The Mathematics of Being Nice

Cooperation in the Prisoner's Dilemma

Graeme Taylor

Edinburgh

October 11, 2007

The Prisoner's Dilemma

The Prisoner's Dilemma

- Two prisoners are independently interrogated by the police for a crime they are guilty of, which they may individually either confess to or deny.

The Prisoner's Dilemma

The Prisoner's Dilemma

- Two prisoners are independently interrogated by the police for a crime they are guilty of, which they may individually either confess to or deny.
- If neither confesses, they receive only shorter sentences of a year for other crimes.

The Prisoner's Dilemma

The Prisoner's Dilemma

- Two prisoners are independently interrogated by the police for a crime they are guilty of, which they may individually either confess to or deny.
- If neither confesses, they receive only shorter sentences of a year for other crimes.
- If both confess, each receives the standard sentence of three years.

The Prisoner's Dilemma

The Prisoner's Dilemma

- Two prisoners are independently interrogated by the police for a crime they are guilty of, which they may individually either confess to or deny.
- If neither confesses, they receive only shorter sentences of a year for other crimes.
- If both confess, each receives the standard sentence of three years.
- However, as an incentive to confess, a prisoner who confesses when the other denies will walk free, his partner in crime receiving a four year sentence.

Some Game Theory

Definition

A *two player game in strategic (or normal) form* X, Y, A, B consists of two strategy sets X and Y, corresponding to the players, and functions $A, B : X \times Y \to \mathbb{R}$ representing the *pay-off*. The game is *finite* if both X and Y are finite sets.

Some Game Theory

Definition

A *two player game in strategic (or normal) form* X, Y, A, B consists of two strategy sets X and Y, corresponding to the players, and functions $A, B : X \times Y \to \mathbb{R}$ representing the *pay-off*. The game is *finite* if both X and Y are finite sets.

Definition

A *play* of the game consists of Player 1 choosing a strategy $x \in X$ and Player 2 simultaneously choosing a strategy $y \in Y$.
Player 1 is then awarded $A(x, y)$ in winnings, and Player 2 awarded $B(x, y)$
If $B = -A$, i.e., Player 2 loses whatever Player 1 wins, then the game is *zero-sum*.

Some Game Theory

Definition

The *payoff matrices* for the game with $X = \{x_1, \ldots, x_m\}$, $Y = \{y_1, \ldots, y_n\}$ and payoff functions A, B are given by

$$\overset{\text{A:Player 1}}{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}} \quad \overset{\text{B:Player 2}}{\begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}}$$

Some Game Theory

Definition

The *payoff matrices* for the game with $X = \{x_1, \ldots, x_m\}$, $Y = \{y_1, \ldots, y_n\}$ and payoff functions A, B are given by

$$\begin{array}{cc} \text{A:Player 1} & \text{B:Player 2} \\ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} & \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} \end{array}$$

The $(i,j)^{th}$ entry of each matrix determines the winnings for the corresponding player when Player 1 chooses strategy $x_i \in X$ and Player 2 chooses strategy $y_j \in Y$. As a shorthand, we may describe Player 1 as choosing the row and Player 2 as choosing the column.

Some Game Theory

Definition

A more compact notation is the *bimatrix* form:

$$\begin{bmatrix} (a_{11}, b_{11}) & \cdots & (a_{1n}, b_{1n}) \\ \vdots & & \vdots \\ (a_{m1}, b_{m1}) & \cdots & (a_{mn}, b_{mn}) \end{bmatrix}$$

Now the $(i, j)^{th}$ entry describes the winnings for both players.

The Prisoner's Dilemma in Strategic Form

Conventions

- Each player in the Prisoner's Dilemma has the same strategy set: they may choose to *defect* or *cooperate*.

The Prisoner's Dilemma in Strategic Form

Conventions

- Each player in the Prisoner's Dilemma has the same strategy set: they may choose to *defect* or *cooperate*.
- cooperation is from the perspective of the other prisoner: i.e., defection means working with the police.

The Prisoner's Dilemma in Strategic Form

Conventions

- Each player in the Prisoner's Dilemma has the same strategy set: they may choose to *defect* or *cooperate*.
- cooperation is from the perspective of the other prisoner: i.e., defection means working with the police.
- We consider the payoffs to be the years of freedom over the next four years.

The Prisoner's Dilemma in Strategic Form

Prisoner's Dilemma

Given these conventions, the Prisoner's Dilemma is given by the 2×2 bimatrix

$$\begin{bmatrix} (1,1) & (4,0) \\ (0,4) & (3,3) \end{bmatrix}$$

The Prisoner's Dilemma in Strategic Form

Prisoner's Dilemma

Given these conventions, the Prisoner's Dilemma is given by the 2×2 bimatrix

$$\begin{bmatrix} (1,1) & (4,0) \\ (0,4) & (3,3) \end{bmatrix}$$

These correspond to the choices:

$$\begin{bmatrix} \text{Both confess} & \text{P1 confesses} \\ \text{P2 confesses} & \text{Neither confess} \end{bmatrix}$$

The Prisoner's Dilemma in Strategic Form

Prisoner's Dilemma

Given these conventions, the Prisoner's Dilemma is given by the 2×2 bimatrix

$$\begin{bmatrix} (1,1) & (4,0) \\ (0,4) & (3,3) \end{bmatrix}$$

These correspond to the choices:

$$\begin{bmatrix} \text{Both confess} & \text{P1 confesses} \\ \text{P2 confesses} & \text{Neither confess} \end{bmatrix}$$

i.e.,

$$\begin{bmatrix} \text{Mutual Defection} & \text{P1 Defect, P2 coop} \\ \text{P1 coop, P2 Defect} & \text{Mutual Cooperation} \end{bmatrix}$$

Cooperate or defect?

Both players prefer the mutual cooperation outcome (score 3 each) to the mutual defection outcome (score 1 each): this suggests they should cooperate. But each stands to score even more highly if they defect when the other cooperates (score 4 instead of 3). Worse, if they suspect their opponent will defect in this way, cooperation will lead to a 'sucker payoff' of 0.

Cooperate or defect?

Both players prefer the mutual cooperation outcome (score 3 each) to the mutual defection outcome (score 1 each): this suggests they should cooperate. But each stands to score even more highly if they defect when the other cooperates (score 4 instead of 3). Worse, if they suspect their opponent will defect in this way, cooperation will lead to a 'sucker payoff' of 0.

So which is the rational choice? We adopt the *non-cooperative* view of game theory, as advanced by John Nash: each player is unable to trust the other, and thus acts to prevent any risk of exploitation by an untrustworthy opponent. This motivates a definition of rationality as follows:

Nash Equilibrium

Definition

A pair of strategies $(\mathbf{p}, \mathbf{q}) \in X \times Y$ is a 2-player *Nash equilibrium* for the game given by strategic form X, Y, A, B if neither player gains by unilaterally deviating from the equilibrium.

Nash Equilibrium

Definition

A pair of strategies $(\mathbf{p}, \mathbf{q}) \in X \times Y$ is a 2-player *Nash equilibrium* for the game given by strategic form X, Y, A, B if neither player gains by unilaterally deviating from the equilibrium.

Theorem

(Nash's theorem) Any finite n-player game in strategic form has a Nash equilibrium.

Nash Equilibrium

Definition

A pair of strategies $(\mathbf{p}, \mathbf{q}) \in X \times Y$ is a 2-player *Nash equilibrium* for the game given by strategic form X, Y, A, B if neither player gains by unilaterally deviating from the equilibrium.

Theorem

(Nash's theorem) Any finite n-player game in strategic form has a Nash equilibrium.

Definition

If a game has a Nash equilibrium, then the *rational* behaviour for each player is to play their equilibrium strategy.

Mutual Defection is rational

Theorem

Mutual defection is a Nash equilibrium for the Prisoner's Dilemma.

Mutual Defection is rational

Theorem

Mutual defection is a Nash equilibrium for the Prisoner's Dilemma.

Proof.

Under mutual defection, each player scores 1. If either player wishes to deviate from this strategy pair, their only option is to move to cooperation. Then they receive a payoff of 0, and their opponent scores 4. Thus neither player benefits from an individual variation of their strategy. So mutual defection is a Nash equilibrium. □

Are people rational?

- A recent experiment with university students revealed that defection rates of non-economics majors was under 40%.

Are people rational?

- A recent experiment with university students revealed that defection rates of non-economics majors was under 40%.
- Economics majors defected 60% of the time in the standard game.

Are people rational?

- A recent experiment with university students revealed that defection rates of non-economics majors was under 40%.
- Economics majors defected 60% of the time in the standard game.
- When given the opportunity to make (non-binding) deals with the other participants before play, both categories dropped to a defection rate of around 30%.

Are people rational?

- A recent experiment with university students revealed that defection rates of non-economics majors was under 40%.
- Economics majors defected 60% of the time in the standard game.
- When given the opportunity to make (non-binding) deals with the other participants before play, both categories dropped to a defection rate of around 30%.
- So people cooperate far more than game theory predicts. How can we resolve these discrepancies?

Explaining Cooperation

- The model is right, but for Nash's theory of rationality, people are irrational!

Explaining Cooperation

- The model is right, but for Nash's theory of rationality, people are irrational!
- The theory is correct, but applied to the wrong model- additional factors alter the pay-off:

Explaining Cooperation

- The model is right, but for Nash's theory of rationality, people are irrational!
- The theory is correct, but applied to the wrong model- additional factors alter the pay-off:
  - Defection carries an additional cost: guilt, fear of later repercussions.
  - Cooperation carries an additional reward: evolutionary psychology / morality drives us to seek the best group outcome despite personal risk.

Explaining Cooperation

- The model is right, but for Nash's theory of rationality, people are irrational!
- The theory is correct, but applied to the wrong model- additional factors alter the pay-off:
  - Defection carries an additional cost: guilt, fear of later repercussions.
  - Cooperation carries an additional reward: evolutionary psychology / morality drives us to seek the best group outcome despite personal risk.
  - Result is a payoff matrix where mutual defection is no longer the equilibrium; then Nash's model of rationality is compatible with cooperation.

Recovering Cooperation

A simple change of the game fundamentally alters rational behaviour, and hence offers some clue as to motives that may influence choices in the standard Prisoner's Dilemma. By playing a larger game, the *Iterated Prisoner's Dilemma*, consisting of several rounds of the Prisoner's Dilemma, actions in a given round will have repercussions in future play and hence for your long-term score. Thus participants have an incentive to cooperate early to build trust and benefit from mutual cooperation later, and thus to cooperate in any single play of the Prisoner's Dilemma.

The Tit-for-Tat Strategy

In fact, a strategy along this line turns out to offer greater expected payoff than persistent defection; cooperation *can* emerge from a non-cooperative game.

The Tit-for-Tat Strategy

In fact, a strategy along this line turns out to offer greater expected payoff than persistent defection; cooperation *can* emerge from a non-cooperative game.

Definition

In an iterated game in which each round consists of a play of the prisoners' dilemma against the same opponent, the *Tit-for-Tat* strategy is:

- In round 1, cooperate.
- In round n for $n \geq 2$, play your opponent's strategy from round $n - 1$.

Evolutionary Stable Strategies

- In analogy to Nash Equilibria for strategic form games, a strategy for an iterated game is *evolutionary stable* if in a population of such strategies, defection to a new strategy reduces your score.

Evolutionary Stable Strategies

- In analogy to Nash Equilibria for strategic form games, a strategy for an iterated game is *evolutionary stable* if in a population of such strategies, defection to a new strategy reduces your score.
- In the IPD, the evolutionary stable strategy is still **always defect**.

Evolutionary Stable Strategies

- In analogy to Nash Equilibria for strategic form games, a strategy for an iterated game is *evolutionary stable* if in a population of such strategies, defection to a new strategy reduces your score.
- In the IPD, the evolutionary stable strategy is still **always defect**.
- However, evolutionary stability is not the same as 'best'/rational!

Evolutionary Stable Strategies

- In analogy to Nash Equilibria for strategic form games, a strategy for an iterated game is *evolutionary stable* if in a population of such strategies, defection to a new strategy reduces your score.
- In the IPD, the evolutionary stable strategy is still **always defect**.
- However, evolutionary stability is not the same as 'best'/rational!
- For instance, **Tit-for-Tat** will lose individual rounds to **always defect**, but will score higher overall in a tournament, provided it gets to play against itself.

Niceness

- **Tit-for-Tat** will never defect first.
- A strategy with this property is described as being *nice*.

Niceness

- **Tit-for-Tat** will never defect first.
- A strategy with this property is described as being *nice*.
- Two nice strategies will both score the maximum possible when they play against each other, settling into a virtuous circle of mutual cooperation.

Niceness

- **Tit-for-Tat** will never defect first.
- A strategy with this property is described as being *nice*.
- Two nice strategies will both score the maximum possible when they play against each other, settling into a virtuous circle of mutual cooperation.
- But niceness alone is not enough- **always cooperate** is the nicest possible strategy, but is easily exploited by opponents which are not nice.

Retaliation and Forgiveness

- **Tit-for-Tat** will always punish any defection by its opponent, by ceasing to cooperate.

Retaliation and Forgiveness

- **Tit-for-Tat** will always punish any defection by its opponent, by ceasing to cooperate.
- However, the opponent can regain the trust of **Tit-for-Tat** and re-establish mutual cooperation within a single turn, by switching back to cooperation. This property is described as *forgiveness*.

Retaliation and Forgiveness

- **Tit-for-Tat** will always punish any defection by its opponent, by ceasing to cooperate.
- However, the opponent can regain the trust of **Tit-for-Tat** and re-establish mutual cooperation within a single turn, by switching back to cooperation. This property is described as *forgiveness*.
- Forgiveness prevents excessive retaliation or revenge-seeking, thus maximising the potential for high scoring through mutual cooperation.

Retaliation and Forgiveness

- **Tit-for-Tat** will always punish any defection by its opponent, by ceasing to cooperate.
- However, the opponent can regain the trust of **Tit-for-Tat** and re-establish mutual cooperation within a single turn, by switching back to cooperation. This property is described as *forgiveness*.
- Forgiveness prevents excessive retaliation or revenge-seeking, thus maximising the potential for high scoring through mutual cooperation.
- **Tit-for-Tat** is the most forgiving strategy (other than **always cooperate**).

Trigger strategies

Tit-for-Tat's main vulnerability is to **random** strategies; nonetheless, it won the first IPD tournaments in the early 80s, and dominated for the following twenty years. It is the simplest example of a *trigger* strategy, as well as the simplest nice, forgiving strategy. Numerous variations have been proposed, however:

Trigger strategies

Tit-for-Tat's main vulnerability is to **random** strategies; nonetheless, it won the first IPD tournaments in the early 80s, and dominated for the following twenty years. It is the simplest example of a *trigger* strategy, as well as the simplest nice, forgiving strategy. Numerous variations have been proposed, however:

Example

The **Tit-for-Two-Tats** strategy, which defects only after two consecutive defections by an opponent, would have scored more highly against the population of the very first competition: but despite being sent as an example strategy to the participants, no-one entered it!

Trigger strategies

Tit-for-Tat's main vulnerability is to **random** strategies; nonetheless, it won the first IPD tournaments in the early 80s, and dominated for the following twenty years. It is the simplest example of a *trigger* strategy, as well as the simplest nice, forgiving strategy. Numerous variations have been proposed, however:

Example

The **Tit-for-Two-Tats** strategy, which defects only after two consecutive defections by an opponent, would have scored more highly against the population of the very first competition: but despite being sent as an example strategy to the participants, no-one entered it!

Example

The **Grim strategy** couples niceness with a complete lack of forgiveness: the trigger of a single defection forces it to defect in all future rounds. However, this can score more highly against **random** depending on the probability of defection.

Beyond Tit-for-Tat

A 20th anniversary tournament attracted 223 entries: and the top three places all went to programs from the University of Southampton. How did they beat **Tit-for-Tat**?

Beyond Tit-for-Tat

A 20th anniversary tournament attracted 223 entries: and the top three places all went to programs from the University of Southampton. How did they beat **Tit-for-Tat**?

- Teams were allowed to enter multiple programs.

Beyond Tit-for-Tat

A 20th anniversary tournament attracted 223 entries: and the top three places all went to programs from the University of Southampton. How did they beat **Tit-for-Tat**?

- Teams were allowed to enter multiple programs.
- **Master** and **slave** programs were designed to make a pre-set sequence of moves for the first few rounds.

Beyond Tit-for-Tat

A 20th anniversary tournament attracted 223 entries: and the top three places all went to programs from the University of Southampton. How did they beat **Tit-for-Tat**?

- Teams were allowed to enter multiple programs.
- **Master** and **slave** programs were designed to make a pre-set sequence of moves for the first few rounds.
- If a suitable sequence was identified, **Slave** programs would switch to **always cooperate** and **Master** programs to **always defect**.

Beyond Tit-for-Tat

A 20th anniversary tournament attracted 223 entries: and the top three places all went to programs from the University of Southampton. How did they beat **Tit-for-Tat**?

- Teams were allowed to enter multiple programs.
- **Master** and **slave** programs were designed to make a pre-set sequence of moves for the first few rounds.
- If a suitable sequence was identified, **Slave** programs would switch to **always cooperate** and **Master** programs to **always defect**.
- **Master** programs climbed to top of leaderboard: but at the expense of **Slave** programs which sunk to the bottom- average performance was worse than **Tit-for-Tat**.

Lessons Learnt

- Be *nice*...

Lessons Learnt

- Be *nice*...
- ...but only as long as everyone else is.

Lessons Learnt

- Be *nice*...
- ...but only as long as everyone else is.
- Punish betrayal relentlessly...

Lessons Learnt

- Be *nice*...
- ...but only as long as everyone else is.
- Punish betrayal relentlessly...
- ...until they see the error of their ways, then show *forgiveness*.

Lessons Learnt

- Be *nice*...
- ...but only as long as everyone else is.
- Punish betrayal relentlessly...
- ...until they see the error of their ways, then show *forgiveness*.
- But to really get ahead, you need a fall guy!

Talking of falling...

Figure: The Forth Rail Bridge

Help raise money for NCH by sponsoring me to abseil off the Forth Bridge!

Thanks!

Proofs, further reading, and MATLAB code to run your own tournaments can be found at

http://maths.straylight.co.uk