

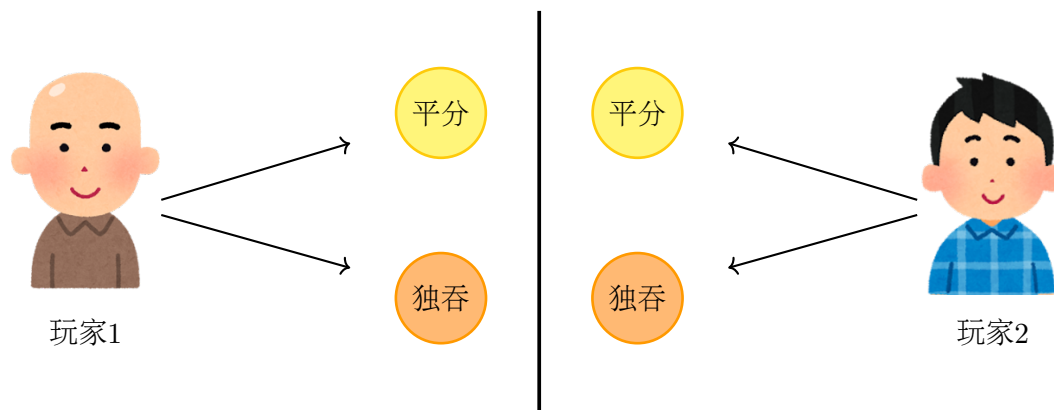
MAT120: 论证 第18讲讲义

数学与游戏

上节课我们学习了数学与社会选择，其中的主要问题是，一个群体应该如何把许多个体偏好整合成一个集体决策。这一讲我们仍然处在“数学与社会”这一单元之中，但现在我们将从投票转向策略。博弈论的核心问题是：在情境规则既定、而且其他人也在进行策略性思考的前提下，我采取什么行动才是最好的？我们将从一个游戏节目例子开始，然后介绍玩家、策略、结果和收益这些基本术语。之后我们会学习几个著名例子，包括囚徒困境、平均值游戏、两性之战以及霍特林法则。最后，我们将定义纳什均衡，这是初等博弈论中最主要的稳定性概念。

1 引言

Golden Balls 是一档英国电视节目，播出时间为2007 年到2009 年。在节目中，一组参赛者要通过一系列不同的挑战进行竞争，而这些挑战会让组内成员被逐步淘汰。过程中，参赛者们会不断累积奖金，通常相当于数万美元。到了比赛最后一轮，几乎所有参赛者都已经被淘汰，只剩下两个人。此时他们会面对最后一个游戏：两位选手面前各有两个 *golden ball*。这些球里面分别写着“split” 和“steal” 两个词。玩家可以看见自己那对golden ball 里面写的是什么，但看不到对方球里的内容：



接下来，这两位玩家必须争夺全部奖金。节目给出的规则如下：

1. 如果两位玩家都选择平分，那么他们就平均分掉这笔钱；
2. 如果玩家1 选择平分，而玩家2 选择独吞，那么玩家2 拿走全部奖金；
3. 如果玩家1 选择独吞，而玩家2 选择平分，那么玩家1 拿走全部奖金；
4. 如果两位玩家都选择独吞，那么他们两人都一无所获。

有趣的是，玩家随后还被允许彼此交谈，并尝试制定一种策略。通常会发生两种情况之一：

- 两位玩家会达成一致，都选择“平分”球，因为这符合他们共同的最佳利益；
- 两位玩家会达成一致，都选择“平分”，但随后其中一人或两人都会背弃协议，实际上选择“独吞”球。

然而，在其中一期节目中，有一位参赛者设计出了一个非常出色的策略：玩家2 坚称，无论如何自己都会选择独吞 球。然后他又声称，在节目结束后，他会把实际奖金和另一位参赛者平分。经过一番来回争论之后，玩家1 选择了平分球，而令人惊讶的是，玩家2 最后也选择了平分。

为了看出这个策略的高明之处，我们来把各种可能性拆开分析。为方便起见，假设奖金池大约等于\$20000。按照规则，这个游戏有四种可能结果：

		玩家2	
		平分	独吞
玩家1	平分	玩家1 得到\$10000 玩家2 得到\$10000	玩家1 得到\$0 玩家2 得到\$20000
	独吞	玩家1 得到\$20000 玩家2 得到\$0	玩家1 得到\$0 玩家2 得到\$0

在这种情形下，玩家2 非常准确地理解了游戏规则。他知道自己不能信任玩家1 一定会选择“平分”球，因此玩家2 有可能最后什么也得不到。为了把这种被背叛的风险降到最低，玩家2 决定坚持说自己会选择独吞球，这实际上把一半的可能结果排除了：

		玩家2	
		平分	独吞
玩家1	平分	玩家1 得到\$10000 玩家2 得到\$10000	玩家1 得到\$0 玩家2 得到\$20000
	独吞	玩家1 得到\$20000 玩家2 得到\$0	玩家1 得到\$0 玩家2 得到\$0

→

		玩家2	
		平分	独吞
玩家1	平分	玩家1 得到\$0 玩家2 得到\$20000	
	独吞	玩家1 得到\$0 玩家2 得到\$0	

通过向玩家1 声称自己会在节目结束后分给他一部分钱，玩家2 实际上是在对玩家1 说：

“如果你选择‘独吞’，那么你得到任何钱的概率是0%。但是，如果你选择‘平分’，那么你在节目结束后得到一半奖金的概率就不是零。”

通过采用这种策略，玩家2 试图把玩家1 选择“独吞”球的可能性降到最低。这个策略奏效了，因此，玩家2 在最后确信钱不会被抢走之后，决定真正选择平分。

乍一看，这似乎是一个关于信任、道德或人类心理的问题。从某种意义上说，它确实同时涉及这些方面。但从另一种意义上看，它又要精确得多：这是一个数学问题，研究当结果依赖于双方决策时，理性的玩家应该如何行动。这正是博弈论的出发点。

博弈论

博弈论是对**策略**进行数学研究的学科。它研究的是：当一个理性决策者的结果不仅依赖于自己的选择，也依赖于其他决策者的选择时，他应该如何行动。

对于博弈论的核心问题，一个最初的表述可以是：

在游戏规则既定的情况下，我采取什么行动才是最好的？

但这还不完整。现实中，策略还应该受其他玩家打算怎么做的影响。因此，更完整的问题是：

在游戏规则既定、并且其他玩家也在形成各自策略的情况下，我采取什么行动才是最好的？

2 博弈的语言

在博弈论中，博弈 并不仅仅限于棋盘游戏或电子游戏。它指的是任何一种情形，在这种情形中，有若干决策者，而他们的结果彼此相互依赖。经济学、政治谈判、生物竞争、价格战、投票和军事规划中都能找到这样的例子。

什么是博弈？

博弈是对策略性互动的一种数学模型。它告诉我们：

- 谁是玩家，
- 每个玩家可以作出哪些选择，
- 以及在每一种选择组合下，每个玩家会得到什么。

2.1 关键术语

以下四个词构成了初等博弈论的核心词汇。

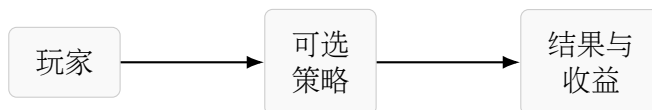
关键术语

- **玩家：** 作出决策的人或群体。
- **策略：** 一种选择，更一般地说，是一种行动计划。
- **结果：** 所有玩家都作出选择之后所发生的事情。
- **收益：** 一个数字，用来描述结果对每个玩家来说有多好。

“收益”这个词不应被理解得太狭隘。在某些情况下，它字面上就是金钱；但很多时候，它只是“一个描述结果有多值得追求的数字”。在某些模型里，收益可能表示利润；在另一些模型里，它可能表示幸福、服刑年数、政治优势、适应度，或任何其他衡量偏好的量。

2.2 把博弈可视化

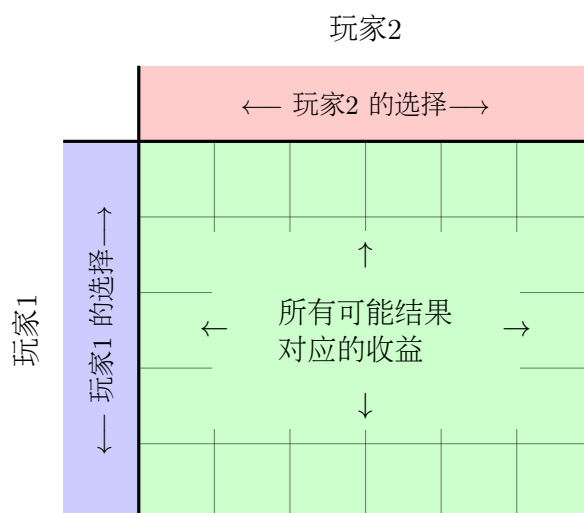
在基本层面上，一个博弈可以被看作一个包含三个阶段的过程：



策略上的困难出现在中间这一阶段。每位玩家都必须在预判其他人可能会做什么、不会做什么的同时，选择自己的策略。

2.3 收益矩阵

对于两个玩家的小型博弈来说，把所有可能结果整理进一个收益矩阵 很方便。它是一个信息表格，记录了博弈中每一种可能结果所对应的收益。从某种意义上说，它包含了这个博弈中所有可能发生的事情。示意图如下：



这里，行对应玩家1 的选择，列对应玩家2 的选择。中间绿色区域是一个网格，记录了每一对选择所产生的收益。

如何读取收益矩阵。 如果某个格子里写着 (a, b) ，那么 a 是玩家1 的收益， b 是玩家2 的收益。第一个数字总是属于行玩家，第二个数字总是属于列玩家。

实际上，我们在引言中已经看到了收益矩阵的一个原型。用不那么漂亮的记号写出来，平分/独吞博弈对应的收益矩阵如下：

		玩家2	
		平分	独吞
玩家1	平分	(10000, 10000)	(0, 20000)
	独吞	(20000, 0)	(0, 0)

2.4 支配策略

有时，一个玩家会有一种策略，不管其他玩家做什么，这种策略都是最好的。这种策略称为支配策略。它是一种支配所有其他备选方案的策略。

支配策略

如果对于某个玩家来说，无论其他玩家作出什么可能的选择，这一策略带给他的收益都至少不低于其他任何可用策略，那么这个策略就称为**支配策略**。如果它总是严格更好，那么我们就称它为**严格支配策略**。

在上面的 *Golden Balls* 形式收益矩阵中，**独吞** 对每一位玩家来说都是支配策略，因为：

- 如果另一位玩家选择平分，那么独吞的收益比平分更大；
- 如果另一位玩家选择独吞，那么独吞与平分给出相同的收益。

因此，从这个形式矩阵的狭义视角来看，“理性”的行动似乎就是独吞。这并不意味着现实中的人类总会这么做。它只是意味着，在这个数学模型内部，独吞支配平分。

3 博弈的例子

现在我们来回顾一些经典的博弈例子。它们是：

1. 囚徒困境，
2. 平均值游戏，
3. 两性之战，
4. 霍特林法则。

正如我们将看到的，每一个博弈都展示了一种不同类型的策略问题。

3.1 例1：囚徒困境

假设你和你的朋友被捕了。警方知道发生了一起犯罪，但他们不知道究竟是谁扮演了主要角色。因此，你们被分开审讯，而你被给出了如下交易：

- 如果你告密，而对方保持沉默，那么你会被释放，而对方会被判10 年监禁；
- 如果你保持沉默，而对方告密，那么你会被判10 年监禁，而对方会被释放；
- 如果你们两人都保持沉默，那么你们都会被判2 年监禁；
- 如果你们两人都告密，那么你们都会被判5 年监禁。

这两名嫌疑人事先不允许沟通。这个博弈的矩阵如下：

		你的朋友	
		不告密	告密
你	不告密	(2, 2)	(10, 0)
	告密	(0, 10)	(5, 5)

这里的数字记录的是服刑年数，因此数字越小越好。

让我们考虑你可选择的两种行动：

1. 如果你的朋友决定不告密，那么你就被限制在收益矩阵的第一列中。在这里，你可以选择不告密，从而坐牢2 年；或者选择告密，从而坐牢0 年。这两种选择中，告密显然更合理。
2. 如果你的朋友决定告密，那么你就被限制在收益矩阵的第二列中。此时，你可以选择不告密，从而坐牢10 年；或者选择告密，只坐5 年。从你的角度看，这种情况下告密也更合理。

在这两种可能情形中，选择告密实际上都对你更有利。换句话说，告密是一个支配策略。

囚徒困境的结论

告密 对两位玩家来说都是一个支配策略。因此，自然的策略预测是两位玩家都会告密，尽管如果他们能够 somehow 协调保持沉默，那么两人都会更好。

这就是“囚徒困境”中的“困境”。根据个人激励，两位玩家都应该互相告密，尽管那里其实存在一个更好的替代结果。

3.2 例2：平均值游戏

假设一个班里的每个学生都选择一个介于0 和100 之间的实数。然后计算全班平均数，而获胜者是那个所选数字最接近该平均数的 $\frac{2}{3}$ 的学生。这里最好的数字应该选什么呢？

乍一看，似乎很多数字都可能是合理的猜测。然而，考虑下面这条推理链：

- 如果每个人都选100，那么平均数的 $\frac{2}{3}$ 也只有大约66.7。所以选择100 显然太大了。事实上，任何大约高于66.6 的数字看起来都不妙，因为这样并不会让你更有机会接近真正的目标值。
- 如果我们已经意识到这一点，我们也就明白班上其他人也可能会想到这一点。因此，我们可以得出结论：不会有人选择大于66.6 的数字。
- 这样一来，我们的问题就缩减成了：大家都在0 到66.6 之间选数的问题。再次注意到，如果每个人都选择最大值66.6，那么平均数就是66.6，而所需的目标数就会是44.4。于是我们得出结论，选择大于44.4 的数没有意义。
- 进一步地，我们会想到，其他所有参与这个游戏的人也可能意识到了这一点，这意味着不会有人选择大于44.4 的数字。

我们可以一遍又一遍重复这一推理，不断把理想猜测往下拉低，直到最后得出结论：最好的猜测其实就是选择数字0。

迭代推理

平均值游戏是**迭代式策略推理** 的一个很好例子。你不只是问“我应该做什么？”你还会问：“既然其他玩家也在问完全同样的问题，那我应该做什么？”

3.3 例3：两性之战

一个经典的协调博弈是两性之战。想象一对已婚夫妇在决定约会之夜要做什么：妻子更想去看芭蕾演出，而丈夫更想去看拳击比赛。不过，两人都宁愿一起去同一个地方，而不是各去各的地方。这个情形的一个典型收益矩阵如下：

		丈夫	
		芭蕾	拳击
妻子	芭蕾	(2, 1)	(0, 0)
	拳击	(0, 0)	(1, 2)

注意，这个收益矩阵中的具体数字本身并不是特别重要，只要偏好顺序满足：

- 妻子更喜欢一起去看芭蕾，其次是一起去看拳击，最差是分开行动；
- 丈夫更喜欢一起去看拳击，其次是一起去看芭蕾，最差是分开行动。

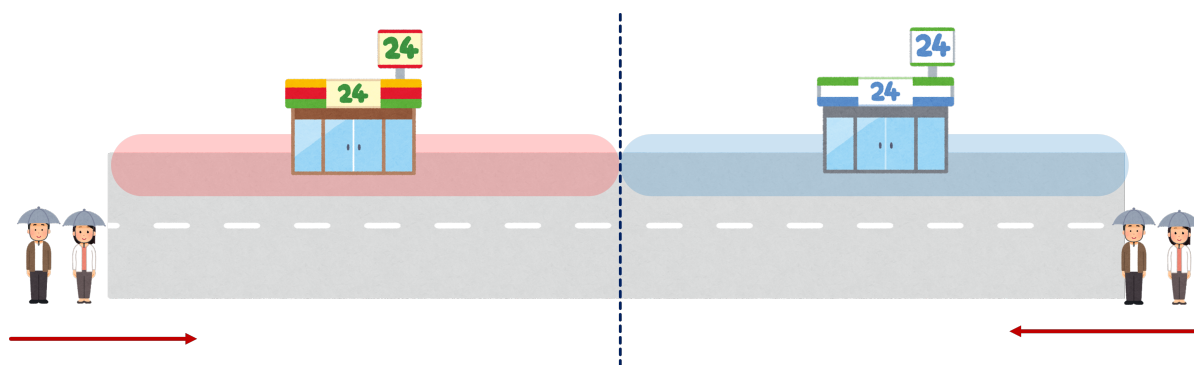
与囚徒困境不同，没有哪一位玩家拥有支配策略。例如，从妻子的角度看：

- 如果预计丈夫会选择芭蕾，那么妻子的最佳应对就是芭蕾；
- 如果预计丈夫会选择拳击，那么妻子的最佳应对就是拳击。

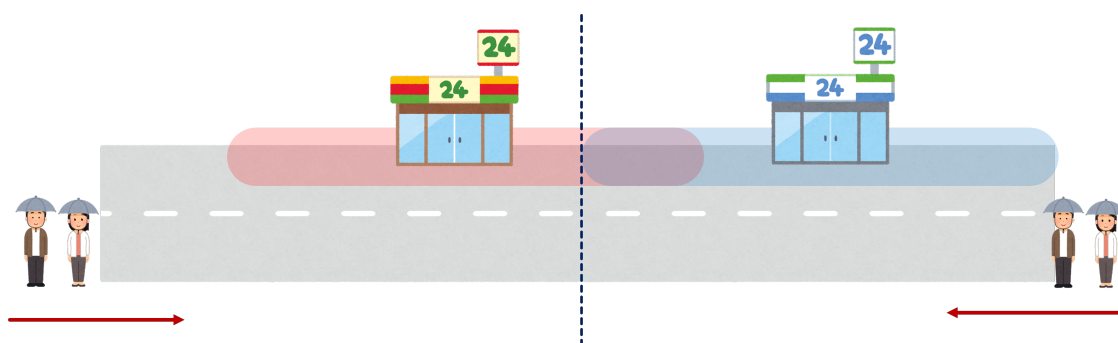
当然，对丈夫来说也是如此。正如我们在这里看到的，主要的困难在于协调，而不是策略支配。

3.4 例4：霍特林法则

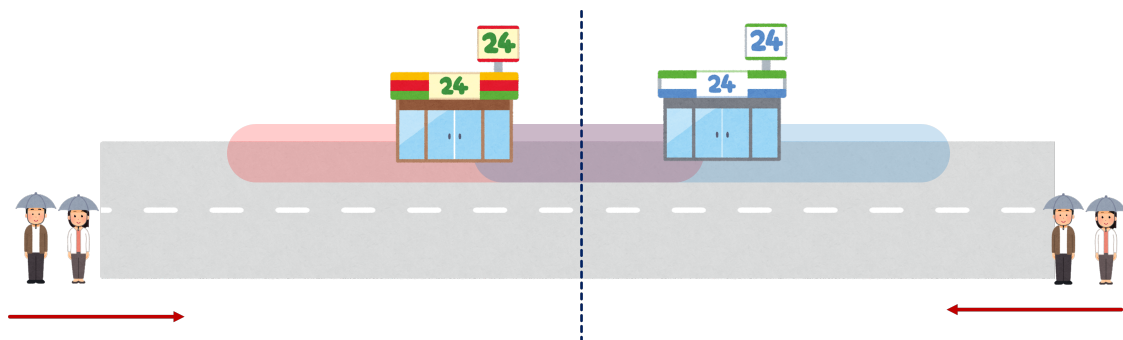
假设两家便利店正在决定要把店开在同一条长街的什么位置。假设这条街只有两个入口：一个在左边，一个在右边。起初，为了公共便利，似乎最理想的情况是一家店靠近左侧，另一家店靠近右侧。这样，从左边进入的顾客会进入第一家便利店，而从右边进入的顾客会进入第二家便利店：



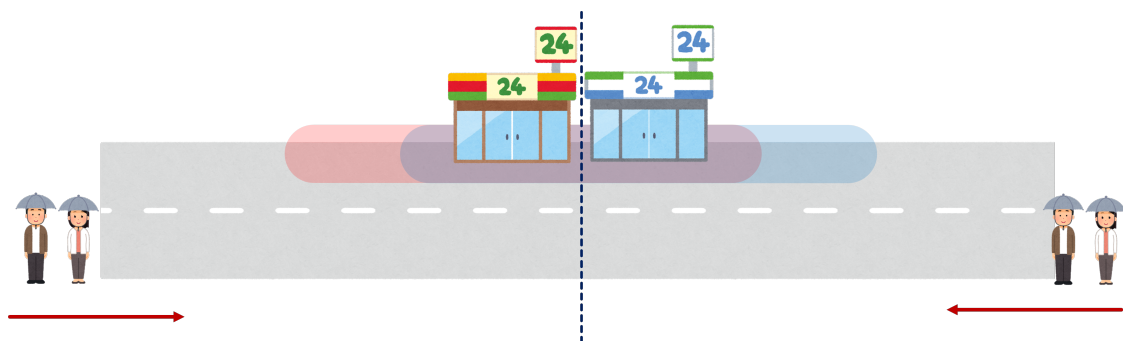
每家店都想尽可能多地吸引顾客。左边那家店意识到，如果自己向另一家店靠近一些，那么就有机会吸引一部分从右边进入街道的人，从而抢走2号店的一部分顾客基础。而且这样做并不会冒着失去从左边进入顾客的风险，因为他们仍然会先看到1号店。



当然，2号店也可能得出同样的结论。无论是先行动，还是出于反制，2号店也可能向1号店靠近，并有可能占走1号店的一部分顾客基础：



最终，通过不断重复这样的推理，两家店都会不可避免地收敛到街道中央。这是街道上唯一“稳定”的位置，因为无论哪一家店再往别处移动，都会把一部分优势让给另一家店：



最后，两家店都大约保证了街道50% 的覆盖率，但代价是它们都聚集在中间附近。因此，这个均衡对商店来说是稳定的，但却以牺牲公共便利为代价。

4 纳什均衡

初等博弈论中的一个核心思想是均衡。在日常语言中，均衡是一个系统的稳定点。在博弈论中，标准的均衡概念以著名博弈论学家John Nash 的名字命名，因为正是他引入了这一均衡概念。

纳什均衡

一个**纳什均衡** 是指这样一种情形：

- 每一位玩家都已经选择了一个策略，
- 每位玩家的策略都是对其他玩家策略的最佳应对，
- 并且任何玩家都无法仅靠自己 改变策略来改善结果。

简单地说，在别人都保持不变的情况下，每个人都已经在做自己所能做的最好选择。

4.1 最佳应对

为了真正理解纳什均衡，我们首先需要“最佳应对”这个概念。假设另一位玩家的选择暂时固定不变。在你的可选策略中，有些会给你比其他策略更大的收益。最佳应对就是，在这一固定假设下，使你自己收益最大的那些策略之一。

因此，在收益矩阵中寻找纳什均衡的通常步骤是：

1. 固定玩家2 的选择，在该列中找出玩家1 的最佳应对；
2. 对每一列都重复这一过程；
3. 然后固定玩家1 的选择，在该行中找出玩家2 的最佳应对；
4. 对每一行都重复这一过程；
5. 任何一个两位玩家同时都在作最佳应对的格子，就是一个纳什均衡。

4.2 一个 3×3 的完整例子

考虑下面这个收益矩阵。我们用星号* 标出玩家1 的最佳应对，用匕首号† 标出玩家2 的最佳应对。

	选择1	选择2	选择3
选择1	3, 7 [†]	3, 2	1, 1
选择2	2, 4	2, 2	*4, 7 [†]
选择3	*6, 2	*5, 4 [†]	3, 3

下面是逐步操作过程：

步骤1： 依次查看矩阵的每一列，用* 标出玩家1 的最佳应对。

- 在第1 列中，第一坐标是3, 2, 6，所以最佳应对是选择3。
- 在第2 列中，第一坐标是3, 2, 5，所以最佳应对仍然是选择3。
- 在第3 列中，第一坐标是1, 4, 3，所以最佳应对是选择2。

步骤2： 逐行给玩家2 的最佳应对加上†。

- 在第1 行中，第二坐标是7, 2, 1，所以最佳应对是选择1。
- 在第2 行中，第二坐标是4, 2, 7，所以最佳应对是选择3。
- 在第3 行中，第二坐标是2, 4, 3，所以最佳应对是选择2。

步骤3： 找到交点。同时带有星号和匕首号的格子是：

(选择2, 选择3) 和 (选择3, 选择2)。

因此，这个博弈有两个 纳什均衡。

重要观点

一个博弈可能：

- 恰好有一个纳什均衡，
- 有多个纳什均衡，
- 完全没有纯策略纳什均衡。

所以目标绝不是先假设均衡一定存在且唯一。目标是仔细分析这个矩阵。

4.3 本讲例子中的纳什均衡

现在我们可以从正确的角度重新看待本讲前面那些例子。

- 在**Golden Balls** 中，形式收益矩阵在(独吞, 独吞) 处有一个纳什均衡。由于独吞对两位玩家都是支配策略，这个均衡是立刻得到的。
- 在**囚徒困境** 中，纳什均衡同样是(告密, 告密)，原因相同。
- 在**平均值游戏** 中，均衡预测是所有人都选择0。
- 在**两性之战** 中，有两个 纯策略纳什均衡：(芭蕾, 芭蕾) 和(拳击, 拳击)。
- 在**霍特林法则** 中，当两家商店都位于街道中部区域时，就会出现均衡。

最重要的教训是：纳什均衡并不一定是道德上最好的结果，也不一定是社会上最好的结果，甚至不一定是最令人愉快的结果。它只是一个稳定的 结果。

5 练习

问题1：基本词汇热身

从下面的词语中填入正确的空格。

玩家，策略，收益，结果，纳什均衡，支配策略

- (a) 一个衡量某个结果对玩家来说有多好的数字，叫作_____。
- (b) 在博弈中的一种选择或行动计划，叫作_____。
- (c) 没有人能够仅靠单独改变自己的选择而改善结果的情形，叫作_____。
- (d) 无论其他玩家做什么都最好的策略，叫作_____。
- (e) 每个人都已经作出选择之后所发生的事情，叫作_____。

问题1 的解答

- (a) 收益
(b) 策略
(c) 纳什均衡
(d) 支配策略
(e) 结果

问题2：读取收益矩阵

考虑下面这个收益矩阵。

		玩家2	
		左	右
玩家1	上	(4, 3)	(1, 5)
	下	(6, 2)	(2, 4)

第一个数字是玩家1 的收益，第二个数字是玩家2 的收益。

- (a) 如果玩家1 选择上，而玩家2 选择右，结果是什么？
- (b) 如果玩家2 选择左，哪一个选择给玩家1 带来更大的收益？
- (c) 如果玩家1 选择下，哪一个选择给玩家2 带来更大的收益？
- (d) 玩家1 是否有一个支配策略？请简要说明。
- (e) 如果存在纳什均衡，请找出来。

问题2 的解答

- (a) 结果是(1,5)。因此玩家1 得到1，玩家2 得到5。
- (b) 如果玩家2 选择左，那么玩家1 比较的是4 和6。所以玩家1 应该选择下。
- (c) 如果玩家1 选择下，那么玩家2 比较的是2 和4。所以玩家2 应该选择右。
- (d) 有。玩家1 的支配策略是下，因为如果玩家2 选择左，那么 $6 > 4$ ；如果玩家2 选择右，那么 $2 > 1$ 。
- (e) 纳什均衡是(下, 右)，因为下是玩家1 对右的最佳应对，而右是玩家2 对下的最佳应对。

问题3：囚徒困境的推理

两名嫌疑人必须各自选择**沉默** 或**告密**。他们的收益如下表所示。

	朋友：沉默	朋友：告密
你：沉默	(2, 2)	(0, 10)
你：告密	(10, 0)	(5, 5)

- (a) 如果你的朋友保持沉默，你应该选择什么？
- (b) 如果你的朋友选择告密，你应该选择什么？
- (c) 你的支配策略是什么？
- (d) 由于对称性，你朋友的支配策略是什么？
- (e) 纳什均衡是什么？

问题3 的解答

- (a) 如果你的朋友保持沉默，你应该选择**告密**，因为 $10 > 2$ 。
- (b) 如果你的朋友选择告密，那么你仍然应该选择**告密**，因为 $5 > 0$ 。
- (c) 你的支配策略是**告密**。
- (d) 由于对称性，你朋友的支配策略也是**告密**。
- (e) 纳什均衡是(**告密**, **告密**)。