第二讲：完全信息博弈

兴军亮
Junliang.Xing@ia.ac.cn

2018年3月25日

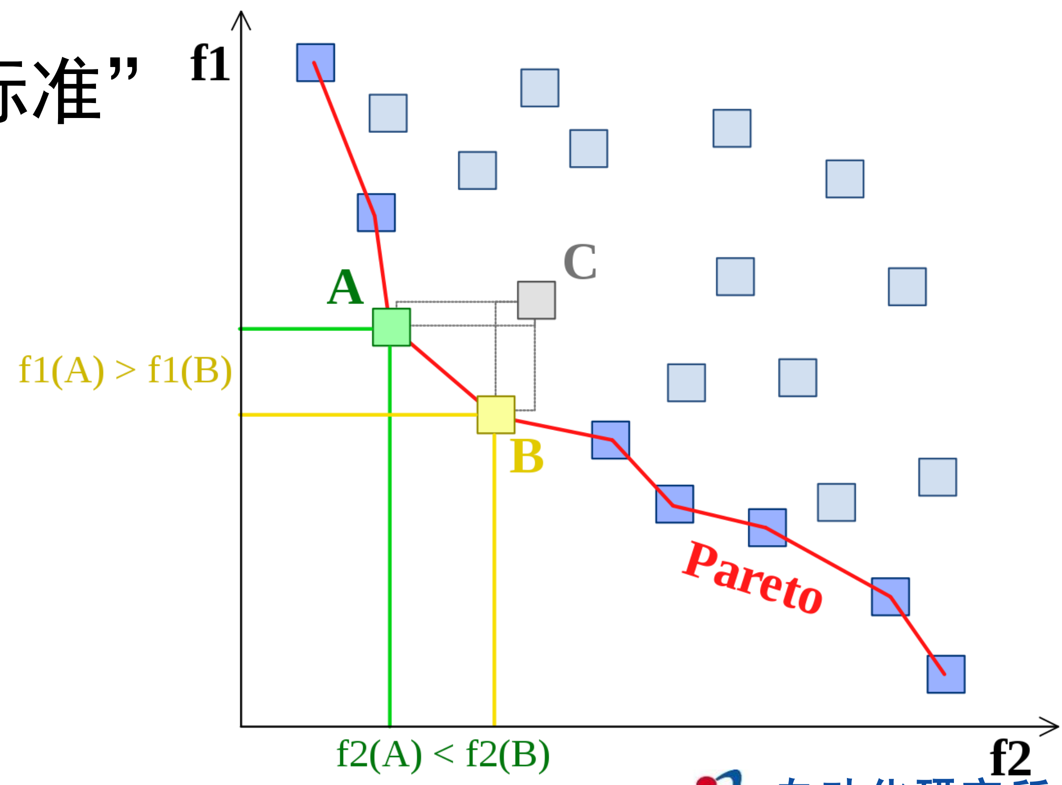
中国科学院大学
University of Chinese of Academy of Sciences
CASIA
自动化研究所
Institute of Automation

本讲提纲

- 极大极小值原理
- 完全信息静态博弈
- 完全信息动态博弈
- 游戏博弈案例探讨

研究博弈的目标：帕雷托最优

- 帕雷托效益/最优（ Pareto Efficiency/Optimality ）
  - 资源分配的一种理想状态：给定固有的一群人和可分配的资源，如果从一种分配状态到另一种状态的变化中，在没有使任何人境况变坏的前提下，使得至少一个人变得更好（帕雷托改善，Pareto Improvement）。
- 只是理想状态的“最低标准”
  - 未达到一定不理想
  - 达到了不一定“最理想”
- 完备且充分竞争的市场
  - 交换最优
  - 生产最优
  - 产品混合最优



二人零和博弈

- 两人博弈中，双方的损益之和为零或一个常数，即一方有所得，另一方必有所失。
  - 帕雷托改善（Pareto Improvement）
  - 帕雷托效益/最优（ Pareto efficiency/optimality ）
  - 非合作博弈，零和的结果出现帕雷托最优
  - 实际案例：分蛋糕、围棋、赌博、双寡头市场竞争等
- 思考：如何分蛋糕？
- 另一个示例：红蓝双方博弈
  - 右图策略式表示有什么特点?
  - 红方动作：1和2
  - 蓝方动作：A、B和C
  - 最终双方的胜率是什么?

红方 \ 蓝方	A	B	C
1	3, -3	-1, 1	2,-2
2	1, -1	2, -2	-2,2

博弈的核心是搜索算法！

- 极大极小搜索：搜索策略保证己方收益最大，对方最小！
  - 最大化最差情况（最小收益）下的收益（maximin）
  - 最小化最差情况（最大损失）下的损失（minimax）
- 极大极小值（maximin value）：在不考虑对方行动的情况下，己方能够获得的最大收益值；即对方知道己方行动情况下，能够迫使对方获得的最小收益值。
- 形式化：

$$v^* = \max_{a_i} \min_{a_{-i}} v_i(a_i, a_{-i})$$

- 适用性：静态博弈、动态博弈、单次博弈、多次博弈等

极大极小搜索示例

- 同样以下图博弈为例：红方的极大极小值是什么？

蓝方 / 红方	A	B	C
1	3, -3	-1, 1	2,-2
2	1, -1	2, -2	-2,2

- 一次博弈的结果是什么？
- 多次博弈的结果是什么？
- 动态博弈的结果是什么？

极大极小搜索的理论保证

- 极大极小值原理：亦称为博弈论基本定理，是关于如下最大最小不等式等号成立条件的定理：

  若$X \subseteq \mathbb{R}^n, Y \subseteq \mathbb{R}^m$为紧致凸集，$f: X \times Y \to \mathbb{R}$为连续的的凸-凹函数，则：

$$\max_{y \in Y} \min_{x \in X} f(x, y) = \min_{x \in X} \max_{y \in Y} f(x, y)$$

- 最早证明：1928年，冯·诺依曼
- 证明方法：利用初等概率和拓扑学的工具
- 后续证明：凸性的论证和支撑超平面的概念；完全代数方法；不动点理论或迭代程序；凸集理论等。

原理的策略式表示和证明

- 纯策略（pure strategy）两人博弈表示：矩阵对策$G = \{S_1, S_2, R\}$，己方策略集$S_1 = \{\alpha_1, \dots, \alpha_m\}$，对方策略集$S_2 = \{\beta_1, \dots, \beta_n\}$, 支付函数$R = \left(r_{ij}\right)_{m\times n}$
- 混合策略（mixed strategy）扩充：$G^* = \{S_1^*, S_2^*, E\}$，己方混合策略集$S_1^* = \{\mathbf{x} \in \mathbb{R}^m | \sum_{i=1}^n x_i = 1, x_i \geq 0\}$, 对方混合策略集$S_2^* = \left\{\mathbf{y} \in \mathbb{R}^n | \sum_{j=1}^n y_j = 1, y_j \geq 0\right\}$，支付函数$E(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\top A\mathbf{y}$。
- 最优混合策略：$G^* = \{S_1^*, S_2^*, E\}$是$G = \{S_1, S_2, R\}$的混合扩充，如果$\max_{\mathbf{x}\in S_1^*} \min_{\mathbf{y}\in S_2^*} E(\mathbf{x}, \mathbf{y}) = \min_{\mathbf{y}\in S_2^*} \max_{\mathbf{x}\in S_1^*} E(\mathbf{x}, \mathbf{y})$，则称混合局势$(\mathbf{x}^*, \mathbf{y}^*)$为双方的最优混合策略。

原理的策略式表示和证明

- 策略式表示下的极大极小值定理三种等价形式
  - 形式1：对任一纯策略博弈$G=\{S_1,S_2,R\}$，一定存在混合策略意义下的解；
  - 形式2：对任意A，都有$\max_{\mathbf{x}}\min_{y}\mathbf{x}^{\top}A\mathbf{y}=\min_{\mathbf{y}}\max_{\mathbf{x}}\mathbf{x}^{\top}A\mathbf{y}$；
  - 形式3：$\max_{\mathbf{x}\in S_1^*}\min_{1\le j\le n}\sum_{i=1}^{m}r_{ij}x_i=\min_{\mathbf{y}\in S_2^*}\max_{1\le i\le m}\sum_{j=1}^{n}r_{ij}y_j$；
- 极大极小值定理的集中几种典型证明方法：
  - 用数学归纳法证明
  - 用凸集分离定理证明
  - 用单纯形理论证明
  - 用Brouwer不动点定理证明

罗长童, 最小最大值定理的几种典型证法，吉林建筑工程学院学报，vol. 20, no. 3, pp. 54-58, 2003.

极大极小值定理的一般形式

- 推广意义下的极大极小值定理是研究在特定条件下的如下不等式中等号成立的条件：

$$f^* := \inf_{x\in X} \sup_{y\in Y} f(x,y) \leq \sup_{y\in Y} \inf_{x\in X} g(x,y) =: g_*$$

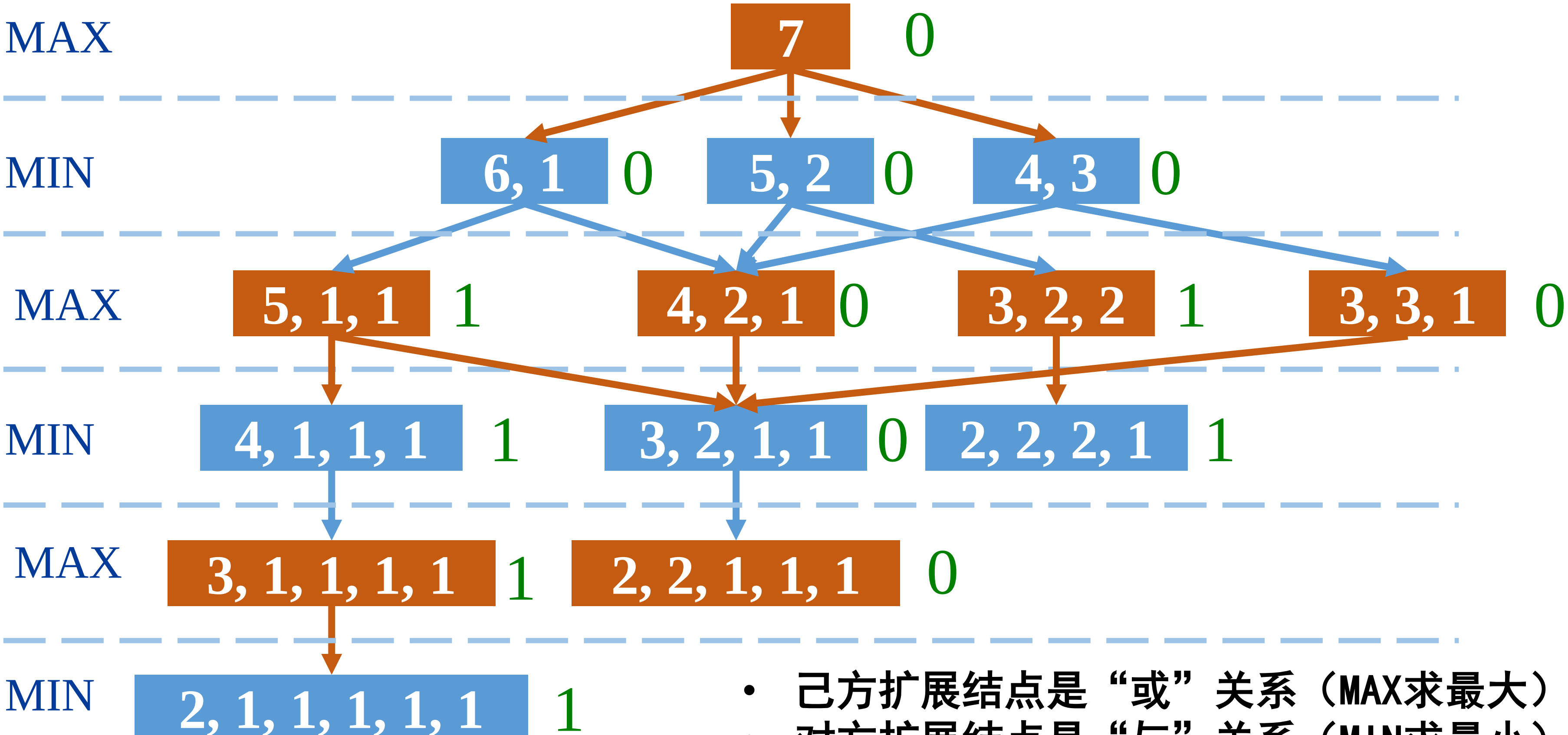
- 一个极大极小定理一般涉及三个假设条件
  - 集合X和Y的空间结构
  - 函数$f(x,y)$和$g(x,y)$的连续性
  - 函数$f(x,y)$和$g(x,y)$的凹凸性：关键条件
- 极大极小值定理实例
  - Sion's minimax theorem: $\min_{x\in X} \sup_{y\in Y} f(x,y) \leq \sup_{y\in Y} \min_{x\in X} f(x,y)$
  - Parthasarathy's theorem: $\max_{\mu\in M_X} \inf_{\lambda\in A_Y} k(\mu,\lambda) \leq \inf_{\lambda\in A_Y} \max_{\mu\in M_X} k(\mu,\lambda)$

极大极小搜索的应用

- 极大极小搜索策略：在考虑双方对弈若干步之后，利用局面评估函数，从可能步骤中选相对好的一步来走，即在有限的搜索范围内进行搜索求解。
- 局面评估函数：对比赛出现的各个局面进行评估
  - 评估的目的：对后面状态提前考虑，指导下一步走法
  - 评估的方法：对己方有利，函数为正；对对方有利，函数为负
  - 评估的标准：由于对立性，只能选择一方作为评估的标准方
- Max结点和Min结点
  - 己方所处结点称为Max结点，选择分支评估函数最大的子结点
  - 对方所处结点称为Min结点，选择分支评估函数最小的子结点
  - Max结点和Min结点交替出现

极大极小搜索应用示例：分硬币游戏

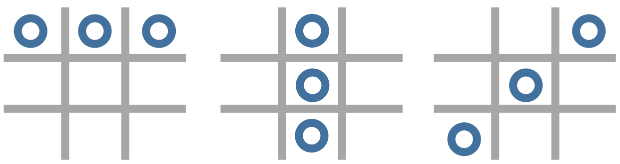
- 分硬币游戏：有7枚硬币，只能将已分好的一堆硬币分成个数不等的两堆，当每堆只有1枚或2枚硬币则不能再分，红蓝双方轮流进行，直到谁不能再分，则为输。

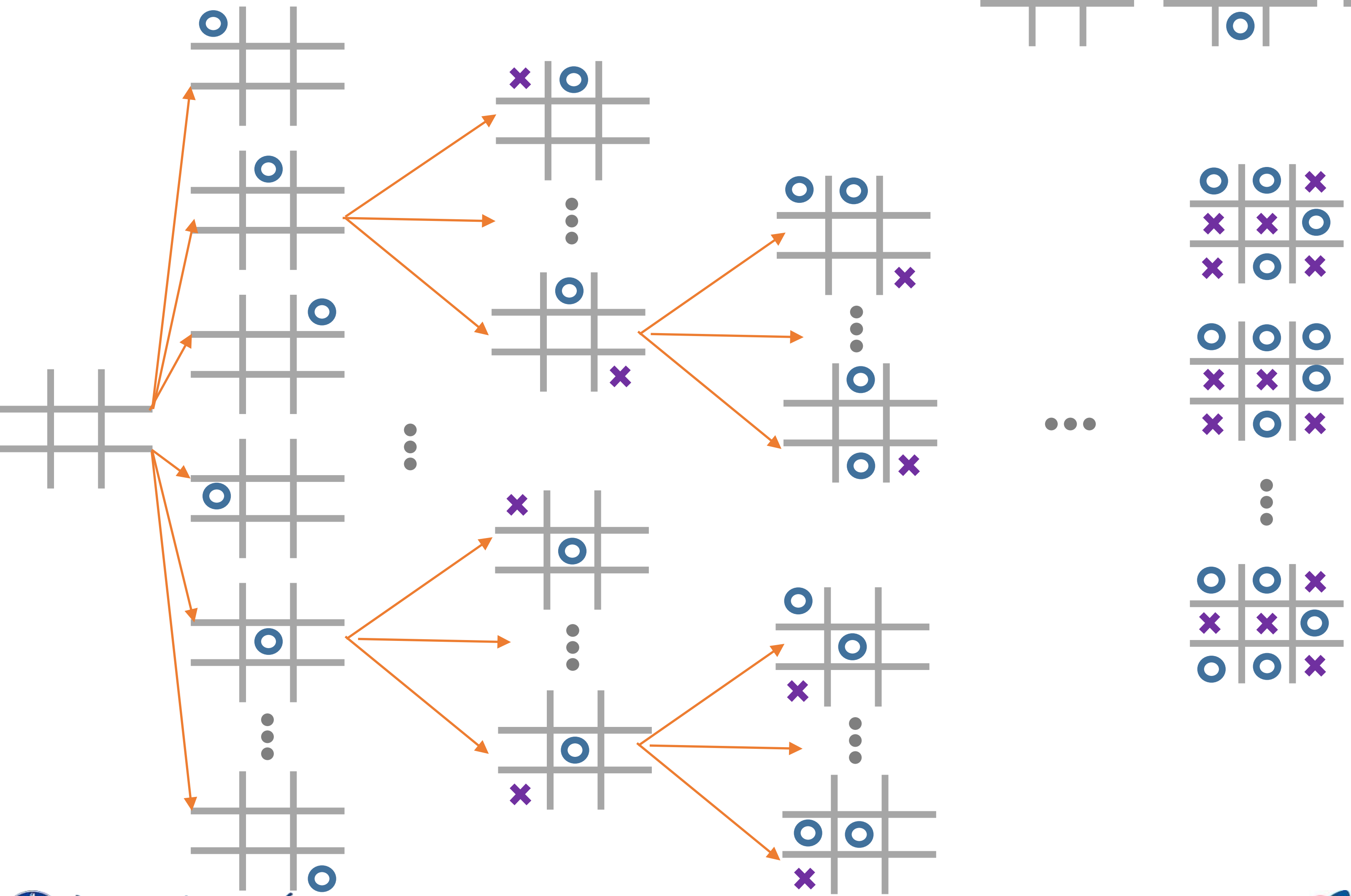


- 己方扩展结点是“或”关系（MAX求最大）
- 对方扩展结点是“与”关系（MIN求最小）

极大极小搜索应用示例：井字棋AI

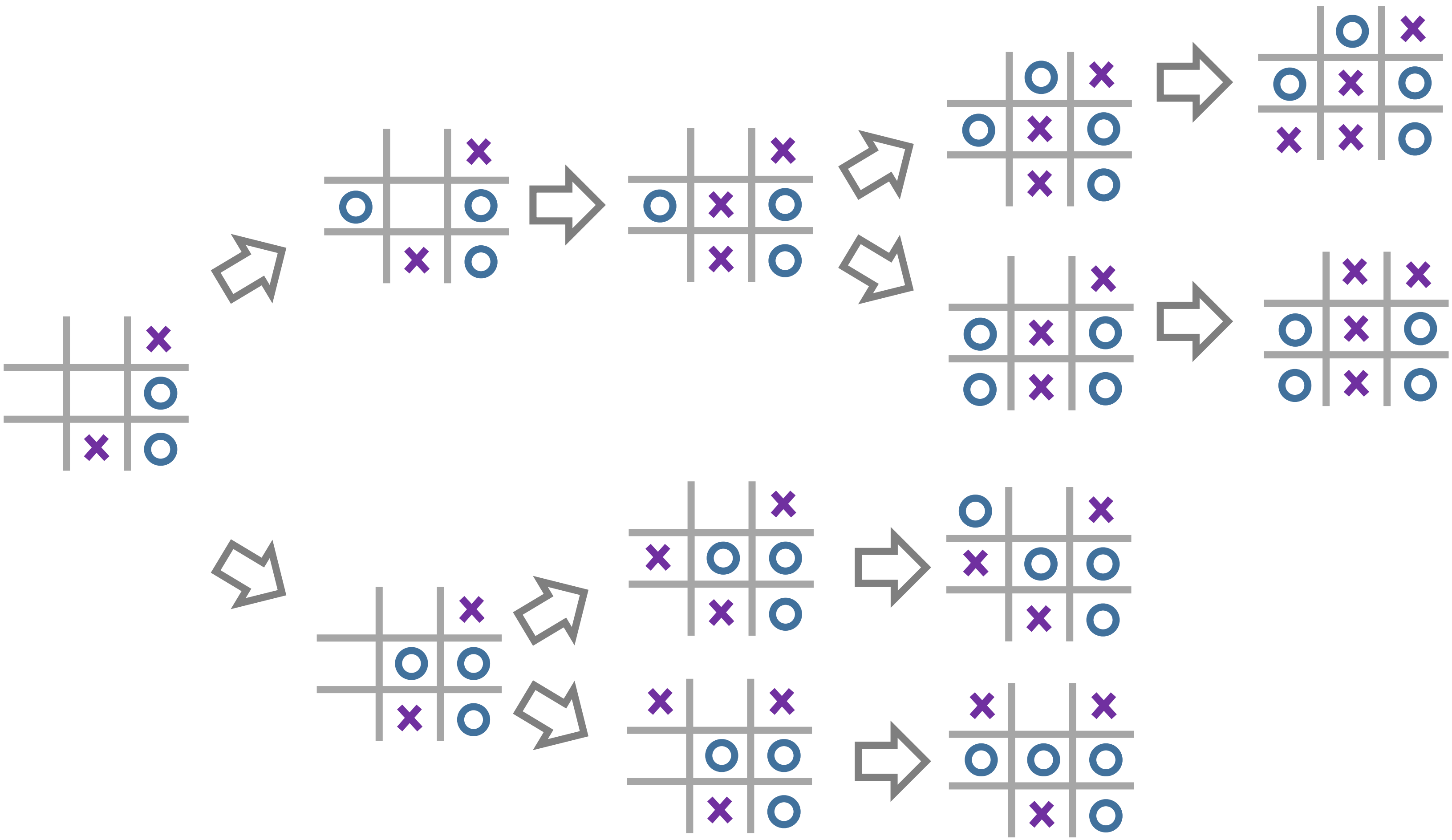
- 井字棋游戏：三子最先成线者胜





极大极小搜索应用示例：井字棋AI

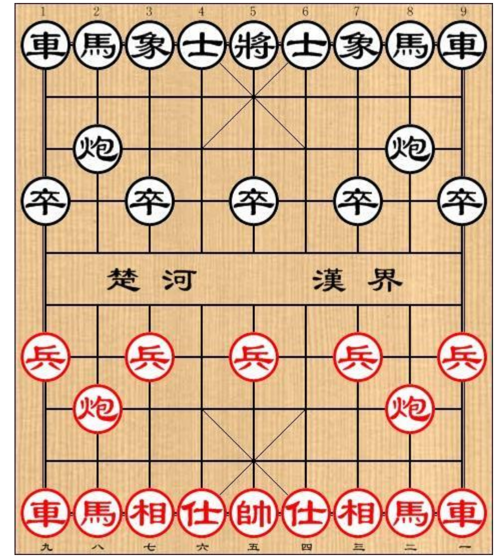
- 井字棋游戏：中间对抗过程举例



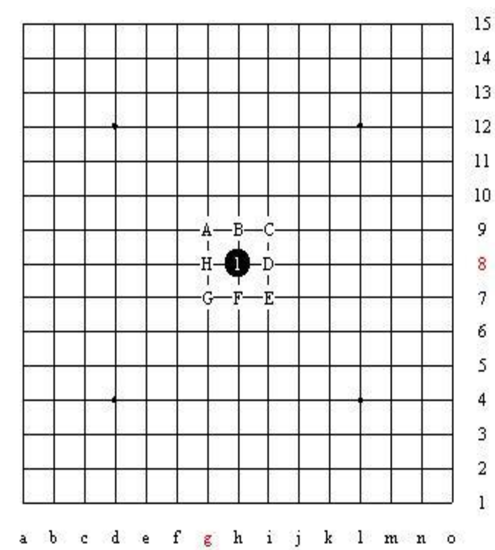
问题：如何设计损益函数？

极大极小搜索应用示例：中国象棋和五子棋

- 中国象棋AI搜索复杂度
  - 棋盘大小：9×10
  - 一盘棋平均走50步，每一步大概有20多种走法，总状态数约为10的48次方；

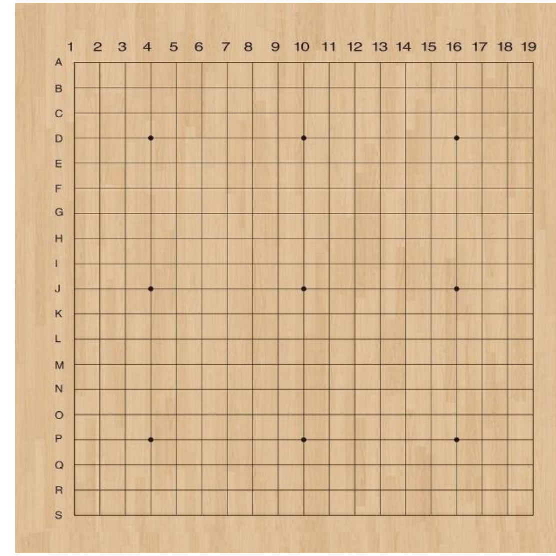


- 五子棋AI的搜素复杂度
  - 棋盘大小15×15
  - 每一步有几十种走法，总状态数约为10的105次方



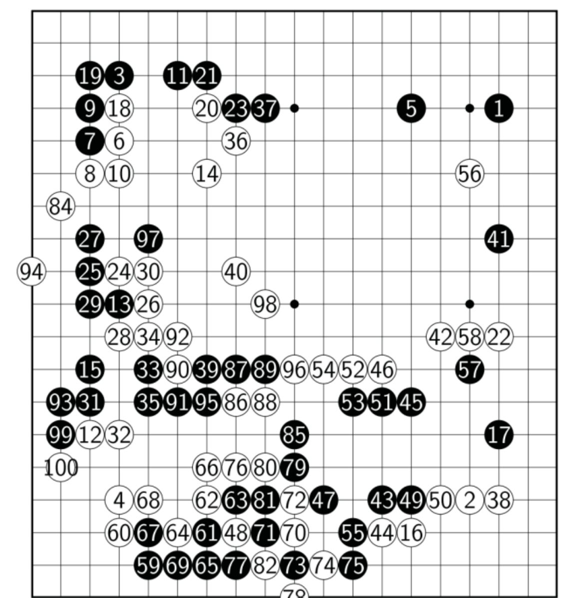
极大极小搜索应用示例：围棋AI

- 围棋AI的搜素复杂度
  - 棋盘大小19×19
  - 每一步有数百种走法，总状态数约为10的172次方



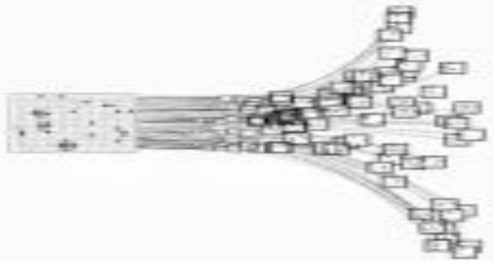
- 假设1微秒走一步，象棋、五子棋、围棋博弈分别需要10^{34}、10^{93}和10^{154}年

- 结论：无论是象棋、五子棋和围棋，都无法穷举搜索！



极大极小搜索应用示例：围棋AI

- 国际象棋AI和围棋AI中的博弈树搜索复杂度对比



定义和表示

- 完全信息静态博弈：博弈过程中参与人预先知道关于博弈的所有信息，但是需要同时行动或者预先不知晓对方行动的博弈。
- 完全信息静态博弈的三要素
  - **局中人**：是博弈的行为主体，可能是自然人，也可能是企业、团体、群体，甚至可以是虚拟的参与人、无形的自然。有时候为了分析方便，参与人每个可能信息状态都可以看到代理人。
  - **策略空间**：参与人i可选择的行动策略集合记为S_i，则每一个选择的策略$s_i \in S_i$，$\mathbf{s} = (s_i, \ldots, s_n)$称为n个参与人的策略组合（strategy profile），所有参与人的策略及S组合就构成了博弈的策略空间$\mathbf{S} = \underset{1 \le i \le n}{\times} S_i$。
  - **损益函数**：或者称为收益函数、支付函数等，是博弈的局中人最关心的。

定义和表示

- 完全信息静态博弈是最为简单的一种博弈类型，通常使用策略型博弈表示，也称为标准式表示，需要同时表示出完全信息静态博弈的三要素：
  - 博弈的参与人集合：$i \in \Gamma, \Gamma = (1,2,\ldots,n)$
  - 每个参与人的策略空间：$S_i, i = 1,2,\ldots,n$
  - 每个参与人的支付函数：$u_i(s_1,\ldots,s_i,\ldots,s_n)$

囚徒B / 囚徒A	坦白	抵赖
坦白	-5, -5	0, -10
抵赖	-10, 0	-1, -1

- 上述表示用于离散策略空间的完全信息静态博弈表示。一般地，完全信息静态博弈可以形式化如下形式：

$$\mathrm{G} = (\{S_i\}_{i=1}^{n}, \{\mu_i(s_1,\ldots,s_n)\})$$

其中：S_i可以是受某些因素影响的连续函数，
$u_i(s_1,\ldots,s_i,\ldots,s_n)$ 也可以是定义在连续空间上的函数。

一些基本概念

- 损益函数的等价
  - 损益函数反映了局中人在一定条件下对于博弈的偏好
  - 局中人偏好如果发生改变，损益函数一定发生了变化
  - 损益函数改变后这种偏好可以不变，这样变化前后的损益函数就认为是等价的。
- 博弈的等价：如果两个博弈的损益函数是等价的，那么这两个博弈就是等价的。

博弈的等价

两个博弈G和$\hat{G}$等价的充要条件是：对于参与人集合中的每个人i和策略集中的任意两个策略p和q，如果参与人i在博弈G中对于p偏好胜过q，就等价于在博弈$\hat{G}$中对于p的偏好胜过q，那么我们就说博弈G和$\hat{G}$等价。

求解：占优策略均衡

- 优策略：博弈参与者的策略一般依赖于其他参与者的策略。在特定情况下，如果参与者的最优策略不依赖于其他参与者的策略选择，这样的最优策略称为优策略。

（严格）优策略和（严格）劣策略：

给定一个博弈的标准式描述$G = \{S_1, \dots, S_n; \mu_1, \dots, \mu_n\}$，令$s_i^*$是参与者$i$的某个策略，$s_i'$是参与者$i$除了$s_i^*$的任意策略，$s_{-i} = (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n)$是其他参与者的任意策略组合，如果$\mu_i(s_i^*, s_{-i}) \geq \mu_i(s_i', s_{-i})$，$s_i^*$称为参与者$i$的**优策略**（**Dominant Strategy**），其他的s_i'称为i的**劣策略**（**Dominated Strategy**），如果$\mu_i(s_i^*, s_{-i}) > \mu_i(s_i', s_{-i})$，$s_i^*$称为$i$的**严格优策略**（**StrictlyDominant Strategy**），其他的其他的s_i'称为i的**严格劣策略**（**Strictly Dominated strategy**）。

求解：占优策略均衡

- 占优策略均衡：
  - 在一个博弈中，如果一个人有优策略，那么根据理性人假设，他肯定会选择优策略，甚至不用知道其他参与者是否会理性。
  - 在一个博弈中，如果所有人都有优策略，那么所有人都会选择各自的优策略，从而得到唯一的均衡。
  - 所有人的优策略集合就构成了一个**占优策略均衡**（**Dominant-Strategy Equilibrium**）。

占优策略均衡

给定一个博弈的标准式描述$G=\{S_1,\ldots,S_n;\mu_1,\ldots,\mu_n\}$中，如果对于所有的$i$，$s_i^*$是$i$的占优策略，那么策略组合$s^*=(s_1^*,\ldots,s_n^*)$称为占优策略均衡。

占优策略均衡举例

- 囚徒博弈
  - 囚徒A占优策略：坦白
  - 囚徒B占优策略：坦白
  - 占优均衡：（坦白，坦白）
- 买家市场房地产开发博弈
  - 开发商A占优策略：开发
  - 开发商B占优策略：开发
  - 占优均衡：（开发，开发）
- 达到占优均衡的条件
  - 只要求每个参与人是理性的即可
  - 不要求参与人知道其他人是理性的

囚徒A \ 囚徒B	坦白	抵赖
坦白	-5, -5	0, -10
抵赖	-10, 0	-1, -1

开发商A \ 开发商B	开发	不开发
开发	4, 4	8, 0
不开发	0, 8	0, 0

求解：重复剔除的占优均衡

- 智猪博弈回顾
  - 对于小猪：存在优策略“不动”
  - 对于大猪：不存在优策略
  - 如何求取这个博弈的均衡?

大猪 \ 小猪	行动	不动
行动	5, 1	4, 4
不动	9, -1	0, 0

- 一种直观的思路
  - 首先找到某个参与人劣策略
  - 剔除这个劣策略得到新博弈
  - 对新博弈继续重复上述过程

重复剔除劣策略

大猪 \ 小猪	行动	不动
行动	5, 1	4, 4
不动	9, -1	0, 0

求解：重复剔除的占优均衡

- 相对严格劣策略和严格优策略

在一个标准博弈中，令s_i'和s_i''是参与人i可选择的两个策略（即$s_i' \in S_i$，$s_i'' \in S_i$），如果对于任意的其他人的策略组合s_{-i}，参与人i从选择s_i'得到支付严格小于从选择s_i''得到的支付，即：

$$u_i(s_i', s_{-i}) < u_i(s_i'', s_{-i}), \qquad \forall s_{-i}$$

我们说策略s_i'严格劣于策略s_i''（s_i' is strictly dominated by s_i''）。或者说，s_i'是相对于s_i''的劣策略，s_i''是相对于s_i'的优策略。

- 相对弱优策略和弱劣策略

如果对于任意的其他人的策略组合s_{-i}，$u_i(s_i', s_{-i}) \leq u_i(s_i'', s_{-i})$，且对于某些$s_{-i}$，严格不等式成立，那么说，策略$s_i'$弱优于策略$s_i''$（$s_i'$ is weakly dominated by s_i''），或者称s_i''是相对于s_i'的弱优策略。

求解：重复剔除的占优均衡

- 在一个标准博弈中，如果策略组合$s^* = (s_1^*, \ldots, s_n^*)$是重复剔除劣策略后剩下的策略组合，那么就称该策略组合为重复剔除的占优策略。

- 对于一个标准博弈，如果它存在唯一的重复剔除的占优均衡，那么就称改博弈是重复剔除占优可解的（dominance solvable）。

- 举例：

红方 \ 蓝方	L	M	R
U	1, 0	1, 2	0, 1
D	0, 3	0, 1	2, 0

红方 \ 蓝方	L	M
U	1, 0	1, 2
D	0, 3	0, 1

红方 \ 蓝方	L	M
U	1, 0	1, 2

红方 \ 蓝方	M
U	1, 2

求解：纳什均衡

- 市场萎靡下的房地产开发博弈
  - 两个开发商都开发：都损失3
  - 一个开发商开发：收益1
  - 都不开发：收益为0

开发商A \ 开发商B	开发	不开发
开发	-3, -3	1, 0
不开发	0, 1	0, 0

- 该博弈不存在占优策略均衡和重复剔除的占优均衡，但存在均衡：纳什均衡！

纳什均衡

在一个标准式博弈$G=\{S_1,\dots,S_n;\mu_1,\dots,\mu_n\}$中，对于一个策略组合$s^*=(s_1^*,\dots,s_n^*)$，如果对于每一个参与人$i$，如果$s_i^*$是给定其他参与人策略$s_{-i}^*=(s_1^*,\dots,s_{i-1}^*,s_{i+1}^*,\dots,s_n^*)$的情况下第$i$个人的最优策略，即：$u_i(s_i^*,s_{-i}^*)\ge u_i(s_i,s_{-i}^*)$，$\forall s_i\in S$，$\forall i$，那么策略组合$s^*=(s_1^*,\dots,s_n^*)$称为纳什均衡。

求解：强纳什均衡和弱纳什均衡

- 强纳什均衡：在一个纳什均衡中，如果给定其他参与人的策略，每一参与人的最优选择都是唯一的。即：s_i^*是一个强纳什均衡，当前仅当对于所有的i，$s_i' \neq s_i^*$，$u_i(s_i^*, s_{-i}^*) > u_i(s_i', s_{-i}^*)$。

- 弱纳什均衡：在一个纳什均衡中，如果给定其他参与人的策略，每一参与人的最优选择不是唯一的。

纳什均衡的应用举例：

- 求取右图抽象博弈的纳什均衡
  - 对于简单的两人有限博弈，可以根据纳什均衡定义设计方法求取
- 具体方法：不断固定一方的策略，求取另一方的最优策略，迭代完成后最优策略的交集就是纳什均衡。

蓝方 / 红方	L	C	R
U	0, 4	4, 0	5, 3
M	4, 0	0, 4	5, 3
D	3, 5	3, 5	6, 6

蓝方 / 红方	L	C	R
U	0, 4	4, 0	5, 3
M	4, 0	0, 4	5, 3
D	3, 5	3, 5	6, 6

给定红方策略，求蓝方最优策略

蓝方 / 红方	L	C	R
U	0, 4	4, 0	5, 3
M	4, 0	0, 4	5, 3
D	3, 5	3, 5	6, 6

给定蓝方策略，求红方最优策略

求解：混合策略纳什均衡

- 纯策略（Pure Strategy）和混合策略（Mixed Strategy）
  - 在一个标准式博弈中，如果一个策略规定参与人在一个给定的信息情况下只选择一种特定的行动，我们就称该策略为纯策略。
  - 如果一个策略规定参与人在给定信息情况下以某种概率分布随机地选择不同的行动，我们就称该策略为混合策略。

混合策略（Mixed Strategy）

在一个标准式博弈$G=\{S_1,\ldots,S_n;\mu_1,\ldots,\mu_n\}$中，假定参与人$i$有$K$个纯策略：$S_i=\{s_{i,1}^*,\ldots,s_{i,K}^*\}$，那么，概率分布$\sigma_i=(\sigma_{i1},\ldots,\sigma_{iK})$称为$i$的一个混合策略，这里$\sigma_{ik}=\sigma(s_{ik})$是$i$选择$s_{ik}$的概率，对于所有的$k=1,\ldots,K$，$0\le\sigma_{ik}\le 1$，$\sum_i^K\sigma_{ik}=1$。

求解：混合策略纳什均衡

- 混合策略博弈要素的表示
  - 参与人的混合策略：$\sigma_i = (\sigma_{i1}, \dots, \sigma_{i2}, \dots, \sigma_{ik})$
  - 参与人i混合策略空间：Σ_i，其中$\sigma_i \in \Sigma_i$
  - 混合策略组合：$\sigma = (\sigma_1, \dots, \sigma_i, \dots, \sigma_n)$
  - 混合策略组合空间：$\Sigma = \times_i \Sigma_i$
  - 参与人i期望支付函数：$v(\sigma) = v(\sigma_i, \sigma_{-i}) = \sum_s \left(\prod_j^n \sigma_j(s_j)\right) u(s)$
- 混合策略纳什均衡

混合策略纳什均衡

在一个标准式博弈$G = \{S_1, \dots, S_n; \mu_1, \dots, \mu_n\}$中，如果对于所有的i$= 1, \dots, n$，混合策略组合$\sigma^* = \{\sigma_1^*, \dots, \sigma_i^* \dots, \sigma_n^*\}$都满足：

$$v(\sigma_i^*, \sigma_{-i}^*) \geq v(\sigma_i, \sigma_{-i}^*), \qquad \forall \sigma_i \in \Sigma_i$$

那么，该混合策略组合是一个纳什均衡。

求解：混合策略纳什均衡

- 税收监督博弈
  - 税务机关检查概率为Θ
  - 纳税人逃税的概率为γ
  - 如何求取其纳什均衡?

税收机关 \ 纳税人	逃税	不逃税
检查	a-C+F, -a-F	a-C, -a
不检查	0, 0	a, -a

给定逃税概率γ，税收机关选择检查($\Theta = 1$)和不检查($\Theta = 0$)的期望收益分别为：

$$\pi_G(1,\gamma) = (a - C + F)\gamma + (a - C)(1 - \gamma) = \gamma F + a - C$$

$$\pi_G(0,\gamma) = 0\gamma + a(1 - \gamma) = a(1 - \gamma)$$

令$\pi_G(1,\gamma) = \pi_G(0,\gamma)$，得到：$\gamma^* = C/(a + F)$

给定检查概率Θ，税收人选择逃税($\gamma = 1$)和不逃税($\gamma = 0$)的期望收益分别为：

$$\pi_p(\theta,1) = (-a - F)\theta + 0(1 - \theta) = (-a - F)\theta$$

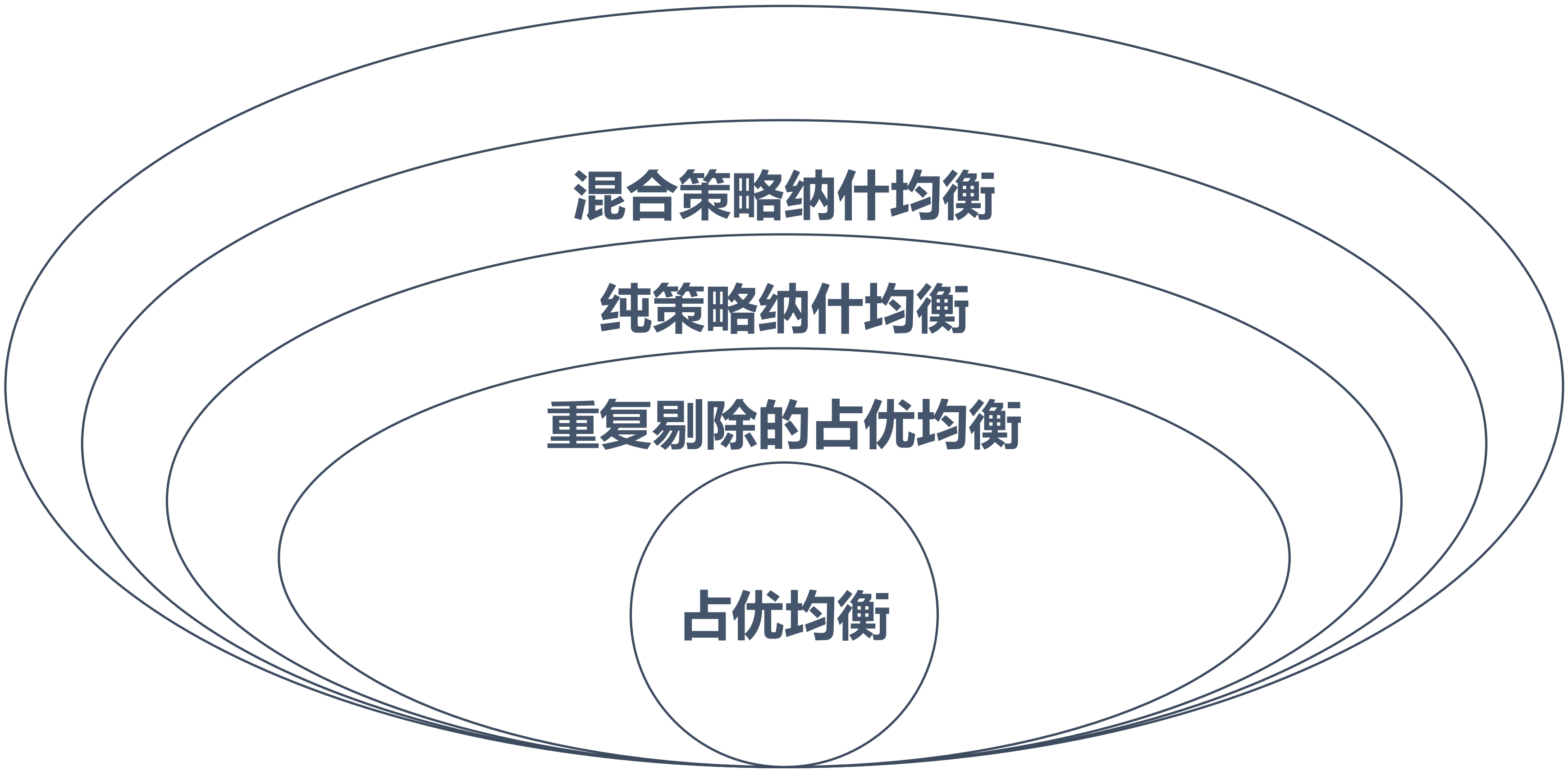
$$\pi_p(\theta,0) = -a\theta + (-a)(1 - \theta) = -a$$

令$\pi_p(\theta,1) = \pi_p(\theta,0)$，得到：$\theta^* = a/(a + F)$

混合策略纳什均衡：$\gamma^* = C/(a + F)$，$\theta^* = a/(a + F)$

纳什均衡的一些讨论

- 四种不同均衡的关系



纳什均衡的一些讨论

- 纳什均衡的存在性

纳什均衡存在性定理一（Nash, 1950）

每一个有限博弈至少存在一个纳什均衡（纯策略的或者混合策略的）。

纳什均衡存在性定理二（Debreu, 1952; Glickskerg, 1952; Fan, 1952）

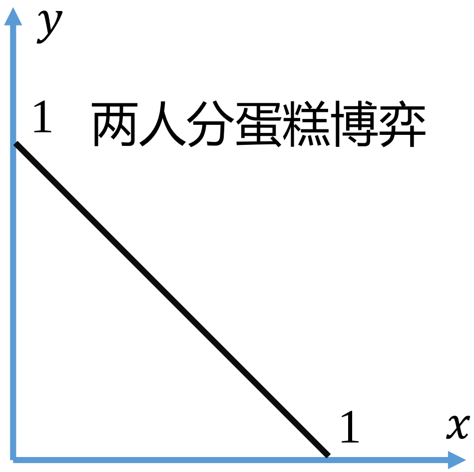
在n人策略式博弈中，如果每个参与人的纯策略空间S_i是欧式空间上一个非空的、闭的、有界的凸集，支付函数$u_i(s)$是连续的且对s_i是拟凹的，那么存在一个纯策略纳什均衡。

纳什均衡存在性定理三（Glickskerg, 1952）

在n人策略式博弈中，如果每个参与人的纯策略空间S_i是欧式空间上一个非空的、闭的、有界的凸集，支付函数$u_i(s)$是连续的，那么存在一个混合策略纳什均衡。

关于纳什均衡的一些讨论

- 纳什均衡的多重性：相对于纳什均衡的存在性，纳什均衡的多重性有时候更为棘手
- 一个博弈纳什均衡可能出现的数目
  - 不存在纳什均衡
  - 有唯一的纳什均衡
  - 有多个纳什均衡
  - 有无数个纳什均衡
- 当一个博弈有多个纳什均衡时
  - 如何保证纳什均衡出现?
  - 如果保证某个纳什均衡出现?



分配规则：
- 每人说出想要得到的蛋糕数量；
- 两人所要数量和不超过蛋糕总量得到所要数量；
- 两人所要数量和超过蛋糕总量则什么都得不到；

定义和表示

- 完全信息动态博弈：博弈过程中参与人预先知道关于博弈的所有信息，参与人的行动有先后顺序，且后行动者能够观察到先行动者行动的博弈。
- 完全信息动态博弈的展开式表示：
  - 参与人集合：$i \in \Gamma, \Gamma = (1, 2, \ldots, n)$，用$N$表示自然
  - 参与人的行动顺序：谁在什么时候行动
  - 参与人的行动空间：每次行动时，有什么选择
  - 信息集合：每次行动结束，参与人知道些什么
  - 效益函数：每次行动时，参与人知道什么
  - 外生事件（即自然选择）的概率分布

展开式表示举例

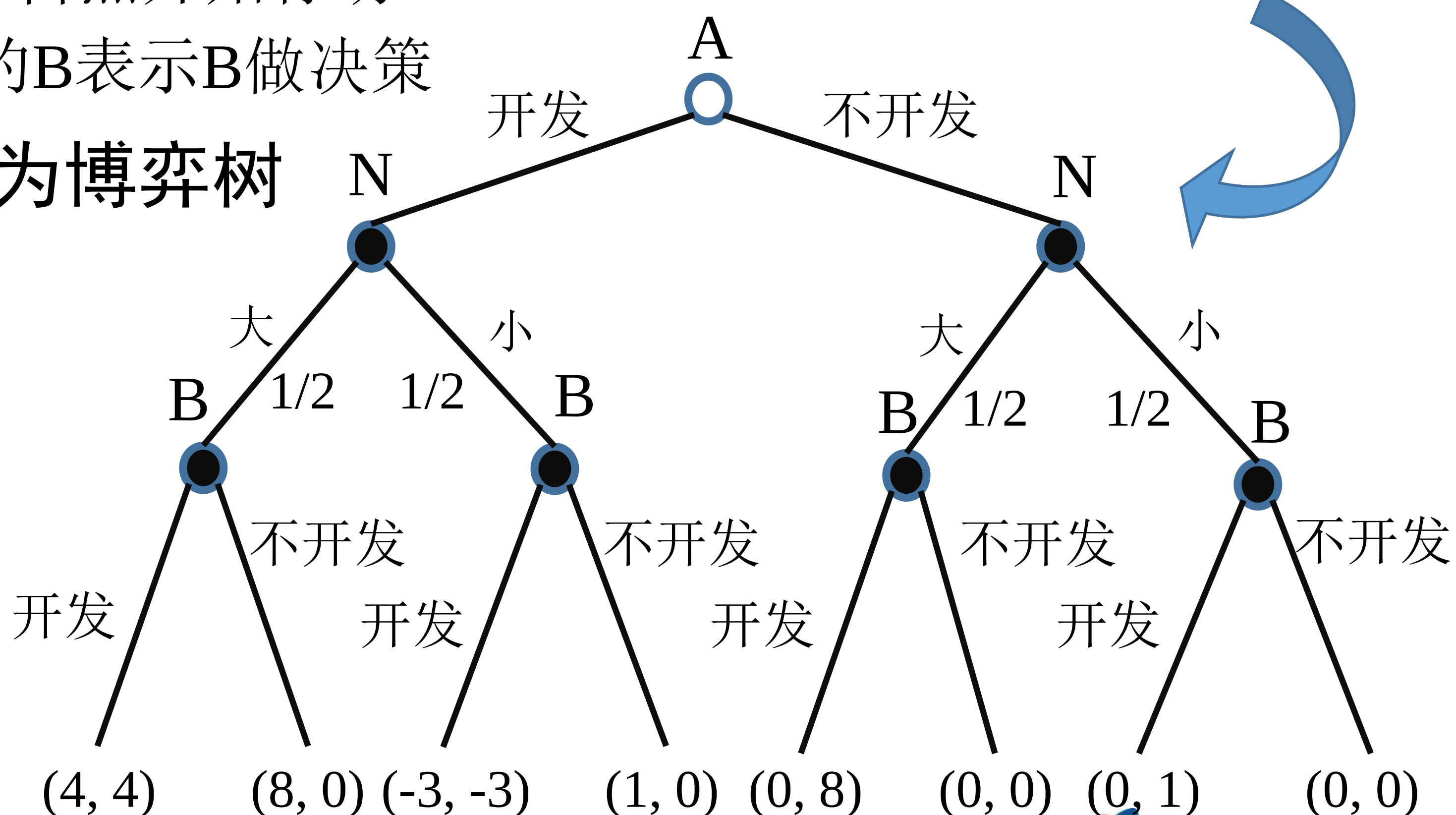
- 房地产开发博弈
  - 博弈从空心圆点开始
  - 空心点旁的A表示A先决策
  - N结点表示自然开始行动
  - 实心圆旁的B表示B做决策
- 该展开式称为博弈树

市场需求大

开发商B / 开发商A	开发	不开发
开发	4, 4	8, 0
不开发	8, 1	0, 0

市场需求小

开发商B / 开发商A	开发	不开发
开发	-3, -3	1, 0
不开发	0, 1	0, 0

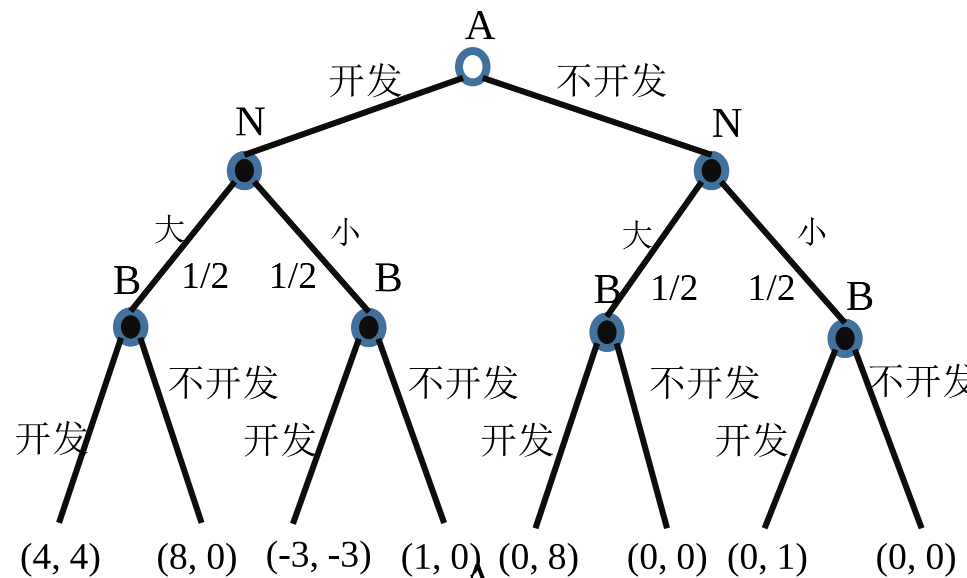


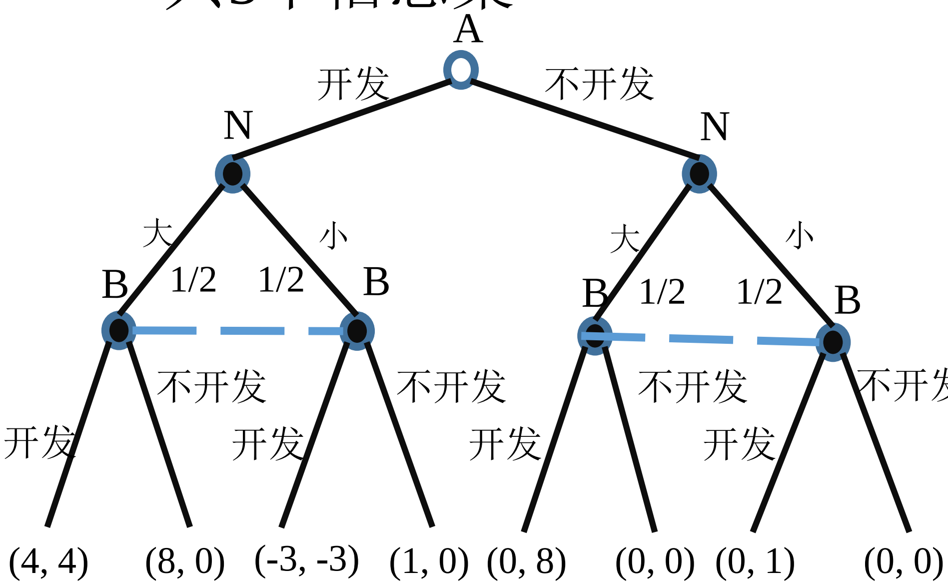
博弈树定义回顾

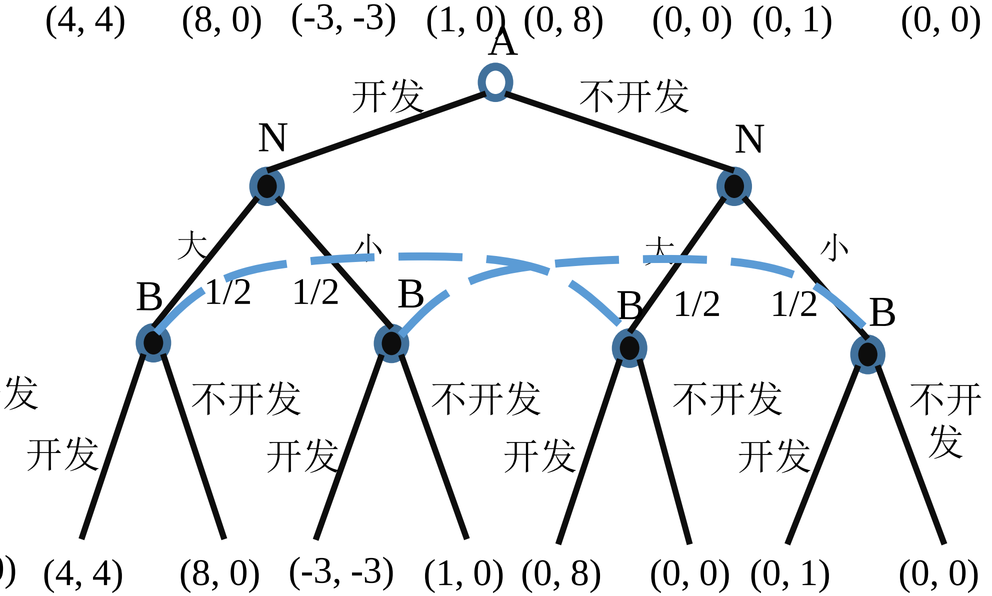
- 博弈树：树中每一对结点有且仅由一个树支与其相连，树的根是博弈的起始结点。
  - 结点：$x \in X \doteq \{x\}$，关系符号$\prec$表示结点的先后顺序关系。具体包括机会结点、决策结点和终结点；
  - 路径：结点x之前的全部结点集合称为x的前列集，又称为x的路径，记为$P(x)$;
  - 后续结点：结点x之后的所有结点的集合称为x的后续结点集，记为$T(x)$;
  - 跟随：当结点y位于结点x的路径上，即$y \in P(x)$，则称x跟随y;
  - 树枝：结点x到其直接跟随结点y的连线，树枝代表一个可供参与人选择的行动策略或者时间；
  - **信息集**：博弈决策结点集合的一个子集，且满足：1）每个决策结点都是同一个参与人的决策结点；2）该参与人知道博弈进入该节点子集，但是不知道自己具体在哪一个决策节点。

房地产开发博弈中的不同信息集

- B知道A和N的选择
  - 共7个信息集
- B知道A不知道N的选择
  - 共5个信息集
- B知道N不知道A的选择
  - 共5个信息集



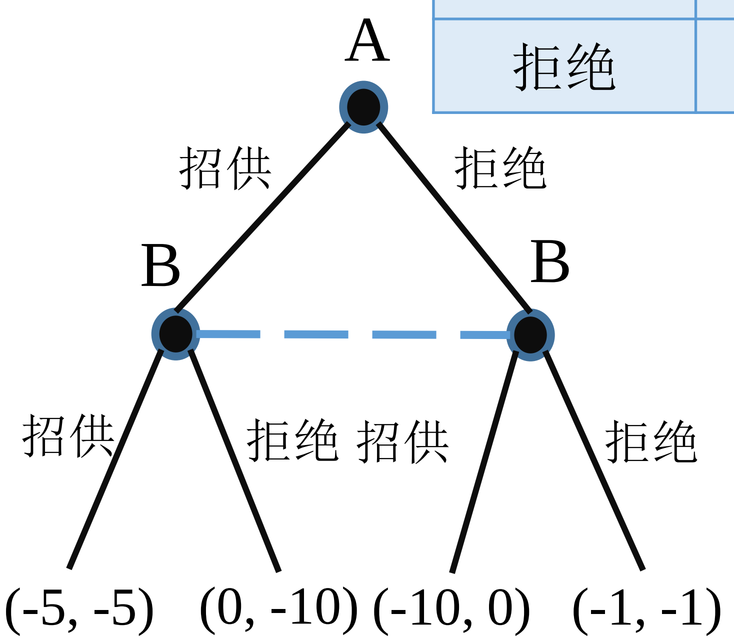


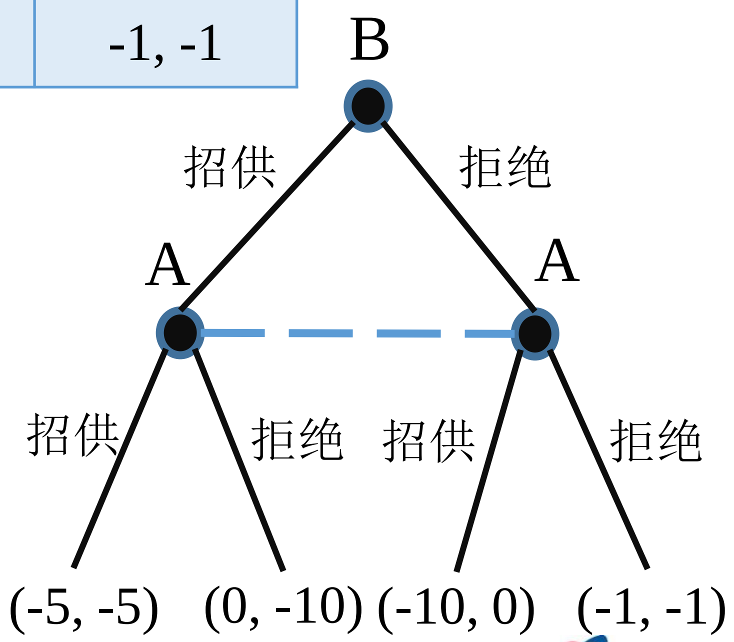


展开式表示用于表示静态博弈

- 借助于信息集的概念，可以用展开式表示来表示静态博弈。以囚徒困境博弈为例：
  - 第二个决策的人不知道第一个人的决策

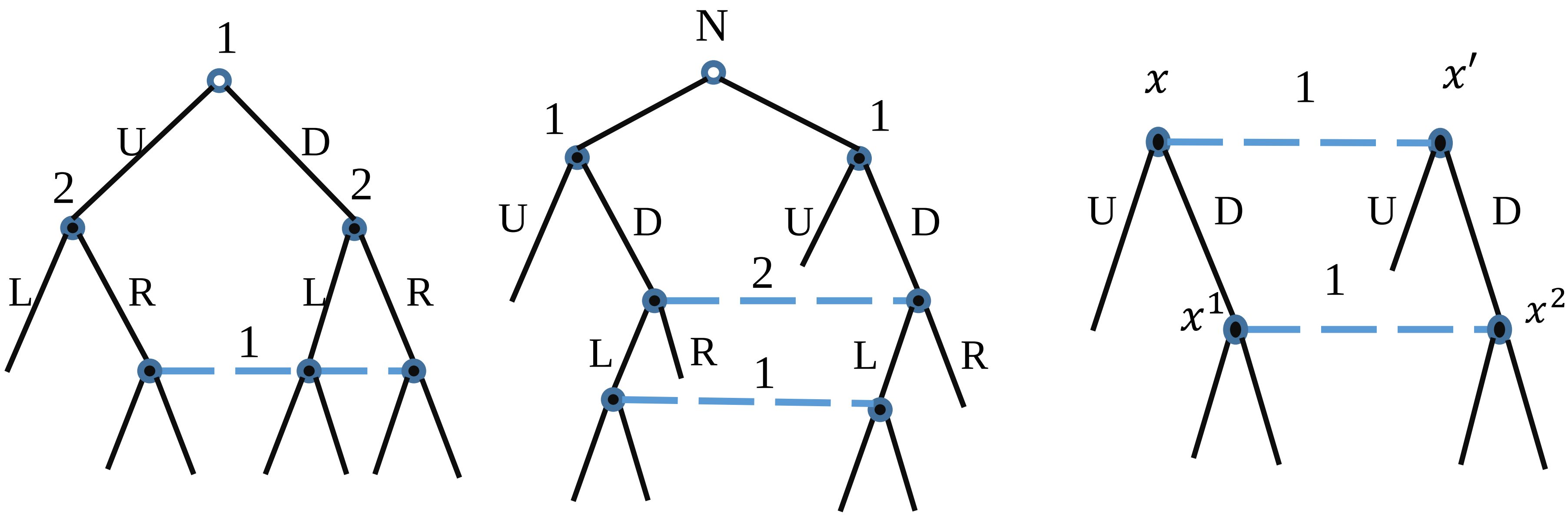
囚徒B / 囚徒A	招供	拒绝
招供	-5, -5	-10, 0
拒绝	0, -10	-1, -1





完美信息的定义和判断

- 完美信息：与信息集相关的一个概念，它指的是没有参与人会忘记自己以前知道的事情，所有参与人都知道自己以前的选择。下图1和图2是两个不完美信息的例子：



- 上图3是保证完美信息的条件：如果x^1和x^2属于同一个信息集，x是x^1的一个前列结，那么一定存在x^2的一个前列结x''和x属于同一个信息集。

扩展式表示到策略式表示的转换

- 扩展式表示的形式化
  - 同样用S_i表示纯策略，u_i表示支付函数
  - H_i：第i个参与人的信息集
  - $A_i \doteq \cup_{h_i \in H_i} A(h_i)$：第$i$个参与人的行动集合，$A(h_i)$是在信息集$h_i$下的行动集合
  - 参与人的一个纯策略是从信息集集合H_i到行动集合A_i的一个映射，用$S_i: H_i \to A_i$表示，其中$h_i \in H_i$，$S_i(h_i) \in A(h_i)$，参与人的纯策略空间$S = \{S_i\}$，$S_i = \underset{h_i \in H_i}{\times} A(h_i)$)，纯策略的个数：$\#S_i = \prod_{h_i \in H_i} \#A(h_i)$
  - 每一个纯策略组合决定了博弈树上的一个路径，比如（开发，{不开发，开发}）决定了博弈的路径为A→开发→ B →不开发 →（0, 1）

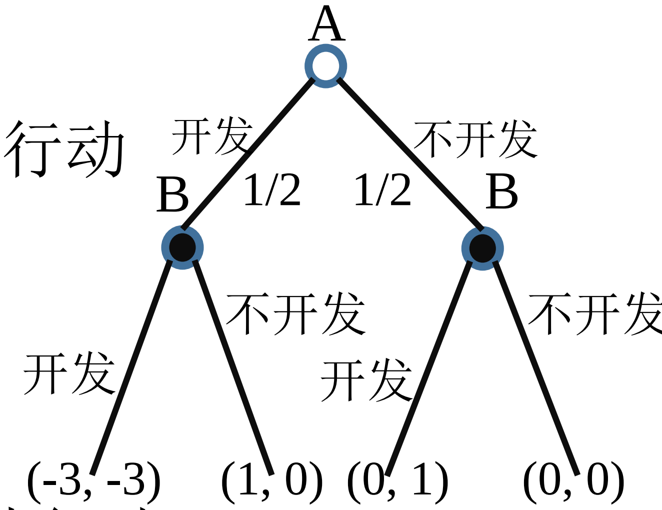
扩展式博弈的纳什均衡

- 通过将扩展式博弈转换为策略式表示，每一个策略组合（也就是博弈树上的一条路径）决定了一个支付向量$u(u_1, \ldots, u_n)$，如果策略组合s^*对于所有的i，能够最大化$u_i(s_i, s_{-i}^*)$（或期望值），即：

$$s_i^* \in \underset{s_i \in S_i}{\operatorname{argmax}}\, u_i(s_i, s_{-i}^*), \qquad \forall i$$

- 运用类似上述的构造方式，对于扩展式表示混合策略（行为策略），同样可以构造出扩展式博弈的混合式策略。

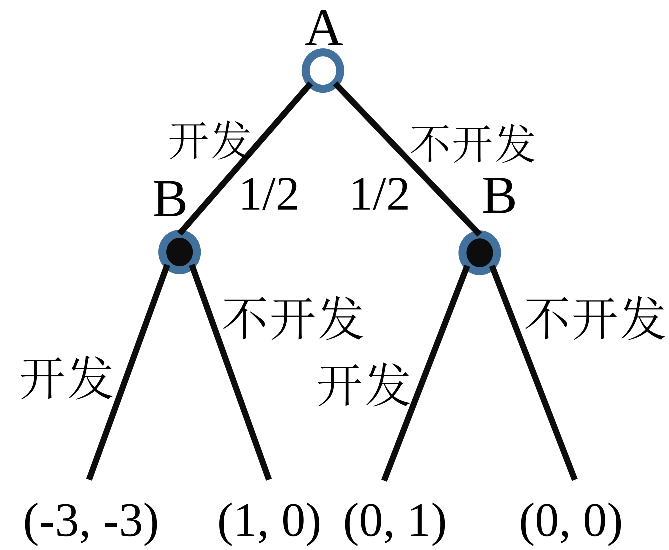
博弈标准式表示和扩展式表示之间的联系



- 卖家市场的房地产开发博弈
  - 开发商A先做行动，B看到A行动后开始行动
  - 如右图扩展式表示，完美信息博弈
- 利用信息集的概念进行构造
  - A有一个单节点信息集，两个可选行动
  - B有两个双节点信息集，每个有两个可选行动
- A的纯策略集很简单：（开发、不开发）
- B的纯策略集有四个
  - 不管A开发还是不开发，我都开发
  - A开发我开发，A不开发我不开发
  - A开发我不开发，A不开发我开发
  - 不管A开发还是不开发，我都不开发

博弈标准式表示和扩展式表示之间的联系

- A的2个纯策略：开发，不开发
- 将B的4个纯策略简记为
  - {开发，开发}
  - {开发，不开发}
  - {不开发，开发}
  - {不开发，不开发}



- 那么A和B之间博弈的标准式如下

开发商B / 开发商A	{开发，开发}	{开发，不开发}	{不开发，开发}	{不开发，不开发}
开发	-3, -3	-3, -3	0, 0	1, 0
不开发	0, 1	0, 0	0, 1	0, 0

扩展式表示博弈的纳什均衡

- 扩展式博弈纳什均衡的存在性

有限完美信息博弈纳什均衡（Zermelo, 1913; Kuhn, 1953）

每一个有限完美信息博弈有一个纯策略纳什均衡。

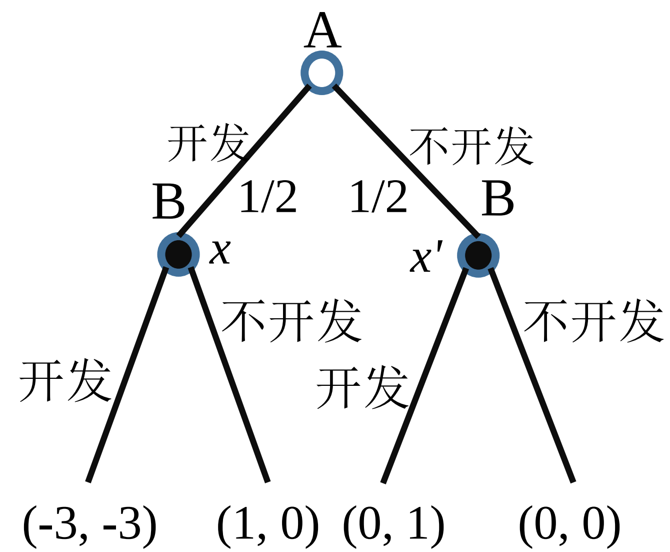
证明：使用动态规划逆向归纳法。
博弈有限：存在一个最后决策结点的集合（倒数第二个结点）
在最后决策结点参与人决策，最大化自己的收益；
给定这个收益，倒数第二个决策结点参与人决策，最大化自己的收益；
如此重复，直到初始结点；
这个倒推过程结束，得到一个路径，给出每个参与人的特定策略，所有这些策略构成了一个纳什均衡。

子博弈精炼纳什均衡

- 动态博弈的扩展式表示都可以转换为策略式表示，从而转换为静态博弈。
- 完全信息静态博弈的纳什均衡因此同样适用于动态博弈。
- 但是这样得到的（多个）纳什均衡往往不是合理的均衡。
    - 纳什均衡会有多个均衡
    - 未考虑到参与者行动的先后顺序，使得很多均衡不合理
- 需要寻求能够改进（perfecting）和精炼（refining）纳什均衡概念，从而得到更为合理的博弈解。

子博弈精炼纳什均衡

- 以右图的房地产开发为例：
  - 首先转化为下图的策略式表示
  - 该策略式表示有三个纳什均衡
    - （不开发，{开发，开发}）
    - （开发，{不开发，不开发}）
    - （开发，{不开发，开发}）



只有（开发，{不开发，开发}）是唯一的合理均衡！

开发商B / 开发商A	{开发，开发}	{开发，不开发}	{不开发，开发}	{不开发，不开发}
开发	-3, -3	-3, -3	1, 0	1, 0
不开发	0, 1	0, 0	0, 1	0, 0

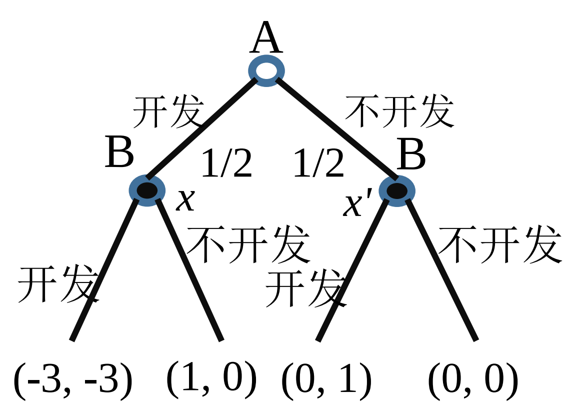
子博弈精炼纳什均衡

- 子博弈：原博弈的一部分，本身可以作为一个独立的博弈进行分析。

子博弈

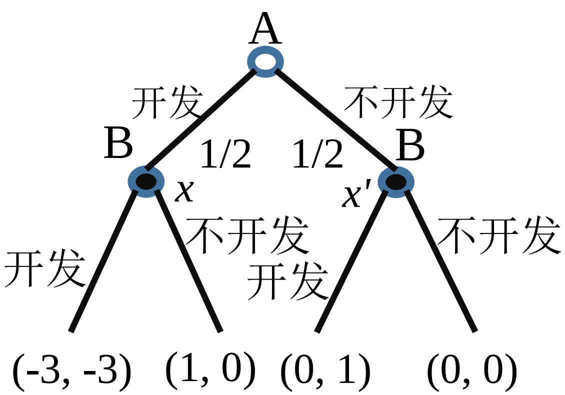
一个扩展式表述博弈的子博弈G由一个决策结x和所有该决策结的后续结$T(x)$（包括终结点）组成，它满足下列条件：1）x是一个单节点信息集，即$h(x) = \{x\}$；2）对于所有的$x' \in T(x)$，如果$x'' \in h(x')$，那么$x'' \in T(x)$。

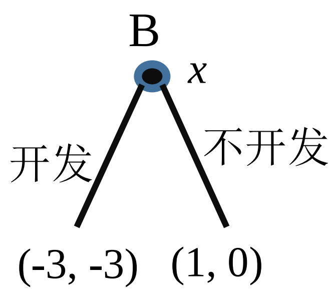
- 任何博弈本身都可以称为自身的一个子博弈

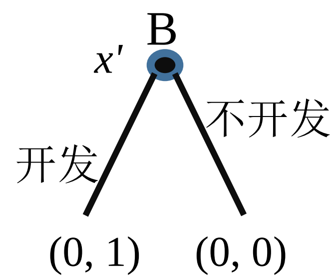


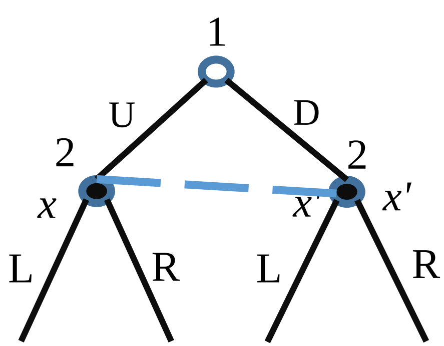
子博弈精炼纳什均衡

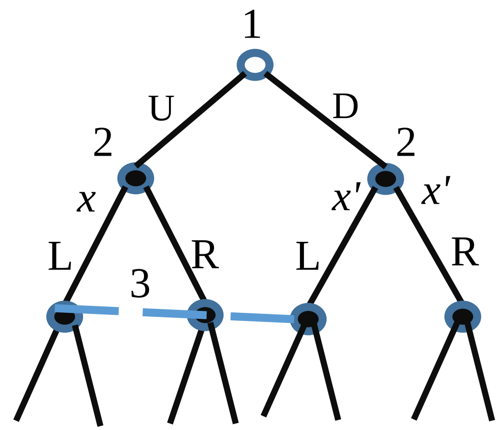
- 子博弈的两个条件分析
  - 条件一：一个子博弈必须从一个单结信息集开始
  - 条件二：子博弈信息集和支付向量都直接继承自原博弈











子博弈精炼纳什均衡

- 子博弈纳什均衡：一个策略组合是子博弈精炼纳什均衡，当且仅当它在每一个子博弈（包括原博弈）上都构成一个纳什均衡。

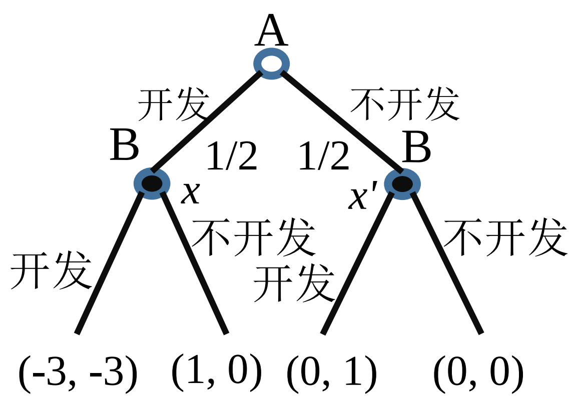
子博弈精炼纳什均衡（Selten, 1965）

扩展式描述博弈的策略组合$s^* = (s_1^*, \dots, s_i^*, \dots, s_n^*)$是一个子博弈精炼纳什均衡，如果：1）它是原博弈的纳什均衡；2）它在每一个子博弈上给出纳什均衡。

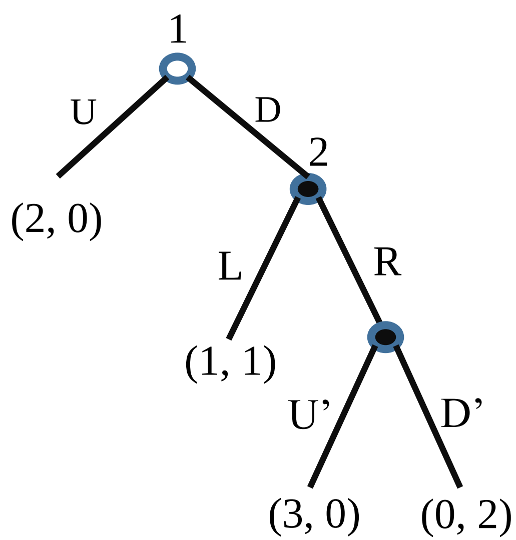
- 在每一个子博弈上给出纳什均衡：无论过去发生什么事情，参与人在每一个决策结上都应该最优化自己的决策。

子博弈精炼纳什均衡

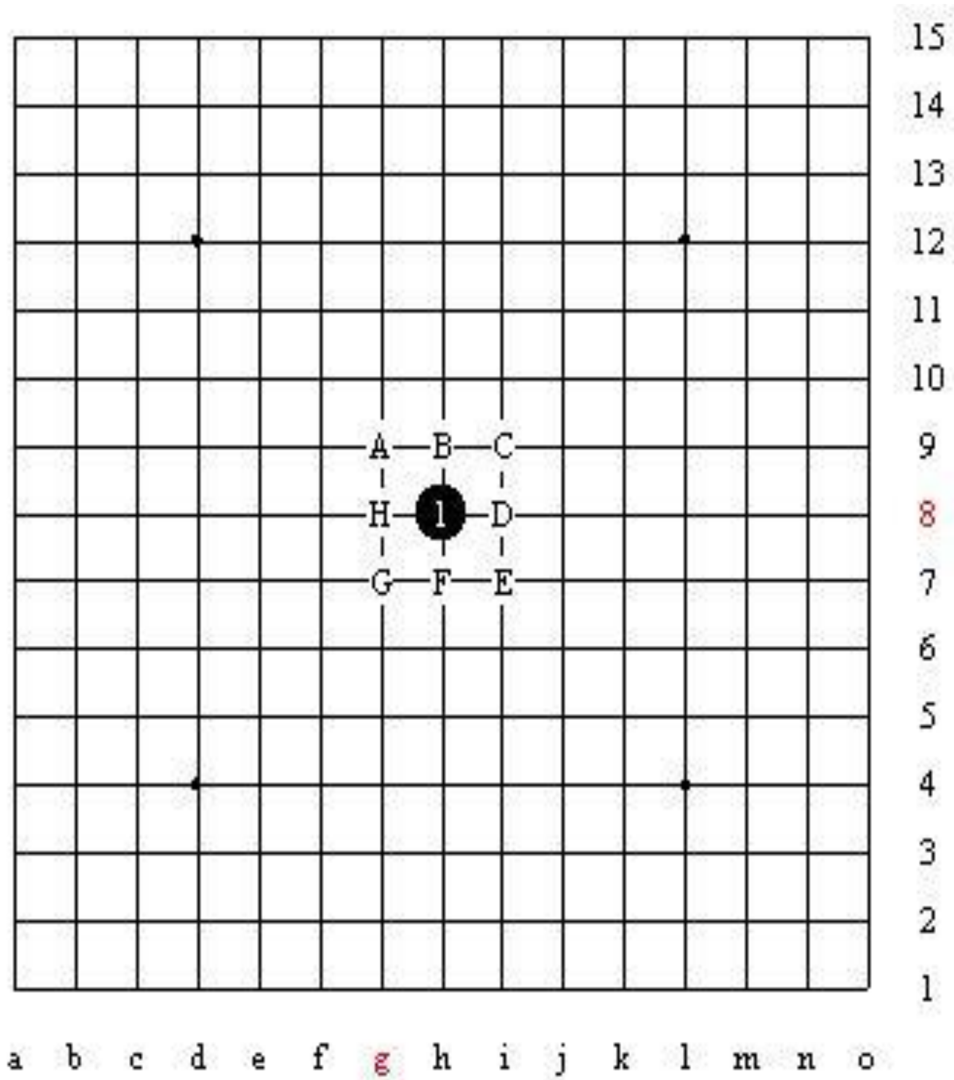
- 使用逆向归纳法求解子博弈精炼纳什均衡
  - 从博弈的最后一个决策结出发，该决策结上行动的参与人做最优决策
  - 倒回到第二个决策结，找出倒数第二个决策者的最优决策
  - 不断重复直到初始决策结



- 举例：
  - 房地产开发博弈
  - 三阶段行动博弈



课堂研讨：五子棋AI设计



本次课程作业

- 作业内容：利用极大极小搜索实现一个零和博弈的AI，比如五子棋、中国象棋等。
  - 参考程序：可参考GitHub上相关程序，但需在实验报告中说明
  - 评分准则：综合考虑所设计AI的智能性、效率以及程序规范性
- 提交时间：2018年4月20日17:00之前
- 提交方法：在课程网站上提交，同时提交电子版Word实验报告和源代码到助教邮箱（peixi.peng@ia.ac.cn）
- 邮件发送规范
  - 邮件主题：博弈论第二次作业_学号_姓名
  - 附件名称：博弈论第二次作业_学号_姓名.zip

感谢聆听！

兴军亮

Junliang.Xing@ia.ac.cn

2018年3月25日

中国科学院大学
University of Chinese of Academy of Sciences
CASIA
自动化研究所
Institute of Automation