

社会计算

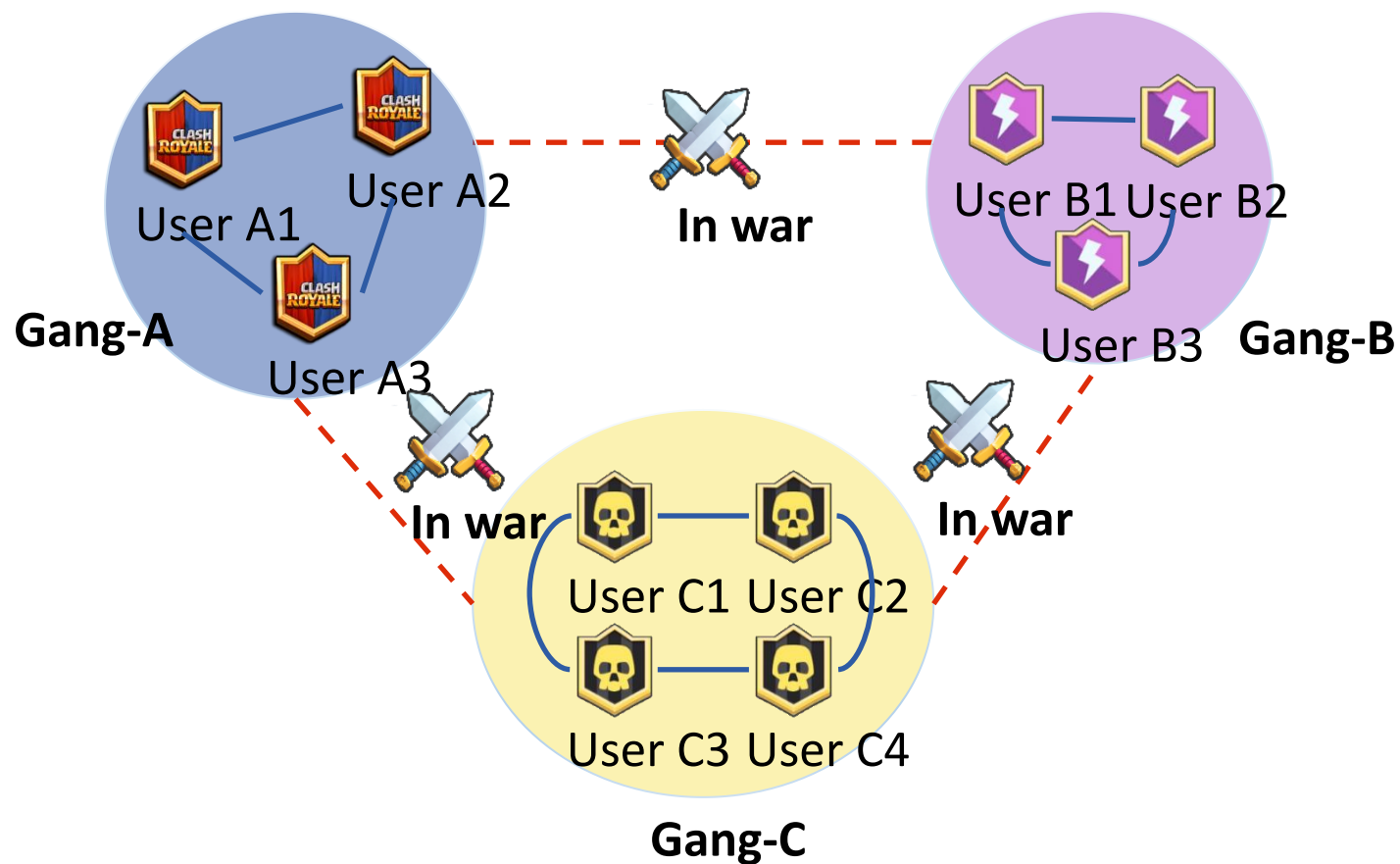


第三节 网络博弈

徐童 2024.3.11

- **更广泛的网络符号性问题**

- 上一节课我们提到，可以通过对正关联节点的归并，来简化非完全图的平衡性判断
- 这种抽象的过程可以揭示不同粒度主体间的合作/竞争关系
 - 现实网络中，不仅个体间存在友好/敌对，个体形成的团体同样有着类似的关系



- **对立环境下的利益与决策**

- 当我们所处的社会网络存在竞争与对立时，我们的决策将影响我们的收益
 - 如何描述博弈的过程？如何判定最优的行为策略？
 - 博弈策略在进化过程中起到何种作用？



- 博弈论概述

- 几类典型的博弈类型
- 纳什均衡
- 混合策略博弈

- 进化博弈论

- 网络流量的博弈模型

- **引子：从田忌赛马说起**
- 小时候都学过的课文——《田忌赛马》？
 - 通过调换三匹马的顺序，2:1反败为胜
 - 由此可以看出，不同的策略对应着不同的结果（收益）

义务教育教科书四年级上册

对策

——田忌赛马



城厢小学：马冬梅

- **引子：从田忌赛马说起**

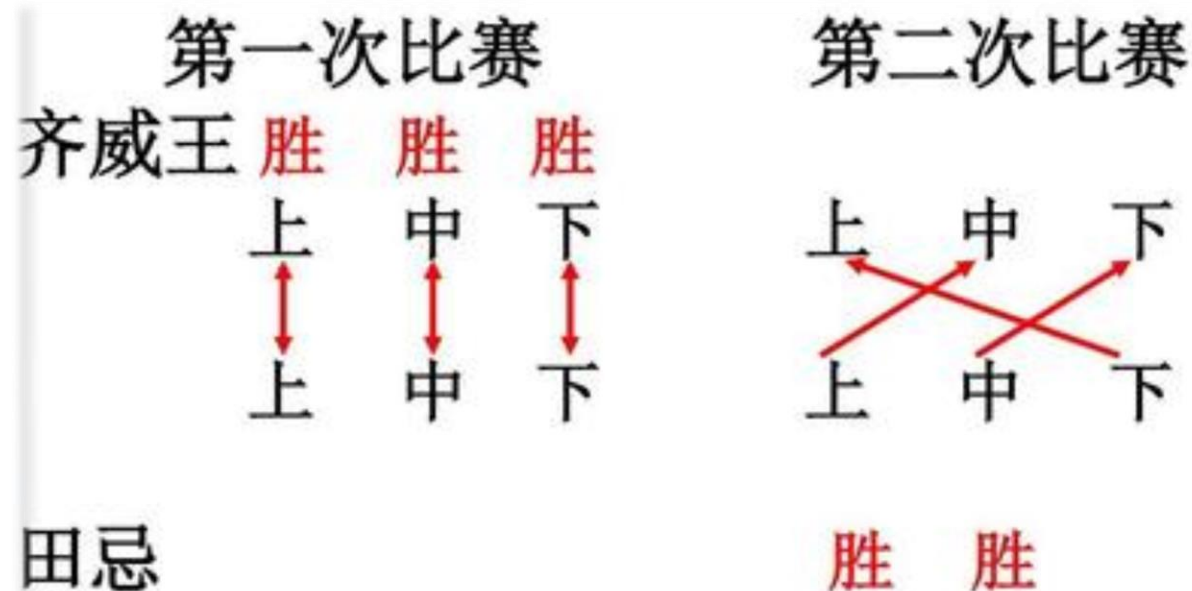
- 在日常生活中，学习田忌赛马的智慧，可以解决许多实际问题（手动狗头）



Arman

三年级时和同学打了一架，我叫来了六年级的哥哥，他把他初二的哥哥叫来了。我也把高一的表哥叫来了，结果他叫来了他上高三的哥！这时候我们以为必输无疑，好在那时表哥已经学会了田忌赛马这篇课文，最后由我对战同学高三的哥。

- **引子：从田忌赛马说起**
- 思考：在这个故事中，包含着哪些要素？
 - **参与人**：齐王、田忌 (+孙臧)
 - **策略集**：赛马的顺序
 - **回报**：比赛的结果 (胜负)



- **引子：从田忌赛马说起**

- 思考：在这个故事中，包含着哪些要素？

- **参与人**：齐威王，田忌

- **策略集**：

1：上中下， 2：上下中， 3：中下上，

4：中上下， 5：下上中， 6：下中上

- **回报**：

-对于策略组 $(1, 1)$ ，田忌三个都输 (-3) ；对于策略组 $(5, 1)$ ，田忌赢二输一 $(+1)$ ；等等

在这个故事中，齐王的策略一直为第 1 种，而田忌的策略由第 1 种转为了第 5 种，从而反败为胜

- **引子：从田忌赛马说起**
- 思考：田忌赛马教会了我们什么？
 - 在对方策略固定的情况下，我们可以见招拆招，选择对应策略进行压制
- 然而在现实中，更多的时候，双方并不知道对方将要采取何种策略
 - 在此情况下，如何根据双方的潜在收益进行预判？



• **第一个例子：复习考试 or 准备报告？**

• 学生的日常：Deadline是第一生产力

• 不过，偶尔也会出现玩砸的情况：事情做不完了，必须有所取舍

• 假如，你需要完成考试和报告两个任务，具体情况如下：

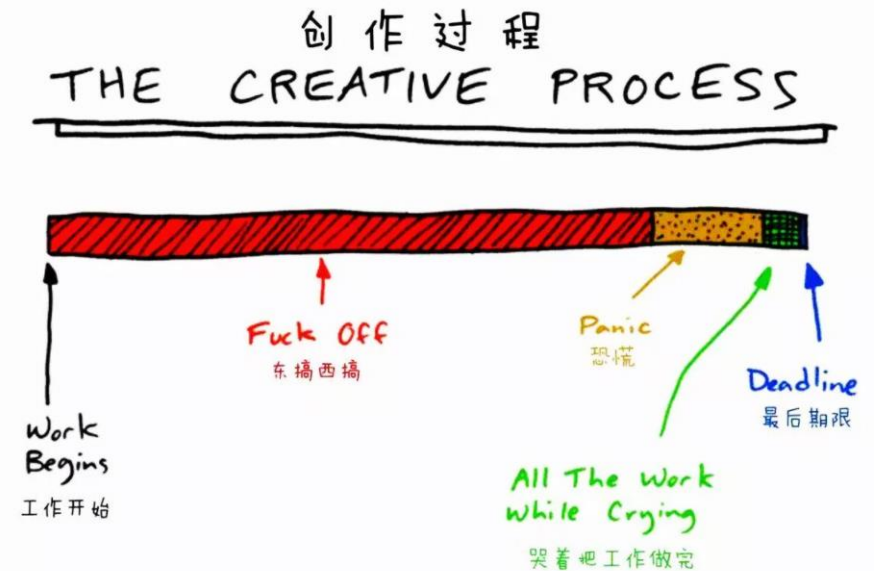
• 考试：复习 = 92分，没复习 = 80分

• 报告，需要和搭档合伙完成，其中

➢ 如果两人共同完成，100分

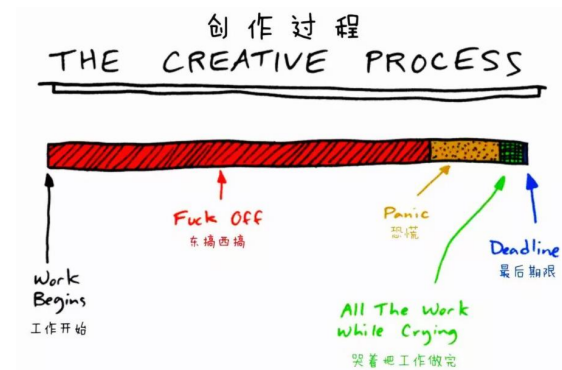
➢ 如果单独一人完成，92分

➢ 如果两人都撂挑子，84分



- **第一个例子：复习考试 or 准备报告？**
- 此时，对于每个学生来说有两种策略：复习迎考 or 准备报告
- 如何计算各种策略的收益：
 - 如果两个人都复习迎考，则收益均为 $(92+84)/2 = 88$ 分
 - 如果两个人都准备报告，则收益均为 $(80+100)/2 = 90$ 分
 - 如果一个人准备报告，一个人复习考试，则：
 - 准备报告的人收益为 $(80+92)/2 = 86$ 分
 - 复习考试的人收益为 $(92+92)/2 = 92$ 分

- **第一个例子：复习考试 or 准备报告？**
- 我们将这个过程套入“田忌赛马”总结出的三要素中，定义博弈的基本要素
 - **参与者**：至少双方，在本案例中是“你”和“你的搭档”
 - **策略集**：每个参与者都有一组如何行为的备选项，在本案例中是“复习”或“准备报告”
 - **收益**：每个策略行为-组合的选择都会导致参与人获得一定收益，在本案例中是分数



• **第一个例子：复习考试 or 准备报告？**

• 从而，我们可以获得如下收益矩阵：

- 其中，每个小方格对应着一个横轴-纵轴组成的策略组合
- 方格中的两个数字，前者对应着 “You” 的收益，后者对应着 “Your Partner” 的收益
 - 注意：参与者可能在同一策略组合中的收益不一致（如下图对角线，双方地位不对等）

		Your Partner	
		<i>Presentation</i>	<i>Exam</i>
You	<i>Presentation</i>	90, 90	86, 92
	<i>Exam</i>	92, 86	88, 88

Figure 6.1: Exam or Presentation?

- **第一个例子：复习考试 or 准备报告？**
- 基于收益矩阵，此时应该如何决策？
 - 从平均分来看，理想情况是大家商议共同准备报告，都可以获得90分的平均分
 - 然而，如果双方没有足够互信，可能因为**白嫖现象**而导致合作失败
 - 如何从收益出发，判断双方的行为并给出合理决策？

		Your Partner	
		<i>Presentation</i>	<i>Exam</i>
You	<i>Presentation</i>	90, 90	86, 92
	<i>Exam</i>	92, 86	88, 88

Figure 6.1: Exam or Presentation?

- **第一个例子：复习考试 or 准备报告？**

- 博弈中的参与人是如何决策的？一般而言，应该满足一下三个基本假设：

- ① 参与人对于收益矩阵有着充分、完整的了解，且所关心的事务只包含在收益矩阵中
- ② 参与人必须满足理性假设
 - ✓ 追求自身收益尽可能的最大化
 - ✓ 同时，相信其他参与人也会持有同样逻辑
- ③ 参与人的决策是相互独立的
 - ✓ 没有商量 / 串通等行为



- **第一个例子：复习考试 or 准备报告？**
- 基于上述的假设，我们来复盘一下面临决策时的收益问题
 - 如果你知道你的搭档将复习考试，那么由于 $88 > 86$ ，复习考试对你而言更合理
 - 如果你知道你的搭档将准备报告，那么由于 $92 > 90$ ，复习考试对你而言依然更合理
 - 总结起来，其实你只有一种最佳选择：复习考试

		Your Partner	
		<i>Presentation</i>	<i>Exam</i>
You	<i>Presentation</i>	90, 90	86, 92
	<i>Exam</i>	92, 86	88, 88

Figure 6.1: Exam or Presentation?

• **第一个例子：复习考试 or 准备报告？**

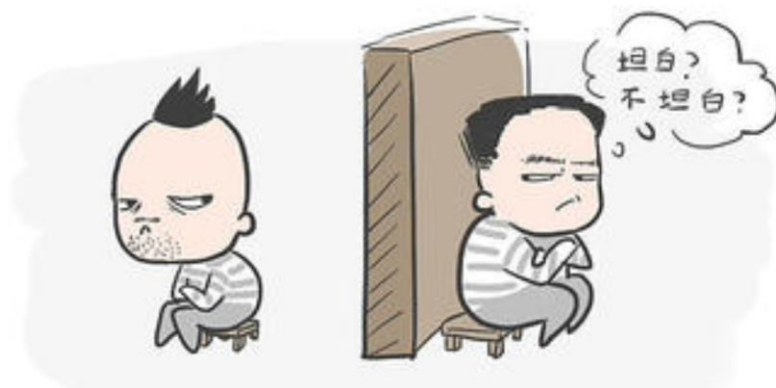
- 通过上述复盘，我们发现在这个博弈中，有“**严格占优策略**”存在：
 - 对于一个参与人（A）来说，如果存在一种策略，无论另一个参与人（B）选择哪种策略，该策略都是最佳选择，那么这个策略就是A的**严格占优策略**
 - 在本案例中，“复习考试”就是**双方的**严格占优策略

		Your Partner	
		<i>Presentation</i>	<i>Exam</i>
You	<i>Presentation</i>	90, 90	86, 92
	<i>Exam</i>	92, 86	88, 88

Figure 6.1: Exam or Presentation?

- **经典的囚徒困境**

- 事实上，类似的情况可见于著名的“囚徒困境”博弈
- 假如有两个嫌疑犯被警方抓获并分开审讯，并被告知如下事项
 - 如果你坦白，而另外一人抵赖，则你马上释放；另外一人将承担全部罪行，会被判刑10年
 - 如果你们都坦白，你们的罪行将被证实。但由于你们有认罪的表现——判刑4年
 - 如果你们都不坦白，那么没有证据证明你们抢劫，我们将以拒捕罪控告你们——判刑1年。



- 经典的囚徒困境

- 根据上文的信息，我们可以得到如下的收益矩阵

		Suspect 2	
		<i>NC</i>	<i>C</i>
Suspect 1	<i>NC</i>	-1, -1	-10, 0
	<i>C</i>	0, -10	-4, -4

Figure 6.2: Prisoner's Dilemma

- 显然，在这个案例中，“坦白”对于两个囚犯来说都是严格占优策略
 - 尽管抵赖对于两个人来说实际上收益会更高
 - 事实上，面对个体的私利，建立合作是一件非常困难的事情

- 经典的囚徒困境
- 囚徒困境现象，在日常生活中非常普遍



军备竞赛与“东亚怪物房”



“多考一分，干掉千人”

- 经典的囚徒困境

- 从现实来看，囚徒困境往往会导致整体收益的降低

- 例如，“服用兴奋剂”问题

		Athlete 2	
		<i>Don't Use Drugs</i>	<i>Use Drugs</i>
Athlete 1	<i>Don't Use Drugs</i>	3, 3	1, 4
	<i>Use Drugs</i>	4, 1	2, 2

Figure 6.3: Performance-Enhancing Drugs

- 兴奋剂会导致双方的收益都降低（考虑到身体的损害）
- 然而，相比于不使用兴奋剂的劣势，人们还是会选择使用兴奋剂

如何破局？

- 经典的囚徒困境

- 可以通过调整收益，诱导决策向更为温和的方向发展

- 例如，我们将复习考试-准备报告问题的收益矩阵修正如下：

		Your Partner	
		<i>Presentation</i>	<i>Exam</i>
You	<i>Presentation</i>	98, 98	94, 96
	<i>Exam</i>	96, 94	92, 92

- 在这个版本中，考试变得更为容易：复习 = 100分，不复习 = 96分

- 此时，“准备报告”取代“复习考试”成为新的严格占优策略，合作成为首选

- **经典的囚徒困境**

- 另一方面，囚徒困境的存在根源来自于参与方的互不信任

- 如果能够实现参与方之间某种形式的信息沟通，
则囚徒困境不攻自破
- 感兴趣的同学可以看看课外小视频：Wil Aime的
“囚徒困境”（太长了就不在课上放了）
- 法文原版：

<https://www.bilibili.com/video/BV11v411x7vu>



• 最佳应对与占优策略

• 经过第一个案例的洗礼，我们试图给出策略选择的符号化定义

- 假设， S 是参与人1的一个策略， T 是参与人2的一个策略
- 在收益矩阵中，我们采用 $P_1(S, T)$ 表示参与人1通过决策组 (S, T) 获得的收益， $P_2(S, T)$ 表示参与人2通过决策组 (S, T) 获得的收益
- 在此情况下，如果参与人1采用策略 S 所获得的收益大于等于其他任何决策，即

$$P_1(S, T) \geq P_1(S', T)$$

- 则有参与人1的策略 S 是参与人2的策略 T 的 最佳应对（由于等号，可能不唯一）
- 如果将“ \geq ”改为“ $>$ ”，则是 严格最佳应对（不一定存在，但存在一定唯一）

- **最佳应对与占优策略**
- 通过对于“最佳应对”的定义，我们可以进一步明确“占优策略”的概念
 - 参与人1的占优策略，是指该策略针对参与人2的任何一个策略，都是最佳应对
 - 同理，参与人1的严格占优策略，是指对参与人2的任何一个策略，都是严格最佳应对
 - 显然，如果严格占优策略存在，则可以预期参与人将采用这一策略



• **最佳应对与占优策略**

- 在先前的“囚徒困境”例子中，对于两个参与人，都可以找到严格占优策略
- 然而，现实中某些案例并不是所有人都存在严格占优策略
 - 例如，两家公司要规划一种新产品。根据目标市场情况，可以采用“廉价”或“高档”两种策略，对应的收益矩阵如下表所示：

		Firm 2	
		<i>Low-Priced</i>	<i>Upscale</i>
Firm 1	<i>Low-Priced</i>	.48, .12	.60, .40
	<i>Upscale</i>	.40, .60	.32, .08

Figure 6.5: Marketing Strategy

• **最佳应对与占优策略**

• 然而，现实中某些案例并不是所有人都存在严格占优策略

• 在下表中，我们惊讶地发现：

➤ 对于公司 1 而言，无论公司 2 采取何种措施，“廉价”都是严格占优策略

➤ 但对于公司 2 而言，采取与公司 1 相反的策略才是最佳应对 (未达成一致性!)

		Firm 2	
		<i>Low-Priced</i>	<i>Upscale</i>
Firm 1	<i>Low-Priced</i>	.48, .12	.60, .40
	<i>Upscale</i>	.40, .60	.32, .08

Figure 6.5: Marketing Strategy

• **最佳应对与占优策略**

• 这种情况下，公司 2 应该如何决策呢？

- 我们已知，如果某个参与人具有严格占优策略，那么可以预期参与人将采用这一策略
- 相应的，由于公司 1 具有严格占优策略，那么公司 2 只要相应地采用其最佳应对即可
 - 因此，可以预期公司 1 将采用“廉价”策略，而公司 2 将相应地选择“高档”策略

		Firm 2	
		<i>Low-Priced</i>	<i>Upscale</i>
Firm 1	<i>Low-Priced</i>	.48, .12	.60, .40
	<i>Upscale</i>	.40, .60	.32, .08

Figure 6.5: Marketing Strategy

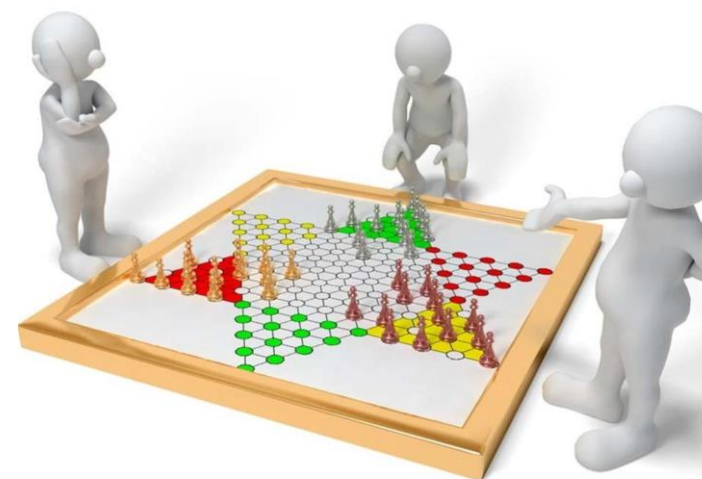
- **最佳应对与占优策略**

- 我们基于这个案例，给出简单博弈的行为推理原则

- 如果两个参与人都有严格占优策略，则可以预计他们均会采取严格占优策略；
- 如果只有一个参与人有严格占优策略，则这个参与人会采取严格占优策略，而另一方会采取此策略的最佳应对（一定会有！）

那么问题来了：

如果两个人都没有严格占优策略呢？



• 纳什均衡

- 现实中真的存在双方都没有严格占优策略的情况吗？还真有
 - 同样，我们假设存在这样两个公司，试图争取同一类市场中的目标客户。
 - 其中，公司 1 规模较小，往往难以单独获取客户源，只能依附于公司 2
 - 根据不同客户带来的收益，我们可以获得如下的收益矩阵：

		Firm 2		
		<i>A</i>	<i>B</i>	<i>C</i>
Firm 1	<i>A</i>	4, 4	0, 2	0, 2
	<i>B</i>	0, 0	1, 1	0, 2
	<i>C</i>	0, 0	0, 2	1, 1

Figure 6.6: Three-Client Game

• 纳什均衡

- 针对如下收益矩阵，我们困惑地发现，两个公司都不具有“严格占优策略”
 - 对于公司 1 而言，其最佳应对是公司 2 选择哪个客户，它跟着选择哪个客户
 - 对于公司 2 而言，情况较为复杂：
 - 当公司 1 选择客户 A 时，它也应该选择客户 A；
 - 而当公司 1 选择客户 B/C 时，它的选择应该相反
 - 某种意义上是“红海”和“蓝海”的区别

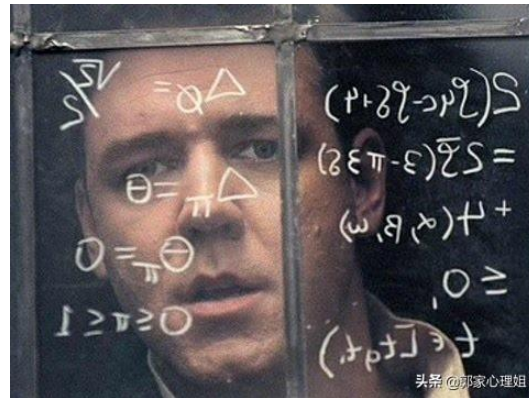
		Firm 2		
		A	B	C
Firm 1	A	4, 4	0, 2	0, 2
	B	0, 0	1, 1	0, 2
	C	0, 0	0, 2	1, 1

Figure 6.6: Three-Client Game

• **纳什均衡**

- 在这种情况下，如何判断两个公司的策略行为呢？
 - 我们发现，两个公司的策略之间存在的明确的对应关系
 - 对于公司 1 的任何一种策略，公司 2 都有策略与之成对
 - 1950年，约翰·纳什提出了重要的“纳什均衡”原则

如果参与人 1 与 参与人 2
 分别选择了策略 S 与策略 T
 且 S 与 T 互为**最佳应对**
 则策略组 (S, T) 为一个**纳什均衡**



• 纳什均衡

• 回到我们先前的案例中来

- 显然，在该案例中，当公司 1/2 选择客户B/C的时候，它们的策略并不是互为最优应对
 - 公司 1 选 B，公司 2 的最佳应对为 C，但此时公司 1 的最佳应对也为 C，反之亦然
- 只有当公司 1 和 公司 2 都选择客户 A 的时候，它们的策略才是互为最优应对
 - 因此，本案例中只存在一组纳什均衡情况

即 (A, A)

		Firm 2		
		A	B	C
Firm 1	A	4, 4	0, 2	0, 2
	B	0, 0	1, 1	0, 2
	C	0, 0	0, 2	1, 1

Figure 6.6: Three-Client Game

- **纳什均衡**
- 造成纳什均衡的原因是：在这一均衡状态下，任何参与人都没有动机（理性的/收益的理由）去更换另一种策略
 - 互为最佳应对，意味着谁也不可能单方面通过改变策略获得额外好处
 - 尽管两人都做出改变是有可能都获得更好的收益的
 - 纳什均衡是一种**信念上**的均衡

Nash Equilibrium



• **多重均衡与协调博弈**

- 从上面的例子可以看出，当双方都不存在严格占优策略的时候，我们可以通过寻找纳什均衡解来判定双方的策略
- 然而，如果纳什均衡解不止一组呢？
 - 协调博弈：存在多种纳什均衡的情况，需要通过协调来进行决策
 - 例如，选择何种软件制作报告？PPT or Keynote？
 - 这种情况下，其实与收益无关

		Your Partner	
		<i>PowerPoint</i>	<i>Keynote</i>
You	<i>PowerPoint</i>	1, 1	0, 0
	<i>Keynote</i>	0, 0	1, 1

Figure 6.7: Coordination Game

- **多重均衡与协调博弈**
- 当面对与收益无关的协调博弈时，有时需要外部规定或者约定俗成来解决问题
- 例如，两辆车相向而行，如何避让？
 - 通过法律规定所有车辆靠左/靠右行驶的方式来规避相撞的可能（左右无伤大雅）



• **多重均衡与协调博弈**

- 通过对于收益矩阵进行简单调整，可以得到更多变形情况
 - 例如，同样还是两组纳什均衡情况，但此时收益是不平衡的
 - 此时，存在一组纳什均衡解，使得双方参与人的收益都上升
 - 例如，如果你在电二楼工作，中午去本科生/研究生食堂的距离是相近的
 - 但是，如果你是在西区活动中心听报告，显然本科生食堂更为接近
- 这种情况下的决策是显而易见的

		Your Partner	
		<i>PowerPoint</i>	<i>Keynote</i>
You	<i>PowerPoint</i>	1, 1	0, 0
	<i>Keynote</i>	0, 0	2, 2

Figure 6.8: Unbalanced Coordination Game

- **多重均衡与协调博弈**

- 由基本协调博弈衍生出的“猎鹿博弈”，受到了极大的关注
 - 表面上看，猎鹿博弈与收益不平衡的协调博弈是一致的
 - 其区别主要在于不合作情况下的收益
 - 按照先前的思路，我们发现这里同样存在两个纳什均衡，且共同猎鹿收益更高
 - 因此，最后的决策应该是显而易见的

		Hunter 2	
		<i>Hunt Stag</i>	<i>Hunt Hare</i>
Hunter 1	<i>Hunt Stag</i>	4, 4	0, 3
	<i>Hunt Hare</i>	3, 0	3, 3

Figure 6.10: Stag Hunt

• **多重均衡与协调博弈**

• 由基本协调博弈衍生出的“猎鹿博弈”，受到了极大的关注

- 然而，相比于基本的不平衡协调博弈，猎鹿博弈的区别体现在“对于追求高收益的惩罚”
 - 如果双方合作猎鹿或者各自猎兔，都能获得相当的收益
 - 但如果一方合作而另一方不合作，则选择合作的那一方会收到更重的惩罚（收益为0）
 - 某种意义上是对“囚徒困境”的延伸，即选择合作时需要考虑潜在的负面效应

		Hunter 2	
		<i>Hunt Stag</i>	<i>Hunt Hare</i>
Hunter 1	<i>Hunt Stag</i>	4, 4	0, 3
	<i>Hunt Hare</i>	3, 0	3, 3

Figure 6.10: Stag Hunt

- **多重均衡与协调博弈**

- 由基本协调博弈衍生出的“猎鹿博弈”，受到了极大的关注

- 事实上，类似的现象在现实中屡见不鲜

- 提醒我们：试图通过合作获得更大收益时，要小心合作伙伴的**背刺**

- *刘皇叔点了个赞*

- *大魏吴王孙十万骂骂咧咧地退出了直播间*



• **多重均衡与协调博弈**

- 通过对于收益矩阵进行简单调整，可以得到更多变形情况
 - 另一种变形的情况更为复杂，也更能体现“协调性”
 - 此时，不同参与人的偏好是不一致的（书中将其称为“性别战”）
 - 因此，必须通过协调，做出偏向一方的决策
 - 例如，电视遥控器由谁负责掌管？（非常重要）

		Your Partner	
		<i>PowerPoint</i>	<i>Keynote</i>
You	<i>PowerPoint</i>	1, 2	0, 0
	<i>Keynote</i>	0, 0	2, 1

Figure 6.9: Battle of the Sexes

• **多重均衡与协调博弈**

- “性别战” 没有显见的策略组，需要通过协调仍可达成一致
- 现实中，还存在着另一类常见的“反协调”活动：鹰-鸽博弈
 - 例如，如果两只动物要分配一块食物，可以选择“分享”或“争夺”
 - 如果“分享”，则各得一半；如果“争夺”，则食物被毁
 - 如果一方“争夺”而另一方“分享”
 则“争夺”一方可以获得大部分食物

		Animal 2	
		<i>D</i>	<i>H</i>
Animal 1	<i>D</i>	3, 3	1, 5
	<i>H</i>	5, 1	0, 0

Figure 6.12: Hawk-Dove Game

• **多重均衡与协调博弈**

• 而现实中，还存在着另一类常见的“反协调”活动

- 显然，在这一收益矩阵下，存在着 (D, H) 和 (H, D) 两个纳什均衡组合
- 在这种情况下，我们很难基于纳什均衡判断参与者的行为
 - 可见，纳什均衡有助于缩小参与者可能参与策略的范围，但未必能给出唯一预测

• 鹰-鸽博弈在现实中很常见

- 例如，对立的两个国家往往处于一方攻，一方守的态势
- 此类博弈也被称为“懦夫博弈”
 - 两辆车相向而行，看谁因惧怕先改道

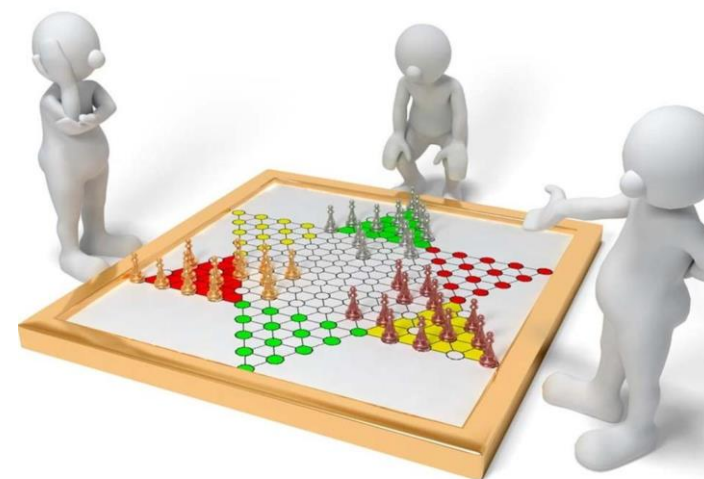
		Animal 2	
		<i>D</i>	<i>H</i>
Animal 1	<i>D</i>	3, 3	1, 5
	<i>H</i>	5, 1	0, 0

Figure 6.12: Hawk-Dove Game

- **多重均衡与协调博弈**

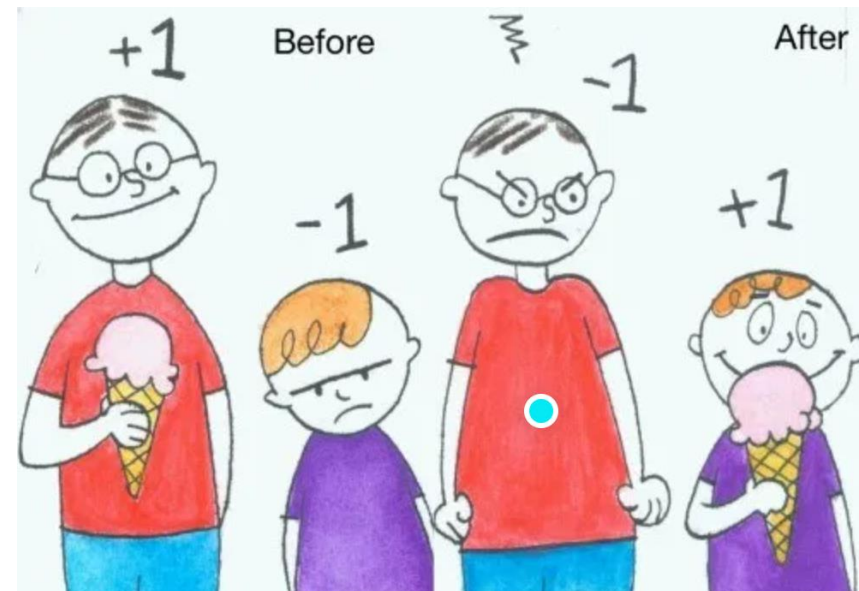
- 最后，我们进一步完善先前所给出的简单博弈行为推理原则

- 如果两个参与人都有严格占优策略，则可以预计他们均会采取严格占优策略；
- 如果只有一个参与人有严格占优策略，则这个参与人会采取严格占优策略，而另一方会采取此策略的最佳应对（一定会有！）
- 如果双方都不存在严格占优策略，则寻找纳什均衡
 - 如果只有一个纳什均衡，则该均衡对应决策结果
 - 如果存在多个纳什均衡，则需要额外信息辅助推断
 - ◆ 例如协调博弈（性别战），鹰-鸽博弈等
 - ◆ 再次强调：纳什均衡有助于缩小范围，但不能保证有结果



• **零和博弈与混合策略**

- 上文中介绍了多种博弈形态，其中参与者基本上都能获得收益
 - 区别仅在于收益的多少，最差情况是收益为 0
- 但现实中，往往存在着更为残酷的所谓 **“零和博弈”**
 - 收益不会凭空增多，你的收益来自于我的损失
 - 例如，期货等投资的“零和游戏”性
 - *股票算不算零和游戏？*



• **零和博弈与混合策略**

- 我们通过简单的硬币配对，来了解最基本的零和博弈问题
 - 甲乙各持一枚硬币，同时选择手中硬币的正反面
 - 如果他们选择的朝向相同，则乙赢得甲的硬币；反之，甲将赢得乙的硬币
 - 由此，可以得到右下角的收益矩阵
 - 显然，这个收益矩阵是不存在纳什均衡的
 - 甲希望硬币朝向相反，乙希望相同，两者永远对立
 - 现在，所有的行为推理原则都失效了，怎么办？

		Player 2	
		<i>H</i>	<i>T</i>
Player 1	<i>H</i>	-1, +1	+1, -1
	<i>T</i>	+1, -1	-1, +1

Figure 6.14: Matching Pennies

• **零和博弈与混合策略**

• 如何应对这种情况？是时候考虑更多因素了

- 一个有趣的问题：当你玩剪子-包袱-锤游戏的时候，三种手势的概率是一致的吗？
- 现在问题来了，当考虑不同策略的概率时，会给行为判断带来哪些变化？
 - 我们还是回到抛硬币的二元选择环境下
 - 此时，你的“策略”可以视作概率加成下的两种选择

• 由此引出新问题：混合策略

- 相对于不考虑概率的“**纯策略**”而言

		Player 2	
		<i>H</i>	<i>T</i>
Player 1	<i>H</i>	-1, +1	+1, -1
	<i>T</i>	+1, -1	-1, +1

Figure 6.14: Matching Pennies

• **零和博弈与混合策略**

• 混合策略的核心思想：引入随机性，参与人以一定概率在策略中进行选择

• 例如，在抛硬币问题中，我们约定

• 参与人 1 的策略是概率 p ，指以概率 p 执行 H，以概率 $1-p$ 执行 T

• 参与人 2 的策略是概率 q ，指以概率 q 执行 H，以概率 $1-q$ 执行 T

• 在此情况下，我们计算收益时，需要考虑概率的加成

• 假如参与人 1 选择 H，则收益为

➤ $(-1) \cdot q + (1)(1-q) = 1 - 2q$

• 反之，参与人 1 选择 T，则收益为

➤ $(1) \cdot q + (-1)(1-q) = 2q - 1$

		Player 2	
		<i>H</i>	<i>T</i>
Player 1	<i>H</i>	-1, +1	+1, -1
	<i>T</i>	+1, -1	-1, +1

Figure 6.14: Matching Pennies

• 零和博弈与混合策略

• 混合策略的核心思想：引入随机性，参与人以一定概率在策略中进行选择

• 在这一情况下，我们根据概率 q 的取值，即可判断参与人 1 选择各种策略收益更高

➤ 若 $q > 0.5$ ，则 $1 - 2q < 2q - 1$

✓ 此时，参与人 1 应选择策略 T，收益更高

✓ 事实上， $q > 0.5$ 意味着参与人 2 更倾向于选择策略 H，因此参与人 1 确实应该选 T

➤ 若 $q < 0.5$ ，则 $1 - 2q > 2q - 1$ ，参与人 1 应选择策略 H

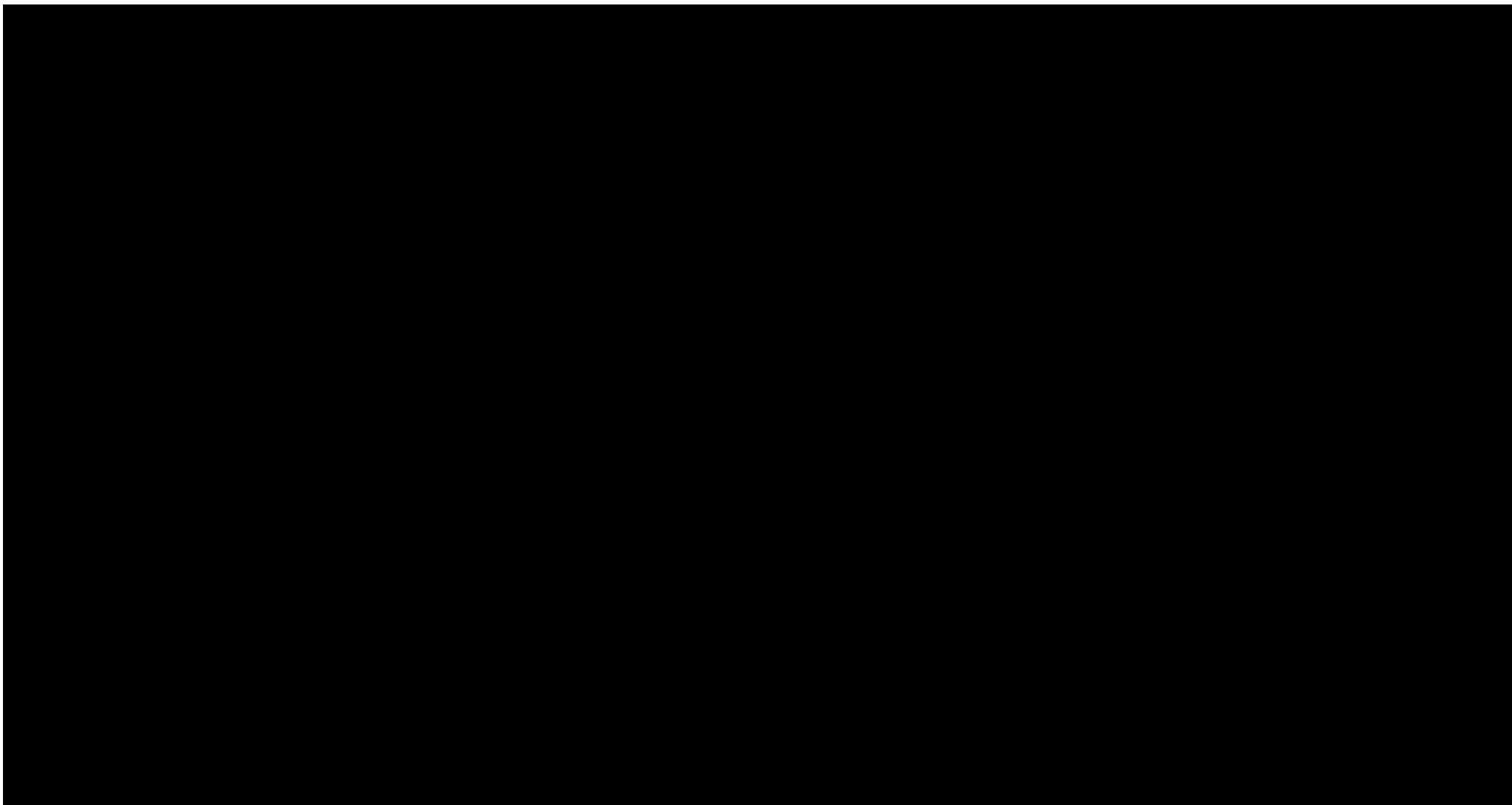
		Player 2	
		H	T
Player 1	H	-1, +1	+1, -1
	T	+1, -1	-1, +1

Figure 6.14: Matching Pennies

- **零和博弈与混合策略**

- 混合策略的核心思想：引入随机性，参与人以一定概率在策略中进行选择
 - 从上面的分析中可以看出，参与人 1 做决策的关键，在于掌握参与人 2 策略对应的概率
 - 因此，通过战术误导对方错判概率，可能将成为在零和博弈中致胜的关键
 - 放松时刻：电影《赌博默示录》的经典片段
 - 玩战术的心都脏 (🐱)

- **零和博弈与混合策略**



- **零和博弈与混合策略**
- 混合策略的核心思想：引入随机性，参与人以一定概率在策略中进行选择
 - 另：一类特殊情况：当 $q = 0.5$ 的时候，此时无论选择策略 H 或 T，期望收益并没有区别
 - 因此，当参与人 2 以 0.5 为概率时，无论参与人 1 如何决策都一样，反之亦然
 - 由此可见， $(0.5, 0.5)$ 是互为最佳对应的一种情况
 - 求解混合策略均衡的原理：**无差异原理**
 - 一对混合策略互为最佳对应的必要条件是，它们分别使对方在两个纯策略上得到的回报无差异

- **尾声：帕累托最优与社会最优**
- 先前的博弈讨论中，我们的目标一直在于考虑个体收益的最优
- 然而，如果将所有参与人视作一个群体，则未必获得了群体的最优
 - 例如，“复习 or 报告”的案例中，个体实现了收益最大（其实是损失最小）
 - 但群体的收益实际上仍有进一步提升的空间

		Your Partner	
		<i>Presentation</i>	<i>Exam</i>
You	<i>Presentation</i>	90, 90	86, 92
	<i>Exam</i>	92, 86	88, 88

Figure 6.1: Exam or Presentation?

• 尾声：帕累托最优与社会最优

• 帕累托最优

- 一组策略被称为“帕累托最优”，当且仅当不存在其他策略，使得能够保障所有参与者至少获得和当前一样高的收益前提下，使至少一个参与者的收益提升
- 例如，下面的收益矩阵中，(92, 86) 是一个帕累托最优，因为找不到其他的策略组，使得在参与人 1 的收益 (92) 不降低的前提下，提升参与人 2 的收益 (86)

		Your Partner	
		<i>Presentation</i>	<i>Exam</i>
You	<i>Presentation</i>	90, 90	86, 92
	<i>Exam</i>	92, 86	88, 88

Figure 6.1: Exam or Presentation?

• **尾声：帕累托最优与社会最优**

• 帕累托最优

- 有趣的是，下方收益矩阵中有三组帕累托最优的策略组，而仅有的一组非帕累托最优的情况，对应着唯一的纳什均衡
- 事实上，虽然人们会意识到有更好的（帕累托最优的）选择，但在缺乏有约束力的合作协议情况下，人们很难维持这种方案 → **不信任导致社会收益降低/代价提升！**

		Your Partner	
		<i>Presentation</i>	<i>Exam</i>
You	<i>Presentation</i>	90, 90	86, 92
	<i>Exam</i>	92, 86	88, 88

Figure 6.1: Exam or Presentation?

- **尾声：帕累托最优与社会最优**

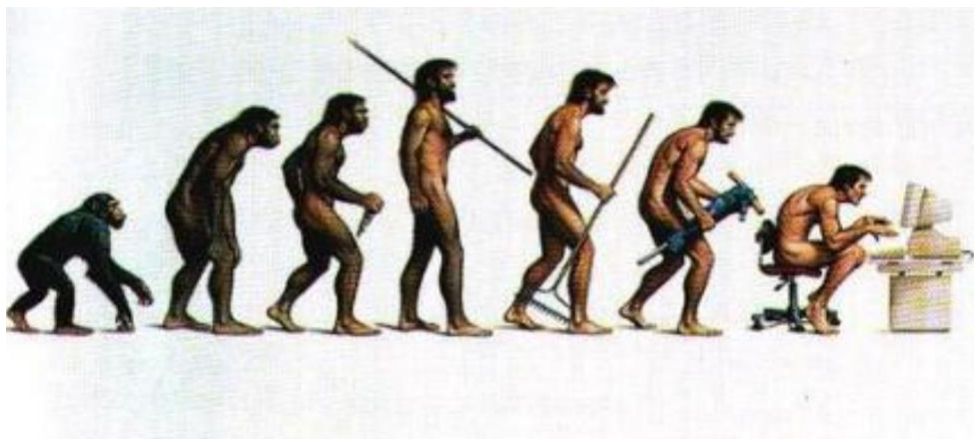
- 社会最优

- 相比于“帕累托最优”，社会最优相对简单粗暴：直接将收益求和并寻找最大值
 - 例如，前述收益矩阵中社会收益最高的情况即“ $90+90=180$ ”的情况
- 社会最优一定是帕累托最优！（想想看，为什么？ Tips：反证法）
- 社会最优有可能和纳什均衡是一致的（例如，P26的另一版本收益矩阵，如下）

		Your Partner	
		<i>Presentation</i>	<i>Exam</i>
You	<i>Presentation</i>	98, 98	94, 96
	<i>Exam</i>	96, 94	92, 92

- 博弈论概述
 - 几类典型的博弈类型
 - 纳什均衡
 - 混合策略博弈
- **进化博弈论**
- 网络流量的博弈模型

- **从博弈到进化博弈**
- 在上一章中，我们讨论了不同博弈中行为推断问题
- 从更长远的视角来看，我们意识到，由于收益不同，部分策略可能逐渐会遭到摒弃，而另一些策略则可能持续存在
 - 例如，不同生物在进化中体现出“物竞天择、适者生存”的现象



- **从博弈到进化博弈**

- 如何从博弈论的角度看待和衡量这种策略的可持续性问题？

- 我们以某个甲虫的种群为例，讨论不同类型甲虫的可持续性问题（这不是生物课）

- 由于基因突变，部分甲虫的体型明显增大

- 大体态的甲虫需要更多营养，否则难以维持生命

- 单纯从这一点来看，大体态甲虫似乎更不可持续

- 那么，从博弈视角来看待这一问题呢？



- 从博弈到进化博弈

- 如何从博弈论的角度看待和衡量这种策略的可持续性问题？

- 当考虑甲虫之间争夺食物的问题时，我们遵循以下原则：

- ① 如果两只甲虫大小相当，则它们平分食物

- ② 如果一只大甲虫和一只小甲虫争夺食物，则大甲虫获得更多食物

- ③ 大甲虫需要消耗更多食物，因此收益更低

- 由此，我们得到右下角所示的收益矩阵

		Beetle 2	
		<i>Small</i>	<i>Large</i>
Beetle 1	<i>Small</i>	5, 5	1, 8
	<i>Large</i>	8, 1	3, 3

Figure 7.1: The Body-Size Game

• 从博弈到进化博弈

- 如何从博弈论的角度看待和衡量这种策略的可持续性问题？
 - 某种意义上说，这一博弈过程与我们开头讨论的“复习-报告”案例是类似的
 - 虽然小体态甲虫可以生活的更为滋润，但为了生存下去，进化为大甲虫更具有优势
 - 进化性军备竞赛的体现
 - 例如，树木的高度增长，根系的向下延伸， etc.
 - 有趣的案例：《瘟疫公司》，传染性/致命性

		Beetle 2	
		<i>Small</i>	<i>Large</i>
Beetle 1	<i>Small</i>	5, 5	1, 8
	<i>Large</i>	8, 1	3, 3

Figure 7.1: The Body-Size Game

- 从博弈到进化博弈

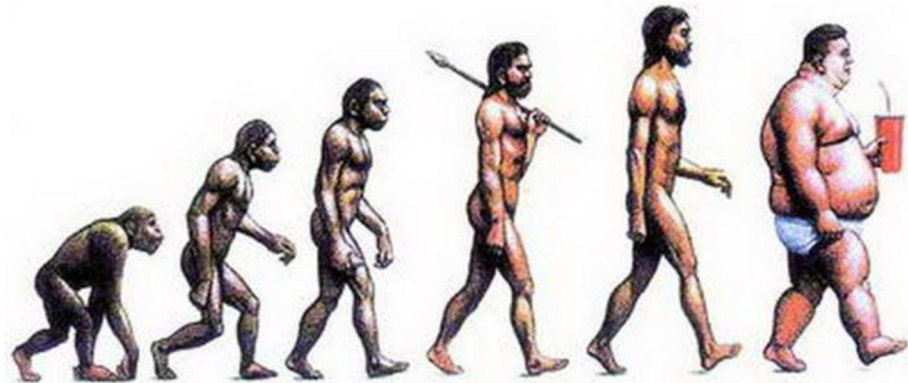
- 更进一步，我们知道这种突变在开始只会以一个很小的比重出现

- 假定存在一个小正数 x ，策略 T 以 x 占比“入侵”策略 S ，表现为：

- ① 生物总体中有 x 占比的生物体采用策略 T

- ② 其余 $1-x$ 占比的生物体采用策略 S

- 假如当 x 处于一定范围内，策略 S 的适应性（收益）严格高于策略 T ，则策略 S **进化稳定**



• 从博弈到进化博弈

• 回到我们先前的甲虫案例中来

• 假定在这一案例中，大体态以占比 x 入侵小体态，则有：

• 一只小甲虫觅食的期望收益是： $5(1-x) + 1x = 5-4x$

• 一只大甲虫觅食的期望收益是： $8(1-x) + 3x = 8-5x$

• 由于约定 x 是一个小正数，则 $5-4x < 8-5x$

• 此时，小体态的策略是不稳定的

• 因此，小体态的甲虫群落会被淘汰

		Beetle 2	
		<i>Small</i>	<i>Large</i>
Beetle 1	<i>Small</i>	5, 5	1, 8
	<i>Large</i>	8, 1	3, 3

Figure 7.1: The Body-Size Game

- 从博弈到进化博弈

- 回到我们先前的甲虫案例中来

- 反之，当我们假定是小体态以占比 x 入侵大体态时，则有：

- 一只小甲虫觅食的期望收益是： $5x + 1(1-x) = 1 + 4x$

- 一只大甲虫觅食的期望收益是： $8x + 3(1-x) = 3 + 5x$

- 由于约定 x 是一个小正数，则 $1 + 4x < 3 + 5x$

- 由此可见，大体态的策略是进化稳定的

- 相应的此，小体态的“入侵”很难成功

		Beetle 2	
		<i>Small</i>	<i>Large</i>
Beetle 1	<i>Small</i>	5, 5	1, 8
	<i>Large</i>	8, 1	3, 3

Figure 7.1: The Body-Size Game

- **从博弈到进化博弈**
- 事实上，这个结果并不出乎意料
 - 当大甲虫入侵小甲虫群体时，由于其强有力的争夺食物能力，小甲虫很难驱赶它们
 - 反过来，当小甲虫入侵大甲虫群体时，它们往往无法争得食物，反而会被驱赶
 - 总结起来，小体态的“生物适应能力”较差
 - 物竞天择，适者生存!



• 从博弈到进化博弈

• 我们将甲虫案例推广到一般化的形式，如右下角收益矩阵

• 假定存在一个小正数 x ，策略 T 以 x 占比“入侵”策略 S，则有：

• 策略 S 的期望收益是： $a(1-x) + bx$

• 策略 T 的期望收益是： $c(1-x) + dx$

• 由于约定 x 是一个小正数，则策略 S 为进化稳定的前提条件是

① $a > c$

② $a = c$ 且 $b > d$

		Organism 2	
		<i>S</i>	<i>T</i>
Organism 1	<i>S</i>	a, a	b, c
	<i>T</i>	c, b	d, d

Figure 7.3: General Symmetric Game

• 从博弈到进化博弈

• 小拓展：进化稳定与纳什均衡有着何种关系？

• 已知策略 T 以 x 占比 “入侵” 策略 S, 策略 S 为进化稳定的前提条件是

① $a > c$

② $a = c$ 且 $b > d$

• 我们有如下结论：如果策略 S 是进化稳定的，则 (S, S) 是一个纳什均衡

• 但反过来，纳什均衡不一定对应进化稳定策略

• 纳什均衡要求 $a \geq c$, 但不关心 b 与 d 的关系

		Organism 2	
		S	T
Organism 1	S	a, a	b, c
	T	c, b	d, d

Figure 7.3: General Symmetric Game

- 博弈论概述
 - 几类典型的博弈类型
 - 纳什均衡
 - 混合策略博弈
- 进化博弈论
- **网络流量的博弈模型**

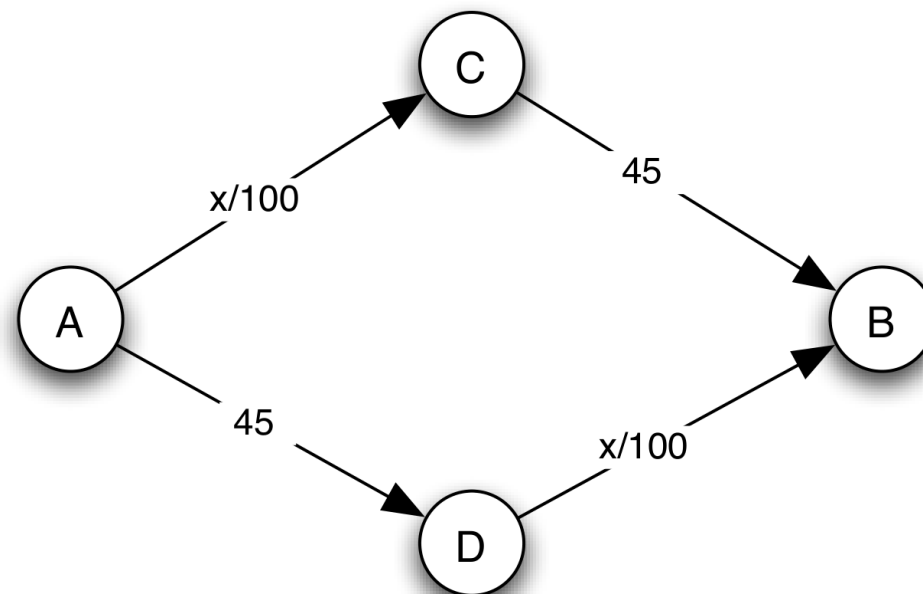
- **一类特殊的网络博弈：流量博弈**
- 今天出门，你堵在哪儿了？
 - 如何在车流量巨大的公路网络中找到合适的路线，对于司机来说无疑是一项挑战
 - 当捷径成为大众选择，捷径 \neq 捷径
 - 节假日高速路免费的后果
 - 事实上，我们可以将路径选择视作一个博弈问题，所有的司机都是参与者，他们的回报就是所需的行程时间



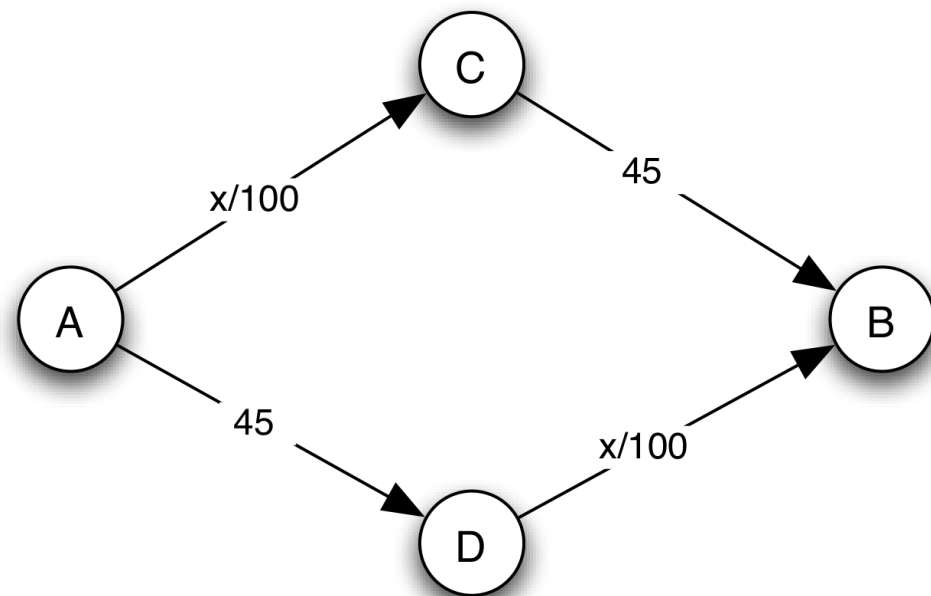
- **一类特殊的网络博弈：流量博弈**

- 我们采用一个简单的例子，来描述网络中的流量博弈

- 假定有 4000 辆车，需要从A到B
 - 参与人：4000位司机
 - 策略：走哪一条路？上面 or 下面
 - 回报：行驶时间（越小越好）

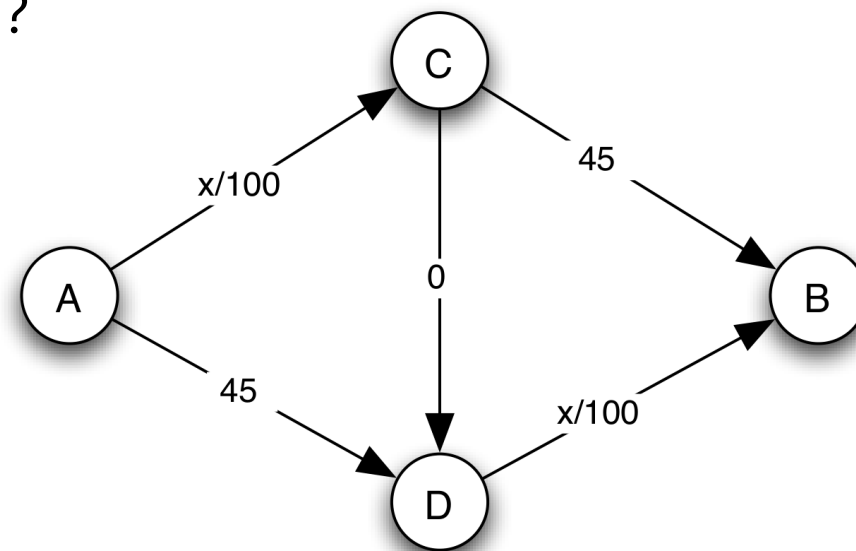


- **一类特殊的网络博弈：流量博弈**
- 我们采用一个简单的例子，来描述网络中的流量博弈
 - 理想的均衡情况是，两条路各有2000位司机通行
 - 此时，所需时长为 $20 + 45 = 65$
 - 而且，这一均衡是可以由司机们自发形成的
 - 如果某一条路车流量大，通行时间长
则司机会自发转到另一条路上
 - 此时的总社会福利为 $4000 * 65 = 260000$



• 布雷斯悖论

- 此时，如果政府出于好意，在C-D之间新修了一条快速路，糟糕的事发生了
 - 为了突出快速路的收益，我们甚至将其时长设为 0
 - 结果此时的均衡状态变成了 $A \rightarrow C \rightarrow D \rightarrow B$ ，所有司机的行驶时间都成了80！
 - 相比于之前的65，新修道路反而延长了通行时间？
 - 更糟的是，此时没有人有改变的动力
 - 如果改走原先的道路，时间为 $85 > 80$
 - 虽然这样做会使大多数人收益，但是我为什么要牺牲我自己呢？



- **一类特殊的网络博弈：流量博弈**
- 回到我们开头的例子，好事一定带来好结果？
 - 节假日高速免费，反而带来了严重拥堵
 - 部分城市将高速路改为公园，反而诱导司机选择公共交通，实际上减少了通行时间
 - 在制定公共政策时，要小心新政策过大的吸引力造成社会资源分配不均，“好心办坏事”



本章小结

网络博弈

- 博弈论概述
 - 几类典型的博弈类型
 - 纳什均衡
- 混合策略博弈
- 进化博弈论
- 网络流量的博弈模型