第3章 完全信息动态博弈

聂辉华 教授
中国人民大学经济学院
www.niehuihua.com
niehuihua(at)vip.163.com

3.1 扩展式、博弈树与信息分割

3.1.1 扩展式表述

成语字典中有“先发制人”。但是，著名企业家段永平说过一句话：敢为天下后。“长江后浪推前浪，前浪死在沙滩上。”如何看待“先发制人”与“敢为人后”的矛盾？

在完全信息静态博弈中，双方同时选择行动。然而，一旦一方先行动，结果可能完全不同。例如，稍微修改一下“性别战”（图 3-1），现在假定女生下午 3 点半下课，而男生下午 4 点半下课。女生可以先去拳击场或者剧院，然后告诉男生自己所在的地方。均衡将会是怎样呢？

标准式难以刻画行动的顺序。为了更好地刻画动态博弈，我们必须引入另一种博弈的表达方式——扩展式（extensive form）。一个完整的扩展式博弈包括 7 个要素：（1）参与人；（2）收益函数；（3）行动顺序；（4）可以行动时的行动集合；（5）参与人行动时的知识，例如男生**是否知道**女生的选择；（6）外生事件的概率分布，借助非策略性参与人“自然”来刻画；（7）博弈结构（1）-（6）为参与人的共同知识。

		女生	
		拳击	芭蕾
男生	拳击	**2，1**	0，0
	芭蕾	0，0	**1，2**

图 3－1 性别战

扩展式通常用博弈树（game tree）来描述。博弈树的主要要素是（1）结（node），包括起始结、后续结、终点结；（2）枝（branch），指特定结上参与人的行动；（3）路径。不妨将上述博弈用博弈树表示如下。

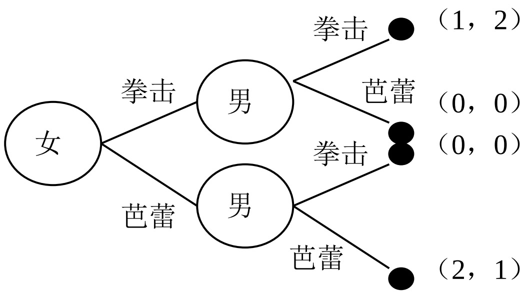


图 3－2 扩展式性别战 I

[注 1] 博弈树不能出现循环的圈，例如图 3－3a。

[注 2] 博弈树的所有后续结只有一个前续结，不能出现如图 3－3b 的情况。

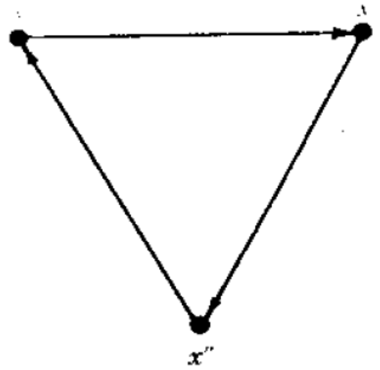


图 3－3a

图 3－3b

3.1.2 时间线表述

另一种简单的表述方式是时间线（time line/timing），它不包括赢利数字，表示的是逻辑时间而非物理时间。时间线广泛应用于信息经济学、契约理论等。例如，图 3－4 表述了序贯协调博弈。

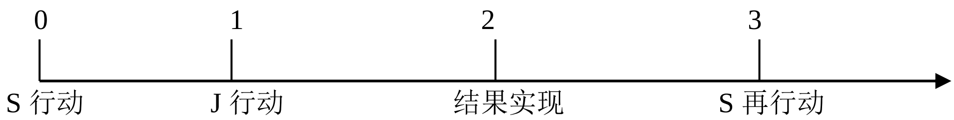


图 3－4 时间线

3.1.3 信息集与信息分割

为了了解谁在什么时候知道什么，我们需要进一步定义信息集和信息分割。

信息集（info set）：参与人某个时间在博弈树中认为有可能是实际结的不同结的集合。

信息分割（info partition）：表示参与人在某个时间能够区分的不同位置，它用信息集的个数来表示。

将图 3－2 修改为图 3－5。女生知道自己的全部行动，但是男生只在女生选择了“上”的时候知道对方的行动，此外就不知道女生到底选择了中还是下。因此，当博弈进行到“中”时，女生知道自己的选择，因此她的信息集是{中}；男生无法区分中还是下，因此其信息集是{中，下}。

由于信息集表示的是参与人对对方行动了解的程度，因此信息集越大，表示“越糊涂”。我们可以将表示信息集的虚线形容为“云”。

[注 1] 一个结不能同时属于一个参与人的两个不同信息集。否则，表示全都属于一个信息集。

[注 2] 信息集仅表示信息的客观描述，不表示任何理性的推理，即它不表示经过逻辑推敲后剔除的环节。

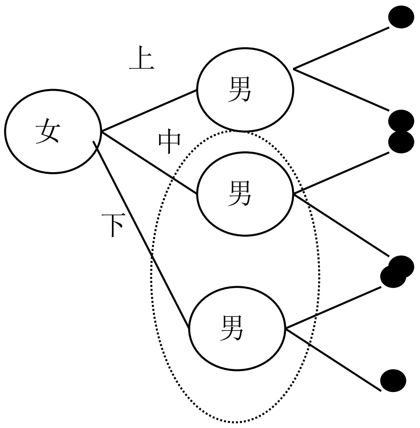


图 3－5 扩展式性别战 II

信息分割具有某种与信息集相反的意思，包含的信息集越多，说明信息越“好”（细）（finer）；而信息集包含的元素越多，表示“越糊涂”。如图 3－5，在第二阶段，女生的信息分割是（{上}，{中}，{下}），因为她清楚自己的每个行动。而男生的信息分割为（{上}，{中，下}），因为他不清楚女生没有选择上时，到底选择了中还是下。总之，信息集包含的元素越多，或信息分割包含的元素越少，就说明信息的质量越差。

[注 1] 信息分割（{上}，{中，下}）和（{上，中}，{下}）不能直接比较。

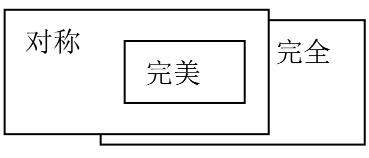
[注 2] 信息多未必是好事（张良脱衣自救），信息少未必是坏事。

3.1.4 信息分类

信息种类	含义
完美（perfect）	每个信息集都是单结的，只存在于完全信息动态博弈中
对称（symmetric）	所有参与人在任何时点都拥有相同的信息集，不存在私人信息
完全（complete）	自然不首先行动，或者后来的行动被所有参与人观察到

[注 1] 完美信息表达了对信息的最强要求。完美信息一定是对称的、完全的，任何不对称或不完全的信息一定不是完美的。例如，斯塔克伯格模型。

[注 2] 对称信息和完全信息不是相互隶属的。有时信息是对称的，但是不完全的。比如一对情侣对于“明天天气怎么样”，既可以表示不完全信息（自然首先行动），也可以表示不对称信息（一方知道天气预报）。如果改为“今天的游玩计划”，它可以是信息完全的（没有自然扰动），并且是对称的（双方均知晓），此时就是完美信息。



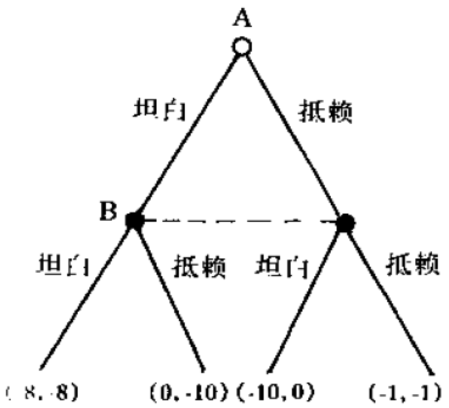
3.1.5 策略式表述与扩展式表述的转换

练习

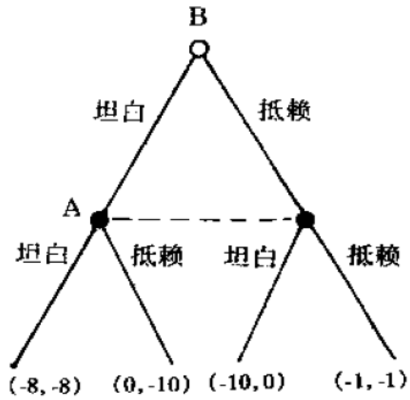
1、用博弈树表示囚徒困境。

1 \ 2	坦白	抵赖
坦白	**−8，−8**	0，−10
抵赖	−10，0	−1，−1

囚徒困境



囚徒困境扩展式 1



囚徒困境扩展式 2

2、用标准式表达序贯的性别战。以 O 表示歌剧，F 表示拳击。

女生 \ 男生	OO	OF	FO	FF
O	2，1	2，1	0，0	0，0
F	0，0	1，2	0，0	1，2

图 3-7 标准式序贯性别战

其中，OO 表示，不管女生选择什么，男生都只选择 O；FO 表示“女生选择 O，男生选择 F；女生选择 F，男生选择 O”。

3.2 子博弈完美纳什均衡：逆向归纳法

3.2.1 纳什均衡的问题

纳什均衡的本质是指给定对方的选择，参与人做出最佳反应，这是一种静态的格局。但是在动态博弈下，先行动的参与人会考虑到后行动的参与人的反应，或者说先行者能够主动影响后来者的选择。当行动顺序很重要时，纳什均衡很难令人满意，粗略地说，纳什均衡的解太多了。由于这种静态和动态的冲突，我们必须对均衡进一步精炼（refine）。

我们将图 2-14 的协调博弈改成一个完全信息动态博弈。如下：

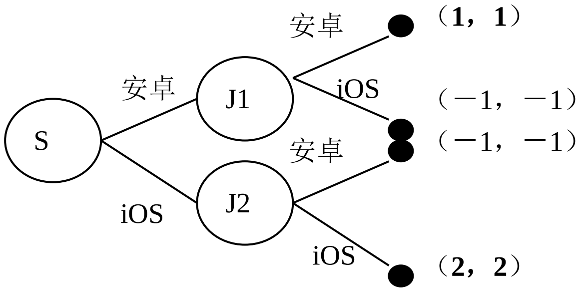


图 3－8 扩展式序贯协调博弈 I

根据我们原先的方法，至少可以得到三个纳什均衡：E1{iOS，（iOS，iOS）}，E2{ iOS，（安卓，iOS）}，E3{安卓，（安卓，安卓）}。[①] 但是我们可以推测，只有 E2 是唯一合理的均衡。就 E1 而言，一旦史密斯选择了安卓，琼斯也只有跟随。就 E3 而言，一旦史密斯选择了 iOS，琼斯也只有选择 iOS。总之，琼斯固定选择某种行动的策略并不可信（noncredible）。技术地说，合理的均衡不仅必须在某个“子博弈”中存在，而且应该在所有可能的路径上都存在，这样地均衡称之为“子博弈完美纳什均衡”。我们称均衡 E1 和 E3 为纳什均衡而非“完美/精炼”纳什均衡。

3.2.2 子博弈完美均衡的定义

A **subgame** *is a game consisting of a node which is a singleton in every player's information partition, that node's successors, and the payoffs at the associated end nodes.*[1]

A strategy profile is a **subgame perfect Nash equilibrium** *if (a) it is a Nash equilibrium for the entire game; and (b) its relevant action rules are a Nash equilibrium for every subgame.*

关于子博弈的定义：

图 3－8 有三个子博弈：（1）整个博弈；（2）从 J1 开始的博弈；（3）从 J2 开始的博弈。简单地说，就是有明确的开始结、**所有的**终点结和相应的收益函数的博弈路径。

[注 1] 任何一个原博弈都可以看作是自身的一个子博弈，即子博弈⊆原博弈。

[注 2] 任何一个博弈必须从一个单结的信息集开始。这意味着，一个完美信息博弈的每个决策结都带动一个子博弈。如图 3－9，x 和 x' 都不能作为子博弈的初始结。

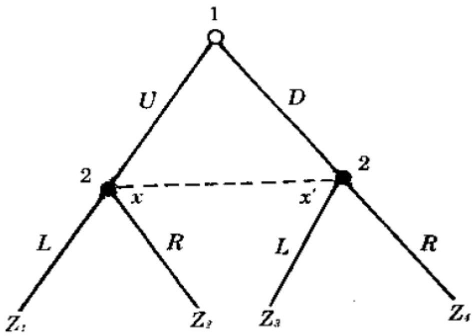


图 3－9

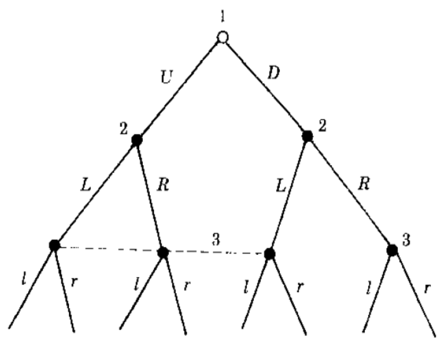


图 3－10

[注 3] 子博弈的信息集和收益向量都直接源于原博弈，即子博弈不能切割原博弈的信息集。如图 3－10，如果结 2 开始一个子博弈，那么 3 的信息集就被切割了，因此 2 不能开始

① 我们从扩展式上看不出三个均衡，可见扩展式也有不如策略式的地方。

一个子博弈。因为在存在切割的情况下，我们无法运用逆向归纳法。

关于子博弈完美纳什均衡（简称 SPE）的定义：

[注 1] 当且仅当一个均衡在所有的子博弈中都是纳什均衡时，该均衡才是子博弈完美纳什均衡。进一步地，SPE 不仅在均衡路径上是最佳的，而且在**非均衡**路径上也是最佳的。

[注 2] 如果原博弈只有一个子博弈，那么纳什均衡和 SPE 是等价的。

[注 3] SPE 概念一方面体现了“序贯理性”（sequential rationality）的思想，即参与人在博弈的每一个点上都重新优化自己的选择（忽视沉没成本），并且把自己在将来会重新优化其选择（理性预期）也考虑进去；另一方面，它可以经受“颤抖的手”（trembling hand）的考验。例如在图 3－8 中，若史密斯不小心选择了“iOS”，那么均衡 E3 显然不成立。

3.2.3 逆向归纳法

在图 3－8 这个两期的动态博弈中，S 先行动，J 后行动，因此 S 在决策时必须考虑到自己的行动对 J 的影响，而 J 的行为又受 S 的影响。因此，我们必须假定博弈进行到最后一个阶段，此时 J 必须做出选择，然后 S 预见到这种选择结果之后才能决定如何在前一个阶段做出选择，这种方法称为“逆向归纳法”（backward induction）。逆向归纳法充分体现了 SPE 的思想。假定有 1、2 两个参与人先后行动，行动空间分别为 A_1 和 A_2 。在第二阶段，给定 1 的行动，2 面临的问题是 $\underset{a_2\in A_2}{Max}\, u_2(a_1,a_2)$ ，解为 $a_2^*=R_2(a_1)$ ，它表示了 2 的最佳策略。在第一阶段，1 预测到 2 的最佳反应，它面临的问题是 $\underset{a_1\in A_1}{Max}\, u_1(a_1,R_2(a_1))$ ，解为 a_1^* 。该博弈的 SPE 为 $(a_1^*,R_2(a_1))$ ，均衡结果为 $(a_1^*,R_2(a_1^*))$ 。[①] 既然 1 和 2 在每一个阶段都选择了自己的最佳决策，那么 SPE 的条件当然满足了。

3.3 应用 I：要挟诉讼模型

我们先介绍一个离散模型——要挟诉讼。这是一个法律问题，也是一个经济学问题。两者结合就是法律经济学（Law and Economics）问题，该领域肇始于 Coase 和 Posner 的开创性贡献。法律经济学有两种思路：用经济学方法解释法律制度；分析法律对经济问题的影响。Shleifer 及其合作者（LLSV）对法律和金融的关系做出了开创性的贡献。

【思考】法学和经济学的相似之处？

[①] 注意“均衡”与“均衡结果”的细微差别。

Players
A plaintiff and a defendant.

The Order of Play

1 The plaintiff decides whether to bring suit against the defendant at cost c.

2 The plaintiff makes a take-it-or-leave-it settlement offer of $s > 0$.

3 The defendant accepts or rejects the settlement offer.

4 If the defendant rejects the offer, the plaintiff decides whether to give up or go to trial at a cost p to himself and d to the defendant.

5 If the case goes to trial, the plaintiff wins amount x with probability γ and otherwise wins nothing.

Payoffs
Figure 4 shows the payoffs. Let $\gamma x < p$, so the plaintiff's expected winnings are less than his marginal cost of going to trial.

注意，要挟诉讼的一个关键假设是，原告不可能真正胜诉，他唯一的目的就是希望通过私了来从被告那里得到一笔赔偿。成本 c 可以理解为搜集证据的费用，而 p 是起诉的成本。

我们画出该博弈的博弈树，如图 3－11。

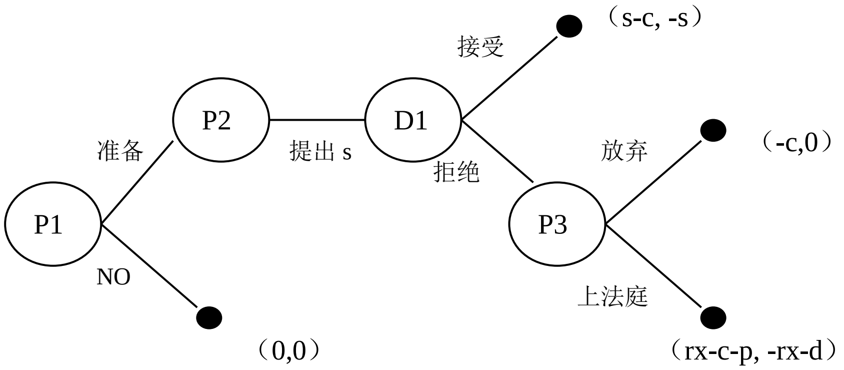


图 3－11 要挟诉讼 I

为了找到 SPE，我们使用逆向归纳法。先从最后一个由 P3 开始的子博弈入手。在 P3，由于$\gamma x < p$，所以$\gamma x - c - p < -c$，即原告会选择放弃（我们在放弃这个枝条上做一个标记）。继续推理，在 D1 点，被告会拒绝赔偿（继续做标记）；在 P2 点，原告的诉讼没有意义了；在 P1 点，原告会选择放弃诉讼（仍然做标记）。最后的均衡是（不诉讼，不赔偿）。我们看到，原告的诉讼威胁是不可置信的（noncredible）。

为了使威胁可置信，关键是在最后一步让对方相信“上法庭”对原告是有利的。例如，原告可以先支付一笔律师费 p，使得$\gamma x - c - p > -c - p$，从而放弃损失更大。如图 3－12。

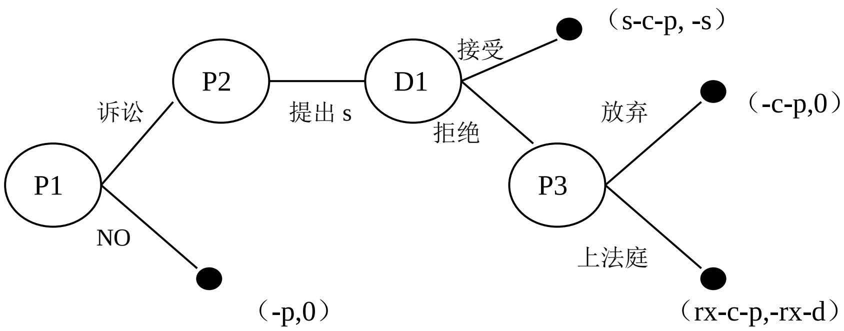


图 3－12 要挟诉讼 II

当$s \ge \gamma x$时，原告就愿意私了。对被告而言，只要$s \le \gamma x + d$，被告一定会同意私了。因此，合理的私了区域为 s∈[c，$\gamma x + d$]。此时，均衡为（诉讼，赔偿）。

既然原告可以使用沉没成本来使自己的行为可置信，被告也可以。比如，被告事前预支律师费 d，这样赔偿金就只能是γx，这对于原告而言是否诉讼已经没有差别，对被告而言是否赔偿也无差异。不管出现何种结果，从社会最优的角度而言律师费 d、p 都是浪费。这可以解释为什么很多大企业都有常设法律顾问。类似地，国防支出也可说是一种社会浪费。

以上分析是在“法治”（rule of law）环境下进行的。如果法官是腐败的，结果会如何呢？一定要注意博弈的制度背景！

当然，逆向归纳法不是完美的。当参与人很多，或者每个参与人有很多次选择机会时，微小的“颤抖”都可能导致 SPE 被破坏。如图 3－13 可以看作一个“接头博弈”，其中的数字表示参与人，A 表示继续，D 表示停止。该博弈的 SPE 是 N 个参与人都选 A。但是假定每个人选择 A 的概率$p<1$，那么其余 N－1 个参与人中任一人“犯错”的概率为$1-p^{n-1}$。当 N 很大时，即使 p 很大，这个概率也不小。因此，要让所有人相信其他人不会犯错误是非常困难的。对于这种“非理性”的行为，我们在“不完全信息动态博弈”中再讲。

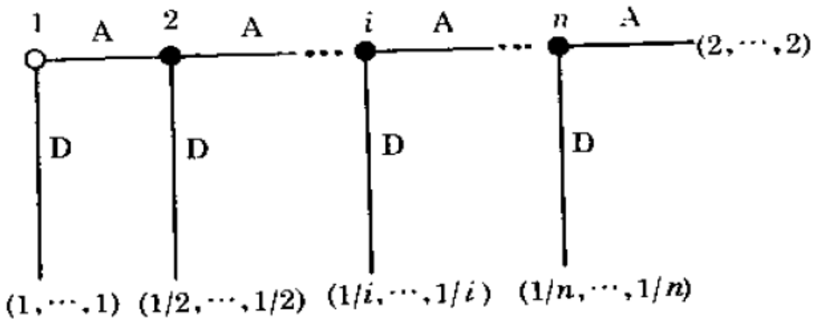


图 3－13 接头博弈

3.4 应用 II：斯塔克博格模型

我们再介绍一个连续模型。（回忆古诺模型的基本含义。）根据 Stackelberg（1934），企业 1、2 先后选择产量q_i，逆需求函数为$p(Q)=a-Q$，利润函数为$\pi_i = q_i[p(Q)-c]$，固定成本为 0。运用逆向归纳法，先考虑企业 2 的最佳反应$R_2(q_1)$应满足

$$\underset{q_2\ge 0}{Max}\,\pi_2(q_1,q_2)=\underset{q_2\ge 0}{Max}\,q_2(a-q_1-q_2-c)$$

解得$R_2(q_1)=\dfrac{a-q_1-c}{2}$。（这一步与古诺模型一样。）

类似地，企业 1 的问题是$\underset{q_1\ge 0}{Max}\,\pi_1(q_1,R_2(q_1))=\underset{q_1\ge 0}{Max}\,q_1(a-q_1-R_2(q_1)-c)$。将$R_2(q_1)$

代入（这一步与古诺模型不同，因为$\frac{dq_1}{dq_2} \neq 0$），解得$q_1^* = \frac{a-c}{2}$，$R_2(q_1^*) = \frac{a-c}{4}$。

另一种解法是：企业 1 的利润函数表达式直接对q_1求导数，同时将$R_2(q_1)$对q_1求导数的结果代入，也可以求出q_1^*。

不妨与古诺模型对比一下：

（1）企业 2 的反应函数都一样，不过在古诺模型中q_1和q_2是同时决定的，因此在均衡状态下$\frac{dq_1}{dq_2} = 0$。但在斯塔克伯格模型的第一步，企业 1 尚未行动时，企业 2 的行动会影响到企业 1 的决策，故$\frac{dq_1}{dq_2} \neq 0$。

（2）总产量：古诺模型中每个企业都生产$\frac{a-c}{3}$，总产量为$\frac{2(a-c)}{3}$。在这里，总产量为$\frac{3(a-c)}{4}$，总产量上升了，因此价格由$\frac{a+2c}{3}$（古诺模型）下降为$\frac{a+3c}{4}$。为什么？

（3）总利润：在古诺模型中，总利润为$(q_1+q_2)[a-(q_1+q_2)-c] = \frac{2(a-c)^2}{9}$；在这里，总利润为$\frac{3(a-c)^2}{16} < \frac{2(a-c)^2}{9}$。

（4）信息：既然企业 1 可以在一开始就选择古诺产量，却没有选择，说明企业 1 现在选择的产量能够获得更多的利润。给定总产量上升，价格下降，这导致企业 2 的利润下降了。这说明：第一，先行动可以成为一种优势；第二，掌握更多信息并不一定是好事。

3.5 承诺

在多阶段动态博弈中，虽然先动者的行动会影响后动者的行动（从而获得一定的先动者优势），但是后动者也可以传递某种信息，来影响先动者的决策。例如，后动者可以对先动者提出威胁（此时他算是实际上的“先动者”吗？）。然而，有些威胁是不可置信的，因为威胁给威胁者自己带来的成本高于其收益。例如，司马相如与卓文君的“私奔博弈”（图 3－14）。

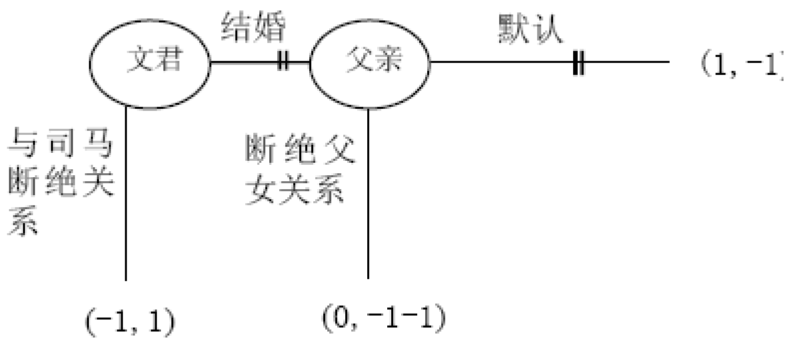


图 3－14 私奔博弈

使威胁置信的办法就是做出承诺（commitment），即通过限制自己的选择集使对方选择自己希望的行动。所谓“置之死地而后生，投之亡地然后存”（孙子兵法）是也。古有“破釜沉舟”，今有“反分裂法”。2005 年 3 月通过的法律使“不放弃使用武力”成为一种可置信的威胁。如图 3－15（董志强，2007）。

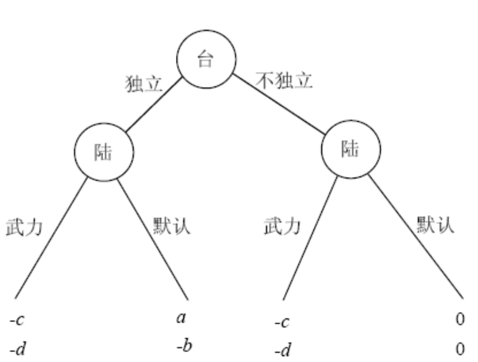


图 3－15a 台海博弈（前）

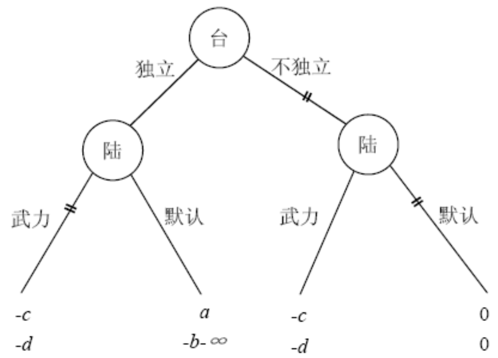


图 3－15b 台海博弈（后）

分析：

- 如果 d>b，那么台湾独立大陆必定默认，台湾不独立大陆也将默认；逆推到第一阶段则台湾必定独立；
- 如果 d<b，那么台湾独立大陆必定动武，台湾不独立大陆将默认；逆推到第一阶段则台湾将不独立（不独立得到 0，独立得到-c）。

这意味着，如果战争的代价超过失去台湾的代价，那么（独立，默认）将是 SPE。因此，台湾方面就会有意提高战争的代价，例如购置军火。通过了《反分裂国家法》之后，情况发生了变化。

- 即使战争的代价超过失去台湾的代价(d>b)，由于反分裂法的存在使得武力是可置信的，结果台海关系的均衡回到了和平的轨道上。所以反分裂法的确是和平促进法。
- 台湾将没有动力去制造一场巨大的战争，因为提高 d 已经没有什么意义。
- 战争力量的对比不是关键。即使大陆的军力处于战争劣势(d>c)，也不会影响博弈的和平均衡路径。

Shepsle（1991）将可信承诺分成两类：“动机意义上的可信承诺”，即声誉机制，它符合参与人的激励相容约束；“强制意义上的可信承诺”，即外在约束保证的承诺。后者又可以分为三种方式：（1）信息的分散化，例如分税制导致中央对于地方的具体事务缺乏准确信息，因此“财政联邦制”是可信的承诺（Qian and Weingnast，1997）；（2）权力的分散化，例如英国“光荣革命”后确立了君主立宪制，结果国王的借债能力反而提高了（North and Weingast，1989）；（3）引入竞争，例如存在大量竞争性企业时，每个企业获取政府补贴的成本就很高，政府不补贴企业的承诺就更可信（Segal，1998）。

思考：还有其它做出可信承诺的办法吗？

真实世界中的案例：公开表态追求女生，“居然之家”承诺“先行赔付”，前线军官切断联系，班主任交出任命班干的权力。

思考题：政府可以和恐怖分子谈判吗？

3.6 有限重复博弈：连锁店悖论

3.6.1 准备知识

贴现率（discount rate）：r，相当于利率。贴现因子（discount factor）：$\delta=\dfrac{1}{1+r}$，且$\delta\in[0,1]$。如果同时考虑博弈结束的概率θ和时间偏好，那么$\delta=\dfrac{1-\theta}{1+r}$。

对于连续时间而言，贴现因子$\delta=e^{-rt}$（因为$\lim\limits_{m\to\infty}(1+\dfrac{1}{m})^m=e$，所以$\lim\limits_{m\to\infty}(1+\dfrac{r}{m})^{tm}=\lim\limits_{m\to\infty}[(1+\dfrac{r}{m})^{m/r}]^{rt}=e^{rt}$）。

平均贴现因子：$1-\delta$，因为$A(1+\delta+\delta^2+...+\delta^n)=T\Rightarrow A=(1-\delta)T$。

[注 1] 贴现因子可以大于 1，表示对未来极度依赖。

动态博弈：
- 序贯博弈（sequential games）：博弈只进行一个回合。
- 重复博弈（repeated games）：阶段博弈（stage game）重复多次。
  - 完全且完美信息重复博弈（动态重复）
  - 完全但不完美信息重复博弈（静态重复）

[注 2] 重复博弈的策略集非常大，即使囚徒困境重复 5 次，每个囚徒的纯策略数量大于 20 亿个，策略组合的数目更大。而且，此时不存在优势策略。

3.6.2 连锁店悖论

现实中存在很多连锁店，例如麦当劳。但是根据有限重复博弈论，连锁店应该被对手逐步击垮，此即“连锁店悖论”（chain-store paradox）（Selten，1978）。如图 3－16。

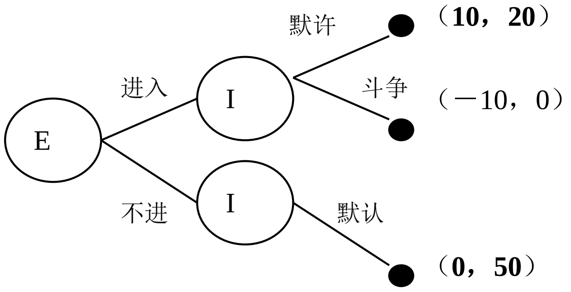


图 3－16 连锁店悖论

此博弈的阶段博弈有一个 SPE——（进入，默许）。现在，假定在位者有 20 个连锁店，那么在位者应该选择斗争吗？答案是否定的。

逆向归纳法：假定进入者已经进入了 19 个店，当然要选择进入第 20 个店，而在位者只有默许。如此逆推，则进入者会进入所有的店，因此连锁店所构成的垄断优势被击破。

[注] 这里没有考虑连锁店的规模经济优势以及声誉优势，这是现实生活中存在连锁店的一些原因。

[思考]为什么麦当劳和肯德基总是“扎堆”出现？

3.6.3 有限重复博弈定理

考虑囚徒困境（或者斯塔克伯格博弈）重复有限次，可以证明，结果仍是（坦白，坦白）（与每个阶段博弈的结果一样）。如何验证？

我们归纳出一个定理：

有限重复博弈定理（Gibbons，1992）：如果阶段博弈或者一个完全且完美信息博弈 G 有唯一的纳什均衡/子博弈完美纳什均衡，那么对任意有限的 T，重复博弈 G（T）有唯一的子博弈完美纳什均衡解，即 G 的均衡结果在每一阶段重复进行。

[注] 如果“唯一性”不能保证，那么此定理不成立。如图 3－17。一次博弈的纳什均衡是（M，L）、（U，M）和（（3/7U，4/7 M），（3/7L，4/7 M）），赢利向量分别是（4，3）、（3，4）和（12/7，12/7）。但是在两阶段博弈中，只要贴现因子大于 7/9，如下策略构成一个 SPE：第一阶段选择（D，R）；如果如此，则第二阶段选择（M，L）；否则第二阶段选择（（3/7U，4/7 M），（3/7L，4/7 M））。这使得帕雷托最优的结果（D，R）可以实现。这说明：有限重复博弈可能存在一种奖惩机制，使得原本不是均衡的解会出现在均衡路径上。

	L	M	R
U	0，0	3，4	6，0
M	4，3	0，0	0，0
D	0，6	0，0	5，5

图 3－17 纳什均衡不唯一的博弈

3.7 无限重复博弈：无名氏定理

3.7.1 菜市场的无限重复博弈

思考：为什么旅游景点的东西又贵又差？为什么农村的欺骗现象比城里少？

无限重复博弈可以改变有限重复博弈的不良结果，比如菜市场的质量博弈，图 3－18。

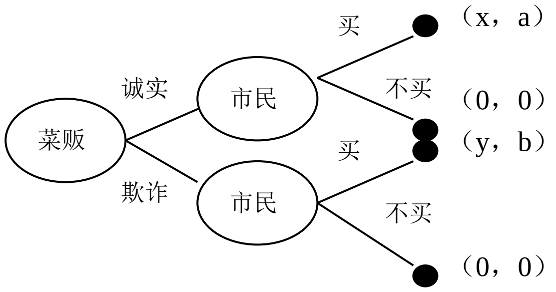


图 3－18 菜市场博弈

分析：

（1）一次博弈：假设 y>x，且顾客短期内无法区分 a、b（经验品或信任品），那么（欺诈，不买）就是唯一的纳什均衡，而且欺诈是菜贩的弱优势策略。

（2）有限重复博弈：结果仍是一样。

（3）无限重复博弈（经济含义）：假定贴现因子为δ，那么菜贩子诚实的充要条件是

$$x + x\delta + ... \geq y + 0$$

$$\frac{x}{1-\delta} \geq y \Rightarrow \delta \geq 1 - \frac{x}{y}$$

如果$y = 2x$，则$\delta \geq 0.5$，且$\frac{d\delta}{dy} > 0$，$\frac{d\delta}{dx} < 0$。这里，δ可以理解为交易重复的概率。进一步解释模型的经济含义，参考张维迎《法律制度的信誉基础》（《经济研究》2002 年第 1 期）。这其实是一个简单的声誉模型（reputation）。在经济学中，“诚实”或“声誉”是理性利益计算的结果，与“品德”无关！

林肯说：“你可欺骗所有人于一时，或欺骗部分人于永远，但你不可能永远欺骗所有的人!”试用重复博弈论分析这句话。如何理解“菩萨心肠，金刚手段”？

3.7.2 无限重复的囚徒困境

将囚徒困境复制如下：

1 \ 2	坦白	抵赖
坦白	**－8，－8**	0，－10
抵赖	－10，0	－1，－1

在无限重复的囚徒困境博弈中，能否得到合作结果取决于囚徒的策略。先考虑囚徒 1 使用“冷酷策略”（grim strategies），即 1 开始选择抵赖（合作），直到一方坦白（背叛），然后永远选择坦白，又称“触发策略”（trigger strategies）。现在考虑囚徒 2 的最佳反应。2 也选择抵赖的必要条件是

$$-1 + \delta(-1) + \delta^2(-1) + ... \geq 0 + \delta(-8) + \delta^2(-8) + ...$$

即$-\frac{1}{1-\delta} \geq -\frac{8\delta}{1-\delta} \Rightarrow \delta^* \geq \frac{1}{8}$。因此，当贴现因子满足条件时，双方都选择冷酷策略构成一个纳什均衡。对于整个博弈而言，它也是 SPE 之一，至少还存在另一个 SPE——双方都永远选择坦白。

现在考虑囚徒 1 选择“针锋相对策略”（tit-for-tat），即 1 开始选择抵赖，在t阶段选择对方在$t-1$阶段的行动，但不持续复仇。这是不是一个纳什均衡呢？考虑到对称性，可以如此总结：

1	2	1	2
抵赖	抵赖	坦白	抵赖
－1	－1	0	－10
－1	－1	－10	0
…	…	…	…
…	…	…	…

阿克塞罗德（Axelrod，1984）在一些博弈论实验中证明，针锋相对策略是团队竞赛中平均收益最高的策略。这一策略依赖于三个条件：代理人首先选择合作（这里是“抵赖”）；代理人“以德报德、以怨报怨”；代理人不持续复仇。然而针锋相对策略不是 SPE。因为一旦对手选择了坦白，此时自己选择坦白是更有利的，而选择上一期对手采取的坦白来应对对方本期的抵赖也不是明智的。

事实上，无限重复博弈的麻烦不是缺少均衡，而是多重均衡。我们将其正式地归纳为“无名氏定理”（folk theorem）：

Theorem 1 (the Folk Theorem)

In an infinitely repeated n-person game with finite action sets at each repetition, any profile of actions observed in any finite number of repetitions is the unique outcome of some subgame perfect equilibrium given

Condition 1: *The rate of time preference is zero, or positive and sufficiently small;*

Condition 2: *The probability that the game ends at any repetition is zero, or positive and sufficiently small; and*

Condition 3: *The set of payoff profiles that strictly Pareto dominate the minimax payoff profiles in the mixed extension of the one-shot game is n- dimensional.*

上述条件必须同时满足。简单地理解，只要参与人的耐心足够大，那么任何满足个人理性的赢利向量都可以通过一个特定的 SPE 得到。这给人类走出“囚徒困境”或寻求合作提供了理论基础。事实上，无限重复博弈可能有无穷多个均衡结果。“一切皆有可能。”

[注 1] 任何一个结果依赖于博弈进行很多次的模型都要假设贴现率（或贴现因子）不是很高（或低）。如菜市场博弈。

[注 2] 逻辑与注 1 类似。注意“博弈将在 T 以前的某个不确定的时点结束”和“博弈结束的概率为 p”是不一样的，前者近似于有限重复博弈。

[注 3] “最小最大”（minimax）赢利指当其他参与人最严厉地惩罚某个参与人时，该参与人能够得到的最高赢利。数学表达为$\underset{s_{-i}}{Min}\,\underset{s_i}{Max}\,\pi_i(s_i,s_{-i})$。直观地说，就是该参与人的保留效用。显然，纳什均衡下的赢利不能低于保留效用。N 维的要求表示 N 人博弈。

与强调最小化违规者收益的“最小最大策略”对应的，是违规者为了保护自己而实行的“最大最小化策略”（maximin strategies）。两者的区别是，minimax 的目的是最严厉地惩罚别人（这可能是损人不利己）；maximin 的目的是最大程度保护自己，或者说是一种“底线思维”。例如：项目 A 有 50%的概率实现年收益 10 万元，50%的概率实现年收益 30 万元；项目 B 有 50%的概率实现年收益 50 万元，50%的概率实现年收益 0 万元（亏损）。风险中性和极端风险规避（底线思维）的选择是相反的。两者都可以是纯策略或者混合策略。

例如，在福利博弈中，贫民的 maximin 策略是工作，minimax 策略为流浪；政府的 minimax 为不管。

		贫民	
		工作	流浪
政府	救济	3，2	－1，3
	不管	－1，1	0，0

图 2－18 福利博弈

		鸟 2	
		老鹰	鸽子
鸟 1	老鹰	－1，－1	2，0
	鸽子	0，2	1，1

图 3－19 鹰鸽博弈

这两种策略都不一定构成均衡，因为它缺乏理性作为共同知识。例如，当一方采取 maximin 策略时，它必须确认对方把“恶意”作为目的，而不是以“利己”为目的。只有在少数零和博弈中，它们是等价的，而且都是纳什均衡。此即“最小最大定理”（von Neumanan，1928）。

3.8 补充内容

3.8.1 演化博弈

经典博弈的假设或许太强了，例如参与人可能是有限理性的（bounded rationality）[①]，或者环境的偶然性可能导致一些怪异行为，此时纳什均衡或许是不恰当的。与此不同，演化博弈论是生物学与经济学的结合，假设低级的理性和环境的变异性，强调均衡的稳定性，其核心概念是“演化稳定策略”（evolutionarily stable strategy，ESS）。相比较而言，经典博弈论强调给定环境下的最优选择，而演化博弈论强调在变异环境下的“适者生存”行为；前者强调个体理性，后者强调群体行为；前者个体可以采取不同策略，后者都采取相同策略。

典型的例子是鹰鸽博弈，如图 3－19。一群鸟在一起生存并争取食物，食物总共有 2 个单位。每种鸟都可以采取两种策略：老鹰或鸽子。鹰鸽博弈与斗鸡博弈非常类似，不过对于生物群体而言，这里没有纯策略均衡，因此也没有纯策略的 ESS。因为无论是全部成为老鹰还是鸽子，都不能在环境中生存下来（注意：此处的策略一旦选定就意味着全体鸟都是一种类型，而不能象斗鸡博弈一样变换行动）。例如，在都是鸽子的情况下，一只老鹰入侵将得到更多的收益。在混合策略 ESS 中，均衡策略是各以 50％的概率扮作老鹰和鸽子（期望收益为 0.5=0.5*2+0.5*-1），这意味着 50％的鸟是老鹰，50％的鸟是鸽子。如果在一个树林中有 60%的鸟是老鹰，那么一只鸟扮作鸽子会有更高的收益。这或许可以解释“生态平衡”。

3.8.2 马尔可夫均衡

对于无限重复博弈而言，几乎不可能找出所有的 SPE，因此我们可以只找出完美马尔可夫均衡（perfect Markov Equilibrium）。在这种均衡中，每个参与人只考虑上一期的行动，而忽略其他历史。PME⊂SPE。PME 在制度经济学或迭代模型中经常使用。

参考文献：

[美]罗伯特·阿克塞尔罗德(Robert Axelrod)，《合作的进化》（The evolution of

① 关于有限理性，可参考两位诺贝尔经济学奖得主 Simon 和 Williamson 的观点。

cooperation），吴坚忠翻译，上海人民出版社 2016 年版。

乔根・W.威布尔，2006 年，《演化博弈论》（Evolutionary Game Theory），上海人民出版社。

“Tit for tat”, Wikipedia, http://en.wikipedia.org/wiki/Tit_for_tat