

Game Theory

7. Repeated Games

Albert-Ludwigs-Universität Freiburg



**UNI
FREIBURG**

Bernhard Nebel and Robert Mattmüller

Summer semester 2019

1 Motivation



Motivation

Repeated
Prisoners
Dilemma

Strategies
and
Preferences
in Infinite
Games

Analysis of
Infinite PD

Punishments
and
Enforceable
Outcomes

- Remember: The *Prisoner's Dilemma* leads to the unsatisfying result because there is neither experience nor future encounters.
- What if the game is played repeatedly?
- Model this as an *extensive game* where in each turn, we repeat a given base game.
- Will *social norms* evolve?
- Will punishments, which can lead to short-term costs, nevertheless be played (are these potential punishments credible threats?)

See: <http://ncase.me/trust/>

Motivation

Repeated
Prisoners
Dilemma

Strategies
and
Preferences
in Infinite
Games

Analysis of
Infinite PD

Punishments
and
Enforceable
Outcomes

2 Repeated Prisoners Dilemma



Motivation

**Repeated
Prisoners
Dilemma**

Strategies
and
Preferences
in Infinite
Games

Analysis of
Infinite PD

Punishments
and
Enforceable
Outcomes

| | <i>C</i> | <i>D</i> |
|----------|----------|----------|
| <i>C</i> | 3,3 | 0,4 |
| <i>D</i> | 4,0 | 1,1 |

(D, D) (i.e. both players defect) is the unique Nash equilibrium, the pair of maximinimizers and the pair of strictly dominating strategies.

So, in a single encounter, there is no argument for rationally playing *C*!

Motivation

Repeated
Prisoners
Dilemma

Strategies
and
Preferences
in Infinite
Games

Analysis of
Infinite PD

Punishments
and
Enforceable
Outcomes

- Assume we play PD a pre-specified number of times k .
- What will be a subgame perfect equilibrium?
- Use backward induction:
 - (D,D) is the NE in the last subgame, since this would be the only NE in the one-shot game.
 - So, (D,D) will also be played in period $k - 1$.
 - ...
 - So, the (only) subgame-perfect equilibrium and the only NE of this repeated game is (D,D), (D,D), ..., (D,D).

Motivation

Repeated
Prisoners
Dilemma

Strategies
and
Preferences
in Infinite
Games

Analysis of
Infinite PD

Punishments
and
Enforceable
Outcomes

- If we play PD infinitely often, we need to solve two problems:
 - 1 How to define a **strategy**?
 - 2 How to define the **payoff** or **preference**?
- How to specify a strategy using only finite resources?
 - In general: One could use an *algorithm*.
 - Usually done in game theory: use **Moore automata**, i.e., finite state automata, where the inputs are actions of the other players, and in each state, a response action to the previous actions is generated.
 - A Nash equilibrium would then be a profile of automata (strategies) such that no deviation would be profitable.

Motivation

Repeated
Prisoners
Dilemma

Strategies
and
Preferences
in Infinite
Games

Analysis of
Infinite PD

Punishments
and
Enforceable
Outcomes

3 Strategies and Preferences in Infinite Games



**UNI
FREIBURG**

Motivation

Repeated
Prisoners
Dilemma

**Strategies
and
Preferences
in Infinite
Games**

Analysis of
Infinite PD

Punishments
and
Enforceable
Outcomes

Using Moore automata, we can specify what to do in response to the new input (action played by others) and the state we are in. Since the automata are finite, we have only finite memory!

- Unconditionally *cooperative*: Always play C.
- Unconditionally *uncooperative*: Always play D.
- *Bipolar*: Start with *D* and then always exchange between C and D.
- *Tit-for-Tat*: Start with *C* and then reply with *C* to each *C* and *D* to each *D*.
- *Grim*: Start with *C*. After any play of *D* play *D* in the future forever.

Motivation

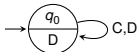
Repeated Prisoners Dilemma

Strategies and Preferences in Infinite Games

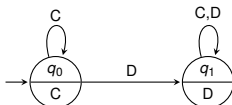
Analysis of Infinite PD

Punishments and Enforceable Outcomes

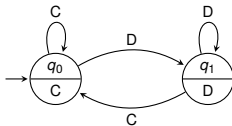
Uncooperative



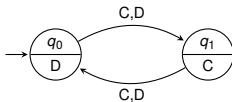
Grim



Tit-for-tat



Bipolar



Motivation

Repeated
Prisoners
Dilemma

Strategies
and
Preferences
in Infinite
Games

Analysis of
Infinite PD

Punishments
and
Enforceable
Outcomes

Player 1 plays Tit-for-tat, player 2 plays bipolar; 4 rounds.



| Round | Action | Utility | Accumulated payoff |
|-------|--------|---------|--------------------|
| 1 | (C,D) | (0,4) | (0,4) |
| 2 | (D,C) | (4,0) | (4,4) |
| 3 | (C,D) | (0,4) | (4,8) |
| 4 | (D,C) | (4,0) | (8,8) |

Motivation

Repeated
Prisoners
Dilemma

Strategies
and
Preferences
in Infinite
Games

Analysis of
Infinite PD

Punishments
and
Enforceable
Outcomes

Preferences over payoffs (1)

How to define the payoff of an infinite game or whether to **prefer** one outcome over another one?

Given two infinite sequences $(v_i^t)_{t=1}^\infty$ and $(w_i^t)_{t=1}^\infty$, we will define when the first is preferred over the second by player i :

$$(v_i^t)_{t=1}^\infty \succsim_i (w_i^t)_{t=1}^\infty$$

- A common method is **discounting** by a *discount factor* $\delta \in (0, 1)$:

$$(v_i^t)_{t=1}^\infty \succsim_i (w_i^t)_{t=1}^\infty \quad \text{iff} \quad \sum_{t=1}^{\infty} \delta^{t-1} (v_i^t - w_i^t) \geq 0$$

- **Overtaking** is an alternative (every period counts the same):

$$(v_i^t)_{t=1}^\infty \succsim_i (w_i^t)_{t=1}^\infty \quad \text{iff} \quad \liminf_{T \rightarrow \infty} \sum_{t=1}^T (v_i^t - w_i^t) \geq 0$$

Motivation

Repeated
Prisoners
Dilemma

Strategies
and
Preferences
in Infinite
Games

Analysis of
Infinite PD

Punishments
and
Enforceable
Outcomes

- Instead one could require that on average the payoff is better, which leads to the **limit of means** criterion:

$$(v_i^t)_{t=1}^\infty \succsim_i (w_i^t)_{t=1}^\infty \quad \text{iff} \quad \liminf_{T \rightarrow \infty} \sum_{t=1}^T (v_i^t - w_i^t) / T \geq 0$$

- In this case, $\lim_{T \rightarrow \infty} \sum_{t=1}^T v_i^t / T$ would be considered the **payoff**, if the limit exists.

Motivation

Repeated
Prisoners
Dilemma

Strategies
and
Preferences
in Infinite
Games

Analysis of
Infinite PD

Punishments
and
Enforceable
Outcomes

- $(1, -1, 0, 0, \dots) \succ (0, 0, 0, 0, \dots)$ for any $\delta \in (0, 1)$ under *discounting*, but not under *limit of means* and *overtaking*.
- $(-1, 2, 0, 0, \dots) \succ (0, 0, 0, 0, \dots)$ under *overtaking*, but for the *limit of means* there is no preference.
- $(0, \dots, 0, 1, 1, 1, \dots) \succ (1, 0, 0, 0, \dots)$, where in the first sequence after m zeros there are only ones, under the *limit of means*. However, for every δ there exists an m^* such that for all $m > m^*$ the preference under *discounting* is the other way around.

Motivation

Repeated
Prisoners
Dilemma

Strategies
and
Preferences
in Infinite
Games

Analysis of
Infinite PD

Punishments
and
Enforceable
Outcomes

Definition (Infinitely repeated game of G)

Let $G = \langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$ be a strategic game with $A = \prod_{i \in N} A_i$. An *infinitely repeated game of G* is an extensive game with perfect information and simultaneous moves $\langle N, H, P, (\succsim_i)_{i \in N} \rangle$ in which

- $H = \{ \langle \rangle \} \cup (\bigcup_{t=1}^{\infty} A^t) \cup A^{\infty}$,
- $P(h) = N$ for all nonterminal histories $h \in H$,
- \succsim_i is a preference relation on A^{∞} that is based on *discounting, limit of means or overtaking*.

Motivation

Repeated
Prisoners
Dilemma

Strategies
and
Preferences
in Infinite
Games

Analysis of
Infinite PD

Punishments
and
Enforceable
Outcomes

4 Analysis of Infinite PD



**UNI
FREIBURG**

Motivation

Repeated
Prisoners
Dilemma

Strategies
and
Preferences
in Infinite
Games

**Analysis of
Infinite PD**

Punishments
and
Enforceable
Outcomes



Let us consider the **Grim** strategy. Is (Grim, Grim) a Nash equilibrium under all the preference criteria?

- *Overtaking and limit of means*: Yes! Any deviation will result in getting ≤ 1 instead of 3 infinitely often.
- *Discounting*: This is a bit more complicated.
 - If a player gets v for every round, he will accumulate the following payoff:

$$v + \delta v + \delta^2 v + \dots = \sum_{i=1}^{\infty} \delta^{i-1} v.$$

- Since we know that $\sum_{i=0}^{\infty} x^i = \frac{1}{1-x}$ (for $0 < x < 1$), we have:

$$\sum_{i=1}^{\infty} \delta^{i-1} v = v \sum_{i=0}^{\infty} \delta^i = \frac{v}{1-\delta}.$$

Motivation

Repeated
Prisoners
Dilemma

Strategies
and
Preferences
in Infinite
Games

Analysis of
Infinite PD

Punishments
and
Enforceable
Outcomes



- Assume now, both players play *Grim* for the first k rounds. Then player 1 deviates and plays D once. In the remainder he must play D in order to get at least 1 in each round.
 - Starting in round k , he receives:

$$4 + \delta + \delta^2 + \dots = 3 + \sum_{i=0}^{\infty} \delta^i = 3 + \frac{1}{1 - \delta}$$

- If he had not deviated, the accumulated payoff starting at round k would have been:

$$3 + 3\delta + 3\delta^2 + \dots = 3 \sum_{i=0}^{\infty} \delta^i = \frac{3}{1 - \delta}.$$

- So, a deviation is profitable if $3 + 1/(1 - \delta) > 3/(1 - \delta)$.
- When $\delta < 1/3$ the deviation is profitable, i.e., for $\delta \geq 1/3$, *Grim* is a NE strategy.

Motivation

Repeated Prisoners Dilemma

Strategies and Preferences in Infinite Games

Analysis of Infinite PD

Punishments and Enforceable Outcomes



- Under which preference criteria is *Tit-for-tat* an equilibrium strategy?
- *Limit of means*: Finitely many deviations do not change the payoff profile in the limit. Infinitely many deviations lead to less payoff. So Tit-for-tat is an NE strategy under this preference criterion.
- *Overtaking*: Even only one deviation leads to a payoff of 5 over two rounds instead of 6. So in no case, a deviation can lead to a better payoff.
- *Discounting*: Deviating only in one move in round k and then returning to be cooperative leads in round k to $4 + 0 + \dots$ instead to $3 + \delta 3 + \dots$
 - The deviation is profitable if $4 > 3 + \delta 3$.
 - This means, we only get a profitable deviation if $\delta < 1/3$; and this is the best case for a deviation!
 - So *Tit-for-tat* is NE if $\delta \geq 1/3$.

Motivation

Repeated Prisoners Dilemma

Strategies and Preferences in Infinite Games

Analysis of Infinite PD

Punishments and Enforceable Outcomes

5 Punishments and Enforceable Outcomes



**UNI
FREIBURG**

Motivation

Repeated
Prisoners
Dilemma

Strategies
and
Preferences
in Infinite
Games

Analysis of
Infinite PD

**Punishments
and
Enforceable
Outcomes**

- Observe that the NE strategies are based on being able to **punish** a deviating player.

Definition (Minmax payoff)

Let $G = \langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$ a strategic game. Player i 's **minimax payoff** in G , also written as $v_i(G)$, is the lowest payoff that the other players can force upon player i :

$$v_i(G) = \min_{a_{-i} \in A_{-i}} \max_{a_i \in A_i} u_i(a_{-i}, a_i).$$

- The idea is that the other players all punish a deviating player in the next round(s) and allow him only to get v_i .

Motivation

Repeated
Prisoners
Dilemma

Strategies
and
Preferences
in Infinite
Games

Analysis of
Infinite PD

Punishments
and
Enforceable
Outcomes



Definition (Feasible Payoff Profile)

Given a strategic game $G = \langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$, a vector $v \in \mathbb{R}^N$ is called **payoff profile** of G if there exists $a \in A$ such that $v = u(a)$. $v \in \mathbb{R}$ is called **feasible payoff profile** if there exists a vector $(\alpha_a)_{a \in A} \in \mathbb{Q}^A$ with $\sum \alpha_a = 1$ and $v = \sum \alpha_a u(a)$.

- **Note:** Such payoffs can be generated in a repeated game by playing β_a rounds a in a set of γ games with $\gamma = \sum_{a \in A} \beta_a$ and $\alpha_a = \beta_a / \gamma$.

Definition (Enforceable payoff)

Given a strategic game $G = \langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$, a payoff profile w with $w_i \geq v_i(G)$ for all $i \in N$ is called **enforceable**. If $w_i > v_i(G)$ for all $i \in N$, it is said to be **strictly enforceable**.

Motivation

Repeated
Prisoners
Dilemma

Strategies
and
Preferences
in Infinite
Games

Analysis of
Infinite PD

Punishments
and
Enforceable
Outcomes



- Using the concept of **enforceable payoffs**, one can construct many different Nash equilibria and payoff profiles for PD!
- E.g., $(31/3, 2)$ is a **feasible payoff profile**, because $4 \times (C, C)$ and $2 \times (D, C)$ leads to $((4 \times 3 + 2 \times 4)/6, (4 \times 3 + 2 \times 0)/6)$.
- Construct two automata that implement this repeated sequence and in case of deviation revert to playing D .
- These two automata implement Nash equilibria strategies, since deviating leads to a payoff of 1 instead of $31/3$ or 2!
- **Folk theorems** stating that all **enforceable outcomes** are reachable have been proven for the general case.

Motivation

Repeated
Prisoners
Dilemma

Strategies
and
Preferences
in Infinite
Games

Analysis of
Infinite PD

Punishments
and
Enforceable
Outcomes



- **Repeated games** are *extensive games* with simultaneous moves, in which a base strategic game is played in each round.
- Strategies are described using finite **Moore automata**.
- For preferences over the payoffs of infinitely repeated games, different **preference criteria** are possible.
- In the repeated *Prisoners Dilemma*, it is possible to play Nash Equilibrium strategies that result in (C,C) sequences.
- In fact, it is possible to achieve any possible **feasible payoff profile** in the **limit of means** criterion.

Motivation

Repeated
Prisoners
Dilemma

Strategies
and
Preferences
in Infinite
Games

Analysis of
Infinite PD

Punishments
and
Enforceable
Outcomes