博弈论基础

姜少峰

北京大学前沿计算研究中心
Center on Frontiers of Computing Studies, Peking University

博弈Game

- 研究如何做决策使收益最大化
- 决策结果不光取决于自己的选择，也取决于参与互动的人的选择
  - 脱胎于游戏；衍生问题：商品定价，拍卖竞价等

何为博弈：引例

- 距离ddl很近了，A有两项需要准备的任务：考试和报告
- 时间有限，只能二选一
  - 如果准备考试，那么预计能得92分，但如果不准备则只有80分
  - 报告是跟B的2人小组作业：如果A和B都充分准备了，则可以得100；只有一个人准备了，得92；否则，得84
- 假设情况对于B也完全一样，且无法就如何选择进行沟通
  - 即自己独立决策时，也假定对方在独立决策
- 目标：A和B各自追求考试和报告的平均得分最大化

- 如果准备考试，那么预计能得92分，但如果不准备则只有80分
- 报告是跟B的2人小组作业：如果A和B都充分准备了，则可以得100；只有一个人准备了，得92；否则，得84

- 如果都准备报告，那么报告都得100，考试都得80，平均分都是90
- 如果都准备考试，那么报告都得84，考试都92，平均分都是88
- 如果A准备考试，B准备报告，
  - A报告92考试92，平均92；B报告92考试80，平均86

	B报告	B考试
A报告	90，90	86，92
A考试	92，86	88，88

左侧A得分，右侧B

- 如何决策？

博弈的基本要素

- 参与人player：A和B
- 策略strategy：准备考试或是报告
- 收益payoff：每个人每个策略的得分

基本假设

- 参与者对博弈中事物的在意程度完全体现在收益上
  - 博弈论框架本身不排除利他主义，这可以在收益上体现
- 参与者都有充分信息：知道所有人的策略集合及收益
  - 不完全信息的博弈？1994 Nobel Prize Economics for (John Harsanyi, 1967)
- 参与者选择自己收益最大化的策略，并假定知道别人（收益最大化的）策略
  - 称为“理性化”模型：个体最在意的事物体现在收益上，并选择最优策略
  - 非理性：经验不足？出错？

分析考试-报告博弈

- 从A的视角分析（B对称）
- 假设B准备考试：
  - A如果准备考试，收益88；准备报告，收益86
- 假设B准备报告：
  - A如果准备考试，收益92；准备报告，收益90
- 综上，不论B采用什么策略，A选“考试”总是占优

	B报告	B考试
A报告	90，90	86，92
A考试	92，86	88，88

左侧A得分，右侧B

严格占优策略

如果一个固定的策略无论别人选什么都占优，则称之为严格占优策略
(strictly dominant strategy)

- 在考试-报告博弈中，A和B的严格占优策略是“准备考试”
- 因此可以预期A和B都准备考试，得到88分平均成绩

	B报告	B考试
A报告	90，90	86，92
A考试	92，86	88，88

左侧A得分，右侧B

有更好的策略吗？

- 都准备考试都得88分
- 如果商量好了准备报告呢？这样都可以平均得90？
  - 在理性假设下不可能：A准备报告，但B如果准备考试会得92，B会“背信弃义”
- 本质问题是大家“只关心自己收益最大化”这条假设
  - 可能的改进方法：两个人关系好，“利他”地设置收益
  - 改变模型：考虑“背信弃义”的社会影响

	B报告	B考试
A报告	90，90	86，92
A考试	92，86	88，88

左侧A得分，右侧B

相关问题：囚徒困境

- 甲乙涉嫌抢劫被抓，但警察暂时没有确凿证据；两人还曾拒捕
- 甲乙分别关押，下面是警察对他们每个人的offer：
  - 如果你坦白而另一人抵赖，则你可以被释放，另一人承担全部罪行，判十年
  - 如果你们都坦白，那么“坦白从宽”，都坐实抢劫但是都只判四年
  - 如果你们都不坦白，那么警方将以轻罪拒捕起诉，判一年
- 两个囚犯都想坦白吗?

	乙抵赖	乙坦白
甲抵赖	-1，-1	-10，0
甲坦白	0，-10	-4，-4

左侧甲收益，右侧乙

- 以甲的视角看:
  - 如果乙坦白，那么甲坦白收益-4，抵赖收益-10 -> 甲应该坦白
  - 如果乙抵赖，那么甲坦白收益0，抵赖收益-1 -> 甲应该坦白
- 坦白是严格占优策略：两个人都坦白，最后收益都是-4
- 理性博弈必然结果：如果抵赖那么都只需判一年，可是另一方会坦白

	乙抵赖	乙坦白
甲抵赖	-1，-1	-10，0
甲坦白	0，-10	-4，-4

左侧甲收益，右侧乙

军备竞赛

- 囚徒困境：个体私利前，合作是困难的
- 服用兴奋剂：
  - 如果我服用了而对手没服用，我可以获得优势，但这会损害身体，并可能DQ
  - 假设兴奋剂较难检测到，并且假设运动员副作用只是小问题

	B未服用	B服用
A未服用	3，3	1，4
A服用	4，1	2，2

左侧A收益，右侧B

- 存在严格占优策略：服用兴奋剂，得到 (2, 2)的收益
- 但是都不服用可以得到(3, 3)

- 这种(2, 2) vs (3, 3)的情形称为军备竞赛
  - 为保持实力相当，大家都会尽力产生更先进的武器，但是力量对比并未改变

- 略微改变收益函数也可以改变囚徒困境
  - 如果准备考试，那么预计能得~~92~~100分，但如果不准备则只有~~80~~96分
  - 如果A和B都准备了，可得100；只有一个人准备了，得92；否则，得84
  - 严格占优策略：准备报告

	B报告	B考试
A报告	98，98	94，96
A考试	96，94	92，92

左侧A得分，右侧B

概念：最佳应对 (best response)

- 依然考虑2个参与人的情形
- 设S是A的一个策略，T是B的一个策略（准备考试/报告）
- $P_A(S,T)$代表A从这组决策得到的收益，$P_B(S,T)$代表B的收益

若A用S达到收益不低于任何策略S’，则称S为对B的T策略的最佳应对

$$\forall S', \qquad P_A(S,T) \geq P_A(S',T)$$

最佳应对可能不唯一；强调唯一性时考虑严格最佳应对：

$$\forall S' \neq S, \qquad P_A(S,T) > P_A(S',T)$$

概念： 占优策略 (dominant strategy)

称S是A的占优策略，如果S对B的每个策略都是最佳应对
称S是A的严格占优策略，如果S对B的每个策略都是严格最佳应对

- 双方有多个占优策略时很难预测对方的行为
  - 严格最优策略的存在规避了这个问题

- 是否有可能无占优策略?

（部分）无占优策略的例子及决策

- 有两家公司，各自规划研发生产一款新品，投放后将直接对立竞争
- 有两个市场：廉价和高档
  - 并假设廉价和高档产品单件利润等同
  - 公司可以选择廉价还是高档，利润等同于销量
- 消费群体有60%倾向于廉价，40%倾向于高档
- A公司强势，如果在同一个市场竞争那么A拿到80%销量（B得到20%）
- 若两公司针对不同市场：各自可以拿到全份额，60%和40%
- 若两公司都针对廉价市场：A收益0.48；B收益0.12
- 若两公司都针对高档市场：A收益0.32，B收益0.08

	B廉价	B高档
A廉价	.48，.12	.60，.40
A高档	.40，.60	.32，.08

	B廉价	B高档
A廉价	.48，.12	.60，.40
A高档	.40，.60	.32，.08

- A公司有严格占优策略：廉价
- B公司无占优策略
  - 若A廉价，B高档是最佳应对
  - 若A高档，B廉价是最佳应对

- 尽管B无占优策略，依然可以决策
  - A有严格占优策略，因此假定A采用该策略
  - 此时B采用最佳应对：高档
  - 注意：A和B依然是同时在预先做决策，即使我们考虑了两步过程
  - 参与者全信息的重要性

两人都无严格占优策略的例子

- 如果两个人都没有严格占优策略，如何预测决策?
- 无严格占优策略的例子：
  - 大小两公司都想与ABC三个大客户之一谈业务
  - 若两公司找同一个客户，大小两公司各分得一半业务
  - 小公司无法单独承担任何一个客户的业务，需要与大公司合作
  - 大公司若单独找B或C，则得到他们全部业务；找A则必须与小公司合作才能承接
  - A是大客户，全盘承接收益8；全盘承B或C收益2

		大公司		
		A	B	C
小公司	A	4，4	0，2	0，2
	B	0，0	1，1	0，2
	C	0，0	0，2	1，1

- 从小公司视角考虑
  - 若大公司选A，严格最佳应对是A
  - 若大公司选B，严格最佳应对是B
  - 若大公司选C，严格最佳应对是C
- 从大公司视角考虑
  - 若小公司选A，严格最佳应对是A
  - 若小公司选B，严格最佳应对是C
  - 若小公司选C，严格最佳应对是B
- 无严格占优策略！

		大公司		
		A	B	C
	A	4，4	0，2	0，2
小公司	B	0，0	1，1	0，2
	C	0，0	0，2	1，1

纳什均衡

对于两个参与人的策略S和T，(S,T)称为一个纳什均衡，如果S和T互为最佳应对

- 采用纳什均衡的人没有动机去改变策略
  - 相应地，如果不是互为最佳应对，那么一方会觉得另一方可能采取其他策略
  - “理性参与人”假设本身不能推出采用纳什均衡的必然性
  - 纳什均衡是一种“信念”上的均衡：若所有参与人都相信大家会采用纳什均衡的策略，那么自己就有动机也采用
- 大小公司博弈：(A,A)是一个纳什均衡（且是唯一纳什均衡）

		大公司		
		A	B	C
小公司	A	4，4	0，2	0，2
	B	0，0	1，1	0，2
	C	0，0	0，2	1，1

多重均衡：基本协调博弈

- A和B为了一个合作项目准备slides，需要决定用PPT还是Keynote
- 然而无法及时联系到另一方（在飞机上做slides）
- 使用哪一个都无所谓，只要使用同一个就更容易合并

	B：PPT	B：Keynote
A：PPT	1，1	0，0
A：Keynote	0，0	1，1

- 纳什均衡：(PPT, PPT)和(Keynote, Keynote)
- 另例：两人在拥挤mall寻找对方，需要决定在南门还是北门等对方

focal point

- Thomas Schelling (1960)提出了focal point作为解决协调问题的框架
- 经常有一些“自然”的原因（可能在收益模型范畴之外）使参与者可以focus on其中一个纳什均衡
  - 例子：自行车迎面相向行驶，如何选择靠左躲避还是靠右躲避?
  - 选左右均可，关键是选择一致；社会因素约定俗成会主动向右行驶
  - 都选择向右走就是focal point

非平衡协调博弈

- 如果已知A和B平时都更喜欢用Keynote

	B：PPT	B：Keynote
A：PPT	1，1	0，0
A：Keynote	0，0	2，2

- 依然是有两个纳什均衡，但是 (Keynote, Keynote)收益更高
  - 在商场碰面，发现北门有一家两个人都喜欢的店
- focal point：多个纳什均衡，选择对双方收益都高的，依然可以协调
  - 不是外部约定俗成，而是使用博弈本身的结构

协调博弈变种：battle of the sexes

- 如果A和B各自喜欢用不同的软件

	B：PPT	B：Keynote
A：PPT	1，2	0，0
A：Keynote	0，0	2，1

- 两个利益冲突的纳什均衡：battle of the sexes
- battle of the sexes：丈夫和妻子选电影，动作片 vs 爱情片
  - 想要协调彼此选择，因此有 (动作，动作)和 (爱情，爱情)的均衡，但是每个均衡只对夫妻一方有较高收益
- 如何解决分歧?

协调博弈变种：猎鹿博弈 (Skyrms, 2003)

- 源于卢梭写的一个故事
  - 猎人A，B外出狩猎，若合作就可以猎到鹿（高收益）
  - 分开狩猎只能分别猎到野兔（低收益）
  - 如果分开狩猎还非要猎鹿，那么就会失败（收益0），但另一方还能猎到野兔

	B猎鹿	B猎兔
A猎鹿	4，4	0，3
A猎兔	3，0	3，3

	B：PPT	B：Keynote
A：PPT	1，1	0，0
A：Keynote	0，0	2，2

- 类似非平衡协调博弈，只是不合作但追求更高受益者会被惩罚
- 决策难点：高收益、高协调失败风险 vs 较低收益、无协调失败风险

与囚徒困境的联系

- 不同点：囚徒困境具有严格占优策略
- 类似点：合作受益，单方面寻求合作有损失

	乙抵赖	乙坦白
甲抵赖	-1，-1	-10，0
甲坦白	0，-10	-4，-4

囚徒困境

- 改变考试-报告博弈的收益结构，会导致类似猎鹿博弈的问题
    - 存在 (报告，报告)和 (考试，考试)两个均衡
    - 一方若试图选择“报告”以谋求高收益，可能反而得到低收益

	B猎鹿	B猎兔
A猎鹿	4，4	0，3
A猎兔	3，0	3，3

	B报告	B考试
A报告	90，90	82，88
A考试	88，82	88，88

鹰鸽博弈：“反协调”的博弈

- A、B两动物决定一块食物在彼此间的分配
- 可选争夺（鹰派策略Hawk）或者分享（鸽派策略Dove）
  - 若都选分享，最后均匀分配食物，各自收益3
  - 若A争夺B分享，A收益5，B收益1
  - 若都选争夺，那么收益0（争夺中践踏食物/食物被趁火打劫/受伤得不偿失）
- 纳什均衡：(D, H)和(H, D)　对称性，很难预测选择哪种策略
- 考试-报告博弈改变收益可以得到鹰鸽博弈
  - 都不想当那个D，但是如果都选了H结果会很坏
- 两人驾车对冲比试胆量，先转向的人输，但都不转向会发生车祸

	B：D	B：H
A：D	3，3	1，5
A：H	5，1	0，0

	B报告	B考试
A报告	90，90	86，92
A考试	92，86	76，76

无纳什均衡的博弈

- 存在没有纳什均衡的博弈：攻防博弈
- 硬币配对博弈
  - A、B各持一枚硬币，同时决定要展示正面还是反面
  - A把自己硬币输给B如果正反一致，否则赢得B的硬币如果不一致

	B正	B反
A正	-1，+1	+1，-1
A反	+1，-1	-1，+1

- 硬币配对博弈是典型的零和博弈：每个结果中参与人总收益=0

- 零和博弈广泛存在于攻防博弈，和一般的双方利益直接对立的博弈中
- 1944年6月的诺曼底登陆
  - 盟军需要决策在诺曼底还是加莱登陆
  - 决策受到德军策略的影响：在诺曼底还是加莱大规模设防

- 硬币配对博弈中不存在一组策略彼此互为最佳应对（无纳什均衡）
  - A知道B的正反，则一定会采取与B的对立策略；类似的B也这么决策

- 将策略引入随机性：变成正反的一个分布
  - 迷惑对手，让对手无法预测自己的行为
  - 可以解决无纳什均衡的问题！

	B正	B反
A正	-1，+1	+1，-1
A反	+1，-1	-1，+1

混合策略：定义

- 混合策略 (mixed strategy)：一组策略的概率混合/随机分布
  - A以概率p选正，1 - p概率选反；B以概率q选正，1 – q概率选反
  - A的混合策略可以用p来简单表示
- 纯策略 (pure strategy)：确定性的策略（p = 0/1, q = 0/1）
- 混合策略的收益：考虑采取纯策略时的期望收益
  - 原因：即使A采用纯策略，B策略的随机性也会使A的收益随机化
  - 例如，A采用正，那么期望收益$-q+1-q=1-2q$
  - 最后，每个人的每个纯策略会对应一个期望收益

	B正	B反	
A正	-1，+1	+1，-1	1-2q
A反	+1，-1	-1，+1	2q-1
	1-2p	2p-1	

混合策略的纳什均衡

- 纳什均衡：一对混合策略(S,T)，使得S和T互为最佳应对（对混合收益）
- 观察：纯策略（确定性，p = 0/1）不可能是纳什均衡
  - 例如A若采用正，那么B应该也采用正（+1），但此时A应采用反，非均衡
- 考察A：采用正，收益1 – 2q；采用反，收益2q – 1
  - 若 $1-2q \neq 2q-1$ ，那么必然采用纯策略更优（不把概率浪费在更坏收益上）
  - 然而，纯策略一定不是纳什均衡的
  - 这推出$1-2q=2q-1$，所以$q=0.5$
  - 对称可得$p=0.5$
- 结论：$p=q=0.5$是纳什均衡
- 一般地：无差异原则

	B正	B反	
A正	-1，+1	+1，-1	1-2q
A反	+1，-1	-1，+1	2q-1
	1-2p	2p-1	

解释

- 考虑A和B根据p和q的混合策略去多次重复进行博弈
  - 无差异规则：双方目标是让自己没有patter可循，只能靠瞎猜
  - 这是最优的，否则会让对方利用自身“弱点”
  - 例如，A超过一半的时间会采用正，那么B可以总是采用正来应对
- p = q = 0.5代表自己的策略不会被“识破”；蒙蔽对方
  - 0.5是巧合，对于其他收益会产生其他结果
  - 蒙蔽对方的结果就是也“蒙蔽自己”

- 宏观视角：两大类群动物，频繁进行1v1的攻防博弈
  - 一类动物是攻击方，另一类动物是防守方
  - 每个个体都只会采用纯策略，此时这种纯策略个体的占比就对应于p和q

零和博弈举例：持球-抛球博弈

- 持球-抛球博弈
  - 若进攻方的持球被阻断，则进攻方收益0
  - 若进攻方持球而防守方防守抛球，进攻方收益5
  - 若进攻方抛球而防守方拦断持球，进攻方收益10

	防守抛球	拦断持球
抛球	0，0	10，-10
持球	5，-5	0，0

- 纯策略是否有纳什均衡?
- 混合策略：进攻方p概率抛球，防守方q概率防守抛球
  - 进攻方抛球期望收益：10 – 10q
  - 进攻方持球期望收益：5q
  - q = 2 / 3，验证：p = 1 / 3
  - 虽然只有1 / 3概率抛球，防守方仍需要2 / 3概率防守抛球
  - 进攻方期望收益10 / 3

零和博弈举例：罚点球博弈

- 罚点球博弈 (Palacios-Huerta 2002)：
  - 主罚队员选择向左还是向右射门，守门员则选择向左还是向右扑救
  - 根据1400多份专业足球比赛罚点球资料分析，得出下列收益矩阵

	守门员L	守门员R
主罚L	.58，-.58	.95，-.95
主罚R	.93，-.93	.70，-.70

- 纯策略纳什均衡？混合策略纳什均衡？主罚p = .39，守门员q = .42
- Palacios-Huerta指出实际统计表明q = .42，p = .40

纯/混合策略同时存在纳什均衡

- 纯策略均衡 及 混合策略均衡 可以同时存在

	B：PPT	B：Keynote
A：PPT	1，1	0，0
A：Keynote	0，0	2，2

- 设A策略p，B策略q
  - A在PPT和Keynote上无差异：$q = 2(1 - q)$，推出$q = 2/3$；类似地$p = 2/3$

非零和博弈的纳什均衡?

- 已经讲过的几个博弈的混合策略纳什均衡?

	乙抵赖	乙坦白
甲抵赖	-1，-1	-10，0
甲坦白	0，-10	-4，-4

	B猎鹿	B猎兔
A猎鹿	4，4	0，3
A猎兔	3，0	3，3

	B：PPT	B：Keynote
A：PPT	1，2	0，0
A：Keynote	0，0	2，1

	B：D	B：H
A：D	3，3	1，5
A：H	5，1	0，0

有关纳什均衡

- Antoine Augustin Cournot 1838： Cournot equilibrium 纯策略纳什均衡
- John von Neumann 1928：minimax定理，证明了双参与者零和博弈纳什均衡的存在性，被认为是game theory的开山之作
- John Forbes Nash 1950, 1951：任何有限博弈都存在纳什均衡

双参与者零和博弈的纳什均衡与线性规划

零和博弈的纳什均衡：从例子说起

	Col：S	Col：T
Row：S	3，-3	-1，1
Row：T	-2，2	1，-1

- 若Row选S概率x_1，选T概率x_2，并且Col知道Row这个策略
- 那么Col选S的期望收益$-3x_1+2x_2$，选T的期望收益x_1-x_2
- 因此Col在最优策略下的收益是$\max\{-3x_1+2x_2,\ x_1-x_2\}$
- 因为是零和博弈，此时Row的收益是
$$-\max\{-3x_1+2x_2, x_1-x_2\} = \min(3x_1-2x_2, -x_1+x_2)$$
- 所以Row应该选$(x_1,x_2) \in \arg\max_{(x_1,x_2)} \min(3x_1-2x_2, -x_1+x_2)$

Col导出的关于Row optimal的LP

- $(x_1, x_2) \in \arg\max_{(x_1,x_2)} \min(3x_1 - 2x_2, -x_1 + x_2)$

$$
\begin{aligned}
&\max z \\
\text{s.t.}\quad & 3x_1 - 2x_2 \geq z \\
& -x_1 + x_2 \geq z \\
& x_1 + x_2 = 1 \\
& x_1, x_2 \geq 0
\end{aligned}
$$

- 最优解是$x_1 = \frac{3}{7}, x_2 = \frac{4}{7}, z = \frac{1}{7}$

Row导出的关于Col optimal的LP

	Col：S	Col：T
Row：S	3，-3	-1，1
Row：T	-2，2	1，-1

- 再来看另一边，若Col选S概率y_1，选T概率y_2且Row知道这个策略
- 那么Row选S收益$3y_1 - y_2$，选T收益$-2y_1 + y_2$
- Row在最优策略下收益$\max\{3y_1 - y_2, -2y_1 + y_2\}$
- 因为零和，所以此时Col的收益是
- $-\max\{3y_1 - y_2, -2y_1 + y_2\} = \min\{-3y_1 + y_2, 2y_1 - y_2\}$
- 因此$(y_1, y_2) \in \arg\max_{(y_1,y_2)} \min\{-3y_1 + y_2, 2y_1 - y_2\}$

- $(y_1, y_2) \in \arg\max_{(y_1,y_2)} \min\{-3y_1 + y_2, 2y_1 - y_2\}$

$$
\begin{aligned}
&\max w \\
\text{s.t.}\quad & -3y_1 + y_2 \geq w \\
& 2y_1 - y_2 \geq w \\
& y_1 + y_2 = 1 \\
& y_1, y_2 \geq 0
\end{aligned}
$$

- 最优解是$y_1 = \frac{2}{7}, y_2 = \frac{5}{7}, w = -\frac{1}{7}$

线性规划的视角看纳什均衡

- 最优解是$x_1 = \frac{3}{7}, x_2 = \frac{4}{7}, z = \frac{1}{7}$
- 最优解是$y_1 = \frac{2}{7}, y_2 = \frac{5}{7}, w = -\frac{1}{7}$
- 你会发现:
  - $w + z = 0$
  - 给定x的最优解，y的最优解是最佳应对（不能继续增大）
  - 给定y的最优解，x的最优解也是最佳应对，因此x和y这组最优解是纳什均衡

- 我们证明两个人分别做这种优化确实可以达到纳什均衡

一般双参与者零和博弈的纳什均衡

- 设R和C分别是Row和Col的收益矩阵

	Col：S	Col：T
Row：S	3，-3	-1，1
Row：T	-2，2	1，-1

→

	Col：S	Col：T
Row：S	3	-1
Row：T	-2	1

R

	Col：S	Col：T
Row：S	-3	-1
Row：T	2	-1

C

- 设$\boldsymbol{x} = (x_1, x_2)$代表Row的混合策略，$\boldsymbol{y} = (y_1, y_2)$代表Col的混合策略
- 我们考虑零和博弈所以有$R + C = 0$

- 采用刚刚的思路：

R

	Col：S	Col：T
Row：S	3	-1
Row：T	-2	1

C

	Col：S	Col：T
Row：S	-3	-1
Row：T	2	-1

- 考虑Col，$\boldsymbol{x}^T C$是一个2维的向量，Col的收益应该是对两维取max
  - 等价于$\max_{\boldsymbol{y}} \boldsymbol{x}^T C\boldsymbol{y}$，并且最优的y可以取成对应两维max的纯策略
- 对Row来说，收益是$-\max_{\boldsymbol{y}} \boldsymbol{x}^T C\boldsymbol{y} = \min_{\boldsymbol{y}} \boldsymbol{x}^T R\boldsymbol{y}$，最后应该追求

$$\boldsymbol{x} = \arg\max_{\boldsymbol{x}} \min_{\boldsymbol{y}} \boldsymbol{x}^T R\boldsymbol{y}$$

- 把$\boldsymbol{x} = \arg\max_{\boldsymbol{x}} \min_{\boldsymbol{y}} \boldsymbol{x}^T R \boldsymbol{y}$ 等价写成LP

$$
\begin{aligned}
& \max z \\
s.t. \quad & \boldsymbol{x}^T R \geq z \mathbf{1}^T \\
& \boldsymbol{x}^T \mathbf{1} = 1 \\
& x_i \geq 0
\end{aligned}
$$

LP(1)

- 考虑LP(1)的dual

$$
\begin{aligned}
& \min z' \\
s.t. \quad & -\boldsymbol{y}^T R^T + z' \mathbf{1}^T \geq \mathbf{0} \\
& \boldsymbol{y}^T \mathbf{1} = 1 \\
& y_j \geq 0
\end{aligned}
$$

DP(1)

$\max \boldsymbol{c}^T \boldsymbol{x}$ s.t. $A\boldsymbol{x} \leq \boldsymbol{b},\ \boldsymbol{x} \geq 0$的dual是$\min \boldsymbol{b}^T \boldsymbol{y}$ s.t. $A^T \boldsymbol{y} \geq \boldsymbol{c},\ \boldsymbol{y} \geq 0$

如何求LP的dual

Constructing the dual LP [edit]

In general, given a primal LP, the following algorithm can be used to construct its dual LP.[1]:85 The primal LP is defined by:

- A set of *n* variables: $x_1, \ldots, x_n$.
- For each variable x_i, a *sign constraint* – it should be either non-negative ($x_i \geq 0$), or non-positive ($x_i \leq 0$), or unconstrained ($x_i \in \mathbb{R}$).
- An objective function: $\text{maximize} \quad c_1 x_1 + \cdots + c_n x_n$
- A list of *m* constraints. Each constraint *j* is: $a_{j1} x_1 + \cdots + a_{jn} x_n \lesseqqgtr b_j$ where the symbol before the b_j can be one of $\geq$ or $\leq$ or $=$.

The dual LP is constructed as follows.

- Each primal constraint becomes a dual variable. So there are *m* variables: $y_1, \ldots, y_m$.
- The sign constraint of each dual variable is "opposite" to the sign of its primal constraint. So "$\geq b_j$" becomes $y_j \leq 0$ and "$\leq b_j$" becomes $y_j \geq 0$ and "$= b_j$" becomes $y_j \in \mathbb{R}$.
- The dual objective function is $\text{minimize} \quad b_1 y_1 + \cdots + b_m y_m$
- Each primal variable becomes a dual constraint. So there are *n* constraints. The coefficient of a dual variable in the dual constraint is the coefficient of its primal variable in its primal constraint. So each constraint *i* is: $a_{1i} y_1 + \cdots + a_{mi} y_m \lesseqqgtr c_i$, where the symbol before the c_i is similar to the sign constraint on variable *i* in the primal LP. So $x_i \leq 0$ becomes "$\leq c_i$" and $x_i \geq 0$ becomes "$\geq c_i$" and $x_i \in \mathbb{R}$ becomes "$= c_i$".

From this algorithm, it is easy to see that the dual of the dual is the primal.

- 考虑LP(1)的dual

$$\begin{aligned} &\min z' \\ s.t.\quad -\boldsymbol{y}^T R^T + z'\mathbf{1}^T &\geq \mathbf{0} \\ \boldsymbol{y}^T\mathbf{1} &= 1 \\ y_j &\geq 0 \end{aligned} \qquad \text{DP(1)}$$

- 令$z'' = -z'$，然后利用$C = -R$得到

$$\begin{aligned} &\max z'' \\ s.t.\quad C\boldsymbol{y} &\geq z''\mathbf{1} \\ \boldsymbol{y}^T\mathbf{1} &= 1 \\ y_j &\geq 0 \end{aligned} \qquad \text{LP(2)}$$

- 最后z''最大值是 $\max\limits_{\boldsymbol{y}} \min\limits_{\boldsymbol{x}} \boldsymbol{x}^T C\boldsymbol{y} = -\min\limits_{\boldsymbol{y}} \max\limits_{\boldsymbol{x}} \boldsymbol{x}^T R\boldsymbol{y}$

$$
\begin{aligned}
& \max z \\
s.t.\quad & \boldsymbol{x}^T R \geq z\mathbf{1}^T \\
& \boldsymbol{x}^T \mathbf{1} = 1 \\
& x_i \geq 0
\end{aligned}
$$

LP(1)

$$
\begin{aligned}
& \max z'' \\
s.t.\quad & C\boldsymbol{y} \geq z''\mathbf{1} \\
& \boldsymbol{y}^T \mathbf{1} = 1 \\
& y_j \geq 0
\end{aligned}
$$

LP(2)

- LP(2)恰好是Col的收益最优LP
- 考虑最优解$(\boldsymbol{x}, z), (\boldsymbol{y}, z'')$ 则 Strong duality of LP: $z = -z''$

strong duality推出纳什均衡

- 下面证最优解$(\boldsymbol{x}, z), (\boldsymbol{y}, z'')$互为最佳应对（也就是纳什均衡）

$$\boldsymbol{x}^T R \geq z\mathbf{1} \Rightarrow \boldsymbol{x}^T R \boldsymbol{y} \geq z \qquad (1)$$

$$\begin{aligned} C\boldsymbol{y} \geq z''\mathbf{1} &\Rightarrow \boldsymbol{x}'^T C\boldsymbol{y} \geq z'', \forall \boldsymbol{x}' \\ &\Rightarrow \boldsymbol{x}'^T R\boldsymbol{y} \leq -z'', \forall \boldsymbol{x}' \end{aligned} \qquad (2)$$

- 如果Col用策略y，那么(2)说明Row不论用什么策略收益都至多z
- 如果Row用x来应对y，Row的收益至少z （说明收益恰好是z）
- 因此x是y的最佳应对
- 同理可证y是x的最佳应对

重要推论

- 任何双参与者的零和博弈都存在纳什均衡
- Minimax Theorem:

$$\max_{\boldsymbol{y}} \min_{\boldsymbol{x}} \boldsymbol{x}^T R \boldsymbol{y} = \min_{\boldsymbol{y}} \max_{\boldsymbol{x}} \boldsymbol{x}^T R \boldsymbol{y}$$

简证： $\mathrm{z}'' = \max_{\boldsymbol{y}} \min_{\boldsymbol{x}} \boldsymbol{x}^T C \boldsymbol{y} = -\min_{\boldsymbol{y}} \max_{\boldsymbol{x}} \boldsymbol{x}^T R \boldsymbol{y} = -z = -\max_{\boldsymbol{x}} \min_{\boldsymbol{y}} \boldsymbol{x}^T R \boldsymbol{y}$

（利用strong duality $z = -z''$）

最后结合$\max_{\boldsymbol{y}} \min_{\boldsymbol{x}} \boldsymbol{x}^T C \boldsymbol{y} = -\min_{\boldsymbol{y}} \max_{\boldsymbol{x}} \boldsymbol{x}^T R \boldsymbol{y}$ 得到结论

纳什均衡都是LP最优解

- 任何纳什均衡都可被LP(1)和LP(2)的最优解找到
- 具体：若$(\boldsymbol{x},\boldsymbol{y})$是纳什均衡，则$(\boldsymbol{x},\boldsymbol{x}^T R\boldsymbol{y})$是LP(1)最优解且$(\boldsymbol{y},-\boldsymbol{x}^T C\boldsymbol{y})$是LP(2)的
- 证：令$z=\boldsymbol{x}^T R\boldsymbol{y}=-\boldsymbol{x}^T C\boldsymbol{y}$
- 根据纳什均衡

$$\boldsymbol{x}^T C\boldsymbol{y}\geq \boldsymbol{x}^T C\boldsymbol{e}_j\ \forall j$$

$$\Rightarrow -\boldsymbol{x}^T R\boldsymbol{y}\geq -\boldsymbol{x}^T R\boldsymbol{e}_j\ \forall j\quad\Rightarrow\quad \boldsymbol{x}^T R\geq z\mathbf{1}$$

- 因此$(\boldsymbol{x},z)$是LP(1)的一组可行解；同理可证$(\boldsymbol{y},-z)$是LP(2)的可行解
- 然而LP(1)和LP(2)是对偶关系，因此z = -z推出两个解都是最优解

更多推论

- 推论：所有纳什均衡(x, y)构成的集合是凸集合
  - 凸集合：若点P和Q在集合里，那么PQ连线的所有点都在集合里
- 推论：Row在所有纳什均衡的收益都是一样的（Col同理）
- 因此，我们可以定义零和博弈的收益是纳什均衡的收益

LP是否可以高效求解?

- 零和博弈的纳什均衡等价于解LP
- 如何解LP?
- 单纯形法 (simplex method)：Danzig 1947
  - 最坏情况是exp(n)的
  - 实际情况表现极佳
  - Spielman and Teng 2004提出了在smoothed analysis下单纯形法是poly(n)的 (Gödel prize)
- 椭圆法（ellipsoid mehtod）：Khachiyan 1979
  - 真正的最坏情况poly(n)
  - 有separation oracle的情况下甚至可以解exp(n)大小的LP
- 求解器广泛存在，尤其是近似求解非常迅速

非零和博弈的纳什均衡的计算?

- Lemke-Howson 1964是一个类似于单纯形法的求纳什均衡的算法
- 根据Lemke-Howson的行为，Papadimitriou 1991提出了PPAD
  - PPAD = Polynomial Parity Argument in a Directed graph
- 双参与者博弈的纳什均衡的精确计算是PPAD-complete的 (Chen, Deng, Teng, 2009)
  - 同时，PTAS不存在除非PPAD在P里
  - Smoothed polynomial精确解不存在除非PPAD = RP
  - 相对误差：$\boldsymbol{x}^T R\boldsymbol{y} \geq (1-\epsilon)\boldsymbol{x}'^T R\boldsymbol{y}$，$\forall \boldsymbol{x}'$
- 很多evidence说明PPAD ≠ P
- 近似算法？ A note on approximate Nash equilibria (Daskalakis et al. 09)

How hard is it to approximate the best Nash equilibrium? (Hazan, Krauthgamer 09) …

多人博弈

- n个参与人$P_1, \dots, P_n$
- 第i个人的收益函数u_i：若第i个人策略是S_i，那么i的收益$u_i(S_1, \dots, S_n)$
- 对于某个策略组S $= (S_1, \dots, S_n)$，令$S_{-i} = (S_1, \dots, S_{i-1}, S_{i+1}, \dots, S_n)$
- 从i的角度讨论，称策略S_i是最佳应对，如果满足对任意策略组S

$$u_i(S_i; S_{-i}) \geq u_i(S_i'; S_{-i}), \qquad \forall S_i'$$

- 如果$(S_1, \dots, S_n)$每个S_i都是最佳应对，那么称这个策略组是纳什均衡的

- 多人的constant-sum博弈也有类似minimax theorem以及poly(n)算法 (Cai, Daskalakis 2011)

多人博弈：混合策略

- 每个参与人i的策略表示成一个m维向量x_i表示一个概率分布
  - m代表可能的策略的个数
- 对策略组$x = (x_1, \ldots, x_n)$参与人i，i的收益函数$u_i(x) \coloneqq \mathbb{E}_{S\sim x}[u_i(S)]$
  - 代表所有人采用所有可能的纯策略组S，根据x定义的概率求期望后，i所得到的期望收益
- 混合策略组$x = (x_1, \ldots, x_n)$是纳什均衡的，若对任何i以及纯策略S_i

$$u_i(x) \geq u_i(S_i; x_{-i})$$

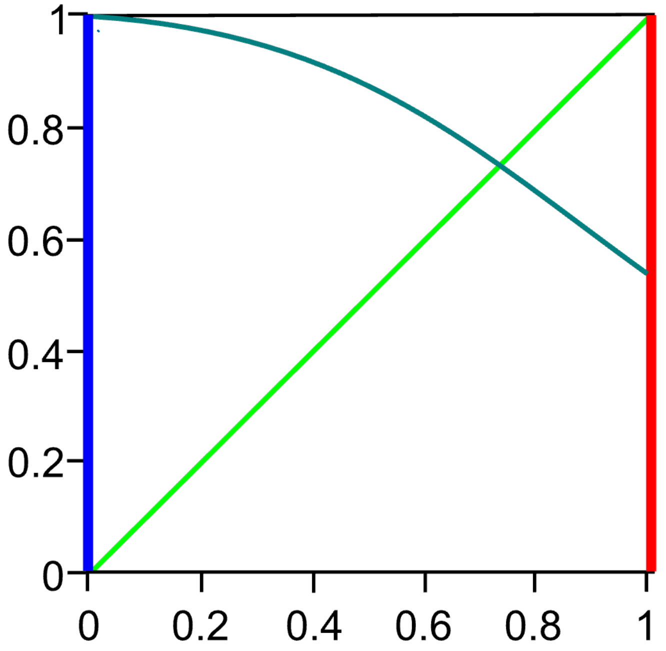
- 给定混合策略组$x = (x_1, \ldots, x_n)$，检查x是否构成纳什均衡是容易的
  - 多项式于输入时间（输入中u函数需要对于每个纯策略组定义！）

一般纳什均衡的存在性

Nash的证明

Brouwer's fixed-point theorem
Let D be a convex, compact subset of Euclidean space. If $f : D \to D$ is continuous then there exists $x \in D$ such that f(x) = x

- 理解：考虑一维，那么D就是一个闭区间，f是闭区间D上的连续函数
  - 必与y = x有交点（中值定理）



- 将中国地图放在国内某个地方，标注的“你在这里”就是不动点
- 地球绕轴自转，自转轴在转的过程中不变
- 桌上一张纸，任意揉皱后放回原来所在地方，必有水平方向重合点

证明思路

- 总体思路：利用不动点定理

1. 设X为所有混合策略组$x = (x_1, \ldots, x_n)$的集合（x_i是第i个人的混合策略分布）
2. 构造一个$f: X \rightarrow X$，并说明f满足不动点定理的要求
3. 得到一个不动点x使得f(x) = x
4. 证明这个不动点对应的混合策略组是纳什均衡的

构造f

Gain: 给定混合策略组$x = (x_1, \dots, x_n) \in X$，对参与者i以及纯策略$S_i$，

$$\mathrm{Gain}_{i;S_i}(x) = \max\{u_i(S_i; x_{-i}) - u_i(x), 0\}$$

- Gain表示相对于混合策略x，参与者i采用纯策略S_i所取得的收益增量
  - 如果增量是负的，那么取0
- 定义$f: X \to X$

给定$x \in X$，$f(x) = y \in X$定义为$y = (y_1, \dots, y_n)$使得对任何$1 \le i \le n$和i的纯策略S_i，

$$y_i(S_i) \coloneqq \frac{x_i(S_i) + \mathrm{Gain}_{i;S_i}(x)}{1 + \sum_{S_i'} \mathrm{Gain}_{i;S_i'}(x)}$$

S_i'是i的纯策略

分母的意义仅仅是归一化，确保对任何i，$\sum_{S_i} y_i(S_i) = 1$

找到不动点x^*

给定$x \in X$，$f(x) = y \in X$定义为$y = (y_1, \dots, y_n)$使得对任何$1 \le i \le n$和i的纯策略S_i，

$$y_i(S_i) \coloneqq \frac{x_i(S_i) + \mathrm{Gain}_{i;S_i}(x)}{1 + \sum_{S_i'} \mathrm{Gain}_{i;S_i'}(x)}$$

Brouwer's fixed-point theorem

Let D be a convex, compact subset of Euclidean space. If $f: D \to D$ is continuous then there exists $x \in D$ such that f(x) = x

- f是连续的，定义域和值域X对应n个概率分布的拼接，是convex compact的
- 可以应用不动点定理，找到一个x^*使$f(x^*) = x^*$

证明：x^*是纳什均衡

Gain: 给定混合策略组$x=(x_1,\dots,x_n)\in X$，对参与者i以及纯策略S_i，

$$\mathrm{Gain}_{i;S_i}(x)=\max\{u_i(S_i;x_{-i})-u_i(x),0\}$$

混合策略组$x=(x_1,\dots,x_n)$是纳什均衡的，若对任何i以及纯策略S_i

$$u_i(x)\ge u_i(S_i;x_{-i})$$

- x^*是纳什均衡的等价于
$$\mathrm{Gain}_{i;S_i}(x^*)=0,\qquad \forall 1\le i\le n,S_i$$
- 反证法。设存在i和S_i使得$\mathrm{Gain}_{i;S_i}(x^*)>0$
- 必然存在S_i'使$u_i(S_i';x_{-i}^*)-u_i(x^*)<0$且$x_i^*(S_i')>0$
  - 原因：$\sum_{S_i'}x_i^*(S_i')[u_i(S_i';x_{-i}^*)-u_i(x^*)]=0$

推出矛盾

- 必然存在S_i'使$u_i(S_i'; x_{-i}^*) - u_i(x^*) < 0$且$x_i^*(S_i') > 0$
- 那么$\mathrm{Gain}_{i;S_i'}(x^*) = 0$
- 设$x^* = f(x^*) = y^*$，那么

$$y_i^*(S_i') = \frac{x_i^*(S_i') + \mathrm{Gain}_{i;S_i'}(x^*)}{1 + \sum_{S_i''} \mathrm{Gain}_{i;S_i''}(x^*)} < x_i^*(S_i')$$

- 这说明$\mathrm{y}^* = f(x^*) \neq x^*$，与$x^*$是不动点的事实矛盾，证毕

不动点定理的证明？

- 经典组合证明：Sperner 's lemma
- https://arxiv.org/pdf/1409.7890.pdf

策略的简化：严格非优策略与纳什均衡

严格非优策略

- 对于参与人i，称纯策略S_i是严格非优的，若存在S_i'使其他参与人任意策略下收益都高于S_i。即$\exists S_i'$对任意策略组S有
$$u_i(S_i';S_{-i}) > u_i(S_i;S_{-i})$$
- 严格非优策略可以“安全”的排除掉
- 作用：可能所有当前策略都不占优，但是存在严格非优策略，可以反复排除，最后得到占优策略

案例：设施选址博弈

- 甲乙公司分别计划在6个地点开设一家店
  - 甲可以选ACE中一个，乙可以选BDF中一个，甲乙独立同时决策
  - 一旦开张，顾客会去离自己较近的店活动，顾客来自于ABCDEF
  - 设顾客在6个点数量分别相等，且收益直接与顾客数量成正比
    - 甲在C开店，乙在B开店，那么B吸引AB顾客，而C吸引CDEF顾客，甲收益4乙收益2

		乙		
		B	D	F
	A	1，5	2，4	3，3
甲	C	4，2	3，3	4，2
	E	3，3	2，4	5，1

- 事实：无占优策略

- 对甲来说，A是严格非优策略：选A的时候总可以选C得到更佳收益
- 对乙来说，F是严格非优策略：总可以选D得到更加收益

- 排除A和F！

乙

		B	D	F
甲	A	1，5	2，4	3，3
	C	4，2	3，3	4，2
	E	3，3	2，4	5，1

- 此时，B和E突然也变成了严格非优！
- 删除后，最后只剩下C和D，也成了甲乙唯一的策略

		乙 B	乙 D
甲	C	4，2	3，3
	E	3，3	2，4

删除严格非优不改变纳什均衡

- Claim：达到纳什均衡的策略组构成的集合在删除非优策略后保持不变
- Fact：任何原博弈的纳什均衡都是简化后的纳什均衡
  - 反证：假设有一个原均衡中的策略S被删除了，这说明存在严格更优的策略S'
  - 说明S不可能是占优策略，矛盾！
- Fact：任何简化后的纳什均衡也是原博弈的纳什均衡
  - 反证：假设存在$(S_1, \dots, S_n)$是简化后博弈的纳什均衡但不是原博弈的纳什均衡
  - 这说明存在比S_i更优的策略S_i'被删除，而S_i'被删除说明有更优的S_i''留下了
  - 这说明S_i''比S_i优，矛盾！

- 因此淘汰严格非优策略可以作为搜索纳什均衡的一种预处理方法

弱非优策略

- 对于参与人i，称S_i是弱非优策略，若存在S_i'使得对于任意其他参与人的任意策略S_i'收益不小于S_i，并有至少一种情况严格更大，即对于任意策略组S有

$$u_i(S_i';S_{-i}) \geq u_i(S_i;S_{-i})$$

且存在一个S有

$$u_i(S_i';S_{-i}) > u_i(S_i;S_{-i})$$

- 猎鹿-猎兔改：猎鹿收益=猎兔
- 此时猎鹿是弱非优，但包含在均衡里，不能删除！

	B猎鹿	B猎兔
A猎鹿	3，3	0，3
A猎兔	3，0	3，3

其他均衡的定义

帕累托 (Pareto) 最优

- 纳什均衡未必是对于群体/所有人的最优选择

> 称一组策略是帕累托最优的， 若不存在其他策略，使得所有参与者收益至少与现在一样高，并至少有一个参与者得到严格更高回报

- 如果现在采用非帕累托最优策略：可以换成帕累托最优，使得无人受损，而且还有人严格受益

帕累托的达成

- 达到帕累托通常需要首先就此种选择达成“有约束力的协议”
  - 否则，当帕累托最优不是纳什均衡时，会有人采用更利己但损害大家的策略

	B报告	B考试
A报告	90，90	86，92
A考试	92，86	88，88

	乙抵赖	乙坦白
甲抵赖	-1，-1	-10，0
甲坦白	0，-10	-4，-4

- 在考试-报告博弈中，(考试，考试)不是帕累托最优
  - 但是如果没有“有约束力的协议”，帕累托最优是无法达成的
- 考试-报告博弈还有其他帕累托最优：(报告，考试)，(考试，报告)
- 考试-报告博弈/囚徒困境，是纳什均衡为唯一非帕累托最优的策略

更强的定义：社会最优 (socially optimal)

称一组策略是社会福利最大化 (social welfare maximizer)或称社会最优，如果该策略使所有参与人的收益和最大化

- 社会最优推出帕累托最优
- 考试-报告博弈的社会最优是180
  - 但是不同人的收益加和未必总是有意义的
- 社会最优与纳什均衡无直接对应关系
  - 报告-考试博弈1中社会最优不是纳什均衡
  - 在博弈2中却是纳什均衡

	B报告	B考试
A报告	90，90	86，92
A考试	92，86	88，88

报告-考试博弈1

	B报告	B考试
A报告	98，98	94，96
A考试	96，94	92，92

报告-考试博弈2

进化博弈论

- 进化论：适应性强的个体繁衍更多后代，对应的基因频率增大
- 博弈观点：性状没有绝对的好坏；好坏是相对于其他个体/环境互动

- 有一类甲虫，有一种突变使得身体显著增大：大甲虫 vs 小甲虫
  - 大甲虫：需要更多食物才能存活（负面变异）
- 考虑生物之间的相互作用：
  - 当与其他甲虫争夺食物时，大甲虫产生优势
  - 若参与争夺的两只甲虫大小相等，则平分食物；否则大甲虫得到大多数食物
  - 大甲虫得到同样食物，对健康收益较少（新陈代谢高）

	小	大
小	5，5	1，8
大	8，1	3，3

- 与先前博弈的区别：甲虫无法“选择”策略；策略是写在基因里的
- 进化稳定策略（evolutionarily stable strategy）
  - 一种由基因决定的策略，一旦在一个种群中盛行，则“倾向于保持下去”
  - 一个策略是进化稳定的，若整个种群采取该策略时，任何采取不同策略的小规模入侵群体在经过足够多代遗传后会最终消亡

进化稳定策略

“适应力”是指一个个体与另一个随机个体争斗后得到的期望收益

称策略T以x程度入侵策略S，若此时x比例个体使用T策略，剩下1-x用S策略

称策略S是进化稳定的，若存在正数y，使任意策略T以x ($x < y$)程度侵略S时所有S个体的适应力严格大于T个体

- 策略“小”是进化稳定的吗?
  - 验证定义：考虑一个正数x，1-x比例采用小，剩下x比例采用大（侵略者）
  - “小”甲虫对于一个随机个体的期望收益是多少?

$$5(1-x)+1\cdot x=5-4x$$

  - “大”甲虫对于一个随机个体的期望收益是多少?

$$8(1-x)+3x=8-5x$$

  - 当x < 3时大甲虫适应力更好，故不存在一个y使任何x < y都有小甲虫适应力强
  - “小”不是进化稳定的!

	小	大
小	5，5	1，8
大	8，1	3，3

- “大”的进化稳定性?
  - 验证定义：考虑一个正数x，1-x比例采用大，剩下x比例采用小（侵略者）
  - “小”甲虫对于一个随机个体的期望收益是多少?

$$5x + 1 \cdot (1 - x) = 4x + 1$$

  - “大”甲虫对于一个随机个体的既往收益是多少?

$$8x + 3(1 - x) = 3 + 5x$$

  - 对任何x > 0都有大甲虫适应力更佳，所以“大”确实是进化稳定的

	小	大
小	5，5	1，8
大	8，1	3，3

大甲虫进化稳定的直观解释

- 考虑大甲虫刚刚“入侵”小甲虫种群（刚刚产生突变，x很小）
  - 起初大甲虫很少遇见其他大甲虫（因为考虑随机遭遇）
  - 遇小甲虫有极大优势
  - 大甲虫开始蔓延，打破小甲虫的进化稳定性
- 相应的，若小甲虫入侵大甲虫：
  - 经常碰上大甲虫产生极大劣势，因此大甲虫维持进化稳定

其他视角的解释

- 与囚徒困境的联系：
  - 大甲虫产生后会迅速取代小甲虫，但最后的结局是大甲虫占多数后，竞争激烈，只能得到3的收益
  - 大甲虫 -> 军备竞赛/兴奋剂
  - 囚徒困境描述的是结果，但是早期采用新军备的人确实是获利的；囚徒困境的产生也是不可逆的
- 自然选择的视角：
  - 自然选择通常是向着增大适应性的方向进行的
  - 大甲虫被选择出来，但是适应性却降低了？
  - “环境”的恶化：其他大甲虫的竞争也是一种环境

	小	大
小	5，5	1，8
大	8，1	3，3

一些真实例子

- 树的高度 (Falster, Westoby, 2003; Iwasa, Cohen, Leon, 1985)
  - 如果树都等高，那么阳光资源可以平分；增高突变会有优势
  - 不会无限增高：“足够高”；再高也没有更多太阳
- 噬菌体（病毒）的变异 (Turner, Chao 1999)
  - 最初病毒：类型A
  - 变种：类型B可以在噬菌体A存在的情况下利用A产生的化学物质，得到更强的“适应性”，但是若只有B则“适应性”降低
  - 囚徒困境：都是A适应性更强，但是B适应性更弱，但是B会取代A
  - 基于实验观察得出的收益矩阵：

	A类	B类
A类	1.00, 1.00	0.65, 1.99
B类	1.99, 0.65	0.83, 0.83

进化稳定的一般条件

- 若S是进化稳定的：
  - 1-x比例用S，x比例用T
  - 用S的期望收益：$a(1-x)+bx$；T的$c(1-x)+dx$
  - 对任意x > 0有$a(1-x)+bx > c(1-x)+dx$，推出$a > c$
  - 若$a = c$，那么需要$b > d$

	S	T
S	a, a	b, c
T	c, b	d, d

在2参与人2策略对称博弈中，S是进化稳定的充要条件是i) $a > c$或者ii) $a = c$且$b > d$

- 直观解释：
  - a > c对应S遇到T时需要有更好的收益
  - a = c时S想稳定则必须保证S与T接触时比T与T接触时有优势

进化稳定与纳什均衡的联系

在2参与人2策略对称博弈中，S是进化稳定的充要条件是i) $a > c$或者ii) $a = c$且$b > d$

- 关系：若S是进化稳定的，则(S, S)是纳什均衡（对应$a \geq c$）
  - 反过来不对：

	B猎鹿	B猎兔
A猎鹿	4，4	0，~~3~~4
A猎兔	~~3~~4，0	3，3

	S	T
S	a, a	b, c
T	c, b	d, d

- 严格纳什均衡：(S, S)互为唯一的严格最优应对（推出a > c）
  - 关系：(S, S)是严格纳什均衡的，推出S进化稳定的
- 进化稳定：在群体上考察，个体不需要“全信息”
  - 可以“随大流”：后来者观察前者的成功经验来效仿（而不知道额外信息）

进化稳定的混合策略

- 与纯策略纳什均衡一样，可能出现无纯进化稳定策略的情况

	B：D	B：H
A：D	3，3	1，5
A：H	5，1	0，0

	S	T
S	a,a	b,c
T	c,b	d,d

- (D, D)和(H, H)都不是纳什均衡，推出D和H都不是进化稳定的
  - 少量H在大量D的群体里占优；而少量D在大量H里可以保持隔离，坐观H互殴
- 混合策略：个体以某概率p选择S以1-p概率选择T
- p策略针对q策略的期望收益：

$$V(p,q) = pqa + p(1-q)b + (1-p)qc + (1-p)(1-q)d$$

称混合策略p是进化稳定的，若存在y > 0使任何q策略的x程度的入侵 (x < y)后，p策略个体的适应力严格好于q策略的。

混合策略的进化稳定性

称混合策略p是进化稳定的，若存在y > 0使任何q策略的x程度的入侵 (x < y)后，p策略个体的适应力严格好于q策略的。

- 适应力：个体的适应力是相对于群体的随机个体的期望收益
  - 两个随机性：i) 个体策略的随机性p（或q），ii) 随机个体（x vs 1 – x）
- 对y > 0，x < y，p是进化稳定的仅当
$$(1-x)V(p,p)+xV(p,q)>(1-x)V(q,p)+xV(q,q)$$
- 观察：若p是进化稳定的，那么当x趋近于0时，$V(p,p) \geq V(p,q)$
- 结论：说明p是p的最佳应对，所以(p, p)是纳什均衡的

	S	T
S	a, a	b, c
T	c, b	d, d

举例

- 鹰鸽博弈的混合进化稳定策略

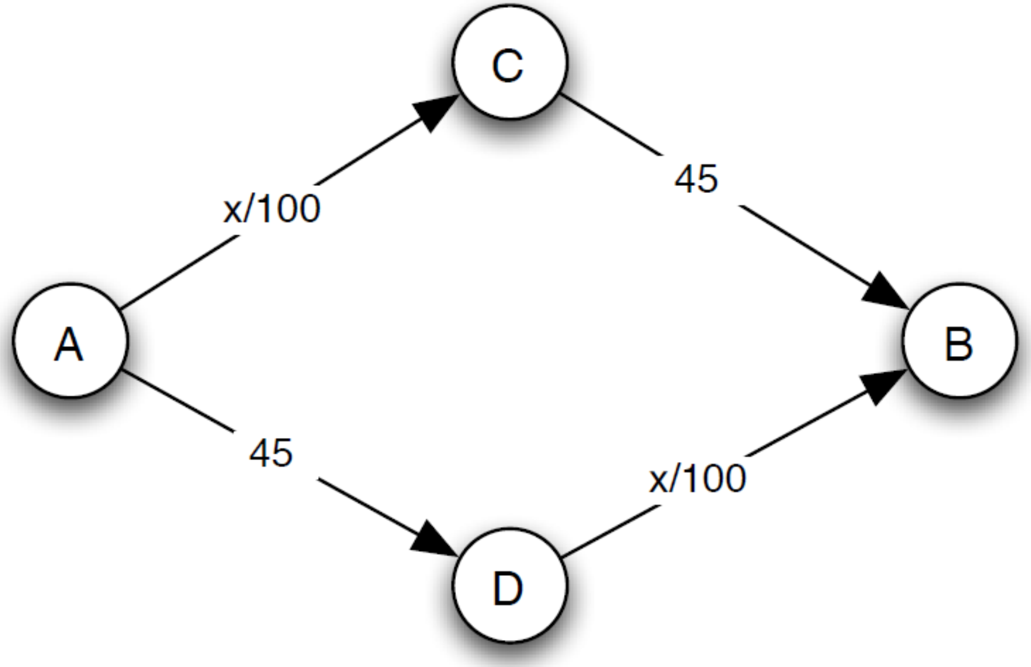
	B：D	B：H
A：D	3，3	1，5
A：H	5，1	0，0

- 纳什均衡：$\left(\frac{1}{3},\frac{1}{3}\right)$；对p = 1 / 3，所有的q都有一样的收益
- 因此V(p, p) = V(p, q)；p进化稳定条件：
$$(1-x)V(p,p)+xV(p,q)>(1-x)V(q,p)+xV(q,q)$$
- 需要检查$V(p,q)>V(q,q)$
  - $V(p,q)=4q+1/3\,,V(q,q)=6q-3q^2$
  - $V(p,q)-V(q,q)=\frac{1}{3}(3q-1)^2>0$

博弈论应用：网络流量的博弈论模型

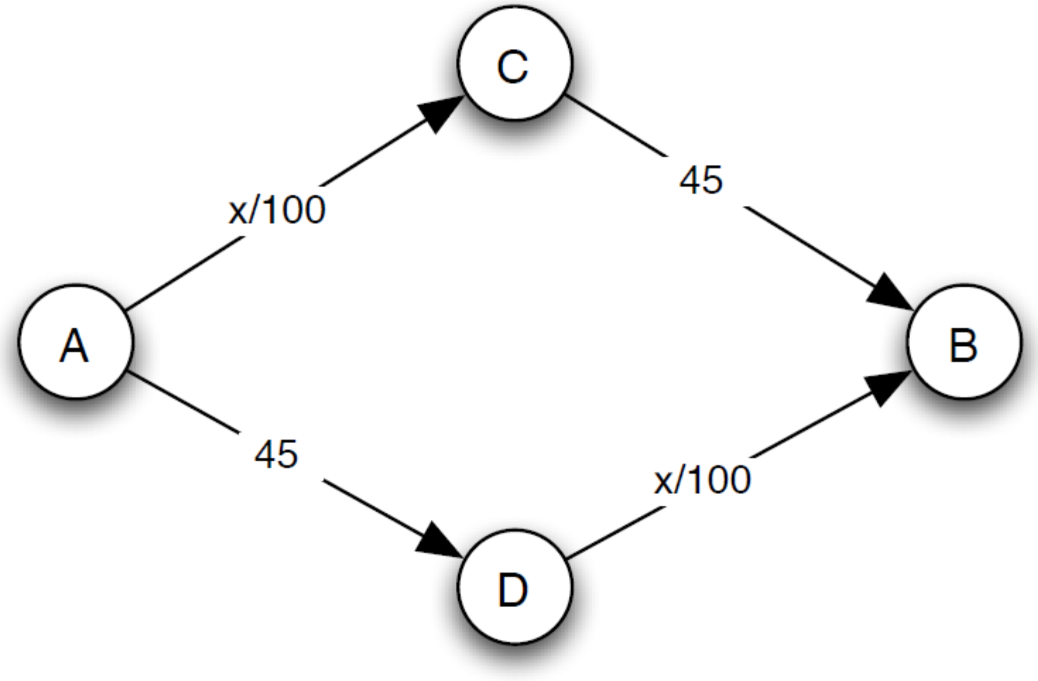
公路网络与通行时间

- 用有向图表示公路网络
  - 有向边是公路，节点是公路出入口；边上标示通行时间，是车数量x的函数
  - 有A、B两个特殊节点：所有人的目的都是从A开车到B



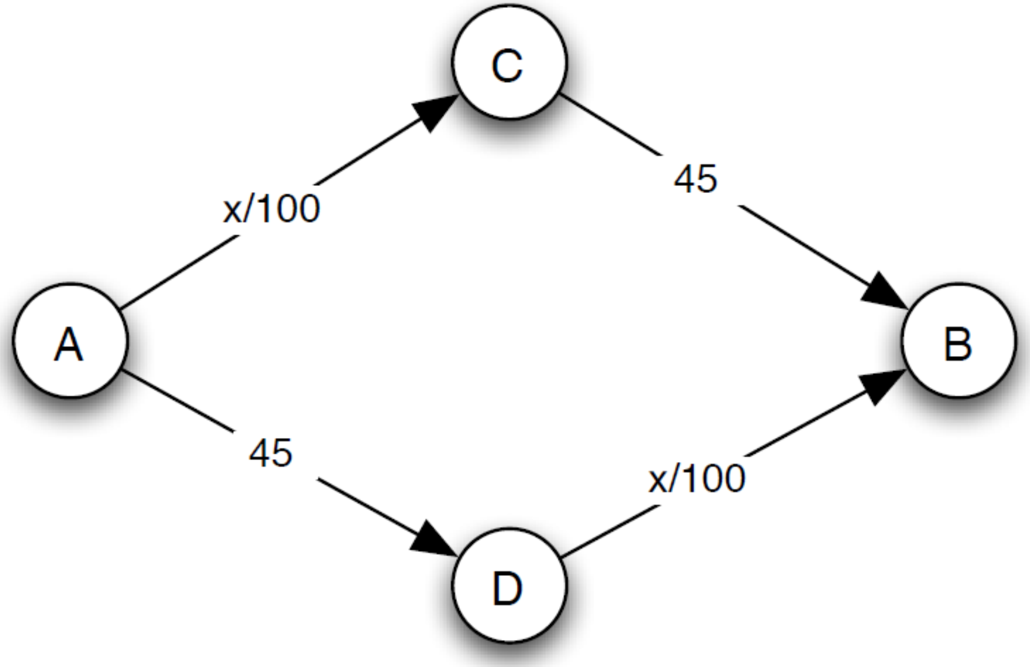
4000辆车：都走一边和平均走两边各需要多长时间?

- 每个司机都是一个博弈的参与者，决策应该走哪边（C vs D）
- 建模收益：每个人的收益是所花时间取负



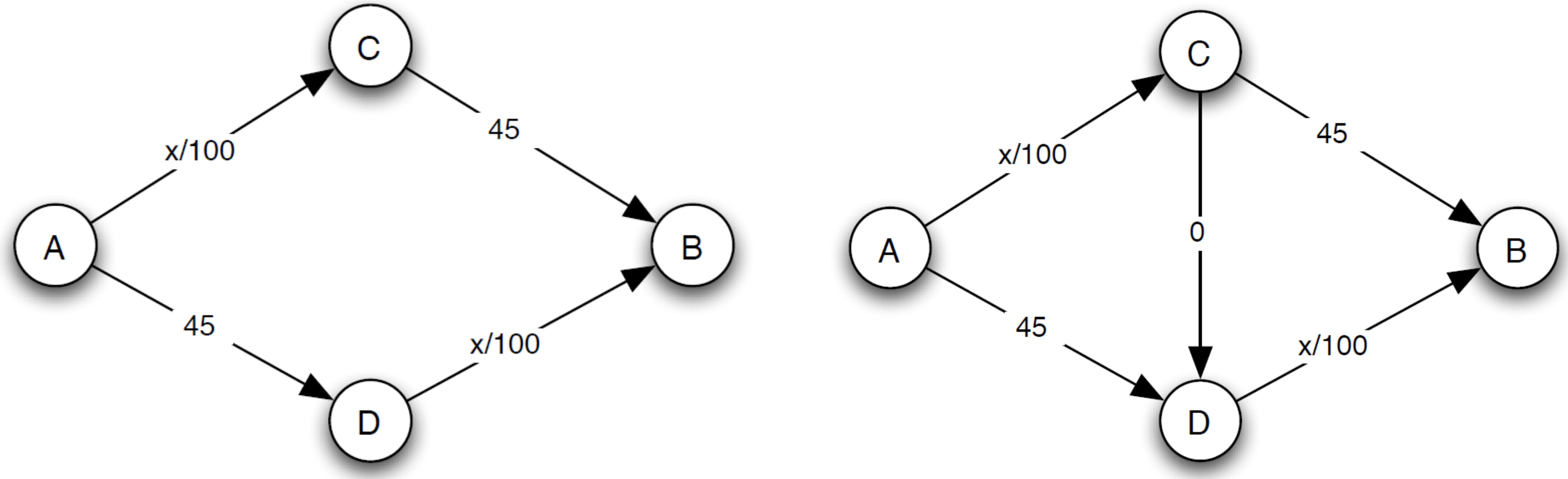
是否有占优策略?

- 无占优策略
- 有纳什均衡：有相等数量的司机选择C和D（2000 vs 2000），用时65
  - 所有的纳什均衡都是这种“平衡”分配的情况（why?）



增加资源

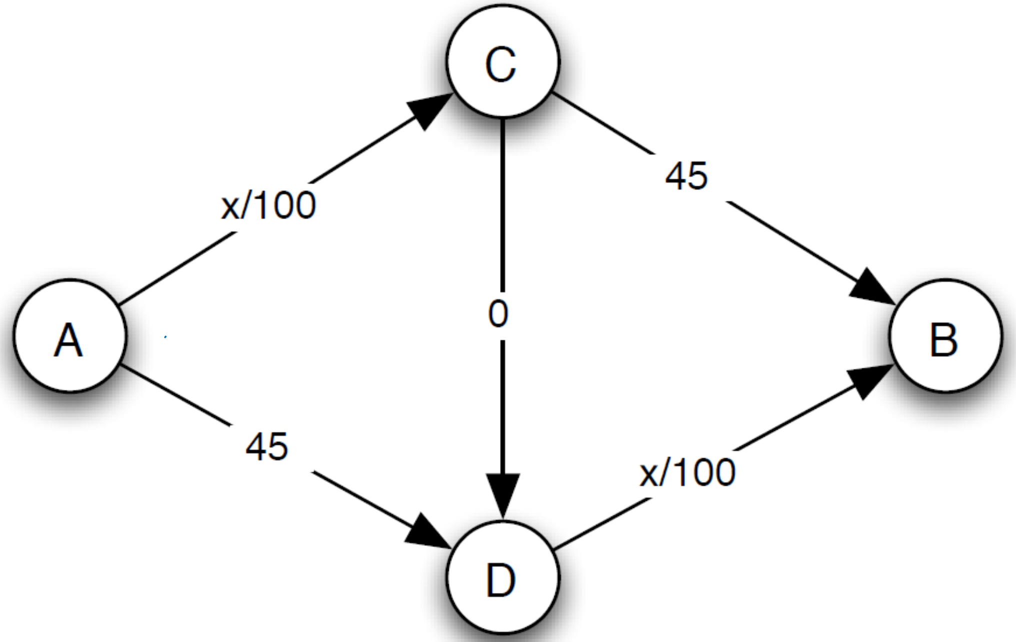
- 在C和D之间修建一条“高速路”？



- 多了这条路，用时应该更短了才对?

布雷斯悖论 (Braess's Paradox)

- 唯一纳什均衡：所有人都走CD，总用时80，比之前的65要坏！



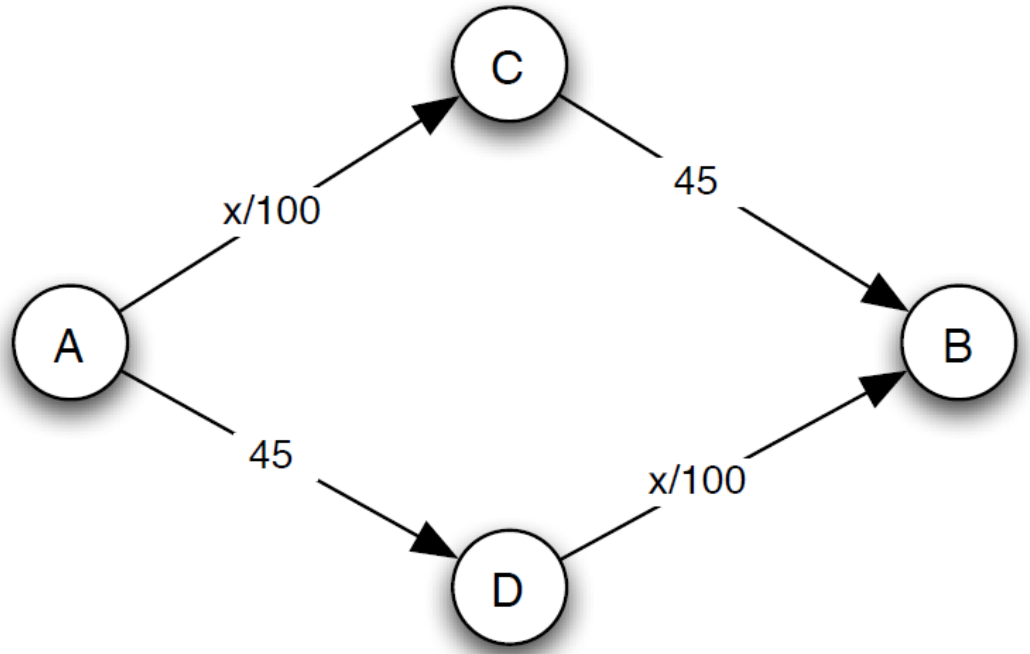
- CD边具有“虹吸”效应
  - 强迫所有人都需要选CD

- 这种增加资源反而劣化收益的现象最早由Dietrich Braess 1968提出
- 韩国首尔曾将6车道公路改为公园，却改善了交通 (Baker 2009)

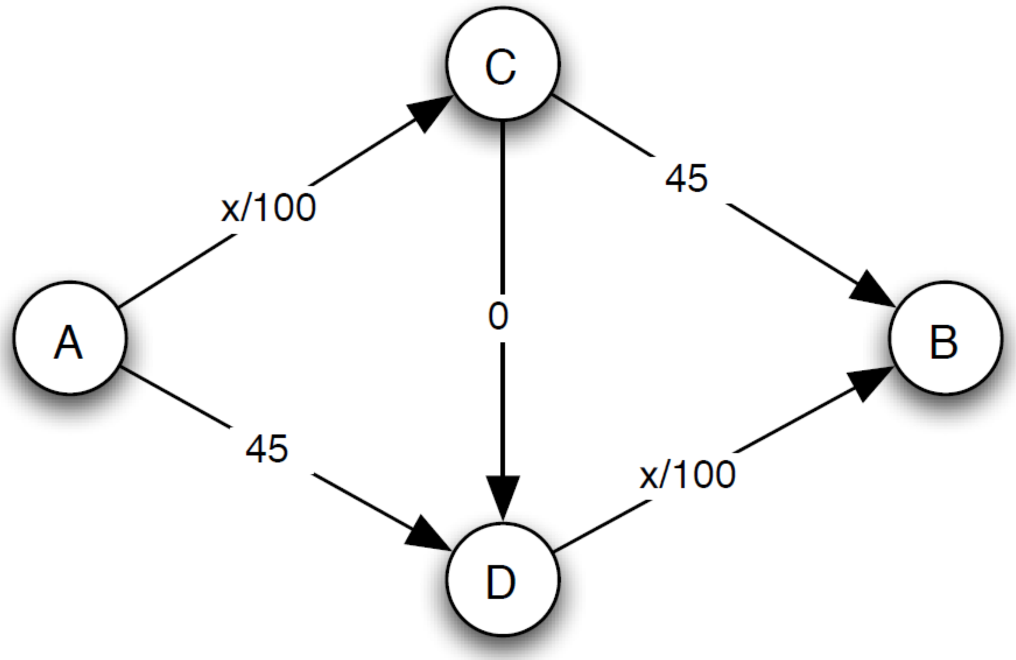
布雷斯悖论中纳什均衡的社会成本

称一组策略是社会福利最大化 (social welfare maximizer)或称社会最优，如果该策略使所有参与人的收益和最大化

- Gap：纳什均衡的收益除以社会最优收益

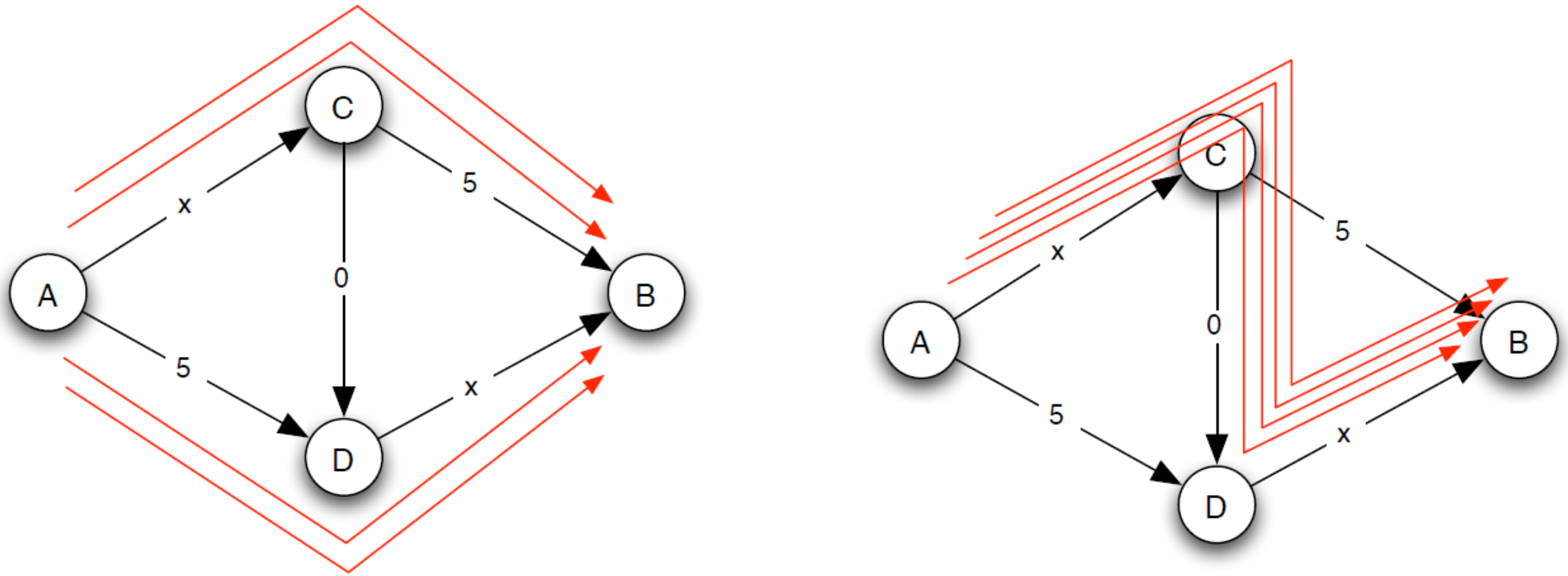


均衡收益是社会最优



均衡收益80，社会最优65，gap $\frac{80}{65} \approx 1.23$

- 考虑一般的有向图网络，每个司机可以有不同的起点和终点
- 每条边e有一个时间函数$T_e(x)$代表x辆车同时选择该路导致的通行时间
- 现假定$T_e(x) = a_e x + b_e$都是线性函数
- “流量模式”（也就是策略）是指每个司机对于路径的选择
- 某个流量模式的社会成本定义为所有司机的通行时间的总和



(a) *The social optimum.* (b) *The Nash equilibrium.*

Figure 8.4: A version of Braess's Paradox: In the socially optimal traffic pattern (on the left), the social cost is 28, while in the unique Nash equilibrium (on the right), the social cost is 32.

布雷斯悖论的社会代价分析

首先：如何找到纳什均衡?

- 该博弈的纳什均衡一定存在（纳什证明对于一般有限博弈均存在均衡）
- 一般而言计算纳什均衡是困难的；但是对于此种特殊博弈有特别算法

使用local search找到纳什均衡

1. 从某个（任意）流量模式A开始
2. 若P均衡则停止；否则可找到司机i，使i切换路线可以严格提高收益
3. 令i切换到新的路径，得到新的流量模式A’，重复步骤2-3

从某个（任意）解开始，每步局部调整直到不能改进的方法叫做local search

Local search的分析

- 正确性：如果停止，那么找到的一定是纳什均衡（why?）
- 难点：分析运行时间，即为什么一定会终止?
  - 为什么不可能在两个解之间“反复横跳”，一直运行下去?
- 一般方法：找到一种度量progress的函数，称为“势能函数”
  - 证明势能函数有下界，并证明local search每步都会严格降低势能
  - 问题：选用社会成本作为势能函数合理吗?

某个流量模式的社会成本定义为所有司机的通行时间的总和

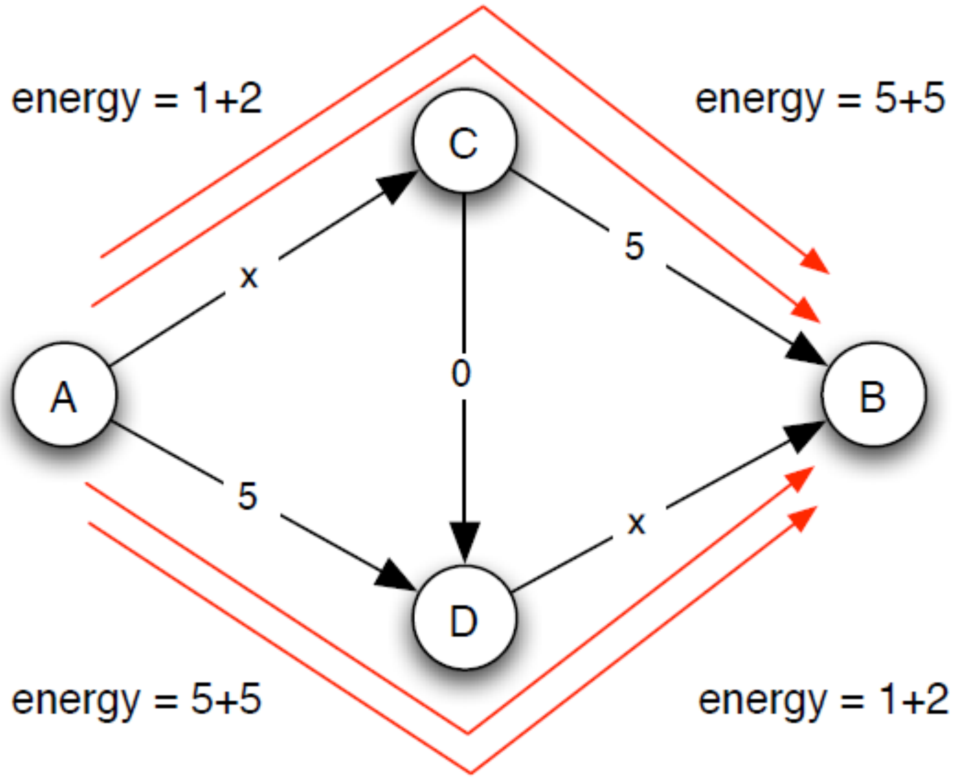
定义势能函数

- 若某条边e现有x辆车，则e的势能定义为

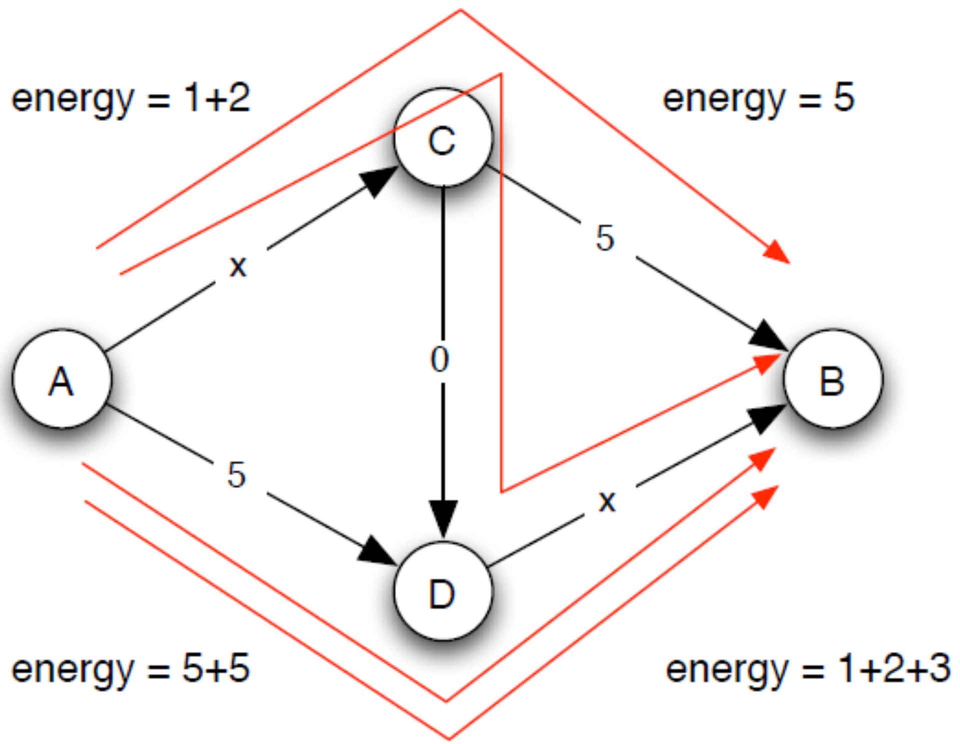
$$\Delta_e(x) \coloneqq T_e(1) + \cdots T_e(x)$$

- 回忆：$T_e(x)$代表e上有x辆车时的通行时间
- 补充定义$\Delta_e(0) = 0$
- 考虑系统的总势能$\sum_e \Delta_e$

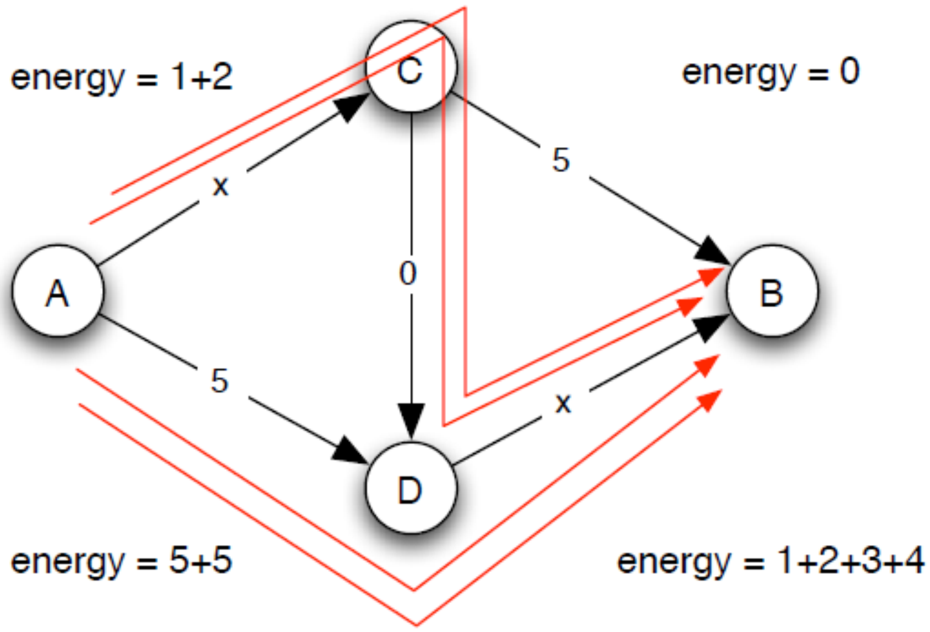
- 怎么理解该势能函数?
  - 一般而言未必有“物理意义”，而更多是来自算法分析中的量/计算步骤
  - 并不是e上通行的总时间，因为总时间是$xT_e(x)$
  - 这里类似于某种“累计”时间



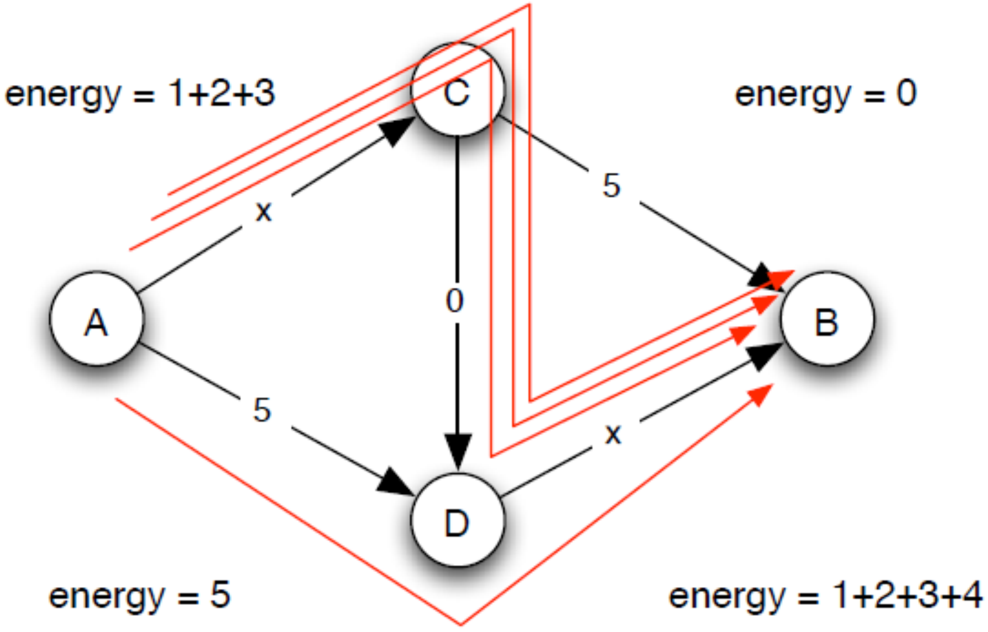
初始流量模式，势能=26



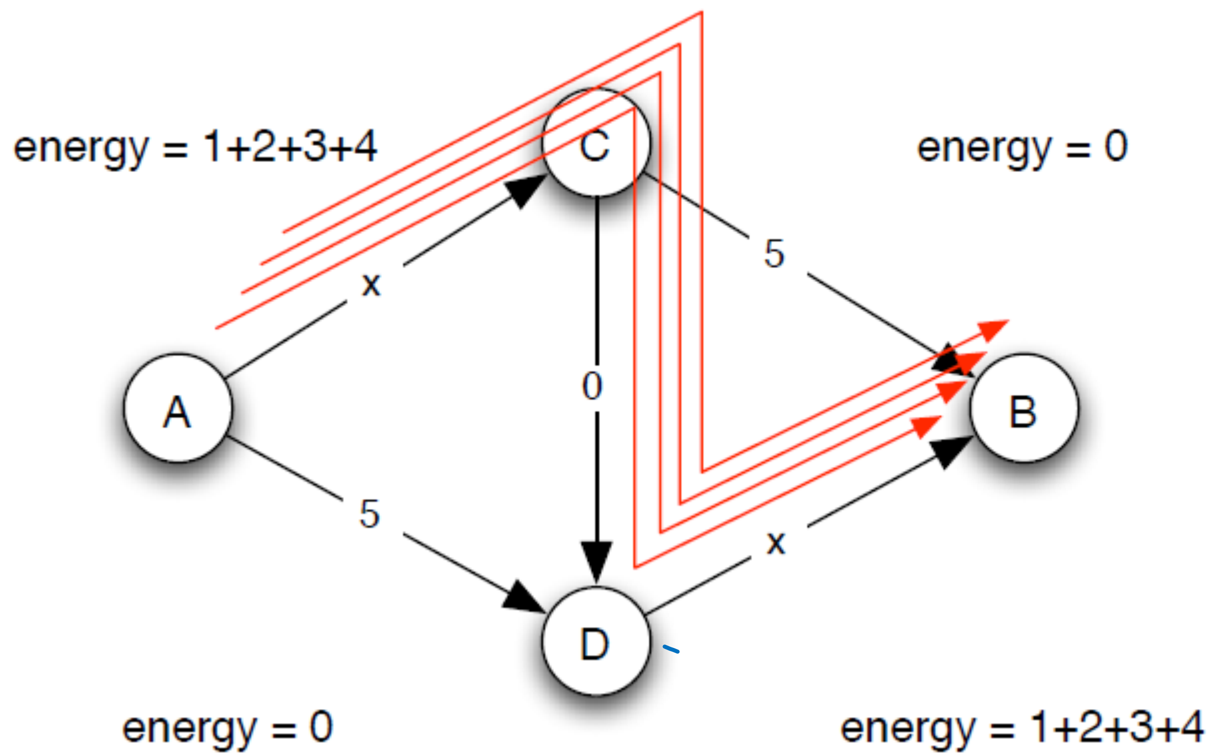
一步local search，势能=24



两步local search，势能=23



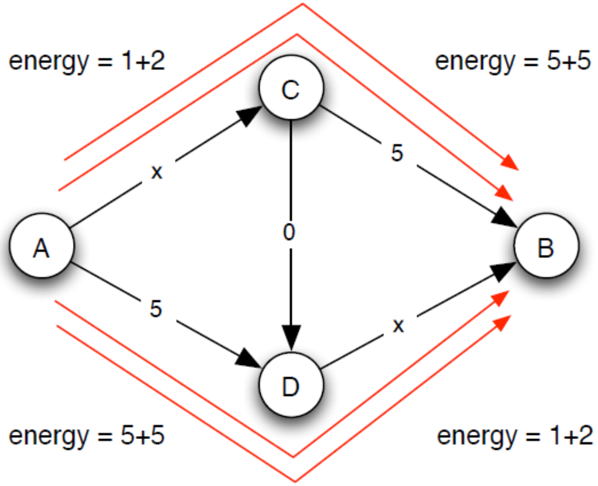
三步local search，势能=21

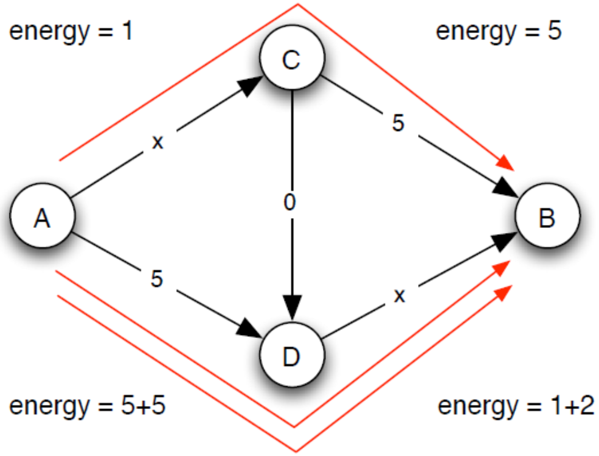


四步local search，达到纳什均衡，势能=20

分析势能变化

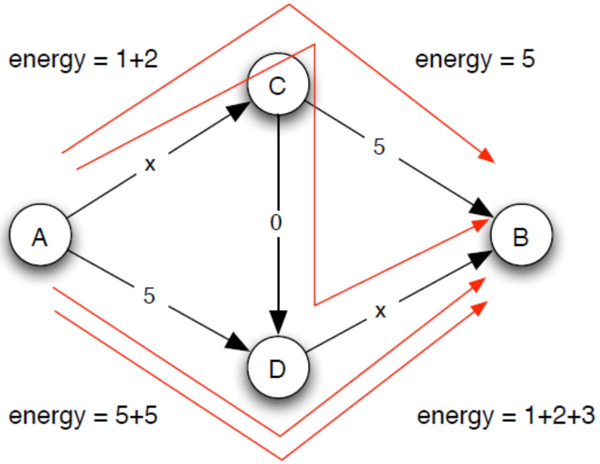
- 假设司机i从A路径换到了B路径
- 理解成先离开整个系统（离开A），然后重新加入到B
- 离开A：
  - 在A上的每条边e，设i离开前有x_e辆车
  - 离开后总势能减小$\sum_{e \in A} T_e(x_e)$，这个减少量恰是i离开前的总用时
- 加入B：
  - 在B上每条边e，设i加入后有x_e'辆车
  - 加入后总势能增加$\sum_{e \in B} T_e(x_e')$，这个增加量恰是i加入后的总用时
- 总变化恰好等于收益的变化量，一定是< 0的（算法更新规则保证的）
- 因此可以推出算法一定会停止





(a) *The potential energy of a traffic pattern not in equilibrium.*

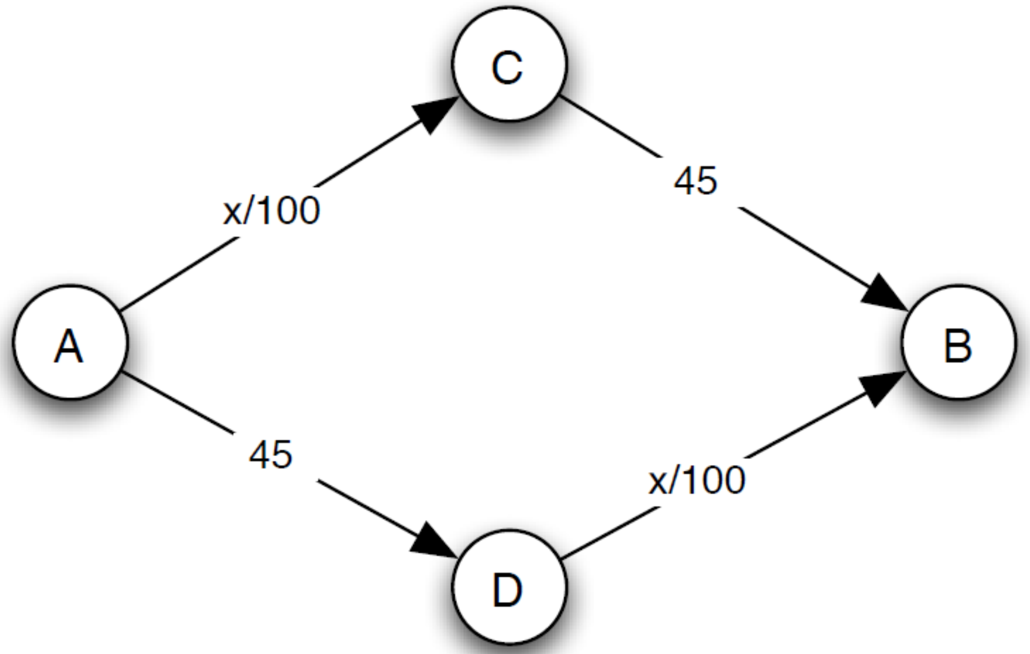
(b) *Potential energy is released when a driver abandons their current path.*



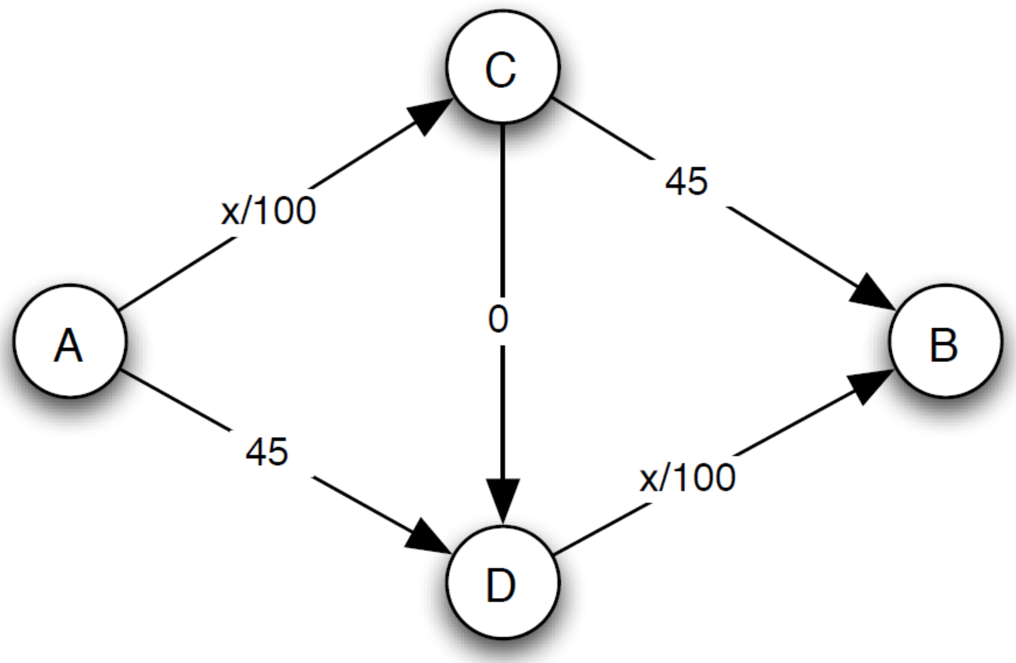
布雷斯悖论中纳什均衡的社会成本

称一组策略是社会福利最大化 (social welfare maximizer)或称社会最优，如果该策略使所有参与人的收益和最大化

- Gap：纳什均衡的收益除以社会最优收益



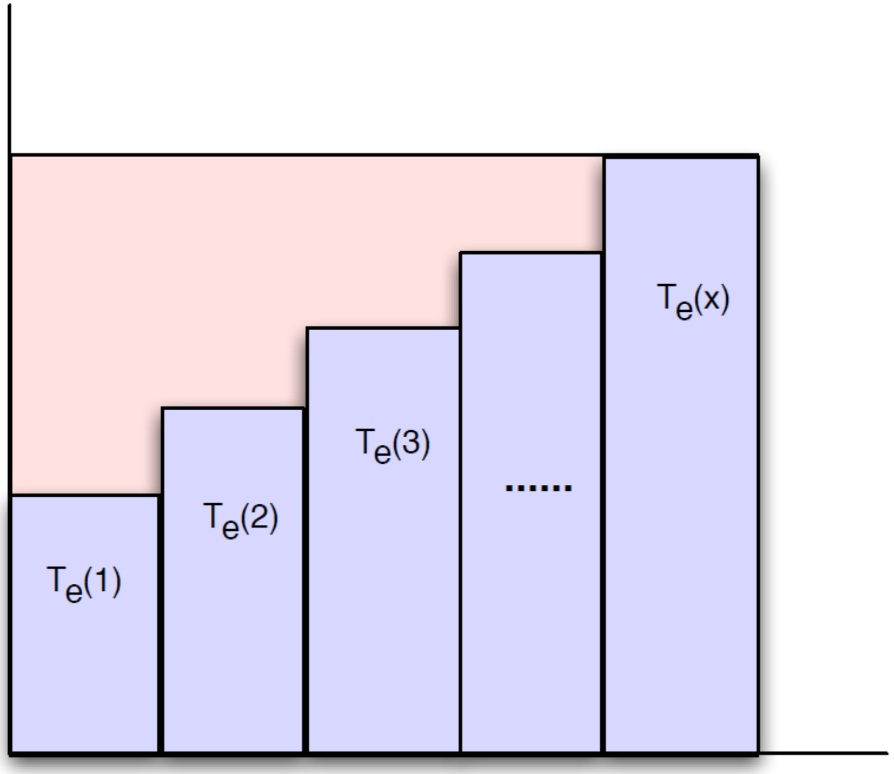
均衡收益是社会最优



均衡收益80，社会最优65，gap $\frac{80}{65} \approx 1.23$

将势能函数与社会代价联系起来

- 考虑任意一个流量模式；回忆$T_e(x) = a_e x + b_e$
- 考虑边e，设当前有x_e辆车经过e，则e对于社会代价的贡献是$x_e T_e(x_e)$
- 回忆势能函数：$\Delta_e(x_e) := T_e(1) + \cdots T_e(x_e)$



1. 可以看出对每个e，$\Delta_e(x_e) \leq x_e T_e(x_e)$
2. 同时有$\Delta_e(x_e) \geq 0.5 x_e T_e(x_e)$

- 已证：$0.5\ x_eT_e(x_e) \le \Delta_e(x_e) \le x_eT_e(x_e)$
- SC(Z)代表Z上的social cost（社会代价）
- 对e求和，可得对于流量模式Z，有

$$0.5\ \mathrm{SC}(Z) \le \Delta(Z) \le \mathrm{SC}(Z)$$

- 取最优社会代价的Z，然后运行local search，记找到的纳什均衡为Z'

$$\mathrm{SC}(Z') \le 2\Delta(Z') \le 2\Delta(Z) \le 2\mathrm{SC}(Z)$$

- 至此我们证明了线性T_e的纳什均衡的**社会代价至多是2倍**

Roughgarden, Tardos JACM 02

- 考虑了相似的问题：每个人i不只能选择一条路，而是可以选择多条，最后只要自己的给定总流量v_i从某个起点s_i发到了t_i即可
- 如果每条边的代价函数$T_e(x)$是线性的，那么纳什均衡的社会代价至多是4/3倍，并且有例子说明4/3也是必须的
- 证明技术也可以用于我们讨论的情形（司机只能选一条路），并得到4/3，我们看到的2的证明是简化版本

- 一般地，均衡下的社会收益与最优社会收益的gap叫做price of anarchy
  - 研究肇始于[Koutsoupias-Papadimitriou, 1999]