CS711: Introduction to Game Theory and Mechanism Design

Teacher: Swaprava Nath

Strategy, Rationality, Common Knowledge

Quick Recap

Game theory

- Analytical approach for predicting reasonable outcome
- Fundamental building blocks: **players**, **strategies**, **utilities**
- Difference between **action** and **strategy**
- Key assumptions: **rationality** and **intelligence**

Strategy of the game of Chess

- History: von Neumann and Morgenstern, Theory of Games and Economic Behavior, 1944
- Schematic description of chess
  - Two player game: White and Black – 16 pieces each
  - Every piece has some legal moves – **actions**
  - The game progresses with each player taking turns and making legal moves – starts with White
  - Ends at
    - Win for White, if White captures the Black King
    - Win for Black, if Black captures the White King
    - Draw – if Black has no legal move but the King is not in check, both players agree to a draw, a board position where no player can win, ...
- In the game of chess,
  - Does White have a winning strategy? – a strategy with which White wins irrespective of Black's strategies
  - Does Black have a winning strategy?
  - Or neither is true?
- What is a **strategy**?

Game Situation

- Board Position is different from **Game Situation**
- More than one sequence of moves can lead to the same board position
- Denote a board position by x_k, set of all possible board positions X

Definition (Game Situation)

A *game situation* in chess is a finite sequence $(x_0, x_1, x_2, \ldots, x_K)$ of board positions, $x_k \in X, k = 0, \ldots, K$, such that

- x_0 is the opening board position
- even k, $x_k \to x_{k+1}$ is achieved by a single action of White
- odd k, $x_k \to x_{k+1}$ is achieved by a single action of Black

Set of all game situations: H

Graphical Interpretation

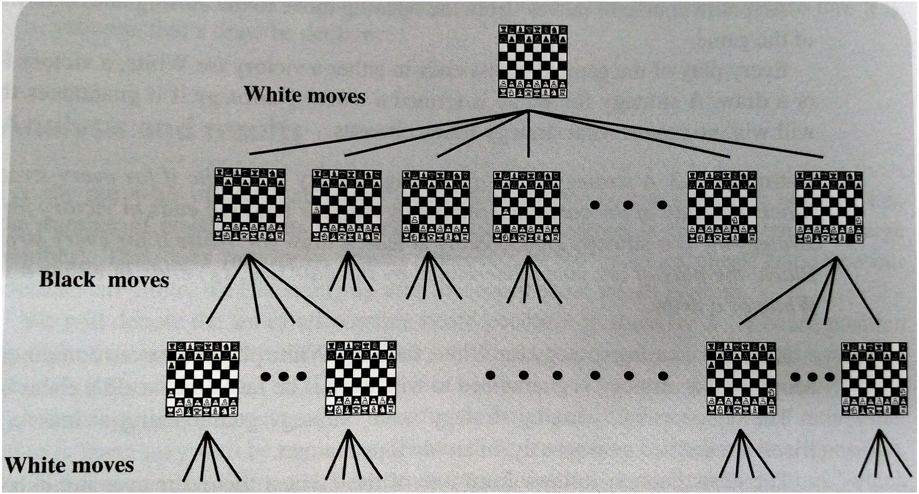
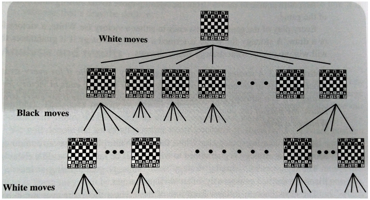


Image courtesy: Maschler et al., Game Theory.

Illustration



- Game tree lists all possible game situations
- Every *vertex* is a **game situation**
- Could have repeated board positions
- The immediate children – consequence of the **actions** of the player
- **strategy**: mapping from game situation to action
- *plan of action* in a given *game situation*
- The complete plan is a strategy

Strategy

Definition (Strategy)

A *strategy* for White is a function s_W that associates every game situation $(x_0, x_1, \dots, x_K) \in H$, where K is even, with a board position x_{K+1}, such that the transition $x_K \to x_{K+1}$ can be accomplished by a single legal move of White. Similarly, A *strategy* for Black is a function s_B that associates every game situation $(x_0, x_1, \dots, x_K) \in H$, where K is odd, with a board position x_{K+1}, such that the transition $x_K \to x_{K+1}$ can be accomplished by a single legal move of Black.

- $(x_0, x_1, \dots, x_K)$ denotes a node in the game tree
- strategy maps this to an action – contrast with prisoner's dilemma
- strategy pair (s_W, s_B) determines an **outcome**

$$x_1 = s_W(x_0),\ x_2 = s_B(x_0, x_1),\ \dots, x_{2k+1} = s_W(x_0, \dots, x_{2k})$$

- entire course of moves – one **play** of the game
- this is a finite game – where does the game end? can the players guarantee a given end?

Winning Strategy

- Every play ends in either (a) *win for White*, (b) *win for Black*, or (c) *draw*
- A **winning strategy** for W is a strategy which makes W win irrespective of the strategy chosen by B

Definition (Winning Strategy)

A strategy s^*_W is a *winning strategy for W* if for every strategy s_B of B, the play of the game determined by (s^*_W, s_B) ends in a victory for W.
A strategy s'_W is a *strategy guaranteeing at least a draw for W* if for every strategy s_B of B, the play of the game determined by (s'_W, s_B) ends in either a victory for W or a draw.

- the winning strategy or strategy guaranteeing at least a draw for B is analogous
- not obvious if such a strategy exists – this is a property of the mappings

An Early Result of Game Theory

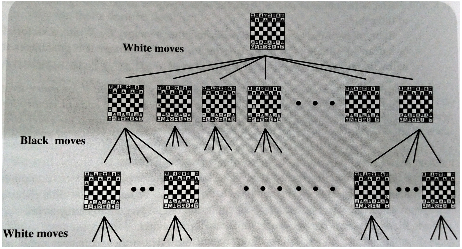
Theorem (von Neumann, 1928)

In chess, one and only one of the following statements must be true:

1. *White has a winning strategy*
2. *Black has a winning strategy*
3. *Each of the players has a strategy guaranteeing at least a draw*

- applies to every game of chess
- clearly no two events can happen together
- this is an exhaustive list – nothing apart from this happens, and exactly one of them is true
- **significant:** it is not known
  - which of the three is true
  - what is the winning/guaranteeing a draw strategy
- chess will be a boring game if the answers were known

Proof



Every *vertex* x is a **game situation**, i.e., $x \in H$
$\Gamma(x)$: subtree rooted at x
$\Gamma(x_0)$: whole game that starts from the initial position
n_x: number of vertices in $\Gamma(x)$
y is a child vertex of x, i.e., $y \in \Gamma(x) \setminus \{x\}$
$\Gamma(y)$ is a subtree of $\Gamma(x)$, $n_y < n_x$
if $n_x = 1$, then x is a terminal vertex – strategy of the player $= \emptyset$

$$\mathcal{F} = \{\Gamma(x) : x \in H\}$$

collection of all subgames defined by the subtrees of the game of chess

Proof (Contd.)

The following lemma rephrases the earlier theorem

Lemma

Every game in $\mathcal{F}$ satisfies one and only one of the following statements:

1. *White has a winning strategy*
2. *Black has a winning strategy*
3. *Each of the players has a strategy guaranteeing at least a draw*

induction on n_x, the number of vertices in $\Gamma(x)$
suppose x is such that $n_x = 1$, leaf vertex
W King is removed, B wins, $\emptyset$ is the winning strategy for B
B King is removed, W wins, $\emptyset$ is the winning strategy for W
both Kings are on board and game has ended implies a draw

Proof (Contd.)

suppose x is a vertex with $n_x > 1$
induction hypothesis: for all vertices y satisfying $n_y < n_x$, one and only one of (1), (2), (3) holds for $\Gamma(y)$

WLOG, assume W moves first in $\Gamma(x)$
consider any board position y reachable from x, i.e., $y \in \Gamma(x) \setminus \{x\}$, $n_y < n_x$ and the induction hypothesis holds
denote by $C(x)$ the vertices reachable from x via one move by W

(i) if $\exists y_0 \in C(x)$ s.t. alternative (1) is true in $\Gamma(y_0)$, then (1) is true in $\Gamma(x)$ as well: W picks the action to reach y_0 augmented with the winning strategy at y_0

(ii) if $\forall y \in C(x)$, alternative (2) is true in $\Gamma(y)$, then (2) is true in $\Gamma(x)$ as well: B identifies which action was taken by W (hence which vertex y is reached) and pick the winning strategy from there

Proof (Contd.)

(iii) else

- (i) does not hold, i.e., W does not have a winning strategy in any $y \in C(x)$, since induction hypothesis holds for every $y \in C(x)$, either B has a winning strategy or both have a strategy of guaranteeing at least a draw in $\Gamma(y)$
- (ii) does not hold, i.e., $\exists y_0 \in C(x)$ where B does not have a winning strategy in $\Gamma(y_0)$, but since (i) does not hold either, W does not have a winning strategy in $\Gamma(y_0)$, by induction hypothesis, both players have a strategy to guarantee at least a draw

in this case neither W nor B can guarantee a win, but both can guarantee at least a draw

W can pick action to reach y_0 and pick the strategy to guarantee at least a draw

B can watch W pick an action that reaches $y \in C(x)$ and then pick the strategy that ensures either win or at least a draw

This concludes the proof

Exercise: prove this theorem when the length of the game is infinite (ex. 1.3, MSZ book)

Game Representations

- Normal form / strategic form – appropriate for single shot games
- Extensive form – appropriate for sequential games

The setting of normal form game representation

- $N = \{1, 2, \ldots, n\}$ – set of **players**
- S_i: set of **strategies** of player i, $s_i \in S_i$
- set of **strategy profiles** $S = \bigtimes_{i \in N} S_i$
- a strategy profile $s = (s_1, s_2, s_3, \ldots, s_n) \in S$
- $s_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$
- $S_{-i} = \bigtimes_{j \neq i} S_j$
- $u_i : \bigtimes_{i \in N} S_i \to \mathbb{R}$ – **utility function** of player i
- NFG representation is a ordered tuple $\langle N, (S_i)_{i \in N}, (u_i)_{i \in N} \rangle$
- if S_i is finite – the game is called a **finite game**

Example: Rock-Paper-Scissor

1$\backslash$**2**	Rock	Paper	Scissor
Rock	0,0	-1,1	1,-1
Paper	1,-1	0,0	-1,1
Scissor	-1,1	1,-1	0,0

- $N = \{1, 2\}$
- $S_1 = S_2 = \{R, P, S\}$
- $u_1(R, R) = 0, u_1(R, P) = -1, u_1(R, S) = 1$
- $u_1(P, R) = 1, u_1(P, P) = 0, u_1(P, S) = -1$
- $u_1(S, R) = -1, u_1(S, P) = 1, u_1(S, S) = 0$

Players' Knowledge and Behavior

Definition (Rationality)

A player is *rational* if she picks actions to **maximize** her utility

Definition (Intelligence)

A player is *intelligent* if she knows the rules of the game perfectly and pick an action considering that there are other rational and intelligent players in the game.

Definition (Common Knowledge)

A fact is a *common knowledge* if

1. All players know the fact, and
2. All players know that all other players know the fact, and
3. All players know that all other players know that all other players know the fact, and ... ad infinitum.

Implication of Common Knowledge

- Isolated island – three blue-eyed individuals (eyes can be either blue or black)
- Assume they do not talk about their eye color and there is no reflecting media
- One day a sage comes to the island and says "Blue-eyed people are bad for the island and must leave. There is at least one blue-eyed person in this island"
- Assume that the sage's statements cannot be disputed – if a person realizes that his eye color is blue, he leaves at the end of the day
- common knowledge percolates to the outcome in the following way

- If there were only one blue-eyed person, he would have seen that the other two had black eyes, realized that his eye color is blue (since sage is always correct), leaves at the end of day one– every other player understands this and stays back

Common Knowledge (Contd.)

- if there were two blue-eyed persons, then both of them will see one blue and one black eyed person, hope that he is not blue-eyed and wait till the second day if the other blue-eyed person leaves on day one
- when it does not happen, he realizes that both of them had blue eyes, so they both leave at the end of day two, the third player understands this and does not leave

- since there are three blue-eyed persons, then extending the same argument, we see that every player will wait till day three if anyone leaves
- when nobody left on day two, it becomes clear that all of them had blue eyes, and they all leave at the end of day three

Assumption

The fact that all players are rational and intelligent is a **common knowledge**

Domination

1\2	L	M	R
U	1,0	1,3	3,2
D	-1,6	0,5	5,3

- Will a rational player 2 ever play R?

Definition (Dominated Strategy)

A strategy $s_i' \in S_i$ of player i is **strictly dominated** if there exists another strategy s_i of i such that for every strategy profile $s_{-i} \in S_{-i}$ of the other players

$$u_i(s_i, s_{-i}) > u_i(s_i', s_{-i}).$$

A strategy $s_i' \in S_i$ of player i is **weakly dominated** if there exists another strategy s_i of i such that for every strategy profile $s_{-i} \in S_{-i}$ of the other players

$$u_i(s_i, s_{-i}) \geqslant u_i(s_i', s_{-i}),$$

and there exists some $\tilde{s}_{-i} \in S_{-i}$ such that

$$u_i(s_i, \tilde{s}_{-i}) > u_i(s_i', \tilde{s}_{-i}).$$

Domination (Contd.)

Definition (Dominant Strategy)

A strategy s_i is **strictly (weakly) dominant strategy** for player i if s_i strictly (weakly) dominates all other $s_i' \in S_i \setminus \{s_i\}$.

Definition (Dominant Strategy Equilibrium)

A strategy profile (s_i^*, s_{-i}^*) is a **strictly (weakly) dominant strategy equilibrium (SDSE (WDSE))** if s_i^* is a strictly (weakly) dominant strategy for every $i, i \in N$.

	D	E
A	$5,5$	$0,5$
B	$5,0$	$1,1$
C	$4,0$	$1,1$