Peking University

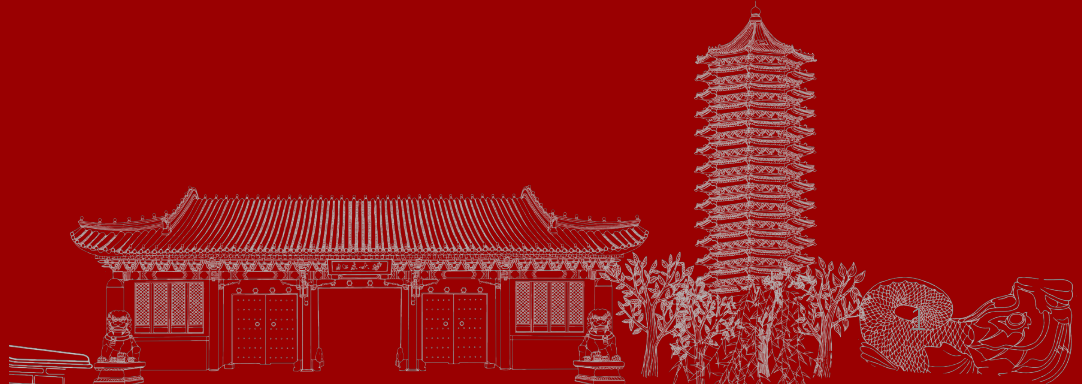
北京大学
PEKING UNIVERSITY

人工智能引论

24. 多智能体 – 博弈论

授课教师 连宙辉

2023年6月1日



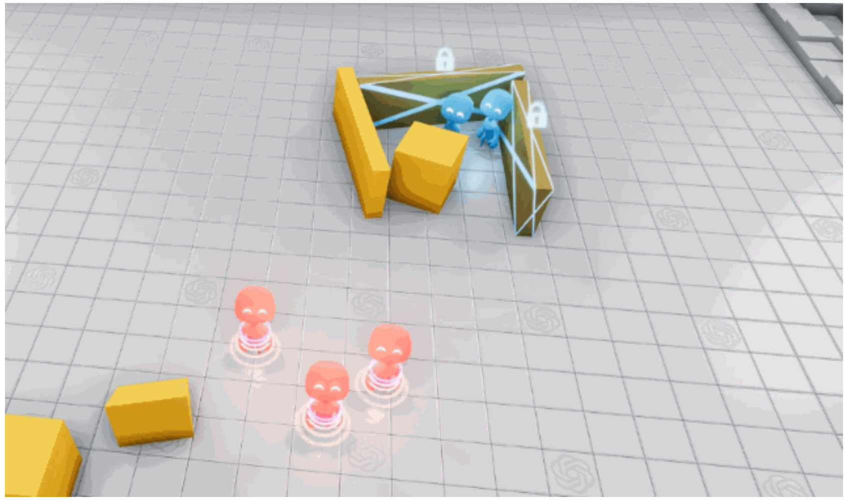
本章内容

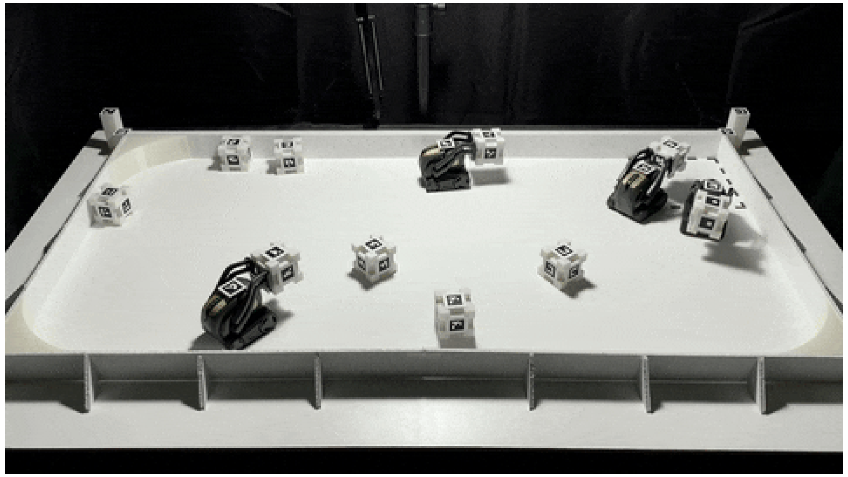
- 多智能体概览
- 博弈论基础
  - 非合作博弈：纳什均衡
  - 合作博弈：公平性公理与沙普利值
  - 机制设计：拍卖

多智能体系统 (Multi-Agent System, MAS)

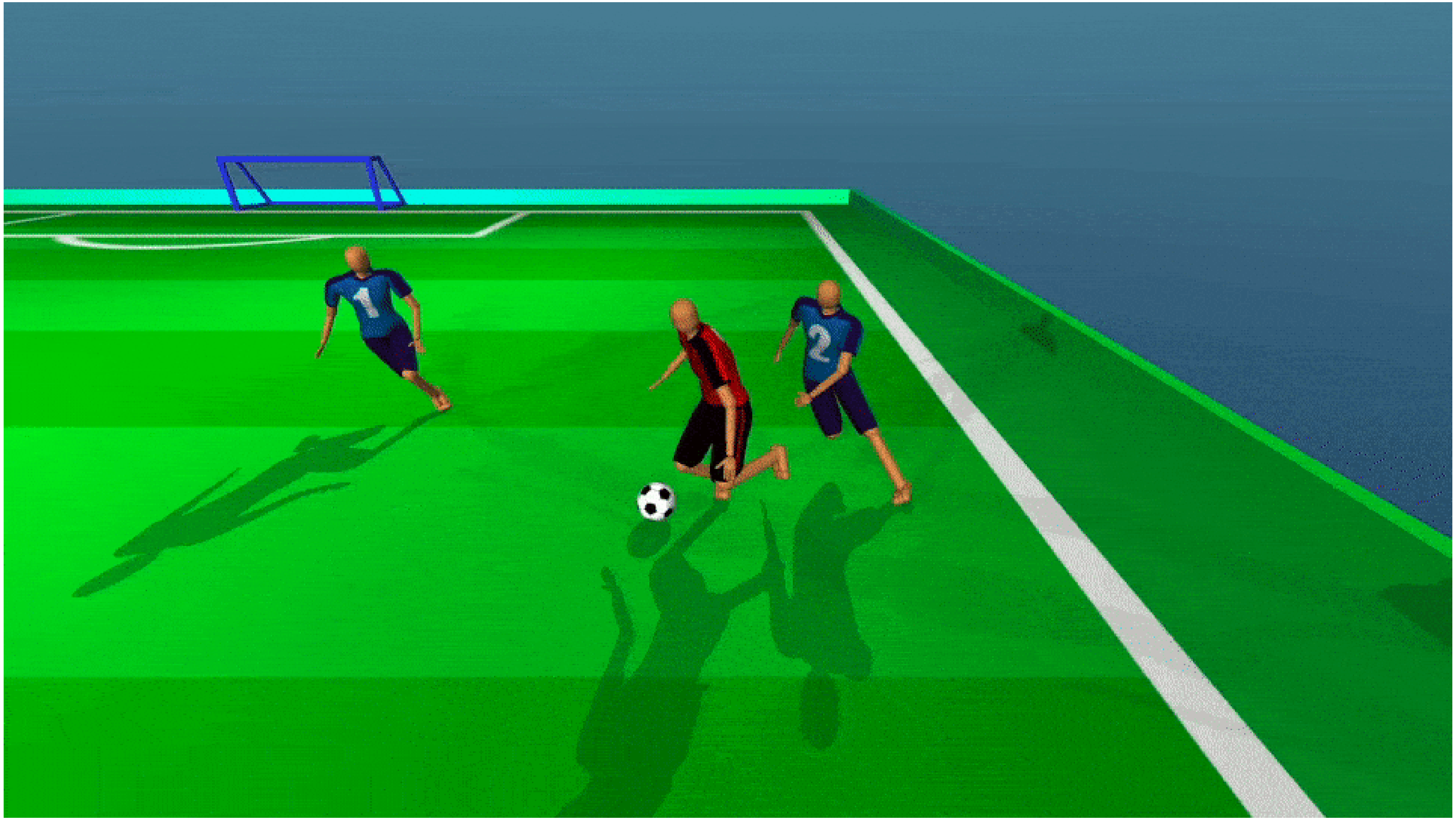
北京大学 PEKING UNIVERSITY

- 一个环境中交互的多个智能体组成的计算系统
  - 每个智能体既是行动者，同时也是决策者
  - 每个智能体都在尝试获得最大的收益
  - 多个智能体可以追求一个共同目标
  - 每个智能体可以有自己的目标，这些目标可能不同甚至冲突





多智能体系统 (Multi-Agent System, MAS)



[From motor control to team play in simulated humanoid football]

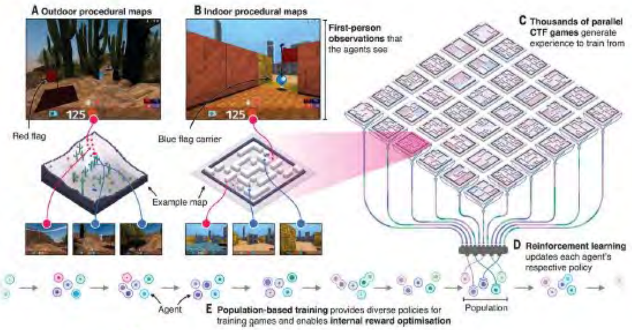
多智能体任务

图片来源：人工智能微课十讲

多智能体任务

AlphaGo系列工作

Capture-the-flag



OpenAI Five: Dota 2 5V5

绝悟王者荣耀AI

2016年1月

2018年7月

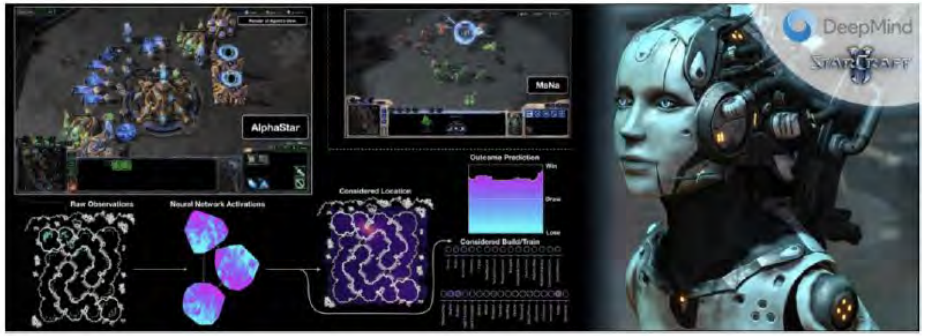
2019年1月

2019年4月

2019年7月

2019年9月

2020年11月



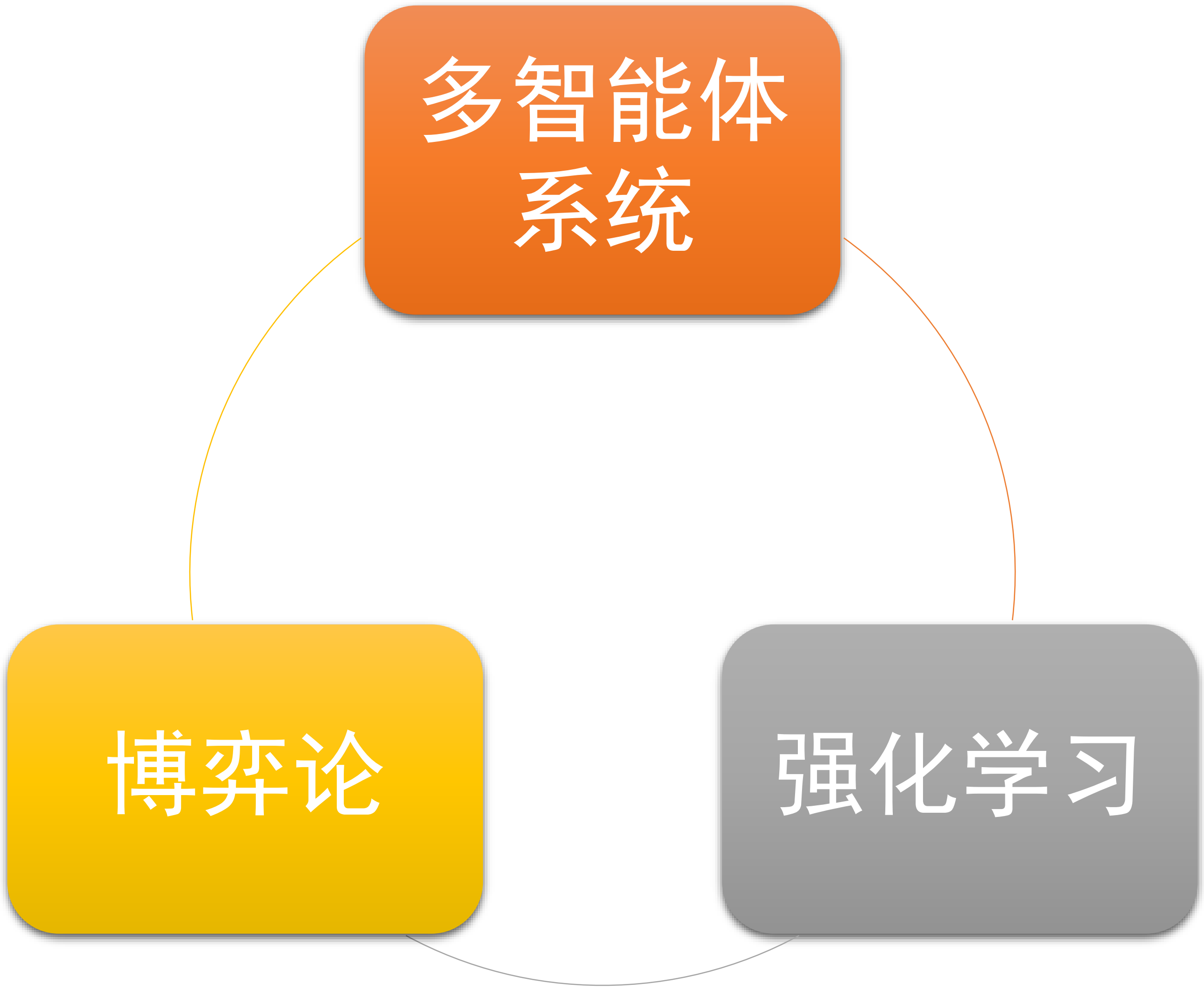
AlphaStar:星际争霸全局游戏

Pluribus德州扑克

Hide and Seek

图片来源：《多智能体系统》

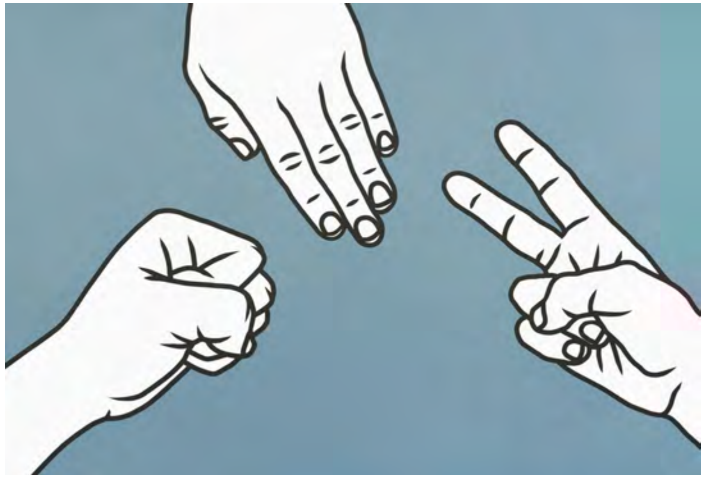
多智能体系统的三大要素



博弈论 (Game Theory)

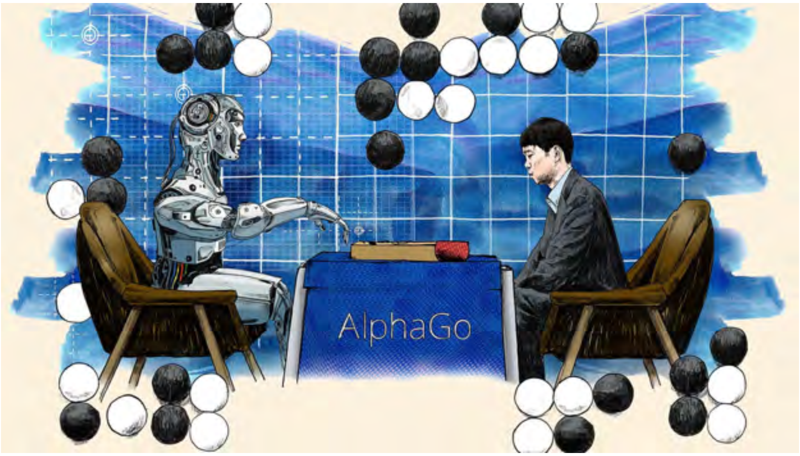
北京大学 PEKING UNIVERSITY

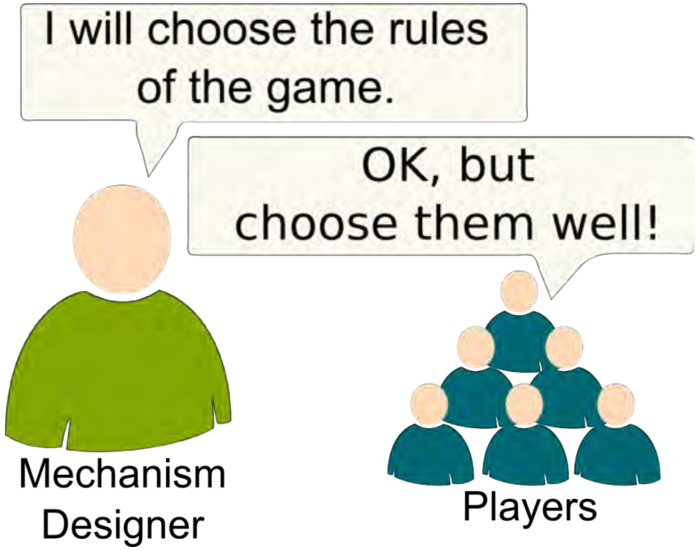
- 研究多个参与者（决策主体）间如何进行交互决策，以及这种决策的均衡问题
  - 参与者之间决策与行为形成相互影响的关系
  - 每个参与者都是理性的，且对收益有充分的认识
  - 每个参与者的决策过程需要考虑其他主体的行动策略
  - 每个参与者都希望其偏好获得最大的满足



博弈论 (Game Theory)

- 博弈论的应用
  - 智能体设计 (Agent Design)
    - 使用博弈论来构建构建智能体的决策，并计算决策效用
  - 机制设计 (Mechanism Design)
    - 定义环境的规则，在每个智能体都在做理性决策时，最大化所有智能体整体的收益





合作博弈与非合作博弈

- 参与者行为交互过程中，是否能够达成一个具有约束力的协议
- 非合作博弈 (Non-cooperative Game)
  - 零和博弈或负和博弈
  - 强调参与者个人行为，智能体之间互相独立，
  - 不假定存在有约束力的协议，每个智能体自行决定是否与其他智能体竞争或合作，以满足个体利益最大化

合作博弈与非合作博弈

北京大学 PEKING UNIVERSITY

- 参与者行为交互过程中，是否能够达成一个具有约束力的协议
- 非合作博弈 (Non-cooperative Game)
- 合作博弈 (Cooperative Game)
  - 正和博弈
  - 强调参与者集体或联盟，研究参与者如何形成联盟，以及如何分配收益
  - 假定存在有约束力的协议，并保证协议能够实施

静态博弈与动态博弈

- 参与者行动是否有先后顺序
- 静态博弈 (Static Game)
  - 同时决策，同时行动
  - 或者不同时行动，但是后行动者不知道先行动者的行动
- 动态博弈 (Dynamic Game)
  - 行动有先后，后行动者可以观察到先行动者的动作

完全信息博弈与不完全信息博弈

- 参与者是否获取参与者的特征，策略空间及回报函数等信息
- 完全信息博弈 (Complete Information Game)
  - 不存在不确定因素
- 不完全信息博弈 (Incomplete Information Game)
  - 存在不确定因素

零和博弈与一般和博弈

- 博弈双方的决策损失与收益是否对应
- 零和博弈
  - 一方获益，另一方必受损
- 一般和博弈（非零和博弈）
  - 双方可能同时获益或受损

单步博弈与重复博弈

- 整个博弈问题需要做多少次决策。
- 单步博弈
  - 每个智能体只做一次决策，最大化单次收益
- 重复博弈
  - 多个或多次单步博弈符合，最大化长期收益
  - 每个参与者会考虑当前行为对其他参与者决策的影响

博弈问题的分类

博弈论

协议	信息	顺序	收益	决策次数
• 合作博弈 • 非合作博弈	• 完全信息博弈 • 不完全信息博弈	• 静态博弈 • 动态博弈	• 零和博弈 • 一般和博弈	• 单次博弈 • 重复博弈

博弈问题的要素

- 参与者（player）或智能体
- 策略（strategy）和策略集 (strategy set)
  - 策略：参与者选择行动（action）
  - 策略集：参与者所有策略的集合，或行动空间
  - **纯策略** (pure strategy)：确定性的选择一个动作
  - **混合策略** (mixed strategy)：随机性的策略，按给定概率选择某个动作（纯策略）
  - 策略组合 (strategy profile): 所有参与者选择的策略的总和
- 收益函数（payoff function）
  - 每个参与者在每种行动组合下各自的效用
  - 对于混合策略，需要计算期望收益

正则形式博弈（Normal-Form Game）

- 描述博弈的一种方式
- 参与者集合 $N = \{1,2,\ldots n\}$
- 策略集合 $S = S_1 \times S_2 \times \cdots \times S_n$
- 收益函数 $u_i: S \to \mathbb{R}$
  - 对于双人博弈，收益函数通常写成矩阵 (payoff matrix)

正则形式博弈（Normal-Form Game）

北京大学 PEKING UNIVERSITY

- 收益矩阵
  - 行/列对应两个参与者的行为
  - 矩阵元素代表相关参与者的收益

	A: 剪刀	A: 石头	A: 布
B: 剪刀	A=0, B=0	A=1, B=-1	A=-1, B=1
B: 石头	A=-1, B=1	A=0, B=0	A=1, B=-1
B: 布	A=1, B=-1	A=-1, B=1	A=0, B=0

例子：猜拳

囚徒困境 (Prisoner’s Dilemma)

- 两个嫌犯 A 和 B 被抓获，并被单独审问，互相无法交流。
- 检查官对两个人分别提出了一个建议：如果一方检举另一方为主犯，则他可以被释放，而被指为主犯的人获刑10年；但是如果双方同时检举对方为主犯，则他们都将获刑5年。
- 同时AB知道：如果双方都保持沉默，目前检察官的证据只够让他们获刑1年
- 是否检举?

	A:**检举**	A: 沉默
B:检举	$A=-5,B=-5$	$A=-10,B=0$
B: 沉默	$A=0,B=-10$	$A=-1,B=-1$

囚徒困境 (Prisoner’s Dilemma)

- 囚徒困境是一种**完全信息静态非合作博弈**
  - 对不对?

	A:**检举**	A: 沉默
B:检举	$A=-5, B=-5$	$A=-10, B=0$
B: 沉默	$A=0, B=-10$	$A=-1, B=-1$

囚徒困境 (Prisoner’s Dilemma)

北京大学 PEKING UNIVERSITY

- **占优策略**（Dominant Strategy）
  - 对某参与者来说，如果对于其他参与者的每个策略选择，策略 s 总是好于策略 s' ，此时对于该参与者， s **强占优**（strong dominate）于 s'
  - 如果 s 在至少1个策略组合上占优，其他组合上**不弱于** s'，则s **弱占优**（weak dominate）于 s'
- 理性参与者：总是选择占优策略，避免占劣策略

	A:**检举**	A: 沉默
B:检举	$A=-5, B=-5$	$A=-10, B=0$
B: 沉默	$A=0, B=-10$	$A=-1, B=-1$

因徒困境 (Prisoner's Dilemma)

- 占优策略均衡（Dominant Strategy Equilibrium）：所有参与者都选择了占优策略的情况
  - 对A和B来说，检举是该博弈的占优策略
  - **均衡**，表示没有参与者想改变自己的策略，因为任意改变都会降低收益
- 因徒困境：双方都选择了占优策略，但是结果并不是最优

	A:**检举**	A: 沉默
B:检举	$A=-5, B=-5$	$A=-10, B=0$
B: 沉默	$A=0, B=-10$	$A=-1, B=-1$

帕累托最优（Pareto Optimality）

- 如果没有其他结果可以在不损害他人利益的情况下，使一个参与者变得更好，那么这个结果是帕累托最优
- 囚徒困境：
  - A、B同时选择“沉默”是帕累托最优
  - 占优策略均衡不是帕累托最优

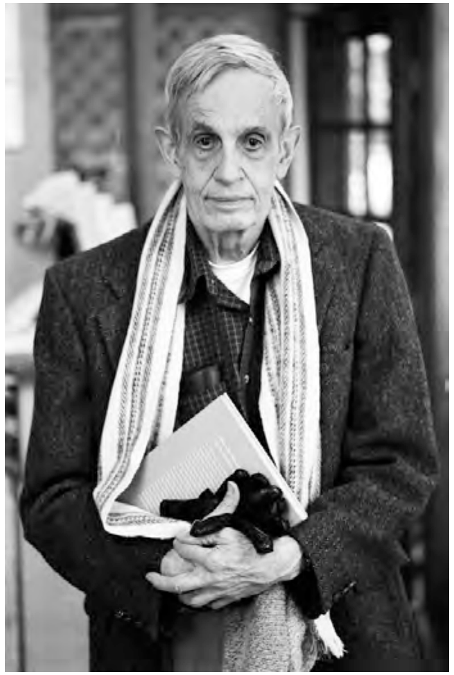
	A:检举	A: 沉默
B:检举	$A=-5, B=-5$	$A=-10, B=0$
B: 沉默	$A=0, B=-10$	$A=-1, B=-1$

囚徒困境 (Prisoner’s Dilemma)

- 现实例子：
  - 军备竞赛
  - 气候政策
  - 公地悲剧 (Tragedy of Commons)
  - 内卷
  - ……

纳什均衡（Nash Equilibrium）

- 在某个**策略组合**下，如果**其他参与者保持不变**，任何一个参与者都无法通过**单方面改变自己的策略**来而**获得更高的收益**，那么这个策略组合为**纳什均衡点**
- 纳什均衡是博弈中的一个**局部稳定点**
- 一个博弈中可能包括多个纳什均衡点，也可能一个也没有
- 占优策略均衡一定是纳什均衡



约翰·福布斯·纳什
诺贝尔经济学奖（1994）

纳什均衡（Nash Equilibrium）

- 占优策略?
- 帕累托最优?
- 纳什均衡?

	A: x	A: y
B: u	$A=-5, B=-5$	$A=-10, B=0$
B: v	$A=0, B=-10$	$A=-1, B=-1$

纳什均衡（Nash Equilibrium）

北京大学 PEKING UNIVERSITY

- 占优策略?
- 帕累托最优?
- 纳什均衡?

	A: x	A: y
B: u	$A = 10, B = 10$	$A = 0, B = 0$
B: v	$A = 0, B = 0$	$A = 1, B = 1$

纳什均衡（Nash Equilibrium）

- 占优策略?
- 帕累托最优?
- 纳什均衡?

	A: x	A: y
B: u	$A = 1, B = -1$	$A = -1, B = 1$
B: v	$A = -1, B = 1$	$A = 1, B = -1$

纳什均衡的存在性

- 定理（纳什，1950）：
  - 在任何有限博弈（即参与者和策略集合均有限的博弈问题）中，都存在至少一个纳什均衡
  - 下面例子的纳什均衡点：A、B都以0.5的概率选择一个动作
    - 思考：为什么这个是纳什均衡点？是否可能有其他的纳什均衡点？

	A: x	A: y
B: u	$A=1, B=-1$	$A=-1, B=1$
B: v	$A=-1, B=1$	$A=1, B=-1$

如何计算纳什均衡?

- 遍历检查所有可能策略组合?
  - 穷举法，通常不可行
- 短视最佳反应（Myopic Best Response）
  迭代最佳反应（Iterative Best Response）
  - 从随机组合出发
    - 如果某个参与者没有处于最优选择，则调整他的选择为最优选择
    - 重复上述过程
  - 类比：循环坐标下降
  - 注意：可能不收敛

	A: x	A: y
B: u	$A=-5, B=-5$	$A=-10, B=0$
B: v	$A=0, B=-10$	$A=-1, B=-1$

如何计算纳什均衡?

- 遍历检查所有可能策略组合?
  - 穷举法，通常不可行
- 短视最佳反应（Myopic Best Response）
迭代最佳反应（Iterative Best Response）
  - 从随机组合出发
    - 如果某个参与者没有处于最优选择，则调整他的选择为最优选择
    - 重复上述过程
  - 类比：循环坐标下降
  - 注意：可能不收敛

	A: x	A: y
B: u	$A=1, B=-1$	$A=-1, B=1$
B: v	$A=-1, B=1$	$A=1, B=-1$

双人零和博弈（Zero-Sum Game)

- 所有参与者的效用之和为 0 （或为一常数）
- 例子：双指猜拳
  - 两个人 (O, E) 同时出1或2根手指，如果总数是奇数，则O获胜，否则E获胜，胜者获得等于手指数的奖励，败者得到相同数量的惩罚

  - 只考虑纯策略下的纳什均衡点？ ✗

	O: *one*	O: *two*
E: *one*	$E=+2, O=-2$	$E=-3, O=+3$
E: *two*	$E=-3, O=+3$	$E=+4, O=-4$

双人零和博弈（Zero-Sum Game）

- 所有参与者的效用之和为 0
- 例子：双指猜拳
  - 两个人 (O, E) 同时出1或2根手指，如果总数是奇数，则O获胜，否则E获胜，胜者获得等于手指数的奖励，败者得到相同数量的惩罚
  - 混合策略下的纳什均衡点?

	O: *one*	O: *two*
E: *one*	$E=+2, O=-2$	$E=-3, O=+3$
E: *two*	$E=-3, O=+3$	$E=+4, O=-4$

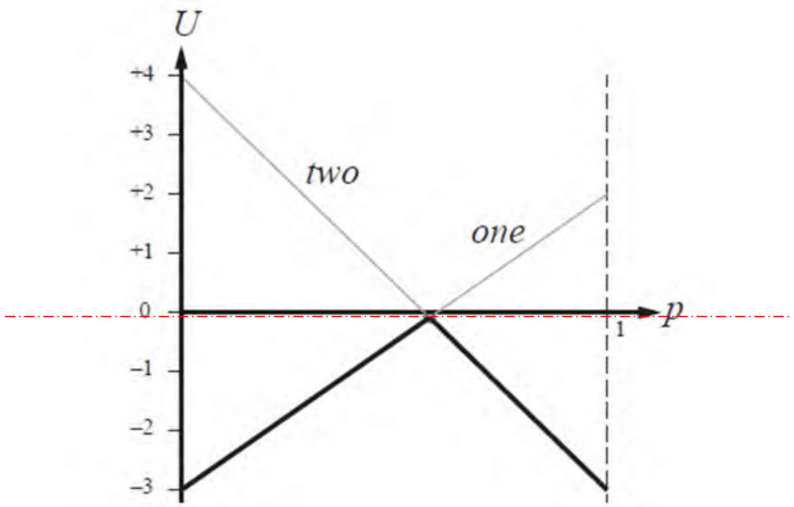
双人零和博弈（Zero-Sum Game)

- 求双指猜拳问题在混合策略下的纳什均衡点
  - 假设 E 选择 one 的概率是 p, O 选择 one 的概率是 q
  - 对于 **O**，在已知 E 的策略参数 p 的情况下，使得 E 收益最小的选择为

$$\arg\min\big(2p-3(1-p),-3p+4(1-p)\big)$$

  - p 的选择应使上式最大，如下图，可有 $p=7/12$, 收益为 $-1/12$

$-\frac{1}{12}$



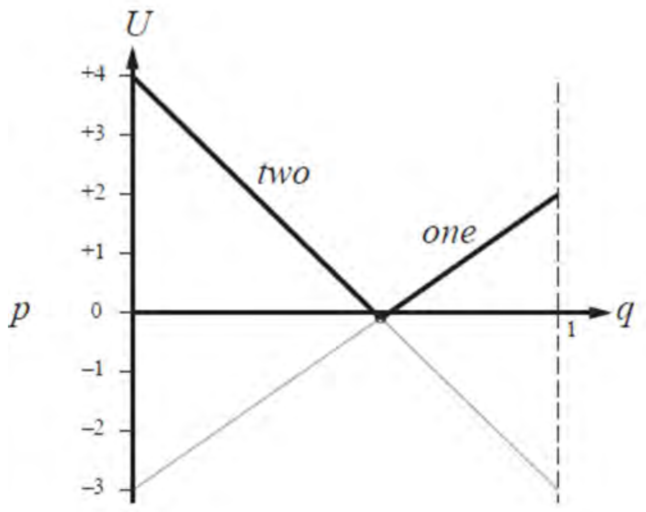
	O: *one*	O: *two*
E: *one*	$E=+2, O=-2$	$E=-3, O=+3$
E: *two*	$E=-3, O=+3$	$E=+4, O=-4$

双人零和博弈（Zero-Sum Game）

- 求双指猜拳问题在混合策略下的纳什均衡点
  - 假设 E 选择 one 的概率是 p, O 选择 one 的概率是 q
  - 对于 **E**，在已知 O 的策略参数 q 的情况下，使得 E 收益最大的选择为

$$\arg\max\big(2q-3(1-q),-3q+4(1-q)\big)$$

  - q 的选择应使上式最小，如下图，同样可有 $q=7/12$, E的收益为 $-1/12$



- $p=7/12\,,q=7/12$ 是一个纳什均衡
- 此时 O 的收益是多少?

	O: *one*	O: *two*
E: *one*	$E=+2, O=-2$	$E=-3, O=+3$
E: *two*	$E=-3, O=+3$	$E=+4, O=-4$

双人零和博弈（Zero-Sum Game)

- 求双指猜拳问题在混合策略下的纳什均衡点
  - 假设 E 选择 one 的概率是 p, O 选择 one 的概率是 q
  - 记 $\boldsymbol{p} = [p, 1-p]^T, \boldsymbol{q} = [q, 1-q]^T$ 为一组混合策略，定义

$$V_{\text{maximin}} = \max_{p} \min_{q} \boldsymbol{p}^T M \boldsymbol{q} \qquad V_{\text{minimax}} = \min_{q} \max_{p} \boldsymbol{p}^T M \boldsymbol{q}$$

- 从 E 的视角看：
  - V_{maximin}代表最好的最差情况， V_{minimax} 代表最差的最好情况
  - $V_{\text{maximin}} \leq V_{\text{minimax}}$

	O: *one*	O: *two*
E: *one*	$E=+2, O=-2$	$E=-3, O=+3$
E: *two*	$E=-3, O=+3$	$E=+4, O=-4$

极小化极大理论 (Minimax Theorem)

- 定理（冯诺伊曼，1928）：对任意双人零和博弈，记 M 为收益矩阵，p_1 为玩家 1 的混合策略，p_2 为玩家 2 的混合策略。若记

$$V_{\text{maximin}} = \max_{p_1} \min_{p_2} p_1^T M p_2 \quad (1)$$

$$V_{\text{minimax}} = \min_{p_2} \max_{p_1} p_1^T M p_2 \quad (2)$$

则有：$V_{\text{maximin}} = V_{\text{minimax}}$

此时(1)的解 p_1^* 和 (2) 的解 p_2^* 构成纳什均衡

约翰·冯·诺伊曼

合作博弈

- 问题关注的基本单元是参与者的组合，以及组合整体的收益
  - 非合作博弈关注对象是每个参与者本身
- 每个参与者仍然是理性的，并且试图最大化自己的收益

合作博弈例子

- 某公司四个股东 A, B, C, D 各有 45%， 25%， 15%， 15% 的股份，他们需要投票决定是否以及如何对今年公司的100万收益进行分红。要通过分配方案至少需要 51% 的股份。如果未能通过，则每个人的分红都是 0 。
- 什么样的分配方案可能获得通过?

具有可转移效用的合作博弈

- 可转移效用 (transferable utility)
  - 所有参加者都认同某种收益形式（例如现金），收益可以在参加者之间自由转移
- 具有可转移效用的合作博弈，可定义为 $G = (N, v)$
  - 参加者集合 $N = \{1,2,\ldots,n\}$
  - 特征函数 (characteristic function) v
    - 对 N 的每个子集 $C \subseteq N$
    - $v(C)$ 给出 C 如果一起工作所能获得的总收益
    - $v(\emptyset) = 0$

合作博弈例子

北京大学 PEKING UNIVERSITY

- 某公司四个股东 A, B, C, D 各有 45%， 25%， 15%， 15% 的股份，他们需要投票决定是否以及如何对今年公司的100万收益进行分红。要通过分配方案至少需要 51% 的股份。如果未能通过，则每个人的分红都是 0 。
- 合作博弈问题： $G = (N, v)$

- $N = \{A, B, C, D\}$
- $v(\{A\}) = v(\{B\}) = v(\{C\}) = v(\{D\}) = 0$
- $v(\{A, B\}) = 100,\ v(\{A, C\}) = 100,\ v(\{A, D\}) = 100$
- $v(\{B, C, D\}) = 100,\ v(\{A, B, C, D\}) = 100$

合作博弈的关键概念

- $C \subseteq N$ 一般被称为**联盟**（coalition）
  - N 本身也是联盟 → 大联盟 (Grand coalition)
- **划分** (partition)：每个参加者必须且仅加入一个联盟，构成的组合
- 每种划分构成一个**联盟结构**（coalition structure）

$N = \{1,2,3\}$ ⇨

可能的联盟: $\{1\},\{2\},\{3\},\{1,2\},\{1,3\},\{2,3\},\{1,2,3\}$

可能的联盟结构: $(\{1\},\{2\},\{3\}),(\{1,2,3\})$

$(\{1\},\{2,3\}),(\{2\},\{1,3\}),(\{3\},\{1,2\})$

合作博弈的关键概念

- 给出合作博弈问题： $G = (N, v)$， 参与者需要决定
  - 选择加入哪个联盟
  - 选择如何分配联盟整体获得的 $v(C)$
  - 结果表示为 $(CS, \boldsymbol{x})$
    - CS: 联盟结构， $\boldsymbol{x}$: 收益向量 (payoff vector), $\boldsymbol{x} = (x_1, \ldots, x_n)$

例如： $\left(\{\{1\},\{2,3\}\},(4,5,5)\right)$

1 选择不结盟，获益4

2,3 选择组成联盟，总获益10，且他们选择平分

合作博弈的分类

- 超可加性博弈（Superadditive Games）
  - $v(C \cup D) \geq v(C) + v(D)$ 对任意 $C, D \subseteq N$ 成立
  - 合作的收益不低于单干的收益

- 是否超可加性 ➔ 参加大联盟 N 总是最优决策?

合作博弈的分类

- 凸博弈（Convex Games）
  - $v(C \cup D) \geq v(C) + v(D) - v(C \cap D)$ 对任意 $C, D \subseteq N$ 成立
  - 超可加性博弈的特例

- 简单博弈（Simple Games）
  - $v(C) \in \{0, 1\}$, 对任意 $C \subseteq N$ 成立
  - 常用于投票问题

合作博弈的目标

- 合作博弈主要需要解决的问题是如何**分配**可能的收益使得组成的联盟的每个人收益最大化
- 收益向量 (payoff vector) $x = (x_1, x_2, \dots, x_n)$
- 理想的收益向量应满足以下条件
  - $\sum_{i=1}^{n} x_i = v(N)$：联盟的所有收益都必须分配
  - $x_i \geq v(\{i\})$：符合**个体理性**（individual rationality），即每个参与者参与合作的收益必须不低于不参与合作的收益

合作博弈的目标

- 什么是一个好的分配方案?
- 稳定：每个成员的收益足够好，从而不会离开而加入其他的联盟
- 公平：每个成员的收益反应了他们的贡献

关于分配方案

- 合作博弈 (N, v) 的核（core）
  - 由收益向量 $x = \{x_1, \ldots, x_n\}$ 构成的集合
  - 对每个可能的联盟 $C \subseteq N$，x 必须满足

$$x(C) := \sum_{i \in C} x_i \geq v(C)$$

关于分配方案

$$x(C) \coloneqq \sum_{i\in C} x_i \geq v(C), \qquad \forall C \subseteq N$$

- 博弈问题的核体现了分配方案的稳定性
  - 对于核以外的向量 x'，存在 $C \subseteq N$ 使得 $x(C) < v(C)$
  - 此时C的成员将倾向于不加入大联盟
- 博弈问题的核非空，表示大联盟可以成立
  - 检查核非空性可能是一个NP完全问题
  - 凸博弈问题一定有非空核

合作博弈例子

- 某公司四个股东 A, B, C, D 各有 45%， 25%， 15%， 15% 的股份，他们需要投票决定是否以及如何对今年公司的100万收益进行分红。要通过分配方案至少需要 51% 的股份。如果未能通过，则每个人的分红都是 0 。
- 合作博弈问题： $G = (N, v)$

- $N = \{A, B, C, D\}$
- $v(\{A\}) = v(\{B\}) = v(\{C\}) = v(\{D\}) = 0$
- $v(\{A, B\}) = 100,\ v(\{A, C\}) = 100,\ v(\{A, D\}) = 100$
- $v(\{B, C, D\}) = 100,\ v(\{A, B, C, D\}) = 100$

收益向量 $x = (??)$ 属于核

合作博弈例子

- 博弈问题 $G = (N, v)$

$$N = \{1,2\}$$

$$v(\{1\}) = v(\{2\}) = 5$$

$$v(\{1,2\}) = 20$$

收益向量 $x = (??)$ 属于核

合作博弈例子

- 博弈问题 $G = (N, v)$

$$N = \{1,2\} \qquad \begin{gathered} v(\{1\}) = v(\{2\}) = 5 \\ v(\{1,2\}) = 20 \end{gathered}$$

- $x = (6,14)$ 是核中的的收益向量，所以大联盟能够成立
- 但是这种分配明显不公平
- 核的概念只能表明什么时候可以形成大联盟，但是不处理分配公平性问题

分配方案的公平性

- 边际贡献（marginal contribution）

$$\mathrm{mc}_i(\mathrm{C}) = \nu(C \cup \{i\}) - \nu(C)$$

  - 即参与者 i 加入联盟 C $(i \notin C)$ 带来的额外收益
  - 直觉上，边际贡献更大的参与者应该获得更多收益分配

分配方案的公平性

- 无效参与者（dummy player）：对联盟没有贡献的参与者
  - 对任意 $C \subseteq N - \{i\}$, $\mathrm{mc}_i(C) = 0$
- 对称参与者（symmetric players)：对任意联盟做出相同贡献的参与者
  - 对任意 $C \subseteq N - \{i, j\}$, $\mathrm{mc}_i(C) = \mathrm{mc}_j(C)$

公平性公理

- 假设$\phi_i(G)$是问题 $G = (N, v)$ 的一个分配方案，公平性公理要求：
  - 所有收益必须得到分配 $\sum_{i \in N} \phi_i(G) = v(N)$
  - 无贡献者不分配收益： $\phi_i(G) = 0, \forall i \in \text{dummy}$
  - 对称参与者应得到相同收益： $\phi_i(G) = \phi_j(G), \forall \text{ symmetric } (i, j)$
  - 分配方案是线性可加的：
    - 对博弈问题 $G = (N, v), G' = (N, v')$ ，定义 $G'' = G + G' = (N, v + v')$
    - 则 $\phi_i(G + G') = \phi_i(G) + \phi_i(G'), \forall\, i \in N$

沙普利值（Shapley Value）

- 直接按照 $\mathrm{mc}_i(N)$ 比例进行分配可能是不合理的
  - 每个参与者的贡献可能跟加入联盟的顺序有关
  - “压垮骆驼的最后一根稻草”
- 沙普利值（Shapley Value）
  - 考察所有可能的入顺序，对应边际贡献的平均值
  - 记收益向量为 $\phi(G) = \left(\phi_1(G), \phi_2(G), \dots, \phi_n(G)\right)$

$$\phi_i(G) = \sum_{C \subseteq N \setminus \{i\}} \frac{|C|!\,(n - |C| - 1)!}{n!} \mathrm{mc}_i(C)$$

- 沙普利值（Shapley Value）是满足公平性公理的唯一解

劳埃德·沙普利

机制设计

- 通过设计合适的规则，使一组智能体自发的完成特定的目标
  - 如何评估一个规则的实际效果？
  - 如何某个目标自动的设计规则？
- 例子：
  - 通过设计拍卖规则，使每个参与者在过程中暴露自己对标的的真实估价（真值暴露）
    - 使标的价格能够达到所有参与者对其估价的最高值
    - 避免参与者合谋达到不公平的协约
    - 在某些情况下，避免参与者随意报价

重要机制举例：拍卖

- 拍卖（Auction）
  - 竞标人（bidder）
  - 私有价值（private value）：竞拍人对标的的估价
  - 竞标人依次出价，每个参与者出价不会高于私有价值

- 增价拍卖（ascending-bid auction）
  - 也叫英式拍卖（English auction）
  - 从最低出价 $b_{\min}$ 开始，每轮出价不低于增量 d
  - 最后出价者胜出

重要机制举例：拍卖

- 增价拍卖的问题
  - 不能完全暴露真值
  - 特定情况下，不能保证最终价格是最优
    - 例如增量过大情况，可能导致估价最高的人无法参与最后一轮竞标
  - 特定情况下，参加者可能会合谋
    - 1999年德国拍卖手机频段的问题

重要机制举例：拍卖

- 密封投标拍卖（Sealed-bid auction）
  - 不公开每个参与者的竞标价格
  - 每个参与者的占优策略是猜测其他参与者的报价，并报一个稍高的价格
  - 需要花费成本去猜测其他参与者的报价

- 次价密封投标拍卖（sealed-bid second-price auction）
  - 维克里拍卖(Vickrey auction)
  - 中标者支付第二高的价格，而非自己的价格
  - 真值暴露，占优策略是直接按私有价值出价，避免猜测成本
  - 广泛应用于分布式人工智能系统

重要机制举例：拍卖

- 学校要给宿舍安装空调，但是数量有限，同时有希望给最需要的宿舍安装。为此让每个宿舍给空调报价，价高者得到空调。
- VCG拍卖：
  - 每个宿舍上报价格 v_i
  - 按照空调数量分配给报价最高的宿舍
  - 学校根据报价和分配情况，计算胜者的损失

重要机制举例：拍卖

- VCG (Vickrey-Clarke-Groves) 拍卖
  - 按照获胜者给失败者带来的损失进行计费
  - 最大化社会效用
- 例如：学校要给宿舍安装空调，但是数量有限，同时有希望给最需要的宿舍安装。为此让每个宿舍给空调报价，价高者得到空调。
  - 有什么问题？
    - 参与者有动机随意报高价
    - 无法衡量真实需求

本章内容

- 多智能体概览
- 博弈论基础
  - 非合作博弈：纳什均衡
  - 合作博弈：公平性公理与沙普利值
  - 机制设计：拍卖

谢谢

北京大學
PEKING UNIVERSITY