 4.5. *A function $f_\epsilon(x)$ is a smooth approximation of a non-smooth function $f(x)$ if:*

1. *$f_\epsilon(x)$ is continuous*
2. *$f_\epsilon(x) \approx f(x)$ as $\epsilon \to 0$ for all x.*
3. *$f_\epsilon(x)$ is differentiable.*

Let us define the following functions:

$$S(x) = \frac{(x)^2}{2\epsilon} + \frac{\epsilon}{2}$$
$$P(x) = \frac{-(|x|-1)^2}{2\epsilon} + (1 - \frac{\epsilon}{2})$$

Then we can redefine $|p_a|, |p_b|$ such that we have:

$$|p_a| \approx \tilde{p_a} = \begin{cases} S(p_a) & |p_a| \leq \epsilon \\ |p_a| & \epsilon < |p_a| < 1-\epsilon \\ P(p_a) & 1-\epsilon \leq |p_a| \leq 1 \end{cases} \tag{4.3.3}$$

$$|p_b| \approx \tilde{p_b} = \begin{cases} S(p_b) & |p_b| \leq \epsilon \\ |p_b| & \epsilon < |p_b| < 1-\epsilon \\ P(p_b) & 1-\epsilon \leq |p_b| \leq 1 \end{cases} \tag{4.3.4}$$

Where ϵ is a suitably small positive constant.

Figure 4.5: $\tilde{p_a}$ and $\tilde{p_b}$ for varying values of ϵ

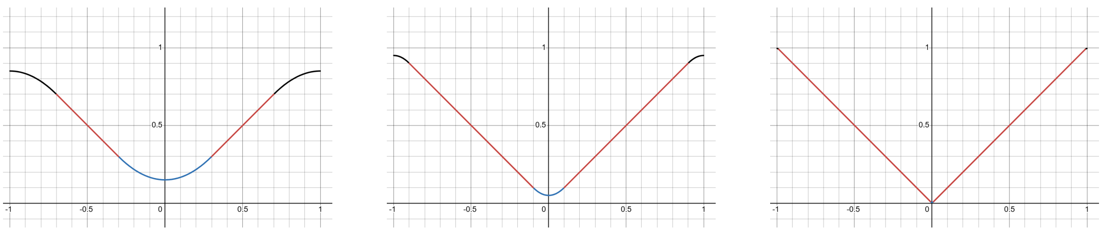


Figure 4.6: $\epsilon = 0.3$ Figure 4.7: $\epsilon = 0.1$ Figure 4.8: $\epsilon = 0.01$

Note that:

$$\tilde{p_a}, \tilde{p_b} \in [0,1] \quad \forall p_a, p_b \quad \text{respectively.} \tag{4.3.5}$$

Proposition 4.6. *$\tilde{p_a}$ is a smooth approximation of $|p_a|$.*

Proof. It is clear that $S(p_a), |p_a|$ and $P(p_a)$ are all continuous in their defined ranges. We inspect the continuity of these functions at $|p_a| = \epsilon$ and $|p_a| = 1-\epsilon$:

$$|\mathbf{p_a}| = \epsilon : \quad S(x) = \frac{|\epsilon|^2}{2\epsilon} + \frac{\epsilon}{2} = \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

$$|\mathbf{p_a}| = \mathbf{1} - \epsilon : \quad P(x) - \frac{(|1-\epsilon|-1)^2}{2\epsilon} + 1 - \frac{\epsilon}{2} = -\frac{\epsilon^2}{2\epsilon} + 1 - \frac{\epsilon}{2} = -\frac{\epsilon}{2} + 1 - \frac{\epsilon}{2} = 1 - \epsilon$$

Therefore, the function is continuous on the torus.

By definition, $\tilde{p_a} \to |p_a|$ as $\epsilon \to 0 \implies \tilde{p_a} \approx |p_a|$ as $\epsilon \to 0 \quad \forall p_a \in [-1,1]$.

We show that $\tilde{p_a}$ is differentiable in the ranges $0 < |p_a| \leq \epsilon, \quad \epsilon < |p_a| < 1-\epsilon$

and $1-\epsilon \leq |p_a| < 1$:

$$\begin{aligned}
\mathbf{0} < |\mathbf{p_a}| \leq \epsilon: &\quad S'(p_a) = \frac{2p_a}{2\epsilon} \\
\epsilon < |\mathbf{p_a}| < \mathbf{1}-\epsilon: &\quad |p_a|' = \begin{cases} 1, & \epsilon < p_a < 1-\epsilon \\ -1, & -\epsilon > p_a > -(1-\epsilon) \end{cases} \\
\mathbf{1}-\epsilon \leq |\mathbf{p_a}| < \mathbf{1}: &\quad P'(p_a) = \begin{cases} \frac{1-p_a}{\epsilon}, & 1-\epsilon \leq p_a < 1 \\ \frac{-p_a-1}{\epsilon}, & -(1-\epsilon) > p_a \geq -1 \end{cases}
\end{aligned}$$

Using these, we can inspect the LHD and RHD at the points $|p_a| = 0, \quad |p_a| = \epsilon, \quad |p_a| = 1-\epsilon$ and $|p_a| = 1$. We leave this to the reader, but summarize the results in Table 4.1.

p_a	**RHS Equation**	**LHS Equation**	**Limit of** $\tilde{p_a}'$
0	$\frac{p_a^2}{2\epsilon} + \frac{\epsilon}{2}$	$\frac{(-p_a)^2}{2\epsilon} + \frac{\epsilon}{2}$	0
ϵ	p_a	$\frac{p_a^2}{2\epsilon} + \frac{\epsilon}{2}$	1
$-\epsilon$	$\frac{(-p_a)^2}{2\epsilon} + \frac{\epsilon}{2}$	$-p_a$	-1
$1-\epsilon$	$-\frac{(p_a-1)^2}{2\epsilon} + 1 - \frac{\epsilon}{2}$	p_a	1
$-(1-\epsilon)$	$-p_a$	$-\frac{(-p_a-1)^2}{2\epsilon} + 1 - \frac{\epsilon}{2}$	-1
1	$-\frac{(-p_a-1)^2}{2\epsilon} + 1 - \frac{\epsilon}{2}$	$-\frac{(p_a-1)^2}{2\epsilon} + 1 - \frac{\epsilon}{2}$	0

Table 4.1: Comparison of the Right-Hand Side (RHS) and Left-Hand Side (LHS) Equations and their Derivatives

Hence, $\tilde{p_a}$ is differentiable everywhere in [-1,1], on the torus. So we have shown $\tilde{p_a}$ is a smooth approximation of $|p_a|$. □

Similarly, we can prove the case for p_b and rewrite our Utility function as:

$$U(p_a, p_b) = \begin{bmatrix} \tilde{p_a}\tilde{p_b}(P+R-T-S) + \tilde{p_a}(T-R) + \tilde{p_b}(S-R) + R \\ \tilde{p_b}\tilde{p_a}(P+R-T-S) + \tilde{p_b}(T-R) + \tilde{p_a}(S-R) + R) \end{bmatrix}, \tag{4.3.6}$$

where:

$$T > R > P > S, \quad |p_a|, |p_b| \in [0,1].$$

Theorem 4.7. *With the Utility function above (4.3.6), (D,D) is the unique Nash Equilibrium of the probabilistic continuous Prisoner's Dilemma on a torus.*

Proof. Differentiating 4.3.6 we get:

$$\frac{dU}{d\tilde{p_a}} = \begin{bmatrix} \tilde{p_b}(P+R-T-S) + T - R, \\ \tilde{p_b}(P+R-T-S) + S - R \end{bmatrix} = \begin{bmatrix} (1-\tilde{p_b})(T-R) + \tilde{p_b}(P-S), \\ (1-\tilde{p_b})(S-R) + \tilde{p_b}(P-T) \end{bmatrix}$$

$$\frac{dU}{d\tilde{p_b}} = \begin{bmatrix} \tilde{p_a}(P+R-T-S)+S-R, \\ \tilde{p_a}(P+R-T-S)+T-R \end{bmatrix} = \begin{bmatrix} (1-\tilde{p_a})(S-R)+\tilde{p_a}(P-S), \\ (1-\tilde{p_a})(T-R)+\tilde{p_a}(P-T) \end{bmatrix}$$

Using the fact that:

$$(1-|\tilde{p_a}|) \geq 0$$
$$(1-|\tilde{p_b}|) \geq 0$$

along with 4.2.3, 4.2.4 and 4.3.5 gives:

$$\frac{dU}{d\tilde{p_a}} = \begin{bmatrix} \frac{d\alpha}{d\tilde{p_a}} > 0 \\ \frac{d\beta}{d\tilde{p_a}} < 0 \end{bmatrix} \qquad \frac{dU}{d\tilde{p_b}} = \begin{bmatrix} \frac{d\alpha}{d\tilde{p_b}} < 0 \\ \frac{d\beta}{d\tilde{p_b}} > 0 \end{bmatrix}$$

This is the same result as the probabilistic Utility function on the square, and thus is proof that there is indeed a unique Nash Equilibrium at (D,D) with this Utility function. □

Chapter 5

Repeated Games

5.1 About Repeated Games

So far we have only considered a game where the players meet with no knowledge of how the other will play, compete once, and then never meet again. This means that each player has no prior knowledge of how their opponent will play. But what if they play against one another multiple times? This gives a player the opportunity to learn about the other's strategy and behaviour.

A possible motivation behind repeated games is it could force a **socially optimal** result. In single-round games like the discrete Prisoner's Dilemma, we often see a lack of cooperation in favour of looking out for one's own interests. In a repeated game, players build a **reputation**. Their could be an incentive to build a 'good' reputation by cooperating to gain your opponent's trust, leading to a better **total payoff**.

We see this idea fail in the Chainstore Paradox, which we explore later in the chapter, a phenomenon where repetition encourages bad behaviour in a player.

Now let us introduce the features of a repeated game.

Definition 5.1. *A finite* ***repeated game*** *is a game* $\langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$ *together with a number* $M \in \mathbb{N}$.

Here M is the predetermined number of rounds of the game. We can choose whether the players know this R or not and this choice can effect their strategy, as we see later in the Chainstore Paradox. At each round $m \in \{1, \ldots, M\}$, each player $n \in N$ has a sequence $\{a_1^n, a_2^n, ..., a_{m-1}^n\}$ of their actions in the previous rounds. This **reputation sequence** is known to all players, and can influence their decisions. This gives rise to complex strategies; a player's strategy on round k may be a function of all players reputation sequences.

Now that we have the definition of repeated games, we can apply it to the Prisoner's Dilemma. Suppose we repeat the game a finite number of times. We want to know that the dilemma still arises, that is that double cooperation yields the greatest joint payoff but a player can betray the other to gain a higher individual payoff. This condition is discussed in [8].

Suppose we play 2 rounds of the Prisoner's Dilemma. Consider if Player 1 plays $\{C, D\}$ and Player 2 plays $\{D, C\}$. Then both player's **total payoff** across the two rounds is $S + T$. We do not want this strategy to be better than both players repeatedly cooperating (this should always be the highest joint payoff strategy). So as well as our previous payoff conditions we impose the additional condition $2R > S + T$, so that the dilemma persists.

5.2 Effective Rules in the Repeated Prisoner's Dilemma

Suppose we want to come up with a set of instructions for a computer to play the repeated Prisoner's Dilemma. For example, an instruction could be *Defect twice after opponent Defects.* We call a set of these instructions a **decision rule** or just a **rule**. The rule is an algorithm for playing the game that a computer could understand and follow.

We wish to discuss what properties make a rule in the repeated Prisoner's Dilemma successful. To measure success, we can make the rules play against one another and test their payoffs over multiple rounds. In fact, this was done in 1980 by Robert Axelrod, when he ran a tournament for players to submit decision rules to play against one another.

In order to better analyse possible strategies, we introduce categories of rules, as Axelrod did in his original paper [9]. The definitions for rules in Axelrod are descriptive rather than mathematical, so we alter them here for rigour.

Definition 5.2. *A rule is **nice** if it will not be the first to Defect.*

This means that a nice rule in the Prisoner's Dilemma will start by cooperating every turn until their opponent Defects. When the opponent defects, they may Cooperate or Defect depending on their own rule. Perhaps surprisingly, in the Axelrod tournament, many of the successful rules were nice, including the winner of the tournament, suggesting this is a desirable quality in a rule. Note that a nice strategy playing against a nice strategy will lead to (C, C) being played every turn, and final payoff of nR for both players (where n is the number of rounds of the game).

We have seen that niceness is a successful quality for a rule, however we want to be careful of being *too* cooperative. A rule that is too kind can easily be taken advantage of. An example of a nice decision rule that can be taken advantage of is 'Always Cooperate', where the player Cooperates every turn no matter the response of their opponent. This rule can be exploited if their opponent, noticing the player is cooperative, Defects every turn. We can reason that although niceness is an important quality, it must be balanced with an idea of punishment if the opponent Defects. This leads us to the next type of strategy: retaliatory.

Axelrod calls this type *provokable*, and defines it to be a rule that immediately Defects after an 'uncalled for' defection from the opponent. We wish to be more precise with the definition. We also call it *retaliatory* to distinguish it from Axelrod's definition. First we make precise the notion of an 'uncalled for' defection.

Definition 5.3. *A defection is **justified** if it comes immediately after the opponent defected. A defection is **unjustified** if it is not justified.*

By this definition, defecting on the very first move is an unjustified defection. The only other type of unjustified defection is defecting immediately after an opponent cooperates.

Definition 5.4. *A rule is* ***retaliatory*** *if it immediately Defects after an unjustified defection from the opponent.*

Note that this definition does not tell us how long a retaliatory rule will Defect after an unjustified defection. A rule of defecting once after the opponent defects, and defecting forever after the opponent Defects are both retaliatory rules to different degrees.

Axelrod claimed that this is also a desirable characteristic in a rule, and was common in the top competitors of the Prisoner's Dilemma tournament. Consider a rule that is both nice and retaliatory. This rule will Cooperate until their opponent Defects, then they will Defect the following round. After this round, the strategy is essentially just retaliatory for the remaining rounds, as after the first defection niceness no longer matters.

We have now defined a rule that takes revenge on defection so as not to be taken advantage of. However long-term retaliation can force the game into a loop of mutual defection, which hurts both players. We would like to create a way to leave this loop and resume cooperation. This is where the idea of a forgiving rule is introduced. Axelrod defines forgiveness to be propensity to Cooperate in the moves after the other player has defected [9]. Again we introduce our own, more rigorous definition.

Definition 5.5. *A rule is* ***forgiving*** *if it resumes cooperation if the opponent Cooperates again.*

Note that in this definition, the rule does not have to resume cooperation instantly, just at some point. If a rule never retaliates, we say it is still forgiving.

Now that we have introduced key characteristics of rules, we can investigate how they perform against one another.

5.3 The Rules

In this section we introduce 6 rules for the repeated Prisoner's Dilemma. The effectiveness of each rule depends on the nature of interactions and the opponent's rule. The table below summarizes the key characteristics of the rules we will introduce [9]:

Strategy	**Cooperates Initially**	**Retaliates**	**Forgives**	**Vulnerable to Exploitation?**
Tit-for-Tat (TFT)	Yes	Yes	Yes	Yes (against Always Defect)
Grim Trigger (GT)	Yes	Yes	No	Yes (to noise)
Generous Tit-for-Tat (GTFT)	Yes	Yes	Yes	Somewhat
Adaptive	Yes	Yes (after n defection)	No	Somewhat
Always Defect (AD)	No	N/A	No	No (but ineffective long-term)
Always Cooperate (AC)	Yes	No	Always	Yes

Table 5.1: Comparison of Different Strategies in the Repeated Prisoner's Dilemma

Tit-for-Tat

The Tit-for-Tat (TFT) rule operates based on the principle of reciprocity. The rule starts by cooperating ("C") on the first round. In the subsequent rounds, it copies the opponent's last move. If the opponent Defects, TFT retaliates immediately but resumes cooperation once the opponent does. We see from our categories that TFT is nice, retaliatory and forgiving! It will never Defect first and does not hold grudges, resuming cooperation when the opponent also resumes. But it also punishes bad behaviour with defection. The pseudocode for the TFT rule is as follows:

Algorithm 1 Tit-for-Tat (TFT)

```
OpponentLastMove ← None                              ▷ Initial move unknown
while Game is ongoing do
    if OpponentLastMove is None then
        MyMove ← "C"                                 ▷ Start with cooperation
    else
        MyMove ← OpponentLastMove                    ▷ Copy opponent's last move
    end if
    Play(MyMove)
    OpponentLastMove ← GetOpponentMove()             ▷ Update opponent's move
end while
```

Always Defect

The Always Defect (AD) rule is a simple yet *highly exploitative* approach, as it always chooses defection ("D"). This is a neither nice nor forgiving rule which never Cooperates, regardless of the opponent's behaviour. It is the most selfish rule, consistently seeking the highest possible individual payoff per round. Whilst AD does well if the opponent Cooperates often, when playing against a retaliatory rule AD suffers from loops of mutual defection. The pseudocode for Always Defect is as follows:

Algorithm 2 Always Defect (AD)

```
while Game is ongoing do
    MyMove ← "D"                                     ▷ Always Defect
    Play(MyMove)
end while
```

Adaptive

The Adaptive strategy provides a middle ground between Grim Trigger and Tit-for-Tat, allowing limited defection tolerance before switching permanently, it is a semi-retaliatory approach that starts by cooperating and defects only after detecting an opponent's defection. Unlike Grim Trigger, which never forgives, Adaptive allows a single mistake before permanently shifting to defection. This makes it more flexible in noisy environments while still discouraging repeated defection. The pseudocode for Adaptive is as follows:

Algorithm 3 Adaptive

```
DefectionCount ← 0                                  ▷ Track opponent's "D" count
ForgivenessProbability ← 0.x                        ▷ Probability of forgiving "D"
BetrayalThreshold ← n                               ▷ Permanent defection threshold
while Game is ongoing do
    if PermanentlyDefect then
        MyMove ← "D"                                ▷ Always defect if threshold was exceeded
    else
        if DefectNextRound then
            DefectNextRound ← False                 ▷ Reset the flag after execution
            if DefectionCount > BetrayalThreshold then
                MyMove ← "D"
                PermanentlyDefect ← True            ▷ Switch to permanent defection
            else
                MyMove ← "D" with probability (1 − ForgivenessProbability)
                MyMove ← "C" otherwise
            end if
        else
            if LastOpponentMove = "D" then
                MyMove ← "D" with probability (1 − ForgivenessProbability)
                MyMove ← "C" otherwise
            else
                MyMove ← "C"                        ▷ Default to cooperation
            end if
        end if
    end if
    Play(MyMove)
    if GetOpponentMove() = "D" then
        DefectionCount ← DefectionCount + 1
        DefectNextRound ← True                      ▷ Delay defection to the next round
    end if
    LastOpponentMove ← GetOpponentMove()
end while
```

Always Cooperate

The Always Cooperate (AC) rule is a nice, non-retaliatory rule that always chooses cooperation, regardless of the opponent's actions. Unlike other strategies, AC does not track previous moves, making it a memoryless approach. AC excels if the opponent Cooperates often, but is very easily exploited by highly defecting rules like AD, since it never punishes defection. The pseudocode is as follows:

Algorithm 4 Always Cooperate (AC)

```
while Game is ongoing do
    MyMove ← "C"                                ▷ Always Cooperate
    Play(MyMove)
end while
```

Grim Trigger

The Grim Trigger (GT) rule is a retaliatory rule that enforces strict punishment after an opponent Defects. The rule starts with Cooperation ("C"). If the opponent ever Defects ("D"), the *DefectedOnce* flag is permanently set to True. Once *DefectedOnce* is True, GT never Cooperates again. This makes GT an extreme form of retaliation, as it does not allow for forgiveness or recovery from mistakes. While effective in discouraging defection, GT suffers in noisy environments where accidental defections can lead to permanent punishment, and can easily find itself in a mutual defection loop. The pseudocode is as follows:

Algorithm 5 Grim Trigger (GT)

```
DefectedOnce ← False
while Game is ongoing do
    if DefectedOnce then
        MyMove ← "D"                ▷ Always Defect after first defection
    else
        MyMove ← "C"                ▷ Cooperate initially
    end if
    Play(MyMove)
    if GetOpponentMove() = "D" then
        DefectedOnce ← True         ▷ Opponent defected, switch permanently
    end if
end while
```

Generous Tit-for-Tat

Generous Tit-for-Tat (GTFT) is a nice, forgiving rule that adapts TFT to be more merciful, only retaliating against an unjustified defection some of the time. This strategy starts with cooperation and mimics the opponent's last move, just like standard Tit-for-Tat (TFT). However, GTFT introduces less retaliation: with probability $p = 0.5$, it chooses to Cooperate even after an opponent Defects. This reduces retaliation loops and allows reconciliation, leading to more stable long-term cooperation. The pseudocode for GTFT is as follows:

Algorithm 6 Generous Tit-for-Tat (GTFT)

```
OpponentLastMove ← None
ForgivenessProbability ← 0.5
while Game is ongoing do
    if OpponentLastMove is None then
        MyMove ← "C"                                   ▷ Start with cooperation
    else
        if OpponentLastMove = "D" and Random() < ForgivenessProbability then
            MyMove ← "C"                               ▷ Occasionally forgive a defection
        else
            MyMove ← OpponentLastMove
        end if
    end if
    Play(MyMove)
    OpponentLastMove ← GetOpponentMove()
end while
```

Now we have defined these 6 rules, in the next section we analyse their efficacy against one another in a tournament.

5.4 Evaluating Strategic Performance

In order to systematically analyse the performance of different rules in the repeated Prisoner's Dilemma, we simulate repeated interactions between two players following distinct rules. Each round, the players choose to either Cooperate (C) or Defect (D) based on their predefined decision-making rules. Their decisions determine the payoffs they receive, following a standard payoff matrix:

		Player B	
		C	D
Player A	C	(3,3)	(0,5)
	D	(5,0)	(1,1)

In the tournament code, rules are paired against one another in repeated Prisoner's Dilemma games, and their cumulative payoffs are tracked. At the end of the tournament, we rank rules based on their total performance, providing insight into which behaviours lead to higher payoffs in the long run. The following pseudocode details the structure of the tournament simulation, outlining how the matchups are organized and scored.

To understand the effectiveness of different rules, it is not enough to rely on theoretical analysis alone. The complexity of the Prisoner's Dilemma stems from its long-term interactivity, where different rules may achieve high scores in the short term, but expose their strengths and weaknesses in the long-term game. Therefore, the tournament simulation is a key method to verify the effectiveness of the rule. By running Python code to play different rules against each other in a repeated Prisoner's Dilemma tournament, we can see the viability of long-term cooperation, the effects of retaliation, and whether tolerance of defection contributes to overall gain.

In order to ensure the comprehensiveness of the evaluation, we conducted the experiment in a circular tournament. This study was inspired by Axelrod's 1984 tournament, where he compared the performance of multiple rules through round-robin experiments. In the tournament, each rule plays multiple rounds against all other rules and is ranked cumulatively based on the score of each game. This approach allows us to fairly compare the performance of different rules in various interactive environments and analyse which rules are better in the long run [9].

Axelrod concluded in his tournament that reciprocity rules (such as TFT) are generally more successful than purely exploitative rules [9]. In fact, TFT was the winning rule in Axelrod's tournament. Our tournament not only validates Axelrod's research, but further tests the performance of Adaptive and Forgiving rules. The heatmaps below show the average performance of different rules in the tournament. Each cell corresponds to the total score a rule (row) earned when playing against another rule (column). The colour gradient helps visualize high-scoring matchups (red) and low-scoring ones (blue), indicating how cooperative or exploitative interactions influenced overall performance.

Table 5.1 shows the key trends in performance. When the number of turns is set to 1000, the rules with "revenge" and "cooperation" features excelled. Tit-for-Tat, GrimTrigger, and Adaptive (with $\mathbb{P}$(forgiveness) = 0.5 and betrayal threshold set to 1) received very similar scores, with GenerousTit-for-Tat (with $\mathbb{P}$(forgiveness) = 0.5) close behind, and AlwaysDefect coming in at the bottom of the overall score. This result is consistent with Axelrod's famous conclusion that in long-term games that repeat the Prisoner's Dilemma, strategies that are nice, retaliatory and forgiving have better adaptability and omit stable results. AlwaysDefect has a short-term advantage, but it is difficult to accumulate a high score against an opponent who can retaliate.

However, when the game is reduced to a short number of rounds, AlwaysDefect jumps to the top of the leaderboard. We can see this in Table 5.2, where we ran the tournament with only 10 rounds. A reason for this is that there is not enough time for retaliatory rules to punish defection in the long-term. Additionally, in the last few rounds defection becomes even stronger, because there is not enough turns left to punish. This phenomenon arises from something called **backwards induction**, which we look at further in Section 5.5. In cases where the retaliation cycle is short, the defection strategy can quickly exploit more cooperative opponents in a few rounds, thus achieving a high ranking.

Tournament Results (Round 1000):	
TitForTat:	12998 pts
GrimTrigger:	12998 pts
Adaptive:	12997 pts
GenerousTitForTat:	12510 pts
AlwaysCooperate:	12000 pts
AlwaysDefect:	10988 pts

Table 5.2: Tournament results after 10 rounds. In the short term, defecting strategies such as AlwaysDefect gain an early advantage, while cooperative strategies perform similarly.

Tournament Results (Round 10):	
AlwaysDefect:	138 pts
TitForTat:	128 pts
GrimTrigger:	128 pts
Adaptive:	127 pts
GenerousTitForTat:	125 pts
AlwaysCooperate:	120 pts

Table 5.3: Tournament results after 1000 rounds. Over the long term, cooperative strategies like TitForTat and GrimTrigger dominate, while AlwaysDefect performs poorly due to lack of sustained cooperation.

The contrast between these two results highlights an important feature of the repeated Prisoner's Dilemma: the higher the number of turns, the more the Cooperate-punish mechanism moderates, making rules like Tit-for-Tat, which are both friendly and punitive, outperform rules of pure betrayal. When the game turns are few, the threat of retalliation cannot significantly offset the benefit of betrayal, leaving AlwaysDefect room to dominate. It can be said that Axelrod's core conclusion applies mainly to scenarios with more rounds, higher value of future interactions, or tournaments with uncertain endings. On a very limited number of occasions, a rule that consistently Defects is more likely to profit from last minute exploitation, thus leapfrogging cooperative and trigger strategies on the leaderboard.

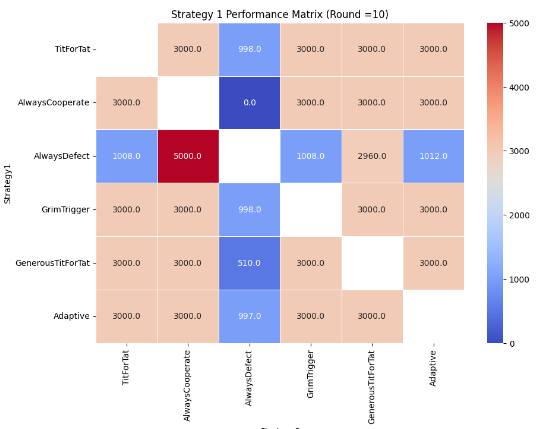


Figure 5.1: Strategy Performance after 1000 Rounds. TFT achieves the highest score, benefiting from long-term cooperation, while defecting strategies perform worse over time.

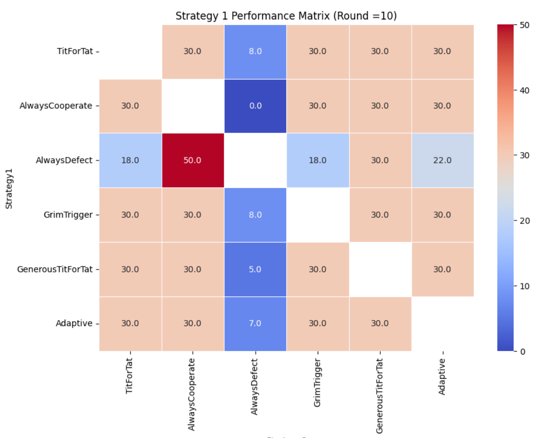


Figure 5.2: Strategy Performance after 10 Rounds. Short-term interactions play a key role, with defecting strategies like AlwaysDefect showing an early advantage.

Figure 5.3: Comparison of Strategy Performance Heatmaps

5.5 The Chainstore Paradox

The Chainstore Paradox is a Game Theory problem proposed by German economist Reinhard Selten in 1978 [10]. The paradox examines how a large monopoly responds to the entry behaviour of potential competitors.

The classic setup involves a finite number N of markets (for example, 20 towns in Selten's original example). In each market, competitors decide whether to enter the market in order. Every competitor knows the previous historical data. The payoffs are as follows:

		Challenger	
		Stay Out	**Enter**
Chainstore	**Accomm**	(5,1)	(2,2)
	Fight	(5,1)	(0,0)

There are two strategies for the chainstore:

- **Accommodate**: The chainstore accepts the challenger's entry and coexists peacefully.
- **Fight**: The chainstore takes aggressive competition measures (such as large price cuts, increased marketing investment, etc.) to pressure the challenger out of the market.

There are two strategies for the challenger:

- **Enter**: The challenger enters the market.
- **Stay Out**: The challenger does not enter the market and chooses to develop in other markets.

In a one-round game, this is very simple, with two Nash Equilibria at (accommodate, in) and (fight, out). When this is extended to a repeated game, it becomes significantly more complex, as the chainstore builds a reputation that influences the choices of future challengers.

We see from the payoff table that it is always best for the chainstore if a challenger chooses out, as it prefers to have no competition. Thus a potential strategy for the chainstore would be to deter future shops from entering by playing aggressive early on. If the chainstore starts by fighting all competitors, in subsequent turns each challenger might choose out to get the safe payoff of 1, rather than risking the lowest payoff of 0.

While intuition suggests that playing aggressively early on would benefit the chainstore, backwards induction tells us this is not the case.

Backwards Induction

Consider the final round of the game, where the last challenger chooses in or out. If the chainstore knows this is the final round, it has no reason to intimidate future competitors, and should focus on maximising their payoff on the last turn. They will *always* get a higher payoff by choosing to accommodate rather than fight, so in the last turn will accommodate. Knowing this, the challenger should play in to get their best payoff of 2. Now we consider the second-to-last turn. No matter the outcome of this turn, by the above logic the last turn will always be (accommodate, in), so again the chainstore does not need to deter future players. Thus the chainstore should also choose to accommodate in the second-to-last turn.

Inductively, the argument continues, finally arriving at the conclusion that the chainstore should play cooperatively every turn to maximise their payoff. It does not matter if there are 20 competitors or 1 million, as long as the chainstore knows when the game ends, they should accommodate the competitor from the beginning by backwards induction. However, the Chainstore Paradox says that this is not the best outcome for the chainstore [10]. The paradox arises when we instead consider **deterrence theory**.

Deterrence Theory

Suppose the chainstore tries out another strategy where it attempts to deter competitors early on by playing aggressive. After hearing of the backwards induction strategy, the chainstore agrees to accommodate the competitor in the last 3 rounds. In the first $n-3$ rounds (where n is the total number of rounds) the chainstore fights the challengers, in hope that this will deter future competitors from entering the market. Seeing repeated fights from the chainstore, it is likely that many of the challengers will choose the safe option of not entering the market. The backwards induction theory would have awarded the chainstore a total payoff of $2n$. Playing this strategy, the last 3 rounds will yield the chainstore a payoff of at least 6. Suppose the deterrence works, and in the first $n-3$ rounds, no challenger enters the market. Then the chainstore gains total payoff at least $5(n-3)+6 = 5n-9$. Note that this is better than the backwards induction payoff of $2n$ as long as $n > 3$ (which it usually is when this game is presented). In fact, even if only k are deterred where $k > \frac{1}{5}(2n-6)$, this is still a better payoff than by induction theory, since the chainstore gains a total payoff of at least $5k+6 > 2n$. Applying this to the classic example of $n = 20$, playing the deterrence strategy is better than playing the induction strategy as long as at least 7 competitors are deterred over 17 rounds. This seems like a pretty reasonable assumption.

In conclusion, the paradox arises because backwards induction dictates that the chainstore should always Cooperate, whilst a higher payoff can be yielded by playing deterrence.

There are a handful of ways this paradox can be resolved. One way is by introducing imperfect information; if the chainstore does not know how many rounds of the game there are, backwards induction becomes impossible [11]. The chainstore will be more likely to fight in the final rounds, because they do not know if there are future competitors to deter. We could even extend this to infinite rounds, which also breaks backwards induction, see [12].

Here we introduce non-rational strategies into the chainstore game to show that a higher payoff can be achieved by the chainstore.

We test the competing ideas in the paradox using Python. We set up a variety of "chainstore strategies" (AlwaysFight, AlwaysAccommodate, GrimTriggerChain, TitForTatChain) and "challenger strategies" (AlwaysEnter, GrimTrigger, TitForTat). Over the course of 20 rounds, the two sides play one another and we get the following payoff heatmap for the chainstore shown below in Figure 5.4.

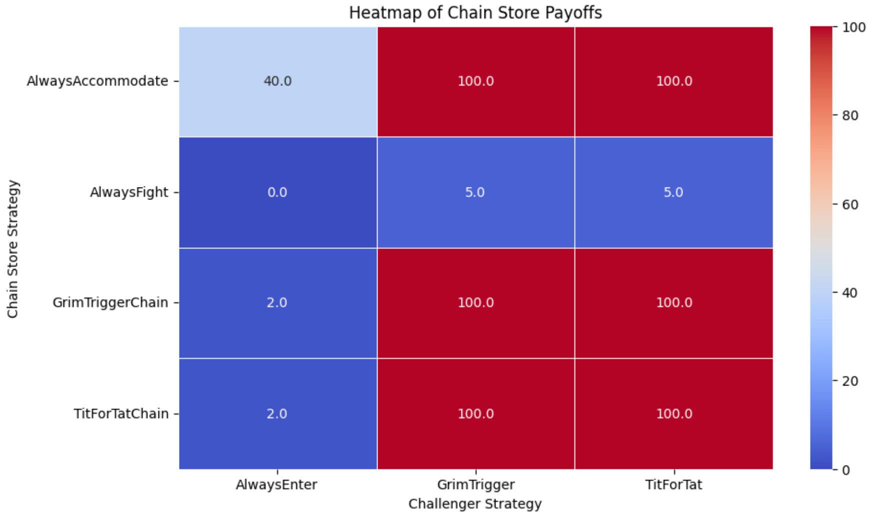


Figure 5.4: Chainstore Heatmap

This result contradicts the idea that backwards induction is the best strategy, with multiple strategies scoring higher. We can see that for some challenger strategies (notably GrimTrigger, TitForTat), chainstores can obtain high monopoly profits (such as 100) through tough strategies (such as GrimTriggerChain and TitForTatChain). This is because we introduce trigger (e.g. GrimTrigger/TitForTat) non-strictly rational strategies. This model is perhaps more accurate to the real world situation, where incomplete information and behavioural pnnces come into play. This suggests that the Chainstore Paradox does not tend to arise in the real world, when there are many strategies that outperform backwards induction [13].

Chapter 6

Repeated Games on a Torus

6.1 Setup

We conclude our report by combining the application of Game Theory on strategy spaces with the process of repeated games to explore repeated discrete games on a torus. This framework has important applications in evolutionary biology, where it models competition between traits over time, to see which traits persist [14]. Note that here the payoffs are discrete and the torus is now interpreted as the *setting* of the game.

We set up the game by creating an $n \times n$ grid (in our case 20×20). We connect the horizontal and vertical edges of the grid so that we have a torus containing 400 *cells*, each containing a player, denoted Player[i,j]. Each player is randomly assigned an attribute, **C** or **D**, corresponding to whether they **Cooperate** or **Defect** in the Prisoner's Dilemma.

In round one, each player plays the Prisoner's Dilemma game against their *neighbours* who we define as the players directly adjacent to them - above, below, left and right. The payoff from each game is then summed to determine the player's total score for that round.

Each player's score will then be compared to the scores of its neighbours. Whichever player has the highest score will take over the cell for the the next round. If their is a tie between the Player and one of its neighbours, the player will keep their cell. If there is a tie between multiple neighbours, one of the neighbours will be chosen at random to take over the cell. This process simulates the evolutionary dynamics of the chosen traits over time.

We begin with the following payoff structure for the Prisoner's Dilemma:

		Player B	
		C	D
Player A	C	(3,3)	(0,5)
	D	(5,0)	(1,1)

where $T = 5, \quad R = 3, \quad P = 1$ and $S = 0$.

We assume a random, equal distribution of Cooperators and Defectors and visualise the evolution of strategies on the torus over time:

Here green cells represent Cooperators and red cells represent Defectors. We see from Figure 6.1 that the dominance of the Defectors is clear, as they take

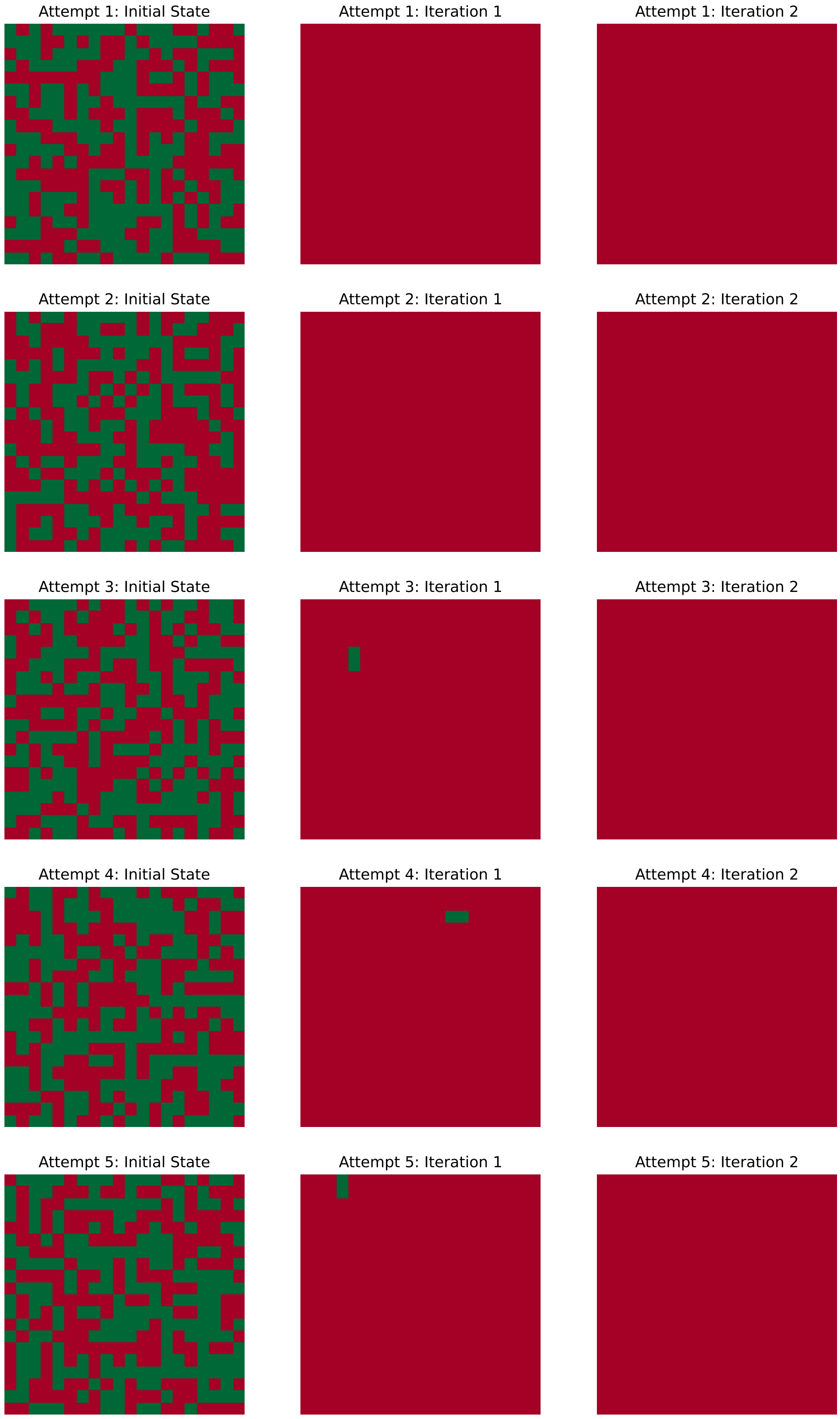


Figure 6.1: Iterations of Classical Discrete Prisoner's Dilemma on a Torus 53

over the entire torus within 2 iterations. This aligns with the Nash Equilibrium, which illustrates that selfish players always have a higher payoff by defecting than cooperating. As each player engages in 4 games of the Prisoner's Dilemma in each round, Defectors are expected to dominate due to their higher payoffs.

We can manipulate two key factors to further explore this setup:

1. The initial distribution of players,

2. The payoff values T,R,P,S.

6.2 Changing distributions

Clearly, skewing the distribution in favour of Defectors wouldn't change much, but how about skewing it in favour of Cooperators?

Below illustrates two games in the 20x20 grid where 95% of the initial positions are filled by Cooperators and only 5% by Defectors.

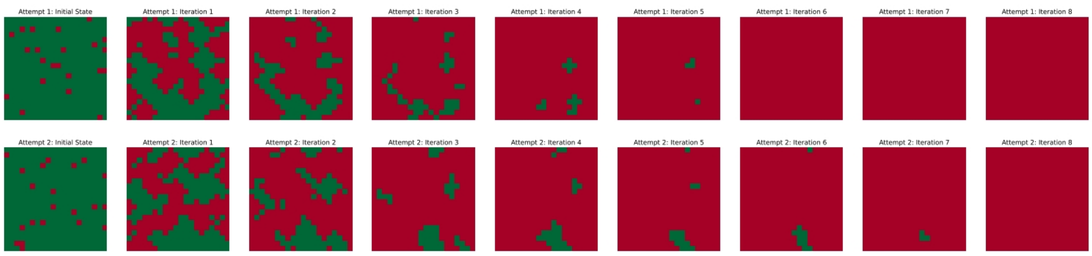


Figure 6.2: Prisoner's Dilemma on a Torus with Skewed Distributions

Incredibly, Figure 6.2 demonstrates that within 8 iterations the torus is entirely dominated by Defectors.

In our current set up if a Cooperator is a neighbour with a Defector, then that Defector needs only one more cooperative neighbour to take over both cooperative cells.

$$3R + S = 3(3) + 0 = 9 < 2T + 2P = 2(5) + 2(1) = 12.$$

The only exception is if the Cooperator has a cooperative neighbour who is themselves entirely surrounded by Cooperators. Then there is a random chance that a Defector or Cooperator will take over the cell:

$$4R = 4(3) = 12 = 2(5) + 2(1) = 2T + 2P.$$

Note that green refers to Cooperator payoffs and red refers to Defector payoffs.

6.3 Changing Payoffs

Altering payoff values reveals interesting patterns that take place. In Figure 6.3 we return to an even distribution of Cooperators and Defectors, but modify the payoffs:

$$T = 5, \quad R = 3, \quad P = 0.5, \quad S = 0.$$

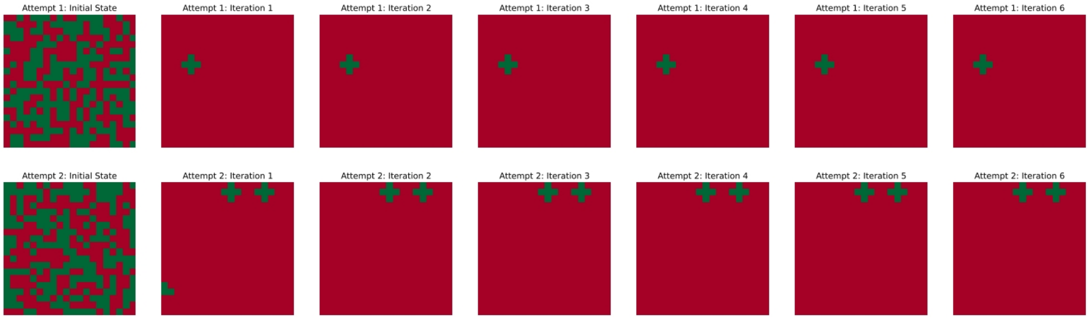


Figure 6.3: Iterations of Discrete Prisoner's Dilemma on a Torus with Altered Payoffs

This small change brings rise to cooperative structures which appear to survive indefinitely against the Defectors surrounding it. We investigate this further.

Definition 6.1. *The **faces** of a group of Cooperators are the outermost set of Cooperators in the North, East, South and West directions.*

Definition 6.2. *A Basic Cooperative Compound (**BCC**) is a group of Cooperators on the $n \times n$ grid of a torus which satisfies the following:*

1. *Only one protruding North, East, South and West face.*

2. *If a quadrilateral structure, its dimensions are $x \times y$ or $y \times x$ where $x \in [2, n], y \in [3, n]$, and $x, y \in \mathbb{N}^*$.*

3. *If not a quadrilateral structure, its faces are connected by a staircase pattern where necessary, consisting of alternating horizontal and vertical segments, each of length 1.*

   - *4 Staircases: faces are of length $a, b, c, d \in \mathbb{N}^*$ where $a, b, c, d \in [1, n-2]$ a, b, c, d.*
   - *2 Staircases: Same as 4 Staircases but one face must be of at least length* 5.
   - *1 Staircase: Same as 4 staircases but one face must be of at least length* 2 *and another of at least length* 4.

Definition 6.3. *A Stable Basic Cooperative Compound (**SBCC**) is a BCC where:*

1. *All protruding faces are of length 1 or 2.*

2. *The four protruding faces are all connected by a staircase patten with horizontal and vertical components of length 1.*

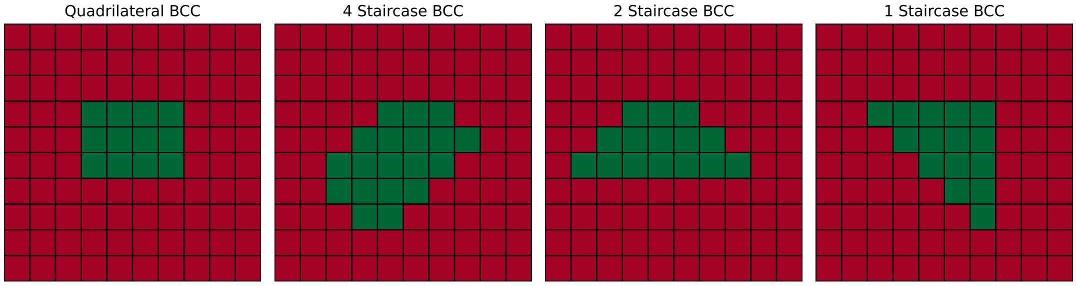


Figure 6.4: Examples of BCCs

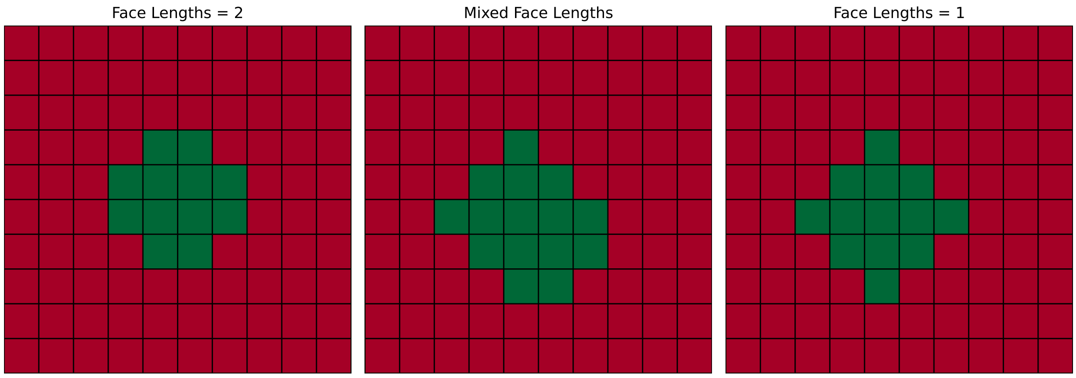


Figure 6.5: Examples of SBCCs

Let us explore how these compounds act in a torus of Defectors. We begin by defining a payoff structure μ, which satisfies:

$$4T > 3T + P > 4R > 2T + 2P > 3R + S > T + 3P > 2R + 2S > R + 3S > 4P > 4S.$$

For the following Propositions we discuss heuristic proofs.

Proposition 6.4. *For the payoff structure μ, every SBCC surrounded by at least 2 layers of Defectors will remain the same shape infinitely.*

Discussion: By the way which we have defined an SBCC, every Cooperator that has a defective neighbour will have a cooperative neighbour who themselves has 4 cooperative neighbours. Therefore, each Cooperator will be *defended* from the Defector by a score of $4R$. By μ, a Defector can then only take control of that cell if it has at least 3 cooperative neighbours ($3T + P > 4R$). However, because there are at least two layers of Defectors, each Defector will always have at least 2 defective neighbours. Additionally, each Cooperator with a defective neighbour will have 2 cooperative and 2 defective neighbours. Hence, they will have a score of $2R + 2S < T + 3P$ and so will never take over a Defector. Thus, the shape will remain constant indefinitely.

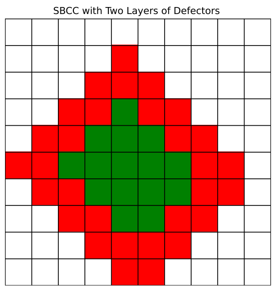


Figure 6.6: SBCC surrounded by two layers of defectors

Definition 6.5. *On the outside layer of a BCC we define* ***face pieces*** *to be any Cooperator between the first and last Cooperator of a face. We call the other Cooperators* ***corner pieces****.*

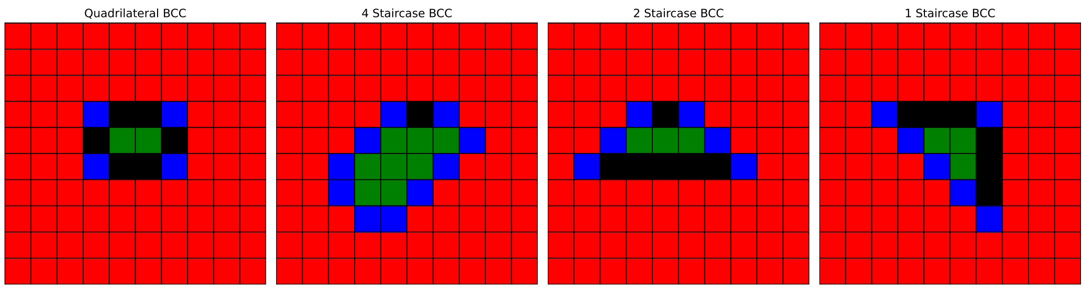


Figure 6.7: Illustration of face pieces (black) and corner pieces (blue) on our BCC examples.

Proposition 6.6. *For the payoff structure μ, every BCC in an infinite grid of Defectors will eventually become an SBCC.*

Discussion: Any face piece of our BCC will have 3 cooperative neighbours and thus a score of $3R + S$. Meanwhile, its defective neighbour will have 3 defective neighbours giving it a score of $T + 3P$. As $3R + S > T + 3P$, the face piece will take over its defective neighbour for the following round. Almost every corner piece will have a cooperative neighbour who themselves has 4 cooperative neighbours. As the BCC is surrounded by at least two layers of Defectors, we follow the same logic as in the last proof to conclude neither corner piece nor neighbouring Defector will take each other over. One exception is a corner piece at the meeting of two faces. They neighbour two Cooperators, each with 3 co-operative neighbours giving them a score of $2R + 2S$ and they are defended by a score of $3R + S$. Their neighbouring Defectors have only one cooperative neighbour and thus a score of $T + 3P$. Since $3R + S > T + 3P > 2R + 2S$, nobody takes over any squares. The other exception is a corner piece with two faces of length one and one face of length $n \geq 4$. As illustrated in the example below, in the first iteration it will be defended by a Cooperator who is neighbours with three Cooperators ($3R + S$), and thus taken over by a Defector with a score of $2T + 2P > 3R + S$. However, in the next iteration, the shape will have a flat face and the Cooperator that was initially defending will be a face piece. Thus it will change back into a Cooperator and remain that way. This means in the long term, no Cooperators become Defectors. This process will keep iterating until there are no face pieces remaining, at which point we will have an SBCC. See Figure 6.8.

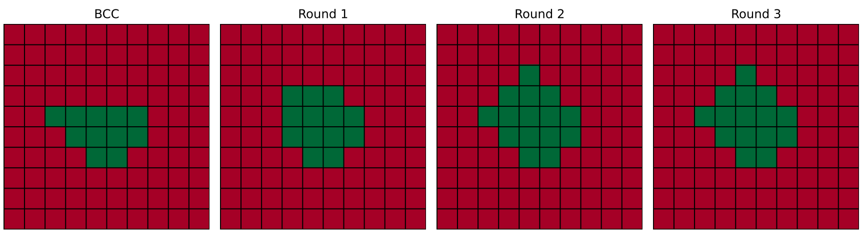


Figure 6.8: Development of a BCC to an SBCC

Corollary 6.7. *For the payoff structure μ, every BCC in an $n \times n$ torus, such that it is surrounded by Defectors up to and including at least two layers surrounding its corresponding SBCC, will survive infinitely.*

Discussion: We have heuristically shown that every BCC in an infinite grid of defectors will eventually become an SBCC and that every SBCC that is surrounded by at least two layers of defectors on a torus will remain the same shape infinitely. We have also noted that over many iterations, no Cooperators permanently become Defectors. This means the SBCC will always be larger than the BCC and thus Proposition 6.6 holds as long as it is surrounded by Defectors up to and including at least two layers surrounding its corresponding SBCC. From here, the Corollary follows.

We stress that this is not an exhaustive list of groups of cooperators that will survive infinitely, just an example of some.

6.4 Taking this further

There are many directions we could further research this specific game setup. We could investigate whether any groups of Cooperators survive infinitely for different payoff structures, and how changing the initial distribution of players might affect this.

We could also attempt to introduce continuity into the game. What if we used the ideas from Section 4.2 and had a continuous score of Defection for each player? We could then calculate their expected return for each game and use this. Perhaps we would change the game so that if one player loses to a neighbour, the new level of defection is halfway between the two original levels of defection.

Or perhaps we could investigate how the evolution of strategies changes when instead of Selfish players, we use Altruistic players.

For full details about the Python implementation, including the rule codes, tournament simulation, data visualization, etc., please refer to the GitHub repository at: `https://github.com/KaiyiQian/Game-Theory-Math-Project.git`.

This repository contains the complete code for Chapters 5 and 6, enabling replication of the tournament results, modifications to rule behaviour, and further experimentation with different parameters.

Bibliography

[1] Ken Binmore. *Playing For Real: A Text On Game Theory*. Oxford Univ. Press, 2013.

[2] Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory*. MIT Press, Cambridge, Mass., 1994.

[3] Kaushik Basu. The traveler's dilemma. *Scientific American*, 296:90–95, 06 2007.

[4] J. F. Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36:48–49, 01 1950.

[5] L.E.J. Brouwer. Über abbildung von mannigfaltigkeiten. *Mathematische Annalen*, 71(1):97–115, 1911.

[6] Tony Carbery. General topology, 2024. University of Edinburgh, Course Notes.

[7] Shizuo Kakutani. A generalization of brouwer's fixed point theorem, 1941.

[8] Tom Verhoeff. A continuous version of the prisoner's dilemma. Technical report, Technische Universiteit Eindhoven, 1993.

[9] Robert Axelrod. Effective choice in the prisoner's dilemma. *Journal of Conflict Resolution*, 24(1):3–25, 1980.

[10] Reinhard Selten. The chain store paradox. *Theory and Decision*, 9:127–159, 1978.

[11] David M. Kreps and Robert B. Wilson. Reputation and imperfect information. *Journal of Economic Theory*, 27:253–279, 1982.

[12] Drew Fudenberg and David K. Levine. Reputation and equilibrium selection in games with a patient player. *Econometrica*, 57(4):759–778, 1989.

[13] Robert J. Aumann. Backward induction and common knowledge of rationality. *Games and Economic Behavior*, 8(1):6–19, 1995.

[14] Martin Nowak and Robert May. Evolutionary games and spatial chaos. *Nature*, 359:826–829, 10 1992.