



# 高级人工智能

沈华伟

[shenhuawei@ict.ac.cn](mailto:shenhuawei@ict.ac.cn)

中国科学院计算技术研究所

2020.12.14

# 课程内容

## ■ 博弈

- 基本概念
- 纳什均衡
- 机制设计

## ■ 两个经济学的应用

- 拍卖
- 讨价

# 一般意义的博弈

## ■ 日常生活中随处可见“博弈”

- 赌博
- 棋类游戏：象棋、围棋
- 田径运动
- 篮球运动
- .....

## ■ 如何在这样的“博弈”中获胜呢？

- 博弈一般包含运气、技术和策略
- 策略是为了获胜所需要的一种智力技巧
  - 是对于如何最好地利用身体、工具等技巧的一种算计
    - 例如：篮球的挡拆战术、犯规战术

# 策略博弈

- 策略本质上涉及与他人的相互影响
  - 其他人在同一时间、对同一情形也在进行类似的思考
- 博弈论就是分析这样的交互式决策，是关于相互作用情况下的理性行为的科学
- 理性行为
  - 明白自己的目的和偏好，同时了解自己行动的限制和约束，以精心策划的方式选择自己的最佳行动
  - 博弈论对理性行为赋予的新含义：与其他同样具有理性的决策者进行相互作用
- 在博弈中真的总能获胜吗？有必胜策略吗？
  - 注意：对手和你一样聪明的

# 博弈的类型

## ■ 静态博弈 vs. 动态博弈

- 静态博弈：所有局中人同时进行策略选择，譬如剪刀-石头-布
- 动态博弈：局中人按照顺序进行策略选择，譬如下棋

## ■ 竞争博弈 vs. 合作博弈

- 竞争博弈：炒股
- 合作博弈：结盟

## ■ 完全信息博弈 vs. 不完全信息博弈

- 完全信息博弈：每个局中人对于所有局中人的策略及其效用充分了解；反之，称之为不完全信息博弈

# 博弈案例

## ■ 二人分配问题

- 两个人分一个东西（譬如分蛋糕），设计什么样的机制以保证尽可能等分

- 一个人切，另一个人选



# 博弈案例

## ■ 田忌赛马

- （田）忌数与齐诸公子驰逐重射。孙子见其马足不甚相远，马有上、中、下辈。于是孙子谓田忌曰：“君弟重射，臣能令君胜。”田忌信然之，与王及诸公子逐射千金。及临质，孙子曰：“**今以君之下驷与彼上驷，取君上驷与彼中驷，取君中驷与彼下驷。**”既驰三辈毕，而田忌一不胜而再胜，卒得王千金。



	第一场	第二场	第三场	获胜方
齐王	上	中	下	
田忌 1	上	中	下	齐王
田忌 2	上	下	中	齐王
田忌 3	中	上	下	齐王
田忌 4	中	下	上	齐王
田忌 5	下	上	中	田忌
田忌 6	下	中	上	齐王

# 博弈的要素

## ■ 局中人（Player）

- 在博弈中有权决定自己行动方案的博弈参加者
- 局中人不一定是具体的人
  - 如球队、军队、企业
- 博弈中利益完全一致的参与者只能看成一个局中人
  - 如桥牌中的南北方和东西方

## ■ 重要假设：局中人是自私的理性人

- 不存在侥幸心理
- 不可能利用其它局中人的失误来扩大自己的利益
- 以最大化个人利益为目的



# 博弈的要素

## ■ 策略集合(Strategy Set)

- 策略：博弈中可供局中人选择的行动方案
- 参加博弈的局中人 $i$ 的策略集合记为 $A_i$
- 所有局中人的策略形成的策略组，称为局势，记作 $S$
- 多人博弈中假定有 $n$ 个局中人，每个局中人从自己的策略集合中选择一个策略 $s_i$ ， $s_i \in A_i$ ，这样就形成了一个局势 $S = \{s_1, s_1, \dots, s_n\}$

## ■ 田忌赛马中田忌的策略集合

- {上中下、上下中、中上下、中下上、下上中、下中上}

# 博弈的要素

- 效用函数(Payoff)
  - 通常用 $U$ 来表示
  - 对每个参与博弈的局中人，都有一个相应的效用函数
  - 效用函数在静态博弈中一般是局势的函数
  - 在动态博弈中效用函数可能是局势的函数，也可能还有其它因素，比如时间
  - 每个局中人的目的都是最大化自己的效用

# 囚徒困境

- 局中人
  - 两个囚徒
- 策略
  - 抗拒
  - 坦白
- 效用函数矩阵

		囚徒B	
		抗拒	坦白
囚徒A	抗拒	-1,-1	-10,0
	坦白	0,-10	-3,-3

# 性别之战

- 局中人
  - 夫妻双方
- 策略
  - 看韩剧、看体育
- 效用函数矩阵

		妻子	
		韩剧	体育
丈夫	韩剧	1,2	0,0
	体育	0,0	2,1

# 剪刀-石头-布(Rock-paper-scissors)

- 局中人
  - 两个玩家
- 策略
  - 剪刀、石头、布
- 效用函数矩阵

		玩家二		
		剪刀	石头	布
玩家一	剪刀	0,0	-1,1	1,-1
	石头	1,-1	0,0	-1,1
	布	-1,1	1,-1	0,0

# 纳什均衡

## ■ 最佳应对

- 假设 $s$ 是局中人1的一个选择策略， $t$ 是局中人2的一个选择策略； $U_1(s, t)$ 是局中人1从这组决策中获得的收益， $U_2(s, t)$ 是局中人2从这组决策中获得的收益
- 针对局中人2的策略 $t$ ，若局中人1用策略 $s$ 产生的收益大于或等于其任何其他策略，则称策略 $s$ 是局中人1对局中人2的策略 $t$ 的**最佳应对**
  - $U_1(s, t) \geq U_1(s', t)$ ， $s'$ 是局中人1除 $s$ 外的其它策略
- 如果一个局中人的某个策略对其它局中人的任何策略都是最佳应对，那么这个策略就是该局中人的占优策略

# 纳什均衡

## ■ 纳什均衡

- 定义：如果一个局势下，每个局中人的策略都是相对其他局中人**当前策略**的最佳应对，则称该局势是一个纳什均衡

## ■ 纳什均衡就是博弈的一个均衡解

## ■ 是一个僵局

- 即给定其他人不动，没有人有动的积极性
- 谁动谁吃亏

# 纳什均衡

## ■ 例子：囚徒困境

### □ 纳什均衡：双方都坦白

- 一方保持策略不变（坦白），另一方如果改变策略（抗拒），其效用会降低（从-3变成-10）

		囚徒B	
		抗拒	坦白
囚徒A	抗拒	-1,-1	-10,0
	坦白	0,-10	-3,-3



# 纳什均衡

## ■ 例子：性别之战

### □ 纳什均衡1：夫妻双方都同意看韩剧

- 妻子保持策略不变（看韩剧），丈夫如果改变策略（看体育），其效用会降低（从1变成0）
- 丈夫保持策略不变（看韩剧），妻子如果改变策略（看体育），其效用会降低（从2变成0）

### □ 纳什均衡2：夫妻双方都同意看体育

		妻子	
		韩剧	体育
丈夫	韩剧	1,2	0,0
	体育	0,0	2,1

# 纳什均衡

- 例子：剪刀-石头-布
  - 不存在纯策略的纳什均衡

		玩家二		
玩家一		剪刀	石头	布
	剪刀	0,0	-1,1	1,-1
	石头	1,-1	0,0	-1,1
	布	-1,1	1,-1	0,0

# 混合策略纳什均衡

## ■ 混合策略

- 每个局中人以某个概率分布在其策略集合中选择策略

## ■ 混合策略下的纳什均衡

- 定义和纯策略纳什均衡一致：基于最佳应对定义
- 必要条件：给定其他局中人的策略选择概率分布的情况下，当前局中人选择任意一个（纯）策略获得的期望效用相等

# 混合策略纳什均衡

## ■ 例子：剪刀-石头-布

- 玩家一的策略选择分布记为  $p = \{p_1, p_2, 1 - p_1 - p_2\}$ ，玩家二的策略选择分布记为  $q = \{q_1, q_2, 1 - q_1 - q_2\}$
- 假设玩家一的策略分布不变，玩家二策略选择的效用为
  - 剪刀：  $0 * p_1 + (-1) * p_2 + 1 * (1 - p_1 - p_2) = 1 - p_1 - 2p_2$
  - 石头：  $1 * p_1 + 0 * p_2 + (-1) * (1 - p_1 - p_2) = 2p_1 + p_2 - 1$
  - 布：  $(-1) * p_1 + 1 * p_2 + 0 * (1 - p_1 - p_2) = p_2 - p_1$
- 令玩家二的各个策略的效用相等，得到  $p_1 = p_2 = \frac{1}{3}$
- 同理可得  $q_1 = q_2 = \frac{1}{3}$

## ■ 剪刀-石头-布的混合纳什均衡态

- 每个玩家各以1/3的概率选择剪刀、石头和布
- 期望收益为0

		玩家二		
		剪刀	石头	布
玩家一	剪刀	0,0	-1,1	1,-1
	石头	1,-1	0,0	-1,1
	布	-1,1	1,-1	0,0

# 纳什定理

- 任何有限博弈都至少存在一个纳什均衡
  - 不一定是纯策略纳什均衡，例如剪刀-石头-布
- 寻找博弈的纳什均衡是困难的
  - 至少从算法角度来讲是这样

# 社会最优

## ■ 帕累托最优

- 以意大利经济学家维尔弗雷多·帕累托的名字命名
- 对于一组策略选择（局势），若不存在其他策略选择使所有参与者得到至少和目前一样高的回报，且至少一个参与者会得到严格较高的回报，则这组策略选择为帕累托最优

## ■ 社会最优

- 使参与者的回报之和最大的策略选择（局势）
- 社会最优的结果一定也是帕累托最优的结果
- 帕累托最优不一定是社会最优

# 社会最优示例

## ■ 囚徒困境案例

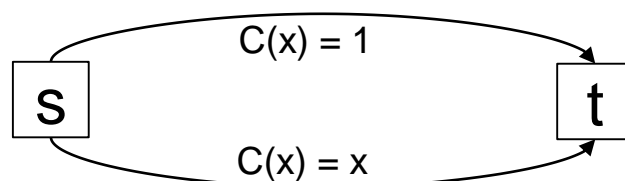
		囚徒B	
		抗拒	坦白
囚徒A	抗拒	-1,-1	-10,0
	坦白	0,-10	-3,-3

帕累托最优的决策组合一共有3个，分别是（坦白，抗拒），（抗拒，坦白）和（抗拒，抗拒），纳什均衡策略组合（坦白，坦白）不是帕累托最优，社会最优策略组合是（抗拒，抗拒）

# 社会最优示例

## ■ 案例

- 从源点s到目标点t有两条通路，第一条的代价恒为1，第二条的代价和选择该路径的人数呈正比



- 纯策略

- 所有人选择第一条路径：总代价为1
- 所有人选择第二条路径：总代价为1

- 混合策略

- 以概率 $x$ 选择第一条路，以概率 $1 - x$ 选择第二条
- 期望代价： $x + (1 - x)^2 = x^2 - x + 1$

- 最优策略是： $x = 1/2$



# 机制设计

- 如何设计一个博弈，使其达到预期结果（譬如，实现社会最优）？
  - 设计游戏规则
    - 次价密封报价拍卖
  - 设计效用函数
    - 诉讼费用、股票印花税
  - 确定哪些信息是私有信息（不完全信息博弈）
    - 密封报价拍卖、公开拍卖
    - 多人协同给图片打标签
  - 静态博弈，还是动态博弈

# 机制设计

## ■ 案例1

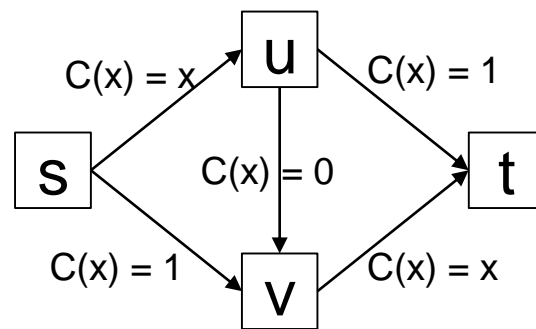
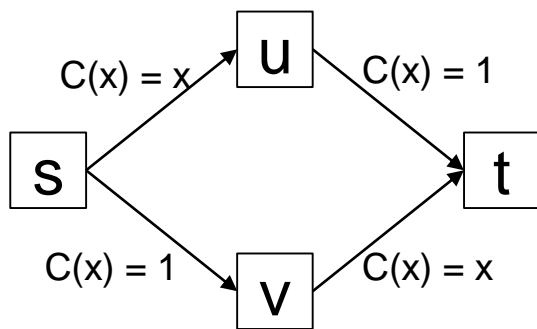
- 二人博弈游戏：轮流从1-4中选择一个数，哪个人选择一个数字后使得之前的所有数字之和等于50则获胜
- 后手必胜策略
  - 无论对方选什么，都凑够5的倍数
  - 后手报数后的数字总和为5, 10, 15, ..., 45, 50

## ■ 案例2

- 二人分蛋糕

# 机制设计的失败案例

## ■ 案例3



## ■ 左图的情形

□ 期望代价是：  $1/2 + 1$

## ■ 在 $u$ 和 $v$ 之间修一条代价为0的高速路，会提高社会效用吗？

□ 社会最优解：以概率  $1/2$  走路径  $s \rightarrow u \rightarrow t$ ，以概率  $1/2$  走路径  $s \rightarrow v \rightarrow t$ ，以概率  $0$  走路径  $s \rightarrow u \rightarrow v \rightarrow t$ ，此时的期望代价为  $1/2 + 1$

□ 纳什均衡解：以概率  $1$  走路径  $s \rightarrow u \rightarrow v \rightarrow t$ ，此时的期望代价为  $2$

课间休息

# 课程内容

## ■ 博弈

- 基本概念
- 纳什均衡
- 机制设计

## ■ 应用案例

- 拍卖
- 讨价

# 经济市场

- 解决稀有资源的分配问题
- 一般市场
  - 多个卖家、多个买家
- 讨价（Bargaining）
  - 多个卖家、一个买家
- 拍卖（Auction）
  - 一个卖家、多个买家

# 拍卖

## ■ 拍卖活动

- 一个卖家向一群买家拍卖一件商品的活动
- 拍卖的基本假设
  - 每个竞争者对被拍卖的商品有各自的估值
  - 这个估值是竞拍者对商品实际所值的估计
  - 如果商品售价不高于这个估值，竞拍者会购买，否则不会购买

# 拍卖

## ■ 拍卖类型

### □ 增价拍卖，又称英式拍卖

- 拍卖者逐渐提高售价，竞拍者不断退出，直到只剩一位竞拍者，该竞拍者以最后的报价赢得商品

### □ 减价拍卖，又称荷式拍卖

- 拍卖者逐渐降低售价，直到有竞拍者出价购买

### □ 首价密封报价拍卖

- 竞拍者同时向拍卖者提交密封报价，拍卖者同时打开这些报价，出价最高的竞拍者以其出价购买该商品

### □ 次价密封报价拍卖

- 竞拍者同时向拍卖者提交密封报价，出价最高的竞拍者赢得商品但以第二高出价购买该商品

### □ 双方出价

- 股票市场



# 首价密封报价拍卖

## ■ 纳什均衡

- 每个竞拍者的报价低于其对商品的估价

## ■ 解读

- 共有 $n$ 个竞拍者，竞拍者 $i$ 的估价记为 $v_i$ ，报价记为 $b_i$ ，其他竞拍者的估价服从 $[a, b]$ 区间上的均匀分布，且诚实出价
- $b_i < a$ 时，竞标失败，收益为0
- 竞拍者 $i$ 赢得竞拍的概率为  $\left(\frac{b_i - a}{b - a}\right)^{n-1}$
- 竞拍者的期望收益是  $f(b_i) = (v_i - b_i) \left(\frac{b_i - a}{b - a}\right)^{n-1}$

# 首价密封报价拍卖

## ■ 解读（续）

### □ 期望收益

$$f(b_i) = (v_i - b_i) \left( \frac{b_i - a}{b - a} \right)^{n-1}$$

### □ 期望收益关于报价 $b_i$ 的梯度为

$$\begin{aligned} f'(b_i) &= - \left( \frac{b_i - a}{b - a} \right)^{n-1} + (n-1)(v_i - b_i) \left( \frac{b_i - a}{b - a} \right)^{n-2} \frac{1}{b - a} \\ &= \left( \frac{b_i - a}{b - a} \right)^{n-2} \left( \frac{-nb_i + a + (n-1)v_i}{b - a} \right) \end{aligned}$$

### □ 最优报价为

$$b_i^* = \frac{a + (n-1)v_i}{n}$$

- ✓ 最优报价低于估价
- ✓ 竞拍者越多，报价越接近于估价

# 次价密封报价拍卖

## ■ 纳什均衡

- 每个竞拍者会倾向于采用其对商品的估价进行报价

## ■ 解读

- 给定一个竞拍者，其估价记为 $v$ ，报价记为 $b$ ，其他竞拍者的最高报价记为 $b^*$
- 理性行为假设下，报价不会高于估价，即 $b \leq v$
- 此时，根据 $b^*$ 的取值有三种情形
  - $b^* > v$ ：收益为0；将报价从 $b$ 提高到 $v$ ，收益不变
  - $b^* < b$ ：收益为 $v - b^*$ ；将报价从 $b$ 提高到 $v$ ，收益不变
  - $b \leq b^* \leq v$ ：收益为0；将报价从 $b$ 提高到 $v$ ，收益变为 $v - b^*$

# 课程内容

## ■ 博弈

- 基本概念
- 纳什均衡
- 机制设计

## ■ 应用案例

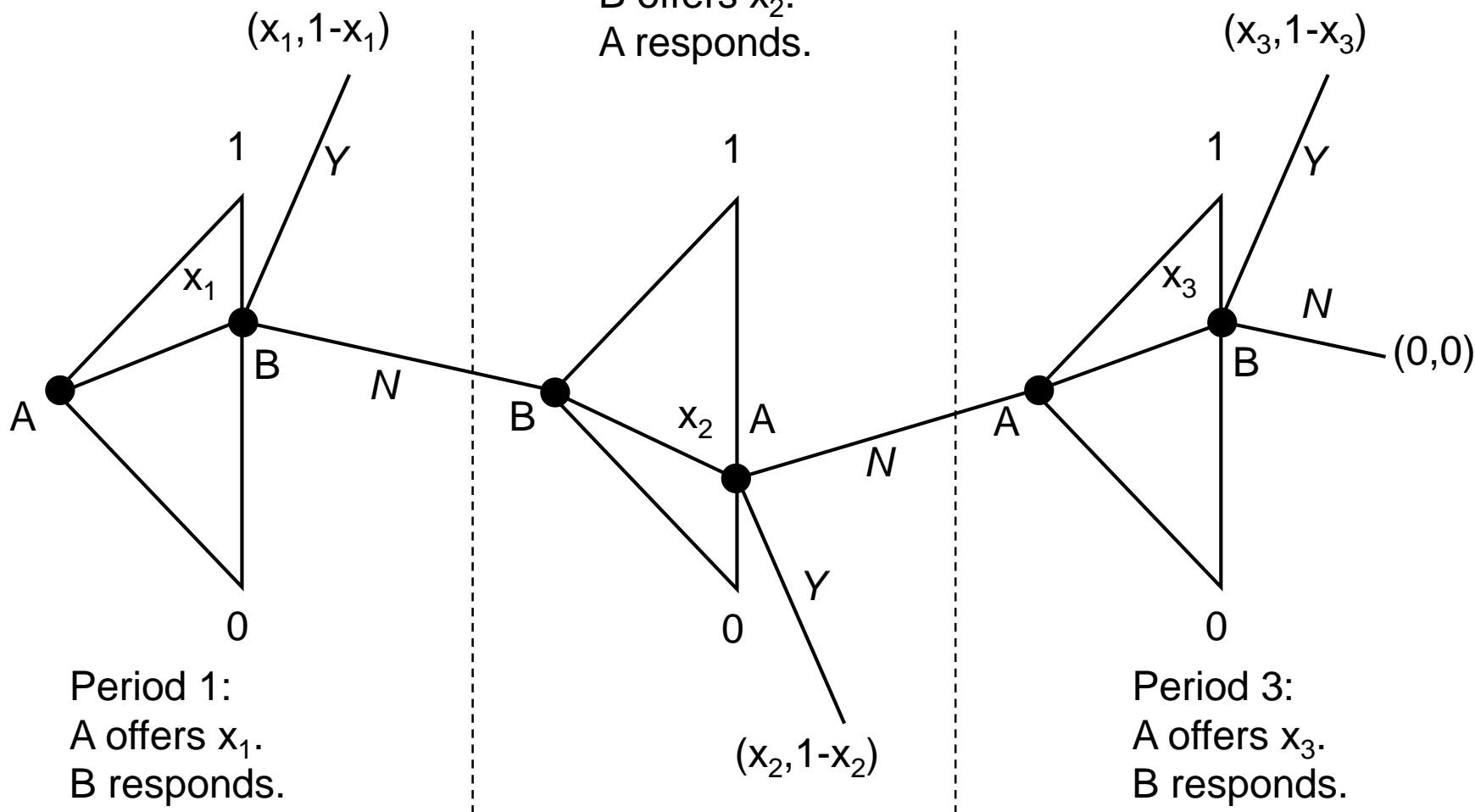
- 拍卖
- 讨价

# 讨价

- 卖家和买家之间的博弈
- 讨价的对象是双方对商品估价之差
  - 假设所有因素都体现在估价中
    - 时间、情感、眼缘等
  - 例子：
    - 衣服进价80，标价200
    - 卖家对衣服的估价在80和200之间，譬如120
    - 买家的估价假如为160
    - 讨价的对象是双方的估价之差，即 $160 - 120 = 40$
- 后续的讨论中，将讨价对象视为整体1
  - 卖家的估价为0，买家的估价为1

# 讨价

## ■ 讨价的博弈过程



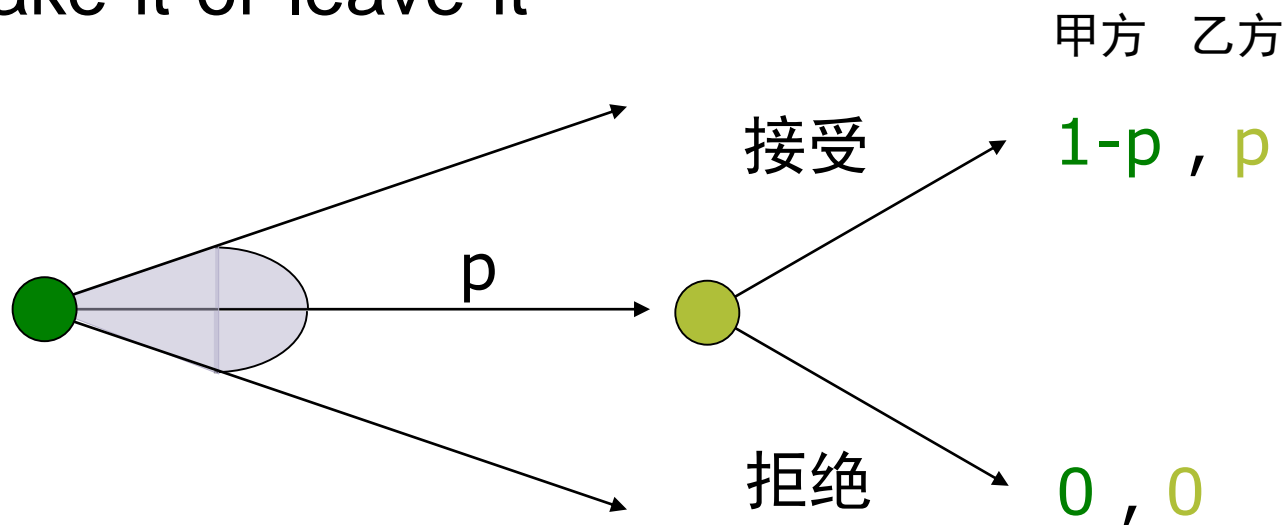
# 讨价

## ■ 场景1

- Take-it-or-leave-it: 无商谈余地
- 一方报价，另一方要么接受报价达成交易，要么交易失败
  - 两个人商量吃蛋糕，一方提出切分比例，另一方如果不同意，双方就都不吃
  - 美国参议院：民主党提出增加财政预算到某个值，共和党要么同意，要么拒绝（但不能提新的方案）
- 通过[回滚\(rollback\)](#)求解纳什均衡

# 讨价

## ■ Take-it-or-leave-it



## ■ 过程

- 阶段1：甲方提出，按照 $1-p$ 和 $p$ 的比例进行分配
- 阶段2：只要 $p$ 大于0，乙方则会接受

甲方（分配方案提出者）得到几乎所有收益



# 讨价

## ■ 常见的讨价情形

- Take-it-or-counteroffer: 要么接受，要么还价

## ■ 过程

- 第一阶段：甲方报价：  $1-p, p$
- 第二阶段：乙方要么接受报价，要么还价  $\delta * (1 - q), \delta * q$
- 第三阶段：甲方决定要么接受乙方的还价，要么交易失败

## ■ 约束条件

- 时间成本：  $\delta$ 刻画可用于分配的总收益随时间衰减 ( $0 \leq \delta \leq 1$ )

## ■ 例子：NBA劳工谈判，分配一个会融化的蛋糕

# 讨价

## ■ Take-it-or-counteroffer过程推演

- 第一阶段之后等同于take-it-or-leave-it讨价
- 假如第一阶段乙方没有接受甲方的报价，那么在接下来的take-it-or-leave-it过程中，甲方的收益将趋近于0
- 因此，甲方在第一阶段报价时，分配给乙方的收益不少于乙方拒绝报价后所得到的收益

$$p \geq \delta * (1 - q) \approx \delta$$

甲方第一轮报价中，乙方获得的收益

甲方第一轮报价被拒绝后，乙方报价时将获得的收益

# 讨价

## ■ Take-it-or-counteroffer过程推演的启示

- 在时间成本约束下，甲乙双方会尽可能在第一轮达成交易，以使共同分割的收益总和最大
- 甲方在第一轮报价时，需根据时间成本来决定报价
- 乙方获得收益依赖于对时间成本的容忍度
- 最终的分配比例是

甲方： $1 - \delta$

乙方： $\delta$

蛋糕融化得越慢，乙方分的越多

# 讨价

- 先发优势，还是后发制人？
- 当时间成本较高（即 $\delta$ 较小）时，甲方有先发优势
  - 例如：炎热的夏天，蛋糕融化得快
- 当时间成本较低（即 $\delta$ 较大）时，乙方可后发制人
  - 例如：寒冷的冬天，蛋糕融化得慢

启示：博弈规则决定最终的结果

# 讨价

## ■ 打官司

- 原告诉被告，要求赔偿100万
- 诉讼费原告和被告各支付10万

## ■ 情形1

- 双方各自认为自己胜诉的概率为1/2
- 开启诉讼
  - 原告收益：50万-10万=40万；被告收益：-50万-10万=-60万
- 可以达成庭外和解：譬如被告支付50万给原告
  - 原告能接受的最低价是：40万
  - 被告能提供的最高价是：60万
  - 讨价分配的“蛋糕”大小为20万

# 讨价

## ■ 情形2

- 双方各自认为自己胜诉的概率为 $3/4$

- 开启诉讼

  - 原告预期收益： $75\text{万}-10\text{万}=65\text{万}$ ；

  - 被告预期收益： $-25\text{万}-10\text{万}=-35\text{万}$

- 无法达成和解

  - 原告能接受的最低价是：65万

  - 被告能提供的最高价是：35万

- 讨价分配的蛋糕大小是： $-30\text{万}$

- 假如诉讼费是30万呢？

# 讨价

## ■ 小结

- 博弈规则决定博弈结果
- 各自的“底牌”是对方报价的依据
- 讨价的蛋糕大小由双方的底牌决定

# 课后作业

## ■ 海盗分金币

- 问题描述：有5个海盗抢到了100个金币，经过激烈争论，就如何分配达成以下协议：
  - 抽签决定每个人提分配方案的顺序
  - 抽到1号签的海盗首先提出自己的分配方案，然后所有人表决（包括方案提出者），当且仅当**半数或超过半数**以上的人同意的时候，才按照他提出的方案执行，否则他会被扔进海里
  - 1号海盗的方案如果未被通过，那么2号海盗提自己的方案，规则和上述一样，直到某个方案通过
- 给出最终的分配方案
  - 提示：从后往前回滚



下课