Lecture 9: Repeated Game

Yi, Yung (이융)
KAIST, Electrical Engineering
http://lanada.kaist.ac.kr
yiyung@kaist.edu

Basic Terminologies

- Dynamic/Static Game
  - Game in which we have sequence of moves or not
- Complete/Incomplete Information
  - Games in which the strategy space and player's payoff functions are common knowledge, or not
- Perfect/Imperfect Information
  - Each move in the game the player with the move knows the full *history* of the game thus far.
  - At some move the player with the move does not know the history of the game.
  - Typically used for dynamic games

Prisoner's Dilemma

	Cooperate	Defect
Cooperate	$1, 1$	$-1, 2$
Defect	$2, -1$	$0, 0$

- What about playing this game iteratively?
- Iterative Prisoner's Dilemma
- A special form of dynamic games
- Difference from the earlier extensive form game, or stackelberg game?

Motivation

	Cooperate	Defect
Cooperate	$1, 1$	$-1, 2$
Defect	$2, -1$	$0, 0$

- Consider this strategy S:
  - Start with C, and choose C as long as the other player chooses C
  - If in any period the other player chooses D, then choose D in *every subsequent period*
- Outcome example
  - (C,C), (C,C), (C,C) … (C,D) (D,D) (D,D) …
  - Why player 2 can choose D in (C,D)? For a short-term gain
  - The strategy S means that I will punish you if you defect!
- If a player value the present more highly than the future, she may or may not choose defect.
  - How *patient* is a player?

What is Repeated Game?

	L_2	R_2
L_1	1,1	5,0
R_1	0,5	4,4

	L_2	R_2
L_1	1,1	5,0
R_1	0,5	4,4

	L_2	R_2
L_1	1,1	5,0
R_1	0,5	4,4

………

- A game is repeated multiple times, say 10 times.
  - 한번만 할 것이라면, L_i를 선택하겠지만, 여러 번 할 것이라면, R_i를 선택하여서, "협조"를 구해보는 것도…
  - 만약, 협조하다가 배신하면, 내가 너를 "응징"하리…

What is Repeated Game?

	L_2	R_2
L_1	1,1	5,0
R_1	0,5	4,4

	L_2	R_2
L_1	1,1	5,0
R_1	0,5	4,4

	L_2	R_2
L_1	1,1	5,0
R_1	0,5	4,4

………

- In ongoing relationships, the promise of future rewards and the threat of future punishments may sometimes provide incentives for good behavior (i.e., cooperation) today. (Nobel prize!)

- T: Period of a repeated game
  - Finite case
  - Infinite case

Terminologies

- **Repeated games**: given a simultaneous-move game G, a repeated game of G is an extensive game with perfect information and simultaneous moves in which a history is a sequence of action profiles in G. I will denote the repeated game, if repeated T times, as G^T.
- G is often called a **stage game**, and G^T is called a **supergame**.

Payoffs of a repeated game

- A player gets a payoff from each stage game, so her total payoff from the supergame is the discounted sum of the payoffs from each stage game.
- Let's call a sequence (with T periods) of action profiles as $(a_1, a_2, \ldots a_T)$, then a player i's total payoff from this sequence, when her discount factor is δ, is

$$u_i(a_1) + \delta u_i(a_2) + \delta^2 u_i(a_3) + \ldots + \delta^{T-1} u(a_T) = \sum_{t=1}^{T} \delta^{t-1} u_i(a_t).$$

- If the sequence is infinite, then the discounted sum is $\sum_{t=1}^{\infty} \delta^{t-1} u_i(a_t)$.

Discounting: Rationale

- Discounting: people value \$100 tomorrow less than \$100 today (Why? Interest rate; the world may end tomorrow.)
- The **discount factor** δ $(0 \leq \delta \leq 1)$ denotes how much a future payoff is valued at the current period, or how patient a player is.
  - If a player has a δ of .8, then \$100 tomorrow is equivalent to \$80 today for her.
- Or, I continue this game with probability δ and end this game with probability $1 - \delta$

Dividing Cases

- Finite repeated game
  - (Case 1) A stage game has unique NE
  - (Case 2) A stage game has multiple NEs
- (Case 3) Infinite repeated game
- In this lecture, in most of examples,
  - Use Prisoner's dilemma as a stage game
- Assumptions
  - <u>Perfect monitoring</u>: At each period, the outcomes of all past periods a re observed by all players

Finite Repeated Game:

Unique NE in the stage game

lanada

Finitely-repeated PD

	C	D
C	2,2	0,3
D	3,0	1,1

- What happens if this game was played $T < \infty$ times?
- We first need to decide what the equilibrium notion is. Natural choice, subgame perfect Nash equilibrium (SPE).
- Recall: SPE $\Longleftrightarrow$ backward induction.
- Therefore, start in the last period, at time T. What will happen?

KAIST

Cont'd

- In the last period, "defect" is a dominant strategy regardless of the history of the game. So the subgame starting at T has a dominant strategy equilibrium: (D, D).
- Then move to stage $T-1$. By backward induction, we know that at T, no matter what, the play will be (D, D). Then given this, the subgame starting at $T-1$ (again regardless of history) also has a dominant strategy equilibrium.
- With this argument, we have that there exists a unique SPE: (D, D) at each date.
- In fact, this is a special case of a more general result.

(D,D,D,D): SPNE Checking

	C	D
C	2,2	0,3
D	3,0	1,1

Period:	1	…	$t-1$	t	$t+1$	…	T
(s_1, s_2):	a^1	…	a^{t-1}	(C, X)	(D, D)	…	(D, D)
Relation between player 1's payoffs:	||	…	||	∧	∧|	…	∧|
(s_1', s_2):	a^1	…	a^{t-1}	(D, X)	$(D, ?)$	…	$(D, ?)$

Thus, we have ...

Theorem

Consider repeated game $G^T(\delta)$ for $T < \infty$. Suppose that the stage game G has a unique pure strategy equilibrium a^*. Then G^T has a unique SPE. In this unique SPE, $a^t = a^*$ for each $t = 0, 1, ..., T$ regardless of history.

Proof: The proof has exactly the same logic as the prisoners' dilemma example. By backward induction, at date T, we will have that (regardless of history) $a^T = a^*$. Given this, then we have $a^{T-1} = a^*$, and continuing inductively, $a^t = a^*$ for each $t = 0, 1, ..., T$ regardless of history.

- In this case, no cooperation appear even if we repeat the game.

Infinite Repeated Game:

What happens if we don't know when the game would end?

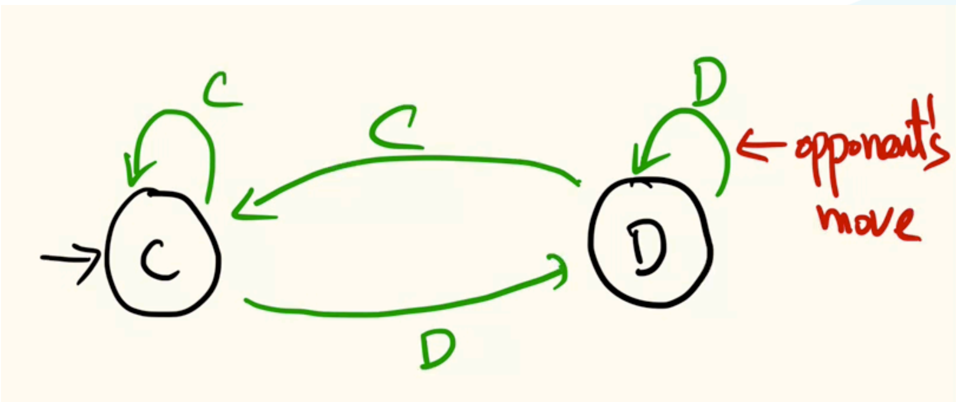
Intro

- An infinitely repeated game usually has more SPNEs than a finitely repeated game, and it may have multiple SPNEs even if the stage game has a unique NE.
- But playing the NE strategies in each stage game, regardless of history, is still a SPNE in the infinitely repeated game.
  - So each player choosing defection is a SPNE in infinitely repeated PD as in finitely repeated PD.
  - But there are other SPNEs in infinitely repeated PD.
- Our interest: Strategies inducing cooperation is SPNE or not

Strategy: Tit-for-tat

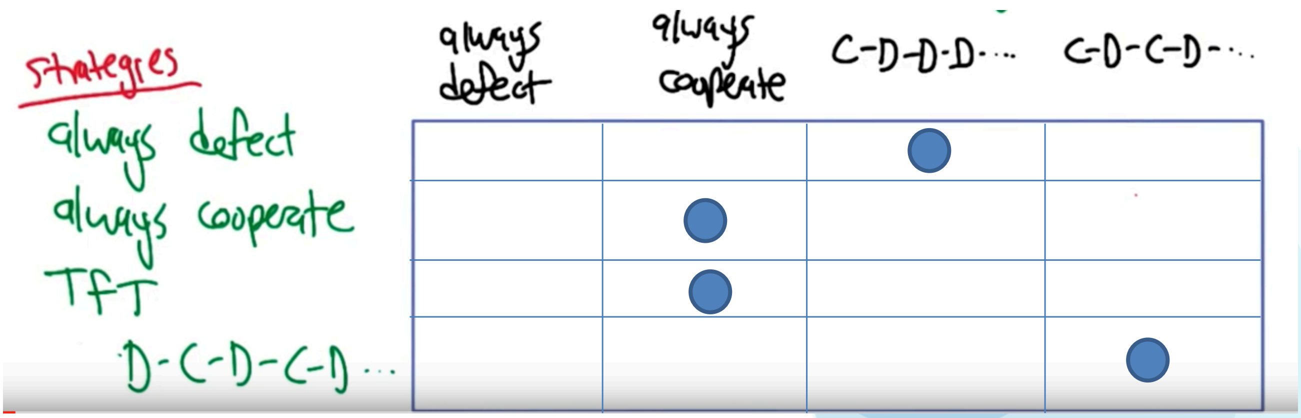
	C	D
C	2,2	0,3
D	3,0	1,1

- TFT strategy
  - Play C first,
  - Then, do whatever the other play did in the previous period



What does TFT do when playing against these strategies?

	C	D
C	2,2	0,3
D	3,0	1,1



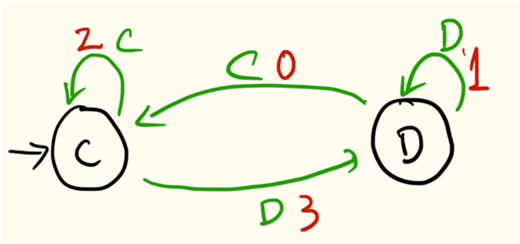
Which is better response to TFT?

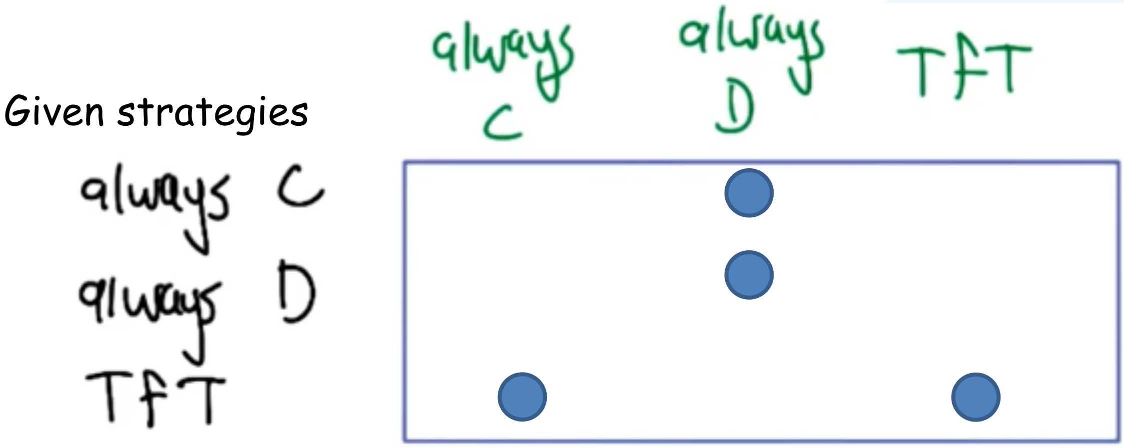
	C	D
C	2,2	0,3
D	3,0	1,1

- Always D = D, D, D, D, …
  - Payoff = $3 + \delta/(1-\delta)$
  - Good for low γ
- Always C = C, C, C, C, …
  - Payoff = $2 * 1/(1-\delta)$
  - Good for high γ
- For what value of γ, are they equally good? 1/2

Best Response to Some Strategies ($\gamma > \frac{1}{2}$)

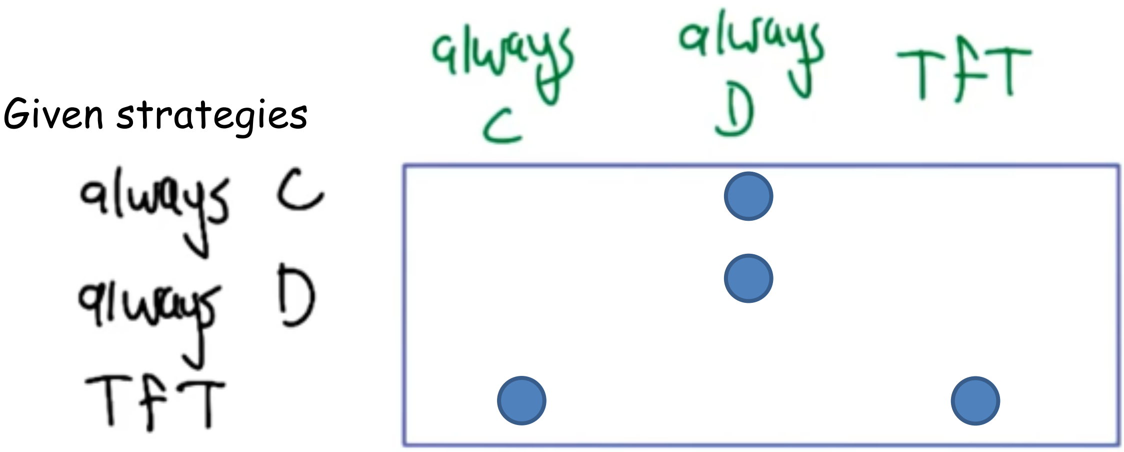
- State: opponent choice
- Edge: my choice
- Edge label: my payoff





What are NEPs? Is there cooperation in some NEP?

Cooperation in Infinite Repeated Game

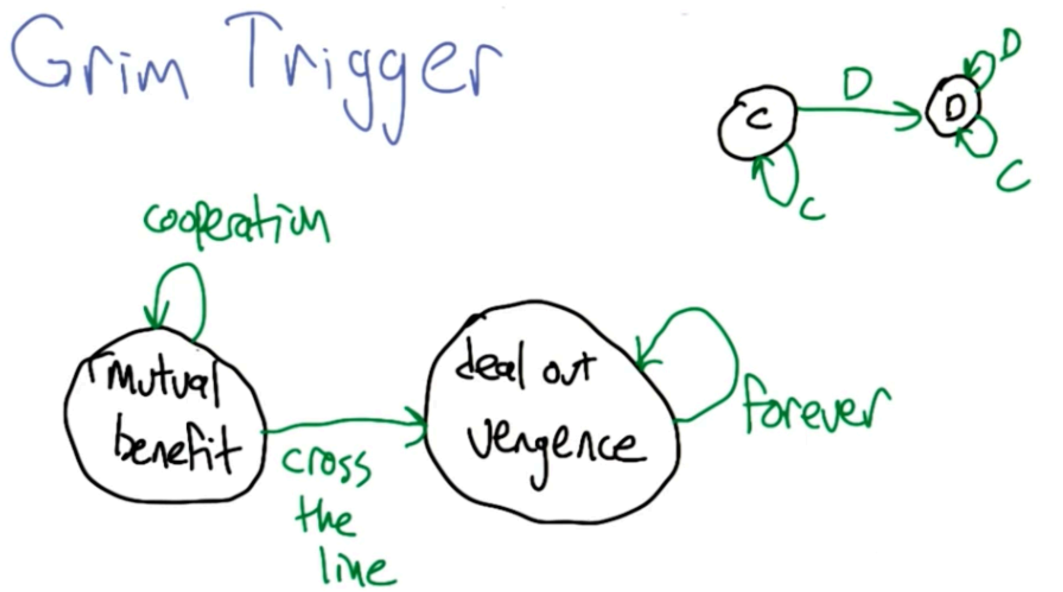


What are NEPs? Is there cooperation in some NEP? Why?

TFT is still a NEP for small δ?

Strategy: Grim Trigger (GT)

- Start with C
- Then, play C if opponent has played D, and play D otherwise
  - Draconian policy

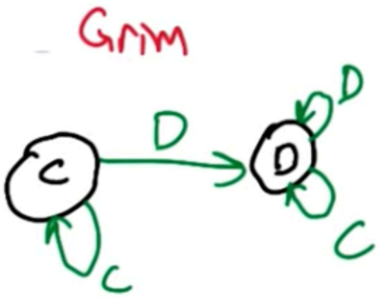


Grim Trigger: NEP

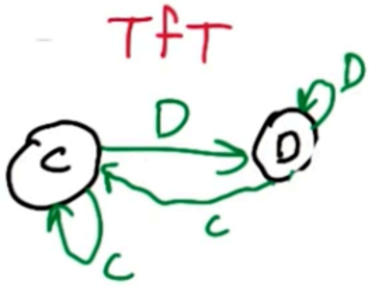
- Under what conditions of δ, (GT,GT) is a NEP?

Incredible Threat (NE but not SPNE)

- Consider a strategy (GT, TFT) as follows:



vs.



- Is this NE?

- Is this SPNE?

Checking SPNE Easily: One Deviation Property

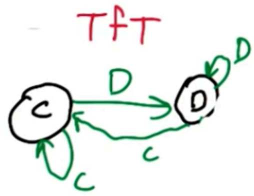
- **One-deviation property**: no player can increase her payoff by changing her action at the start of any subgame in which she is the first-mover, given the other players' strategies *and* the rest of her own strategy.

- A strategy profile in an infinitely repeated game with a discount factor less than 1 is a SPNE if and only if it satisfies the one-deviation property.

TFT: SPNE

	C	D
C	2,2	0,3
D	3,0	1,1

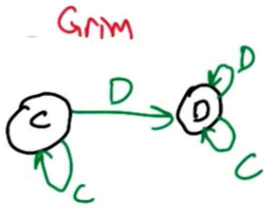
- Under what δ, (TFT,TFT) is a SPNE?



Grim Trigger Strategy: SPNE

	C	D
C	2,2	0,3
D	3,0	1,1

- Under what δ, (GT,GT) is a SPNE?



Summary

Finite Repeated Game:

Multiple NEs in the stage game

Message

- SPNE can include a strategy (at some stage game) that is NOT NE of the associated stage game.

- Generally, many SPNEs

Example: Extended PD

	Cooperate	Defect	Punish
Cooperate	4,4	0,5	0,0
Defect	5,0	1,1	0,0
Punish	0,0	0,0	3,3

- Cooperate = Quiet (묵비권),
- Defect (배반) = Fink (고자질)

- What are the NEs? (D,D) and (P,P)
- Play twice, i.e., T=2
- We will see
  - Even for the known ends, still cooperation helps
  - “Like to sustain (C,C)”, which is not an NE of the one-shot game

Strategy “Yung” that is SPNE

- Strategy “Yung”
  - Play C and then
  - Play P if (C,C), and Play D otherwise

- Is “Yung” a strategy?
- Is “Yung” a SPNE? Yes!

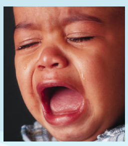
How to check a strategy is SPNE?

- One deviation principle, i.e.,
- "Assuming that other players are playing Yung, what happens if I deviate?"
- If I get a larger payoff (by deviation), than Yung is not an SPNE .
- If I get a smaller payoff (by deviation), than Yung is an SPNE.

Lanada

What happens if I deviate?

- I deviate?
  - In other words, I don't play C, but D (no reason to play P)
  - Why? Temptation to cheat because of an increasing payoff present (현실에 눈이 어두워서…)
- If I play C (i.e., playing Yung)
  - C → 4 (C,C) + 3 (P,P) = 7
- If I deviate and play D (i.e., playing some other strategy)
  - D → 5 (D,C) + 1 (D,D) = 6
- Temptation to cheat (5-4 = 1) < reward – punishment (3-1 = 2) → I should not have deviated
- (C,C) is reward, and (B,B) is punishment
- Yung is a SPNE
  - C,P,P,P,P….



KAIST

One Deviation Property

- **One-deviation property**: no player can increase her payoff by changing her action at the start of any subgame in which she is the first-mover, given the other players' strategies *and* the rest of her own strategy.

- One-deviation property of SPNE of finite horizon games: A strategy profile in an extensive game with perfect information and a finite horizon is a SPNE if and only if it satisfies the one-deviation property.

One Deviation Property

- One-deviation property of SPNE of finite horizon games: A strategy profile in an extensive game with perfect information and a finite horizon is a SPNE if and only if it satisfies the one-deviation property.
- A strategy profile in an infinitely repeated game with a discount factor less than 1 is a SPNE if and only if it satisfies the one-deviation property.
- **One-deviation property**: no player can increase her payoff by changing her action at the start of any subgame in which she is the first-mover, given the other players' strategies *and* the rest of her own strategy.