

CHAPTER 1

Games in Normal Form

Game theory studies what happens when self-interested agents interact. What does it mean to say that agents are self-interested? It does not necessarily mean that they want to cause harm to each other, or even that they care only about themselves. Instead, it means that each agent has his own description of which states of the world he likes—which can include good things happening to other agents—and that he acts in an attempt to bring about these states of the world.

The dominant approach to modeling an agent's interests is *utility theory*. This theoretical approach quantifies an agent's degree of preference across a set of available alternatives, and describes how these preferences change when an agent faces uncertainty about which alternative he will receive. Specifically, a *utility function* is a mapping from states of the world to real numbers. These numbers are interpreted as measures of an agent's level of happiness in the given states. When the agent is uncertain about which state of the world he faces, his utility is defined as the expected value of his utility function with respect to the appropriate probability distribution over states.

When agents have utility functions, acting optimally in an uncertain environment is conceptually straightforward—at least as long as the outcomes and their probabilities are known to the agent and can be succinctly represented. However, things can get considerably more complicated when the world contains *two or more* utility-maximizing agents whose actions can affect each other's utilities. To study such settings, we must turn to *noncooperative* game theory.

The term “noncooperative” could be misleading, since it may suggest that the theory applies exclusively to situations in which the interests of different agents conflict. This is not the case, although it is fair to say that the theory is most interesting in such situations. By the same token, in Chapter 8 we will see that *coalitional game theory* (also known as *cooperative game theory*) does not apply only in situations in which agents' interests align with each other. The essential difference between the two branches is that in noncooperative game theory the basic modeling unit is the individual (including his beliefs, preferences, and possible actions) while in coalitional game theory the basic modeling unit is the group. We will return to that later in Chapter 8, but for now let us proceed with the individualistic approach.

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

FIGURE 1.1: The TCP user's (aka the Prisoner's) Dilemma.

1.1 EXAMPLE: THE TCP USER'S GAME

Let us begin with a simpler example to provide some intuition about the type of phenomena we would like to study. Imagine that you and another colleague are the only people using the internet. Internet traffic is governed by the TCP protocol. One feature of TCP is the *backoff* mechanism; if the rates at which you and your colleague send information packets into the network causes congestion, you each back off and reduce the rate for a while until the congestion subsides. This is how a correct implementation works. A defective one, however, will not back off when congestion occurs. You have two possible strategies: C (for using a correct implementation) and D (for using a defective one). If both you and your colleague adopt C then your average packet delay is 1 ms. If you both adopt D the delay is 3 ms, because of additional overhead at the network router. Finally, if one of you adopts D and the other adopts C then the D adopter will experience no delay at all, but the C adopter will experience a delay of 4 ms.

These consequences are shown in Figure 1.1. Your options are the two rows, and your colleague's options are the columns. In each cell, the first number represents your payoff (or, the negative of your delay) and the second number represents your colleague's payoff.¹

Given these options what should you adopt, C or D ? Does it depend on what you think your colleague will do? Furthermore, from the perspective of the network operator, what kind of behavior can he expect from the two users? Will any two users behave the same when presented with this scenario? Will the behavior change if the network operator allows the users to communicate with each other before making a decision? Under what changes to the delays would the users' decisions still be the same? How would the users behave if they have the opportunity to face this same decision with the same counterpart multiple times? Do answers to these questions depend on how rational the agents are and how they view each other's rationality?

¹A more standard name for this game is the Prisoner's Dilemma; we return to this in Section 1.3.1.

Game theory gives answers to many of these questions. It tells us that any rational user, when presented with this scenario once, will adopt D —regardless of what the other user does. It tells us that allowing the users to communicate beforehand will not change the outcome. It tells us that for perfectly rational agents, the decision will remain the same even if they play multiple times; however, if the number of times that the agents will play this is infinite, or even uncertain, we may see them adopt C .

1.2 DEFINITION OF GAMES IN NORMAL FORM

The normal form, also known as the strategic or matrix form, is the most familiar representation of strategic interactions in game theory. A game written in this way amounts to a representation of every player's utility for every state of the world, in the special case where states of the world depend only on the players' combined actions. Consideration of this special case may seem uninteresting. However, it turns out that settings in which the state of the world also depends on randomness in the environment—called Bayesian games and introduced in Chapter 7—can be reduced to (much larger) normal-form games. Indeed, there also exist normal-form reductions for other game representations, such as games that involve an element of time (extensive-form games, introduced in Chapter 4). Because most other representations of interest can be reduced to it, the normal-form representation is arguably the most fundamental in game theory.

Definition 1.2.1 (Normal-form game). *A (finite, n -person) normal-form game is a tuple (N, A, u) , where:*

- N is a finite set of n players, indexed by i ;
- $A = A_1 \times \dots \times A_n$, where A_i is a finite set of actions available to player i . Each vector $a = (a_1, \dots, a_n) \in A$ is called an action profile;
- $u = (u_1, \dots, u_n)$ where $u_i : A \mapsto \mathbb{R}$ is a real-valued utility (or payoff) function for player i .

A natural way to represent games is via an n -dimensional matrix. We already saw a two-dimensional example in Figure 1.1. In general, each row corresponds to a possible action for player 1, each column corresponds to a possible action for player 2, and each cell corresponds to one possible outcome. Each player's utility for an outcome is written in the cell corresponding to that outcome, with player 1's utility listed first.

1.3 MORE EXAMPLES OF NORMAL-FORM GAMES

1.3.1 Prisoner's Dilemma

Previously, we saw an example of a game in normal form, namely, the Prisoner's (or the TCP user's) Dilemma. However, it turns out that the precise payoff numbers play a limited role. The

4 ESSENTIALS OF GAME THEORY

	C	D
C	a, a	b, c
D	c, b	d, d

FIGURE 1.2: Any $c > a > d > b$ define an instance of Prisoner’s Dilemma.

essence of the Prisoner’s Dilemma example would not change if the -4 was replaced by -5 , or if 100 was added to each of the numbers.² In its most general form, the Prisoner’s Dilemma is any normal-form game shown in Figure 1.2, in which $c > a > d > b$.³

Incidentally, the name “Prisoner’s Dilemma” for this famous game-theoretic situation derives from the original story accompanying the numbers. The players of the game are two prisoners suspected of a crime rather than two network users. The prisoners are taken to separate interrogation rooms, and each can either “confess” to the crime or “deny” it (or, alternatively, “cooperate” or “defect”). If the payoff are all nonpositive, their absolute values can be interpreted as the length of jail term each of prisoner will get in each scenario.

1.3.2 Common-payoff Games

There are some restricted classes of normal-form games that deserve special mention. The first is the class of *common-payoff games*. These are games in which, for every action profile, all players have the same payoff.

Definition 1.3.1 (Common-payoff game). *A common-payoff game is a game in which for all action profiles $a \in A_1 \times \cdots \times A_n$ and any pair of agents i, j , it is the case that $u_i(a) = u_j(a)$.*

Common-payoff games are also called *pure coordination games* or *team games*. In such games the agents have no conflicting interests; their sole challenge is to coordinate on an action that is maximally beneficial to all.

As an example, imagine two drivers driving towards each other in a country having no traffic rules, and who must independently decide whether to drive on the left or on the right. If the drivers choose the same side (left or right) they have some high utility, and otherwise they have a low utility. The game matrix is shown in Figure 1.3.

²More generally, under standard utility theory games are insensitive to any positive affine transformation of the payoffs. This means that one can replace each payoff x by $ax + b$, for any fixed real numbers $a > 0$ and b .

³Under some definitions, there is the further requirement that $a > \frac{b+c}{2}$, which guarantees that the outcome (C, C) maximizes the sum of the agents’ utilities.

	Left	Right
Left	1, 1	0, 0
Right	0, 0	1, 1

FIGURE 1.3: Coordination game.

1.3.3 Zero-sum Games

At the other end of the spectrum from pure coordination games lie *zero-sum games*, which (bearing in mind the comment we made earlier about positive affine transformations) are more properly called *constant-sum games*. Unlike common-payoff games, constant-sum games are meaningful primarily in the context of two-player (though not necessarily two-strategy) games.

Definition 1.3.2 (Constant-sum game). *A two-player normal-form game is constant-sum if there exists a constant c such that for each strategy profile $a \in A_1 \times A_2$ it is the case that $u_1(a) + u_2(a) = c$.*

For convenience, when we talk of constant-sum games going forward we will always assume that $c = 0$, that is, that we have a zero-sum game. If common-payoff games represent situations of pure coordination, zero-sum games represent situations of pure competition; one player's gain must come at the expense of the other player. The reason zero-sum games are most meaningful for two agents is that if you allow more agents, any game can be turned into a zero-sum game by adding a dummy player whose actions do not impact the payoffs to the other agents, and whose own payoffs are chosen to make the sum of payoffs in each outcome zero.

A classical example of a zero-sum game is the game of *Matching Pennies*. In this game, each of the two players has a penny, and independently chooses to display either heads or tails. The two players then compare their pennies. If they are the same then player 1 pockets both, and otherwise player 2 pockets them. The payoff matrix is shown in Figure 1.4.

The popular children's game of Rock, Paper, Scissors, also known as Rochambeau, provides a three-strategy generalization of the matching-pennies game. The payoff matrix of this zero-sum game is shown in Figure 1.5. In this game, each of the two players can choose either rock, paper, or scissors. If both players choose the same action, there is no winner and the utilities are zero. Otherwise, each of the actions wins over one of the other actions and loses to the other remaining action.

	Heads	Tails
Heads	1, -1	-1, 1
Tails	-1, 1	1, -1

FIGURE 1.4: Matching Pennies game.

1.3.4 Battle of the Sexes

In general, games tend to include elements of both coordination and competition. Prisoner's Dilemma does, although in a rather paradoxical way. Here is another well-known game that includes both elements. In this game, called *Battle of the Sexes*, a husband and wife wish to go to the movies, and they can select among two movies: "Lethal Weapon (LW)" and "Wondrous Love (WL)." They much prefer to go together rather than to separate movies, but while the wife (player 1) prefers LW, the husband (player 2) prefers WL. The payoff matrix is shown in Figure 1.6. We will return to this game shortly.

1.4 STRATEGIES IN NORMAL-FORM GAMES

We have so far defined the actions available to each player in a game, but not yet his set of *strategies* or his available choices. Certainly one kind of strategy is to select a single action and play it. We call such a strategy a *pure strategy*, and we will use the notation we have already developed for actions to represent it. We call a choice of pure strategy for each agent a *pure-strategy profile*.

	Rock	Paper	Scissors
Rock	0, 0	-1, 1	1, -1
Paper	1, -1	0, 0	-1, 1
Scissors	-1, 1	1, -1	0, 0

FIGURE 1.5: Rock, Paper, Scissors game.

		Husband	
		LW	WL
Wife	LW	2, 1	0, 0
	WL	0, 0	1, 2

FIGURE 1.6: Battle of the Sexes game.

Players could also follow another, less obvious type of strategy: randomizing over the set of available actions according to some probability distribution. Such a strategy is called a mixed strategy. Although it may not be immediately obvious why a player should introduce randomness into his choice of action, in fact in a multiagent setting the role of mixed strategies is critical.

We define a mixed strategy for a normal-form game as follows.

Definition 1.4.1 (Mixed strategy). *Let (N, A, u) be a normal-form game, and for any set X let $\Pi(X)$ be the set of all probability distributions over X . Then the set of mixed strategies for player i is $S_i = \Pi(A_i)$.*

Definition 1.4.2 (Mixed-strategy profile). *The set of mixed-strategy profiles is simply the Cartesian product of the individual mixed-strategy sets, $S_1 \times \cdots \times S_n$.*

By $s_i(a_i)$ we denote the probability that an action a_i will be played under mixed strategy s_i . The subset of actions that are assigned positive probability by the mixed strategy s_i is called the *support* of s_i .

Definition 1.4.3 (Support). *The support of a mixed strategy s_i for a player i is the set of pure strategies $\{a_i | s_i(a_i) > 0\}$.*

Note that a pure strategy is a special case of a mixed strategy, in which the support is a single action. At the other end of the spectrum we have *fully mixed strategies*. A strategy is fully mixed if it has full support (i.e., if it assigns every action a nonzero probability).

We have not yet defined the payoffs of players given a particular strategy profile, since the payoff matrix defines those directly only for the special case of pure-strategy profiles. But the generalization to mixed strategies is straightforward, and relies on the basic notion of decision theory—*expected utility*. Intuitively, we first calculate the probability of reaching each outcome

8 ESSENTIALS OF GAME THEORY

given the strategy profile, and then we calculate the average of the payoffs of the outcomes, weighted by the probabilities of each outcome. Formally, we define the expected utility as follows (overloading notation, we use u_i for both utility and expected utility).

Definition 1.4.4 (Expected utility of a mixed strategy). *Given a normal-form game (N, A, u) , the expected utility u_i for player i of the mixed-strategy profile $s = (s_1, \dots, s_n)$ is defined as*

$$u_i(s) = \sum_{a \in A} u_i(a) \prod_{j=1}^n s_j(a_j).$$

CHAPTER 2

Analyzing Games: From Optimality To Equilibrium

Now that we have defined what games in normal form are and what strategies are available to players in them, the question is how to reason about such games. In single-agent decision theory the key notion is that of an *optimal strategy*, that is, a strategy that maximizes the agent's expected payoff for a given environment in which the agent operates. The situation in the single-agent case can be fraught with uncertainty, since the environment might be stochastic, partially observable, and spring all kinds of surprises on the agent. However, the situation is even more complex in a multiagent setting. In this case the environment includes—or, in many cases we discuss, consists entirely of—other agents, all of whom are also hoping to maximize their payoffs. Thus the notion of an optimal strategy for a given agent is not meaningful; the best strategy depends on the choices of others.

Game theorists deal with this problem by identifying certain subsets of outcomes, called *solution concepts*, that are interesting in one sense or another. In this section we describe two of the most fundamental solution concepts: Pareto optimality and Nash equilibrium.

2.1 PARETO OPTIMALITY

First, let us investigate the extent to which a notion of optimality can be meaningful in games. From the point of view of an outside observer, can some outcomes of a game be said to be better than others?

This question is complicated because we have no way of saying that one agent's interests are more important than another's. For example, it might be tempting to say that we should prefer outcomes in which the sum of agents' utilities is higher. However, as remarked in Footnote 2 earlier, we can apply any positive affine transformation to an agent's utility function and obtain another valid utility function. For example, we could multiply all of player 1's payoffs by 1,000—this could clearly change which outcome maximized the sum of agents' utilities.

Thus, our problem is to find a way of saying that some outcomes are better than others, even when we only know agents' utility functions up to a positive affine transformation. Imagine

10 ESSENTIALS OF GAME THEORY

that each agent's utility is a monetary payment that you will receive, but that each payment comes in a different currency, and you do not know anything about the exchange rates. Which outcomes should you prefer? Observe that, while it is not usually possible to identify the best outcome, there *are* situations in which you can be sure that one outcome is better than another. For example, it is better to get 10 units of currency A and 3 units of currency B than to get 9 units of currency A and 3 units of currency B, regardless of the exchange rate. We formalize this intuition in the following definition.

Definition 2.1.1 (Pareto domination). *Strategy profile s Pareto dominates strategy profile s' if for all $i \in N$, $u_i(s) \geq u_i(s')$, and there exists some $j \in N$ for which $u_j(s) > u_j(s')$.*

In other words, in a Pareto-dominated strategy profile some player can be made better off without making any other player worse off. Observe that we define Pareto domination over strategy profiles, not just action profiles.

Pareto domination gives us a partial ordering over strategy profiles. Thus, in answer to our question before, we cannot generally identify a single “best” outcome; instead, we may have a set of noncomparable optima.

Definition 2.1.2 (Pareto optimality). *Strategy profile s is Pareto optimal, or strictly Pareto efficient, if there does not exist another strategy profile $s' \in S$ that Pareto dominates s .*

We can easily draw several conclusions about Pareto optimal strategy profiles. First, every game must have at least one such optimum, and there must always exist at least one such optimum in which all players adopt pure strategies. Second, some games will have multiple optima. For example, in zero-sum games, *all* strategy profiles are strictly Pareto efficient. Finally, in common-payoff games, all Pareto optimal strategy profiles have the same payoffs.

2.2 DEFINING BEST RESPONSE AND NASH EQUILIBRIUM

Now we will look at games from an individual agent's point of view, rather than from the vantage point of an outside observer. This will lead us to the most influential solution concept in game theory, the *Nash equilibrium*.

Our first observation is that if an agent knew how the others were going to play, his strategic problem would become simple. Specifically, he would be left with the single-agent problem of choosing a utility-maximizing action. Formally, define $s_{-i} = (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n)$, a strategy profile s without agent i 's strategy. Thus we can write $s = (s_i, s_{-i})$. If the agents other than i (whom we denote $-i$) were to commit to play s_{-i} , a utility-maximizing agent i would face the problem of determining his best response.

Definition 2.2.1 (Best response). *Player i 's best response to the strategy profile s_{-i} is a mixed strategy $s_i^* \in S_i$ such that $u_i(s_i^*, s_{-i}) \geq u_i(s_i, s_{-i})$ for all strategies $s_i \in S_i$.*

The best response is not necessarily unique. Indeed, except in the extreme case in which there is a unique best response that is a pure strategy, the number of best responses is always infinite. When the support of a best response s^* includes two or more actions, the agent must be indifferent among them—otherwise, the agent would prefer to reduce the probability of playing at least one of the actions to zero. But thus *any* mixture of these actions must also be a best response, not only the particular mixture in s^* . Similarly, if there are two pure strategies that are individually best responses, any mixture of the two is necessarily also a best response.

Of course, in general an agent will not know what strategies the other players will adopt. Thus, the notion of best response is not a solution concept—it does not identify an interesting set of outcomes in this general case. However, we can leverage the idea of best response to define what is arguably the most central notion in noncooperative game theory, the Nash equilibrium.

Definition 2.2.2 (Nash equilibrium). *A strategy profile $s = (s_1, \dots, s_n)$ is a Nash equilibrium if, for all agents i , s_i is a best response to s_{-i} .*

Intuitively, a Nash equilibrium is a *stable* strategy profile: no agent would want to change his strategy if he knew what strategies the other agents were following.

We can divide Nash equilibria into two categories, strict and weak, depending on whether or not every agent's strategy constitutes a *unique* best response to the other agents' strategies.

Definition 2.2.3 (Strict Nash). *A strategy profile $s = (s_1, \dots, s_n)$ is a if, for all agents i and for all strategies $s'_i \neq s_i$, $u_i(s_i, s_{-i}) > u_i(s'_i, s_{-i})$.*

Definition 2.2.4 (Weak Nash). *A strategy profile $s = (s_1, \dots, s_n)$ is a if, for all agents i and for all strategies $s'_i \neq s_i$, $u_i(s_i, s_{-i}) \geq u_i(s'_i, s_{-i})$, and s is not a strict Nash equilibrium.*

Intuitively, weak Nash equilibria are less stable than strict Nash equilibria, because in the former case at least one player has a best response to the other players' strategies that is not his equilibrium strategy. Mixed-strategy Nash equilibria are necessarily always weak, while pure-strategy Nash equilibria can be either strict or weak, depending on the game.

2.3 FINDING NASH EQUILIBRIA

Consider again the Battle of the Sexes game. We immediately see that it has two pure-strategy Nash equilibria, depicted in Figure 2.1.

We can check that these are Nash equilibria by confirming that whenever one of the players plays the given (pure) strategy, the other player would only lose by deviating.

	LW	WL
LW	(2, 1)	0, 0
WL	0, 0	(1, 2)

FIGURE 2.1: Pure-strategy Nash equilibria in the Battle of the Sexes game.

Are these the only Nash equilibria? The answer is no; although they are indeed the only pure-strategy equilibria, there is also another mixed-strategy equilibrium. In general, it is tricky to compute a game's mixed-strategy equilibria. This is a weighty topic lying outside the scope of this booklet (but see, for example, Chapter 4 of Shoham and Leyton-Brown [2008]). However, we will show here that this computational problem is easy when we know (or can guess) the *support* of the equilibrium strategies, particularly so in this small game. Let us now guess that both players randomize, and let us assume that husband's strategy is to play LW with probability p and WL with probability $1 - p$. Then if the wife, the row player, also mixes between her two actions, she must be indifferent between them, given the husband's strategy. (Otherwise, she would be better off switching to a pure strategy according to which she only played the better of her actions.) Then we can write the following equations.

$$\begin{aligned}
 U_{\text{wife}}(\text{LW}) &= U_{\text{wife}}(\text{WL}) \\
 2 * p + 0 * (1 - p) &= 0 * p + 1 * (1 - p) \\
 p &= \frac{1}{3}
 \end{aligned}$$

We get the result that in order to make the wife indifferent between her actions, the husband must choose LW with probability $1/3$ and WL with probability $2/3$. Of course, since the husband plays a mixed strategy he must also be indifferent between his actions. By a similar calculation it can be shown that to make the husband indifferent, the wife must choose LW with probability $2/3$ and WL with probability $1/3$. Now we can confirm that we have indeed found an equilibrium: since both players play in a way that makes the other indifferent between their actions, they are both best responding to each other. Like all mixed-strategy equilibria, this is a weak Nash equilibrium. The expected payoff of both agents is $2/3$ in this equilibrium, which means that each of the pure-strategy equilibria Pareto-dominates the mixed-strategy equilibrium.

	Heads	Tails
Heads	1, -1	-1, 1
Tails	-1, 1	1, -1

FIGURE 2.2: The Matching Pennies game.

Earlier, we mentioned briefly that mixed strategies play an important role. The previous example may not make it obvious, but now consider again the Matching Pennies game, reproduced in Figure 2.2. It is not hard to see that no pure strategy could be part of an equilibrium in this game of pure competition. Therefore, likewise there can be no strict Nash equilibrium in this game. But using the aforementioned procedure, the reader can verify that again there exists a mixed-strategy equilibrium; in this case, each player chooses one of the two available actions with probability $1/2$.

We have now seen two examples in which we managed to find Nash equilibria (three equilibria for Battle of the Sexes, one equilibrium for Matching Pennies). Did we just luck out? Here there is some good news—it was not just luck.

Theorem 2.3.1 (Nash, 1951). *Every game with a finite number of players and action profiles has at least one Nash equilibrium.*

The proof of this result is somewhat involved, and we do not discuss it here except to mention that it is typically achieved by appealing to a *fixed-point theorem* from mathematics, such as those due to Kakutani and Brouwer (a detailed proof appears, for example, in Chapter 3 of Shoham and Leyton-Brown [2008]).

Nash's theorem depends critically on the availability of mixed strategies to the agents. (Many games, such as Matching Pennies, have only mixed-strategy equilibria.) However, what does it mean to say that an agent plays a mixed-strategy Nash equilibrium? Do players really sample probability distributions in their heads? Some people have argued that they really do. One well-known motivating example for mixed strategies involves soccer: specifically, a kicker and a goalie getting ready for a penalty kick. The kicker can kick to the left or the right, and the goalie can jump to the left or the right. The kicker scores if and only if he kicks to one side and the goalie jumps to the other; this is thus best modeled as Matching Pennies. Any pure strategy on the part of either player invites a winning best response on the part of the other player. It is only by kicking or jumping in either direction with equal probability, goes the argument, that the opponent cannot exploit your strategy.

14 ESSENTIALS OF GAME THEORY

Of course, this argument is not uncontroversial. In particular, it can be argued that the strategies of each player are deterministic, but each player has uncertainty regarding the other player's strategy. This is indeed a second possible interpretation of mixed strategies: the mixed strategy of player i is everyone else's assessment of how likely i is to play each pure strategy. In equilibrium, i 's mixed strategy has the further property that every action in its support is a best response to player i 's beliefs about the other agents' strategies.

Finally, there are two interpretations that are related to learning in multiagent systems. In one interpretation, the game is actually played many times repeatedly, and the probability of a pure strategy is the fraction of the time it is played in the limit (its so-called *empirical frequency*). In the other interpretation, not only is the game played repeatedly, but each time it involves two different agents selected at random from a large population. In this interpretation, each agent in the population plays a pure strategy, and the probability of a pure strategy represents the fraction of agents playing that strategy.