

of technical difficulty. I recommend Chapter Three of Kreps 1991 as the best source for further reading. Chapters Two and Thirteen in Luce and Raiffa 1957 are also a good source, although they include material that is quite dated. I learned from Savage 1972, one of the original sources. Savage 1972 presents the mathematics demonstrating the recovery of both a utility function and a subjective probability distribution from a set of preferences over actions. Savage also provides a good discussion of the motivation of utility theory and of many misinterpretations of the theory. *The New Palgrave: Utility and Probability* (Eatwell, Milgate, and Newman 1990) is a useful place to begin further reading.

Jackman 1993, Riker 1990, and Zagare 1990 are recent justifications of the assumption of rational choice in political science. The deterrence example traces originally back to Ellsberg 1960; my presentation here is different and more general. The voting example is quite well known in political science. Original sources are Downs 1957 and Riker and Ordeshook 1968. Aldrich 1993 is a recent survey on why people vote.

The work on violations of utility theory is found in psychology. Three initial sources are Quattrone and Tversky 1988, Kahneman and Tversky 1979, and Tversky and Kahneman 1981. The first uses political science examples, the second is the most technical and detailed, and the third is the shortest. Machina 1987 and Machina 1989 are accessible introductions to non-expected utility theories of rational choice.

Chapter Three

Specifying a Game

Game theory examines decisionmaking in situations where the decisions of several actors produce the final outcome. Consequently, each actor's decision depends upon the other actors' decisions. There is a very strong parallel between game theory and utility theory. The other actors' decisions in game theory correspond to the states of the world in decision theory. Actors attempt to establish the likely moves of the other actors and use their best available responses. Before we can analyze games, we need a formal structure to specify the players' interrelated decisions.

This chapter presents three different ways to specify a game. How do we go from an informal understanding of a situation to the formal statement of a game? The first section discusses deterrence in international crises and presents a game that captures some important aspects of such crises. This example introduces the basic elements of a game in the extensive form. The extensive form of a game details the players' choices, the order and consequences of their choices, how they evaluate those consequences, and what they know when they choose. Extensive forms can be reduced to the strategic form of a game through the use of strategies. A player's strategy for a game is a complete plan to play the game. In the strategic form, all players simultaneously select a strategy for the game. The interaction of their strategies determines the outcome.

Formalizing a Situation: Deterrence in the Cuban Missile Crisis

Political situations can present actors with many decisions under a variety of conditions. Before we can analyze those situations formally, we need to specify what choices the players have. Consider the Cuban missile crisis. Almost every analysis of the Crisis focuses on the three central decisions that Allison (1971, 39) identifies: "Why did the Soviet Union attempt to place offensive missiles in Cuba? Why did the United States choose to respond to the Soviet missile emplacement with a blockade of Cuba? Why did the Soviet Union decide to withdraw the missiles?"

But these questions are just the specific realizations of three general questions that arise in an explanation of any international crisis. We treat a crisis as a contest between two actors, a challenger and a defender. Why did the challenger attempt to change the status quo? Why did the other side respond to the

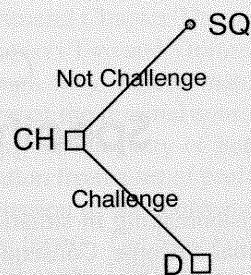


Figure 3.1 Challenger's First Decision in Deterrence Game

challenge with a threat? Finally, why did one side or the other back down short of war, or if the crisis ended in war, why did both sides choose war instead of backing down? The choices of challenging the status quo and of resisting the challenger are necessary for a crisis. If the challenger does not attempt to change the status quo or if that attempt is not resisted, there is no crisis. Allison's three questions are just the specific examples in the Cuban missile crisis of these three general questions.

Models should strive to capture the general characteristics of a situation, as opposed to the specific decisions of a particular event. After the fact, one can always construct a model that reflects what happened. Game theory models are more illuminating when they address the general characteristics of recurrent situations in politics. For example, we study models of deterrence, rather than models of the Cuban missile crisis. The Cuban missile crisis will never occur again, but unfortunately military threats and deterrence may be with us forever.

The crisis began with the Soviet decision to place offensive missiles in Cuba. When this decision was made is not clear yet, but it certainly occurred before Kennedy's decision to blockade Cuba. It preceded the decision of the United States, and the decision of the Soviet Union was known to the United States when it decided on its response. We can model the initial decision in many ways. The simplest way is a simple binary choice between challenging or not challenging the status quo. We could create several types of challenges. In the Cuban missile crisis, these other challenges could include the deployment of Soviet ground troops in Cuba or a direct move against West Berlin. But those options are specific to the Cuban missile crisis; if we are constructing a model of international crises in general, we characterize the options simply as a provocative challenge or a nonprovocative challenge. For now, we keep our model simple and just examine two options, challenging or not challenging the status quo. Figure 3.1 represents this choice as a box called a **node**, labeled CH for the challenger, with two **branches** representing the two choices coming out of the box. The Not Challenge action ends the model with no crisis, an **outcome** labeled SQ, for status quo. The Challenge option leads to a decision

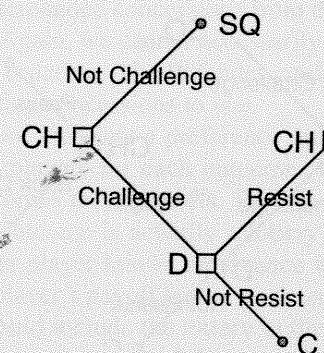


Figure 3.2 Challenger's First Decision and Defender's Decisions in Deterrence Game

by the defender, labeled D. This figure gives the model for the challenger's opening move.

The second decision in the crisis is up to the defender. It can resist the challenge, or it can make concessions to satisfy the challenger. Resistance could take one of many forms. In the Cuban missile crisis, for example, the United States could have resisted the challenge of the Soviet Union in a number of ways. It could have launched an air strike on the missile sites, invaded Cuba with ground troops, or applied diplomatic pressure through the United Nations. There were advocates of each of these options on the ExCom (short for Executive Committee) that discussed U.S. options. If the defender decides not to resist, concessions could take many possible forms. All of the challenger's demands could be granted, or negotiations could be opened after some initial concessions. But for simplicity, I allow the defender only two options, either resisting the challenge by making a threat or not resisting the challenge and granting sufficient concessions to the challenger to satisfy its demands. In the latter case, the crisis ends with an outcome I call C, for concessions. The defender makes this new move only after learning what the challenger did in its initial move. Adding the new move, we have Figure 3.2 of the first two decisions.

The decision by the United States to blockade Cuba parallels the defender's decision in the calculus of deterrence in Chapter Two. There, the challenger's response was treated as a lottery, in which the probability q gave the defender's confidence that the challenger would press its threat. But the challenger decided whether to press the threat in its own decision. The defender's confidence about the challenger's willingness to press its threat should reflect what the defender knows about the challenger's motivation to press the threat. That being the case, we need to include the challenger's final decision in the model.

The last choice in our model rests with the challenger, who must decide whether it will press the threat if the defender resists. We simplify this choice

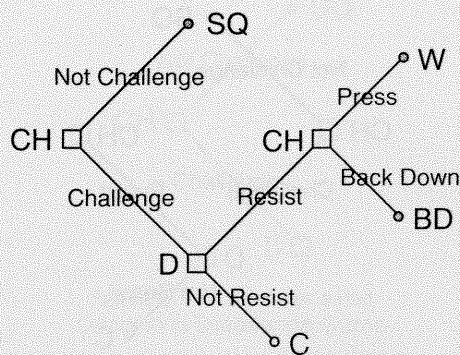


Figure 3.3 All Three Choices
in Deterrence Game

into one between pressing the threat, which will cause a war, or backing down, which will end the crisis with no change in the status quo other than its effect on the players' reputations. Again, this choice in the real world is more complex than the choice in the model. The challenger could reduce its demands or make new threats, perhaps including third parties in the crisis. Like the defender in the second move, the challenger knows the prior decisions in the game when it must choose if it must make this final move. I label the last outcomes W, for war, and BD, for the challenger's backing down.

Figure 3.3 gives the complete sequence of decisions in this simple model of deterrence. It has four possible outcomes: the status quo (SQ), concessions by the defender to the challenger (C), the challenger's backing down (BD), and war, (W). As a next step, we need the actors' preferences over these outcomes. In this case, we can make some assumptions about which outcomes each actor prefers. All challengers should prefer concessions to the status quo and prefer the status quo to backing down. Having its demands granted changes the status quo in ways the challenger prefers—otherwise, it would not make those demands. And the deterrence literature generally sees backing down from a threat as more costly than not making the threat in the first place—the challenger may have to make concessions in the face of the defender's counterthreat; moreover, the credibility of its threats in the future is reduced (or at least many scholars think so). However, we cannot say how the challenger ranks war in its preference ordering. It could prefer war to concessions—that is, the challenger's demands could be simply an excuse to start a war—or it could prefer backing down to war, a true bluff.

Similarly, we can make some assumptions about the defender's preferences. It prefers the backing down outcome to the status quo and prefers the latter to making concessions itself. If the challenger backs down, future demands are unlikely either because the credibility of such threats has been reduced and or

because the defender demanded concessions from the challenger to guarantee against future threats. Again, we cannot necessarily say where war falls in the defender's preference ordering. The defender could prefer war to any other outcome or prefer making concessions to war.

Finally, we need more than just a preference order over the outcomes; we need utilities for both players for each outcome. We call the players' utilities for particular outcomes their **payoffs**. Payoffs can be expected utilities. For instance, the war outcome is actually a lottery over all the possible outcomes of a war. Neither player knows in advance which side will win if war breaks out. Thus each player's payoff for the war outcome is an expected utility over that lottery. We could include the lottery in the game itself as a move by **Chance**, or **Nature**. Moves from one of Chance's, or Nature's, nodes are given by fixed and known probabilities of selecting each action. If the players choose the war outcome, they then play the war lottery. Instead, we generally do not include this lottery in the game because there are no other moves after it. We just give each side's expected utility for the war lottery as its payoff for the war outcome.

Because the players may face risks in their choices, we need the utilities they hold for the outcomes in order to judge what risks they will accept. The risks in the game arise from the other player's future actions. If the challenger threatens the status quo, will the defender give in or resist? Any challenger is happy to challenge if it knows that the defender will always grant its demands. All challengers prefer concessions to the status quo by assumption. Similarly, the defender probably does not know whether or not the challenger will go to war if the defender resists the initial threat. How can we represent these uncertainties in a game? In the calculus of deterrence, the probabilities in the lotteries represented these uncertainties. Typically, we assume that all players know the other players' payoffs, all players know that everyone else knows that information, and so on. Any information that all players know, that all players know that all players know, and so on, is **common knowledge**.

There is a clever way to represent uncertainty about payoffs. What if there were two types of defenders, one type that preferred war to making concessions (resolute) and one that preferred making concessions to war (irresolute)? If the challenger did not know what type of defender it was facing, it would have to judge whether it was worth the risk of encountering the resolute defender to gain the benefit of challenging the irresolute defender. Formally, we represent the challenger's uncertainty by beginning the game with a chance move that determines whether or not the defender is resolute. The defender is informed of the outcome of this chance move, but the challenger is not. As shown in Figure 3.4, Chance's opening move (labeled C) produces a tree with two main branches, the top one for resolute defenders and the bottom one for irresolute defenders. Payoffs in the bottom branch of Figure 3.4 are marked with an asterisk to remind us that they involve the irresolute defender so we

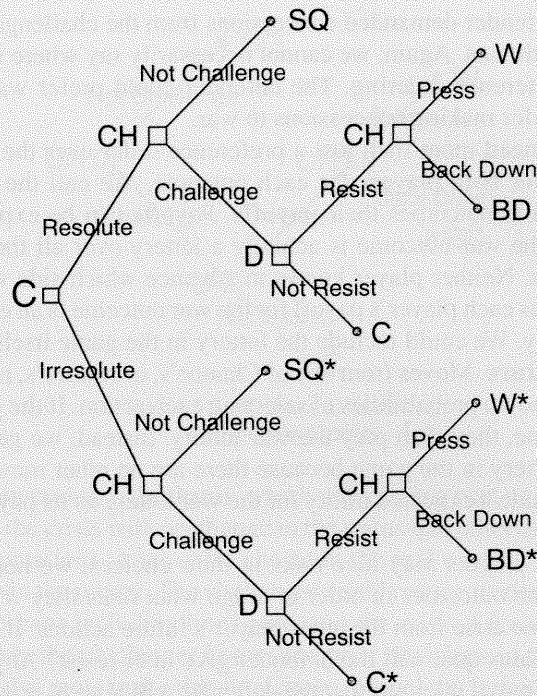


Figure 3.4 Deference Game with Two Types

cannot assume the defender's payoffs are the same as the parallel outcomes in the upper branch.

How can we represent the challenger's uncertainty then? We link together the corresponding nodes for the challenger in the two branches, as shown in Figure 3.5. Using a broken line, we link the challenger's two initial nodes and its two final nodes to show that the challenger cannot tell which of these two nodes it is at when it must choose. We call these sets of choice nodes **information sets** because they detail what information the moving player has about prior moves when he or she must choose. Here, the challenger does not know the result of the chance move that determines the defender's payoffs. We do not link the defender's two nodes into an information set because the defender does know the result of the chance move, that is, it does know its own payoffs.

Much of what actually occurs in a crisis is excluded in the model in Figure 3.5. The complex interaction of threats and offers is omitted. Instead, I have focused on the bare essentials of a crisis: a challenge, a response, and a final decision. Where there are many options in an actual crisis, I have simplified the decisions to binary choices. I have created a final choice for the challenger between war and backing down that may not exist in an actual crisis. But these simplifications establish a central point about modeling—every argument is

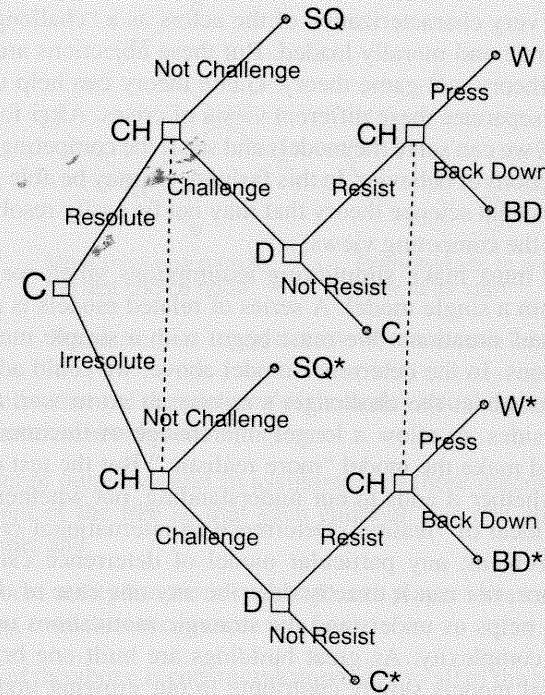


Figure 3.5 Deference Game with Uncertainty about Two Types of Defenders

a simplification of actual events. A formal model forces us to choose our assumptions and makes us aware that other reasonable alternative models exist. *The single most important decision in modeling is the design of the game.* The game states what choices we believe the actors see in the situation, what they understand about their choices, what consequences they believe can follow from their decisions, and how they evaluate those consequences. Careful thought is needed about what assumptions characterize a situation when developing a model. There are no correct models floating around out there; there are always many possible models. Formal theory alerts us to the plethora of possible models, allows us to capture the differences between them, and permits us to see if those differences have important consequences for the behavior those models predict.

For example, the deterrence model above contains several assumptions that many scholars would find objectionable. The final choice between war and backing down is artificial. The restriction to binary choices unrealistically limits the actors' range of options. The challenger is uncertain about the defender's resolve, but the defender knows the challenger's willingness to go to war. Some

would find the very characterization of the actors as a “challenger” and “defender” inaccurate and morally loaded. But these objections are questions in social science theory, not game theory. Game theory can help us understand what evidence separates these different views of crises. After formalizing the different views, we can solve the models and see if the competing views lead to different predictions of behavior. In this fashion, we may be able to resolve differences about social science theory that may not be easily resolved by verbal descriptions of the competing views.

Because we must make simplifying assumptions when we model, truth rarely lies within a single model. A series of related models is needed to understand political situations. We must begin with a simple model, and then add complications. In the deterrence model above, we could add uncertainty for the defender about the challenger’s resolve to go to war, add more options for both sides, or allow a longer interchange of threats. Any of these additions would make the model “more realistic.” But the test of an individual model is whether it adds to our understanding, not whether it appeals to some abstract ideal of “reality.” Deterrence in international crises is a very complicated problem; any particular model of deterrence cannot make us “feel” deterrence, nor can it exactly describe any one case of deterrence. Instead, a model helps us understand the strategic motivations underlying one aspect of that complexity. As great buildings are built one brick at a time, so do individual models slowly contribute to our growing understanding of politics.

Games in Extensive Form

How do we formally define a game? Extensive-form games are the basic formalization of game theory. To state a game in the extensive form, we must specify the actors, or players, and specify what decisions they face, in what order, by which player, under what conditions, and with what result.

Definition: An n -person game in extensive form is

1. a finite **game tree** composed of **nodes** and **branches** where each node of the tree is one move of the game or an endpoint of the game and the branches connect the nodes;
2. a division of the nodes over the players, Chance, and the endpoints of the game, with one and only one player, Chance, or an endpoint assigned to each node (this division is called a **partition** of the nodes);
3. a probability distribution for each chance move;

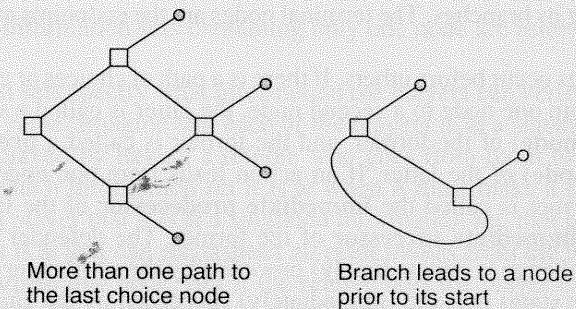


Figure 3.6 Two Things Not Allowed in Game Trees

4. a refinement of the partition of the nodes into the player sets into **information sets** for each player;
5. a set of outcomes and an assignment of those outcomes to each endpoint of the tree so that each endpoint has one and only one outcome; and
6. a set of utility functions such that each player i has a utility function, u_i , over the outcomes.

All of the above is **common knowledge** to all of the players—all players are aware of it, all players are aware that all other players know it, and so on ad infinitum.

A **game tree** consists of a series of **nodes** linked in sequence. Each node has a number of **branches**, which lead to other nodes. The nodes represent decisions and the branches the actions that could be chosen at each decision. Some nodes have no branches because they give an endpoint of the game; we call them **terminal nodes**. The other nodes are called **choice nodes** or **choice points** because a player has a choice at that point in the game (including chance moves, as choices by Chance, or Nature). In Figure 3.5, there are a total of fifteen nodes. Eight are terminal nodes, four each in the top and bottom parts of the game. There are seven choice nodes, four assigned to the challenger, two to the defender, and one to Chance (the initial move). The branches give the players’ choices from each choice node and the possible results of the chance move at the beginning of the game. I labeled each of the branches to help us see what actions they correspond to. The branches cannot “grow” back on themselves. No branch can lead to a node prior to the node it starts at. There is one and only one path through the tree to each node in the game. Figure 3.6 shows two examples of nodes and branches that are not allowed. The entire specification of what choices are available from each node and when the game ends is called a game tree because it resembles a tree. Each node spreads additional

choices along its branches. The terminal nodes are the endpoints of each branch system.

Some nodes occur before others. If there is a path of choices or chance moves that leads from one node to a second node, the latter is called a **successor** (or **succeeding node**) of the former, and the former is called a **predecessor** (or **preceding node**) of the latter. If an action leads from one node to a second node, the former is called the **immediate predecessor** of the latter, and the latter is the **immediate successor** of the former. The defender's decision to resist the challenge is (immediately) preceded by the challenger's decision to challenge the status quo and (immediately) succeeded by the challenger's decision to press the threat within each part of the game. But the challenger's decision to challenge when the defender is resolute does not precede the defender's decision to resist when it is not resolute because there is no path of choices from the former node to the latter node. The complete sequence of moves that precedes a node is called the **history** of the game up to that point. Every node has a unique history, which summarizes all prior moves because there is one and only one path of actions and chance moves to each node in the game. The history of the deterrence game up to the point where a resolute defender must choose to resist a challenge is "Resolute, Challenge". We specify both the chance move and the challenger's decision to challenge in the history.

The nonterminal nodes are divided among the players, with one and only one player assigned to each node. To allow for chance moves, we consider Chance, sometimes referred to as Nature, to be a player. Chance moves are nodes where Chance moves instead of a player. At each node assigned to Chance, there is a separate probability distribution that selects its move. The probabilities attached to a chance move are known to all the players at the start of the game. Chance moves allow us to include random elements in a game. The deterrence game included a chance move to determine whether the defender was resolute. I earlier discussed including another chance move to represent the lottery that the war outcome represents. Chance moves can be placed at any point in the game.

Each player's nodes are further divided into **information sets**. Information sets express a player's knowledge of prior moves when it must decide. When a player reaches an information set with more than one node (for an example, see the sets designated by broken lines in Figure 3.5), it knows only that it must make a decision and that it is at one of the nodes in that information set. Information sets with multiple nodes reflect the player's ignorance of prior moves in the game tree. It cannot confirm which of several prior moves were made. Identical moves must follow from each node in an information set. Otherwise, the player could distinguish the nodes in an information set by examining its available choices. Information sets form the actual choice points in the game because they summarize when a player chooses and what it knows at that point

in the game. Information sets containing only one node are referred to as **singletons**.

Information sets specify what information the players can verify about prior moves in the game. Imagine that each of the players is in a separate room and the game is played by having a referee who goes from room to room to tell each player which information set it is at when it must make a choice. If the referee tells a player that it is at a singleton information set, that player knows which node it is at and can reconstruct all prior moves from the history of the game that leads to that node. But if the referee tells the player that it is at an information set with multiple nodes, the player cannot reconstruct the history of the game as different histories lead to different nodes in that information set. Of course, the player may make inferences about the history of the game, but it cannot confirm that those inferences are correct. The term *information sets*, then, is an odd name for the element of a game that summarizes the players' ignorance about the play of the game.

The terminal nodes indicate the end of the game, and an **outcome** is assigned to each terminal node. Each player has a utility function over all the outcomes. A player's utility function evaluates the desirability of the outcomes for that player. Often, we just give the players' utility evaluation of the outcome for each terminal node, rather than describing the outcome itself. If so, we call those utilities **payoffs**.

Finally, we assume that players know the game they are playing. An aspect of the game is **common knowledge** if all players know it, all players know the other players know it, and so on. We assume that the extensive form of the game is common knowledge. The players can use their knowledge of the game to anticipate other players' moves and form expectations about the future of the game when they face decisions. The assumption that the game is common knowledge eliminates the infinite-regress problem of "They think that we think that they know that, etc." Any piece of information known to a player that is not common knowledge is that player's **private information**.

The assumption that the game is common knowledge appears quite restrictive. One might think it prevents us from analyzing situations where the players face fundamental uncertainties about their situation. But the deterrence game shows us that we can represent such uncertainties in a game. The defender's payoffs are its private information; the challenger does not know the defender's payoffs. By setting up chance moves and information sets carefully, we can model uncertainty about a game without violating the assumption that the game is common knowledge.

At this point, a simple example may help clarify the definition of an extensive form. Consider the game of Matching Pennies. Two players each have a penny. Secretly, each player chooses to place the penny either heads up or tails up. The pennies are then revealed. If the pennies match, the first player receives both; if they do not match, the second player receives both. Figure 3.7

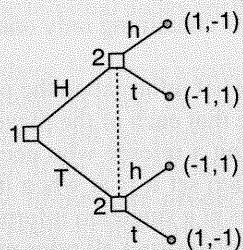


Figure 3.7 Extensive Form of Matching Pennies

gives the extensive form of Matching Pennies. This figure allows me to introduce the conventions I use in this book for extensive-form games. The boxes denote choice nodes with the moving player listed before each node. The game begins with Player 1 choosing between heads and tails, abbreviated as H and T in the diagram. Player 2 also chooses between heads and tails, denoted as h and t. For convenience, I write Player 2's move as succeeding Player 1's move. Because their moves are simultaneous, I could write Player 2's move first. There are two possible nodes for Player 2, one for Player 1's selecting heads and the other for Player 1's selecting tails. The actions are listed next to each branch. Terminal nodes are indicated by the filled-in circles and are followed by the players' payoffs for that outcome, given as (Player 1's payoff, Player 2's payoff). Although Player 2 has two choice nodes, she cannot confirm Player 1's move when she must choose; the players' moves are simultaneous. We represent the simultaneous moves by linking Player 2's two moves in one information set, represented by the broken line connecting her two nodes. This information set denotes that Player 2 does not know Player 1's move when Player 2 must choose her move. This is how simultaneous moves are represented in extensive-form games. Note that both of Player 2's nodes must have the same actions available. Otherwise, they could not be linked in one information set.

I adopt the following conventions for more complicated games. Boxes denote choice nodes, and filled-in circles denote terminal nodes. Chance's nodes are preceded by a capital C. When I use abbreviations for the players' actions, I distinguish the labels as follows: Player 1's actions are given in CAPITALS, Player 2's in lower case, Player 3's (if needed) in *CAPITAL ITALICS*, and Player 4's in *lower-case italics*. If there are more than two players, I continue writing their payoffs in sequence, separated by commas, in the parentheses. I refer to Player 1 as "he," Players 2 and 3 as "she," and Player 4 as "he" for convenience. When I talk about players generically, I use "it." I use some abstract games that have no reference to politics to illustrate points about game theory. In these games, I generally call the moves Up and Down, abbreviated

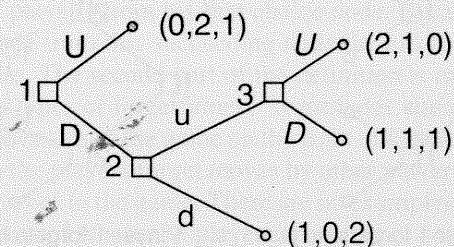


Figure 3.8 Exercise 3.2

U and D or u and d; Left and Right, abbreviated L and R or l and r; or Front and Back, abbreviated F and B or f and b.

Exercise 3.1: Draw the extensive form of the following bargaining game. The players are trying to divide \$6. Player 1 can first offer either \$4 or \$2 to Player 2. Player 2 then chooses to either accept or reject this first offer. If she rejects Player 1's offer, they flip a coin. If the coin comes up heads, Player 1 receives \$3 and Player 2 receives \$1. If the coin comes up tails, Player 1 receives \$1 dollar and Player 2 receives \$3 (i.e., if they do not agree, it costs them \$2 to flip the coin).

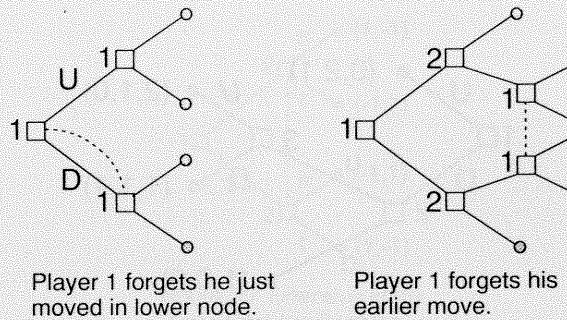
Exercise 3.2: Describe the game in Figure 3.8 in words.

Exercise 3.3: Draw the game tree for the first two moves of tic-tac-toe, one move for each player.¹ (Hint: Be sure to exploit the symmetry of the game; for example, there are only three possible initial moves, a corner box, the middle box, or one of the boxes in the middle of the sides.)

Three terms are used to describe the information the players may have when playing a game.

Definition: A game is played under **perfect information** if all information sets are singletons. A game is played under **complete information** if all the players' payoffs are common knowledge. A game is played under **incomplete information** if some player's payoff is its private information.

Perfect information means that all players' information sets contain only one node, and so all players know the history of the game whenever they make



Player 1 forgets he just moved in lower node.

Player 1 forgets his earlier move.

Figure 3.9 Game Trees That Violate Perfect Recall

a move. Among common parlor games, chess and checkers are played under perfect information, while bridge and poker are not. In bridge and poker, there is a chance move, the deal, at the start of the game that is not completely revealed to the players until the end of play. Perfect information covers situations where the players know everything that has happened before each move in a game. Complete information means that all players know one another's payoffs. The deterrence game is played under incomplete information because the challenger does not know the result of the chance move when it moves. We model that incomplete information with imperfect information. The information sets with multiple nodes reflect the uncertainties the players face. As in the deterrence game, uncertainty can be introduced into a game by an initial chance move followed by imperfect information about the result of that chance move.

We also assume that the players remember their prior moves and any information that they knew at an earlier node. This assumption is known as perfect recall.

Definition: A game is played under **perfect recall** when

1. no information set in the game contains both a node and one of its predecessors; and
2. for any two nodes x and x' in an information set for player i , if x'' is a predecessor of x in an information set for player i , then there exists a predecessor node of x' , y , in the same information set as x'' such that player i takes the same action from y to x' as it does from x'' to x . (Note: nodes x'' and y could be the same node.)

The first part of the definition requires players to distinguish current moves from future moves. The game fragment on the left of Figure 3.9 violates this condition: Player 1 forgot that he just moved if he chooses D. Player 1's in-

formation set in this game connects his first choice node with his choice node that results if he plays D from his first choice node. Because he cannot distinguish his first node from this succeeding node, he has forgotten that he chose D at his first node. The second part of the definition is more complicated and excludes several types of forgetfulness. It excludes situations where a player forgets earlier moves, as in the game on the right of Figure 3.9. It also excludes situations where the players forget moves by other players (including Chance) that they knew earlier in the game. There are no examples of games in political science that I know of where the players do not have perfect recall. Bridge is an example of a game without perfect recall if you think of it as a two-player game between the two teams of players, North-South and East-West. During the bidding, each "player" temporarily forgets what cards it has seen after a hand bids. After North bids, then the player North-South forgets what cards are in North's hand when South must bid.

Extensive forms can be inconvenient to write out for some games. If the players have an infinite array of moves, drawing an infinite number of branches could take some time. For example, the range of possible policies that could be adopted on an issue is often modeled as the points along a line, as opposed to a finite set. Those who make policy choose from an infinite number of possible policies. Games with an infinite but well-specified set of moves are often described in **time line form**. The time line form of a game parallels the extensive form. It specifies the order of the players' decision nodes, what choices are available to a player at each choice node, what information a player has when he or she must move, how the choices produce the outcomes, and the players' utility functions over the outcomes. A time line form lists the nodes in order, specifying who moves, the available choices, and what information is available to that player at that node. Terminal nodes give the outcomes and the players' utility functions.

The idea of time captured in both time line and extensive forms is decision time, rather than strict chronological time. There is no assumption that nodes are equally spaced in time or even that the nodes must occur after their predecessors. Moves can be simultaneous, as in the Matching Pennies game. These forms strive to capture the decision environment of choices. What do the players know? How do their decisions interact to produce outcomes? Although temporal sequence can affect what the players know and how their choices lead to outcomes, what matters is the decision environment, not the temporal sequence.

Games in Strategic Form

Even very simple games quickly become quite complicated when expressed in extensive form. Consequently, games are often analyzed in a reduced form,

called the strategic form (also known as the normal form). In the strategic form, we reduce each player's choice to the selection of a complete plan for playing the game. The choice of a strategy is assumed to be made before the game begins.

Definition: A strategy S_i for player i assigns one move to each of i 's information sets in a given game.

A strategy is a complete plan for playing a game for a particular player. It specifies what moves that player would make in any situation. Imagine that the players are scheduled to play the game at a set time. However, one player has a pressing commitment at the time they are to play. The referee asks that player for a comprehensive plan to play his or her position. That plan must specify a move for all possible moves that the player could make, even those that the player thinks will not occur. Then the referee can simply play for the missing player by following the player's strategy. Players can make probabilistic moves by specifying a probability distribution over the available actions at an information set. Strategies that do not include any probabilistic moves are called **pure strategies**.

We use players' strategies to reduce a game to an interaction of their pure strategies. Once we know the players' pure strategies, we can mechanically trace out their interactions in the extensive form and derive the outcome of the game. It is as if all players mailed in their strategies at the start of the game. We just open the envelopes and follow their instructions to determine the outcome of the game.

Definition: An **n-person game in strategic form** is an n -dimensional array of all the players' pure strategies with each entry of the array filled with the players' utilities for the outcome (which could be a probability distribution of outcomes) that results from that combination of strategies.

Example: To find the strategic form of Matching Pennies, we specify each player's pure strategies. Each player has one information set with two possible moves. Thus each player can have only two possible strategies, heads or tails. These two strategies create a two-by-two table, shown in Figure 3.10. To find the outcomes that result from each pair of the players' strategies, we trace out the result of the game if the players play those strategies. In the upper left-hand cell, Player 1 plays heads and Player 2 also plays heads. According to the rules of the game, Player 1 wins both pennies, and Players 1 and 2 have payoffs 1 and -1 , respectively, for this outcome. The figure gives the complete strategic form.

		Player 2	
		h	t
Player 1	H	(1,-1)	(-1,1)
	T	(-1,1)	(1,-1)

Figure 3.10 Strategic Form of Matching Pennies

Exercise 3.4:

- Draw the extensive form of Matching Pennies if Player 1 must reveal his move before Player 2 chooses hers.
- Find the strategic form of this game. (Hint: Does Player 2 still have only two possible strategies?)

More complicated extensive forms create larger strategic forms. Consider the deterrence game discussed at the beginning of the chapter and presented as a tree in Figure 3.5. Begin with the challenger's strategies. The challenger has two moves at each of its two information sets. It thus has four possible strategies. I list these strategies by its move at its first node, followed by its move at its second node: Not Challenge, Press; Not Challenge, Back Down; Challenge, Press; and Challenge, Back Down. The defender also has two moves at each of its two information sets, giving it four possible strategies. Call these strategies: Always Resist; Resist If Resolute, Do Not Resist If Irresolute; Do Not Resist If Resolute, Resist If Not Resolute; and Never Resist. Recall that the defender is resolute (or resolved—I use the two terms interchangeably) in the upper branch of the tree. Its second strategy is to play Resist in the upper branch and Not Resist in the lower branch. I have summarized this strategy as Resist If Resolute, Do Not Resist If Not Resolute, to give a better intuitive flavor for what the strategy directs the player to do.

I did not specify two elements of the extensive form of the deterrence game in my description: the probabilities in the chance move at the beginning of the game that determined whether the defender was resolute or not, and the players' utilities for the outcomes. Assume the probabilities are a $\frac{1}{2}$ chance that the defender is resolute and a $\frac{1}{2}$ chance it is not resolute. For the players' utilities, I write $u_{CH}(O)$ and $u_D(O)$ for the challenger's and defender's utilities, respectively, for outcome O .

The strategic form of the deterrence game requires a four-by-four array because each player has four possible strategies. Figure 3.11 shows this array. We trace the consequences of each pair of strategies to determine the players'

Challenger

	Challenge, Press	Challenge, Back Down	Do Not Challenge, Press	Do Not Challenge, Back Down
Always Resist	$\frac{1}{2}u_{ch}(W) + \frac{1}{2}u_{ch}(W^*),$ $\frac{1}{2}u_b(W) + \frac{1}{2}u_b(W^*)$	$\frac{1}{2}u_{ch}(BD) + \frac{1}{2}u_{ch}(C^*),$ $\frac{1}{2}u_b(BD) + \frac{1}{2}u_b(C^*)$	$\frac{1}{2}u_{ch}(SQ) + \frac{1}{2}u_{ch}(SQ^*),$ $\frac{1}{2}u_b(SQ) + \frac{1}{2}u_b(SQ^*)$	$\frac{1}{2}u_{ch}(SQ) + \frac{1}{2}u_{ch}(SQ^*),$ $\frac{1}{2}u_b(SQ) + \frac{1}{2}u_b(SQ^*)$
Resist if Resolute, Do Not Resist if Irresolute	$\frac{1}{2}u_{ch}(W) + \frac{1}{2}u_{ch}(C^*),$ $\frac{1}{2}u_b(W) + \frac{1}{2}u_b(C^*)$	$\frac{1}{2}u_{ch}(BD) + \frac{1}{2}u_{ch}(C^*),$ $\frac{1}{2}u_b(BD) + \frac{1}{2}u_b(C^*)$	$\frac{1}{2}u_{ch}(SQ) + \frac{1}{2}u_{ch}(SQ^*),$ $\frac{1}{2}u_b(SQ) + \frac{1}{2}u_b(SQ^*)$	$\frac{1}{2}u_{ch}(SQ) + \frac{1}{2}u_{ch}(SQ^*),$ $\frac{1}{2}u_b(SQ) + \frac{1}{2}u_b(SQ^*)$
Do Not Resist if Resolute, Resist if Irresolute	$\frac{1}{2}u_{ch}(C) + \frac{1}{2}u_{ch}(W^*),$ $\frac{1}{2}u_b(C) + \frac{1}{2}u_b(W^*)$	$\frac{1}{2}u_{ch}(C) + \frac{1}{2}u_{ch}(BD^*),$ $\frac{1}{2}u_b(C) + \frac{1}{2}u_b(BD^*)$	$\frac{1}{2}u_{ch}(SQ) + \frac{1}{2}u_{ch}(SQ^*),$ $\frac{1}{2}u_b(SQ) + \frac{1}{2}u_b(SQ^*)$	$\frac{1}{2}u_{ch}(SQ) + \frac{1}{2}u_{ch}(SQ^*),$ $\frac{1}{2}u_b(SQ) + \frac{1}{2}u_b(SQ^*)$
Never Resist	$\frac{1}{2}u_{ch}(C) + \frac{1}{2}u_{ch}(C^*),$ $\frac{1}{2}u_b(C) + \frac{1}{2}u_b(C^*)$	$\frac{1}{2}u_{ch}(C) + \frac{1}{2}u_{ch}(C^*),$ $\frac{1}{2}u_b(C) + \frac{1}{2}u_b(C^*)$	$\frac{1}{2}u_{ch}(SQ) + \frac{1}{2}u_{ch}(SQ^*),$ $\frac{1}{2}u_b(SQ) + \frac{1}{2}u_b(SQ^*)$	$\frac{1}{2}u_{ch}(SQ) + \frac{1}{2}u_{ch}(SQ^*),$ $\frac{1}{2}u_b(SQ) + \frac{1}{2}u_b(SQ^*)$

Defender

Figure 3.11 Strategic Form of the Deterrence Game with Uncertainty about Two Types of Defenders

expected utility for those strategies, including the effects of the chance move on the outcome in the expected utilities. For instance, what expected utilities result when the challenger plays Challenge, Press, and the defender plays Resist If Resolute, Do Not Resist If Not Resolute (the second cell down in the first column)? The result of the chance move affects the outcome, both by changing the move the defender makes in response to the challenge and by changing the outcome itself. If the defender is resolved, it resists the challenge, leading the challenger to press the challenge, producing an outcome of W. If the defender is not resolved, it does not resist the challenge, leading to the C* outcome. Each of these possibilities occurs with a probability of 1/2. The challenger's expected utility for this pair of strategies is 1/2u_{CH}(W) + 1/2u_{CH}(C*); the defender's expected utility is 1/2u_D(W) + 1/2u_D(C*).

The figure gives the strategic form for the deterrence game. It may strike you as strange that I differentiate between the Not Challenge, Press strategy and the Not Challenge, Back Down one for the challenger. After all, the game is over if the challenger does not challenge in its first move. These two strategies always produce the same outcome. These strategies are equivalent because they have the same consequences regardless of the defender's strategy.

Definition: Two strategies s_i and s'_i for player i are **equivalent** iff they lead to the same probability distribution over outcomes for all pure strategies of i's opponents.

Because equivalent strategies always have the same consequences, a set of equivalent strategies can be collapsed into just one strategy. This strategy represents the entire set of equivalent strategies. One can then reduce a strategic-form game into a reduced strategic form by collapsing all equivalent strategies into one. In the deterrence example, I separate the equivalent strategies of the challenger for the sake of completeness. I prefer seeing all the feasible strategies. As we will see later, it makes little difference whether equivalent strategies are collapsed for the purpose of analyzing a game.

A game in strategic form consists of the following:

1. a set of n players, indexed from 1 to n;
2. n sets of pure strategies S_i, one for each player; and
3. n payoff functions M_i, one for each player.

The payoff to player i from the strategies s₁, s₂, ..., s_n is written M_i(s₁; s₂; ...; s_n). All the players' payoffs from the strategies s₁, s₂, ..., s_n is written M(s₁; s₂; ...; s_n) = [M₁(s₁; s₂; ...; s_n), M₂(s₁; s₂; ...; s_n), ..., M_n(s₁; s₂; ...; s_n)]. The payoff functions incorporate both strategy choices leading to outcomes and the utility evaluation of those outcomes.

Exercise 3.5: Find the strategic forms of the games in

- a) Exercises 3.1.
b) Exercises 3.2.

The above notation can be used to define strategic-form games for games with infinite pure strategies. Such games cannot be written as an array. Instead, we define strategic-form games with infinite strategy sets by using payoff functions. These functions give each player's payoff as a function of the strategies that the players select.

Classical game theory analyzes strategic-form games, but extensive-form games are more fundamental than strategic forms. Strategic forms are reductions of extensive forms. Many commonly used strategic-form games assume a particular extensive form that does not capture the sequence of strategic interaction. The extensive form underlying all strategic-form games is such that the players select their strategies simultaneously. If the players do not select their moves simultaneously in the extensive form, they can react to earlier moves of other players. Extensive forms show us the sequence of moves clearly. Strategic forms submerge this sequence in the set of strategies. When players have sequential moves, their strategies must specify multiple responses to the other players' earlier moves. As in Exercise 3.4, changing the sequence and information of the players changes the strategies they can play and thus changes the strategic form.

Occasionally, different extensive forms reduce to the same strategic form. Consider the two extensive-form games in Figure 3.12. Both of these games have the strategic form in Figure 3.13. The strategic form here has eliminated a critical difference between the two games. In the extensive-form game on the right-hand side of Figure 3.12, Player 2 knows what Player 1 has done when she must decide on her move. In the other game, she does not know. The strategic-form array in Figure 3.13 does not show this difference. It is true that since both games produce the same consequences, you can argue that Player 2's move at the upper node in the extensive-form game on the left-hand side of Figure 3.12 is a nonmove because the payoffs are 1 for both players regardless of what Player 2 chooses if her upper node is reached. However, as we will see later, Player 2's ignorance of Player 1's move when she must choose in the game on the left can change what moves are strategically defensible.

The choice of extensive form is probably the most critical step in any game theory model; it determines the players' choices and how the players interact strategically. All too often, scholars have assumed that strategic-form games are fundamental, but one should always remember that they are derived from a more detailed account of the game.

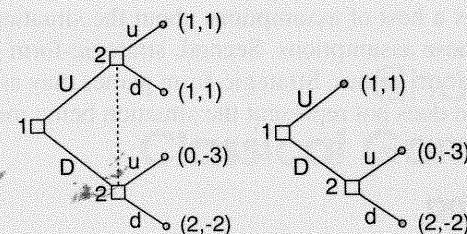


Figure 3.12 Two Different Extensive Forms That Produce the Same Strategic Form

		Player 2	
		u	d
Player 1	U	(1,1)	(1,1)
	D	(0,-3)	(2,-2)

Figure 3.13 Strategic Form of the Games in Figure 3.12

Review

This chapter has explained how games are specified. The extensive form is the basic specification of a game. It consists of nodes, divided into choice and terminal nodes. The choice nodes are divided among the players, with one player for each, and Chance is considered a player. An outcome is assigned to each terminal node. There are moves from each choice node that lead to other nodes. All of Chance's moves have assigned probabilities for each move. Each player has a utility function over the outcomes. Information sets may collect several of a player's nodes to signify that the player cannot verify which node it is at when it must move. Finally, the game is assumed to be common knowledge among the players.

Games in the strategic form have the players choose strategies at the beginning of the game, and their chosen strategies produce payoffs for the players. Strategies are complete plans to play the game; they must specify a move for every information set that the player has in the extensive form. The payoffs combine both the outcome that a set of chosen strategies produces and the utilities the players have for that outcome. Strategic forms are commonly given in an array of strategies and payoffs.

I stress two main substantive points about using game models in this chapter. First, the design of a game is the most important stage in modeling a situation.

The game captures a host of assumptions about the situation; think carefully when you make those assumptions. Second, strategic-form games are reductions of extensive-form games. Strategic-form games may assume a particular extensive form that does not represent the situation being modeled.

Further Reading

Most of the textbooks reviewed in the Further Reading section of Chapter One have useful chapters on the definition of games. Kreps's (1990) Chapter Eleven is probably the best because he takes time to discuss how a simple game attempts to capture the logic of a duopoly. However, the reader should expect to find parts of this chapter to be highly abstract. Luce and Raiffa (1957) present an accessible, but at times dated, discussion of the parts of the definition of an extensive form.

The discussion of the deterrence game draws on the literature on modeling deterrence as an extensive-form game. I recommend Wagner 1989 as a first source for modeling deterrence as an extensive-form game. Wagner (1983) discusses the misuse of games as representations of arguments in international politics.

Chapter Four Classical Game Theory

This chapter presents the classical approach to game theory and its basic results. In classical game theory, we analyze how players choose their strategies given the strategic form. It may seem strange that I begin by analyzing games in the strategic form when I have argued that the extensive form is fundamental. I do this for two reasons. First, the concept of equilibrium arose from games of pure competition, zero-sum games, in the strategic form. Second, the intuition behind the idea of an equilibrium is easily explained in the strategic form. Chapter Five begins the analysis of equilibria in extensive forms. That analysis assumes the reader is already familiar with equilibrium in classical game theory.

The central idea of this chapter is Nash equilibrium. In a Nash equilibrium, neither player has an incentive to change its strategy unilaterally. This idea comes from two-person, zero-sum games. The pure competition of such games drives the players to adopt strategies such that neither can exploit the other. The equilibria of two-person, zero-sum games have several nice properties. The notion of Nash equilibrium generalizes this concept to the non-zero-sum setting, where the players have some complementary interests. However, moving this idea to non-zero-sum games comes at a cost. It eliminates some of the nice properties of equilibrium in two-person, zero-sum games.

This chapter moves between two-person, zero-sum games and non-zero-sum games with two or more players to explain the idea of equilibrium in classical game theory. The chapter begins with the definition of a zero-sum game and a discussion of the distinction between cooperative and noncooperative game theory. It then discusses Nash equilibrium in two-person, zero-sum and two-person, non-zero-sum games. I use two-person games to illustrate Nash equilibrium in non-zero-sum games for convenience; the concept of a Nash equilibrium for n -person, non-zero-sum games is identical. The discussion of Nash equilibrium introduces three important ideas: dominated strategies, best replies, and mixed strategies. One strategy dominates another strategy for a player when the first is always at least as good as and sometimes better than the second. A player's best reply to a particular strategy of the other player is the strategy that gives the first player its highest payoff against the particular strategy of the other player. The notion of mixed strategies adds the idea that the players may have incentives to make their actions unpredictable to the other player.

The chapter returns to two-person, zero-sum games to state the Minmax Theorem. This theorem proves the general existence of equilibria in those games and the desirable properties those equilibria have. I next discuss the weaknesses of the Nash equilibrium concept. Nash equilibria of non-zero-sum (no matter how many players) games do not have all of the desirable properties of the equilibria of two-person, zero-sum games. Nash equilibria assume that players share a common conjecture about how the game will be played; the consequences of this idea and some principles for selecting among Nash equilibria are discussed. Rationalizable strategies provide an idea of what could happen in a game when the players lack a common conjecture. I compare these strategies, which are produced by the repeated elimination of dominated strategies, to Nash equilibria.

I end the chapter with a pair of examples about elections and a brief discussion of cooperative game theory. The first example is a simple model of political reform in democracies. The second example presents the spatial model of candidate competition as a two-person, zero-sum game. It presents the Median Voter Theorem, the germinal result in formal models of elections and legislatures. The Nash bargaining solution and a brief introduction to n -person game theory represent cooperative game theory. The Nash bargaining solution is my first discussion of bargaining. It is closely tied to the first noncooperative model of bargaining I present in Chapter Five, the Rubinstein bargaining model. These two models together demonstrate some of the differences between cooperative and noncooperative game theory. I provide a brief introduction to the ideas of n -person game theory. They have been important in political science. I do not cover them in detail because to do so would lead me far from my focus on noncooperative game theory. The chapter closes with a review of the central ideas of classical game theory.

I am quite brief in my coverage of the topics in this chapter. Entire textbooks have been written about each of the three branches of classical game theory: two-person, zero-sum games; two-person, non-zero-sum games; and n -person games. This chapter strives to acquaint the reader with the basic logic of the classical theory as a basis for understanding noncooperative game theory. I also expect that you will be able to solve simple strategic-form games after completing this chapter. I do not expect that you will emerge from this chapter as an expert on classical game theory. My discussion here focuses on explaining the basic concepts and on showing how they are used to solve simple games.

Defining the Terms of Classical Game Theory

Two-person, zero-sum games are games of pure competition between the players.

Definition: A game is **zero-sum** iff

$$\sum_{i=1}^n M_i = 0,$$

that is, if and only if the sum of the payoffs to all players equals zero.

In a zero-sum game, everything won by one player must be lost by another player. With only two players, this is a game of pure competition. The players have no interest in communicating or coordinating their strategies to their mutual benefit because there is no mutual benefit in these games. In a two-person, zero-sum game, the payoff to Player 2 is just the opposite of Player 1's payoff. We simplify the strategic form by writing down only Player 1's payoff, with the assumption that Player 2's payoff is the opposite of Player 1's payoff. Zero-sum games can have more than two players. In such games two or more players can have an interest in coordinated action to take advantage of another player (or several other players).

Two-person, zero-sum games are poor models of most social phenomena. Almost all interesting social phenomena create mixed motives for the parties involved. There are very few situations in which two actors have directly opposed interests—perhaps only candidate competition and the military maneuvers of warfare. Even in those situations, it may be inaccurate to describe the situation as zero-sum. It may appear at first that two-candidate races are zero-sum because one must win and the other lose. But the candidates may have to answer to constituencies other than just the voters. It may be in their interest to take divergent positions on the issues to satisfy their parties, independent of any benefits they derive from their parties in this election. They may have aspirations to higher office in future elections. Such aspirations could create motivations beyond vote maximization in this election. Non-zero-sum games capture these mixed motivations. They represent situations in which the sides have both competitive and cooperative interests. No longer is one player's gain automatically the other player's loss.

Because non-zero-sum games create the possibility of cooperative behavior between the players, we must consider the players' ability to coordinate their strategies through communication. Non-zero-sum games fall into two classes, based on the enforceability of agreements.

Definition: In a **cooperative** game, players can make binding agreements before and during the play of the game, and communication between the players is allowed. In a **noncooperative** game, binding agreements cannot be made by the players, although communication may or may not be allowed.

The difference between cooperative and noncooperative games is not the players' behavior in these games. Rather, the terms under which the players can make arrangements in their common interest separates cooperative and noncooperative games. Cooperative games allow the players to coordinate their strategies through binding agreements. In noncooperative games, the players must enforce any coordination of strategy through the game itself. An agreement can be enforced only through the interests of the agreeing parties.

One of the great disputes in game theory concerns the primacy of noncooperative games. Harsanyi and Selten (1988) contend that noncooperative games are more fundamental than cooperative games because they explicitly model the means of enforcement of agreements. Cooperative games, they argue, make it too easy for the players to make agreements; the players can bind themselves to agreements that may not be enforceable. Noncooperative games force us to consider how collaboration among players is implemented in the game and what incentives the players have to violate such agreements. I agree with this argument. Three of the key questions game theory should address is when, how, and why the players cooperate to their mutual benefit. The enforcement of agreements (or alternatively why people honor agreements at a short-run cost) is a critical question, and cooperative game theory assumes away that critical question. Consequently, this book focuses on noncooperative game theory. Nevertheless, cooperative game theory has been important in the development of game theory and its application to political science. I introduce some important ideas from cooperative game theory at the end of this chapter. I also provide further sources on cooperative game theory at the end of this chapter.

A traditional distinction between cooperative and noncooperative games is that the latter do not allow any communication between the players. Each player must determine its own strategy on its own. Over time, game theorists have found the question of whether communication is allowed less important than whether agreements are binding. If agreements cannot be enforced by an outside agency, then communication is irrelevant in some situations. Noncooperative game theorists argue that communication should be explicitly modeled in the game. Early discussions of communication in games allowed the players to discuss anything openly before playing. But interest has grown in analyzing the strategic effects of communication. To do so, the messages have to be incorporated as moves within the game itself. Allowing the players to say anything was simply too ill defined to evaluate the strategic role of communication. It is common now to model communication explicitly in noncooperative games. Still, such highly formalized communication is not as rich as language. In Chapter Eight, I provide several examples of the formal analysis of communication within a game.

		Player 2	
		s_1	s_2
Player 1	S_1	-1	2
	S_2	4	3

Figure 4.1 A Two-Person, Zero-Sum Game

Domination, Best Replies, and Equilibrium

How should the players play a game? Sometimes, one strategy is always better for a player than any of its other strategies. Consider the two-person, zero-sum game in Figure 4.1. Regardless of which strategy Player 2 plays, Player 1 is always better off playing S_2 . If Player 2 plays s_1 , Player 1's payoff is -1 if he plays S_1 and 4 if he plays S_2 . If Player 2 plays s_2 , Player 1's payoff is 2 if he plays S_1 and 3 if he plays S_2 .

Because S_2 is always better for Player 1 than S_1 , we say that S_2 dominates S_1 .

Definition: A strategy S_1 **strongly (or strictly) dominates** another strategy S_2 for Player 1 iff

$$M_1(S_1; s_j) > M_1(S_2; s_j) \text{ for all } s_j.$$

A strategy S_1 **weakly dominates** another strategy S_2 for Player 1 iff

$$M_1(S_1; s_j) \geq M_1(S_2; s_j) \text{ for all } s_j$$

and

$$M_1(S_1; s_j) > M_1(S_2; s_j) \text{ for some } s_j.$$

We call S_2 in this situation a **weakly dominated strategy**. If S_1 strongly dominates all other strategies S_i , we call S_1 a **dominant strategy**. If both players have dominant strategies, the resulting equilibrium is called a **dominant strategy equilibrium**.

There are similar definitions of domination for Player 2. Strongly dominated strategies are always inferior. A player always does better by playing the strategy that dominates it. A player should never play a strongly dominated strategy. A weakly dominated strategy is never better and sometimes worse than the strategy that dominates it.

		Player 2	
		s_1	s_2
Player 1	S_1	(5,5)	(-10,10)
	S_2	(10,-10)	(-5,-5)

Figure 4.2 Prisoner's Dilemma

If a player has a dominant strategy, it should always play that strategy because it does worse by playing any other strategy. If both players have dominant strategies, then we have a dominant strategy equilibrium, a very strong prediction for the outcome of a game. Both players have a strong incentive to play their dominant strategies, regardless of what the other player does.

Example: A classic two-person, non-zero-sum game is Prisoner's Dilemma. Prisoner's Dilemma is typically presented with the following cute story. Two criminals are arrested, but the district attorney does not have enough evidence to convict either of them for serious charges unless one or both confess to the crime. The district attorney separates the two and makes the following offer to each: "If you confess and your partner does not, I will grant you immunity, and you will walk out free. However, if your partner squeals, and you don't, I'm going to throw the book at you. If neither of you confesses, then I'll have to settle for misdemeanor charges, which will get you each a brief prison term. If you both confess, I'll get you both on felony charges, but I'll argue for shorter sentences than if you do not confess and your partner does. Think about it and tell me what you want to do."

Figure 4.2 gives the payoff matrix for Prisoner's Dilemma (let S_1 be no confession and S_2 be a confession). A complete equilibrium must specify both players' strategies. I adopt the notation of listing the players' strategies sequentially, with semicolons separating their strategies. When the players must specify multiple moves in their strategies, I use commas to separate the different moves in a player's strategy.

Prisoner's Dilemma has a dominant strategy equilibrium, $(S_2; s_2)$: both crooks should squeal. But both players are better off at $(S_1; s_1)$ than at $(S_2; s_2)$. A deal to play $(S_1; s_1)$ would be in both players' interest. But their self-interest leads both players to cheat on such a deal. This is the dilemma in the Prisoner's Dilemma. Communication alone cannot solve the dilemma. Even if both criminals are dragged off to the interrogation rooms screaming, "I'll

a)		Player 2	
		s_1	s_2
Player 1	S_1	2	3
	S_2	0	1

b)		Player 2	
		s_1	s_2
Player 1	S_1	(3,4)	(5,-1)
	S_2	(3,-2)	(-5,-4)

c)		Player 2		
		s_1	s_2	s_3
Player 1	S_1	(1,1)	(-2,0)	(4,-1)
	S_2	(0,3)	(3,1)	(5,4)
	S_3	(1,5)	(4,2)	(5,2)

Figure 4.3 Exercise 4.1

never squeal," each is always better off confessing, regardless of what the other does. The players need a binding agreement or some way to enforce a deal to solve the Prisoner's Dilemma. In Chapter Nine, I discuss the play of Prisoner's Dilemma when the game is played repeatedly. Iterated Prisoner's Dilemma is a very different game from Prisoner's Dilemma.

Exercise 4.1: Find any dominated strategies in the games in Figure 4.3. Are there any dominant strategy equilibria in these games?

Unfortunately, most games do not have a dominant strategy equilibrium. The game in Figure 4.1 does not; although S_2 is a dominant strategy for Player 1, Player 2 does not have a dominant strategy. If Player 1 plays S_1 , Player 2 prefers playing s_1 , giving her a payoff of 1. (Recall that Player 2's payoff is the opposite of Player 1's payoff in a two-person, zero-sum game.) If Player 1 plays S_2 , Player 2 prefers playing s_2 , giving her a payoff of -3 . The strategy that is best for Player 2 depends upon what strategy Player 1 plays. The strategy that is best for a player against a particular strategy of the other player is the first player's best reply to that particular strategy.

Definition: A strategy S_i is a **best reply** to strategy s_j for Player 1 iff for all other strategies S available to Player 1,

$$M(S_i; s_j) \geq M(S; s_j).$$

A strategy S_i is a **strict best reply** to strategy s_j for Player 1 iff for all other strategies S available to Player 1,

		Player 2	
		s_1	s_2
Player 1	S_1	(1,4)	(0,2)
	S_2	(-1,0)	(5,1)

Figure 4.4 A Two-Person, Non-Zero-Sum Game without Dominant Strategies

$$M(S_i; s_j) > M(S; s_j).$$

There are parallel definitions for Player 2. A dominant strategy is a strict best reply against all of the other player's strategies.

Return to Figure 4.1. Player 1 has a dominant strategy, S_2 . Player 2's best reply to Player 1's dominant strategy is s_2 . The pair of strategies $(S_2; s_2)$ is stable. Neither player wants to change his or her strategy given that they know what strategy the other player is playing. Assuming that both players know the game and reason about how the other will play the game, Player 2 should anticipate that Player 1 will play S_2 . She should then play s_2 .

In some games, neither player has a dominant strategy. Neither player has a dominant strategy in the two-person, non-zero-sum game in Figure 4.4. S_1 and S_2 are Player 1's best replies to s_1 and s_2 , respectively. Similarly, s_1 and s_2 are Player 2's best replies to S_1 and S_2 , respectively. Each player's best strategy depends upon the other player's chosen strategy.

We define a Nash equilibrium as a situation where each player's strategy is a best reply to the other player's strategy. When both players are playing best replies against each other's strategies, they have no incentive to change their strategies. The game in Figure 4.4 has two Nash equilibria in pure strategies, $(S_1; s_1)$ and $(S_2; s_2)$. S_1 is Player 1's best reply to s_1 , and s_1 is Player 2's best reply to S_1 . Similarly, S_2 is Player 1's best reply to s_2 , and s_2 is Player 2's best reply to S_2 .

Definition: A pair of strategies S_i and s_j forms a **Nash equilibrium** iff the strategies are best replies to each other. Alternatively, a pair of strategies forms a Nash equilibrium iff

$$M_1(S_i; s_j) \geq M_1(S; s_j) \text{ for all } S \neq S_i$$

and

$$M_2(S_i; s_j) \geq M_2(S_i; s) \text{ for all } s \neq s_j.$$

A Nash equilibrium is stable because neither player has an incentive to deviate unilaterally from its equilibrium strategy. Either player reduces its payoff if it defects from its equilibrium strategy. However, this observation does not imply that an equilibrium is the best outcome for either player. Nor are equilibria "fair" in any common meaning of "fairness." Instead, Nash equilibrium is a minimal condition for a solution to a game if the players can correctly anticipate each other's strategies. Nash equilibria entail stable, mutual anticipations of other players' strategies. If such anticipations exist, neither player has an incentive to change its strategy unilaterally. To do so would reduce its payoff.

Nash equilibria in pure strategies can be found easily from the strategic form. Player 1 can choose the row of the strategic form, and Player 2, the column. Sometimes Players 1 and 2 are called Row and Column for this reason. Within a given column, that is, for a fixed strategy for Player 2, Player 1's best reply is the strategy that produces the largest payoff for him in that column. Within a given row, Player 2's best reply is the strategy that produces the largest payoff for her in that row. Nash equilibria can be found, then, by searching for outcomes where the first payoff is the greatest in its column and the second payoff is the greatest in its row. The strategies that produce such outcomes form a Nash equilibrium in pure strategies.

Exercise 4.2: Find the Nash equilibria in pure strategies for the games in Figure 4.5.

In a two-person, zero-sum game, the players' payoff at an equilibrium must be the greatest in its column and the least in its row. Player 1 has control over the row chosen and can benefit by defection if some outcome in the same column has a greater payoff. Conversely, Player 2 wishes to minimize Player 1's payoff (recall that $M_2 = -M_1$ in a zero-sum game) and has control over the column chosen within a given row. Equilibria of zero-sum games are sometimes called minmax points because they are simultaneously row minima and column maxima.

Exercise 4.3: Find the equilibria of the two-person, zero-sum games in Figure 4.6.

Mixed Strategies

Some games do not have Nash equilibria in pure strategies. Figure 4.7 gives the strategic form of the Matching Pennies game from Chapter Three (using the convention for payoffs in zero-sum games). Matching Pennies does not have a Nash equilibrium in pure strategies. Player 1's best replies to h and t are H and T , respectively; Player 2's best replies to H and T are t and h , respectively.

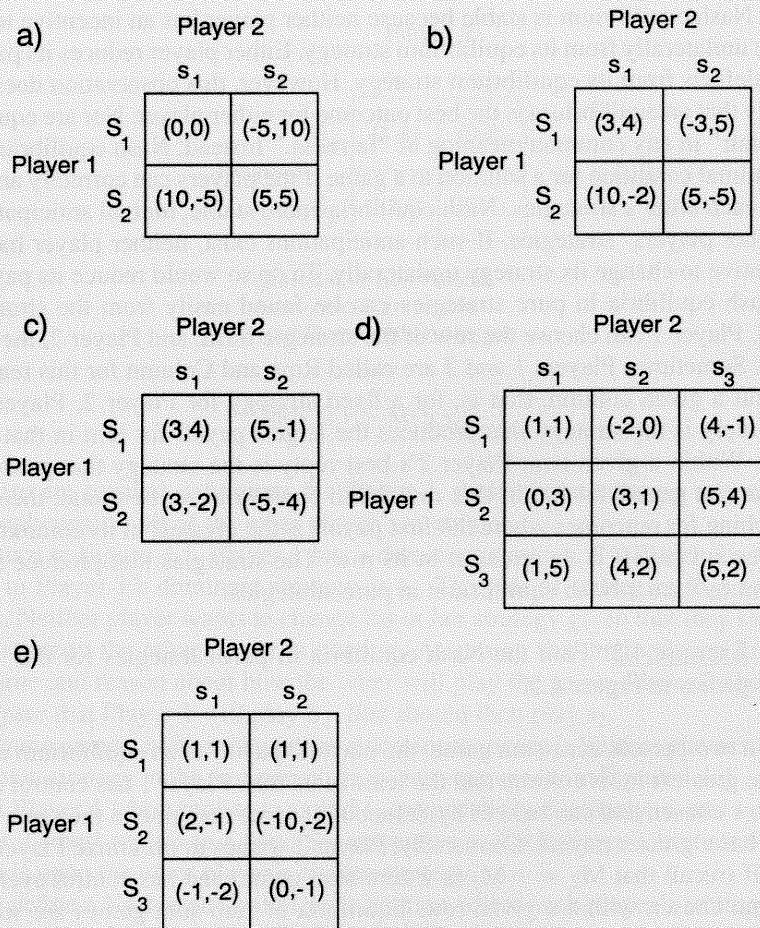


Figure 4.5 Exercise 4.2

Put simply, Player 1 wants to match Player 2's strategy, and Player 2 wants to avoid matching Player 1's strategy. How should they play this game?

Each player wants to make the other player uncertain about which strategy he or she will play. If a player can predict whether the other player will play heads or tails, the first player can take advantage of that prediction to win the game. We use probabilities to represent degrees of uncertainty about what strategy the other player will play.

Definition: A **mixed strategy** for a player is a probability distribution on the set of its pure strategies. Mixed strategies are denoted by (p_1S_1, \dots, p_nS_n) , where p_i is the probability of playing strategy S_i where the player has n pure strategies and

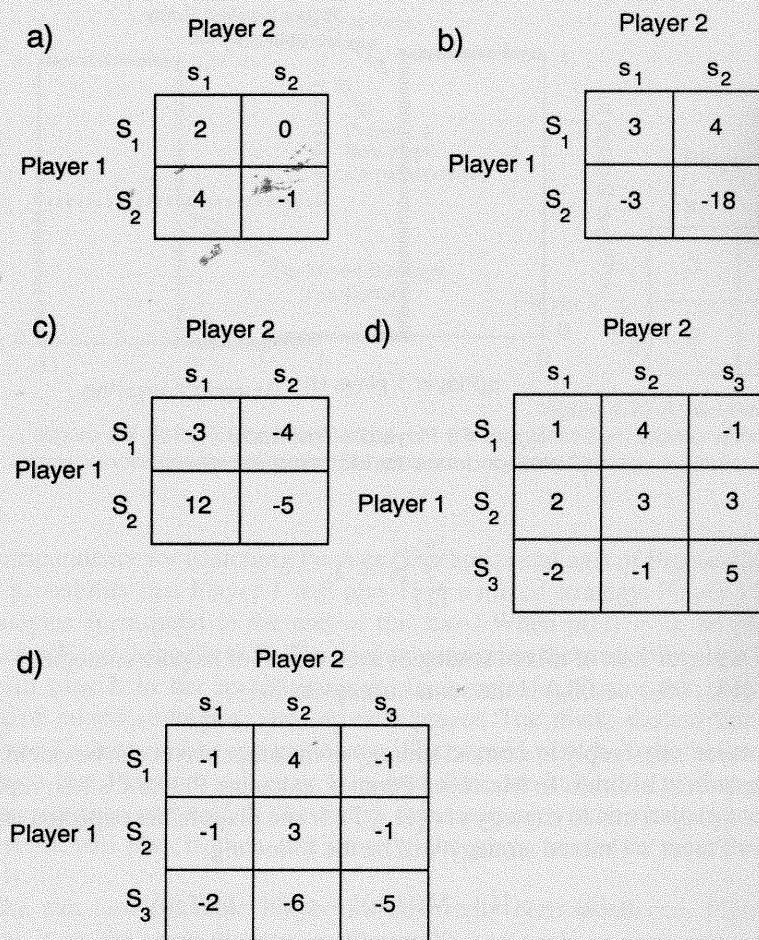


Figure 4.6 Exercise 4.3

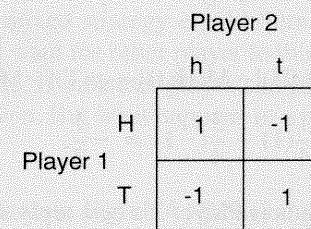


Figure 4.7 Matching Pennies