

## 第6章 重复博弈与合作行为

张维迎 教授  
北京大学光华管理学院

### 序贯博弈与重复博弈

- 序贯博弈(sequential game): 参与人在前一个决策点的选择决定随后的子博弈的结构, 因此, 从后一个决策点开始的子博弈不同于从前一个决策点开始的子博弈, 或者说, 同样结构的子博弈只出现一次;
- 重复博弈(repeated game): 同样结构的博弈重复多次, 其中的每次博弈被称为“阶段博弈”(stage game).

### 重复博弈的三个特征

- 阶段博弈之间没有物质上的联系, 也就是说, 前一阶段的博弈不改变后一阶段的结构;
- 所有参与人观察到博弈过去的历史;
- 参与人的总支付(报酬)是所有阶段博弈支付的贴现值之和;

### 重复博弈和信誉问题

- 如果博弈不是一次的, 而是重复进行的, 参与人过去行动的历史是可以观察到的, 参与人就可以将自己的选择依赖于其他人之前的行动, 因而有了更多的战略可以选择, 均衡结果可能与一次博弈大不相同。
- 重复博弈理论的最大贡献是对人们之间的合作行为提供了理性解释; 在囚徒困境中, 一次博弈的唯一均衡是不合作(即坦白)。但如果博弈无限重复, 合作就可能出现。

### 囚徒困境博弈

		B	
		合作	不合作
A	合作	3, 3	-1, 4
	不合作	4, -1	0, 0

### 重复博弈与战略空间的扩展

- 假定上属博弈重复多次或无限次; 那么, 每个参与人有多个可以选择的战略: 仅举几例:
- All-D: 不论过去什么发生, 总是选择不合作;
- All-C: 不论过去什么发生, 总是选择合作;
- 合作-不合作交替进行;
- tit-for-tat: 从合作开始, 之后每次选择对方前一阶段的行动;
- trigger strategies: 从合作开始, 一直到有一方不合作, 然后永远选择不合作。

## “囚徒困境”的一般表示

	合作	不合作
合作	T, T	S, R
不合作	R, S	P, P

满足:  $R > T > P > S$ ;  $(S+R) < T+T$

## 支付函数

- 双方都不合作:

$$V(\text{all-d, all-d}) = P + \delta P + \delta^2 P + \delta^3 P + \dots = P \frac{1}{1-\delta}$$

- 对  $\delta$  的解释:

- 贴现率;
- 博弈继续的概率;
- 二者的结合;
- 一般化: 未来收益的重要程度

## 无名氏定理(Folk Theorem)

- 在无限次重复博弈中, 如果参与人对未来足够重视 ( $\delta$  足够大), 那么, 任何程度的合作都可以通过一个特定的子博弈精炼纳什均衡得到。
- 这里“合作程度”定义为整个博弈中合作出现的频率。
- 50年代就人所共知, 但无人有发明权;

## Tit-for-tat

- 纳什均衡, 但不是精炼纳什均衡:

$$V(\text{TFT, TFT}) = T + \delta T + \delta^2 T + \delta^3 T + \dots = T \frac{1}{1-\delta}$$

$$V(\text{All-D, TFT}) = T + \delta P + \delta^2 P + \delta^3 P + \dots = T + P \frac{\delta}{1-\delta}$$

## Why Not A Perfect NE

- 假定A在t=5的时候, 没有合作。根据TFT战略, 在t=6, B应该选择惩罚(不合作)。B会这样吗?
- 如果B相信A采取的是TFT战略, 那么: 如果B对A实施惩罚, 预期的收入流为:  

$$R, S, R, S, R, \dots$$
- 反之, 如果B原谅A,

$$T, T, T, T, T, \dots$$

## Axelrod (1984)

- Tit-for-tat 是成功率最高的战略

## Trigger strategies

- 如果未来足够重要，精炼纳什均衡：

$$V(\text{合作, trigger}) = T + \delta T + \delta^2 T + \delta^3 T + \dots = T \frac{1}{1-\delta}$$

$$V(\text{不合作, trigger}) = R + \delta P + \delta^2 P + \delta^3 P + \dots = R + P \frac{\delta}{1-\delta}$$

## 合作的条件

- 如果下列条件满足，合作就是均衡结果：

$$T \frac{1}{1-\delta} \geq R + P \frac{\delta}{1-\delta}$$

$$\delta \geq \frac{R-T}{R-P}$$

## 解释

- R-T可以理解为不合作的诱惑；
- R-P是合作的剩余（利益）；
- 条件说明：
  - 给定未来的重要程度，不合作的一次性诱惑（R-T）相对于合作带来的利益（R-P）越小，合作的可能性越大；
  - 给定不合作的诱惑和合作带来的利益，未来越重要，合作的可能性越大；

## 行为的信息传递

- 假如欺骗两次才被发现：

$$V(\text{不合作, trigger})$$

$$= R + \delta R + \delta^2 P + \delta^3 P + \delta^4 P + \dots = R(1+\delta) + P \frac{\delta^2}{1-\delta}$$

$$\delta \geq \sqrt{\frac{R-T}{R-P}} > \frac{R-T}{R-P}$$

## 含义

- 欺骗行为越难以被发现，欺骗发生的可能性越大；或者说，合作越困难；
- 在前面的例子，R=4, T=3, P=0. 如果欺骗一次就被发现，只要  $\delta \geq \frac{4-3}{4-0} = 0.25$  合作就会出现；而如果欺骗两次才被发现，只有当  $\delta \geq 0.5$  时，合作才可能出现；
- 一般地，欺骗行为越不容易被发现，合作越困难。

## 惩罚与合作

- Abreu(1986)：最大合作战略是使用最严厉的可信惩罚（the strongest credible punishment）；
- 维护合作并不需要无限期的惩罚；只要惩罚期足够长就可以了；
- 萝卜加大棒（stick and carrot）：从合作开始，一直合作直到：如果有任何一方在t期不合作，在t+1期，前期合作者选择“不合作”来实施惩罚，前期不合作者选择合作；如果该合作的没有合作或者该惩罚的没有惩罚，在t+2期继续按照上述t+1期的战略规定博弈；否则，合作恢复。

## 解释

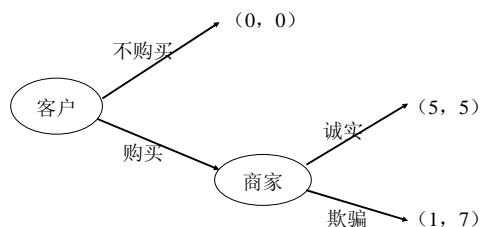
- 在合作子博弈，合作的条件是：

$$T + \delta T \geq R + \delta S \Rightarrow \delta \geq \frac{R - T}{T - S}$$

- 在非惩罚子博弈，合作的条件是：

$$S + \delta T \geq P + \delta S \Rightarrow \delta \geq \frac{P - S}{T - S}$$

## 不可信的惩罚：垄断厂家



## 不确定性下的最优处罚

- 在确定的情况下，惩罚越严厉，越有助于合作。因为均衡情况下，欺骗从来不会发生，所以惩罚实际上是没有成本的。
- 但在不确定的情况下，即使每个人都选择合作，“坏结果”也会出现。如果坏结果总是触发惩罚，就会冤枉好人，过重的惩罚反到导致不合作；但如果总是原谅，合作也不会发生。

## 多重交易关系与合作行为

交易关系I

	合作	不合作
合作	3, 3	-1, 4
不合作	4, -1	0, 0

交易关系II

	合作	不合作
合作	5, 5	0, 9
不合作	9, 0	4, 4

## 合作条件

- 在交易关系I中，合作要求  $\delta \geq 0.25$
- 在交易关系II中，合作要求  $\delta \geq 0.8$
- 如果两种交易在同样的两个人之间进行，只要  $\delta \geq 0.25$ ，两种交易中都会合作；
- 设想实际的  $\delta = 0.6$ ，那么如果两种交易发生在不同的人之间，合作只出现在市场I，不会出现于市场II。但如果两人之间同时存在两种交易，合作存在于两个市场。

## 社会关系与合作行为

- 更一般地讲，市场交易常常镶嵌在复杂的社会关系中。这种关系可以提高交易的合作程度。这也是人们愿意发展社会关系的原因。

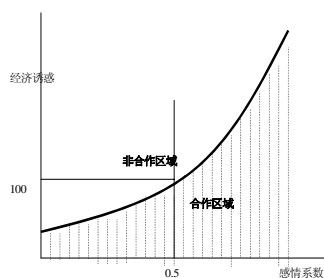
## 举例

- 在交易关系II中，假定除了这一交易外，当事人双方之间还存在另外的社会关系，这一社会关系对每一方的现值是V。如果交易中出现欺骗，这一价值就不存在。合作条件为：

$$\delta \geq \frac{4 - V}{5 - V}$$

## 应用

- 家庭血缘关系；
- 朋友关系；
- 同学关系；
- 老乡关系（保姆市场）；



## 家族企业的困境：“杀熟”

- 如果惩罚对惩罚者本身的损害太大，惩罚就是不可信的。这与投鼠忌器是一个道理，你讨厌老鼠，但是你没有办法，因为你心疼那个盘子，那个器皿。投鼠忌器在我们企业内部很多环节都会发生，导致惩罚不可信，所以对方就不会太注重信誉。家族成员有时候比非家族成员更不可信任，更不守规矩，就是这个道理。
- 朋友专门骗朋友，是“杀熟”。问题也与惩罚的不可信有关。

## 第三方实施的惩罚

- 前面假定了固定的一对参与人进行重复博弈，对不合作的惩罚是由“受害人”本人实施的，称为“second-party enforcement”，或者“personal enforcement”；
- 但更经常的情况是参与人不固定的情况。此时，惩罚要由第三方实施(third-party enforcement)；
- 问题是第三方惩罚欺骗别人的同时，自己可能失去合作带来的好处。“second-order prisoner's dilemma”。

## 长期参与人与不固定的短期参与人

- 最简单的例子是厂家与消费者之间的博弈：每个消费者一般只购买一次，而厂家重复出售产品给众多的消费者。
- 此时，只要消费者足够多，并且每个消费者能观察到前一个消费者购买的产品质量，合作仍然可以出现。

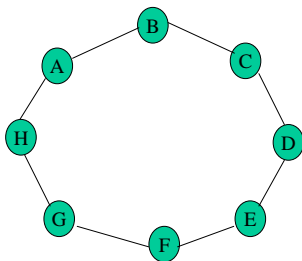
## 产品质量博弈

	高质量	低质量
购买	1, 1	-1, 2
不购买	0, 0	0, 0

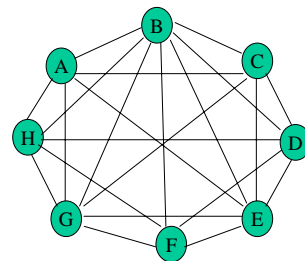
## Klein-Leffler Model

- 如果  $\delta \geq 0.5$  下列战略组合构成一个精炼纳什均衡：厂家开始生产高质量；继续生产高质量，除非曾经生产过低质量；如果上一期生产了低质量，之后永远生产低质量；消费者：第一个消费者购买；只要低质量事件没有发生过，之后的消费者继续购买；但一旦发现低质量，之后的消费者不再购买；
- 结果：（购买，高质量）
- 解释连锁店

## 社团内的博弈



## 社团内的博弈



## Third-party enforcement problem

- 如果A欺骗了B，其他成员是否应该惩罚A？
- 如果C惩罚A，C就失去了与A合作的机会，C为什么要替B惩罚A？

## 社会规范：集体抵制

- Boycott:** 每个人都应该诚实；都有责任惩罚骗过的人；不参与惩罚的人应该受到惩罚；如：假定A在t期欺骗了B，C在t+1期就不应该与A合作，否则，D在t+2期就不应该与C合作；如果C在t+1期与B合作，而D在t+2期又与C合作，F在t+3期就不应该与D合作，如此等等；
- （美国对不参与伊拉克战争的国家的态度）；

## 敌友规则

- 开始把所有的人当朋友；t期的朋友关系继续保持到t+1期，当只当他在t期不曾骗过任何人并不曾与你的敌人合作；
- 朋友的朋友是朋友；
- 朋友的敌人是敌人；
- 敌人的朋友是敌人；
- 敌人的敌人未必是朋友。

## PNE

- Milgrom, North and Weingast (1990), Econ and Poli; Kandori (1992), Review of Econ Stud; Mahoney and Sanchirico (2003), Cal L R;
- 如果每个人的行为是公共信息，合作是一个精炼纳什均衡。
- 信息问题为理解法律的出现提供了一个原因。法律机关的一个功能是通过集中化信息使得信誉机制可以更好得发挥作用。

## 有限次博弈

- 在前面的完全信息博弈中，只要博弈的次数是有限的，合作就不会出现。但如果参与人的类型是不完全信息，只要博弈进行的足够长，合作仍然可能出现。

## 信誉的条件

- 重复博弈；
- 足够耐心；
- 相对确定环境；
- 欺骗可以被观察到；
- 受骗人有积极性惩罚。

## 眼前利益与长远利益



## 传统社会的个人信誉

- 在传统社会，人们常年生活在封闭的村庄，村民之间彼此非常熟悉，欺骗行为很容易识别，人们之间的口头交流足以使任何欺骗行为广为而知，每个人的历史都存储在别人的脑海里，对欺骗行为的惩罚即使不能施加于欺骗者本人，也可以通过家庭成员而实现，前面讲的四个条件基本是可以得到满足。因此，即使没有法律，村民之间也可以建立起高度的信任，欺骗行为很少发生。

## 商业社会的问题

- 现在社会被称为“匿名社会”（anonymous society），与乡村社会不同，居民的流动性大，交易双方通常并不认识，相互之间也缺少如乡村社会中存在的其他制约关系，使得受害人的惩罚措施受到很大限制；开放的社会也使得人们较不在乎闲言碎语的议论。凡此种种，使得传统的以个人为基础的信誉机制失灵，这也是都市社会犯罪率高的一个重要原因。但是，西方市场经济之所以能发达达到今天的程度，除了较完善的司法制度之外，是因为他们在更高的形态上复制出了前面讲的四个条件。