Prisoner's Dilemma

Theme and Variations

A.C.

January 21, 2025

Outline

Nash Equilibrium

We say that players in a non-cooperative game are in a **Nash Equilibrium** if neither can improve their score by adopting a new strategy.

Nash Equilibrium

We say that players in a non-cooperative game are in a **Nash Equilibrium** if neither can improve their score by adopting a new strategy.

- Players can communicate. Their strategies are assumed transparent to each other (e.g., by declaration).

Nash Equilibrium

We say that players in a non-cooperative game are in a **Nash Equilibrium** if neither can improve their score by adopting a new strategy.

- Players can communicate. Their strategies are assumed transparent to each other (e.g., by declaration).
- Players needn't cooperate. It's still a Nash Equilibrium if both players changing strategy *together* yields improvement.

Prisoner's Dilemma

	Cooperate	Defect
Cooperate	3, 3	0, 5
Defect	5, 0	1, 1

- Played as a **stage game**; both players act at the same time

Prisoner's Dilemma

	Cooperate	Defect
Cooperate	$3, 3$	$0, 5$
Defect	$5, 0$	$1, 1$

- Played as a **stage game**; both players act at the same time
- Nash equilibrium on D-D

Prisoner's Dilemma

	Cooperate	Defect
Cooperate	3, 3	0, 5
Defect	5, 0	1, 1

- Played as a **stage game**; both players act at the same time
- Nash equilibrium on D-D
- Optimal move is D, independent of opponent choice

Homo Economicus

Iterated Prisoner's Dilemma

What happens if we play several rounds of prisoner's dilemma? A tension between developing good will (cooperating) and spending it (defecting)?

Simple rules:

- N rounds of play.
- Same reward-space in every round.
- Final score is sum of individual rounds.
- Every player has perfect recall.

Rational Play

- Final round reduces to plain PD;
  - D-D

Rational Play

- Final round reduces to plain PD;
  - D-D
- Penultimate round reduces to plain PD, followed by D-D.
  - D-D, D-D

Rational Play

- Final round reduces to plain PD;
  - D-D
- Penultimate round reduces to plain PD, followed by D-D.
  - D-D, D-D
- round before that reduces to plain PD, followed by D-D, D-D.
  - D-D, D-D, D-D

Rational Play

- Final round reduces to plain PD;
  - D-D
- Penultimate round reduces to plain PD, followed by D-D.
  - D-D, D-D
- round before that reduces to plain PD, followed by D-D, D-D.
  - D-D, D-D, D-D
- . . .

Rational Play

- Final round reduces to plain PD;
  - D-D
- Penultimate round reduces to plain PD, followed by D-D.
  - D-D, D-D
- round before that reduces to plain PD, followed by D-D, D-D.
  - D-D, D-D, D-D
- ...
- ...

Rational Play

- Final round reduces to plain PD;
  - D-D
- Penultimate round reduces to plain PD, followed by D-D.
  - D-D, D-D
- round before that reduces to plain PD, followed by D-D, D-D.
  - D-D, D-D, D-D
- ...
- ...
- D-D,D-D,D-D,...,D-D

Greater Fools

- Iterated PD still yields universal defection under optimal, rational play.

Greater Fools

- Iterated PD still yields universal defection under optimal, rational play.
- Irrational partners could cooperate.
  - We might want to (sometimes) cooperate as well, depending on partner.

Infinite Round Prisoner's Dilemma

What happens if we play forever?

- ∞ rounds of play.
- Same reward-space in every round.
- Every player has perfect recall.
- Final score is the *discounted* sum of individual rounds.
  - $s = \sum \gamma^n s_n$ for a fixed discount factor $\gamma \in [0, 1)$

Infinite Round Prisoner's Dilemma

What happens if we play forever?

- ∞ rounds of play.
- Same reward-space in every round.
- Every player has perfect recall.
- Final score is the *discounted* sum of individual rounds.
  - $s = \sum \gamma^n s_n$ for a fixed discount factor $\gamma \in [0, 1)$
  - Encodes a bird-in-the-hand sensibility, or stochastic early stopping

Infinite Round Prisoner's Dilemma

What happens if we play forever?

- ∞ rounds of play.
- Same reward-space in every round.
- Every player has perfect recall.
- Final score is the *discounted* sum of individual rounds.
  - $s = \sum \gamma^n s_n$ for a fixed discount factor $\gamma \in [0, 1)$
  - Encodes a bird-in-the-hand sensibility, or stochastic early stopping
  - Also, it makes the sum converge.

Defect Always

- Let player A plan to always defect.

Defect Always

- Let player A plan to always defect.
- Player B obviously has no incentive to cooperate, will also plan to defect.

Defect Always

- Let player A plan to always defect.
- Player B obviously has no incentive to cooperate, will also plan to defect.
- By symmetry, A also has no incentive to cooperate.

Defect Always

- Let player A plan to always defect.
- Player B obviously has no incentive to cooperate, will also plan to defect.
- By symmetry, A also has no incentive to cooperate.

Defect-always is still *an* equilibrium.

Grim Trigger

- Let players A and B plan to cooperate until defect.
  - After first defection, each player plans to revert to defect-always.

Grim Trigger

- Let players A and B plan to cooperate until defect.
  - After first defection, each player plans to revert to defect-always.
- First defector scores a C-D bonus, but all future rounds will be D-D.

Grim Trigger

- Let players A and B plan to cooperate until defect.
  - After first defection, each player plans to revert to defect-always.
- First defector scores a C-D bonus, but all future rounds will be D-D.
- If γ is large enough, nobody will ever want to defect.
  - Can apply OSDP if more rigor is desired.

Grim Trigger

- Let players A and B plan to cooperate until defect.
  - After first defection, each player plans to revert to defect-always.
- First defector scores a C-D bonus, but all future rounds will be D-D.
- If γ is large enough, nobody will ever want to defect.
  - Can apply OSDP if more rigor is desired.

Grim Trigger is an always-cooperating equilibrium.

Credibility

Cooperation in ∞-PD possible only with credible retaliation.

Credibility

Cooperation in ∞-PD possible only with credible retaliation.

- Strategies are immutable.
  - Or at least, we have no incentive to mutate our strategies.
- Strategies are published.

Credibility

Cooperation in ∞-PD possible only with credible retaliation.

- Strategies are immutable.
  - Or at least, we have no incentive to mutate our strategies.
- Strategies are published.

How realistic is that, anyways?

Credibility II: Roko's Basilisk

Credibility III: Declaratory Strategy

Outline

Tournament Format

Practical experiment run in 1980 by Robert Axelrod.

- Field of 14 prisoner's dilemma actors.
- Opponent strategy was *not* announced.
  - A few actors tried to infer opposing strategies from observed play.
- Each actor plays every other actor, plus a `Random` round and a mirror round.
  - 200 round iterated PD per matchup
- Actor with the highest average score wins.

DOWNING

- Probably most sophisticated entrant.

DOWNING

- Probably most sophisticated entrant.
- Maintains a belief score that cooperation will be punished
  - Defects if it believes it can get away with it.
  - Cooperates otherwise

DOWNING

- Probably most sophisticated entrant.
- Maintains a belief score that cooperation will be punished
  - Defects if it believes it can get away with it.
  - Cooperates otherwise
- Defects for first two rounds (minimum), to discover if it can extort

DOWNING

- Probably most sophisticated entrant.
- Maintains a belief score that cooperation will be punished
  - Defects if it believes it can get away with it.
  - Cooperates otherwise
- Defects for first two rounds (minimum), to discover if it can extort
- Rank 10 with an average score of 391.

Grim Trigger

```
procedure FRIEDMAN(history)
    if D ∈ history then
        return D
    return C
```

- Co-operates well with other "Nice" strategies.

Grim Trigger

```
procedure FRIEDMAN(history)
    if D ∈ history then
        return D
    return C
```

- Co-operates well with other "Nice" strategies.
- Over-defends itself against extortion strategies.

Grim Trigger

procedure FRIEDMAN(history)
    if $D \in$ history **then**
        return D
    return C

- Co-operates well with other "Nice" strategies.
- Over-defends itself against extortion strategies.
- Rank 7 with an average score of 473.

Grim Trigger

procedure FRIEDMAN(history)
  if $D \in$ history **then**
    return D
  return C

- Co-operates well with other "Nice" strategies.
- Over-defends itself against extortion strategies.
- Rank 7 with an average score of 473.
- Scores poorly without the benefit of a declaratory strategy.

Tit-for-Tat

procedure TitForTat(history)
  if history[-1] $=$ D **then**
    return D
  return C

- Co-operates well with other "Nice" strategies.

Tit-for-Tat

procedure TitForTat(history)
  if history[-1] = D **then**
    return D
  return C

- Co-operates well with other "Nice" strategies.
- Effectively defends itself against extortion strategies.

Tit-for-Tat

```
procedure TitForTat(history)
    if history[-1] = D then
        return D
    return C
```

- Co-operates well with other "Nice" strategies.
- Effectively defends itself against extortion strategies.
- Won Axelrod's tournament with an average score of 504.

Tournament Results

The key features of good performance in Axelrod's tournament were

1. Niceness: early cooperation paid enormous dividends.
2. Forgiveness: reverting to a cooperative posture allowed "resets" after a limited defection.

Obvious question: why is this so out of line with our theoretical, defection-happy discussions of PD?

Founder Effects

Axelrod's tournament took place within a context, or meta-game.

Founder Effects

Axelrod's tournament took place within a context, or meta-game.

Huge "nice" population (8 of 12 strategies) reinforced itself, out-competing extortionist strategies.

Founder Effects

Axelrod's tournament took place within a context, or meta-game.

Huge "nice" population (8 of 12 strategies) reinforced itself, out-competing extortionist strategies.

Nice, forgiving population produces a **founder effect**.

The Red Queen

The goal of Axelrod's tournament is not to maximize score, it's to maximize probability of winning (or surviving).

Mixed Signals

Problem variant: Cooperate/Defect decision is randomly flipped with some probability p.

What happens?

TFT

C-C

TFT

C-C, C-C

TFT

C-C, C-C, C-C

TFT

C-C, C-C, C-C, D-C

TFT

C-C, C-C, C-C, D-C, C-D

TFT

C-C, C-C, C-C, D-C, C-D,D-C

TFT

C-C, C-C, C-C, D-C, C-D,D-C,. . .

TFT

C-C, C-C, C-C, D-C, C-D,D-C,...

- Cooperate-then-defect chain continues until an error or error(s) puts us into C-C again.

TFT

C-C, C-C, C-C, D-C, C-D,D-C,. . .

- Cooperate-then-defect chain continues until an error or error(s) puts us into C-C again.
- Harder to fall out of the non-cooperation loop than into it.

TF2T

Simple robustification strategy: retaliate only on two defects:

TF2T

Simple robustification strategy: retaliate only on two defects: C-C

TF2T

Simple robustification strategy: retaliate only on two defects: C-C, C-C

TF2T

Simple robustification strategy: retaliate only on two defects:
C-C, C-C, C-C

TF2T

Simple robustification strategy: retaliate only on two defects:
C-C, C-C, C-C, D-C

TF2T

Simple robustification strategy: retaliate only on two defects:
C-C, C-C, C-C, D-C, C-C

TF2T

Simple robustification strategy: retaliate only on two defects:
C-C, C-C, C-C, D-C, C-C, C-C

TF2T

Simple robustification strategy: retaliate only on two defects:
C-C, C-C, C-C, D-C, C-C, C-C,. . .

- Very difficult to fall into accidental conflict.

TF2T

Simple robustification strategy: retaliate only on two defects:
C-C, C-C, C-C, D-C, C-C, C-C,. . .

- Very difficult to fall into accidental conflict.
- No harder to fall out of it.

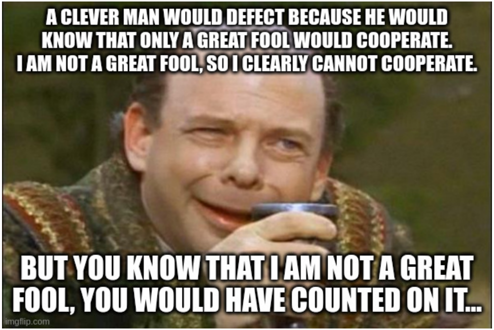
TF2T

Simple robustification strategy: retaliate only on two defects:
C-C, C-C, C-C, D-C, C-C, C-C,. . .

- Very difficult to fall into accidental conflict.
- No harder to fall out of it.
- Can be exploited by “predatory” strategies.

Outline

Super-Rationality



Mutual Interest

$$R(\text{player}) \to R(\text{player}) + \alpha R(\text{other})$$

- For a sufficiently large α, cooperation becomes obvious.

Mutual Interest

$$R(\text{player}) \to R(\text{player}) + \alpha R(\text{other})$$

- For a sufficiently large α, cooperation becomes obvious.
- *Common interest is common*.

  See, e.g., kin selection.

Fin

The source for this presentation is hosted at
`https://github.com/alan-christopher/edu`.

Questions?