INVASION DYNAMICS IN EVOLUTIONARY GAME THEORY

AGASTYA MEHTA

Abstract. This paper investigates the emergence and persistence of cooperation in infinitely repeated Prisoner's Dilemma games through the lens of evolutionary game theory. It formalises the conditions under which cooperation can invade and resist invasion in environments where defection is strictly dominant. The study quantifies the performance of deterministic and stochastic strategies, in both mistake-free conditions and mistake-prone environments, using Markov chains to compute long-run probability distributions and expected payoffs. The paper explains why and how strategies like Tit-for-Tat can initiate cooperation, and why Win-Stay-Lose-Shift is evolutionarily stable in symmetric environments.

Introduction

The *Prisoner's Dilemma*, the quintessential introductory game in game theory, presents a striking implication: rational players would intentionally choose not to cooperate, despite mutual cooperation leaving both better off.

Yet, countless real-world interactions - natural, social, and economic - can be deconstructed into mechanisms similar to *Prisoner's Dilemma* games. From microbial symbiosis to oligopolies, we can empirically infer that not only has cooperative behaviour evolved, but it has also become evolutionarily entrenched. This presents a fundamental paradox: the emergence and persistence of cooperation within environments where defection is strictly dominant over cooperation. Resolving this paradox has profound implications on our understanding of the dynamics of trust, competition, and cooperation. Thus, this paper aims to formally examine the conditions under which cooperation emerges and persists in such environments and the replicator dynamics associated with this process. Thereby, it aims to systematically integrate frameworks to prepare a rigorous explanation behind the aforementioned examination.

This paper is most suited for curious mathematicians and social scientists with a keen interest in game theory and behavioural tendencies. Familiarity with linear algebra and basic calculus would help the reader understand the underlying math behind the reasoning, however, the paper is structured such that the broader intuition remains accessible to a passionate reader without a rigorous math background. Reading the examples will be helpful in this regard. Those intrigued by the question the paper attempts to address will, it is hoped, find this to be an engaging exploration.

1. Replicator Equation

In the context of evolutionary dynamics, this equation models exponential growth by measuring time as a continuous quantity. It assumes that cell division occurs in an exponentially distributed time period. Let $x(t)$ be the abundance of a strategy at a time t, and the strategy grows at a continuous rate r. Then, its differential equation is given by:

Date: July 12, 2025.

$$\frac{dx}{dt} = rx. \tag{1.1}$$

The solution for $x(t)$ by separating variables is as follows:

$$\int \frac{dx}{x} = \int r\,dt$$
$$\log x = rt + C$$
$$x(t) = e^{rt+C}$$

Let $x(0) = x_0$, thus:

$$x_0 = e^C$$

$$x(t) = x_0 e^{rt}. \tag{1.2}$$

Upon incorporating death at a continuous rate d, the differential equation becomes:

$$\dot{x} = \frac{dx}{dt} = (r-d)x \tag{1.3}$$

where the effective growth rate is $(r-d)$.

This equation highlights that for $r > d$, i.e. $\frac{r}{d} > 1$, the population grows indefinitely. For $r < d$, i.e. $\frac{r}{d} < 1$, the population converges to 0. If $r = d$, i.e. $\frac{r}{d} = 1$, the population remains constant, yet this equilibrium is unstable. Any uneven perturbation in r or d will either lead to exponential expansion, for:

$$\dot{x} = \lim_{(r-d)\to 0^+} (r-d)\,x$$

or decay, for:

$$\dot{x} = \lim_{(r-d)\to 0^-} (r-d)\,x$$

as $\dot{x} = 0$ precisely at $\frac{r}{d} = 1$.

The study of evolutionary game theory emphasises the study of competitive advantages of strategies represented by relative population dynamics, rather than absolute population growth. This involves a normalisation exercise using the logistic equation:

$$\dot{x} = rx(1 - \frac{x}{K}) \tag{1.4}$$

where K is the maximum carrying capacity of the environment. The logistic equation limits the population growth to K, with the rate of growth declining as $x \to K$. Normalisation implies $K = 1$, whereby:

$$\dot{x} = rx(1-x). \tag{1.5}$$

The "relative abundance" of a strategy is interpreted as its frequency and its "growth" as its fitness. Henceforth, the replicator equation will imply this normalised equation that models the change in the relative proportion ($\dot{x}$) of a strategy with respect to time, with x denoting its relative abundance.

The motivation behind normalisation is that it turns absolute growth rates into relative competition dynamics and connects frequencies to the probability of encountering each strategy in the population, linking directly to the computation of expected payoffs as described in the following chapters. To intuitively understand the motive:

Example 1.1. Consider a competitive market and equate payoffs to market share gain. The absolute sales growth of all the companies is less relevant than their relative growth compared to the industry.

1.1. **Natural Selection.** For an environment consisting only of species A and B having relative proportions (frequencies) of x and y, respectively:

$$x + y = 1. \tag{1.6}$$

Thus, the average fitness of the population is the following:

$$\phi = ax + by \tag{1.7}$$

where a and b represent the fitness of x and y, respectively. The replicator equations for the growth in the frequencies of A and B, respectively, are the following:

$$\dot{x} = x(a - \phi) \tag{1.8}$$

$$\dot{y} = y(b - \phi). \tag{1.9}$$

Thus, for the frequency of A to grow, its fitness must exceed the average fitness ϕ of the population, i.e. $a > \phi$, or else its relative abundance decays. Similarly, for B to grow, $b > \phi$.

Selection can be described among n different types. For $i = 1, \dots, n$, let $x_i \in [0, 1]$ denote the frequency of strategy i. Thus, the relative abundance of the population can be described by the probability vector $\vec{x} = (x_1, x_2, \dots, x_n)$, for

$$\sum_{i=1}^{n} x_i = 1.$$

Definition 1.1. ***Replicator Equation (Normalised):*** The *replicator equation* describes how the relative abundance (frequency) of strategies in a population changes over time, depending on the difference between a strategy's fitness and the average fitness of the population.[5]

Formally, for a population comprising n strategies, let x_i be the frequency of strategy i, f_i its fitness, and ϕ the average fitness of the population given by:

$$\phi = \sum_{j=1}^{n} x_j f_j$$

Then, the replicator equation is:

$$\dot{x}_i = x_i \left(f_i - \phi \right)$$

The interpretation of fitness as payoffs bridges evolutionary dynamics to evolutionary game theory. Rephrasing natural selection in game-theoretic terms, strategies earning a higher payoff than the average payoff of the population tend to increase their frequency, while those earning below-average payoffs get outcompeted.

Thus, "the key idea of evolutionary game theory is to consider a population consisting of A and B players and to equate payoff with fitness".[5, p. 49]

We require the framework of evolutionarily stable strategies to understand the mechanics of evolutionary stability under replicator dynamics.

2. NASH EQUILIBRIUM AND EVOLUTIONARILY STABLE STRATEGIES

2.1. **Nash equilibrium.**

Definition 2.1. ***Nash Equilibrium:*** A *Nash equilibrium* is a strategy profile such that no player can increase their payoff by unilaterally deviating from their strategy, given the strategies chosen by the other players.[6] Formally, for a strategy profile $s^* = (s_1^*, s_2^*, \ldots, s_n^*)$, a Nash equilibrium holds if for every player i,

$$u_i\left(s_i^*, s_{-i}^*\right) \geq u_i\left(s_i, s_{-i}^*\right) \qquad \forall\, s_i \in S_i$$

where u_i is the payoff function for player i, s_{-i}^* denotes the strategies of all other players, and S_i is the set of available strategies for player i.

Further, a strategy profile is considered a mixed-strategy Nash equilibrium when the fitness of all strategies is equal. For $i = 1, \ldots, n$,

$$f_1 = f_2 = f_3 = \cdots = f_n$$

where f_i represents the fitness of the strategy i.

It must be noted that there may be more than one Nash equilibrium in a game. Considering the payoff matrix between two strategies A and B as follows:

$$\begin{pmatrix} 10 & 6 \\ 7 & 2 \end{pmatrix} \tag{2.1}$$

where 10 is the payoff for A playing against A, 6 when A plays against B, 7 when B plays against A and 2 when B plays against B.

When player 2 chooses strategy A, player 1 obtains a higher payoff selecting strategy A since $10 > 7$. When player 2 chooses strategy B, player 1 still obtains a higher payoff by selecting strategy A since $6 > 2$. The same logic symmetrically applies to player 2. Neither player gains from switching from strategy A, and thus neither deviates.

This indicates that since A strictly dominates B for both players, (A, A) is the unique Nash equilibrium of this game.

Its most popular example is the Prisoner's Dilemma, the protagonist of the next section. Further,

Definition 2.2. ***Subgame Perfect Nash Equilibrium*** A Nash equilibrium is subgame perfect if it is a Nash equilibrium in every subgame of the extensive game.[3]

This definition is particularly used in the context of repeated games with fixed rounds and is relevant to Subsection 4.1.

2.2. **ESS.** Consider strategies A and B having the payoff matrix:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{2.2}$$

Then, by equating payoff with fitness, their expected frequencies are:

$$f_A = ax + by \tag{2.3}$$

$$f_B = cx + dy \tag{2.4}$$

where x is the frequency of the strategy A, y is the frequency of the strategy B, a is the payoff when A plays against A, b when A plays against B, c when B plays against A and d when B plays against B.

This assumes that the probability of interacting with an A player is the frequency of A in the population. Similarly, the probability of interacting with a B player is its frequency. This kind of selection is called frequency-dependent selection.

Consider a population of A players. In this population, players of strategy B are introduced, having an infinitesimal initial frequency of ϵ, where $\epsilon \to 0$. Thus, A's frequency is $1 - \epsilon$.

For the population of A players to successfully resist invasion by B players, its fitness must exceed B's fitness. That is,

$$a(1-\epsilon) + b\epsilon > c(1-\epsilon) + d\epsilon. \tag{2.5}$$

Taking the limit $\epsilon \to 0$,

$$a > c$$

If $a = c$,

$$b\epsilon > d\epsilon$$

$$b > d$$

Thus, a strategy is considered evolutionarily stable if it can successfully resist the invasion of another strategy. Mathematically, it is an ESS if

$$a > c \tag{2.6}$$

$$\text{OR } a = c \text{ and } b > d \tag{2.7}$$

if condition 1 fails.

Resultantly, the fitness of B is less than the average fitness of the pure A player population, and thus gets outcompeted and cannot invade. To gain an intuitive understanding, consider the following example:

Example 2.1. Consider ESS to be a home-field advantage test and A and B to be two strategies adopted by football teams. Payoffs represent an increase in fans. If A beats B when B visits A, then the second team cannot invade the home fanbase of the first team. If the outcome is a tie at the first team's home ground, then A must beat B more often at B's home turf to remain dominant.

Definition 2.3. ***Evolutionarily Stable Strategy:*** A strategy is an *evolutionarily stable strategy* (ESS) if, when it is common in the population, no mutant strategy can invade it.[4] Formally, given strategies S_k and S_j and their payoffs $E(S_k, S_j)$ and $E(S_j, S_k)$ respectively, this implies that the strategy S_k is an ESS if:

$$E(S_k, S_k) > E(S_j, S_k)$$

or if

$$E(S_k, S_k) = E(S_j, S_k) \text{ and } E(S_k, S_j) > E(S_j, S_j)$$

Definition 2.4. ***Weak Evolutionarily Stable Strategy:*** A *weak evolutionarily stable strategy* (weak ESS) is a strategy such that no alternative mutant strategy can invade it.[5]

Formally, for strategies S_k and S_j with payoff function $E(S_k, S_j)$, strategy S_k is a weak ESS if for all $S_j \neq S_k$:

$$E(S_k, S_k) \geq E(S_j, S_k) \tag{2.8}$$

$$\text{and if } E(S_k, S_k) = E(S_j, S_k), \text{ then } E(S_k, S_j) \geq E(S_j, S_j) \tag{2.9}$$

Thus, ESS is a subset of weak ESS from Definition 2.3 and Definition 2.4. In such a condition, an infinitesimal proportion of the invading strategy persists but does not grow until it disappears as a result of stochastic fluctuations (faults).

2.3. **ESS vs. Nash equilibria in Two-Strategy Games.** Systematising the notation of Definition 2.1, a strategy profile is a Nash equilibrium if:

$$E(S_k, S_k) \geq E(S_j, S_k) \; \forall j. \tag{2.10}$$

A weak ESS naturally implies a Nash equilibrium, from Definition 2.1 and Definition 2.4.

$$E(S_k, S_k) > E(S_j, S_k) \tag{2.11}$$

in which case, the strategy S_k is an ESS, following from Definition 2.3. Thus, a strict Nash equilibrium implies an ESS.

However, a Nash equilibrium may not be evolutionarily stable when

$$E(S_k, S_k) = E(S_j, S_k) \text{ and } E(S_k, S_j) < E(S_j, S_j) \tag{2.12}$$

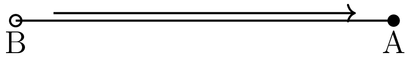
Thus, a Nash equilibrium does not imply the presence of an ESS.

$$\text{Strict Nash} \rightarrow \text{ESS} \rightarrow \text{Weak ESS} \rightarrow \text{Nash}. \tag{2.13}$$

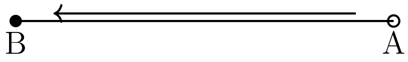
The concept of an ESS was developed because certain Nash equilibria were not evolutionarily meaningful: while some can resist invasion, other equilibria described by the condition (2.12) cannot. This makes the ESS a stricter and more relevant concept than the Nash equilibrium for evolutionary conditions.

Evolutionary game dynamics of two strategies can be represented by points on a simplex S_2, a closed interval $[0, 1]$. Consider the payoff matrix of the two strategies, A and B, to be the matrix (2.2).

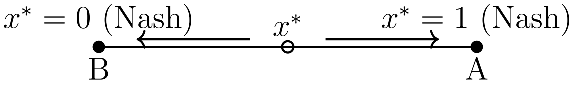
Case 1: When strategy A is dominant, i.e. $a \geq c, b \geq d$ where at least one inequality is strict, then the Nash equilibrium is (A, A) and the equilibrium is fixed at the corner point $x^* = 1$. A is the ESS that can resist invasion from B.


Case 1: A dominates

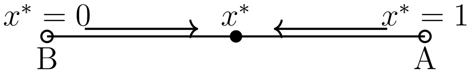
Case 2: When strategy B is dominant, i.e. $a \leq c, b \leq d$ where at least one inequality is strict, then the Nash equilibrium is (B, B) and the equilibrium is fixed at the corner point $x^* = 0$. B is the ESS that can resist invasion from A.


Case 2: B dominates

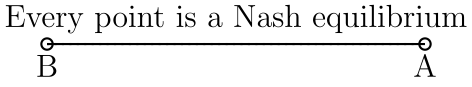
Case 3: When A and B are bistable, i.e. $a > c, b < d$, the best response for A is A, and the best response for B is B. Thus, the selection dynamics of the population depends on the initial condition. An unstable equilibrium in $[0, 1]$, given by $x^* = \frac{d-b}{a-b-c+d}$ exists. If the initial condition $x(0) > x^*$, the system converges to $x^* = 1$, an all-A population. If $x(0) < x^*$, the system converges to $x^* = 0$, an all-B population. The corner points $x^* = 1$ and $x^* = 0$ represent the two Nash equilibria for the strategy profile. However, neither pure strategy is evolutionarily stable.


Case 3: Bistability, unstable equilibrium at x^*

Case 4: When A and B coexist, i.e. $a < c, b > d$, the best response for A is B and the best response for B is A. The system will converge to an interior stable equilibrium $x^* = \frac{d-b}{a-b-c+d}$. In this equilibrium, the payoffs to A and B are equal. Thus, from Definition 2.1, the given condition represents a Nash equilibrium. Yet neither pure strategy is evolutionarily stable.


Case 4: Coexistence, stable equilibrium at x^*

Case 5: When A and B are neutral, i.e. $a = c, b = d$, selection does not change the composition of the population and every point on the simplex is an equilibrium. Thus, the initial condition determines the equilibrium. The system possesses a Nash equilibrium at every point on the simplex. However, neither strategy is evolutionarily stable.


Case 5: Neutral drift (every point is a Nash equilibrium)

To illustrate these concepts and understand why the distinction between Nash equilibria and ESS matters in evolutionary contexts, we now examine a fundamental game that appears throughout biology and economics: the *Prisoner's Dilemma.*

3. The *Prisoner's Dilemma*

3.1. An Example, and its Definition.

An example is perhaps the best way to understand the *Prisoner's Dilemma*.

Example 3.1. Consider a one-shot game played between two rational firms in duopolistic competition, selling a homogeneous product, and thus not having an inherent pricing advantage against each other. They have two choices, either to cooperate (C) by maintaining current prices or to defect (D) by actively undercutting prices.

Considering payoffs to represent their relative market share gain compared to their current and prospective competitors, let the payoff matrix for each be:

$$\begin{pmatrix} 0 & -12 \\ 10 & -2 \end{pmatrix}. \tag{3.1}$$

Here, 0 represents the payoff when both cooperate, $+10$ when the firm defects while its competitor cooperates, -12 when the firm cooperates following its competitor's defection, and -2 when both defect.

This follows logically from the assumptions. If one firm cuts prices while the other does not, demand shifts from the cooperating firm to the defecting firm, benefiting it to the detriment of its competitor. The competitor becomes significantly weakened and more vulnerable to competition. When both decide to defect by cutting prices, the market for the good experiences a price race to the bottom, where both are equally worse off. This may make both vulnerable to competition. On the other hand, when both decide to cooperate, both are equally well off.

An intuitive exercise to find the most probable outcome would be to address the choice of cooperation or defection from the perspective of a firm. The response of the firm to the cooperation of its competitor would be defection, since $10 > 0$. Its response to the defection of its competitor would also be to defect, since $-2 > -12$. Thus, both rational firms would choose to defect, making (D, D) the unique and singular outcome of the game.

This is a strict Nash equilibrium since neither firm benefits from switching strategies: both choose only to defect. The strategy of pure defection is therefore also an ESS as shown by (2.13). This is a classic example of the Prisoner's Dilemma, a Case 1/Case 2 game.

Definition 3.1. ***The Prisoner's Dilemma:*** The *Prisoner's Dilemma* is a game in which each player has two strategies: "cooperate" and "defect", and the following inequalities hold for the payoffs:

$$T > R > P > S$$
$$2R > T + S$$

where:

$T =$ temptation payoff for unilateral defection

$R =$ reward for mutual cooperation

$P =$ punishment for mutual defection

$S =$ sucker's payoff for cooperating while the other defects

In such a game, mutual defection is the unique Nash equilibrium, even though mutual cooperation would yield a higher payoff for both players.[6]

In our example,

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix} = \begin{pmatrix} 0 & -12 \\ 10 & -2 \end{pmatrix}.$$

This leads to a collectively suboptimal and Pareto inefficient outcome as the Nash equilibrium outcome leaves both players worse off than if they had both cooperated.

3.2. **Why use the *Prisoner's Dilemma* in this context?** The *Prisoner's Dilemma* is a remarkably intuitive game that has direct real-world applications in various domains of study, particularly in competitive environments. Its conclusion, that short-term incentives drive mutual defection, is corroborated by biological, economic, and social evidence. It structurally incentivises defection while having a Pareto-inefficient Nash equilibrium. It is the only game of its class for which defection strictly dominates cooperation, yet mutual cooperation offers a higher payoff than mutual defection for both the players.

Unlike cooperative games like the *Stag-Hunt*, where multiple Pareto-efficient equilibria exist and mutual best responses are cyclic, the *Prisoner's Dilemma* has an isolated, strict Nash equilibrium which directly undermines cooperation.

It is precisely this perceived inexplicability of the evolution of cooperation in the face of short-term incentives for mutual defection that is the focus of this exposition.

3.3. **Biological Implications.** This game demonstrates a fundamental paradox in evolutionary biology, as it implies that populations of defectors are evolutionarily stable and strictly dominant vis-à-vis cooperating mutants.

"Evolutionary progress, the construction of new features, often requires the cooperation of simpler parts that are already available. But cooperation is always vulnerable to exploitation by defectors." [5, p. 72]

"Natural selection favours defection. In a mixed population of cooperators and defectors, the latter will always have the higher payoff. In the framework of evolutionary game theory, therefore, defectors reproduce faster and outcompete cooperators. Under natural selection, the average fitness of the population continuously declines. When the population consists only of defectors, the average fitness has reached its minimum. Specific mechanisms are needed for natural selection to favour cooperation." [5]

Replications of the Prisoner's Dilemma are played out in countless interactions of players in nature, as players naturally prioritise short-term incentives.

However, consider the example of the evolution of mitochondria.

Example 3.2. Individual bacteria had short-term incentives to defect by taking resources without contributing toward energy production, yet the bacteria cooperated by providing ATP (energy) to early eukaryotes, who in turn cooperated by providing protection and nutrients to them. This symbiotic relationship was so successful that mitochondria are now essential organelles in nearly all eukaryotic cells, demonstrating how cooperation can become evolutionarily entrenched.

This suggests that evolutionary progress has been driven by mechanisms that have enabled cooperation to persist with defection, rather than the strict natural selection of defection even in conditions corresponding with the *Prisoner's Dilemma.*

But how has cooperation prevailed in this ocean of defection?

4. COOPERATION IN REPEATED GAMES

Example 4.1. The development of cooperation is exemplified by the institution of trade on credit. Consider a buyer and a seller of a good. The buyer is unable to make the payment for the requisite quantity at the time of purchase, and the seller extends unsecured credit to the buyer which is repaid in the short term. The buyer gains the highest immediate payoff if it chooses to defect and default on the payment. The seller obtains the highest immediate payoff if it chooses to defect on the offer for credit and demand immediate settlement from the buyer. However, since the buyer has to continually interact with sellers, a default can affect its reputation and creditability for future interactions with other sellers, which in turn may force it into unfavourable terms. The seller, on the other hand, would lose out on a prospective sticky customer and thus would bear an opportunity cost if it chose to defect.

Complex, long-run natural and social interactions are inherently repeated, rather than isolated games. This necessitates extending the analysis to iterated games where the possibility of direct reciprocity alters evolutionary outcomes.

"Direct reciprocity means that individuals use their own experience to decide whether to cooperate with another person."[8] The direct reciprocity associated with repeated games incentivises cooperation.

4.1. Questioning the Persistence of Cooperation.

Consider a Prisoner's Dilemma (T, R, P, S) game repeated m times. There are two strategies that a player can select: GRIM and ALLD. GRIM involves cooperation with the other player until the other player defects. Following this first defection, GRIM permanently switches to defection. In other words, GRIM does not forgive defection. ALLD involves defecting regardless of the action of the other player.

The payoff matrix associated with this interaction is

$$\begin{pmatrix} mR & S+(m-1)P \\ T+(m-1)P & mP \end{pmatrix} \tag{4.1}$$

where mR is the payoff when GRIM plays against GRIM, since both players cooperate for all m rounds and receive R each round. $S+(m-1)P$ is the payoff when GRIM plays against ALLD, as GRIM cooperates while ALLD defects in the first round, resulting in the player incurring the sucker's payoff of S. Following this, both defect for the next $m-1$ rounds and receive P for each round. $T+(m-1)P$ is the payoff when ALLD plays against GRIM, as ALLD defects in the first round while GRIM cooperates, resulting in the player receiving the highest payoff T. Following this, both players defect for the next $m-1$ rounds and receive P for each round. mP is the payoff when ALLD plays against ALLD, as both defect for all m rounds and receive P each round.

If

$$\begin{aligned} E(GRIM, GRIM) &> E(ALLD, GRIM) \\ mR &> T+(m-1)P \end{aligned}$$

i.e. if

$$m > \frac{T-P}{R-P} \tag{4.2}$$

then (GRIM, GRIM) is the strict Nash equilibrium. This is when cooperation over m rounds pays more than the reward of defecting on the first round but receiving penalized payoffs for all subsequent rounds. Definition 2.3 and inequality (4.2) imply that GRIM is an ESS and ALLD will not be able to invade a population of GRIM. This explains the persistence of cooperation when faced with purely defective strategies once cooperation in the form of a GRIM population is established.

However, $(\mathrm{ALLD}, \mathrm{ALLD})$ is also a strict Nash equilibrium with ALLD as an ESS, since $mP > S + (m-1)P$ because $P > S$. However, this equilibrium does not account for the emergence of cooperation in a defecting population.

As Nowak highlights in [5, p. 76], a major problem arises when both players know exactly how many rounds the game lasts. Since direct reciprocity does not add weight to the decision to cooperate in the last round, neither player has an incentive to do so. The last round thus becomes a classic one-shot *Prisoner's Dilemma*. The GRIM strategy with a modification — a last round defection — is strictly dominant against the original GRIM and is thus the ESS:

$$\begin{pmatrix} mR & S+(m-1)R \\ T+(m-1)R & mR+P \end{pmatrix}. \tag{4.3}$$

The same analysis can be extended to the second-last round. Since the last round will be a one-shot *Prisoner's Dilemma* with a fixed (D, D) outcome, neither player has an incentive to cooperate in the penultimate round. Thus, the modification of the GRIM strategy that defects in the last two rounds strictly dominates the modification that defects in only the last round. Backward induction implies that the ESS reaches ALLD from GRIM as each variant is replaced by another variant that defects one round before it. No cooperative strategy is part of the subgame perfect Nash equilibrium, i.e. SPNE (Definition 2.2), as it is rational to defect in every round.

Formalising this backward induction argument, let H_t be the set of all histories up to round t, where a history $h_{t-1} \in H_{t-1}$ records all actions taken from round 1 to $t-1$. For a strategy profile $\sigma = (\sigma_1, \sigma_2)$ and history h_{t-1}:

$$V_t(\sigma, h_{t-1}) = \text{the expected payoff from round t onward for player 1} \tag{4.4}$$

For the m^{th} round,

$$V_m(\sigma^*, h_{m-1}) = P \ \forall h_{m-1} \tag{4.5}$$

where σ^* is the SPNE. For earlier rounds, $t < m$:

$$V_t(\sigma^*, h_{t-1}) = P + V_{t+1}(\sigma^*, h_t) = P + P + \cdots + P = (m-t+1)P. \tag{4.6}$$

Any deviation from σ^* $\forall t$, yields at most:

$$\text{any unilateral deviation} + (m-t)P = S + (m-t)P < P + (m-t)P = (m-t+1)P$$

since the deviating player receives the sucker's payoff if they unilaterally deviate from the SPNE strategy of defection. Thus, σ^*, i.e. ALLD is the only strict Nash equilibrium throughout the game. This threatens the explanation for the persistence of cooperation in fixed repeated games.

4.2. **Infinitely Repeated Games.** Nowak, in [5, p. 76], notes that in the case of both natural and human interactions the players do not know how many rounds the game will last. With an uncertain number of rounds, defection becomes more costly as premature defection risks forfeiting future payoffs if the game continues longer than anticipated.

"An infinitely repeated game is a game in which the same stage game is played an infinite number of times, or when each round continues with a fixed probability after the previous round, independently of history." [6]

Consider a repeated Prisoner's Dilemma where the probability that another round will be played is w and the number of rounds is $\bar{m}$. The expected number of rounds is:

$$\bar{m} = \frac{1}{1-w}. \tag{4.7}$$

The greater the w, the higher the payoff for cooperation as w discounts future payoffs and makes immediate payoffs less salient. There is no GRIM variant that can simply defect on the last move in this game. Thus, GRIM is an ESS when:

$$\bar{m} > \frac{T-P}{R-P}. \tag{4.8}$$

However, GRIM is not ideal since it is unforgiving once defected against. This may risk a significant payoff opportunity cost in the case of premature defection. Which strategy has the highest average payoff in its interactions with other strategies in a strategy profile in an infinitely repeated game? Does the pursuit of this inquiry hint toward the emergence of cooperation?

5. Axelrod's Tournaments, Stochastic Strategies and their Payoffs

Robert Axelrod, a political scientist at the University of Michigan, invited people to submit strategies for the repeated Prisoner's Dilemma game, where every strategy would be played against every other strategy and the payoff for each strategy was summed and ranked.[1] The tournaments assumed deterministic and error-free conditions. The winning strategy was Tit-For-Tat (TFT). TFT starts with C and then replicates the opponent's action in the previous round. TFT is a "nice" strategy – it is never the first to defect and never earns a payoff that is higher than its opponent. However, the payoff sum over all matches is the highest for TFT, implying that its success is due to high average payoffs per round. TFT is an ESS against ALLD when condition (4.2) is satisfied for $\bar{m}$. TFT betters GRIM as it forgives defection.

However, TFT has a flaw: its performance is significantly hampered by mistakes. "Real-world situations are permeated by mistakes. A "trembling hand" can lead to a misimplementation of one's own action. A "fuzzy mind" can cause a misinterpretation of the opponent's move." [5, p. 79]

Consider a game between two TFT players. One mistake can move the game from mutual cooperation to alternate cooperation and defection. The mistake being at $t = 2$,

$$TFT : CCCC\dot{D}CDCDC\ldots \tag{5.1}$$

$$TFT : CCCCCDCDCD\ldots \tag{5.2}$$

A second mistake can lead to mutual defection.

$$TFT : CCCCC\dot{D}CD\dot{D}D\ldots$$
$$TFT : CCCCCCDCDD\ldots$$

In the limit as $\bar{m} \to \infty$, with two TFT players having a probability ϵ (taking $\epsilon \to 0$) of making mistakes, the game approximates a Markov process, preventing the system from remaining at (C, C). Therefore, the payoff for the two players is:

$$E(TFT, TFT) = \bar{m}(\frac{R + P + T + S}{4}) < \bar{m}R. \tag{5.3}$$

Thus, TFT cannot correct mistakes. This reduces its average payoff considerably. Since Axelrod's tournaments did not factor mistakes they did not expose TFT's "Achilles Heel". [5, p. 79]

Unlike deterministic strategies such as TFT, GRIM, ALLD and ALLC, which assign a binary probability to C or D, memory-1 stochastic strategies (M1SS) assign a probability $p \in [0, 1]$ for C after the cooperation of the opponent and $q \in [0, 1]$ for C after the opponent's defection. M1SS can assign a non-zero probability for q, thus helping to correct mistakes.

The payoff matrix for a repeated Prisoner's Dilemma between two M1SS is a Markov chain for the four states in order, (C, C), (C, D), (D, C), (D, D). Consider two M1SS, $S(p_1, q_1)$ and $S(p_2, q_2)$. The Markov chain is defined by the transition matrix $M = [a_{ij}]$, with each entry a_{ij} indicating the probability of shifting from state i to j, as being:

$$M = \begin{pmatrix} p_1p_2 & p_1(1-p_2) & (1-p_1)p_2 & (1-p_1)(1-p_2) \\ q_1p_2 & q_1(1-p_2) & (1-q_1)p_2 & (1-q_1)(1-p_2) \\ p_1q_2 & p_1(1-q_2) & (1-p_1)q_2 & (1-p_1)(1-q_2) \\ q_1q_2 & q_1(1-q_2) & (1-q_1)q_2 & (1-q_1)(1-q_2) \end{pmatrix}. \tag{5.4}$$

Let $\vec{x_t}$ denote the probability distribution of the four states of the game at time t, implying the time spent in each state by the system. Therefore,

$$\vec{x_t} = \begin{pmatrix} \pi_{CC} \\ \pi_{CD} \\ \pi_{DC} \\ \pi_{DD} \end{pmatrix} \tag{5.5}$$

where π_{CC}, π_{CD}, π_{DC}, and π_{DD} represent the long-term probabilities that the system is in state (C, C), (C, D), (D, C), and (D, D), respectively. Therefore, at time $t + 1$, the distribution evolves as:

$$\vec{x_{t+1}} = M\vec{x_t}. \tag{5.6}$$

The probability distribution vector of a given round is transformed after each round by the transition matrix and thus is dependent only on the probability distribution vector of the previous round, highlighting its Markov memory-1 characteristic.

Definition 5.1. ***Eigenvector and Eigenvalue:*** A non-zero vector $\vec{x}$ is called an *eigenvector* of a square matrix M if there exists a scalar λ (called the *eigenvalue*) such that:

$$M\vec{x} = \lambda\vec{x}$$

Definition 5.2. ***Regular Matrix:*** A real square matrix M is said to be *regular* if there exists a positive integer k such that all entries of M^k are strictly positive.

The Perron-Frobenius Theorem states:

Theorem 5.3. *Let M be a real, non-negative square matrix of size $n \times n$. If M is regular, i.e. there exists some power k such that all entries of M^k are positive, then:*

1. *M has a unique largest real eigenvalue λ_{max} (called the Perron-Frobenius eigenvalue).*
2. *This eigenvalue is simple (has algebraic multiplicity 1) and positive.*
3. *All other eigenvalues λ of M satisfy $|\lambda| < \lambda_{max}$.*
4. *There exists a unique (up to scaling) positive left-hand eigenvector $\vec{x}$ associated with λ_{max}.*

[2][7].

.

For any probability distribution $\vec{x_t}$, $\sum_{i=1}^{4} x_{i,t} = 1 \; \forall \; t$ since all states are mutually exclusive and collectively exhaustive. Thus, $\lambda = 1$, proven by Theorem 5.3.

The components of $\vec{x}$ represent the stationary long-run distribution of the Markov chain, for the limit $t \to \infty$. Thus,

$$\vec{x} = M\vec{x}. \tag{5.7}$$

The Markov chain converges to this $\vec{x_t}$ if and only if all eigenvalues other than $\lambda = 1$ satisfy $|\lambda| < 1$.

And therefore, the expected average payoff per round for $t \to \infty$ is:

$$E[S(p_1, q_1),\; S(p_2, q_2)] = \begin{pmatrix} R & S & T & P \end{pmatrix} \vec{x} \tag{5.8}$$

Let:

$$r_1 = p_1 - q_1 \tag{5.9}$$

$$r_2 = p_2 - q_2. \tag{5.10}$$

Each measures the responsiveness of the strategy to the action of the opponent.

Let s_1 and s_2 denote the long-run probabilities of cooperation for players 1 and 2.

$$\begin{aligned} s_1 &= \pi_{CC}p_1 + \pi_{CD}q_1 + \pi_{DC}p_1 + \pi_{DD}q_1 \\ s_2 &= \pi_{CC}p_2 + \pi_{CD}q_2 + \pi_{DC}p_2 + \pi_{DD}q_2 \end{aligned} \tag{5.11}$$

Thus,

$$\vec{x} = \begin{pmatrix} s_1 s_2 \\ s_1(1 - s_2) \\ s_2(1 - s_1) \\ (1 - s_1)(1 - s_2) \end{pmatrix}. \tag{5.12}$$

By substituting equation (5.12) into equation (5.11),

$$s_1 = s_1 s_2 p_1 + s_1(1 - s_2)q_1 + (1 - s_1)s_2 p_1 + (1 - s_1)(1 - s_2)q_1 \tag{5.13}$$

$$s_2 = s_1 s_2 p_2 + s_1(1 - s_2)q_2 + (1 - s_1)s_2 p_2 + (1 - s_1)(1 - s_2)q_2. \tag{5.14}$$

Taking s_1 and simplifying the equation (5.13),

$$s_1 = s_1 s_2 p_1 + s_1 q_1 - s_1 s_2 q_1 + s_2 p_1 - s_1 s_2 p_1 + q_1 - s_1 q_1 - s_2 q_1 + s_1 s_2 q_1.$$

$$s_1 = s_2 p_1 - s_2 q_1 + q_1 = s_2 r_1 + q_1 \tag{5.15}$$

Similarly,

$$s_2 = s_1 r_2 + q_2 \tag{5.16}$$

Substituting equation (5.16) into equation (5.15),

$$s_1 = s_1 r_2 r_1 + r_1 q_2 + q_1$$

$$s_1 (1 - r_2 r_1) = r_1 q_2 + q_1$$

$$s_1 = \frac{r_1 q_2 + q_1}{1 - r_2 r_1}. \tag{5.17}$$

Similarly,

$$s_2 = \frac{r_2 q_1 + q_2}{1 - r_2 r_1}. \tag{5.18}$$

Thus, from equations (5.8) and (5.12), the per round payoff is:

$$E(S_1, S_2) = R s_1 s_2 + S s_1 (1 - s_2) + T (1 - s_1) s_2 + P (1 - s_1)(1 - s_2). \tag{5.19}$$

For equations (5.17) and (5.18) to hold, $|r_1 r_2| < 1$. Or else, it implies infinite ($|r_1 r_2| = 1$) or negative ($|r_1 r_2| > 1$) probabilities for s_1 and s_2, which is not possible. This probabilistic intuition is theoretically supported by Theorem 5.3 to necessitate $|r_1 r_2| < 1$ for the Markov chain to be regular, which is, in turn, required for convergence to a unique stable probability distribution.

Therefore, this condition is necessary for convergence to a unique stationary and stable probability distribution eigenvector having $\lambda = 1$.

For deterministic strategies having $|r_1 r_2| = 1$, in a mistake-free environment, convergence is either trivial (such as in a game between TFT and TFT, there is a stable state (C, C)) or is cyclic with known periodicity.

6. The Emergence of Cooperation

Let a strategy be described by its parameters $S = (p_{CC}, p_{CD}, p_{DC}, p_{DD})$.

Using the method mentioned in the previous chapter for the computation of payoffs for interactions between stochastic strategies, Nowak, in [5, p. 84], investigates the evolution of cooperation through in silico evolution that factored in mistakes, taking n strategies. In most cases, the n strategies converge to the one closest to ALLD. However, he noted that in cases where n included a strategy close to TFT (1, 0, 1, 0), this strategy increased in frequency versus the strategy closest to ALLD once almost all other strategies disappeared, because its average payoff is higher than ALLD's above a certain frequency of TFT players in the population. This threshold depends on the particular payoff matrix of the game. This is because:

$$E(TFT, TFT) = \bar{m} R > E(ALLD, ALLD) = \bar{m} P. \tag{6.1}$$

TFT is, in turn, replaced by more cooperative strategies such as the "Generous" Tit-for-Tat (1, 1/3, 1, 1/3). This is because:

$$E(GTFT, GTFT) > E(TFT, TFT) = \bar{m}(\frac{R+P+T+S}{4}). \tag{6.2}$$

GTFT forgiveness of defections with a probability $\frac{1}{3}$, results in a significantly higher payoff when it plays against itself, compared to TFT's when it plays against itself, since GTFT can make up for mistakes. Thus, cooperation evolves further from TFT to GTFT. This highlights TFT's second weakness: its vulnerability toward a drift toward a greater degree of cooperation.

However:

$$E(GTFT, ALLD) < E(TFT, ALLD) \tag{6.3}$$

as stochastically forgiving defections results in a greater S for GTFT against ALLD compared to TFT against ALLD. Thus, the minimum frequency threshold for GTFT to be an ESS in an ALLD environment is significantly higher than the same for TFT.

Therefore, TFT is needed to initiate cooperation. But once cooperation is established, GTFT shall replace TFT.

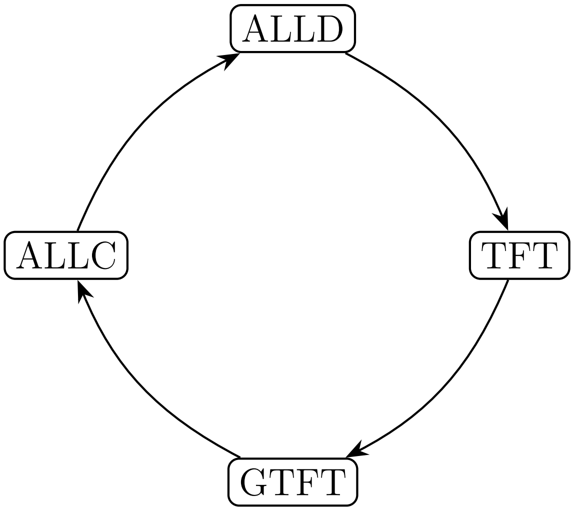
GTFT is defined as the following:

$$GTFT : p = 1 \text{ and } q = \min\{1 - \frac{T-R}{R-S}, \frac{R-P}{T-P}\} \tag{6.4}$$

where q is the highest probability of forgiveness that is an ESS. Increasing q invites invasion by ALLD. q is derived by maximising $E(q)$ over q in [0,1]. Thus, GTFT provides the highest payoff for a strategy that resists invasion by ALLD.

Resultantly, the strategy profile evolves from defective strategies to cooperative strategies. This causes the q for GTFT to reduce over generations, resulting in a cooperative drift from GTFT to ALLC. Thus, GTFT is by definition vulnerable to cooperative drift, which invites defectors, initiating an evolutionary race toward ALLD.

The Chain of Evolution



Thus, while TFT explains the emergence of cooperation against ALLD, the subsequent chain of evolution does not explain its persistence.

Example 6.1. This natural sequence of events occurs in the financial markets.

(1) In a prosperous economic environment, institutions with greater risk tolerance (analogous to cooperators) gain market share against more conservative peers. As a result, the degree of forgiveness (risk appetite) in the economy increases.
(2) This makes the environment increasingly fragile. When a shock or a wave of defaults occurs, the overly forgiving institutions are unable to absorb the losses.
(3) Post-collapse, risk-averse institutions (defectors) who survive this dominate by gaining market share from their risk-tolerant competitors.
(4) As market confidence returns, cautiously cooperative institutions (TFTs) begin outperforming defectors by selectively raising risk tolerance. This plays a role in initiating a prosperous economic environment.
(5) The cycle repeats.

7. Win-Stay-Lose-Shift, The Gamechanger

Until now, we have analysed four cooperative strategies: ALLC, GRIM, TFT and GTFT. None is evolutionarily stable, as each has its weaknesses.

(1) ALLC is vulnerable to defectors.
(2) GRIM is vulnerable to ALLD in repeated games with fixed rounds and is weakened due to its punitive defective stance in infinitely repeated games.
(3) TFT is vulnerable to mistakes and an eventual cooperative drift as it evolves into GTFT.
(4) GTFT is naturally vulnerable to a cooperative drift as its frequency in the environment increases.

Fortunately, a semi-cooperative Win-Stay, Lose-Shift (WSLS) strategy also evolved from ALLC in the in-silico evolution experiment, described deterministically by $S = (1, 0, 0, 1)$. This strategy can be considered "out-of-the-box" since it responds to the outcome of the previous round rather than action of the opponent, unlike ALLC, GRIM, TFT and GTFT.

WSLS cooperates when the previous outcome is a (C, C) or (D, D) and defects when the previous outcome is a (D, C) or (C, D). This implies that for the "winning" outcomes, i.e. T or R, the strategy "stays" with its action while for the "losing" outcomes i.e. S or P, it "shifts" its action.

7.1. **Assessing WSLS' Evolutionary Stability.** The first step is to evaluate the stability of WSLS against invasion by defective strategies.

$E(WSLS, WSLS) = \bar{m}R$ while $E(ALLD, WSLS) = \bar{m}(\frac{T+P}{2})$ for $\bar{m} \to \infty$.

Thus, from (2.10), WSLS is an ESS against ALLD only if:

$$R > \frac{(T+P)}{2} \tag{7.1}$$

Therefore, WSLS $(1, 0, 0, 1-\epsilon)$, taking $\epsilon \to 0$, is an ESS against ALLD. For $R < \frac{(T+P)}{2}$, WSLS $(1, 0, 0, x)$ for $x < \frac{R-P}{T-R}$ is an ESS against ALLD. Thus, WSLS overcomes ALLC's weakness of susceptibility to defective strategies.

WSLS is also capable of correcting its mistakes. For a game between two WSLS players, both start with a sequence of cooperative decisions resulting in (C, C) until one defects by mistake. Following this defection, the other player also defects, since (D, C) is losing for them, while the first player repeats their defection since (D, C) is winning for them, resulting in a (D, D) outcome in the subsequent round. As this outcome is considered “losing” for both, both players shift to cooperation, i.e. (C, C).

$$\begin{aligned} WSLS &: CCCCCC\dot{D}DCC\ldots \\ WSLS &: CCCCCCCDCC\ldots \end{aligned}$$

Thereby, WSLS overcomes TFT’s susceptibility to mistakes.

Further, for a game between two players, one adopting WSLS and the other adopting ALLC, both start by a sequence of cooperative decisions resulting in (C, C) until the WSLS player defects by mistake. Resultantly, the outcome is (D, C) which is in the favour of the WSLS player. This is considered a “winning” outcome and thus, the WSLS player will continue to defect for the rest of the rounds. On the other hand, ALLC does not reciprocate the defection. Thereby WSLS is resistant to cooperative drift, further outperforming TFT and bettering GTFT.

$$\begin{aligned} WSLS &: CCCCCC\dot{D}DDD\ldots \\ ALLC &: CCCCCCCCCC\ldots \end{aligned}$$

Therefore, the WSLS strategy is evolutionarily stable against deterministic defective strategies once established, and thus ensures the persistence of cooperation.

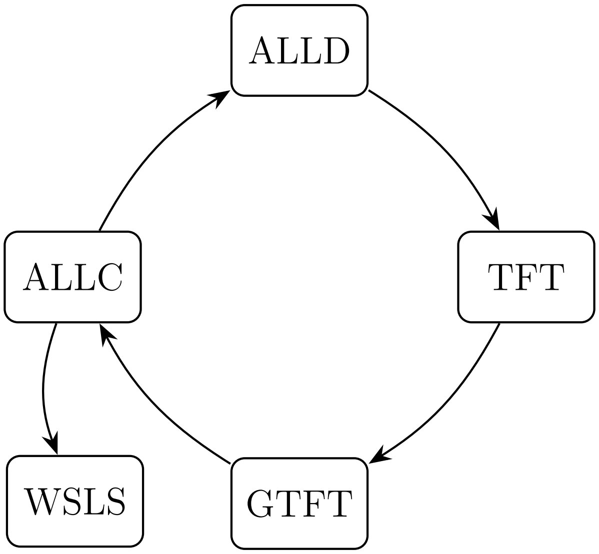
However, WSLS evidently cannot initiate cooperation, as for that, $E(ALLD, ALLD) < E(ALLD, WSLS)$. But:

$$\begin{aligned} E(ALLD, ALLD) &= \bar{m}P \\ E(WSLS, ALLD) &= \frac{\bar{m}(S+P)}{2} \end{aligned}$$

$$\therefore E(ALLD, ALLD) > E(WSLS, ALLD). \tag{7.2}$$

A diagrammatic representation of the evolution and persistence of cooperation is:

The Emergence and Persistence of Cooperation



7.2. **WSLS' Vulnerability to Asymmetry, and an Insight into Rationality.** WSLS is an unfathomably robust strategy in the standard infinitely repeated *Prisoner's Dilemma* with symmetric preferences when the condition (7.1) is satisfied, apart from its inability to initiate cooperation.

However, WSLS is vulnerable to asymmetric preferences. Let this table be an asymmetric *Prisoner's Dilemma* game between players A and B:

Outcome	**A's Payoff**	**B's Payoff**
(C,C)	3	2
(C,D)	0	5
(D,C)	5	0
(D,D)	1	1

In this distinct game, let the payoff threshold for "winning" outcomes for A be $E(A,B) > 1$ and for B be $E(B,A) \geq 1$.

If a sequence of (C,C) outcomes is followed by a mistake that generates a (C,D) outcome, A shifts from C to D as it is "losing" for A, while B will stay with D as it is "winning" for B. This results in (D,D). A shifts once again from D to C, however, B does not, as the payoff for (D,D) meets its requisite threshold.

This is followed by similar alternating sequence of C and D as in TFT's case (5.1). This occurrence results in:

$$E(A,B) = \bar{m}(\frac{R+T+P+S}{4}) \tag{7.3}$$

which is equivalent to $E(TFT,TFT)$ in symmetric games, i.e. condition (5.3).

Such asymmetric games arise, either from unequal payoffs or from unequal payoff-based definitions of "winning" outcomes for the two players. These differences reflect the revealed preference of the two players: preferences inferred from observed historical behaviour of each player rather than assumed symmetric payoff matrices. The Revealed Preference Hypothesis empirically investigates subjective rationality. Thus, for real-life infinitely repeated *Prisoner's Dilemma* games, WSLS may not remain as resolute a strategy as it does under the assumption of objective rationality, i.e. consistent payoffs and preference orderings.

Exceedingly rarely, yet possibly, in an infinitely repeated symmetric *Prisoner's Dilemma*, WSLS may witness a cooperative drift in case the probability of mistakes occurring in games between WSLS and ALLC (ϵ) exceeds the probability of mistakes occurring in games between WSLS and WSLS. This, in turn, initiates the chain of evolution until a strategy similar to WSLS breaks out from a cooperative strategy profile.

8. Appendix: An Example of Long-run Payoff Calculations: Axelrod's Tournaments

The strategy profile S for the environment comprises $(TFT, ALLD, ALLC, GTFT, WSLS)$, where each is defined deterministically as $(1,0,1,0), (0,0,0,0), (1,1,1,1), (1,\frac{1}{3},1,\frac{1}{3})$, and $(1,0,0,1)$, respectively. This environment will not be susceptible to mistakes. Let:

$$S = \begin{pmatrix} 1/5 \\ 1/5 \\ 1/5 \\ 1/5 \\ 1/5 \end{pmatrix} \tag{8.1}$$

Let the *Prisoner's Dilemma* payoff matrix for this environment be:

$$\begin{pmatrix} 3 & 0 \\ 5 & 1 \end{pmatrix} \tag{8.2}$$

8.1. **Payoff Table.** Thus, the payoff table for every strategy is:

Table 1. (Axelrod's PD: T=5, R=3, P=1, S=0)

vs	**TFT**	**ALLC**	**ALLD**	**GTFT**	**WSLS**
TFT (1,0,1,0)	$3\bar{m}$	$3\bar{m}$	$1\bar{m}$	$3\bar{m}$	$3\bar{m}$
ALLC (1,1,1,1)	$3\bar{m}$	$3\bar{m}$	$0\bar{m}$	$3\bar{m}$	$3\bar{m}$
ALLD (0,0,0,0)	$1\bar{m}$	$5\bar{m}$	$1\bar{m}$	$2.33\bar{m}$	$3\bar{m}$
GTFT (1,1/3,1,1/3)	$3\bar{m}$	$3\bar{m}$	$0.67\bar{m}$	$3\bar{m}$	$3\bar{m}$
WSLS (1,0,0,1)	$3\bar{m}$	$3\bar{m}$	$0.5\bar{m}$	$3\bar{m}$	$3\bar{m}$

8.2. **Using the Replicator Equation.** The average payoffs, i.e. the fitness f for each strategy is:

$$\begin{aligned}
f_{TFT} &= \frac{1}{5} * (3+3+1+3+3)\bar{m} = 2.6\bar{m} \\
f_{ALLC} &= \frac{1}{5} * (3+3+0+3+3)\bar{m} = 2.4\bar{m} \\
f_{ALLD} &= \frac{1}{5} * (1+5+1+2.33+3)\bar{m} = 2.46\bar{m} \\
f_{GTFT} &= \frac{1}{5} * (3+3+0.67+3+3)\bar{m} = 2.54\bar{m} \\
f_{WSLS} &= \frac{1}{5} * (3+3+0.5+3+3)\bar{m} = 2.5\bar{m}
\end{aligned}$$

Therefore, the average fitness ϕ for the strategy profile, from Definition 1.1, is:

$$\phi = \frac{1}{5} * (2.6+2.4+2.46+2.54+2.5)\bar{m} = 2.5\bar{m} \tag{8.3}$$

By inserting the calculated inputs into the replicator equation, from Definition 1.1:

$$\begin{aligned}
\dot{x}_{TFT} &= \frac{1}{5} * (2.6 - 2.5) = 0.02 \\
\dot{x}_{ALLC} &= \frac{1}{5} * (2.4 - 2.5) = -0.02 \\
\dot{x}_{ALLD} &= \frac{1}{5} * (2.46 - 2.5) = -0.008 \\
\dot{x}_{GTFT} &= \frac{1}{5} * (2.54 - 2.5) = 0.008 \\
\dot{x}_{WSLS} &= \frac{1}{5} * (2.5 - 2.5) = 0
\end{aligned}$$

This corroborates with the results of Axelrod's Tournaments, where in a symmetric infinitely repeated *Prisoner's Dilemma* without mistakes, TFT emerged as the winner by having the greatest differential for frequency with respect to time.

Acknowledgement

This paper was written under the Euler Circle Independent Research Paper Writing class. I would like to thank Dr. Simon Rubinstein-Salzedo for this opportunity and for pointing to resources used to study the topic. I would like to thank my mentor Zarif Ahsan for his suggestions and assurance as I wrote the paper.

References

[1] Robert Axelrod and William D. Hamilton. The evolution of cooperation. *Science*, 211(4489):1390–1396, 1981.

[2] Georg Frobenius. Über matrizen aus nicht negativen elementen. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pages 456–477, 1912.

[3] Michael Maschler, Eilon Solan, and Shmuel Zamir. *Game Theory*. Cambridge University Press, Cambridge, 2013. See Chapter 4, Section 4.3, p.142 for the definition of subgame perfect equilibrium.

[4] John Maynard Smith. *Evolution and the Theory of Games*. Cambridge University Press, 1982.

[5] Martin A Nowak. *Evolutionary Dynamics: Exploring the Equations of Life*. Harvard University Press, 2006.

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

[7] Oskar Perron. Zur theorie der matrizen. *Mathematische Annalen*, 64(2):248–263, 1907.

[8] Laura Schmid, Krishnendu Chatterjee, Christian Hilbe, and Martin A. Nowak. A unified framework of direct and indirect reciprocity. *Nature Human Behaviour*, 5(10):1292–1302, October 2021.