博弈论与信息经济学

4. 完全信息动态博弈（一）

深圳大学经济学院 会计学学术学位硕士研究生 专业选修课（2023-2024）
主讲：黄嘉平　中国经济特区研究中心讲师　工学博士 经济学博士
办公室：粤海校区汇文楼1510　　　Email: huangjp@szu.edu.cn

基本概念

扩展式博弈

Extensive-form game

- 扩展式博弈是在标准式博弈的基础上加入了参与人的行动顺序的博弈
- 扩展式博弈和标准式博弈都包含的要素：
  1. 参与人集合 N
  2. 支付函数 $\{v_i(\cdot)\}_{i \in N}$
- 扩展式博弈包含的独特要素：
  3. 参与人的行动顺序
  4. 参与人能够行动时的行动集合（行动集合并不一定是固定的）
  5. 参与人能够行动时所具有的知识（例如，我是否知道在我之前行动的参与人的策略选择）
  6. 外生事件的概率分布（存在不确定性时，“自然”如何做出选择）
- 我们假设上述 1–6 为共同知识

博弈树

Game tree

- 描述参与人行动顺序和行动集合的一种简洁方法是利用决策树，这里成为博弈树（game tree）
- 右上的博弈树描述了一种信任博弈：
  假设参与人 1 和 2 分别为闲鱼平台的上的买家和买家。买家首先选择是否信任卖家（T为信任，即代表下单），之后卖家选择是否诚信交易（C为诚信交易）

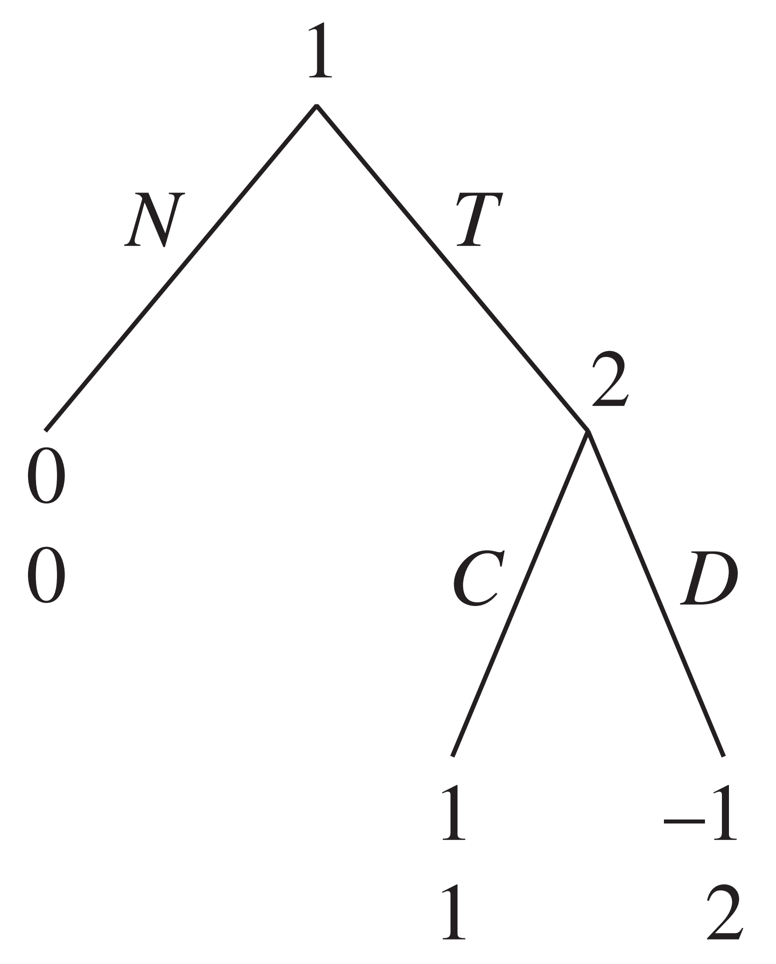
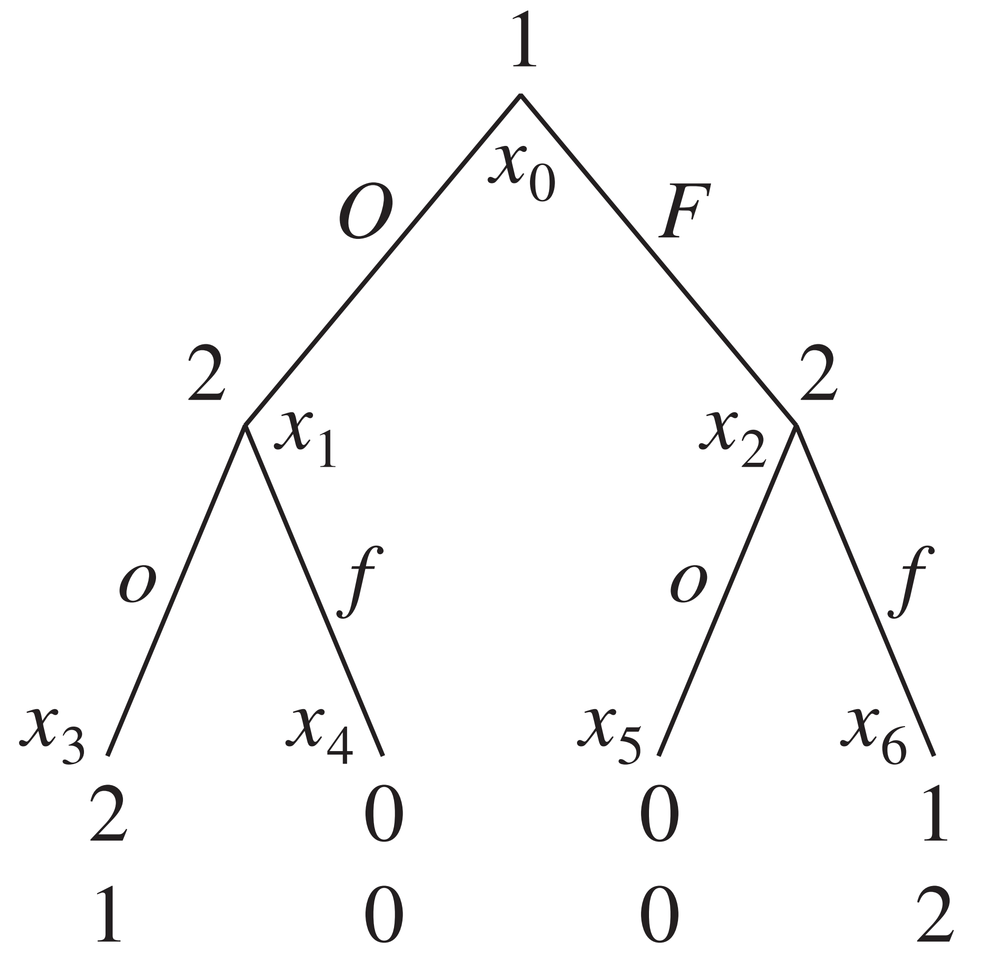


FIGURE 7.1 A trust game.

博弈树包括以下要素

- 节点（node）集合：X
- 节点间的先行（precedence）关系：$x \in X, x' \in X$，则 $x > x'$ 代表 x 先行于 x'
  - 每个节点只有一个直接先行节点（predecessor）
  - 先行关系是<u>可传递的</u>（$x > x', x' > x'' \Rightarrow x > x''$）、<u>非对称的</u>（$x > x' \Rightarrow \neg(x' > x)$）、<u>不完全的</u>（并不是任意一对节点都可以排序）
  - 存在一个根（root）节点 x_0，它是所有其他节点的先行节点
  - 不先行于任何其他节点的节点称为终点（terminal node），终点的集合记作 $Z \subset X$
- 每个终点对应一个结果（参与人的回报组合）
- 每个非终点节点 x 对应一个参与人 $i(x)$ 和行动集合 $A_i(x)$，或者对应“自然”



Sequential-move BoS

- 参与人集合 $N = \{1,2,3,4\}$
- 参与人 3 无法选择行动，因此称为虚拟参与人
- 终点集合为 $Z = \{x_4, x_5, x_6, x_7, x_8\}$
- 支付函数可以表达为 $v_i(x),\ x \in Z$

  例如：$v_1(x_5) = 1$，$v_3(x_8) = 8$

- 每个节点对应的参与人

  $$i(x_0) = 1,\ i(x_1) = 2,\ i(x_3) = i(x_4) = 4$$

- 至此，我们并没有明确参与人行动时所具有的知识。例如参与人 4 在节点 x_2 时是否知道先行参与人 2 在节点 x_1 的决策结果?

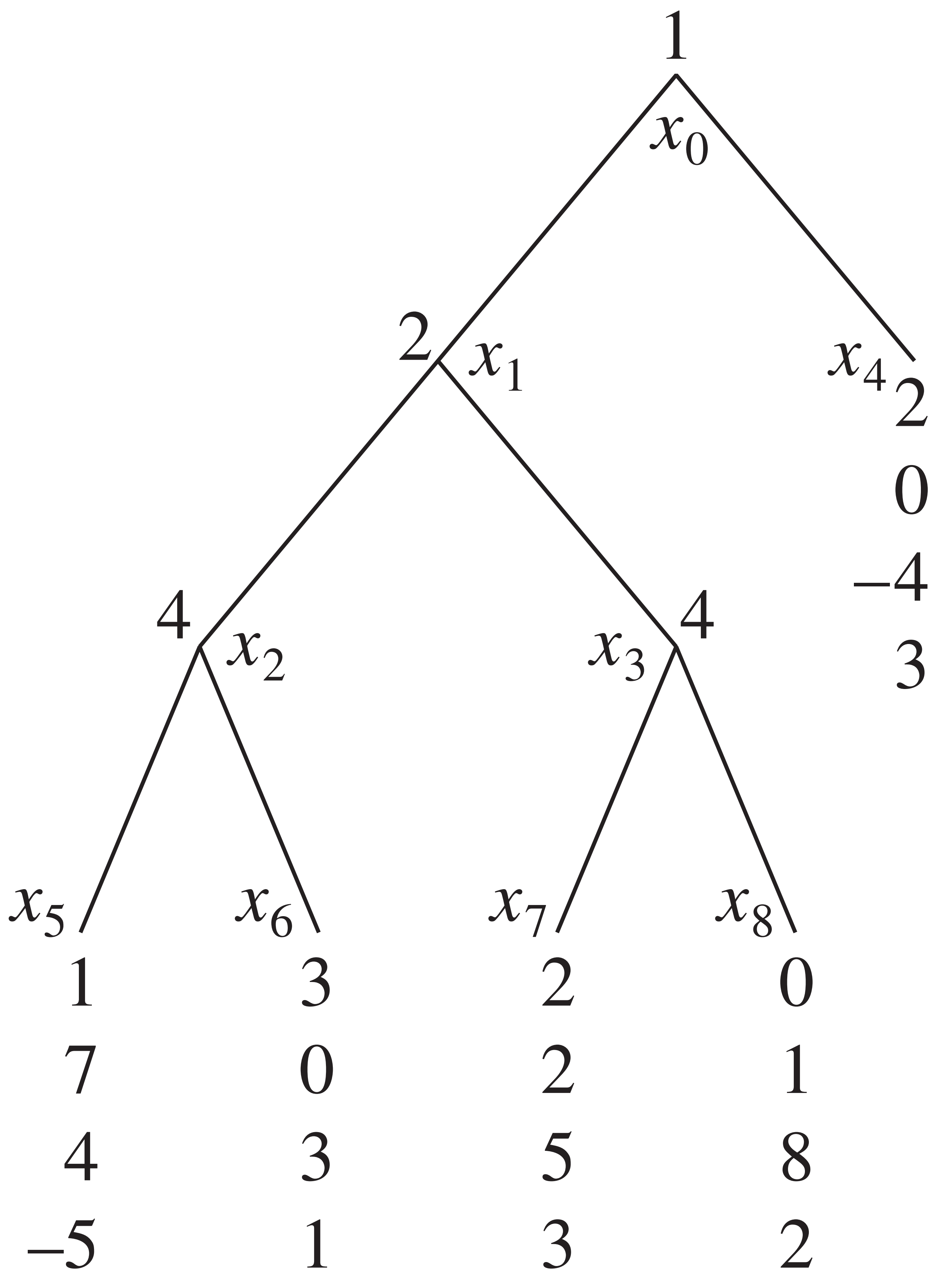


FIGURE 7.3 A game tree with a “dummy” player.

信息集

Information set

- 参与人如果对博弈进程了解得非常详细，代表他知道自己行动时位于博弈树中的哪个节点上。反之，他将无法判断自己所处的位置
- 我们用信息集来描述参与人所具有的知识

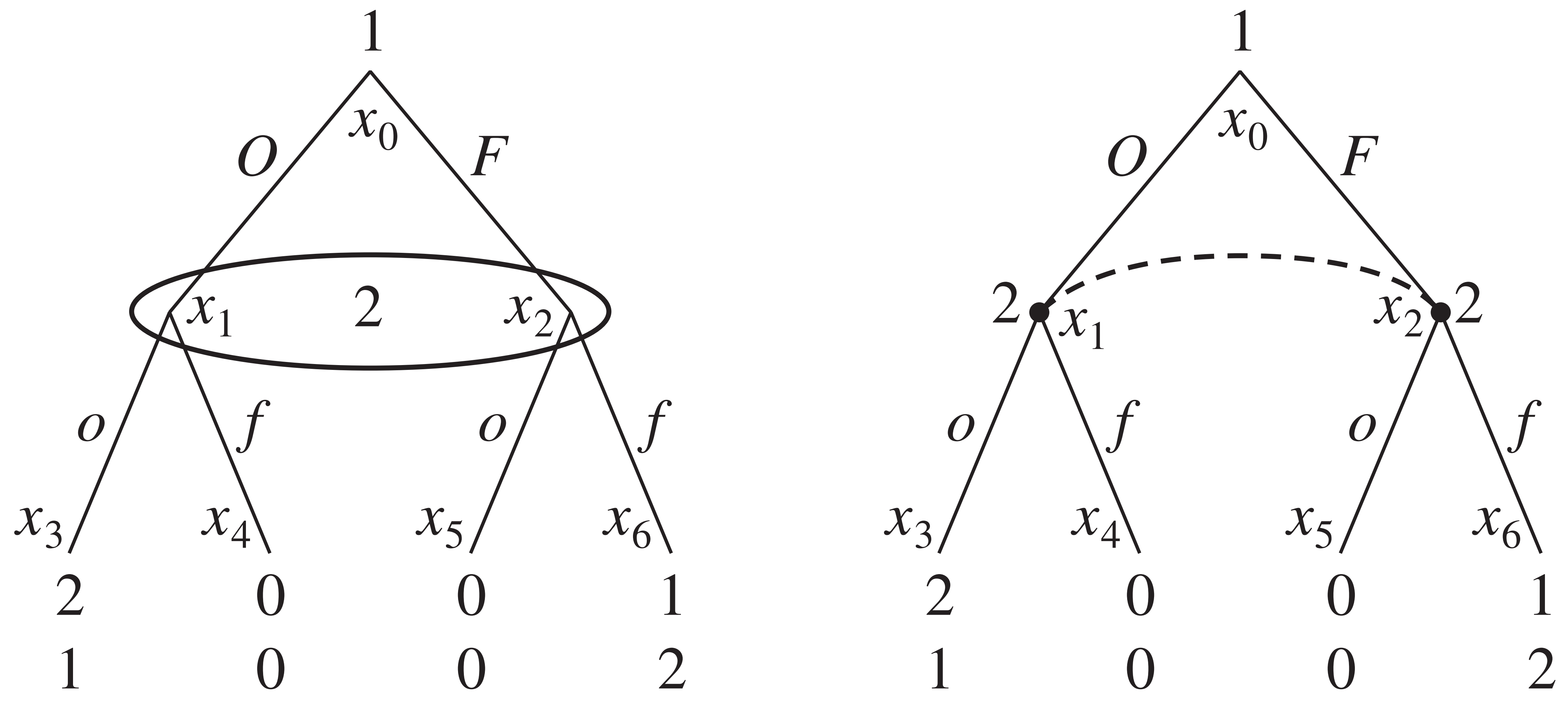


FIGURE 7.4 The simultaneous-move Battle of the Sexes game.

信息集（information set）：参与人 i 拥有信息集的族 H_i，其中的信息集将 i 的行动节点进行分割。信息集 $h \in H_i$ 满足下列性质：

1. 如果 h 只包含一个节点 x，即 $h = \{x\}$，则在 x 行动的参与人 i 知道自己在 x 处
2. 如果 $x \neq x'$, $x \in h$, $x' \in h$，则在 x 行动的参与人 i 不知道自己是在 x 处还是在 x' 处
3. 如果 $x \neq x'$, $x \in h$, $x' \in h$，则 $A_i(x) = A_i(x')$

- 右上的例子展示了两种表达信息集的方法

静态博弈的扩展式表达

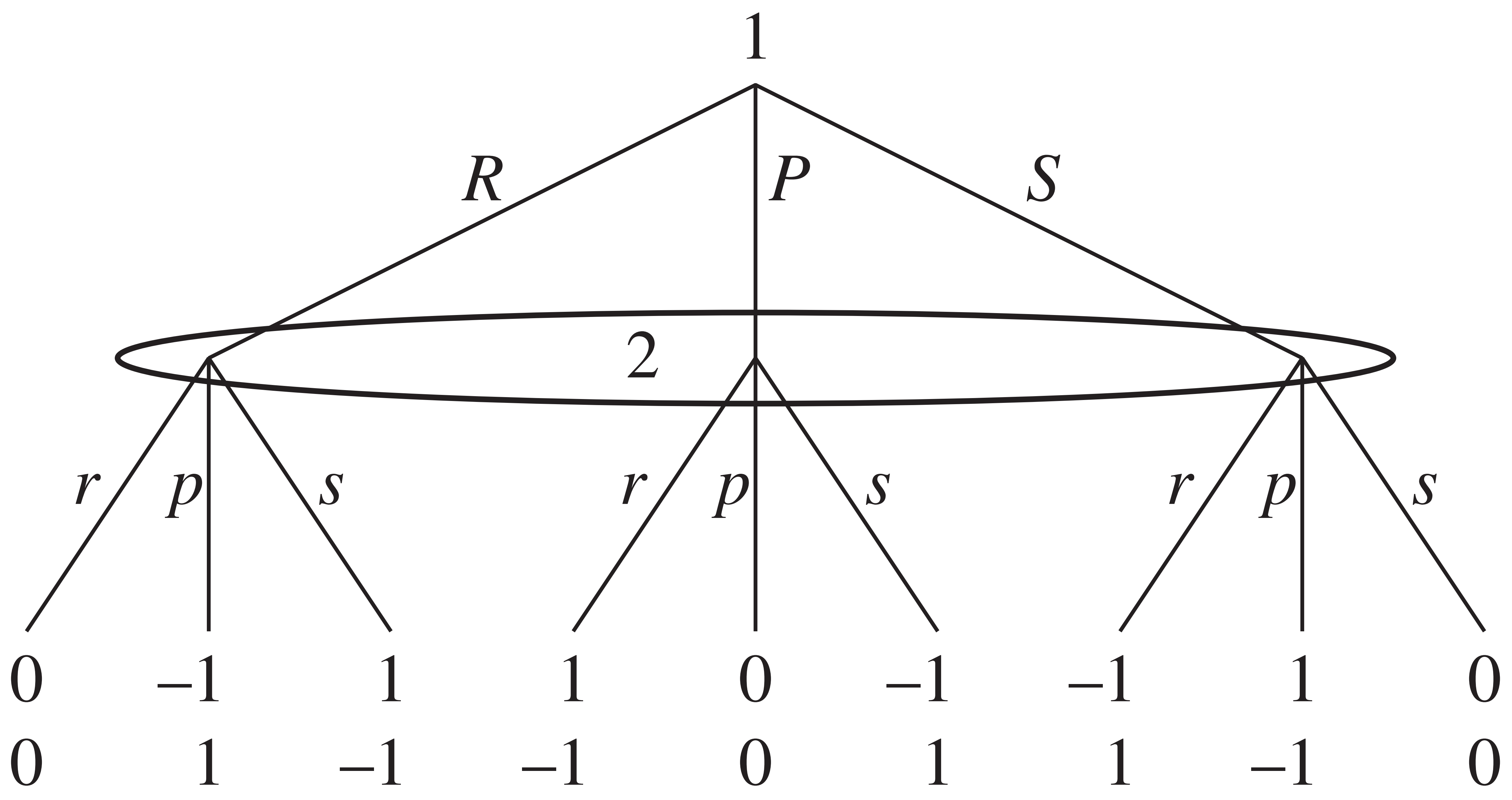


FIGURE 7.5 Game tree of rock-paper-scissors.

石头剪子布博弈

		参与人 2		
		R	*P*	*S*
	R	0, 0	–1, 1	1, –1
参与人 1	*P*	1, –1	0, 0	–1, 1
	S	–1, 1	1, –1	0, 0

完美信息

Perfect information

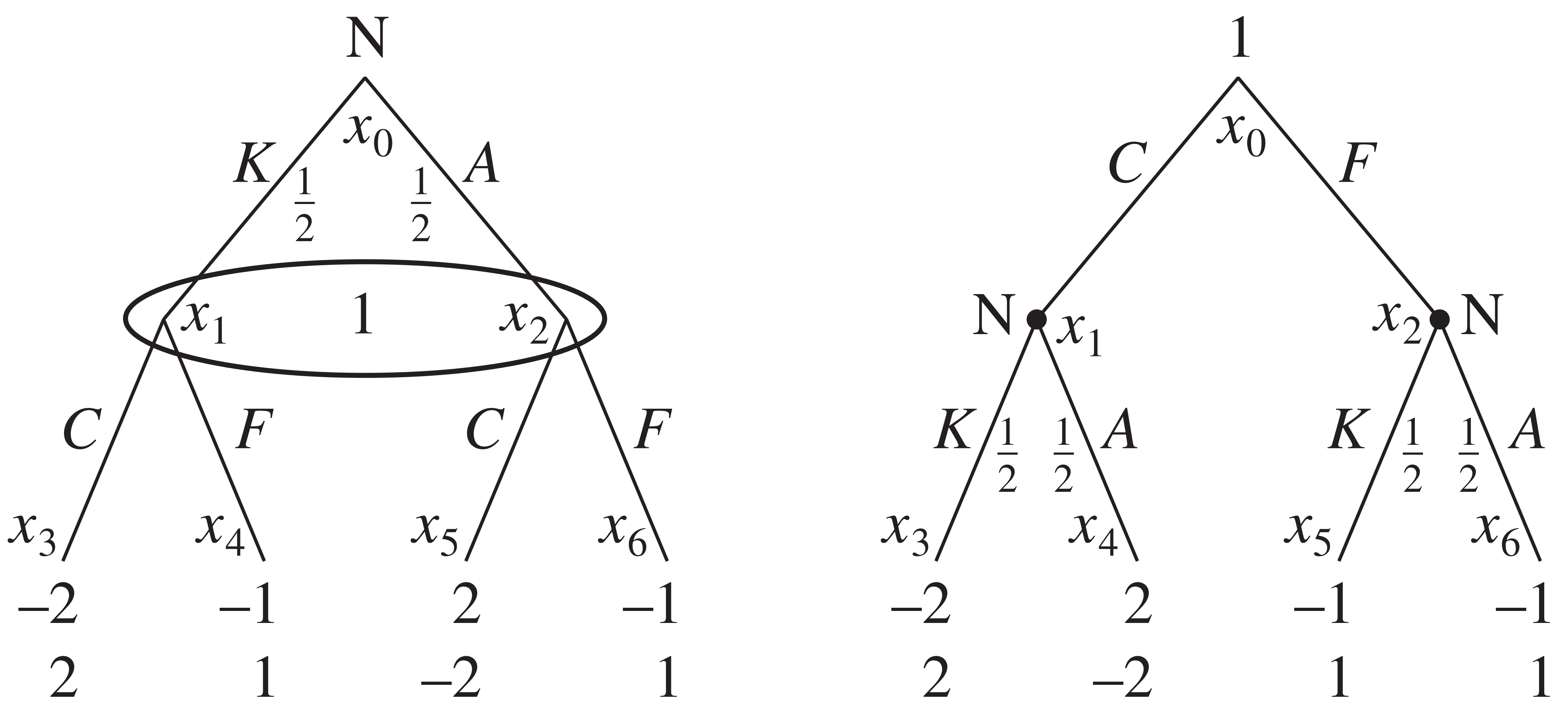
完全信息：每个参与人都知道所有参与人的行动集合和支付函数，且这一点是共同知识

在完全信息博弈（games of complete information）中，

- 所有信息集都是单点，且没有“自然”参与的博弈称为**完美信息博弈**（games of perfect information）
- 部分信息集包含多个节点，或有“自然”参与的博弈称为**不完美信息博弈**（games of imperfect information）

$\Rightarrow$ 每个同时行动（静态）博弈都是不完美信息博弈

- 比大小：
  - 一副扑克牌中仅包含相同数量的 K 和 A
  - 参与人 1 从中抽一张，但不允许看牌
此时参与人可以选择叫牌（C）或放弃（F）
  - 如果选择放弃，则参与人 1 付 1 元给参与人 2
  - 如果选择叫牌，
    - 如果手牌为 K，则参与人 1 付 2 元给参与人 2
    - 如果手牌为 A，则参与人 2 付 2 元给参与人 1



这两个博弈树是等价的，因此我们在定义不完美信息时要考虑“自然”

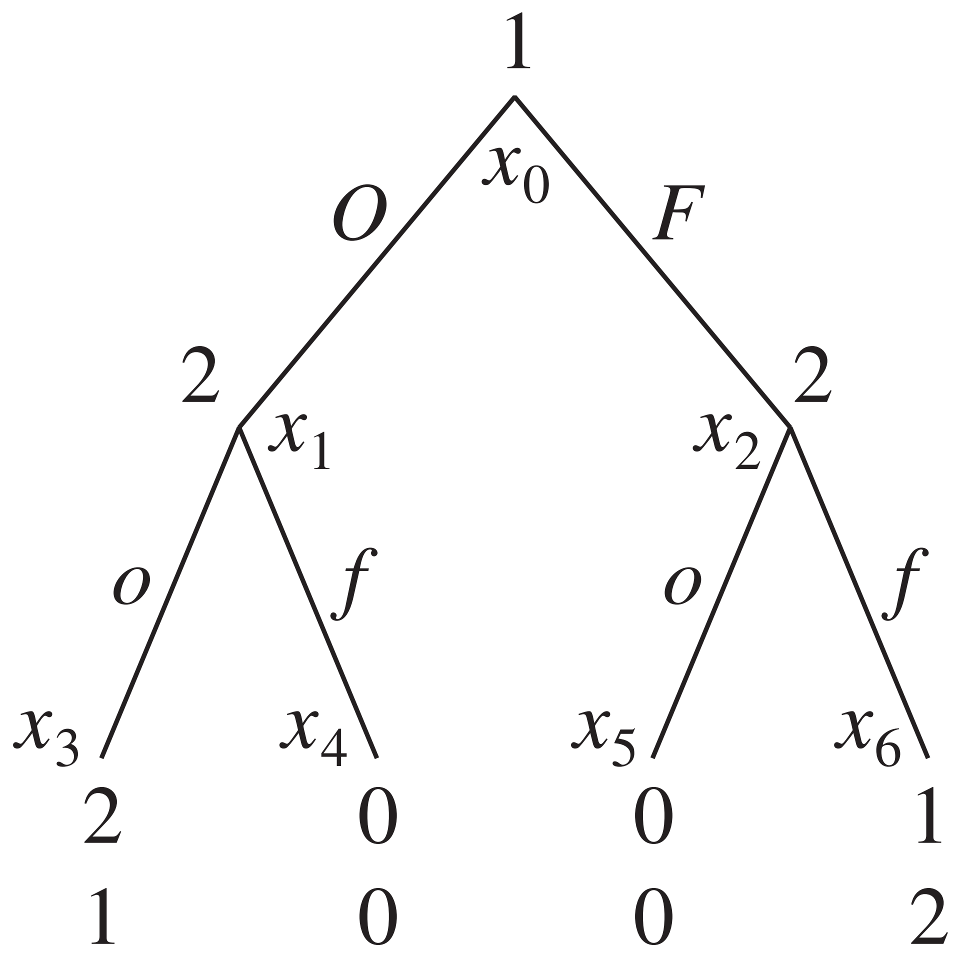
纯策略

扩展式博弈中的纯策略：参与人 i 的纯策略是一个完整的行动计划，包括参与人 i 在他的每个信息集上所选择的纯行动

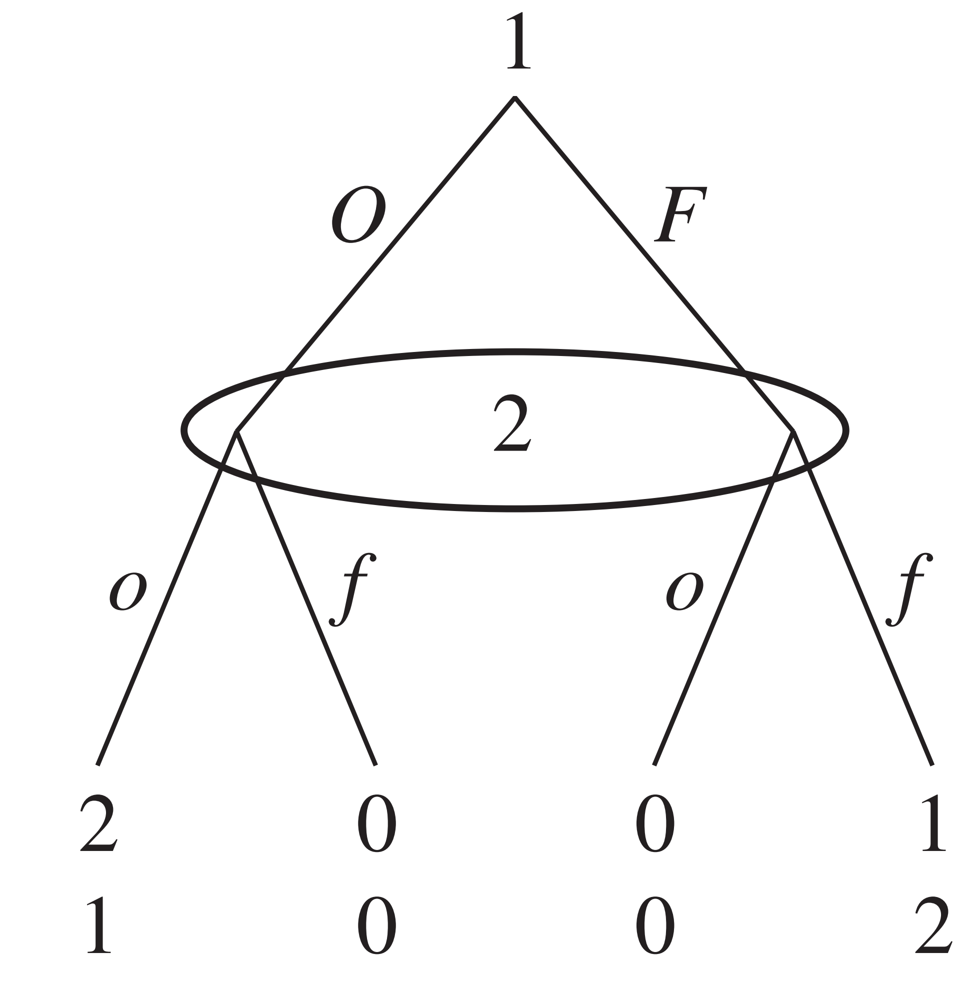
- 在右上的序贯行动 BoS 中，参与人 1 的纯策略是 O 和 F，即 $S_1 = \{O, F\}$；参与人 2 的纯策略是：
  - 在信息集 (x_1, x_2) 上选择 $(o, o), (o, f), (f, o), (f, f)$，可以简写成 $S_2 = \{oo, of, fo, ff\}$
- 在右下的同时行动 BoS 中，参与人 2 只有一个信息集，因此他的纯策略只有 o 和 f

参与人 i 的**纯策略**是一个函数 $s_i : H_i \to A_i$，它给每一个信息集 $h \in H_i$ 赋予一个行动 $A_i(h) \in A_i$。参与人 i 的所有纯策略的集合写作 S_i

- 纯策略的个数：当参与人 i 有 k 个信息集，且信息集 j 中包含 m_j 个行动时，参与人 i 共有 $m_1 \times m_2 \times \cdots \times m_k$ 个纯策略



Sequential-move BoS



Simultaneous-move BoS

混合策略与行为策略

Mixed versus behavioral strategies

参与人 i 的**混合策略**是关于他的纯策略 $s_i \in S_i$ 的概率分布

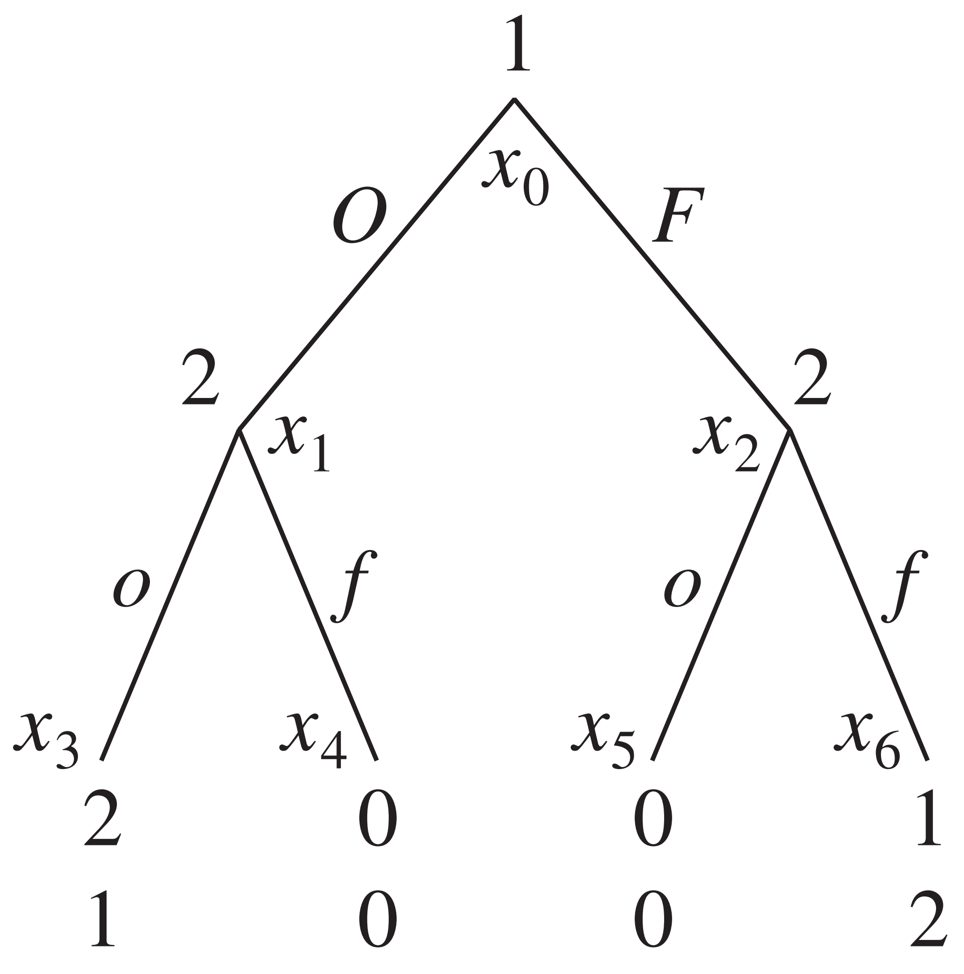
- 选择一个特定的混合策略，等于在博弈开始之前随机选择一个完整的行动计划，然后全部执行。但在序贯博弈中，参与人可能想在到达特定的信息集时再随机选择行动。

  例如在序贯行动 BoS 中，参与人 2 可能会考虑：“当参与人 1 选择 O 时，我选择 f；当参与人 1 选择 F 时，我依照 $(2/3, 1/3)$ 的概率选择 (o, f)”

  混合策略无法描述这种行动计划，因此我们需要定义下面的行为策略

参与人 i 的**行为策略**（behavioral strategy）赋予每个信息集 $h \in H_i$ 一个 $A_i(h)$ 上的独立的概率分布，记作 $\sigma_i : H_i \to \Delta A_i(h)$。其中，$\sigma_i(a_i(h))$ 代表参与人 i 在信息集 h 处选择行动 $a_i(h)$ 的概率

- 在序贯行动 BoS 中，参与人 2 的混合策略是概率分布 $(p_{oo}, p_{of}, p_{fo}, p_{ff})$，而行为策略是在信息集 x_1 和 x_2 的两个概率分布 $(\sigma_2(o_{x_1}), \sigma_2(f_{x_1}))$ 和 $(\sigma_2(o_{x_2}), \sigma_2(f_{x_2}))$



Sequential-move BoS

混合策略与行为策略间的转换

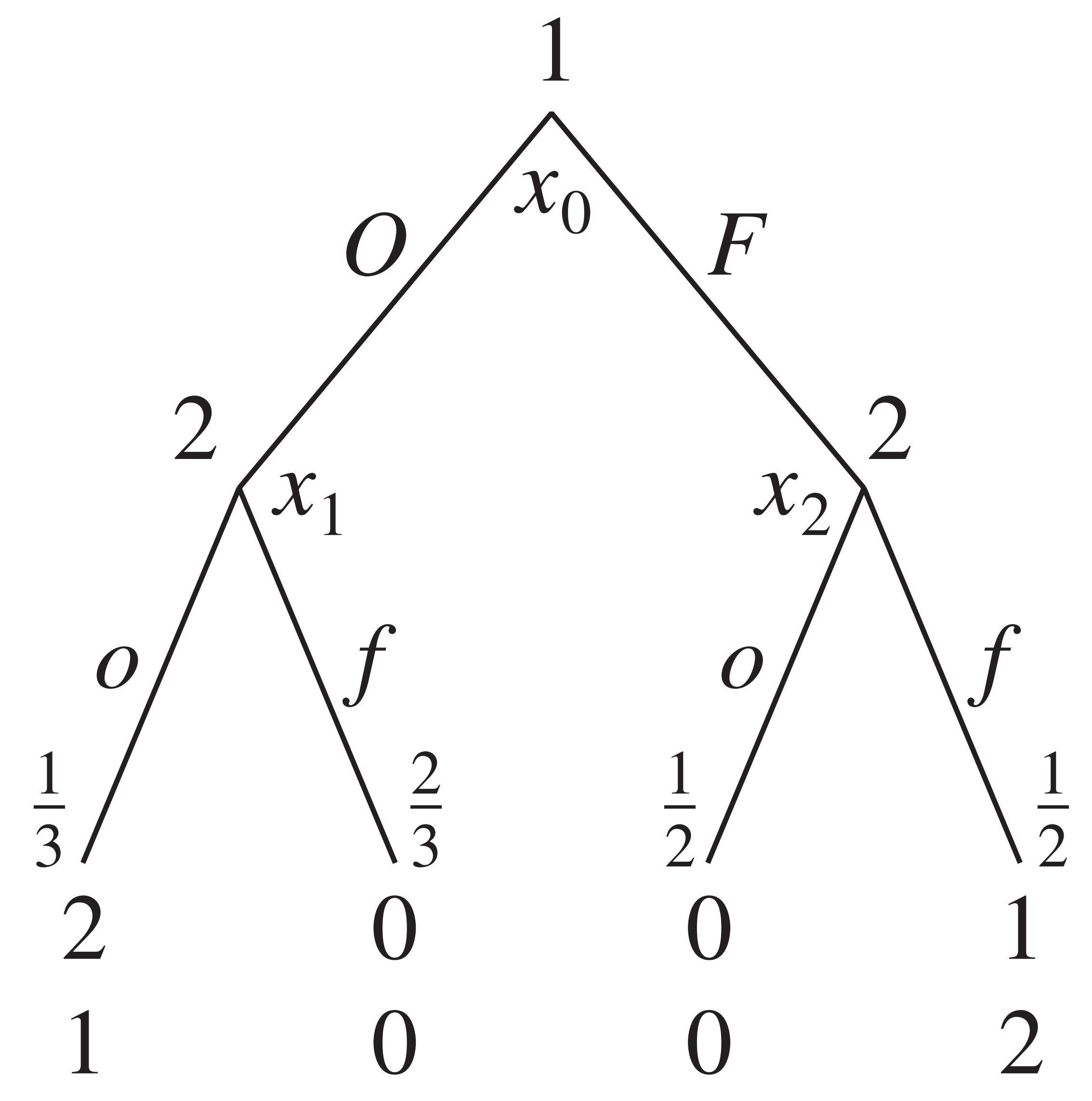
- Q1：给出一个混合策略时，是非存在一个回报相同的行为策略?
  - 当参与人 2 的混合策略是 $(p_{oo}, p_{of}, p_{fo}, p_{ff})$ 时，与之回报相同的行为策略是

$$\sigma_2(o_{x_1}) = \Pr(o\,|\,O) = p_{oo} + p_{of}\,，\ \sigma_2(f_{x_1}) = \Pr(f\,|\,O) = p_{fo} + p_{ff}$$
$$\sigma_2(o_{x_2}) = \Pr(o\,|\,F) = p_{oo} + p_{fo}\,，\ \sigma_2(f_{x_2}) = \Pr(f\,|\,F) = p_{of} + p_{ff}$$

- Q2：给出一个行为策略时，是非存在一个回报相同的混合策略?
  - 右图给出了一个参与人 2 的行为策略，与之回报相同的混合策略满足

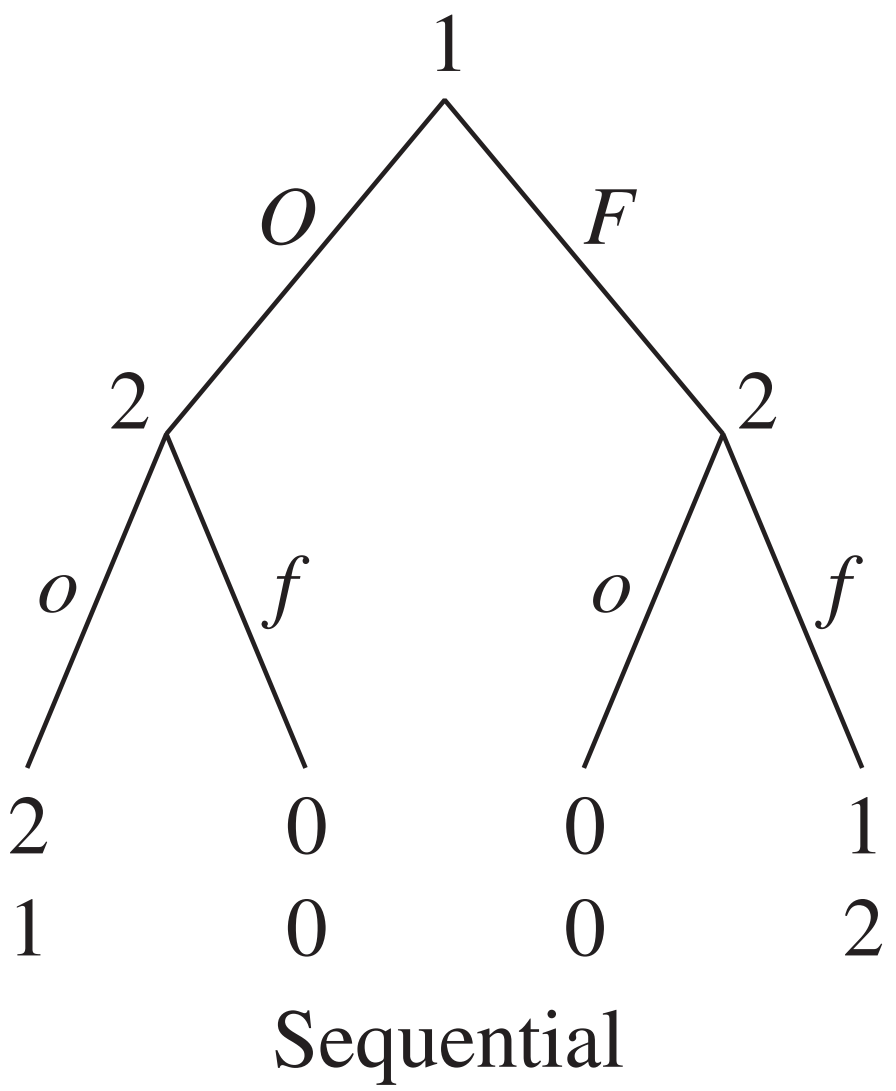
$$\Pr(o\,|\,O) = \frac{1}{3} = p_{oo} + p_{of}\,，\ \Pr(f\,|\,O) = \frac{2}{3} = p_{fo} + p_{ff}\,,$$
$$\Pr(o\,|\,F) = \frac{1}{2} = p_{oo} + p_{fo}\,，\ \Pr(f\,|\,F) = \frac{1}{2} = p_{of} + p_{ff}\,,$$

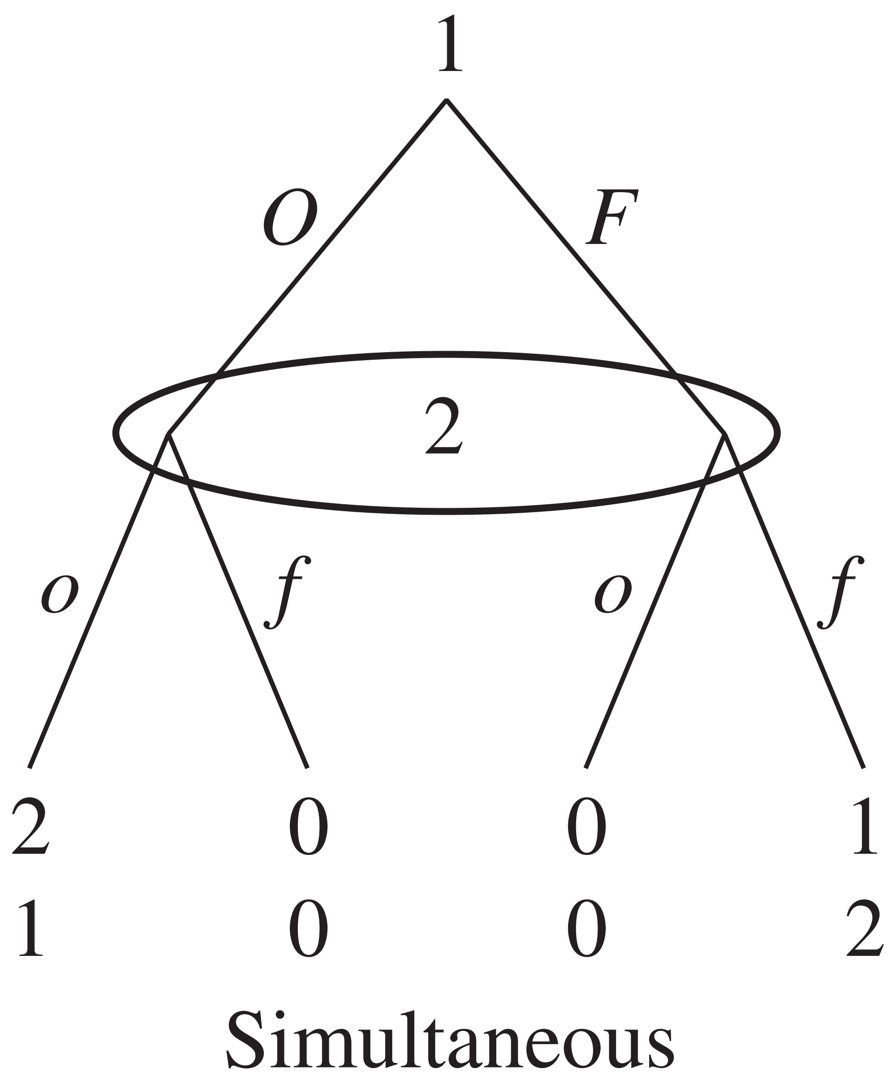
  因为 $p_{oo} + p_{of} + p_{fo} + p_{ff} = 1$，上面四个等式退化为两个，因此存在无限多个回报相同的混合策略（例如 (1/3, 0, 1/6, 1/2) 和 (1/6, 1/6, 1/3, 1/3)）



- 当参与人不会忘记任何已知信息时，我们称该博弈为**完美回忆博弈**（**games of perfect recall**）。在完美回忆博弈中，给定对手的策略时，同一期望回报既可以由混合策略产生，也可以用行为策略产生

扩展式博弈的标准式表达





任意扩展式博弈都有唯一的标准式表达
（将扩展式博弈中的纯策略用作标准式博弈中的纯策略）

⇒ 通过标准式表达可以找到纳什均衡

Sequential-move BoS

		参与人 2			
		oo	*of*	*fo*	*ff*
参与人 1	*O*	2, 1	2, 1	0, 0	0, 0
	F	0, 0	1, 2	0, 0	1, 2

Simultaneous-move BoS

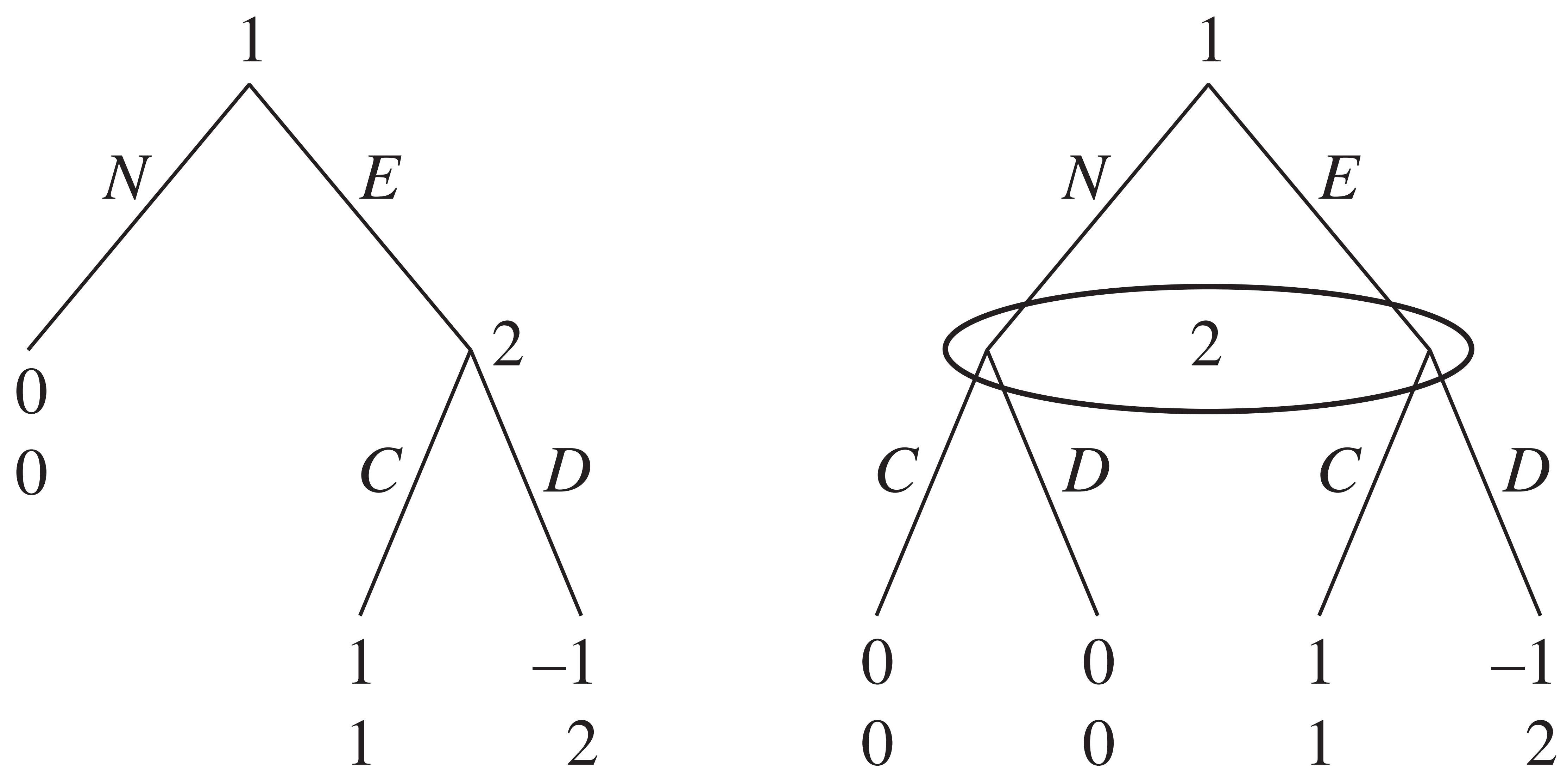
		参与人 2	
		o	*f*
参与人 1	*O*	2, 1	0, 0
	F	0, 0	1, 2

标准式博弈的扩展式表达

参与人 2

参与人 1		*C*	*D*
	N	0, 0	0, 0
	E	1, 1	–1, 2

标准式博弈可以有不同的扩展式表达



Sequential-move BoS

		参与人 2			
		oo	*of*	*fo*	*ff*
参与人 1	*O*	**2, 1**	**2, 1**	0, 0	0, 0
	F	0, 0	1, 2	0, 0	**1, 2**

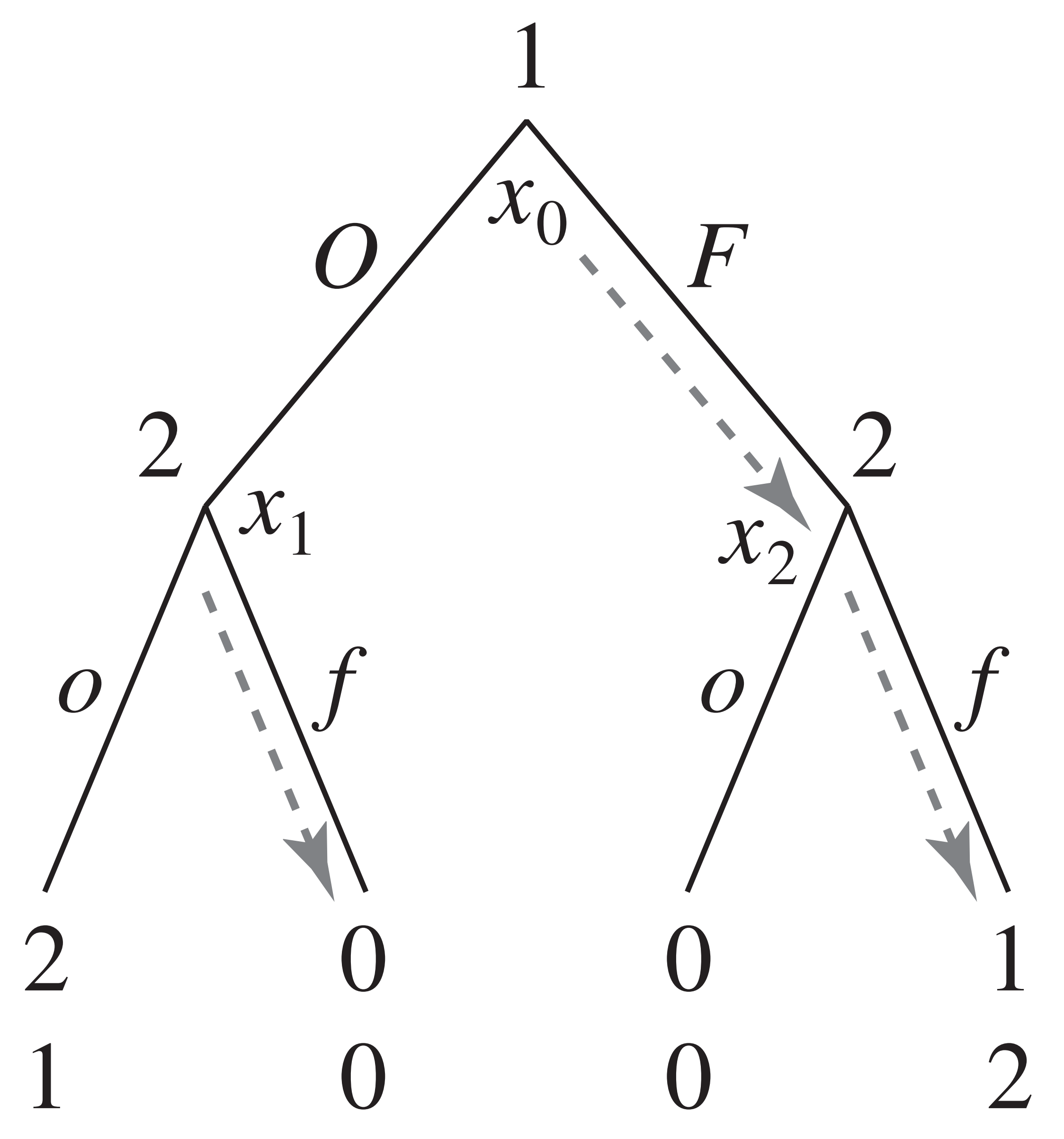
纳什均衡和博弈路径

Nash equilibrium and paths of play

- 序贯行动 BoS 的标准式表达中的纯策略纳什均衡为：$(O, oo), (O, of), (F, ff)$
- (O, oo) 和 (O, of) 会产生相同的回报，那么这两个均衡有什么不同呢?

  ⇒ 两者的不同是参与人 2 在“参与人 1 选择 F” 时（即信息集 x_2）选择的策略不同，而这个信息集在博弈中是无法到达的

令 $\sigma^* = (\sigma_1^*, \ldots, \sigma_n^*)$ 为扩展式博弈中的行为策略纳什均衡。当一个信息集在 σ^* 下可以以正概率到达时，我们称之为**在均衡路径上（on the equilibrium path）**；当信息集不可能到达时，称之为**偏离均衡路径（off the equilibrium path）**



- 从右图可以看出，到达信息集 x_2 的概率为 100%，因此它在均衡路径上，而 x_1 则偏离了均衡路径。参与人 2 在 x_1 上选择 f 的作用是迫使参与人 1 选择 F， 这在博弈论中称为“威胁”
- 但是这种威胁是不可信的，因为如果参与人 1 错误的选择了 O，那参与人 2 不会选择 f
  造成这种奇怪现象的原因是标准式表达失去了动态博弈中关于行动顺序的信息

练习：否决权（veto power）

- 两个参与人需要对三个候选人 a, b, c 进行投票
- 参与人 1 的偏好是 $a \succ_1 b \succ_1 c$，参与人 2 的偏好是 $c \succ_2 b \succ_2 a$
- 参与人 1 首先行动，从三人中选出一人进行否决；然后参与人 2 行动，从剩下的两人中选出获胜者

- 回答下面的问题：
  1. 画出这个博弈的博弈树（定义适当的支付函数）
  2. 写出每个参与人的纯策略集合
  3. 找到全部纯策略纳什均衡

练习：蜈蚣博弈（centipede game）

- 一个盒子中有6元钱。两个参与人进行下面的游戏：
  - 行动顺序是参与人 1 → 参与人 2 → 参与人 1 → 参与人 2
  - 在每一回合，能够行动的参与人可以选择收钱（G）或继续（S）。如果选择 G，则自己可以获得盒子中金额的 2/3，对手获得 1/3，游戏结束
  - 在前三回合，能够行动的参与人如果选择 S，则盒子中的金额会变成 3/2 倍，游戏进入下一回合。在第四回合，如果参与人 2 选择 S，则盒子中的金额依然会变成 3/2 倍，然后自己获得翻倍后金额的 1/3，参与人 1 获得 2/3，游戏结束

- 回答下面的问题：
  1. 画出博弈树
  2. 有多少终点？有多少信息集？
  3. 每个参与人有多少纯策略？
  4. 找到全部纯策略纳什均衡，以及它们对应的回报
  5. 如果参与人 2 在第四回合选择 S 时，两人平分盒子中翻倍后的金额，纯策略纳什均衡会发成变化吗？