

动态博弈与数据定价

2024-2025 学年短学期 数据要素市场

吴一航

yhwu_is@zju.edu.cn

浙江大学计算机科学与技术学院

2024 年 7 月 10 日

一个简单的例子

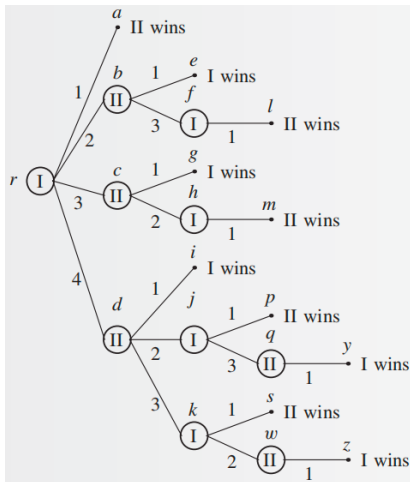
考虑下图所示的简单博弈，博弈有两个参与人，参与人 1 首先行动。每个参与人一次抓取一个方格，条件如下：

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 谁就输了

问：参与人 1 和参与人 2 有制胜策略吗？谁有制胜策略？

一个简单的例子 (Cont'd)



- ❶ 整个博弈的进行流程可以表达为一棵树
- ❷ 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
- ❸ 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？
- ❹ 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

为了完整描述这个博弈，除了树之外，我们还需要指明非叶子节点将要选择行动的参与人，以及每个叶子节点对应的结果。这样的博弈称为扩展式博弈。

定义 (扩展式博弈)

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

- ① I 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 $(V_i)_{i \in I}$ 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

一般而言，在叶子节点 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 (Cont'd)

我们将前面的例子形式化表达：

- ① $I = \{1, 2\}$
- 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 $u(a) = u(l) = u(m) = u(p) = u(s) = 2 \text{ 胜}$,
 $u(e) = u(g) = u(i) = u(y) = u(z) = 1 \text{ 胜}$

扩展式博弈的定义 (Cont'd)

我们将前面的例子形式化表达：

- ① $I = \{1, 2\}$
- 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

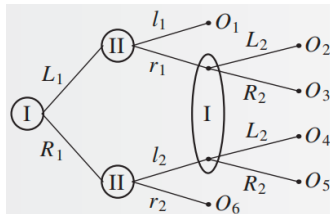
扩展式博弈的定义 (Cont'd)

为了接下来讨论的方便，我们需要引入一些记号和定义：

- $J(x)$ 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 $J(v_i) = i$
- $C(x)$ 表示非叶节点 x 的所有子节点的集合
- $A(x)$ 表示参与人 $J(x)$ 在节点 x 处所有可能行动的集合
- 参与人 i 的**策略 (strategy)** 是一个函数 s_i ，将每一个自己做决策的节点 x 映射到 $A(x)$ 中的一个行动，或者对等的， $C(x)$ 中的一个节点。我们用 S_i 表示参与人 i 的所有可能策略的集合
- 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
- 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)**

策略式博弈和扩展式博弈

		Player II			
		$l_1 l_2$	$l_1 r_2$	$r_1 l_2$	$r_1 r_2$
Player I	$L_1 L_2$	O_1	O_1	O_2	O_2
	$L_1 R_2$	O_1	O_1	O_3	O_3
	$R_1 L_2$	O_4	O_6	O_4	O_6
	$R_1 R_2$	O_5	O_6	O_5	O_6



- 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
- 从扩展式博弈到策略式博弈的表达是唯一的，但从策略式博弈到扩展式博弈的表达是不唯一的

冯·诺依曼定理

为了描述接下来的关键定理，我们需要首先给出制胜策略、至少保证和局的策略的形式化定义：

定义

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

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

如果满足下式，则称参与人 1 的策略 s_1 是至少保证和局的策略：

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

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

冯·诺依曼定理 (Cont'd)

定理 (冯·诺依曼定理, 1928)

在任何完全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

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 $n_x = 1$ ，说明 x 是叶节点，如果 x 的结果是 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

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 $n_x = 1$ ，说明 x 是叶节点，如果 x 的结果是 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

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 $n_x = 1$ ，说明 x 是叶节点，如果 x 的结果是 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
 - ① 如果存在 $y \in C(x)$ ，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 ，然后按照 y 的制胜策略行动

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 $n_x = 1$ ，说明 x 是叶节点，如果 x 的结果是 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
 - ① 如果存在 $y \in C(x)$ ，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 ，然后按照 y 的制胜策略行动
 - ② 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

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 $n_x = 1$ ，说明 x 是叶节点，如果 x 的结果是 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
 - ① 如果存在 $y \in C(x)$ ，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 ，然后按照 y 的制胜策略行动
 - ② 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
 - ③ 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

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 $2K$ 。

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 ， $1 \leq k \leq K$ ，我

们用 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：

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

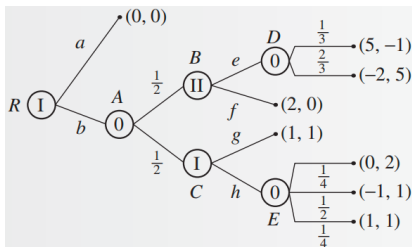
“参与人 1 没有制胜策略”可以表达为：

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

这等价于 $\forall a_1 \exists b_1 \forall a_2 \exists b_2 \cdots \forall a_K \exists b_K (\neg F)$ ，即“参与人 2 有至少保证和局的策略”。因此我们证明了，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，从而命题得证

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



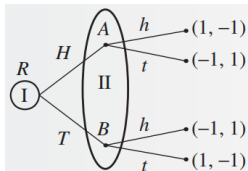
- 随机行动需要加入一个新的参与人，叫做“自然”，用 0 表示
- 还需要加入 p_x 来表示自然在节点 x 处抽签带来的可能结果的概率向量

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)

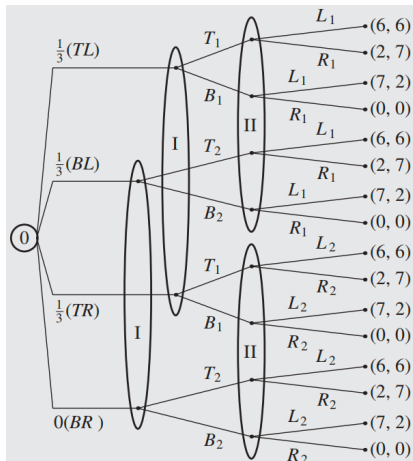
猜硬币博弈是一个两人博弈，每个参与人选择出示正面还是反面， H 表示硬币正面， T 表示硬币反面。如果两个参与人的选择一致，则参与人 2 给参与人 1 一美元；如果两个参与人的选择不一致，则参与人 1 给参与人 2 一美元。这一博弈用扩展式博弈表达如下页图所示。



- 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
- 完美信息博弈信息集都是单点集

完美信息博弈 (Cont'd)

事实上，之前斗鸡博弈相关均衡的例子也可以表达为不完全信息博弈



- 第一步仲裁人随机选择（这一步的随机结果是自然选择的）
- 接下来 1 和 2 知道了自己的推荐，但不知道对方的推荐，从而产生了对应的信息集

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，这样的选择叫做**均衡精炼 (equilibrium refinements)**。

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，这样的选择叫做**均衡精炼 (equilibrium refinements)**。

在扩展式博弈中，每个策略向量 σ 定义了从根节点到叶子节点的一条路径。当策略向量 σ 是纳什均衡时，如此得到的路径称为均衡路径。

定义 (子博弈完美均衡)

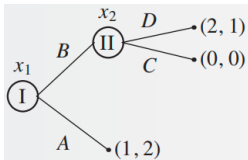
在扩展式博弈 Γ 中，一个策略向量 σ^* 是子博弈完美均衡 (**subgame perfect equilibrium**)，如果对于博弈的任意子博弈 $\Gamma(x)$ ，局限在那个子博弈的策略向量 σ^* 是 $\Gamma(x)$ 的纳什均衡：对每个参与人 i ，每个策略 σ_i 和子博弈 $\Gamma(x)$ ，

$$u_i(\sigma^* | x) \geq u_i(\sigma_i, \sigma_{-i}^* | x)$$

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？考虑下图所示二人博弈：

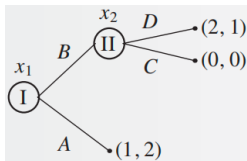


		Player II	
		C	D
Player I	A	1, 2	1, 2
	B	0, 0	2, 1

- 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)
- (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）

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？考虑下图所示二人博弈：



		Player II	
		C	D
Player I	A	1, 2	1, 2
	B	0, 0	2, 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)
- (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）
- 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

前面的例子中, (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.

子博弈完美均衡的充分条件

前面的例子中, (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. 我们有如下定理:

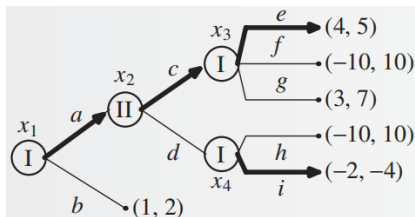
定理

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.

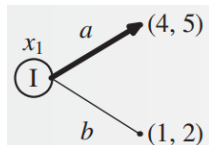
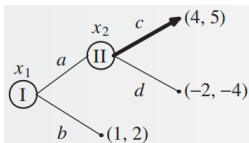
- 定理是显然的, 因为如果 σ^* 不是子博弈完美均衡, 那么在某个子博弈 $\Gamma(x)$ 上存在有利可图的偏离, 并且这个偏离产生的概率不为 0, 因此也可以带来全局的有利可图的偏离
- 推论: 完全混合的纳什均衡是子博弈完美均衡

逆向归纳法

如何找到完美信息博弈的子博弈完美均衡？我们可以使用逆向归纳法：



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 (Cont'd)

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：

定理

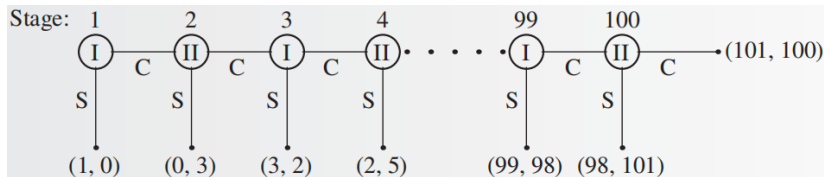
每个完美信息扩展式博弈都至少有一个子博弈完美纯策略均衡。

- 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
- 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）

接下来我们来看一个例子

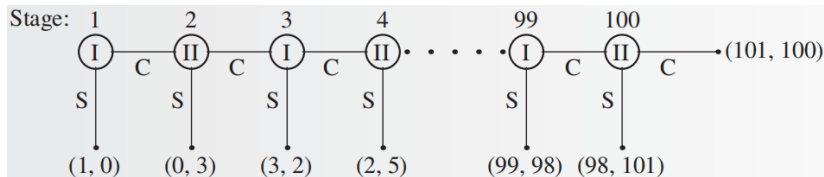
逆向归纳法的局限

蜈蚣博弈：两个参与人依次行动：在奇数轮 $t = 1, 3, \dots, 99$ ，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 $t = 2, 4, \dots, 100$ ，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：



逆向归纳法的局限

蜈蚣博弈：两个参与人依次行动：在奇数轮 $t = 1, 3, \dots, 99$ ，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 $t = 2, 4, \dots, 100$ ，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：



- 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

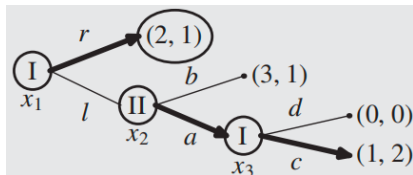
以上例子表明，逆向归纳不足以描述理性：逆向归纳忽视了到达这一节点的历史，但历史本身提供了关于其他参与人行为的信息

- 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了继续，那么我也应该继续？

正向归纳法

以上例子表明，逆向归纳不足以描述理性：逆向归纳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了继续，那么我也应该继续？



逆向归纳法的解为 (rc, a) ，但是如果参与人 I 选择了 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：完美均衡

产量领导模型（斯塔克尔伯格模型）

我们来看一个实际市场中的子博弈完美均衡的应用，即产量领导模型（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。

产量领导模型（斯塔克尔伯格模型）

我们来看一个实际市场中的子博弈完美均衡的应用，即产量领导模型（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，用 $c_1(y_1)$ 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）（Cont'd）

因此厂商 1 的利润最大化决策为：

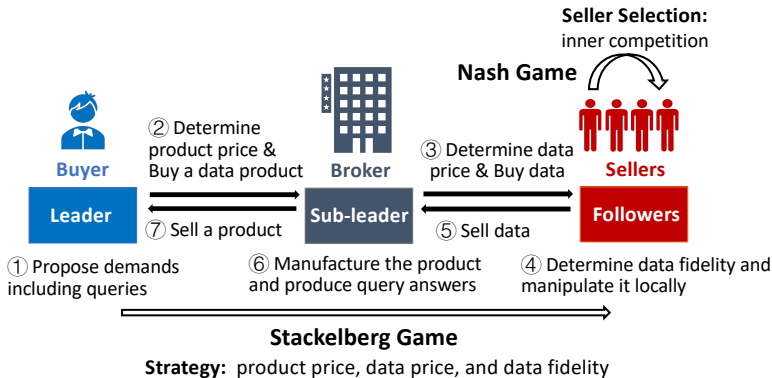
$$\begin{aligned} \max_{y_1} \pi_1(y_1, y_2) &= p(y_1 + y_2)y_1 - c_1(y_1) \\ \text{s.t. } y_2 &= \arg \max_{y_2} \pi_2(y_1, y_2) = p(y_1 + y_2)y_2 - c_2(y_2) \end{aligned}$$

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

- 因此我们应当先求解厂商 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
- 具体的计算留作习题

产量领导模型的应用

价格领袖模型（斯塔克尔伯格模型）可以自然地建模市场中存在行为先后顺序的情况，例如在数据市场中，经常是买家发起购买数据的请求，因此可以建模如下¹：



¹Yuran Bi, Chen Zhao, Junyi Zhao, Jinfei Liu, Kui Ren, Li Xiong. (2024) Share: Stackelberg-Nash based Data Markets. The annual IEEE International Conference on Data Engineering.

产量领导模型的应用 (Cont'd)

核心流程:

- 1 买家作为领导者, 首先提出购买需求和他希望的购买价格, 目标是自己的效用最大化 $\Phi(p^M, \tau) = u(q^D, score) - p^M q^M$
- 2 中间商决定数据的出售价格, 目标是最大化自己的利润 $\Omega(p^M, p^D, \tau) = p^M q^M - p^D q^D - c(N, score)$
- 3 卖家需要根据数据的出售价格和自己的隐私成本决定最优的隐私系数, 最大化利润 $\Psi_i(p^D, \tau_i) = p^D q_i^D - l_i(\tau_i)$

产量领导模型的应用 (Cont'd)

核心流程:

- ① 买家作为领导者, 首先提出购买需求和他希望的购买价格, 目标是自己的效用最大化 $\Phi(p^M, \tau) = u(q^D, score) - p^M q^M$
- ② 中间商决定数据的出售价格, 目标是最大化自己的利润 $\Omega(p^M, p^D, \tau) = p^M q^M - p^D q^D - c(N, score)$
- ③ 卖家需要根据数据的出售价格和自己的隐私成本决定最优的隐私系数, 最大化利润 $\Psi_i(p^D, \tau_i) = p^D q_i^D - l_i(\tau_i)$

讨论:

- 当然, 数据卖家具有较大的市场势力, 也可能成为价格领袖
- 然而价格领袖模型中市场是完全信息的, 并且定价并不完全针对数据的特性

不完全信息博弈

现实中的博弈不一定是完全信息的：斗地主我们不知道对手的手牌，厂商竞争互相之间的实力也并非完全已知。因此我们需要引入不完全信息博弈 (**game with incomplete information**)。

不完全信息博弈

现实中的博弈不一定是完全信息的：斗地主我们不知道对手的手牌，厂商竞争互相之间的实力也并非完全已知。因此我们需要引入不完全信息博弈 (game with incomplete information)。

	fight	not fight
fight	(1, -2)	(2, -1)
not fight	(-1, 2)	(0, 0)

	fight	not fight
fight	(-2, 1)	(2, -1)
not fight	(-1, 2)	(0, 0)

假设你和另一个同学互相之间闹矛盾，你们可以选择打架或者不打架，但是你们不知道对方的更强还是自己更强，因此出现了上面两个表格的情况：左边的表格代表你更强，右边的表格代表对方更强

- 如果你知道自己更强，那结局就是 (*fight*, *not fight*)，这是严格占优策略，反之应当是 (*not fight*, *fight*)
- 因此如果你们两者谁更强这一事实已知，那么你们不会打架，一方会直接投降
- 现实中存在冲突是因为你们两者谁更强这一事实未知

不完全信息海萨伊博弈

海萨伊模型包含两个要素：一是参与人可能参与的博弈，称为“状态博弈”，第二个是参与人拥有的关于状态博弈和其他参与人的信念

定义 (不完全信息海萨伊博弈)

不完全信息海萨伊博弈是有序五元组 $(I, (T_i)_{i \in I}, p, S, (s_t)_{t \in \times_{i \in I} T_i})$ ，其中

- ① I 是参与人的有限集合
- ② T_i 是参与人 i 类型的有限集合，类型向量集合表示为 $T = \times_{i \in I} T_i$
- ③ $p \in \Delta(T)$ 是类型向量集合上的概率分布，对每个参与人 $i \in N$ 和每个类型 $t_i \in T_i$, $p(t_i) = \sum_{t_{-i} \in T_{-i}} p(t_i, t_{-i}) > 0$
- ④ S 是自然状态集，也被称为**状态博弈 (state games)**，每个自然状态 $s \in S$ 是向量 $s = (I, (A_i)_{i \in I}, (u_i)_{i \in I})$ ，其中 A_i 是参与人 i 的非空行动集合， $u_i : T \times (\times_{i \in I} A_i) \rightarrow \mathbb{R}$ 是参与人 i 的效用函数
- ⑤ $s_t = (I, (A_i(t_i))_{i \in I}, (u_i(t_i))_{i \in I}) \in S$ 是状态向量 $t \in T$ 的状态博弈，因此参与人 i 在状态博弈 s_t 中的行动集只取决于他的类型 t_i ，独立于其他参与人的类型

不完全信息海萨伊博弈 (Cont'd)

不完全信息博弈按照以下顺序进行（可以视为扩展式博弈）：

- ① 随机行动根据概率分布 p 选择类型向量 $t = (t_1, \dots, t_n)$
- ② 每个参与人 i 知道自己的类型 t_i ，但不知道其他参与人的类型 t_{-i}
- ③ 参与人同时选择自己的行动：每个参与人 i 知道自己的类型 t_i ，选择行动 $a_i \in A_i(t_i)$
- ④ 每个参与人 i 得到收益 $u_i(t; a)$ ，其中 $a = (a_1, \dots, a_n)$ 是所有参与人的行动向量

不完全信息海萨伊博弈 (Cont'd)

不完全信息博弈按照以下顺序进行（可以视为扩展式博弈）：

- ① 随机行动根据概率分布 p 选择类型向量 $t = (t_1, \dots, t_n)$
- ② 每个参与人 i 知道自己的类型 t_i ，但不知道其他参与人的类型 t_{-i}
- ③ 参与人同时选择自己的行动：每个参与人 i 知道自己的类型 t_i ，选择行动 $a_i \in A_i(t_i)$
- ④ 每个参与人 i 得到收益 $u_i(t; a)$ ，其中 $a = (a_1, \dots, a_n)$ 是所有参与人的行动向量

代入前面的例子，“上帝”首先随机选择你和对手的强弱类型，你们只知道自己的类型，不知道对方，然后你们选择自己的类型下可以选择的行动，最后根据你们的行动和强弱类型得到收益

- 需要注意的是，收益函数在结果出来之前是未知的，因为你不确定对手的类型，例如因为你不知道对手强弱，因此不知道两个人打起来是谁胜谁负，但是行动集只与自己的类型有关，与对手无关

不完全信息海萨伊博弈 (Cont'd)

- 在海萨伊模型中，类型集合是所有参与人之间的共同知识，并且所有人类型上的先验分布是共同的（因此还有更一般的信念空间）
- 参与人 i 在知道了自己的类型 i 之后，对于其他人的类型 t_{-i} 的信念是条件概率 $p(t_{-i} | t_i)$
 - 这被称为一阶信念 (**belief**)，事实上还有高阶的信念，例如我认为你是类型 1 的概率为 p ，是类型 2 的概率为 $1 - p$ ，这是一阶信念，接下来我可以分别算出你在这两种类型下你的信念，然后求期望，这是二阶信念，即我认为你认为我的类型是什么的概率
- 在经济学文献中，在每个人类型被指派之前的阶段被称为事前阶段 (**ex ante**)，在每个人类型被指派之后的阶段被称为事中阶段 (**interim**)

不完全信息博弈的均衡

我们同样利用纳什均衡来描述不完全信息博弈下博弈的解，在此之前我们先要定义不完全信息博弈中的策略

- 参与人 i 的纯策略是函数 $s_i : T_i \rightarrow \bigcup_{t_i \in T_i} A_i(t_i)$ ，满足 $s_i(t_i) \in A_i(t_i)$ ，即 $s_i(t_i)$ 是策略 s_i 为类型 t_i 的参与人 i 指定的行动
- 混合策略是纯策略上的概率分布，即 $\sigma_i : T_i \rightarrow \bigcup_{t_i \in T_i} \Delta(A_i(t_i))$ ，满足 $\sigma_i(t_i) = (\sigma_i(t_i; a_i))_{a_i \in A_i(t_i)} \in \Delta(A_i(t_i))$ ，也就是说 $\sigma_i(t_i; a_i)$ 是类型 t_i 的参与人 i 选择行动 a_i 的概率
- 如果参与人的行为策略向量是 $\sigma = (\sigma_1, \dots, \sigma_n)$ ，随机行动选择的类型向量是 $t = (t_1, \dots, t_n)$ ，那么每个行动向量 $(a_1, \dots, a_n)$ 被选择的概率是 $\prod_{i \in I} \sigma_i(t_i; a_i)$ ，因此参与人 i 的期望收益是

$$U_i(t; \sigma) = \mathbf{E}_\sigma[u_i(t)] = \sum_{a \in \times_{i \in I} A_i} \prod_{i \in I} \sigma_i(t_i; a_i) u_i(t; a)$$

- 参与人 i 的事前阶段收益为 $U_i(\sigma) = \sum_{t \in T} p(t) U_i(t; \sigma)$
- 参与人 i 的事中阶段收益为 $U_i(\sigma | t_i) = \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) U_i(t; \sigma)$

不完全信息博弈的均衡 (Cont'd)

不难证明事前阶段期望收益与事中阶段期望收益的关系

$$U_i(\sigma) = \sum_{t_i \in T_i} p(t_i) U_i(\sigma | t_i)$$

事实上，我们现在可以定义两类纳什均衡：第一类是事前均衡，没有人在知道自己类型之前愿意偏离自己的策略，第二类是事中均衡，没有人在自己的类型确定之后愿意偏离自己的策略

定义

策略向量 $\sigma^* = (\sigma_1^*, \dots, \sigma_n^*)$ 是不完全信息博弈的

- ① **纳什均衡 (Nash equilibrium)**，如果对每个参与人 i 和每个策略 σ_i ，有 $U_i(\sigma^*) \geq U_i(\sigma_i, \sigma_{-i}^*)$
- ② **贝叶斯均衡 (Bayesian equilibrium)**，如果对每个参与人 i ，每个类型 t_i 以及每个可能行动 $a_i \in A_i(t_i)$ ，有 $U_i(\sigma^* | t_i) \geq U_i((a_i, \sigma_{-i}^*) | t_i)$

回忆 σ_i 是函数以及混合策略纳什均衡等价定义可知上述定义合理性

不完全信息博弈的均衡 (Cont'd)

因为对于任意有限扩展式博弈都存在混合策略纳什均衡，而不完全信息博弈可以视为扩展式博弈，因此不完全信息博弈一定存在混合策略纳什均衡；而不完全信息博弈也可以表示为等价的策略式代理人博弈，因此不完全信息博弈一定存在贝叶斯均衡

定理

有限不完全信息博弈一定存在纳什均衡和贝叶斯均衡。

不完全信息博弈的均衡 (Cont'd)

因为对于任意有限扩展式博弈都存在混合策略纳什均衡，而不完全信息博弈可以视为扩展式博弈，因此不完全信息博弈一定存在混合策略纳什均衡；而不完全信息博弈也可以表示为等价的策略式代理人博弈，因此不完全信息博弈一定存在贝叶斯均衡

定理

有限不完全信息博弈一定存在纳什均衡和贝叶斯均衡。

下面这一定理非常重要，统一了不完全信息博弈的均衡概念

定理

有限不完全信息博弈中，每个贝叶斯均衡都是一个纳什均衡，每个纳什均衡都是一个贝叶斯均衡。

根据两个均衡之间的关联直接可以证明。关于不完全信息博弈，还有一个重要的直观在于可以解释混合策略的合理性：混合策略可以视为在每个类型上选择纯策略，但类型是不确定的。每个人在选择上添加随机性，实际上就是为了隐藏自己的类型，因此混合策略是合理的

两人冲突例子的计算

在本小节开头的例子中，参与人 I 的类型有两种：强或弱，我们假设强和弱的概率相等，均为 $1/2$ ，参与人 II 只有一种类型，现在我们要计算这个博弈的均衡。注意均衡策略是一个从自己的类型出发的函数，因此每种类型的策略我们都需要假设出来：设 I 强的时候选择 **fight** 的概率为 x ，于是 **not fight** 的概率为 $1 - x$ ；弱的时候选择 **fight** 的概率为 y ，**not fight** 的概率为 $1 - y$ 。II 的策略是选择 **fight** 的概率为 q ，**not fight** 的概率为 $1 - q$ 。

博弈开始后，参与人 I 知道了自己的类型，如果他知道自己强，事实上 **fight** 是占优策略，因此 $x = 1$ 。如果自己弱，那么可以算出自己的期望收益为 $-3yq + 2y - q$ ，求导即可得到对于每个 q 的最优反应。对于 II，他不知道 I 的类型，只知道 I 分别以 $1/2$ 的概率是强和弱，因此他的期望收益为 $1/2(q - y - 1)$ ，因此最优反应直接得到是 $q = 1$ ，代回弱的情况下的最优反应，解得 $y = 0$ 。因此如果 I 是强的，两个人会真的打起来，并且 I 获胜；如果 I 是弱的，I 会向 II 投降。

第一价格拍卖的对称均衡

在第一价格拍卖中，尽管诚实报价并非最优，但在一定的假设下我们可以求出贝叶斯纳什均衡。假设有 N 个拍卖参与人，所有参与人对物品的价值都有相同的先验分布 $X \sim F$ （相应的密度函数为 f ）。显然我们会猜想这种情况会有一个对称均衡：即每个参与人的出价都等于 $\beta(x)$ ，其中 x 是每个参与人自己的私有价值——这是因为在这一拍卖中每个人都是对称的，所以我们有理由相信有一个所有人策略都一致的均衡。

注意到只有参与人 i 提交的报价最高才能赢得拍卖：设参与人 i 的私人价值是 x ，报价为 b ，假设其他人都使用对称均衡，那么当 $b > \max_{j \neq i} \beta(X_j) = \beta(\max_{j \neq i} X_j)$ 时（等号是因为我们有理由相信 β 是递增的），参与人 i 会赢得拍卖。设 $Y_1 = \max_{j \neq i} X_j$ （实际上这就是一个次序统计量），假设 $Y_1 \sim G$ ，对应的密度函数为 g ，那么参与人 i 在拍卖中的期望收益等于

$$G(\beta^{-1}(b)) \times (x - b)$$

第一价格拍卖的对称均衡 (Cont'd)

关于 b 求最大化一阶条件可以得到

$$\frac{g(\beta^{-1}(b))}{\beta'(\beta^{-1}(b))}(x - b) - G(\beta^{-1}(b)) = 0$$

我们知道均衡下 $b = \beta(x)$ ，因此上式可以进一步改写为

$$G(x)\beta'(x) + g(x)\beta(x) = xg(x)$$

这等价于 $\frac{d}{dx}(G(x)\beta(x)) = xg(x)$ ，然后利用 $\beta(0) = 0$ 并两边积分可得

$$\beta(x) = \frac{1}{G(x)} \int_0^x yg(y)dy = \mathbf{E}[Y_1 \mid Y_1 < x]$$

当然上面的推导是假定存在这样的对称均衡得到的，我们还需要验证这一策略的确构成对称均衡，验证在此略去。

收入等价原理

由此我们可以看到在对称均衡下第一价格拍卖中，估价为 x 的竞拍者的期望支付为

$$\beta(x) \cdot G(x) = G(x) \cdot \mathbf{E}[Y_1 \mid Y_1 < x]$$

而在第二价格拍卖中，估价为 x 的竞拍者的期望支付为赢得拍卖的概率乘以自己是最高估价时第二高估价的期望值，即

$$G(x) \cdot \mathbf{E}[Y_1 \mid Y_1 < x]$$

我们发现，二者的表达式完全一致。事实上，这并不是巧合，而是下面这一定理的一个特例：

定理 (收入等价原理)

假设估价是独立同分布的，并且所有的竞拍者都是风险中性的，那么只要估价为 0 的竞拍者的期望支付为 0，任何标准拍卖（即报价最高者得到物品）的对称递增均衡都会导致卖家获得相等的期望收入。

孔雀开屏

事实上，在前面的例子中，我们发现当参与人 I 类型是强的时候，两个人真的会产生冲突——但这并不是我们想看到的：如果 I 是强的，他应当建议对手投降，这样对双方都会更好。因此我们引入**信号博弈 (signaling game)**，这是一种不完全信息博弈，其中参与人可以发送信号来影响其他参与人的信念，例如此时 I 可以发送信号告诉 II 自己是强的，从而避免冲突。

孔雀开屏

事实上，在前面的例子中，我们发现当参与人 I 类型是强的时候，两个人真的会产生冲突——但这并不是我们想看到的：如果 I 是强的，他应当建议对手投降，这样对双方都会更好。因此我们引入**信号博弈 (signaling game)**，这是一种不完全信息博弈，其中参与人可以发送信号来影响其他参与人的信念，例如此时 I 可以发送信号告诉 II 自己是强的，从而避免冲突。

然而对手可以不相信你的说辞，因此你需要给出一些实际的证明，例如你可以举起很重的杠铃。自然界中的孔雀开屏就是这个道理：雄性孔雀希望通过自己的开屏来显示自己的强壮，从而吸引异性。



信息设计

通常而言，改变人们行动的方式可以大致分为以下两类：

① 机制设计 (mechanism design)

- 拍卖：二价拍卖是诚实报价占优的
- 折扣/消费券：对消费的激励
- 工作合同设计：例如较高的绩效提成等

信息设计

通常而言，改变人们行动的方式可以大致分为以下两类：

① 机制设计 (mechanism design)

- 拍卖：二价拍卖是诚实报价占优的
- 折扣/消费券：对消费的激励
- 工作合同设计：例如较高的绩效提成等

② 信息设计 (information design) 或劝说 (persuasion)

- 兵者，诡道也
- 新闻媒体报道，广告，博客...
- 策略性信息泄露：例如孔雀开屏，“仅剩最后一件”等

信息设计

通常而言，改变人们行动的方式可以大致分为以下两类：

① 机制设计 (mechanism design)

- 拍卖：二价拍卖是诚实报价占优的
- 折扣/消费券：对消费的激励
- 工作合同设计：例如较高的绩效提成等

② 信息设计 (information design) 或劝说 (persuasion)

- 兵者，诡道也
- 新闻媒体报道，广告，博客...
- 策略性信息泄露：例如孔雀开屏，“仅剩最后一件”等

需要注意的是，信号博弈是想要通过行动让他人知道我的真实信息，例如孔雀开屏，或者公司选在高价写字楼中办公等；但一般的信息设计并不要求完全展示真实信息，而是要通过操纵信息达到自身的目的，例如接下来的例子

导师如何写推荐信



假设导师要向企业的人事写推荐信推荐自己的学生

- $1/3$ 的学生是优秀的， $2/3$ 的学生是普通的
- 对于人事而言，招一个优秀学生的的好处是 $1 + \varepsilon$ ，招一个普通学生的的好处是 -1 ，不招人的好处是 0
- 显然，先验来看，如果没有导师推荐信区分学生水准，人事不会选择招聘这个学校的学生，因为期望收益是

$$1/3 \times (1 + \varepsilon) + 2/3 \times (-1) = 1/3\varepsilon - 1/3 < 0$$

- 对于导师而言，一个学生被录用的效用为 1 ，否则为 0

导师如何写推荐信 (Cont'd)

显然，导师现在具有信息优势，那么他如何利用这个信息优势来提高自己的效用呢？

- 导师写推荐信声称每个学生都很优秀
 - 显然，这样下去人事会知道这个导师的推荐信毫无效力，写了等于没写，因此保持不招聘的选择，导师效用为 0

导师如何写推荐信 (Cont'd)

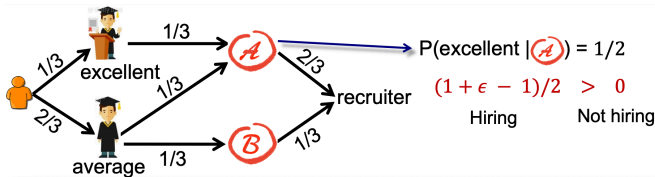
显然，导师现在具有信息优势，那么他如何利用这个信息优势来提高自己的效用呢？

- 导师写推荐信声称每个学生都很优秀
 - 显然，这样下去人事会知道这个导师的推荐信毫无效力，写了等于没写，因此保持不招聘的选择，导师效用为 0
- 导师诚实地写推荐信（将信息完全暴露）
 - 这样人事会知道这个导师的推荐信是可信的，因此会选择招聘这个学校的学生，因为招聘到的每个学生都能带来正的效用，此时导师效用为 $1/3 \times 1 + 2/3 \times 0 = 1/3$

导师如何写推荐信 (Cont'd)

显然，导师现在具有信息优势，那么他如何利用这个信息优势来提高自己的效用呢？

- 导师写推荐信声称每个学生都很优秀
 - 显然，这样下去人事会知道这个导师的推荐信毫无效力，写了等于没写，因此保持不招聘的选择，导师效用为 0
- 导师诚实地写推荐信（将信息完全暴露）
 - 这样人事会知道这个导师的推荐信是可信的，因此会选择招聘这个学校的学生，因为招聘到的每个学生都能带来正的效用，此时导师效用为 $1/3 \times 1 + 2/3 \times 0 = 1/3$
- 导师提供有噪声的信息，此时导师效用为 $2/3$



贝叶斯劝说

基本模型：

- 两个参与人：劝说人（persuader，或 sender）和决策者（decision maker，或 receiver）
- 决策者希望选择行动 $i \in [n] = \{1, \dots, n\}$ ， $\theta \in \Theta$ 表示自然的状态
 - 决策者的收益 $r(i, \theta)$
 - 劝说人的效用 $s(i, \theta)$
- 两个参与人都知道 θ 的先验分布，但劝说人具有信息优势：他知道真实的 θ
- 劝说人希望策略性地显示 θ 的信息从而“劝说”决策者做出有利于劝说人的决策

贝叶斯劝说

基本模型：

- 两个参与人：劝说人（persuader，或 sender）和决策者（decision maker，或 receiver）
- 决策者希望选择行动 $i \in [n] = \{1, \dots, n\}$ ， $\theta \in \Theta$ 表示自然的状态
 - 决策者的收益 $r(i, \theta)$
 - 劝说人的效用 $s(i, \theta)$
- 两个参与人都知道 θ 的先验分布，但劝说人具有信息优势：他知道真实的 θ
- 劝说人希望策略性地显示 θ 的信息从而“劝说”决策者做出有利于劝说人的决策

在刚刚的例子中：

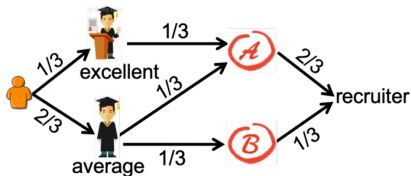
- 劝说人是导师，决策者是人事
- 决策者的决策 $i \in \{\text{招聘}, \text{不招聘}\}$ ，自然状态是 $\theta \in \{\text{优秀}, \text{普通}\}$

完全隐藏信息和完全显示信息都不一定对劝说人最优，那么我们要考虑显示部分信息的机制

信号

定义

一个信号机制 (signaling scheme) 是一个映射 $\pi : \Theta \rightarrow \Delta(\Sigma)$, 其中 Σ 是信号集合. π 可以完全被 $\{\pi(\sigma | \theta)\}_{\theta \in \Theta, \sigma \in \Sigma}$ 所决定, 其中 $\pi(\sigma | \theta)$ 是 θ 下信号 σ 的概率



- 例如左图中,
 $\Theta = \{\text{优秀, 普通}\}$,
 $\Sigma = \{A, B\}$, $\pi(A, \text{普通}) = 1/2$
- 注意 π 也是共同知识

在如上定义下, 决策者接收到信号 σ 之后, 可以计算出 θ 的后验概率分布

$$p(\theta | \sigma) = \frac{p(\theta)\pi(\sigma | \theta)}{\sum_{\theta' \in \Theta} p(\theta')\pi(\sigma | \theta')}$$

例如 $p(\text{优秀} | A) = 1/2$

决策者的效用分析

除了有利于劝说人，我们还想知道信号机制对决策者的效用有何影响

- 在接收到信号 σ 之后，决策者的期望效用是

$$\begin{aligned}
 R(\sigma) &= \max_{i \in [n]} \left[\sum_{\theta \in \Theta} p(\theta | \sigma) r(i, \theta) \right] \\
 &= \max_{i \in [n]} \left[\sum_{\theta \in \Theta} r(i, \theta) \cdot \frac{p(\theta) \pi(\sigma | \theta)}{\sum_{\theta' \in \Theta} p(\theta') \pi(\sigma | \theta')} \right]
 \end{aligned}$$

- 而发出信号 σ 的概率为 $p(\sigma) = \sum_{\theta' \in \Theta} p(\theta') \pi(\sigma | \theta')$ ，因此决策者的期望效用为

$$\begin{aligned}
 R &= \sum_{\sigma \in \Sigma} p(\sigma) R(\sigma) \\
 &= \sum_{\sigma \in \Sigma} \sum_{\theta' \in \Theta} p(\theta') \pi(\sigma | \theta') \max_{i \in [n]} \left[\sum_{\theta \in \Theta} r(i, \theta) \cdot \frac{p(\theta) \pi(\sigma | \theta)}{\sum_{\theta' \in \Theta} p(\theta') \pi(\sigma | \theta')} \right] \\
 &= \sum_{\sigma \in \Sigma} \max_{i \in [n]} \left[\sum_{\theta \in \Theta} r(i, \theta) p(\theta) \pi(\sigma | \theta) \right]
 \end{aligned}$$

决策者的效用分析 (Cont'd)

定理

决策者的期望效用在任意信号机制下都是不差于无信号机制的效用的。

决策者的效用分析 (Cont'd)

定理

决策者的期望效用在任意信号机制下都是不差于无信号机制的效用的。

$$\begin{aligned}
 R &= \sum_{\sigma \in \Sigma} \max_{i \in [n]} \left[\sum_{\theta \in \Theta} r(i, \theta) p(\theta) \pi(\sigma | \theta) \right] \\
 &\geq \max_{i \in [n]} \sum_{\sigma \in \Sigma} \sum_{\theta \in \Theta} r(i, \theta) p(\theta) \pi(\sigma | \theta) \\
 &= \max_{i \in [n]} \sum_{\theta \in \Theta} r(i, \theta) p(\theta) \sum_{\sigma \in \Sigma} \pi(\sigma | \theta) \\
 &= \max_{i \in [n]} \sum_{\theta \in \Theta} r(i, \theta) p(\theta)
 \end{aligned}$$

其中第一行是任意信号机制下的期望效用，最后一行是没有任何信号机制下（因此只有先验分布）的期望效用，故得证。

最优劝说的信号设计

接下来我们需要研究最优的信号机制 $\pi(\sigma | \theta)$ ，使得劝说人的效用最大化。第一个问题就是，信号的设计方式可以非常多，信号的个数甚至都不确定，下面的定理解决了这一问题

定理

使用 n （即决策者的行动个数）个信号就可以导出最优信号机制，其中每个信号 σ_i 对应一个最优的决策者行动 i 。

最优劝说的信号设计

接下来我们需要研究最优的信号机制 $\pi(\sigma | \theta)$ ，使得劝说人的效用最大化。第一个问题就是，信号的设计方式可以非常多，信号的个数甚至都不确定，下面的定理解决了这一问题

定理

使用 n （即决策者的行动个数）个信号就可以导出最优信号机制，其中每个信号 σ_i 对应一个最优的决策者行动 i 。

在信号 σ 下，决策者选择最优的行动

$$i^* = \arg \max_{i \in [n]} \sum_{\theta \in \Theta} r(i, \theta) p(\theta | \sigma)$$

命题：如果两个信号 σ 和 σ' 对应的最优行动相同，那么对任意的 $\lambda \in [0, 1]$ ， $\lambda\sigma + (1 - \lambda)\sigma'$ 对应的最优行动也是相同的

- 这些信号下最优行动相同，因此劝说人能获得的效用相同
- 因此可以把所有最优行动为 i^* 的信号合并成一个 σ_{i^*}

最优劝说的信号设计 (Cont'd)

命题：如果两个信号 σ 和 σ' 对应的最优行动相同，那么对任意的 $\lambda \in [0, 1]$ ， $\lambda\sigma + (1 - \lambda)\sigma'$ 对应的最优行动也是相同的

$$\sum_{\theta \in \Theta} r(i^*, \theta) p(\theta | \sigma) \geq \sum_{\theta \in \Theta} r(i, \theta) p(\theta | \sigma), \forall i \in [n]$$

$$\sum_{\theta \in \Theta} r(i^*, \theta) p(\theta | \sigma') \geq \sum_{\theta \in \Theta} r(i, \theta) p(\theta | \sigma'), \forall i \in [n]$$

将第一个不等式乘以 λ ，第二个不等式乘以 $1 - \lambda$ ，相加得到

$$\sum_{\theta \in \Theta} [\lambda p(\theta | \sigma) + (1 - \lambda) p(\theta | \sigma')] r(i^*, \theta)$$

$$\geq \sum_{\theta \in \Theta} [\lambda p(\theta | \sigma) + (1 - \lambda) p(\theta | \sigma')] r(i, \theta), \forall i \in [n]$$

因此 $\lambda\sigma + (1 - \lambda)\sigma'$ 对应的最优行动也是相同的

最优劝说的线性规划

已知先验分布 $p(\theta)$ ，决策者的收益 $r(i, \theta)$ ，劝说人的效用 $s(i, \theta)$ ，我们的目标是：设计劝说人的 n 个信号 $(\sigma_i)_{i \in [n]}$ 使得劝说人的效用最大化，根据贝叶斯公式，问题转为设计信号机制 $\pi(\sigma_i | \theta)$

最优劝说的线性规划

已知先验分布 $p(\theta)$ ，决策者的收益 $r(i, \theta)$ ，劝说人的效用 $s(i, \theta)$ ，我们的目标是：设计劝说人的 n 个信号 $(\sigma_i)_{i \in [n]}$ 使得劝说人的效用最大化，根据贝叶斯公式，问题转为设计信号机制 $\pi(\sigma_i | \theta)$

$$\begin{aligned}
 & \max_{\pi} \sum_{\theta \in \Theta} \sum_{i=1}^n s(i, \theta) \pi(\sigma_i | \theta) p(\theta) \\
 & \text{s.t.} \sum_{\theta \in \Theta} r(i, \theta) \pi(\sigma_i | \theta) p(\theta) \geq \sum_{\theta \in \Theta} r(j, \theta) \pi(\sigma_i | \theta) p(\theta), \forall i, j \in [n] \\
 & \sum_{i=1}^n \pi(\sigma_i | \theta) = 1, \forall \theta \in \Theta \\
 & \pi(\sigma_i | \theta) \geq 0, \forall \theta \in \Theta, i \in [n]
 \end{aligned}$$

其中第一个约束表明， σ_i 下的最优行动是 i ，后面两个约束是普通的正则性约束。事实上第一个约束会让我们想起相关均衡的约束，因此劝说类似于向对方提供信号 σ_i 以推荐行动 i ，对方不愿意偏离。

信息市场

数据市场与贝叶斯劝说相关的一个重要话题是信息市场，即市场上的信息交易。例如社交平台出售用户画像、移动公司等出售用户信息供银行分析贷款风险等。

信息市场

数据市场与贝叶斯劝说相关的一个重要话题是信息市场，即市场上的信息交易。例如社交平台出售用户画像、移动公司等出售用户信息供银行分析贷款风险等。

- 信息市场与贝叶斯劝说的关联是什么？
 - 在贝叶斯劝说中，劝说人部分地显示信息，以期望对方做出自己想要的选择，提高自己的效用
 - 在信息市场中，我们显示信息以求获得利润，但是两者信息差等其他设定没有区别
 - 因此信息市场模型相较于贝叶斯劝说，只需要将两者的角色变换为卖家和买家，卖家的效用变换为利润即可

信息市场

数据市场与贝叶斯劝说相关的一个重要话题是信息市场，即市场上的信息交易。例如社交平台出售用户画像、移动公司等出售用户信息供银行分析贷款风险等。

- 信息市场与贝叶斯劝说的关联是什么？
 - 在贝叶斯劝说中，劝说人部分地显示信息，以期望对方做出自己想要的选择，提高自己的效用
 - 在信息市场中，我们显示信息以求获得利润，但是两者信息差等其他设定没有区别
 - 因此信息市场模型相较于贝叶斯劝说，只需要将两者的角色变换为卖家和买家，卖家的效用变换为利润即可
- 利润最大化的情况是什么
 - 信息市场中卖家能赚取的利润实际上就是提供信息之后买家的效用与只有先验分布时买家的效用之差，即信息带给买家的额外效用
 - 因此最大的利润来源于卖家提供完整信息的情况，收益为

$$\sum_{\theta \in \Theta} p(\theta) \cdot \max_{i \in [n]} r(i, \theta) - \max_{i \in [n]} \sum_{\theta \in \Theta} r(i, \theta) p(\theta)$$

信息定价模型

但前面的定价模型并不符合实际：现实中我们可能不知道买家对信息的估值，因此我们重新建模如下：

- 参与人：信息的卖家和买家，分别对应于原先的劝说人和决策者
- 买家有一个决策任务（拥有更多信息完成任务更出色），可能的行动是 $i \in [n]$
- 买家有效用函数 $u(i, \theta; t)$ ，其中 $\theta \sim \mu$ 是自然状态， $t \sim f$ 是买家的类型，不同类型有不同效用函数
 - 其中 u, μ, f 都是共同知识， θ 和 t 是独立的
 - 卖家知道具体的 θ ，但不知道卖家具体的 t ；买家知道自己的 t ，但不知道具体的 θ

信息定价模型

但前面的定价模型并不符合实际：现实中我们可能不知道买家对信息的估值，因此我们重新建模如下：

- 参与人：信息的卖家和买家，分别对应于原先的劝说人和决策者
- 买家有一个决策任务（拥有更多信息完成任务更出色），可能的行动是 $i \in [n]$
- 买家有效用函数 $u(i, \theta; t)$ ，其中 $\theta \sim \mu$ 是自然状态， $t \sim f$ 是买家的类型，不同类型有不同效用函数
 - 其中 u, μ, f 都是共同知识， θ 和 t 是独立的
 - 卖家知道具体的 θ ，但不知道卖家具体的 t ；买家知道自己的 t ，但不知道具体的 θ

因此定价机制包含三步：第一收集买家的私人信息，第二出售信息，第三向买家收费。但这个过程可能是多轮的，并且买家可能误报自己的类型

- 显示原理：存在等价的激励相容的直接机制
- 因此我们的目标转向计算最优的激励相容机制

最优信息定价机制

Expected revenue

$$\begin{aligned}
 \max \quad & \sum_t f(t) \cdot x_t \\
 \text{s.t.} \quad & \sum_i \left[\sum_{\theta} \mu(\theta) \pi_t(\sigma_i, \theta) u(i, \theta; t) \right] - x_t \\
 & \geq \sum_i \max_j \left[\sum_{\theta} \mu(\theta) \pi_{t'}(\sigma_i, \theta) u(j, \theta; t) \right] - x_{t'}, \quad \text{for } t' \neq t \\
 & \sum_i \left[\sum_{\theta} \mu(\theta) \pi_t(\sigma_i, \theta) u(i, \theta; t) \right] - x_t \geq \max_i \sum_{\theta} \mu(\theta) u(i, \theta; t), \quad \text{for } t \\
 & \sum_{\theta} \mu(\theta) \pi_t(\sigma_i, \theta) u(i, \theta; t) \geq \sum_{\theta} \mu(\theta) \pi_t(\sigma_i, \theta) u(j, \theta; t), \quad \text{for } i \neq j, t \\
 & \sum_i \pi_t(\sigma_i, \theta) = 1, \quad \text{for } \theta, t \\
 & \pi_t(\sigma_i, \theta) \geq 0, \quad \text{for } t, \sigma_i, \theta
 \end{aligned}$$

我们仍然是设计信息机制 $\pi(\sigma_i, \theta)^2$ ，其中最后三条约束与最优劝说一致，重点在于前两条。第一条是激励相容约束，即诚实显示自己的类型是最优的；第二条是个人理性约束，即任何类型的参与人参与市场会比不参与好，或者说卖家收取的价格不超过买家因为购买信息而获得的效用的提升。

²图中的 $\pi(\sigma | \theta)$ 其实就是 $\pi(\sigma_i | \theta)$

其它相关话题

① Prediction Markets

- 有需要预测的信息，例如天气、股票的走势等，可以通过期权等汇集市场信息来实现
- 机器学习也可以实现预测，但机器学习基于历史数据，而预测市场基于人们的信念

② Scoring Rules

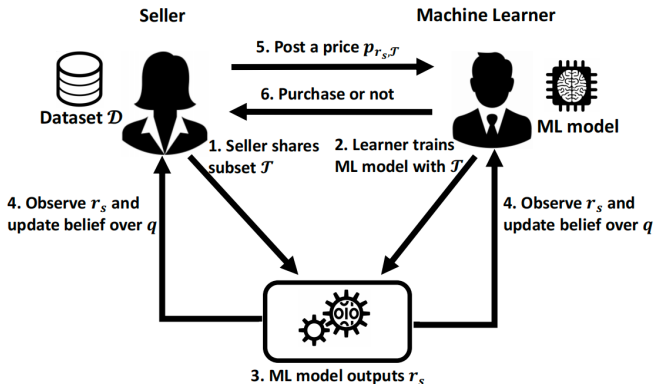
- 关于一件事情的预测，我们可以向专家寻求建议，但专家可能不会诚实地告诉我们结果
- 因此我们需要设计对专家意见的评分，给予专家和评分相关的奖励来激励专家诚实提供建议

③ Peer Prediction

- 通过 crowd sourcing 获得了很多信息，但无法判断信息的准确性，和 ground truth 比较是困难的
- 解决方案：很多人完成同一个任务，那就用他们完成的结果互相评价
- 应用：没有老师的课怎么评价学生的作业？数据提供者如何如实提供数据？

信号博弈在数据市场中的应用

“商品试用”的经济学³:



³Chen, J., Li, M., & Xu, H. (2022). Selling Data To a Machine Learner: Pricing via Costly Signaling. International Conference on Machine Learning.