Ch.11 Repeated Games

October 13, 2023

Repeating A Zero-Sum Game

Repeating the Prisoners' Dilemma

Infinite Repetitions

- Consider a 0-sum game.
- Only the payoffs of row player (I) are shown.

	t_1	t_2
s_1	1	0
s_2	0	1

- NE: $(0.5s_1 \bigoplus 0.5s_2, 0.5t_1 \bigoplus 0.5t_2)$
- equilibrium payoffs = (0.5, 0.5)
- I's security strategy: $0.5s_1 \bigoplus 0.5s_2$
  I's security payoff = 0.5
- What happens if this stage game is repeated twice?

Game Forms

- The stage game repeated twice:

	t_1t_1	t_1t_2	t_2t_1	t_2t_2
s_1s_1	2	1	1	0
s_1s_2	1	2	0	1
s_2s_1	1	0	2	1
s_2s_2	0	1	1	2

- If I Plays a mixed strategy (0, 0.5, 0.5, 0), his exp. payoff is 1 no matter what II plays.
- How can II reduce I's exp. payoff below 1?

What II can do?

- I plays $(0.5 s_1 s_2 \bigoplus 0.5 s_2 s_1)$
- In the 1st stage game, no matter what II plays, I's exp. payoff $=$ 0.5
- Observing I's move in the 1st stage, II knows exactly what I's move will be in the 2nd stage, and hence can play accordingly to make I receives 0.
- What went wrong?

A New Strategic Form

- II's choice in the 2nd stage is much more complicated than the previous strategic form allows.
- List II's all choices in the 2nd stage, including the one contingent on the history of the 1st stage.
- The same for I.
- Def: history $h_{ij} \equiv (s_i, t_j)$
- Def: $H = \{h_{ij}\}_{i,j=1,2}$
- Def: $S = \{s_1, s_2\}$, $T = \{t_1, t_2\}$
- I's pure strategy in the 2nd stage is a function $f : H \to S$

I's Strategy in the 2nd Stage, $f : H \to S$

- I's choice in the 2nd stage depends on the history.
- Different history may lead I's to act differently.
- 4 history, 2 choices for each history
  $\Rightarrow$# of I's pure strategies $= 2^4 = 16$
- For instance, f_{1212}:
  $$\begin{array}{cccc} h_{11} & h_{12} & h_{21} & h_{22} \\ \downarrow & \downarrow & \downarrow & \downarrow \\ s_1 & s_2 & s_1 & s_2 \end{array}$$
- Together with 2 choices in the 1st stage, I's pure strategy in the repeated game, (s, f) has 32 possibilities.

The Size of the Game

- The proper strategic form is a 32*32 matrix.
- For I, what's the difference between (s_1, f_{1111}) and (s_1, f_{1112})?
- Given I plays s_1 in the 1st stage, only 2 histories are meaningful: h_{11}, h_{12}
- Only the 1st two numbers of f's subscript matter.
- (s_1, f_{1111}), (s_1, f_{1112}), (s_1, f_{1121}), (s_1, f_{1122}) mean the same to I.
- This group of 4 can be reduced to 1 strategy.
- Other groups of 4 can be reduced as well $\Rightarrow$8 pure strategies for I

The Size of the Game

- 8 pure strategies for I: (s_i, F)
- Given I plays s_1 in the 1st stage, there are only 2 meaningful histories depending on II's behavior in the 1st stage
- I has 4 contingent strategy in the 2nd stage, for instance:

$$F_{11} \quad \begin{array}{cc} t_1 & t_2 \\ \downarrow & \downarrow \\ s_1 & s_1 \end{array}$$

- For 2 stages, I has 8(=2*4) pure strategies.

Security Strategy

- Once we have the 8*8 strategic form, we can find one of I's security strategies to be: playing (s_1, F_{11}), (s_1, F_{22}), (s_2, F_{11}), (s_2, F_{22}) with equal probabilities.
- I's security level is 1.
- Given I uses this mixed strategy, can II lower down I's exp. payoff?
- In the 1st stage, I plays s_1 and s_2 with the same probability. I's exp. payoff $=$0.5 no matter what II plays.
- Suppose I plays s_1 in the 1st stage, for II, I will play F_{11} and F_{22} with equal probabilities in the 2nd stage.
- That is, for II, I will play s_1 and s_2 with equal probabilities again.
- II can do nothing but allow I's exp. payoff to be 0.5 in the 2nd stage.
- Repeating a 0-sum game twice, each player's security level simply doubles so long as they play the security strategy in the stage game twice.

Prisoners' Dilemma

- Consider the following stage game:

	d	h
d	2,2,	-1,3
h	3,-1	0,0

- h is each player's dominant strategy, but the outcome of (h, h) is Pareto inferior to the outcome of (d, d).
- What happens if the stage game is repeated twice?

Repeated an Indefinite Number of Times

- A judge will allow the game to continue to the next stage with p. $2/3$.
- The Grim strategy calls for d to be played as long as the opponent reciprocates by playing d also.
- Is (Grim, Grim) a NE?
- Given I plays Grim, II has 2 possible meaningful choices:
  - play Grim, i.e. always play d
  - play d in the 1st n stages, h in the $n+1$th stage, then h forever after

II's Payoff in Each Case

- If II plays Grim, he'll receive:

$$C = 2 + 2(\frac{2}{3}) + ...2(\frac{2}{3})^{n-1} + 2(\frac{2}{3})^{n} + 2(\frac{2}{3})^{n+1} + 2(\frac{2}{3})^{n+2} + ...$$

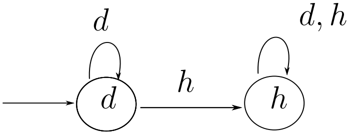
- If II switches to h in the $n+1$th stage, he'll receive:

$$D = 2 + 2(\frac{2}{3}) + ...2(\frac{2}{3})^{n-1} + 3(\frac{2}{3})^{n} + 0(\frac{2}{3})^{n+1} + 0(\frac{2}{3})^{n+2} + ...$$

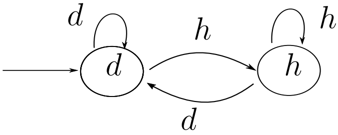
$$\begin{aligned} C - D &= (2-3)(\frac{2}{3})^{n} + 2(\frac{2}{3})^{n+1} + ... \\ &= (\frac{2}{3})^{n}[-1 + 2 * \frac{2}{3} + ...] > 0 \end{aligned}$$

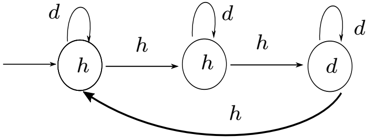
- II's Grim is the best response to I's Grim, and vice versa. (Grim, Grim) is a NE.

- What happens if the stage game is repeated infinitely?
- We'll focus on strategies that can be represented by finite automata.
- For instance, Grim:



- Tit-for-Tat:





- Tweedledum:
- I plays Tit-for-Tat and II plays Tweedledum

	-1	0	0	3	-1	...
Tit-for-Tat	d	h	h	h	d	...
Tweedledum	h	h	h	d	h	...

$$\begin{aligned} \text{I's discounted payoff} &= -1+0+0+3\delta^3+\ldots \\ &= (-1+3\delta^3)+\delta^4(-1+3\delta^3)+\ldots \\ &= (-1+3\delta^3)\left(\frac{1}{1-\delta^4}\right) \end{aligned}$$

- $\delta \to 1$, I's payoff $\to \infty$
- $\max$ payoff $\Longleftrightarrow \max(1-\delta)$payoff

$$\lim_{\delta\to 1}(-1+3\delta^3)\left(\frac{1-\delta}{1-\delta^4}\right) = (-1+0+0+3)\lim_{\delta\to 1}\frac{-1}{-4\delta^3} = \frac{-1+0+0+3}{4}$$

- I cares about the long run average.

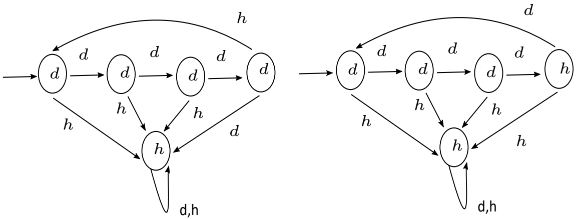
Is (Tit-for-Tat, Tit-for-Tat) a NE?

- Suppose I plays Tit-for-Tat.
- If II also plays Tit-for-Tat, II will receive 2 in every stage.
- If II plays d for some time, then switches to h, he will receive 3($+1$) in that stage.
  - If II continues playing h, he will meet I's h and receive only 0(-2).
  - To stop losing a payoff of 2, II has to start to play d when I plays h, and II's payoff in this stage is -1(-3).
- II's initial gain of 1 is smaller than the later losses.
- It does not pay for II to switch h.
- II's Tit-for-Tat is the best response to I's Tit-for-Tat, and vice versa.

NE

- Is (always d, always d) a NE?
- Is (always h, always h) a NE?
- If players are patient, there are at least 2 possible equilibrium long-run average payoffs: (0,0) and (2,2).
- Is there any other possibility?

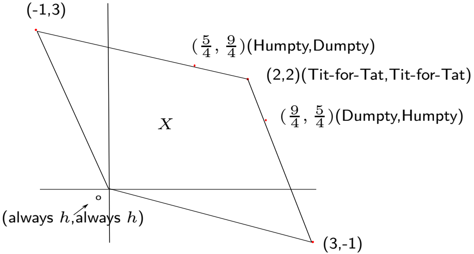
- Humpy and Dumpty



- Given I plays Humpty, if II plays Dumpty, II's average payoff = (3*2+3)/4=9/4.
- If II deviates,
  - $(d, d) \to (d, h)$, II gains +1 at that stage, but bears the loss of -2,-2,-3,-2,...
  - $(d, h) \to (d, d)$, II gains -1 at that stage, but bears the loss of -2,-2,-2,-3,...
  - It is not worth for II to deviate.
- Given II plays Dumpty, if I plays Humpty, I's average payoff = (3*2-1)/4=5/4.
  - It is not worth for I to deviate.

Payoffs

- Payoffs of profile of pure strategies: (2,2), (-1,3), (3,-1), (0,0)
- Players can coordinate to achieve a convex combination of these 4 points.
- Convex Hull $X = \{p_1(2,2) + p_2(-1,3) + p_3(3,-1) + p_4(0,0) : p_i \geq 0, \sum p_i = 1.\}$

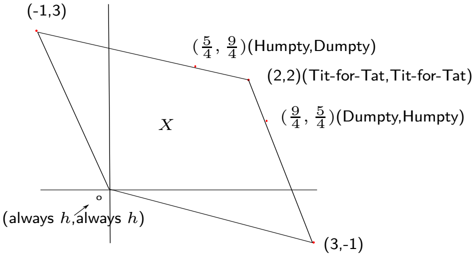


- At least 4 equilibrium payoffs, what more?

Punishments

- The more severe the punishment for a deviating move is, the more likely is a coordination to achieve.
- If II uses t to punish I, II expects I to play $r_1(t)$ where $r_1(.)$ is I's best response function.
- I's payoff will be $\pi_1(r_1(t), t) = \max_s \pi_1(s, t)$.
- The most severe t will yield the lowest π_1: $\min_t \max_s \pi_1(s, t) \equiv \overline{m}_1$.
- I's minimax value $\overline{m}_1$ is I's payoff when punished most severely.
- Similarly, we define II's minimax value $\overline{m}_2 \equiv \min_s \max_t \pi_2(s, t)$.
- For Prisoners' Dilemma, $r_1(d) = r_1(h) = h$.
  - $\pi(h, d) = 3$, $\pi(h, h) = 0$, $\overline{m}_1 = 0$.
  - Similarly, $\overline{m}_2 = 0$.

- Folk Theorem: Consider that players only use finite automaton, and players care about the long-run average payoffs. The outcomes corresponding to NE in pure strategies of the infinitely repeated game are dense in the set: $\{x : x \in X, x \geq (\overline{m}_1, \overline{m}_2)\}$.
- Consider $x = \frac{1}{3}(3,-1) + \frac{1}{3}(-1,3) + \frac{1}{3}(0,0) = (\frac{2}{3}, \frac{2}{3})$.



- To achieve x, I and II will play (h,d), (d,h) and (h,g) cyclically. If one deviates, the other will play h forever.
  - Deviation can only lower the long-run average payoff.