博弈论第十四讲

授课时间: 2021年12月17日 授课教师: 张家琳 记录人: 王一鸣 吴凌轩

1 无秩序代价 (Price of Anarchy)

质量作为一项关乎每个人利益的事项，同样会在市场竞争和公共管理方面存在博弈，也会存在好的纳什均衡和坏的纳什均衡。均衡虽然能够让博弈进入稳定状态，但是均衡状态下社会福利可能很糟糕。所以定义均衡的无秩序代价（PoA）来衡量博弈中**均衡的质量**。

定义 1 (无秩序代价, POA). 博弈的无秩序代价为最坏均衡的社会代价与最小社会代价的比值。

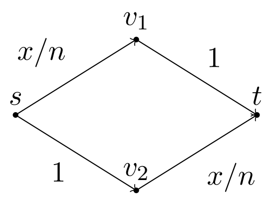
本章中，我们主要讨论纯策略纳什均衡的PoA。

1.1 布雷斯悖论(Braess's paradox)

布雷斯悖论是指在一个交通网络上增加一条路段反而使网络上的旅行时间增加。这一附加路段不但没有减少交通延滞，反而降低了整个交通网络的服务水准。这一悖论可以解释现有主要道路关闭后交通反而改善的例子，这种现象主要源于纳什均衡并不一定使社会最优化。

给出如下两个网络图的例子，分别计算其PoA。

例 1 n 位参与者在如下的有向图上进行 s 到 t 的路由博弈：



各边的延迟函数为：

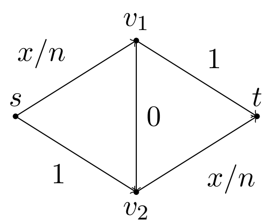
$$\begin{cases} d_{s\to v_1}(x) = x/n \\ d_{v_2\to t}(x) = x/n \\ d_{s\to v_2}(x) = 1 \\ d_{v_1\to t}(x) = 1 \end{cases}$$

最优局面：一半人选择 $s \to v_1 \to t$，另一半人选择 $s \to v_2 \to t$，社会代价为 $3n/2$。

最坏PNE：一半人选择$s \to v_1 \to t$，另一半人选择 $s \to v_2 \to t$，社会代价为 $3n/2$。

故PoA $= 1$。

例 2 在例1的基础上增加v_1到v_2的一条边：



各边的延迟函数为：

$$\begin{cases} d_{s\to v_1}(x) = x/n \\ d_{v_2\to t}(x) = x/n \\ d_{s\to v_2}(x) = 1 \\ d_{v_1\to t}(x) = 1 \\ d_{v_1\to v_2}(x) = 0 \end{cases}$$

最优局面：一半人选择 $s \to v_1 \to t$，另一半人选择 $s \to v_2 \to t$，社会代价为 $3n/2$。

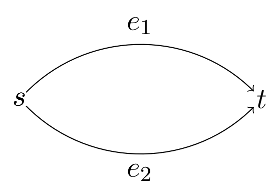
最坏PNE：全部选择 $s \to v_1 \to v_2 \to t$，社会代价为 $2n$。

故PoA $= 4/3$。

1.2 Pigou 网络

首先考虑特殊情况下的Pigou 网络形式。

例 3 n 位参与者在如下的有向图上进行 s 到 t 的路由博弈：



•各边的延迟函数为：

$$\begin{cases} d_{e_1}(x) = 1 \\ d_{e_2}(x) = x/n \end{cases}$$

最优局面：一半人选择 e_1，另一半人选择 e_2，社会代价为 $3n/4$。

最坏PNE：全部选择 e_2，社会代价为 n。

故PoA $= 4/3$。

•各边的延迟函数（非线性变种）为：

$$\begin{cases} d_{e_1}(x) = 1 \\ d_{e_2}(x) = (x/n)^p \end{cases}$$

最优局面：分配 $(1-\epsilon)n$ 的交通流到 e_2，其中 $\epsilon \to 0$。此时能够使得两条边的代价基本均衡，社会代价为 $\epsilon n + (1-\epsilon)n(1-\epsilon)^p = [\epsilon + (1-\epsilon)^{p+1}]n$

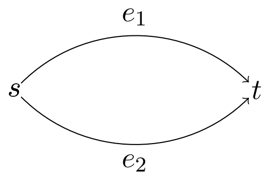
最坏PNE：全部选择 e_2，社会代价为 n。

当 $p \to \infty$时，PoA $= \frac{1}{\epsilon+(1-\epsilon)^{p+1}} \to \infty$。

1.3 Pigou-like 网络

将1.2中的网络模型拓展到一般情形。

例 4 在如下的有向图上，r单位的交通流量进行 s 到 t 的路由博弈：



各边的延迟函数为:

$$\begin{cases} d_{e_1}(x) = c(r) \\ d_{e_2}(x) = c(x) \end{cases}$$

其中，c 为给定的单调非降函数: $\mathbb{R} \to \mathbb{R}$。$r$为参与者数量。

首先考虑线性形式：$c(r) = ar + b, c(x) = ax + b$。显然，参与者的占有策略，即最坏PNE仍然是全部选择e_2，即最坏PNE为$rc(r) = r(ar + b)$。而最优局面为:

$$\min_x \{(r - x) \cdot (ar + b) + x \cdot (ax + b)\}$$

显然，$x = r/2$时取得最小值，即最优局面的社会代价为$\frac{3}{4}ar^2 + rb$。

此时，

$$\mathrm{PoA} = \frac{r(ar+b)}{\frac{3}{4}ar^2 + rb} = \frac{ar+b}{\frac{3}{4}ar+b} \leq \frac{4}{3}$$

将代价函数继续拓展至一般多项式形式，即$c(x) = \sum_{i=0}^{d} a_i x^i$，其中 d 为阶数。对于不同阶数的多项式形式，PoA值如下表所示：

Description	Typical Representative	Price of Anarchy
Linear	$ax + b$	$\frac{4}{3}$
Quadratic	$ax^2 + bx + c$	$\frac{3\sqrt{3}}{3\sqrt{3}-2} \approx 1.6$
Cubic	$ax^3 + bx^2 + cx + d$	$\frac{4\sqrt{4}}{4\sqrt{4}-3} \approx 1.9$
Quartic	$ax^4 + \dots$	$\frac{5\sqrt{5}}{5\sqrt{5}-4} \approx 2.2$
Polynomials of degree $\leq d$	$\sum_{i=0}^{d} a_i x^i$	$\frac{(d+1)\sqrt{(d+1)}}{(d+1)\sqrt{(d+1)}-d} \approx \frac{d}{\ln d}$

对于特定的代价函数集合$\mathcal{C}$，下面定理可以将计算PoA变成很简单的问题。考虑代价函数属于$\mathcal{C}$的所有网络，本来我们需要对这些可能的网络进行搜索，但是

定理 2 (自私路由的严格PoA界). *在所有代价函数集合为$\mathcal{C}$的网络中，和$Pigous$示例类似的网络的PoA最大。*

所以我们只需要在类似Pigou示例的简单网络中考察这个问题就好了。例如当$\mathcal{C}$是系数非负的线性函数，我们只需要在类Pigou网络中求解，结合上面的例子，立刻知道PoA至多为$\frac{4}{3}$。

2 扩展形式的博弈 (Extensive Form Game)

定义 3 (**扩展形式的博弈**). *扩展形式通过*树*来描述博弈。每个节点（称作决策节点）表示博弈进行中的每一个可能的状态。博弈从唯一的初始节点开始，通过由参与者决定的路径到达终端节点，此时博弈结束，参与者得到相应的收益。每个非终端节点只属于一个参与者；参与者在该节点选择其可能的行动，每个可能的行动通过边从该节点到达另一个节点。*

互动中，一个参与者可以在博弈中多次行动，并且在不同的状态中可以做出不同的行为。

例 5 一个双人博弈有选手 1 和 2。每个非终端节点上的数字表示该节点所属的参与者。终端节点上的数字表示参与者的收益。图片里每个边上的符号是这个边所代表的行动的名字。

初始节点属于参与者1，表示该参与者先动。

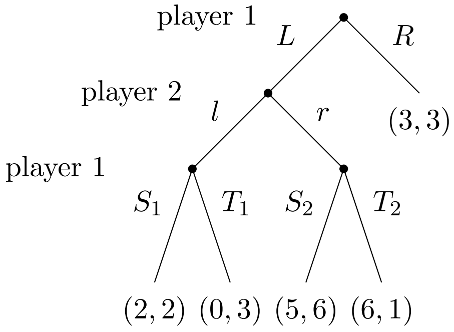


图 1: 扩展形式完美信息博弈树

该博弈树清楚地表示了参与者 1 先行动，参与者 2 观察到参与者 1 的行动，并在当前以当前状态为根节点的子树上继续当前博弈。然而，一些博弈并不是这样。参与者并不是一直能观察到另一个人的选择（例如，参与者同时行动或者其他参与者行动被隐藏）。此时，我们可以定义 *信息集* (information set) 如下：

定义 4 (信息集). *信息集是决策节点的集合：每个节点都属于一个参与者，且参与者无法区分信息集里的多个节点。*

在此基础上，我们可以得出 *完美信息博弈* (perfect information game) 和 *非完美信息博弈* (imperfect informationgame) 的定义如下：

定义 5 (完美信息博弈). *完美信息的博弈是指在博弈的任何阶段，每个参与者都清楚博弈之前发生的所有行动，即每个信息集都是一个单元素集合。*

定义 6 (非完美信息博弈). *存在参与人的信息集是多节点信息集的博弈。*

此时，玩家 i 的策略可以被表示为 $S_i : H_i \to \mathcal{A}_i$。其中 H_i 为玩家 i 的信息集，$\mathcal{A}_i$ 为玩家 i 所有可能的行动集合。

定义 7 (子博弈完美 (subgame perfect)). *若一个纳什均衡在每个子博弈中都是纳什均衡，则该均衡是子博弈完美的。*

此处子博弈指拓展形式博弈树的子树，且所有包含该子树节点的信息集也都完整地处于该子树中。我们可以使用 *逆向归纳* 的方法求解子博弈完美均衡。

对于例 5 的完美信息博弈，我们进行逆向归纳有:

- 第 3 回合，参与者 1 两个子博弈中分别选择占优策略 S_1, T_2；
- 第 2 回合，参与者 2 选择 l, r 的结果分别被归约为 $(2, 2)$, $(6, 1)$，参与者 2 选择占优策略 l；

- 第 1 回合，同理可得，参与者 1 会选择占优策略 R 。

综上所述，例 5 中的博弈有子博弈完美纳什均衡 玩家1选择$< R >$。

实际上所有的完美信息博弈都可以用逆向归纳的方式得到类似先手必胜或者后手必胜的结论。但是对实际的博弈问题，如围棋，虽然肯定存在先手或后手的必胜策略，但逆向归纳的复杂度过高，导致我们无法真的计算出来必胜策略。

例 6 对于**例 5** 中的博弈，我们可以将其转换为非完美信息博弈：参与者 1 不知道参与者 2 在上一轮的选择，即存在多节点的信息集（使用虚线相连表示）。

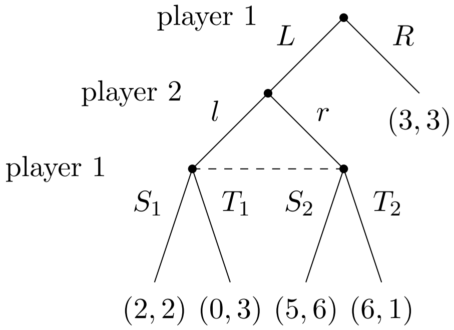


图 2: 扩展形式非完美信息博弈树

我们仍然使用逆向归纳法：

1. 首先，我们仍然逆向寻找最小单位的子博弈（该子博弈不会切分信息集），得到子博弈如下：

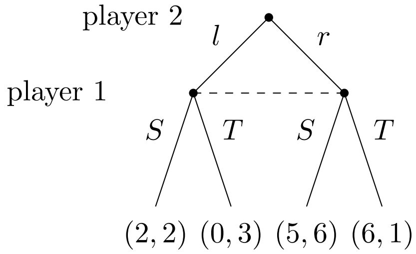


图 3: 非完美信息子博弈树

我们可以将该博弈使用表格表示，有：

player 2 \ player 1	S	T
l	2 2	0 3
r	5 6	6 1

图 4: 非完美信息子博弈表格表示

易得，此时有子博弈纳什均衡：对于玩家 1 有 $\mathbf{supp}(S_1) = \{S, T\}$，$D_1 = (\frac{1}{3}, \frac{2}{3})$；对于玩家 2 有 $\mathbf{supp}(S_2) = \{l, r\}$，$D_2 = (\frac{1}{3}, \frac{2}{3})$。在此均衡下，玩家 1 的期望收益是$1/9 \times 2 + 2/9 \times 0 + 2/9 \times 5 + 4/9 \times 6 = 4$，玩家 2 的期望收益是$1/9 \times 2 + 2/9 \times 3 + 2/9 \times 6 + 4/9 \times 1 = 8/3$。

2. 此时，博弈树可以规约为

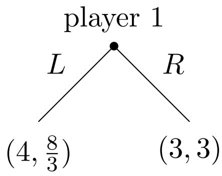


图 5: 规约博弈树

易得，此时玩家 1 会选择占优策略 L。

综上所述，该非完美信息博弈有子博弈完美纳什均衡如下：

- 对于玩家 1 有 $\mathbf{supp}(S_1) = \{< L, S >, < L, T >\}$，$D_1 = (\frac{1}{3}, \frac{2}{3})$；
- 对于玩家 2 有 $\mathbf{supp}(S_2) = \{l, r\}$，$D_2 = (\frac{1}{3}, \frac{2}{3})$，