▶30

博弈论的应用

在上一章，我们描述了博弈论中的许多重要的概念，并通过一些例子对它们进行了说明。在本章，我们将考察博弈论中的 4 个非常重要的问题——合作问题、竞争问题、共存问题和承诺问题——以搞清楚在各种策略互动过程中，它们是如何发挥作用的。

为了做到这一点，我们首先要创建一个重要的分析工具，即*最优反应曲线*，它可以用来解决博弈论中的均衡问题。

30.1　最优反应曲线

考虑一个双人博弈，假定你是其中的一个参与人。对于另一个参与人的任何选择，你的*最优反应*是使你的收益实现最大化。如果最大化收益的选择不止一个，那么，你的最优反应就是所有这些选择的集合。

举例来说，考察表 30.1 所示的博弈，我们可以用它来阐明纳什均衡的概念。如果列参与人选择“左”，行参与人的最优反应是选择“上”；如果列参与人选择“右”，那么，行参与人的最优反应就是选择“下”。同样，列参与人对于行参与人选择“上”的最优反应是选择“左”，而对于行参与人选择“下”的最优反应是选择“右”。

表 30.1　一个简单的博弈

		列参与人	
		左	右
行参与人	上	2，1	0，0
	下	0，0	1，2

我们可以将上述情形表述为下面的表格形式：

列参与人的选择	左	右
行参与人的最优反应	上	下
行参与人的选择	上	下
列参与人的最优反应	左	右

注意,如果列参与人认为行参与人将选择“上”,那么,列参与人就会选择“左”;而如果行参与人认为列参与人将选择“左”,那么,行参与人就会选择“上”。因此,在这里,每一个参与人都在针对另一方的选择作出最优的反应,从这层意义上看,选择组合(上,左)是相互一致的。

考虑一个一般的双人博弈,其中,行参与人可能选择 r_1, …, r_R,列参与人可能选择 c_1, …, c_C。对于行参与人的每一个选择 r, $b_c(r)$ 表示列参与人的最优反应。相应地,对于列参与人的每一个选择 c, $b_r(c)$ 表示行参与人的最优反应。由此,纳什均衡是使得以下两个式子成立的一个策略组合(r^*, c^*):

$$c^* = b_c(r^*)$$
$$r^* = b_r(c^*)$$

纳什均衡的概念体现了“相互一致性”的思想。如果行参与人预期列参与人将选择“左”,那么,他就会选择“上”;而如果列参与人预期行参与人将选择“上”,他就会选择“左”。因此,在纳什均衡中,相互一致的是参与人的信念和行动。

注意,在某些情况下,其中的一个参与人也许在几个最优反应之间无差异。这就是我们只要求“c^* 是列参与人的最优反应之一,r^* 是行参与人的最优反应之一”的原因。如果对于每一个选择,存在唯一的最优反应,那么,最优反应曲线就可以由最优反应函数表示。

这种考虑纳什均衡概念的方式使得我们很容易看到,纳什均衡只是第 28 章所描述的古诺均衡的一般形式。在古诺模型中,选择变量是产量,它是一个连续变量。古诺均衡具有这样的性质,给定其他厂商的选择,每一家厂商都会选择使利润最大化的产量。

第 28 章还介绍了伯特兰均衡,它是一个定价策略的纳什均衡。给定对其他厂商将要作出选择的预期,每一家厂商都会选择使利润最大化的价格。

这些例子展示了最优反应曲线是如何使前文提到的模型一般化的,并且,它还提供了一种相对简单的求解纳什均衡的方法。这些性质使得最优反应曲线成为求解博弈均衡的一种非常有用的工具。

30.2 混合策略

接下来,我们将利用最优反应曲线来分析表 30.2 所示的博弈。

表 30.2 求解纳什均衡

		列参与人	
		左	右
行参与人	上	2, 1	0, 0
	下	0, 0	1, 2

我们对于寻求混合策略均衡与纯策略均衡一样感兴趣,所以,我们令 r 表示行参与人选择“上”的概率,那么,$(1-r)$ 就表示他选择“下”的概率。同样,令 c 表示列参与人选择“左”的概率,那么,$(1-c)$ 就表示他选择“右”的概率。当 r 和 c 等于 0 或者 1 时,相应的

策略就是纯策略。

我们计算当行参与人按概率 r 选择“上”，而列参与人按概率 c 选择“左”时，行参与人的期望收益。请看以下的排列：

组　合	概　率	行参与人的收益
上，左	rc	2
下，左	$(1-r)c$	0
上，右	$r(1-c)$	0
下，右	$(1-r)(1-c)$	1

为了计算行参与人的期望收益，我们对第 3 列中行参与人的各项收益进行加权平均，权数就是第 2 列中各项收益出现的概率，答案是

$$\text{行参与人的期望收益} = 2rc + (1-r)(1-c)$$

经过重新整理，我们有

$$\text{行参与人的期望收益} = 2rc + 1 - r - c + rc$$

现在，假定 r 增加了 Δr，那么，行参与人的收益会如何变化呢？

$$\text{行参与人的收益变化} = 2c\Delta r - \Delta r + c\Delta r = (3c-1)\Delta r$$

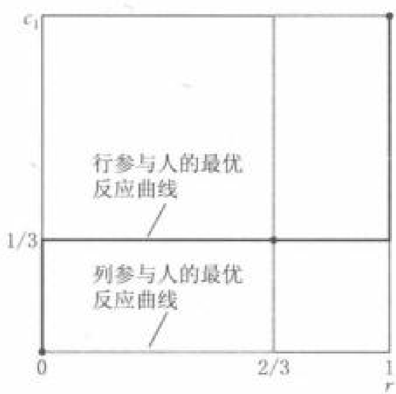
当 $3c>1$ 时，上式取正值；当 $3c<1$ 时，上式取负值。因此，当 $c>1/3$ 时，行参与人会提高 r 值；而当 $c<1/3$ 时，行参与人会降低 r 值；当 $c=1/3$ 时，他对于任意的 $0\leqslant r\leqslant 1$ 无差异。

同样，列参与人的收益由下式给出：

$$\text{列参与人的期望收益} = cr + 2(1-c)(1-r)$$

当 c 增加 Δc 时，列参与人的收益变化如下：

$$\text{列参与人的收益变化} = r\Delta c + 2r\Delta c - 2\Delta c = (3r-2)\Delta c$$



这两条曲线分别体现了行参与人和列参与人对应于对方选择时的最优反应。曲线的交点就是纳什均衡。在这种情况下，博弈存在 3 个均衡，其中，两个是纯策略均衡，一个是混合策略均衡。

图 30.1　最优反应曲线

因此，当 $r>2/3$ 时，列参与人会提高 c 值；当 $r<2/3$ 时，他会降低 c 值；而当 $r=2/3$ 时，他对于任意的 $0\leqslant c\leqslant 1$ 无差异。

这里，我们可以利用这个信息来绘制最优反应曲线。先从行参与人开始。如果列参与人选择 $c=0$，那么，行参与人就会使 r 值尽可能得小，所以，$r=0$ 就是 $c=0$ 时行参与人的最优反应。并且，$r=0$ 一直都是行参与人的最优反应，直到 $c=1/3$ 时为止。当 $c=1/3$ 时，位于 0 和 1 之间的任意 r 值都是最优反应。对于所有的 $c>1/3$，行参与人的最优反应是 $r=1$。

图 30.1 显示的是这些曲线。不难发现，它们相交于 3 个点：(0, 0)、(2/3, 1/3) 和 (1, 1)，这 3 个点分别对应于该博弈的 3 个纳什均衡。在这些均衡中，两个是纯策略均衡，一个是混合策略均衡。

30.3 合作博弈

利用上一节介绍的工具，我们现在就可以考察第一类博弈，即协调博弈。在这类博弈中，当参与人能够协调他们之间的策略时，他们的收益就会实现最大化。实际中的问题是要创建一种能够实现这种协调的机制。

性别战

有关协调博弈的经典例子是性别战。在这个博弈中，一个男孩和一个女孩决定约会看电影，但他们事前并没有安排好看哪一部。更糟糕的是，他们都忘记了对方的电话号码，所以，他们没有办法协调他们的约会，而只能猜测对方喜欢看哪一部影片。

男孩想看最新推出的动作片，而女孩更愿意欣赏文艺片，但是，他们都宁愿看同一部电影也不愿意放弃约会。表 30.3 显示了与这些偏好一致的收益。请注意协调博弈的定义性特征：参与人在他们协调行动时获得的收益要大于不协调时的收益。

表 30.3 性别战

		女孩	
		动作片	文艺片
男孩	动作片	2，1	0，0
	文艺片	0，0	1，2

这个博弈的纳什均衡是怎样的呢？幸运的是，这恰好就是上一节中用来阐明最优反应曲线的博弈。我们已经知道，这个博弈存在 3 个均衡：双方都选择动作片，双方都选择文艺片，或者双方分别按 2/3 的概率选择各自偏好的影片。

由于这些都是可能的均衡，所以，仅仅依据上述的分析，我们并不能确定究竟哪种均衡会发生。一般地，我们要依靠有关该博弈的正式描述以外的因素，来解决这个问题。举例来说，假定放映文艺片的影院距离其中的一个参与人较近。于是，双方都有理由相信观看文艺片将是最终的均衡。

当参与人完全有理由相信，其中的一个均衡相对于其他均衡更为“自然”时，这个均衡就称作这个博弈的聚点。

囚徒困境

我们在上一章广泛讨论的囚徒困境也是一个协调博弈。回顾这个博弈：两个囚徒或者选择坦白，从而招供另一方；或者选择抵赖，拒绝承认犯罪。这个博弈的收益如表 30.4 所示。

表 30.4 囚徒困境

		参与人 B	
		坦白	抵赖
参与人 A	坦白	−3，−3	0，−6
	抵赖	−6，0	−1，−1

囚徒困境的显著特征是,坦白是一个占优策略,即使按总收益的标准,协调(双方都选择抵赖)是更优越的选择。协调能够使囚犯获得最大化的收益,但问题是,不存在一种使它能够在一次性博弈中实现的简易方式。

解决囚徒困境的一种方式是,通过增加新的选择以扩展原先的博弈。在上一章,我们看到,无限重复的囚徒困境博弈通过针锋相对的策略,是能够实现协调博弈的结果的,在这种博弈中,参与人通过将来的行动来奖励合作和惩罚不合作。这里的额外策略性因素是,今天拒绝合作就会招致明天的惩罚。

解决囚徒困境的另一种方式是增加缔结合约的可能性。举例来说,参与人双方可以签署一份合同,表示他们将恪守合作性策略。如果任何一方违背合同,违约方将支付罚金或者接受其他某种方式的惩罚。合同在实现各种结果方面是非常有用的,但它却要依赖于能够强制执行这种合同的法律体制的存在。这对于商业谈判的情形是有意义的,但在其他场合,如军事博弈或国际谈判,它就不再是一个恰当的假设。

保证博弈

考虑20世纪50年代美国和苏联之间的军备竞赛,当时,这两个国家都可以选择生产核导弹,也可以选择不生产。表30.5显示了与这些策略相对应的收益。最优的结果是双方都不生产核导弹,此时的收益为(4, 4)。但是,如果一方不生产核导弹而另一方生产,那么,生产一方的收益变为3,而不生产一方的收益变为1。双方同时生产核导弹时的收益为(2, 2)。

表30.5 军备竞赛

		苏联	
		不生产	生产
美国	不生产	4, 4	1, 3
	生产	3, 1	2, 2

不难发现,这里存在两个纯策略纳什均衡:(不生产,不生产)和(生产,生产)。但是,(不生产,不生产)对于双方都是一个较好的选择。问题在于,任何一方都不知道对方将会作出的选择。在承诺不生产以前,每一方都想得到对方不会生产的保证。

获得这种保证的一条途径是,其中的一方先采取行动,并接受公开的检查。注意,这可以是一种单边的行动,至少在它相信博弈中的收益时是如此。如果一方宣称它已经停止部署核导弹,并给予另一方充分的证据来证明它的选择,那么,可以确信无疑的是,另一方也会停止部署核导弹。

懦夫博弈

有关协调博弈的最后一个例子是在电影中非常流行的汽车博弈(懦夫博弈)。两个年轻人分别从一条街的两头,驾车笔直地驶向对方。第一个转向的人会颜面尽失成为懦夫;如果没有人转向,那么,他们最终会撞在一起。表30.6显示了这个博弈的可能收益。

表 30.6　懦夫博弈

		列参与人	
		转向	不转向
行参与人	转向	0，0	−1，1
	不转向	1，−1	−2，−2

这里存在两个纯策略纳什均衡：（行转向，列不转向）和（行不转向，列转向）。列参与人偏好第一个均衡，而行参与人则偏好第二个均衡，但每一个均衡都要比撞车好。注意这个博弈与保证博弈之间的差异；在保证博弈中，与做不同的事情相比较，博弈双方做相同事情（同时生产核导弹，或者，同时不生产核导弹）时的境况较好。而在这里，博弈双方做相同事情（直线驾驶或转向）时的境况要劣于做不同事情时的境况。

每一个参与人都知道，如果他能够承诺直线驾驶，对方就会因惧怕而放弃直线驾驶。但是，每一个参与人也都知道，两车相撞是一件很疯狂的事情。因此，参与人如何坚持他们所偏好的均衡呢？

一个重要的策略是作出某种承诺。假定行参与人在出发前，很夸张地将汽车的方向盘锁住。于是，列参与人就会意识到，行参与人现在只有直线驾驶而别无其他选择，此时，他将选择转向。当然，如果双方都把方向盘上了锁，结果将是灾难性的！

如何协调

如果你是协调博弈的一个参与人，那么，你也许愿意与对方在双方都偏爱的均衡处合作（保证博弈）；或者与对方在其中一方偏爱的均衡处合作（性别战）；或者采取有别于均衡策略的策略（囚徒困境）；或者作出某种选择从而实现你所偏爱的结果（懦夫博弈）。

在保证博弈、性别战和懦夫博弈中，这一点可以通过某一方先采取行动，并承诺选择某个特定的策略来实现。然后，对方就可以观察到第一个参与人的选择，并相应地作出反应。在囚徒困境中，这种策略并不起作用：如果博弈的一方选择抵赖，那么，另一方的最佳选择也是选择抵赖。除序贯行动以外，声誉和缔结合同是“解决”囚徒困境的主要方式。

30.4　竞争博弈

与合作相对应的另一个极端是竞争。它就是有名的零和博弈，之所以称作零和博弈，是因为在这种博弈中，博弈一方的收益等于另一方的损失。

实际上，大多数体育竞技项目都是零和博弈：一个组得 1 分等价于另一个组失 1 分。由于参与人之间的利益是完全相反的，这类博弈中的竞争非常激烈。

我们通过一个足球比赛的例子，来阐明零和博弈是如何进行的。行参与人主罚点球，列参与人防守。行参与人可以踢向球门的左方，也可以踢向球门的右方；列参与人可能偏好球门的某一方，从而扑向球门的左方或球门的右方，以拦截来球。

我们用预期的百分点数来表示这些策略产生的收益。很显然，如果列参与人扑错了方向，行参与人得分的可能性就较大一些。另一方面，这个博弈并非完全对称，这是因为，行参与人可能善于朝某个方向踢；同样，列参与人可能善于按某个方向防守。

我们假定,如果行参与人踢向球门的左方,那么,当列参与人扑向右方时,行参与人将在80%的时间内得分,当列参与人扑向左方时,行参与人将在50%的时间内得分;如果行参与人踢向球门的右方,那么,当列参与人扑向左方时,行参与人将在90%的时间内得分,当列参与人扑向右方时,行参与人将在20%的时间内得分。这些收益列示在表30.7中。

表30.7　足球赛中的罚点球得分

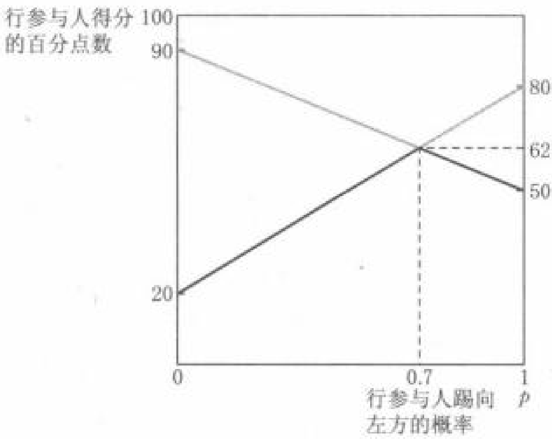
		列参与人	
		扑向左方	扑向右方
行参与人	扑向左方	50, −50	80, −80
	扑向右方	90, −90	20, −20

注意,每一方格中的总收益为零,这显示参与人的得分完全相反。行参与人试图最大化他的期望收益,列参与人也想使自己的期望收益最大化——也就是说,最小化行参与人的收益。

很明显,如果列参与人知道行参与人踢球的方向,他就拥有极大的优势。意识到这一点,行参与人会竭力使得列参与人猜不透自己的意图。具体地,他有时会踢向他擅长的一方,有时会踢向他不擅长的一方。这就是说,他采取的是一种混合策略。

如果行参与人选择踢向左方的概率为 p,那么,当列参与人扑向左方时,行参与人的期望收益为 $50p+90(1-p)$;当列参与人扑向右方时,行参与人的期望收益为 $80p+20(1-p)$。行参与人想使这个期望收益尽可能得大,而同时,列参与人却想使这个期望收益尽可能得小。

举例来说,假定行参与人选择踢向左方的概率是50%。当列参与人扑向左方时,行参与人的期望收益为 $50\times1/2+90\times1/2=70$;当列参与人扑向右方时,行参与人的期望收益为 $80\times1/2+20\times1/2=50$。



这两条直线表示的是行参与人的期望收益,它是 p 的函数,而 p 则是行参与人踢向左方的概率。无论行参与人选择怎样的 p 值,列参与人将竭力使行参与人的收益最小化。

图30.2　行参与人的战略

当然,上述的推理过程也同样适用于列参与人。如果列参与人相信行参与人踢向左方的概率为50%,那么,他会扑向右方,因为这是使行参与人的期望收益最小化的选择(从而也是使列参与人的期望收益最大化的选择)。

图30.2显示了对应于不同的概率选择 p,行参与人的期望收益。这只要把两个函数 $50p+90(1-p)$ 和 $80p+20(1-p)$ 绘制在图上即可。由于这两个表达式都是 p 的线性函数,所以它们的图形都是直线。

行参与人意识到,列参与人总是在试图最小化他的期望收益。因此,对于任意的 p,他有望获得的最佳收益是这两个策略

所给出的收益中的最小值。在图30.2中，我们用深色线段来表示这些最小值。

这些最小收益的最大值出现在什么位置呢？很明显，它位于深色线段的顶点，或者等价地，它位于这两条直线相交的地方。用代数方法求解下式，我们就可以得到这个 p 值：

$$50p+90(1-p)=80p+20(1-p)$$

经过验证，$p=0.7$。

因此，如果行参与人踢向左方的概率为 70%，并且，列参与人相应地作出最优反应，那么，行参与人将获得的期望收益是 $50\times0.7+90\times0.3=62$。

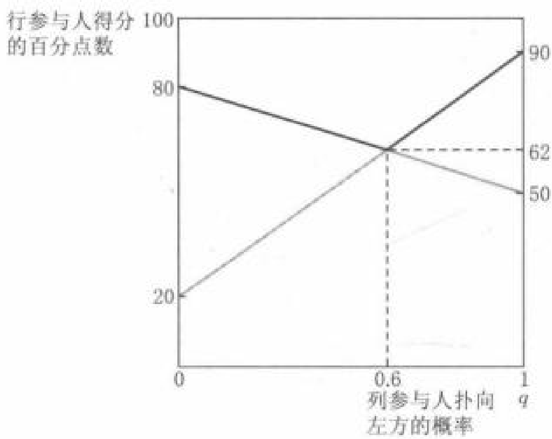
列参与人的情况又如何呢？对于它的选择，我们可以进行相似的分析。假定列参与人选择扑向左方的概率为 q，从而其扑向右方的概率为 $(1-q)$。由此，当行参与人踢向左方时，行参与人的期望收益是 $50q+80(1-q)$；当行参与人踢向右方时，行参与人的期望收益是 $90q+20(1-q)$。对于任意的 q，列参与人将使得行参与人的收益最小化。但是，列参与人同时意识到，行参与人想最大化这个收益。

因此，如果列参与人选择扑向左方的概率是 1/2，他意识到，行参与人在踢向左方时获得的期望收益为 $50\times1/2+80\times1/2=65$，在踢向右方时获得的期望收益为 $90\times1/2+20\times1/2=55$。这种情况下，行参与人当然会选择踢向左方。

我们可以把这两个收益函数绘制在图 30.3 中，它们与前面的图形类似。从列参与人的角度看，两条直线的较高部分是重要的，这是因为，这部分线段反映了对于每个 q，行参与人的最优选择。因此，图 30.3 将这部分线段加深了颜色。与前文一样，我们可以找到列参与人的最佳 q 值——在该点，行参与人的最大收益取最小值。也就是说，下面的式子成立：

$$50q+80(1-q)=90q+20(1-q)$$

这意味着 $q=0.6$。



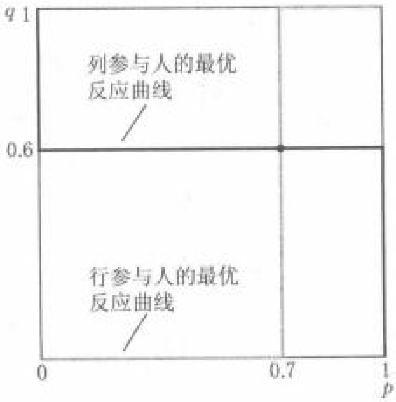
这两条直线表示的是行参与人的期望收益，它是 q 的函数，而 q 则是列参与人扑向左方的概率。无论列参与人选择怎样的 q 值，行参与人将竭力最大化自己的收益。

图 30.3 列参与人的策略

迄今为止，我们已经计算出了这两个参与人的均衡策略。行参与人应该按概率 0.7 踢向左方，列参与人应该以概率 0.6 扑向左方。这些概率值使得不论对方采取什么行动，行参与人和列参与人都得到相等的收益，这是因为，我们是通过使对方选择两个策略所给出的收益相等，来确定这些概率值的。

因此，当行参与人选择 $p=0.7$ 时，列参与人在扑向左方和扑向右方之间无差异，或者，基于此，他可以选择按任意的概率 q 扑向左方。特别地，按概率 0.6 扑向左方是列参与人最乐意做的事情。

类似地，如果列参与人按概率 0.6 扑向左方，那么，行参与人在踢向左方和踢向右方之间无差异，或者，他可以选择这两个策略的任何混合策略。特别地，他愿意按 0.7 的概率



图中显示的是行参与人和列参与人的最优反应曲线,它们分别是 p 和 q 的函数,其中,p 是行参与人踢向左方的概率,q 是列参与人扑向左方的概率。

图 30.4 最优反应曲线

踢向左方。因此,这些选择是一个纳什均衡:给定对方的选择,每一个参与人的选择都是最优的。

在均衡状态,行参与人在 62%的时间内能够得分,而在 38%的时间内不能得分。如果对方作出最优的反应,那么,这就是行参与人的最优选择。

如果列参与人没有作出最优的反应,情况又会如何呢?行参与人能够做得更好吗?为了回答这个问题,我们可以使用本章开始所介绍的最优反应曲线。我们已经看到,当 p 小于 0.7 时,列参与人将扑向左方;当 p 大于 0.7 时,他将扑向右方。类似地,当 q 小于 0.6 时,行参与人将踢向左方;当 q 大于 0.6 时,他将踢向右方。

图 30.4 显示了这两条最优反应曲线。注意到,它们相交于 $p=0.7$ 且 $q=0.6$ 的点。有关最优反应曲线的一件令人愉快的事情是,对于对方采取的任何选择,不论是最优的还是非最优的,最优反应曲线告诉每一个参与人他应该采取的行动。作为对方一个最优选择的唯一最优反应出现在这两条曲线相交的地方——纳什均衡。

30.5 共存博弈

前面,我们将混合策略解释为参与人的随机选择。在罚点球博弈中,如果行参与人的策略是按概率 0.7 踢向左方,按概率 0.3 踢向右方,那么,我们认为行参与人将"混合他的策略",即他会在 70%的时间内踢向左方,在 30%的时间内踢向右方。

但是,这里还存在另外一种解释。假定一组踢球者和防守者随机地组合在一起,并且,有 7 成的踢球者总是踢向球门的左方,有 3 成的踢球者总是踢向球门的右方。于是,从防守者的角度看,这类似于面临单独一个按这些概率随机选择方向的踢球者。

对于罚点球博弈,这种表述并不是非常有说服力,但对于动物的行为,它却是一个相当合理的描述。这里的观点是,各种各样的行为都是基于遗传指令的,并且,进化所选择的种群混合比例在进化动力方面保持稳定。最近,生物学家开始把博弈论视作研究动物行为的一种不可或缺的工具。

有关动物互动的一个最著名的例子是鹰-鸽博弈。它并不是指老鹰和鸽子之间的博弈(这样的结果是完全可以预测的),而是指涉及显示两种行为的单一物种的博弈。

考虑豺狗的情形。当两只豺狗同时遇到一块食物时,它们必须决定是争斗还是分享食物。争斗是鹰派的策略:一方将获胜,另一方将落败。分享食物则是鸽派的策略:在对方也是鸽派时,这个策略将发挥很好的功效;但当对方是鹰派时,分享食物的提议则会遭到拒绝,并且,鸽派参与人将一无所获。

表 30.8 显示了这个博弈可能的收益组合。

表 30.8 鹰-鸽博弈

		列	
		鹰派	鸽派
行	鹰派	−2, −2	4, 0
	鸽派	0, 4	2, 2

如果两只豺狗都采取鸽派策略,它们最终将获得(2, 2)。如果一方采取鹰派策略,另一方采取鸽派策略,那么,鹰派豺狗将获得全部的收益。但如果双方都采取鹰派策略,两只豺狗最终都会落得严重受伤的下场。

很显然,双方都采取鹰派策略不可能是一个均衡,这是因为,如果一方转向鸽派策略,它最终获得的收益就不是−2 而是 0;如果两只豺狗都采取鸽派策略,那么,某一方转向鹰派策略就是值得的。因此,在均衡状态,一定存在鹰派和鸽派这两种类型的某个混合比例。问题是,我们可以预期哪种类型的混合比例呢?

假定鹰派的比例是 p。于是,一个鹰派遇见另一个鹰派的概率是 p,而遇见一个鸽派的概率是 $1-p$。基于此,鹰派的期望收益等于

$$H=-2p+4(1-p)$$

鸽派的期望收益等于

$$D=2(1-p)$$

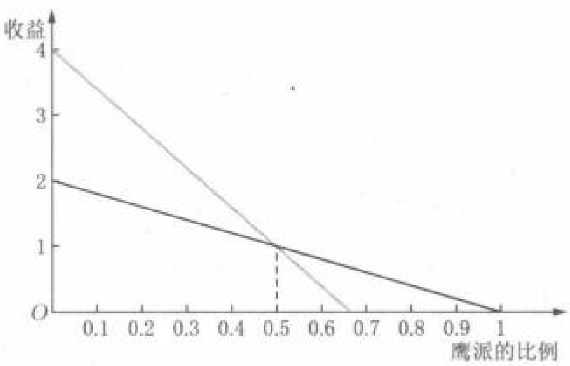
假定具有较高收益的类型的繁殖速度更快一些,并且,它会将其采取鹰派策略或鸽派策略的倾向遗传给后代。那么,我们就可以预期,如果 $H>D$,种群中鹰派的比例就将上升;如果 $H<D$,鸽派的数量就将上升。

种群处于均衡状态的唯一途径是这两种类型的收益相等。这要求下面的等式成立,即

$$H=-2p+4(1-p)=2(1-p)=D$$

求解上式,我们得到 $p=1/2$。

我们已经发现,鹰派和鸽派之间 50 - 50 的混合比例是一个均衡。它在某种程度上是一个稳定的均衡吗?鹰派和鸽派的收益都是种群中鹰派的比例 p 的函数,我们将这两个函数绘制在图 30.5 中。注意,如果 $p>1/2$,鹰派的收益小于鸽派的收益,所以,我们可以预期,鸽派的繁殖速度较快,从而使我们回到 50 - 50 比例的均衡状态。类似地,当 $p<1/2$ 时,鹰派的收益大于鸽派的收益,使得鹰派的繁殖速度更快一些,最终还是会返回均衡。



鹰派的收益表示为浅色线段,而深色线段则代表的是鸽派的收益。当 $p>1/2$ 时,鹰派的收益小于鸽派的收益,反之亦然,这显示了均衡是稳定的。

图 30.5 鹰-鸽博弈的收益

上述分析表明,$p=1/2$ 不仅是一个

均衡,它还是一个在进化动力下稳定的均衡。这种分析引出了一个称作进化稳定策略(evolutionarily stable strategy, ESS)的概念。[①]特别地,可以证明一个 ESS 就是一个纳什均衡,即使它源于完全不同的考虑。

纳什均衡概念旨在处理精明的、理性的个人之间以下的行为问题,即他们都在努力设计一种与对方可能选择的最优策略相适应的策略。ESS 则是用来模型化进化动力下的动物行为的,这里,具有较高适应性收益的策略的复制速度较快一些。但是,ESS 均衡也是纳什均衡的事实,为博弈论中的这个特殊概念为何如此引人注目提供了另一个注解。

30.6 承诺博弈

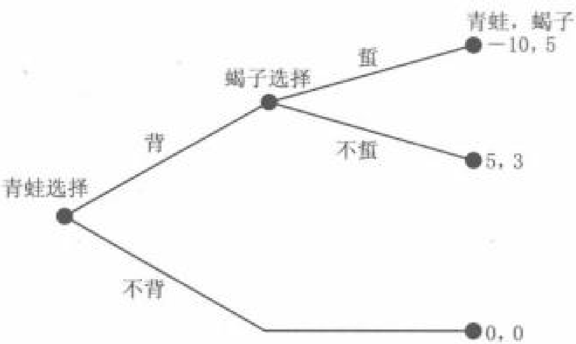
前文有关合作博弈和竞争博弈的例子关注的都是同时行动的博弈。在这类博弈中,每个参与人必须在不知道对方所作选择(不论他正在选择还是已经选择了)的情况下,作出自己的选择。事实上,在合作博弈和竞争博弈中,如果一方知道另一方的选择,博弈就会变得毫无价值。

在本节,我们将注意力转向序贯行动的博弈。这类博弈的一个重要的策略问题是承诺。为了看清楚承诺是如何起作用的,我们回顾本章前面讨论的懦夫博弈。在懦夫博弈中,如果一个参与人能够强迫自己直线驾驶,那么,另一个参与人的最优选择就是转向。在保证博弈中,如果其中的一方先采取行动,结果就会变得对双方更有利。

注意,对于另一个参与人而言,这个被承诺的选择必须同时是不可撤销的和可观察的。不可撤销性是承诺的部分内容,而如果要说服另一方改变行为,那么,可观察性就是关键的要素。

青蛙和蝎子

我们从青蛙和蝎子的寓言故事开始分析。青蛙和蝎子站在一条河岸上,商量一个过河的办法。蝎子说:"我想到了!我爬到你的背上,你背我过河。"青蛙说:"如果你蜇我怎么办?"蝎子说:"我怎么会这样做呢?这样做只会使我们都淹死。"



如果青蛙选择背蝎子过河,蝎子将选择蜇青蛙,双方最终都会被淹死。

图 30.6 青蛙和蝎子

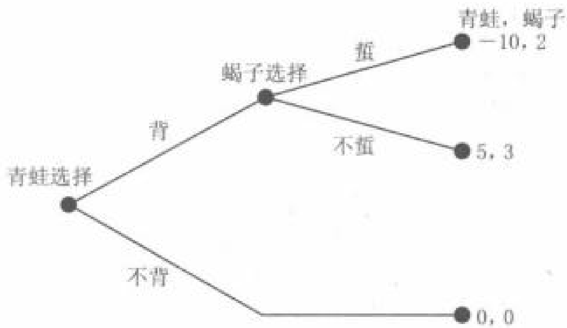
青蛙觉得蝎子的话有道理,所以,它让蝎子爬上了自己的后背,开始过河。游到中途,在河水最深的地方,蝎子蜇了青蛙一下。青蛙疼得叫了起来:"你为什么要蜇我?现在,我们都要死掉了。""唉,"蝎子在不断沉入河底时应道,"这就是我的本性呀。"

我们从博弈论的角度来考察青蛙和蝎子的问题。图 30.6 显示了一个序贯博弈,这个博弈的收益与上述寓言故事一致。我们从博弈树的底部开始分析。如

① 参看约翰·梅纳德·史密斯(John Maynard Smith):《进化和博弈理论》,剑桥大学出版社 1982 年版。

果青蛙拒绝蝎子的建议，双方都一无所获。考虑蝎子没蜇青蛙的博弈分枝，我们看到，如果青蛙背蝎子过河，它会因做这件好事而得到收益 5，蝎子也会因渡过河而得到收益 3。在青蛙被蜇的博弈分枝中，青蛙的收益为－10，蝎子的收益为 5，蝎子的这部分收益代表它因实现自然本能而获得的满足。

从博弈的最后一个阶段开始分析是最佳的：蝎子面临蜇还是不蜇的选择。如果蜇，蝎子将获得较高的收益，因为“蜇是它的本性”。因此，青蛙的理性选择应该是拒绝背蝎子过河。不幸的是，青蛙并不清楚蝎子的收益；很显然，它认为蝎子的收益是图 30.7 所示的形式。唉，青蛙犯的这个错误是致命的！



依据这些收益，如果青蛙选择背蝎子过河，蝎子是不会蜇青蛙的，所以，双方能够安全地渡过河。

图 30.7 青蛙和蝎子

一只聪明的青蛙能够合计出某种办法，使得蝎子作出不蜇的承诺。例如，它可以缚住蝎子的尾巴，或者，它可以雇用一名“职业杀手”，在出现意外时对蝎子家族实施报复。不论采取何种策略，青蛙的关键问题是要使得蝎子蜇的成本更高而不蜇时获得的奖励更多，从而改变蝎子的收益。

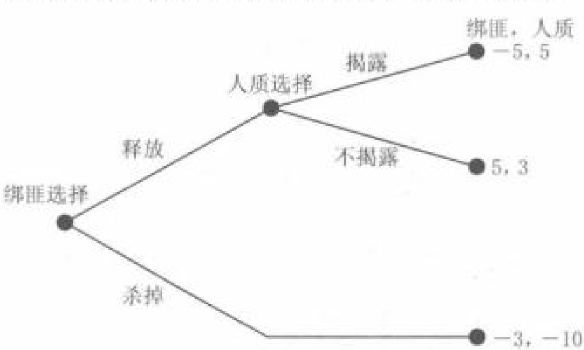
善意的绑匪

在世界上的某些地区，绑架并勒索赎金是一笔很大的交易。在哥伦比亚，据预测每年大约发生 2 000 起绑架勒索案。在前苏联，绑架勒索案从 1992 年的 5 起上升到 1999 年的 105 起。许多受害人都是西方的商人。

有些国家，如意大利，颁布了禁止支付赎金的法律。它们的理由是，如果受害人的家庭或者雇主能够承诺不支付赎金，那么，绑架者最初就不会有动机绑架受害人。

当然，问题是一旦绑架已经发生，受害人的家庭将偏好于向绑架者支付赎金，即使这样做是违法的。因此，惩罚向绑架者支付赎金的行为并不是一个有效的承诺工具。

假定几个绑匪绑架了一名人质，但接下来他们发现他们得不到任何赎金。他们应该释放人质吗？当然，人质会发誓不揭露绑匪的身份。但是，人质会信守这个承诺吗？一旦人质获得释放，他就不再有遵守承诺的激励——他将努力惩罚这些绑匪。因此，即使绑匪想释放人质，但因担心暴露身份，他们也不会这样做。



绑架者愿意释放人质，但如果他这样做，人质将揭露他的身份。

图 30.8 绑架博弈

图 30.8 显示了这个博弈可能的收益。绑匪会因杀掉人质而感到沮丧，从而得到收益－3。当然，人质的境况将更加糟糕，他此时的收益为－10。如果人质获得释放，并且对绑匪的身份守口如瓶，那么，人质获得收益 3，绑匪获得收益 5。但是，如果人质揭露了绑匪的身份，则他的收益为 5，绑匪的收益为－5。

现在轮到人质面临一个承诺问题:他如何能够使绑匪相信他不会违背自己的诺言而揭露绑匪的身份呢?

人质需要合计出一个办法,以改变博弈的收益。具体地说,他需要找到一个方法,使得如果他揭露绑匪的身份,他就要承担某种成本。

托马斯·谢林(Thomas Schelling)是马里兰大学的一位经济学家,他曾经对动态博弈的策略分析作过广泛的研究,他建议,人质可以让绑匪对他们的不雅行为拍照,并保留这些照片。这有效地改变了人质将来揭露绑匪身份时的收益,这是因为,绑匪此时可以曝光这些照片。

这种策略就是所谓的"交换人质"。在中世纪,当两位国王想保证某个条约的实施时,他们就会把自己的家人当作人质交换。如果任何一方违背条约,他的人质就会被杀害。任何一方都不想牺牲自己的家人,所以,每一位国王都具有恪守条约的激励。

在绑架的情形中,令人难堪的照片如果曝光就会增加人质所承担的成本,从而保证他遵守协议,不揭露绑架者的身份。

当力量成为弱势时

我们的下一个例子来自动物世界的心理学。我们发现,猪群会迅速地建立起一种支配-从属关系,在这种关系下,支配猪倾向于支配从属猪。

一些心理学家考察过这样一种情形:一个猪圈里有两只猪,一只是支配猪,另一只是从属猪。①猪圈的一头装有一个控制杆,通过它可以将一部分食物释放到安装在猪圈另一头的食槽里。这些心理学家感兴趣的是,哪只猪将会按控制杆,哪只猪将会吃到食物?

令人感到有些奇怪的是,实验结果表明,支配猪按控制杆,从属猪等待进食。从属猪会吃掉大部分食物,而支配猪则以最大可能的速度奔向食槽,最终,它只能吃到一小部分食物。表 30.9 显示了阐明这个问题的一个博弈。

表 30.9 智猪博弈

		支配猪	
		不按控制杆	按控制杆
从属猪	不按控制杆	0, 0	4, 1
	按控制杆	0, 5	2, 3

从属猪会比较(0, 4)和(0, 2)这两组收益,并足够理性地得到这样的结论,即不按控制杆优于按控制杆。给定从属猪不按控制杆,支配猪别无选择,只有按控制杆。

如果支配猪不吃掉全部食物,并奖励从属猪按控制杆的行为,它就能实现一个较好的结果。问题是两只猪之间没有缔结合约,支配猪还是忍不住作一个贪婪者。

如同善意绑匪的例子,支配猪也要面临一个承诺问题。它只须作出不吃掉全部食物

① 这个例子最早出自鲍德温和米斯(Baldwin and Meese)的《基于操作性条件反射的猪的社会行为》(《动物行为》,1979 年)。我这里引用的是约翰·梅纳德·史密斯(John Maynard Smith)在《进化和博弈理论》(剑桥大学出版社,1982 年)一书中的描述。

的承诺，它的境况就会得到改善。

储蓄和社会保障

承诺问题不仅仅局限于动物的世界，它在经济政策中也有所体现。

退休储蓄是一个有趣而适宜的例子。每个人口头上都称赞储蓄是一个不错的想法。不幸的是，实际中很少有人会这样做。不愿意储蓄的部分原因是，他们意识到社会不会让他们挨饿，所以，他们将来极有可能被救助。

为了明确地描述两代人之间的博弈，我们考虑可供老年人选择的两个策略：储蓄或者挥霍。类似地，年轻人也可以选择两个策略：赡养老人或者为自己退休进行储蓄。表30.10显示了这个博弈的可能的收益矩阵。

表 30.10 两代人之间有关储蓄的冲突

		年轻人	
		赡养	不赡养
老年人	储蓄	3，−1	1，1
	挥霍	2，−1	−2，−2

如果老年人储蓄，并且年轻人赡养老人，那么，老年人的收益为3，年轻人的收益为−1。如果老年人挥霍，年轻人赡养老人，最终，老年人获得收益2，年轻人得到收益−1。

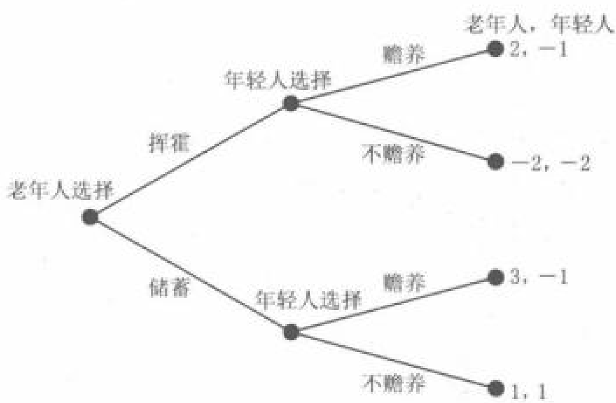
如果年轻人不赡养自己的老人，并且老年人储蓄，那么，老年人的收益为1，年轻人的收益为1。最后，如果老年人挥霍，年轻人不赡养老人，那么，最终老年人因挨饿而获得收益−2，年轻人因必须探视照顾而得到收益−2。

不难发现，这个博弈存在两个纳什均衡。如果老年人储蓄，那么，年轻人的最优选择是不赡养老人。但是，如果老年人挥霍，那么，年轻人赡养老人就是最优的。当然，给定年轻人赡养老人，老年人的最优选择是挥霍。

但是，上述分析忽略了博弈的时间结构：老年人的（较少）优势之一是先采取行动。如果我们画出博弈树，博弈的收益如图30.9所示。

如果老年人储蓄，年轻人将会不赡养他们，所以，老年人的最终收益为1。如果老年人挥霍，他们知道年轻人不能忍受眼看着他们挨饿，所以，他们的最终收益为3。因此，老年人的明智选择是挥霍，因为他们清楚，将来他们一定会得到救助。

当然，大多数发达国家现在都拥有一套类似于美国的社会保障计划，这种计划强迫每一代人都进行退休



知道年轻人将赡养他们，老年人会选择挥霍。子博弈完美均衡是（赡养，挥霍）。

图 30.9 扩展形式的储蓄博弈

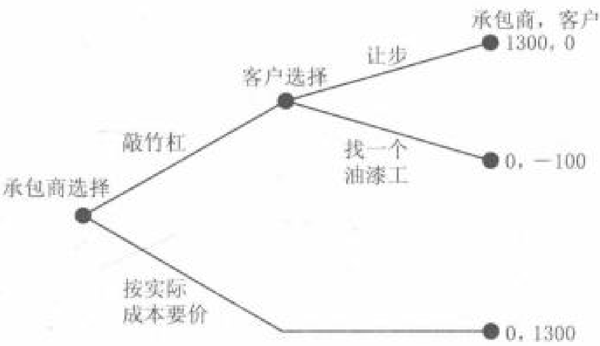
储蓄。

敲竹杠

我们考虑以下的策略互动。你雇用一家承包商建造一座仓库。当建造计划被批准并且建筑几乎完工时,你注意到建筑的颜色非常糟糕,所以,你要求承包商更换颜色,这种颜色修改只涉及较少的费用。但是,承包商却说:"要更换颜色,请支付 1 500 美元。"

你意识到,延迟工期直至找到一个油漆工的成本至少等于承包商的开价,并且,你特别喜欢新颜色,所以,抱怨几句后,你还是会支付这笔费用。恭喜你,你被敲竹杠了。

当然,承包商并不是这类博弈中唯一找麻烦的一方,客户也可以延迟对承包商的支付"敲竹杠",从而给承包商带来诸多困扰。



承包商为改颜色索要一个较高的价格,因为客户除了接受没有其他选择。

图 30.10　敲竹杠问题

图 30.10 显示了这个敲竹杠问题的博弈树。我们假定,客户为粉刷新油漆需要支付 1 500 美元,粉刷油漆的实际成本是 200 美元。我们从博弈树的顶端开始分析,如果承包商要价 1 500 美元,他将实现净利润1 300美元,客户的净效用为零。

如果客户寻找另一个油漆工,那么,他要支付油漆工 200 美元,还要承担时间成本 1 400 美元。仓库粉刷成他喜欢的颜色,这对他的价值是1 500美元;他还需要支付直接成本和误工成本,总共是 1 600 美元,最终,他获得的净收益为−100 美元。

如果承包商向客户实际要价 200 美元,他能够实现盈亏平衡;客户支付 200 美元的费用,并同时获得价值 1 500 美元的东西,那么,他的净收益就为 1 300 美元。

我们发现,承包商的最优选择是敲竹杠,客户的最优选择是让步。但是,一位明智的客户将意识到,类似更换颜色的要求在任何工程中都会出现。有鉴于此,客户就不愿意雇用背负敲竹杠名声的承包商,这种声誉当然对承包商是不利的。

厂商如何解决敲竹杠问题呢?基本的答案是缔结合同。正式地,承包商会提供一份合同,详细规定哪种类型的修改是适当的要求,以及它们的成本是如何确定的。有时,这些合同还包括仲裁条款或其他纠纷解决程序。总之,大量的时间、精力和金钱都花费在起草合同上,以避免敲竹杠问题。

但是,合同并不是唯一的解决方案。解决这类问题的另一种方式是承诺。例如,承包商也许会缴纳保证金,以保证及时地完成工程。而且,合同中通常有某些客观条款规定了建筑完工需要满足的条件。

另一个重要的因素是声誉。很明显,一家总是在竭力敲诈客户的承包商具有不良的声誉。这个客户不会再雇用他,他当然也一定不会获得好的评价。我们可以在重复博弈的框架下考察这种声誉效应,在重复博弈下,承包商今天敲竹杠,将来他就要付出代价。

30.7 讨价还价

经典的讨价还价问题是分配货币。两个参与人共同拥有 1 美元，现在，他们想分配这 1 美元，他们应该怎么做?

这个问题没有答案，因为这里只有很少量的信息，还不足以创建一个合理的模型。将讨价还价问题模型化所面临的一个挑战，是要找到问题的其他一些角度，从这些角度出发，参与人可以进行谈判。

一种解决方案是纳什讨价还价模型，它采用一种公理化的方式，规定一个合理的讨价还价解所应该具备的某些性质，并且，它还证明只有一种结果满足这些公理。

最终的结果依赖于参与人厌恶风险的程度，以及不存在讨价还价时将会发生的情况。不过，对这个模型的完整描述已经超出了本书的范围。

一种替代的方案是鲁宾斯坦讨价还价模型，它考虑的是一系列选择，并求解出子博弈完美均衡。幸运的是，我们可以利用简单的例子，来阐明这个模型的基本内容。

两个参与人，艾丽丝和鲍勃，分配 1 美元。他们同意最多用 3 天的时间协商分配问题。第 1 天，艾丽丝给出一个报价；第 2 天，鲍勃可以接受也可以拒绝这个报价，如果他拒绝，他要提出一个新报价；第 3 天，艾丽丝提出最终的报价。如果他们不能在 3 天之内达成协议，那么，双方都将一无所获。

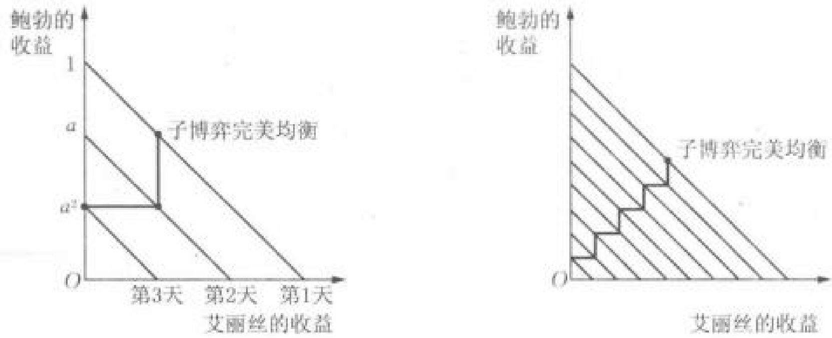
我们假定艾丽丝和鲍勃具有不同的耐性：艾丽丝对将来收益的贴现因子是 α/每天；鲍勃对将来收益的贴现因子是 β/每天。最后，我们假定，如果一方在两个报价之间无差异，他将选择对方最偏好的报价。这里的思想是，对方会提供某个任意小的数量使得这个参与人严格地偏好某个选择，并且，依据上述假定，我们就可以将这个“任意小的数量”近似为零。可以证明，这种讨价还价博弈存在唯一的子博弈完美均衡。

我们从博弈的末端即博弈结束的前一天开始分析。这一天，艾丽丝可以向鲍勃提供一个“要么接受要么拒绝”的报价。很明显，艾丽丝此时的最优选择是向鲍勃提供一个他可以接受的最小数量，依据假定，这个数量等于零。因此，如果博弈实际上只持续 3 天，艾丽丝将得到 1 美元，鲍勃将一无所获(即，一个任意小的数量)。

现在，我们转向博弈的上一个阶段，此时，鲍勃要给出一个报价。这一天，鲍勃应该意识到，在博弈的下一个阶段，艾丽丝拒绝鲍勃的报价就能保证自己获得 1 美元。下一个时期的 1 美元在这个时期对艾丽丝价值 α 美元，所以，任何低于 α 美元的报价一定会遭到拒绝。与下一个时期的 0 美元相比，鲍勃当然偏好现在的 $1-\alpha$ 美元，所以，他应该理性地向艾丽丝提供 α 美元，艾丽丝自然也会接受这个报价。因此，如果博弈在第二个阶段结束，艾丽丝将获得 α 美元，鲍勃将获得 $1-\alpha$ 美元。

现在，我们再回到博弈的第一个阶段。这一天，艾丽丝要提出报价，她意识到，鲍勃只要等到第二天，就可以获得 $1-\alpha$ 美元。因此，为了避免拖延，艾丽丝必须向鲍勃提供一个至少等于这个数额的现值的报价，所以，她向鲍勃提供 $\beta(1-\alpha)$ 美元。鲍勃发现，这个数额恰好是他可以接受的，博弈到此结束。最终的结果是，博弈在第一个阶段结束，艾丽丝获得 $1-\beta(1-\alpha)$ 美元，鲍勃得到 $\beta(1-\alpha)$ 美元。

图 30.11 中的第一个图形演示的是 $\alpha=\beta<1$ 的情形。最外面的一条斜线显示的是第 1 天的可能收益模式,即所有的收益满足 $x_A+x_B=1$。移向原点的第二条斜线显示的是,博弈在第 2 天结束时的所有收益的现值:$x_A+x_B=\alpha$。距离原点最近的斜线表示的是,博弈在第 3 天结束时的所有收益的现值;这条线的方程是 $x_A+x_B=\alpha^2$。图中的直角形路径描述的是每个时期最低可以接受的分配额,由它们可以得到最终的子博弈完美均衡。图 30.11 中的第二个图形显示了相同的进程在包含更多个谈判阶段时的形式。



图中的粗线将子博弈的均衡结果连接在一起。粗线最外端的点表示的是子博弈完美均衡。

图 30.11　一个讨价还价博弈

很自然,我们会考虑无穷的情形,也就是说,无限博弈下的情况是怎样的呢?在这种情况下,可以证明,子博弈完美均衡的分配额是

$$\text{艾丽丝获得的收益}=\frac{1-\beta}{1-\alpha\beta}$$

$$\text{鲍勃获得的收益}=\frac{\beta(1-\alpha)}{1-\alpha\beta}$$

注意,如果 $\alpha=1$ 且 $\beta<1$,那么,艾丽丝将得到全部的收益。

最后通牒博弈

鲁宾斯坦讨价还价模型是如此优美,以至于经济学家们急着在实验室对它作出检验。但是,他们发现,优美并不意味着准确。朴素学科(即非经济学专业)并不擅长于预测未来。

此外,其他一些因素也可能引起麻烦。为了看清楚这一点,我们考虑上述讨价还价模型的一个只包括一个阶段的变形。艾丽丝和鲍勃分配 1 美元,艾丽丝给出一个报价,如果鲍勃接受,博弈就此结束。问题是,艾丽丝的报价应该是多少?

依据理论分析,她应该为自己保留 99 美分,而只分给鲍勃 1 美分。鲍勃意识到 1 美分总比一无所获要好,所以,他接受了这个报价,艾丽丝高兴地回到家,并庆幸自己读过经济学。

但是,事实并非如此。一种更有可能的结果是,鲍勃非常厌恶这微不足道的 1 美分,他拒绝艾丽丝的报价。最终,艾丽丝一无所获。艾丽丝在考虑到这种可能性后,会给出一个对鲍勃更有吸引力的报价。在实际的实验中,美国大学生给出的平均报价是 45 美分,并且,在大多数情况下,这个报价都能够被接受。

报价的一方是理性的,这是因为,给定观察到的报价被拒绝的频率,45 美分相当接近

使期望收益最大化的水平。实际上，与理论预测相违背的是报价受盘人的行为，因为他们会拒绝小额的报价，即使这样做会使他们的境况恶化。

对此存在许多解释。一种观点认为，一个太少的报价违背了行为的社会准则。事实上，经济学家已经在最后通牒博弈中发现了非常重要的跨文化差异。另外一种前后一致的观点是，作为对小额报价的报复，受盘人可以从伤害报价人的行为中获得某些效用。毕竟，与失去一小笔金钱相比，回击对方所获得的满足感还是相当有诱惑力的。

小　结

1. 一个参与人的最优反应曲线给出了他的最优选择，这个最优选择是其他参与人可能作出的选择的函数。
2. 在双人博弈中，纳什均衡是指这样一组策略组合，两个策略分别对应于不同的参与人，并且，每个策略都是一方相对于另一方选择的最优反应。
3. 混合策略纳什均衡涉及对几个纯策略的随机选择。
4. 普通的合作博弈包括：性别战，博弈双方想做同一件事情而不是不同的事情；囚徒困境，最终的占优策略使双方都受到伤害；保证博弈，只要参与人相信对方会采取合作的策略，他们都会选择合作；懦夫博弈，博弈双方会避免采取相同的策略。
5. 在双人零和博弈中，一方的收益是另一方收益的相反数。
6. 进化博弈关注的是人口繁殖下的稳定结果。
7. 在序贯博弈中，参与人交替采取行动。因此，每一个参与人都必须推测对方对自己的选择所作出的反应。
8. 在许多序贯博弈中，承诺是一个重要的问题。这里，找到一种方法，强制承诺采取某个特定的策略是非常重要的。

复习题

1. 在一个双人博弈纳什均衡中，每一个参与人都在针对什么作出最优的反应？在一个占优策略均衡中，每一个参与人又都在针对什么作出最优的反应？

2. 在有关混合策略的章节中，考虑行参与人和列参与人的最优反应。它们会产生最优反应函数吗？

3. 在一个合作博弈中，如果博弈双方作出相同的选择，那么，结果对于他们都令人满意。这个结论是否正确？

4. 本章正文指出，在均衡状态，行参与人在62%的时间内会得分。这个数值是如何得到的？

5. 承包商说，他打算“降低要价，并在以后的修改要求中寻求补偿”。他的本意是什么？