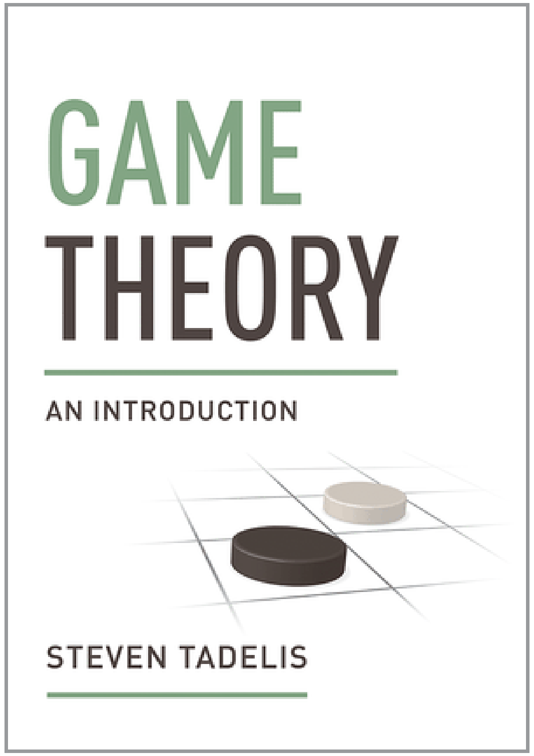
課程名稱：賽局理論與應用

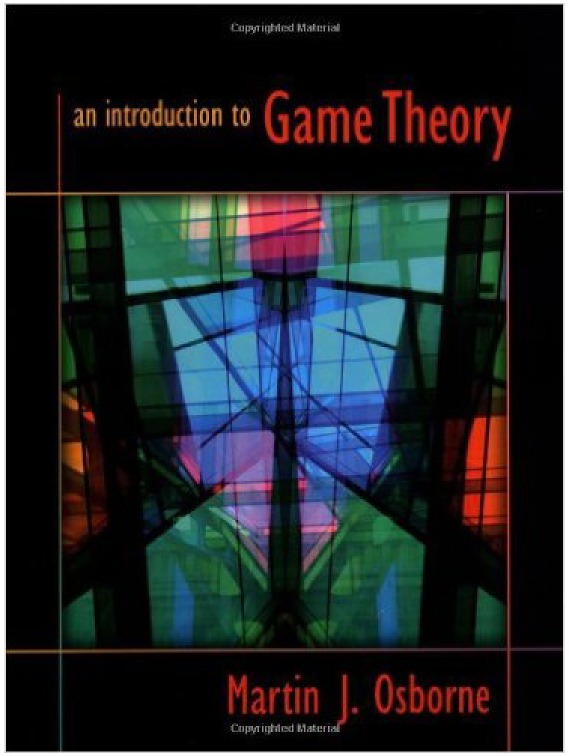
Game Theory and Applications

Topic of This Lecture: Cooperation by Infinitely Repeated Games

授課教師: 周孜燦
國立中山大學電機系
聯絡方式: ztchou@ee.nsysu.edu.tw

References for This Chapter





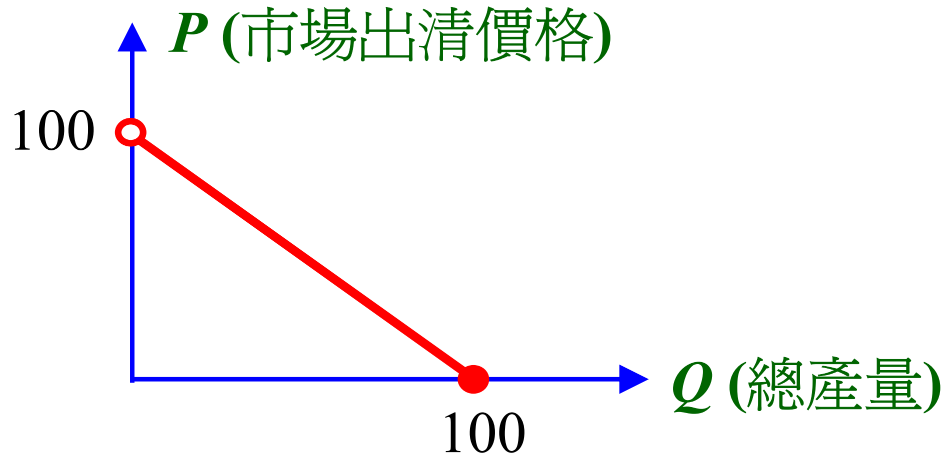
- Steven Tadelis, *Game Theory: An Introduction*, **Chapter 10**, Princeton University Press, 2013.
- Martin J. Osborne, *An Introduction to Game Theory*, **Chapters 14 and 15**, Oxford University Press, 2003.
- 這二本書，網路上皆找得到電子檔

回顧：古諾雙寡頭產量賽局

Antoine Cournot
(1801 – 1877) 古諾，
法國經濟學家

假設某個產品目前僅有二個廠商 F_1 和 F_2 在生產。假設生產該產品的成本爲 10。假設 F_1 打算生產 q_1 個，F_2 打算生產 q_2 個。假設根據專家估計及市場經驗，要將 $q = q_1 + q_2$ 個產品銷售一空，售價應爲 $P(q) = \max\{100 - q, 0\}$。那麼 q_1 和 q_2 的值應該如何設定才能使得 F_1 和 F_2 都最大化自己的 payoff？

註：假設 $\boldsymbol{q_1}$ 和 $\boldsymbol{q_2}$ 的值皆可爲實數。



合謀產量與均衡產量

◆ 假設廠商 F_i 和 F_j 分別生產 q_i 和 q_j 個產品，那麼 F_i 的 payoff 爲 $u_i(q_i,q_j) = \underbrace{(100 - q_i - q_j)}_{\text{售價}} q_i - \underbrace{10q_i}_{\text{成本}} = (90 - q_i - q_j)q_i$. 若廠商間自由競爭，皆採用 Nash 均衡策略，那麼每個廠商將生產 30 個產品。此時每個產品的出清價格爲 40, 每個廠商的獲利爲 900.

註：如果忘了如何求出 Nash 均衡產量，請自行複習先前的教材。

◆ 假設 F_i 公司派人去遊說 F_j 公司，希望二家公司合起來只生產 q 個產品 (聯合壟斷), 那 q 的值 (壟斷產量) 應該多少最好呢？

要讓 $u(q) = (100 - q)q - 10q$ 的值最大，令 $\dfrac{du(q)}{dq} = 90 - 2q = 0$

$\Rightarrow$ 當 $q = 45$ 時，二家公司合起來總共賺 $u(q) = 2025$.

◆ 因此若廠商合謀，各自生產壟斷產量的一半：22.5 個 (合謀產量)，此時每個產品的出清價格爲 55, 每個廠商的獲利爲 1012.5.

Quantity Competition：Prisoner's Dilemma Game

		F_2	
		合謀產量 22.5（合作）	均衡產量 30（背叛）
F_1	合謀產量 22.5（合作）	1012.5, 1012.5	843.75, 1125
	均衡產量 30（背叛）	1125, 843.75	900, 900

◆ 我們剛才看到：在合謀產量下和在均衡產量下，產品的價格差很多。爲避免損害消費者權益，很多國家都有制訂反壟斷法 (台灣稱爲公平交易法)：若廠商「簽約」共同制訂產品的價格或產量，政府將苛以重罰.

◆ 因此，如上表所示，廠商間的產量競爭賽局將有如囚犯困境賽局：如果雙方能透過「放話（遊說）」達成合作，彼此皆可獲利 1012.5。但若雙方互相背叛，那麼彼此的獲利都只有 900。

◆ 我們當然知道「放話」是無效的，除非雙方至少進行二次賽局，而且另一個不同的賽局有二個 payoff 不同的 NE。但現實情況裡頭，公司間的產量競爭，例如「面板」廠商間的競爭，通常會一直持續，而且這些廠商不太可能「一起」改玩別的賽局。那這樣廠商之間還有辦法達成合作嗎？

Finitely Repeated Game

Stage (*T*-1)		player 2	
		c (合作)	d (背叛)
player 1	C (合作)	4, 4	−1, 5
	D (背叛)	5, −1	**1, 1**

Stage *T*		player 2	
		c (合作)	d (背叛)
player 1	C (合作)	4, 4	−1, 5
	D (背叛)	5, −1	**1, 1**

[定理一] 假設一個多階段賽局 $G(T,\delta)$ 的每一個階段皆是相同的 normal form game G, 且 G 只有一個 Nash 均衡 s. 那麼 strategy profile $s^* = \underbrace{s \cdots s}_{\text{重複 } T \text{ 次}}$ 爲多階段賽局 $G(T,\delta)$ 唯一的 SPNE。

[證明] 用倒推法（backward induction）證明。前一章，我們已證明 players 在第 T 階段必須採用 Nash 均衡策略 s. 由於每個 player 在第 T 階段所能獲得的 payoff 已經固定， 無法改變，因此 players 無法根據第 $T-1$ 階段的賽局結果來決定第 T 階段的行動；如此一來，players 在第 $T-1$ 階段所能採取的最佳策略只剩 Nash 均衡策略 s. 依此類推，$s^* = \underbrace{s \cdots s}_{\text{重複 } T \text{ 次}}$ 爲賽局 $G(T,\delta)$ 唯一的 SPNE。◆ 這意味即使囚犯困境賽局重複 T 次，players 也不會合作

Cooperation in a Infinitely Repeated Game？

經濟日報

ECONOMIC DAILY NEWS

友達奇美涉壟斷　南韓開罰

各面板廠遭罰金額

廠商	罰款金額（億韓元）
三星電子及其台、日分公司	972.9

根據定理一，如果二人賽局 *G* 只有一個 Nash 均衡，且雙方合作並非 Nash 均衡，那麼即使賽局 *G* 重複進行 *T* 次，players 之間仍然不可能合作。那麼在重複賽局中，players 要合作，似乎只剩一個可能性：*T* 必須爲無限大！

就產量競爭來說，以面板產業爲例，台灣最大的廠商包含<u>友達</u>和<u>奇美</u>（後來和<u>群創</u>合併）。若沒意外，奇美和友達將會一直重複進行類似囚犯困境的產量賽局。這樣雙方就會達成合作嗎？有很高的可能性！因爲我們多次在媒體上看到面板廠商壟斷市場的報導。因此我們接下來要介紹無限重複賽局

Total Payoff of a Player

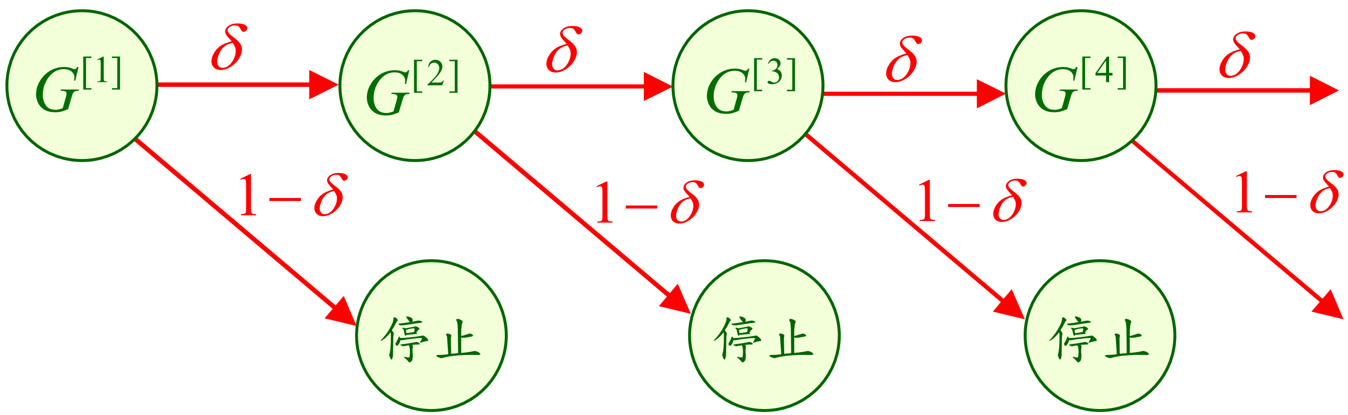
		player 2	
		c (合作)	d (背叛)
player 1	C (合作)	4, 4	−1, 5
	D (背叛)	5, −1	1, 1

◆ 我們用符號 $G(\delta)$ 表示相同的 normal-form game G 一直重複進行，永遠都不會停下來。我們稱這樣的賽局爲無限重複賽局。爲了解說方便，我們用符號 $G^{[k]}$ 表示第 k 階段的 normal-form game。當然 $G^{[k]}=G$. 假設 player i 預期可在 $G^{[k]}$ 期間獲得 payoff $u_i^{[k]}$, 那麼 player i 預期可獲得的 total payoff 現值爲 $u_i = u_i^{[1]} + u_i^{[2]}\delta + \cdots + u_i^{[T]}\delta^{T-1} + \cdots = \sum_{k=1}^{\infty} u_i^{[k]}\delta^{k-1}$ ....(1)

其中 $0<\delta<1$ 表示 payoff 的 discount rate（折現率）.

◆ 註：我們希望 $\delta \neq 0$, 否則 $u_i = u_i^{[1]}$, 這相當於只有一階段的賽局。另方面，我們要求 $\delta \neq 1$, 否則以無限囚犯困境賽局爲例，player 1 即使永遠選擇背叛，每階段的 payoff 都大於 0，total payoff 即可無限大，不會有誘因想要合作！

Another Interpretation of Discount Rate

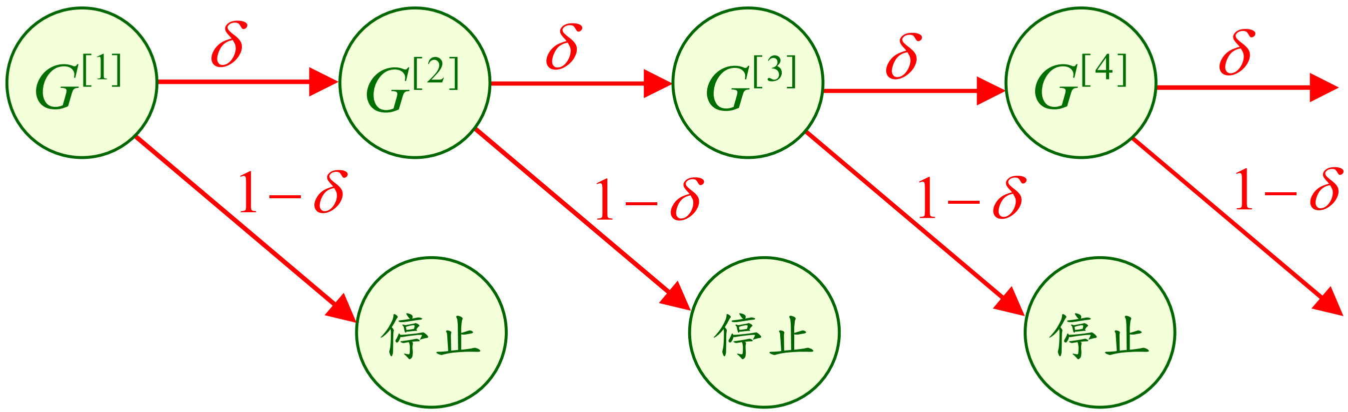


◆ δ 的另一種解釋：也許有人認爲「在現實生活中，賽局不可能永遠一直進行的阿！」. 因此，如上圖，我們可以假設折現率爲 1，然後把 δ 解釋爲「在每個階段的賽局結束後，players 有 δ 的機率繼續進行下個階段的賽局，有 $1-\delta$ 的機率結束賽局」. 在這種情況下，player i 可獲得的 total payoff 的期望值爲 $u_i = u_i^{[1]} + \left[\delta u_i^{[2]} + (1-\delta)\times 0\right] + \left[\delta^2 u_i^{[3]} + \delta(1-\delta)\times 0\right] + \cdots$

$= \sum_{k=1}^{\infty} u_i^{[k]} \delta^{k-1}$ .... (2). 我們可發覺公式 (1) 和 (2) 完全相同。

◆ 在這種解釋之下，賽局永遠一直持續進行的機率爲 $\lim_{T\to\infty} \delta^T = 0,$ 這表示重複賽局必定會在有限時間內結束 (只是沒人知道何時會結束)。

Average Payoff in a Round



◆ 我們剛才看到，無限重複賽局 $G(\delta)$ 可解釋爲：每個階段的賽局 G 結束後，players 有 δ 的機率繼續進行下個階段的賽局 G, 有 $1-\delta$ 的機率結束賽局。在這種情況下，令 T 表示賽局總共進行的階段個數。那麼 T 爲隨機變數，且我們有 $\Pr[T=k]=\delta^{k-1}(1-\delta)$. 所以平均而言，賽局總共進行的次數爲 $E[T]=\sum_{k=1}^{\infty} k\times\Pr[T=k]=\sum_{k=1}^{\infty} k\delta^{k-1}(1-\delta)=\dfrac{1}{1-\delta}$.

◆ 這意味平均而言，player i 可在每個階段獲得的 payoff 期望値爲

$$v_i=\frac{u_i}{E[T]}=(1-\delta)\sum_{k=1}^{\infty} u_i^{[k]}\delta^{k-1}.$$

History of a Repeated Game

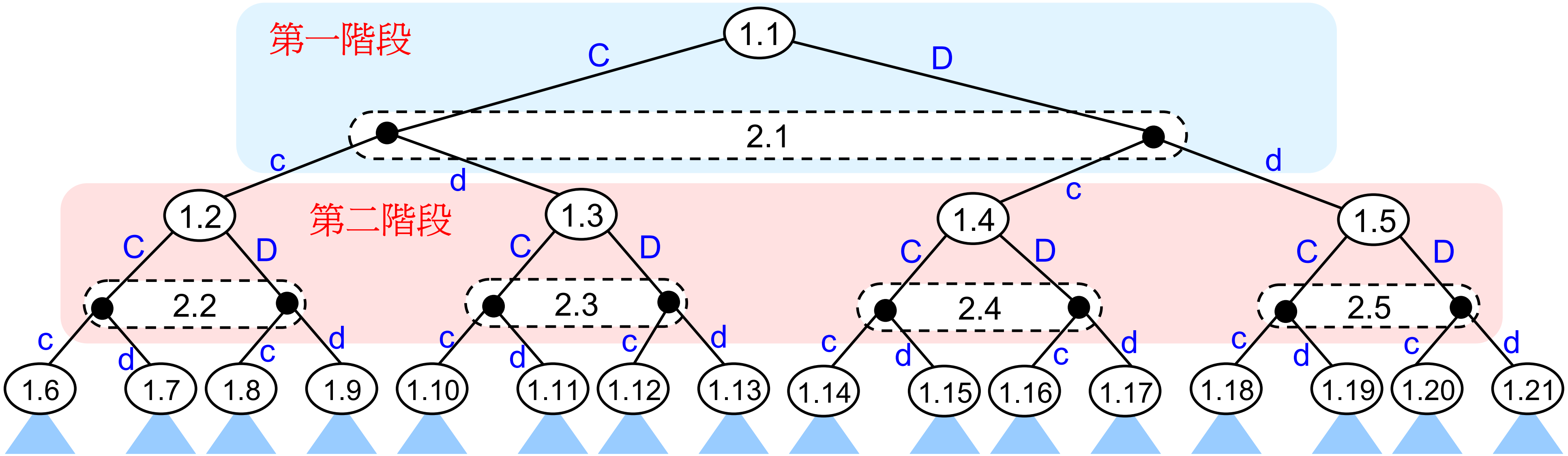
Stage 1		player 2	
		c (合作)	d (背叛)
player 1	C (合作)	4, 4	−1, 5
	D (背叛)	5, −1	**1, 1**

Stage 2		player 2	
		c (合作)	d (背叛)
player 1	C (合作)	4, 4	−1, 5
	D (背叛)	5, −1	**1, 1**

◆ 給定 n 個 players，令 S_i 表示 player i 的 strategy set。在本章中，我們假設 players 在每個階段的賽局 G 皆採取 pure strategy，並且假設 G 裡頭存在有純策略 Nash 均衡，以簡化問題的分析。令 $H = S_1 \times \cdots \times S_n$. 假設 players 在第 k 個階段所採取的 strategy profile $s^{[k]} = \left(s_1^{[k]}, \ldots, s_n^{[k]}\right) \in S$. 那麼我們說 players 在前 k 個階段的歷史為 $h^{[k]} = s^{[1]}s^{[2]} \cdots s^{[k]} \in H^k$.

◆ 以二階段囚犯局為例，$S_1 = \{C, D\}, S_2 = \{c, d\}$, $S = \{(C,c),(C,d),(D,c),(D,d)\}$. 假設在第一階段，player 1 採用 C, player 2 採用 d, 則 $s^{[1]} = (C,d) \in H$. 假設在第二階段，player 1 採用 D, player 2 採用 c, 則 $s^{[2]} = (D,c) \in H$. 那麼 players 在前 2 個階段的歷史 $h^{[2]} = s^{[1]}s^{[2]}$ 為 $((C,d),(D,c)) \in H^2$.

Description of a Strategy



◆ 如上圖，我們可用 extensive-form game 的型式來表達無限重複賽局 $G(\delta)$. 在無限重複賽局 $G(\delta)$ 所對應的 game tree 裡頭，root 到達某個 info. set 只有唯一的一條路徑，且該路徑對應到唯一的一段歷史。舉例來說，只有 players 的歷史爲 (Cd) 才可到達 info. sets 1.3 或 2.3，只有 players 的歷史爲 (Cd, Dc) 才可到達 info. set 1.12.

◆ 在 extensive-form game 裡頭，player 的策略的定義是：描述 player 在所有屬於他的每一個 info. set 上所做的選擇。既然 info. set 和 players 的歷史有一對一的對應，因此 player i 在第 k 階段賽局的策略爲 $s_i^{[k]} : H^{k-1} \rightarrow S_i$.

Playing NE in Each Stage Constitutes an SPNE

[定理二] 給定 n 個 players 的無限重複賽局 $G(\delta)$, 令 $s^* = (s_1^*, s_2^*, \ldots, s_n^*)$ 爲賽局 G 的 Nash 均衡。令 $H^{[1..\infty]} = \bigcup_{k=1}^{\infty} H^k$ 表示 $G(\delta)$ 中所有可能的歷史結果。假設在 $G(\delta)$ 中，對所有 $h \in H^{[1..\infty]}$, player i 的策略皆爲 $s_i(h) = s_i^*$. 那麼 strategy profile $\left(s_1(h), s_2(h), \ldots, s_n(h)\right)$ 爲賽局 $G(\delta)$ 的 SPNE 之一。

註： $G(\delta)$ 的 SPNE 未必只有一個，至少因爲賽局 G 的 NE 未必只有一個。

[證明] 因爲每位 player 的策略都和歷史無關，所以每位 player 都不用考慮在目前階段所採取的行動會對未來產生任何影響。在任何一個階段的賽局 G 中，給定其他 players 的策略爲 s_{-i}^* 的情況下，player i 的最佳策略爲 s_i^*. 因此 strategy profile $\left(s_1(h), \ldots, s_n(h)\right) = \left(s_1^*, \ldots, s_n^*\right)$ 爲賽局 $G(\delta)$ 的 SPNE。

◆ 這意味在無限囚犯困境賽局中，每位 player 都採取「永遠背叛」的策略爲 SPNE .... 那 players 在無限囚犯困境賽局中到底有沒有辦法合作呢？

Carrots in a Infinitely Repeated Game

在「有限」階段的賽局裡頭，最後一個階段的賽局必須同時有「胡蘿蔔」和「棒子」才能促使 players 合作。在 T 階段的囚犯困境賽局裡頭（假設 $T < \infty$），由於第 T 個階段的賽局只有一個 NE：互相背叛，也就是說，最後階段的賽局喪失了胡蘿蔔，只剩棒子， 所以 players 無法合作。

然而在無限囚犯困境賽局裡頭，如果 δ 夠大，想像 $\delta = 1$，那就沒有「最後一個階段」的存在：在第 T 個階段的賽局結束之後，雙方還可以進行下個階段的賽局。在下個階段，players 還可以繼續合作，那麼「永遠有機會繼續合作」就形成了胡蘿蔔，用以促進合作。

Sticks in a Infinitely Repeated Game

Case (a)：雙方永遠合作

階段	1	2	3	4	5	6	7	…
player 1	C	C	C	C	C	C	C	…
payoff	4	4	4	4	4	4	4	…
player 2	c	c	c	c	c	c	c	…
payoff	4	4	4	4	4	4	4	…

Case (b)：player 2 在第 3 階段率先背叛

階段	1	2	3	4	5	6	7	…
player 1	C	C	C	D	D	D	D	…
payoff	4	4	−1	1	1	1	1	…
player 2	c	c	d	d	d	d	d	…
payoff	4	4	5	1	1	1	1	…

多階段賽局除了要有「胡蘿蔔」以吸引 players 合作之外，還要有「棒子」才能防止 players 背叛。囚犯困境賽局裡頭的棒子就是「背叛」這個選項。

上圖顯示在無限囚犯困境賽局裡頭，殘酷觸發策略利用「棒子」來防止 players 背叛的情況：假設 players 一開始合作。假設如 case (b) 所示，player 2 在第三階段率先背叛。對照 case (a)，選擇背叛可以讓 payoff 增加 1（從 4 變成 5）。然而一旦 player 2 開始背叛，player 1 從此以後都選擇背叛，這樣 player 2 在往後的每個階段 payoff 都會損失 3（從 4 降成 1）…..可預期，當 $\delta \to 1$，這個損失就會 $\to \infty$，具有相當大的嚇阻作用。

Failure of Backward Induction

按「胡蘿蔔與棒子」的原理，對於二人無限重複囚犯困境賽局 $G(\delta)$, 我們可設計「殘酷觸發策略 (grim-trigger strategy) $s_{GT}=(s_1,s_2)$」, 運作方式如下：

◆ Player 1 的策略 s_1: $s_1^{[1]}=C$. 也就是第一階段採取合作。對於第 $k>1$ 階段，若 $h^{[k-1]}=\underbrace{((C,c),(C,c),\ldots,(C,c))}_{(C,c)\text{ 重複出現 } k-1 \text{ 次}}$, 則 $s_1(h^{[k-1]})=C$; 否則 $s_1(h^{[k-1]})=D$.

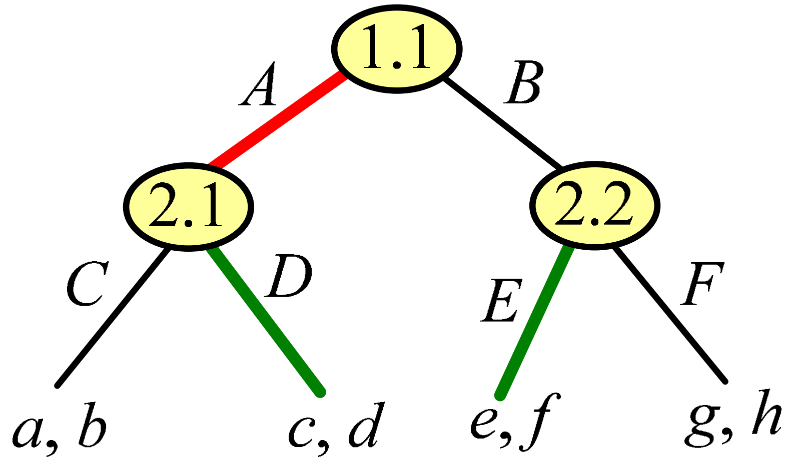
也就是說，如果之前的歷史是雙方一直合作則繼續合作；否則若 player 1 或 player 2 其中一方曾經背叛，就選擇背叛。

◆ Player 2 的策略 s_2: 和 player 1 相同，意即 $s_2^{[1]}=C$, 且對於 $k>1$, 若 $h^{[k-1]}=\underbrace{((C,c),(C,c),\ldots,(C,c))}_{(C,c)\text{ 重複出現 } k-1 \text{ 次}}$, 則 $s_2(h^{[k-1]})=C$; 否則 $s_2(h^{[k-1]})=D$.

◆ 如何證明雙方採用殘酷觸發策略爲 SPNE？ 在有限賽局中，我們採用 backward induction (倒推法) 來證明。然而在無限重複賽局裡頭，並沒有「最後階段」的存在，那要怎麼證明？ 此時我們需要「一階段偏離原則」

One-Stage Deviation Principle (1/9)

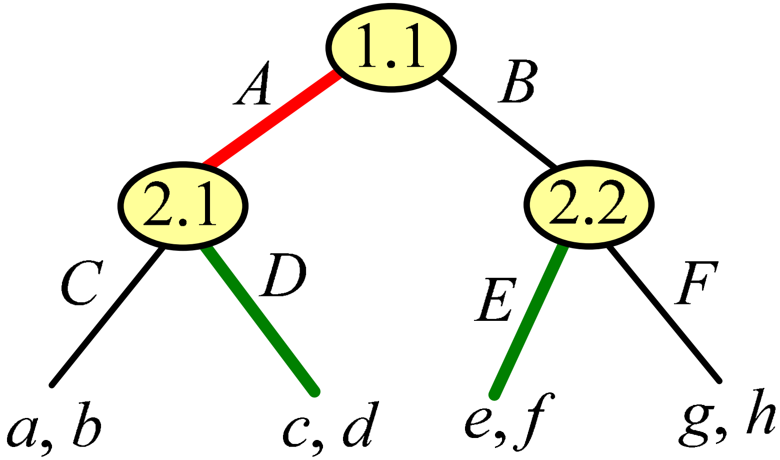
[定理三] 對於 perfect-information extensive-form game 而言，一個 strategy profile 為 SPNE ⇔ 沒有單獨一個 player 能夠只在某一個屬於他的 info. set 上面改變選擇而獲得更高的 payoff。



(解說 ⇒) 若 strategy profile 為 SPNE，表示每個 player 都在所有 subgames 裡頭採用 NE 策略，當然沒有 player 能夠單獨改變他的選擇而獲得更高利益.

（解說 ⇐）如圖，給定 strategy profile $s = (s_1; s_2) = (A; DE)$, 如果 player 1 無法「單獨」在 info. set 1.1 上面改變選擇而提高 payoff，那麼 $c \geq e$ 必須成立；如果 player 2 無法「單獨」在某一個 info. set（2.1 或 2.2）上面改變選擇而提高 payoff，那麼 $d \geq b,\ f \geq h$ 必須成立。

One-Stage Deviation Principle (2/9)



$G_{2.1}$

player 2	
C	**D**
a, b	**c, d**

$G_{2.2}$

player 2	
E	**F**
e, f	**g, h**

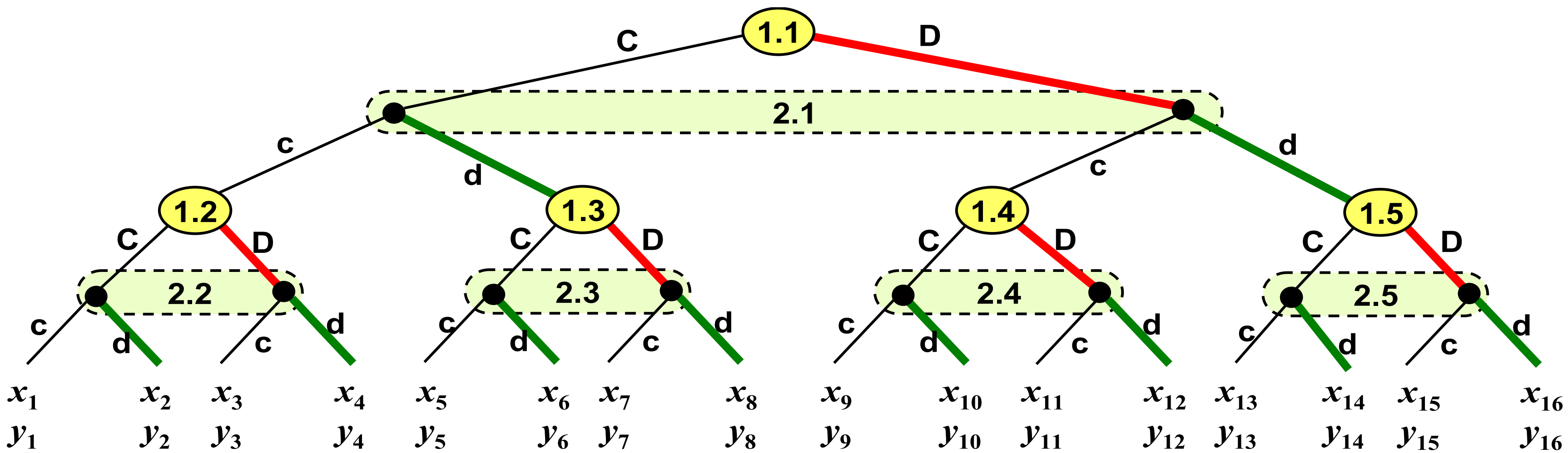
$G_{1.1}$

		player 2			
		CE	**CF**	**DE**	**DF**
player 1	**A**	**a, b**	**a , b**	**c, d**	**c, d**
	B	**e, f**	**g, h**	**e, f**	**g, h**

給定 info. set $i.s$, 令 $G_{i.s}$ 表示 root 爲 $i.s$ 的 subtree 所對應的 normal-form subgame。我們要檢驗 strategy profile $(s_1; s_2) = (A; DE)$ 是否使得 players 在每個 subgame 皆採用 NE 策略。我們剛才推導出 $c \ge e,\ d \ge b,\ f \ge h.$ 對於 $G_{2.1}$, 由於 $d \ge b$, 所以 player 2 採用 D 爲 NE 策略。對於 $G_{2.2}$, 由於 $f \ge h$, 所以 player 2 採用 E 爲 NE 策略。觀察 $G_{1.1}$, 我們有 $DE \in BR_2(A) = \{DE, DF\}$, $A = BR_1(DE)$, 因此 (A, DE) 爲 NE. 所以 $(A; DE)$ 爲整個 game tree 的 SPNE.

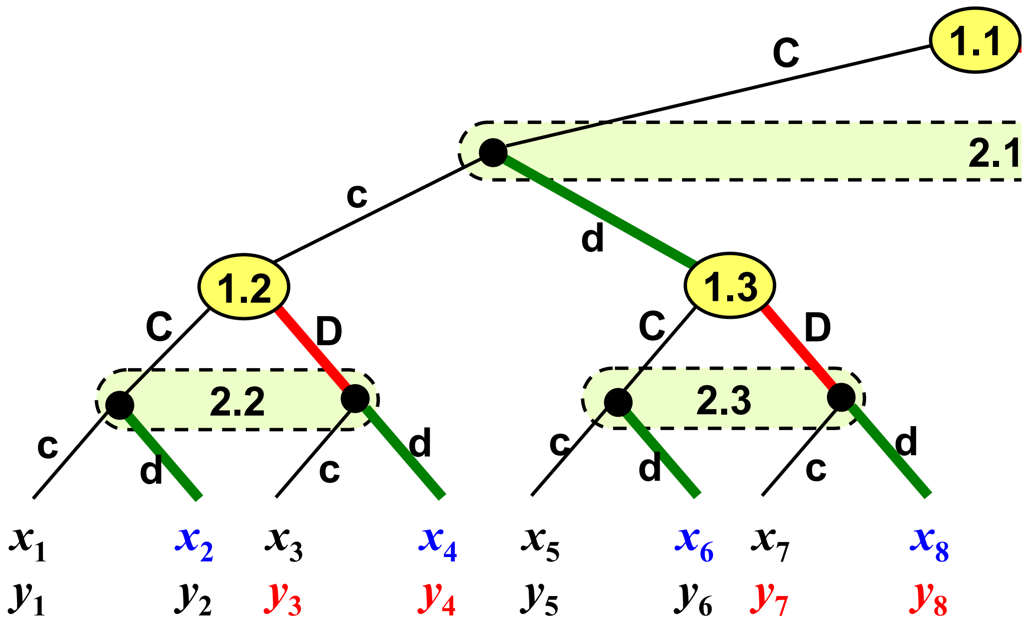
One-Stage Deviation Principle (3/9)

[定理四] 對於有限多階段賽局，一個 strategy profile 為 SPNE ⇔ 沒有 player 能只在某一個階段改變選擇而獲得更高的 payoff，不論該階段之前的歷史為何



(解說 ⇒) 按 SPNE 定義，當然沒有 player 能夠單獨在某個階段改變選擇而獲得更高的利益. (解說 ⇐) 如圖，我們以二階段囚犯賽局為範例來解說。假設 player 1 無法只在某一個階段改變選擇而獲得更高的 payoff，「不論該階段之前的歷史為何」.....所以 player 1 有可能處在 info. set 1.1, 1.2, 1.3, 1.4 或 1.5 其中一個。例：如果之前的歷史為 ∅, 那麼 player 1 處在 info. set 1.1; 如果之前的歷史為 Cc, 那麼 player 1 處在 info. set 1.2. 所以此一定理相當於是說：每個 player i 都無法單獨在某一個 info. set $i.s$ 改變選擇而獲得更高 payoff

One-Stage Deviation Principle (4/9)

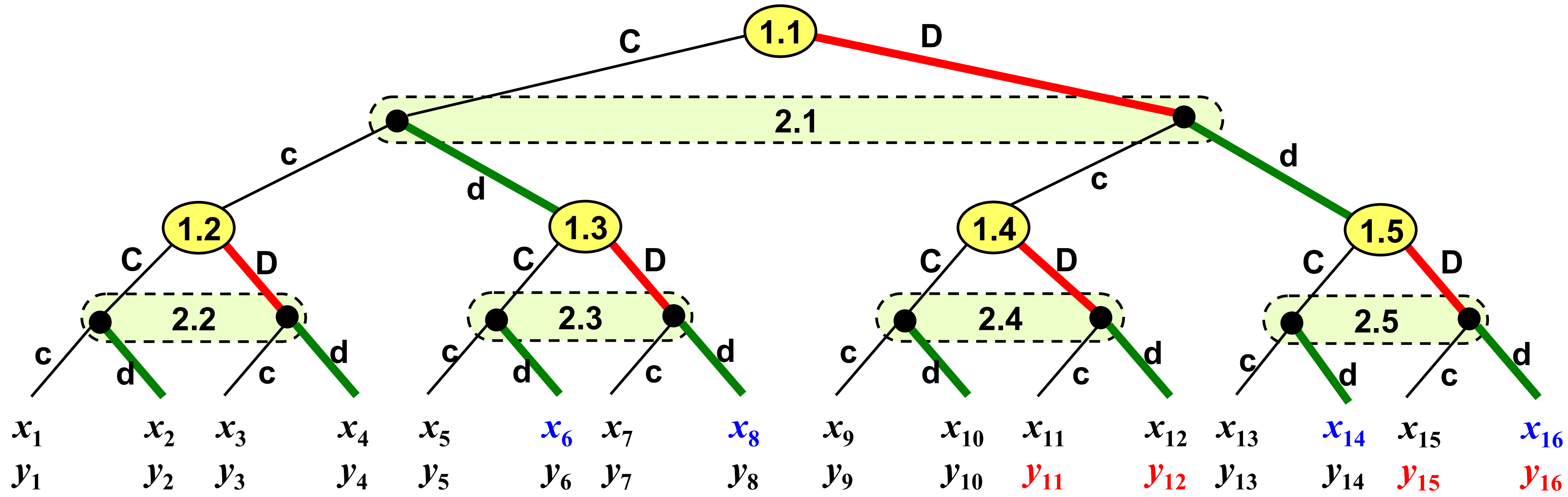


$G_{1.2}$

		player 2	
		c	d
player 1	C	x_1, y_1	x_2, y_2
	D	x_3, y_3	x_4, y_4

令 $\Gamma_{i.s}$ 表示 root 爲 info. set $i.s$ 的 subtree, 而 $G_{i.s}$ 表示 subtree $\Gamma_{i.s}$ 所對應的 normal-form subgame。給定 game tree 的 strategy profile $s = (s_1; s_2) = (D^{[1]}D^{[2]};\ d^{[1]}d^{[2]})$, 假設 player 1 無法單獨在 info. set 1.2 改變選擇而獲得更高 payoff, 則 $x_4 \geq x_2$. 假設 player 2 無法單獨在 info. set 2.2 改變選擇而獲得更高 payoff, 則 $y_4 \geq y_3$. 那麼，對 $G_{1.2}$ 來說, $BR_1(d) = D$, $BR_2(D) = d$. 所以 (D,d) 爲 $G_{1.2}$ 的 NE. 同理可推，$x_8 \geq x_6$, $y_8 \geq y_7$; $x_{12} \geq x_{10}$, $y_{12} \geq y_{11}$; $x_{16} \geq x_{14}$, $y_{16} \geq y_{15}$; 且 $G_{1.3}$, $G_{1.4}$, $G_{1.5}$ 的 NE 皆爲 (D,d).

One-Stage Deviation Principle (5/9)



假設 player 1 無法單獨在 1.1 改變選擇而獲得更高的 payoff; 因此 $x_{16} \geq x_8$.

假設 player 2 無法單獨在 2.1. 改變選擇而獲得更高的 payoff; 因此 $y_{16} \geq y_{12}$.

參考上圖，令 $D_{1.1}$ 表示 player 1 在 info. set 1.1 採用 D. 我們有

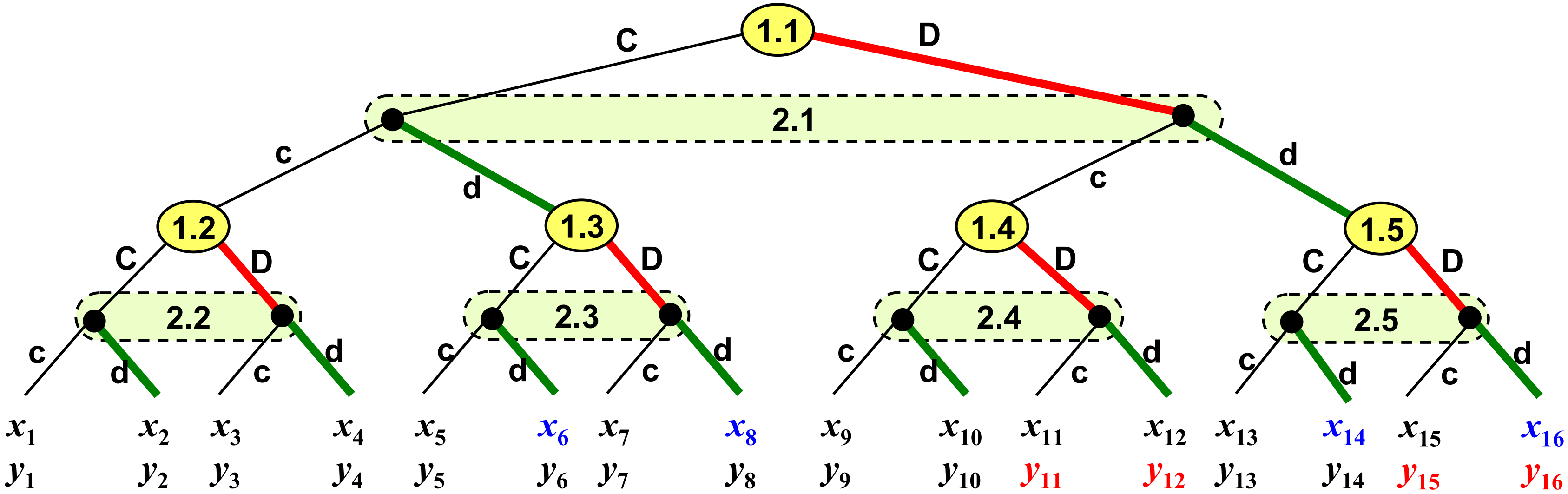
$d_{2.1}d_{2.2}d_{2.3}d_{2.4}d_{2.5} \in BR_2(D_{1.1}D_{1.2}D_{1.3}D_{1.4}D_{1.5})$ 因爲 $y_{16} = \max\{y_{11}, y_{12}, y_{15}, y_{16}\}$,

且 $D_{1.1}D_{1.2}D_{1.3}D_{1.4}D_{1.5} \in BR_1(d_{2.1}d_{2.2}d_{2.3}d_{2.4}d_{2.5})$ 因爲 $x_{16} = \max\{x_6, x_8, x_{14}, x_{16}\}$.

亦即 players 1 和 2 在 $G_{1.1}$ 採用 NE 策略。結論：players 1 和 2 在所有 subgames 皆採用 NE 策略. 所以 $(D^{[1]}D^{[2]}; d^{[1]}d^{[2]})$ 爲整個 game tree 的 SPNE.

One-Stage Deviation Principle (6/9)

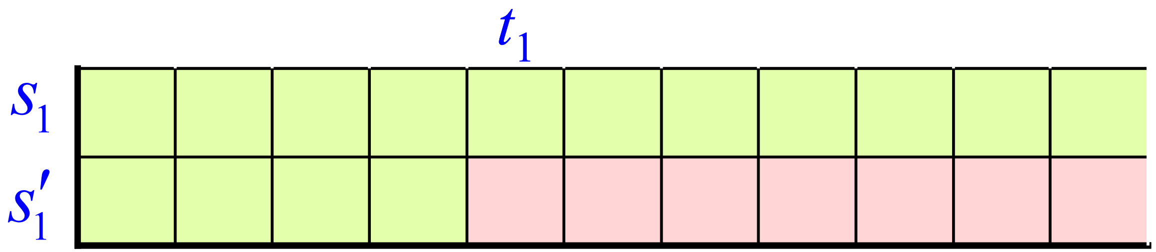
[引伸四] 對於有限多階段賽局，一個 strategy profile 爲 SPNE ⇔ 沒有單獨一個 player 能夠透過有限階段的改變而獲得更高的 payoff



(解說 ⇐) 所謂的「有限階段」當然包含至少一個階段。根據定理四：給定 strategy profile $s=(s_1;s_2)$, 若每個 player 都無法單獨透過一階段的改變來提高 payoff，那麼 $(s_1;s_2)$ 爲 SPNE。(解說 ⇒) 以二階段囚犯賽局爲例，假設 $(s_1;s_2)=(D^{[1]}D^{[2]};\ d^{[1]}d^{[2]})\ =(D_{1.1}D_{1.2}D_{1.3}D_{1.4}D_{1.5};\ d_{2.1}d_{2.2}d_{2.3}d_{2.4}d_{2.5})$. 從定理四的證明可知 $s_1 \in BR_1(s_2)$. 這意味：只要 s_2 不變，那麼不論賽局歷史爲何、不論 player 1 如何改變 s_1, player 1 都無法提高 payoff；同理可推 player 2 的情況。

One-Stage Deviation Principle (7/9)

[定理五] 對於無限重複賽局，一個 strategy profile 爲 SPNE $\Leftrightarrow$ 沒有單獨一個 player 能只在某一個階段改變選擇而獲得更高的 payoff，不論該階段之前的歷史爲何。

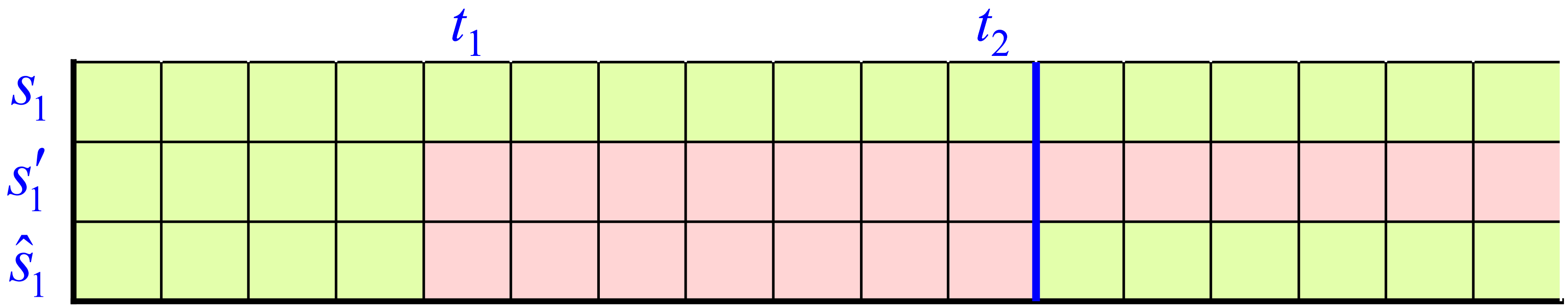


（解說 $\Leftarrow$ ）採用反證法。假設 $s=(s_1;s_2)$ 爲無限重複賽局 $G(\delta)$ 的一個 strategy profile，滿足一階段偏離原則，但卻不是 SPNE。由於 s 不爲 SPNE，所以我們假設：在 s_2 不變的情況下，player 1 有另一個策略 s_1' 可獲得比 s_1 更高的 payoff。因此我們有

$$u_1(s_1';s_2)-u_1(s_1;s_2)>2\varepsilon \text{ ....(1)}$$

其中 $\varepsilon>0$. 而且，如上圖，我們假設從第 t_1 階段開始，s_1' 採用的選擇就和 s_1 不完全一樣 (因此用不同顏色表示)，且不一樣的次數無限多；否則按引伸四，公式 (1) 不成立。

One-Stage Deviation Principle (8/9)



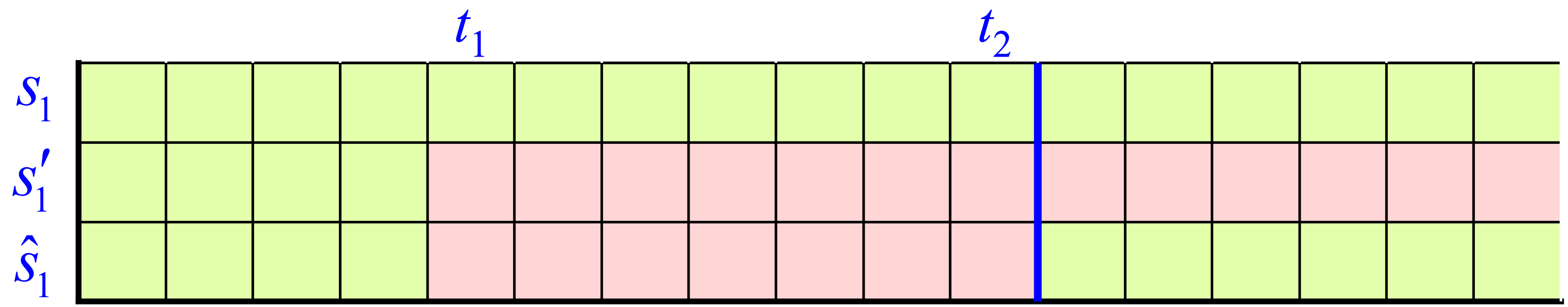
現在，我們爲 player 1 製造的另一個策略 $\hat{s}_1$: 在 $t_1 \sim t_2$ 階段期間都採用和 s_1' 一樣的選擇；在其餘階段則採取和 s_1 一樣的選擇. 令 $u_1^{[k]}(s_1)$ 表示 player 2 的策略爲 s_2 的情況下，player 1 採取 s_1 在第 k 階段所獲得的 payoff。所以

$$\left|u_1(s_1';s_2) - u_1(\hat{s}_1;s_2)\right| \le \left|u_1^{[t_2+1]}(s_1') - u_1^{[t_2+1]}(\hat{s}_1)\right|\delta^{t_2} + \left|u_1^{[t_2+2]}(s_1') - u_1^{[t_2+2]}(\hat{s}_1)\right|\delta^{t_2+1} + \cdots$$

$\le \dfrac{c\delta^{t_2}}{1-\delta}$, 其中 $c = 2\left|\max\limits_{x\in A_1, y\in A_2}\{u_1(x,y)\}\right|$ 且 A_i 表示每回合的 normal form game 裡頭 player i 的 action set. 以右圖的囚犯困境賽局來說, $A_1 = \{C, D\}$, $A_2 = \{c, d\}$, $\max\limits_{x\in A_1, y\in A_2}\{u_1(x,y)\} = 5.$

		player 2	
		c (合作)	d (背叛)
player 1	C (合作)	4, 4	−1, 5
	D (背叛)	5, −1	1, 1

One-Stage Deviation Principle (9/9)



剛才，我們推導出 $\left|u_1(s_1';s_2)-u_1(\hat{s}_1;s_2)\right| \le \dfrac{c\delta^{t_2}}{1-\delta}$. 我們可調整 t_2 的值使得 $\dfrac{c\delta^{t_2}}{1-\delta}<\varepsilon$, 例如：設定 $t_2=\left\lfloor \ln_\delta \dfrac{c(1-\delta)}{\varepsilon}-1 \right\rfloor$. 此時我們有

$u_1(s_1';s_2)-u_1(\hat{s}_1;s_2)<\varepsilon$ ....(2) 且 $u_1(s_1';s_2)-u_1(\hat{s}_1;s_2)>-\varepsilon$ ....(3).

透過 $-1\times(2)+(1)$, 我們可得 $u_1(\hat{s}_1;s_2)-u_1(s_1;s_2)>\varepsilon$ ....(4)

◆ 由於從第 t_2+1 階段開始，s_1 和 $\hat{s}_1$ 在每階段都採用相同的選擇，因此讓我們考慮我們 $1\sim t_2$ 階段的有限賽局。我們假設 $s=(s_1;s_2)$ 滿足一階段偏離原則，所以在 s_2 不改變的情況下，player 1 無法在此一有限賽局透過有限階段的改變而提高 payoff。但這與 (4) 矛盾，所以假設錯誤。得證。

SPNE：Grim-Trigger Strategy (1/2)

		player 2	
		c (合作)	d (背叛)
player 1	C (合作)	4, 4	−1, 5
	D (背叛)	5, −1	1, 1

					k			
s_2	c	c	c	d	d	d	d	d
s_1 payoff	C 4	C 4	D 5	D 1	D 1	D 1	D 1	D 1
s'_1 payoff	C 4	C 4	D 5	D 1	C −1	D 1	D 1	D 1

◆ 要證明殘酷觸發策略 $s_{GT} = (s_1; s_2)$ 爲 SPNE，就必須證明沒有單獨一個 player（假設是 player 1）能夠在某一個階段（假設爲第 $k \geq 2$ 階段）改變選擇而獲得更高的 payoff，不論第 k 階段之前的歷史爲何。

◆ 由於在殘酷觸發策略裡頭，每個 player 所採取的行動只和上一階段賽局結果有關，因此我們可以用「歷史的最後一階段」對歷史做分類。

Case 1. 假設第 k 階段之前的歷史是以 (C,d), (D,c), 或 (D,d) 作結尾。右表是第 $k-1$ 階段賽局結果爲 (D,d) 的例子。我們用符號 s_1' 表示「player 1 在第 k 階段改用 C, 但從第 $k+1$ 階段開始，仍然遵守殘酷觸發策略」. 對照 s_1 和 s_1', 我們可發現：player 1 改用 s_1' 無法提高 payoff。

SPNE：Grim-Trigger Strategy (2/2)

					k			
s_2	c	c	c	c	c	c	c	c
s_1 payoff	C 4	C 4	C 4	C 4	C 4	C 4	C 4	C 4
s_2	c	c	c	c	c	d	d	d
s'_1 payoff	C 4	C 4	C 4	C 4	D 5	D 1	D 1	D 1

Case 2. 假設第 k 階段之前的歷史是以 (C,c) 作結尾。

我們用符號 s_1' 表示「player 1 在第 k 階段改用 D, 但從第 $k+1$ 階段開始，仍然遵守殘酷觸發策略」. 對於 $(s_1';s_2)$, players 雙方自第 $k+1$ 階段起，都會採用「背叛」. 令 $v_1^{[k]}(s_1)$ 表示 player 1 在採用 s_1 的情況下，自第 k 個階段起的 average payoff（in a round）. 所以 $v_1^{[k]}(s_1) = 4$, $v_1^{[k]}(s_1') = (1-\delta)(5+\delta+\delta^2+\cdots) = 5-4\delta$. 按一階段偏離原則，若殘酷觸發策略爲 SPNE, 則需滿足條件 $4 \ge 5-4\delta$.

$\Rightarrow$ 所以當 $\delta \ge 1/4$, 殘酷觸發策略爲無限囚犯困境賽局之 SPNE。

注意事項

◆ 我們剛才在使用一階段偏離原則時，我們只檢查「player 2 策略不變，但 player 1 改變某一個階段」的情況。事實上，我還應該額外檢查「player 1 策略不變，但 player 2 改變某一個階段」的情況。

◆ 我們之所以可以只檢查「player 2 策略不變，但 player 1 改變某一個階段」的情況，這是因爲 player 1 和 player 2 的角色是對稱的，可以互換，意即 (1) player 1 和 player 2 採用相同的策略，且 (2) 每階段的 normal form game 的 payoff 矩陣爲對稱矩陣。

◆ 給定 payoff 矩陣 $\mathbf{U} = \begin{bmatrix} (a_{11}, b_{11}) & \cdots & (a_{1n}, b_{1n}) \\ \vdots & \ddots & \vdots \\ (a_{m1}, b_{m1}) & \cdots & (a_{mn}, b_{mn}) \end{bmatrix}$, 若 $m = n$ 且

$$\begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix}^T = \begin{bmatrix} b_{11} & \cdots & b_{n1} \\ \vdots & \ddots & \vdots \\ b_{1n} & \cdots & b_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix},$$ 則 $\mathbf{U}$ 爲對稱矩陣.

Win-Stay, Lose-Shift Strategy

		player 2	
		c (合作)	**d (背叛)**
player 1	**C (合作)**	***R*, *R***	***S*, *T***
	D (背叛)	***T*, *S***	***P*, *P***

		player 2	
		c (合作)	**d (背叛)**
player 1	**C (合作)**	**4, 4**	**−1, 5**
	D (背叛)	**5, −1**	**1, 1**

階段	1	2	3	4	5	6	…
player 1	**C**	**C**	**D**	**D**	**C**	**C**	**…**
payoff	**4**	**4**	**5**	**1**	**4**	**4**	**…**
player 2	**c**	**c**	**c**	**d**	**c**	**c**	**…**
payoff	**4**	**4**	**-1**	**1**	**4**	**4**	**…**

◆ 除了殘酷觸發策略之外，是否還有其他的 SPNE 策略？讓我們看一下「贏則繼續，輸則改變 (win-stay lose-shift，簡寫成WSLS)」策略。

◆ 如左上表的 payoff 矩陣所示，只要滿足 $T > R > P > S$, 那麼此一賽局便是囚犯困境賽局。我們用 R 表示「reward 獎勵」, T 表示「temptation 背叛的誘惑」, S 表示「sucker 傻子」, 而 P 表示「punishment 受懲罰」.

◆ WSLS策略：第一階段採用 C. 若上階段的 payoff 爲 R 或 T, 則此階段採用和上階段相同行動; 若上階段的 payoff 爲 S 或 P, 此階段採用和上階段不同行動.

右表中，players 皆採用 WSLS 策略，只是 player 1 在第三階段選錯選成 D.

SPNE：Win-Stay, Lose-Shift Strategy (1/2)

				k			
s_2	c	c	d	d	c	c	c
s_1 payoff	C 4	C 4	C -1	D 1	C 4	C 4	C 4
s_2	c	c	d	d	d	c	c
s'_1 payoff	C 4	C 4	C -1	C -1	D 1	C 4	C 4

第 ***k*–1** 階段的賽局結果為 ***(C, d)***

				k			
s_2	c	c	c	d	c	c	c
s_1 payoff	C 4	C 4	D 5	D 1	C 4	C 4	C 4
s_2	c	c	c	d	d	c	c
s'_1 payoff	C 4	C 4	D 5	C -1	D 1	C 4	C 4

第 ***k*–1** 階段的賽局結果為 ***(D, c)***

要證明雙方採用 WSLS 策略 $s_{WSLS}=(s_1;s_2)$ 為 SPNE，就必須證明沒有單獨一個 player（假設是 player 1）能夠在某一個階段（假設為第 $k\geq 2$ 階段）改變選擇而獲得更高的 payoff，不論第 k 階段之前的歷史為何。

Case 1. 假設第 k 階段之前的歷史是以 (C,d) 或 (D,c) 作結尾。

左表是第 $k-1$ 階段賽局結果為 (C,d) 的例子。觀察左表下方，我們用符號 s_1' 表示「player 1 在第 k 階段改用 C, 但從第 $k+1$ 階段開始，仍然遵守 WSLS 規則」的策略。對照 s_1 和 s_1', 我們可發現：player 1 改用 s_1' 無法提高 payoff。

SPNE：Win-Stay, Lose-Shift Strategy (2/2)

				k			
s_2	c	c	c	c	c	c	c
s_1 payoff	C 4	C 4	C 4	C 4	C 4	C 4	C 4
s_2	c	c	c	c	d	c	c
s'_1 payoff	C 4	C 4	C 4	D 5	D 1	C 4	C 4

第 $k-1$ 階段的賽局結果為 (C, c)

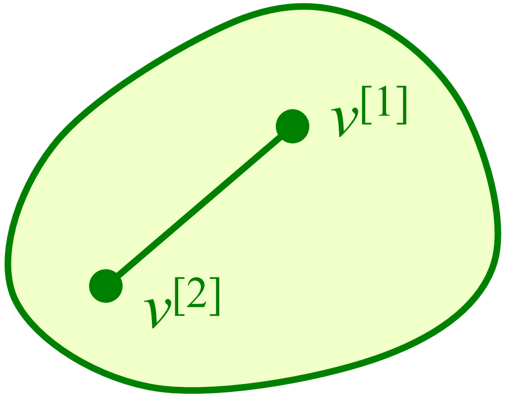
				k			
s_2	c	c	d	c	c	c	c
s_1 payoff	C 4	C 4	D 1	C 4	C 4	C 4	C 4
s_2	c	c	d	c	d	c	c
s'_1 payoff	C 4	C 4	D 1	D 5	D 1	C 4	C 4

第 $k-1$ 階段的賽局結果為 (D, d)

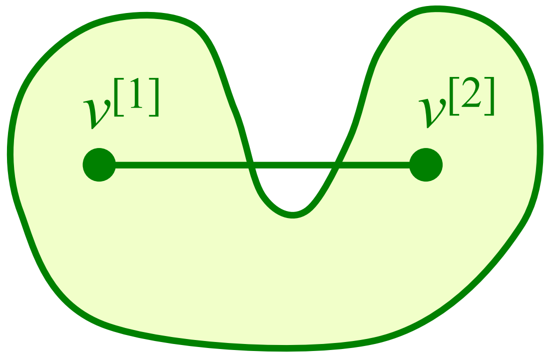
我們要證明沒有單獨一個 player（假設是 player 1）能夠在某一個階段（假設為第 k 階段）改變選擇而獲得更高的 payoff，不論第 k 階段之前的歷史為何。

Case 2. 假設第 k 階段之前的歷史是以 (C,c) 或 (D,d) 作結尾. 考慮左表，令 s_1' 表示「player 1 在第 k 階段改用 D, 但從第 $k+1$ 階段開始，仍然遵守 WSLS 規則」的策略. 那麼 $v_1^{[k]}(s_1) = (1-\delta)(4+4\delta+4\delta^2+4\delta^3+\cdots)$, $v_1^{[k]}(s_1') = (1-\delta)(5+\delta+4\delta^2+4\delta^3+\cdots)$. 按定理五，若 WSLS 策略為 SPNE，則需滿足 $4+4\delta \geq 5+\delta$. [結論] 當 $\delta \geq 1/3$, WSLS 策略為無限囚犯困境賽局之 SPNE。

Convex Set (凸集合)



convex set（凸集合）



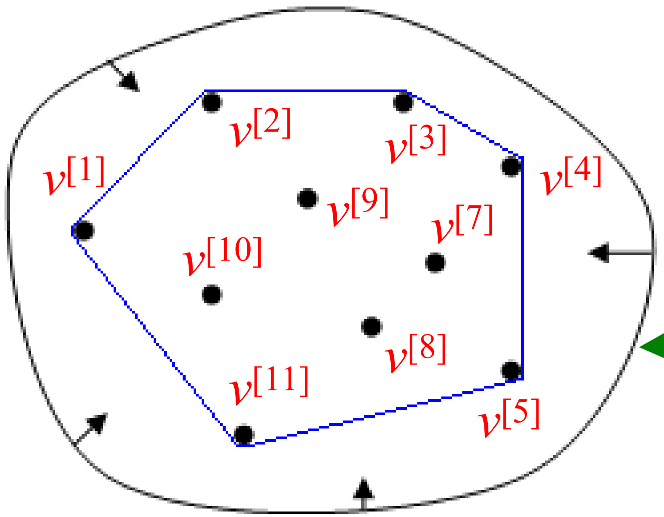
concave set（凹集合）

◆ 對於無限囚犯困境賽局，我們已經看到三種 SPNE，分別是雙方同時採用永遠背叛、殘酷觸發策略、或 WSLS 策略。我們不禁好奇：到底有幾種不同的 SPNE？答案是無窮多. Why？讓我們先看 convex set 和 convex hull 的定義.

◆ 若二維平面上的一個集合 $C \subseteq \mathbb{R}^2$ 被稱爲是 convex set，表示給定 C 裡頭的任意二個點 $v^{[1]}$ 和 $v^{[2]}$, 那麼線段 $\overline{v^{[1]}v^{[2]}} = \left\{ v = \lambda v^{[1]} + (1-\lambda)v^{[2]} \;\middle|\; 0 \le \lambda \le 1 \right\}$ 也必須落在 C 裡頭。注意：一個「線段」也是一個 convex set。

註：如果一個集合不是 convext set，我們稱它爲 concave set。

Convex Hull (最小凸多邊形)



這條具有可縮性的黑色線圈所圍出來的區域爲包含 V 的 convex set。我們將線圈一直往內縮，直到線圈碰到 V 中的點爲止。此時藍色線圈所圍起來的區域稱爲「由 V 所構成的 convex hull」

◆ 給定 n 個點所構成的集合 $V=\{v^{[1]},\ldots,v^{[n]}\}$, 我們說集合 $CH(V)$ 爲由 V 所構成的 convex hull，表示 $CH(V)$ 是包含 V 的最小 convex set。

◆ $CH(V)$ 的正式數學定義爲：

$$CH(V)=\left\{ v \;\middle|\; v=\sum_{i=1}^{n}\alpha_i v^{[i]},\ \text{其中 } v^{[i]}\in V,\ \alpha_i\geq 0,\ \text{且 } \sum_{i=1}^{n}\alpha_i=1 \right\}.$$

◆ 也就是：給定 $CH(V)$ 裡頭的一個點 v, 必定存在 n 個非負實數 $\alpha_1,\ \alpha_2,\cdots,\ \alpha_n$ 使得 $\sum_{i=1}^{n}\alpha_i=1$ 且 $\sum_{i=1}^{n}\alpha_i v^{[i]}=v.$

符號介紹

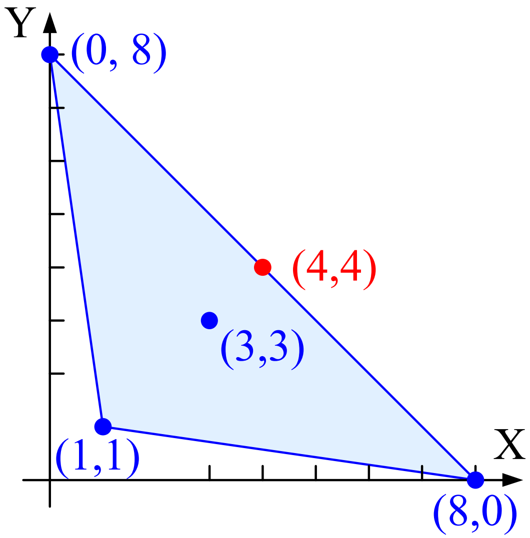
		player 2	
		B_1 = c	B_2 = d
player 1	A_1 = C	3, 3	0, 8
	A_2 = D	8, 0	1, 1

		player 2				
		B_1	$\cdot\cdot$	B_j	$\cdot\cdot$	B_m
player 1	A_1					
	$\vdots$					
	A_i			$u(A_i, B_j) = v^{[i,j]}$		
	$\vdots$					
	A_n					

◆ 給定二個 players 的 normal form game G, 我們假設 player 1 的 strategy set 爲 $A=\{A_1,A_2,\ldots,A_n\}$ 且 player 2 的 strategy set 爲 $B=\{B_1,B_2,\ldots,B_m\}$.

◆ 對於賽局 G, 假設 player 1 和 player 2 分別採用行動 A_i 和 B_j, 那麼此一賽局所產生的 payoff profile 爲 $u(A_i,B_j)$. 爲了讓符號表達更簡單，我們用 $v^{[i,j]}$ 表示 $u(A_i,B_j)$. 例：左表中，$u(A_1,B_2)=(0,8)=v^{[1,2]}$.

讓 Payoff 逼近凸邊形內某一點 (1/2)：合作方法



		player 2	
		$B_1 = \text{c}$	$B_2 = \text{d}$
player 1	$A_1 = \text{C}$	3, 3	0, 8
	$A_2 = \text{D}$	8, 0	1, 1

雙方永遠合作

階段	1	2	3	4	5	6	7	8	…
player 1	D	C	D	C	D	C	D	C	…
payoff	8	0	8	0	8	0	8	0	…
player 2	c	d	c	d	c	d	c	d	…
payoff	0	8	0	8	0	8	0	8	…

給定 payoff matrix 裡頭的所有 payoff profiles $V = \{(3,3),\ (8,0),\ (0,8),\ (1,1)\}$, 左上圖藍色區域即爲 $CH(V)$. 給定 $CH(V)$ 裡頭的點，比如 (4,4), 必定存在非負實數 $\alpha_1 \sim \alpha_4$ 使得 $\sum_{i=1}^{4} \alpha_i = 1$ 且 $\alpha_1 \times u(C,c) + \alpha_2 \times u(D,c) + \alpha_3 \times u(C,d) + \alpha_4 \times u(D,d) = (4,4)$. 例：$\alpha_1 = \alpha_4 = 0$ 且 $\alpha_2 = \alpha_3 = 1/2$. 那麼，在無限囚犯困境賽局中，如果在奇數階段，雙方採取 (D,c); 在偶數階段，雙方採取 (C,d), 那麼當 $\delta \to 1$, players 平均每個階段的 payoff profile 也就趨近於 (4,4).

讓 Payoff 逼近凸邊形內某一點 (2/2)：懲罰措施

Case (a)：雙方永遠合作

階段	1	2	3	4	5	6	7	8	…
player 1	**D**	**C**	**D**	**C**	**D**	**C**	**D**	**C**	**…**
payoff	**8**	**0**	**8**	**0**	**8**	**0**	**8**	**0**	**…**
player 2	**c**	**d**	**c**	**d**	**c**	**d**	**c**	**d**	**…**
payoff	**0**	**8**	**0**	**8**	**0**	**8**	**0**	**8**	**…**

Case (b)：**player 1** 在第 **6** 階段率先背叛

階段	1	2	3	4	5	6	7	8	…
player 1	**D**	**C**	**D**	**C**	**D**	**D**	**D**	**D**	**…**
payoff	**8**	**0**	**8**	**0**	**8**	**1**	**1**	**1**	**…**
player 2	**c**	**d**	**c**	**d**	**c**	**d**	**d**	**d**	**…**
payoff	**0**	**8**	**0**	**8**	**0**	**1**	**1**	**1**	**…**

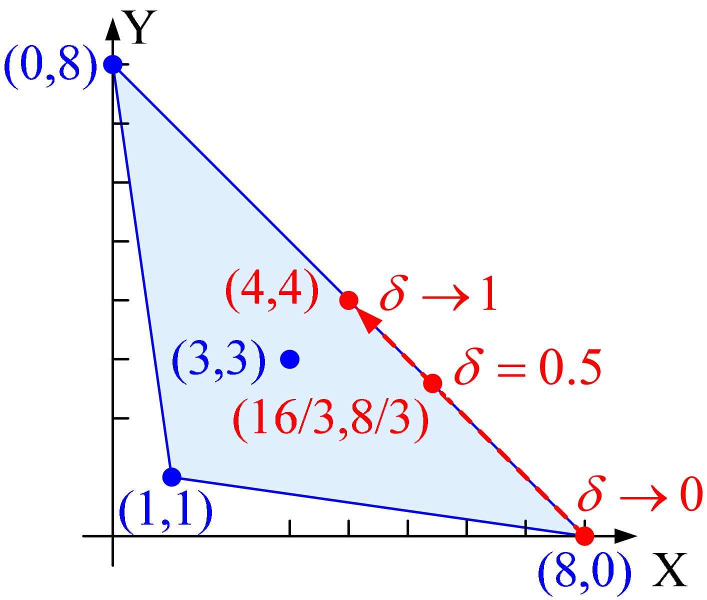
[結論] 當 $\delta \to 1$, 採用殘酷觸發策略便可逼使 players 合作. [解說] 假設 $\delta \to 1$. 令 (v_1^*, v_2^*) 表示 $CH(V)$ 裡頭的任意一點, (d_1, d_2) 表示 NE 的 payoff profile. 例: $(v_1^*, v_2^*) = (4,4)$, $(d_1, d_2) = (1,1)$. 左表顯示，在無限重複賽局中，若 players 合作，則 $(v_1, v_2) \to (v_1^*, v_2^*)$. 假設 player 1 在第 k 階段背叛 (此策略稱爲 s_1'); 即使背叛的 payoff b_1 大於 v_1^*, 但雙方從此翻臉，採用 NE 策略；此後 player 1 每階段的 payoff 僅有 d_1, 因此 $v_1^{[k]}(s_1') = (1-\delta)b_1 + \delta d_1 \to d_1$. 同理，假設 player 2 在第 k 階段背叛 (此策略稱爲 s_2')，且背叛的 payoff 爲 b_2. 我們可推得 $v_1^{[k]}(s_2') = (1-\delta)b_2 + \delta d_2 \to d_2$. 因此，只要 $\delta \to 1$ 且 $(v_1^*, v_2^*) > (d_1, d_2)$, players 便願意合作

Folk Theorem (無名氏定理)

[定理六] 給定 2 個 players 的 normal-form game G, 令 A 和 B 分別表示 player 1 和 player 2 的 strategy set。令 (d_1, d_2) 表示賽局 G 的 NE 所對應的 payoff profile。令 $V = \left\{ v^{[i,j]} = u(A_i, B_j) \middle| A_i \in A,\ B_j \in B \right\}$ 且 $(v_1^*, v_2^*) \in CH(V)$. 那麼在無限重複賽局 $G(\delta)$ 裡頭，只要 $\delta \to 1$ 且 $(v_1^*, v_2^*) > (d_1, d_2)$, 必然存在一個 SPNE 使得 players 願意合作，且 $(v_1, v_2) \to (v_1^*, v_2^*)$, 其中 v_i 表示 player i 平均每階段所能獲得的 payoff。

[解說] 因爲 $(v_1^*, v_2^*) \in CH(V)$, 所以必然存在實數 $\alpha_{i,j} \in [0,1]$ 使得 $(v_1^*, v_2^*) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{i,j} v^{[i,j]}$. 那麼對於 $G(\delta)$, 我們可製造出一個 SPNE，如下：players 雙方按 $\alpha_{i,j}$ 的比例採用 strategy profile (A_i, B_j). 若其中一方背叛, 那麼雙方從此展開 NE 策略。此一定理被稱爲無名氏定理是因爲：沒人知道這定理是誰發明的。Steven Tadelis 猜測此定理是由 James Friedman（現 耶魯大學教授）於 1971 年所提出。

Verify the Folk Theorem via an Example (1/4)



雙方永遠合作

階段	1	2	3	4	5	6	7	8	…
player 1	**D**	**C**	**D**	**C**	**D**	**C**	**D**	**C**	…
payoff	**8**	**0**	**8**	**0**	**8**	**0**	**8**	**0**	…
player 2	**c**	**d**	**c**	**d**	**c**	**d**	**c**	**d**	…
payoff	**0**	**8**	**0**	**8**	**0**	**8**	**0**	**8**	…

[驗證] 如果雙方合作，那麼 player 1 平均每一階段所能獲得的 payoff 爲

$$v_1 = (1-\delta)(8+8\delta^2+8\delta^4+\cdots) = (1-\delta)\times\frac{8}{1-\delta^2} = \frac{8}{1+\delta};$$

- player 2 平均每一階段所能獲得的 payoff 爲

$$v_2 = (1-\delta)(8\delta+8\delta^3+8\delta^5+\cdots) = (1-\delta)\times\frac{8\delta}{1-\delta^2} = \frac{8\delta}{1+\delta}.$$

- 我們有 $v_1 + v_2 = 8,\ \lim_{\delta\to 1} v_1 = \lim_{\delta\to 1}\frac{8}{1+\delta} = 4,$ 且 $\lim_{\delta\to 1} v_2 = \lim_{\delta\to 1}\frac{8\delta}{1+\delta} = 4.$

Verify the Folk Theorem via an Example (2/4)

Case (a)：player 1 在第 6 階段率先背叛（改用 D）

階段	1	2	3	4	5	6	7	8	9	…
player 1	D	C	D	C	D	D	D	D	D	…
payoff	8	0	8	0	8	1	1	1	1	…
player 2	c	d	c	d	c	d	d	d	d	…
payoff	0	8	0	8	0	1	1	1	1	…

Case (b)：player 2 在第 $k = 7$ 階段改變選擇

階段	1	2	3	4	5	6	7	8	9	…
player 1	D	C	D	C	D	D	D	D	D	…
payoff	8	0	8	0	8	1	8	1	1	…
player 2	c	d	c	d	c	d	c	d	d	…
payoff	0	8	0	8	0	1	0	1	1	…

- 我們要證明沒有單獨一個 player 能夠在某一個階段（假設爲第 $k \geq 2$ 階段）改變選擇而獲得更高的 payoff，不論第 k 階段之前的歷史爲何。
- 我們假設：在第 k 階段，player 1 遵守 SPNE 殘酷觸發策略 s_1, 而 player 2 不遵守 SPNE 策略（註：因爲 payoff 矩陣對稱 且 players 採用相同策略，所以不用檢查 player 2 遵守 s_2, 而 player 1 不遵守 s_1 的情況）
- Case 1. 假設第 k 階段之前的歷史爲其中一方（假設是 player 1）曾經背叛。上圖爲 $k = 7$ 的例子。對照左右二圖，player 2 在第 k 階段原本應該採用 d; 若改用 c, 當然 payoff 會下降。

Verify the Folk Theorem via an Example (3/4)

Case (a)：雙方永遠合作

階段	1	2	3	4	5	6	7	8	…
player 1	D	C	D	C	D	C	D	C	…
payoff	8	0	8	0	8	0	8	0	…
player 2	c	d	c	d	c	d	c	d	…
payoff	0	8	0	8	0	8	0	8	…

Case (b)：player 2 在第 $k = 6$ 階段率改變選擇

階段	1	2	3	4	5	6	7	8	…
player 1	D	C	D	C	D	C	D	D	…
payoff	8	0	8	0	8	3	1	1	…
player 2	c	d	c	d	c	c	d	d	…
payoff	0	8	0	8	0	3	1	1	…

- 觀察左圖，若 players 雙方策略不變，那麼每個 player 都有 high-payoff 階段 (此時 payoff 爲 8) 和 low-payoff 階段 (此時 payoff 爲 0).
- Case 2. 假設第 k 階段前，players 雙方未曾背叛，且第 k 階段爲 player 2 的 high-payoff 階段。右圖爲 player 爲 player 2 且 $k = 6$ 的例子。左圖顯示，若雙方合作，則 $v_2^{[k]}(s_2) = (1-\delta)(8 + 8\delta^2 + 8\delta^4 + \cdots) = 8/(1+\delta)$; 若 player 2 在第 k 階段改用 c (我們稱此策略爲 s_2'), 那麼 $v_2^{[k]}(s_2') = (1-\delta)(3 + \delta + \delta^2 + \cdots)$ $= 3 - 2\delta$. 若殘酷觸發策略爲 SPNE，則需滿足條件 $8/(1+\delta) \ge 3 - 2\delta$. 自己驗證：不論 $0 < \delta < 1$ 的值爲何，$8/(1+\delta) \ge 3 - 2\delta$ 都成立。

Verify the Folk Theorem via an Example (4/4)

Case (a)：雙方永遠合作

階段	1	2	3	4	5	6	7	8	…
player 1	D	C	D	C	D	C	D	C	…
payoff	8	0	8	0	8	0	8	0	…
player 2	c	d	c	d	c	d	c	d	…
payoff	0	8	0	8	0	8	0	8	…

Case (b)：player 2 在第 $k=5$ 階段率改變選擇

階段	1	2	3	4	5	6	7	8	…
player 1	D	C	D	C	D	D	D	D	…
payoff	8	0	8	0	1	1	1	1	…
player 2	c	d	c	d	d	d	d	d	…
payoff	0	8	0	8	1	1	1	1	…

◆ 觀察左圖，若 players 雙方策略不變，那麼每個 player 都有 high-payoff 階段 (此時 payoff 為 8) 和 low-payoff 階段 (此時 payoff 為 0).

◆ Case 3. 假設第 k 階段前，players 雙方未曾背叛，且第 k 階段為 player 2 的 low-payoff 階段。上圖為 player 為 player 2 且 $k=5$ 的例子。上圖顯示，若雙方合作，則 $v_2^{[k]}(s_2)=(1-\delta)(8\delta+8\delta^3+8\delta^5+\cdots)=8\delta/(1+\delta)$; 若 player 2 在第 k 階段改用 d (我們稱此策略為 s_2'), 那麼 $v_2^{[k]}(s_2')=(1-\delta)(1+\delta+\delta^2+\cdots)$ $=1$. 若殘酷觸發策略為 SPNE，則需滿足條件 $8\delta/(1+\delta)\geq 1$; 這意味 $\delta \geq 1/7$.

◆ 綜合 Cases 1, 2, 3 可得：當 $\delta \geq 1/7 \approx 0.143$ 時，殘酷觸發策略為 SPNE.

自我評量 (1/3)

		player 2	
		c (合作)	d (背叛)
player 1	C (合作)	R, R	S, T
	D (背叛)	T, S	P, P

1. 對於上表所示之 payoff matrix，我們知道：只要滿足 $T > R > P > S$, 那麼此一賽局便是囚犯困境賽局。對於無限囚犯困境賽局，讓我們考慮 tit-for-tat（可譯爲「以牙還牙」或「模仿貓」）策略，其規則如下：player i 在第一階段選擇合作。此後，player i 在第 $k \geq 2$ 階段採用對手在第 $k-1$ 階段所採取的行動。證明：只有剛好 $\delta = \dfrac{T-R}{R-S} = \dfrac{P-S}{T-P}$ 時，「雙方採用模仿貓策略」才會是無限囚犯困境賽局的 SPNE。

提示：如果「雙方採用模仿貓策略」爲 SPNE，就必須滿足「沒有單獨一個 player（假設是 player 1）能夠在某一個階段（假設爲第 $h \geq 2$ 階段）改變選擇而獲得更高的 payoff，不論第 h 階段之前的歷史爲何。考慮第 h 階段之前的歷史是以情況 (C,c), (C,d), (D,c), 或 (D,d) 作結尾。

自我評量 (2/3)

2. 假設某個產品目前只有 **A** 和 **B** 二家公司在生產。假設 **A** 和 **B** 生產該產品的成本皆爲 **30**。當 **A** 和 **B** 同時進行爲期一個月的量產賽局時，若 **A** 生產 $\boldsymbol{q_1}$ 個，**B** 生產 $\boldsymbol{q_2}$ 個，那麼根據專家估計，要將 $\boldsymbol{q = q_1 + q_2}$ 個產品銷售一空，產品售價應爲 $\boldsymbol{P(q) = \max\{130 - q, 0\}}$。假設 **A** 公司派人去 **B** 公司遊說，希望彼此都生產壟斷產量 $\boldsymbol{q^*}$ 的一半（稱之爲「合謀產量」）。論證：是否存在一個 **SPNE** 可以使得 **A** 和 **B** 長期（可能持續好幾年）合作，生產合謀產量？

[提示] 令 q^* 表示壟斷產量，那麼 $q^* = \arg\max_{0 \le q \le 130} \{(130 - q)q - 30q\}. \Rightarrow q^* = 50.$ 所以players 雙方合作時，每個player都生產合謀產量 $q_1^* = q_2^* = q^*/2 = 25.$ 此時每個player的payoff爲 1250. 我們現在要檢驗「殘酷觸發策略是否爲SPNE」. 我們考慮第 h 階段之前的歷史是以情況「(C,c)」或「不是(C,c)」作結尾。我們示範以 (C,c) 作結尾，另一情況較爲簡單，自己練習。假設player 2的策略不改，但player 1在第 h 階段改變選擇，不配合生產「合謀產量」. 那麼player 1此時的最佳選擇是$q_1' = \arg\max_{0 \le q_1 \le 130} \{(130 - q_1 - 25)q_1 - 30q_1\}. \Rightarrow q_1' = 37.5.$ 此時player 1的payoff爲1406.25. 然而，此後雙方展開報復，都使用NE策略，此時每個player生產$\hat{q}_1 = 100/3$個，payoff爲1111.11。所以如果殘酷觸發策略是SPNE，那麼條件 $1250 \ge (1-\delta)\left[1406.25 + 1111.11(\delta + \delta^2 + \cdots)\right]$ 必須成立 $\Rightarrow$ 只要 $\delta \ge 0.529$ 就成立

自我評量 (3/3)

3. 承上題，後來 **A** 公司發明新的生產技術，能使得每個產品的生產成本降為 **10**。此時 **A** 公司再度派人去遊說 **B** 公司，希望從下個月的月初開始，**A** 公司生產「壟斷產量 s^* 的 **75%**」， **B** 公司生產「壟斷產量 s^* 的 **25%**」**.** 注意：假設 **A** 和 **B** 總共生產 s^* 個產品可以讓二家公司合起來賺最多，那麼 s^* 便稱為壟斷產量。此時 $s^* \neq q^*$（因為 **A** 和 **B** 公司的生產成本已經不同）。論證：是否存在一個 **SPNE** 使得 **B** 公司願意長期（可能持續好幾年）配合 **A** 公司的主張？

[提示] 令 s 表示所有players的產量總和，那麼所有players的payoff總和為
$U(s) = (130 - s)s - 0.75s \times 10 - 0.25s \times 30.$ 那麼壟斷產量 $s^* = \arg\max_{0 \le s \le 130} \{U(s)\}.$

$\Rightarrow s^* = 57.5.$ .....如果players都遵守A公司的主張，那麼player 1應該生產$57.5 \times 0.75 = 43.125$ 個, 而player 2應該生產$57.5 \times 0.25 = 14.375$ 個。

[注意] 在此題中，由於player 1和player 2的成本不同，所以在應用「一階段偏離原則」時，須檢查player 1策略不變，player 2是否能單獨在第 h 階段改變選擇而提高 payoff；同時還須檢查player 2策略不變，player 1是否能單獨在第 h 階段改變選擇而提高 payoff