



# 基础博弈论

主讲：李林静 研究员

中国科学院自动化研究所  
多模态人工智能系统全国重点实验室  
互联网大数据与安全信息学研究中心(iBASIC)

中国科学院大学人工智能学院

北京·雁栖湖

2023-4-11

# 课程内容

## ■ 博弈案例

- 智猪博弈、公地悲剧、田忌赛马
- 拍卖

## ■ 博弈论

- 选择与效用
- 完全信息静态、完全信息动态
- 不完全信息静态、不完全信息动态
- 合作博弈论

## ■ 博弈论与人工智能

# 概念总结

- 均衡
- 博弈矩阵
- 优势策略、劣势策略
- 混合策略
- 最优反应
- 证明某策略组合是均衡的方法
- 计算混合策略的方法

# 课程内容

## ■ 博弈案例

- 智猪博弈、公地悲剧、田忌赛马
- 拍卖

## ■ 博弈论

- 选择与效用
- 完全信息静态、完全信息动态
- 不完全信息静态、不完全信息动态
- 合作博弈论

## ■ 博弈论与人工智能

# 博弈求解—非合作博弈

静态

动态

完全信息

( 初始条件可知性 )

完全信息静态博弈

解概念：纳什均衡  
典型博弈：GANs

不完全信息

不完全信息静态博弈

解概念：贝叶斯纳什均衡  
典型博弈：暗标拍卖

完全信息动态博弈

解概念：子博弈完美纳什均衡  
典型博弈：象棋、围棋

不完全信息动态博弈

解概念：完美贝叶斯均衡  
典型博弈：德州扑克

完全但不完美信息动态博弈

解概念：完美贝叶斯均衡  
典型博弈：星际

不完全且不完美信息动态博弈

解概念：完美贝叶斯均衡  
典型博弈：麻将

完美信息

不完美信息

海萨尼转换



完全但不完美信息动态博弈

# 完全信息静态博弈—纳什均衡

完全信息静态博弈  $\mathcal{G} = (\mathcal{I}, \mathcal{A}, u)$  , 策略型 (Strategic Form)

- $\mathcal{I} = \{1, 2, \dots, I\}$ , 参与人集合
- 智猪博弈 ,  $\mathcal{I} = \{\text{大猪、小猪}\}$
- 公地悲剧 ,  $\mathcal{I} = \{1、2、3、\dots、N\}$
- 田忌赛马 ,  $\mathcal{I} = \{\text{齐威王、田忌}\}$

# 完全信息静态博弈—纳什均衡

完全信息静态博弈  $\mathcal{G} = (\mathcal{I}, \mathcal{A}, \mathbf{u})$

- $\mathcal{I} = \{1, 2, \dots, I\}$ , 参与人集合
- $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times \dots \times \mathcal{A}_I$ , 行动(Action)空间,  $\mathcal{A}_i$ 为  $i$  全部可能行动的集合
- 智猪博弈,  $\mathcal{A}_{\text{大}} = \mathcal{A}_{\text{小}} = \{\text{等待、按钮}\}$ 
$$\mathcal{A} = \mathcal{A}_{\text{大}} \times \mathcal{A}_{\text{小}} = \{(\text{等待、按钮}), (\text{等待、等待}), (\text{按钮、按钮}), (\text{按钮、等待})\}$$
- 公地悲剧,  $\mathcal{A}_i = \{0, 1, 2, 3, \dots, 96\}$ 
$$\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times \dots \times \mathcal{A}_N = \{0, 1, 2, 3, \dots, 96\}^N$$
- 田忌赛马,  $\mathcal{A}_{\text{王}} = \{\text{ABC、ACB、BAC、BCA、CAB、CBA}\}$ 
$$\mathcal{A}_{\text{忌}} = \{\text{abc、acb、bac、bca、cab、cba}\}$$
$$\mathcal{A} = \mathcal{A}_{\text{王}} \times \mathcal{A}_{\text{忌}} = \{(\text{ABC、abc}), (\text{ABC、acb}), \dots, (\text{CBA、cba})\}$$

# 完全信息静态博弈—纳什均衡

## 完全信息静态博弈 $\mathcal{G} = (\mathcal{I}, \mathcal{A}, \mathbf{u})$

➤  $\mathcal{I} = \{1, 2, \dots, I\}$ , 参与人集合

➤  $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times \dots \times \mathcal{A}_I$ , 行动(Action)空间,  $\mathcal{A}_i$  为  $i$  全部可能行动的集合

➤ 纯策略, 参与人  $i$  的纯策略  $s_i$  即其可选的行动  $s_i \in \mathcal{A}_i$

➤ 纯策略组合,  $\mathbf{s} = (s_i, s_{-i}) = (s_1, s_2, \dots, s_I)$ ,  $s_i \in S_i$ , 纯策略空间  $S_i = \mathcal{A}_i$

$$\mathcal{S} = S_1 \times S_2 \times \dots \times S_I$$

➤ 智猪博弈,  $S_{\text{大}} = S_{\text{小}} = \{\text{等待、按钮}\}$

$$\mathcal{S} = S_{\text{大}} \times S_{\text{小}} = \{(\text{等待、按钮}), (\text{等待、等待}), (\text{按钮、按钮}), (\text{按钮、等待})\}$$

➤ 公地悲剧,  $S_i = \{0, 1, 2, 3, \dots, 96\}$ ,  $\mathbf{s} = (s_i, s_{-i}) = (s_1, s_2, \dots, s_I)$

$$\mathcal{S} = S_1 \times S_2 \times \dots \times S_N = \{0, 1, 2, 3, \dots, 96\}^N$$

➤ 田忌赛马,  $S_{\text{王}} = \{\text{ABC、ACB、BAC、BCA、CAB、CBA}\}$

$$S_{\text{忌}} = \{\text{abc、acb、bac、bca、cab、cba}\}$$

$$\mathcal{S} = S_{\text{王}} \times S_{\text{忌}} = \{(\text{ABC、abc}), (\text{ABC、acb}), \dots, (\text{CBA、cba})\}$$

# 完全信息静态博弈—纳什均衡

完全信息静态博弈  $\mathcal{G} = (\mathcal{I}, \mathcal{A}, \mathbf{u})$

- $\mathcal{G} = \{1, 2, \dots, I\}$ , 参与人集合
- $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times \dots \times \mathcal{A}_I$ , 行动(Action)空间,  $\mathcal{A}_i$  为  $i$  全部可能行动的集合
  - 纯策略, 参与人  $i$  的纯策略  $s_i$  即其可选的行动  $s_i \in \mathcal{A}_i$
  - 纯策略组合,  $\mathbf{s} = (s_i, s_{-i}) = (s_1, s_2, \dots, s_I)$ ,  $s_i \in S_i$ , 纯策略空间  $S_i = \mathcal{A}_i$

$$\mathcal{S} = S_1 \times S_2 \times \dots \times S_I$$

- 混合策略, 参与人  $i$  的混合策略  $\sigma_i$  为  $\mathcal{A}_i$  上的概率分布
- 混合策略空间,  $\Sigma_i = \Delta(\mathcal{A}_i)$ , 全体分布构成的集合
- 混合策略组合,  $\boldsymbol{\sigma} = (\sigma_i, \sigma_{-i}) = (\sigma_1, \sigma_2, \dots, \sigma_I)$ ,  $\sigma_i \in \Sigma_i$

$$\boldsymbol{\Sigma} = \Sigma_1 \times \Sigma_2 \times \dots \times \Sigma_I$$

|       |     | $q_1$        | $q_2$        | $q_3$        | $q_4$        | $q_5$        | $q_6$        |
|-------|-----|--------------|--------------|--------------|--------------|--------------|--------------|
|       |     | abc          | acb          | bac          | bca          | cab          | cba          |
| $p_1$ | ABC | <u>3, -3</u> | 1, -1        | 1, -1        | 1, -1        | <u>-1, 1</u> | 1, -1        |
| $p_2$ | ACB | 1, -1        | <u>3, -3</u> | 1, -1        | 1, -1        | 1, -1        | <u>-1, 1</u> |
| $p_3$ | BAC | 1, -1        | <u>-1, 1</u> | <u>3, -3</u> | 1, -1        | 1, -1        | 1, -1        |
| $p_4$ | BCA | <u>-1, 1</u> | 1, -1        | 1, -1        | <u>3, -3</u> | 1, -1        | 1, -1        |
| $p_5$ | CAB | 1, -1        | 1, -1        | 1, -1        | <u>-1, 1</u> | <u>3, -3</u> | 1, -1        |
| $p_6$ | CBA | 1, -1        | 1, -1        | <u>-1, 1</u> | 1, -1        | 1, -1        | <u>3, -3</u> |

$$\begin{aligned}\boldsymbol{\sigma} &= (\sigma_{\text{王}}, \sigma_{\text{忌}}) = (\mathbf{p}, \mathbf{q}) \\ &= ([p_1, p_2, p_3, p_4, p_5, p_6], [q_1, q_2, q_3, q_4, q_5, q_6])\end{aligned}$$

$$\mathbf{1}^T \mathbf{p} = 1, \mathbf{p} \geq \mathbf{0} \quad \mathbf{1}^T \mathbf{q} = 1, \mathbf{q} \geq \mathbf{0}$$

# 完全信息静态博弈—纳什均衡

完全信息静态博弈  $\mathcal{G} = (\mathcal{I}, \mathcal{A}, \mathbf{u})$

➤  $\mathcal{I} = \{1, 2, \dots, I\}$ , 参与人集合

➤  $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times \dots \times \mathcal{A}_I$ , 行动(Action)空间,  $\mathcal{A}_i$  为  $i$  全部可能行动的集合

➤ 纯策略, 参与人  $i$  的纯策略  $s_i$  即其可选的行动  $s_i \in \mathcal{A}_i$

➤ 纯策略组合,  $s = (s_i, s_{-i}) = (s_1, s_2, \dots, s_I)$ ,  $s_i \in S_i$ , 纯策略空间  $S_i = \mathcal{A}_i$

$$\mathcal{S} = S_1 \times S_2 \times \dots \times S_I$$

➤ 混合策略, 参与人  $i$  的混合策略  $\sigma_i$  为  $\mathcal{A}_i$  上的概率分布

➤ 混合策略组合,  $\sigma = (\sigma_i, \sigma_{-i}) = (\sigma_1, \sigma_2, \dots, \sigma_I)$ ,  $\sigma_i \in \Sigma_i$ , 混合策略空间  $\Sigma_i = \Delta(\mathcal{A}_i)$

$$\Sigma = \Sigma_1 \times \Sigma_2 \times \dots \times \Sigma_I$$

➤  $\mathbf{u} = (u_1, u_2, \dots, u_I)$ ,  $u_i(s)$  为参与人  $i$  在纯策略组合  $s$  上的效用

|    |    | 小猪   |       | $) = p$ |       |              |              |              |              |              |
|----|----|------|-------|---------|-------|--------------|--------------|--------------|--------------|--------------|
|    |    | 等待   | 按钮    | $q_1$   | $q_2$ | $q_3$        | $q_4$        | $q_5$        | $q_6$        |              |
| 大猪 | 等待 | 0, 0 | 9, -1 | abc     | acb   | bac          | bca          | cab          | cba          |              |
|    | 按钮 | 4, 4 | 5, 1  | $p_1$   | ABC   | <u>3, -3</u> | 1, -1        | 1, -1        | -1, 1        |              |
|    |    |      |       | $p_2$   | ACB   | 1, -1        | <u>3, -3</u> | 1, -1        | 1, -1        | -1, 1        |
|    |    |      |       | $p_3$   | BAC   | 1, -1        | -1, 1        | <u>3, -3</u> | 1, -1        | 1, -1        |
|    |    |      |       | $p_4$   | BCA   | -1, 1        | 1, -1        | 1, -1        | <u>3, -3</u> | 1, -1        |
|    |    |      |       | $p_5$   | CAB   | 1, -1        | 1, -1        | 1, -1        | -1, 1        | <u>3, -3</u> |
|    |    |      |       | $p_6$   | CBA   | 1, -1        | 1, -1        | -1, 1        | 1, -1        | <u>3, -3</u> |

# 完全信息静态博弈—纳什均衡

完全信息静态博弈  $\mathcal{G} = (\mathcal{I}, \mathcal{A}, \mathbf{u})$

|       |     | $q_1$        | $q_2$        | $q_3$        | $q_4$        | $q_5$        | $q_6$        |
|-------|-----|--------------|--------------|--------------|--------------|--------------|--------------|
|       |     | abc          | acb          | bac          | bca          | cab          | cba          |
| $p_1$ | ABC | <u>3, -3</u> | 1, -1        | 1, -1        | 1, -1        | <u>-1, 1</u> | 1, -1        |
| $p_2$ | ACB | 1, -1        | <u>3, -3</u> | 1, -1        | 1, -1        | 1, -1        | <u>-1, 1</u> |
| $p_3$ | BAC | 1, -1        | <u>-1, 1</u> | <u>3, -3</u> | 1, -1        | 1, -1        | 1, -1        |
| $p_4$ | BCA | <u>-1, 1</u> | 1, -1        | 1, -1        | <u>3, -3</u> | 1, -1        | 1, -1        |
| $p_5$ | CAB | 1, -1        | 1, -1        | 1, -1        | <u>-1, 1</u> | <u>3, -3</u> | 1, -1        |
| $p_6$ | CBA | 1, -1        | 1, -1        | <u>-1, 1</u> | 1, -1        | 1, -1        | <u>3, -3</u> |

$$\bar{u}_\text{主}[\mathbf{p}, \mathbf{q}] = \mathbf{p}^T \mathbf{A} \mathbf{q}$$

对混合策略  $\sigma$ ，期望效用为

$$\bar{u}_\text{忌}[\mathbf{p}, \mathbf{q}] = \mathbf{q}^T \mathbf{B}^T \mathbf{p}$$

$$\bar{u}_i(\sigma) = \sum_{j_1 j_2 \cdots j_I} \sigma_{1j_1} \sigma_{2j_2} \cdots \sigma_{Ij_I} u(s_{1j_1}, s_{2j_2}, \dots, s_{Ij_I})$$

# 完全信息静态博弈—纳什均衡

## 纳什均衡(NE)

- Nash Equilibrium
- 策略组合 $\sigma^*$ 构成博弈 $\mathcal{G}$ 的纳什均衡, 若在此策略组合上, 所有参与人均没有单方面改变自己策略的激励(Incentive), 即有

$$\bar{u}_i(\sigma_i^*, \sigma_{-i}^*) \geq \bar{u}_i(s_i, \sigma_{-i}^*), \forall s_i \in S_i, \forall i \in \mathcal{I}$$

均衡效用

任意更换己方纯策略

对手策略不变

|    |    | 小猪   |       |
|----|----|------|-------|
|    |    | 等待   | 按钮    |
| 大猪 | 等待 | 0, 0 | 9, -1 |
|    | 按钮 | 4, 4 | 5, 1  |

■ 状态: (按钮, 等待)

- 给定大猪去“按钮”, 小猪的最优策略是“等待”
- 给定小猪只“等待”, 大猪的最优策略是“按钮”
- 大猪、小猪 均不会改变自己的策略

$$(s_1^*, s_2^*) = (32, 32)$$

牧民1

$$100 - (s_1 + s_2) - s_1 - 4 = 0$$

$$100 - (s_1 + 32) - s_1 - 4 = 0$$

$$s_1^* = 32$$

牧民2

$$100 - (s_1 + s_2) - s_2 - 4 = 0$$

$$100 - (s_2 + 32) - s_2 - 4 = 0$$

$$s_2^* = 32$$

# 完全信息静态博弈—纳什均衡

纳什均衡(NE)

- Nash Equilibrium
- 策略组合 $\sigma^*$ 构成博弈 $\mathcal{G}$ 的纳什均衡, 若在此策略组合上, 所有参与人均没有单方面改变自己策略的激励(Incentive), 即有

$$\bar{u}_i(\sigma_i^*, \sigma_{-i}^*) \geq \bar{u}_i(s_i, \sigma_{-i}^*), \forall s_i \in S_i, \forall i \in \mathcal{I}$$

均衡效用

任意更换己方纯策略

对手策略不变

给定对手策略 $\sigma_{-i}$ , 参与人 $i$ 的最优反应策略 $\sigma_i^*$ 可表示为

$$\sigma_i^* \in BR_i(\sigma_{-i}), \forall i \in \mathcal{I}$$

对任意策略组合 $\sigma$ , 定义最优反应映射(可能多值)

$$BR(\sigma) = (\sigma_1^*, \sigma_2^*, \dots, \sigma_I^*), \sigma_i^* \in BR_i(\sigma_{-i}), \forall i \in \mathcal{I}$$

- $BR(\text{按钮}, \text{等待}) = (\text{按钮}, \text{等待}) \quad (s_1^*, s_2^*) = (48, 48) - (s_1, s_2)/2$
- $BR(\text{按钮}, \text{按钮}) = (\text{等待}, \text{等待}) \quad (32, 32) = (48, 48) - (32, 32)/2$
- $BR(\text{等待}, \text{等待}) = (\text{按钮}, \text{等待}) \quad (32, 32) = (48, 48) - (32, 32)/2$
- $BR(\text{等待}, \text{按钮}) = (\text{等待}, \text{等待})$

# 完全信息静态博弈—纳什均衡

纳什均衡(NE)

- Nash Equilibrium
- 策略组合 $\sigma^*$ 构成博弈 $\mathcal{G}$ 的纳什均衡, 若在此策略组合上, 所有参与人均没有单方面改变自己策略的激励(Incentive), 即有

$$\bar{u}_i(\sigma_i^*, \sigma_{-i}^*) \geq \bar{u}_i(s_i, \sigma_{-i}^*), \forall s_i \in S_i, \forall i \in \mathcal{I}$$

均衡效用

任意更换己方纯策略

对手策略不变

给定对手策略 $\sigma_{-i}$ , 参与人 $i$ 的最优反应策略 $\sigma_i^*$ 可表示为

$$\sigma_i^* \in BR_i(\sigma_{-i}), \forall i \in \mathcal{I}$$

对任意策略组合 $\sigma$ , 定义最优反应映射(可能多值)

$$BR(\sigma) = (\sigma_1^*, \sigma_2^*, \dots, \sigma_I^*), \sigma_i^* \in BR_i(\sigma_{-i}), \forall i \in \mathcal{I}$$

纳什均衡为参与人最优反应映射的不动点

以拓扑学中布劳威尔(Brouwer)不动点定理为基础  
在很弱的条件下就可以证明混合策略纳什均衡存在

$\epsilon$ -NE

➤ 单方面改变带来的收益增加不超过 $\epsilon$

# 完全信息静态博弈—纳什均衡



单一纯策略均衡(坦白、坦白)，  
“坦白”是优势策略  
双方可预期，稳定

# 完全信息静态博弈—纳什均衡

|   |  |   |
|---|--|---|
|   | 乙  |   |
| 甲 | 坦白                                       | 不坦白                                     |
|   | 坦白      5年, 5年 ← 0年, 8年<br>↓           ↑ | 不坦白    8年, 0年 ← 1年, 1年<br>↑           ↓ |

单一纯策略均衡(坦白、坦白)，  
“坦白”是优势策略  
双方可预期，稳定

|   |                                      |                      |
|---|--------------------------------------|----------------------|
|   | 乙                                    |                      |
| 甲 | left                                 | Right                |
|   | Up      1, 3 ← 0, 0<br>↑           ↓ | Down     2, 2 → 0, 0 |

双纯策略均衡，  
双方偏好均衡点不同，可能震荡  
 $(2/5, 3/5; 2/3, 1/3)$ 构成混合策略  
 $3p_1 = 2p_2 = 2 - 2p_1, p_1 = 2/5$   
 $q_1 = 2p_2 = 1 - p_2, p_2 = 1/3$

# 完全信息静态博弈—纳什均衡

|   |  |   |
|---|--|---|
|   | 乙  |   |
| 甲 | 坦白                                       | 不坦白                                     |
|   | 坦白      5年, 5年 ← 0年, 8年<br>↓           ↑ | 不坦白    8年, 0年 ← 1年, 1年<br>↑           ↓ |

单一纯策略均衡(坦白、坦白)，  
“坦白”是优势策略  
双方可预期，稳定

|   |                                      |                                       |
|---|--------------------------------------|---------------------------------------|
|   | 乙                                    |                                       |
| 甲 | left                                 | Right                                 |
|   | Up      1, 3 ← 0, 0<br>↑           ↓ | Down     2, 2 → 0, 0<br>↓           ↑ |

双纯策略均衡，  
双方偏好均衡点不同，可能震荡  
 $(2/5, 3/5; 2/3, 1/3)$ 构成混合策略  
 $3p_1 = 2p_2 = 2 - 2p_1, p_1 = 2/5$   
 $q_1 = 2q_2 = 1 - q_2, q_2 = 1/3$

|   |                                      |                                       |
|---|--------------------------------------|---------------------------------------|
|   | 乙                                    |                                       |
| 甲 | left                                 | Right                                 |
|   | Up      2, 1 → 0, 2<br>↑           ↓ | Down     1, 2 ← 3, 0<br>↓           ↑ |

无纯策略均衡  
 $(2/3, 1/3; 3/4, 1/4)$ 构成混合策略  
 $p_1 + 2p_2 = 2p_1, p_1 = 2p_2, p_2 = 1/3$   
 $2q_1 = q_1 + 3q_2, q_1 = 3q_2, q_2 = 1/4$

# 不完全信息静态博弈——贝叶斯纳什均衡

## 博弈描述

- ✓ 完全信息，对参与人的描述就是其可能的行动  $\mathcal{A}_i$
- ✓ 不完全信息，各方存在只有自己知道的参数，称为类其型  $\theta_i \in \Theta_i$
- ✓ 全体参与人类型向量  $\theta$  的先验分布已知

## 策略空间

- ✓ 完全信息，纯策略空间  $S_i =$  行动集合  $\mathcal{A}_i$
- ✓ 不完全信息，考虑参与人的类型，策略为从类型到行动的映射

$$s_i: \Theta_i \rightarrow \mathcal{A}_i$$

- ✓ 纯策略空间  $S_i$  包含全部可能映射 (混合策略空间结构复杂)

两人各发一张牌，比大小定胜负

- ✓ 底牌，不完全信息
- ✓ 综合考虑己方底牌，猜测对手底牌，决定 弃牌/开牌
- ✓ 贝叶斯公式，后验概率

# 不完全信息静态博弈—贝叶斯纳什均衡

## 贝叶斯纳什均衡 (BNE)

- ✓ Bayesian-Nash Equilibrium
- ✓ 根据己方类型，计算条件期望效用
- ✓ 依据条件期望效用，各方没有单方面改变策略的激励(NE条件)

$$\mathbb{E}_{\theta_{-i}|\theta_i}[u_i(s_i^*(\theta_i), s_{-i}^*(\theta_{-i}), \theta_i, \theta_{-i})] \geq \mathbb{E}_{\theta_{-i}|\theta_i}[u_i(s_i(\theta_i), s_{-i}^*(\theta_{-i}), \theta_i, \theta_{-i})],$$

$\forall s_i \in S_i, \forall i \in J$

决策考虑效用  
的条件期望

策略依赖参与人类型

效用依赖参与人类型

- 完全信息静态博弈，可以引入虚拟类型，转化为不完全信息静态博弈
- 纯策略同样为选择行动 (根据类型)
- 混合策略则等价为贝叶斯纳什均衡

# 纳什均衡—贝叶斯纳什均衡

- $N=2$  公地悲剧
- 牧民1的成本  $c=4$
- 牧民2引进了新技术
  - 掌握新技术,  $c^L = 2$
  - 尚未掌握进行技术,  $c^H = 4$
- 牧民2是否完全掌握了新技术, 是私有信息
  - 牧民1认为牧民2掌握了新技术的概率是1/2

# 纳什均衡—贝叶斯纳什均衡

■ 牧民1的期望收益：

$$\begin{aligned}\mathbb{E}[u_1(c_1, c_2)] &= \frac{1}{2} u_1(c, cH) + \frac{1}{2} u_1(c, cL) \\ &= \frac{1}{2} q_1(100 - q_1 - q_2^H - 4) + \frac{1}{2} q_1(100 - q_1 - q_2^L - 4)\end{aligned}$$

$$4q_1 + q_2^L + q_2^H = 96 \times 2$$

# 纳什均衡—贝叶斯纳什均衡

## ■ 牧民2的期望收益：

- 掌握了新技术

$$\mathbb{E}[u_2(c_1, c_2)] = u_2(c, c^L) = q_2^L(100 - q_1 - q_2^L - 2)$$

$$q_1 + 2q_2^L = 98$$

- 尚未掌握新技术

$$\mathbb{E}[u_2(c_1, c_2)] = u_2(c, c^H) = q_2^H(100 - q_1 - q_2^H - 4)$$

$$q_1 + 2q_2^H = 96$$

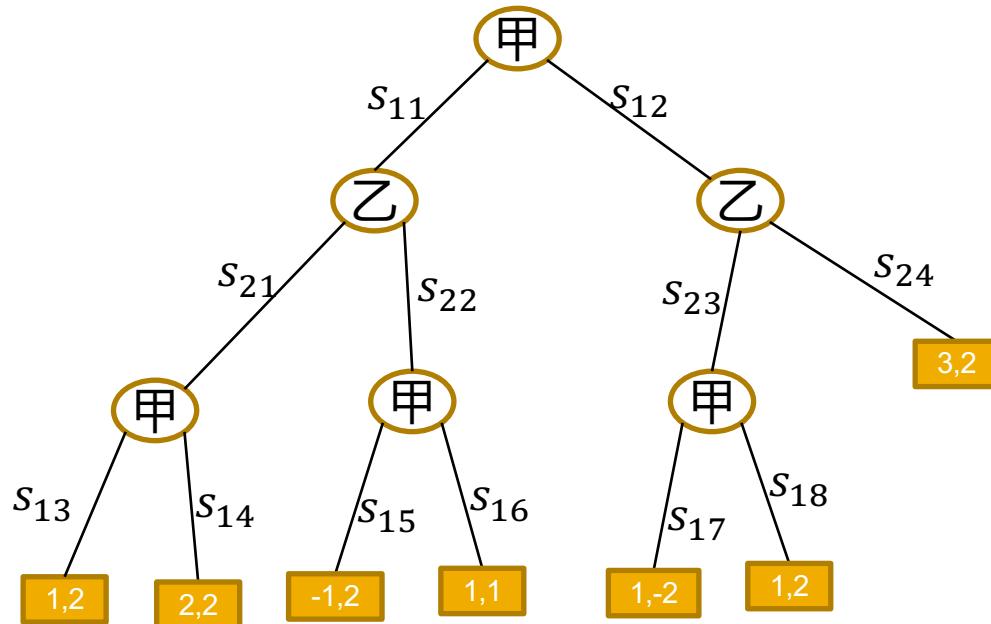
$$4q_1 + q_2^L + q_2^H = 96 \times 2$$

$$q_1 = \frac{95}{3} \approx 31.67, q_2^L \approx 33.17, q_2^H \approx 32.17$$

# 完全信息动态博弈—子博弈完美均衡

动态博弈用扩展型(Extensive Form) 表示

- 完全信息动态博弈的扩展型为一棵博弈树
- 叶节点表示各方的收益、其它节点(信息集)表示参与人
- 边表示参与人使用的策略



动态引入的问题

- ✓ 相机(Contingent)决策
- ✓ 承诺/威胁可信性问题
- ✓ 纳什均衡条件不足以排除不可信的承诺/威胁

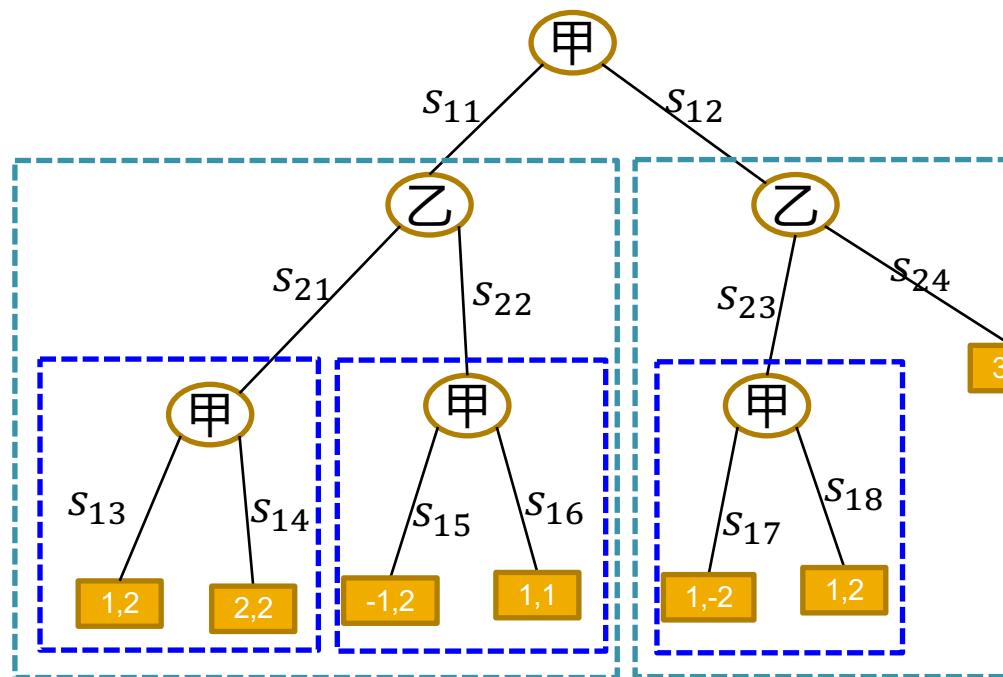
# 完全信息动态博弈—子博弈完美均衡

## 子博弈(Subgame)

- ✓ 每个信息集开启一个子博弈
- ✓ 子博弈为嵌套结构

## 子博弈完美均衡 (SPNE)

- ✓ Subgame Perfect Nash Equilibrium
- ✓ 策略组合在每一个子博弈上均构成SPNE
- ✓ 子博弈为静态则策略组合只需满足NE条件



## 计算方法

- ✓ 逆向递归 (Backward Induction, Zermelo)
- ✓ 起始层为个人决策优化

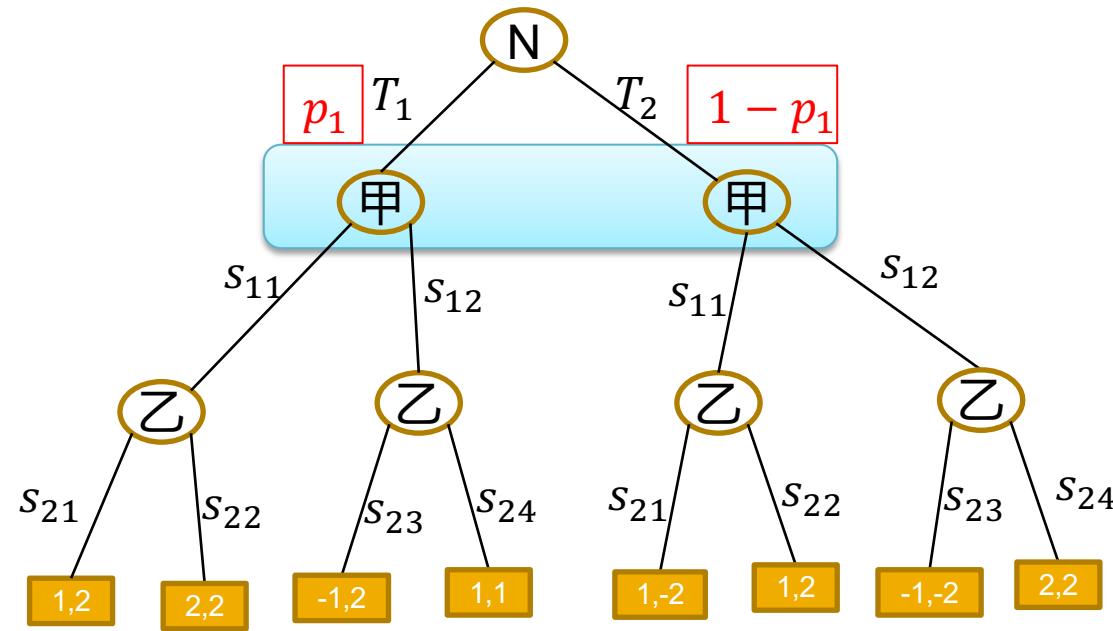
# 不完全信息动态博弈—完美贝叶斯均衡

不完全信息使得扩展型中存在多节点的信息集

➤ 不能确定博弈的路径，SPNE无法应用，BI也无法用于计算

➤ 利用贝叶斯公式，计算每条博弈路径的后验概率

➤ 在后继博弈上，保持策略和后验概率一致，即相互确定



完美贝叶斯均衡 (PBE)

- ✓ Perfect Bayesian Equilibrium
- ✓ PBE包括策略组合 $s^*$ 和后验概率 $p^*$ 两个部分( $s^*, p^*$ )，先验给定
- ✓ 在各后继博弈上，按照贝叶斯公式， $s^*$ 和 $p^*$ 满足一致性条件
  - 给定 $s^*$ ，后验概率为 $p^*$
  - 给定 $p^*$ ，最佳策略为 $s^*$

# 博弈求解—非合作博弈

静态

完全信息

完全信息静态博弈

解概念：纳什均衡  
典型博弈：GANs

不完全信息

不完全信息静态博弈

解概念：贝叶斯纳什均衡  
典型博弈：暗标拍卖

动态

完全信息动态博弈

解概念：子博弈完美纳什均衡  
典型博弈：象棋、围棋

不完全信息动态博弈

解概念：完美贝叶斯均衡  
典型博弈：德扑、星际

特殊结构动态博弈

重复博弈(Repeated Game)

- ✓ 阶段(Stage)博弈完全相同
- ✓ 阶段博弈结构相同，**状态**变化 (围棋)
- ✓ 马尔科夫完美均衡

学习  
Learning

演化博弈(Evolutionary Game)

- ✓ 智能体(Agent)

智能  
Intelligence

# 课程内容

## ■ 博弈案例

- 智猪博弈、公地悲剧、田忌赛马
- 拍卖

## ■ 博弈论

- 选择与效用
- 完全信息静态、完全信息动态
- 不完全信息静态、不完全信息动态
- 合作博弈论

## ■ 博弈论与人工智能

# 合作博弈

- 解概念：根据目标对合作创造的“收益”进行有效率/公正/公平的分配
- 基本思路：给各方分配“收益”，任何一方及小集团都不能通过自己单干和小集团合作而获得更多的“收益”
- 存在问题：组合优化
- 代表性解概念：
  - Shapley Value

# 例：公平分摊—团购

- 单价：10元，两件95折，三件88折
- A买两件：19元
- B买一件：10元
- A、B拼团：26.4元

A: 17.6元, B: 8.8元 ?

# Shapley Value

## 博弈设定

- $S = \{1, 2, \dots, N\}$  是所有博弈方构成的集合
- 某个小集团  $s \subseteq S$  合作时的收益记为  $v(s)$
- 某参与方  $i$  单干的收益则为  $v(i)$
- $v(\emptyset) = 0$
- 合作能够创造更多的收益

$$v(s_1 \cup s_2) \geq v(s_1) + v(s_2) \text{ 超可加}$$

$$v(s_1 \cup s_2) \leq v(s_1) + v(s_2) \text{ 次可加}$$

# Shapley Value

## ■ 设 $\varphi$ 是对合作收益 $v$ 的一种分配

- $\varphi_i(v)$ 是给参与方 $i$ 分配的收益
- $\varphi_s(v)$ 是给小集团 $s$ 分配的收益

## ■ 公平分配的必要条件

- 对称性， $\varphi$ 与参与方的标号 $i$ 无关 (匿名性)
- 完全性，合作收益全部分配给参与方

$$\sum_{i \in S} \varphi_i(v) = v(S)$$

- 可加性，参与多起合作得到的分配，等于单个合作得到的分配之和

$$\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2)$$

# Shapley Value

## ■ 唯一解

$$\varphi_i(v) = \sum_{i \in s \subseteq S} \frac{(|s|-1)!(N-|s|)!}{N!} [v(s) - v(s \setminus \{i\})]$$

# 例：公平分摊—团购

- 单价：10元，两件95折，三件88折
- A买两件：19元
- B买一件：10元
- A、B拼团：26.4元

$$\varphi_i(v) = \sum_{i \in s \subseteq S} \frac{(|s|-1)!(N-|s|)!}{N!} [v(s) - v(s \setminus \{i\})]$$

- $S = \{A, B\}, N=2$
- $v(A) = 19, v(B) = 10, v(A, B) = 26.4$
- $\varphi_A(v) = \frac{(1-1)!(2-1)!}{2!} (19 - 0) + \frac{1}{2} (26.4 - 10) = 17.7 > 17.6$
- $\varphi_B(v) = \frac{1}{2} (10 - 0) + \frac{1}{2} (26.4 - 19) = 8.7 < 8.8$

# 例：公平分摊—团购

- 单价：10元，两件95折，三件88折
- A买两件：19元，收益1元
- B买一件：10元，收益0元
- A、B拼团：26.4元，总收益3.6元

- $S = \{A, B\}, N=2$   $\varphi_i(v) = \sum_{i \in s} \frac{(|s|-1)!(N-|s|)!}{N!} [v(s) - v(s \setminus \{i\})]$
- $v(A) = 1, v(B) = 0, v(A, B) = 3.6$
- $\varphi_A(v) = \frac{1}{2}(1 - 0) + \frac{1}{2}(3.6 - 0) = 2.3 = 20 - 17.7$
- $\varphi_B(v) = \frac{1}{2}(0 - 0) + \frac{1}{2}(3.6 - 1) = 1.3 = 10 - 8.7$

# Shapley Value

$$\varphi_i(v) = \sum_{i \in s \subseteq S} \frac{(|s|-1)!(N-|s|)!}{N!} [v(s) - v(s \setminus \{i\})]$$

- 参与方*i*分配到的收益是其参与合作带来的收益增量的加权和
- $N!$  全部可能小集团数目
- $(|s| - 1)!$  小集团*s*里面去掉*i*的可能数目
- $(N - |s|)!$  小集团*s*之外参与方的可能数目

# Shapley Value 的等价形式

原始形式：

$$\varphi_i(v) = \sum_{i \in s \subseteq S} \frac{(|s|-1)!(N-|s|)!}{N!} [v(s) - v(s \setminus \{i\})]$$

基于排列的等价形式：

$$\varphi_i(v) = \frac{1}{N!} \sum_{\sigma \in \pi(S)} [v([\sigma]_i) - v([\sigma]_i \setminus \{i\})]$$

- $\pi(S)$  博弈集合  $S$  的全部可能排列
- $[\sigma]_i$  在排列  $\sigma$  中，排在参与方  $i$  以后的小集团（包含参与方  $i$ ）

# Shapley Value 的等价形式

## 基于回归的等价形式：

Shapley Value的计算可转化为带权线性回归问题

$$\begin{aligned} & \min_{\varphi_0, \varphi_1, \dots, \varphi_N} \sum_{S \subseteq S, 0 < |S| < N} \mu(s) \left[ \varphi_0 + \sum_{i \in S} \varphi_i - v(S) \right]^2 \\ & s.t. \varphi_0 = v(\emptyset), \varphi_0 + \sum_{i \in S} \varphi_i = v(S) \end{aligned}$$

- $\mu(s) = \frac{N-1}{\binom{N}{|S|}|S|(N-|S|)}$ ,  $\mu(\emptyset) = \mu(S) = \infty$

转化为向量形式

$$\begin{aligned} & \min_{\varphi_0, \varphi_1, \dots, \varphi_N} \sum_z p(z) [v(\mathbf{0}) + z^T \boldsymbol{\varphi} - v(z)]^2 \\ & s.t. \mathbf{1}^T \boldsymbol{\varphi} = v(\mathbf{1}) - v(\mathbf{0}) \end{aligned}$$

- $z \sim p(z) \propto \mu(z)$ ,  $z \in \{0,1\}^N$  为二元向量
- $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_N) \in \mathbb{R}^N$

# Shapley Value 的近似计算

## 1. 蒙特卡罗采样法 $O(n)$

$$\varphi_i(v) = \frac{1}{n} \sum_{\sigma \in \Pi \subset \pi(S)} [v([\sigma]_i) - v([\sigma]_i \setminus \{i\})]$$

- $\Pi \subset \pi(S)$  从全排列中均匀采样n个排列
- 无偏估计，方差较高
- 可采用对偶采样减少方差

# Shapley Value 计算

## 2. 核方法

基于回归等价形式，采样n个样本

$$\begin{aligned} & \min_{\varphi_0, \varphi_1, \dots, \varphi_N} \sum_z p(z) [\nu(\mathbf{0}) + z^T \boldsymbol{\varphi} - \nu(z)]^2 \\ & s.t. \quad \mathbf{1}^T \boldsymbol{\varphi} = \nu(\mathbf{1}) - \nu(\mathbf{0}) \end{aligned}$$

$$\begin{aligned} & \min_{\varphi_0, \varphi_1, \dots, \varphi_N} \frac{1}{n} \sum_{i=1}^n [\nu(\mathbf{0}) + z_i^T \boldsymbol{\varphi} - \nu(z_i)]^2 \\ & s.t. \quad \mathbf{1}^T \boldsymbol{\varphi} = \nu(\mathbf{1}) - \nu(\mathbf{0}) \end{aligned}$$

采用拉格朗日乘数法解约束问题

$$L(\boldsymbol{\varphi}, \lambda) = \boldsymbol{\varphi}^T \mathbf{A} \boldsymbol{\varphi} - 2 \boldsymbol{\varphi}^T \mathbf{b} + \frac{1}{n} \sum_{i=1}^n (\nu(z_i) - \nu(\mathbf{0}))^2 + 2\lambda(\mathbf{1}^T \boldsymbol{\varphi} - \nu(\mathbf{1}) + \nu(\mathbf{0}))$$

$$\text{其中 } \mathbf{A} = \frac{1}{n} \sum_{i=1}^n z_i z_i^T, \mathbf{b} = \frac{1}{n} \sum_{i=1}^n z_i (\nu(z_i) - \nu(\mathbf{0}))$$

采用KKT条件解得：

$$\boldsymbol{\varphi} = \mathbf{A}^{-1} \left( \mathbf{b} - \mathbf{1} \frac{\mathbf{1}^T \mathbf{A}^{-1} \mathbf{b} - \nu(\mathbf{1}) + \nu(\mathbf{0})}{\mathbf{1}^T \mathbf{A}^{-1} \mathbf{1}} \right)$$

- 有偏估计，方差较低

# Shapley Value 计算

## 3. 神经网络优化方法

基于回归等价形式，使用神经网络 $\phi(x; \theta)$ 近似 $\varphi(x)$

$$\begin{aligned} & \min_{\varphi_0, \varphi_1, \dots, \varphi_N} \sum_z p(z) [\nu(\mathbf{0}) + z^T \boldsymbol{\varphi} - \nu(z)]^2 \\ & s.t. \quad \mathbf{1}^T \boldsymbol{\varphi} = \nu(\mathbf{1}) - \nu(\mathbf{0}) \end{aligned}$$

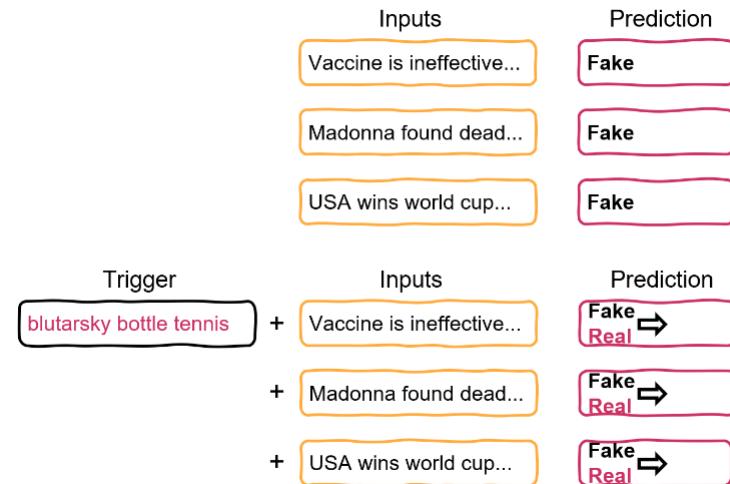
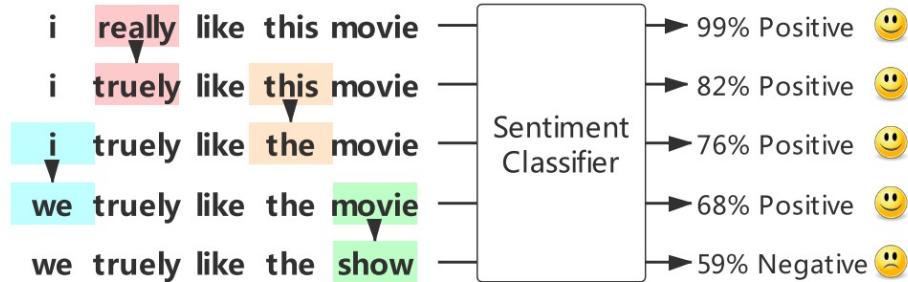
$$\theta^* = \arg \min L(\theta)$$

$$= \arg \min \mathbb{E}_{p(x)} \mathbb{E}_{p(z)} [\nu_x(z) - \nu_x(\mathbf{0}) - z^T \phi(x; \theta)]^2$$

- $p(x)$  从数据集中采样
- $p(z) \propto \mu(z)$
- 可证明  $\phi(x; \theta^*) = \varphi(x)$

# Shapley Value 在对抗攻击中的应用

## 文本对抗攻击



## 两段式攻击

- 单词重要性排序
- 单词扰动（替换，增加，删除）

# Shapley Value 在对抗攻击中的应用

## 单词重要性度量

- $s(x_i) = f(x) - f(x \setminus x_i)$
- $s(x_i) = \left| \frac{\partial f}{\partial x} \right|_p$

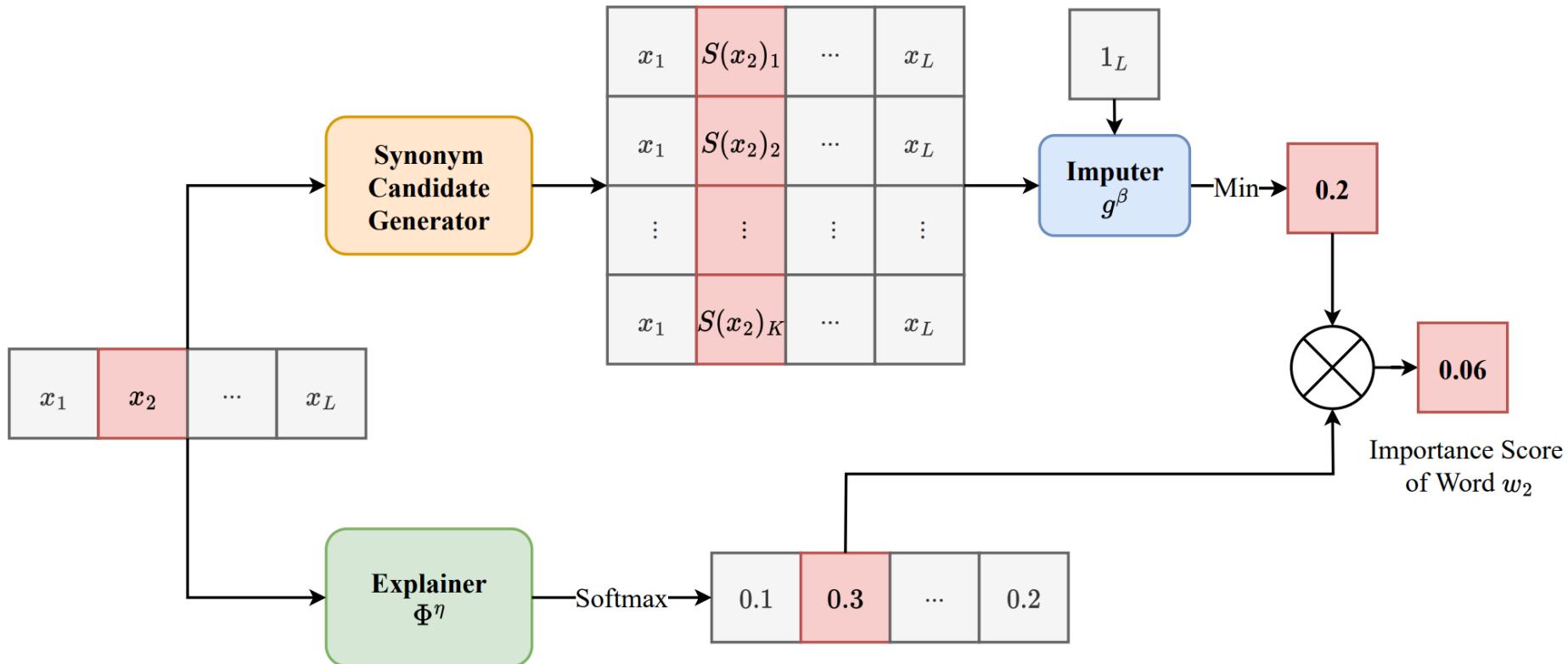
基于启发式规则，缺乏理论支持

将文本分类建模成合作博弈

- 每个单词视作参与者
  - 将模型输出的概率值视作合作产生的收益
- 使用Shapley Value作为单词重要性判据

# Shapley Value 在对抗攻击中的应用

## 融合Shapley Value的单词重要性度量

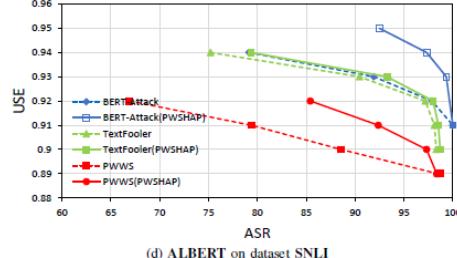
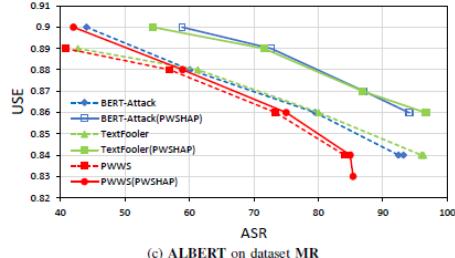
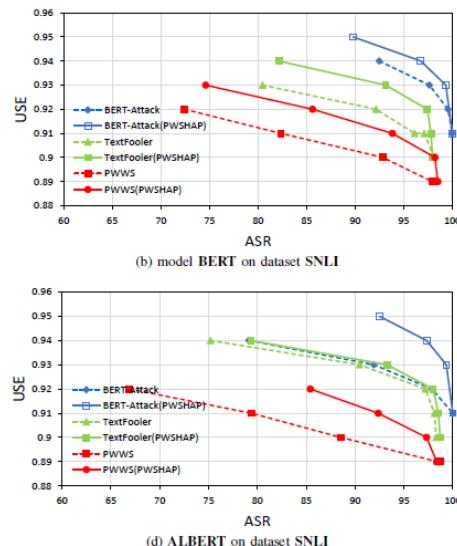
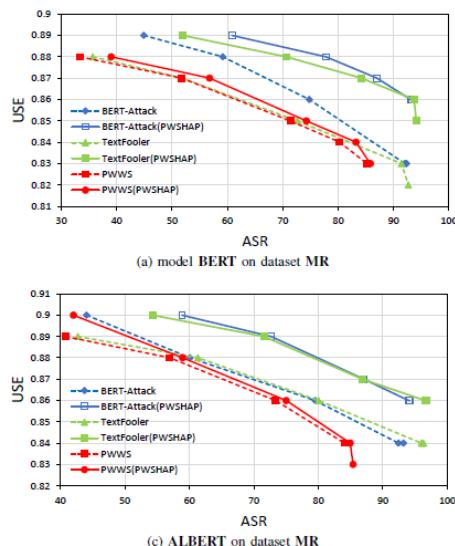


# Shapley Value 在对抗攻击中的应用

| Model       | Attack      | WIR              | ASR                 | PPL                   | USE                        | QRY              | PTB                 |
|-------------|-------------|------------------|---------------------|-----------------------|----------------------------|------------------|---------------------|
| BERT        | Text Fooler | delete<br>PWSHAP | 92.7<br><b>94.1</b> | 222.6<br><b>187.2</b> | 0.82<br><b>0.85</b>        | 106<br><b>76</b> | 18.6<br><b>14.1</b> |
|             | PWWS        | PWWS<br>PWSHAP   | 85.2<br><b>85.8</b> | 208.1<br>211.5        | <b>0.83</b><br><b>0.83</b> | 146<br><b>40</b> | <b>15.2</b><br>15.3 |
|             | BERT-Attack | unk<br>PWSHAP    | 90.2<br><b>92.6</b> | <b>144.9</b><br>156.8 | 0.83<br><b>0.86</b>        | 136<br><b>90</b> | 15.5<br><b>11.6</b> |
| ALBERT      | Text Fooler | delete<br>PWSHAP | 96.3<br><b>96.8</b> | 213.0<br><b>181.9</b> | 0.84<br><b>0.86</b>        | 91<br><b>72</b>  | 16.4<br><b>14.0</b> |
|             | PWWS        | PWWS<br>PWSHAP   | 84.7<br><b>85.4</b> | <b>200.7</b><br>208.6 | <b>0.84</b><br>0.83        | 148<br><b>43</b> | <b>15.2</b><br>16.3 |
|             | BERT-Attack | unk<br>PWSHAP    | 93.2<br><b>94.2</b> | 174.1<br><b>161.5</b> | 0.84<br><b>0.86</b>        | 125<br><b>93</b> | 15.3<br><b>12.3</b> |
| Distil BERT | Text Fooler | delete<br>PWSHAP | 95.2<br><b>96.3</b> | 196.4<br><b>179.2</b> | 0.85<br><b>0.86</b>        | 80<br><b>61</b>  | 15.2<br><b>12.8</b> |
|             | PWWS        | PWWS<br>PWSHAP   | 88.2<br><b>89.1</b> | 205.8<br>203.1        | <b>0.84</b><br>0.84        | 141<br><b>35</b> | 14.6<br><b>14.5</b> |
|             | BERT-Attack | unk<br>PWSHAP    | 94.8<br><b>95.6</b> | 168.5<br><b>157.3</b> | 0.85<br><b>0.87</b>        | 103<br><b>77</b> | 13.0<br><b>11.3</b> |

| Model       | Attack      | WIR              | ASR                        | PPL                 | USE                        | QRY                | PTB                      |
|-------------|-------------|------------------|----------------------------|---------------------|----------------------------|--------------------|--------------------------|
| BERT        | Text Fooler | delete<br>PWSHAP | 98.2<br><b>98.7</b>        | 66.6<br><b>57.3</b> | 0.92<br><b>0.95</b>        | 592<br><b>238</b>  | 7.6<br><b>4.2</b>        |
|             | PWWS        | PWWS<br>PWSHAP   | 97.7<br><b>96.5</b>        | 58.9<br><b>58.3</b> | <b>0.94</b><br><b>0.94</b> | 1428<br><b>150</b> | <b>4.3</b><br><b>4.3</b> |
|             | BERT-Attack | unk<br>PWSHAP    | 98.5<br><b>99.8</b>        | 57.9<br><b>52.4</b> | 0.92<br><b>0.95</b>        | 619<br><b>198</b>  | 6.2<br><b>2.7</b>        |
| ALBERT      | Text Fooler | delete<br>PWSHAP | 98.4<br><b>99.1</b>        | 63.2<br><b>56.2</b> | 0.93<br><b>0.95</b>        | 580<br><b>224</b>  | 6.5<br><b>3.9</b>        |
|             | PWWS        | PWWS<br>PWSHAP   | 96.5<br><b>97.2</b>        | 57.6<br><b>59.0</b> | <b>0.95</b><br>0.94        | 1430<br><b>151</b> | <b>3.6</b><br>4.3        |
|             | BERT-Attack | unk<br>PWSHAP    | 99.9<br><b>100</b>         | 55.1<br><b>51.4</b> | 0.94<br><b>0.96</b>        | 519<br><b>154</b>  | 3.9<br><b>2.2</b>        |
| Distil BERT | Text Fooler | delete<br>PWSHAP | <b>99.8</b><br><b>99.8</b> | 58.6<br><b>53.7</b> | 0.94<br><b>0.96</b>        | 370<br><b>138</b>  | 4.5<br><b>2.7</b>        |
|             | PWWS        | PWWS<br>PWSHAP   | <b>99.5</b><br><b>99.5</b> | <b>55.9</b><br>56.4 | <b>0.95</b><br><b>0.95</b> | 1385<br><b>95</b>  | <b>3.2</b><br>3.3        |
|             | BERT-Attack | unk<br>PWSHAP    | 99.7<br><b>99.9</b>        | 53.3<br><b>51.0</b> | 0.95<br><b>0.96</b>        | 356<br><b>137</b>  | 3.2<br><b>1.9</b>        |

| Model       | Attack      | WIR              | ASR                        | PPL                 | USE                        | QRY              | PTB               |
|-------------|-------------|------------------|----------------------------|---------------------|----------------------------|------------------|-------------------|
| BERT        | Text Fooler | delete<br>PWSHAP | 98.2<br><b>97.9</b>        | 38.6<br>39.1        | <b>0.90</b><br><b>0.90</b> | 41<br>43         | 7.1<br>7.2        |
|             | PWWS        | PWWS<br>PWSHAP   | 98.2<br><b>98.5</b>        | <b>43.5</b><br>48.3 | <b>0.89</b><br><b>0.89</b> | 175<br><b>28</b> | <b>7.3</b><br>9.1 |
|             | BERT-Attack | unk<br>PWSHAP    | <b>100</b><br><b>100</b>   | 32.0<br>31.5        | <b>0.91</b><br><b>0.91</b> | 66<br><b>56</b>  | <b>6.1</b><br>6.4 |
| ALBERT      | Text Fooler | delete<br>PWSHAP | 98.3<br><b>98.7</b>        | 40.9<br><b>40.0</b> | <b>0.90</b><br><b>0.90</b> | 47<br><b>42</b>  | 7.7<br>7.4        |
|             | PWWS        | PWWS<br>PWSHAP   | <b>98.7</b><br>98.5        | 41.7<br>60.4        | <b>0.89</b><br><b>0.89</b> | 173<br><b>40</b> | 7.0<br><b>6.7</b> |
|             | BERT-Attack | unk<br>PWSHAP    | <b>100</b><br>99.9         | 32.2<br>36.0        | <b>0.91</b><br>0.90        | 66<br><b>57</b>  | 6.1<br><b>5.7</b> |
| Distil BERT | Text Fooler | delete<br>PWSHAP | <b>98.6</b><br>98.3        | 40.0<br>36.9        | 0.90<br><b>0.91</b>        | 41<br>37         | 7.1<br>6.4        |
|             | PWWS        | PWWS<br>PWSHAP   | <b>98.3</b><br><b>98.3</b> | 42.7<br>43.1        | <b>0.88</b><br><b>0.88</b> | 174<br><b>24</b> | <b>7.3</b><br>7.5 |
|             | BERT-Attack | unk<br>PWSHAP    | <b>99.9</b><br><b>99.9</b> | 32.1<br><b>30.9</b> | <b>0.91</b><br><b>0.91</b> | 65<br><b>49</b>  | 6.0<br><b>5.7</b> |



# Shapley Value 在对抗攻击中的应用

| WIR    | Examples   | Labels                                 |
|--------|--|--|
| delete | a mature (grownup), deeply felt fantasy of a director (rector)'s travel through 300 years of russian history (bygone).   | Positive 100%<br>↓<br>Negative 97%     |
| PWSHAP | a mature, deeply felt fantasy (illusion) of a director's travel through 300 years of russian history.  | Positive 100%<br>↓<br>Negative 96%     |
| delete | LIGHTS OF NEW YORK was the first "all-taking (opted)" feature film, coming in at a brisk (briskly) 57 minutes and directed by Bryan Foy (of the famous vaudeville family).The story has two dopey barbers (hairdressers) (Cullen Landis, Eugene Palette) yearning for a chance at "big city life" and getting involved with gangsters (mobsters) and bootleg booze (intoxicating). One of the guys gets framed for the murder of a cop but is saved at the last minute by a gun moll (Gladys Brockwell). Much of the story takes place in a night club called The Night Hawk, which is run by a crook named Hawk (Wheeler Oakman) who has his eye on a pretty chorine (Helene Costello) who is the girl friend of Landis. Costello gets to do a brief dance, and we hear Harry Downing (made up to resemble Ted Lewis) sing "At Dawning) in his best Al Jolson style. The acting ranges from good (Palette and Brockwell) to awful (Oakman)... | Positive 100%<br>↓<br>Negative 58%     |
| PWSHAP | LIGHTS OF NEW YORK was the first "all-taking" feature film, coming in at a brisk 57 minutes and directed by Bryan Foy (of the famous vaudeville family).The story has two dopey barbers (Cullen Landis, Eugene Palette) yearning for a chance at "big city life" and getting involved with gangsters and bootleg booze. One of the guys gets framed for the murder of a cop but is saved at the last minute by a gun moll (Gladys Brockwell). Much of the story takes place in a night club called The Night Hawk, which is run by a crook named Hawk (Wheeler Oakman) who has his eye on a pretty chorine (Helene Costello) who is the girl friend of Landis. Costello gets to do a brief dance, and we hear Harry Downing (made up to resemble Ted Lewis) sing "At Dawning) in his best Al Jolson style. The acting ranges (fluctuates) from good (Palette and Brockwell) to awful (Oakman)...   | Positive 100%<br>↓<br>Negative 54%     |
| delete | <b>Premise:</b> A black man in a blue suit is talking on a cellphone while smoking<br><b>Hypothesis:</b> The man (boy) is talking to his friend (girlies) on the phone   | Neutral 100%<br>↓<br>Contradiction 70% |
| PWSHAP | <b>Premise:</b> A black man in a blue suit is talking on a cellphone while smoking<br><b>Hypothesis:</b> The man is talking to his friend on the phone (handset)   | Neutral 100%<br>↓<br>Contradiction 64% |

# 课程内容

## ■ 博弈案例

- 智猪博弈、公地悲剧、田忌赛马
- 拍卖

## ■ 博弈论

- 选择与效用
- 完全信息静态、完全信息动态
- 不完全信息静态、不完全信息动态
- 合作博弈论

## ■ 博弈论与人工智能

# 博弈思想与博弈论简史

- 春秋时代，《孙子兵法》
- 1512，马基雅维利，《君主论》
- ~ 1830，克劳塞维茨，《战争论》

- 1713, Waldegrave, 二人博弈Min-Max混合策略解
- 1838, Cournot, 双寡头垄断, 产量竞争模型(静态)
- 1883, Bertrand, 双寡头垄断, 价格竞争模型(动态)

- 1913, Zermelo, 棋类必胜(不败)策略, 逆向递归
- 1921, Borel, 混合策略定义, Min-Max解
- 1928, Von Neumann & Morgenstern, Min-Max解

# 博弈思想与博弈论简史

- 1944 , Von Neumann & Morgenstern, 现代起点  
《博弈论与经济行为》
- 1950, Nash, 非合作博弈, 纳什均衡及存在性  
Rand公司, 博弈实验, 战略、冲突分析
- 1952, Shapley , 合作博弈, Core、Shapley值

- 1965, Selten , 动态博弈, 子博弈完美均衡
- 1967, Harsanyi, 不完全信息博弈, 贝叶斯纳什均衡
- 1972, Smith, 进化博弈论, 进化稳定策略
- 1982, Kreps & Wilson, 序贯均衡
- 1991, Fudenberg & Tirole , 完美贝叶斯均衡

# 博弈论成就认可

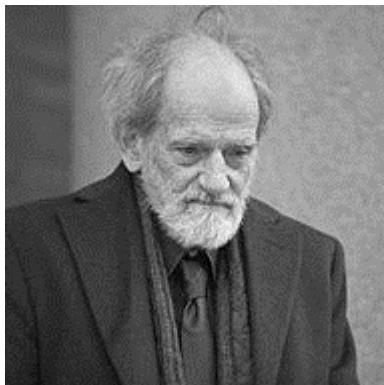
## ■ 博弈论解析模型1950—1990基本成型

- 高峰 ~~1970
- 1994—2020, 9次20人Nobel经济学奖

## ■ 博弈论行为模型与实验验证1990—今

- 羊群效应、嫉妒、利他、公平、锚定效应、框架
- 2002—2020, 3次6人Nobel经济学奖

# 博弈求解—合作博弈



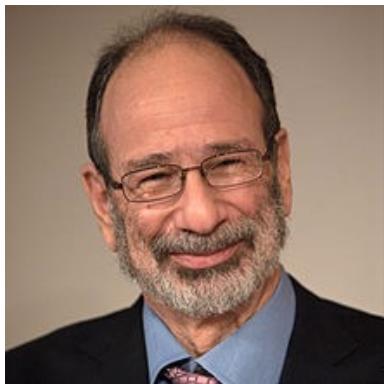
**Lloyd Shapley**  
1923—2016

Nobel Laureate, 2012  
合作博弈论

Core、Stable Matching  
Kernel、Nucleolus  
Shapley值

One of the giants of  
*game theory* and economic theory

美国经济学会，2016



**Alvin E. Roth**  
1951—

Nobel Laureate, 2012  
合作博弈论

Stable Matching  
(择偶、择校、器官捐献)

美国科学院院士，2013  
美国经济学会会长，2017—

给定规则，决定如何合作

# 博弈求解——非合作博弈

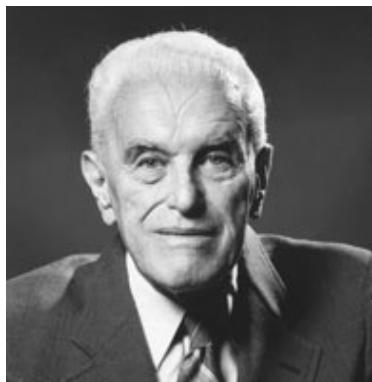


**John Nash**

1928—2015

Nobel Laureate, 1994  
博弈论

纳什均衡 (非合作)  
纳什解 (合作)

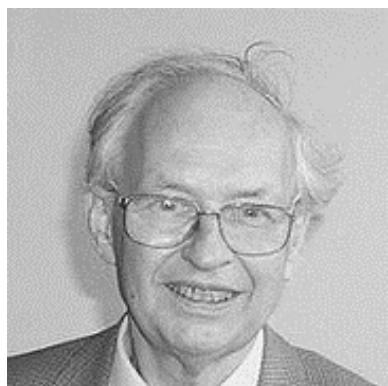


**John Harsanyi**

1920—2010

Nobel Laureate, 1994  
非合作博弈论

不完全信息动态博弈  
海萨尼转换



**Reinhard Selten**

1930—2015

Nobel Laureate, 1994  
非合作博弈论

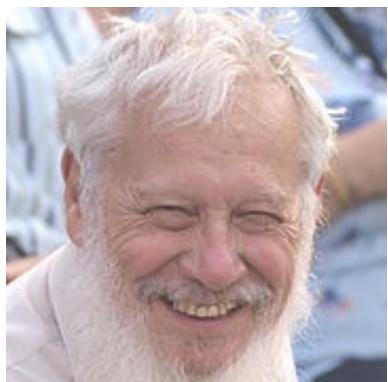
子博弈完美均衡  
颤抖手均衡



**Thomas Schelling**  
1921—2016

Nobel Laureate, 2005  
非合作博弈论

焦点(Focal Point)  
承诺(commitment)  
军备控制、核武器消减



**Robert Aumann**

1930—

Nobel Laureate, 2005  
非合作博弈论

共同知识、相关均衡  
重复博弈分析



**Jean Tirole**  
1953—

Nobel Laureate, 2014  
非合作博弈论

完美贝叶斯均衡、机制  
垄断、规制、创新

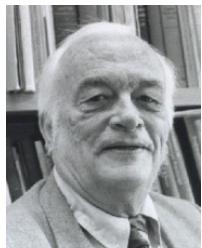
给定规则，决定如何对抗

# “逆向博弈论”

1996：机制设计理论

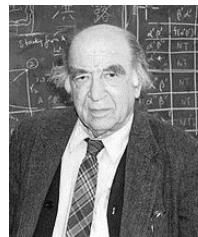


**James Mirrlees**  
1936—2018



**William Vickrey**  
1914—1996

2007：机制设计理论



**Leonid Hurwicz**  
1917—2008



**Eric Maskin**  
1950—



**Roger Myerson**  
1951—

2020：拍卖理论



**Paul Milgrom**  
1948—



**Robert Wilson**  
1914—1996

2001：信息经济学



**George Akerlof**  
1940—



**Michael Spence**  
1943—



**Joseph Stiglitz**  
1943—

2016：合同理论



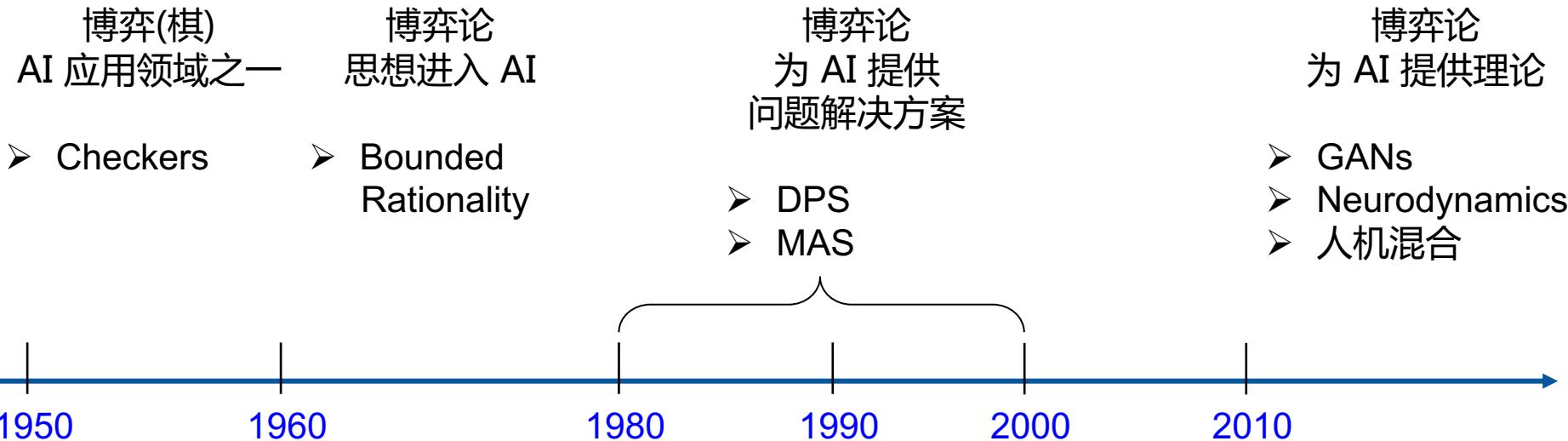
**Oliver Hart**  
1948—



**Bengt Holmström**  
1949—

根据结果，设计能实现该结果的博弈规则

# 人工智能与博弈论



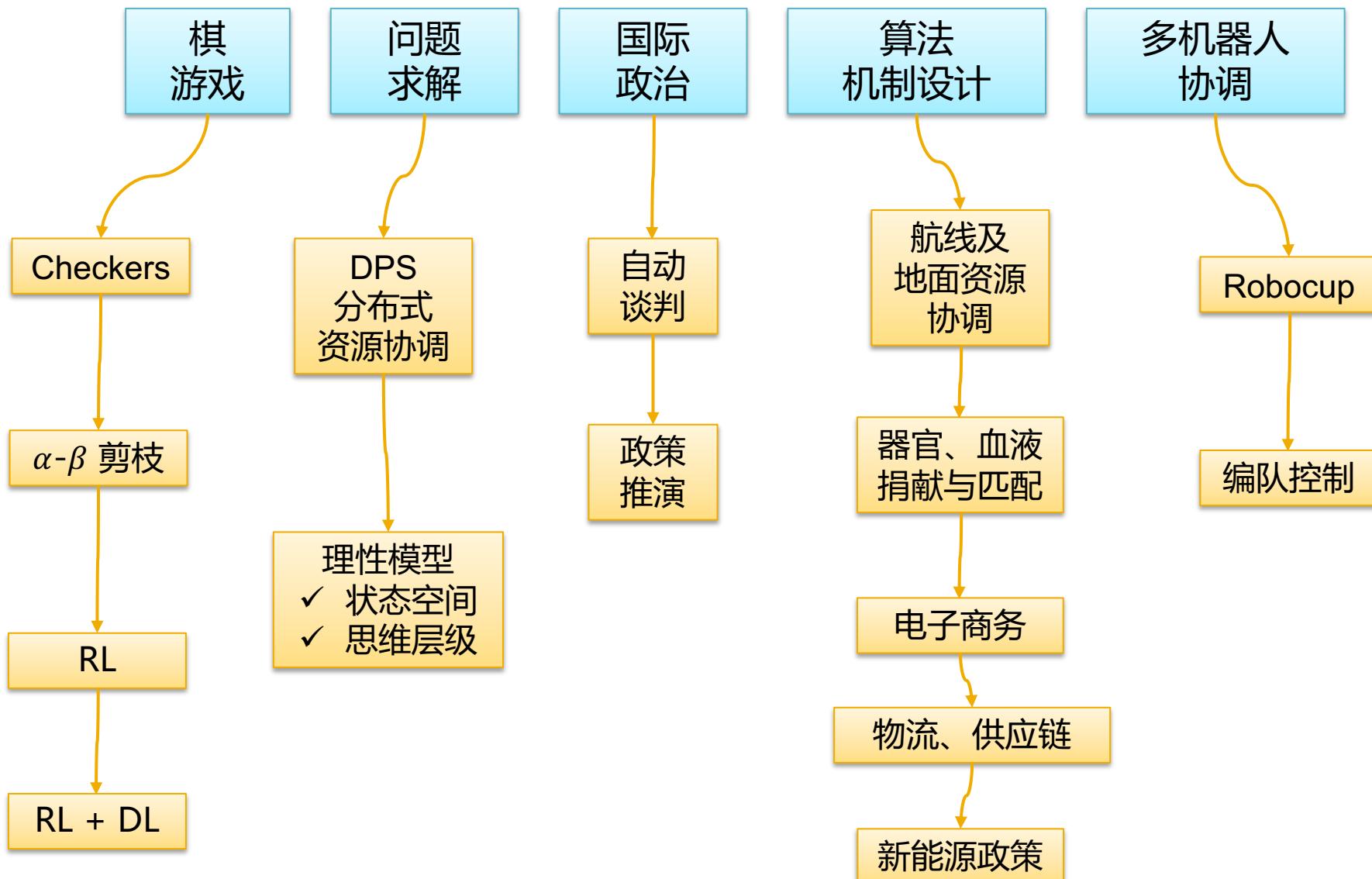
搜索  
 $\alpha$ - $\beta$  剪枝

算法博奕论

本质：博弈论+CS+AI

- 国际象棋
- 算法机制设计
- 均衡计算
- State, Automata
- Recursive depth
- 深度学习
- 强化学习
- AlphaX
- DeepStack
- Libratus

# 博弈 + AI 代表性成果



- 计算博弈
  - 均衡概念偏存在性，缺乏可计算性和实战性
  - 现有框架实战效果突出，尚未在理论上经过均衡检验
- 棋类，完全信息，Zermelo定理保证必胜(不败)策略存在
  - ◆ AlphaX系列框架在长期是否学习到？ 算力优势？
- 扑克类，AI计算 $\epsilon$ -NE
  - ◆ 效果有大幅提升，离人还有距离？ 多NE？ 不完全信息，NE可用？
- 即时策略类
  - ◆ 效果尚一般
- Self-play/Self-improve 框架？
  - ◆ 适用完全信息博弈，N=2 (N>3, 策略空间进一步复杂，合作、小集团)
  - ◆ 算力消耗(有意义、无意义)

## ■ GANs (完全信息静态博弈)

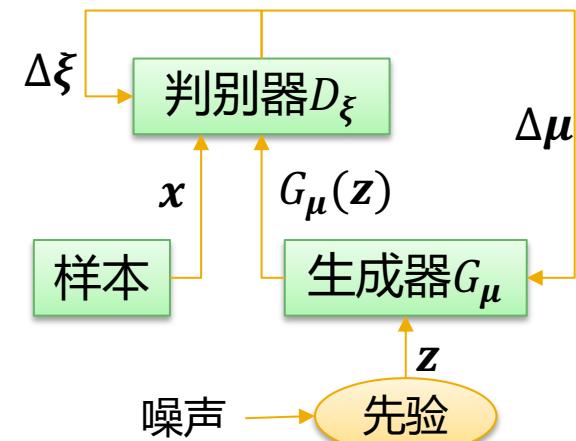
$$\min_{\mu} \max_{\xi} \mathbb{E}_x [\phi(D_\xi(x))] + \mathbb{E}_z [\phi(1 - D_\xi(G_\mu(z)))]$$

$D$  和  $G$  具有无限表达能力，收敛到最优均衡

- ✓ 生成器学到样本真实分布
- ✓ 判别器以1/2概率随机判断

- 均衡点不存在，对某些  $\phi$
- 均衡点存在，训练在均衡点附近震荡
- 均衡点非最优，生成分布离样本分布较远

利用少量  $D_\xi$  和  $G_\mu$  的混合(Mixture)，可收敛到  $\epsilon$ -纳什均衡，但实际性能无保证



$D_\xi$  和  $G_\mu$  表达能力有限 (NN+BP)

- Mode Collapse
- Unstable
- Vanishing Gradient

# ChatGPT—LM因果缺陷

- “表观” 推理能力，4.0 相对 3.5 有巨大提升

## 爱因斯坦之谜

- 在一条街上,有5座房子,喷了5种颜色
- 每个房子里住着不同国家的人
- 每个人喝不同的饮料,抽不同牌子的香烟,养不同的宠物

- 1、英国人住红色房子
- 2、瑞典人养狗
- 3、丹麦人喝茶
- 4、绿色房子在白色房子左面
- 5、绿色房子主人喝咖啡
- 6、抽Pall Mall 香烟的人养鸟
- 7、黄色房子主人抽Dunhill 香烟
- 8、住在中间房子的人喝牛奶
- 9、挪威人住第一间房
- 10、抽Blends香烟的人住在养猫的人隔壁
- 11、养马的人住抽Dunhill 香烟的人隔壁。
- 12、抽Blue Master的人喝啤酒。
- 13、德国人抽Prince香烟。
- 14、挪威人住蓝色房子隔壁。
- 15、抽Blends香烟的人有一个喝水的邻居。

哪个国家的人养鱼 ?

# ChatGPT—LM因果缺陷

- 3.5 无法推断
- 4.0 知道问题叫“爱因斯坦之谜”，但不能正确推理

1. 挪威人住在第一间房子。
2. 由于挪威人住在蓝色房子隔壁，所以第二间房子是蓝色的。
3. 由于红色房子住着英国人，且绿色房子在白色房子左边，因此第三间房子是红色的，英国人住在这里。
4. 因此，第四间房子是绿色的，第五间房子是白色的。
5. 住在中间房以英国人喝牛奶。  
**无蕴含关系**
6. 绿色房子主人四间房子的人喝咖啡。
7. **丹麦人喝茶，所以丹麦人住在第二间房子。**
8. 抽Blue Master的人喝啤酒，所以挪威人抽Blue Master香烟并且喝啤酒。
9. 黄色房子主人抽Dunhill香烟，所以第一间房子是黄色的，挪威人抽Dunhill香烟。
10. 德国人抽Prince香烟，所以德国人住在第四间房子。
11. **瑞典人养狗，所以瑞典人住在第五间房子。**  
**前后矛盾**
12. 抽Pall Mall香烟的人养鸟，所以抽Pall Mall香烟并且养鸟。
13. 养马的人住抽Dunhill香烟的人隔壁，所以在第二间房子的丹麦人养马。
14. 抽Blends香烟的人住在养猫的人隔壁，所以德国人抽Blends香烟，**瑞典人养猫。**
15. 抽Blends香烟的人有一个喝水的邻居，所以第五间房子的瑞典人喝水。
16. 现在我们已经知道了所有房子的详细信息，唯一没有提到宠物的是第四间房子的德国人，因此我们可以得出结论：**德国人养鱼。**

答案正确

# ChatGPT—LM因果缺陷

## ● 怎么确定红色房子一定在中间呢？

你是对的。

在我之前的回答中，我没有考虑到所有可能的情况。现在让我们重新整理信息并考虑可能的房子顺序。

我们知道：挪威人住在第一间房子。住在中间房子的人喝牛奶。绿色房子在白色房子左面。挪威人住在蓝色房子隔壁，所以第二间房子是蓝色的。现在，我们有两种可能的房子顺序：

A) 红色房子在中间：黄、蓝、红、绿、白。

B) 红色房子在最右边：黄、蓝、绿、白、红。

由于抽Blends香烟的人住在养猫的人隔壁，那么养猫的人必须住在第四间房子。现在，我们已经知道：

挪威人住在第一间房子，抽Dunhill香烟。丹麦人住在第二间房子，喝茶，养马。第三间房子的人喝牛奶，抽Blends香烟。第四间房子的人喝咖啡，养猫。第五间房子是白色的。我们还需要找出其他信息：

英国人住红色房子。根据现有信息，英国人只能住在第三间房子，所以第三间房子是红色的。

由于英国人抽Pall Mall香烟并养鸟，那么第三间房子的人养鸟。

绿色房子在白色房子左面，所以第四间房子是绿色的，第五间房子是白色的。这意味着瑞典人住在第五间房子里，并养狗。抽Blue Master香烟的人喝啤酒。因为瑞典人养狗，所以瑞典人抽Blue Master香烟并喝啤酒。德国人抽Prince香烟。由于其他房子的香烟品牌已知，德国人只能住在第四间房子，抽Prince香烟。现在，所有的房子和居民都已经确定：

所以，**挪威人养鱼。**

答案错误

# ChatGPT—LM因果缺陷

## ● 到底是德国人还是挪威人养鱼？

抱歉给你带来了困扰。我的错误。请允许我再次审查线索并给出正确的答案。

挪威人住在第一间房子。住在中间房子的人喝牛奶。绿色房子在白色房子左面。挪威人住在蓝色房子隔壁，所以第二间房子是蓝色的。现在，我们有两种可能的房子顺序：

- A) 红色房子在中间：黄、蓝、红、绿、白。
- B) 红色房子在最右边：黄、蓝、绿、白、红。

我们来分析其他线索，找出正确的顺序：

由于挪威人住在第一间房子，且黄色房子主人抽Dunhill香烟，所以第一间房子是黄色的，挪威人抽Dunhill香烟。因为养马的人住抽Dunhill香烟的人隔壁，所以第二间房子的人养马。绿色房子主人喝咖啡，所以第四间房子是绿色的，且第四间房子的人喝咖啡。绿色房子在白色房子左面，所以第五间房子是白色的。因此，养猫的人住在第四间房子。养马的人住抽Dunhill香烟的人隔壁。我们已经知道第二间房子的人养马。抽Blue Master香烟的人喝啤酒。由于瑞典人是唯一剩下的人，瑞典人必须抽Blue Master香烟并喝啤酒。德国人抽Prince香烟。因此，德国人住在第四间房子。

所以，是挪威人养鱼。再次为之前的错误向您道歉。

答案错误

# ChatGPT—LM因果缺陷

设已知的公式集为{P,  $(P \wedge Q) \rightarrow R, (S \vee T) \rightarrow R, T$ }求证结论R

我们可以通过演绎推理来证明结论R。具体步骤如下：

1. 由第一个公式P, 得到P为真。
2. 由第二个公式 $(P \wedge Q) \rightarrow R$ 和已知P, 得到 $Q \rightarrow R$ , 即如果Q为真, 则R为真。
3. 由第四个公式T和已知 $(S \vee T) \rightarrow R$ , 得到 $(S \vee T)$ 为真, 即S为假。
4. 由第三个公式 $(S \vee T) \rightarrow R$ 和已知S为假, 得到 $T \rightarrow R$ , 即如果T为真, 则R为真。
5. 由第四步得到的 $T \rightarrow R$ , 结合第四个公式T, 得到R为真。

因此, 结论R为真。

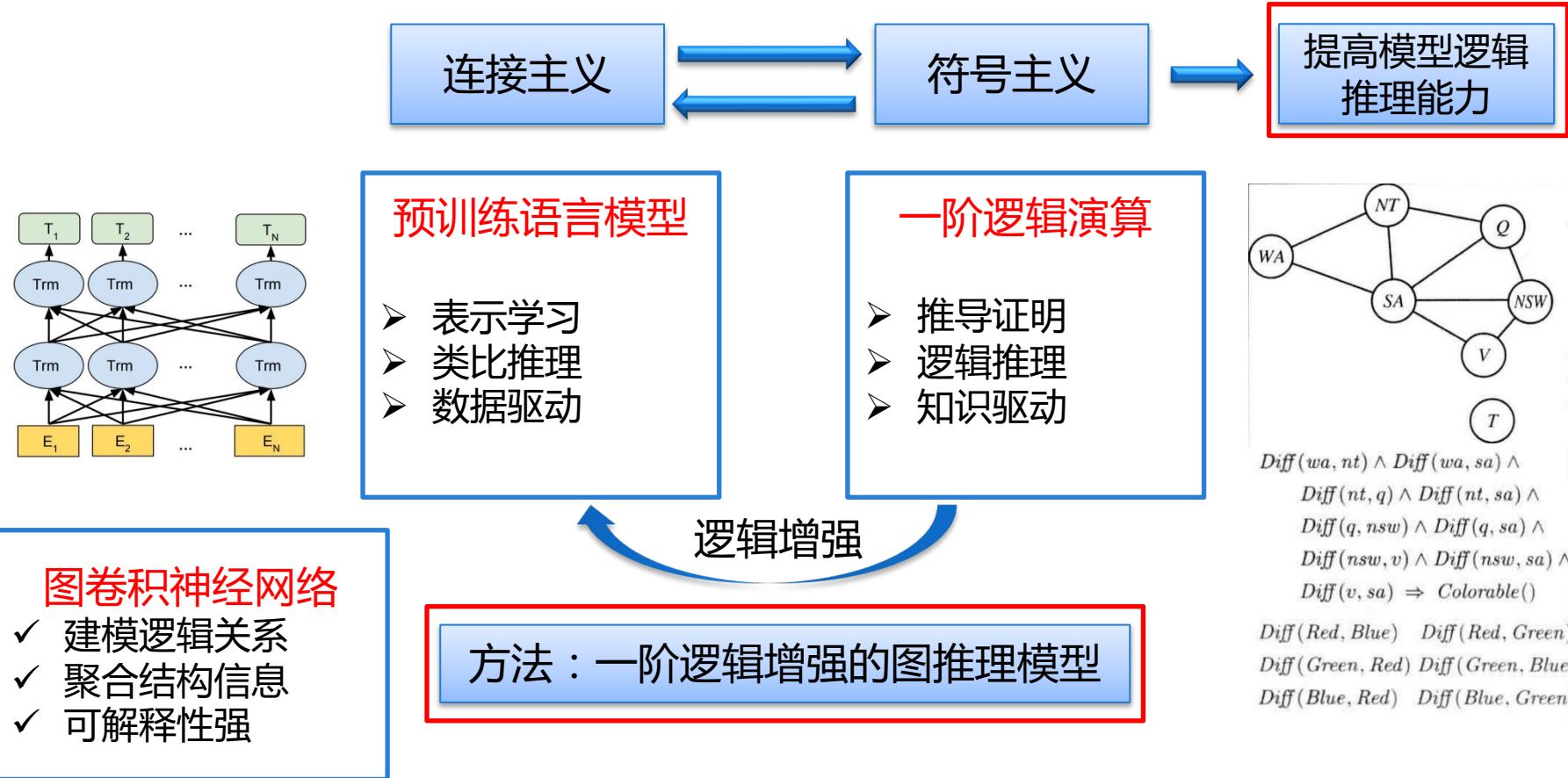
结论正确

➤  $T \rightarrow S \vee T$

➤  $S \vee T \rightarrow R$

➤ AR结构性问题

# FOL 增强推理



# 新架构

➤ Next-word ?

自回归

➤ Transformer ?

➤ Neural Network

➤ 万能逼近器

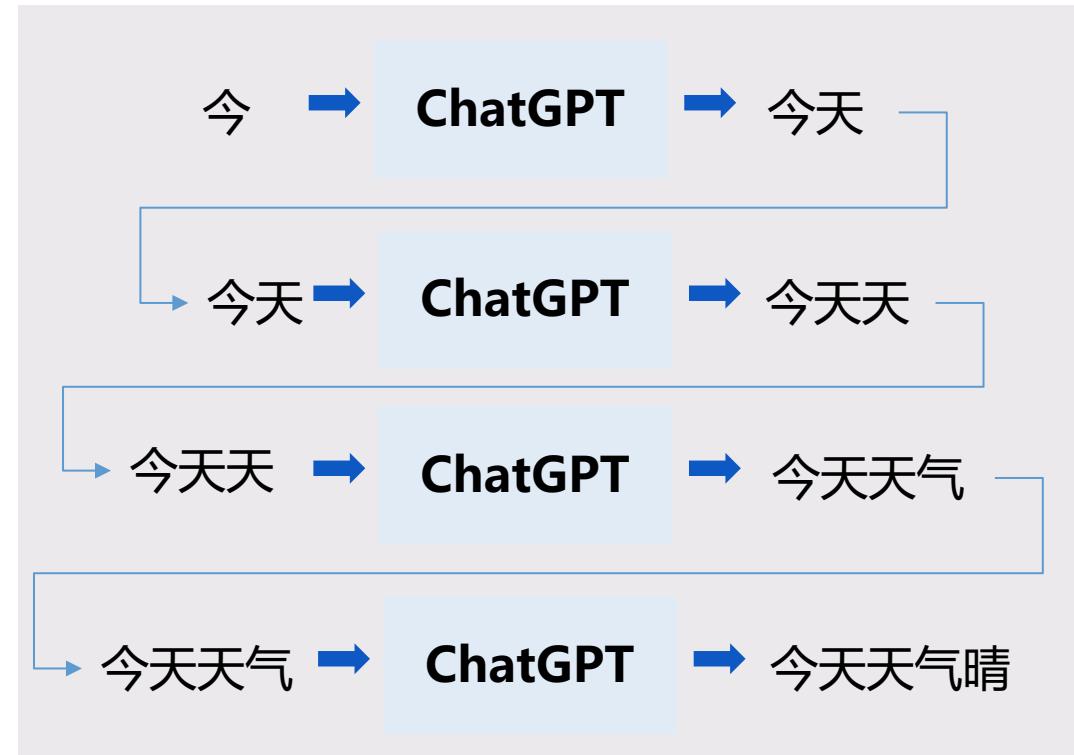
➤ 网络结构、神经元数量跟  
近似性能的关系？

➤ 大小模型混合架构

✓ 长期记忆

✓ 4K、32K

✓ SNN



# 后ChatGPT时代的人工智能

- AI研究脱离解决具体问题、越来越抽象化？
- AI牛顿定理/相对论/量子力学
- AI中的不可能定理？
  - 智能水平 vs 算力

# 博弈论有不有用？

✓ 如果各方想使用均衡策略，均衡点怎么达到？

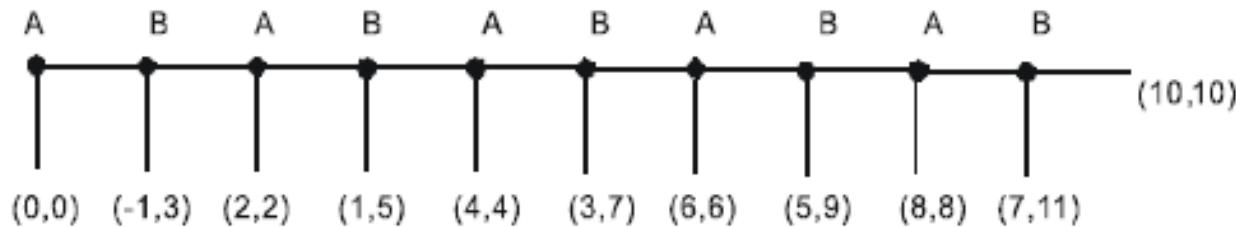
✓ 参与人能不能算出均衡？

✓ 贝叶斯公式所有人都知道吗？

✓ 动态博弈分析的困难

✓ 偏好改变

✓ BI不是唯一方法



R. Rosenthal, 蜈蚣博弈

# 博弈论有不有用？

- 现实博弈问题中，各方采用均衡策略吗？

中美关税战场景

| GDP/进出口/福利 |                            | 美国   |      |          |             |             |
|------------|----------------------------|------|------|----------|-------------|-------------|
|            |                            | 0关税  | 维持现状 | 500亿，25% | + 2000亿，10% | + 2000亿，25% |
| 中国         | 0关税                        | ?, ? | ?, ? | ?, ?     | ?, ?        | ?, ?        |
|            | 维持现状                       | ?, ? | ?, ? | ?, ?     | ?, ?        | ?, ?        |
|            | 500亿<br>25%                | ?, ? | ?, ? | ?, ?     | ?, ?        | ?, ?        |
|            | + 600亿<br>(25%、20%、10%、5%) | ?, ? | ?, ? | ?, ?     | ?, ?        | ?, ?        |
|            | + 750亿<br>(5%、10%)         | ?, ? | ?, ? | ?, ?     | ?, ?        | ?, ?        |

问题1：中国“硬碰硬”的经济合理性？ 问题2：美国“多轮加增”的经济合理性？

Nash 均衡

每一轮博弈，中美两国的策略均构成 Nash 均衡

# 博弈论有不有用？

- 现实博弈问题中，各方采用均衡策略吗？

战略层

输入：各方可能的策略



- ✓ 0关税
- ✓ 保持现状
- ✓ 加征25%关税
- ✓ 加征10%关税
- ✓ 加征25%关税

输出：博弈结果

✓ 事中

- 给定对手策略，我方最优策略
- 预测各方的策略选择

✓ 事后

- 策略检讨、评估

方法：博弈论

✓ 第一步：计算得失

- (5%关税, 15%关税)
- (5%效用, 15%效用)

✓ 第二步：均衡分析

- Nash均衡
- 其它可能的精炼

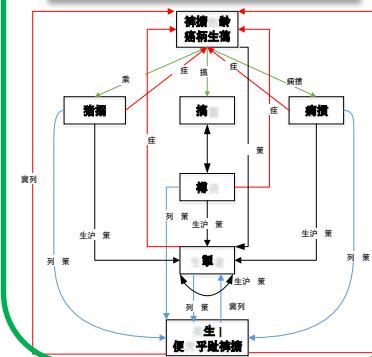
战术层

公司  
企业  
个人  
行为  
难以  
精确  
模拟

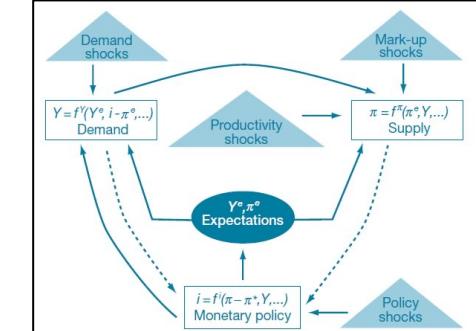
投入-产出模型

| 产品部门               | 代码 | 产出        |          | 投入        |        |         |          |          |                |
|--------------------|----|-----------|----------|-----------|--------|---------|----------|----------|----------------|
|                    |    | 农业        | 煤炭开采和洗选业 | 石油和天然气开采业 | 金属矿采选业 | 非金属矿采选业 | 食品制造业    | 烟草业      | 服装、皮革、羽绒制品及制鞋业 |
| 农业                 | 01 | 46365.66  | 3551.07  | 61        | 73933  | 17267   | 10651.05 | 11937.94 | 29652.27       |
| 木材加工及木制品业          | 02 | 87.513    | 13017.9  | 186562    | 61060  | 61815   | 433356   | 57995    | 36631          |
| 石油和天然气开采业          | 03 | 973       | 1967     | 32213     | 30022  | 63101   | 27101    | 73114    | 0              |
| 金属矿采选业             | 04 | 0         | 0        | 0         | 97363  | 0       | 0        | 0        | 0              |
| 非金属矿采选业            | 05 | 83659     | 29987    | 2554      | 10269  | 772670  | 11809    | 0        | 0              |
| 食品制造业及烟草加工业        | 06 | 15717.68  | 3949     | 14926     | 6182   | 0       | 19537192 | 314969   | 2419154        |
| 纺织业                | 07 | 168326    | 2314     | 64736     | 10786  | 14732   | 137156   | 32590683 | 18930159       |
| 服装、皮革、羽绒制品及制鞋业     | 08 | 45061     | 129043   | 125971    | 71164  | 78637   | 10708    | 253921   | 838711         |
| 木材加工及木制品业          | 09 | 15813     | 26197    | 39189     | 31884  | 81784   | 119565   | 73140    | 112318         |
| 石油和天然气开采业          | 10 | 0         | 0        | 0         | 1113   | 7660    | 30500    | 15110    | 10709          |
| 金属矿采选业             | 11 | 205637    | 44671    | 68739     | 98260  | 47869   | 291693   | 264175   | 166911         |
| 化学工业               | 12 | 194245.02 | 938833   | 49073     | 723344 | 1154184 | 4769907  | 9173823  | 4319468        |
| 非金属矿物制品业           | 13 | 909991    | 322616   | 123297    | 109609 | 163207  | 828731   | 162063   | 162063         |
| 金属冶炼及压延加工业         | 14 | 321205    | 165929   | 59926     | 315200 | 113736  | 258710   | 98978    | 133017         |
| 金属制品业              | 15 | 838205    | 751646   | 251288    | 298020 | 132655  | 109167   | 161846   | 264132         |
| 通用、专用设备制造业         | 16 | 2345401   | 1171976  | 83770     | 596951 | 938387  | 812784   | 1133826  | 281395         |
| 交通运输设备制造业          | 17 | 1255199   | 327627   | 280907    | 265198 | 30075   | 382724   | 182769   | 96916          |
| 电气机械、计算机及其他电子设备制造业 | 18 | 250985    | 817549   | 165914    | 112813 | 37881   | 172298   | 294181   | 121415         |
| 仪器仪表及文化办公用机械制造业    | 19 | 109511    | 11132    | 120318    | 98095  | 10076   | 14070    | 158153   | 105235         |
| 其他制造业              | 20 | 109511    | 361825   | 333218    | 59995  | 65666   | 94916    | 86123    | 101314         |
|                    | 21 | 250985    | 165202   | 40948     | 64316  | 91772   | 197258   | 273933   | 21526          |

CGE模型



DSGE模型



# 博弈论有不有用？

- 现实博弈问题中，各方采用均衡策略吗？

|    |                            | 美国             |                 |                 |                 |                 |
|----|----------------------------|----------------|-----------------|-----------------|-----------------|-----------------|
|    |                            | GDP%           | 0关税             | 维持现状            | 500亿，25%        | + 2000亿，10%     |
| 中国 | 0关税                        | 0.3903, 0.1539 | -0.1648, 0.3818 | -0.1720, 0.3841 | -0.1897, 0.3907 | -0.2162, 0.4005 |
|    | 维持现状                       | 0.5551,-0.2280 | 0, 0            | -0.0072, 0.0023 | -0.0249, 0.0088 | -0.0513, 0.0187 |
|    | 500亿<br>25%                | 0.5650,-0.2483 | 0.0098,-0.0203  | 0.0027,-0.0180  | -0.0150, 0.0115 | -0.0415,-0.0016 |
|    | + 600亿<br>(25%、20%、10%、5%) | 0.5717,-0.2663 | 0.0166,-0.0384  | 0.0094,-0.0361  | -0.0083,-0.0295 | -0.0548,-0.0197 |
|    | + 750亿<br>(5%、10%)         | 0.5782,-0.2803 | 0.0231,-0.0523  | 0.0159,-0.0500  | -0.0017,-0.0434 | -0.0282,-0.0336 |
|    |                            |                |                 |                 |                 |                 |

→ 双方占优策略

随着关税制裁力度的升级，双方的贸易收益不断恶化。

在每一轮博弈中，制裁均构成中美双方的占优策略，从经济的角度，这解释了为何双方选择“硬碰硬”。