中国科学院大学·人工智能学院·本科生专业课

博弈论

上课时间：每周一晚上18:10-19:50

上课地点：玉泉路校区 教学楼阶一1

授课团队：兴军亮（教师）、徐航（助教）

上一讲内容回顾：完全信息动态博弈

- **信息不完全和不完美**
- **海萨尼转换**
- **定义及表示**
- **贝叶斯纳什均衡**

知识点

能力线

- **问题的相互转换能力**：信息不完全与不完美
- **贝叶斯学派思维方式**：概率思维看待问题，将无解问题到有解

- **不完全信息静态博弈**：如何估计对手类型？

① 获取对手的类型分布

② 期望均衡收益最大化

价值面

2023春季学期 · 本科生专业课 · 《博弈论》

第五讲：不完全信息动态博弈

兴军亮

授课时间：2024年3月25日

联系方式：jlxing@nlpr.ia.ac.cn

第五讲 不完全信息动态博弈

内容提纲

1. 博弈类型引入与介绍
2. 博弈类型定义及表示
3. 博弈均衡分析与求解
4. 博弈应用案例与讲解

第五讲 不完全信息动态博弈

内容提纲

回顾：常见的博弈类型

行动次序 / 信息	静态	动态
完全信息	完全信息静态博弈 纳什均衡	完全信息动态博弈 子博弈精练纳什均衡
不完全信息	不完全信息动态博弈 贝叶斯均衡	不完全信息动态博弈 精炼贝叶斯均衡

重复博弈
Repeated Games

随机博弈
Stochastic Game

回顾：完全信息静态博弈的纳什均衡

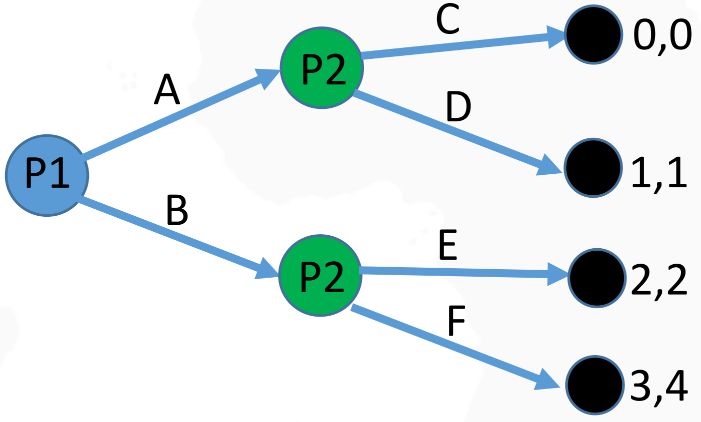
- 固定一方的策略，求取另一方的最优反应，最优反应的交集就是纯策略纳什均衡

囚徒B 囚徒A	坦白	抵赖
坦白	2,2	0,3
抵赖	3,0	1,1

0,0	0,0	0,0	0,0
0,0	0,0	-1,1	-1,1
0,0	1,-1	0,0	-1,1
0,0	1,-1	1,-1	0,0

回顾：完全信息动态博弈的纳什均衡

- 完全信息动态博弈一般用扩展式表示
  - 扩展式表示可以转化为矩阵式表示
  - 然后用前面的方法求纳什均衡



扩展式表示

S_1 ={A,B}

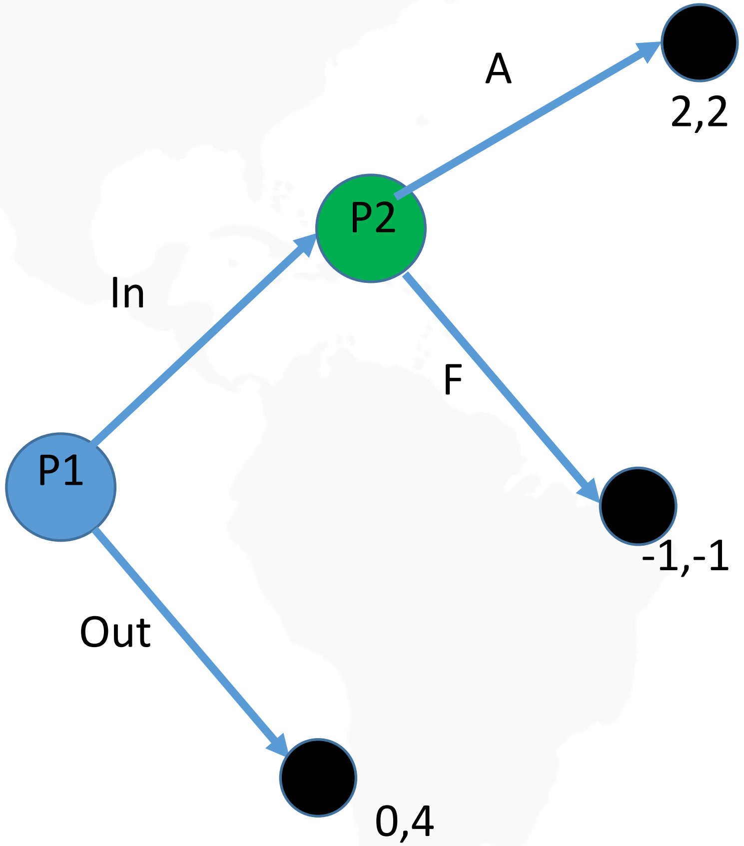
S_2 ={CE, CF, DE, DF}

	CE	CF	DE	DF
A	0,0	0,0	1,1	1,1
B	2,2	3,4	2,2	3,4

矩阵式表示

回顾：完全信息动态博弈的纳什均衡

- 需要剔除不合理的解

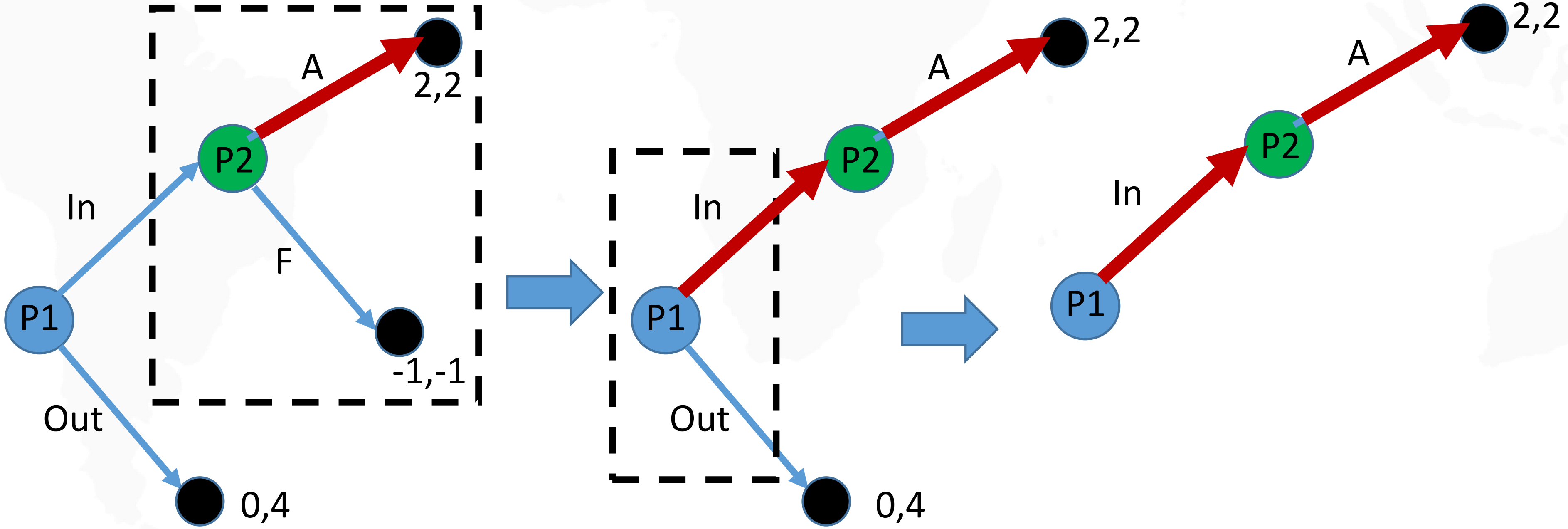


	A	F
In	2,2	-1,-1
Out	0,4	0,4

这两个纳什均衡
都合理吗？
Out/F不合理

回顾：完全信息动态博弈的纳什均衡

- 子博弈精炼纳什均衡（ Subgame Perfect Equilibrium ）
  - 子博弈可以粗略看作是博弈树的一棵子树
  - 每个子博弈都要满足纳什均衡条件
- 用逆向归纳法（Backward Induction）求解
  - 反方向不断求取各玩家的最优策略

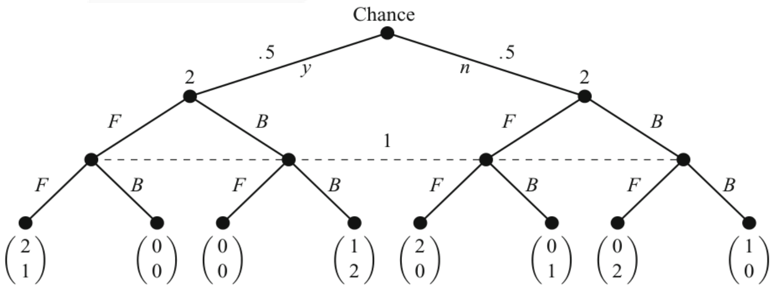


回顾：不完全信息静态博弈的纳什均衡

- 贝叶斯纳什均衡
  - A BNE is a set of strategies, one for each type of player, such that no type has incentive to change his or her strategy given the beliefs about the types and what the other types are doing.
- 求解示例：P1（男方）不知道P2（女方）是否想和他出去
  - P1一种类型，P2两种类型 y 和 n；P1知道P2属于 y 和 n 的概率各为0.5

P1 \ P2	*F*	*B*
F	2, 1	0, 0
B	0, 0	1,2

P1 \ P2	*F*	*B*
F	2, 0	0, 2
B	0, 1	1,0

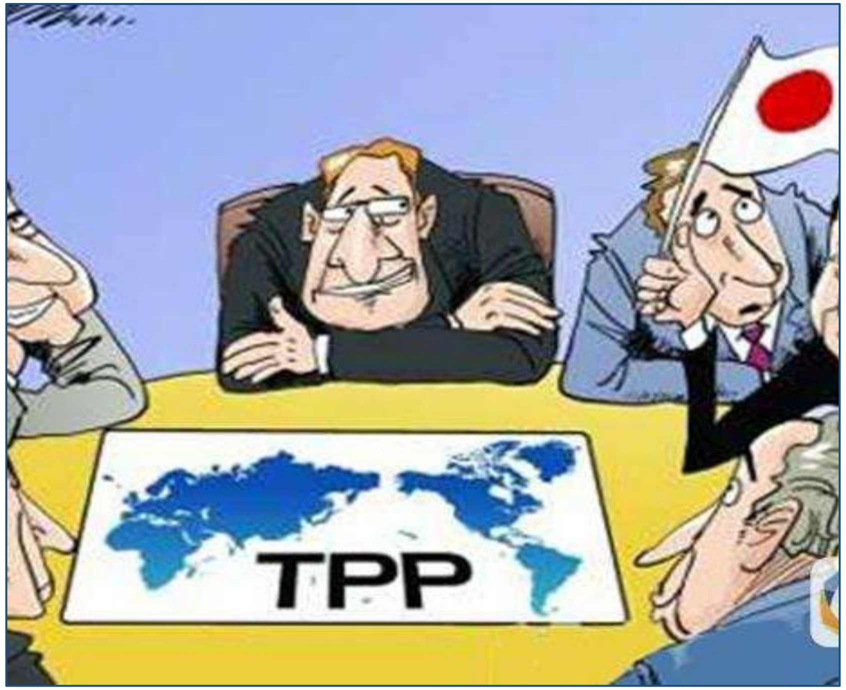


⇒

P1 \ P2	F^yF^n	F^yB^n	B^yF^n	B^yB^n
F	2, 0.5	1, 1.5	$\boxed{1,0}$	0,1
B	0, 0.5	0.5,0	0.5,1.5	1,1

不完全信息动态博弈的引入

- 实际博弈问题往往不仅存在不完全信息，而且会持续多个阶段。博弈参与者会在观察到先行动者所选择动作来推断其类型或修正对其类型的先验信息（概率分布），然后选择自己的最优行动。



如何根据对方行动修正对其类型的先验信息？

贝叶斯法则（Bayesian's Rule）

- 基本思想：当面临不确定性时，人们对某件事情发生的可能性有一个初步判断，然后人们会根据新到的**信息**来修正这一判断。
  - 修正之前的初步判断为**先验概率**，修正之后的判断为**后验概率**。
  - 贝叶斯法则是根据新信息从先验概率得到后验概率的基本方法。

Likelihood：似然函数
How probable is the evidence given that our hypothesis is true?

Prior：先验概率
How probable was our hypothesis before observing the evidence?

$$P(H|e) = \frac{P(e|H)P(H)}{P(e)}$$

Posterior：后验概率
How probable is our hypothesis given the observed evidence?

Marginal：边际分布
How probable is the new evidence under all possible hypotheses?

贝叶斯法则（Bayesian’s Rule）

- 贝叶斯法则举例分析：好人坏人的例子
  - 我们把所有的人分为好人（GP）和坏人（BP），所有的事分为好事（GT）和坏事（BT）
  - 那么一个人干好事的概率 p(GT) 等于：他是好人的概率p(GP)乘以好人做好事的概率 p(GT|GP) ，加上他是坏人的概率p(BP) 乘以坏人做好事的概率 p(GT|BP)

$$p(GT) = p(GT|GP)p(GP) + p(GT|BP)p(BP)$$

  - 假设我们观测到一个人干了一件好事，那么这个人是好人的后验概率是：

$$p(\mathrm{GP}|\mathrm{GT}) = \frac{p(\mathrm{GT}|\mathrm{GP})p(\mathrm{GP})}{p(\mathrm{GT})}$$

贝叶斯法则（Bayesian's Rule）

- 贝叶斯法则举例分析：好人坏人的例子
  - 假设我们认为一个人是好人的先验概率是1/2，那么观察到他干了一件好事之后，我们如何修正他是好人的先验概率依赖于 我们认为这件好事好到了什么程度。
  - 第一种情况：这是一件很好的好事，好人一定干，坏人不会干：
    - $p(\mathrm{GT|GP})=1$; $p(\mathrm{GT|BP})=0$; $p(GP|GT)=\frac{p(\mathrm{GT|GP})p(\mathrm{GP})}{p(\mathrm{GT})}=\frac{1\times1/2}{1\times1/2+0\times1/2}=1$
    - 尽管我们原来认为这个人是好人的概率是0.5，但观测到他干了这件好事后，我们会得出结论：他肯定是个好人。
  - 第二种情况：这是一件一般的好事，好人会干，坏人也会干：
    - $p(\mathrm{GT|GP})=1$; $p(\mathrm{GT|BP})=1$; $p(GP|GT)=\frac{p(\mathrm{GT|GP})p(\mathrm{GP})}{p(\mathrm{GT})}=\frac{1\times1/2}{1\times1/2+1\times1/2}=1/2$
    - 我们对他的看法不变

贝叶斯法则（Bayesian’s Rule）

- 贝叶斯法则举例分析：好人坏人的例子
  - 第三种情况：好人肯定会干，坏人可能干也可能不干：
    - $p(\mathrm{GT}|\mathrm{GP})=1, p(\mathrm{GT}|\mathrm{BP})=1/2$；$p(GP|GT)=\frac{p(\mathrm{GT}|\mathrm{GP})p(\mathrm{GP})}{p(\mathrm{GT})}=\frac{1\times1/2}{1\times1/2+1/2\times1/2}$=2/3
    - 我们认为他是好人的可能性增加了，但仍有1/3的概率为坏人。
  - 同理，当我们观测到一个人干了一件坏事
    - 如果我们相信好人绝对不会干坏事，只有坏人才干坏事
    - $p(\mathrm{BT}|\mathrm{GP})=0, p(\mathrm{BT}|\mathrm{BP})=p$；$p(GP|BT)=\frac{p(\mathrm{BT}|\mathrm{GP})p(\mathrm{GP})}{p(\mathrm{BT})}=\frac{0\times1/2}{0\times1/2+p\times1/2}=0$
  - 我们改变对一个人的看法
    - 不仅依赖于我们认为他是好人还是坏人的先验概率
    - 还依赖于好人干好事、坏人干好事的条件概率

贝叶斯法则（Bayesian's Rule）

- 贝叶斯法则的应用（以不完全信息博弈为例）
  - $\{\theta^k\}_{k=1}^K$为参与人i可能的类型集合，$\{a^h\}_{h=1}^H$为可能的行动集合
  - i属于θ^k的先验概率$p(\theta^k) \geq 0, \sum_{k=1}^K p(\theta^k) = 1$
  - 给定i属于θ^k，其选择a^h的条件概率为$p(a^h|\theta^k)$，选择a^h的边缘概率为$p(a^h) = \sum_{k=1}^K p\left(a^h \middle| \theta^k\right) p(\theta^k)$（全概率公式）
  - 贝叶斯法则推导过程是：假如观测到了参与人i选择了行动a^h，那么参与人i属于类型θ^k的后验概率为：

$$p\left(\theta^k \middle| a^h\right) = \frac{p(\theta^k, a^h)}{P(a^h)} = \frac{p\left(a^h \middle| \theta^k\right) p(\theta^k)}{\sum_{k=1}^K p(a^h|\theta^k) p(\theta^k)}$$

第五讲 不完全信息动态博弈

内容提纲

1 博弈类型引入与介绍

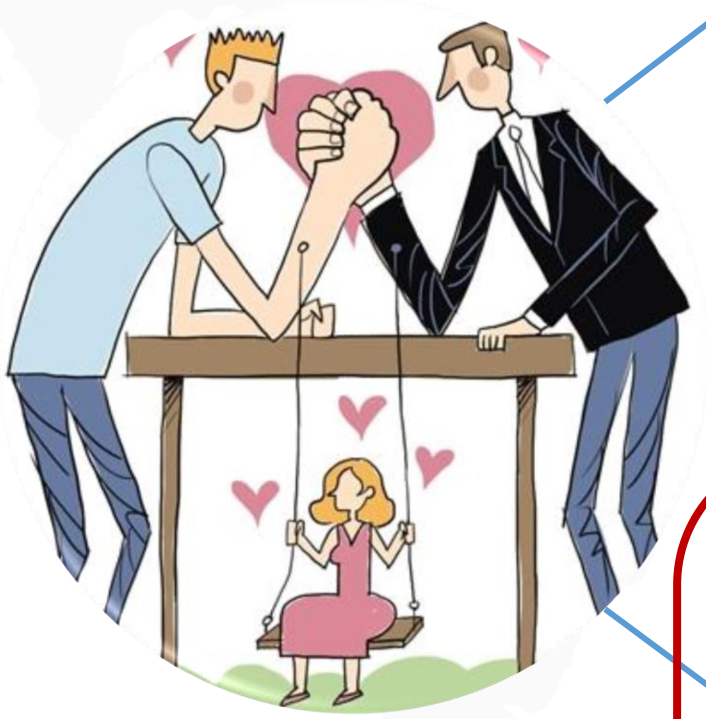
2 **博弈类型定义及表示**

3 博弈均衡分析与求解

4 博弈应用案例与讲解

不完全信息动态博弈形式化定义

- 不完全信息博弈的分类和特点



不完全信息静态博弈

- 别名：静态贝叶斯博弈
- 定义与表示：同时、策略式、类型
- 示例与分析：竞争、拍卖等
- 均衡与求解：贝叶斯纳什均衡

不完全信息动态博弈

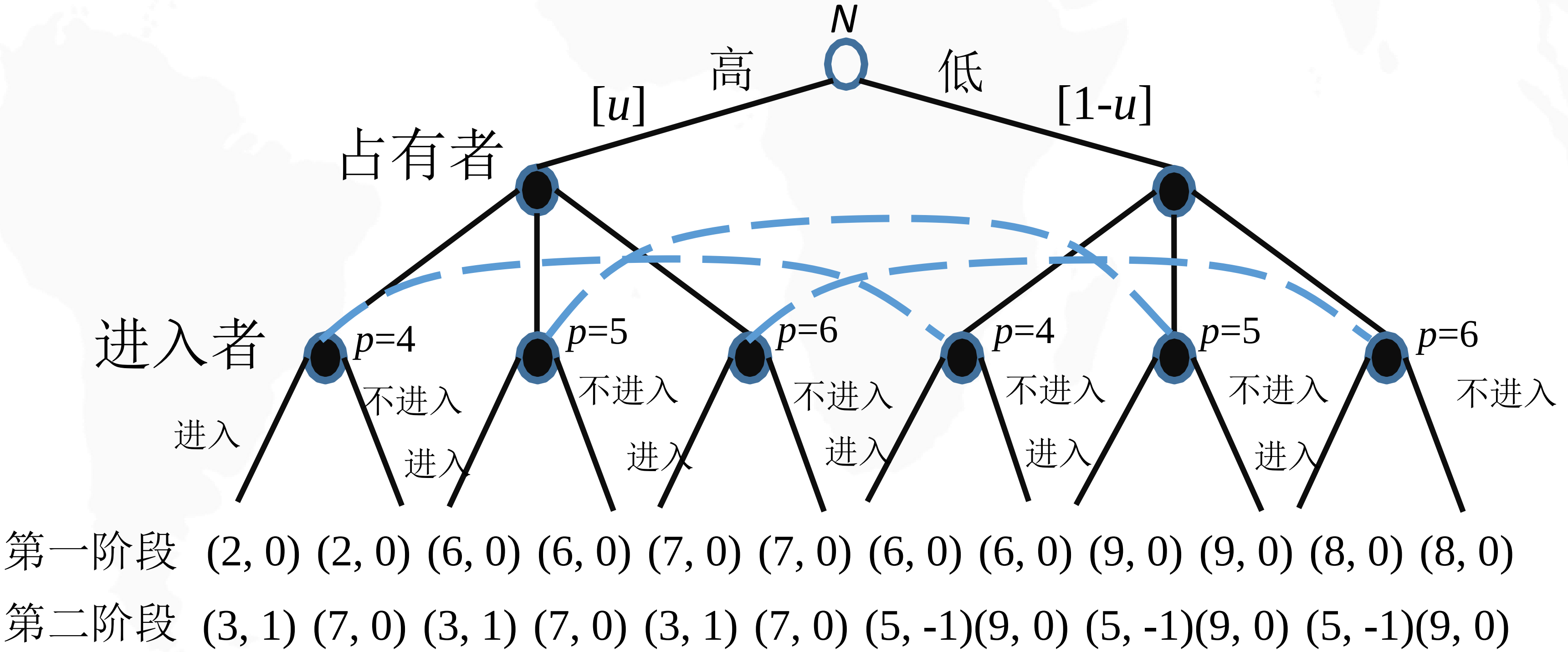
- 别名：动态贝叶斯博弈
- 定义与表示：顺序、展开式、类型
- 示例与分析：扑克、多轮谈判等
- 均衡与求解：精炼贝叶斯均衡

不完全信息动态博弈形式化定义

- 不完全信息动态博弈：博弈过程中参与人预先不知道关于博弈的所有信息（最终体现在支付/偏好等），参与人的行动有先后顺序，且后行动者能够观察到先行动者行动的博弈。
- 不完全信息动态博弈的展开式表示：
  - **参与人集合：**$i \in \mathrm{N}, \mathrm{N} = \{1,2, \ldots, n\}$，**用**$N$**表示自然**
  - **参与人的行动顺序：谁在什么时候行动**
  - **参与人的行动空间：每次行动时，有什么选择**
  - **信息集合：每次行动结束，参与人知道些什么**
  - **效益函数：博弈结束后，参与人知道将得到什么**
  - **外生事件（即自然选择）的概率分布**

举例：不完全信息下的市场进入阻挠博弈

- 博弈描述：
  - 两阶段博弈：第一阶段占有者垄断市场，进入者考虑是否进入； 第二阶段如果进入者进入，两者进行库诺特博弈，否则占有者继续垄断。占有者有高和低两种成本，有三种价格选择 $p = 4, 5, 6$，高低成本对应利润分别为2, 6, 7 和6, 9, 8。进入者进入成本为2，生产成本为高成本。



举例：不完全信息下的市场进入阻挠博弈

- 沿用静态博弈的思路：
  - 对于占有者而言，第二阶段收益均相同。从第一阶段结果来看，如果占有者是高成本的，它的最优选择是 $p=6$；如果占有者是低成本的，它的最优选择是 $p=5$。
  - 对于进入者而言，如果占有者是高成本的，它的最优选择是进入，如果占有者是低成本的，它的最优选择是不进入；而只有进入者认为占有者是高成本的概率u大于1/2时，才会选择进入。
  - 与静态博弈不同的是，在观察到占有者第一阶段的价格选择后，进入者可以修正对占有者的先验概率u，因为占有者的价格选择可能包含着有关其成本函数的信息。
  - 因此，**尽管$p=6$对高成本占有者是有利可图的，但这将导致进入者推断占有者是高成本的，进入者将选择进入。因此高成本的占有者在第一阶段就不会选择$p=6$。**
- 因此，采用静态分析的方法会得到不合理的均衡。

不完全信息动态博弈

- 基本思路
  - 参与人的类型，只有参与人自己知道。
  - 后行动者可以通过观察先行动者所选择的行动来推断其类型或修正对其类型的先验信念（概率分布），选择自己的最优行动。
  - 先行动者预测到自己的行为将被后行动者所利用，就会设法选择传递对自己最有利的信息。
  - 参与人双方都只知道对方行动的类型以及采取这样行动的概率，然后双方不断试探，根据对方的反应调整自己的策略，最终取得最优决策。总之，**不完全信息动态博弈不仅是参与人选择行动的过程，而且是参与人不断修正信念的过程。**
- 直观例子：黔驴技穷、信号传递

不完全信息动态博弈形式化定义

- 静态贝叶斯博弈的等价执行时间顺序

1. 自然赋予博弈各方类型向量$\Theta = \{\theta_1, \ldots, \theta_n\}$
2. 参与人得知自己的类型，但不知其他人的类型
3. 第一个行动者根据类型行动，传递出某种信息
4. 下一个行动者根据上个行动者的行动更新对其类型的先验推断，得到后验推断，然后做出行动
5. 下来的行动者不断重复上述过程直到博弈终止

第五讲 不完全信息动态博弈

内容提纲

1 博弈类型引入与介绍

2 博弈类型定义及表示

3 博弈均衡分析与求解

4 博弈应用案例与讲解

精炼贝叶斯均衡

- 精炼贝叶斯均衡的相关概念
  - 假定博弈有n个参与人，参与人i的类型为θ_i，$p(\theta_{-i}|\theta_i)$是参与人i认为其他$n-1$个参与人的类型$\theta_{-i}=(\theta_1,\ldots,\theta_{i-1}，\theta_{i+1}，\ldots,\theta_n)$的先验概率。
  - 令$S_i=\{s_i(\theta_i)\}$是参与人的策略空间，s_i是一个特定的策略。
  - 令$a_{-i}^h=\left(a_1^h,\ldots,a_{i-1}^h,a_{i+1}^h,\ldots,a_n^h\right)$是在第$h$个信息集上参与人$i$观测到的其他个$n-1$参与人的行动组合，这个行动组合是策略组合$s_{-i}=(s_1,\ldots,s_{i-1},s_{i-2},\ldots,s_n)$的一部分。
  - 令$\tilde{p}_i\left(\theta_{-i}\middle|a_{-i}^h\right)$是在观测到$a_{-i}^h$的情况下参与人 i认为其他$n-1$个参与人属于类型 θ_{-i}的后验概率，$\tilde{p}_i$是所有后验概率的集合 即 $\tilde{p}_i$ 包含了参与人 i在每一个信息集上的后验概率。
  - 令$u_i(s_i,s_{-i},\theta_i,\theta_{-i})$是参与人$i$的效用函数。

精炼贝叶斯均衡

- 精炼贝叶斯均衡定义

精炼贝叶斯均衡

精炼贝叶斯均衡是一个策略组合$s^*(\theta)=\left(s_1^*(\theta_1),\dots,s_n^*(\theta_n)\right)$和一个后验概率组合 $\tilde{p}=(\tilde{p}_1,\dots,\tilde{p}_n)$，满足：

（P）对于所有的参与人i，在每一个信息集合h，

$$s_i^*(s_{-i},\theta_i)\in\max_{s_i}\sum_{\theta_{-i}}\tilde{p}_i\left(\theta_{-i}\middle|a_{-i}^h\right)u_i(s_i,s_{-i},\theta_i,\theta_{-i})$$

（B）$\tilde{p}_i\left(\theta_{-\mathrm{i}}\middle|a_{-i}^h\right)$是使用贝叶斯法则从先验概率$p_i(\theta_{-i}|\theta_i)$、观测到的$a_{-i}^h$和最优策略$s_{-i}^*(*)$得到的（在可能的情况下）。

注：上述定义写出的是纯策略精炼贝叶斯纳什均衡的定义，对于混合策略也同样适用。

精炼贝叶斯均衡

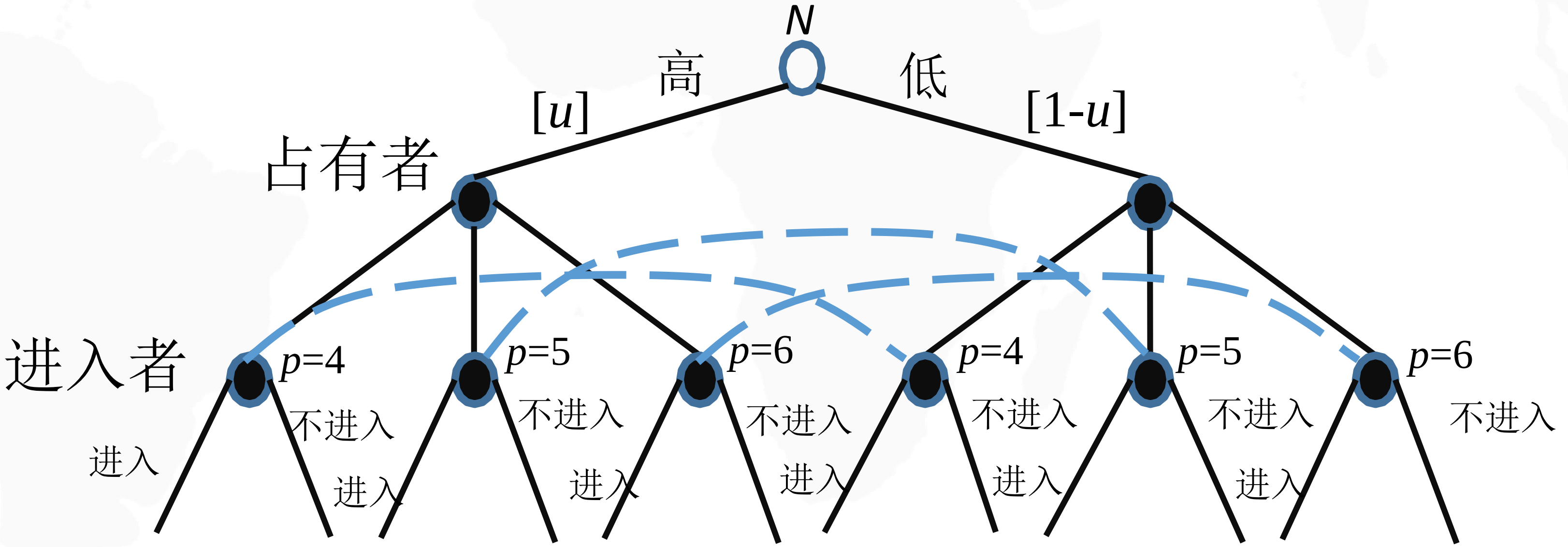
- 关于贝叶斯纳什均衡的一些解释
  - （P）是精炼条件（perfectness condition），给定其他参与人的策略s_{-i}和参与人的后验概率 $\tilde{p}_i(\theta_{-i} \mid a^h_{-i})$，每个参与人$i$的策略在所有从信息集 h开始的后续博弈上都是最优的，即所有参与人都是序贯理性的（sequentially rational）。
  - （B）对应的是贝叶斯法则的运用。如果参与人是多次行动的，修正概率涉及贝叶斯法则的重复运用。其中策略本身是不可观测的，因此参与人i只能根据观测到的行动组合a_{-i}修正概率，但它假定所观测到的行动是最优策略s^*_{-i}规定的行动。
  - “在可能的情况下”：如果a_{-i}不是均衡策略下的行动，即a_{-i}是一个零概率事件。根据贝叶斯法则，任何$\tilde{p}_i\left(\theta_{-i} \mid a^h_{-i}\right)$都是允许的，只要它与均衡策略相容。

精炼贝叶斯均衡

- 贝叶斯纳什均衡定义的关键
  - 精炼贝叶斯均衡是均衡策略和均衡信念的结合：给定信念$\tilde{p} = (\tilde{p}_1, \ldots, \tilde{p}_n)$，策略$s^* = (s_1^*, \ldots, s_n^*)$是最优的；给定策略$s^* = (s_1^*, \ldots, s_n^*)$，信念 $\tilde{p} = (\tilde{p}_1, \ldots, \tilde{p}_n)$ 是使用贝叶斯法则从均衡策略和所观测到的行动得到的。
  - 因此，精炼贝叶斯均衡是一个不动点，后验概率依赖于策略，策略依赖于后验概率，因此完全信息博弈中使用的逆向归纳法（backward induction）在不完全信息中不再适用（如果我们不知道先行者如何选择，我们就不可能知道后行动者应该如何选择），必须使用前向法（forward manner）进行贝叶斯修正。

示例回顾：不完全信息市场进入阻挠博弈

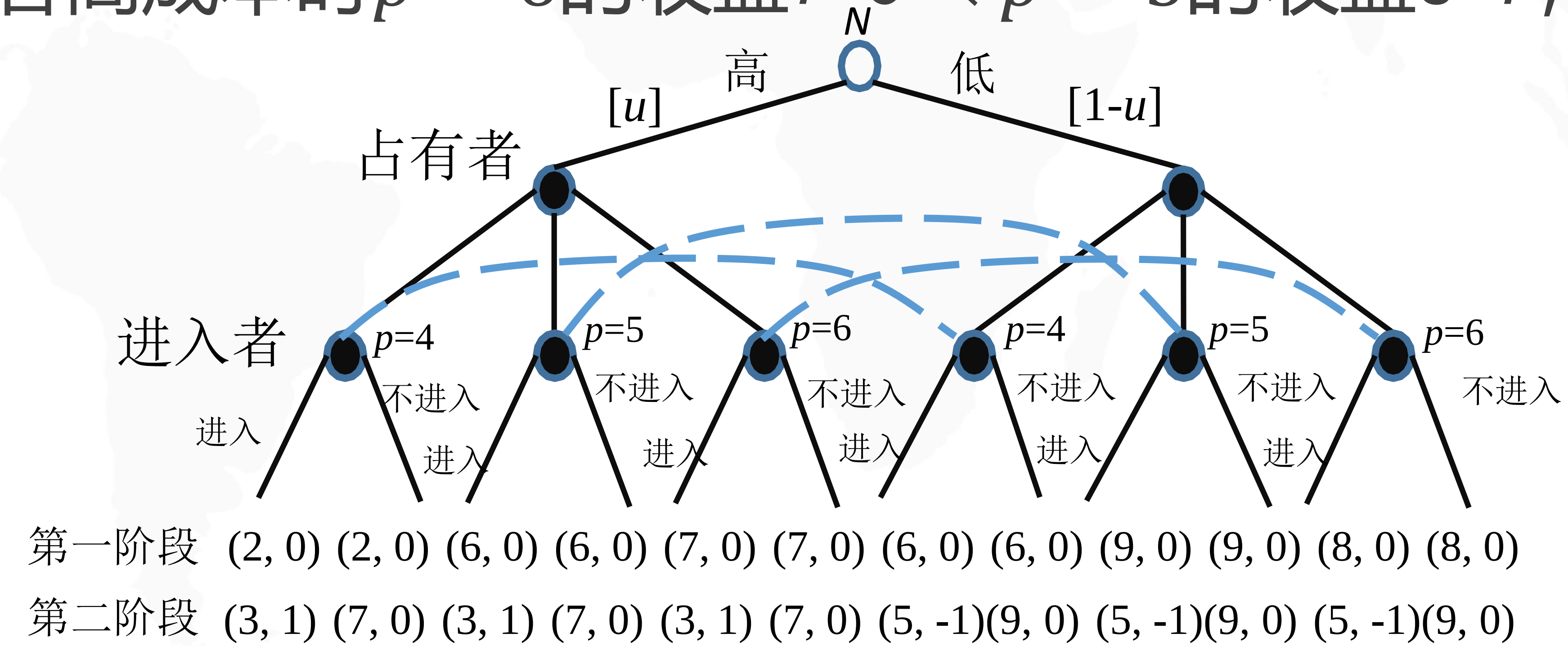
- 占有者有两个潜在类型：高成本和低成本；进入者只有一个类型：进入成本为2；因此只有进入者需要修正信念。
- 令$\tilde{u}(p)$是进入者在观测到占有者的价格选择后认为占有者是高成本的后验概率（注意：这里p代表占有者的价格而非概率）。



第一阶段	(2, 0)	(2, 0)	(6, 0)	(6, 0)	(7, 0)	(7, 0)	(6, 0)	(6, 0)	(9, 0)	(9, 0)	(8, 0)	(8, 0)
第二阶段	(3, 1)	(7, 0)	(3, 1)	(7, 0)	(3, 1)	(7, 0)	(5, -1)	(9, 0)	(5, -1)	(9, 0)	(5, -1)	(9, 0)

示例回顾：不完全信息市场进入阻挠博弈

- 首先证明：不论u为多少，占有者选择单阶段最优垄断价格（即高成本时选择$p = 6$，低成本时选择$p = 5$）并不是精炼贝叶斯均衡：因为即使在高成本时选择$p = 5$可以获得更大收益
- 证明：由于进入者看到$p = 6$得知高成本，看到$p = 5$得知低成本，故占有者高成本时$p = 6$的收益7+3 ＜ $p = 5$的收益6+7，得证。



示例回顾：不完全信息市场进入阻挠博弈

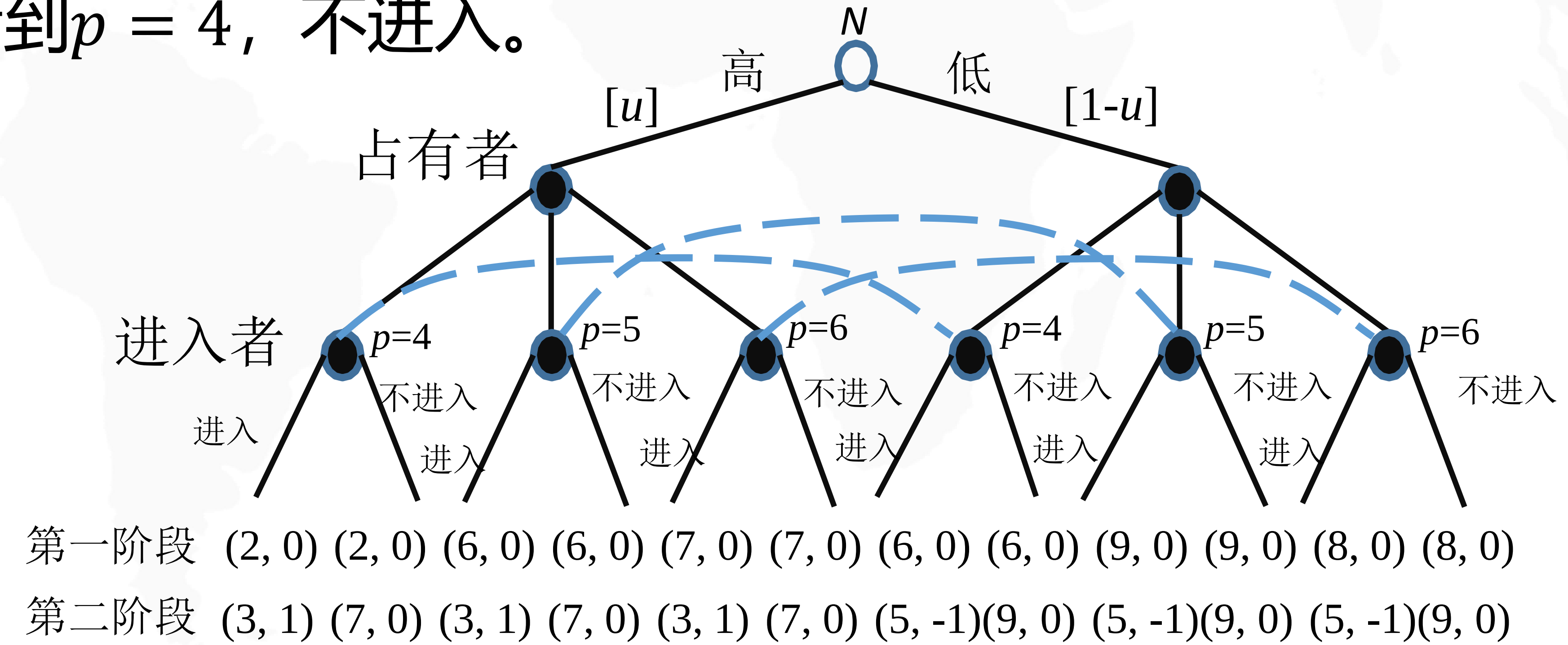
- 下来证明：当$u < 1/2$时，精炼贝叶斯均衡是：不论成本高低，占有者选择$p = 5$；当观测到$p = 6$时，进入者选择进入，否则不进入。
- 证明：首先占有者无论成本高低，选择$p = 5$都是最优的，因为占有者选择 $p = 5$ 时，进入者不能从观测到的价格得到任何新的信息，即 $\tilde{u}(5) = \frac{1\times u}{(1\times u)+1\times(1-u)} = u < \frac{1}{2}$，进入者选择进入的期望收益是 $u \times 1 + (1-u) \times (-1) = 2u - 1 < 0$ ，选择不进入的期望收益是0，因此进入者的最优策略是不进入，**此时收益6+7**（高成本时）或**9+9**（低成本时）**均最大**。
- 上述均衡称为**混同均衡**，因为两类占有者选择相同的价格，用于隐藏自己是高成本这个事实。

示例回顾：不完全信息市场进入阻挠博弈

- 最后证明：当$u \geq 1/2$时，精炼贝叶斯均衡是：低成本在位者选择$p=4$，高成本占有者选择$p=6$；当观测到$p=4$时，进入者选择不进入，否则进入。
- 证明：首先不同类型占有者选择相同价格$p=5$，因为$\tilde{u}(5) \geq 1/2$，进入者将进入。下来考虑低成本占有者：选择$p=4$，进入者不进入，占有者收益为$6+9$；选择$p=5$，进入者进入，占有者收益为$9+5$；故$p=4$最优；考虑高成本占有者：选择$p=4$，进入者不进入，占有者收益$2+7$；选择$p=6$，占有者收益$7+3$。最后考虑进入者后验概率和策略：看到$p=6$，进入；看到$p=4$，不进入。
- 上述均衡称为**分离均衡**，因为两类占有者选择不同的价格，用于证明自己真实的成本。

示例回顾：不完全信息市场进入阻挠博弈

分离均衡证明小结：首先不同类型占有者选择相同价格$p=5$，进入者将进入。下来考虑低成本占有者：选择$p=4$，进入者不进入，占有者收益$6+9$；选择$p=5$，进入者进入，占有者收益$9+5$；因此$p=4$最优；最后考虑高成本占有者：选择$p=4$，进入者不进入，占有者收益$2+7$；选择$p=6$，进入者收益$7+3$。最后考虑进入者后验概率和策略。看到$p=6$，选择进入；看到$p=4$，不进入。



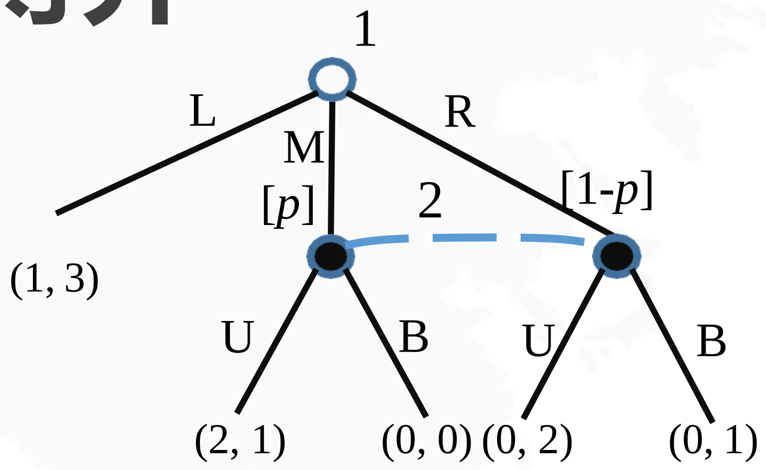
第五讲 不完全信息动态博弈

内容提纲

不完美信息博弈的精炼贝叶斯均衡

- 因为不完全信息博弈能够通过海萨尼转换为不完美信息博弈，因此精炼贝叶斯均衡也适用于不完美信息博弈。
- 贝叶斯法则在不完美信息动态博弈中的运用
  - 在不完全信息博弈中，参与人根据观测到其他参与人的行动和其他参与人的最优策略，使用贝叶斯法则修正对其他参与人类 型的信念。
  - 在不完美信息博弈中，参与人观测不到其他参与人的行动，通过观测到博弈是否进入自己的信息集，修正自己处于该信息集 的每一个决策结的概率。
  - **尽管贝叶斯法则在非均衡路径上没有定义，但如何规定非均衡路径上的后验概率是至关重要的。**

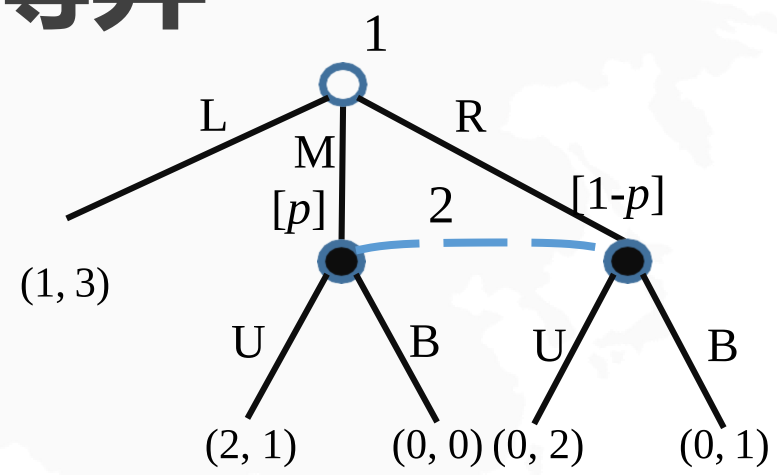
不完美信息博弈举例1：两人博弈



- 博弈规则如右图：
  - 参与人1首先行动，选择L、M或R；
  - 选择L，博弈结束；选择M或R，参与人2选择U或B
- 不完美信息：参与人2在作出自己决策时（U or B）时并不知道参与人 1 选择了M或R，尽管他知道 L没有被选择。
- 均衡分析
  - 两个纯策略纳什均衡 (L, B) 和 (M, U)
  - 这两个同时也是子博弈精炼纳什均衡
  - 但(L, B)却不是合理的均衡，需要通过精炼贝叶斯均衡剔除

参与人2 / 参与人1	U	B
L	1, 3	1, 3
M	2, 1	0, 0
R	0, 2	0, 1

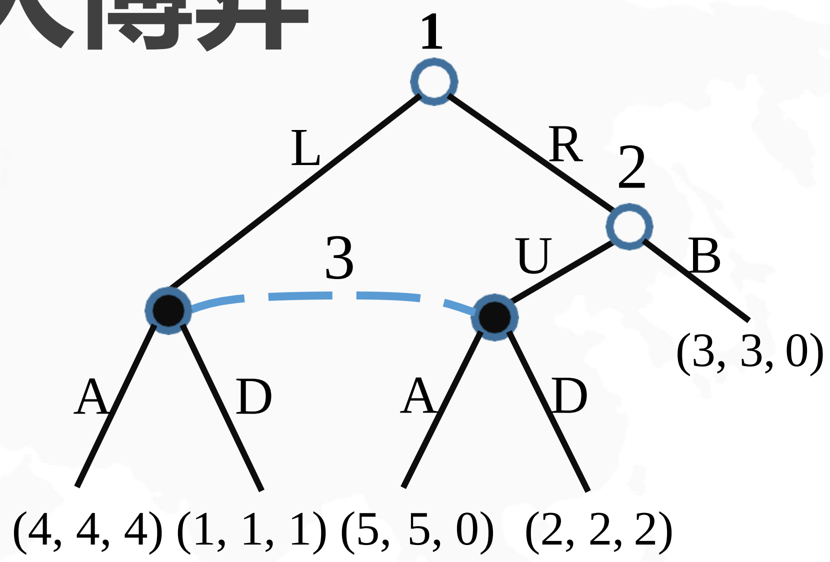
不完美信息博弈举例1：两人博弈



- 博弈规则如右图：
- (M, U)和(L, B)都是子博弈精炼纳什均衡
- 通过精炼贝叶斯均衡剔除(L, B)
  - 当博弈进入到参与人2的信息集时，必须有一个参与人1选择M或者R的概率分布，假设其为p和$1-p$：
    - 参与人2选择U收益：$p \times 1 + (1-p) \times 2 = 2 - p$
    - 参与人2选择B收益：$p \times 0 + (1-p) \times 1 = 1 - p$
    - 由于$2 - p > 1 - p$，故参与人2一定会选择U，因此贝叶斯均衡(L, B)依赖于一个不可置信的威胁：(L, U)
  - 给定参与人2的最优策略U，参与人1的最优策略是M，因此

    (M, U) 是一个精炼贝叶斯均衡。

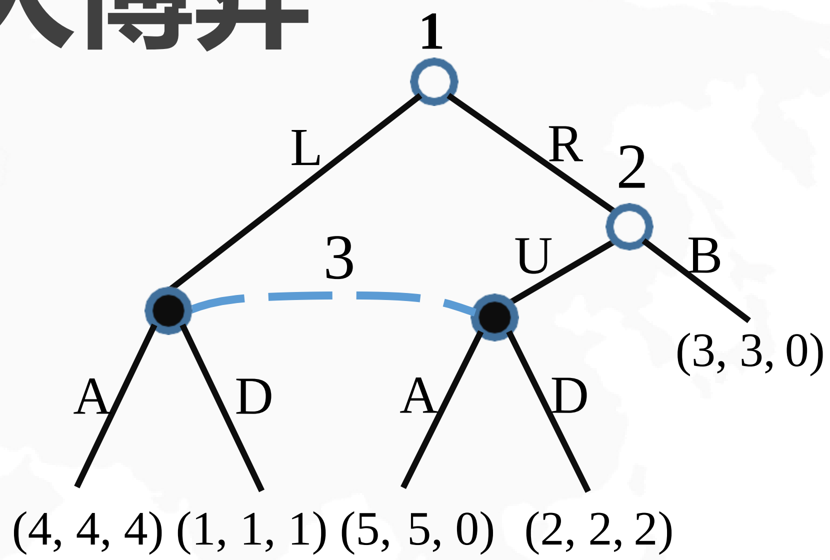
不完美信息博弈举例2：三人博弈



- 三个参与人，$i =$ 1, 2, 3；
  - 1首先行动，选择L或R；
  - 1选择R，2选择U或B（结束）；
  - 1选择L，或2选择U，3选择A或D。
- 不完美信息：参与人3的信息集
- 均衡分析
  - 两个纳什均衡(L, B, A)和(R, B, D)
  - 同时也是子博弈精炼纳什均衡
  - 但(L, B, A)却不是合理的均衡
  - 需要通过精炼贝叶斯均衡剔除

3	A		D	
1 \ 2	U	B	U	B
L	4,4,4	4,4,4	1,1,0	1,1,0
R	5,5,0	3,3,0	2,2,2	3,3,0

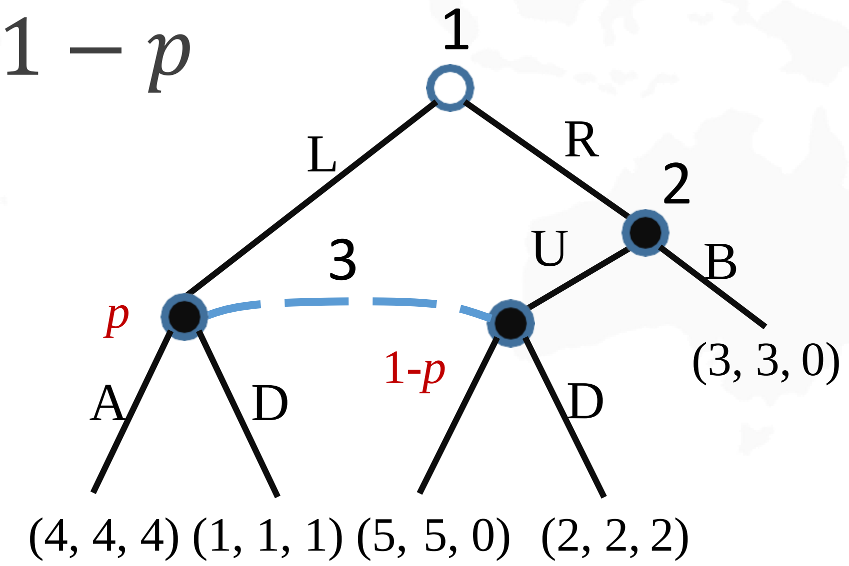
不完美信息博弈举例2：三人博弈



- 三个参与人，$i = 1, 2, 3$；
  - 1首先行动，选择L或R；
  - 1选择R，2选择U或B（结束）；
  - 1选择L，或2选择U，3选择A或D。
- 利用精炼贝叶斯均衡检查(L, B, A)这个均衡解是否合理
  - 我们不可能使用子博弈精炼均衡剔除(L, B, A)，因为唯一的子博弈是博弈本身，只能使用精炼贝叶斯均衡剔除(L, B, A)。
  - 均衡(L, B, A)中2的选择B不在均衡路径上，需要利用精炼贝叶斯均衡检查其合理性：给定3选择A，2的最优选择是U而不是B
  - 给定2选择U而不是B，1应该选择R而不是L，进而3应该选择D
  - 假设3选择D，2应该选择B，进而1应该选择R。这样得到(R, B, D)是该博弈的另一个纳什均衡，也需要检查下其是否合理。

不完美信息博弈举例2：三人博弈

- 精炼贝叶斯均衡要求
  - 1）每一个参与人的信息集上有一个概率分布；
  - 2）给定概率分布和其他参与人选择，每个参与人的策略最优；
  - 3）概率分布是使用贝叶斯法则从最优策略和观测到的行动得到的（在可能的情况下）；
- 参与人1和2都只有一个单结信息集，在该决策结他们决策的概率为1；参与人3的信息集有两个决策结：p和$1-p$
- （L, B, A）及相关联的$p=1$
  - 不满足（2）：B不是最优
  - 不满足（3）：进而B的$p=1$不合理
- 大家检查下均衡（R, B, D; $p<2/5$）

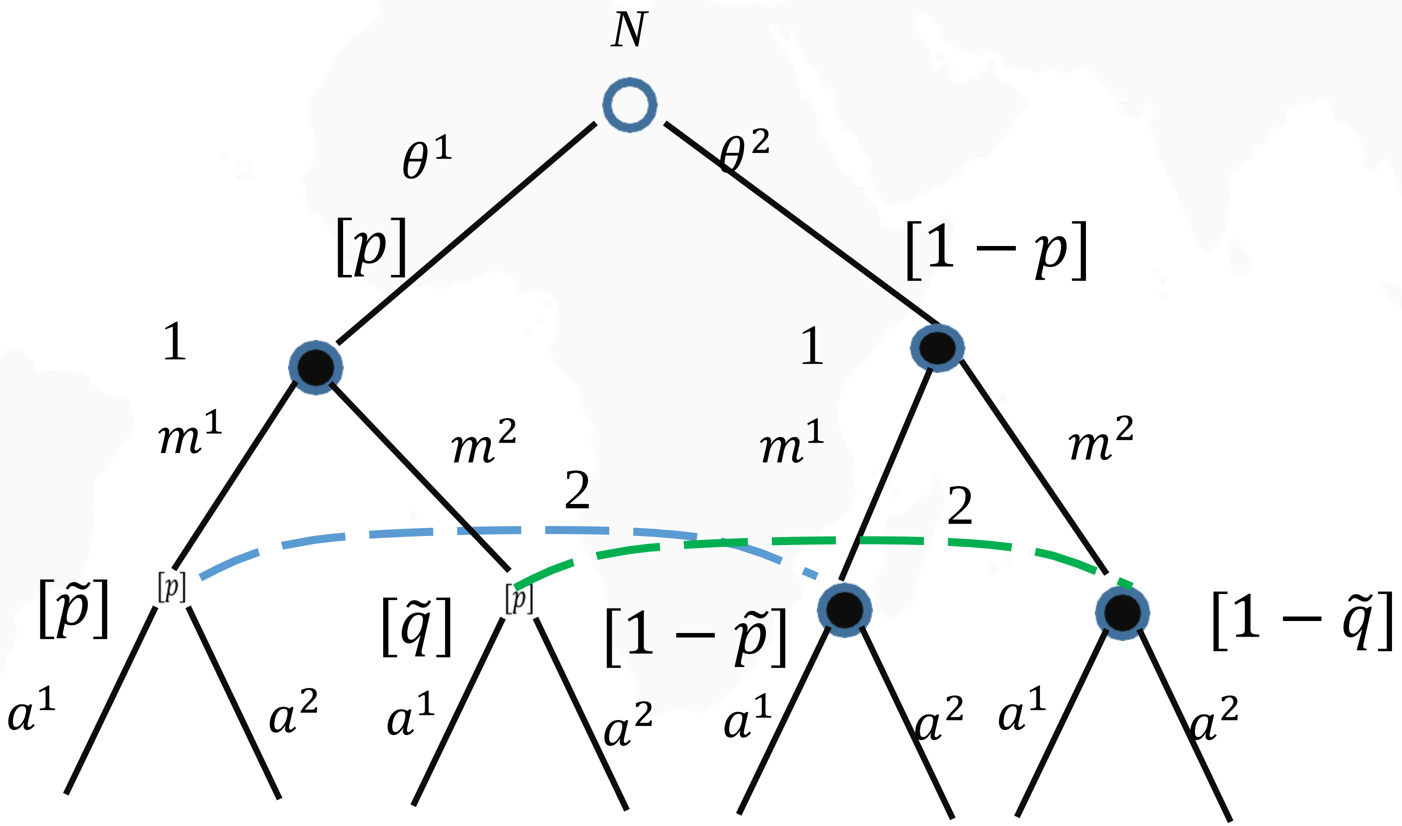


不完美信息博弈举例3：信号传递博弈

- 博弈描述：两个参与人，参与人1称为信号的发送者，参与人2称为信号的接收者；参与人1的类型是私人信息，参与人2的类型是公共信息。
  - 首先，“自然”选择参与人1的类型$\theta \in \Theta$，$\Theta = \{\theta^1, \cdots, \theta^K\}$是参与人1的类型空间，也是参与人1的私有信息，参与人2不知道具体取值，但是参与人2知道参与1属于类型θ的先验概率是$p = p(\theta)$, 且$\sum_k p(\theta^k) = 1$。
  - 然后，参与人1在观测到类型θ之后选择发出信号$m \in M$，这里$M = \{m^1, \cdots, m^J\}$是信号空间。
  - 接下来，参与人2 观测到参与人1的信号m（但不知道具体类型θ），使用贝叶斯法则，从先验概率$p = p(\theta)$得到后验概率$\tilde{p} = \tilde{p}(\theta|m)$，然后选择行动$a \in A$，其中$A = \{a^1, \cdots, a^H\}$是参与人2的行动空间。
  - 最后，得到支付函数$u_1(m, a, \theta)$和$u_2(m, a, \theta)$。

不完美信息博弈举例3：信号传递博弈

- 信号传递博弈的执行顺序示例
  - 具体赋值：$K = J = H = 2, \tilde{p} = \tilde{p}(\theta^1|m^1), \tilde{q} = \tilde{p}(\theta^1|m^2),$
  - 扩展式表示:



不完美信息博弈举例3：信号传递博弈

- 令$m(\theta)$是参与人1的类型依存策略，$a(m)$是参与人2的行动策略（信号依存行动）。这里允许混合策略，即：参与人1以某种概率随机地选择不同信号，参与人2以某种概率随机选择行动。

信号传递博弈的精炼贝叶斯均衡

信号传递博弈的精炼贝叶斯均衡是一个策略组合$(m^*(\theta), a^*(m))$和后验概率$\tilde{p}(\theta|m)$，满足：

（P$_1$）$a^*(m) \in \arg\max_a \Sigma_\theta \, \tilde{p}(\theta|m) u_2(m, a, \theta)$；

（P$_2$）$m^*(\theta) \in \arg\max_m u_1(m, a^*(m), \theta)$；

（B）$\tilde{p}(\theta|m)$是参与人2使用贝叶斯法则从先验概率$p(\theta)$、观测信号m和参与人1的最优策略$m^*(\theta)$到的（在可能的情况下）。

不完美信息博弈举例3：信号传递博弈

信号传递博弈的精炼贝叶斯均衡

信号传递博弈的精炼贝叶斯均衡是一个策略组合$(m^*(\theta), a^*(m))$和后验概率$\tilde{p}(\theta|m)$，满足：

（P_1）$a^*(m) \in \arg\max_a \Sigma_\theta\ \tilde{p}(\theta|m) u_2(m, a, \theta)$；

（P_2）$m^*(\theta) \in \arg\max_m u_1(m, a^*(m), \theta)$；

（B）$\tilde{p}(\theta|m)$是参与人2使用贝叶斯法则从先验概率$p(\theta)$、观测信号m和参与人1的最优策略$m^*(\theta)$到的（在可能的情况下）。

说明

- P_1和P_2等价于精炼贝叶斯均衡中的P条件；
- P_1指给定后验概率，参与人2对参与人1的信号做出的最优反应；
- P_2是指预测到参与人2的最优反应，参与人1将做出的决策；
- B中“在可能的情况下”与之前定义同，限制了贝叶斯条件。

不完美信息博弈举例3：信号传递博弈

- 信号传递博弈可能的精炼贝叶斯均衡的三类均衡

分离均衡	混同均衡	准分离均衡
• 不同类型的发送者以确定性的概率选择发送不同的信号	• 不同类型的发送者选择相同的信号/没有任何类型选择其他类型的信号	• 一些类型的发送者发随机选择信号，另一些类型的发送者选择特定信号

不完美信息博弈举例3：信号传递博弈

- 信号传递博弈中不同类型的发送者（参与人1）以1的概率选择不同的信号，或者说没有任何类型选择与其他类型相同的信号。在分离均衡下，信号准确地揭示出类型。

信号传递博弈的分离均衡

假定$K = J = 2$(即只有两个类型两个信号)，那么分离均衡意味着： 如果m^1是类型 θ^1的最优选择，那么m^1就不可能是类型θ^2的最优选择，并且，m^2一定是类型θ^2的最优选择。即：

$$u_1(m^1, a^*(m), \theta^1) > u_1(m^2, a^*(m), \theta^1)$$
$$u_1(m^2, a^*(m), \theta^2) > u_1(m^1, a^*(m), \theta^2)$$

因此，后验概率是：

$$\tilde{p}(\theta^1|m^1) = 1, \qquad \tilde{p}(\theta^1|m^2) = 0,$$
$$\tilde{p}(\theta^2|m^1) = 0, \qquad \tilde{p}(\theta^2|m^2) = 1$$

不完美信息博弈举例3：信号传递博弈

- 不同类型的发送者选择相同的信号，或者说，没有任何类型选择其它类型不同的信号，因此接收者不修正先验概率。

信号传递博弈的混同均衡

假定m^j是均衡策略，那么，

$$u_1\left(m^j, a^*(m), \theta^1\right) \geq u_1(m, a^*(m), \theta^1)$$
$$u_1\left(m^j, a^*(m), \theta^2\right) \geq u_1(m, a^*(m), \theta^2)$$
$$\tilde{p}\left(\theta^k | m^j\right) \equiv p(\theta^k)$$

- 从贝叶斯分析的角度：各个分布可能得到的观测样本是相同的，那么采样对分布的估计就没有什么帮助，只能是原先认为是什么分布，现在仍然维持原判。

不完善信息博弈举例3：信号传递博弈

- 一些类型的发送者随机选择信号，另外一些类型的发送者选择特定的信号。

信号传递博弈的准分离均衡

假定类型θ^1的发送者随机地选择m^1或者m^2，而类型θ^2的发送者以1的概率选择m^2，如果这个策略组合是均衡策略组合，那么

$$u_1(m^1, a^*(m), \theta^1) = u_1(m^2, a^*(m), \theta^1)$$
$$u_1(m^1, a^*(m), \theta^2) < u_1(m^2, a^*(m), \theta^2)$$

后验概率

$$\tilde{p}(\theta^1|m^1) = \frac{\alpha \times p(\theta^1)}{\alpha \times p(\theta^1) + 0 \times p(\theta^2)} = 1;$$
$$\tilde{p}(\theta^1|m^2) = \frac{(1-\alpha) \times p(\theta^1)}{(1-\alpha) \times p(\theta^1) + 1 \times p(\theta^2)} < p(\theta^1),$$
$$\tilde{p}(\theta^2|m^2) = \frac{1 \times p(\theta^1)}{(1-\alpha) \times p(\theta^1) + 1 \times p(\theta^2)} > p(\theta^2)$$

本讲内容小结

- **贝叶斯定理**
- **定义及表示**
- **精炼贝叶斯纳什均衡**
- **不完美信息/信号博弈**

知识点

能力线

- **知识综合运用能力**：信息不完全+动态依赖性
- **贝叶斯思维再强化**：概率意义上的最优估计模型和方法

- **不完全信息动态博弈**：估计信号传达的信息

① 不断观察，迭代修正

② 无关路径，也要理性

价值面

本次课程作业

- 作业内容：寻找存在动态贝叶斯博弈应用示例的不完全信息动态博弈习题两道，给出习题描述、问题分析、求解过程和最终结果。
  - 评分准则：选取习题的典型性、分析求解过程的完整性以及最终结果的指导性意义。
- 截止时间：下次上课前（2024年4月1日）
- 提交方法：在Sep课程网站上提交，同时提交电子版到助教邮箱（xuhang2020@ia.ac.cn）
- 邮件发送规范
  - 邮件主题：博弈论第**五**次作业_**学号**_**姓名**
  - 附件名称：博弈论第**五**次作业_**学号**_**姓名**.docx

感谢大家认真听讲

兴军亮

jlxing@nlpr.ia.ac.cn

2024年3月25日