ZJU 2024–2025 学年春夏学期

计算经济学讨论班

Lec 3: 扩展式博弈

吴一航 yhwu_is@zju.edu.cn

浙江大学计算机科学与技术学院

CONTENT

目录

引入：一个简单的例子

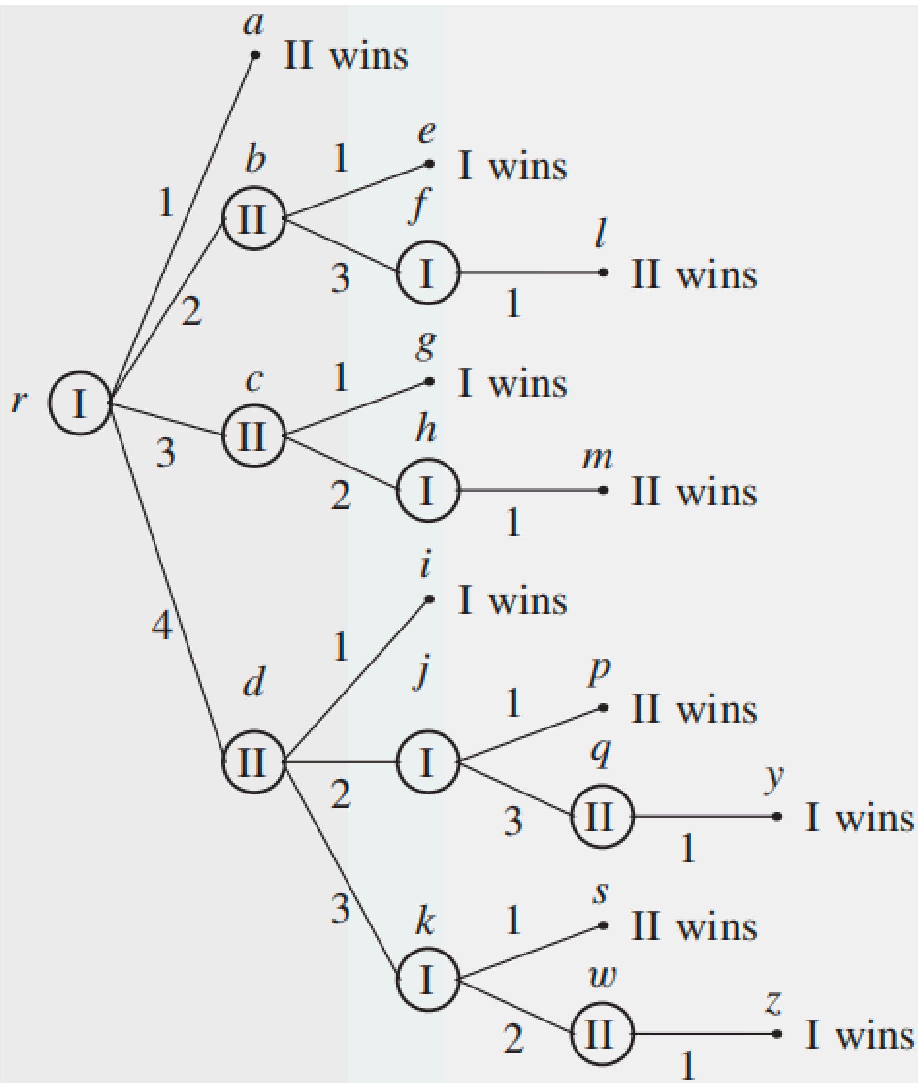
考虑如下所示的简单博弈，博弈有两个参与人，参与人 1 首先行动。每个参与人一次抓取一个方格，条件如下：

2	4
1	3

1. 只有当某个方格未被抓取时，参与人才能抓取该方格
2. 如果方格 2 或 3 已经被抓取，那么参与人不能抓取方格 4
3. 当方格 1 被抓取时，博弈结束，谁抓到方格 1 谁就输

问：参与人 1 和参与人 2 有制胜策略吗？谁有制胜策略？

引入：一个简单的例子



1. 整个博弈的进行流程可以表达为一棵树；
2. 参与人 1 有制胜策略：先抓取 4，然后如果 2 选择 2 则抓取 3，否则抓取 2；
3. 这一博弈如何用策略式博弈表达？用策略式博弈表达会损失什么信息？
4. 如果只有一棵光秃秃的树，整个博弈的信息能被完整表达吗？

扩展式博弈的定义

定义

一个**扩展式博弈（extensive-form game）**是一个有序的七元组 $\Gamma = \left(N, V, E, x^0, (V_i)_{i\in N}, O, u\right)$，其中：

1. N 是参与人的有限集合；
2. (V, E, x^0) 是一棵树，其中 V 是节点的有限集合，E 是边的有限集合，x^0 是根节点；
3. $(V_i)_{i\in N}$ 是非叶节点的集合，i 表示对应节点选择行动的参与人；
4. O 是博弈的所有可能结果；
5. u 是一个函数，将树的每一个叶子节点映射到 O 中的一个结果。

通常，叶子节点 x 处的结果 $u(x)$ 就是参与人到达 x 时的效用。

扩展式博弈的定义

将前面的例子形式化表达：

1. $N = \{1, 2\}$；
2. V 和 E 如图，$x^0 = r$；
3. $V_1 = \{r, f, h, j, k\}$，$V_2 = \{b, c, d, q, w\}$；
4. $O = \{1\text{胜}, 2\text{胜}\}$；
5. $u(a) = u(l) = u(m) = u(p) = u(s) = 2\text{胜}$，$u(e) = u(g) = u(i) = u(y) = u(z) = 1\text{胜}$。

我们研究的扩展式博弈都是有限博弈，即树的节点数量有限。无限的情况可能有两种，一是博弈的深度有界，但一个参与人可能在某一点处有无数个可能行动；二是博弈的深度无界，即博弈可能永远不会结束

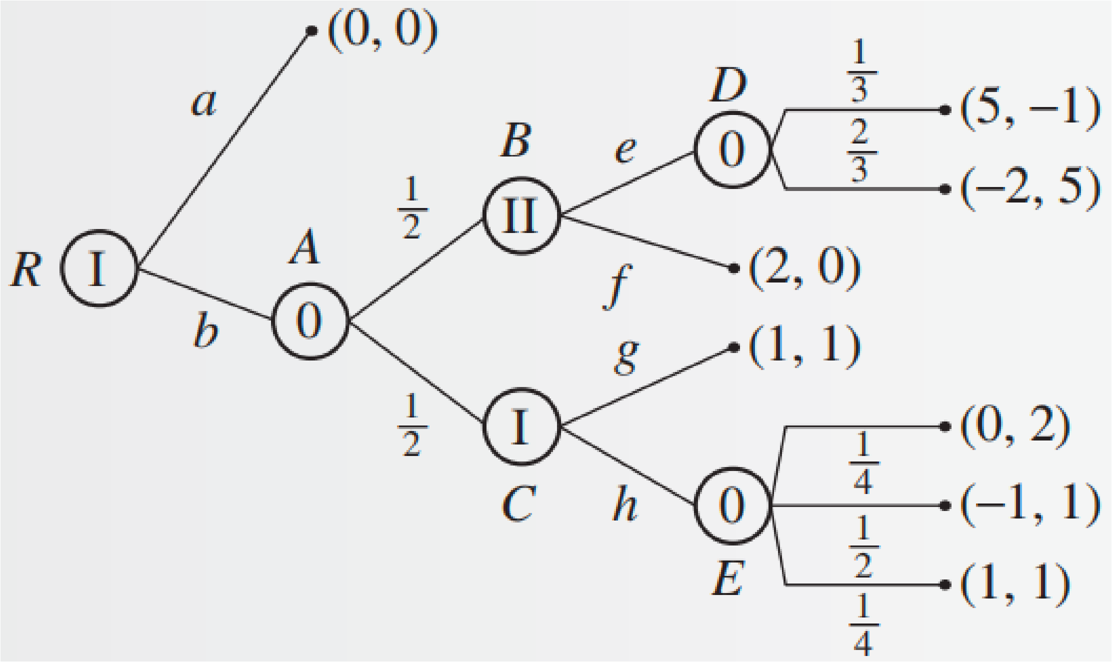
扩展式博弈的定义

为了接下来讨论的方便，需要引入一些记号和定义：

1. $J(x)$ 表示非叶节点 x 处的决策者，例如 V_i 中任意节点 v_i 的决策者就是 $J(v_i) = i$；
2. $C(x)$ 表示非叶节点 x 的所有子节点的集合；
3. $A(x)$ 表示参与人 $J(x)$ 在节点 x 处所有可能行动的集合；
4. 从 x 出发的子博弈 $\Gamma(x)$ 是一个扩展式博弈，只是定义中的参数需要限制在 x 为根的子树上；
5. 从根节点 x^0 处 $J(x^0)$ 行动开始，到一个叶子节点 x 结束的节点序列称为博弈的一个**展开（play）**。

包含随机行动的博弈

在之前的例子中，从一个状态过渡到另一个状态的行动都是由参与人完成的，这样的模型对国际象棋是合适的，但是对大富翁这样的游戏就不合适了，因为每次掷骰子的结果是随机的，状态的改变可能取决于随机过程：



- 随机行动需要加入一个新的参与人：“自然”，用 0 或 N 表示；
- 还需要加入 p_x 来表示自然在节点 x 处抽签带来的可能结果的概率向量。

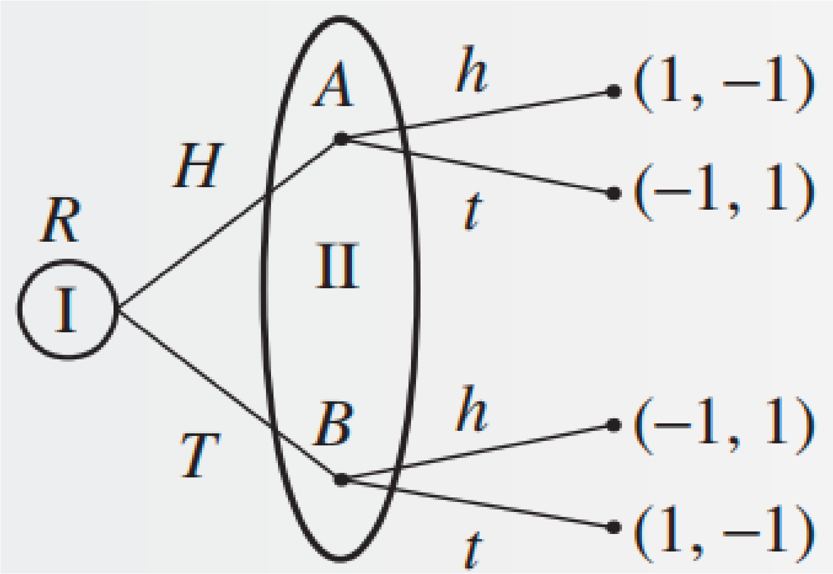
完美信息博弈

如果每个参与人在选择行动时，都清楚地知道他位于博弈树的哪个节点上，那么这个博弈就是**完美信息博弈（game with perfect information）**，例如国际象棋。但很多博弈不符合这一条件，例如斗地主，你不知道其他玩家的手牌，也可能忘了别人或者自己之前的行动。

例

猜硬币博弈是一个两人博弈，每个参与人选择出示正面还是反面，H 表示硬币正面，T 表示硬币反面。如果两个参与人的选择一致，则参与人 2 给参与人 1 一美元；如果两个参与人的选择不一致，则参与人 1 给参与人 2 一美元。这一博弈用扩展式博弈表达如下页图所示。

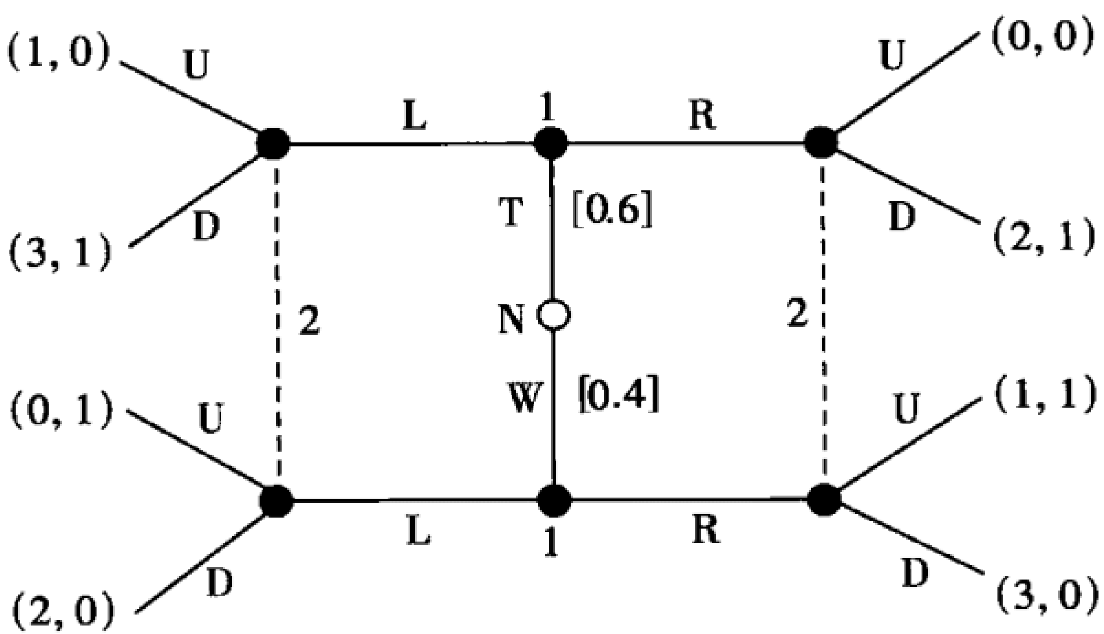
完美信息博弈



1. 圈起来的两个点构成一个**信息集（information set）**，表示参与人在这两个节点上无法区分；
2. 完美信息博弈信息集都是单点集。

以上只是表达信息集的一种方式，在另外一些教材中，同一信息集中的节点可能用虚线连接。

下图也是一种表达扩展式博弈的形式。自然首先行动并选择了参与人 1 的**类型（type）**或**私人信息（private information）**。参与人 1 知道自己是“强”（T）的类型的概率为 0.6，而参与人 1 知道自己是“弱”（W）的类型的概率为 0.4。参与人 1 可以选择向左（L）或向右（R）行动。参与人 2 可以观察到参与人 1 的行为，但不能观察到他的类型，且参与人 2 可以选择向上（U）或向下（D）行动。



子博弈

在定义了完美信息的概念后，我们可以严格定义子博弈的概念：

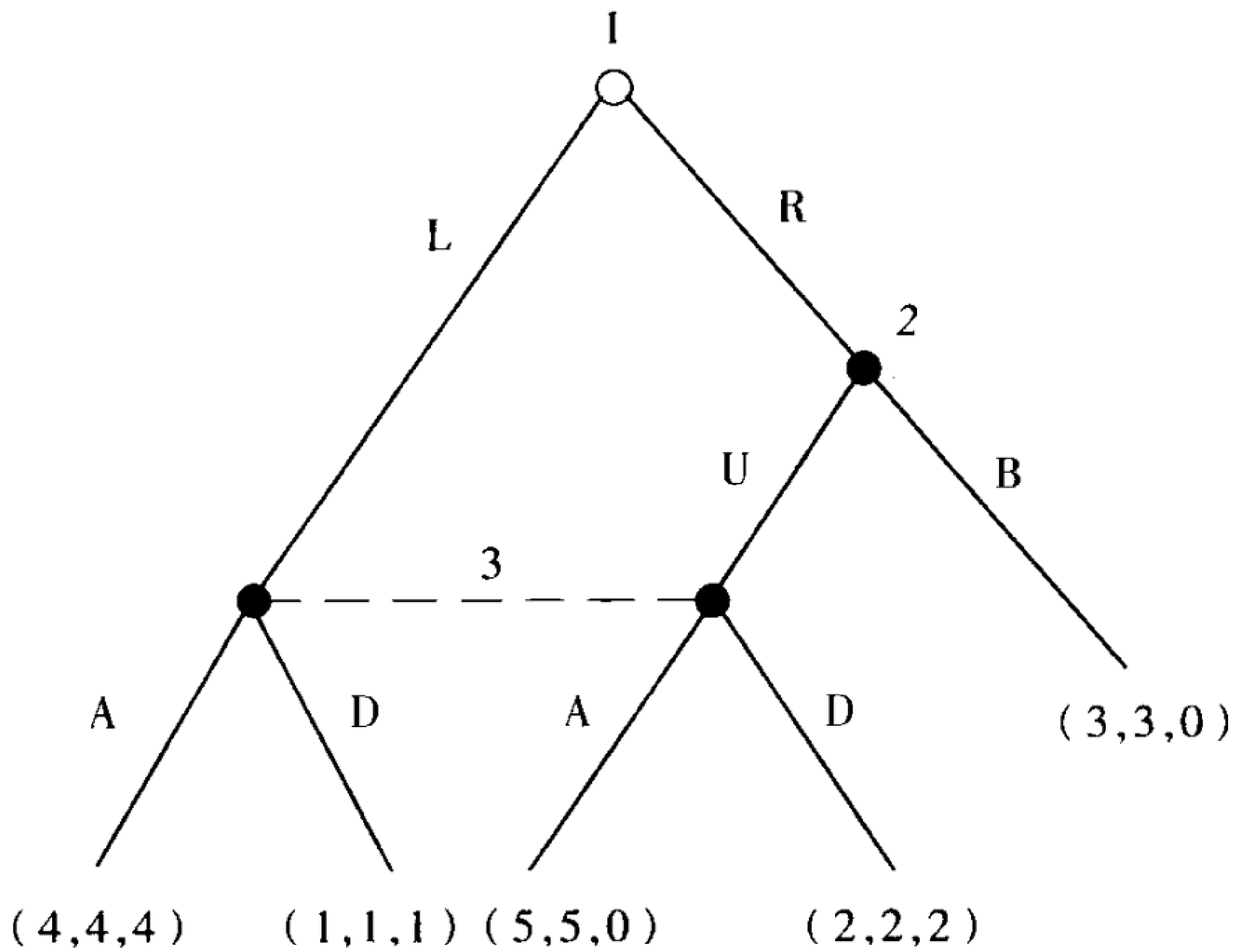
定义

一个扩展式博弈的子博弈 G 由一个节点 x 和所有该节点的后继节点 $T(x)$ 组成，并且满足如下两个条件：

1. x 是一个单节点信息集，即 $h(x) = \{x\}$；
2. 对于所有 $x' \in T(x)$，如果 $x'' \in h(x')$，那么 $x'' \in T(x)$。

两个要求都是非常自然的：第一个要求可以考虑囚徒困境，如果没有单点信息集的限制，囚徒困境也可以出现子博弈的概念，显然不合理；第二个要求意味着子博弈不能切割原博弈的信息集，否则参与人在原博弈中不知道的信息在子博弈中就变成知道的信息。

子博弈



根据子博弈的定义，上图中的子博弈只有原博弈自身。

完美记忆博弈

几乎所有经济文献中的博弈都是**完美记忆博弈（games of perfect recall）**：所有参与人都不会忘记曾经知道的任何信息，清楚他们前面所选择的行动。

- 为了更加规范地表述这一点，首先要求如果 x 和 x' 是在同一信息集中，则它们两者任何一个都不会是另一个的前续节点；
- 但这并不足以保证参与人永远不会忘记他所知道的信息，如图所示：

- 为了排除这种情况，要求如果 $x'' \in h(x')$，x 是 x' 的一个前续节点，并且同一参与人 i 在 x 和 x'（因而也在 x''）上采取行动，那么存在一个节点 $\hat{x}$（有可能就是 x 本身）位于与 x 同样的信息集当中，且 x 是 x' 的一个前续节点，在 x 点所采取的行动到达 x' 的路径与在 x 点所采取的行动到达 x'' 的路径是一样的；
- 直观上，节点 x' 和 x'' 可以用参与人所不具有的信息加以区分，因而当参与人位于信息集 $h(x)$ 时，他不可能已经拥有这些信息：x' 和 x'' 必须与在 $h(x)$ 的同一行动相一致，因为参与人可以回想起他曾经采取过的行为。

CONTENT

目录

行为策略

为了描述扩展式博弈的策略，首先需要一些定义：

- 用 H_i 表示参与人 i 的信息集的集合；
- 用 $A_i = \bigcup_{h_i \in H_i} A(h_i)$ 表示参与人 i 所有可选择行动的集合；
- 参与人 i 的纯策略就是映射 $s_i : H_i \to A_i$，且对所有 $h_i \in H_i$，$s_i(h_i) \in A(h_i)$；
- 参与人 i 的纯策略空间 S_i 就是所有这类 s_i 的空间。

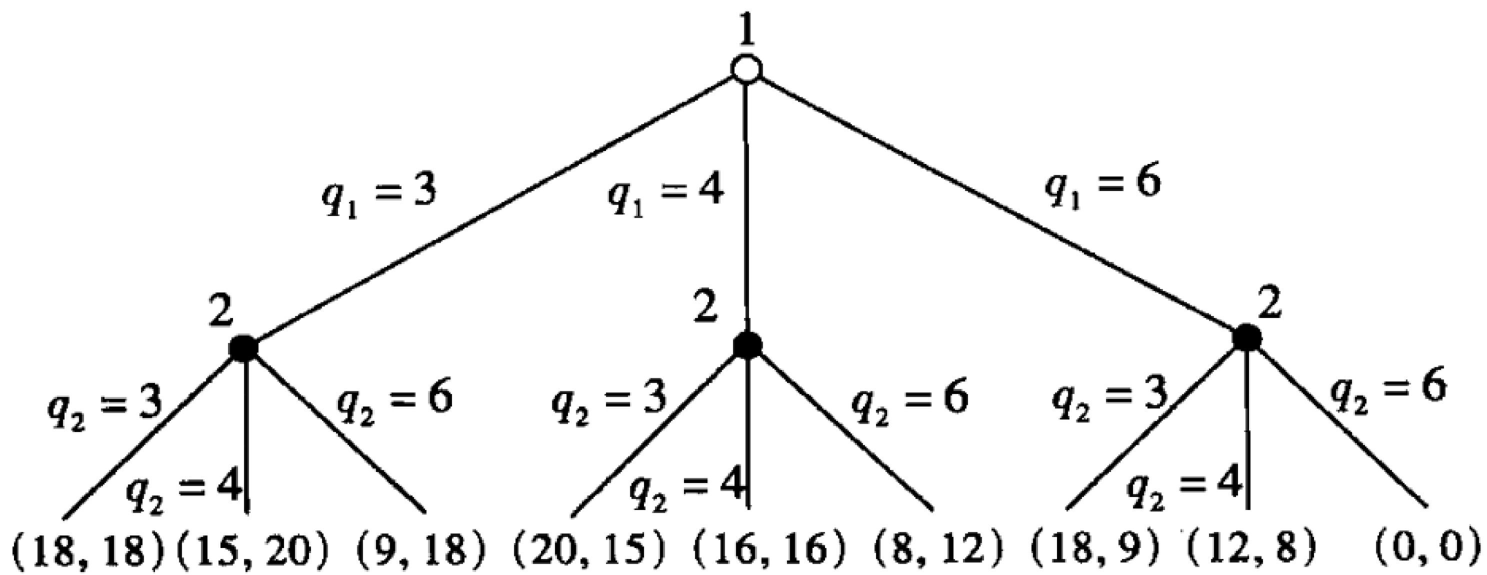
由于每一个纯策略都是从信息集到行动集的映射，我们可以把 S_i 写成每一信息集 h_i 下的行动空间的笛卡儿乘积的形式：

$$S_i = \times_{h_i \in H_i} A(h_i)$$

也就是说，每个纯策略是一个 $|H_i|$ 元组，其中每个元素是 $A(h_i)$ 的一个元素，即对应一个信息集中的行动。

行为策略

下图给出了一个扩展式博弈的例子。参与人 1 拥有一个信息集和三种行动，因而他有三个纯策略。参与人 2 有三个信息集，分别对应于参与人 1 的三种可能的行动，且参与人 2 在每一个信息集下也有三种可能的行动，因而参与人 2 有 27 种纯策略。



更一般地，参与人 i 的纯策略的个数 $\#S_i$ 等于

$$\prod_{h_i \in H_i} \#A(h_i)$$

给定每一参与人 i 的一个纯策略及自然的行动概率分布，可以计算结局的概率分布，从而对每一组合 s 给出期望收益 $u_i(s)$。

定义了每一纯策略的收益后，接下来可以定义扩展式博弈的纳什均衡 s^*：每一个参与人 i 的策略 s_i^* 在给定其竞争对手的策略 s_{-i}^*；下能最大化其期望收益。

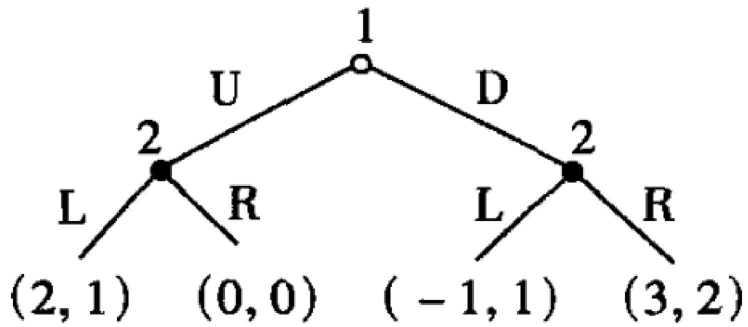
- 注意，由于这一纳什均衡定义在检验参与人 i 是否愿意偏离其当前策略的前提是保持其竞争对手的策略不变，因而这就好像各个参与人同时选择他们的策略一样；
- 但这不意味着在纳什均衡中，参与人必须同时选择行动。例如如果参与人 2 在上页图博弈中的固定策略是 $s_2 = (4, 4, 3)$，那么当参与人 1 认为参与人 2 的策略是固定的时候，他并不是假设参与人 2 的行动不受他自己行动的影响，而是认为参与人 2 会以 s_2 所确定的方式来对参与人 1 的行动作出反应；
- 根据最优反应的定义不难验证博弈的纳什均衡是 $(6, (4, 4, 3))$。

行为策略

下一步将定义扩展式博弈混合策略及混合策略均衡。这些策略被称做**行为策略（behavior strategy）**，以区别于策略式下的混合策略。

- 以 $\Delta(A(h_i))$表示 $A(h_i)$ 的概率分布；
- 参与人 i 的一个行为策略用 b_i 来表示，就是笛卡儿乘积 $\times_{h_i \in H_i} \Delta(A(h_i))$ 的一个元素；
- 也就是说，一个行为策略就确定了在每一 h_i 下各个行动的概率分布，同时此概率分布在不同的信息集下又是相互独立的；
- 注意，纯策略是一种特殊的行为策略，即在每一信息集上的分布都是退化的行为策略；
- 策略组合 $b = (b_1, ..., b_n)$ 生成结果的概率分布，从而得到了各个参与人的期望收益。行为策略的纳什均衡就是这样一种组合，即没有参与人可以通过运用不同的行为策略来增加他的期望收益。

扩展式博弈的策略式表述



	(L, L)	(L, R)	(R, L)	(R, R)
U	2, 1	2, 1	0, 0	0, 0
D	–1, 1	3, 2	–1, 1	3, 2

- 从扩展式博弈到策略式博弈的策略，指的是在不同的信息集上参与人的策略（例如图中参与人 2 的策略，括号中第一项是对应一个策略集上的选择，第二项对应第二个策略集）；
- 扩展式博弈到策略式博弈的表达唯一，但策略式博弈到扩展式博弈的表达不唯一（例如静态博弈两人行动“顺序”可交换）；
- 一种类比：一个纯策略是一本指南书，它的每一页告诉你在某个信息集中应如何进行选择。策略空间 S_i 就像这些书组成的图书馆，一个混合策略则是这些书的一个概率度量，即从图书馆中选择图书的一个随机方式。与之相比较，一个给定的行为策略则只是单一的一本书，但它在每一页里都给出一个随机行动选择。

库恩定理

读者可能有一种直觉：扩展式博弈的混合策略和行为策略有一个对应关系，因为本质上最终都是决定了到达各个节点的概率：

Kuhn（库恩）定理

在完美记忆博弈中，混合策略与行为策略等价，即能得到同样的概率分布结局。

- 更准确地说，每一个混合策略均衡都与它生成的唯一行为策略等价，每一个行为策略也等价于每一个生成它的混合策略（接下来将具体解释“生成”的含义）；
- 由于我们考虑的都是完美记忆博弈，故此后的“策略”既可以是混合策略，也可以是行为策略，当然行为策略表达更为简便。

库恩定理

任何一个策略式表的混合策略 σ_i 都会生成唯一的行为策略 b_i：用 $R_i(h_i)$ 表示参与人 i 在没有排除 h_i 下的纯策略集合，因而对所有的 $s_i \in R_i(h_i)$，存在一个参与人 i 的竞争对手到达 h_i 的策略组合 s_{-i}。如果 σ_i 使得 $R_i(h_i)$ 中某些 s_i 具有正概率，则定义 b_i 分配给 $a_i \in A(h_i)$ 的概率为：

$$b_i(a_i \mid h_i) = \frac{\sum_{s_i \in R(h_i) \text{且} s_i(h_i)=a_i} \sigma_i(s_i)}{\sum_{s_i \in R(h_i)} \sigma_i(s_i)}$$

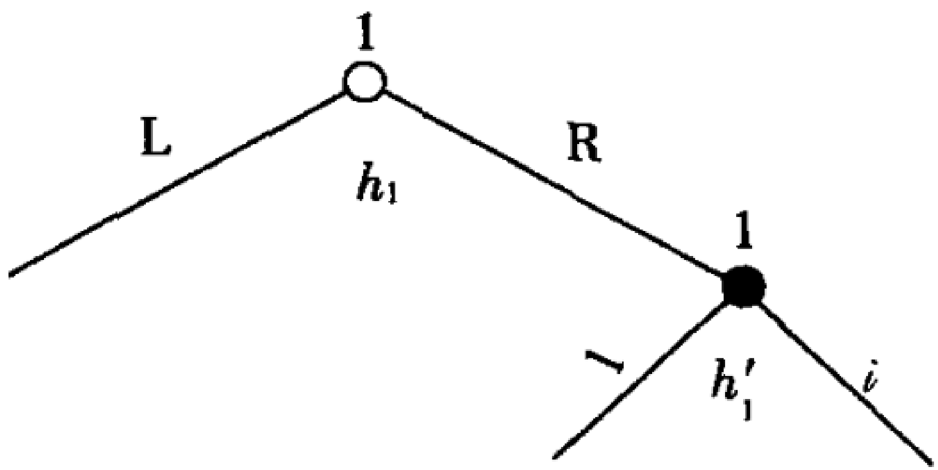
如果 σ_i 使得所有 $s_i \in R(h_i)$ 的概率都为 0，则

$$b_i(a_i \mid h_i) = \sum_{s_i(h_i)=a_i} \sigma_i(s_i)$$

以上两种情况都能保证 $b_i(\cdot \mid \cdot)$ 非负且 $\sum_{a_i \in A(h_i)} b_i(a_i \mid h_i) = 1$

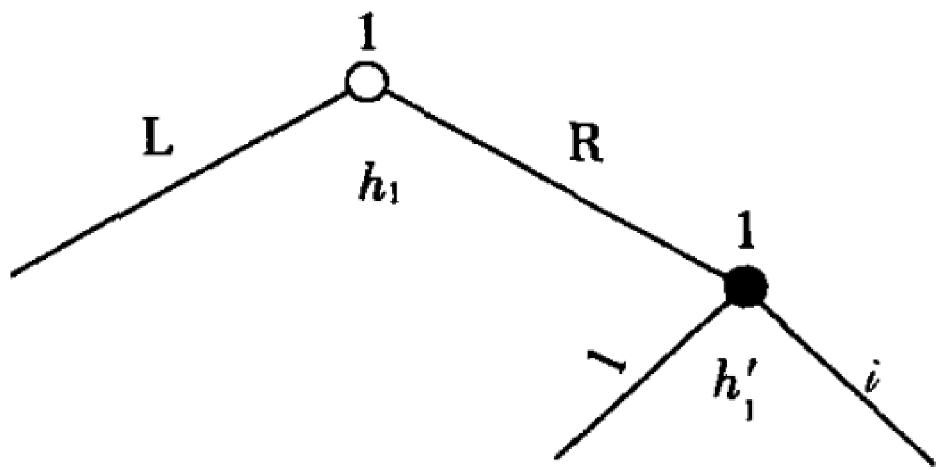
库恩定理

下面给出一些例子。下图中同一个参与人两次采取行动，考察其混合策略 $\sigma_1 = \left(\frac{1}{2}(L,l), \frac{1}{2}(R,r)\right)$。



库恩定理

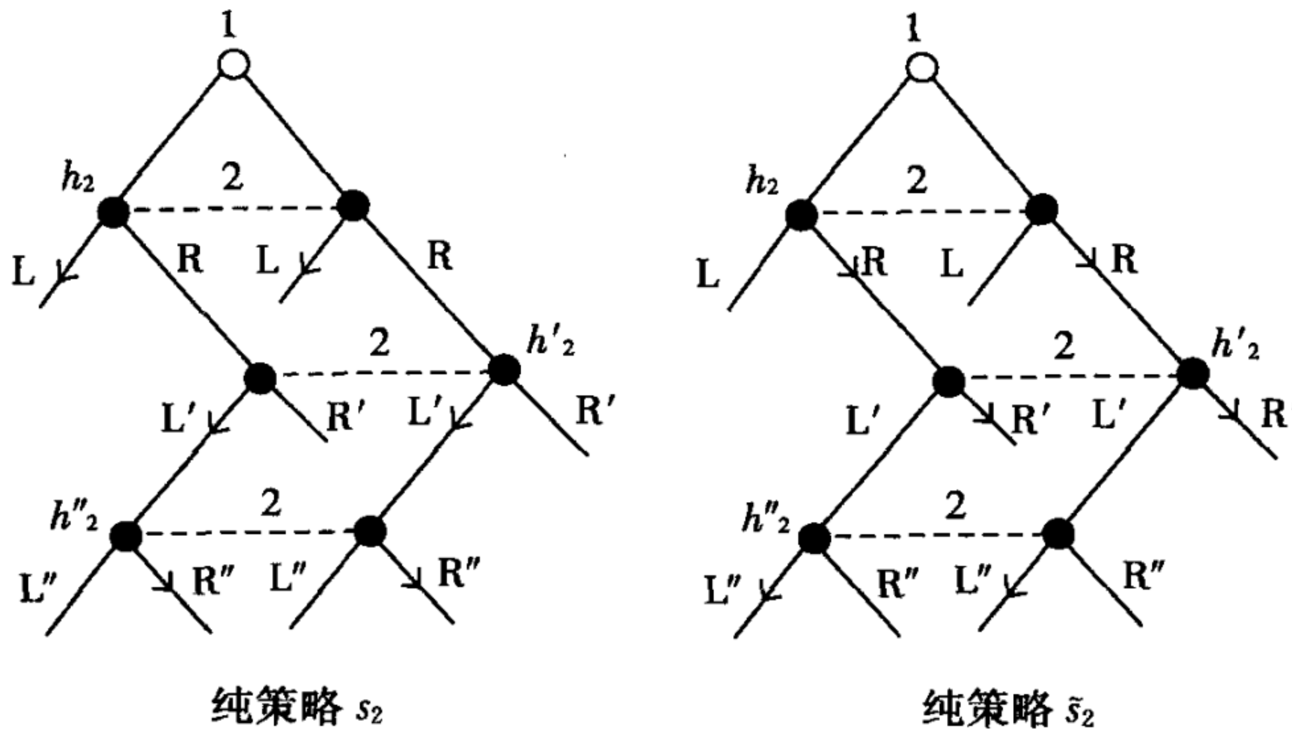
下面给出一些例子。下图中同一个参与人两次采取行动，考察其混合策略 $\sigma_1 = \left(\frac{1}{2}(L,l), \frac{1}{2}(R,r)\right)$。



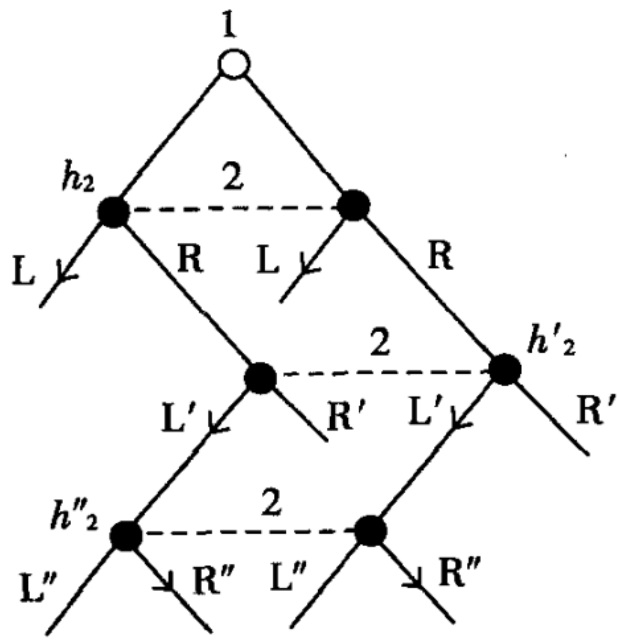
- 这一策略生成的行为策略是在信息集 $h_1{}'$ 以概率 1 选择行动 r，因为只有 $(R,r) \in R_1(h_i{}')$；
- 这里就体现了之前定义集合 R_i 的合理性。

库恩定理

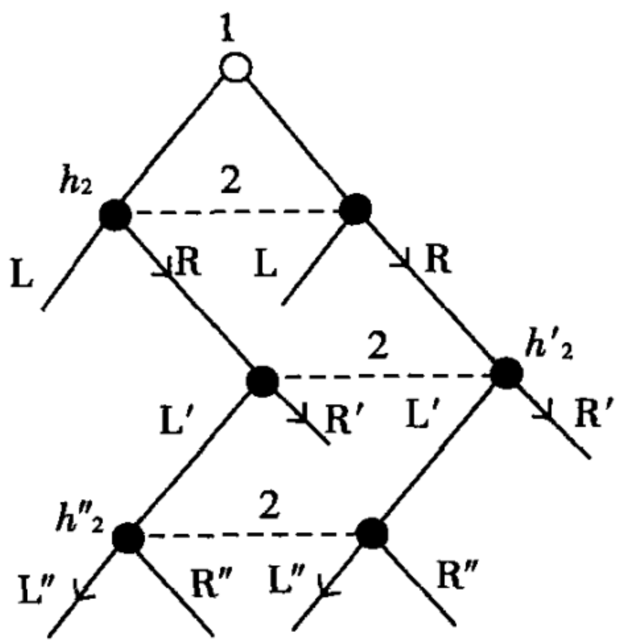
再考虑下图中的例子，参与人 2 的策略 σ_2 分别使得 $s_2 = (L, L', R'')$ 和 $\bar{s}_2 = (R, R', L'')$ 的概率为 $\frac{1}{2}$。



再考虑下图中的例子，参与人 2 的策略 σ_2 分别使得 $s_2 = (L, L', R'')$ 和 $\bar{s}_2 = (R, R', L'')$ 的概率为 $\frac{1}{2}$。



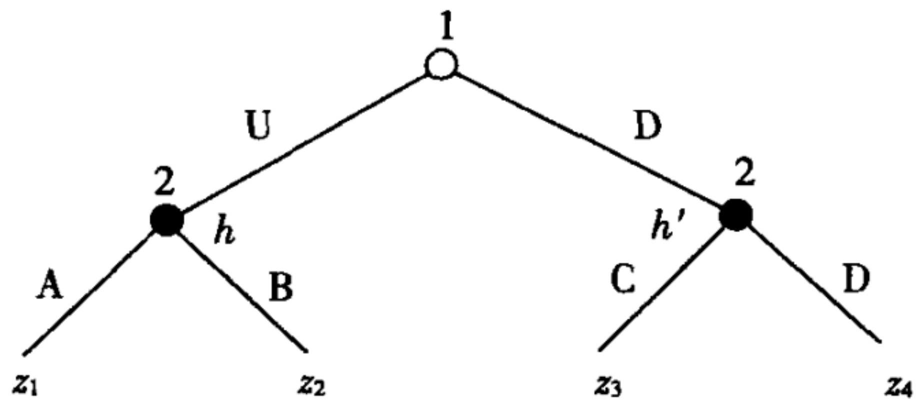
纯策略 s_2



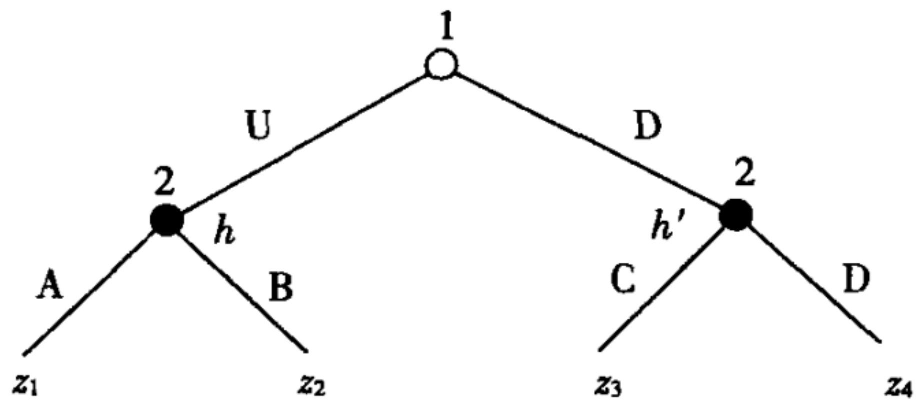
纯策略 $\bar{s}_2$

等价的行为策略为：$b_2(L \mid h_2) = b_2(R \mid h_2) = \frac{1}{2}$；$b_2(L' \mid h_2{}') = 0, b_2(R' \mid h_2{}') = 1$；$b_2(L'' \mid h_2{}'') = b_2(R'' \mid h_2{}'') = \frac{1}{2}$。

然而，许多不同的混合策略可以生成同一个行为策略：考虑行为策略 $b_2 = \{(\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2})\}$，即在每一个信息集上参与人 2 都以概率 $\frac{1}{2}$ 选择行动 L 或 R。



然而，许多不同的混合策略可以生成同一个行为策略：考虑行为策略 $b_2 = \{(\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2})\}$，即在每一个信息集上参与人 2 都以概率 $\frac{1}{2}$ 选择行动 L 或 R。



令 $\sigma_2 = (\sigma_{21}, \sigma_{22}, \sigma_{23}, \sigma_{24})$ 为参与人 2 的混合策略（四项分别对应 $(L, L), (L, R), (R, L), (R, R)$）。则有

$$\sigma_{21} + \sigma_{22} = \frac{1}{2}, \sigma_{23} + \sigma_{24} = \frac{1}{2}, \sigma_{21} + \sigma_{23} = \frac{1}{2}, \sigma_{22} + \sigma_{24} = \frac{1}{2}$$

会发现 $\sigma_2 = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ 和 $\sigma_2 = (\frac{1}{2}, 0, 0, \frac{1}{2})$ 都能生成 b_2。

下面讨论扩展式博弈纳什均衡的存在性。首先对于完美记忆有限扩展式博弈，根据库恩定理和纳什定理可得必然存在混合策略纳什均衡和行为策略纳什均衡。下面的问题是，是否存在纯策略纳什均衡呢？我们从最经典的国际象棋有无制胜策略的问题开始。为了描述接下来的关键定理，首先给出如下定义：

定义

令 Γ 为一个扩展式博弈，参与人为 1 和 2，结果集为 $O = \{1\text{胜}, 2\text{胜}, \text{和局}\}$。若满足下式，则称参与人 1 的策略 s_1 是制胜策略：

$$u(s_1, s_2) = 1\text{胜}, \forall s_2 \in S_2$$

若满足下式则称参与人 1 的策略 s_1 是至少保证和局的策略：

$$u(s_1, s_2) \in \{1\text{胜}, \text{和局}\}, \forall s_2 \in S_2$$

同样地，我们可以定义参与人 2 的制胜策略和至少保证和局的策略。有了如上定义，可以给出 Zermelo 定理：

Zermelo（策梅洛）定理

在任何完美信息的有限二人扩展式博弈中，如果结果集为 $O = \{1\text{胜}, 2\text{胜}, \text{和局}\}$，那么下面三种情况有且仅有一种成立：

1. 参与人 1 有制胜策略；
2. 参与人 2 有制胜策略；
3. 参与人 1 和 2 都有至少保证和局的策略。

- 事实上，根据接下来的定理证明，我们只需要要求博弈的深度有限即可，但如果深度无限则定理可能不成立；
- 国际象棋（50 回合内必须分胜负）等博弈都符合这个定理的前提，因此三个条件中必有一种成立，但我们至今未知哪一种成立，也不知道如何计算这些策略。

Zermelo 定理的第一种证明

使用数学归纳法，对子博弈 $\Gamma(x)$ 的节点数 n_x 进行归纳：

1. $n_x = 1$，x 是叶节点，若结果是 1 胜，则 1 有制胜策略，2 胜则 2 有制胜策略，和局则 1 和 2 都有至少保证和局的策略；
2. 假设 $n_x > 1$，根据归纳假设，任意满足 $n_y < n_x$ 的子博弈 $\Gamma(y)$ 都满足定理的结论。不失一般性，假设 $J(x) = 1$，则 x 出发能到达的任何一个局势都满足 $n_y < n_x$，我们用 $C(x)$ 表示参与人 1 在 x 处一步能到达的所有局势；
   - 若存在 $y \in C(x)$，使得 $\Gamma(y)$ 中 1 有制胜策略，则 1 在 x 有制胜策略：在 x 处选择到达 y，然后按 y 的制胜策略行动；
   - 如果对任意 $y \in C(x)$，$\Gamma(y)$ 中 2 都有制胜策略，则 2 在 x 有制胜策略；
   - 否则，对任意 $y \in C(x)$，要么 $\Gamma(y)$ 中 1 和 2 有至少保证和局的策略，要么 $\Gamma(y)$ 中 2 有制胜策略，此时 1 和 2 不可能有制胜策略（为什么），1 可以选择 1 和 2 有至少保证和局的策略的 y 保证和局。

Zermelo 定理的第二种证明

对于有限博弈，存在一个自然数 K，使得在博弈的任何展开中，两个参与人的行动次数都不超过 K。如果实际展开深度不到 $2K$，可以增加“什么都不做”的行动，使得展开深度为 $2K$。

对任一 k，$1 \leqslant k \leqslant K$，用 a_k 表示参与人 1 的第 k 步行动，b_k 表示参与人 2 的第 k 步行动。我们用 F 表示 $2k$ 步后参与人 1 获胜，那么 $\neg F$ 表示 $2k$ 步后参与人 2 获胜或者和局。于是参与人 1 有制胜策略’‘可以表达为：

$$\exists a_1 \forall b_1 \exists a_2 \forall b_2 \ldots \exists a_K \forall b_K (F)$$

“参与人 1 没有制胜策略”可以表达为：

$$\neg(\exists a_1 \forall b_1 \exists a_2 \forall b_2 \ldots \exists a_K \forall b_K (F))$$

Zermelo 定理的第二种证明

这等价于 $\forall a_1 \exists b_1 \forall a_2 \exists b_2 \ldots \forall a_K \exists b_K (\neg F)$，即“参与人 2 有至少保证和局的策略”。

由此证明了如果参与人 1 没有制胜策略，则参与人 2 有至少保证和局的策略，同理如果参与人 2 没有制胜策略，则参与人 1 有至少保证和局的策略，命题得证。

Zermelo 定理的推广

事实上策梅洛定理可以进一步理解为，完美信息的有限二人扩展式博弈如果满足定理条件，必定存在纯策略纳什均衡。例如参与人 1 有必胜策略时，1 的必胜策略和 2 的任意纯策略组成一个纯策略纳什均衡，因为双方都不会偏离这个策略。

运用与第一种证明类似的方法，可以证明 Zermelo 定理的推广：

Zermelo（策梅洛）定理的推广（Kuhn, 1953）

有限完美信息博弈存在纯策略纳什均衡。

具体的操作需要从次终点节点（即直接后继节点是叶节点）开始，每次选择最大化自己收益的行动，直到根节点。在之后的讨论中，这一方法被称为**逆向归纳法（backward induction）**。

CONTENT

目录

子博弈完美均衡

当一个博弈存在不止一个均衡时，我们希望基于合理的选择标准选择一些均衡，而剔除另一些均衡，这样的一个选择叫做**均衡精炼（equilibrium refinements）**。

定义

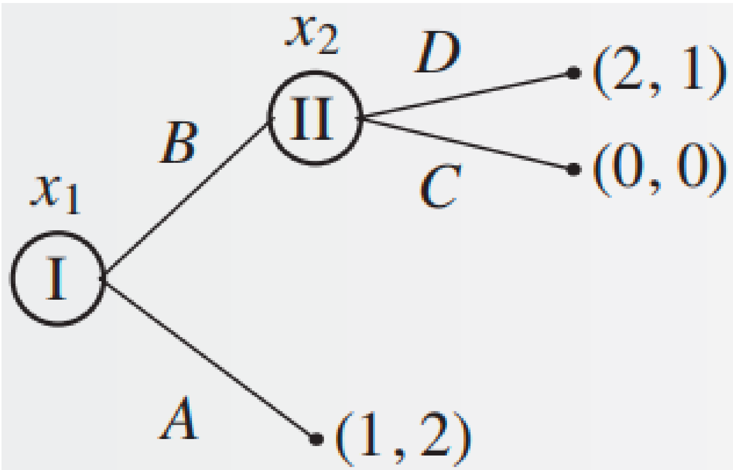
在扩展式博弈 Γ 中，一个策略向量 σ^* 是**子博弈完美均衡（subgame perfect equilibrium）**，如果对于博弈的任意子博弈 $\Gamma(x)$，局限在那个子博弈的策略向量 σ^* 是 $\Gamma(x)$ 的纳什均衡：对每个参与人 i，每个策略 σ_i 和子博弈 $\Gamma(x)$，

$$u_i(\sigma^* \mid x) \geqslant u_i(\sigma_i, \sigma^*_{-i} \mid x)$$

这一定义是很直观的，因为如果某个子博弈 $\Gamma(x)$ 上参与人存在有利可图的偏离，那么全局来看这也是一个有利可图的偏离。

子博弈完美均衡的例子

是否存在不是子博弈完美均衡的纳什均衡？考虑下图二人博弈：



		Player II: C	Player II: D
Player I	A	1, 2	1, 2
	B	0, 0	2, 1

1. 这一博弈有两个纯策略纳什均衡：(A, C) 和 (B, D)，参与人 I 更偏好 (B, D)，参与人 II 更偏好 (A, C)；
2. (A, C) 不是子博弈完美均衡，因为在 x_2 处参与人 II 存在有利可图的偏离：选择 D 而不是 C（因此子博弈完美均衡的确是纳什均衡的精炼）；
3. 在 (A, C) 下，I 不会偏离均衡，是因为 II **威胁** I：如果你选择 B，我就选择 C，然而这个威胁显然是不可置信的，因为如果 I 选择 B，那么 II 还是选择 D 更有利。

子博弈完美均衡的充分条件

例子中 (A, C) 能作为均衡，或者说 C 这一被 D 占优的策略可以成为均衡，是因为 (A, C) 到不了真正要选择 C, D 的 x_2 点。

用 $P_\sigma(x)$ 表示当实施策略向量 σ 时，博弈展开将造访节点 x 的概率。有如下定理：

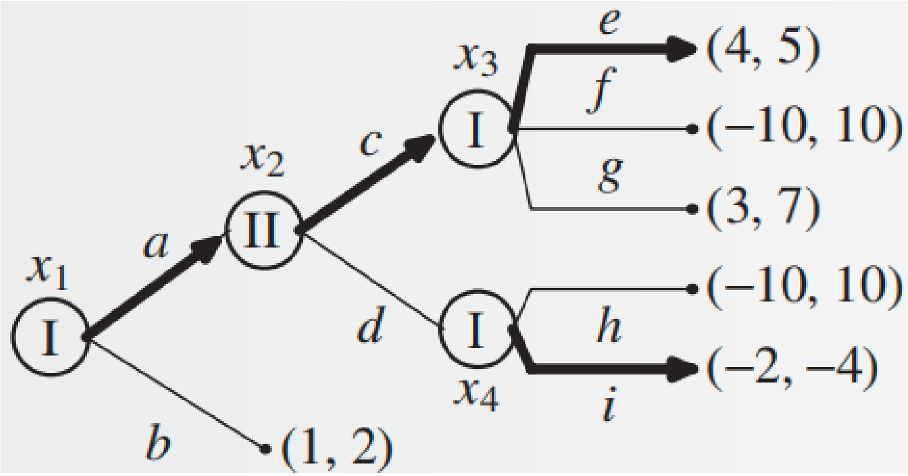
定理

令 σ^* 是扩展式博弈 Γ 的纳什均衡，如果对所有 x 都有 $P_{\sigma^*}(x) > 0$，那么 σ^* 是子博弈完美均衡。

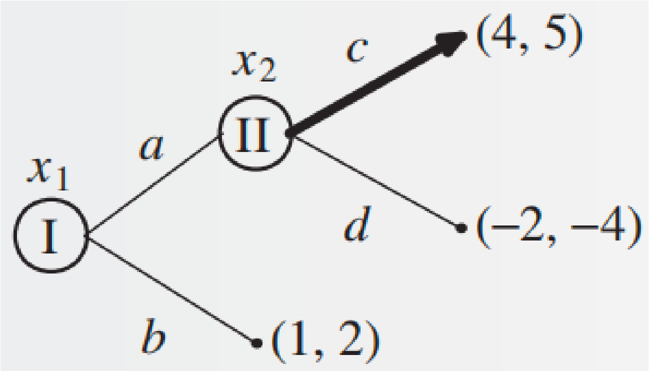
- 定理是显然的，因为如果 σ^* 不是子博弈完美均衡，那么在某个子博弈 $\Gamma(x)$ 上存在有利可图的偏离，并且这个偏离产生的概率不为 0，因此也可以带来全局的有利可图的偏离；
- 推论：完全混合的纳什均衡是子博弈完美均衡。

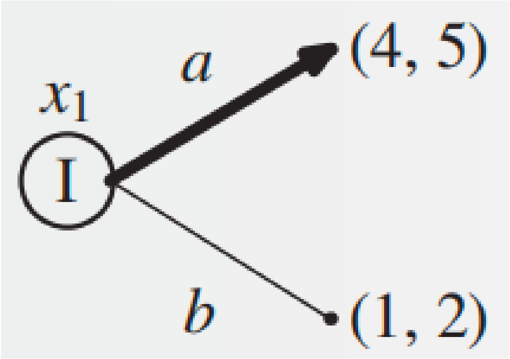
逆向归纳法

如何找到完美信息博弈的子博弈完美均衡？可使用逆向归纳法：



从最小的子博弈出发，即 $\Gamma(x_3)$ 和 $\Gamma(x_4)$，选择图中加粗的策略（子博弈的均衡），然后将均衡结果替代子博弈，逐步向上推导到根节点即可（因此子博弈完美均衡是 (ae, c)）





逆向归纳法

这一方法称为**逆向归纳法（backward induction）**，该方法的应用保证了每一个子博弈都使用了均衡策略，并且每一步都能做出选择，由此可以得到 Zermelo 定理的拓展：

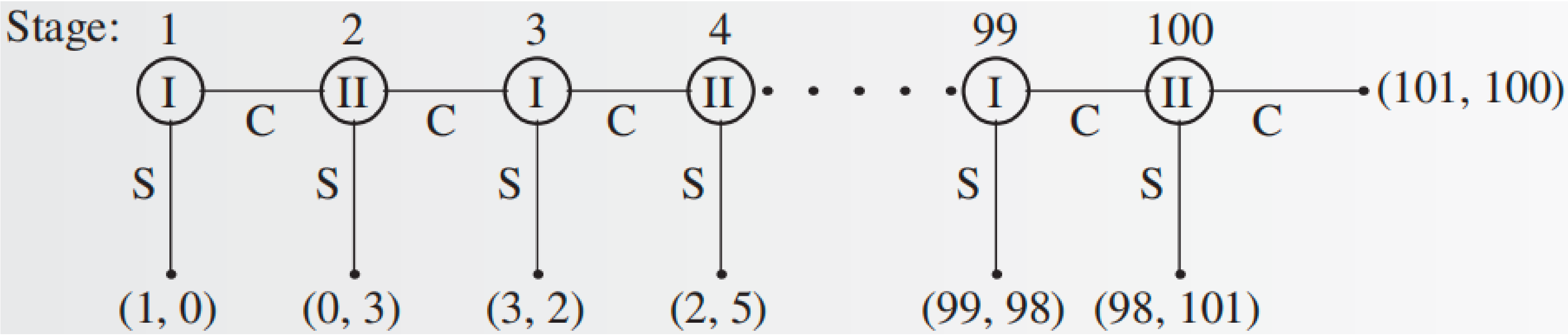
定理

每个有限完美信息扩展式博弈都至少有一个子博弈完美纯策略均衡。

- 石头剪刀布没有纯策略纳什均衡，因为这并不是完美信息的扩展式博弈；
- 然而，逆向归纳法存在局限性：
  - 有时候不是子博弈完美均衡的均衡可能更好；
  - 重复囚徒困境有限轮，逆向归纳法会得到两个囚徒仍然在每一轮都选择承认（因此我们需要新的博弈建模方式描述人们在长期关系中会合作这一事实）。

逆向归纳法的局限

接下来来看一个例子，称之为**蜈蚣博弈**：两个参与人依次行动：在奇数轮 $t = 1, 3, ..., 99$，参与人 1 选择停止博弈（S）或者继续博弈（C），如果他在第 t 轮选择停止，收益为 $(t, t-1)$，否则继续博弈；在偶数轮 $t = 2, 4, ..., 100$，参与人 2 选择停止博弈（S）或者继续博弈（C），如果他在第 t 轮选择停止，收益为 $(t-2, t+1)$，否则继续。如果最初 99 轮没人停止，那么 100 轮后博弈结束，双方收益为 $(101, 100)$。下图表现了为什么这一博弈被称为蜈蚣博弈：

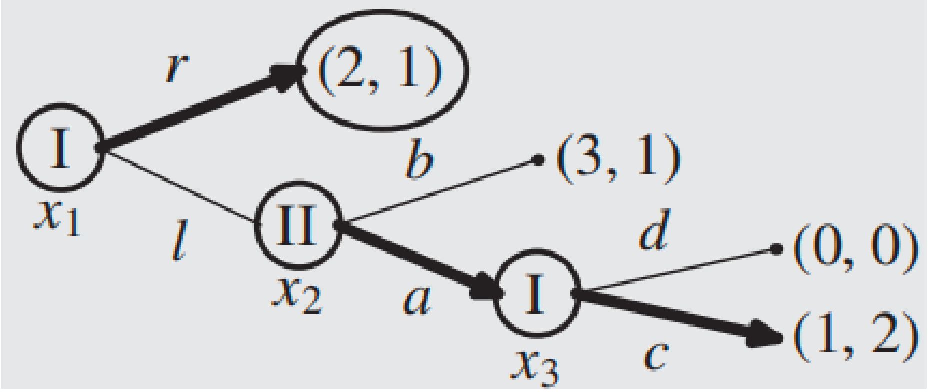


显然，逆向归纳法的结果是参与人在第一轮就要选择停止，但现实中通常双方都会试探前进一段才会结束。

正向归纳法

以上例子表明，逆向归纳不足以描述理性：它忽视了到达这一节点的历史，但历史本身提供了关于其他参与人行为的信息：

- 重复囚徒困境中，对方前几个阶段都选择不承认，是否在提醒我也应该选择不承认？
- 蜈蚣博弈到第二阶段意味着对方选择继续，那我也应该继续？



逆向归纳法解为 (rc, a)，但若参与人 1 选择 l，这意味着什么？

- I 不理性，II 担心选了 a 后 I 会选择 d 得到更差的结果；
- 或许选 l 是 I 精心设计的让 II 觉得 I 不理性的坑？然后 I 就可以拿到 3 的收益了；
- 颤抖手均衡等角度改进子博弈完美均衡。

产量领导模型（斯塔克尔伯格模型）

经济学中的子博弈完美均衡最基本的应用就是产量领导模型（或称**斯塔克尔伯格（Stackelberg）**模型）。斯塔克尔伯格模型常用于描述有一家厂商处于支配地位或充当自然领导者的行业。例如 IBM 是具有支配地位的行业，通常观察到的其它小企业的行为模式是等待 IBM 宣布新产量然后调整自己的产量决策，此时 IBM 就是斯塔克尔伯格领导者，其它厂商是跟随者。

设厂商 1 是领导者，选择产量 y_1；厂商 2 是跟随者，选择产量 y_2。用 $p(y_1+y_2)$ 表示总产量为 y_1+y_2 时的市场价格，$c_1(y_1)$ 和 $c_2(y_2)$ 表示厂商 1 和 2 在生产 y_1 和 y_2 单位商品时的成本。

产量领导模型（斯塔克尔伯格模型）

因此厂商 1 的利润最大化决策为:

$$\max_{y_1} \pi_1(y_1, y_2) = p(y_1 + y_2)y_1 - c_1(y_1)$$

$$\text{s.t.} y_2 = \arg\max_{y_2} \pi_2(y_1, y_2) = p(y_1 + y_2)y_2 - c_2(y_2)$$

- 这是一个**双层优化问题（bi-level optimization problem）**，即优化的约束条件是另一个优化问题：厂商 1 在做决策时，他知道厂商 2 会根据他的决策做出最优反应；
- 因此应当先求解厂商 2 的最优反应函数 $y_2^* = f_2(y_1)$，然后将其代入厂商 1 的利润函数中，求解厂商 1 的最优产量 y_1^*；
- 或者说，厂商 1 计算出自己任意产量下对方的最优选择，在此基础上看自己的最优选择，因此这是逆向归纳法的应用，解得的是斯塔克尔伯格博弈的子博弈完美均衡。