UNIVERSITY OF OREGON

Repeated Games

Dante Yasui

Winter 2024

EC327 Game Theory

rewrite HW problem
like Godzilla

- So far, we have only seen games as either **one-shot** simultaneous or **finitely sequential**
- However, these representations can only do so much to represent the many complicated social interactions in which **repeated interactions** between the same players matter
- Specifically, in the Strategic Moves section, we discussed how reputation could play a role, but only allowed it to show up in the single-shot game through changing the payoffs

- In games with **finite** numbers of player actions, we can always use backwards induction to find equilibria
- But often players *do not know* when certain social interactions will end, and so it won't be reasonable to assume that they can backwards induct

- When games are repeated over time, we will use *discount rates* to represent how patient players are
- We can combine this with probability that a game will end at each stage in what we will call an effective rate of return

General Repeated Prisoners' Dilemma

		Column	
		Defect	Cooperate
Row	Defect	D,D	H,L
	Cooperate	L,H	C,C

What ordering of payoffs D, H, L, and C make this a **Prisoners' Dilemma**?

a) $C > D > H > L$

b) $H > D > C > L$

c) $H > C > D > L$

d) $C > H > L > D$

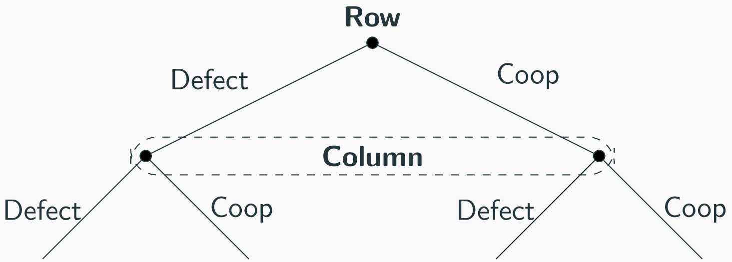
$H > L \Rightarrow$ incentive to cheat

$C > D \Rightarrow$ Pareto optimal C, C

$D > L \Rightarrow$ deviation hurts opponent

General Prisoners’ Dilemma

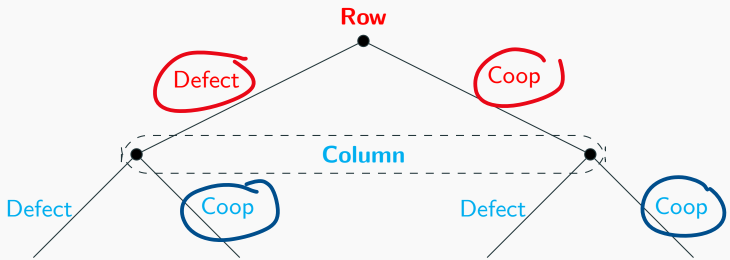
A single-stage prisoners’ dilemma in extensive form:



Row	D	H	L	C
Col	D	L	H	C

A **two-stage** prisoners' dilemma in mixed extensive form:

stage 1



2

	Def.	Coop.
Def.	$D+\delta D$, $D+\delta D$	$D+\delta H$, $D+\delta L$
Coop.	$D+\delta L$, $D+\delta H$	$D+\delta C$, $D+\delta C$

	Def.	Coop.
Def.	$H+\delta D$, $L+\delta D$	$H+\delta H$, $L+\delta L$
Coop.	$H+\delta L$, $L+\delta H$	$H+\delta C$, $L+\delta C$

	Def.	Coop.
Def.	$L+\delta D$, $H+\delta D$	$L+\delta H$, $H+\delta L$
Coop.	$L+\delta L$, $H+\delta H$	$L+\delta C$, $H+\delta C$

	Def.	Coop.
Def.	$C+\delta D$, $C+\delta D$	$C+\delta H$, $C+\delta L$
Coop.	$C+\delta L$, $C+\delta H$	$C+\delta C$, $C+\delta C$

$H+\delta D$ $\quad$ $C+\delta D$

Recall that δ is the subjective discount rate from stage to stage

Test Your Understanding

What should you do in the two-stage Prisoners' Dilemma if your opponent plays *Coop* in stage 1 and *Coop* in stage 2?

a) *Coop* in stage 1, *Coop* in stage 2

b) *Coop* in stage 1, *Defect* in stage 2

c) *Defect* in stage 1, *Coop* in stage 2

d) *Defect* in stage 1, *Defect* in stage 2

$\mathbb{T}$-stage repeated Prisoners' Dilemma

A complete strategy in a $\mathbb{T}$-stage repeated game will look like:

$$S^{\mathbb{T}}_{t=1} = \begin{cases} \text{In stage } t = 1 & \text{take action } A_0 \\ \text{In stage } t > 1 & \begin{cases} \text{If history so far was } h_t, \text{ take action } A_t(h_t) \\ \text{Else if history was } h'_t, \text{take action } A_t(h'_t) \\ \dots \end{cases} \end{cases}$$

We can see that the number of possible strategies increases exponentially as $\mathbb{T}$ gets larger

$\mathbb{T}$-stage repeated Prisoners' Dilemma

Suppose that $\mathbb{T}$ is a very large number, but we have played to the very last stage of a repeated Prisoners' Dilemma with that many stages:

		Column	
		Defect	Cooperate
Row	Defect	$\text{Tot.}^R + D$, $\text{Tot.}^C + D$	$\text{Tot.}^R + H$, $\text{Tot.}^C + L$
	Cooperate	$\text{Tot.}^R + L$, $\text{Tot.}^C + H$	$\text{Tot.}^R + C$, $\text{Tot.}^C + C$

Let Tot.^R and Tot.^C represent the total payoffs that both players have earned over stages 0 to $\mathbb{T} - 1$

$\mathbb{T}$-stage repeated Prisoners' Dilemma

$\mathbb{T}-1$

		Column	
		Defect	Cooperate
Row	Defect	Tot.$^{R}_{T-1} + D$, Tot.$^{C} + D$	Tot.$^{R} + H$, Tot.$^{C} + L$
	Cooperate	Tot.$^{R} + L$, Tot.$^{C} + H$	Tot.$^{R} + C$, Tot.$^{C} + C$

Notice that the equilibrium of this subgame is still *Defect*, *Defect* because Tot.^{R} and Tot.^{C} are already decided by prior actions.

Repeated Prisoners' Dilemma with Uncertain Second Stage

Suppose that the first stage of the game is the Trenches Game:

		German soldiers	
		Kill	Miss
Allied Soldiers	Kill	4, 4	8, 2
	Miss	2, 8	6, 6

But with probability p, the game repeats in the second round and with probability $1 - p$, it ends after the first round

Repeated Prisoners' Dilemma with Uncertain Second Stage

Consider the following strategy:

$$\begin{cases} \text{In stage } 1 & : \textit{Miss} \\ \text{In stage } 2 & : \begin{cases} \textit{Miss} \text{ if the other player Missed in stage 1} \\ \textit{Kill} \text{ if the other player Killed in stage 1} \end{cases} \end{cases}$$

Let's call this strategy *Punisher* because it starts off friendly, but will try to punish someone who defects in the first round by defecting in the second round.

Suppose you are playing against a *Punisher* in this game.

		German soldiers	
		Kill	Miss
Allied Soldiers	Kill	4, 4	8, 2
	Miss	2, 8	6, 6

- What is your **expected utility** of playing *Kill*, *Kill*?

$8 + 4p$

Repeated Prisoners' Dilemma with Uncertain Second Stage

Suppose you are playing against a *Punisher* in this game.

		German soldiers	
		Kill	Miss
Allied Soldiers	Kill	4, 4	8, 2
	Miss	2, 8	6, 6

What is your **expected utility** of playing *Miss, Kill*?

$\underline{6 + 8p}$

What about from playing *Kill, Miss?*

Kill, Kill

$8 + 2p \leq 8 + 4p$

Would you rather defect earlier or later?

Repeated Prisoners' Dilemma with Uncertain Second Stage

Suppose you are playing against a *Punisher* in this game.

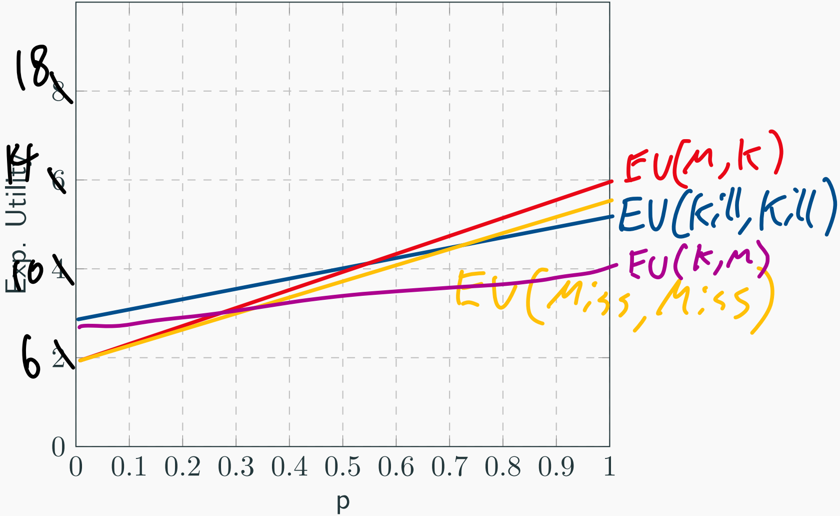
		German soldiers	
		Kill	Miss
Allied Soldiers	Kill	4, 4	8, 2
	Miss	2, 8	6, 6

What is your **expected utility** of playing *Miss, Miss*?

$6 + 6p$

Repeated Prisoners' Dilemma with Uncertain Second Stage

Let's graph your expected utilties of each strategy:



Repeated Prisoners' Dilemma with Uncertain Second Stage

Now you should be getting some of the intuition for how cooperative equilibria might be achieved.

- We need the payoffs of the last period to be uncertain (or never reached)
- If trying to cheat a *Punisher* or *Grim Trigger* strategy, it is better to start cheating them sooner rather than later
- In order for the equilibrium to have both players always cooperating, defecting in at least one period must not be a dominant strategy

Back to the General Form Prisoners' Dilemma

		Column	
		Defect	Cooperate
Row	Defect	D,D	H,L
	Cooperate	L,H	C,C

Now let's suppose that this game is repeated for an **infinite number of stages**.

Extending Plays to Infinity

Suppose the game is in the 'good' equilibrium where all players always play *Cooperate*.

- What is **present value** from this equilibrium?

$$
\begin{aligned}
pv\left(\{C\}_{t=1}^{\infty}\right) &= C + \delta C + \delta^2 C + \delta^3 C + \ldots \\
&= C \sum_{t=1}^{\infty} \delta^{t-1} \\
&= C \frac{1}{1-\delta}
\end{aligned}
$$

Extending Plays to Infinity

Let's extend the *Punisher* strategy we had from the two-stage game into the *Grim Trigger strategy* of the general infinite horizon game:

$$\begin{cases} \text{In stage } 1 & : \textit{Cooperate} \\ \text{In stage } t \geq 2 & : \begin{cases} \textit{Cooperate} \text{ if only cooperation has happened so far} \\ \textit{Defect} \text{ if anyone has } \textit{ever} \text{ Defected in the past} \end{cases} \end{cases}$$

Grim Trigger SPNE in Repeated PD

Is both players playing *Grim Trigger* stable?

- Does a player have an incentive to *Defect* against Grim Trigger:

$$pv(Always\ Coop) \geq pv(Defect\ once) \geq pv(Defect \geq 1)$$

$\delta C + \delta^2 C + \dots$

$$C + \delta C + \delta^2 C + \dots \geq H + \delta D + \delta^2 D + \dots$$

$$\underbrace{C}_{t=1} + C \sum_{t=2}^{\infty} \delta^t \geq \underbrace{H}_{t=1} + D \sum_{t=2}^{\infty} \delta^t$$

$$C + C\delta \sum_{t=2}^{\infty} \delta^{t-1} \geq H + D\delta \sum_{t=2}^{\infty} \delta^{t-1}$$

$$C + \frac{\delta C}{1-\delta} \geq H + \frac{\delta D}{1-\delta} \Rightarrow \frac{\delta C - \delta D}{1-\delta} \geq H - C$$

$$\delta \geq \frac{H-C}{H-D}$$

$$\frac{\delta C}{1-\delta} \geq H + \frac{\delta D}{1-\delta} - C$$

$$\frac{\delta C}{1-\delta} - \frac{\delta D}{1-\delta} \geq H - C$$

$$\frac{\delta}{1-\delta}\left(C - D\right) \geq H - C$$

$$\delta\left(C - D\right) \geq (1-\delta)(H-C)$$

$$(H-C) - \delta(H-C)$$

$$\delta\left(\cancel{C} - D + H - \cancel{C}\right) \geq H - C$$

$$\delta \geq \frac{H-C}{H-D}$$

$$C - D$$

Grim Trigger SPNE in Repeated PD

How do we interpret this statement:

$$\text{Cooperation is stable when } \delta \geq \frac{H-C}{H-D}$$

- Recall that the definition of the Prisoners' Dilemma was that $H > C > D > L$
- So this means $\frac{H-C}{H-D}$ is positive and less than 1
- As the $H - C$, the relative benefit of defecting increases, it gets harder to sustain cooperation
- It also gets harder to sustain cooperation as the relative penalty of defecting, $H - D$, shrinks

$0 \leq \frac{H-C}{H-D} \leq 1$

Other Strategies in Repeated Games

So far we've only looked at one example of a type of strategy in repeated game, *Grim Trigger*.

- Can you think of some others?
    - Recall that a complete strategy for a repeated game needs:
        - An initial move at $t = 1$
        - A plan of action for *every* possible history in *every* later stage $t \geq 2$
    - Ideally you would be able to tell a computer how to implement your strategy

Other Strategies in Repeated Games

Telling a computer how to implement strategies is exactly what Robert Axelrod did in a famous tournament in 1980.

- He invited people to submit their programs which would play 200 rounds of the prisoners' dilemma against each other
- The winning program was the one which had the highest total score after playing 200 rounds against all other programs
- What types of strategies do you think would succeed?

An Unexpected Winner

The winning program was named `TIT FOR TAT`

Surprisingly, it was fairly simple:

$$\begin{cases} \text{In stage } 1 & : \textit{Cooperate} \\ \text{In stage } t \geq 2 & : \begin{cases} \text{repeat what the other player did in } t-1 \end{cases} \end{cases}$$

Tit-for-Tat

Like *Grim Trigger*, *Tit-for-Tat* can punish other players for defecting.

- If a player plays *Defect*, it will copy them with *Defect* next round

But unlike *Grim Trigger* it has a short memory; or is very forgiving

- If the player who defected goes back to playing cooperatively, *Tit-for-Tat* will go back to cooperating too

Axelrod's Tournament

If you want to learn more:

- Read the original paper:
  Axelrod, Robert; Hamilton, William D. (27 March 1981), "The Evolution of Cooperation" (PDF), *Science*, 211 (4489): 1390–96
- The 1984 Book *The Evolution of Cooperation*, Basic Books
- Run the tournament yourself in python!
  `https://github.com/Axelrod-Python/Axelrod`
- Play this fun and short web game!
  `https://ncase.me/trust/`

Other Repeated Games

A More Complicated Game

		Player 2		
		x	y	z
	x	5, 5	2, 7	1, 3
Player 1	y	7, 2	3, 3	0, 1
	z	3, 1	1, 0	2, 2

What are the **pure strategy Nash equilibria** of the one-shot game?

(y, y) (z, z)

Repeated Game with 3 strategies per period

Now suppose that this game is played repeatedly an infinite number of times.

- Can we do better than the single period equilibrium?

Grim Trigger

Player 1

$$\begin{cases} t = 0 & \text{Play } x \\ t > 0 & \begin{cases} \text{Play } x \text{ if only } x \text{ has been played} \\ \text{Play } y \text{ if anything other than } x \text{has been played} \end{cases} \end{cases}$$

Player 2

- $EV_{Coop} = \frac{5}{1-\delta}$
- $EV_{Cheat} = 7 + \frac{3\delta}{1-\delta}$

$$\frac{5}{1-\delta} \geq 7 + \frac{3\delta}{1-\delta}$$

$$5 + \frac{5\delta}{1-\delta} \geq 7 + \frac{3\delta}{1-\delta}$$

$$-5 \qquad\qquad -7$$

$$\frac{\delta}{1-\delta}(5-3) \geq 7-5$$

$$(\delta/1-\delta)(2) \geq 2$$

$$\frac{\delta}{1-\delta} \geq 1 \Rightarrow \delta \geq 1-\delta$$

$$\delta \geq 1/2$$

Grim Trigger

Solve for the value of δ for which this is a **SPNE**

Tit-for-Tat with extra forgiveness

Player 1

$$\begin{cases} t=0 & \text{Play } x \\ t>0 & \text{Play Player 2's strategy from } t-1 \\ & \begin{cases} \text{If P2 didn't play } x \text{ in } t-2 \text{ and P2 did play } x \text{ in } t-1, \\ \text{ play P2's strategy from } t-1 \\ \text{Else play } y \text{ forever} \end{cases} \end{cases}$$

Player 2

- $EV_{Coop} = \frac{5}{1-\delta} = 5 + 5\delta + \cancel{\frac{5\delta^2}{1-\delta}}$
- $EV_{Cheat} = 7 + 2\delta + \frac{5\delta^2}{1-\delta}$

$7 + 2\delta + \cancel{\frac{5\delta^2}{1-\delta}}$

$5 + 5\delta \geq 7 + 2\delta$

$3\delta \geq 2$

$\delta \geq 2/3$

Tit-for-Tat with extra forgiveness

Solve for the value of δ for which this is a **SPNE**

Tit-for-Tat vs Grim Trigger

Tit-for-Tat is more forgiving

- If $\delta \geq 1/2$, equilibrium is supported by Grim Trigger
- If $\delta \geq 2/3$, equilibrium is supported by Tit-for-Tat

A reciprocating cooperation strategy

Player 1 $\begin{cases} t = 0 & \text{Play } y \\ t > 0 & \begin{cases} \text{Play } y \text{ if } t \text{ is even} \\ \text{Play } x \text{ if } t \text{ is odd} \\ \text{Play } z \text{ forever} \\ \quad \text{if P2 played } y \\ \quad \text{when } t \text{ is even} \end{cases} \end{cases}$

Player 2 $\begin{cases} t = 0 & \text{Play } x \\ t > 0 & \begin{cases} \text{Play } y \text{ if } t \text{ is odd} \\ \text{Play } x \text{ if } t \text{ is even} \\ \text{Play } z \text{ forever} \\ \quad \text{if P2 played } y \\ \quad \text{when } t \text{ is odd} \end{cases} \end{cases}$

$EU_{coop} = 2 + 7\delta + 2\delta^2 + 7\delta^3 + \ldots$

$EU_b = 7 + 3\delta + \frac{2\delta^2}{1-\delta}$

$\delta \geq 1/4$

A reciprocating cooperation strategy

Solve for the value of δ for which this is a **SPNE**

When can cooperation be achieved?

With all of these different ways of achieving repeated cooperation, you might be wondering if there is a way to tell what strategies can actually work

Folk Theorem

Any strategy is a potential SPNE for a **repeated** stage game if:

- Both agents are sufficiently patient and far-sighted (high enough δ)
- The payoffs from the cooperative strategy profile satisfy the two properties:
  - **Individually Rational:** the payoffs to each agent (weakly) exceed their minimax payoffs in the stage game
  - **Feasibility:** the payoffs are weighted averages of the payoffs found in the stage game

Folk Theorem with 3x3 Repeated Game Example

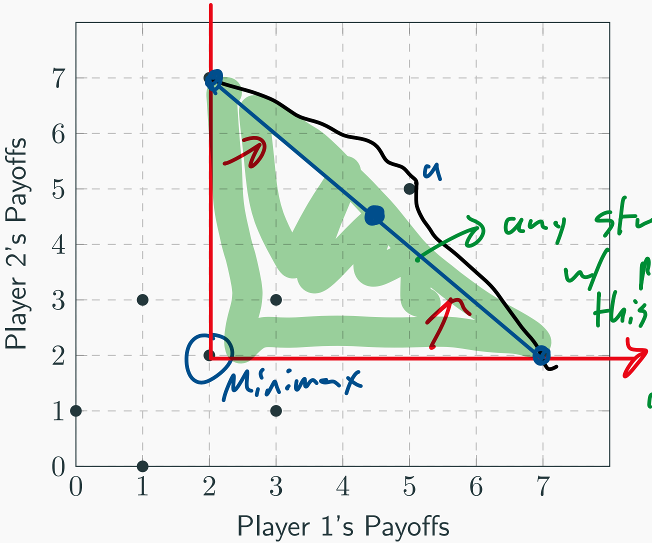
		Player 2		
		x	y	z
Player 1	x	5, 5	2, 7	1, 3
	y	7, 2	3, 3	0, 1
	z	3, 1	1, 0	2, 2

The Minimax equilibrium is (z,z)

Maximin

- it *minimizes* the *maximum* payoff that your opponent could get
- The Minimax payoffs in this stage game are (2, 2)
- Intuitively, this is the *safe* option: you can always fall back on it if cooperation fails

Folk Theorem with 3x3 Repeated Game Example payoffs



any strategy profile w/ payoff in this region are sustainable in rep. cooperative eqm.

Folk Theorem with 3x3 Repeated Game Example

- The shaded region of the graph shows us all of the strategy profiles which could be sustained by the **Folk Theorem**
- This shows us why that strategy profile of alternating between (x, y) and (y, x) worked:
    - even though getting 2 on even or odd periods was no better than the Minimax payoffs, because you could alternate with the higher payoff of 7 you could do better as long as you are patient enough
    - this mix between (2, 7) and (7, 2) is *within the convex hull* of sustainable payoffs

- As you can probably tell, there are an infinite number of strategy profiles which can achieve cooperation
  - We could allow for mixed strategies, which would work similar to the alternating example we saw
  - The Folk Theorem tells us that all we need is for all players to be patient enough
  - and also that the past plays are common knowledge

Importance of the Folk Theorem

Why does this matter for real life?

- Most strategic interactions in your life are repeated
    - Sharing chores with your roommates
    - Interacting in class with me every week
    - Being nice to the barista at your regular cafe

Importance of the Folk Theorem

- Even when you don't repeatedly interact with the same exact people, you still see cooperative outcomes
- **Institutions, Reputations, and Social Structures** all serve to allow for past interactions to be common knowledge
- The history of humanity is built on how we arrange our strategic interactions in ways so that people are incentivized to play nice with others

Importance of the Folk Theorem

Some caveats:

- People can't know exactly when the game will end; if they don't have any incentive from future cooperative gains, they will always defect
  - Institutions have to *seem* like they are infinitely lived (compared to finitely lived humans)
- Cooperative equilibria must be better than peoples' outside options
  - If you make your institution too costly for people to engage with, they will opt out
- People need to be patient enough to make cooperation worth it