The Evolution of Cooperation

Agastya Mehta
agastyamehta28@gmail.com

July 8, 2025

Introduction

- Why study the evolution of cooperation?
- Under which lens can cooperation be analysed?
- Ultimately, how does cooperation emerge (and persist)?

Evolutionarily Stable Strategies

Definition

A strategy is an evolutionarily stable strategy (ESS) if, when it is common in the population, no mutant strategy can invade it. 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

$$E(S_k, S_k) = E(S_j, S_k) \;\&\; E(S_k, S_j) > E(S_j, S_j)$$

Evolutionary Stable Strategies

Consider strategies P and Q having the payoff matrix:

	P	Q
P	a	b
Q	c	d

Evolutionary Stable Strategies

Consider strategies P and Q having the payoff matrix:

	P	Q
P	a	b
Q	c	d

For a population of P $(1-\epsilon)$ players to resist the invasion of Q (ϵ) players, P's payoffs must exceed Q's:

$$a(1-\epsilon) + b\epsilon > c(1-\epsilon) + d\epsilon$$

Evolutionary Stable Strategies

Consider strategies P and Q having the payoff matrix:

	P	Q
P	a	b
Q	c	d

For a population of P $(1-\epsilon)$ players to resist the invasion of Q (ϵ) players, P's payoffs must exceed Q's:

$$a(1-\epsilon) + b\epsilon > c(1-\epsilon) + d\epsilon$$

Taking the limit $\epsilon \to 0$:

$$a > c$$

For $a = c$,

$$b > d$$

The Lens: The *Prisoner's Dilemma*

Definition

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:

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

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix}$$

where:

$$T = \textit{temptation payoff for unilateral defection}$$
$$R = \textit{reward payoff for mutual cooperation}$$
$$P = \textit{punishment for mutual defection}$$
$$S = \textit{sucker's payoff for unilateral cooperation}$$

Why the *Prisoner's Dilemma*?

In a *Prisoner's Dilemma* game, $c > a$ and $d > b$.

Defection is the ESS. Cooperation is structurally disincentivised. Yet, mutual cooperation offers a higher payoff than mutual defection.

+ Unlike other games (like *Stag-Hunt*, *Halk-Dove*), it has an isolated, strictly dominant (Nash) equilibrium directly undermining cooperation.

Why the *Prisoner's Dilemma*?

In a *Prisoner's Dilemma* game, $c > a$ and $d > b$.

Defection is the ESS. Cooperation is structurally disincentivised. Yet, mutual cooperation offers a higher payoff than mutual defection.

+ Unlike other games (like *Stag-Hunt*, *Halk-Dove*), it has an isolated, strictly dominant (Nash) equilibrium directly undermining cooperation.

= Finding the mechanism driving cooperation's evolution in the *Prisoner's Dilemma* is like finding the solution to the hardest problem of the pset.

Repeated Games

Interactions are not one-shot *Prisoner's Dilemmas*, they are long-run, having future consequences.
"Direct reciprocity means that individuals use their own experience to decide whether to cooperate with another person."

Repeated Games

Interactions are not one-shot *Prisoner's Dilemmas*, they are long-run, having future consequences.
"Direct reciprocity means that individuals use their own experience to decide whether to cooperate with another person."

Example

Consider a buyer and a seller. The buyer demands short-term unsecured credit.
The buyer has two options, considering credit is extended:

1. *Default on payment, i.e. maximum short-term gain (defect)*
2. *Settle the transaction (cooperate)*

The seller has two options:

1. *Demand immediate settlement (defect)*
2. *Extend credit (cooperate)*

Repeated Games with Fixed Rounds

Consider two strategies, GRIM and ALLD. GRIM cooperates until the opponent defects and does not forgive this defection. ALLD always defects. Their payoff matrix for fixed m rounds is:

	$GRIM$	$ALLD$
$GRIM$	mR	$S + (m-1)P$
$ALLD$	$T + (m-1)P$	mP

Repeated Games with Fixed Rounds

Consider two strategies, GRIM and ALLD. GRIM cooperates until the opponent defects and does not forgive this defection. ALLD always defects. Their payoff matrix for fixed m rounds is:

	$GRIM$	$ALLD$
$GRIM$	mR	$S+(m-1)P$
$ALLD$	$T+(m-1)P$	mP

This makes GRIM an ESS against ALLD, for $mR > T+(m-1)P$. Right?

Repeated Games with Fixed Rounds

Consider two strategies, GRIM and ALLD. GRIM cooperates until the opponent defects and does not forgive this defection. ALLD always defects. Their payoff matrix for fixed m rounds is:

	$GRIM$	$ALLD$
$GRIM$	mR	$S+(m-1)P$
$ALLD$	$T+(m-1)P$	mP

This makes GRIM an ESS against ALLD, for $mR > T+(m-1)P$. Right?

Wrong. What prevents either player from defecting on the last round? Not direct reciprocity.

Infinitely Repeated Games

But do most games have fixed rounds? Do players know how many times they will interact with each other?

Infinitely Repeated Games

But do most games have fixed rounds? Do players know how many times they will interact with each other?

Definition

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.

This would make GRIM an ESS against ALLD. But is GRIM vulnerable to other strategies?

Infinitely Repeated Games

But do most games have fixed rounds? Do players know how many times they will interact with each other?

Definition

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.

This would make GRIM an ESS against ALLD. But is GRIM vulnerable to other strategies?

What are the evolutionary dynamics of infinitely repeated games?

Axelrod's Tournaments

Robert Axelrod, a political scientist, sought an answer. He called for strategies for a mistake-free infinitely repeated *Prisoner's Dilemma* game.

Each strategy is played against every other, and are then ranked on basis of average payoffs.

Axelrod's Tournaments

Robert Axelrod, a political scientist, sought an answer. He called for strategies for a mistake-free infinitely repeated *Prisoner's Dilemma* game.

Each strategy is played against every other, and are then ranked on basis of average payoffs.

The winner: Tit-for-Tat (TFT).

Axelrod's Tournaments

Robert Axelrod, a political scientist, sought an answer. He called for strategies for a mistake-free infinitely repeated *Prisoner's Dilemma* game.

Each strategy is played against every other, and are then ranked on basis of average payoffs.

The winner: Tit-for-Tat (TFT).

TFT vs. ALLD is very similar to GRIM vs. ALLD, but it beats GRIM by not holding grudges.

TFT is an ESS against ALLD for $\bar{m}R > T + (\bar{m} - 1)P$.

Axelrod's Tournaments

Robert Axelrod, a political scientist, sought an answer. He called for strategies for a mistake-free infinitely repeated *Prisoner's Dilemma* game.

Each strategy is played against every other, and are then ranked on basis of average payoffs.

The winner: Tit-for-Tat (TFT).

TFT vs. ALLD is very similar to GRIM vs. ALLD, but it beats GRIM by not holding grudges.

TFT is an ESS against ALLD for $\bar{m}R > T + (\bar{m} - 1)P$.

But what makes TFT special?

The Emergence of Cooperation

The payoff matrix between ALLD and TFT:

	$ALLD$	TFT
$ALLD$	$\bar{m}P$	$T + (\bar{m} - 1)P$
TFT	$S + (\bar{m} - 1)P$	$\bar{m}R$

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

The Emergence of Cooperation

The payoff matrix between ALLD and TFT:

	$ALLD$	TFT
$ALLD$	$\bar{m}P$	$T + (\bar{m} - 1)P$
TFT	$S + (\bar{m} - 1)P$	$\bar{m}R$

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

For $\bar{m} \to \infty$, $\bar{m}P - (S + (\bar{m} - 1)P) \to 0$, i.e. $a = c$.

The Emergence of Cooperation

The payoff matrix between ALLD and TFT:

	$ALLD$	TFT
$ALLD$	$\bar{m}P$	$T+(\bar{m}-1)P$
TFT	$S+(\bar{m}-1)P$	$\bar{m}R$

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

For $\bar{m} \to \infty$, $\bar{m}P - (S + (\bar{m}-1)P) \to 0$, i.e. $a = c$.

Good news: $\bar{m}R > T + (\bar{m}-1)P$, i.e. $d > b$

$\implies$ TFT invades ALLD!

TFT's Achilles Heel

We have not considered mistakes until now.

TFT's Achilles Heel

We have not considered mistakes until now.

When we do:

$$TFT : CCCC\dot{D}CDCDC\ldots$$
$$TFT : CCCCCDCDCD\ldots$$

TFT's Achilles Heel

We have not considered mistakes until now.

When we do:

$$\begin{aligned} TFT &: CCCC\dot{D}CDCDC\ldots \\ TFT &: CCCCCDCDCD\ldots \end{aligned}$$

Even worse, for two mistakes:

$$\begin{aligned} TFT &: CCCC\dot{D}CD\dot{D}DD\ldots \\ TFT &: CCCCCDCDDD\ldots \end{aligned}$$

TFT's Achilles Heel

We have not considered mistakes until now.

When we do:

$$TFT : CCCC\dot{D}CDCDC\ldots$$
$$TFT : CCCCCDCDCD\ldots$$

Even worse, for two mistakes:

$$TFT : CCCC\dot{D}CD\dot{D}DD\ldots$$
$$TFT : CCCCCDCDDD\ldots$$

For $\bar{m} \to \infty,\ E(TFT, TFT) = \bar{m}(\frac{R+P+T+S}{4}) < \bar{m}R$

Consequence

TFT becomes vulnerable to *cooperative drift*.

1. "Generous" Tit-for-Tat (GTFT) forgives defections probabilistically. It cannot invade ALLD, but evolves from TFT in a cooperative environment as it is able to correct mistakes.

Consequence

TFT becomes vulnerable to *cooperative drift*.

1. "Generous" Tit-for-Tat (GTFT) forgives defections probabilistically. It cannot invade ALLD, but evolves from TFT in a cooperative environment as it is able to correct mistakes.

2. The probability of forgiveness begins to approach 1, as a greater degree of forgiveness beats forgiveness, in a cooperative environment. This weakens the barrier of protection against defective invasion. Until it falls.

Consequence

TFT becomes vulnerable to *cooperative drift*.

1. "Generous" Tit-for-Tat (GTFT) forgives defections probabilistically. It cannot invade ALLD, but evolves from TFT in a cooperative environment as it is able to correct mistakes.
2. The probability of forgiveness begins to approach 1, as a greater degree of forgiveness beats forgiveness, in a cooperative environment. This weakens the barrier of protection against defective invasion. Until it falls.
3. Defective strategies like ALLD take advantage of this forgiveness and cooperation fades.

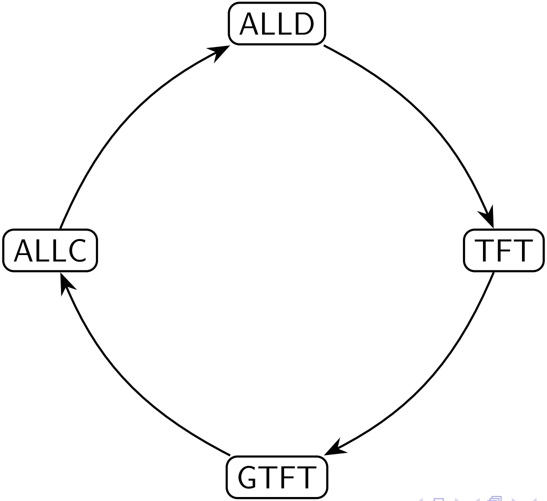
The Applicability of this Study

This natural sequence of events plays out in the financial markets.

1. In a prosperous economic environment, institutions with greater risk tolerance (cooperators) gain market share against more conservative peers. As a result, the system-wide degree of forgiveness (risk appetite) rises.
2. This makes the environment fragile. When a wave of defaults occurs, the overly forgiving institutions are unable to absorb the losses.
3. Post-collapse, risk-averse institutions (defectors) dominate. As market confidence returns, cautiously cooperative institutions (TFTs) begin outperforming defectors by raising risk tolerance.
4. The cycle repeats.

Result

The Chain of Evolution



Thank you!