인공지능의 기초

게임 이론

Games

- The decision-making process in situations where outcomes depend upon choices made by one or more players
  - Game is not used in conventional sense
  - Game describes any situation involving positive or negative outcomes determined by the players' choices and/or chances

- Requirements for a game
  - At least 2 players
  - Actions available for each player
  - Payoffs to each player for those actions

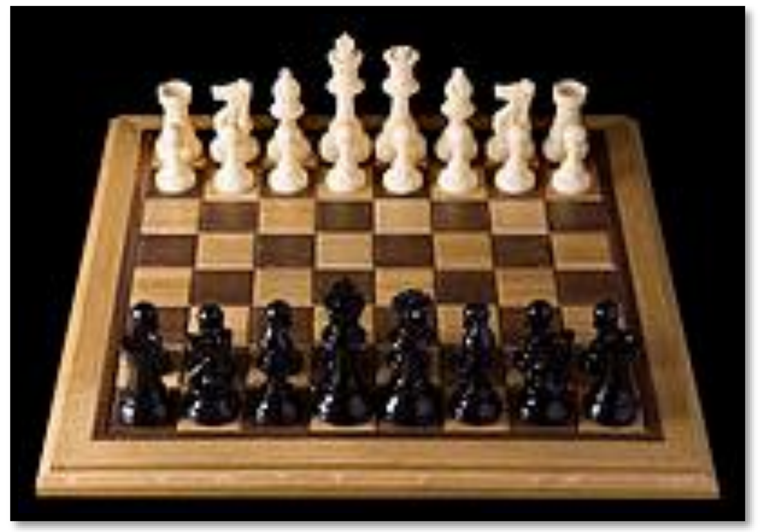
MEMO

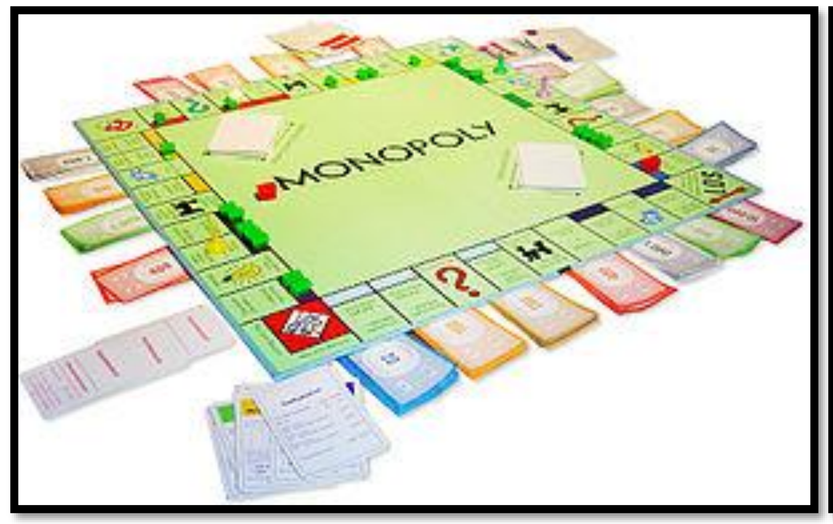
Games

Assumptions in Game Theory:

1) All players want to maximize their utility
2) All players are rational
3) It is common knowledge that all players are rational

Examples?





MEMO

[Toumas Sandholm]

Prisoner's Dilemma

- **Two prisoners P_1, P_2 being interrogated, can either stay silent or implicate the other one**
  - If both stay silent, each sentenced to a year in jail
  - If only one implicates another, he goes free and other gets 5 years in jail
  - If both implicate each other, both get 3 years
  - Even though (Silent/Silent) is best for both, each one strictly benefits from implicating the other, regardless of other's actions

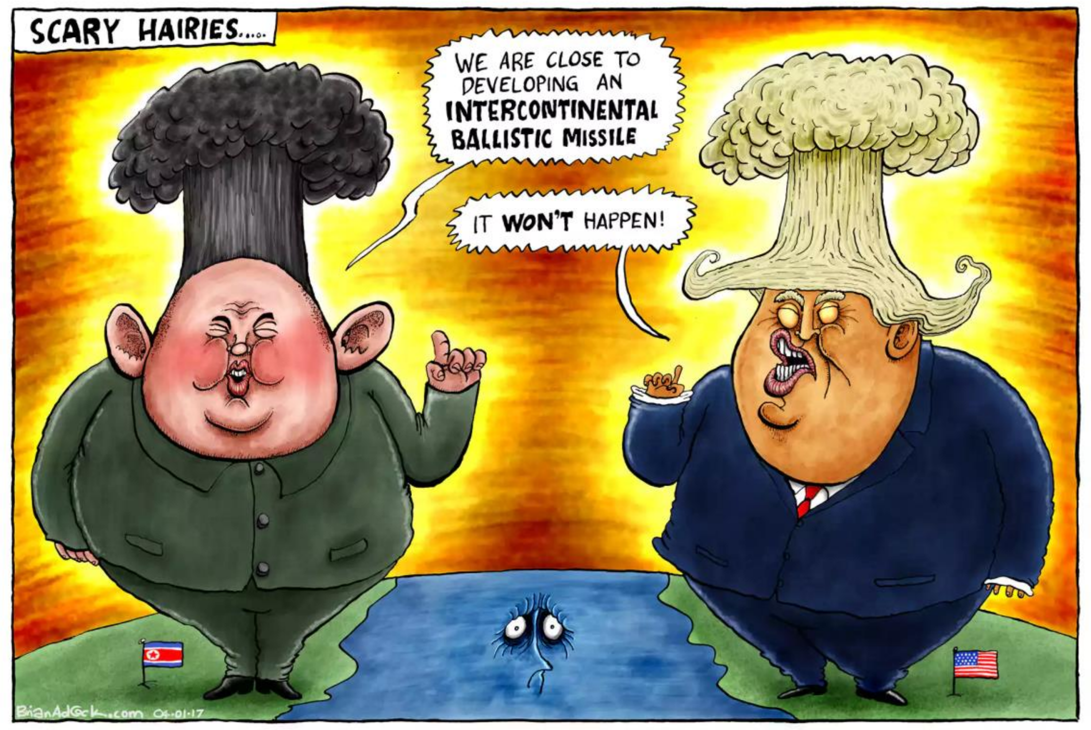
$$\begin{array}{ccc} & & P_2\text{'s action} \\ & & \begin{array}{cc} \text{Silent} & \text{Implicate} \end{array} \\ P_1\text{'s action} & \begin{array}{c} \text{Silent} \\ \text{Implicate} \end{array} & \begin{pmatrix} -1,-1 & -5,0 \\ 0,-5 & -3,-3 \end{pmatrix} \end{array}$$

MEMO

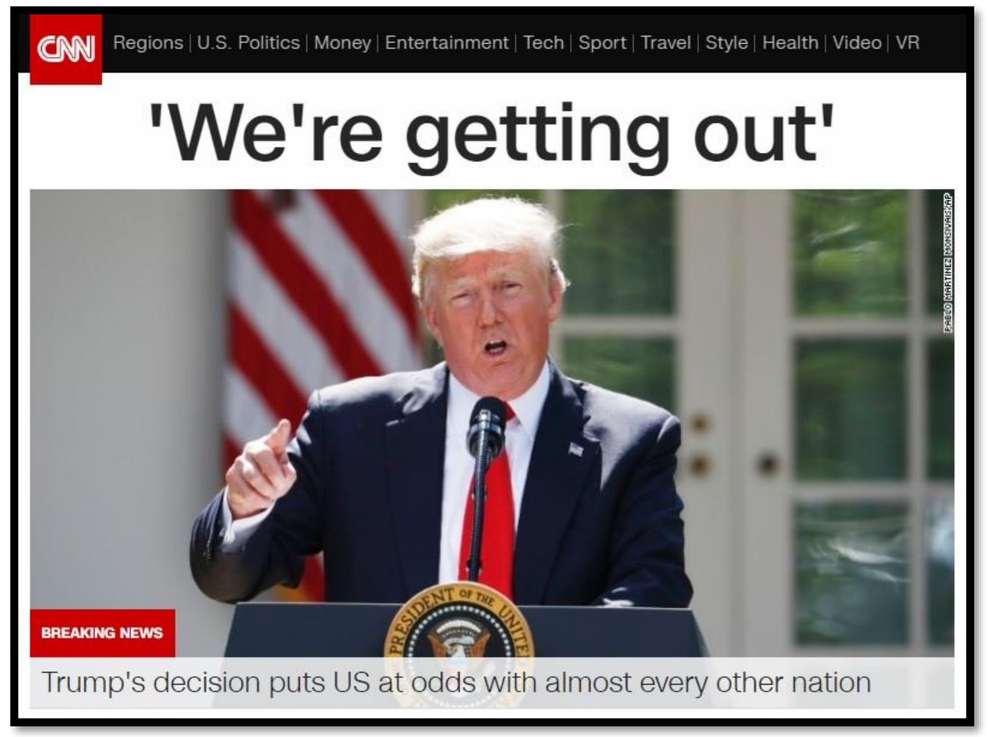
Understanding Dilemma

- **Related to the tragedy of the commons**
  - Defection is a dominant strategy
  - But the players can do much better by cooperating

- **e.g. Nuclear arms race**
  - Cooperate = destroy arsenal
  - Defect = build arsenal



- **e.g. Climate change**
  - Cooperate = curb CO_2 emissions
  - Defect = do not curb



MEMO

[Ariel Procaccia]

Normal Form

- Games in normal form defined by $(\boldsymbol{N}, \boldsymbol{A}, \boldsymbol{u})$
  - N: # players, each indexed by i
  - $\mathrm{A} = \mathrm{A}_1 \times \cdots \times \mathrm{A}_\mathrm{N}$: a set of actions, where A_i is a finite set of actions available to i
  - $\mathrm{u}: \mathrm{A} \to \mathbb{R}^\mathrm{N}$: utility function that maps each set of action $\mathrm{a} \in \mathrm{A}$ to
    a set of utilities N, one for each agent
  - e.g. $\mathrm{u}_\mathrm{i}(\mathrm{a})$: the utility of player i for action a

MEMO

Prisoner's Dilemma in Normal Form

Games in normal form defined by $(\mathbf{N}, \mathbf{A}, \mathbf{u})$

- $\mathrm{N} = 2$: # players
- $\mathrm{A} = \mathrm{A}_1 \times \cdots \times \mathrm{A}_\mathrm{N} = \{\mathrm{S}, \mathrm{I}\} \times \{\mathrm{S}, \mathrm{I}\}$: a set of actions
- $\mathrm{u}: \mathrm{A} \to \mathbb{R}^\mathrm{N}$: utility function

$$
\begin{aligned}
u(a) &= \begin{bmatrix} -1 \\ -1 \end{bmatrix} \quad \text{if } a = (S, S) \\
&= \begin{bmatrix} -5 \\ 0 \end{bmatrix} \quad \text{if } a = (S, I) \\
&= \begin{bmatrix} 0 \\ -5 \end{bmatrix} \quad \text{if } a = (I, S) \\
&= \begin{bmatrix} -3 \\ -3 \end{bmatrix} \quad \text{if } a = (I, I)
\end{aligned}
$$

$$
\begin{array}{c c}
 & \begin{array}{cc} \text{Silent} & \text{Implicate} \end{array} \\
\begin{array}{c} \text{Silent} \\ \text{Implicate} \end{array} & \begin{pmatrix} -1, -1 & -5, 0 \\ 0, -5 & -3, -3 \end{pmatrix}
\end{array}
$$

MEMO

Zero-Sum Game

Game where no wealth is created or destroyed

- The two-player game (i.e. $N = 2$) is zero-sum if

$$u_1(a) = -u_2(a), \qquad \forall a \in A$$

- Player 1 is trying to maximize $u_1(a)$ and player 2 is trying to maximize $u_2(a) = -u_1(a)$ (i.e. minimize $u_1(a)$)
- e.g. rock-paper-scissors

$$\begin{array}{c c} & \begin{array}{ccc} \text{R} & \text{P} & \text{S} \end{array} \\ \begin{array}{c} \text{R} \\ \text{P} \\ \text{S} \end{array} & \begin{pmatrix} 0,0 & -1,1 & 1,-1 \\ 1,-1 & 0,0 & -1,1 \\ -1,1 & 1,-1 & 0,0 \end{pmatrix} \end{array}$$

- e.g. Prisoner's Dilemma is NOT!

MEMO

Pure and Mixed Strategies

- **A strategy for player $\boldsymbol{i}$, denoted $\boldsymbol{s_i : a_i \rightarrow [0, 1]}$ is a probability distribution over actions**
  - $s_i(a_i)$ denotes the probability that player i takes action a_i
  - A strategy profile $s = (s_1, \dots, s_N)$ is a set of strategies for all players
  - The support of a strategy s_i is the set of actions that have non-zero probability
  - A strategy is pure if players use the same strategy for every time
  - A strategy is mixed if players introduce a randomness in their choice of strategy

MEMO

Nash Equilibrium

- **Best response: a best strategy that a player can play given the strategies of all opponents**
  - Let $s_{\neg i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_N)$ a strategy profile except player i
  - The best response for player i given strategy profile $s_{\neg i}$ is the strategy s_i^* such that $u(s_i^*, s_{\neg i}) \geq u(s_i, s_{\neg i})$ for all possible s_i

- **A strategy profile $\boldsymbol{s}$ is a Nash equilibrium if $\boldsymbol{s_i}$ is a best response to $\boldsymbol{s_{\neg i}}$ for all players $\boldsymbol{i = 1, \ldots, N}$**
  - No agent gains an advantage by switching their strategy
  - Can be one or more Nash equilibria for a game

- **What is Nash equilibrium for zero-sum game?**

MEMO

Two-player Zero-Sum Game

Zero-sum game can be represented by the matrix

$$
\begin{array}{cc} & \begin{array}{ccc} \text{R} & \text{P} & \text{S} \end{array} \\ \begin{array}{c} \text{R} \\ \text{P} \\ \text{S} \end{array} & \begin{pmatrix} 0,0 & -1,1 & 1,-1 \\ 1,-1 & 0,0 & -1,1 \\ -1,1 & 1,-1 & 0,0 \end{pmatrix} \end{array} \Rightarrow \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}
$$

- Let's generalize $\begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix}$
- Player 1 chooses the row while player 2 chooses the column
- x_{ij}: the payoff (utility) for player 1 when player 1 select i-th row and player 2 selects j-th column

MEMO

Pure Maxmin Strategies

$$\begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix}$$

The maxmin strategy for player 1

- Sees what her worst outcome for each row, and then selects the row where her worst outcome is the best
- $v_i^1 = \min\{x_{i1}, \ldots, x_{in}\}$ is the worst payoff that P1 selects the action i
- The maxmin strategy for P1 is to choose $v^1 = \max\{v_i^1, \ldots, v_m^1\}$

Similarly, the minmax strategy for player 2

- Looks for the largest entry in each column and then chooses the column to minimize the largest entry
- $v_j^2 = \max\{x_{1j}, \ldots, x_{mj}\}$ is the worst payoff for P2 with action j
- The minmax strategy for P2 is to choose $v^2 = \min\{v_i^2, \ldots, v_m^2\}$

MEMO

Example of Pure Maxmin Strategies

$$\begin{pmatrix} 2 & 0 & 1 \\ 4 & -3 & 2 \\ 1 & -2 & -2 \end{pmatrix}$$

The maxmin strategy for player 1

- $v_1^1 = 0, v_2^1 = -3, v_3^1 = -2 \implies v^1 = 0$: P1's maxmin strategy is a_1
- since $v_i^1 = \min\{x_{i1}, \ldots, x_{in}\}, v^1 = \max\{v_i^1, \ldots, v_m^1\}$

The maxmin strategy for player 2

- $v_1^2 = 4, v_2^2 = 0, v_3^2 = 2 \implies v^2 = 0$: P1's minmax strategy is b_2

The maxmin solution

- If $v^1 = v^2 = v$, then v is called value of the game
- v and (a_i, b_j) is a maxmin (i.e. minmax) solution

MEMO

Example of Pure Maxmin Strategies

$$\begin{pmatrix} 2 & 0 & 1 \\ 4 & -3 & 2 \\ 1 & -2 & -2 \end{pmatrix}$$

- **Meaning of the maxmin solution ($v^1 = v^2 = v = 0$, (a_1, b_2))**
  - Regardless of P2's strategy, P1's strategy (a_1) guarantees a payoff v
  - Regardless of P1's strategy, P2's strategy (b_2) guarantees a payoff $-v$
  - The maxmin solution is called a game saddle point, because it's minimum of its row, and the maximum of its column
- **(a_i, b_j) is a maxmin solution to a zero-sum game iff (a_i, b_j) is a Nash equilibrium**

MEMO

Multi-player Games

A game having several players

- Players can be independent opponents or teams...
- Quickly difficult...

Nash equilibrium

- Games with several players have a stable solution provided that coalitions between players are disallowed
- If cooperation between players is allowed, then the game becomes more complex
- No good general theory has yet been developed

MEMO

Non-Zero-Sum Games

- In non-zero-sum games, there is no universally accepted solution
  - In zero-sum games, completely competitive players are assumed
  - But, here non-strictly competitive (have both competitive and cooperative elements)
- Often consider some special cases only
  - Prisoner's dilemma
  - Stag hunt
  - Coordination game
  - Game of Chicken

MEMO

Stag Hunt

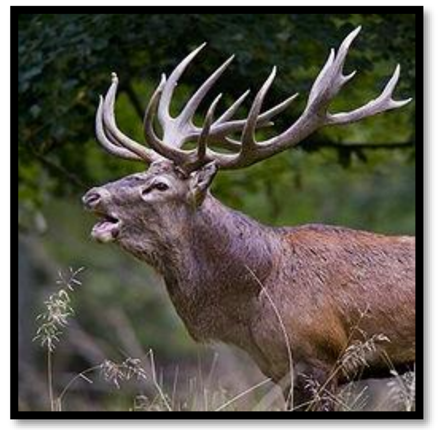
- **a.k.a assurance game, coordination game, trust dilemma**
  - Described by Jean-Jacques Rousseau
  - Each of two hunters individually choose to hunt a stag or a hare
  - Each player chooses without knowing the choice of the other
  - Stag hunt can succeed by the cooperation of two players
  - Hare hunt can be individually, but a hare is worth less than a stag
  - $a > b \geq d > c$

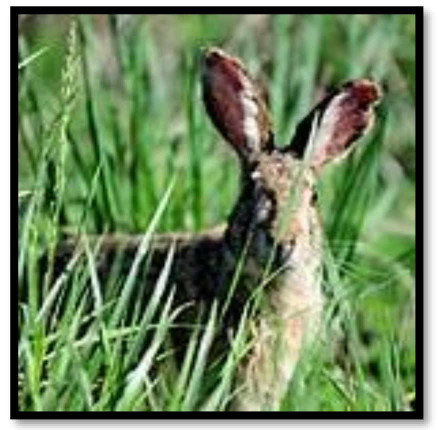
P_2's action

P_1's action

$$\begin{array}{cc} & \begin{array}{cc}\text{Stag} & \text{Hare}\end{array} \\ \begin{array}{c}\text{Stag}\\ \text{Hare}\end{array} & \begin{pmatrix} a,a & c,b \\ b,c & d,d \end{pmatrix}\end{array}$$

Generic symmetric stag hunt





$$\begin{array}{cc} & \begin{array}{cc}\text{Stag} & \text{Hare}\end{array} \\ \begin{array}{c}\text{Stag}\\ \text{Hare}\end{array} & \begin{pmatrix} 2,2 & 0,1 \\ 1,0 & 1,1 \end{pmatrix}\end{array}$$

Stag hunt example

MEMO

Battle of the Sexes

An example of two-player coordination game

- The husband would prefer to go to the football game, while the wife would rather go to the ballet
- If they cannot communicate, where should they go?

Is there a Nash equilibrium?

- (ballet, ballet) or (football, football)

$$\begin{array}{ccc} & & \text{Wife} \\ & & \begin{array}{cc}\text{Ballet} & \text{Football}\end{array} \\ \text{Husband} & \begin{array}{c}\text{Ballet} \\ \text{Football}\end{array} & \begin{pmatrix} 2,3 & 0,0 \\ 0,0 & 3,2 \end{pmatrix} \end{array} \quad \text{or} \quad \begin{array}{ccc} & & \text{Wife} \\ & & \begin{array}{cc}\text{Ballet} & \text{Football}\end{array} \\ \text{Husband} & \begin{array}{c}\text{Ballet} \\ \text{Football}\end{array} & \begin{pmatrix} 2,3 & 0,0 \\ 1,1 & 3,2 \end{pmatrix} \end{array}$$

MEMO

Game of Chicken

- a.k.a hawk–dove game or snowdrift game
  - It is to both players' benefit if one player yields

- Is there a Nash equilibrium?
  - (straight, chicken) or (chicken, straight)

$$\begin{array}{c} \\ \text{Straight} \\ \text{Chicken} \end{array} \begin{array}{c} \begin{array}{cc} \text{Straight} & \text{Chicken} \end{array} \\ \begin{pmatrix} \text{Crash, Crash} & \text{Win, Loss} \\ \text{Loss, Win} & \text{Tie, Tie} \end{pmatrix} \end{array} \qquad \begin{array}{c} \\ \text{Straight} \\ \text{Chicken} \end{array} \begin{array}{c} \begin{array}{cc} \text{Straight} & \text{Chicken} \end{array} \\ \begin{pmatrix} -10,-10 & 1,-1 \\ -1,1 & 0,0 \end{pmatrix} \end{array}$$

MEMO

Game of Chicken

- **Anti-coordinate game: it is mutually beneficial for the two players to choose different strategies**
  - Model of escalated conflict in humans or animals

- **Hawk-dove game**
  - Use threat displays (play Dove) or physically attack each other (play Hawk)
- **How are the players to decide what to do?**
  - Pre-commitment or threats

$$\begin{array}{c c} & \begin{array}{cc} \text{Straight} & \text{Chicken} \end{array} \\ \begin{array}{c} \text{Straight} \\ \text{Chicken} \end{array} & \begin{pmatrix} \text{Crash, Crash} & \text{Win, Loss} \\ \text{Loss, Win} & \text{Tie, Tie} \end{pmatrix} \end{array}$$

$$\begin{array}{c c} & \begin{array}{cc} \text{Straight} & \text{Chicken} \end{array} \\ \begin{array}{c} \text{Straight} \\ \text{Chicken} \end{array} & \begin{pmatrix} -10,-10 & 1,-1 \\ -1,1 & 0,0 \end{pmatrix} \end{array}$$

MEMO