

---

# CS 6001 : GAME THEORY AND ALGORITHMIC MECHANISM DESIGN

---

**Amit Rajaraman**

Last updated August 6, 2022

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Chess . . . . .	3
1.2	Normal Form Games . . . . .	4
1.3	Equilibria . . . . .	5
1.4	Matrix Games . . . . .	8

## §1. Introduction

In typical linear programming, we have an objective that a *single* individual is trying to maximize (subject to some constraints). In game theory, we typically study systems where there are *multiple* individuals with objectives that they are trying to maximize, but each individual can only set some of the variables (the set of variables is shared). An optimal solution for one individual might not be the optimal solution for another. As a result, it is now better to look at *equilibria* instead of pure black-and-white optimality.

These equilibria are what we shall study.

The course is broadly split in two parts:

- *Game theory*, where we study games, interactions between agents who want to maximize their utilities. This is a predictive approach. We wish to find the most probable outcomes or responses of the agents/players.
- *Mechanism design*, where we try to design a game with desirable outcomes. This is a prescriptive approach. Given a reasonable outcome, we wish to build a game that yields this as a probable outcome.

The above is essentially analysis versus synthesis.

A well-known example of game theory is the *Prisoner's Dilemma*, which we slightly rephrase in the following.

**Example.** Two kingdoms can each decide to dedicate all their resources to either agriculture or warfare. If both kingdoms choose warfare, they each earn one unit of joy. If one chooses agriculture and the other chooses warfare, the former earns nothing while the latter earns six units of joy. Finally, if both choose agriculture, they each earn five units of joy.

This can be represented by the following matrix.

A, B	Agriculture	War
Agriculture	5, 5	0, 6
War	6, 0	1, 1

The above matrix is called a *game matrix*. Each of the two numbers in the cells are referred to as the *utilities* of the respective kingdoms.

Suppose that kingdom A decides to invest in agriculture. In this case, note that irrespective of what the other kingdom decides to choose, A can increase its payoff by switching to investing in warfare. Consequently, it makes more sense to invest in warfare (since the game is symmetric, this is true for both players). Even though both kingdoms heading to war is not the most profitable outcome, it appears to be the most likely outcome.

A *game* is a formal representation of the strategic interaction between multiple agents called *players*. The choices available to the players are called *actions*. A mapping from the state of the game to the set of actions is referred to as a *strategy*.

Depending on the context, games can be represented in different ways, normal form, extensive form, repeated form, stochastic form, etcetera.

*Game theory* is the formal study of the strategic interactions between players who are rational and intelligent.

A player is rational if they pick actions to maximize their payoff.

A player is intelligent if they know the rules of the game perfectly and picks actions assuming that the remaining players are also rational and intelligent. This also assumes that the player has sufficient computational ability to find the “optimal” action.

Let us now look at an example of mechanism design.

**Example.** Suppose we want to split a cake in two parts (for two children, say) in an “envy-free” fashion. That is, neither child would prefer the other piece of cake. We do not see the children’s preferences, so we do not even know what a fair division might involve. This well-known problem has a well-known solution – make one child cut the cake and the other choose the piece. Why does this work? The first child splits it in a way that is exactly half from their perspective, and they are indifferent to the two pieces. The second child on the other hand gets a larger piece in their perspective.

## 1.1. Chess

The reader is no doubt familiar with the rules of chess. It has two players, *White* and *Black*, with 16 pieces each. Each piece has some legal moves (the players’ actions are these moves). The game starts with White and players take turns. White wins if they capture Black’s king, and Black wins they capture White’s king. There are a couple of conditions under which a draw can occur, which we do not detail.

There are numerous natural questions that arise. First and foremost, does White (or Black) have a winning strategy? A winning strategy is a plan of moves such that it wins irrespective of the moves performed by Black. Alternatively, is it possible to guarantee a draw? It is possible for none of these exist.

What is a strategy? Denote a board position by  $x_k$ . A *game situation* is a finite sequence  $(x_0, \dots, x_k)$ , such that  $x_0$  is the opening board position and  $x_k \rightarrow x_{k+1}$  for even (resp. odd)  $k$  is created by a single action of White (resp. Black).

This set of game situations can be naturally represented by a tree, referred to as the *game tree*. The nodes are labelled with board positions  $(x_i)$ , and arrows between nodes are labelled with actions. A *strategy* is a mapping from game situation to action, which describes what action to take at every vertex of this game tree. This is something of a contingency plan for every possible situation.

A strategy pair  $(s_W, s_B)$  which describes strategies for both players determines an outcome, also called a *play* of the game. This describes a path through the game tree.

All leaves in the game tree correspond to either White’s victory, Black’s victory, or a draw.

A *winning strategy* for White is a strategy  $s_W^*$  such that for any strategy  $s_B$ ,  $(s_W^*, s_B)$  ends in a win for White. Similarly, we can define a *strategy guaranteeing at least a draw for White*, denoted  $s_W'$ .

It is not immediately obvious if such strategies exist.

**Theorem 1.1** (von Neumann, 1928). In chess, exactly one of the following statements is true.

- (1) White has a winning strategy.
- (2) Black has a winning strategy.
- (3) Each player has a strategy guaranteeing a draw.

If any such strategy was known, the game would become boring.

*Proof.* Each vertex of the game tree is a game situation. Denote by  $\Gamma(x)$  the subtree rooted at  $x$  (including  $x$  itself) and by  $n_x$  the number of vertices in  $\Gamma(x)$ .  $n_x = 1$  implies that  $x$  is a terminal vertex.

Using induction on  $n_x$ , we prove that one of the three statements must hold at any game situation. The theorem is clearly true for  $n_x = 1$ . Suppose  $x$  is a vertex with  $n_x > 1$ . By the inductive hypothesis, for all  $y \in \Gamma(x) \setminus \{x\}$ , the statement holds ( $n_y < n_x$ ). Let  $C(x)$  be the vertices reachable from  $x$  in one step, and assume wlog that it is White’s turn.

- (a) If there exists  $y_0 \in C(x)$  such that (1) is true, then (1) is true at  $x$  as well.
- (b) If (2) is true for all  $y \in C(x)$ , then Black will win irrespective of White’s move.

- (c) Otherwise, because (a) does not hold, White does not have a winning strategy for any  $y \in C(x)$ . As a result, for every  $y \in C(x)$ , either Black has a winning strategy or both have a draw-guaranteeing strategy. Because (b) does not hold, there is some  $y \in C(x)$  where Black does not have a winning strategy. By the preceding argument, both players must have a draw-guaranteeing strategy at this node.

■

*Remark.* The above proof is technically incorrect since the game tree for chess is infinite, so  $\Gamma(x)$  need not be general. This issue can be fixed through some clever manipulation of the “cycles” in the game tree to make the graph finite, but we do not go into more detail on this.

## 1.2. Normal Form Games

Normal form is a representation technique for games. The set  $N = \{1, \dots, n\}$  is the set of players.  $S_i$  is the set of strategies for player  $i$ . A particular strategy is denoted  $s_i \in S_i$ . The set of “strategy profiles” is  $S = \times_{i \in N} S_i$ , with specific elements  $s = (s_1, \dots, s_n) \in S$ . A strategy profile without  $i$  is  $s_{-i} = (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n)$ .  $u_i : S \rightarrow \mathbb{R}$  is the utility function of player  $i$ .

The *normal form game* (NFG) representation of a game is the tuple  $\langle N, (S_i)_{i \in N}, (u_i)_{i \in N} \rangle$ . This is the representation we shall use for the remainder of the course.

If  $S_i$  is finite for all  $i \in N$ , the game is said to be a *finite* game.

As mentioned earlier, a player is rational if they pick actions that maximize their utility. A player is intelligent if they know the rules of the game perfectly, and picks actions assuming that all other players are rational and intelligent.

A fact is said to be *common knowledge* if

1. all players know the fact,
2. the fact that “all players know the fact” is also common knowledge.

**Example.** There is an isolated island (with a hundred people, say) where all inhabitants have eye color either blue or black. There is no reflecting surface on the island (people cannot figure out their own eye color) and nobody can communicate with each other.

One day, a truth-telling god comes to the island and declares that all blue-eyed people are bad for the island and must leave as soon as possible. He also says that there is at least one blue-eyed person on the island. The inhabitants, being deeply devout, do listen to him and leave at the end of the day if they discover that their eyes are blue. In this setting, the fact that there is at least one blue-eyed person on the island is common knowledge.

If there was only one blue-eyed person, he would see that all other people have black eyes. Because the god said that there is a blue-eyed person, he infers that he must be the only blue-eyed person and leaves at the end of the first day.

If there were two, then on the second day everyone would notice that all people remain on the island, so they would infer that there are at least two blue-eyed people on the island. If one of the inhabitants sees that exactly one of the other four people is blue-eyed, then he, along with the other blue-eyed person, leaves on the second day.

This goes on, and it is seen that if there are exactly  $n$  blue-eyed people, then all of these  $n$  people leave at the end of the  $n$ th day.

Now, we discuss the concept of *domination* in NFGs. Consider the following game matrix:

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

Observe that player 2 has no reason to ever play R. Indeed, irrespective of what player 1 chooses, they can increase their payoff by switching to M instead. In such a scenario, we say that R is dominated by M.

**Definition 1.1.** A strategy  $s'_i \in S_i$  of player  $i$  is said to be *strictly dominated* if there exists another strategy  $s_i \in S_i$  such that for every strategy profile  $s_{-i} \in S_{-i}$ ,

$$u_i(s_i, s_{-i}) \geq u_i(s'_i, s_{-i}).$$

A strategy  $s'_i \in S_i$  of player  $i$  is said to be *strictly dominated* if there exists another strategy  $s_i \in S_i$  such that for every strategy profile  $s_{-i} \in S_{-i}$ ,

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

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

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

So, in the earlier example R is strictly dominated and D is weakly dominated.

There is no reason for a rational player to ever play a (weakly or strictly) dominated strategy.

**Definition 1.2.** A strategy  $s_i \in S_i$  is *strictly (weakly) dominant* if it strictly (weakly) dominates all  $s'_i \in S_i \setminus \{s_i\}$ .

Recall the agriculture/defense example we looked at towards the beginning of this section. In this, warfare strictly dominates agriculture, which is precisely what we said there, albeit in more explicit terms.

Let us give another example of this.

**Example.** Suppose there are two players having values  $v_1, v_2$  respectively. Each player can choose a number in  $[0, M]$ , where  $M \gg v_1, v_2$ . The player who quotes the larger number wins the object (with ties being broken in favour of player 1, say), and pays the losing player's chosen number. The utility of the winning player is their value  $v_i$  minus their payment (the amount bid by the other player), and the utility of the losing player is 0.

As a NFG representation,  $N = \{1, 2\}$ ,  $S_1 = S_2 = [0, M]$ , and

$$u_1(s_1, s_2) = \begin{cases} v_1 - s_2, & s_1 \geq s_2, \\ 0, & \text{otherwise,} \end{cases} \text{ and } u_2(s_1, s_2) = \begin{cases} v_2 - s_1, & s_1 < s_2, \\ 0, & \text{otherwise,} \end{cases}$$

It turns out that the strategy  $s_i$  where player  $i$  chooses  $v_i$  is a weakly dominant strategy! Let us check this for player 1.

Suppose that  $s'_1 > v_1$  is in  $S_1$  and  $s_2 \in S_2$ . We would like to show that  $u_1(s_1, s_2) \geq u_1(s'_1, s_2)$ . If  $s_2 > s'_1$ , both payoffs are zero. If  $s_2 \leq s_1 < s'_1$ , then both payoffs are equal to the same value. The remaining case is when  $s_1 \leq s_2 \leq s'_1$ . The payoff for  $s_1$  is zero, but the payoff for  $s'_1$  is non-positive since we are paying more than we value the item.

### 1.3. Equilibria

**Definition 1.3.** A strategy profile  $(s_1^*, s_2^*, \dots, s_n^*)$  is a *strictly (weakly) dominant strategy equilibrium* if each  $s_i^*$  is a strictly (weakly) dominant strategy for each  $i$ .

We abbreviate the above as SDSE or WDSE.

No rational player would play dominated strategies, so we can eliminate dominated strategies one-by-one. A point of note here is that after eliminating a strategy, we get a reduced game with fewer strategies, and this game may have dominated strategies that were not there earlier.

For strictly dominated strategies, the order of elimination does not matter. For weakly dominated strategies however, some reasonably outcomes may be eliminated (since this effectively says how we break ties).

**Example.** Consider the following.

	L	C	R
T	1,2	2,3	0,3
M	2,2	2,1	3,2
B	2,1	0,0	1,0

Right off the bat, it is seen that T, B, and C are weakly dominated strategies. Suppose we start by imposing that player 1 does not play T. If we do this, then R becomes weakly dominated as well, so it makes sense to eliminate it. Similarly, we can go on to eliminate B and C. Finally, the payoff of (M, L) is 2, 2.

On the other hand, if we eliminate strategies in the order of B, L, C, T, then the final remaining strategies are (M, R), which gives a payoff of 3, 2, which is not the same as the previous 2, 2!

Dominant strategies (and dominant strategy equilibria) need not exist in games in general, as can be seen in the following game.

	L	R
L	1, 1	0, 0
R	0, 0	1, 1

As a result, dominance is not enough to give a reasonable outcome, so we must give a more refined notion.

**Definition 1.4 (Nash Equilibrium).** A strategy profile  $(s_i^*, s_{-i}^*)$  is a *pure strategy Nash equilibrium* (PSNE) if for all  $s_i \in S_i$ ,

$$u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*).$$

That is, fixing the remaining players' strategies, no player can increase their payoff by moving to another strategy. Unilateral deviation is not beneficial.

In the above example, (L, L) and (R, R) are both PSNEs.

PSNEs need not exist either!

**Example.** The following game has no PSNE.

	L	R
L	-1, 1	1, -1
R	1, -1	1, -1

**Definition 1.5.** A *best response* of player  $i$  against a strategy profile  $s_{-i}$  is a strategy that gives the maximum utility. That is,

$$B_i(s_{-i}) = \{s_i \in S_i : u_i(s_i, s_{-i}) \geq u_i(s'_i, s_{-i}) \text{ for all } s'_i \in S_i\}.$$

So, a PSNE is a strategy profile  $(s_i^*, s_{-i}^*)$  such that  $s_i^* \in B_i(s_{-i}^*)$  for all  $i \in N$ .

A PSNE gives some sort of stability. Once there, no rational player has a reason to change their strategy.

One of our biggest assumptions thus far is that all players are rational and intelligent. There are, however, other types of rationality.

One is risk-aversion, where each player makes pessimistic estimates of others. This worst case optimal choice is called a max-min strategy.

**Definition 1.6.** A strategy  $s_i^{\max\min}$  is a *max-min strategy* for player  $i$  if

$$s_i^{\max\min} \in \arg \max_{s_i \in S_i} \min_{s_{-i} \in S_{-i}} u_i(s_i, s_{-i}).$$

The *max-min value* is defined by

$$\underline{v}_i = \max_{s_i \in S_i} \min_{s_{-i} \in S_{-i}} u_i(s_i, s_{-i}).$$

We have that for any  $t_{-i} \in S_{-i}$ ,

$$u_i(s_i^{\max\min}, t_{-i}) = \underline{v}_i.$$

**Theorem 1.2.** Any dominant strategy is a maxmin strategy.

*Proof.* Let  $s_i^*$  be a dominant strategy for player  $i$ . We have that for any  $s_{-i} \in S_{-i}$  and  $s'_i \in S_i \setminus \{s_i^*\}$ ,

$$u_i(s_i^*, s_{-i}) \geq u_i(s'_i, s_{-i}).$$

In particular, when we set  $s_{-i}$  as any  $s_{-i}^{\min}(s_i^*) \in \arg \min_{s_{-i} \in S_{-i}} u_i(s_i^*, s_{-i})$ , we get

$$\min_{s_{-i} \in S_{-i}} u_i(s_i^*, s_{-i}) = u_i(s_i^*, s_{-i}^{\min}(s_i^*)) \geq u_i(s'_i, s_{-i}^{\min}(s'_i)) \geq \min_{s_{-i} \in S_{-i}} u_i(s'_i, s_{-i})$$

That is,

$$s_i^* \in \arg \max_{s_i \in S_i} \min_{s_{-i} \in S_{-i}} u_i(s_i, s_{-i})$$

■

**Theorem 1.3.** Every PSNE  $s^* = (s_1^*, \dots, s_n^*)$  of an NFG satisfies

$$u_i(s^*) \geq \underline{v}_i$$

for all  $i \in N$ .

*Proof.* We have

$$\begin{aligned} u_i(s_i^*, s_{-i}^*) &= \max_{s_i \in S_i} u_i(s_i, s_{-i}^*) && \text{(by definition of PSNE)} \\ &\geq \max_{s_i \in S_i} \min_{s_{-i} \in S_{-i}} u_i(s_i, s_{-i}) = \underline{v}_i. && \blacksquare \end{aligned}$$

What happens to stability and security when games are eliminated? Recall that depending on the order in which dominant strategies are eliminated, the final value can change.

**Example.** The game matrix involved in this example is very similar to that in Section 1.3.

	L	C	R
T	1,2	2,3	0,3
M	2,2	2,1	3,2
B	2,0	0,0	1,0

The initial maxmin values for the two players are 2 for player 1 and 0 for player 2, but after we eliminate B, the values go to 2 and 2.

It is not a coincidence that the maxmin value of 2 for player 1 is unchanged.

**Theorem 1.4.** Consider an NFG  $G = \langle N, (S_i)_{i \in N}, (u_i)_{i \in N} \rangle$ . Let  $\hat{s}_j \in S_j$  be a dominated strategy, and  $\hat{G}$  be the residual game after removing  $\hat{s}_j$ . The maxmin value of player  $j$  in  $\hat{G}$  is equal to that in  $G$ .

The idea is that the eliminated strategy cannot be the unique maxmin strategy in  $G$  since it is dominated.

*Proof.* We are done if we show that there is a maxmin strategy in  $G$  in  $S_j \setminus \{\hat{s}_j\}$  (Why?).

Let  $\hat{s}_j$  be dominated by  $t_j \in S_j \setminus \{\hat{s}_j\}$ . Then, for all  $s_{-j} \in S_{-j}$ ,

$$u_j(t_j, s_{-j}) \geq u_j(\hat{s}_j, s_{-j}).$$

We have

$$\min_{s_{-j} \in S_{-j}} u_j(t_j, s_{-j}) \geq \min_{s_{-j} \in S_{-j}} u_j(\hat{s}_j, s_{-j}),$$

so

$$\max_{s_j \in S_j \setminus \{\hat{s}_j\}} \min_{s_{-j} \in S_{-j}} u_j(s_j, s_{-j}) \geq \min_{s_{-j} \in S_{-j}} u_j(\hat{s}_j, s_{-j}),$$

completing the proof – there is another strategy whose value is at least that of  $\hat{s}_j$ . ■

To summarize,

- eliminating strictly dominated strategies has no effect on the PSNEs,
- eliminating weakly dominated strategies may make the set of PSNEs smaller, but does not add new PSNEs, and
- the maxmin value is unaffected by eliminating (strictly or weakly) dominated strategies.

## 1.4. Matrix Games

**Definition 1.7.** A *matrix game* or *two player zero-sum game* is a normal form game  $\langle N, (S_i)_{i \in N}, (u_i)_{i \in N} \rangle$  with  $N = \{1, 2\}$  and  $u_1 + u_2 \equiv 0$ .



**Example.** Examples of matrix games are the following.

	L	C	R
T	3, -3	-5, 5	-2, 2
M	1, -1	4, -4	1, -1
B	6, -6	-3, 3	-5, 5

	L	R
L	-1, 1	1, -1
R	1, -1	-1, 1

A point to note now that we shall examine in more detail in some time is that the first game has PSNEs while the second does not.

A matrix game can be represented by a single *utility matrix*, by considering the utilities of only one of the players, say player 1. Player 2's utilities are then just the negative of the matrix.

Given a utility matrix  $u$ ,  $u_1 \equiv u$  and  $u_2 \equiv -u$ .

What are the PSNEs, if any, of a matrix game?

**Definition 1.8.** A *saddle point* of a matrix  $A$  is an index  $A_{ij}$  that is the largest in the  $i$ th row and the smallest in the  $j$ th column.

**Theorem 1.5.** In a matrix game with utility matrix  $u$ ,  $(s_1^*, s_2^*)$  is a PSNE if and only if it is a saddlepoint.

*Proof.* Indeed,  $(s_1^*, s_2^*)$  is a saddle point iff

$$\begin{aligned} u(s_1^*, s_2^*) &\geq u(s_1, s_2^*) \text{ for all } s_1 \in S_2 \text{ and} \\ u(s_1^*, s_2^*) &\leq u(s_1^*, s_2) \text{ for all } s_2 \in S_2, \end{aligned}$$

which is equivalent to being a PSNE since  $u_1 \equiv u$  and  $u_2 \equiv -u$ . ■

**Definition 1.9.** Given a two player game with utility matrix  $u$ , define the *maxmin* value  $\underline{v}$  by

$$\underline{v} = \max_{s_1 \in S_1} \min_{s_2 \in S_2} u(s_1, s_2)$$

and the *minmax* value  $\bar{v}$  by

$$\bar{v} = \min_{s_2 \in S_2} \max_{s_1 \in S_1} u(s_1, s_2).$$

The above can be slightly rephrased to say

$$\begin{aligned} \underline{v} &= \max_{s_1 \in S_1} \min_{s_2 \in S_2} u_1(s_1, s_2) \\ \bar{v} &= \max_{s_2 \in S_2} \min_{s_1 \in S_2} u_2(s_1, s_2). \end{aligned}$$

**Lemma 1.6.** For matrix games,  $\bar{v} \geq \underline{v}$ .

*Proof.* Let  $\bar{v}$  and  $\underline{v}$  be attained by  $(s_1, s_2)$  and  $(t_1, t_2)$  respectively. Then,

$$\bar{v} = u(s_1, s_2) \stackrel{(1)}{\geq} u(t_1, s_2) \stackrel{(2)}{\geq} u(t_1, t_2) = \underline{v},$$

where (1) is because  $(s_1, s_2)$  is a minmax strategy and (2) is because  $(t_1, t_2)$  is a maxmin strategy. ■

Going back to Section 1.4, check that in the first game,  $\bar{v} = 1 = \underline{v}$ , while in the second game,  $\bar{v} = 1$  and  $\underline{v} = -1$ .

**Theorem 1.7.** A matrix game has a PSNE iff  $\bar{v} = \underline{v}$ .

*Proof.* We wish to show that a utility matrix  $u$  has a saddle point iff its maxmin and minmax values are equal.

Suppose that  $(s_1^*, s_2^*)$  is a saddle point. Then,

$$u(s_1^*, s_2^*) \geq \max_{s_1 \in S_1} u(s_1, s_2^*) \geq \min_{s_2 \in S_2} \max_{s_1 \in S_1} u(s_1, s_2) = \bar{v}.$$

An identical argument for player 2 (keeping in mind that  $u_2 \equiv -u$ !) yields that  $\underline{v} \geq u(s_1^*, s_2^*)$ .

Combining the two,

$$\underline{v} \geq u(s_1^*, s_2^*) \geq \bar{v}.$$

Recalling Lemma 1.6, we must have that  $\bar{v} = \underline{v}$ .

Now, suppose that  $\bar{v} = \underline{v}$ . There is a strategy  $(s_1^*, s_2^*)$  that is both a maxmin and minmax strategy (Why?).

For any  $s_2 \in S_2$ ,

$$u(s_1^*, s_2) \geq \min_{t_2 \in S_2} u(s_1^*, t_2) = \max_{t_1 \in S_1} \min_{t_2 \in S_2} u(t_1, t_2) = \underline{v}.$$

Similarly, for any  $s_1 \in S_1$ ,

$$u(s_1, s_2^*) \leq \max_{t_1 \in S_1} u(t_1, s_2^*) = \min_{t_2 \in S_2} \max_{t_1 \in S_1} u(t_1, t_2) = \bar{v}.$$

The two equations above imply that  $(s_1^*, s_2^*)$  is a saddle point, so we are done. ■