



# 无秩序代价/自私路由/选址博弈

2024-2025 学年春夏学期计算经济学讨论班

郑涵文

zhwen@zju.edu.cn

2025 年 5 月 16 日



## Introduction

## 自私路由博弈

## 超额配置与单元自私路由

## 均衡的层级与存在性

## 选址博弈、平滑博弈

## 网络代价分摊博弈

## 最优反应动力学

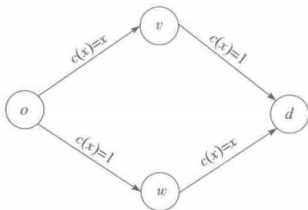
## 总结

# Introduction

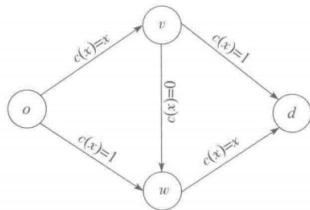
我们知道博弈中的“均衡”这一概念，并认为参与者没有单独偏离均衡的动机，但是均衡显然不一定是整体（社会）收益最优（代价最小）的，比如很经典的囚徒困境问题，（不承认，不承认）是最优解，但不是均衡。我们今天就要研究均衡和全局最优解之间的关系，并研究一些模型，期待这些模型的均衡能够接近全局最优解。

# 布雷斯悖论

我们回顾布雷斯悖论：一个单位的交通流（例如一群司机），从起点  $o$  出发向终点  $d$ 。每个司机都希望自己的时间尽可能短。对于网络来说，每条边都有一个花费函数。我们首先看初始网络，每个司机都有两条路径可以选择：走上面或是走下面。由于上面和下面两条路是等价的，所以有唯一的均衡：一半司机（0.5 单位的交通流）走上面，另一半走下面。这样每个司机的花费都是  $0.5 + 1 = 1.5$ 。如果我们改进网络，司机能够瞬间从  $v$  传送到  $w$ ，那么此时  $o \rightarrow v \rightarrow w \rightarrow d$  对每个司机来说都是占优策略，在这个均衡下每个人花费的时间是 2。而事实上像之前一样的话，最好的情况是每个人都花费  $3/2$ ，所以可以发现这个瞬间传送没有任何好处。



a) 初始网络



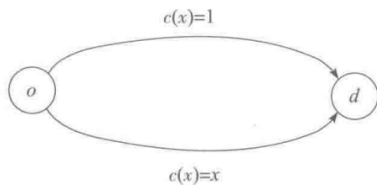
b) 改进网络

# 布雷斯悖论

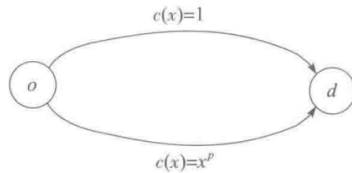
如果网络长成改进网络这个样子，那么在没有任何合作和第三方建议的情况下，司机们最终会收敛到花费为 2 的占优策略均衡上。但如果有人来维护秩序，就像交警和红绿灯那样，他们就可以把司机分流，最终使得每个司机都获益（花费降低到  $3/2$ ）。这就是有秩序和无秩序的区别。我们定义这个博弈下的**无秩序代价 (POA)** 为：均衡情况下所用的时间与最小的平均时间的比值，在这个例子中就是  $2/(3/2) = 4/3$ 。

# 庇古示例

图中左边的网络是一个更加简单的自私路由网络，它的 POA 是  $4/3$ 。均衡是所有人选择下面的边，而最优情况可以通过计算发现是上下各走一半的人。对于布雷斯示例和上述庇古示例 POA 都是  $4/3$ ，这还可以接受，说明我们的均衡不算太差。但是对于图中右边的庇古示例的非线性变种就不一样了。当  $p$  很大时，均衡没有变化，但是最优解变得好得多： $p \rightarrow \infty$  时下面的司机几乎不费时间就抵达了终点，所以分配  $1 - \epsilon$  的司机走下面，剩下的司机走上面，可以发现总时间也是趋于 0 的。因此当  $p \rightarrow \infty$  时 POA 无界。



a) Pigou 示例



b) Pigou 示例的一种非线性情况

图 11.2 Pigou 示例及其非线性变种

# POA 与线性代价函数

观察前面的几个例子，我们提出猜想：代价函数的高非线性是实现较小 POA 的唯一障碍。如果代价函数的非线性不高，无论网络怎么复杂，POA 都能够接近 1。我们将规范描述此问题并给出定理。

考虑以下模型：有向图  $G = (V, E)$ ，每条边  $e$  有一个非负的单调递增的代价函数  $c_e(x)$ ，表示流量为  $x$  时的代价。自然地，我们总是假设代价函数非负、连续、非递减。我们首先给出一个结论的非正式版本：

## 定理 1.1

在所有代价函数集合为  $\mathcal{C}$  的网络中，和庇古示例类似的网络的 POA 最大。

我们之后会对“和庇古示例类似的网络”进行精确的定义。这个定理的要点在于：最坏情况的例子总是很简单的那些网络，和 POA 大小的相关性不在于网络的复杂度而是代价函数。

# POA 与线性代价函数

举个例子，当  $C$  是系数非负的仿射代价函数集时，定理 1.1 就表明，我们前面讨论的庇古示例带来的 POA 是最大的，为  $4/3$ 。下表展示了代价函数为系数非负多项式时的情况：

表 11.1 如果自私路由网络代价函数为系数非负的多项式函数，且阶数最高为  $p$ ，其最坏情况下的 POA 如下所示

表达式	典型示例	无秩序代价
线性函数	$ax + b$	$4/3$
二次函数	$ax^2 + bx + c$	$\frac{3\sqrt{3}}{3\sqrt{3}-2} \approx 1.6$
三次函数	$ax^3 + bx^2 + cx + d$	$\frac{4\sqrt[3]{4}}{4\sqrt[3]{4}-3} \approx 1.9$
四次函数	$ax^4 + bx^3 + cx^2 + dx + e$	$\frac{5\sqrt[4]{5}}{5\sqrt[4]{5}-4} \approx 2.2$
阶数 $\leq p$	$\sum_{i=0}^p a_i x^i$	$\frac{(p+1)\sqrt[p]{p+1}}{(p+1)\sqrt[p]{p+1}-p} \approx \frac{p}{\ln p}$



# 主要结论的正式表述

首先我们正式定义一下类似庇古网络.

## 类似 Pigou 网络的组成成分

1. 两个顶点  $o$  和  $p$ 。
2. 从  $o$  到  $p$  的两条边，一条边在上面，一条边在下面。
3. 非负的交通流率  $r$ 。
4. 下面那条边的代价函数  $c(\cdot)$ 。
5. 上面那条边的代价函数处处相等且为  $c(r)$ 。



# 主要结论的正式表述

类庇古网络的 POA 很好算，因为下面那条边永远是占优策略，因此均衡的总时间为  $r \cdot c(r)$ 。  
 而最优解可以写成  $\inf_{0 \leq x \leq r} \{x \cdot c(x) + (r-x) \cdot c(r)\}$ ，因而 POA 可以写成  $\sup_{x \geq 0} \left\{ \frac{r \cdot c(r)}{x \cdot c(x) + (r-x) \cdot c(r)} \right\}$ 。

令  $\mathcal{C}$  表示任意非负、连续、非减的代价函数集合。庇古界定义为类庇古网络下，边代价函数属于  $\mathcal{C}$  的情况下 POA 的最大值：

$$\alpha(\mathcal{C}) = \sup_{c \in \mathcal{C}} \sup_{r \geq 0} \sup_{x \geq 0} \left\{ \frac{r \cdot c(r)}{x \cdot c(x) + (r-x) \cdot c(r)} \right\}$$

至此我们能够给出定理 1.1 的正式表述。

## 定理 1.2

对于所有代价函数的集合  $\mathcal{C}$  及代价函数属于  $\mathcal{C}$  的自私路由网络，其 POA 至多是  $\alpha(\mathcal{C})$ 。

# 超额配置

超额配置的动机在于，通信网络中增加额外的容量不会很花钱，虽然超额配置使得网络没有被完全利用，但是考虑到未来的需求的增长，以及网络冗余时较低的延迟和掉包率，超额配置具有较高的性价比。

我们考虑代价函数为

$$c_e(x) = \begin{cases} \frac{1}{u_e - x} & x < u_e \\ +\infty & x \geq u_e \end{cases}$$

参数  $u_e$  表示边  $e$  的容量。只有在流量接近容量时，代价会变得很大，否则代价的变化非常平稳且较小。超额配置就在于把  $u_e$  设置的比较大，使得流量很难接近  $u_e$ ，因此带来比较小的代价。

我们定义一个参数  $\beta$ ，如果说在上述代价函数下，对于网络中所有边都有  $f_e \leq (1 - \beta)u_e$ ，其中  $f$  是网络的均衡分流，我们就称这样的网络是  $\beta$ -超额配置的。这种情况下，代价函数事实上和低阶非负多项式类似。所以根据定理 1.2，这样的网络 POA 比较小。

## 结论

$\beta$ -超额配置的网络的最坏情况 POA 是  $\frac{1}{2}(1 + \sqrt{\frac{1}{\beta}})$ 。

# 资源增广界

我们现在讨论在对代价函数不做任何假设的情况下，任意自私路由网络的一个结论。回忆庇古示例的非线性变种，这种网络的 POA 很容易构造出无界的情况。因此这里研究的点不在于 POA。

对于图 11.2b，如果网络只有一个单位的流且  $p$  非常大，那么均衡代价为 1，最优分流代价几乎为 0；但是如果我们把网络的流量变成两个单位，结果就有变化了：由于下面边的限制，不能够给它超过 1 的流量，因此最优情况是给下面的边提供  $(1 - \epsilon)$  的流量，给上面的边提供  $(1 + \epsilon)$  的流量。这个时候，最优分流的代价超过了只传输一个单位的流量的均衡分流代价。

我们事实上能够得到对所有自私路由网络都成立的结论：

## 定理 2.1

在所有自私路由网络中，若  $r > 0$ ，则交通流率为  $r$  的均衡分流代价至多等于交通流率为  $2r$  的最优分流的代价。



# 单元自私路由

之前我们研究了自私路由的非单元模型，也就是说单个智能体占总流量份额很小，或者说流量是无限可分的情况。现在我们研究单元自私路由，现在每个智能体具有所谓“原子性”，每个智能体都占据交通流相当大的份额，总交通流不再是无限可分的了。

## 定义

单元自私路由网络由有向图  $G = (V, E)$  和  $k$  个智能体构成，其中边的代价函数非负且非减。智能体  $i$  有一个起点  $o_i$  和一个终点  $d_i$ 。每个智能体在一条  $o_i - d_i$  的路径上传输 1 个单位的交通流，同时他还会想方设法使其代价最小化。令  $P_i$  表示图  $G$  中  $o_i - d_i$  的路径。分流表示成向量  $(P_1, \dots, P_k)$  的形式，其中  $P_i \in \mathcal{P}_i$  为智能体  $i$  传输交通流的路径。均衡分流为没有任何一个智能体可以通过单方面改变策略来减少自身代价的分流。

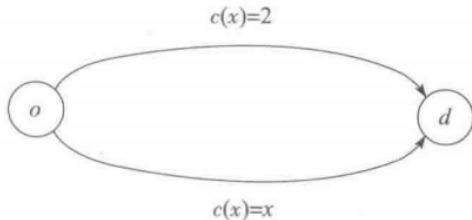
定义 12.2 (均衡分流 (单元)) 若对每个智能体  $i$  和路径  $\hat{P}_i \in \mathcal{P}_i$  有

$$\underbrace{\sum_{e \in P_i} c_e(f_e)}_{\text{改变策略前}} \leq \underbrace{\sum_{e \in \hat{P}_i \cap P_i} c_e(f_e) + \sum_{e \in \hat{P}_i \setminus P_i} c_e(f_e + 1)}_{\text{改变策略后}}$$

单元情况下的均衡和非单元情况下是不同的，因为智能体的大小不能忽略，因此在单个智能体改变策略时，边的流量会产生显著变化，因而导致边的代价改变。

# 单元自私路由

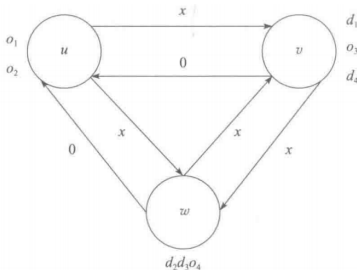
我们以下面的图作为一个例子. 假设有两个智能体, 每个智能体都控制一个单位的流. 最优解是一条边一个智能体, 代价为 3. 同时可以发现这这也是一个均衡. 但是还有另一个均衡: 由于下面的边是占优策略, 所以两个智能体都选择下面的边. 此时均衡代价为 4. 因此我们这里把 POA 的定义扩展到有多个均衡的博弈中: POA 是最坏情况均衡分流的代价除以最优分流的代价, 此处为  $4/3$ .



# 单元自私路由

非单元自私路由和单元自私路由的另一个区别是：后者的 POA 比前者大。为了验证这个区别，考虑有四个智能体的双向三角网络，如图 12.3 所示。每个智能体有两个选择，选一跳的路径和选两跳的路径。在最优分流中，每个智能体都会选一跳的路径来传输。这些一跳的路径就是四条代价函数为  $c(x) = x$  的边，所以最优分流的代价为 4。这个最优分流也是均衡流。另一方面，如果每个智能体都选择两跳的路径来传输，那么我们就得到第二个均衡分流（练习 12.5）。因为前两个智能体的代价为 3，后两个代价为 2，所以此时分流的代价为 10。这个网络的 POA 为  $10/4 = 2.5$ 。

代价函数为仿射函数的单元自私路由网络不会有比这更大的 POA。



# 单元自私路由

## 定理 2.2

在代价函数为仿射函数的单元自私路由网络中, POA 的上界为  $5/2$ .



# 均衡的层级

单元自私路由网络中的均衡分流是纯策略纳什均衡 (PNE) 的一种形式, 因为智能体不会随机选择路径. 但纯策略纳什均衡实际上是很罕见的, 比如石头剪刀布就没有纯策略纳什均衡. 一个更宽松的均衡概念是混合策略纳什均衡 (MNE). 但是事实上即使只有两个智能体, 计算混合策略纳什均衡也是很困难的, 这迫使人们去寻找更容易计算的均衡概念, 这就引出了相关均衡 (CE) 和粗糙相关均衡 (CCE).

# 相关均衡

## 定义：相关均衡

若对于任意一个智能体  $i \in \{1, 2, \dots, k\}$ , 策略  $s_i \in S_i$  及其单方面策略改变  $s'_i \in S_i$ , 都有

$$E_{s \sim \sigma}[C_i(s)|s_i] \leq E_{s \sim \sigma}[C_i(s'_i, s_{-i})|s_i] \quad (1)$$

则称局势集合  $S_1 \times \dots \times S_k$  上的分布  $\sigma$  是代价最小化博弈的一个相关均衡。

要注意的是定义中的  $\sigma$  不一定是联合概率分布。

为解释相关均衡，通常要引入一个信任的第三方。局势的分布  $\sigma$  是公共已知的。信任的第三方会根据  $\sigma$  抽取一个局势  $s$  的样本。对于所有智能体  $i = 1, 2, \dots, k$ , 第三方会私下建议智能体  $i$  选择策略  $s_i$ 。智能体  $i$  可以选择使用建议的策略  $s_i$ , 也可以选择不使用。在做决定的时候, 智能体  $i$  知道: 分布  $\sigma$  以及局势  $s$  中一个元素  $s_i$ 。因此, 在这些已知信息的情况下, 智能体会对局势  $s$  的其他项, 即第三方建议其他智能体的策略  $s_{-i}$  有一个后验估计。相关均衡要求了, 对于每个智能体, 使用第三方建议的策略能够最小化其期望代价。这个期望是在假设其他智能体使用第三方建议的策略的情况下计算的。

# 相关均衡

红绿灯就是一个属于 CE 但不属于 MNE 的很好的示例。考虑以下两个智能体的博弈，矩阵的行列分别表示不同智能体，矩阵内的单元值表示智能体采取一定行动后对应的代价：

	停	走
停	1, 1	1, 0
走	0, 1	5, 5

其中，等红灯代价适中为 1，发生车祸代价很大为 5。这个博弈有两个 PNE，(停，走) 和 (走，停)。令  $\sigma$  为这两个局势的随机均匀分布。注意这并不是整个博弈四个局势上的联合分布，所以它不是 MNE，但是它是一个 CE。比如，考虑行智能体，如果第三方（例如信号灯）建议选择策略“走”（信号灯为绿），那么行智能体就知道列智能体得到的建议是“停”（信号灯为红）。假设列智能体选择了第三方建议的策略——红灯停，那么行智能体的最优策略就是听从第三方的建议——绿灯行。类似的，若行智能体被告知要“停”，那么他就会假设列智能体会“走”，在这个假设下，“停”就是最优策略。

我们可以证明 CE 不是计算困难的，因此我们可以计算 CE 并分析 CE 的 POA 的界。

# 粗糙相关均衡

我们进一步扩大均衡的集合，得到一个“更容易处理”的均衡概念。

## 定义：粗糙相关均衡

若对于任意一个智能体  $i \in \{1, 2, \dots, k\}$ ，及其单方面策略改变  $s'_i \in S_i$ ，都有

$$E_{s \sim \sigma}[C_i(s)] \leq E_{s \sim \sigma}[C_i(s'_i, s_{-i})] \quad (2)$$

则称局势集合  $S_1 \times \dots \times S_k$  上的分布  $\sigma$  是代价最小化博弈上的一个粗糙相关均衡。

当智能体  $i$  考虑策略改变到  $s'_i$  时，他只知道分布  $\sigma$  而不知道实际的  $s_i$ 。换句话说，CCE 只防止无条件的单方面策略改变，而不像定义 13.4 中以  $s_i$  为条件。每一个 CE 也是一个 CCE，因此 CCE 一定存在且容易计算。

# 示例

接下来，我们利用一个实际例子来加强对图 13.1 中 4 个均衡概念的直观理解。考虑有四个智能体的单元自私路由网络 (12.4 节)。网络很简单，有一个公共起点  $o$ ，公共终点  $d$ ，边集  $E = \{0, 1, 2, 3, 4, 5\}$  分别表示 6 条平行的  $o$ - $d$  边。每条边的代价函数为  $c(x) = x$ 。纯策略纳什均衡就是每个智能体选择不同的边。每个智能体在纯策略纳什均衡中的代价为 1。一个明显不是纯策略的混合策略纳什均衡为，每个智能体独立、随机且均匀地选择一条边，每个智能体在混合策略纳什均衡中的期望代价为  $\frac{3}{2}$ 。满足一条边上有两个智能体且有另外两条边上分别有一个智能体的局势有多种，这些局势上的均匀分布就是一个相关均衡（非联合的）。这是因为对所有  $i$ ， $s_i$  和  $s'_i$  来说，式 (13.4) 的两边都是  $\frac{3}{2}$ 。如果上述这样的智能体选边的结果只能是边集  $\{0, 2, 4\}$  或  $\{1, 3, 5\}$ ，那么在这两个边集上的均匀分布就是一个粗糙相关均衡。这是因为对所有  $i$  和  $s'_i$  来说，式 (13.5) 的两边都是  $\frac{3}{2}$ 。注意这不是一个相关均衡，因为智能体  $i$  被建议选择边  $s_i$  时，他可以改变策略到  $s'_i$ ，即旁边的边（即按照边集中的数对 6 取模之后的值计数），从而将他的条件期望代价降到 1。

# 均衡分流的存在性

我们有以下结论：

## 定理 3.1

对于任意的最小单元自私路由网络，都存在至少一个均衡分流。

这里的“最小”是指每个智能体的流量都为 1 的情况。我们接下来将证明上述定理。

## 证明

构造单元自私路由网络中关于分流的一个函数，这个函数将是证明的关键：

$$\Phi(f) = \sum_{e \in E} \sum_{i=1}^{f_e} c_e(i) \quad (3)$$

其中  $f_e$  表示在分流  $f$  中使用边  $e$  的智能体数目。

考虑分流  $f$ ，智能体  $i$  使用  $f$  中的  $o_i - d_i$  路径  $P_i$ ，并可能改变策略选择其他  $o_i - d_i$  路径  $\hat{P}_i$ 。令  $\hat{f}$  表示智能体  $i$  从  $P_i$  变到  $\hat{P}_i$  后的分流。

# 均衡分流的存在性

## 证明 (续)

我们断言：

$$\Phi(\hat{f}) - \Phi(f) = \sum_{e \in \hat{P}_i} c_e(\hat{f}_e) - \sum_{e \in P_i} c_e(f_e) \quad (4)$$

具体来说，在智能体单方面改变策略偏离均衡的情况下， $\Phi$  函数的变化和偏离智能体自身代价的变化相同。因此，只需要一个函数  $\Phi$  就可以同时描述每个智能体的偏离所带来的影响。为证明式 (4)，我们需要考察当智能体  $i$  从  $P_i$  变为  $\hat{P}_i$  时，势函数  $\Phi$  如何变化。如果边  $e$  在  $\hat{P}_i$  中而不在  $P_i$  中，势函数中有关边  $e$  的项就多了一个  $c_e(f_e + 1)$ ；而如果边  $e$  在  $P_i$  中而不在  $\hat{P}_i$  中，势函数中有关边  $e$  的项就少了一个，成为  $c_e(f_e)$ 。因此，式 (4) 左边等于：

$$\sum_{e \in \hat{P}_i \setminus P_i} c_e(f_e + 1) - \sum_{e \in P_i \setminus \hat{P}_i} c_e(f_e) \quad (5)$$

这其实和式 (4) 右边是一样的。

# 均衡分流的存在性

## 证明 (续)

考虑能最小化  $\Phi$  的分流  $f$ . 因为可能的分流的数目是有限的, 所以能最小化  $\Phi$  的分流一定存在. 并且分流  $f$  中任何智能体单方面改变策略都再也不能减少  $\Phi$ . 根据式 (4), 也没有任何智能体能够通过单方面改变策略来减少自身的代价, 所以  $f$  是一个均衡分流.

对于非单元自私路由网络, 可以得到类似的结论.





# 拥塞博弈和势博弈

前面的证明中没使用到网络结构的信息. 这个证明对拥塞博弈依然成立, 拥塞博弈中有一个关于资源 (如前述的边) 的抽象集合  $E$ , 每个资源本身存在一个代价函数, 每个智能体  $i$  都有一个可以自由选择策略集  $S_i \subseteq 2^E$  (如前述的  $o_i - d_i$  路径集合), 每个策略都是一个资源子集.

势博弈是存在势函数  $\Phi$  的博弈, 势博弈有一个性质: 当某个智能体偏离均衡时, 势函数值的变化等于偏离者自身的代价变化. 即对每个局势  $s$ , 智能体  $i$  及其单方面策略改变  $s'_i \in S_i$ , 都有:

$$\Phi(s'_i, s_{-i}) - \Phi(s) = C_i(s'_i, s_{-i}) - C_i(s) \quad (6)$$

直观上来讲, 势博弈的智能体都在无意间共同优化  $\Phi$ . 在这几章中我们只考虑有限势博弈. 定理 3.1 的证明中的等价式 (4) 表明: 所有单元自私路由博弈, 或者更泛化地讲, 所有拥塞博弈都是势博弈.

定理 3.1 证明的最后一段意味着以下结论成立: 所有势博弈都至少有一个 PNE. 势函数是少有的证明 PNE 存在性的通用工具之一.

# 选址博弈、平滑博弈

感觉书上讲得不错，我决定直接对着书讲。



# 网络代价分摊博弈

接下来介绍的这个网络模型是具有正外部性的博弈，也就是说智能体愿意和别的智能体分享同一条边。网络代价分摊博弈 (network cost-sharing game) 产生在图  $G = (V, E)$  上。此图可以是有向图或无向图，并且每条边  $e \in E$  都有一个固定的代价  $\gamma_e \geq 0$ 。图中共有  $k$  个智能体，每个智能体  $i$  都有一个源点  $o_i \in V$  和一个终点  $d_i \in V$ ， $i$  的策略集为图中所有  $o_i - d_i$  路径。此博弈的局势是一个路径向量  $P = (P_1, \dots, P_k)$ 。

我们把  $\gamma_e$  看作构建边  $e$  的代价，这个代价与使用此边的智能体的数目**无关**。智能体的代价则由路径中的每条边来定义。如果多个智能体在的路径中都有边  $e$ ，那么他们一起分担此边的固定总代价  $\gamma_e$  (假设为平均分摊)。智能体  $i$  在局势  $P$  下的代价  $C_i(P)$  为：

$$C_i(P) = \sum_{e \in P_i} \frac{\gamma_e}{f_e} \quad (7)$$

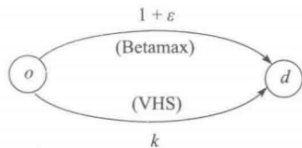
其中  $f_e = |\{j: e \in P_j\}|$  表示其路径中包含边  $e$  的智能体的数目。目标函数为最小化所生成网络的总代价：

$$\text{cost}(P) = \sum_{e \in E, f_e \geq 1} \gamma_e = \sum_{i=1}^k C_i(P) \quad (8)$$

# 网络代价分摊博弈

我们首先通过一个直观的例子来理解这个博弈。有  $k$  个智能体，他们共享一个源点  $o$  和终点  $d$ 。可以将这个例子看作在两个相互竞争的产品技术中做选择。例如，在 20 世纪 80 年代，有两种新兴技术用于电影租赁。技术极客认为 Betamax 更好，他们对应于图中的代价较小的边。VHS 是另一项技术，其在早期占据了大部分的市场份额。由于消费者在选择技术时会与大多数其他人的选择一致，Betamax 虽技术更好，但这并不能弥补市场占有率上的劣势，因此最终慢慢被市场淘汰了。

在图中，所有智能体都选择上面那条边是最优局势，此时总的代价为  $1 + \epsilon$ 。这个策略也是一个 PNE。但遗憾的是，网络中还存在第二个 PNE，即所有智能体都选择下面那条边。由于代价  $k$  被平均分摊，每个智能体的支付是 1。如果一个智能体单方面把策略改变到上面那条边，他会支付那条边上的全部代价  $1 + \epsilon$ ，这是一个更高的代价。这个例子表明，网络代价分摊博弈中的 POA 可以和智能体的数目  $k$  一样高。

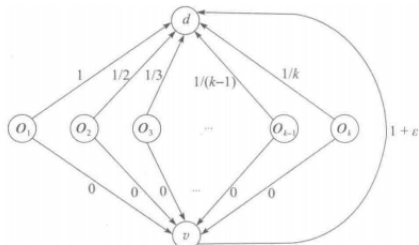


# 退出博弈

接下来我们来看另一个例子，在这个例子中 POA 仍然是非常数级别的。

$k$  个智能体有  $k$  个不同的源点  $o_1, \dots, o_k$ ，但是有相同的终点  $d$ 。他们可以选择首先在集结地  $v$  汇合，然后再共同去  $d$ ，总的代价为  $1 + \epsilon$ 。每个智能体也可以“选择退出”，即单独地直接沿着  $o_i - d$  这条路走。当智能体  $i$  选择退出策略时，他的代价为  $1/i$ 。

最优局势一目了然：如果所有智能体都途径集结点，那么总的代价为  $1 + \epsilon$ ，但这不是 PNE：智能体  $k$  可以通过选择退出策略来略微降低他的支付，这对他来说是一个占优策略。鉴于智能体  $k$  在一个 PNE 中没有使用集结点，智能体  $k-1$  同样也不会使用。重复这个过程，发现每个人都会选择退出，这是唯一的均衡，总代价为调和级数，我们用  $\mathcal{H}_k$  表示。当  $\epsilon \rightarrow 0$  时，POA 趋于  $\mathcal{H}_k$ 。



# 最好情况

上面的例子让我们很挫败，因为即使是只有唯一的均衡，其代价也可能是最优局势的  $\mathcal{H}_k$  倍。不过好在这就是上界了。

## 定理 5.1 (稳定的代价)

在任意含有  $k$  个智能体的网络代价分摊博弈中，总存在一个 PNE，其代价最多为最优局势的  $\mathcal{H}_k$  倍。

所谓稳定的代价就是 POA 的一个“乐观”版本，为**最优均衡**的代价除以最优局势的代价，也就是研究最好情况下，均衡和最优局势的比值。

# 强纳什均衡

我们仍然可以针对最差均衡下的均衡来讨论 POA，只不过这时要把均衡的概念收紧——在强纳什均衡条件下讨论。

回忆一下 VHS 还是 Betamax 这个例子。由于智能体进行单方面策略改变会支付上面边的所有代价  $1 + \epsilon$ ，因此高代价的局势是一个 PNE。如果有两个智能体联合改变到上面的边上会如何？每一个改变的智能体仅仅支付约  $\frac{1}{2}$ 。因此这对两人来说都是一个有利的改变。我们可以得出结论：当允许智能体组成联盟来偏离均衡时，高代价的 PNE 并不能维持。换句话说，在强纳什均衡下，智能体们有了合作和联合的机会，就更有可能靠近最优结果，不过一般来说它也不是最优。在退出博弈中，所有人都选择退出不仅是 PNE 还是一个强纳什均衡。这也就说明，就算是强纳什均衡，POA 也能到达  $\mathcal{H}_k$  的级别。不过我们可以证明不存在比这更差的情况了。

## 定理 5.2

在任意包含  $k$  个智能体的网络代价分摊博弈中，每个强纳什均衡的代价最多为最优局势的  $\mathcal{H}_k$  倍。

# 势博弈

我们提出这样一个问题：这些策略性的智能体能不能自动达到均衡状态？如果可以，哪种学习算法能够**快速收敛**到一个均衡的结果？我们提出动力学系统，它描述当智能体不在均衡状态的行为模式。

首先提出一个最简单的所谓“最优反应动力学”，它和我们以前学习过的局部搜索很类似。

## 最优反应动力学

只要目前的局势  $s$  不是一个 PNE：

- 任意选择一个智能体  $i$  以及一个对  $i$  有利的策略改变  $s'_i$ ，然后更新局势为  $(s'_i, s_{-i})$ 。

上述最优反应动力学可以抽象成在一张有限图上行走，图中每个顶点表示一组策略组合，每条出边刻画了一个有利的策略改变，而那些没有出边的顶点就是 PNE。如果一个博弈不存在 PNE，那么最优反应动力学会一直循环。并且，就算存在 PNE，最优反应动力学也可能一直循环。

但是，对于势博弈来说，由于总策略是有限的，每次改变都会使势函数下降，并且势函数是有最小值的，那就说明这个过程最终经过有限步停止，能够收敛到一个 PNE。



# 近似 PNE 与快速收敛

我们很难研究一个能快速收敛到 PNE 的动力学，但是我们能够期望快速收敛到一个近似 PNE.

## 定义 ( $\varepsilon$ -纯策略纳什均衡)

对  $\varepsilon \in [0, 1]$ , 给定代价最小化博弈的局势  $s$  中, 如果对每一个智能体  $i$  和策略  $s'_i \in S_i$ , 都有

$$C_i(s'_i, s_{-i}) \geq (1 - \varepsilon) \cdot C_i(s) \quad (9)$$

那么称局势  $s$  为一个  $\varepsilon$ -纯策略纳什均衡.

## 定义 ( $\varepsilon$ -行动)

智能体  $i$  存在一个行动  $s'_i$  使得  $C_i(s'_i, s_{-i}) < (1 - \varepsilon) \cdot C_i(s)$ , 则称  $s'_i$  为智能体  $i$  在局势  $s$  中的一个  $\varepsilon$ -行动.

# 近似 PNE 与快速收敛

## $\varepsilon$ -最优反应动力学 (最大回报)

只要当前局势  $s$  不是一个  $\varepsilon$ -PNE:

- 在所有拥有  $\varepsilon$ -行动的智能体中, 用  $i$  表示能获得最大代价下降

$$C_i(s) - \min_{\hat{s}_i \in S_i} C_i(\hat{s}_i, s_{-i})$$

的智能体.

- 并用  $s'_i$  表示  $i$  针对  $s_{-i}$  的一个最优反应.
- 更新局势为  $(s'_i, s_{-i})$ .

# 近似 PNE 与快速收敛

我们有以下结论：

## 定理

考虑单元自私路由博弈，假设：

- ① 所有智能体都有相同的源点和终点。
- ② 对  $\alpha \geq 1$ ，每条边  $e$  的代价函数  $c_e$  满足  $\alpha$ -限界跳跃条件，即对每一条边  $e$  和正整数  $x$  有  $c_e(x+1) \in [c_e(x), \alpha \cdot c_e(x)]$ 。

那么， $\varepsilon$ -最优反应动力学的最大回报变种最多在

$$\frac{k\alpha}{\varepsilon} \ln \frac{\Phi(s^0)}{\Phi_{\min}}$$

轮后收敛到一个  $\varepsilon$ -PNE，其中  $s^0$  是初始局势， $\Phi_{\min} = \min_s \Phi(s)$ 。

# 平滑势博弈

我们研究“快速收敛”的另一概念：快速到达一个局势，此局势虽然不一定收敛到一个近似 PNE，但是此局势下目标函数的值几乎与达到近似 PNE 时一样。换句话说，就是“近似”近似 PNE（在目标函数值的意义上）。我们通过弱化快速收敛的概念，得到一个适用于所有满足  $(\lambda, \mu)$ -平滑条件的博弈的势函数，这就包括了前面所讲的所有单元自私路由博弈和选址博弈。

我们使用最优反应动力学的变种：

## 最优反应动力学（最大回报）

只要当前输出  $s$  不是一个 PNE：

在所有存在有利策略改变的智能体中，用  $i$  表示能够获得最大代价下降

$$C_i(s) - \min_{\hat{s}_i \in S_i} C_i(\hat{s}_i, s_{-i}) \quad (10)$$

的智能体。

用  $s'_i$  表示智能体  $i$  对  $s_{-i}$  的一个最优反应。

更新局势为  $(s'_i, s_{-i})$ 。

# 平滑势博弈

我们在代价最小化博弈中陈述过这个定理，类似的结论同样适用于  $(\lambda, \mu)$ -平滑收益最大化博弈。

## 定理

考虑一个  $(\lambda, \mu)$ -平滑代价最小化博弈 ( $\mu \leq 1$ )，此博弈有一个正的势函数  $\Phi$ ；对于任意局势  $s$ ，有  $\Phi(s) \leq \text{cost}(s)$ 。用  $s^0, \dots, s^T$  表示在最大回报最佳动力学里生成的一系列的局势，用  $s^*$  表示最小代价局势，给定参数  $\eta \in (0, 1)$ 。那么除了最多  $\frac{k}{1-\mu} \ln \frac{\Phi(s^0)}{\Phi_{\min}}$  个局势外，剩余局势  $s_t$  都满足

$$\text{cost}(s_t) \leq \left( \frac{\lambda}{(1-\mu)(1-\eta)} \right) \cdot \text{cost}(s^*) \quad (11)$$

其中  $\Phi_{\min} = \min_s \Phi(s)$ ， $k$  是智能体的数目。

上述定理的大致证明思路是，如果现在是一个不满足 (11) 式的坏状态，那么下一轮里最大回报动力学能够显著降低势函数。根据这一点就能导出所有坏状态数目的上界。

# 总结

- 自私路由博弈、选址博弈 → 平滑博弈
- POA——均衡和最优的关系
- 最优反应动力学——怎么快速收敛到均衡/近似均衡/近似近似均衡

