

重複賽局

Repeated Games

1

重複賽局

- 玩家可能會因為“懲罰”的“威脅”（減少其長期收益, long term payoff）而無法利用其短期利益 (short-term advantage)。

4

智財權保護聲明

- 本影片及教材之內容，僅供修課學生個人使用，未經授課教師同意，不得以任何形式轉載、重製、散布、公開播送、出版或發行本影片之內容。如有侵權行為，需自負法律上之責任。

2

重複賽局

Repeated Games

- 企業之間相互合作及競爭
- 互相有利的交流(favor exchange)
- 動物之間的互惠(Reciprocal altruism)
- 買家和賣家關係

5

重複賽局

Repeated Games

當一組玩家重複互動時，每個玩家都可以在每個時間點，根據其他玩家的先前策略來制定自己的策略。

3

重複賽局

Repeated Games

- 僱傭關係
- 政治聯盟
- 朋友關係

6

a friendship as a repeated game,



7

the Repeated Prisoner's Dilemma



10

信譽 (Reputation)

- 用來描述一個人過去的行動，以影響未來的行為
- 為未來犧牲短期的利益。

8

重複的囚徒困境

- 重複的囚徒困境模擬兩個玩家之間需要合作的賽局。
- 重複的囚徒困境也被稱為“和平的戰爭遊戲”。

11

囚犯困局

	Cooperate 合作	Defect 背叛
Cooperate 合作	2,2	0,3
Defect 背叛	3,0	1,1

9

一回合的囚犯困局賽局

- 在囚徒困境，合作是劣勢策略，背叛是優勢策略， (Cooperating is strictly dominated by defecting)
- 無論其他玩家採取什麼策略，玩家採取背叛總是會獲得更大的收益。

12

兩回合重複的囚犯困境賽局

- 在兩次重複的囚徒困境賽局，玩家都知道每輪賽局最優的策略是告密。
- 唯一可能的納許均衡是要始終告密”。

13

N回合的囚犯困境賽局

- 兩位玩家應該一次又一次採取背叛，無論多少次玩這個賽局。

16

(DD, DD) is Nash Equilibrium

	CC	CD	DC	DD
CC	(-4,-4)	(-7,-3)	(-7,-3)	(-10,-2*)
CD	(-3,-7)	(-6,-6)	(-6,-6)	(-9,-5*)
DC	(-3,-7)	(-6,-6)	(-6,-6)	(-9,-5*)
DD	(-2*,-10)	(-5*,-9)	(-5*,-9)	(-8*,-8*)

14

不定回合的囚犯困境賽局(Indefinite Iteration of PD)

- 只有當玩家玩一個不確定次數的囚徒困境賽局，合作可能是一個平衡。
- 告密或背叛的動機可能因懲罰的威脅而取消。
- 當囚徒困境賽局無限重複，合作可能是一個子賽局完美均衡。
- http://www.youtube.com/watch?v=p3Uos2fzIJ0&feature=player_embedded

17

N回合重複的囚犯困境賽局

- 在 N 次重複的囚徒困境賽局，玩家都知道每輪賽局最優的策略是告密。
- 唯一可能的納許均衡是要始終告密”。

15

為什麼玩家要合作？

- 一次賽局：背叛是嚴格優勢策略
- 無限重複 (Infinitely repeated) 賽局：只要對方能保持合作就可以保持合作。

18

有限重複賽局

Finite Repeated game

- 玩家玩囚犯困境賽局 100次。
- 兩位玩家都知道賽局將玩 100次。
- 兩位玩家在下一個次賽局之前，都觀察以前的賽局。
- 他們可以維持合作嗎？

19

重複賽局

Repeated Game

- Player 1 DD DDD DDD DDDD D ...
- Player 2 DD DDD DDD DDDD D ...

22

Backward Induction

- 玩家將在 賽局 100 採取背叛策略。
- 玩家將在 賽局 99 採取背叛策略，依此類推。
- 納許均衡是玩家總是採取背叛策略

20

嚴峻觸發策略

Grim Trigger Strategy

- 只要其他玩家選擇 C ，則選擇 C 。
- 如果有其他玩家選擇背叛 D，則在隨後的每個時段中選擇背叛 D 。

23

有限次重複囚犯困境的賽局

- 每一個有限次重複的囚犯困境的賽局的納許均衡為在每一個賽局採取 (D, D) 。

21

嚴峻觸發策略
Grim Trigger Strategy

24

嚴峻觸發策略應用在男女關係

- 如果雙方都忠誠且值得信賴，合作就會持續下去。
- 如果一個夥伴欺騙或破壞信任，嚴酷觸發策略將導致另一個夥伴**永久結束關係或不再信任該夥伴**，即使最初的違規者道歉或試圖修復損害。

25

(Player 1 adopt Grim Trigger Strategy)

Player 1 C C D D D D D D D D D D...

Player 2 C D D D D D D D D D D D

28

嚴峻觸發策略應用的歷史事件

- Athenian Response to the Melian Revolt (416 BC)
- The Roman Empire's Punishment of Revolts
- The Treaty of Versailles (1919) - After World War I
- The Soviet Invasion of Hungary (1956) and Czechoslovakia (1968)

26

囚犯困局

	Cooperate 合作	Defect 背叛
Cooperate 合作	2,2	0,3
Defect 背叛	3,0	1,1

29

現代政治中很少使用Grim Trigger

- 雖然一些歷史和地緣政治局勢具有嚴峻觸發戰略的要素（其中一次叛逃會導致長期或永久性後果），
- 但由於存在升級和缺乏靈活性的風險，該戰略的使用很少。

27

嚴峻觸發策略（處罰策略） (Grim Trigger Strategy)

- 玩家1，2合作
- 如果玩家1在每一個時段選擇C，在每一個時段(C, C)，玩家獲得了報酬2。
- 如果玩家1選擇D，在此時段(D,C)獲得了報酬3，和在隨後(D,D)的報酬1。
- 但只要合作的未來報酬(2, 2, 2, ... 和玩家背叛後未來的報酬) \geq (3, 1, 1, ...)，她在每一個時期最好選擇C。

30

怨恨策略 The Grudger strategy

- 怨恨策略是只要別人給予合作，就願意寬恕和合作。
- 如果怨恨者遭遇背叛，在隨後的互動中拒絕與背叛者合作

31

無限次重複賽局

- 若且唯若貼現因子 δ (discount factor) 是至少 $\frac{1}{2}$ ，合作可以持續。
- 如果貼現因子超過 $\frac{1}{2}$ ，無一玩家有動機背叛。

34

$\delta \geq 1/2$, 玩家採取策略 C

- 如果玩家 2 採取嚴峻觸發策略
- 如果玩家 1 永遠採取策略 C，玩家1的報酬 = $2 + 2\delta + 2\delta^2 + 2\delta^3 + \dots = \frac{2}{1-\delta}$
- 如果玩家 1 在第一階段採取策略 D，玩家 1, 2 在第二階段以後都採取 (D, D)
- 玩家1的報酬 = $3 + \delta + \delta^2 + \delta^3 + \dots = 3 + \frac{\delta}{1-\delta}$
 $\frac{2}{1-\delta} \geq 3 + \frac{\delta}{1-\delta} \Rightarrow \delta \geq \frac{1}{2}$

32

囚犯困局的無限次重複賽局 (嚴峻的觸發策略)

- (D, D) 是一個同時行動囚犯困局賽局的惟一納許均衡。
- (C, C) 是玩家採取嚴峻的觸發策略，無限次重複囚犯困局的一個納許均衡。

35

如何求有限幾何級數的和

- $\delta + \delta^2 + \dots + \delta^{n-1} = (\delta - \delta^n) / (1 - \delta)$
- $1 + \delta + \delta^2 + \dots + \delta^{n-1} = (1 - \delta^n) / (1 - \delta)$
- $1 + \delta + \delta^2 + \dots + \delta^{n-1} = 1 / (1 - \delta)$
as $n \rightarrow \infty$

33

無限次重複囚犯困局的賽局

- 無限次重複的囚犯困局賽局產生的納許均衡：(D, D) 或 (C, C)。

36

無限重複囚徒困境 懲罰三個時段D的策略

- **Player 1** C C D D D C C C C C C C C ...
- **Player 2** C D C C C C C C C C C C C C C

37

以牙還牙應用在男女關係

- 如果雙方都忠誠、誠實、透明，他們就能保持合作和互信。
- 如果一個夥伴背叛了信任（例如撒謊或隱瞞某些事情），另一方可能會透過撤回信任或對抗他們來做出反應。
- 在道歉或採取糾正措施（恢復合作）之後，信任可以慢慢重建。這種方法允許錯誤發生，但鼓勵在糾正叛逃後迅速恢復相互尊重和誠實。

40

Tit for Tat



38

tit-for-tat vs. grim trigger strategy

- 以牙還牙(tit-for-tat)仍有合作的餘地
- 嚴峻的觸發策略(grim trigger strategy) 是一個無休止的背叛。

41

以牙還牙 (Tit for Tat)

- “等效報復 equivalent retaliation ”（“以牙還牙”），使用此策略的玩家將首先進行合作，然後對對手的先前策略做出回應。
- 如果對手以前是合作，則玩家採取合作的。如果對手不合作，則玩家採取不合作。

39

現代政治常用以牙還牙的策略

- 以牙還牙是現代政治中常見的策略，尤其是在國際關係中，互惠和報復往往定義國家之間的互動。
- 雖然它在某些情況下可以促進穩定和公平，但也帶來升級和僵局的風險，特別是在高風險衝突中。
- 通常需要外交靈活性和降級機制，以防止針鋒相對的循環失控。

42

以牙還牙應用於現代政治

- US-China Trade War (2018-present)
- **Russia-West Sanctions (2014-present)**
- Israel-Palestine Conflict (Ongoing)
- India-Pakistan Border Tensions (2019 Balakot Airstrikes)
- North Korea-South Korea Propaganda Battles (2010s)

TFT

<i>TFT AllC</i>	<i>TFT AllD</i>	<i>TFT Grim</i>	<i>TFT TFT</i>	<i>TFT Tester</i>
C C	C D	C C	C C	C D
C C	D D	C C	C C	D C
C C	D D	C C	C C	C C
C C	D D	C C	C C	C C
C C	D D	C C	C C	C C
C C	D D	C C	C C	C C
C C	D D	C C	C C	C C
⋮ ⋮	⋮ ⋮	⋮ ⋮	⋮ ⋮	⋮ ⋮

Tit for Tat
(Player 1 adopt Tic-for-Tat)

- Player 1 C D D C C C C C D D D D C...
- Player 2 D D C C C C C D D D D C...

雜訊使維持合作變得困難
Noise Makes it Difficult to Maintain Cooperation

- 兩個玩家都使用TFT。
- 一次事故或誤解可能會導致一連串的報復。
- C C
- C D(Noise)
- D (Retaliate) C
- C D (Retaliate)
- D (Retaliate) C

Tit-For-Two-Tats (TFTT)

- 僅在其他玩家連續兩次出現背叛時進行報復 (Retaliate)
- 可以忍受其他玩家一次的背叛，但易於利用其慷慨

無限次重複賽局

- 評估未來價值對有效的聲譽是必不可少。
- 當考慮是否在一個時期內選擇背叛，玩家們認為玷污自己的名聲將導致未來的損失。

囚犯困局

	Cooperate	Defect
Cooperate	3,3	0,5
Defect	5,0	1,1

49

收益 Payoffs

- 今天玩家 1 從背叛獲得的收益是 2
- 背叛造成的未來的損失是 $2p + 2p^2 + \dots + 2p^n$
- 玩家 1 進行採取合作, 如果
- 今天背叛獲得的收益 \leq 背叛造成的未來的損失
- $2 \leq 2p + 2p^2 + \dots + 2p^n = 2P/(1-p)$
- $p \geq 1/2$

52

無限囚徒困境 Indefinite Prisoner's Dilemma

- 下一個賽局繼續進行的概率為 p 。
- 賽局結束的概率為 $(1-p)$ 。

50

Grim Trigger Equilibrium

- 如果下一個賽局繼續進行的概率 ($p \geq 1/2$) 合作是“可信的”

53

收益 Payoffs

- 如果玩家 2 採取合作策略, 則玩家 1 也採取合作的收益為 $3 + 3p + 3p^2 + \dots + 3p^n$
- 如果玩家 1 第一時段 (the first period) 採取背叛, 兩個玩家都採取嚴峻的觸發策略, 玩家 1 的收益, 則為 $5 + p + p^2 + \dots + p^n$
- $(5 + p + p^2 + \dots + p^n) - 3 + 3p + 3p^2 + \dots + 3p^n$
- $= 2 - (2p + 2p^2 + \dots + 2p^n)$

51

第一次世界大戰中的塹壕戰 Trench warfare in World War I



54

第一次世界大戰中的塹壕戰 Trench warfare in World War I

- 戰爭是殺人，但是和平有時還是會出現，因為敵對塹壕的士兵達成休戰。
- 上午8：00-9：00的時間被視為私人時間。
- 進餐時不射擊。
- 在敵人的補給線上不發射砲彈。
- 但是這種合作是如何實現的？

55

第一次世界大戰中的塹壕戰 Trench warfare in World War I

		German Soldiers	
		Kill	Miss
Allied Soldiers	Kill	(2, 2)	(6, 0)
	Miss	(0, 6)	(4, 4)

58

unexpected peace during war, showing cooperation between soldiers from opposing trenches.



56

這些收益來自何處？

- (Miss, Kill) 的收益 (0, 6) 的收益，因為
- $u_{\text{Allied}} = 4 + 2 * 0 - 4 * 1 = 0,$
- $u_{\text{German}} = 4 + 2 * 1 - 4 * 0 = 6$
- (Kill, Kill) 的收益 (2, 2)，因為
- $u_{\text{Allied}} = 4 + 2 * 1 - 4 * 1 = 2,$
- $u_{\text{German}} = 4 + 2 * 1 - 4 * 1 = 2$

59

第一次世界大戰中的塹壕戰 Trench warfare in World War I

- 我們可以假設每個士兵都重視殺死敵人，但更重視不被殺死。
- 也就是說，士兵的收益是
 $4 \text{ (不被殺死)} + 2 \text{ (敵軍被殺)} - 4 \text{ (本軍被殺)}$

57

(Kill, Kill) is the unique NE

- (Kill, Kill) 是的唯一NE。
- 兩個玩家的嚴格優勢策略都是殺人

60

Trench warfare
infinitely repeated version

- 考慮時段 t 和每個玩家都曾經合作過的先前歷史（即，沒有玩家選擇“Kill”）。

- $4 + 4\delta + 4\delta^2 + \dots = \frac{4}{1-\delta}$

61

阿克塞爾羅德的比賽
Axelrod's Tournament

- 1980年，密歇根大學政治學教授（Robert Axelrod）舉辦了一場針對囚徒困境的各種策略比賽。

64

Trench warfare
infinitely repeated version

- 如果您選擇 Kill，則您的收益為
- $6 + 2\delta + 2\delta^2 + \dots$
 $= 6 + \frac{2\delta}{1-\delta}$

62

阿克塞爾羅德的比賽
Axelrod's Tournament



65

Trench warfare
infinitely repeated version

- $\frac{4}{1-\delta} \geq 6 + \frac{2\delta}{1-\delta}$
- $4 \geq 6(1-\delta) + 2\delta$
- $\delta \geq \frac{1}{2}$
- 如果 $\delta \geq 1/2$ ，則選擇 Miss 是無限重複博弈的 SPNE

63

以牙還牙 (TIT FOR TAT strategy)

- Axelrod 錦標賽的獲勝者是 TIT FOR TAT 策略。
- 該策略在第一個步驟中進行合作，然後執行對手在前一個步驟中所做的一切。
- 當與全背叛策略相匹配時，TIT FOR TAT 策略在第一步之後總是會出現背叛。
- 當與全力合作的策略相匹配時，TIT FOR TAT 總是合作。

66

以牙還牙(TIT FOR TAT strategy)

- 這種策略的好處是既可以與友善的對手進行合作，獲得合作的全部好處，又可以在與有背叛的對手匹配時選擇背叛。
- 當與對手是以牙還牙相匹配時，TIT FOR TAT策略將始終保持合作。

67

全背叛 All-Defect

- 全背叛：此策略在每次every turn 都會存在背叛。
- 這是最安全的策略，因為它無法被利用。
- 但是，它錯過了與準備合作的對手合作獲得更大收益的機會。

70

以牙還牙 (TIT FOR TAT strategy)

- TAT FOR TAT的假設是其對手正試圖最大化自己的得分。
- 當與“RANDOM”之類的盲目策略結合使用的對手時，“TIT FOR TAT”將陷入對手的境地。

68

騙子 Cheater

- 「騙子」通常用於描述社會互動中的參與者，他們利用他人的合作來謀取個人利益，而不為集體福祉做出貢獻

71

以牙還牙 (TIT FOR TAT strategy)

- 不能將TIT FOR TAT稱為“最佳”策略。
- 必須認識到，對於囚徒困境，確實沒有“最佳”策略。
- 為了獲勝，玩家必須弄清楚對手的策略，然後選擇最適合情況的策略

69

全合作 Always cooperate

- 全合作：與自身相匹配時，此策略效果很好。
- 但是，如果對手選擇背叛，那麼這種策略將表現不佳。

72

傻瓜 Suckers

- 個人或實體，他們合作或為他人提供利益而不接受任何回報。
- 「傻瓜」與涉及互惠利他 reciprocal altruists 或 tit-for-tat behaviors 的策略形成鮮明對比行為

73

利他主義者 Altruists

- 利他主義是對他人福祉的無私關心。
- 利他主義涉及造福他人而不期望個人利益或回報的行動或行為。

76

互惠利他主義者 reciprocal altruists

- 互惠利他主義是指個體相互合作和幫助，期望在未來某個時刻得到幫助作為回報的社會行為形式。
- 通常涉及「以牙還牙」的方法，即個人以類似的方式對其他人的行為做出反應。合作是有回報的，不合作可能會導致未來得不到幫助。

74

隨機 Random

- 該策略在50%的時間內進行合作。

77

互惠 Reciprocity

- 「互惠」是與他人交換物品以實現互惠互利的做法。
- 它涉及相互給予和接受或相互合作。
- 互惠並不一定意味著直接的一對一的回應；隨著時間的推移，它可能涉及更廣泛的合作。

75

賽局結束的概率 P vs 玩家的前瞻性 δ

- 下一個賽局繼續進行的概率為 p 。
- 賽局結束的概率為 $(1-p)$ 。
- 隨著 δ 的增加，玩家的前瞻性 forward-looking 會逐漸提高。

78

References

- GTO-5-01: Repeated Games
- <https://www.youtube.com/watch?v=GfvgGUmFh9k&list=PLeY-lFPWgBTiRqTRkhu7bjbPr7ITOG3Jk>
- Game Theory 9. Asymmetric information (part 1: Cheap Talk equilibrium vs. Signalling)
- <https://www.youtube.com/watch?v=EjSI-CMYzkM>