博弈论

第一讲：绪论

黄嘉平

深圳大学中国经济特区研究中心

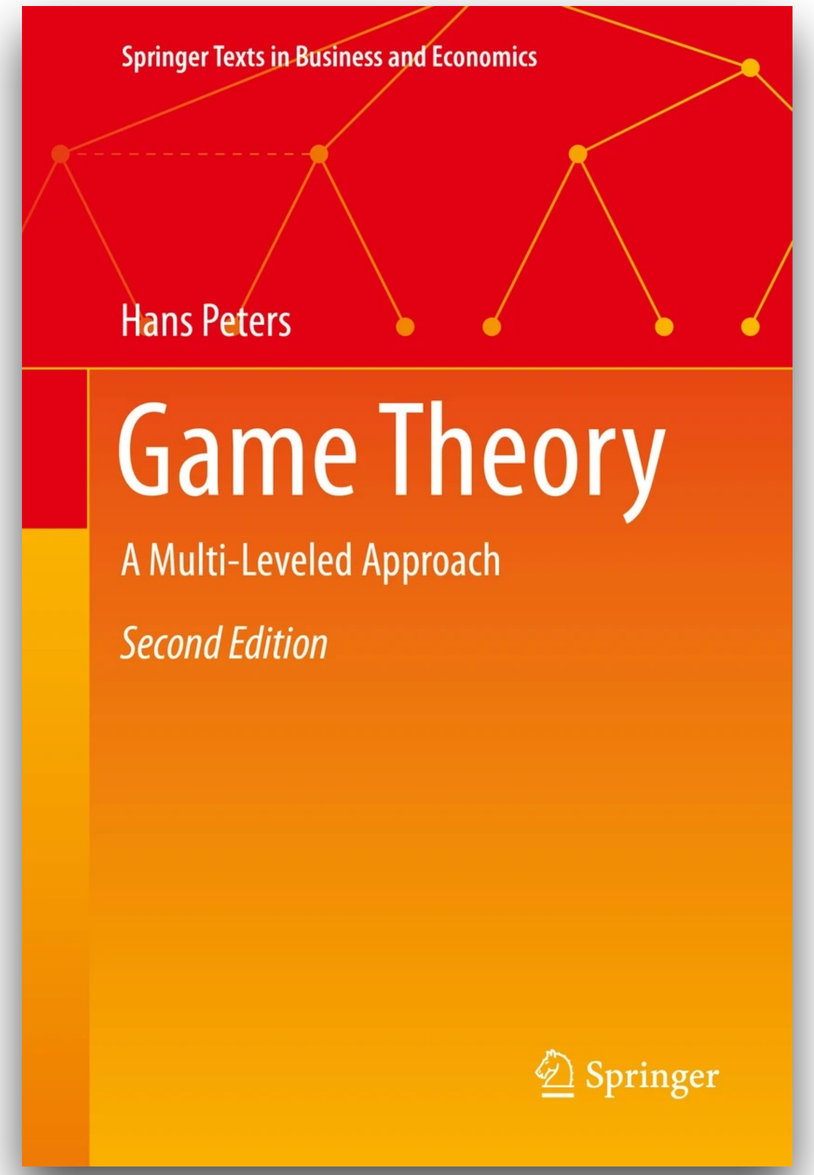
办公地点： 粤海校区汇文楼 1510
丽湖校区守正楼 A 座 3 楼公共办公室

电子邮箱： huangjp@szu

课程主页： https://huangjp.com/GT/

课程基本信息

- 课程网站：`https://huangjp.com/GT/`
- 教学辅助工具：微助教
- 成绩评价：出席情况 10% + 课堂表现 10% + 作业 20% + 期末考核 60%
- 主要参考书：
  Peters, H. (2015). *Game Theory: A Multi-Leveled Approach*, 2nd Edition. Springer.
  https://link.springer.com/book/10.1007/978-3-662-46950-7
  可通过 深圳大学图书馆 > 数据库 > Springer Nature 电子图书 下载 PDF 版本



- 讲义：内容基于 Peters (2015) 的第 1 ～ 6、9、10 章，涵盖非合作博弈与合作博弈的基本内容，可在课程网站查看

课程特点与要求

- 对数学基础的要求不高，认真学习高数和线性代数的同学应该不存在理解障碍。
- 需要一定的经济学知识和洞察力。
- 我们的目的不是学会做题，而是理解方法并能够正确应用。需要有较强的自学能力，愿意主动思考。
- 鼓励在课堂上提问并参与讨论。
- 关于生成式 AI：
  - 不建议作为能辅助学习的工具
  - 禁止在作业和期末报告中使用

学习目标与考核方式间的对应关系

学习目标	考核方式			
	出席	课堂提问	作业	期末考核
Level 1 理解基本概念及思维方式	—	✅	✅	✅
Level 2 会运用所学方法解答教材或教师给出的练习题	—		✅	✅
Level 3 会运用所学方法对实际问题进行建模分析	—			✅

什么是博弈论?

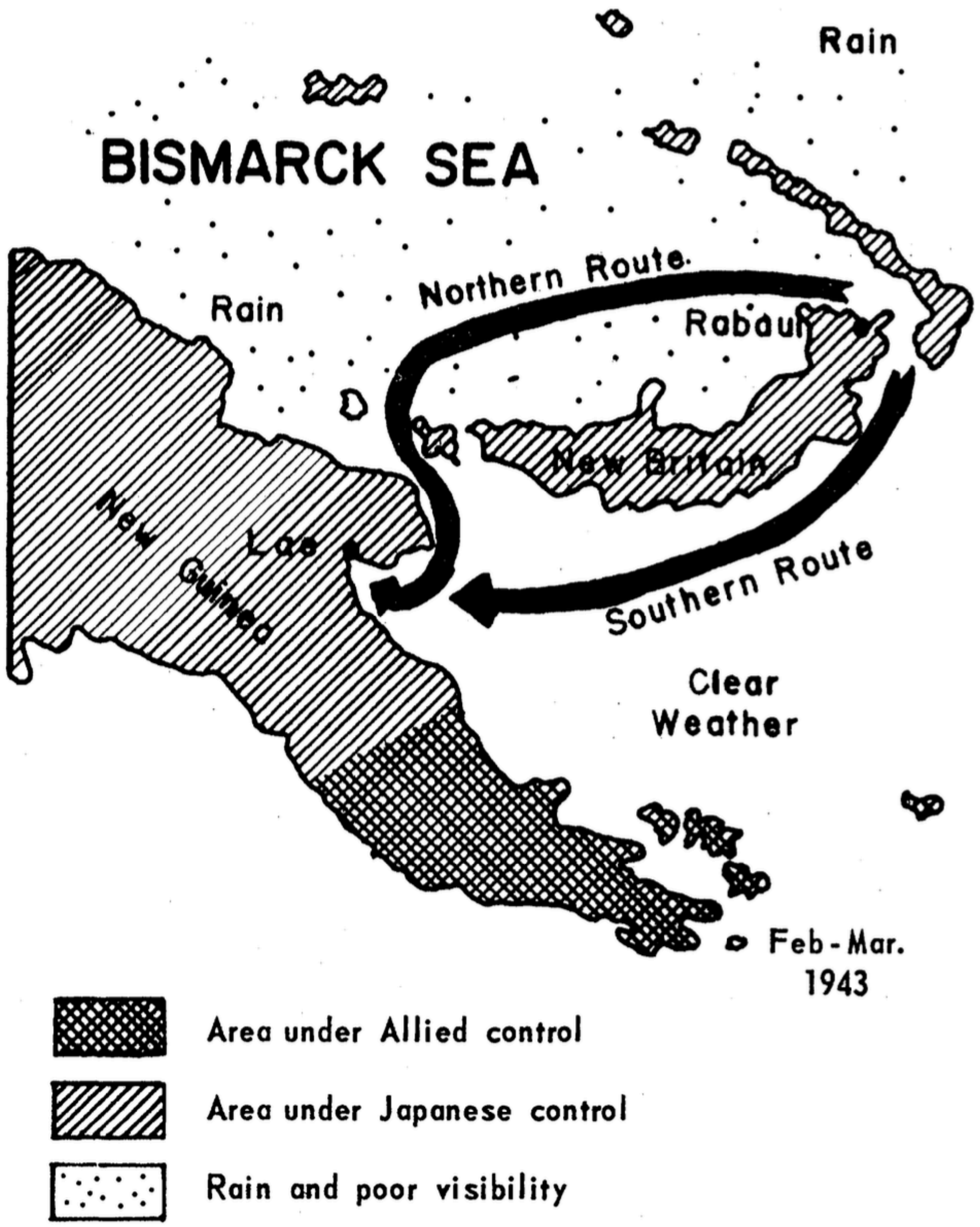
- 博弈论研究的是多人间的竞争与合作
  - 竞争：参与人之间存在潜在的利益冲突，且各自的利益受自身和他人行为的影响 ➡ 非合作博弈
  - 合作：参与人可以通过合作提升集体利益 ➡ 合作博弈
- 目的（博弈的解）
  - 非合作博弈：参与人如何在不同的行动之间作出最佳选择
  - 合作博弈：如何对合作产生的总收益（或总费用）进行分配
- 典型应用方向：
  - 产业组织（由少数企业形成的市场）、数字经济（平台）
  - 机器学习（例如 AlphaGo）
  - 市场设计与匹配（拍卖、器官移植的最优匹配方法等）

一些博弈的例子

零和博弈（zero-sum game）

1. 俾斯麦海海战

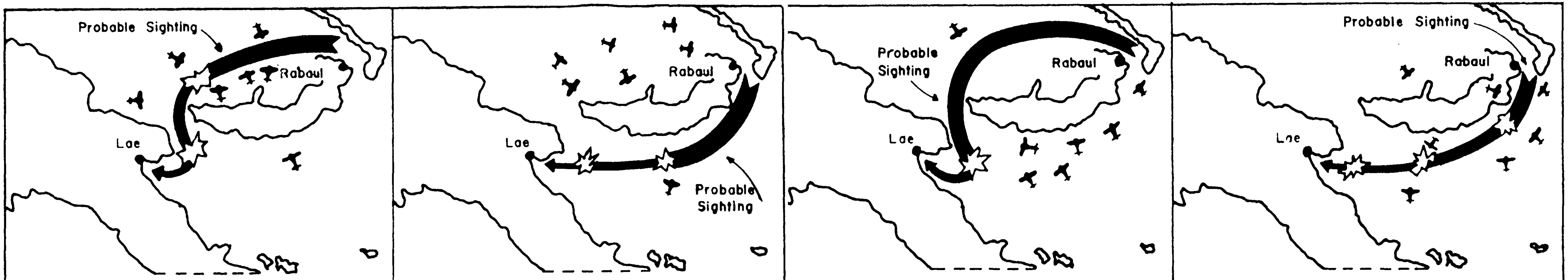
- 1943年的太平洋战争中，日军为争夺对新几内亚岛的控制权，计划从新不列颠岛的拉包尔（Rabaul）运送陆军士兵到新几内亚岛的莱城（Lae）。盟军情报部门事先获取了这一消息，并计划出动轰炸机拦截运输船队。
- 日军有两条行进路线可以选择，分别是北线和南线。两条路线都需要走 3 天。天气预报显示前两天北线海域为雨雾天气，能见度差，南线海域天气持续晴朗。
- 盟军需要选择将主力轰炸机队派往北线或者南线。北线海域因视野不好，预计需要 1 天时间才能发现日军船队的位置。
- 双方都面临同样的决策问题：选择北线还是南线。



图片出处：Haywood Jr., O. G. (1954). Military Decision and Game Theory. Journal of the Operations Research Society of America, 2(4):365-385.

零和博弈（zero-sum game）

1. 俾斯麦海海战



日军：北线
盟军主力：北线

预计轰炸持续时间：2 天
（因天气原因）

日军：南线
盟军主力：北线

预计轰炸持续时间：2 天
（需重新部署轰炸机编队）

日军：北线
盟军主力：南线

预计轰炸持续时间：1 天
（重新部署 + 天气）

日军：南线
盟军主力：南线

预计轰炸持续时间：3 天

图片出处：Haywood Jr., O. G. (1954). Military Decision and Game Theory. Journal of the Operations Research Society of America, 2(4):365-385.

零和博弈（zero-sum game）

1. 俾斯麦海海战

- 这是一个典型的多人决策问题。日军指挥官和盟军指挥官都需要作出自己的选择，但是却无法单独决定战争的结果。日军指挥官希望最大限度的减少损失，而美军指挥官希望最大限度的造成伤害。双方的选择都是**独立**且**同时**进行的。

- 如果我们将持续轰炸天数作为盟军的收益，将该天数的负数作为日军的收益，上述情形可以用一个简单的矩阵表达。

- 右侧的矩阵中只写出了盟军的收益。因为双方的收益之和为零，就没有必要标出日军的收益。

		日军	
		北线	南线
盟军	北线	2	2
	南线	1	3

- 这是一个典型的**二人零和博弈**。例子中的盟军和日军称作**参与人（player）**。盟军为行参与人（row player），因为他可以选择的**行动（action）**体现为矩阵的行。日军为列参与人（column player）。在二人博弈中，我们通常把行参与人称为参与人 1，把列参与人称为参与人 2。

注：也有文献把 player 翻译成 “局中人”。

零和博弈（zero-sum game）

1. 俾斯麦海海战

		日军	
		北线	南线
盟军	北线	2	2
	南线	1	3

- “解”一个博弈意味着对参与人的选择作出合理预测。
- 预测日军指挥官的选择比较容易。对于他而言，无论对手怎样选择，自己选择北线都不会比南线差。我们称北线优于南线，南线劣于北线。日军应当选择北线。
- 盟军指挥官的选择题则相对更加复杂。当对手选择北线时，自己也应该选择北线；而如果对手选择了南线，自己也应该选择南线。这称为**最佳响应（best reply/response）**策略。仅凭这个逻辑是不足以作出判断的，重要的是盟军指挥官也了解对手面临的情况，知道北线是对方的最优选择。因此，此时盟军也应当选择北线。
- （北线，北线）的行动组合是我们对参与人选择的预测，是这个博弈的一个解。而历史上双方的选择也和我们的预测一致，俾斯麦海海战以日军大败而告终。
- 那么日军指挥官的选择错了吗?

零和博弈

2. 猜硬币（matching pennies）

- 两个参与人各有一枚硬币。他们需要同时展示出硬币的一面，如果出现的面相同则参与人 1 会赢得参与人 2 的硬币，反之则参与人 2 赢得参与人 1 的硬币。
- 这个博弈的矩阵表达为右图所示。今后我们将默认参与人 1 为行参与人，矩阵中的数字为参与人 1 的收益。
- 在这个例子里，双方都没有优势策略，因此也没有看起来比较自然的解。
- 解决问题的一个办法是允许参与人依某个概率分布随机选择行动，这被称为**混合策略（mixed strategy）**。混合策略可以理解为其他参与人对自己的选择作出的预测，也可以理解为当博弈反复进行时选择特定行动的频率。

	正	反
正	1	–1
反	–1	1

零和博弈

2. 猜硬币（matching pennies）

	正	反
正	1	–1
反	–1	1

- 令参与人 1 选择正面的概率为 p，参与人 2 选择正面的概率为 q。此时我们可以计算二人的期望收益。
- 参与人 1 的期望收益是

$$\begin{aligned} & p\left[q\cdot 1+(1-q)\cdot(-1)\right]+(1-p)\left[q\cdot(-1)+(1-q)\cdot 1\right] \\ = {} & p(2q-1)-(1-p)(2q-1)=(2p-1)(2q-1) \end{aligned}$$

- 参与人 2 的期望收益是

$$\begin{aligned} & q\left[p\cdot(-1)+(1-p)\cdot 1\right]+(1-q)\left[p\cdot 1+(1-q)\cdot(-1)\right] \\ = {} & q(1-2p)-(1-q)(1-2p)=-(2p-1)(2q-1) \end{aligned}$$

- 一种比较稳妥的选择方式：假设对手会想方设法降低自己的收益，而自己能做的是在此基础上作出最优选择。对于参与人 1 来说这是将最小收益最大化，称作 maximin 策略；而对于参与人 2 来说这是将最大损失最小化，称作 minimax 策略。
- 策略组合 $(p=0{,}5,\ q=0.5)$ 是唯一的 maximin 策略和 minimax 策略的组合，因此可以作为此博弈的解。

非零和博弈

1. 囚徒困境（prisoner’s dilemma）

- 这是一个非常著名的例子。两个人因共同犯罪的嫌疑被捕，但是警方的客观证据不足，需要嫌犯口供才能起诉。因此警方将二人分开审理。
- 两个嫌疑人都有两种选择，一是和同伙串供，通常称为（与同伙）合作（cooperation, C），二是指证同伙犯罪，通常称为背叛（defect, D）。被指证的嫌疑人将被起诉 10 年徒刑，而没有被指证的则只能被起诉 1 年。指证同伙可以因配合警方而获得 1 年减刑。
- 此博弈可以用右侧的矩阵表达。与之前的例子不同的是，此时双方的收益之和不再是零。因此，矩阵的每一个元素都包含两个数字，第一个是参与人 1 的收益，第二个是参与人 2 的收益。

	C	D
C	–1, –1	–10, 0
D	0, –10	–9, –9

非零和博弈

1. 囚徒困境（prisoner’s dilemma）

	C	D
C	−1, −1	−10, 0
D	0, −10	−9, −9

- 对每个参与人来说，无论对方如何选择，自己选择背叛（D）都会带来更高的收益。因此 (D, D) 是此博弈的解。
- 然而，很明显 (C, C) 比 (D, D) 给双方带来的收益都会更高。这说明 (D, D) 的收益不是帕累托最优的。双方虽然都进行了合理的选择，但结果却并不令人满意，因此称之为“困境”。
- (D, D) 的组合同时也是**纳什均衡（Nash equilibrium）**。也就是说，在固定对方策略不变的前提下，双方都无法通过改变策略而提高自己的收益。
- 与囚徒困境博弈类似的例子还有公地悲剧（tragedy of the commons）。例如，几个牧场主共享一块天然草场。每个牧场主都想通过增加 🐑 的数量来提高自己的收益，但是成本（草）确是由大家共同承担的。最极端的情况是草被吃完了，所有牧场主的收益都归零。这里也出现了个人理性决策无法达成社会最优状态的情况。

非零和博弈

2. 性别战（battle of sexes）

- 有一对情侣准备出去约会。他们可以选择去看足球赛或者看芭蕾舞。由于双方的喜好不同，他们只事先约定了约会时间，并没有约定地点。假设在约会当天，出于某些原因（例如其中一人的手机没电了）他们无法联络沟通，只能自己从足球和芭蕾之间作出选择。
- 假设男人更喜欢看足球赛，女人更喜欢看芭蕾舞。两人都希望和对方一起度过这段时间，因此孤身一人时的收益（或效用）为零。我们设男人为参与人 1 （行参与人）。
- 此博弈的矩阵表达如右图所示。
- 是否存在优势策略?
- 是否存在纳什均衡?

	足球	芭蕾
足球	2, 1	0, 0
芭蕾	0, 0	1, 2

非零和博弈

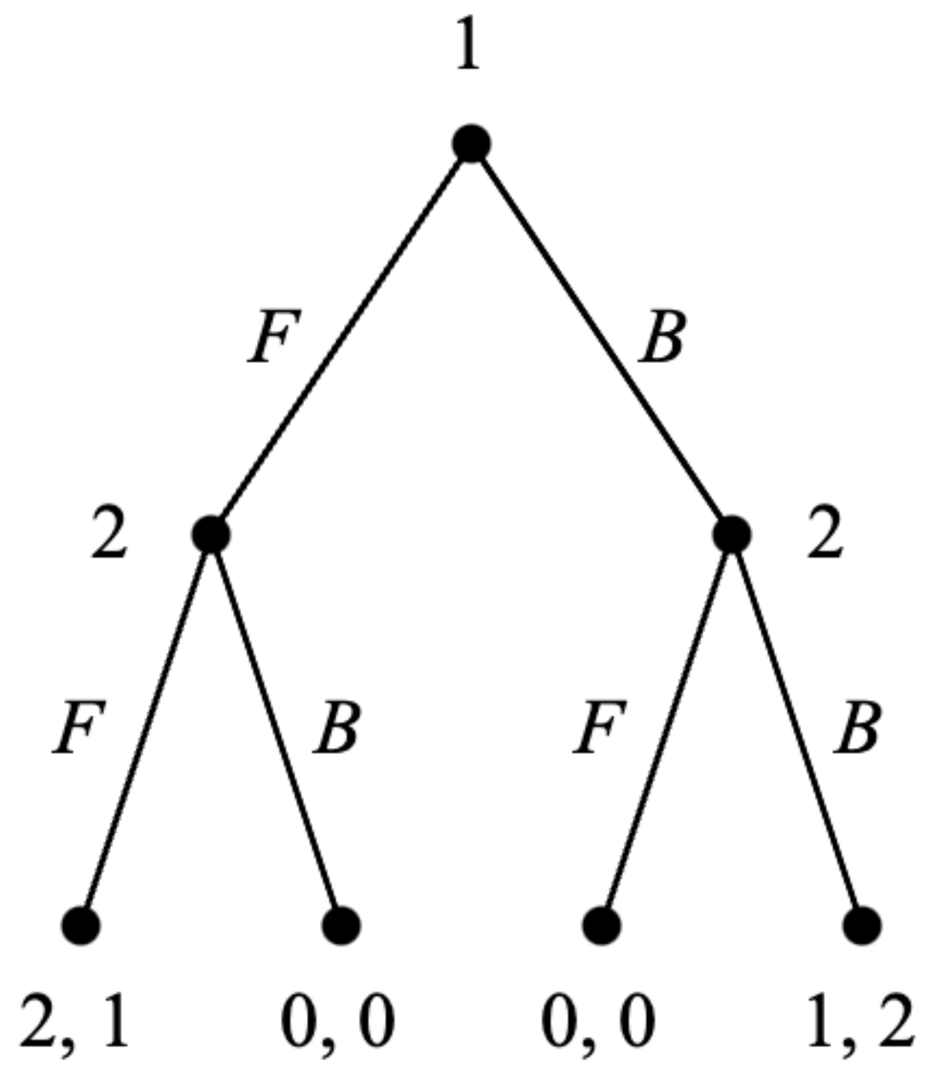
3. 古诺博弈（a Cournot game）

- 这是非常经典的寡头市场模型。假设两家公司生产同一种产品，该产品的市场价格由总产量 Q 决定，即 $p = \max\{1 - Q, 0\}$。假设生产该产品不需要成本。
- 参与人 i $(i = 1, 2)$ 的利润函数是 $K_i(q_1, q_2) = q_i \cdot \max\{1 - q_1 - q_2, 0\}$
- 双方需要决定自己的产量 $q_i \in \mathbb{R}$。
- 双方的行动集（action set）都是无限的，因此无法用矩阵表达。但是我们依然可以想办法计算纳什均衡。
- 确认 $(q_1, q_2) = (1/3, 1/3)$ 是纳什均衡，即双方的策略互为最佳响应。
- 确认这个纳什均衡的收益不是帕累托最优。你更找到它的帕累托改进吗?

扩展式博弈

博弈树

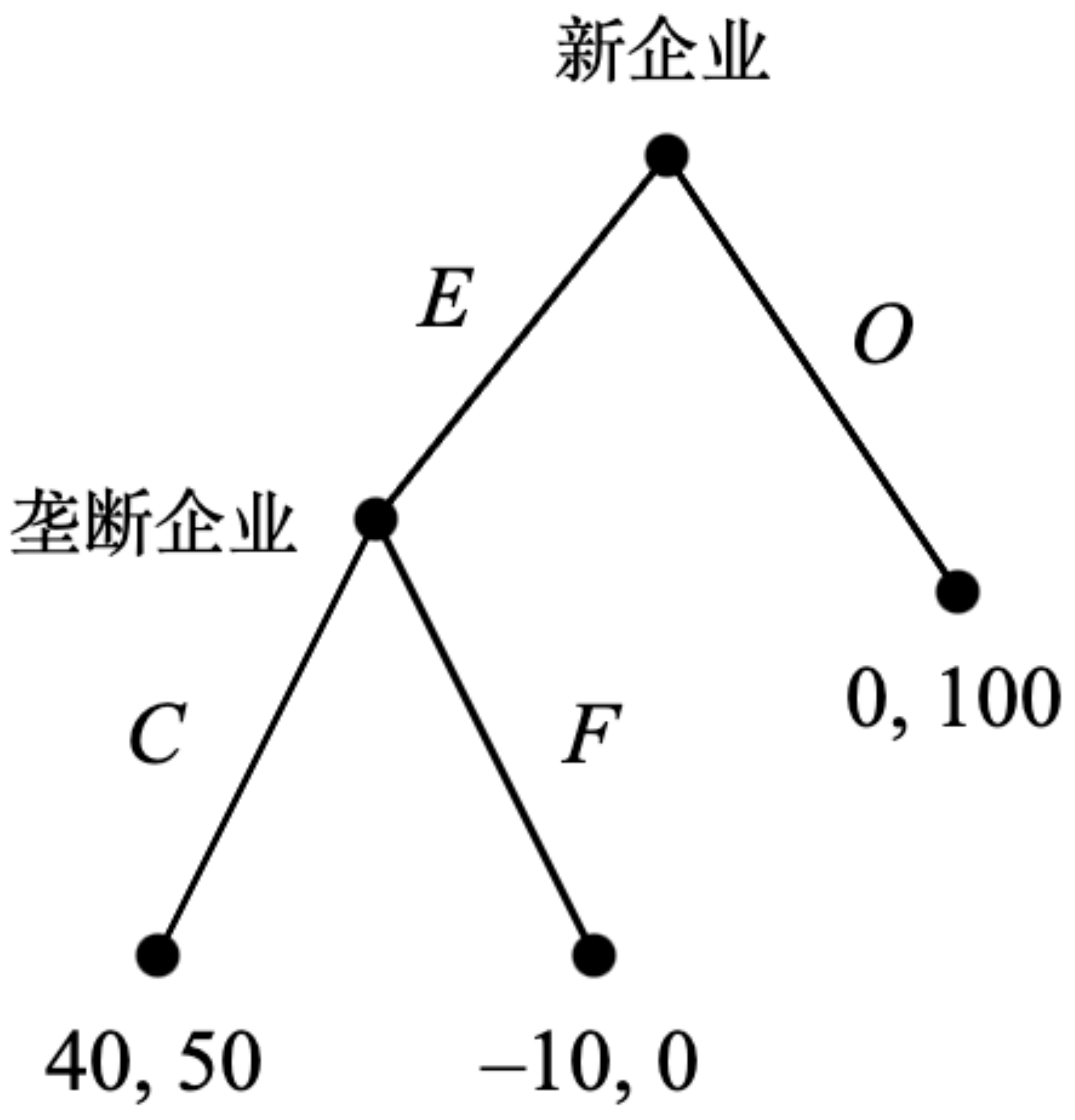
- 前面的例子都是参与人同时进行决策。很多时候，参与人是按照一定的顺序依次进行决策，这类博弈称为**序贯博弈（sequential game）**。我们用扩展式博弈（extensive form game）来描述这类情形。
- 扩展式博弈通常以**博弈树（game tree）**的形式表达。
- 树是图论中的概念，是图（graph）的一种。树由节点（node）和边（edge）组成，但不能包含环（cycle）。节点中有一个根节点（root）和一些终点（end node，或称叶 leaf）。终点以外的节点都是决策节点，即某一参与人需要进行决策。
- 右图是参与人 1 首先行动的序贯性别战博弈的博弈树。



扩展式博弈

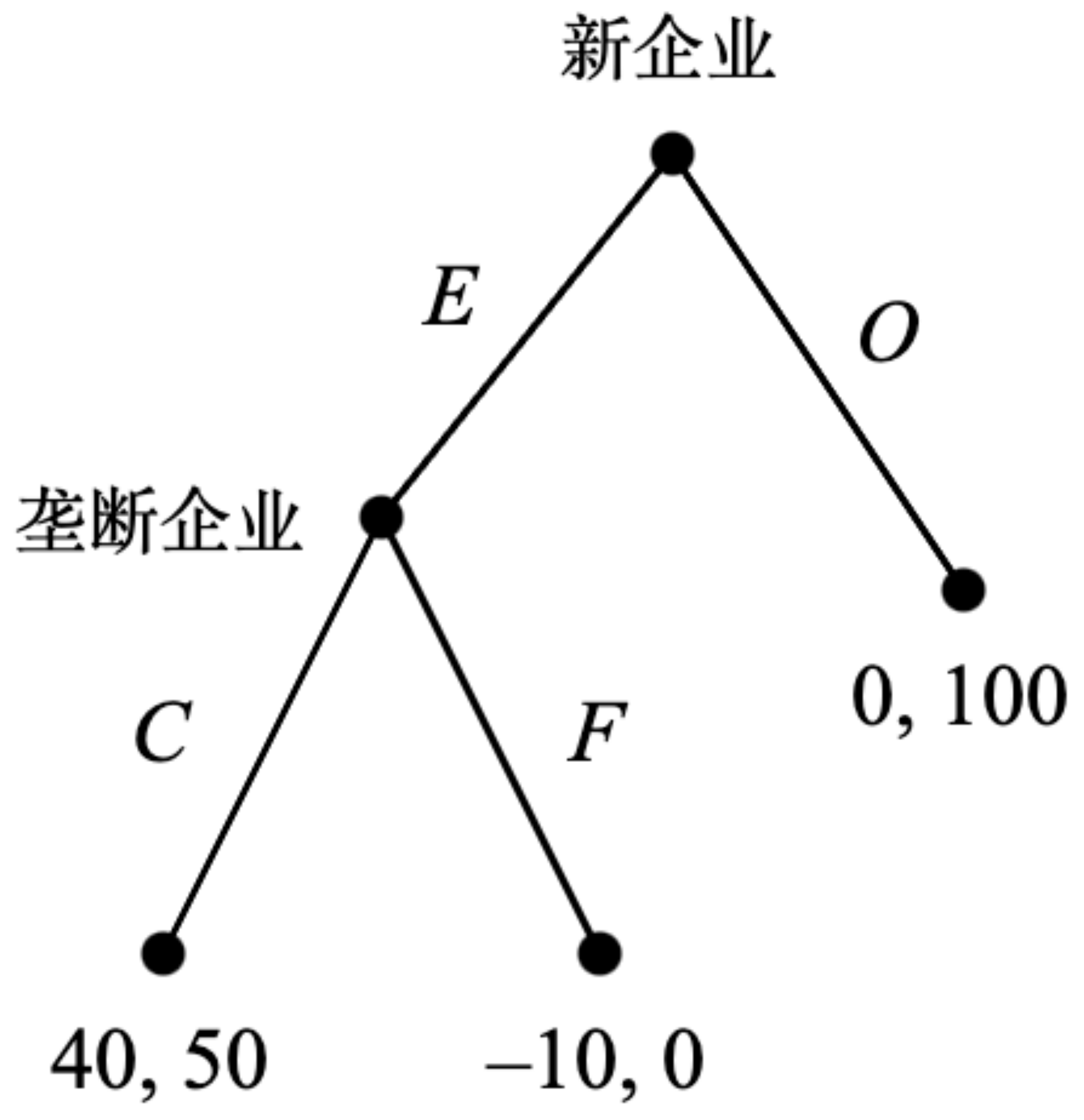
1. 市场进入与威慑（entry deterrence）

- 在产业组织论中有一个古老的问题，就是垄断企业如何应对新企业的出现。常见的策略是以降低价格作为防御手段阻止新企业进入市场。
- 在经典的市场进入模型中，新企业首先选择是否进入该市场，之后垄断企业选择是否开始价格战。假设该市场在垄断价格下的利润为 100，在价格战下为 0。如果垄断企业选择不进行价格战（即共谋，collusion），则双方平分垄断价格下的利润。新企业进入市场的成本是 10。
- 博弈树见右图。



扩展式博弈

1. 市场进入与威慑（entry deterrence）



- 此类博弈可以通过**逆向归纳法（backward induction）**来解。
- 我们从最后一个进行决策的参与人开始，在这个博弈中就是垄断企业。根据收益的比较可知，在新企业选择进入市场时，垄断企业选择共谋（C）要优于选择价格战（F）。
- 新企业也知道垄断企业的判断逻辑，因此也认同上面的结果并将其纳入自己的决策过程。当他进行选择时，他可以认为进入市场（E）带来的收益是 40。因此进入市场（E）是最好的选择。
- 综上所述，逆向归纳法的结果是策略组合 (E, C)，这也称为**子博弈完美纳什均衡（subgame perfect Nash equilibrium）**。

扩展式博弈

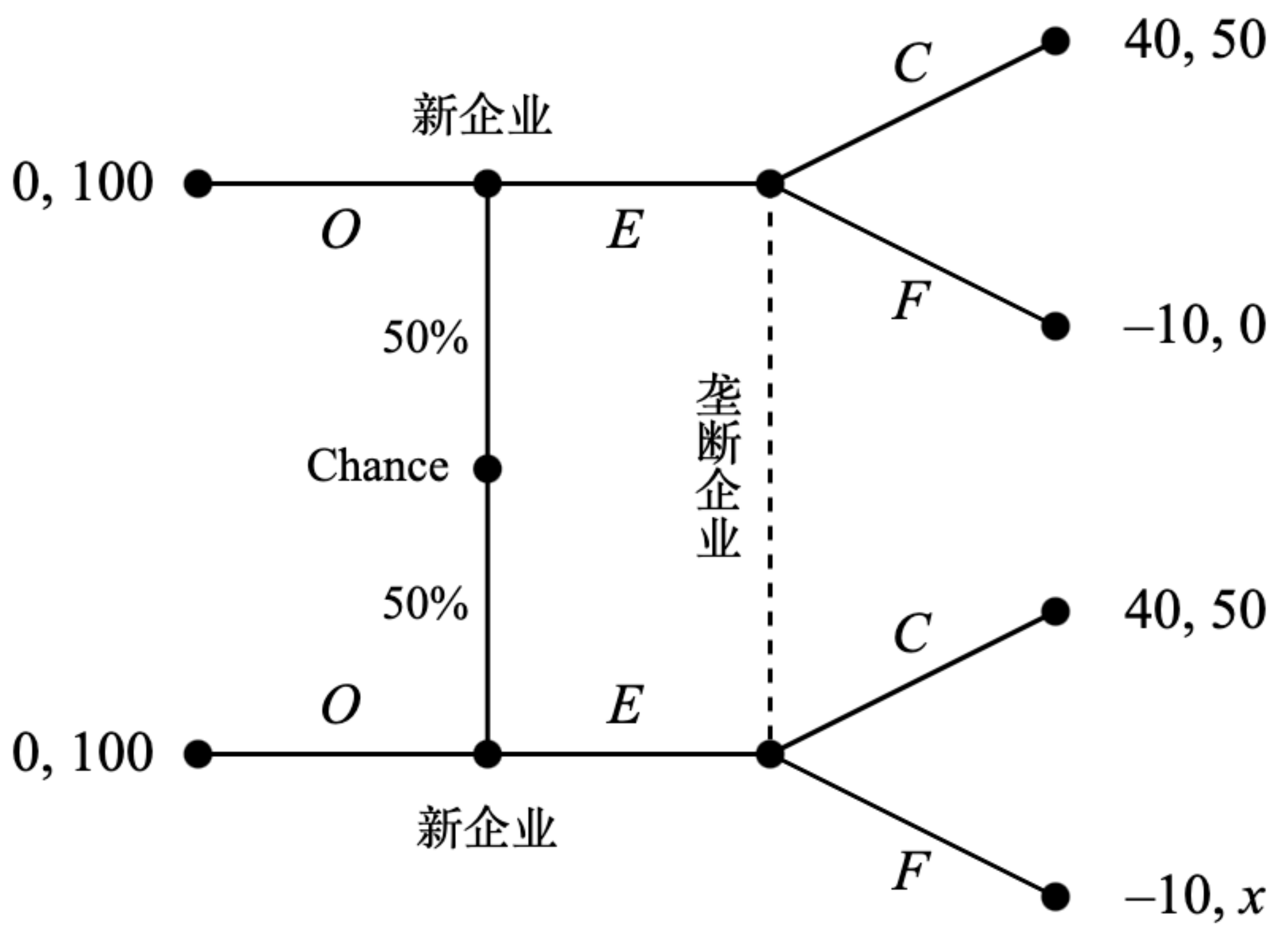
2. 存在不完全信息的市场进入与威慑（entry deterrence with incomplete info.）

- 现在我们假设垄断企业在进行价格战时有 50% 的概率可以获得正收益 x。这可能因为他对新企业了解得不够充分，而新企业可能没有足够能力将价格战进行到底。
- 新企业依然首先进行选择，并且知道垄断企业在选择 F 时的收益是 0 还是 x。因为他非常了解自身的情况。换句话说，这是新企业的私密信息（private information），而垄断企业只掌握了不完全信息（incomplete information）。
- 为了描述这种情形，我们在博弈树中加入随机选择（chance move）节点。也可以理解为决策系统外部存在一个特殊的决策者（称为 nature），它的作用是随机地选择参与人的不同类型（例如新企业的能力），但自身不会产生收益。

扩展式博弈

2. 存在不完全信息的市场进入与威慑（entry deterrence with incomplete info.）

- 博弈以随机选择开始，之后是新企业进行决策，最后是垄断企业。
- 为了表达垄断企业无法区分不同类型的新企业，我们将垄断企业的两个决策节点用虚线连接。以虚线连接在一起的决策节点的集合称为**信息集（information set）**。出于信息集上的决策者作出的选择适用于所有该信息集内的节点。
- 适用于这类博弈的解的概念是**完美贝叶斯均衡（perfect Bayesian equilibrium）**。我们将在后面的课程中介绍。



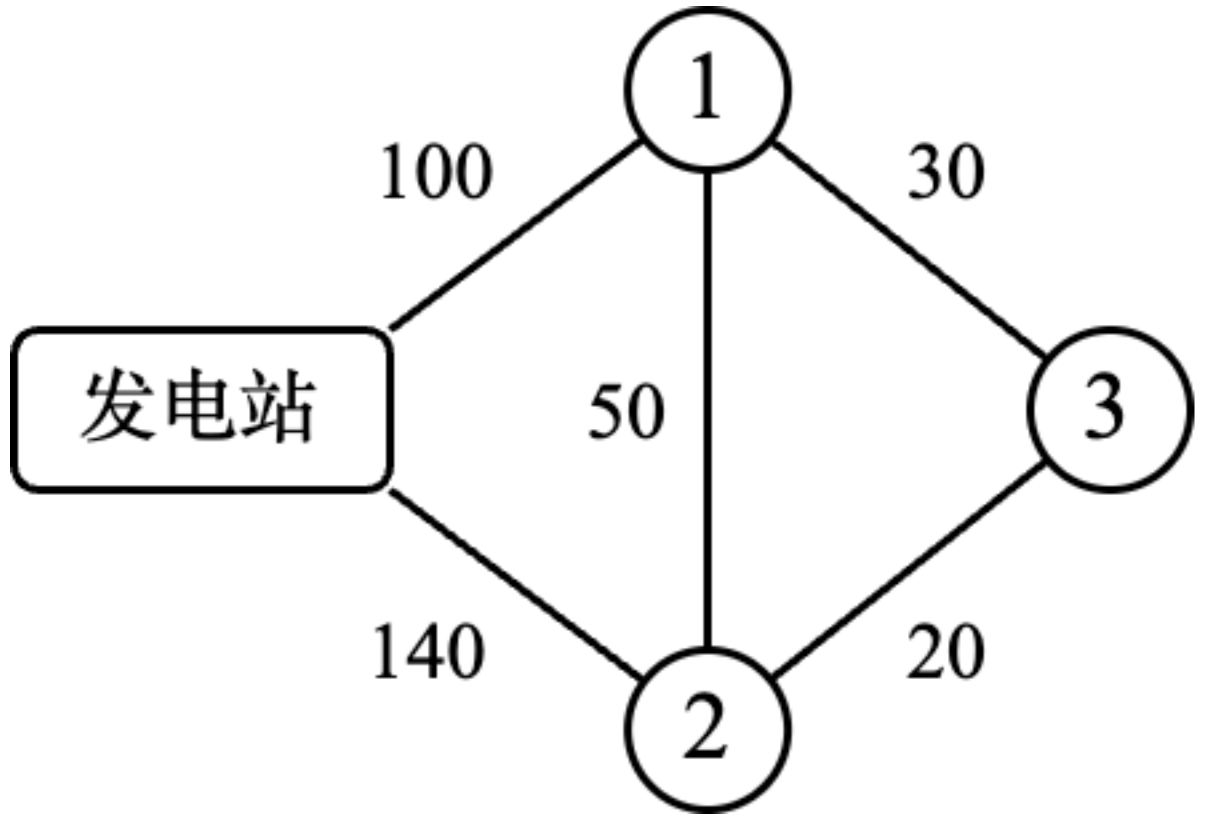
合作博弈

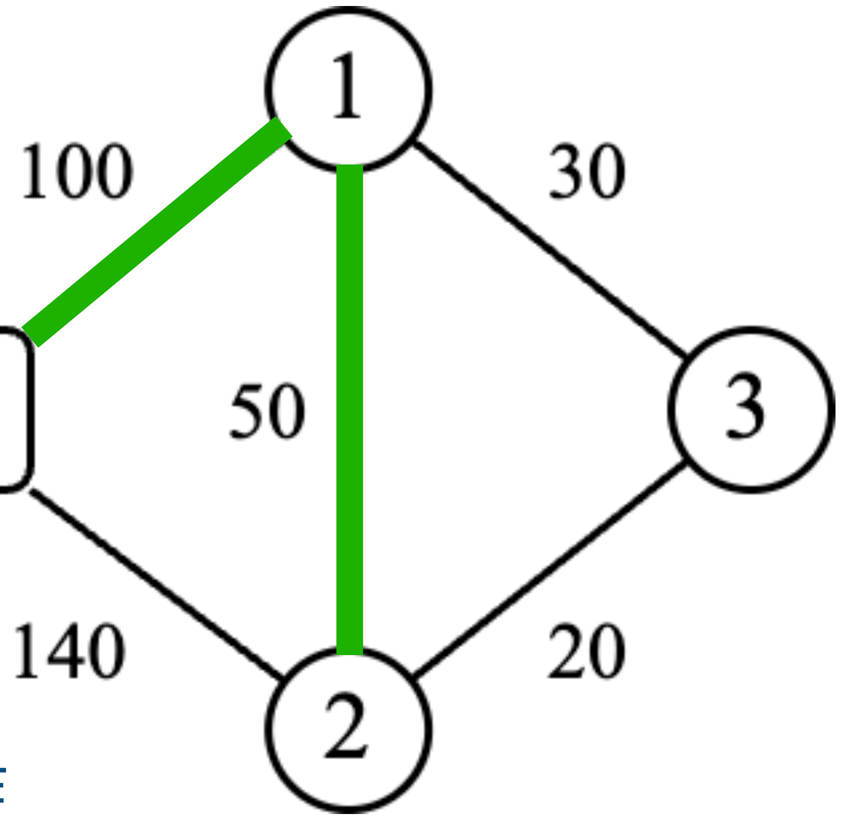
- 前面介绍的都是**非合作博弈（non-cooperative game）**，适用于参与人之间存在利益冲突且无法达成合作协议的情况。
- **合作博弈（cooperative game）**假设参与人之间可以形成有约束力的约定，从而通过协商作出最优决策。因此，我们的关注点不再是参与人选择什么策略，而是参与人之间是否能够形成**联盟（coalition）**，以及如何分配联盟产生的共同收益或成本。
- 由于是否加入联盟也取决于自己可以从该联盟中获得多少收益，因此合作博弈主要讨论联盟成员间的收益分配方法。
- 有时非合作博弈与合作博弈之间的区别并不是非常清晰的，很多应用问题也可以从非合作和合作两个角度进行分析。最容易区分两者的是建模方法。

合作博弈

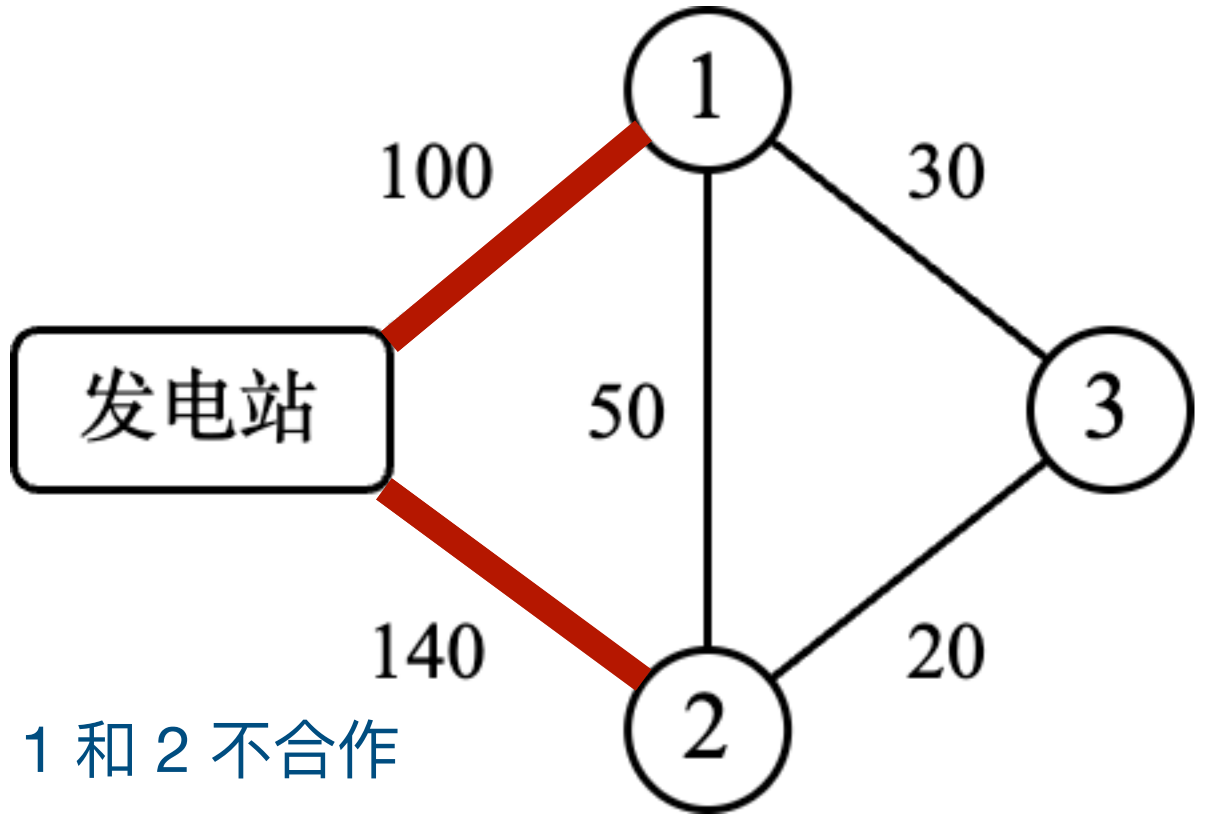
三个城市间的合作

- 城市 1、2、3 共享一座发电站。从发电站到城市间的输电网络和距离（即输电成本）如右图所示。
- 城市间可以通过合作减少输电成本。例如城市 1 和 2 合作时的总成本是 150，而不合作时的总成本是 240。
- 合作博弈的定义包含以下要素：
    - 参与人集合：$N = \{1,2,3\}$
    - 联盟的价值函数：$v(S)$, $S \subseteq N$
      这里我们定义 v 为联盟成员在不合作时的总成本与合作时的总称本之差
    - (N, v) 定义了一个合作博弈





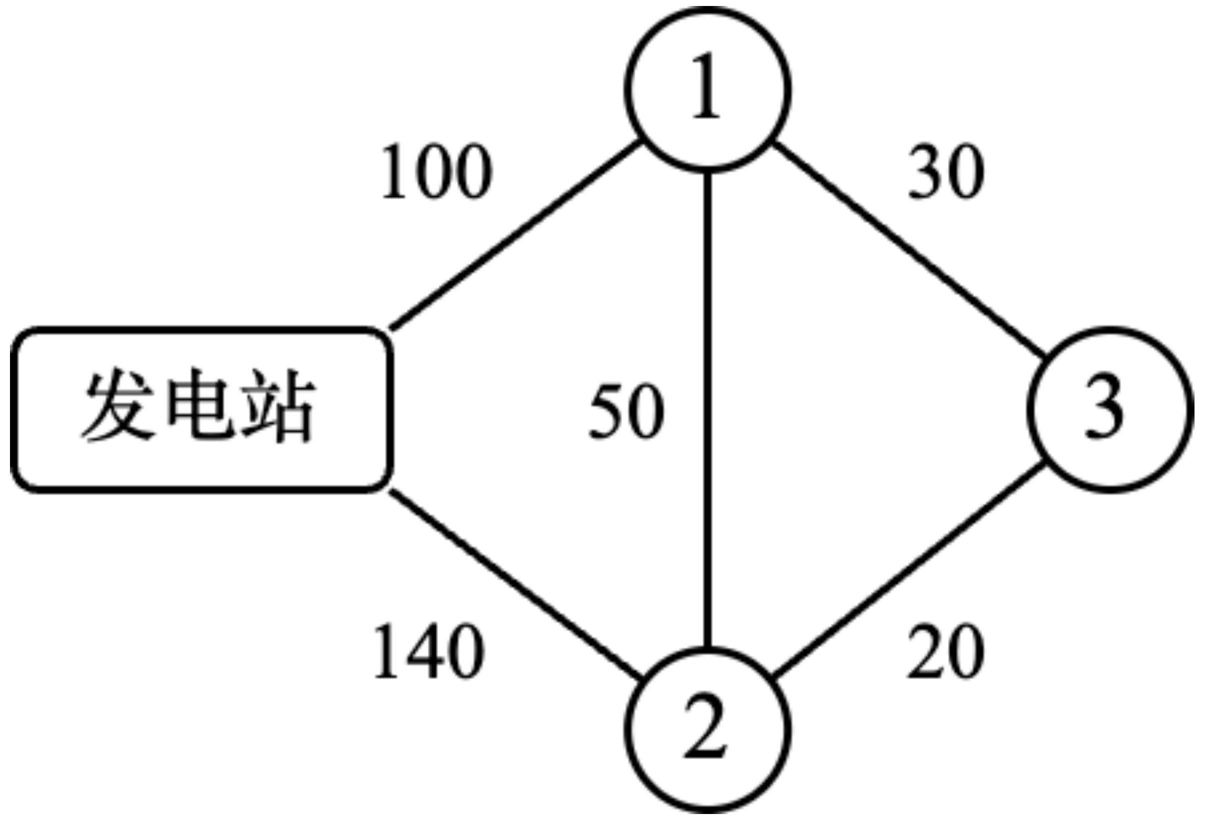
1 和 2 合作



1 和 2 不合作

合作博弈

三个城市间的合作



- 当参与人较少时可以列出所有联盟的价值，如下表所示

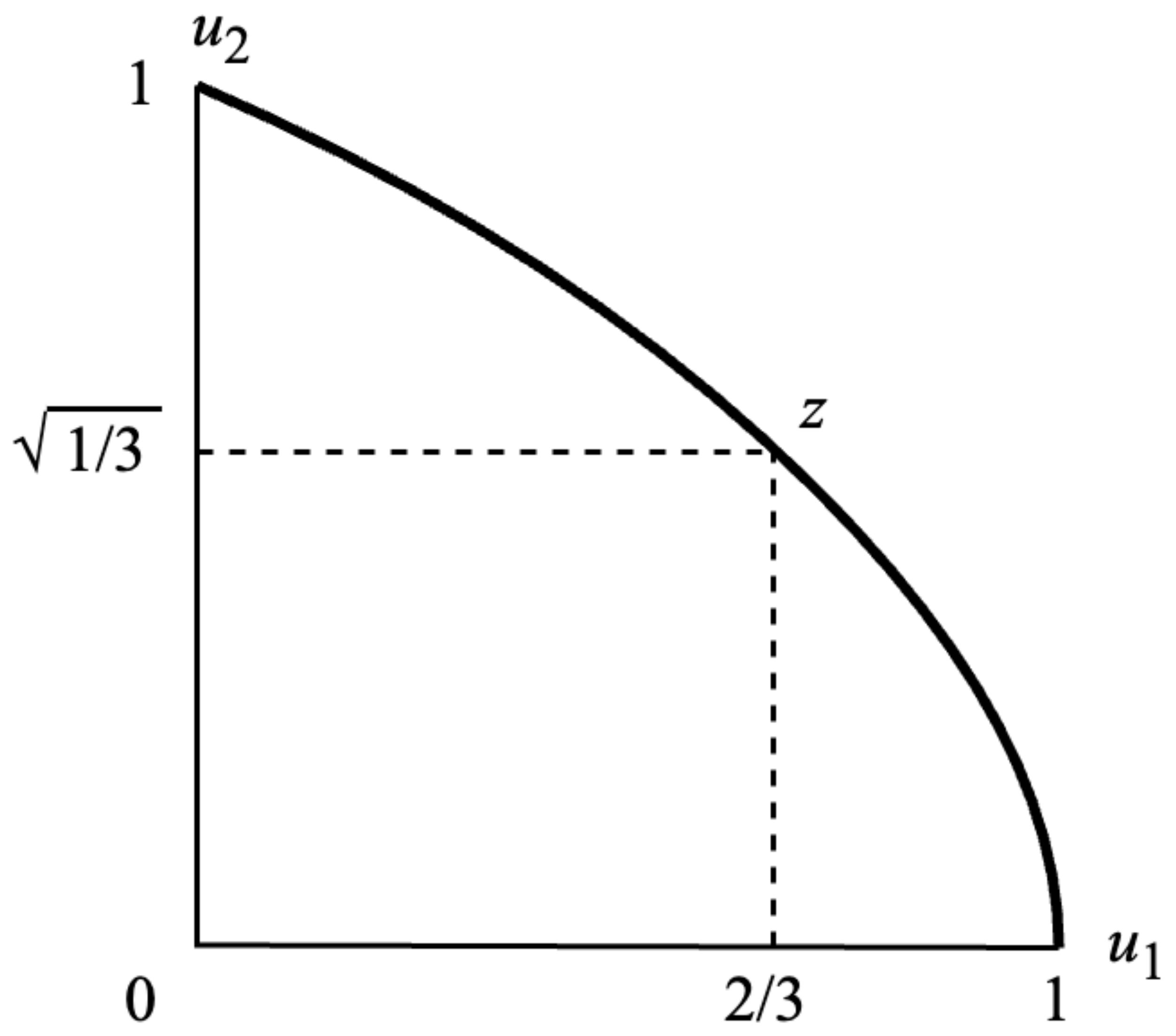
S	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$
$c(S)$	100	140	130	150	130	150	150
$v(S)$	0	0	0	90	100	120	220

- 合作博弈的解是大联盟（grand coalition，即 N）中的一种分配方案 $\boldsymbol{x} = (x_1, x_2, x_3)$，该方案需满足 $x_1 + x_2 + x_3 = v(N) = 220$ 这一基本条件。
- 常用的解包括**核（core）**、**Shapley 值（Shapley value）**和 **nucleolus**。核是集合解，两外两个是单值解。
- 此博弈中，Shapley 值是 $(65, 75, 80)$，nucleolus 是 $(56\frac{2}{3}, 76\frac{2}{3}, 86\frac{2}{3})$。
- 如何判断哪一个解更好呢？学术界通常采用公理化的方式对解进行再定义和解释。

讨价还价博弈

分蛋糕

- 蛋糕指代一单位的完美可分割物品，类似的还有酒类、仓库面积等。
- 考虑两个人分一块蛋糕的情形。可能的结果有两种：一是两人同意以 (α, β) 的比例切蛋糕；二是两人无法达成共识，因此什么也分不到。
- 我们用效用函数表示参与人对蛋糕的偏好。设参与人 1 的效用函数为 $u_1(\alpha) = \alpha$，参与人 2 的效用函数为 $u_2(\beta) = \sqrt{\beta}$。此时，分配 $(\alpha, 1 - \alpha)$ 对应的效用为 $(\alpha, \sqrt{1 - \alpha})$。
- 当 $0 \leq \alpha \leq 1$ 时可以得到有效的效用组合。这对应右图中的粗实线。两人的分配比例之和也可以小于 1，这对应粗实线和两个坐标轴之间的部分。



讨价还价博弈

分蛋糕

- 一个基于合作博弈的解是**纳什谈判解（Nash bargaining solution）**。它是使两人效用之积最大的分配，即

$$\max_{0\le\alpha\le1} \alpha\sqrt{1-\alpha}$$

也就是 $\alpha = 2/3$。

- 为什么不用效用之和的最大化分配呢？纳什的公理化分析或许可以回答这个问题。
- 我们也可以用非合作博弈的方法讨论这个问题。例如鲁宾斯坦的交替报价模型。