3

Preliminaries

The language and tools of analysis that we have developed so far seem to be ideal to depict and analyze a wide variety of decision problems that a rational individual, or an entity with well-defined objectives, could face. The essence of our framework argues that any decision problem is best understood when we set it up in terms of the three elements of which it is made up: the possible actions, the deterministic or probabilistic relationship between actions and outcomes, and the decision maker's preferences over the possible outcomes. We proceeded to argue that a decision maker will choose those actions that are in his best interest.

This framework offers many attractive features: it is precise, well structured, and generally applicable, and most importantly it lends itself to systematic and consistent analysis. It does, however, suffer from one drawback: the world of a decision problem was described as a world in which the outcomes that determine our well-being are consequences of our own actions and some randomness that is beyond our control.

Let's consider for a moment a decision problem that you may be facing now if you are using this text as part of a university course, which you are taking for a grade. It is, I believe, safe to assume that your objective is some combination of learning the material and obtaining a good grade in the course, with higher grades being preferred over lower ones. This objective determines your preferences over outcomes, which are the set of all possible combinations of how much you learned and what grade you obtained. Your set of possible actions is deciding how hard to study, which includes such elements as deciding how many lectures to attend, how carefully to read the text, how hard to work on your problem sets, and how much time to spend preparing for the exams. Hence you are now facing a well-defined decision problem.

To complete the description of your decision problem, I have yet to explain how the outcome of your success is affected by the amount of work you choose to put into your course work. Clearly as an experienced student you know that the harder you study the more you learn, and you are also more likely to succeed on the exams. There is some uncertainty over how hard a given exam will be; that may depend on many random events, such as how you feel on the day of the exam and what mood the professor was in when the exam was written.

Still something seems to be missing. Indeed you must surely know that grades are often set on a curve, so that your grade relies on your success on the exam as an absolute measure of not only how much you got right but also how much the *other students in the class* got right. In other words, if you're having a bad day on an exam, your only hope is that everyone else in your class is having a worse day!

The purpose of this example is to point out that our framework for a decision problem will be inadequate if your outcomes, and as a consequence your well-being, will depend on the choices made by other decision makers. Perhaps we can just treat the other players in this decision problem as part of the randomness of nature: maybe they'll work hard, maybe not, maybe they'll have a bad day, maybe not, and so on. This, however, would not be part of a rational framework, for it would not be sensible for you to treat your fellow players as mere random "noise." Just as you are trying to optimize your decisions, so are they. Each player is trying to guess what others are doing, and how to act accordingly. In essence, you and your peers are engaged in a *strategic environment* in which you have to think hard about what other players are doing in order to decide what is best for you—knowing that the other players are going through the same difficulties.

We therefore need to modify our decision problem framework to help us describe and analyze strategic situations in which players who interact understand their environment, how their actions affect the outcomes that they and their counterparts will face, and how these outcomes are assessed by the other players. It is useful, therefore, to start with the simplest set of situations possible, and the simplest language that will capture these strategic situations, which we refer to as **games.** We will start with **static games of complete information,** which are the most fundamental games, or environments, in which such strategic considerations can be analyzed.

A static game is similar to the very simple decision problems in which a player makes a once-and-for-all decision, after which outcomes are realized. In a static game, a *set of players* independently choose once-and-for-all actions, which in turn cause the realization of an outcome. Thus a static game can be thought of as having two distinct steps:

Step 1: Each player *simultaneously and independently* chooses an action.

By *simultaneously and independently,* we mean a condition broader and more accommodating than players all choosing their actions at the *exact same* moment. We mean that players must take their actions without observing what actions their counterparts take and without interacting with other players to coordinate their actions. For example, imagine that you have to study for your midterm exam *two days before the midterm* because of an athletic event in which you have to participate on the day before the exam. Assume further that I plan on studying *the day before the midterm,* which will be after your studying effort has ended. If I don't know how much you studied, then by choosing my action after you I have no informational advantage over you; it is *as if* we are making our choices simultaneously and independently of each other. This idea will receive considerable attention as we proceed.

Step 2: Conditional on the players' choices of actions, payoffs are distributed to each player.

That is, once the players have all made their choices, these choices will result in a particular outcome, or probabilistic distribution over outcomes. The players have preferences over the outcomes of the game given by some payoff function over outcomes. For example, if we are playing rock-paper-scissors and I draw paper while you simultaneously draw scissors, then the outcome is that you win and I lose, and the payoffs are what winning and losing mean in our context—something tangible, like \$0.10, or just the intrinsic joy of winning versus the suffering of losing.

Steps 1 and 2 settle what we mean by *static*. What do we mean by *complete information*? The loose meaning is that *all players understand the environment they are in*—that is, the game they are playing—in every way. This definition is very much related to our assumptions about rational choice in Section 1.2. Recall that when we had a single-person decision problem we argued that the player must know four things: (1) all his possible actions, A; (2) all the possible outcomes, X; (3) exactly how each action affects which outcome will materialize; and (4) what his preferences are over outcomes. How should this be adjusted to fit a game in which many such players interact?

Games of Complete Information A **game of complete information** requires that the following four components be common knowledge among all the players of the game:

1. all the possible actions of all the players,
2. all the possible outcomes,
3. how each combination of actions of all players affects which outcome will materialize, and
4. the preferences of each and every player over outcomes.

This is by no means an innocuous set of assumptions. In fact, as we will discuss later, they are quite demanding and perhaps almost impossible to justify completely for many real-world "games." However, as with rational choice theory, we use these assumptions because they provide structure and, perhaps surprisingly, describe and predict many phenomena quite well.

You may notice that a new term snuck into the description of games of complete information: *common knowledge*. This is a term that we often use loosely: "it's common knowledge that he gives hard exams" or "it's common knowledge that green vegetables are good for your health." It turns out that *what exactly common knowledge means* is by no means common knowledge. To make it clear,

Definition 3.1 An event E is **common knowledge** if (1) everyone knows E, (2) everyone knows that everyone knows E, and so on *ad infinitum*.

On the face of it, this may seem like an innocuous assumption, and indeed it may be in some cases. For example, if you and I are both walking in the rain together, then it is safe to assume that the event "it is raining" is common knowledge between us. However, if we are both sitting in class and the professor says "tomorrow there is an exam," then the event "there is an exam tomorrow" may not be common knowledge. Despite me knowing that I heard him say it, perhaps you were daydreaming at the time, implying that I *cannot be sure* that you heard the statement as well.

Thus requiring common knowledge is not as innocuous as it may seem, but without this assumption it is quite impossible to analyze games within a structured framework. This difficulty arises because we are seeking to depict a situation in which players can engage in *strategic reasoning*. That is, I want to predict how you will make your choice, given *my belief* that you understand the game. Your understanding incorporates *your belief* about my understanding, and so on. Hence common knowledge will assist us dramatically in our ability to perform this kind of reasoning.

3.1 Normal-Form Games with Pure Strategies

Now that we understand the basic ingredients of a static game of complete information, we develop a formal framework to represent it in a parsimonious and general way, which captures the strategic essence of a game. As with the simple decision problem, the players will have actions from which to choose, and the combination of their choices will result in outcomes over which the players have preferences. For now we will restrict attention to players choosing certain (deterministic) actions that together cause certain (deterministic) outcomes. That is, players will not choose actions stochastically, and there will be no "Nature" player who will randomly select outcomes given a combination of actions that the players will choose.

One of the most common ways of representing a game is described in the following definition of the normal-form game:

A **normal-form game** consists of three features:

1. A set of **players.**
2. A set of **actions** for each player.
3. A **set of payoff functions** for each player that give a payoff value to each combination of the players' chosen actions.

This definition is similar to that of the single-person decision problem that we introduced in Chapter 1, but here we incorporate the fact that many players are interacting. Each has a set of possible actions, the combination (profile) of actions that the players choose will result in an outcome, and each has a payoff from the resulting outcome.

We now introduce the commonly used concept of a **strategy.** A strategy is often defined as *a plan of action intended to accomplish a specific goal.* Imagine a candidate in a local election going to meet a group of potential voters at the home of a neighborhood supporter. Before the meeting, our aspiring politician should have a plan of action to deal with the possible questions he will face. We can think of this plan as a list of the form "if they ask me question q_1 then I will respond with answer a_1; if they ask me question q_2 then I will respond with answer a_2; . . . " and so on. A different candidate may, and often will, have a different strategy of this kind.

The concept of a strategy will escort us throughout this book, and for this reason we now give it both formal notation and a definition:

Definition 3.2 A **pure strategy** for player i is a deterministic plan of action. The set of all **pure strategies** for player i is denoted S_i. A **profile of pure strategies** $s = (s_1, s_2, \ldots, s_n)$, $s_i \in S_i$ for all $i = 1, 2, \ldots, n$, describes a particular combination of pure strategies chosen by all n players in the game.

A brief pause to consider the term "pure" is in order. As mentioned earlier, for the time being and until we reach Chapter 6, we restrict our attention to the case in which players choose deterministic actions. This is what we mean by "pure" strategies: you choose a *certain* plan of action. To illustrate this idea, imagine that you have an exam in three hours, and you must decide how long to study for the exam and how long to just relax, knowing that your classmates are facing the same choice. If, say, you measure time in intervals of 15 minutes, then there are a total of 12 time units in the three-hour window. Your set of pure strategies is then $S_i = \{1, 2, \ldots, 12\}$, where each $s_i \in S_i$ determines how many 15-minute units you will spend studying for the exam. For example, if you choose $s_i = 7$ then you will spend 1 hour and 45 minutes studying and 1 hour and 15 minutes relaxing. An alternative to choosing one of your pure strategies would be for you to choose actions *stochastically.* For example, you can take a die and say "I will roll the die and study for as many 15-minute units as the number on the die indicates." This means that you are stochastically (or randomly) choosing between *any one* of the six pure strategies of studying for 15 minutes, 30 minutes, and so on for up to 1 hour and 30 minutes.

You may wonder why anyone would choose randomly among plans of action. As an example, dwell on the following situation. You meet a friend to go to lunch. Your strategy can be to offer the names of two restaurants that you like and then have your friend decide. But what should you do if he says, "You go ahead and choose"? One option is for you to be prepared with a choice. Another is for you to take out a coin and flip it, so that it is *not you* who is choosing; instead you are randomizing between the two choices.[1] For now, we will restrict attention to pure strategies in which such stochastic play is not possible. That said, stochastic choices play a critical role in game theory. We will introduce *stochastic* or *mixed strategies* in Chapter 6 and continue to use them throughout the rest of the book.

To some extent applying the concept of a strategy or a plan of action to a static game of complete information is overkill, because the players choose actions *once and for all* and *simultaneously.* Thus the only set of relevant plans for player i is the set of his possible actions. This change of focus from actions to strategies may therefore seem redundant. That said, focusing on strategies instead of actions will set the stage for games in which there will be relevance to conditioning one's actions on events that unfold over time, as we will see in Chapter 7. Hence what now seems merely semantic will later be quite useful and important. We now formally define a normal-form game as follows.[2]

Definition 3.3 A **normal-form game** includes three components as follows:

1. A finite **set of players,** $N = \{1, 2, \ldots, n\}$.
2. A **collection of sets of pure strategies,** $\{S_1, S_2, \ldots, S_n\}$.
3. A **set of payoff functions,** $\{v_1, v_2, \ldots, v_n\}$, each assigning a payoff value to each combination of chosen strategies, that is, a set of functions $v_i : S_1 \times S_2 \times \cdots \times S_n \to \mathbb{R}$ for each $i \in N$.

This representation is very general, and it will capture many situations in which each of the players $i \in N$ must simultaneously choose a possible strategy $s_i \in S_i$. Recall again that by *simultaneous* we mean the more general construct in which

1. From my experience, once you offer to take out the coin then your friend is very likely to say, "Oh never mind, let's go to x." By taking out the coin you are effectively telling your friend, "If you have a preference for one of the places, now is your last chance to reveal it." This takes away your friend's option of "being nice" by letting you choose since it is not you who is choosing. I always find this strategy amusing since it works so well.

2. Recall that a finite set of elements will be written as $A = \{a, b, c, d\}$, where A is the set and a, b, c, and d are the elements it includes. Writing $a \in A$ means "a is an element of the set A." If we have two sets, A and B, we define the *Cartesian product* of these sets as $A \times B$. If $a \in A$ and $h \in B$ then we can write $(a, h) \in A \times B$. For more on this subject, refer to Section 19.1 of the mathematical appendix.

each player is choosing a strategy without knowing the choices of the other players. After strategies are selected, each player will realize his payoff, given by $v_i(s_1, s_2, \ldots, s_n) \in \mathbb{R}$, where $(s_1, s_2, \ldots, s_n)$ is the strategy profile that was selected by the agents. Thus from now on the normal-form game will be a triple of sets: $\langle N, \{S_i\}_{i=1}^n, \{v_i(\cdot)\}_{i=1}^n \rangle$, where N is the set of players, $\{S_i\}_{i=1}^n$ is the set of all players' strategy sets, and $\{v_i(\cdot)\}_{i=1}^n$ is the set of all players' payoff functions over the strategy profiles of all the players.[3]

3.1.1 Example: The Prisoner's Dilemma

The Prisoner's Dilemma is perhaps the best-known example in game theory, and it often serves as a parable for many different applications in economics and political science. It is a static game of complete information that represents a situation consisting of two individuals (the players) who are suspects in a serious crime, say, armed robbery. The police have evidence of only petty theft, and to nail the suspects for the armed robbery they need testimony from at least one of the suspects.

The police decide to be clever, separating the two suspects at the police station and questioning each in a different room. Each suspect is offered a deal that reduces the sentence he will get if he confesses, or "finks" (F), on his partner in crime. The alternative is for the suspect to say nothing to the investigators, or remain "mum" (M), so that they do not get the incriminating testimony from him. (As the Mafia would put it, the suspect follows the "omertà"—the code of silence.)

The payoff of each suspect is determined as follows: If both choose mum, then both get 2 years in prison because the evidence can support only the charge of petty theft. If, say, player 1 mums while player 2 finks, then player 1 gets 5 years in prison while player 2 gets only 1 year in prison for being the sole cooperator. The reverse outcome occurs if player 1 finks while player 2 mums. Finally, if both fink then both get only 4 years in prison. (There is some reduction of the 5-year sentence because each would blame the other for being the mastermind behind the robbery.)

Because it is reasonable to assume that more time in prison is worse, we use the payoff representation that equates each year in prison with a value of -1. We can now represent this game in its normal form as follows:

Players: $N = \{1, 2\}$.

Strategy sets: $S_i = \{M, F\}$ for $i \in \{1, 2\}$.

Payoffs: Let $v_i(s_1, s_2)$ be the payoff to player i if player 1 chooses s_1 and player 2 chooses s_2. We can then write payoffs as

$$v_1(M, M) = v_2(M, M) = -2$$
$$v_1(F, F) = v_2(F, F) = -4$$
$$v_1(M, F) = v_2(F, M) = -5$$
$$v_1(F, M) = v_2(M, F) = -1.$$

This completes the normal-form representation of the Prisoner's Dilemma. We will soon analyze how rational players would behave if they were faced with this game.

3. $\{S_i\}_{i=1}^n$ is another way of writing $\{S_1, S_2, \ldots, S_n\}$, and similarly for $\{v_i(\cdot)\}_{i=1}^n$.

3.1.2 Example: Cournot Duopoly

A variant of this example was first introduced by Augustin Cournot (1838). Two identical firms, players 1 and 2, produce some good. Assume that there are no fixed costs of production, and let the variable cost to each firm i of producing quantity $q_i \geq 0$ be given by the cost function, $c_i(q_i) = q_i^2$ for $i \in \{1, 2\}$. Demand is given by the function $q = 100 - p$, where $q = q_1 + q_2$. Cournot starts with the benchmark of firms that operate in a competitive environment in which each firm takes the market price, p, as given, and believes that its behavior cannot influence the market price. Under this assumption, as every economist knows, the solution will be the competitive equilibrium in which each firm produces at a point at which price equals marginal costs, so that the profits on the marginally produced unit are zero. In this particular case, each firm would produce $q_i = 25$, the price would be $p = 50$, and each firm would make 625 in profits.[4]

Cournot then argues that this competitive equilibrium is naive because rational firms should understand that the price is not given, but rather determined by their actions. For example, if firm 1 realizes its effect on the market price, and produces $q_1 = 24$ instead of $q_1 = 25$, then the price will have to increase to $p(49) = 51$ for demand to equal supply because total supply will drop from 50 to 49. The profits of firm 1 will now be $v_1 = 51 \times 24 - 24^2 = 648 > 625$. Of course, if firm 1 realizes that it has such an effect on price, it should not just set $q_1 = 24$ but instead look for the best choice it can make. However, its best choice depends on the quantity that firm 2 will produce—what will that be? Clearly firm 2 should be as sophisticated, and thus we will have to find a solution that considers both *the actions and the counteractions* of these rational and sophisticated firms.

For now, however, let's focus on the representation of the normal form of the game proposed by Cournot. The actions are choices of quantity, and the payoffs are the profits. Hence the following represents the normal form:

Players: $N = \{1, 2\}$.

Strategy sets: $S_i = [0, \infty]$ for $i \in \{1, 2\}$ and firms choose quantities $s_i \in S_i$.

Payoffs: For $i, j \in \{1, 2\}, \quad i \neq j$,

$$v_i(s_i, s_j) = \begin{cases} (100 - s_i - s_j)s_i - s_i^2 & \text{if } s_i + s_j < 100 \\ -s_i^2 & \text{if } s_i + s_j \geq 100. \end{cases}$$

Notice that the payoff function is a little tricky because it has to be well defined for *any* pair of strategies (quantities) that the players choose. We are implicitly assuming that prices cannot fall below zero, so that if the firms together produce a quantity that is greater than 100, the price will be zero (because $p = 100 - s_1 - s_2$) and each firm's payoffs are its costs.

3.1.3 Example: Voting on a New Agenda

Consider three players on a committee who have to vote on whether to remain at the status quo (whatever it is) or adopt a new policy. For example, they could be three

4. Those who have taken a course in microeconomics know that the marginal cost is the derivative of the cost function and hence is equal to $2q_i$. Equating this to the price gives us each firm's supply function, $2q_i = p$ or $q_i = \frac{p}{2}$, and adding up the two supply functions yields the market supply, $q = p$. Equating this to demand yields $p = 100 - p$, resulting in the competitive price of $p = 50$, and plugging this into the supply function yields $q_i = 25$ for $i = 1, 2$.

housemates who currently have an agreement under which they clean the house once every two weeks (the status quo) and they are considering cleaning it every week (the new policy). They could also be the members of the board of a firm who have to vote on changing the CEO's compensation, or they could be a committee of legislators who must vote on whether to adopt new regulations.

Each can vote "yes" (Y), "no" (N), or "abstain" (A). We can set the payoff from the status quo to be 0 for each player. Players 1 and 2 prefer the new policy, so let their payoff value for it be 1, while player 3 dislikes the new policy, so let his payoff from it be -1. Assume that the choice is made by majority voting as follows: if there is no majority of Y over N then the status quo prevails; otherwise the majority is decisive.

We can represent this game in normal form as follows:

Players: $N = \{1, 2, 3\}$.

Strategy sets: $S_i = \{Y, N, A\}$ for $i \in \{1, 2, 3\}$.

Payoffs: Let P denote the set of strategy profiles for which the new agenda is chosen (at least two "yes" votes), and let Q denote the set of strategy profiles for which the status quo remains (no more than one "yes" vote). Formally,

$$P = \left\{ \begin{array}{ll} (Y, Y, N), & (Y, N, Y), \\ (Y, Y, A), & (Y, A, Y), \\ (Y, A, A), & (A, Y, A), \\ (Y, Y, Y), & (N, Y, Y), \\ (A, Y, Y), & (A, A, Y) \end{array} \right\} \quad \text{and}$$

$$Q = \left\{ \begin{array}{llll} (N, N, N), & (N, N, Y), & (N, Y, N), & (Y, N, N), \\ (A, A, A), & (A, A, N), & (A, N, A), & (N, A, A), \\ (A, Y, N), & (A, N, Y), & (N, A, Y), & (Y, A, N), \\ (N, Y, A), & (Y, N, A), & (N, N, A), & (N, A, N), \\ (A, N, N) & & & \end{array} \right\}.$$

Then payoffs can be written as

$$v_i(s_1, s_2, s_3) = \begin{cases} 1 & \text{if } (s_1, s_2, s_3) \in P \\ 0 & \text{if } (s_1, s_2, s_3) \in Q \end{cases} \quad \text{for } i \in \{1, 2\},$$

$$v_3(s_1, s_2, s_3) = \begin{cases} -1 & \text{if } (s_1, s_2, s_3) \in P \\ 0 & \text{if } (s_1, s_2, s_3) \in Q. \end{cases}$$

This completes the normal-form representation of the voting game.

3.2 Matrix Representation: Two-Player Finite Game

As the voting game demonstrates, games that are easy to describe verbally can sometimes be tedious to describe formally. The value of a formal representation is clarity, because it forces us to specify who the players are, what they can do, and how their actions affect each and every player. We could take some shortcuts to make our life easier, and sometimes we will, but such convenience can come at the cost of misspecifying the game. It turns out that for two-person games in which each player has a finite number of strategies, there is a convenient representation that is easy to read.

In many cases, players may be constrained to choose one of a finite number of actions. This is the case for the Prisoner's Dilemma, rock-paper-scissors, the voting game described previously, and many more strategic situations. In fact, even when players have infinitely many actions to choose from, we may be able to provide a good approximation by restricting attention to a finite number of actions. If we think of the Cournot duopoly example, then for any product that comes in well-defined units (a car, a computer, or a shirt), we can safely assume that we are limited to integer units (an assumption that reduces the strategy set to the natural numbers—after all, fractional shirts will not sell very well). Furthermore, the demand function $p = 100 - q$ suggests that flooding the market with more than 100 units will cause the price of the product to drop to zero. This means that we have effectively restricted the strategy set to a finite number of strategies (101, to be accurate, for the quantities $0, 1, \ldots, 100$).

Being able to distinguish games with finite action sets is useful, so we define a finite game as follows:

Definition 3.4 A **finite game** is a game with a finite number of players, in which the number of strategies in S_i is finite for all players $i \in N$.

As it turns out, any two-player finite game can be represented by a matrix that will capture all the relevant information of the normal-form game. This is done as follows:

Rows Each row represents one of player 1's strategies. If there are k strategies in S_1 then the matrix will have k rows.

Columns Each column represents one of player 2's strategies. If there are m strategies in S_2 then the matrix will have m columns.

Matrix entries Each entry in this matrix contains a two-element vector (v_1, v_2), where v_i is player i's payoff when the actions of both players correspond to the row and column of that entry.

As the following examples show, this is a much simpler way of representing a two-player finite game because all the information will appear in a concise and clear way. Note, however, that neither the Cournot duopoly nor the voting example described earlier can be represented by a matrix. The Cournot duopoly is not a finite game (there are an infinite number of actions for each player), and the voting game has more than two players.[5]

It will be useful to illustrate this with two familiar examples.

3.2.1 Example: The Prisoner's Dilemma

Recall that in the Prisoner's Dilemma each player had two actions, M (mum) and F (fink). Therefore, our matrix will have two rows (for player 1) and two columns (for player 2). Using the payoffs for the prisoner's dilemma given in the example above, the matrix representation of the Prisoner's Dilemma is

5. We can represent the voting game using three 3×3 matrices: the rows of each matrix represent the actions of player 1, the columns those of player 2, and each matrix corresponds to an action of player 3. However, the convenient features of two-player matrix games are harder to use for three-player, multiple-matrix representations—not to mention the rather cumbersome structure of multiple matrices.

		Player 2	
		M	F
Player 1	M	$-2, -2$	$-5, -1$
	F	$-1, -5$	$-4, -4$

Notice that all the relevant information appears in this matrix.

3.2.2 Example: Rock-Paper-Scissors

Consider the famous child's game rock-paper-scissors. Recall that rock (R) beats scissors (S), scissors beats paper (P), and paper beats rock. Let the winner's payoff be 1 and the loser's be -1, and let the payoff for each player from a tie (i.e., they both choose the same action) be 0. This is a game with two players, $N = \{1, 2\}$, and three strategies for each player, $S_i = \{R, P, S\}$. Given the payoffs already described, we can write the matrix representation of this game as follows:

		Player 2		
		R	P	S
	R	$0, 0$	$-1, 1$	$1, -1$
Player 1	P	$1, -1$	$0, 0$	$-1, 1$
	S	$-1, 1$	$1, -1$	$0, 0$

Remark Such a matrix is sometimes referred to as a *bi-matrix.* In a traditional matrix, by definition, each entry corresponding to a row-column combination must be a single number, or element, while here each entry has a vector of two elements—the payoffs for each of the two players. Thus we formally have *two matrices,* one for each player. We will nonetheless adopt the common abuse of terminology and call this a matrix.

3.3 Solution Concepts

We have focused our attention on how to describe a game formally and fit it into a well-defined structure. This approach, of course, adds value only if we can use the structure to provide some analysis of what will or should happen in the game. Ideally we would like to be able to either advise players on how to play or try to predict how players will play. To accomplish this, we need some method to *solve* the game, and in this section we outline some criteria that will be helpful in evaluating potential methods to analyze and solve games.

As an example, consider again the Prisoner's Dilemma and imagine that you are player 1's lawyer, and that you wish to advise him about how to behave. The game may be represented as follows:

		Player 2	
		M	F
Player 1	M	$-2, -2$	$-5, -1$
	F	$-1, -5$	$-4, -4$

Being a thoughtful and rational adviser, you make the following observation for player 1: "If player 2 chooses F, then playing F gives you -4, while playing M gives you -5, so F is better." Player 1 will then bark at you, "My buddy will never squeal on me!" You, however, being a loyal adviser, must coolly reply as follows: "If you're right, and player 2 chooses M, then playing F gives you -1, while playing M gives you -2, so F is still better. In fact, it seems like F is always better!"

Indeed if I were player 2's lawyer, then the same analysis would work for him, and this is the "dilemma": each player is better off playing F regardless of his opponent's actions, but this leads the players to receive payoffs of -4 each, while if they could only agree to both choose M, then they would obtain -2 each. Left to their own devices, and to the advocacy of their lawyers, the players should not be able to resist the temptation to choose F. Even if player 1 believes that player 2 will play M, he is better off choosing F (and vice versa).

Perhaps your intuition steers you to a different conclusion. You might want to say that they are friends, having stolen together for some time now, and therefore that they care for one another. In this case one of our assumptions is incorrect: the payoffs in the matrix may not represent their true payoffs, and if taken into consideration, altruism would lead both players to choose M instead of F. For example, to capture the idea of altruism and mutual caring, we can assume that a year in prison for each player is worth -1 to himself and imposes $-\frac{1}{2}$ on the other player's payoff. (You care about your friend, but not as much as you care about yourself.) In this case, if player 1 chooses F and player 2 chooses M then player 1 gets $-3\frac{1}{2}$ $\left(-\frac{1}{2}\right.$ for each of the 5 years player 2 goes to jail, and -1 for player 1's year in jail$\left.\right)$ and player 2 gets $-5\frac{1}{2}$ $\left(-\frac{1}{2}\right.$ for the year player 1 is in jail and -5 for the 5 years he spends in jail$\left.\right)$. The matrix representing the "altruistic" Prisoner's Dilemma is given by the following:

		Player 2: M	Player 2: F
Player 1	M	$-3, -3$	$-5\frac{1}{2}, -3\frac{1}{2}$
	F	$-3\frac{1}{2}, -5\frac{1}{2}$	$-6, -6$

The altruistic game will predict cooperative behavior: regardless of what player 2 does, it is always better for player 1 to play M, and the same holds true for player 2. This shows us that our results will, as they always do, depend crucially on our assumptions.[6] This is another manifestation of the "garbage in, garbage out" caveat—we have to get the game parameters right if we want to learn something from our analysis.

Another classic game is the *Battle of the Sexes,* introduced by R. Duncan Luce and Howard Raiffa (1957) in their seminal book *Games and Decisions.* The story goes as follows. Alex and Chris are a couple, and they need to choose where to meet this evening. The catch is that the choice needs to be made while each is at work, and they have no means of communicating. (There were no cell phones or email in 1957, and even landline phones were not in abundance.) Both players prefer being together over not being together, but Alex prefers opera (O) to football (F), while Chris prefers the opposite. This implies that for each player being together at the venue of choice is

6. Another change in assumptions might be that player 2's brother is a psychopath. If player 1 finks, then player 2's brother will kill him, giving player 1 a utility of, say, $-\infty$ from choosing to fink.

better than being together at the other place, and this in turn is better than being alone. Using the payoffs of 2,1 and 0 to represent this order, the game is summarized in the following matrix:

		Chris O	Chris F
Alex	O	2, 1	0, 0
Alex	F	0, 0	1, 2

What can you recommend to each player now? Unlike the situation in the Prisoner's Dilemma, the best action for Alex depends on what Chris will do and vice versa. If we want to predict or prescribe actions for this game, we need to make assumptions about the *behavior* and the *beliefs* of the players. We therefore need a **solution concept** that will result in predictions or prescriptions.

A solution concept is a method of analyzing games with the objective of restricting the set of *all possible outcomes* to those that are *more reasonable than others*. That is, we will consider some reasonable and consistent assumptions about the behavior and beliefs of players that will divide the space of outcomes into "more likely" and "less likely." Furthermore, we would like our solution concept to apply to a large set of games so that it is widely applicable.

Consider, for example, the solution concept that prescribes that each player choose the action that is always best, regardless of what his opponents will choose. As we saw earlier in the Prisoner's Dilemma, playing F is always better than playing M. Hence this solution concept will predict that in this game both players will choose F. For the Battle of the Sexes, however, there is *no strategy* that is always best: playing F is best if your opponent plays F, and playing O is best if your opponent plays O. Hence for the Battle of the Sexes, this simple solution concept is not useful and offers no guidance.

We will use the term **equilibrium** for any one of the strategy profiles that emerges as one of the solution concept's predictions. We will often think of equilibria as the *actual predictions* of our theory. A more forgiving meaning would be that equilibria are the *likely predictions,* because our theory will often not account for all that is going on. Furthermore, in some cases we will see that more than one equilibrium prediction is possible for the same game. In fact, this will sometimes be a strength, and not a weakness, of the theory.

3.3.1 Assumptions and Setup

To set up the background for equilibrium analysis, it is useful to revisit the assumptions that we will be making throughout:

1. **Players are "rational":** A *rational* player is one who chooses his action, $s_i \in S_i$, to maximize his payoff consistent with his beliefs about what is going on in the game.
2. **Players are "intelligent":** An *intelligent* player knows everything about the game: the actions, the outcomes, and the preferences of all the players.
3. **Common knowledge:** The fact that players are rational and intelligent is common knowledge among the players of the game.

To these three assumptions, which we discussed briefly at the beginning of this chapter, we add a fourth, which constrains the set of outcomes that are reasonable:

4. **Self-enforcement:** Any prediction (or equilibrium) of a solution concept must be *self-enforcing.*

The requirement that any equilibrium must be self-enforcing is at the core of our analysis and at the heart of **noncooperative game theory.** We will assume throughout this book that the players engage in noncooperative behavior in the following sense: each player is in control of his own actions, and he will stick to an action only if he finds it to be in his best interest. That is, if a profile of strategies is to be an equilibrium, we will require each player to be happy with his own choice given how the others make their own choices. As you can probably figure out, the profile (F, F) is self-enforcing in the Prisoner's Dilemma game: each player is happy playing F. Indeed, we will see that this is a very robust outcome in terms of equilibrium analysis.

The requirement of self-enforcing equilibria is a natural one if we take the game to be the complete description of the environment. If there are outside parties that can, through the use of force or sanctions, enforce profiles of strategies, then the game we are using is likely to be an inadequate depiction of the actual environment. In this case we ought to model the third party as a player who has actions (strategies) that describe the enforcement.

3.3.2 Evaluating Solution Concepts

In developing a theory that predicts the behavior of players in games, we must evaluate our theory by how well it does as a *methodological* tool. That is, for our theory to be widely useful, it must describe a method of analysis that applies to a rich set of games, which describe the strategic situations in which we are interested. We will introduce three criteria that will help us evaluate a variety of solution concepts: existence, uniqueness, and invariance.

3.3.2.1 Existence: How Often Does It Apply? A solution concept is valuable insofar as it applies to a wide variety of games, and not just to a small and select family of games. A solution concept should apply generally and should not be developed in an ad hoc way that is specific to a certain situation or game. That is, when we apply our solution concept to different games we require it to result in the *existence* of an equilibrium solution.

For example, consider an ad hoc solution concept that offers the following prediction: "Players always choose the action that they think their opponent will choose." If this is our "theory" of behavior, then it will fail to apply to many—maybe most—strategic situations. In particular when players have different sets of actions (e.g., one chooses a software package and the other a hardware package) then this theory would be unable to predict which outcomes are more likely to emerge as equilibrium outcomes.

Any proposed theory for a solution concept that relies on very specific elements of a game will not be general and will be hard to adapt to a wide variety of strategic situations, making the proposed theory useless beyond the very special situations it was tailored to address. Thus one goal is to have a method that will be general enough to apply to many strategic situations; that is, it will prescribe a solution that will *exist* for most games we can think of.

3.3.2.2 Uniqueness: How Much Does It Restrict Behavior? Just as we require our solution concept to apply broadly, we require that it be meaningful in that it restricts the set of possible outcomes to a smaller set of reasonable outcomes. In fact one might argue that being able to pinpoint a single *unique* outcome as a prediction would be ideal. *Uniqueness* is then an important counterpart to *existence.*

For example, if the proposed solution concept says "anything can happen," then it always exists: regardless of the game we apply this concept to, "anything can happen" will always say that the solution is one of the (sometimes infinite) possible outcomes. Clearly this solution concept is useless. A good solution concept is one that balances existence (so that it works for many games) with uniqueness (so that we can add some intelligent insight into what can possibly happen).

It turns out that the nature of games makes the uniqueness requirement quite hard to meet. The reason, as we will learn to appreciate, lies in the nature of strategic interaction in a noncooperative environment. To foreshadow the reasons behind this observation, notice that a player's best action will often depend on what other players are doing. A consequence is that there will often be several combinations of strategies that will support each other in this way.

3.3.2.3 Invariance: How Sensitive Is It to Small Changes? Aside from existence and uniqueness, a third more subtle criterion is important in qualifying a solution concept as a reasonable one, namely that the solution concept be *invariant* to small changes in the game's structure. However, the term "small changes" needs to be qualified more precisely.

Adding a player to a game, for instance, may not be a small change if that player has actions that can wildly change the outcomes of the game. Thus adding or removing a player cannot innocuously be considered a small change. Similarly if we add or delete strategies from the set of actions that are available to a player, we may hinder his ability to guarantee himself some outcomes, and therefore this too should not be considered a small change to the game. We are left with only one component to fiddle with: the payoff functions of the players. It is reasonable to argue that if the payoffs of a game are modified only slightly, then this is a small change to the game that should not affect the predictions of a "robust" solution concept.

For example, consider the Prisoner's Dilemma. If instead of 5 years in prison, imposing a pain of -5 for the players, it imposed a pain of -5.01 for player 1 and -4.99 for player 2, we should be somewhat discouraged if our solution concept suddenly changed the prediction of what players will or ought to do. Thus *invariance* is a robustness property with which we require a solution concept to comply. In other words, if two games are "close," so that the action sets and players are the same yet the payoffs are only slightly different, then our solution concept should offer predictions that are not wildly different for the two games. Put formally, if for a small enough value $\varepsilon > 0$ we alter the payoffs of every outcome for every player by no more than ε, then the solution concept's prediction should not change.

3.3.3 Evaluating Outcomes

Once we subscribe to any particular solution concept, as social scientists we would like to evaluate the properties of the *solutions,* or predictions, that the solution concept will prescribe. This process will offer insights into what we expect the players of a game to achieve when they are left to their own devices. In turn, these insights can guide us toward possibly changing the environment of the game so as to improve the social outcomes of the players.

We have to be precise about the meaning of "to improve the social outcomes." For example, many people may agree that it would be socially better for the government to take \$10 away from the very rich Bill Gates and give that \$10 to an orphan in Latin America. In fact even Gates himself might have approved of this transfer, especially if the money would have saved the child's life. However, Gates may or may not have *liked* the idea, especially if such government intervention would imply that over time most of his wealth would be dissipated through such transfers.

Economists use a particular criterion for evaluating whether an outcome is *socially undesirable.* An outcome is considered to be socially undesirable if there is a different outcome that would make some people better off *without harming anyone else.* As social scientists we wish to avoid outcomes that are socially undesirable, and we therefore turn to the criterion of *Pareto optimality,* which is in tune with the idea of efficiency or "no waste." That is, we would like all the possible value deriving from a given interaction to be distributed among the players. To put this formally:[7]

Definition 3.5 A strategy profile $s \in S$ **Pareto dominates** strategy profile $s' \in S$ if $v_i(s) \geq v_i(s') \forall\, i \in N$ and $v_i(s) > v_i(s')$ for at least one $i \in N$ (in which case, we will also say that s' is **Pareto dominated** by s). A strategy profile is **Pareto optimal** if it is not Pareto dominated by any other strategy profile.

As social scientists, strategic advisers, or policy makers, we hope that players will act in accordance with the *Pareto criterion* and find ways to coordinate on Pareto-optimal outcomes, or avoid those that are Pareto dominated.[8] However, as we will see time and time again, this result will not be achievable in many games. For example, in the Prisoner's Dilemma we made the case that (F, F) should be considered as a very likely outcome. In fact, as we will argue several times, it is the *only* likely outcome. One can see, however, that it is Pareto dominated by (M, M). (Notice that (M, M) is not the only Pareto-optimal outcome. (M, F) and (F, M) are also Pareto-optimal outcomes because no other profile dominates any of them. Don't confuse Pareto optimality with the best "symmetric" outcome that leaves all players "equally" happy.)

3.4 Summary

- A normal-form game includes a finite set of players, a set of pure strategies for each player, and a payoff function for each player that assigns a payoff value to each combination of chosen strategies.
- Any two-player finite game can be represented by a matrix. Each row represents one of player 1's strategies, each column represents one of player 2's strategies, and each cell in the matrix contains the payoffs for both players.
- A solution concept that proposes predictions of how games will be played should be widely applicable, should restrict the set of possible outcomes to a small set of reasonable outcomes, and should not be too sensitive to small changes in the game.
- Outcomes should be evaluated using the Pareto criterion, yet self-enforcing behavior will dictate the set of reasonable outcomes.

---

7. The symbol ∀ denotes "for all."

8. The criterion is named after the Italian economist Vilfredo Pareto. In general economists and other advocates of rational choice theory view this criterion as noncontroversial. However, this view is not necessarily held by everyone. For example, consider two outcomes: In the first, two players get \$5 each. In the second, player 1 gets \$6 while player 2 gets \$60. The Pareto criterion clearly prefers the second outcome, but some other social criterion with equity considerations may disagree with this ranking.

3.5 Exercises

3.1 **eBay:** Hundreds of millions of people bid on eBay auctions to purchase goods from all over the world. Despite being carried out on line, in spirit these auctions are similar to those that have been conducted for centuries. Is an auction a game? Why or why not?

3.2 **Penalty Kicks:** Imagine a kicker and a goalie who confront each other in a penalty kick that will determine the outcome of a soccer game. The kicker can kick the ball left or right, while the goalie can choose to jump left or right. Because of the speed of the kick, the decisions need to be made simultaneously. If the goalie jumps in the same direction as the kick, then the goalie wins and the kicker loses. If the goalie jumps in the opposite direction of the kick, then the kicker wins and the goalie loses. Model this as a normal-form game and write down the matrix that represents the game you modeled.

3.3 **Meeting Up:** Two old friends plan to meet at a conference in San Francisco, and they agree to meet by "the tower." After arriving in town, each realizes that there are two natural choices: Sutro Tower or Coit Tower. Not having cell phones, each must choose independently which tower to go to. Each player prefers meeting up to not meeting up, and neither cares where this would happen. Model this as a normal-form game and write down the matrix form of the game.

3.4 **Hunting:** Two hunters, players 1 and 2, can each choose to hunt a stag, which provides a rather large and tasty meal, or hunt a hare—also tasty, but much less filling. Hunting stags is challenging and requires mutual cooperation. If either hunts a stag alone, then the stag will get away, while hunting the stag together guarantees that the stag will be caught. Hunting hares is an individualistic enterprise that is not done in pairs, and whoever chooses to hunt a hare will catch one. The payoff from hunting a hare is 1, while the payoff to each from hunting a stag together is 3. The payoff from an unsuccessful stag hunt is 0. Represent this game as a matrix.

3.5 **Matching Pennies:** Players 1 and 2 each put a penny on a table simultaneously. If the two pennies come up the same side (heads or tails) then player 1 gets both pennies; otherwise player 2 gets both pennies. Represent this game as a matrix.

3.6 **Price Competition:** Imagine a market with demand $p(q) = 100 - q$. There are two firms, 1 and 2, and each firm i has to simultaneously choose its price p_i. If $p_i < p_j$, then firm i gets all of the market while no one demands the good of firm j. If the prices are the same then both firms split the market demand equally. Imagine that there are no costs to produce any quantity of the good. (These are two large dairy farms, and the product is manure.) Write down the normal form of this game.

3.7 **Public Good Contribution:** Three players live in a town, and each can choose to contribute to fund a streetlamp. The value of having the streetlamp is 3 for each player, and the value of not having it is 0. The mayor asks each player to contribute either 1 or nothing. If at least two players contribute then the lamp will be erected. If one player or no players contribute then the lamp will not be erected, in which case any person who contributed will not get his money back. Write down the normal form of this game.