

博弈论

第四讲：扩展式博弈（二）

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/>

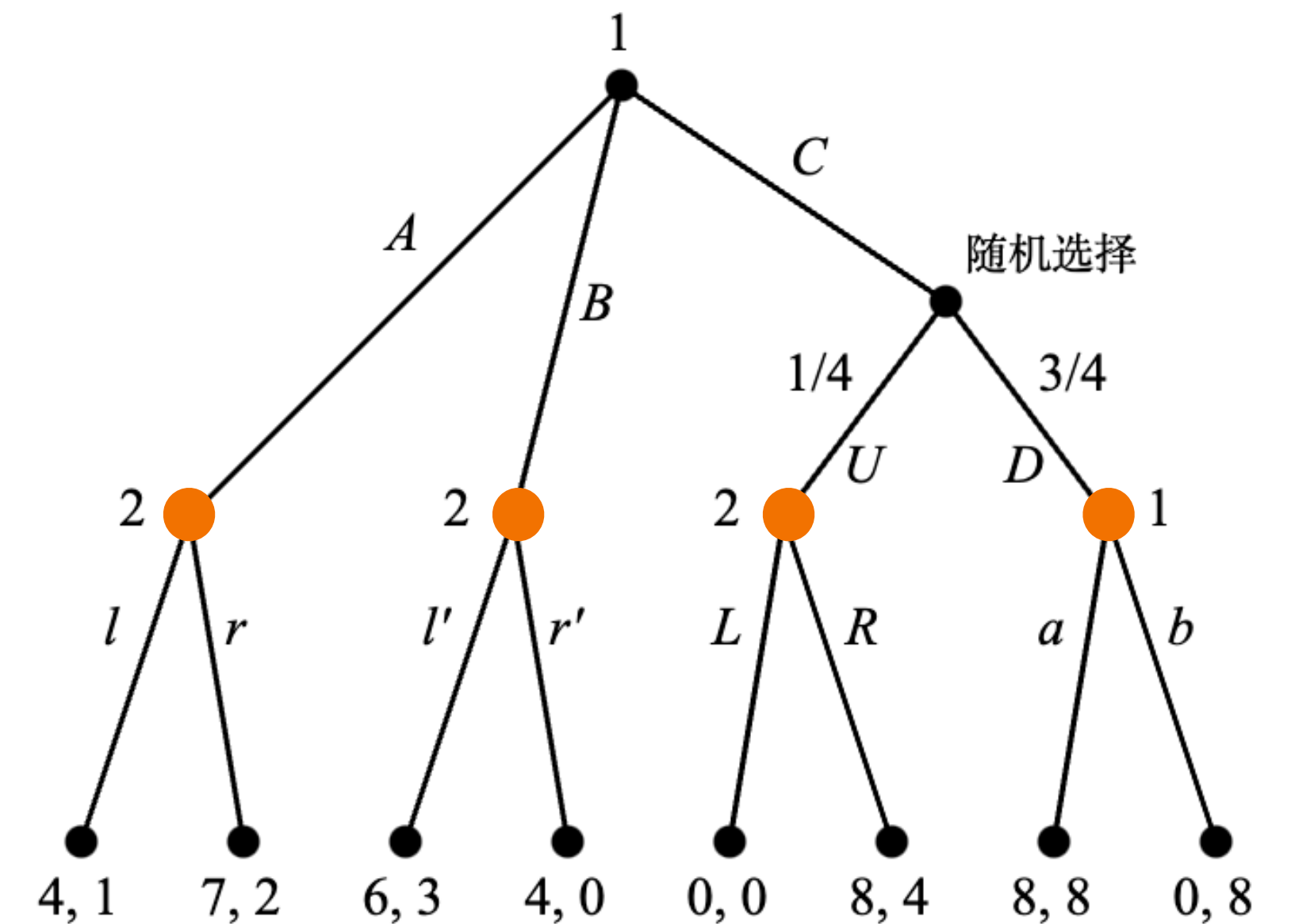
子博弈完美均衡

逆向归纳法

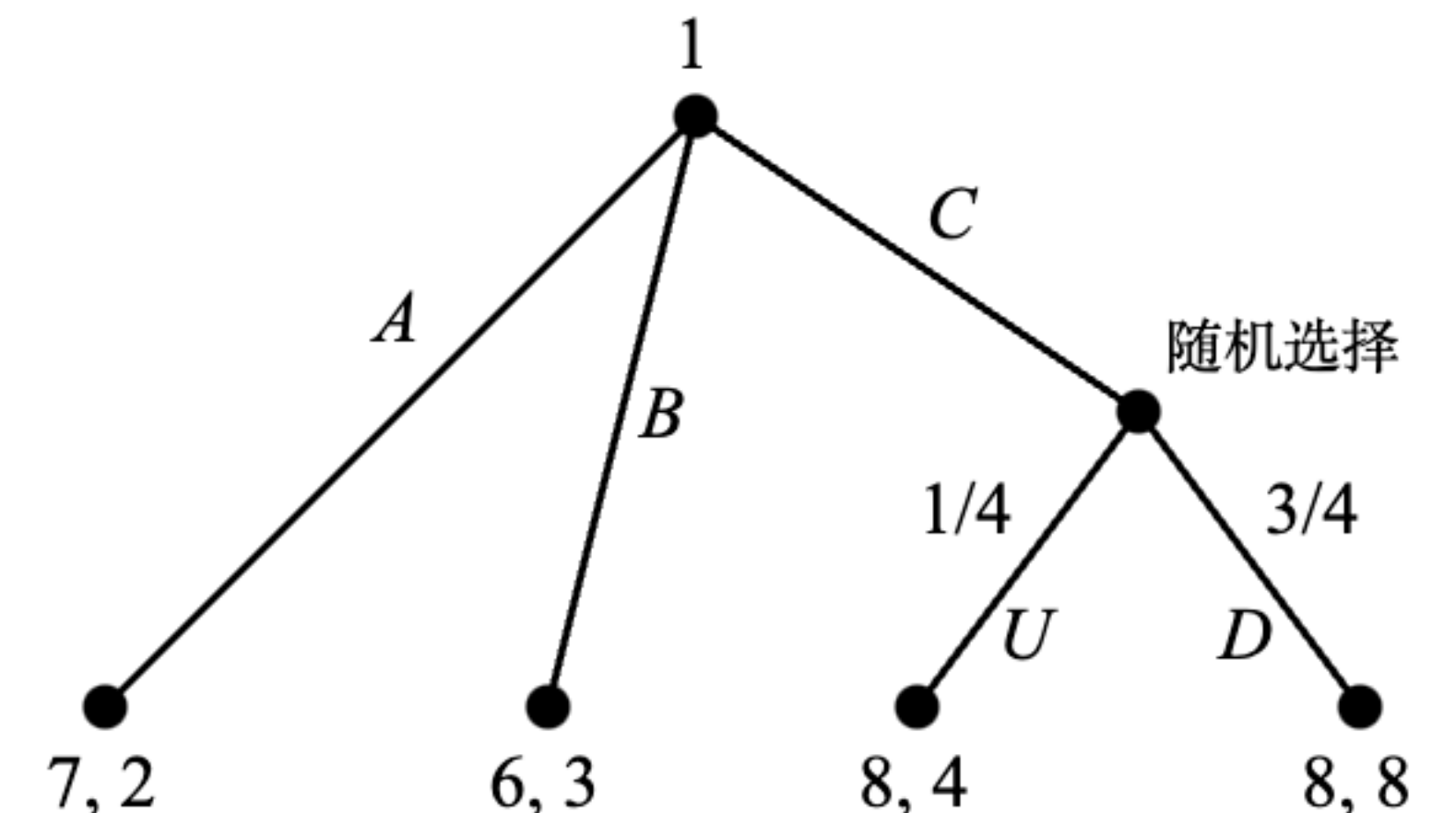
逆向归纳法 (backward induction) 的步骤如下:

- 针对每一个终点前的参与人决策节点 (如右上图中的●), 找出该节点的最优行动, 然后将该节点改为终点, 并赋予最优行动对应的收益向量。
- 通过以上操作我们可以获得一个简化的博弈 (如右下图)。针对新的博弈重复上面的操作, 直至仅剩一个节点。
- 这样获得的策略组合 (在右边的例子中为 $(Ca, r\ell'R)$) 是纯策略纳什均衡。

在完美信息博弈中, 用逆向归纳法一定会得到一个纯策略纳什均衡。因此完美信息博弈至少拥有一个纯策略纳什均衡。

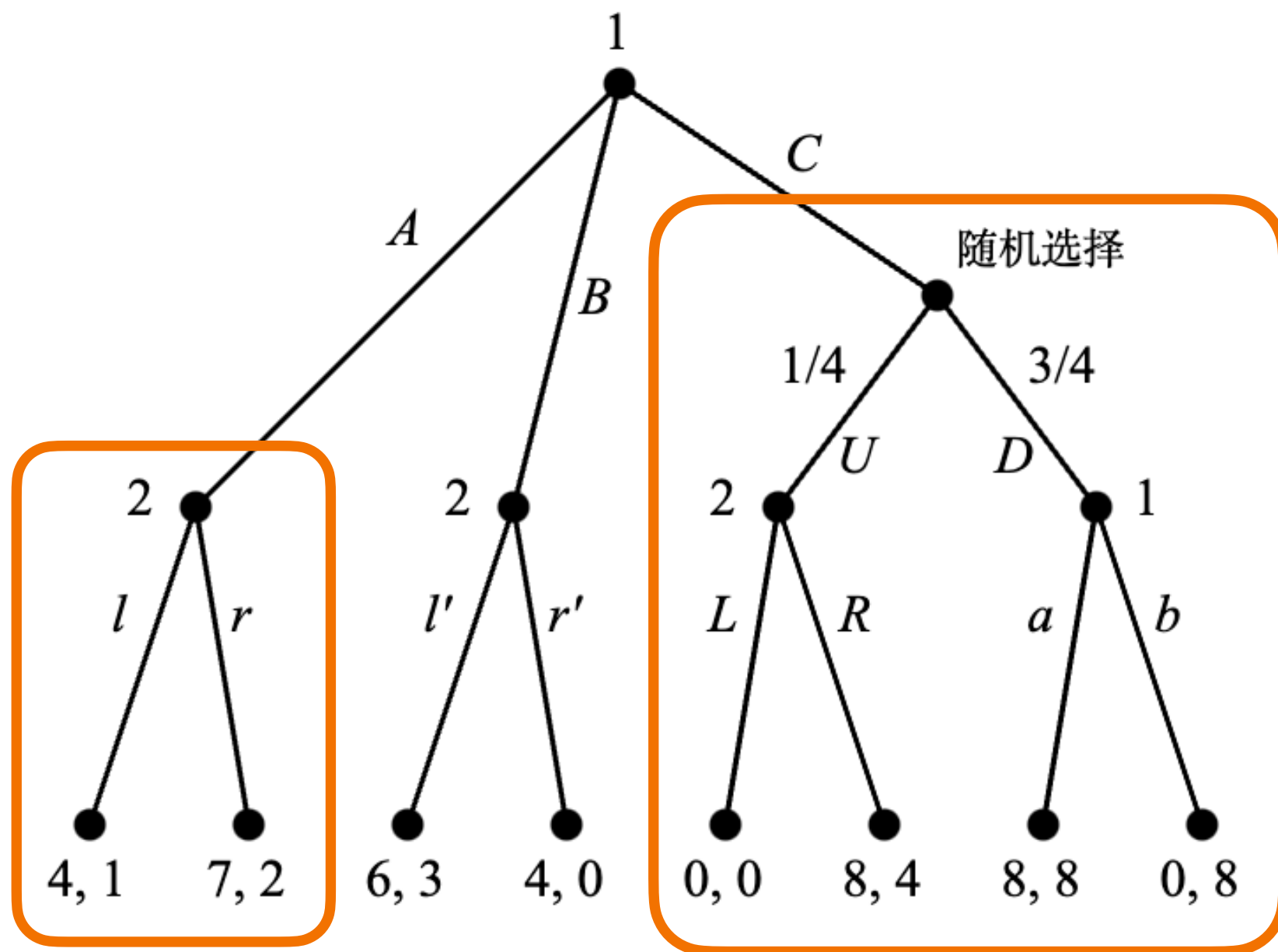


这个博弈有十个纯策略纳什均衡

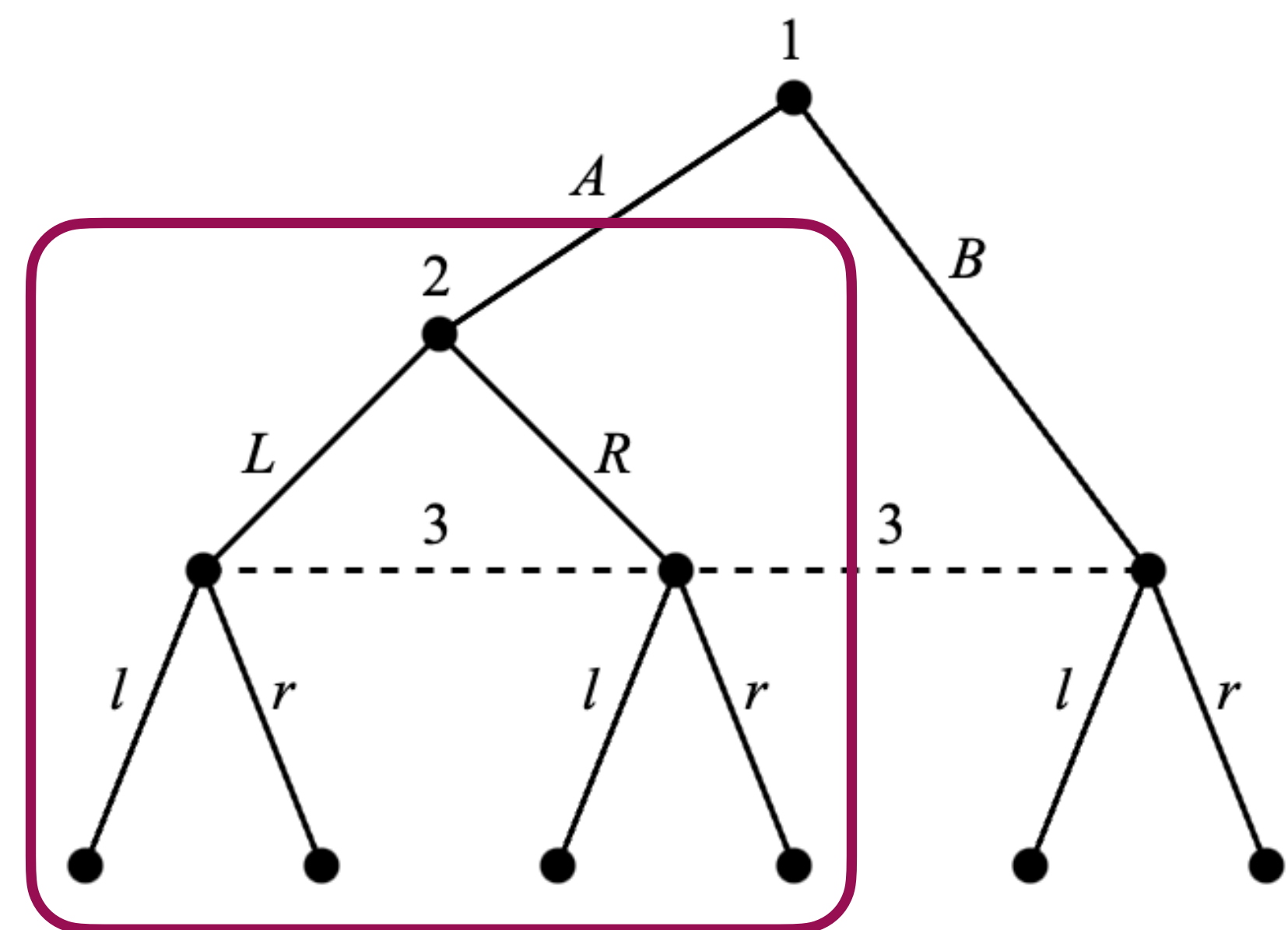


子博弈

子博弈 (subgame) 是博弈树中的一个平凡信息集及其以下的部分，并且这一部分中的任意节点都不会和外部的节点处在同一个信息集中。



子博弈的例子



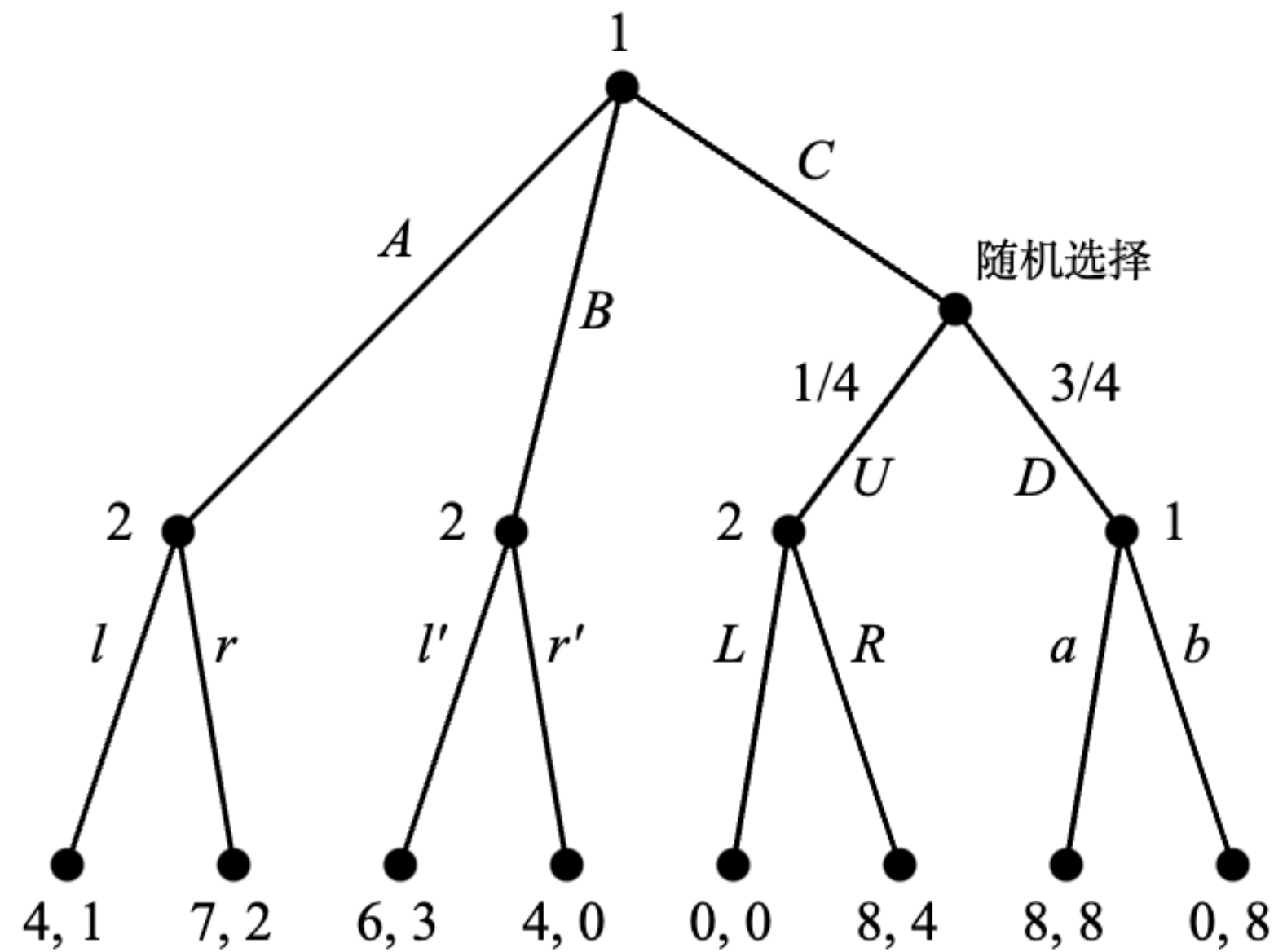
这不是子博弈

子博弈完美均衡

如果一个策略组合能在每个子博弈中形成一个纳什均衡，则称其为子博弈完美均衡（**subgame perfect equilibrium, SPE**）。

- 任意子博弈完美均衡也是纳什均衡。
- 在完美信息博弈中，由逆向归纳法获得的纳什均衡和子博弈完美均衡是一致的。因此完美信息（扩展式）博弈一定存在子博弈完美均衡。

子博弈完美均衡一定程度上解决了纳什均衡过多的问题。



这个博弈共包含六个子博弈（原博弈也是自身的子博弈）。

- 逆向归纳均衡 $(Ca, r\ell'R)$ 在最下面的四个子博弈中都做出了最优选择。
- 随机选择是随机的，因此不存在优劣之说。
- $(Ca, r\ell'R)$ 是原博弈的纳什均衡。

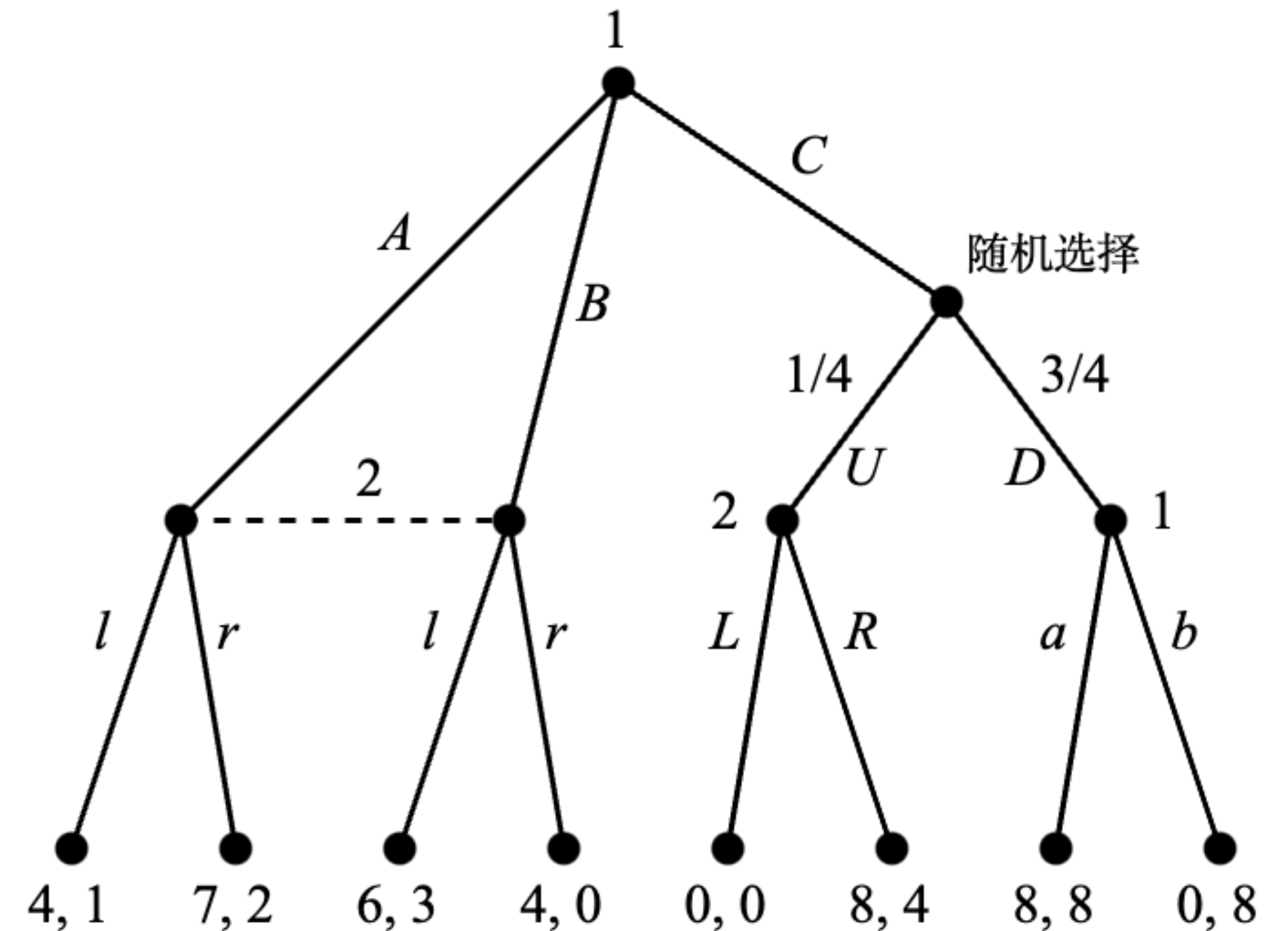
不完美信息博弈的子博弈完美均衡

- 右侧的博弈有四个子博弈。
- 对于最后两个子博弈，最优选择分别是 R 和 a 。
- 忽略由随机选择节点开始的子博弈。
- 已知原博弈共有六个纳什均衡：

$$(Aa, rL), (Ab, rL), (Ba, \ell L), \\ (Bb, \ell L), (Ca, \ell R), (Ca, rR)$$

其中 $(Ca, \ell R)$ 和 (Ca, rR) 是子博弈完美均衡。

- 注意：在这个博弈中无法使用逆向归纳法。



取火柴游戏的子博弈完美均衡策略

这是一个完美信息博弈。虽然很难完整地画出这个博弈的博弈树，但是可以用逆向归纳法进行分析。

- ▶ 桌上有 21 根火柴
- ▶ 两个参与人轮流拿走一部分火柴，但每次只能拿走 1, 2 或 3 根
- ▶ 参与人 1 首先开始
- ▶ 拿走最后一根火柴的参与人获胜

- 最后一轮：如果桌上剩下 1, 2 或 3 根火柴，则下一个取火柴的参与人的最优策略是拿走全部火柴并获胜。
- 如果桌上有 4 根火柴，则下一个取火柴的参与人只能选择拿走 1, 2 或 3 根火柴，而这都会导致对手获胜。因此，参与人如果能够给桌上留下 4 根火柴就能保证自己的胜利。
- 如果桌上有 5, 6 或 7 根火柴，则下一个取火柴的参与人可以保证桌上剩下 4 根火柴，从而保障自己的胜利。
- 如果桌上有 8 根火柴，则下一个取火柴的参与人只能留下 5, 6 或 7 根火柴，而这都会导致对手获胜。因此，参与人如果能够给桌上留下 8 根火柴就能保证自己的胜利。
- 依此类推，参与人应当努力使桌上留下 4, 8, 12, 16 或 20 根火柴。
- 第一轮：第一个行动的参与人只要选择取走 1 根火柴，就能保证自己可以获胜。在后面的回合里，他只需要根据对手的行动调整自己的选择，使桌上依次留下 16, 12, 8, 4 根火柴，并在最后一轮取走所有剩余的火柴。

练习

在 MOBA 类游戏中选择英雄

考虑 MOBA 类游戏（例如王者荣耀）中两队选择英雄的环节（即 ban/pick 中的 pick 环节），并将其简化建模。

- 假设两个队伍分别为参与人 1 和 2。
- 共有四个英雄可供选择，分别为 H_1, H_2, H_3 和 H_4 。参与人对每个英雄的熟练度为

	H_1	H_2	H_3	H_4
参与人 1	1	2	3	4
参与人 2	2	3	4	1

- 参与人 1 首先选择一名英雄，随后参与人 2 选择一名英雄，最后参与人 1 再选择一名英雄。剩下的英雄由参与人 2 获得。
- 参与人的收益是其获得英雄的熟练度之和。

回答下面的问题：

1. 画出博弈树。
2. 找出子博弈完美均衡。
3. 是否存在和子博弈完美均衡结果不同的纳什均衡？

完美贝叶斯均衡

完美贝叶斯均衡

子博弈完美均衡对非平凡信息集不起作用，因此我们需要一个适用于不完美信息博弈的针对纳什均衡的改进，下面介绍的**完美贝叶斯均衡 (perfect Bayesian equilibrium)** 就是其中一种。

对于一个非平凡信息集，在该信息集做决策的参与人的**信念 (belief)** 是关于该信息集内节点的概率分布（概率向量）。可以理解为参与人对信息集内节点的发生概率的预测。

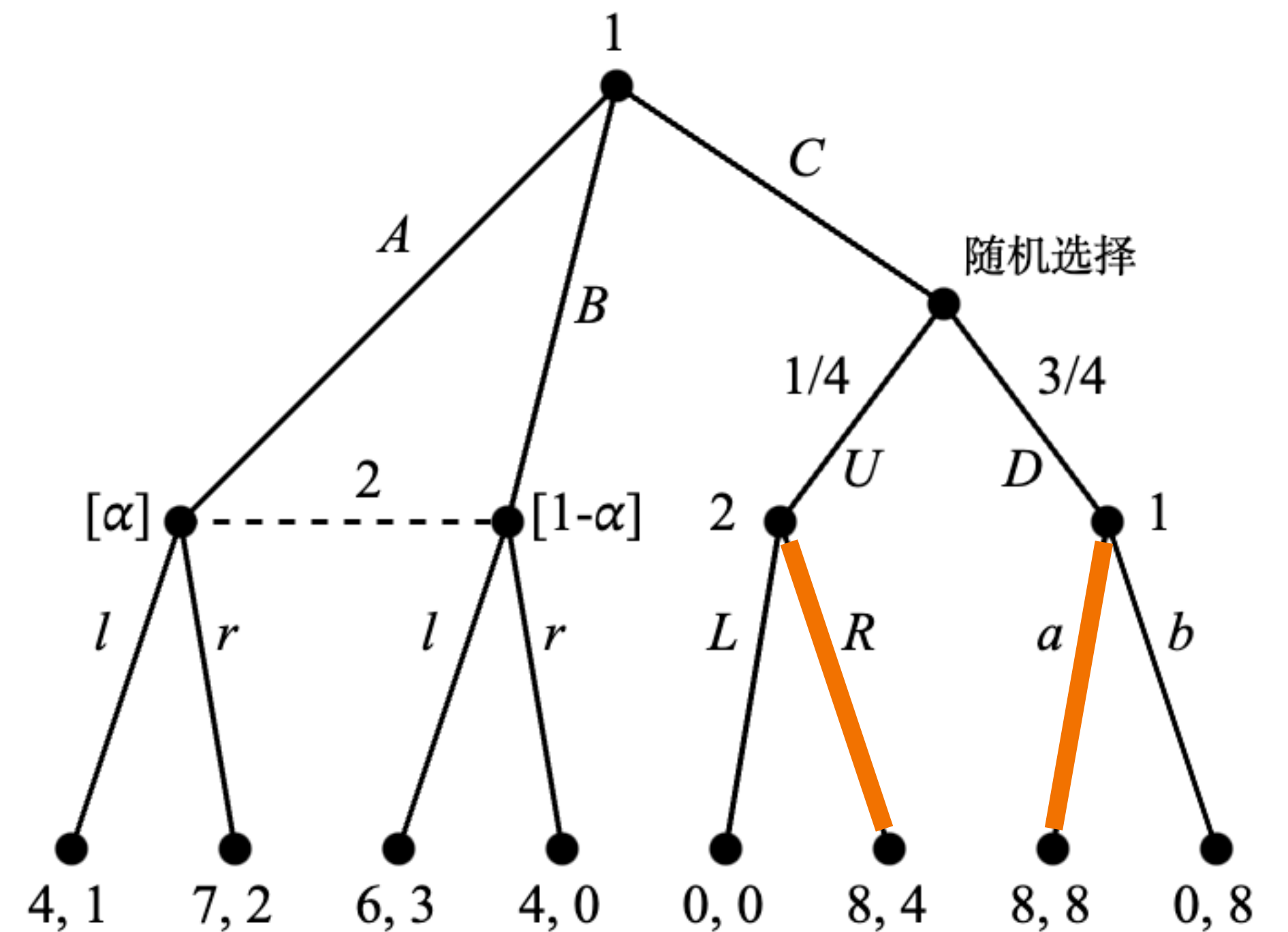
扩展式博弈中的完美贝叶斯均衡是满足下面两个条件的策略与信念的组合。

1. **贝叶斯一致性 (Bayesian consistency)**：信念需和策略保持一致，即在任何可行的情况下依照贝叶斯公式计算信念。
2. **序贯理性 (sequential rationality)**：参与人在各自的信念下选择最优行动，即根据信念计算期望收益，并根据期望收益最大化原则进行决策。

信念

这里我们对右侧博弈中参与人 2 的信息集进行了标注

- 这个信息集共有两个节点。令左侧节点的发生概率为 α ，右侧节点的发生概率为 $1 - \alpha$ ，则参与人 2 在这个信息集上的信念为 $(\alpha, 1 - \alpha)$ 。
- 根据序贯理性，在这个信念下，参与人 2 选择 ℓ 的期望收益是 $1\alpha + 3(1 - \alpha) = 3 - 2\alpha$ ，选择 r 的期望收益是 $2\alpha + 0(1 - \alpha) = 2\alpha$ 。因此，当 $\alpha \leq 3/4$ 时选择 ℓ ，当 $\alpha \geq 3/4$ 时选择 r 。（我们不讨论混合策略，因此在 $\alpha = 3/4$ 时只有两个收益相同的纯策略）
- 同样根据序贯理性，参与人 2 在 U 下面的节点选择 R ，参与人 1 在 D 下面的节点选择 a 。

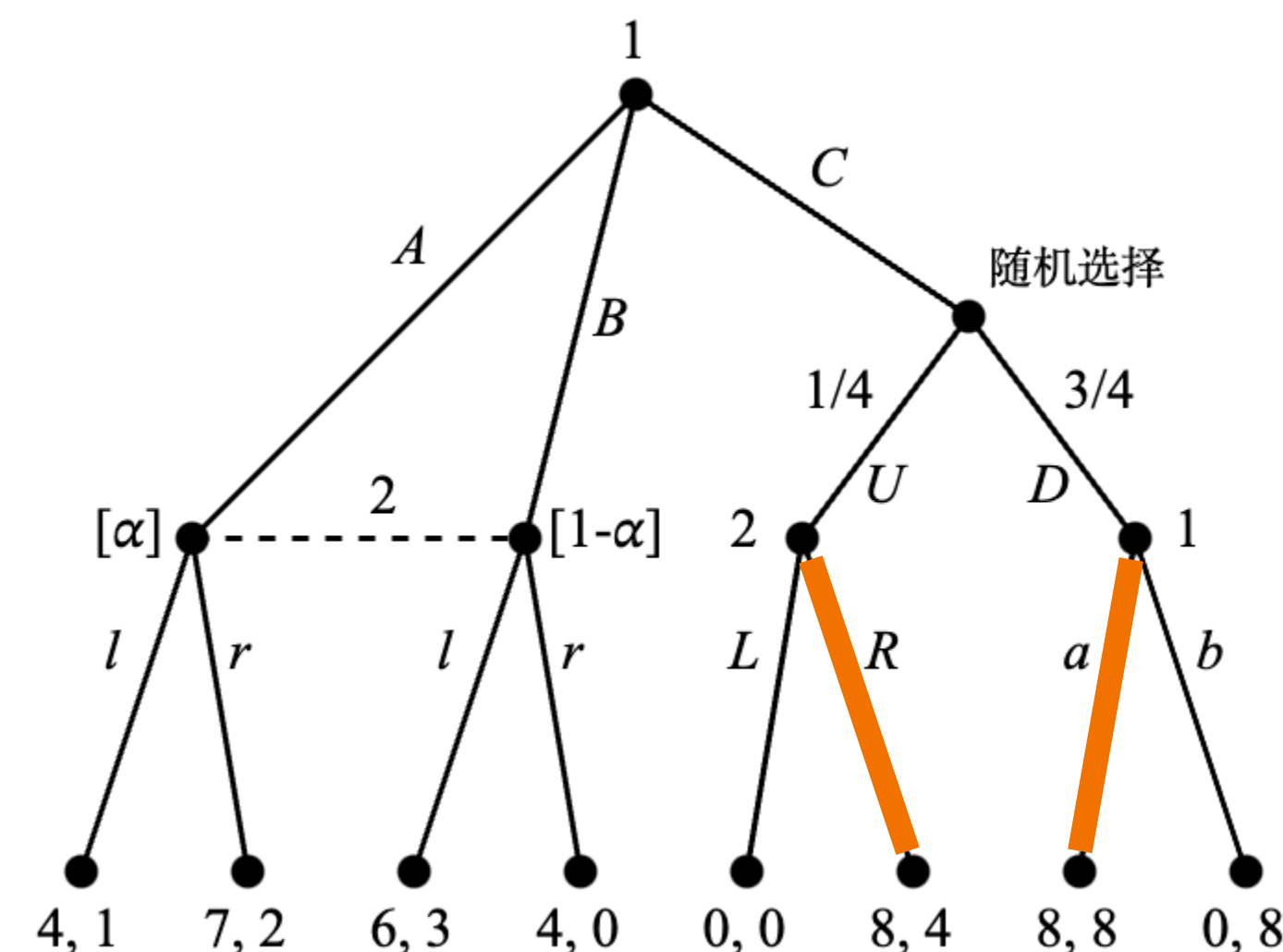


信念的计算

贝叶斯公式：

$$\Pr[P | Q] = \frac{\Pr[P] \times \Pr[Q | P]}{\Pr[Q]}$$

- 那么如何计算 α 呢？根据贝叶斯一致性，我们应该尽可能地利用贝叶斯公式



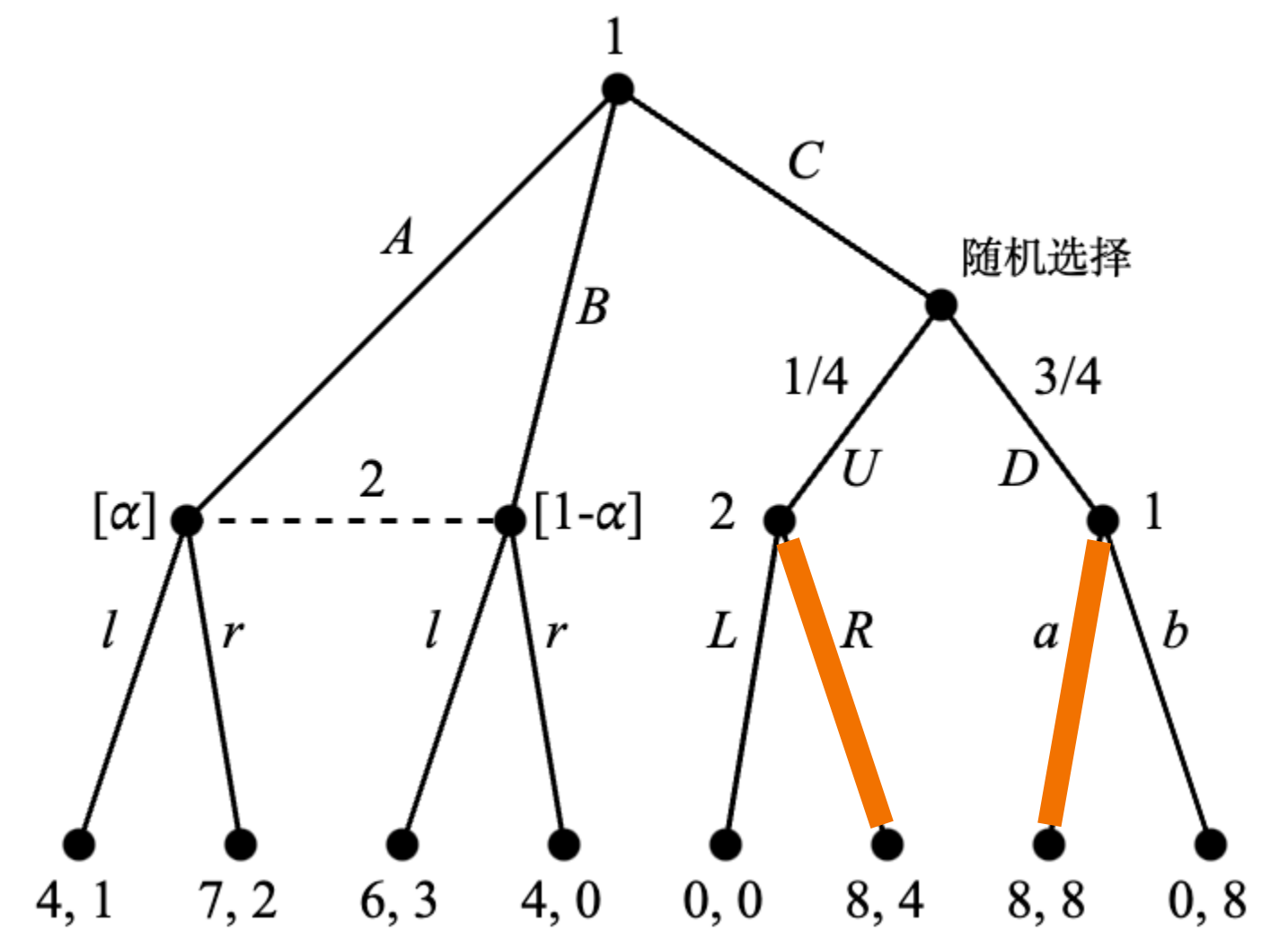
$$\begin{aligned} \alpha &= \Pr[\text{参与人 1 选择了 } A \mid \text{参与人 1 选择了 } A \text{ 或 } B] \\ &= \frac{\Pr[\text{参与人 1 选择了 } A] \times \Pr[\text{参与人 1 选择了 } A \text{ 或 } B \mid \text{参与人 1 选择了 } A]}{\Pr[\text{参与人 1 选择了 } A \text{ 或 } B]} \\ &= \frac{\Pr[\text{参与人 1 选择了 } A]}{\Pr[\text{参与人 1 选择了 } A \text{ 或 } B]} \end{aligned}$$

- α 的取值取决于参与人 1 的策略。如果他的策略让他选择 A ，则 $\alpha = 1$ ，如果他的策略让他选择 B ，则 $\alpha = 0$ 。但如果他的策略让他选择 C ，则上面的表达式分母为零，说明此时贝叶斯公式不适用，而 α 可以取任意值（在贝叶斯公式不适用时，参与人可以随意预测发生概率）。

完美贝叶斯均衡

- 在起点处，参与人 1（在已知参与人 2 在 U 后面会选择 R 的前提下）选择 C 的期望收益是 8，而选择 A 或 B 的期望收益则不会大于 7。因此，参与人 1 的最优策略是 Ca 。
- 在这种情况下，参与人 2 的非平凡信息集无法到达，因此贝叶斯一致性没有约束力。也就是允许参与人 2 拥有任意信念，即 $0 \leq \alpha \leq 1$ 。
- 序贯理性依然有效，因此参与人 2 的最优策略为：当 $0 \leq \alpha \leq 3/4$ 时选择策略 ℓR ，当 $3/4 \leq \alpha \leq 1$ 时选择策略 rR 。
- 综上所述，此博弈的完美贝叶斯均衡是

$$\begin{cases} (Ca, \ell R) & \text{if } 0 \leq \alpha \leq 3/4 \\ (Ca, rR) & \text{if } 3/4 \leq \alpha \leq 1 \end{cases}$$



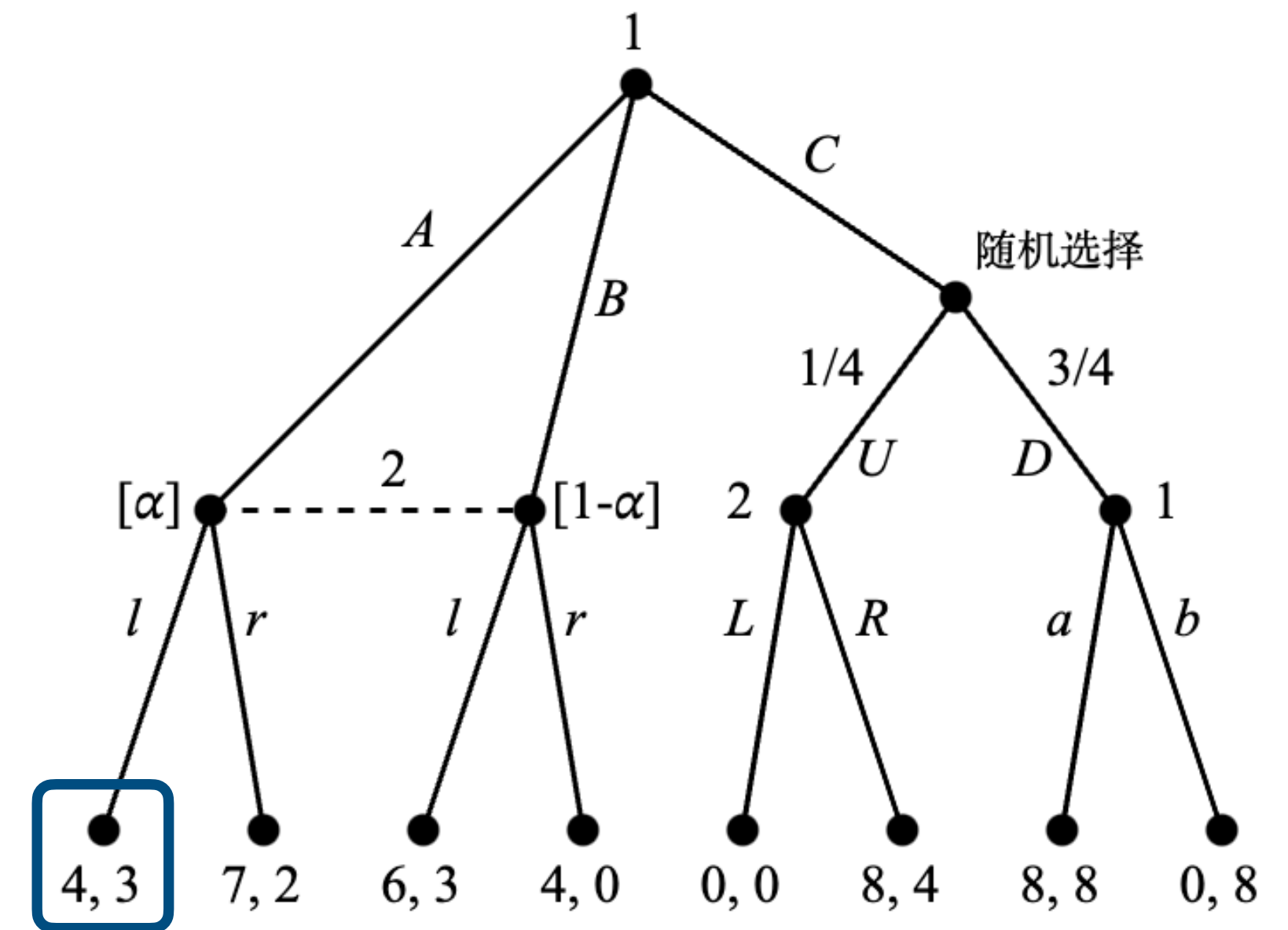
注意：($Ca, \ell R$) 和 (Ca, rR) 都是是这个博弈的子博弈完美均衡，完美贝叶斯均衡在此基础上添加了信念的支撑。

完美贝叶斯均衡和子博弈完美均衡

右图是将前一个博弈中最左侧终点的收益由 $(4, 1)$ 改为 $(4, 3)$ 而获得的博弈。

- 子博弈完美均衡依然是 $(Ca, \ell R)$ 和 (Ca, rR) 。
- 当参与人 2 处于非平凡信息集时，无论是在 A 节点还是 B 节点，她选择 ℓ 的收益都大于选择 r 的收益。因此， rR 并不是最优策略，而子博弈完美均衡无法将其排除。
- 如果考虑完美贝叶斯均衡，那么在信念 $(\alpha, 1 - \alpha)$ 下，选择 ℓ 的期望收益是 $3\alpha + 3(1 - \alpha) = 3$ ，选择 r 的期望收益是 2α ，因此对于任意 $0 \leq \alpha \leq 1$ ，序贯理性都要求参与人 2 选择 ℓ 而非 r 。因此，唯一的完美贝叶斯均衡是

$$(Ca, \ell R) \quad \text{for any } 0 \leq \alpha \leq 1$$



一个三人博弈的分析

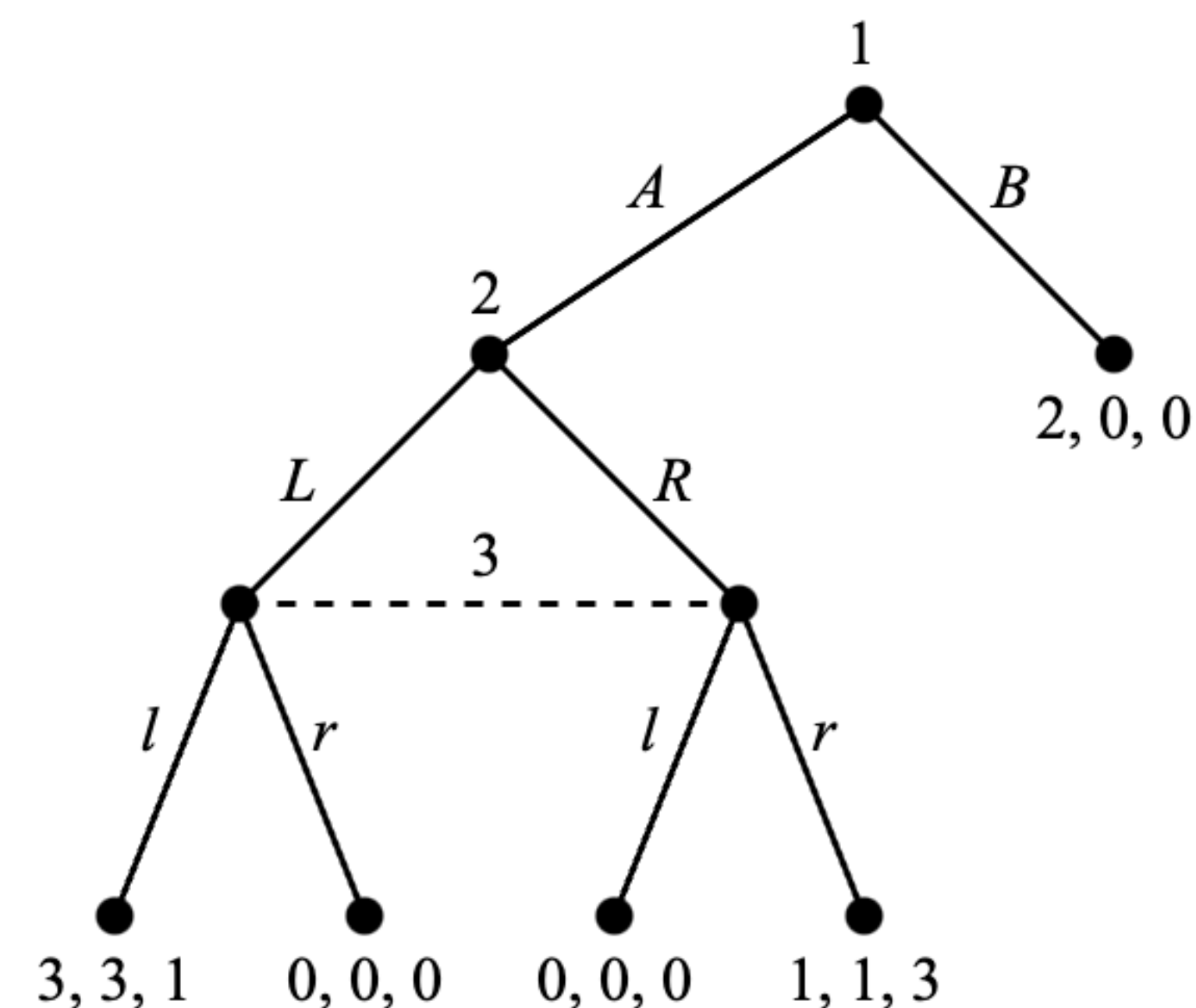
右图中给出了一个三人博弈

- 纳什均衡：

由于有三个参与人，策略式表达需要多个双矩阵

$$1 : A \quad \begin{array}{cc} \ell & r \\ L \left(\underline{3}, \underline{3}, \underline{1} & 0, 0, 0 \right) \\ R \left(0, 0, 0 & \underline{1}, \underline{1}, \underline{3} \right) \end{array} \quad 1 : B \quad \begin{array}{cc} \ell & r \\ L \left(\underline{2}, \underline{0}, \underline{0} & \underline{2}, \underline{0}, \underline{0} \right) \\ R \left(\underline{2}, \underline{0}, \underline{0} & \underline{2}, \underline{0}, \underline{0} \right) \end{array}$$

纯策略纳什均衡为 (A, L, ℓ) , (B, L, r) , (B, R, ℓ) 和 (B, R, r) 。



一个三人博弈的分析

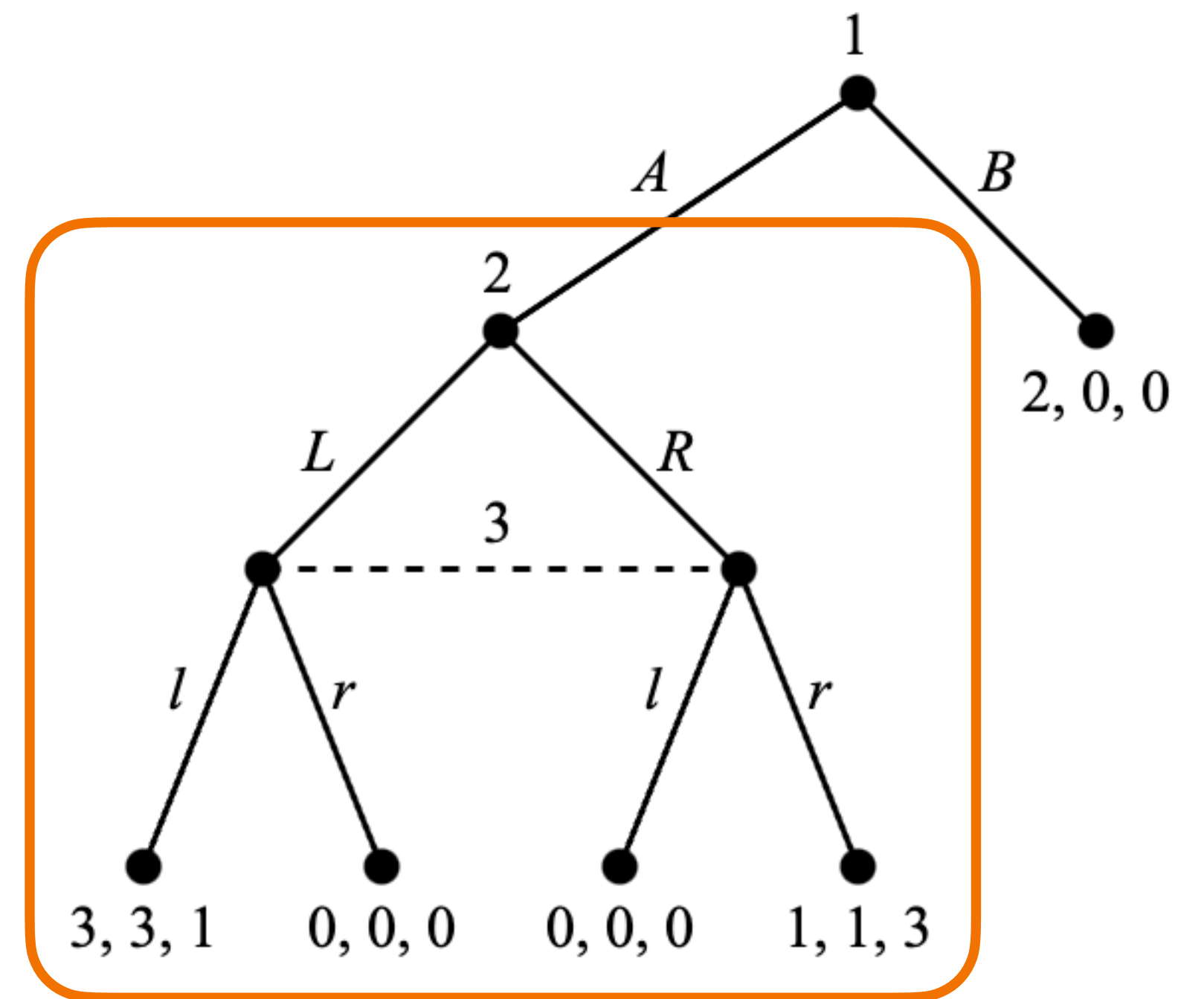
- 子博弈完美均衡：

除了原博弈，还有一个子博弈（如右图）。这个子博弈的纳什均衡是它对应的策略式博弈的纳什均衡，即

$$\begin{array}{cc} & \ell & r \\ L & (\underline{3}, \underline{1} & 0, 0) \\ R & 0, 0 & (\underline{1}, \underline{3}) \end{array}$$

均衡策略为 (L, ℓ) 和 (R, r) 。

在上面的子博弈中，如果参与人 2 和 3 选择 (L, ℓ) ，则参与人 1 应当选择 A ；如果是 (R, r) ，则参与人 1 应该选择 B 。因此，子博弈完美均衡是 (A, L, ℓ) 和 (B, R, r) 。



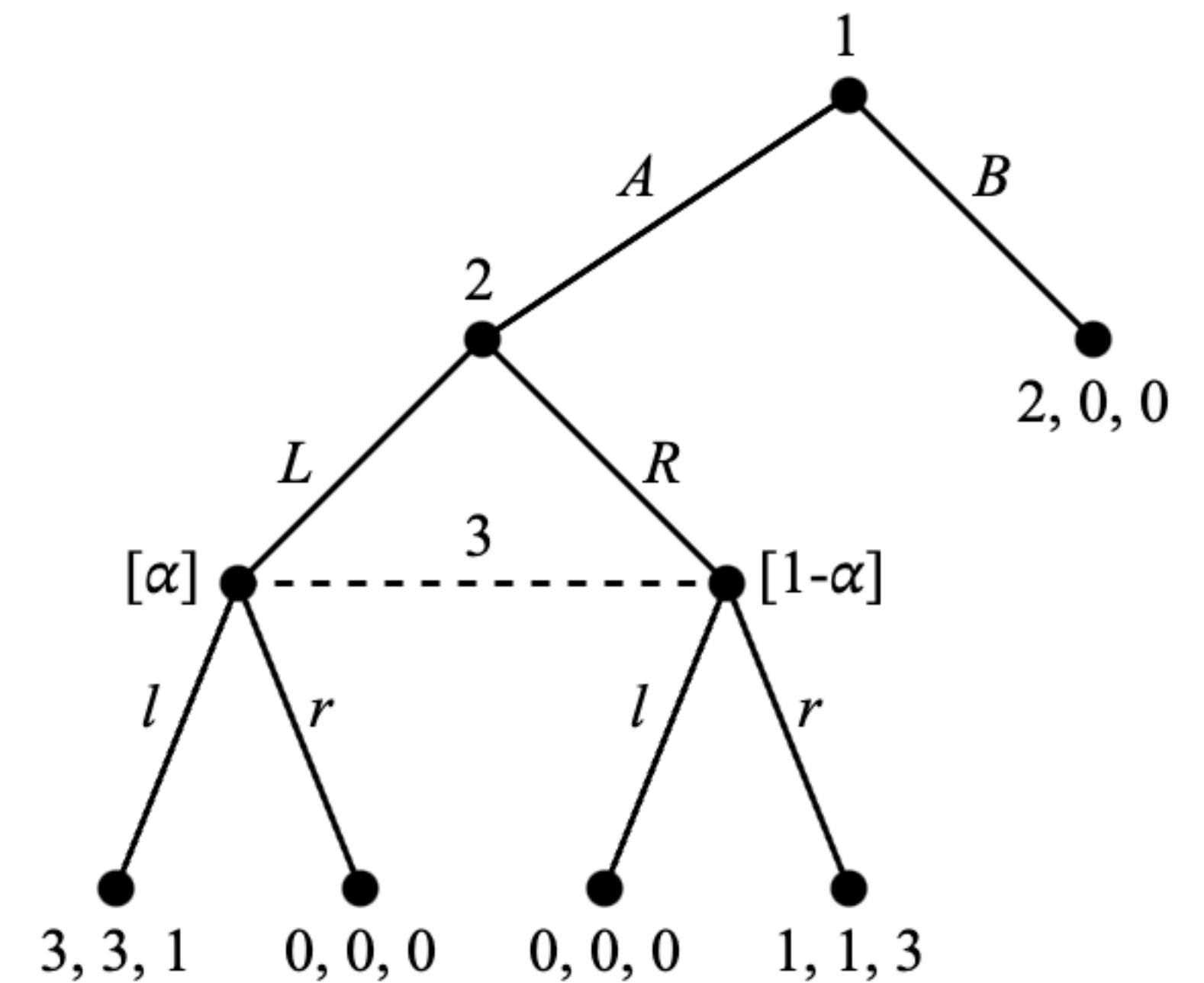
一个三人博弈的分析

- 完美贝叶斯均衡：

令参与人 3 在自己的信息集中的信念为 $(\alpha, 1 - \alpha)$ 。此时行动 ℓ 和 r 的期望收益分别是 α 和 $3 - 3\alpha$ ，根据序贯理性，参与人 3 的最优选择是

$$\begin{cases} r & \text{if } 0 \leq \alpha \leq 3/4 \\ \ell & \text{if } 3/4 \leq \alpha \leq 1 \end{cases}$$

- 如果参与人 3 选择 ℓ ，参与人 2 应当选择 L ，参与人 1 应当选择 A 。此时参与人 3 的信念是 $\alpha = 1$ ，且基于此信念，他的最优选择是 ℓ ，符合贝叶斯一致性（没有自相矛盾）。这样我们就得到了一个完美贝叶斯均衡：策略组合 (A, L, ℓ) 和信念 $\alpha = 1$ 。



一个三人博弈的分析

- 完美贝叶斯均衡：

令参与人 3 在自己的信息集中的信念为 $(\alpha, 1 - \alpha)$ 。此时行动 ℓ 和 r 的期望收益分别是 α 和 $3 - 3\alpha$ ，根据序贯理性，参与人 3 的最优选择是

$$\begin{cases} r & \text{if } 0 \leq \alpha \leq 3/4 \\ \ell & \text{if } 3/4 \leq \alpha \leq 1 \end{cases}$$

- 如果参与人 3 选择 r ，参与人 2 应当选择 R ，参与人 1 应当选择 B 。此时参与人 3 的信念是 $\alpha = 0$ ，且基于此信念，他的最优选择是 r ，符合贝叶斯一致性。这样我们就得到了另一个完美贝叶斯均衡：策略组合 (B, R, r) 和信念 $\alpha = 0$ 。

