

게임이론 Part 3

게임의 기본 개념들 (게임이론, 진화, 그리고 협력)

허준석→이동한→조남운

`mailto:experiment.namun+2016f@gmail.com`

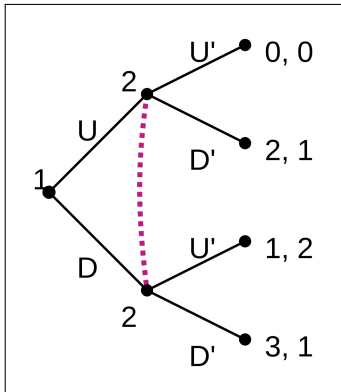
2016년 10월 7일

목차

- 1 전개형 게임 2(Extensive Form Game 2)
- 2 Repeated Game (RG)

정보집합 (Information Set)

- 2번째 노드에서 P_2 가 어디 있는지 모를 때
- 이것은 무엇과 동일한 상황인가?
- 이것을 NFG (Normal Form Game) 으로 나타내본다면?
- 정보집합을 어떻게 그릴 수 있는 조건은?



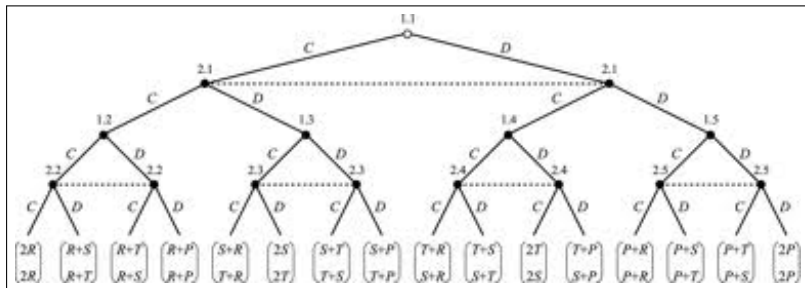
Exercise

- Big John (P_1) vs Little John (P_2)
 - 1 옆의 NFG의 PSNE은 무엇인가?
 - 2 옆의 NFG의 MSNE은 무엇인가?
 - 3 Big John이 First mover일 때 EFG의 SPE (Subgame Perfect Equilibrium)은?
 - 4 Little John이 First mover일 때 EFG의 SPE은?

	C	W
C	5, 3	4, 4
W	9, 1	0, 0

특별한 종류의 전개형 게임

- 매기는 NFG
- 전기의 결과에 의존해 다음 기가 진행
- 전략의 수는?



- URL: 교실상에서 공지

RG에서 균형은 어떻게 찾을까?

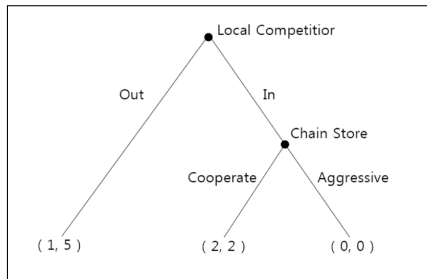
- RG의 전략의 수는 너무 많다.
- 기존의 방법으로는 찾기 힘들다.
- 게다가, RG이 유한하게 전개된다면?
- BI (Backward Induction)에 따른 균형은 뻔한 결과

Finite vs. Infinite

- Finitely repeated game에서 균형은 BI으로
- Infinite repeated game에서 균형은?
- 일단 전략이 다른 방식으로 주어져야

Chain-store Game in Finitely Repeated Game

- 20번에 걸쳐서 순차적으로 이 게임을 한다고 생각해보자. 즉, 1명의 현재 독점자와 20명의 순차적 경쟁자
- BI에 따른 균형은?
- 하지만, chainstore는 협박을 통해 이윤을 늘릴 수 있다! 아마도 BI에 필요한 가정에 문제가 있는 것은 아닐까?



Strategy in Infinitely Repeated Games

- 전략은 algorithm 처럼 주어진다.
- 즉, 상대의 전략에 의존적으로 주어진다.
- 아울러 모든 결과를 정하기 힘들다면, 제한된 기억에 의존한다.
- history matters.

RG의 전략은 history에서 action으로 가는 함수

- RG의 전략들
 - ① Tit-For-Tat (TFT)
 - ② GRIM
 - ③ Pavlov (Win-Stay-Lose-Shift)
 - ④ ALLC / ALLD

Payoff in Infinitely Repeated Games

- 보수는 어떻게 정의되어야 할까?

$$\overbrace{\Pi(s)}^* = \pi_1(s) + \rho \times \pi_2(s) + \rho^t \times \overbrace{\pi_t(s)}^{**} + \dots$$

* 무한반복 게임의 보수

** t 기에 s 전략에 따른 one-shot 보수

- ρ 에 대한 해석:
 - ① 시간 할인 (현재가치로 환산하기)
 - ② 같은 사람과 다음 기에 다시 만날 확률
- $\Pi(s)$ 는 수렴해야 한다.

GRIM (Trigger) Strategy

- GRIM 전략은 다음과 같이 정의된다.
- τ 기에 선수 i 는 다음과 같이 전략을 선택한다.
 - ① plays C at $\tau = 1$
 - ② plays C , τ 이전 시기에 나와 남의 행동이 모두 C 였다면
 - ③ plays D , 다른 모든 경우에
- 왜 'Trigger' strategy인가?



TFT Strategy

- TFT 전략은 다음과 같이 정의된다.
- τ 기에 선수 i 는 다음과 같이 전략을 선택한다.
 - ① plays C at $\tau = 1$
 - ② 바로 전기에 상대의 행동을 plays
- 로버트 액설로드, [협력의 진화]



Back to PD Game!

- One-shot PDG 게임 혹은 finitely RPDG의 결론은?
- 이러한 결론이 바뀌는 것이 가능할까?
- 만일 바뀐다면 어떤 것들이 영향을 주게 되는가?

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

- 게임 상황
 - ① GRIM에 입각한 두 명의 플레이어
 - ② 게임은 무한 반복
- 만일 모두 GRIM을 따라서 움직인다면, 이때 선수들의 보수는?

$$\Pi = 2 + \rho \times 2 + \rho^2 \times 2 + \dots$$

무한등비급수 계산하는 법

- 초항이 a 이고 공비가 ρ 인 어떤 무한등비급수의 합이 S 라고 하자. 즉,

$$S = a + \rho \times a + \rho^2 \times a + \cdots + \rho^t \times a \dots$$

- 이때 이 식은 다시 다음과 같이 정리될 수 있다.

$$S = a + \rho \left(\overbrace{a + \rho^1 \times a + \dots}^{=S} \right).$$

- 따라서, $S = \frac{a}{1 - \rho}$.
- 단, 저 값이 존재하려면 ρ 는 어때야 할까?

만일 플레이어가 GRIM을 따르지 않게 된다면?

- 어떤 시점 t 에서 선수가 GRIM을 따르지 않기로 했다고 쳐보자. 즉, 선수는 앞서 누구도 D 를 하지 않았으나 C 를 하는 대신 D 를 했다고 해보자.

$$\Pi' = \overbrace{2 + \rho \times 2 + \rho^{t-1} \times 2}^{\text{GRIM을 따를 때}} + \overbrace{\rho^t \times 3}^{\text{따르지 않았을 때}} + \overbrace{\rho^{t+1} \times 1 + \dots}^{\text{그 이후}}$$

- 만일 이러한 일탈이 발생하지 않으려면, 어떤 조건이 성립하면 될까?

$$\Pi' > \Pi$$

만일 플레이어가 GRIM을 따르지 않게 된다면?

- 사실상 $t - 1$ 까지는 모두가 GRIM에 충실했으므로 두 사람의 보수가 같을 것이다. 따라서, 이건 그냥 잘라내고 생각을 하자. 그리고 두 사람의 보수를 모두 ρ^t 로 나누어보자.
- 결국 이는 어떤 플레이어가 게임이 시작되는 첫기에 GRIM에 따라서 움직이지 않았을 때의 보수와 이에 충실했을 때의 보수 사이의 비교의 문제가 된다.
- 결국 1회의 PDG와 다른 결과를 낼 가능성을 부여하는 것은 무엇인가?

한가지 이론적인 질문 (Optional)

- 앞서서 반복게임이란 '특수한' 형태의 전개형 게임이라고 언급.
- 그런데, 전개형 게임에서 NE을 걸러내는 틀 중에서 유용하게 사용한 것은 무엇인가?
- 어떤 균형이 SPE가 되려면, 우리가 지나가지 않은 경로에서도 모두 최적으로 행동을 해야 한다. 즉, 만일 지나가지 않은 경로에서 그렇지 않게 행동했을 것이라면 BI에 따라서 원래 플레이어의 결정도 지금과 같지 않을 것이기 때문.
- 이 문제를 앞서에 응용해보자. 플레이어가 C 를 하지 않고 D 를 했을 경우에도 GRIM의 지시에 따르는 것이 플레이어에게 가장 좋은가?

Conscientious GRIM?

- 앞서 정의한 GRIM을 다음처럼 다시 정의해보자.
 - ① plays C at $\tau = 1$
 - ② plays C , τ 이전 시기에 남의 행동이 모두 C 였다면
 - ③ plays D , 다른 모든 경우에
- 이렇게 정의된 전략은 SPE인가?

게임은 반복되어야 한다

- 정해진 횟수의 게임 : 역추론의 문제 발생
→ 무임승차 전략
- 고정된 확률로 게임이 반복
→ 다음회의 보복이 두려워 협조를 한다
→ 다음 회가 발생할 확률이 중요

상호성 가설에 의한 이타적 행위 모형 I

- 무조건 배신 (ALLD): 첫 회부터 배신하는 것으로 시작해서 계속 배신
- 무조건 협조 (ALLC): 첫 회부터 협조하는 것으로 시작해서 계속 배신

	1회	2회	3회	4회	5회	...	n회	...
경기자 I (ALLC)	C	C	C	C	C	...	C	...
경기자 II (ALLC)	C	C	C	C	C	...	C	...

i) ALLC vs ALLC

	1회	2회	3회	4회	5회	...	n회	...
경기자 I (ALLD)	D	D	D	D	D	...	D	...
경기자 II (ALLC)	C	C	C	C	C	...	C	...

ii) ALLD vs ALLC

상호성 가설에 의한 이타적 행위 모형 II

- 조건부 협조: 첫 회에 협조로 시작, 그 다음 회부터는 상대방이 전기에 협조를 한 경우에만 협조. 상대가 한 번이라도 배신하면 그 다음 부터 계속 배신

	1회	2회	3회	4회	5회	...	n회	...
경기자 I (조건부 협조, Grim)	C	C	C	C	C	...	C	...
경기자 II (조건부 협조, Grim)	C	C	C	C	C	...	C	...

i) Grim vs Grim

	1회	2회	3회	4회	5회	...	n회	...
경기자 I (배신)	D	D	D	D	D	...	D	...
경기자 II (조건부 협조, Grim)	C	D	D	D	D	...	D	...

ii) Defect vs Grim

상호성 가설에 의한 이타적 행위 모형 III

- 게임이 반복될 확률이 δ 일 때 총보수

	각회에서 얻어 지는 보수	δ 가 0.8일 때 게임이 이 때까지 지속될 확률	δ 가 0.2일 때 게임이 이 때까지 지속될 확률	반복 확률이 δ 일 때 게임이 이때까지 지속될 확률
1회	1	1	1	1
2회	1	0.8	0.2	δ^1
3회	1	0.8×0.8	0.2×0.2	δ^2
4회	1	$0.8 \times 0.8 \times 0.8$	$0.2 \times 0.2 \times 0.2$	δ^3
\vdots	\vdots	\vdots	\vdots	\vdots
총 보수 즉 각회 보수의 합		$\frac{1}{1-0.8} = 5$	$\frac{1}{1-0.2} = 1.25$	$\frac{1}{1-\delta}$

상호성 가설에 의한 이타적 행위 모형 IV

- $\frac{b-c}{1-\delta} > b \rightarrow \delta > \frac{c}{b}$

조건부협조 (Grim, Trigger) 전략이 효과적이려면 높은 확률로 반복되어야 !!

		경기자 II	
		C	D
경기자 I	C	$b-c, b-c$	$-c, b$
	D	$b, -c$	$0, 0$

(A) 1-Shot Game

		경기자 II	
		C	D
경기자 I	C	$\frac{b-c}{1-\delta}, \frac{b-c}{1-\delta}$	$-c, b$
	D	$b, -c$	$0, 0$

(B) Repeated Game

상호성 가설에 의한 이타적 행위 모형 V

- 루소의 사슴사냥 게임 (stag hunt game):

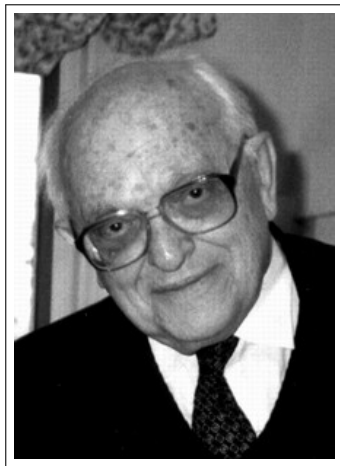
→ 조정게임

→ 죄수의 딜레마 게임이 반복되면 조정게임과 유사한 게임으로 변화

		사냥꾼 II			
		사슴		토끼	
사냥꾼 I	사슴	10,	10	0,	6
	토끼	6,	0	6,	6

Axelrod's Computer Simulation I

- 1980년, 정치학과 교수인 액셀로드는 몇몇 게임이론가들에게 200회 동안 치러질 PDG의 전략의 설계를 부탁.
- 이들을 각각을 토너먼트에 붙이고 점수를 부여. 이 대결의 승자는?
- 생물학자인 Anatol Rapoport의 "TIT FOR TAT (TFT)"
- 이 결과가 발표되자 사람들은 놀랐다. 왜?
제출된 프로그램들 중에서 가장 단순한 4



Axelrod's Computer Simulation II

- 이 사실이 학술지에 발표된 이후, 더 많고 다양한 전문가들에게 의뢰해 같은 모의 실험을 실시
- 두번째 라운드의 결과 역시 TFT가 가장 우수한 성적을 거두었다.
- 이후 RPDG 게임에서 협력을 가져온 좋은 해결책으로 TFT가 부상



Why TFT I

- ① Be nice:
먼저 배신하지는 않는다.
- ② Be provokable:
상대에게 호구가 되지 않는다.
- ③ Don't be envious:
상대보다 많은 보수를 바라지 않는다.
- ④ Don't be too clever:
충분히 예상 가능한 행보를 보여준다.



Why TFT II

- 점수가 매겨진 방식을 생각해보자.
- 사실 TFT는 상대를 완전히 이긴 경우는 별로 없었다. 이겨도 그리 큰 차이로 이기지 않고, 져도 그리 크게 손해보지 않는 것.
- 만일 스포츠의 방식처럼 승자독식 방식이었다면?
- 아울러, 컴퓨터 알고리즘은 실수를 하지 않는다. 하지만 인간은 어떠한가?



Trembling Hand I

- 눈에는 눈, 이에는 이 TFT:

한 번의 실수는 상대방의 다음 전략에 영향을 미쳐 (C,D), (D,C)
쌍이 반복

	1회	2회	3회	4회	5회	...
경기자 I (TFT)	C	C	C	C	C	...
경기자 II (TFT)	C	C	C	C	C	...

i) TFT vs. TFT

	1회	2회	3회	4회	5회	...
경기자 I (TFT)	C	C	D	C	D	...
경기자 II (TFT)	C	D	C	D	C	...

ii) 2회 때 실수로 C 대신 D

Trembling Hand II

- trigger 방아쇠 전략:

한 번의 실수는 영원한 보복을 촉발하여 (D,D) 쌍이 반복

	1회	2회	3회	4회	5회	...
경기자 I (Trigger)	C	C	C	C	C	...
경기자 II (Trigger)	C	C	C	C	C	...

i) Trigger vs Trigger

	1회	2회	3회	4회	5회	...
경기자 I (Trigger)	C	C	D	D	D	...
경기자 II (Trigger)	C	D	C	D	D	...

ii) 2회 때 실수로 C 대신 D

Trembling Hand III

- 'Win-Stay-Lose-Shift' (Pavlov 파브로프 전략):
 - (1) 내가 C 상대방도 C, 내가 만족 다음 기에도 C
 - (2) 내가 D 상대방은 C, 내가 만족 다음 기에도 D
 - (3) 내가 C 상대방은 D, 내가 불만족 전략을 바꿔서 D
 - (1) 내가 D 상대방은 D, 내가 불만족 전략 바꿔서 C

	1회	2회	3회	4회	5회	...
경기자 I (WSLS)	C	C	D	C	C	...
경기자 II (WSLS)	C	D	D	C	C	...

실제 사례 I

- 현실에서 찾을 수 있는 가장 좋은 사례?
- 삼성과 엘지, 진로와 하이트
- 이들은 경쟁관계이면서 협력관계
- 기업간 “짬짜미”는 반복 게임의 좋은 사례



실제 사례 II

- 덩어리가 너무 커서 배신에 따른 타격이 크다면?
- 이 거래들을 여러 단계로 쪼개서, 전번 거래의 정보를 이번 거래에 활용한다.
- 왜 마약거래는 대부분 자잘하게 이뤄지는가?



실제 사례 III

- 큰가시고기 (stickleback fish)의 협력?
- 포식자가 나타났을 때 이를 알아보기 위한 정찰이 필요
- 포식자에 대한 접근을 반복 게임으로 나타낼 수 있다.
- Milinski는 이 점에 착안하여 큰가시고기의 협력 실험을 고안



무한 반복게임은 언제나 좋은 결과를 낳을까?

- IRG 자체가 좋은 결과를 낳는 것은 아니다.
- 사실 IRG는 어떤 균형이든 가능하게 한다.
- 옆의 예를 보자. 1-shot NE은 어디인가?
- 만일 $\rho \geq \frac{1}{2}$ 이고, GRIM을 쓰면 (B, B) 도 균형으로 만들 수 있다!

	A	B	C
A	2, 2	2, 1	0, 0
B	1, 2	1, 1	-1, 0
C	0, 0	0, -1	-1, -1