第八章 重复博弈

重复博弈：

基本博弈（完全信息静态博弈，完全信息动态博弈）重复进行构成的博弈过程。

重复博弈所关心的议题：

将来可信的威胁或承诺如何影响到当前的行动

目的：实现合作，提高收益

A、为何研究重复博弈

- **经济中的长期关系,例如**

  (1)两家企业在一个市场上的长期竞争,

  (2)市场营销中的回头客问题,

  (3) 买卖问题。

  (4)信任、信誉、声誉问题

- 人们的预见性。

  由于人的思维的限制，在短期行为中缺乏默契或合作的关系，但在长期中这样的机会就大得多。即是未来利益对当前行为的制约

B、重复博弈的分类

1、有限次重复博弈：给定一个基本博弈G（可以是静态博弈，也可以是动态博弈），重复进行T次G，记为G(T)。

而G则称为G(T)的“原博弈”。G(T)中的每次重复称为G(T)的一个“阶段”。

2、**无限次重复博弈：**一个基本博弈G一直重复博

弈下去的博弈，记为G(∞).

注：1、无法验证某个重复博弈会一直重复下去。

2、如果主观上认为博弈会不断进行下去，那么博弈就可无限次重复下去。

C、策略、子博弈和均衡路径

- 策略：博弈方在每个阶段针对每种情况如何行为的计划。
- 子博弈：从某个阶段（不包括第一阶段）开始，包括此后所有的重复博弈部分。

因此：动态博弈的分析方法都可用于重复博弈。

- 均衡路径：由每个阶段博弈方的行为组合串联而成。

D、重复博弈的得益

- 有限次重复博弈的总体得益方法之一：

计算重复博弈的“总得益”——博弈方各次重复得益的总和。

- 有限次重复博弈的总体得益方法之二：

计算各阶段的平均得益

E、重复博弈的得益（Cont…）

- 如果重复的时间很长，就应考虑资金的时间价值，此时考虑贴现系数

$$\delta = \frac{1}{1+r}$$

重复T期的重复博弈总得益为：

$$\pi = \pi_1 + \delta\pi_2 + \delta^2\pi_3 + \cdots + \delta^{T-1}\pi_T = \sum_{t=1}^{T}\delta^{t-1}\pi_t$$

重复无限期的重复博弈总得益为：

$$\pi = \pi_1 + \delta\pi_2 + \delta^2\pi_3 + \cdots = \sum_{t=1}^{\infty}\delta^{t-1}\pi_t$$

平均得益：如果一常数$\bar{\pi}$ 作为重复博弈（有限次重复博弈或无限次重复博弈）各个阶段的得益，能产生与得益序列$\pi_1,\pi_2,\cdots$相同的现在值，则称$\bar{\pi}$ 为$\pi_1,\pi_2,\cdots$的平均得益

有限次重复博弈不一定考虑贴现因素

无限次重复博弈必须考虑贴现问题$\bar{\pi}=(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\pi_t$

一、有限重复博弈

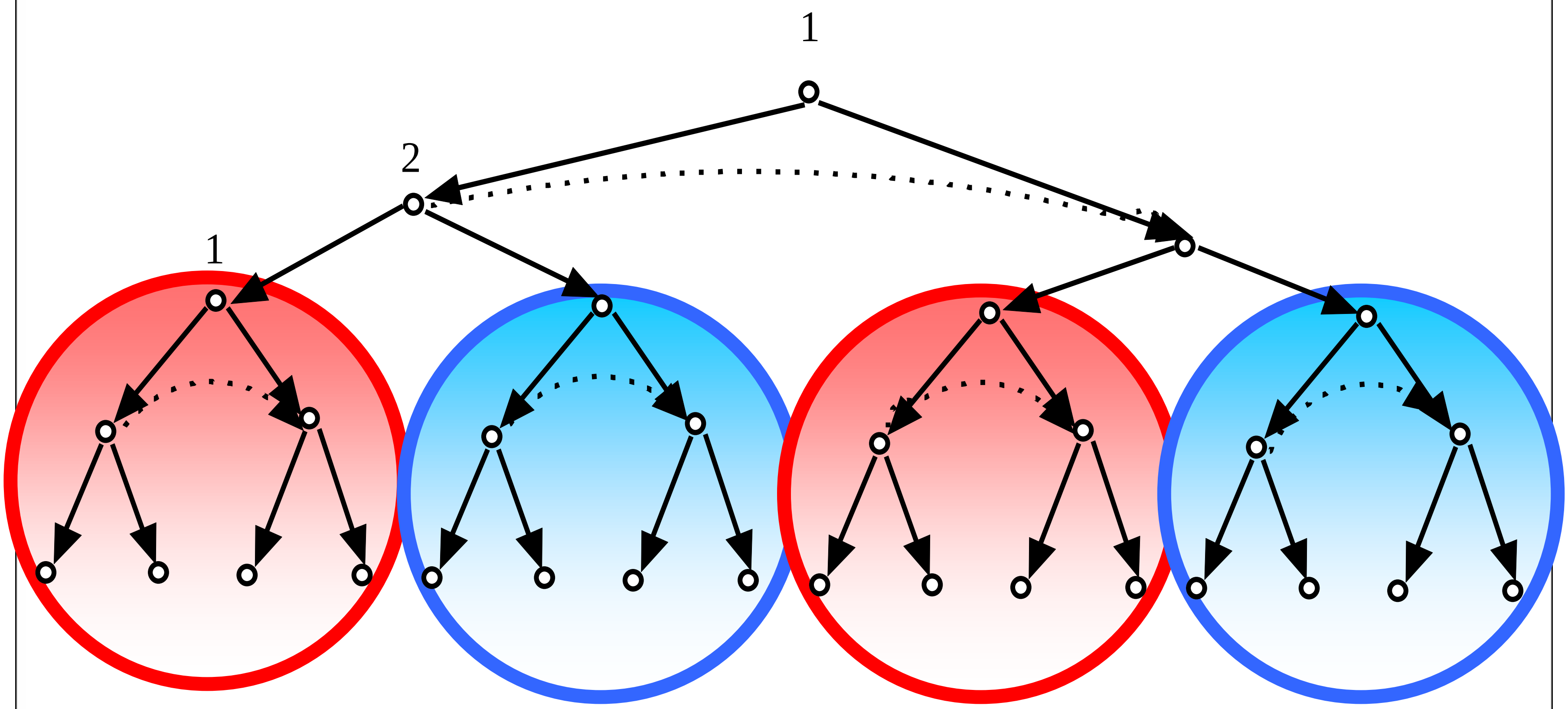
考察下列博弈

		2	
		L	R
1	U	1,1	5,0
	D	0,5	4,4

上述博弈存在唯一的Nash均衡。

将上述博弈重复两次，其中第二次博弈开始时，第一次博弈的结果已知。

两次重复博弈的博弈树



上述重复博弈只存在唯一的Nash均衡：在每次博弈中，参与人1都选择U，参与人2都选择L，即

$$((U, U, U, U, U), (L, L, L, L, L))$$

可以证明：该均衡为精炼Nash均衡。

		2	
		L	R
1	U	1,1	5,0
	D	0,5	4,4

		2	
		L	R
1	U	1,1	5,0
	D	0,5	4,4

		2	
		L	R
1	U	1+1,1+1	5+1,0+1
	D	0+1,5+1	4+1,4+1

		2	
		L	R
1	U	1,1	5,0
	D	0,5	4,4

前面的分析说明：在两次重复博弈中，合作仍无法到达。

同样可证明：在n阶段重复博弈(即博弈重复n次且每次博弈开始时，前面博弈的结果都已知)中，合作同样无法到达。

定理：

如果阶段博弈G有唯一的Nash均衡，则对任意有限的T，重复博弈$G(T)$有唯一的子博弈精炼解，即G的Nash均衡结果在每一个阶段重复进行。

考察下列博弈

1 \ 2	L_2	M_2	R_2
L_1	1,1	5,0	0,0
M_1	0,5	4,4	0,0
R_1	0,0	0,0	3,3

上述博弈存在两个Nash均衡：

$$(L_1, L_2)\text{和}(R_1, R_2)$$

将上述博弈重复两次。

1) 战略：

每个局中人都有

$$3^{\infty} = \infty$$

个战略；

2) 战略组合：

一共存在

$$\infty \times \infty = \infty$$

个战略组合；

3) 均衡：

可以根据以下原则构造均衡：

由第一阶段的结果，预测第二阶段的均衡。

根据上述原则，可构造如下策略：

S_1：第一阶段选择M_1；如第一阶段结果为(M_1，M_2)，则下一阶段选R_1；否则选择L_1。

S_2：第一阶段选择M_2；如第一阶段结果为(M_1，M_2)，则下一阶段选R_2；否则选择L_2。

触发策略(Trigger Strategy)： 两博弈方先试探合作，一旦发现对方不合作则也用不合作报复。

　　触发策略是重复博弈实现合作和高效的关键机制。

在上述策略下，博弈可表示为：

			2	
		L_2	M_2	R_2
	L_1	1+1,1+1	5+1,0+1	0+1,0+1
1	M_1	0+1,5+1	4+3,4+3	0+1,0+1
	R_1	0+1,0+1	0+1,0+1	3+1,3+1

这意味着：合作可以在第一阶段达到

定理：

如果$G = < T, (A_i), (u_i) >$是一个有多个Nash均衡的完全信息静态博弈，则$G(T)$可以存在子博弈精炼解，其中对每一$t<T$，t阶段的结果都不是G的Nash均衡。

上述结论说明：

对将来行动所作的可信威胁或承诺可以影响到当前的行动。

考察下列博弈。

		2			
1		X_2	Y_2	Z_2	P_2	Q_2
	X_1	1,1	5,0	0,0	0,0	0,0
	Y_1	0,5	4,4	0,0	0,0	0,0
	Z_1	0,0	0,0	3,3	0,0	0,0
	P_1	0,0	0,0	0,0	4,1/2	0,0
	Q_1	0,0	0,0	0,0	0,0	1/2,4

如果第一阶段出现(Y_1,Y_2),则第二阶段(Z_1,Z_2);

如果第一阶段出现(Y_1,w),其中$(w \neq Y_2)$,则第二阶段为(P_1,P_2);

如果第一阶段出现(w,Y_2),其中$(w \neq Y_1)$,则第二阶段(Q_1,Q_2);

如果第一阶段出现(w_1,w_2),其中$(w_1 \neq Y_1, w_2 \neq Y_2)$,则第二阶段为$(Z_1,Z_2)$。

		2				
		X_2	Y_2	Z_2	P_2	Q_2
1	X_1	1+3,1+3	5+1/2,0+4	0+3,0+3	0+3,0+3	0+3,0+3
	Y_1	0+4,5+1/2	4+3,4+3	0+4,0+1/2	0+4,0+1/2	0+4,0+1/2
	Z_1	0+3,0+3	0+1/2,0+4	3+3,3+3	0+3,0+3	0+3,0+3
	P_1	0+3,0+3	0+1/2,0+4	0+3,0+3	4+3,1/2+3	0+3,0+3
	Q_1	0+3,0+3	0+1/2,0+4	0+3,0+3	0+3,0+3	1/2+3,4+3

显然，上述策略构成博弈的Nash均衡，且为子博弈精炼Nash均衡。

有限次重复博弈的民间定理

		厂商2 A	厂商2 B
厂商1	A	3，3	1，4
厂商1	B	4，1	0，0

两市场博弈

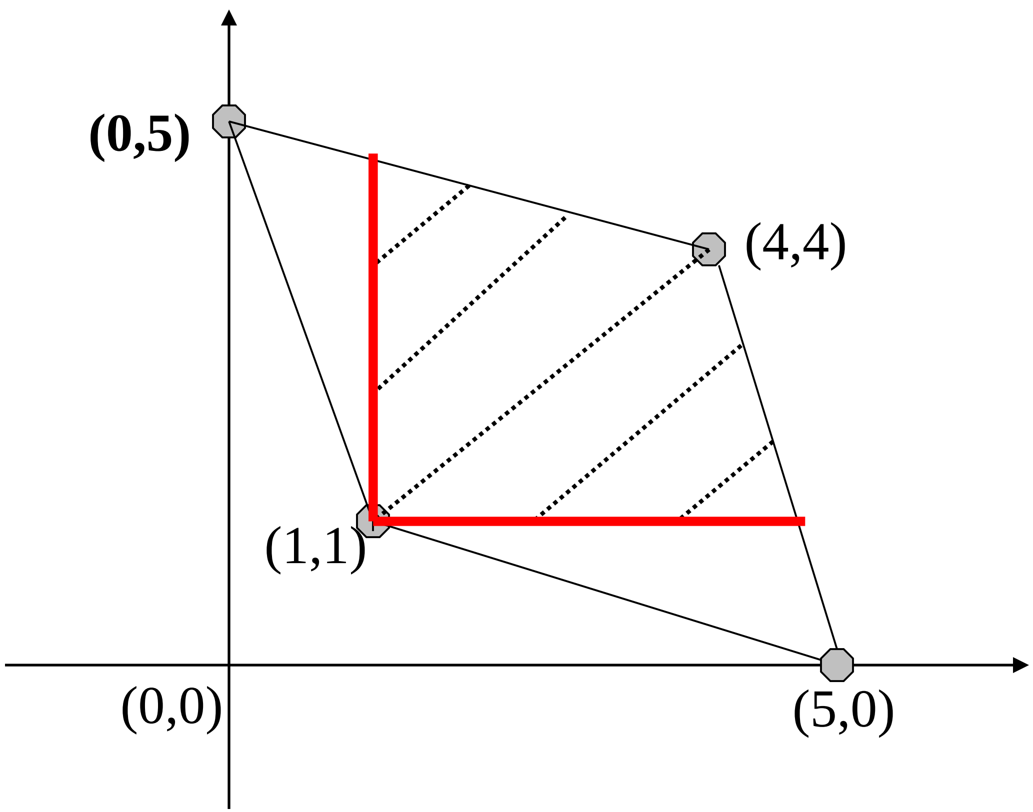
1、个体理性得益：博弈方 i 的最小最大值

$$\underline{v_i} = \min_{\alpha_{-i} \in A_{-i}} \max_{\alpha_i \in A_i} U_i(\alpha_i, \alpha_{-i})$$

2、可实现得益：博弈中所有纯策略组合得益的加权平均数组

民间定理：略

子博弈精炼Nash均衡的可行收益区间



二、 无限重复博弈

- 定义(无限重复博弈)给定一阶段博弈G，令$G(\infty,\delta)$表示相应的无限重复博弈，其中G将无限次低重复进行，且参与人的贴现率为δ。对每个t，之前t-1次阶段博弈的结果在t阶段开始进行前都可以被观测到，每个参与人在$G(\infty,\delta)$中的收益都是该参与人在无限次的阶段博弈中所得受益的现值。

例1：下列博弈重复无限次。

		2	
		L	R
1	U	1,1	5,0
	D	0,5	4,4

对于阶段博弈为上述博弈的有限重复博弈，合作不可能形成。

但对于无限重复博弈，在一定的贴现率下，合作有可能形成。

构造如下触发策略：

S_1：第i阶段选择D；如第i阶段结果为(D, R)，则下一阶段选D；否则以后一直选择U。

S_2：第i阶段选择R；如第i阶段结果为(D, R)，则下一阶段选R；否则以后一直选择L。

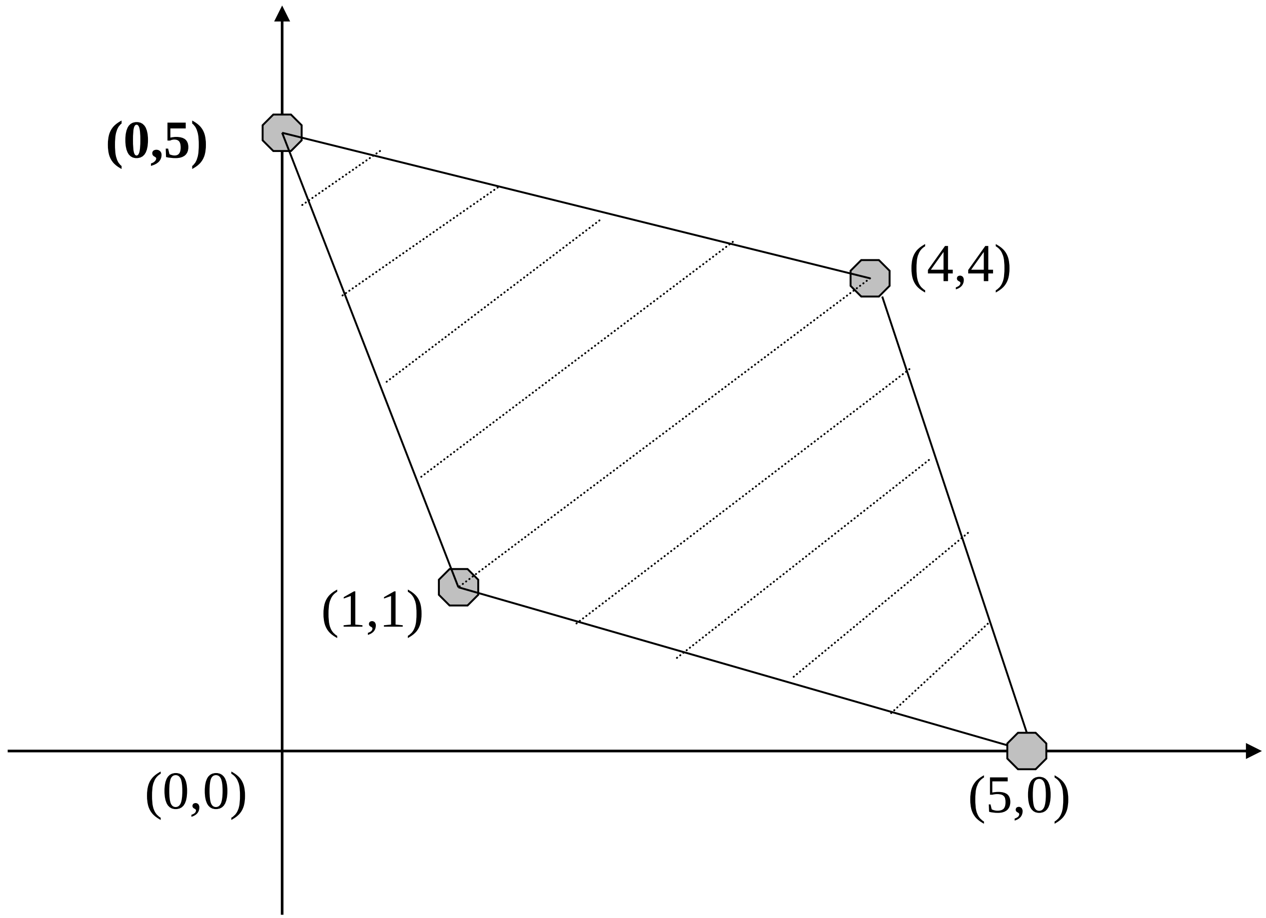
如何证明：在一定的贴现率下，上述触发策略构成Nash均衡？

可行收益

一组收益 $(x_1, x_2, \cdots, x_n)$ 为阶段博弈G的可行收益，如果它们是G的纯战略收益的凸组合(即纯战略收益的加权平均，权重非负且和为1)。

前述阶段博弈的可行收益集合如下图所示。

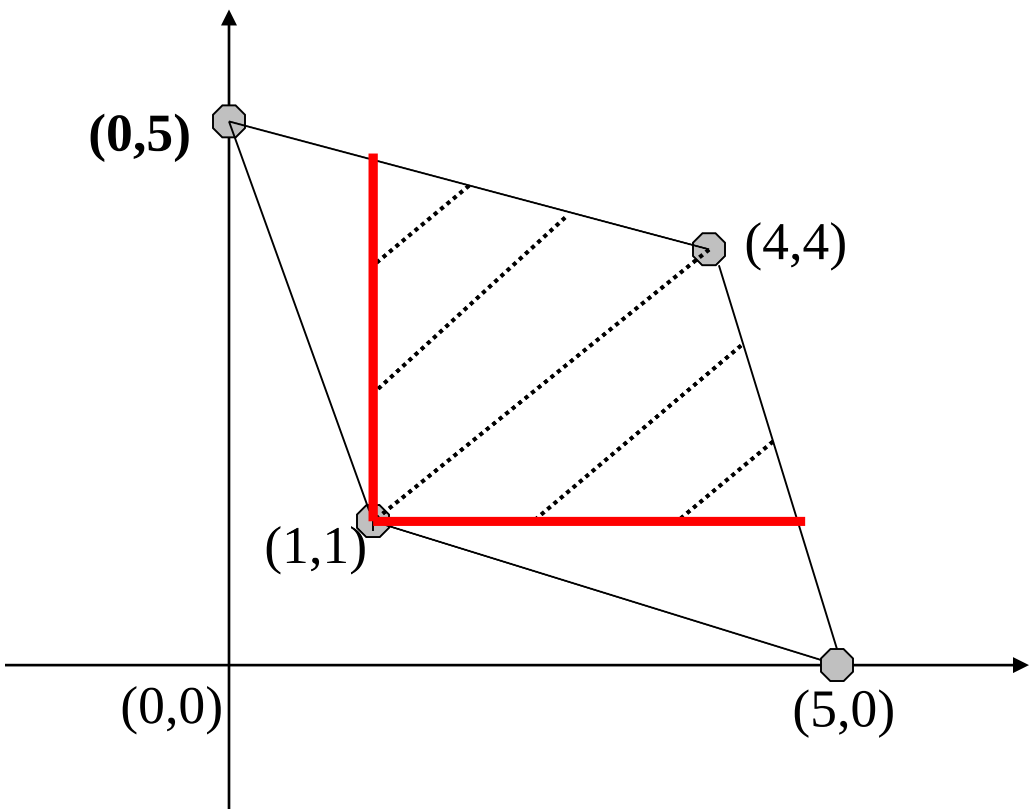
阴影部分为上述博弈的可行收益区间



无名氏定理（folk theorem）

如果博弈重复无限次，或者每次结束的概率足够小，如果δ充分接近1，任何个人理性可行支付向量都可以作为子博弈精炼纳什均衡结果出现

子博弈精炼Nash均衡的可行收益区间



例2：古诺博弈的无限重复博弈

考虑古诺博弈为阶段博弈的无限重复博弈，两企业的贴现率都为δ。

计算两个企业的下述触发战略成为无限重复博弈的Nash均衡时，贴现率δ的值。

触发战略

在第一阶段都生产垄断产量的一半 $q_m/2$ 。

第 t 阶段，如果前面 t-1个阶段两个企业的产量都为$q_m/2$，则生产$q_m/2$；否则，生产古诺产量 q_C 。

当双方都生产$q_m/2$时，每个企业的利润为$(a-c)^2/8$，用$\pi_m/2$来表示。当双方都生产q_C时，每个企业的利润为$(a-c)^2/9$我们用π_C表示。

- 若企业i将在本期生产$q_m/2$，则使企业j本期利润最大化的产量为下式的解

$$\max_{q_j}(a-q_m/2-q_j-c)\cdot q_j$$

- 其解为 $q_j=3(a-c)/8$，其利润水平为

$$\pi_d=9(a-c)^2/64$$

要使两企业采取上述触发战略成为Nash均衡，则必须

$$\frac{1}{1-\delta}\cdot\frac{1}{2}\pi_m \geq \pi_d + \frac{\delta}{1-\delta}\cdot\pi_C$$

代入π_m、π_d和π_C的值，即可得

$$\delta \geq \frac{9}{17}$$

当$\delta \geq \dfrac{9}{17}$时，给定$\delta$，企业将如何行动？

首先计算对任意一个给定贴现率，如果双方都采用触发战略，一旦出现背离就永远转到古诺产出，企业可以达到的利润最大化的产量。显然，该产量处于古诺产出与垄断产出之间。

考虑如下的触发战略：

第一阶段生产 q^* 。在第 t 阶段，如果在此之前的 t-1个阶段两企业的产量都是 q^* ，生产 q^* ；否则，生产古诺产出 q_C 。

若企业都生产q^*，则利润水平分别为

$\pi^* = (a - 2q^* - c)q^*$。

若企业i将在本期生产q^*，则使企业j本期利润最大化的产量为下式的解。

$$\max_{q_j}(a - q^* - q_j - c) \cdot q_j$$

其解为$q_j = (a - q^* - c)/2$，其利润水平为

$$\pi_d^* = (a - q^* - c)^2 / 4$$

要使两企业采取触发战略成为Nash均衡，则必须

$$\frac{1}{1-\delta}\cdot\frac{1}{2}\pi^{*}\geq\pi_{d}^{*}+\frac{\delta}{1-\delta}\cdot\pi_{C}$$

将π^{*}、π_{d}^{*}和π_{C}的值代入上式，即可解得触发战略成为子博弈精炼Nash均衡的q^{*}。

$$q^{*}=\frac{9-5\delta}{3(9-\delta)}(a-c)$$

q^*随δ单调递减，当$\delta=\dfrac{9}{17}$时，

$$q^* = q_m / 2$$

当$\delta = 0$时，

$$q^* = q_C / 2$$

罗伯特·爱克斯罗德实验

罗伯特·爱克斯罗德(政治科学家)，对合作的问题具有研究兴趣。

为了进行关于合作的研究，他组织了一场计算机竞赛。

这个竞赛的思路非常简单：

任何想参加这个计算机竞赛的人都扮演“囚徒困境”案例中一个囚犯的角色。他们把自己的策略编入计算机程序，然后他们的程序会被成双成对地融入不同的组合。分好组以后，参与者就开始玩“囚徒困境”的游戏。他们每个人都要在合作与背叛之间做出选择，并且游戏重复多次。

竞赛的第一个回合交上来的14个程序中包含了各种复杂的策略。但使爱克斯罗德和其他人深为吃惊的是，竞赛的桂冠属于其中最简单的策略：一报还一报(TIT FOR TAT)。这是多伦多大学心理学家阿纳托·拉帕波特提交上来的策略。

一报还一报的策略是这样的：它总是以合作开局，但从此以后就采取以其人之道还治其人之身的策略。也就是说，一报还一报的策略实行了胡萝卜加大棒的原则。

一报还一报的策略永远不先背叛对方，从这个意义上来说它是“善意的”。

一报还一报策略会在下一轮中对对手的前一次合作给予回报（哪怕以前这个对手曾经背叛过它），从这个意义上来说它是"宽容的"。

但一报还一报策略会采取背叛的行动来惩罚对手前一次的背叛，从这个意义上来说它又是“强硬的”。

而且，一报还一报策略的策略极为简单，对手程序一望便知其用意何在，从这个意义来说它又是"简单明了的"。

为了验证上述结果的合理性，爱克斯罗德又举行了第二轮竞赛，特别邀请了更多的人，看看能否从一报还一报策略那儿将桂冠夺过来。这次有62个程序参加了竞赛，结果是一报还一报又一次夺魁。

竞赛的结论无可争议地证明：好人，或更确切地说，具备以下特点的人，将总会是赢家。

1、善意的；

2、宽容的；

3、强硬的；

4、简单明了的。

作业

1. 教材
2. 考虑一个无限次重复博弈的囚徒困境博弈，其中的阶段博弈如下。记δ为贴现因子。试证明：（以牙还牙，以牙还牙）组合是该重复博弈的一个SPNE，当且仅当$y-x=1$和$\delta=1/x$（注意有$1<x<y$）.

		博弈方 2	
博弈方 1	策略	C	D
	C	x, x	0, y
	D	y, 0	1, 1

阶段博弈得益矩阵

- 3.证明“以牙还牙”策略是NE，但不是SPNE.

		2: N	2: C
1	N	a，a	0，c
	C	c，0	b，b

- 其中a<b<c