

CONTENTS

Preface *ix*

Introduction: How Should People Behave in Society? *xv*

PART I

1. Ten Tales of Strategy 3
2. Games Solvable by Backward Reasoning 32
3. Prisoners' Dilemmas and How to Resolve Them 64
4. A Beautiful Equilibrium 102

Epilogue to Part I 137

PART II

5. Choice and Chance 141
 6. Strategic Moves 173
 7. Making Strategies Credible 201
- Epilogue to Part II: A Nobel History 229

PART III

8. Interpreting and Manipulating Information 235
9. Cooperation and Coordination 270
10. Auctions, Bidding, and Contests 301
11. Bargaining 335
12. Voting 359
13. Incentives 386
14. Case Studies 409

Further Reading 443

Workouts 447

Notes 457

Index 473

Preface

WE DIDN'T SET out to write a new book. The plan was simply to revise our 1991 book, *Thinking Strategically*. But it didn't quite turn out that way.

One model for writing a revision comes from Borges's character Pierre Menard, who decides to rewrite Cervantes's *Don Quixote*. After great effort, Menard's revision ends up being word-for-word identical to the original. However, 300 years of history and literature have passed since *Quixote*, including *Quixote* itself. Although Menard's words are the same, his meaning is now entirely different.

Alas, our original text wasn't *Don Quixote*, and so the revision did require changing a few words. In fact, most of the book is entirely new. There are new applications, new developments in the theory, and a new perspective. So much is new that we decided a new title was called for as well. Although the words are new, our meaning remains the same. We intend to change the way you see the world, to help you think strategically by introducing the concepts and logic of game theory.

Like Menard, we have a new perspective. When we wrote *Thinking Strategically*, we were younger, and the zeitgeist was one of self-centered competition. We have since come to the full real-

ization of the important part that cooperation plays in strategic situations, and how good strategy must appropriately mix competition and cooperation.*

We started the original preface with: "Strategic thinking is the art of outdoing an adversary, knowing that the adversary is trying to do the same to you." To this we now add: It is also the art of finding ways to cooperate, even when others are motivated by self-interest, not benevolence. It is the art of convincing others, and even yourself, to do what you say. It is the art of interpreting and revealing information. It is the art of putting yourself in others' shoes so as to predict and influence what they will do.

We like to think that *The Art of Strategy* includes this older, wiser perspective. But there is also continuity. Even though we offer more real-life stories, our purpose remains to help you develop your own ways of thinking about the strategic situations you will face; this is not an airport book offering "seven steps for sure strategic success." The situations you face will be so diverse that you will succeed better by knowing some general principles and adapting them to the strategic games you are playing.

Businessmen and corporations must develop good competitive strategies to survive, and find cooperative opportunities to grow the pie. Politicians have to devise campaign strategies to get elected and legislative strategies to implement their visions. Football coaches plan strategies for players to execute on the field. Parents trying to elicit good behavior from children must become amateur strategists—the children are pros.

Good strategic thinking in such numerous diverse contexts remains an art. But its foundations consist of some simple basic principles—an emerging science of strategy, namely game theory. Our premise is that readers from a variety of backgrounds and occupations can become better strategists if they know these principles.

Some people question how we can apply logic and science to a

* Pursuing this line led one of us to write a book on this idea; see Adam Brandenburger and Barry J. Nalebuff, *Co-opetition* (New York: Doubleday, 1996).

world where people act irrationally. It turns out that there is often method to the madness. Indeed, some of the most exciting new insights have come from recent advances in behavioral game theory, which incorporates human psychology and biases into the mix and thus adds a social element to the theory. As a result, game theory now does a much better job dealing with people as they are, rather than as we might like them to be. We incorporate these insights into our discussions.

While game theory is a relatively young science—just over seventy years old—it has already provided many useful insights for practical strategists. But, like all sciences, it has become shrouded in jargon and mathematics. These are essential research tools, but they prevent all but the specialists from understanding the basic ideas. Our main motive for writing *Thinking Strategically* was the belief that game theory is too interesting and important to leave to the academic journals. The insights prove useful in many endeavors—business, politics, sports, and everyday social interactions. Thus we translated the important insights back into English and replaced theoretical arguments with illustrative examples and case studies.

We are delighted to find our view becoming mainstream. Game theory courses are some of the most popular electives at Princeton and Yale, and most other schools where they are offered. Game theory permeates strategy courses in MBA programs, and a Google search for game theory produces more than 6 million pages. You'll find game theory in newspaper stories, op-eds, and public policy debates.

Of course, much of the credit for these developments belongs to others: to the Economics Nobel Prize Committee, which has awarded two prizes in game theory—in 1994, to John Harsanyi, John Nash, and Reinhard Selten and in 2005, to Robert Aumann and Thomas Schelling;* to Sylvia Nasar, who wrote *A Beautiful*

* There have also been three Nobel Prizes awarded for work in mechanism design and information economics, both of which are closely related to game theory: in 1996, to William Vickrey and James Mirrlees; in 2001, to George Akerlof, Michael Spence, and Joseph Stiglitz; and in 2007, to Leonid Hurwicz, Eric Maskin, and Roger Myerson.

Mind, the best-selling biography of John Nash; to those who made the award-winning movie of the same name; and to all those who have written books popularizing the subject. We might even share a bit of the credit. Since publication, *Thinking Strategically* has sold 250,000 copies. It has been translated into numerous languages, and the Japanese and Hebrew translations have been best sellers.

We owe a special debt to Tom Schelling. His writings on nuclear strategies, particularly *The Strategy of Conflict* and *Arms and Influence*, are justly famous. In fact, Schelling pioneered a lot of game theory in the process of applying it to nuclear conflict. Michael Porter's *Competitive Strategy*, drawing on the lessons of game theory for business strategy, is equally important and influential. An annotated guide to the works of Schelling, Porter, and many others is provided in our Further Reading section.

In this book we do not confine the ideas to any particular context. Instead, we offer a wide range of illustrations for each basic principle. Thus readers from different backgrounds will all find something familiar here. They will also see how the same principles bear on strategies in less familiar circumstances; we hope this will give them a new perspective on many events in news as well as history. We also draw on the shared experience of our readers, with illustrations from, for example, literature, movies, and sports. Serious scientists may think this trivializes strategy, but we believe that familiar examples are an effective vehicle for conveying the important ideas.

The idea of writing a book at a more popular level than that of a course text came from Hal Varian, now at Google and the University of California, Berkeley. He also gave us many useful ideas and comments on earlier drafts. Drake McFeely at W. W. Norton was an excellent if exacting editor for *Thinking Strategically*. He made extraordinary efforts to fashion our academic writing into a lively text. Many readers of *Thinking Strategically* gave us encouragement, advice, and criticism, all of which have helped us when writing *The Art of Strategy*. At the grave risk of omitting some, we must mention ones to whom we owe special thanks. Our coauthors on

other related and unrelated book projects, Ian Ayres, Adam Brandenburger, Robert Pindyck, David Reiley, and Susan Skeath, generously gave us much useful input. Others whose influence continues in this new book include David Austen-Smith, Alan Blinder, Peter Grant, Seth Masters, Benjamin Polak, Carl Shapiro, Terry Vaughn, and Robert Willig. Jack Repcheck at W. W. Norton has been a constantly supportive, understanding, and perceptive editor for *The Art of Strategy*. Our manuscript editors, Janet Byrne and Catherine Pichotta, were generous to our faults. Every time you don't find a mistake, you should thank them.

We owe special thanks to Andrew St. George, book critic for the *Financial Times*. In choosing *Thinking Strategically* as a book he enjoyed reading most in the year 1991, he said: "it is a trip to the gym for the reasoning facilities" (*FT Weekend*, December 7/8, 1991). This gave us the idea of labeling some intriguing challenges we pose to the readers in this edition "Trips to the Gym." Finally, John Morgan, of the University of California, Berkeley, gave us a powerful incentive with the threat, "If you don't write a revision, I will write a competing book." And after we saved him the trouble, he helped us out with many ideas and suggestions.

AVINASH DIXIT

BARRY J. NALEBUFF

October 2007

How Should People Behave in Society?

OUR ANSWER DOES not deal with ethics or etiquette. Nor do we aim to compete with philosophers, preachers, or parents. Our theme, although less lofty, affects the lives of all of us just as much as do morality and manners. This book is about strategic behavior. All of us are strategists, whether we like it or not. It is better to be a good strategist than a bad one, and this book aims to help you improve your skills at discovering and using effective strategies.

Work, even social life, is a constant stream of decisions. What career to follow, how to manage a business, whom to marry, how to bring up children, and whether to run for president are just some examples of such fateful choices. The common element in these situations is that you do not act in a vacuum. Instead, you are surrounded by active decision makers whose choices interact with yours. This interaction has an important effect on your thinking and actions.

To illustrate the point, think of the difference between the decisions of a lumberjack and those of a general. When the lumberjack decides how to chop wood, he does not expect the wood to fight back: his environment is neutral. But when the general tries to cut down the enemy's army, he must anticipate and overcome

resistance to his plans. Like the general, you must recognize that your business rivals, prospective spouse, and even your children are strategic. Their aims often conflict with yours, but they may well coincide. Your own choice must allow for the conflict and utilize the cooperation. This book aims to help you think strategically, and then translate these thoughts into action.

The branch of social science that studies strategic decision making is called *game theory*. The games in this theory range from chess to child rearing, from tennis to takeovers, and from advertising to arms control. As the Hungarian humorist George Mikes expressed it, "Many continentals think life is a game; the English think cricket is a game." We think both are right.

Playing these games requires many different kinds of skills. Basic skills, such as shooting ability in basketball, knowledge of precedents in law, or a blank face in poker, are one kind; strategic thinking is another. Strategic thinking starts with your basic skills and considers how best to use them. Knowing the law, you must decide the strategy for defending your client. Knowing *how well* your football team can pass or run and how well the other team can defend against each choice, your decision as the coach is *whether* to pass or to run. Sometimes, as in the case of nuclear brinkmanship, strategic thinking also means knowing when not to play.

The science of game theory is far from being complete, and much of strategic thinking remains an art. Our ultimate aim is to make you better practitioners of that art, but this requires a good foundation in some elementary concepts and methods of the science. Therefore we mix the two approaches. Chapter 1 begins with examples of the art, showing how strategic issues arise in a variety of decisions. We point out some effective strategies, some less effective ones, and even some downright bad ones that were used by players in these real-life games. These examples begin to suggest a conceptual framework. In the next set of chapters, 2-4, we build this basis for the science using examples, each of which is designed to bring out one principle. Then we turn our attention to more specific concepts and strategies for dealing with particular situations—

how to mix moves when any systematic action can be exploited by the other player, how to change a game to your advantage, and how to manipulate information in strategic interaction. Finally, we take up several broad classes of strategic situations—bargaining, auctions, voting, and the design of incentives—where you can see these principles and strategies in action.

Science and art, by their very nature, differ in that science can be learned in a systematic and logical way, whereas expertise in art has to be acquired by example, experience, and practice. Our exposition of the basic science generates some principles and broad rules—for example, the idea and method of backward reasoning that is developed in chapter 2, and the concept of Nash equilibrium in chapter 4. On the other hand, the art of strategy, in all the varied situations where you may need it, requires you to do more work. Each such situation will have some unique features that you must take into account and combine with the general principles of the science. The only way to improve your skill at this art is the inductive way—by seeing how it has been done before in similar situations. That is exactly how we aim to improve your strategic IQ: by giving numerous examples, including a case study, in each chapter and in a collection of case studies in the final chapter.

The examples range from the familiar, trivial, or amusing—usually drawn from literature, sports, or movies—to the frightening—nuclear confrontation. The former are merely a nice and palatable vehicle for the game-theoretic ideas. As to the latter, at one point in time many readers would have thought the subject of nuclear war too horrible to permit rational analysis. But with the cold war now long over, we hope that the game-theoretic aspects of the arms race and the Cuban missile crisis can be examined for their strategic logic with some detachment from their emotional content.

The case studies are similar to ones you might come across in a business-school class. Each case sets out a particular set of circumstances and invites you to apply the principles discussed in that chapter to find the right strategy for that situation. Some cases are open-ended; but that is also a feature of life. At times there is no

clearly correct solution, only imperfect ways to cope with the problem. A serious effort to think each case through before reading our discussion is a better way to understand the ideas than any amount of reading of the text alone. For more practice, the final chapter is a collection of cases, in roughly increasing order of difficulty.

By the end of the book, we hope that you will emerge a more effective manager, negotiator, athlete, politician, or parent. We warn you that some of the strategies that are good for achieving these goals may not earn you the love of your rivals. If you want to be fair, tell them about our book.

Ten Tales of Strategy

Part I

THE above ten tales are tales of strategy taken with care from the old and often preliminary theories on how best to play the game and will have found similar problems in modern life and will have needed the creative solution often found through trial and error. But before, some of the ancient may be surprising. The purpose is not the primary purpose of the strategy. The purpose is to show that the ancient are powerful, that they are not so much a set of rules, and that the ancient strategy is not so much a set of rules as a set of principles.

The above diagrams are drawn from the ancient of strategy and are not supposed to be the ancient strategy. Think of them as a set of rules and not as a set of rules. They are not so much a set of rules as a set of principles.

1. THE A NUMBER

Below is a diagram of the ancient of strategy and is not supposed to be the ancient strategy. Think of them as a set of rules and not as a set of rules. They are not so much a set of rules as a set of principles.

Below is a diagram of the ancient of strategy and is not supposed to be the ancient strategy. Think of them as a set of rules and not as a set of rules. They are not so much a set of rules as a set of principles.

Ten Tales of Strategy

WE BEGIN WITH ten tales of strategy from different aspects of life and offer preliminary thoughts on how best to play. Many of you will have faced similar problems in everyday life and will have reached the correct solution after some thought or trial and error. For others, some of the answers may be surprising, but surprise is not the primary purpose of the examples. Our aim is to show that such situations are pervasive, that they amount to a coherent set of questions, and that methodical thinking about them is likely to be fruitful.

In later chapters, we develop these systems of thought into prescriptions for effective strategy. Think of these tales as a taste of dessert before the main course. They are designed to whet your appetite, not fill you up.

#1. PICK A NUMBER

Believe it or not, we are going to ask you to play a game against us. We've picked a number between 1 and 100, and your goal is to guess the number. If you guess correctly on the first try, we'll pay you \$100.

Actually, we aren't really going to pay you \$100. It would be

rather costly for us, especially since we want to give you some help along the way. But, as you play the game, we'd like you to imagine that we really are going to give you money, and we'll play the same way.

The chance of getting the number right on the first shot is quite low, only one in a hundred. So to improve your chances, we'll give you five guesses, and after each wrong guess, we'll also tell you if you are too high or too low. Of course, there's a bigger reward for getting the right answer quickly. If you guess correctly on the second try, you'll get \$80. On the third try, the payment is down to \$60, then \$40 for the fourth guess, and \$20 if you get the number on the fifth try. If it takes more than five guesses, the game is over and you get nothing.

Are you ready to play? We are, too. If you are wondering how to play a game with a book, it is a bit of a challenge, but not impossible. You can go to the artofstrategy.info web site and play the game interactively. Or, we can anticipate how you might be playing the game and respond accordingly.

Is your first guess 50? That's the most common first guess and, alas for you, it's too high.

Might your second guess be 25? Following 50, that is what most folks do. Sorry, that's too low. The next step for many is 37. We're afraid that 37 is too low. What about 42? Too low, again.

Let's pause, take a step back, and analyze the situation. This is your fifth guess coming up and your last chance to take our money. You know the number is above 42 and less than 50. There are seven options: 43, 44, 45, 46, 47, 48, and 49. Which of those seven do you think it will be?

So far, you have been guessing in a way that divides the interval into two equal parts and picking the midpoint. This is the ideal strategy in a game where the number has been chosen at random.* You are getting the most information possible from each guess and therefore will converge to the number as quickly as possible.

* The technical term for this search strategy is *minimizing the entropy*.

Indeed, Microsoft CEO Steven Ballmer is said to have used this game as a job interview question. For Ballmer the correct answer was 50, 25, 37, 42, . . . He was interested in seeing if the candidate approached the search problem in the most logical and efficient manner.

Our answer is a bit different. In Ballmer's problem, the number was picked at random, and so the engineer's strategy of "divide the set in two and conquer" was just right. Getting the most information from each guess minimizes the expected number of guesses and therefore leads you to get the most money. In our case, however, the number was *not* picked at random. Remember that we said that we were playing this game as if we actually had to pay you the money. Well, no one is reimbursing us for money that, hypothetically, we would have to pay you. And as much as we like you for having bought our book, we like ourselves even more. We'd rather keep the money than give it to you. So we deliberately picked a number that would be hard for you to find. Think about it for a moment—would it have made any sense for us to have picked 50 as the number? That would have cost us a fortune.

The key lesson of game theory is to put yourself in the other player's shoes. We put ourselves in your shoes and anticipated that you would guess 50, then 25, then 37, then 42. Understanding how you would play the game allowed us to greatly decrease the chance that you would guess our number and thus reduce how much we'd have to pay out.

In explaining all of this to you before the game is over, we've given you a big hint. So now that you understand the real game you are playing, you've got one last guess, for \$20. What number do you pick?

49?

Congratulations. To us, not you. You've fallen right into our trap again. The number we picked was 48. Indeed, the whole speech about picking a number that was hard to find according to the split-the-interval rule was further designed to mislead you. We wanted you to pick 49 so that our choice of 48 would remain safe. Remember our objective is not to give you money.

To beat us at that game, you had to be one step ahead of us. You would have had to think, "They want us to pick 49, so I'm going to pick 48." Of course, if we had thought you would have been so clever, then we would have picked 47 or even 49.

The larger point of this game is not that we are selfish professors or cunning tricksters. Rather, the point is to illustrate as cleanly as possible what makes something a game: you have to take into account the objectives and strategies of the other players. When guessing a number picked at random, the number isn't trying to hide. You can take the engineer's mindset and divide the interval in two and do the best possible. But if you are playing a game, then you have to consider how the other player will be acting and how those decisions will influence your strategy.

#2. WINNING BY LOSING

We admit it: we watched *Survivor*. We would never have made it on the island. If we hadn't starved first, the others would surely have voted us off for being eggheads. For us, the challenge was trying to predict how the game would play out. We weren't surprised when the pudgy nudist Richard Hatch outwitted, outplayed, and outlasted his rivals to become the first champion of the CBS series and earn the million-dollar prize. He was gifted in his ability to act strategically without appearing to be strategic.

Hatch's most cunning ploy was in the last episode. The game was down to three players. Richard's two remaining rivals were 72-year-old retired Navy SEAL Rudy Boesch and 23-year-old river guide Kelly Wiglesworth. For their final challenge, the three of them had to stand on a pole with one hand on the immunity idol. The last one standing would go into the finals. And just as important, the winner would get to choose his or her opponent in the finals.

Your first impression might be that this was just a physical endurance contest. Think again. All three players understood that Rudy was the most popular of the three. If Rudy made it to the finals, Rudy would likely win. Richard's best hope was to go against Kelly in the finals.

There were two ways that could happen. One is that Kelly would win the pole-standing competition and pick Richard. The other is that Richard would win and pick Kelly. Richard could count on Kelly picking him. She was also aware of Rudy's popularity. Her best hope of winning was to get to the finals against Richard.

It would seem that if either Richard or Kelly won the final challenge, each would pick the other as his or her opponent. Hence Richard should try to stay in the game, at least until Rudy had fallen off. The only problem is that Richard and Rudy had a long-standing alliance. If Richard won the challenge and didn't pick Rudy, that would have turned Rudy (and all Rudy's friends) against Richard, and this could have cost him the victory. One of the great twists of *Survivor* is that the ousted contestants vote to determine the final winner. One has to be very careful how one disposes of rivals.

From Richard's perspective, the final challenge could go one of three ways:

- i. Rudy wins. Rudy then picks Richard, but Rudy would be the likely victor.
- ii. Kelly wins. Kelly would be smart enough to know her best hope was to eliminate Rudy and go against Richard.
- iii. Richard wins. If he picks Rudy to go on, Rudy beats him in the finals. If he picks Kelly to go on, Kelly might beat him because Richard would lose the support of Rudy and his many friends.

Comparing these options, Richard does best by losing. He wants Rudy eliminated, but it is better if Kelly does the dirty work for him. The smart money was on Kelly winning the challenge. She had won three of the previous four and as an outdoors guide was in the best shape of the three.

As a bonus, throwing the game saved Richard the trouble of standing on a pole under a hot sun. Early in the competition, host Jeff Probst offered a slice of orange to anyone willing to call it quits. Richard stepped off the pole and took the orange.

Throughout the book, you'll find these asides, which contain what we call a "Trip to the Gym." These trips take a look at more advanced elements of the game that we glossed over. For example, Richard could have tried to wait and see who dropped out first. If Kelly fell early, Richard might have preferred to beat Rudy and choose Kelly than to let Rudy win and have to go against Rudy in the finals. He might also have been concerned that Kelly would be savvy enough to do the same calculation and drop out early. The next chapters will show you how to use a more systematic approach to solve a game. The end goal is to help change the way you approach strategic situations, recognizing that you won't always have time to analyze every possible option.

After 4 hours and 11 minutes, Rudy fumbled when shifting his stance, let go of the immunity idol, and lost. Kelly picked Richard to go on to the finals. Rudy cast the swing vote in his favor, and Richard Hatch became the first *Survivor* champion.

With the benefit of hindsight it may all seem easy. What makes Richard's play so impressive is that he was able to anticipate all the different moves before they happened.* In chapter 2, we'll provide some tools to help you anticipate the way a game will play out and even give you a chance to have a go at another *Survivor* game.

#3. THE HOT HAND

Do athletes ever have a "hot hand"? Sometimes it seems that Yao Ming cannot miss a basket or that Sachin Tendulkar cannot fail to score a century in cricket. Sports announcers see these long streaks of consecutive successes and proclaim that the athlete has a hot hand. Yet according to psychology professors Thomas Gilovich, Robert Vallone, and Amos Tversky, this is a misperception of reality.¹

They point out that if you flip a coin long enough, you will find some very long series of consecutive heads. The psychologists suspect that sports commentators, short on insightful things to say, are just finding patterns in what amounts to a long series of coin

* Richard would have done well to anticipate the consequences of not paying taxes on his \$1 million winnings. On May 16, 2006, he was sentenced to 51 months in prison for tax evasion.

tosses over a long playing season. They propose a more rigorous test. In basketball, they look at all the instances of a player's baskets and observe the percentage of times that player's next shot is also a basket. A similar calculation is made for the shots immediately following misses. If a basket is more likely to follow a basket than to follow a miss, then there really is something to the theory of the hot hand.

They conducted this test on the Philadelphia 76ers basketball team. The results contradicted the hot hand view. When a player made his last shot, he was less likely to make his next; when he missed his previous attempt, he was more likely to make his next. This was true even for Andrew Toney, a player with the reputation for being a streak shooter. Does this mean we should be talking of the "stroboscopic hand," like the strobe light that alternates between on and off?

Game theory suggests a different interpretation. While the statistical evidence denies the presence of streak shooting, it does not refute the possibility that a hot player might warm up the game in some other way. The difference between streak shooting and a hot hand arises because of the interaction between the offensive and defensive strategies. Suppose Andrew Toney does have a truly hot hand. Surely the other side would start to crowd him. This could easily lower his shooting percentage.

That is not all. When the defense focuses on Toney, one of his teammates is left unguarded and is more likely to shoot successfully. In other words, Toney's hot hand leads to an improvement in the 76ers' *team* performance, although there may be a deterioration in Toney's *individual* performance. Thus we might test for hot hands by looking for streaks in team success.

Similar phenomena are observed in many other team sports. A brilliant running back on a football team improves the team's passing game and a great receiver helps the running game, as the opposition is forced to allocate more of its defensive resources to guard the stars. In the 1986 soccer World Cup final, the Argentine star Diego Maradona did not score a goal, but his passes through a ring of West German defenders led to two Argentine goals. The value

of a star cannot be assessed by looking only at his scoring performance; his contribution to his teammates' performance is crucial, and assist statistics help measure this contribution. In ice hockey, assists and goals are given equal weight for ranking individual performance.

A player may even assist himself when one hot hand warms up the other. The Cleveland Cavaliers star LeBron James eats and writes with his left hand but prefers shooting with his right (though his left hand is still better than most). The defense knows that LeBron is right-handed, so they concentrate on defending against right-handed shots. But they do not do so exclusively, since LeBron's left-handed shots are too effective to be left unguarded.

What happens when LeBron spends his off season working to improve his left-handed shooting? The defense responds by spending more time covering his left-handed shots. The result is that this frees his right hand more often. A better left-handed shot results in a more effective right-handed shot. In this case, not only does the left hand know what the right hand is doing, it's helping it out.

Going one step further, in chapter 5 we show that when the left hand is stronger it may even be used *less* often. Many of you will have experienced this seemingly strange phenomenon when playing tennis. If your backhand is much weaker than your forehand, your opponents will learn to play to your backhand. Eventually, as a result of all this backhand practice, your backhand will improve. As your two strokes become more equal, opponents can no longer exploit your weak backhand. They will play more evenly between forehand and backhand. You get to use your better forehand more often; this could be the real advantage of improving your backhand.

#4. TO LEAD OR NOT TO LEAD

After the first four races in the 1983 America's Cup finals, Dennis Conner's *Liberty* led 3-1 in a best-of-seven series. On the morning of the fifth race, "cases of champagne had been delivered

to *Liberty's* dock. And on their spectator yacht, the wives of the crew were wearing red-white-and-blue tops and shorts, in anticipation of having their picture taken after their husbands had prolonged the United States' winning streak to 132 years.² It was not to be.

At the start, *Liberty* got off to a 37-second lead when *Australia II* jumped the gun and had to recross the starting line. The Australian skipper, John Bertrand, tried to catch up by sailing way over to the left of the course in the hopes of catching a wind shift. Dennis Conner chose to keep *Liberty* on the right hand side of the course. Bertrand's gamble paid off. The wind shifted five degrees in *Australia II's* favor and she won the race by one minute and forty-seven seconds. Conner was criticized for his strategic failure to follow *Australia II's* path. Two races later, *Australia II* won the series.

Sailboat racing offers the chance to observe an interesting reversal of a "follow the leader" strategy. The leading sailboat usually copies the strategy of the trailing boat. When the follower tacks, so does the leader. The leader imitates the follower even when the follower is clearly pursuing a poor strategy. Why? Because in sailboat racing (unlike ballroom dancing) close doesn't count; only winning matters. If you have the lead, the surest way to stay ahead is to play monkey see, monkey do.*

Stock-market analysts and economic forecasters are not immune to this copycat strategy. The leading forecasters have an incentive to follow the pack and produce predictions similar to everyone else's. This way people are unlikely to change their perception of these forecasters' abilities. On the other hand, newcomers take the risky strategies; they tend to predict boom or doom. Usually they are wrong and are never heard of again, but now and again they are proven correct and move to the ranks of the famous.

* This strategy no longer applies once there are more than two competitors. Even with three boats, if one boat tacks right and the other tacks left, the leader has to choose which (if either) to follow.

Industrial and technological competitions offer further evidence. In the personal-computer market, Dell is less known for its innovation than for its ability to bring standardized technology to the mass market. More new ideas have come from Apple, Sun, and other start-up companies. Risky innovations are their best and perhaps only chance of gaining market share. This is true not just of high-technology goods. Procter & Gamble, the Dell of diapers, followed Kimberly-Clark's innovation of resealable diaper tape and recaptured its commanding market position.

There are two ways to move second. You can imitate as soon as the other has revealed his approach (as in sailboat racing) or wait longer until the success or failure of the approach is known (as in computers). The longer wait is more advantageous in business because, unlike in sports, the competition is usually not winner-take-all. As a result, market leaders will not follow the upstarts unless they also believe in the merits of their course.

#5. HERE I STAND

When the Catholic Church demanded that Martin Luther repudiate his attack on the authority of popes and councils, he refused to recant: "I will not recant anything, for to go against conscience is neither right nor safe." Nor would he compromise: "Here I stand, I cannot do otherwise."³ Luther's intransigence was based on the divinity of his positions. When defining what was right, there was no room for compromise. His firmness had profound long-term consequences; his attacks led to the Protestant Reformation and substantially altered the medieval Catholic Church.

Similarly, Charles de Gaulle used the power of intransigence to become a powerful player in the arena of international relations. As his biographer Don Cook expressed it, "[de Gaulle] could create power for himself with nothing but his own rectitude, intelligence, personality and sense of destiny."⁴ But above all, his was "the power of intransigence." During the Second World War, as the self-proclaimed leader in exile of a defeated and occupied nation, he held his own in negotiations with Roosevelt and Churchill. In the

1960s, his presidential "Non!" swung several decisions France's way in the European Economic Community (EEC).

In what way did his intransigence give him power in bargaining? When de Gaulle took a truly irrevocable position, the other parties in the negotiation were left with just two options—to take it or to leave it. For example, he single-handedly kept England out of the European Economic Community, once in 1963 and again in 1968; the other countries were forced either to accept de Gaulle's veto or to break up the EEC. De Gaulle judged his position carefully to ensure that it would be accepted. But that often left the larger (and unfair) division of the spoils to France. De Gaulle's intransigence denied the other party an opportunity to come back with a counteroffer that was acceptable.

In practice, this is easier said than done, for two kinds of reasons. The first kind stems from the fact that bargaining usually involves considerations other than the pie on today's table. The perception that you have been excessively greedy may make others less willing to negotiate with you in the future. Or, next time they may be more firm bargainers as they try to recapture some of their perceived losses. On a personal level, an unfair win may spoil business relations, or even personal relations. Indeed, biographer David Schoenbrun faulted de Gaulle's chauvinism: "In human relations, those who do not love are rarely loved: those who will not be friends end up by having none. De Gaulle's rejection of friendship thus hurt France."⁵ A compromise in the short term may prove a better strategy over the long haul.

The second kind of problem lies in achieving the necessary degree of intransigence. Luther and de Gaulle achieved this through their personalities, but this entails a cost. An inflexible personality is not something you can just turn on and off. Although being inflexible can sometimes wear down an opponent and force him to make concessions, it can equally well allow small losses to grow into major disasters.

Ferdinand de Lesseps was a mildly competent engineer with extraordinary vision and determination. He is famous for building the Suez Canal in what seemed almost impossible conditions. He

did not recognize the impossible and thereby accomplished it. Later, he tried using the same technique to build the Panama Canal. It ended in disaster.* Whereas the sands of the Nile yielded to his will, tropical malaria did not. The problem for de Lesseps was that his inflexible personality could not admit defeat even when the battle was lost.

How can one achieve selective inflexibility? Although there is no ideal solution, there are various means by which commitment can be achieved and sustained; this is the topic for chapter 7.

#6. THINNING STRATEGICALLY

Cindy Nacson-Schechter wanted to lose weight. She knew just what to do: eat less and exercise more. She knew all about the food pyramid and the hidden calories in soft drinks. Still, nothing had worked. She had gained forty pounds with the birth of her second child and it just wasn't coming off.

That's why she accepted ABC's offer to help her lose weight. On December 9, 2005, she came into a photographer's studio on Manhattan's West Side, where she found herself changing into a bikini. She hadn't worn a bikini since she was nine, and this wasn't the time to start again.

The setup felt like backstage at the *Sports Illustrated* swimsuit issue shoot. There were lights and cameras everywhere, and all she had on was a tiny lime-green bikini. The producers had thoughtfully placed a hidden space heater to keep her warm. Snap. Smile. Snap. Smile. What in the world was she thinking? Snap.

If things worked out as she hoped, no one would ever see these pictures. The deal she made with ABC *Primetime* was that they would destroy the pictures if she lost 15 pounds over the next two months. They wouldn't help her in any way. No coach, no trainer,

* The Suez Canal is a sea-level passage. The digging was relatively easy since the land was already low-lying and desert. Panama involved much higher elevations, lakes along the way, and dense jungle. Lesseps's attempt to dig down to sea level failed. Much later, the U.S. Army Corps of Engineers succeeded using a very different method—a sequence of locks, using the lakes along the way.

no special diets. She already knew what she had to do. All she needed was some extra motivation and a reason to start today rather than tomorrow.

Now she had that extra motivation. If she didn't lose the promised weight, ABC would show the photos and the videos on prime-time television. She had already signed the release giving them permission.

Fifteen pounds in two months was a safe amount to lose, but it wouldn't be a cakewalk. There was a series of holiday parties and Christmas dinners. She couldn't risk waiting until the New Year. She had to start now.

Cindy knew all about the dangers of being overweight—the increased risk of diabetes, heart attack, and death. And yet that wasn't enough to scare her into action. What she feared more than anything was the possibility that her ex-boyfriend would see her hanging out of a bikini on national TV. And there was little doubt that he would watch the show. Her best friend was going to tell him if she failed.

Laurie Edwards didn't like the way she looked or how she felt. It didn't help that she worked part-time tending bar, surrounded by hot twenty-somethings. She had tried Weight Watchers, South Beach, Slim-Fast, you name it. She was headed in the wrong direction and needed something to help her change course. When she told her girlfriends about the show, they thought it was the stupidest thing she'd ever done. The cameras captured that "what am I doing?" look on her face and a lot more.

Ray needed to lose weight, too. He was a newlywed in his twenties but looked closer to forty. As he walked the red carpet in his racing swimsuit, it wasn't a pretty picture. Click. Smile. Click.

He wasn't taking any chances. His wife wanted him to lose weight and was willing to help. She offered to diet with him. Then she took the plunge. She changed into a bikini, too. She wasn't as overweight as Ray, but she wasn't bikini-ready, either.

Her deal was different from Cindy's. She didn't have to weigh in. She didn't even have to lose weight. The pictures of her in a bikini would only be shown if Ray didn't lose the weight.

For Ray, the stakes had been raised even higher. He was either going to lose the weight or his wife.

All together, four women and one couple bared their soles and much more in front of the cameras. What were they doing? They weren't exhibitionists. The ABC producers had carefully screened them out. None of the five wanted to see these photos appear on TV, and none of them expected they ever would.

They were playing a game against their future selves. Today's self wants the future self to diet and exercise. The future self wants the ice cream and the television. Most of the time, the future self wins because it gets to move last. The trick is to change the incentives for the future self so as to change its behavior.

In Greek mythology, Odysseus wanted to hear the Sirens' songs. He knew that if he allowed his future self to listen to their song, that future self would sail his ship into the rocks. So he tied his hands—literally. He had his crew bind his hands to the mast (while plugging their own ears). In dieting, this is known as the empty-fridge strategy.

Cindy, Laurie, and Ray went one step further. They put themselves in a bind that only dieting would get them out of. You might think that having more options is always a good thing. But thinking strategically, you can often do better by cutting off options. Thomas Schelling describes how the Athenian General Xenophon fought with his back against an impassable ravine. He purposefully set himself up so that his soldiers had no option of retreat.⁶ Backs stiffened, they won.

Similarly, Cortés scuttled his ships upon arrival in Mexico. This decision was made with the support of his troops. Vastly outnumbered, his six hundred soldiers decided that they would either defeat the Aztecs or perish trying. The Aztecs could retreat inland, but for Cortés's soldiers there was no possibility of desertion or retreat. By making defeat worse, Cortés increased his chance of victory and indeed conquered.*

* Cortés was also helped by the Aztecs' misconception that he was Quetzalcoatl, a fair-skinned god.

What worked for Cortés and Xenophon worked for Cindy, Laurie, and Ray. Two months later, just in time for Valentine's Day, Cindy had lost 17 pounds. Ray was down 22 pounds and two belt loops. While the threat was the motivator to get them started, once they got going, they were doing it for themselves. Laurie lost the required 15 pounds in the first month. She kept on going and lost another 13 in month two. Laurie's 28 pounds translated into two dress sizes and over 14 percent of her body weight. Her friends no longer think the ABC show was a stupid idea.

At this point, you shouldn't be surprised to know that one of us was behind the show's design.⁷ Perhaps we should have called this book *Thinning Strategically* and sold many more copies. Alas, not, and we return to study these types of strategic moves in chapter 6.

#7. BUFFETT'S DILEMMA

In an op-ed promoting campaign finance reform, the Oracle of Omaha, Warren Buffett, proposed raising the limit on individual contributions from \$1,000 to \$5,000 and banning all other contributions. No corporate money, no union money, no soft money. It sounds great, except that it would never pass.

Campaign finance reform is so hard to pass because the incumbent legislators who have to approve it are the ones who have the most to lose. Their advantage in fundraising is what gives them job security.* How do you get people to do something that is against their interest? Put them in what is known as the prisoners' dilemma.† According to Buffett:

Well, just suppose some eccentric billionaire (not me, not me!) made the following offer: If the bill was defeated, this person—

* Between 1992 and 2000, Dan Rostenkowski was the only case of an incumbent congressman losing a contest for reelection. The incumbent success rate was 604 out of 605, or 99.8 percent. When Rostenkowski lost, he was under indictment on seventeen counts of extortion, obstruction of justice, and misuse of funds.

† While prisoner's dilemma is the more common usage, we prefer the plural, because unless there are two or more prisoners involved, there is no dilemma.

the E.B.—would donate \$1 billion in an allowable manner (soft money makes all possible) to the political party that had delivered the most votes to getting it passed. Given this diabolical application of game theory, the bill would sail through Congress and thus cost our E.B. nothing (establishing him as not so eccentric after all).⁸

Consider your options as a Democratic legislator. If you think that the Republicans will support the bill and you work to defeat it, then if you are successful, you will have delivered \$1 billion to the Republicans, thereby handing them the resources to dominate for the next decade. Thus there is no gain in opposing the bill if the Republicans are supporting it. Now, if the Republicans are against it and you support it, then you have the chance of making \$1 billion.

Thus whatever the Republicans do, the Democrats should support the bill. Of course, the same logic applies to the Republicans. They should support the bill no matter what the Democrats do. In the end, both parties support the bill, and our billionaire gets his proposal for free. As a bonus, Buffett notes that the very effectiveness of his plan “would highlight the absurdity of claims that money doesn’t influence Congressional votes.”

This situation is called a prisoners’ dilemma because both sides are led to take an action that is against their mutual interest.* In the classic version of the prisoners’ dilemma, the police are separately interrogating two suspects. Each is given an incentive to be the first to confess and a much harsher sentence if he holds out while the other confesses. Thus each finds it advantageous to confess, though they would both do better if each kept quiet.

Truman Capote’s *In Cold Blood* provides a vivid illustration. Richard “Dick” Hickock and Perry Edward Smith have been arrested for the senseless murder of the Clutter family. While there were no witnesses to the crime, a jailhouse snitch had given their

* The active players in the game are the losers, but outsiders can benefit. While incumbent politicians might be unhappy with campaign finance reform, the rest of us would be better off.

names to the police. During the interrogation, the police play one against the other. Capote takes us into Perry's mind:

... that it was just another way of getting under his skin, like that phony business about a witness—"a living witness." There couldn't be. Or did they mean— If only he could talk to Dick! But he and Dick were being kept apart; Dick was locked in a cell on another floor. . . . And Dick? Presumably they'd pulled the same stunt on him. Dick was smart, a convincing performer, but his "guts" were unreliable, he panicked too easily. . . . "And before you left that house you killed all the people in it." It wouldn't amaze him if every Old Grad in Kansas had heard that line. They must have questioned hundreds of men, and no doubt accused dozens; he and Dick were merely two more. . . .

And Dick, awake in a cell on the floor below, was (he later recalled) equally eager to converse with Perry—find out what the punk had told them.⁹

Eventually Dick confessed and then Perry.* That's the nature of the game.

The problem of collective action is a variant of the prisoners' dilemma, albeit one with many more than two prisoners. In the children's story about belling the cat, the mice decide that life would be much safer if the cat were stuck with a bell around its neck. The problem is, who will risk his life to bell the cat?

This is a problem for both mice and men. How can unpopular tyrants control large populations for long periods? Why can a lone bully terrorize a schoolyard? In both cases, a simultaneous move by the masses stands a very good chance of success.

But the communication and coordination required for such action is difficult, and the oppressors, knowing the power of the masses, take special steps to keep it difficult. When the people must act individually and hope that the momentum will build up, the question arises, "Who is going to be the first?" Such a leader

* Although each of the two thought that confession would bring more favorable treatment, in this instance that did not happen—both were sentenced to death.

will pay a high cost—a broken nose or possibly his life. His reward may be posthumous glory or gratitude. There are people who are moved by considerations of duty or honor, but most find the costs exceed the benefits.

Khrushchev first denounced Stalin's purges at the Soviet Communist Party's 20th Congress. After his dramatic speech, someone in the audience shouted out, asking what Khrushchev had been doing at the time. Khrushchev responded by asking the questioner to please stand up and identify himself. The audience remained silent. Khrushchev replied, "That is what I did, too."

Each person acts in his or her self-interest, and the result is a disaster for the group. The prisoners' dilemma is perhaps the most famous and troubling game in game theory, and we return to the topic in chapter 3 to discuss what can be done. We should emphasize right from the start that we have no presumption that the outcome of a game will be good for the players. Many economists, ourselves included, tout the advantages of the free market. The theory behind this conclusion relies on a price system that guides individual behavior. In most strategic interactions, there is no invisible hand of prices to guide the baker or the butcher or anyone else. Thus there is no reason to expect that the outcome of a game will be good for the players or society. It may not be enough to play a game well—you must also be sure you are playing the right game.

#8. MIX YOUR PLAYS

Apparently Takashi Hashiyama has trouble making decisions. Both Sotheby's and Christie's had made attractive offers to be the auction house for the sale of his company's \$18 million art collection. Rather than choose one over the other, he suggested the two of them play a game of Rock Paper Scissors to determine the winner. Yes, Rock Paper Scissors. Rock breaks scissors, scissors cuts paper, and paper covers rock.

Christie's chose scissors and Sotheby's chose paper. Scissors cut paper and so Christie's won the assignment and a nearly \$3 million

commission. With the stakes so high, could game theory have helped?

The obvious point is that in this type of game, one can't be predictable. If Sotheby's had known that Christie's would be playing scissors, then they would have chosen rock. No matter what you choose, there is something else that beats it. Hence it is important that the other side can't predict your play.

As part of their preparation, Christie's turned to local experts, namely the kids of their employees who play the game regularly. According to eleven-year-old Alice, "Everybody knows you always start with scissors." Alice's twin sister, Flora, added her perspective: "Rock is way too obvious, and scissors beats paper. Since they were beginners, scissors was definitely the safest."¹⁰

Sotheby's took a different tack. They thought this was simply a game of chance and hence there was no room for strategy. Paper was as good as anything else.

What is interesting here is that both sides were half right. If Sotheby's picked its strategy at random—with an equal chance of rock, scissors, or paper—then whatever Christie's did would have been equally good. Each option has a one-third chance of winning, a one-third chance of losing, and a one-third chance of a tie.

But Christie's didn't pick at random. Thus Sotheby's would have done better to think about the advice Christie's would likely get and then play to beat it. If it's true that everyone knows you start with scissors, Sotheby's should have started with Bart Simpson's favorite, good old rock.

In that sense, both players got it half wrong. Given Sotheby's lack of strategy, there was no point in Christie's efforts. But given Christie's efforts, there would have been a point to Sotheby's thinking strategically.

In a single play of a game, it isn't hard to choose randomly. But when games get repeated, the approach is trickier. Mixing your plays does not mean rotating your strategies in a predictable manner. Your opponent can observe and exploit any systematic pattern almost as easily as he can exploit an unchanging repetition of a single strategy. It is *unpredictability* that is important when mixing.

It turns out most people fall into predictable patterns. You can test this yourself online where computer programs are able to find the pattern and beat you.¹¹ In an effort to mix things up, players often rotate their strategies too much. This leads to the surprise success of the "avalanche" strategy: rock, rock, rock.

People are also too influenced by what the other side did last time. If both Sotheby's and Christie's had opened with scissors, there would have been a tie and a rematch. According to Flora, Sotheby's would expect Christie's to play rock (to beat their scissors). That should lead Sotheby's to play paper and so Christie's should stick with scissors. Of course, that formulaic approach can't be right, either. If it were, Sotheby's could then play rock and win.

Imagine what would happen if there were some known formula that determined who would be audited by the IRS. Before you submitted a tax return, you could apply the formula to see if you would be audited. If an audit was predicted, but you could see a way to "amend" your return until the formula no longer predicted an audit, you probably would do so. If an audit was unavoidable, you would choose to tell the truth. The result of the IRS being completely predictable is that it would audit exactly the wrong people. All those audited would have anticipated their fate and chosen to act honestly, while those spared an audit would have only their consciences to watch over them. When the IRS audit formula is somewhat fuzzy, everyone stands some risk of an audit; this gives an added incentive for honesty.

The importance of randomized strategies was one of the early insights of game theory. The idea is simple and intuitive but needs refinement to be useful in practice. It is not enough for a tennis player to know that he should mix his shots between the opponent's forehand and backhand. He needs some idea of whether he should go to the forehand 30 percent or 64 percent of the time and how the answer depends on the relative strengths of the two sides. In chapter 5 we develop methods to answer such questions.

We'd like to leave you with one last commentary. The biggest loser in the Rock Paper Scissors game wasn't Sotheby's; it was Mr.

Hashiyama. His decision to deploy Rock Paper Scissors gave each of the two auction houses a 50 percent chance of winning the commission. Instead of letting the two contestants effectively agree to split the commission, he could have run his own auction. Both firms were willing, even eager, to lead the sale with a 12 percent commission.* The winning house would be the one willing to take the lowest fee. Do I hear 11 percent? Going once, going twice, . . .

#9. NEVER GIVE A SUCKER AN EVEN BET

In *Guys and Dolls*, gambler Sky Masterson relates this valuable advice from his father:

One of these days in your travels, a guy is going to show you a brand-new deck of cards on which the seal is not yet broken. Then this guy is going to offer to bet you that he can make the jack of spades jump out of this brand-new deck of cards and squirt cider in your ear. But, son, you do not accept this bet because, as sure as you stand there, you're going to wind up with an ear full of cider.

The context of the story is that Nathan Detroit has offered Sky Masterson a bet about whether Mindy's sells more strudel or cheesecake. Nathan had just discovered the answer (strudel) and is willing to bet if Sky will bet on cheesecake.†

This example may sound somewhat extreme. Of course no one would take such a sucker bet. Or would they? Look at the market for futures contracts on the Chicago Board of Exchange. If

* The standard commission is 20 percent on the first \$800,000 and 12 percent thereafter. Mr. Hashiyama's four paintings sold for a combined \$17.8 million, suggesting a total commission of \$2.84 million.

† We should add that Sky never quite learned his father's lesson. A minute later, he offers to bet that Nathan does not know the color of his own bow tie. Sky can't win. If Nathan knows the color, he takes the bet and wins. As it turns out, Nathan doesn't know the color and thus doesn't bet. Of course, that was the real gamble. Sky is betting that Nathan won't take the offer.

another speculator offers to sell you a futures contract, he will make money only if you lose money.*

If you happen to be a farmer with soy beans to sell in the future, then the contract can provide a hedge against future price movements. Similarly, if you sell soy milk and hence need to buy soy beans in the future, this contract is insurance, not a gamble.

But the volume of the contracts on the exchange suggests that most people buying and selling are traders, not farmers and manufacturers. For them, the deal is a *zero-sum* game. When both sides agree to trade, each one thinks it will make money. One of them must be wrong. That's the nature of a zero-sum game. Both sides can't win.

This is a paradox. How can both sides think that they can outsmart the other? Someone must be wrong. Why do you think the other person is wrong, not you? Let us assume that you don't have any insider information. If someone is willing to sell you a futures contract, any money you make is money they lose. Why do you think that you are smarter than they are? Remember that their willingness to trade means that they think they are smarter than you.

In poker, players battle this paradox when it comes to raising the stakes. If a player bets only when he has a strong hand, the other players will soon figure this out. In response to a raise, most other players will fold, and he'll never win a big pot. Those who raise back will have even stronger hands, and so our poor player will end up a big loser. To get others to bet against a strong hand, they have to think you might be bluffing. To convince them of this possibility, it helps to bet often enough so that you must be bluffing some of the time. This leads to an interesting dilemma. You'd like others to fold against your bluffs and thereby win with weak hands. But that won't lead to high-pot victories. To convince others to raise your bets, you also need to get caught bluffing.

As the players get even more sophisticated, persuading others to take big bets against you becomes harder and harder. Consider the

* Buying stocks is not the same as betting on a futures contract. In the case of stocks, the capital you provide to the firm allows it to grow faster, and thus you and the firm can both win.

following high-stakes game of wits between Erick Lindgren and Daniel Negreanu, two of poker's top-ranked players.

. . . Negreanu, sensing a weak hand, raised him two hundred thousand [dollars]. "I put two hundred and seventy thousand in, so I have two hundred thousand left," Negreanu said. "And Erick looks over my chips and says, 'How much you got left?' And he moves all in"—wagering all he had. Under the special betting rules governing the tournament, Negreanu had only ninety seconds to decide whether to call the bet, and risk losing all his money if Lindgren wasn't bluffing, or to fold, and give up the hefty sum he had already put into the pot.

"I didn't think he could be so stupid," Negreanu said. "But it wasn't stupid. It was like a step above. He knows that I know that he wouldn't do something so stupid, so by doing something so quote-unquote stupid it actually became a great play."¹²

While it is obvious that you shouldn't bet against these poker champions, when should you take a gamble? Groucho Marx famously said that he didn't care to belong to any club that would accept him as a member. For similar reasons, you might not want to take any bet that others offer. You should even be worried when you win an auction. The very fact that you were the highest bidder implies that everyone else thought the item was worth less than you did. The result of winning an auction and discovering you've overpaid is called the winner's curse.

Every action someone takes tells us something about what he knows, and you should use these inferences along with what you already know to guide your actions. How to bid so that you won't be cursed when you win is something we discuss in chapter 10.

There are some rules of the game that can help put you on more equal footing. One way to allow trading with lopsided information is to let the less informed party pick which side of the bet to take. If Nathan Detroit agreed in advance to take the bet whatever side Sky picked, then Nathan's inside information would be of no help. In stock markets, foreign exchange markets, and other financial markets, people are free to take either side of the bet. Indeed, in

some exchanges, including the London stock market, when you ask for a quote on a stock the market maker is required to state both the buying and selling prices *before* he knows which side of the transaction you want. Without such a safeguard, market makers could stand to profit from private information, and the outside investors' fear of being suckered might cause the entire market to fold. The buy and sell prices are not quite the same; the difference is called the bid-ask spread. In liquid markets the spread is quite small, indicating that little information is contained in any buy or sell order. We return to the role of information in chapter 8.

#10. GAME THEORY CAN BE DANGEROUS TO YOUR HEALTH

Late one night, after a conference in Jerusalem, two American economists—one of whom is this book's coauthor—found a taxicab and gave the driver directions to the hotel. Immediately recognizing us as American tourists, the driver refused to turn on his meter; instead, he proclaimed his love for Americans and promised us a lower fare than the meter. Naturally, we were somewhat skeptical of this promise. Why should this stranger offer to charge less than the meter when we were willing to pay the metered fare? How would we even know whether or not we were being overcharged?

On the other hand, we had not promised to pay the driver anything more than what would be on the meter. We put on our game-theory hats. If we were to start bargaining and the negotiations broke down, we would have to find another taxi. But if we waited until we arrived at the hotel, our bargaining position would be much stronger. And taxis were hard to find.

We arrived. The driver demanded 2,500 Israeli shekels (\$2.75). Who knew what fare was fair? Because people generally bargain in Israel, Barry protested and counteroffered 2,200 shekels. The driver was outraged. He claimed that it would be impossible to get from there to here for that amount. Before negotiations could continue, he locked all the doors automatically and retraced the route at breakneck speed, ignoring traffic lights and pedestrians. Were

we being kidnapped to Beirut? No. He returned to the original position and ungraciously kicked us out of his cab, yelling, "See how far your 2,200 shekels will get you now."

We found another cab. This driver turned on his meter, and 2,200 shekels later we were home.

Certainly the extra time was not worth the 300 shekels. On the other hand, the story was well worth it. It illustrates the dangers of bargaining with those who have not yet read our book. More generally, pride and irrationality cannot be ignored. Sometimes, it may be better to be taken for a ride when it costs only two dimes.

There is a second lesson to the story. We didn't really think far enough ahead. Think of how much stronger our bargaining position would have been if we had begun to discuss the price *after* getting out of the taxi. (Of course, for hiring a taxi, this logic should be reversed. If you tell the driver where you want to go before getting in, you may find your taxi chasing after some other customer. Get in first, then say where you want to go.)

Some years after this story was first published, we received the following letter:

Dear Professors,

You certainly don't know my name, but I think you will remember my story. I was a student in Jerusalem moonlighting as a taxi driver. Now I am a consultant and chanced upon your book when it was translated into Hebrew. What you might find interesting is that I too have been sharing the story with my clients. Yes, it was indeed a late night in Jerusalem. As for the rest, well, I recall things a bit differently.

Between classes and working nights as a taxi driver, there was almost no time for me to spend with my new bride. My solution was to have her ride with me in the front seat. Although she was silent, it was a big mistake for you to have left her out of the story.

My meter was broken, but you didn't seem to believe me. I was too tired to argue. When we arrived, I asked for 2,500 shekels, a fair price. I was even hoping you would round the fare up to 3,000. You rich Americans could well afford a 50¢ tip.

I couldn't believe you tried to cheat me. Your refusal to pay a fair price dishonored me in front of my wife. As poor as I was, I did not need to take your meager offer.

Americans think that we should be happy to take whatever crumbs you offer. I say that we should teach you a lesson in the game of life. My wife and I are now married twenty years. We still laugh about those stupid Americans who spent a half an hour riding back and forth in taxis to save twenty cents.

Sincerely,
(name withheld)

Truth be told, we never received such a letter. Our point in creating it was to illustrate a critical lesson in game theory: you need to understand the other player's perspective. You need to consider what they know, what motivates them, and even how they think about you. George Bernard Shaw's quip on the golden rule was to not do unto others as you would have them do unto you—their tastes may be different. When thinking strategically, you have to work extra hard to understand the perspective and interactions of all the other players in the game, including ones who may be silent.

That brings us to one last point. You may be thinking you are playing one game, but it is only part of a larger game. There is always a larger game.

THE SHAPE OF THINGS TO COME

These examples have given us glimpses of principles that guide strategic decisions. We can summarize these principles with a few "morals" from our tales.

Think 48 when you are wondering what the other player is trying to achieve. Recall Richard Hatch's ability to play out all the future moves to figure out what he should do. The story of the hot hand told us that in strategy, no less than in physics, "For every action we take, there is a reaction." We do not live and act in a vacuum. Therefore, we cannot assume that when we change our

behavior everything else will remain unchanged. De Gaulle's success in negotiations suggests that "the stuck wheel gets the grease."* But being stubborn is not always easy, especially when one has to be more stubborn than an obstinate adversary. That stubborn adversary might well be your future self, especially when it comes to dieting. Fighting or dieting with your back up against the wall can help strengthen your resolve.

In Cold Blood and the story of belling the cat demonstrate the difficulty of obtaining outcomes that require coordination and individual sacrifice. In technology races, no less than in sailboat races, those who trail tend to employ more innovative strategies; the leaders tend to imitate the followers.

Rock Paper Scissors points out the strategic advantage of being unpredictable. Such behavior may also have the added advantage that it makes life just a little more interesting. Our taxi rides make it clear that the other players in games are people, not machines. Pride, spite, and other emotions may color their decisions. When you put yourself in others' shoes, you have to take them as they are, not as you are.

We could go on offering more examples and drawing morals from them, but this is not the best way to think methodically about strategic games. That is better done by approaching the subject from a different angle. We pick up the principles—for example, commitment, cooperation, and mixing—one at a time. In each instance, we explore examples that bear centrally on that issue, until the principle is clear. Then you will have a chance to apply that principle in the case study that ends each chapter.

CASE STUDY: MULTIPLE CHOICE

We think almost everything in life is a game, even things that might not seem that way at first. Consider the following question from the GMAT (the test given to MBA applicants).

* You may have heard this expression as the "squeaky wheel"—a stuck wheel needs even more grease. Of course, sometimes it gets replaced.

Unfortunately, issues of copyright clearance have prevented us from reproducing the question, but that shouldn't stop us. Which of the following is the correct answer?

- | | | |
|----------------------|-----------------------|-----------------------|
| a. 4π sq. inches | c. 16 sq. inches | e. 32π sq. inches |
| b. 8π sq. inches | d. 16π sq. inches | |

Okay, we recognize that you're at a bit of a disadvantage not having the question. Still, we think that by putting on your game-theory hat you can still figure it out.

Case Discussion

The odd answer in the series is c. Since it is so different from the other answers, it is probably not right. The fact that the units are in square inches suggests an answer that has a perfect square in it, such as 4π or 16π .

This is a fine start and demonstrates good test-taking skills, but we haven't really started to use game theory. Think of the game being played by the person writing the question. What is that person's objective?

He or she wants people who understand the problem to get the answer right and those who don't to get it wrong. Thus wrong answers have to be chosen carefully so as to be appealing to folks who don't quite know the answer. For example, in response to the question: How many feet are in a mile, an answer of "Giraffe," or even 16π , is unlikely to attract any takers.

Turning this around, imagine that 16 square inches really is the right answer. What kind of question might have 16 square inches as the answer but would lead someone to think 32π is right? Not many. People don't often go around adding π to answers for the fun of it. "Did you see my new car—it gets 10π miles to the gallon." We think not. Hence we can truly rule out 16 as being the correct solution.

Let's now turn to the two perfect squares, 4π and 16π . Assume for a moment that 16π square inches is the correct solution. The problem might have been what is the area of a circle with a radius of 4? The correct formula for the area of a circle is πr^2 . However,

the person who didn't quite remember the formula might have mixed it up with the formula for the circumference of a circle, $2\pi r$. (Yes, we know that the circumference is in inches, not square inches, but the person making this mistake would be unlikely to recognize this issue.)

Note that if $r = 4$, then $2\pi r$ is 8π , and that would lead the person to the wrong answer of b. The person could also mix and match and use the formula $2\pi r^2$ and hence believe that 32π or e was the right answer. The person could leave off the π and come up with 16 or c, or the person could forget to square the radius and simply use πr as the area, leading to 4π or a. In summary, if 16π is the correct answer, then we can tell a plausible story about how each of the other answers might be chosen. They are all good wrong answers for the test maker.

What if 4π is the correct solution (so that $r = 2$)? Think now about the most common mistake, mixing up circumference with area. If the student used the wrong formula, $2\pi r$, he or she would still get 4π , albeit with incorrect units. There is nothing worse, from a test maker's perspective, than allowing the person to get the right answer for the wrong reason. Hence 4π would be a terrible right answer, as it would allow too many people who didn't know what they were doing to get full credit.

At this point, we are done. We are confident that the right answer is 16π . And we are right. By thinking about the objective of the person writing the test, we can suss out the right answer, often without even seeing the question.

Now, we don't recommend that you go about taking the GMAT and other tests without bothering to even look at the questions. We appreciate that if you are smart enough to go through this logic, you most likely know the formula for the area of a circle. But you never know. There will be cases where you don't know the meaning of one of the answers or the material for the question wasn't covered in your course. In those cases, thinking about the testing game may lead you to the right answer.

Games Solvable by Backward Reasoning

IT'S YOUR MOVE, CHARLIE BROWN

In a recurring theme in the comic strip *Peanuts*, Lucy holds a football on the ground and invites Charlie Brown to run up and kick it. At the last moment, Lucy pulls the ball away. Charlie Brown, kicking only air, lands on his back, and this gives Lucy great perverse pleasure.

Anyone could have told Charlie that he should refuse to play Lucy's game. Even if Lucy had not played this particular trick on him last year (and the year before and the year before that), he knows her character from other contexts and should be able to predict her action.

At the time when Charlie is deciding whether or not to accept Lucy's invitation, her action lies in the future. However, just because it lies in the future does not mean Charlie should regard it as uncertain. He should know that of the two possible outcomes—letting him kick and seeing him fall—Lucy's preference is for the latter. Therefore he should forecast that when the time comes, she is going to pull the ball away. The logical possibility that Lucy will let him kick the ball is realistically

irrelevant. Reliance on it would be, to borrow Dr. Johnson's characterization of remarriage, a triumph of hope over experience. Charlie should disregard it, and forecast that acceptance will inevitably land him on his back. He should decline Lucy's invitation.

TWO KINDS OF STRATEGIC INTERACTIONS

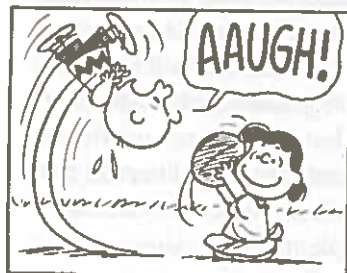
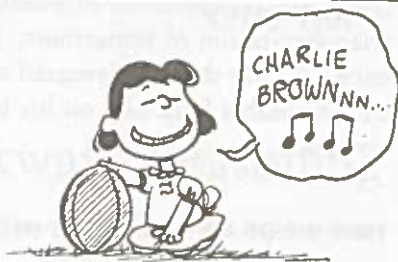
The essence of a game of strategy is the interdependence of the players' decisions. These interactions arise in two ways. The first is *sequential*, as in the Charlie Brown story. The players make alternating moves. Charlie, when it is his turn, must look ahead to how his current actions will affect the future actions of Lucy, and his own future actions in turn.

The second kind of interaction is *simultaneous*, as in the prisoners' dilemma tale of chapter 1. The players act at the same time, in ignorance of the others' current actions. However, each must be aware that there are other active players, who in turn are similarly aware, and so on. Therefore each must figuratively put himself in the shoes of all and try to calculate the outcome. His own best action is an integral part of this overall calculation.

When you find yourself playing a strategic game, you must determine whether the interaction is simultaneous or sequential. Some games, such as football, have elements of both, in which case you must fit your strategy to the context. In this chapter, we develop, in a preliminary way, the ideas and rules that will help you play sequential games; simultaneous-move games are the subject of chapter 3. We begin with really simple, sometimes contrived, examples, such as the Charlie Brown story. This is deliberate; the stories are not of great importance in themselves, and the right strategies are usually easy to see by simple intuition, allowing the underlying ideas to stand out much more clearly. The examples get increasingly realistic and more complex in the case studies and in the later chapters.

PEANUTS

featuring
"Good ol'
Charlie Brown"
by SCHULZ



The First Rule of Strategy

The general principle for sequential-move games is that each player should figure out the other players' future responses and use them in calculating his own best current move. This idea is so important that it is worth codifying into a basic rule of strategic behavior:

RULE 1: Look forward and reason backward.

Anticipate where your initial decisions will ultimately lead and use this information to calculate your best choice.

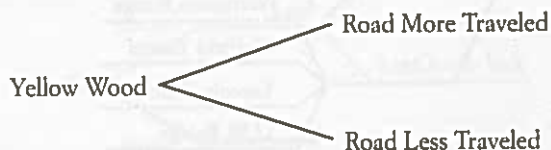
In the Charlie Brown story, this was easy to do for anyone (except Charlie Brown). He had just two alternatives, and one of them led to Lucy's decision between two possible actions. Most strategic situations involve a longer sequence of decisions with several alternatives at each. A tree diagram of the choices in the game sometimes serves as a visual aid for correct reasoning in such games. Let us show you how to use these trees:

DECISION TREES AND GAME TREES

A sequence of decisions, with the need to look forward and reason backward, can arise even for a solitary decision maker not involved in a game of strategy with others. For Robert Frost in the yellow wood:

Two roads diverged in a wood, and I—
I took the road less traveled by,
And that has made all the difference.¹

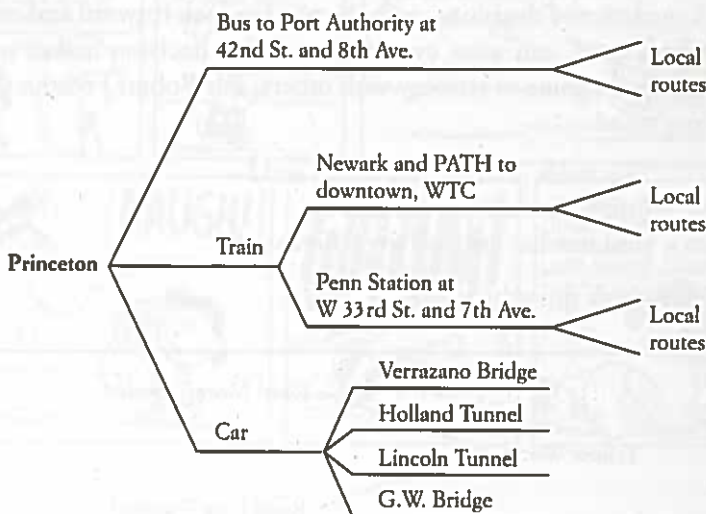
We can show this schematically.



This need not be the end of the choice. Each road might in turn have further branches. The road map becomes correspondingly complex. Here is an example from our own experience.

Travelers from Princeton to New York have several choices. The first decision point involves selecting the mode of travel: bus, train, or car. Those who drive then have to choose among the Verrazano-Narrows Bridge, the Holland Tunnel, the Lincoln Tunnel, and the George Washington Bridge. Rail commuters must decide whether to switch to the PATH train at Newark or continue to Penn Station. Once in New York, rail and bus commuters must choose among going by foot, subway (local or express), bus, or taxi to get to their final destination. The best choices depend on many factors, including price, speed, expected congestion, the final destination in New York, and one's aversion to breathing the air on the New Jersey Turnpike.

This road map, which describes one's options at each junction, looks like a tree with its successively emerging branches—hence the term. The right way to use such a map or tree is not to take the

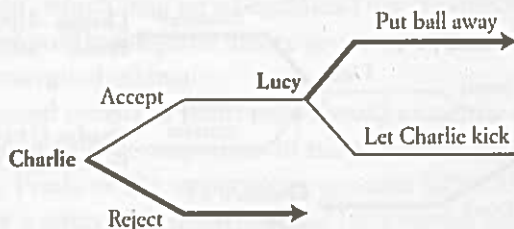


route whose first branch looks best—for example, because you would prefer driving to taking the train when all other things are equal—and then “cross the Verrazano Bridge when you get to it.” Instead, you anticipate the future decisions and use them to make your earlier choices. For example, if you want to go downtown, the PATH train would be superior to driving because it offers a direct connection from Newark.

We can use just such a tree to depict the choices in a game of strategy, but one new element enters the picture. A game has two or more players. At various branching points along the tree, it may be the turn of different players to make the decision. A person making a choice at an earlier point must look ahead, not just to his own future choices but to those of others. He must forecast what the others will do, by putting himself figuratively in their shoes, and thinking as they would think. To remind you of the difference, we will call a tree showing the decision sequence in a game of strategy a *game tree*, reserving *decision tree* for situations in which just one person is involved.

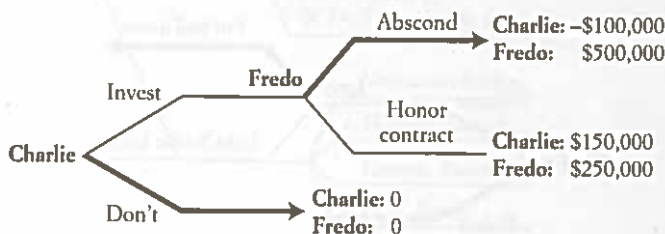
Charlie Brown in Football and in Business

The story of Charlie Brown that opened this chapter is absurdly simple, but you can become familiar with game trees by casting that story in such a picture. Start the game when Lucy has issued her invitation, and Charlie faces the decision of whether to accept. If Charlie refuses, that is the end of the game. If he accepts, Lucy has the choice between letting Charlie kick and pulling the ball away. We can show this by adding another fork along this road.



As we said earlier, Charlie should forecast that Lucy will choose the upper branch. Therefore he should figuratively prune the lower branch of her choice from the tree. Now if he chooses his own upper branch, it leads straight to a nasty fall. Therefore his better choice is to follow his own lower branch. We show these selections by making the branches thicker and marking them with arrowheads.

Are you thinking that this game is too frivolous? Here is a business version of it. Imagine the following scenario. Charlie, now an adult, is vacationing in the newly reformed formerly Marxist country of Freedonia. He gets into a conversation with a local businessman named Fredo, who talks about the wonderful profitable opportunities that he could develop given enough capital, and then makes a pitch: "Invest \$100,000 with me, and in a year I will turn it into \$500,000, which I will share equally with you. So you will more than double your money in a year." The opportunity Fredo describes is indeed attractive, and he is willing to write up a proper contract under Freedonian law. But how secure is that law? If at the end of the year Fredo absconds with all the money, can Charlie, back in the United States, enforce the contract in Freedonian courts? They may be biased in favor of their national, or too slow, or bribed by Fredo. So Charlie is playing a game with Fredo, and the tree is as shown here. (Note that if Fredo honors the contract, he pays Charlie \$250,000; therefore Charlie's profit is that minus the initial investment of \$100,000—that is, \$150,000.)



What do you think Fredo is going to do? In the absence of a clear and strong reason to believe his promise, Charlie should predict that Fredo will abscond, just as young Charlie should have been sure that Lucy would pull the ball away. In fact the trees of the two games are identical in all essential respects. But how many Charlies have failed to do the correct reasoning in such games?

What reasons can there be for believing Fredo's promise? Perhaps he is engaged in many other enterprises that require financing from the United States or export goods to the United States. Then Charlie may be able to retaliate by ruining his reputation in the United States or seizing his goods. So this game may be part of a larger game, perhaps an ongoing interaction, that ensures Fredo's honesty. But in the one-time version we showed above, the logic of backward reasoning is clear.

We would like to use this game to make three remarks. First, different games may have identical or very similar mathematical forms (trees, or the tables used for depictions in later chapters). Thinking about them using such formalisms highlights the parallels and makes it easy to transfer your knowledge about a game in one situation to that in another. This is an important function of the "theory" of any subject: it distills the essential similarities in apparently dissimilar contexts and enables one to think about them in a unified and therefore simplified manner. Many people have an instinctive aversion to theory of any kind. But we think this is a mistaken reaction. Of course, theories have their limitations. Specific contexts and experiences can often add to or modify the prescriptions of theory in substantial ways. But to abandon theory altogether would be to abandon a valuable starting point for thought, which may be a beachhead for conquering the problem. You should make game theory your friend, and not a bugbear, in your strategic thinking.

The second remark is that Fredo should recognize that a strategic Charlie would be suspicious of his pitch and not invest at all, depriving Fredo of the opportunity to make \$250,000. Therefore Fredo has a strong incentive to make his promise credible. As an individual businessman, he has little influence over Freedonia's

weak legal system and cannot allay the investor's suspicion that way. What other methods may be at his disposal? We will examine the general issue of credibility, and devices for achieving it, in chapters 6 and 7.

The third, and perhaps most important, remark concerns comparisons of the different outcomes that could result based on the different choices the players could make. It is not always the case that more for one player means less for the other. The situation where Charlie invests and Fredo honors the contract is better for both than the one where Charlie does not invest at all. Unlike sports or contests, games don't have to have winners and losers; in the jargon of game theory, they don't have to be zero-sum. Games can have win-win or lose-lose outcomes. In fact, some combination of commonality of interest (as when Charlie and Fredo can both gain if there is a way for Fredo to commit credibly to honoring the contract) and some conflict (as when Fredo can gain at Charlie's expense by absconding after Charlie has invested) coexist in most games in business, politics, and social interactions. And that is precisely what makes the analysis of these games so interesting and challenging.

More Complex Trees

We turn to politics for an example of a slightly more complex game tree. A caricature of American politics says that Congress likes pork-barrel expenditures and presidents try to cut down the bloated budgets that Congress passes. Of course presidents have their own likes and dislikes among such expenditures and would like to cut only the ones they dislike. To do so, they would like to have the power to cut out specific items from the budget, or a line-item veto. Ronald Reagan in his State of the Union address in January 1987 said this eloquently: "Give us the same tool that 43 governors have—a line-item veto, so we can carve out the boondoggles and pork, those items that would never survive on their own."

At first sight, it would seem that having the freedom to veto

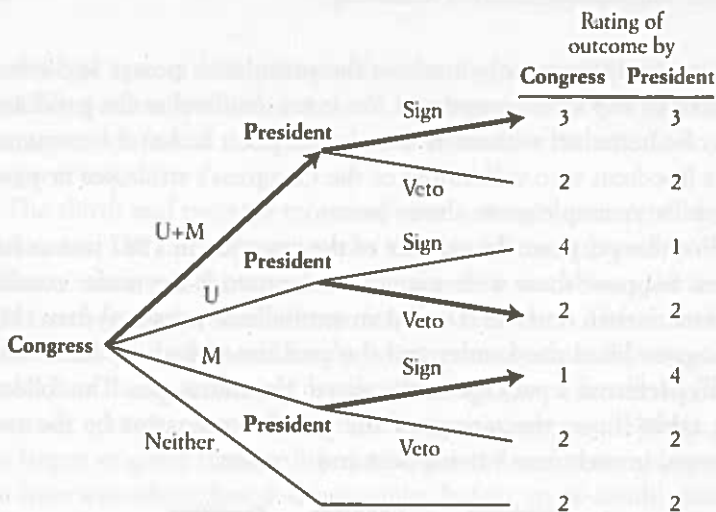
parts of a bill can only increase the president's power and never yield him any worse outcomes. Yet it is possible that the president may be better off without this tool. The point is that the existence of a line-item veto will influence the Congress's strategies in passing bills. A simple game shows how.

For this purpose, the essence of the situation in 1987 was as follows. Suppose there were two items of expenditure under consideration: urban renewal (U) and an antiballistic missile system (M). Congress liked the former and the president liked the latter. But both preferred a package of the two to the status quo. The following table shows the ratings of the possible scenarios by the two players, in each case 4 being best and 1, worst.

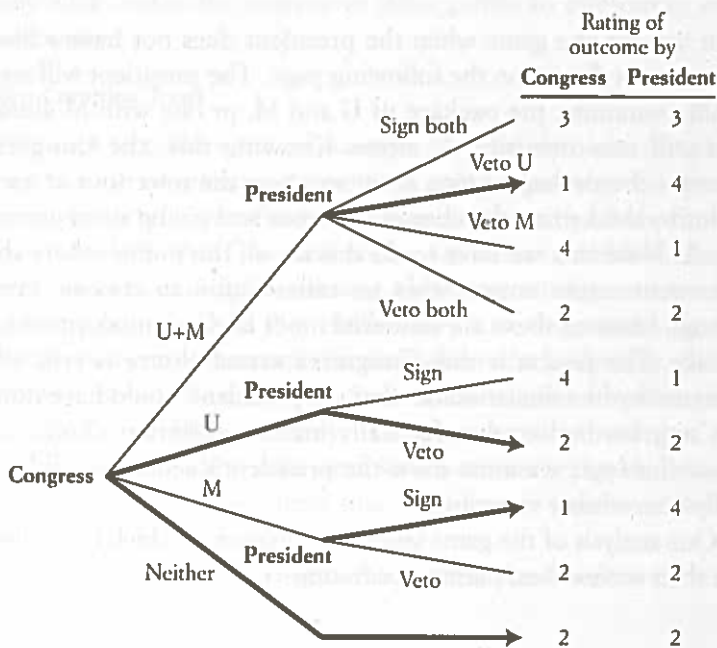
Outcomes	Congress	President
Both U and M	3	3
U only	4	1
M only	1	4
Neither	2	2

The tree for the game when the president does not have a line-item veto is shown on the following page. The president will sign a bill containing the package of U and M, or one with M alone, but will veto one with U alone. Knowing this, the Congress chooses the package. Once again we show the selections at each point by thickening the chosen branches and giving them arrowheads. Note that we have to do this for all the points where the president might conceivably be called upon to choose, even though some of these are rendered moot by Congress's previous choice. The reason is that Congress's actual choice is crucially affected by its calculation of what the president would have done if Congress had counterfactually made a different choice; to show this logic we must show the president's actions in all logically conceivable situations.

Our analysis of the game yields an outcome in which both sides get their second best preference (rating 3).



Next, suppose the president has a line-item veto. The game changes to the following:



Now Congress foresees that if it passes the package, the president will selectively veto U, leaving only M. Therefore Congress's best action is now either to pass U only to see it vetoed, or pass nothing. Perhaps it may have a preference for the former, if it can score political points from a presidential veto, but perhaps the president may equally score political points by this show of budgetary discipline. Let us suppose the two offset each other, and Congress is indifferent with respect to the two choices. But either gives each party only their third-best outcome (rating 2). Even the president is left worse-off by his extra freedom of choice.²

This game illustrates an important general conceptual point. In single-person decisions, greater freedom of action can never hurt. But in games, it can hurt because its existence can influence other players' actions. Conversely, tying your own hands can help. We will explore this "advantage of commitment" in chapters 6 and 7.

We have applied the method of backward reasoning in a game tree to a very trivial game (Charlie Brown), and extended it to a slightly more complicated game (the line-item veto). The general principle remains applicable, no matter how complicated the game may be. But trees for games where each player has several choices available at any point, and where each player gets several turns to move, can quickly get too complicated to draw or use. In chess, for example, 20 branches emerge from the root—the player with the white pieces can move any of his/her eight pawns forward one square or two, or move one of his two knights in one of two ways. For each of these, the player with the black pieces has 20 moves, so we are up to 400 distinct paths already. The number of branches emerging from later nodes in chess can be even larger. Solving chess fully using the tree method is beyond the ability of the most powerful computer that exists or might be developed during the next several decades, and other methods of partial analysis must be sought. We will discuss later in the chapter how chess experts have tackled this problem.

Between the two extremes lie many moderately complex games that are played in business, politics, and everyday life. Two approaches can be used for these. Computer programs are

available to construct trees and compute solutions.³ Alternatively, many games of moderate complexity can be solved by the logic of tree analysis, without drawing the tree explicitly. We illustrate this using a game that was played in a TV show that is all about games, where each player tries to “outplay, outwit, and outlast” the others.

STRATEGIES FOR “SURVIVORS”

CBS’s *Survivor* features many interesting games of strategy. In the sixth episode of *Survivor: Thailand*, the two teams or tribes played a game that provides an excellent example of thinking forward and reasoning backward in theory and in practice.⁴ Twenty-one flags were planted in the field of play between the tribes, who took turns removing the flags. Each tribe at its turn could choose to remove 1 or 2 or 3 flags. (Thus zero—passing up one’s turn—was not permitted; nor was it within the rules to remove four or more at one turn.) The team to take the last flag, whether standing alone or as a part of a group of 2 or 3 flags, won the game.⁵ The losing tribe had to vote out one of its own members, thus weakening it in future contests. In fact the loss proved crucial in this instance, and a member of the other tribe went on to win the ultimate prize of a million dollars. Thus the ability to figure out the correct strategy for this game would prove to be of great value.

The two tribes were named Sook Jai and Chuay Gahn, and Sook Jai had the first move. They started by taking 2 flags and leaving 19. Before reading on, pause a minute and think. If you were in their place, how many would you have chosen?

Write down your choice somewhere, and read on. To understand how the game should be played, and compare the correct strategy with how the two tribes actually played, it helps to focus on two very revealing incidents. First, each tribe had a few minutes to discuss the game among its own members before the play started. During this discussion within Chuay Gahn, one of the members, Ted Rogers, an African American software developer, pointed out, “At the end, we must leave them with four flags.” This is correct: if Sook Jai faces 4 flags, it must take 1 or 2 or 3,

leaving Chuay Gahn to take the remaining 3 or 2 or 1, respectively, at its next turn and win the game. Chuay Gahn did in fact get and exploit this opportunity correctly; facing 6 flags, they took 2.

But here is the other revealing incident. At the previous turn, just as Sook Jai returned from having taken 3 flags out of the 9 facing them, the realization hit one of their members, Shii Ann, a feisty and articulate competitor who took considerable pride in her analytical skills: "If Chuay Gahn now takes two, we are sunk." So Sook Jai's just-completed move was wrong. What should they have done?

Shii Ann or one of her Sook Jai colleagues should have reasoned as Ted Rogers did but carried the logic of leaving the other tribe with 4 flags to its next step. How do you ensure leaving the other tribe with 4 flags at its next turn? By leaving it with 8 flags at its previous turn. When it takes 1 or 2 or 3 out of eight, you take 3 or 2 or 1 at your next turn, leaving them with 4 as planned. Therefore Sook Jai should have turned the tables on Chuay Gahn and taken just 1 flag out of the 9. Shii Ann's analytical skill kicked into high gear one move too late! Ted Rogers perhaps had the better analytical insights. But did he?

How did Sook Jai come to face 9 flags at its previous move? Because Chuay Gahn had taken 2 from 11 at *its* previous turn. Ted Rogers should have carried his own reasoning one step further. Chuay Gahn should have taken 3, leaving Sook Jai with 8, which would be a losing position.

The same reasoning can be carried even farther back. To leave the other tribe with 8 flags, you must leave them with 12 at their previous turn; for that you must leave them with 16 at the turn before that and 20 at the turn before that. So Sook Jai should have started the game by taking just 1 flag, not 2 as it actually did. Then it could have had a sure win by leaving Chuay Gahn with 20, 16, . . . 4 at their successive turns.*

* Does the first mover always have a sure win in all games? No. If the flags game started with 20 flags instead of 21, the second mover would have a sure win. And in some games, for example the simple 3-by-3 tic-tac-toe, either player can ensure a tie with correct play.

Now think of Chuay Gahn's very first turn. It faced 19 flags. If it had carried its own logic back far enough, it would have taken 3, leaving Sook Jai with 16 and already on the way to certain defeat. Starting from any point in the middle of the game where the opponent has played incorrectly, the team with the turn to move can seize the initiative and win. But Chuay Gahn did not play the game perfectly either.*

The table below shows the comparison between the actual and the correct moves at each point in the game. (The entry "No move" means that all moves are losing moves if the opponent plays correctly.) You can see that almost all the choices were wrong, except Chuay Gahn's move when facing 13 flags, and that must have been accidental, because at their next turn they faced 11 and took 2 when they should have taken 3.

Tribe	No. of flags before move	No. of flags taken	Move to put team on path to sure victory
Sook Jai	21	2	1
Chuay Gahn	19	2	3
Sook Jai	17	2	1
Chuay Gahn	15	1	3
Sook Jai	14	1	2
Chuay Gahn	13	1	1
Sook Jai	12	1	No move
Chuay Gahn	11	2	3
Sook Jai	9	3	1
Chuay Gahn	6	2	2
Sook Jai	4	3	No move
Chuay Gahn	1	1	1

* The fates of the two key people were also interesting. Shii Ann made another key miscalculation in the next episode and was voted out, at number 10 among the 16 contestants who started the game. Ted, more quiet but perhaps somewhat more skillful, made it to the last five.

Before you judge the tribes harshly, you should recognize that it takes time and some experience to learn how to play even very simple games. We have played this game between pairs or teams of

students in our classes and found that it takes Ivy League freshmen three or even four plays before they figure out the complete reasoning and play correctly all the way through from the first move. (By the way, what number did you choose when we asked you to initially, and what was your reasoning?) Incidentally, people seem to learn faster by watching others play than by playing themselves; perhaps the perspective of an observer is more conducive to seeing the game as a whole and reasoning about it coolly than that of a participant.

To fix your understanding of the logic of the reasoning, we offer you the first of our Trips to the Gym—questions on which you can exercise and hone your developing skills in strategic thinking. The answers are in the Workouts section in the end of the book.

Now that you are invigorated by this exercise, let us proceed to think about some general issues of strategy in this whole class of games.

WHAT MAKES A GAME FULLY SOLVABLE BY BACKWARD REASONING?

The 21-flags game had a special property that made it fully solvable, namely the absence of uncertainty of any kind: whether about some natural chance elements, the other players' motives and capabilities, or their actual actions. This seems a simple point to make, but it needs some elaboration and clarification.

First, at any point in the game when one tribe had the move, it knew exactly what the situation was, namely how many flags remained. In many games there are elements of pure chance, thrown up by nature or by the gods of probability. For example, in many card games, when a player makes a choice, he/she does not

TRIP TO THE GYM NO. 1

Let us turn the flag game into hot potato: now you win by forcing the other team to take the last flag. It's your move and there are 21 flags. How many do you take?

know for sure what cards the other players hold, although their previous actions may give some basis for drawing some inferences about that. In many subsequent chapters, our examples and analysis will involve games that have this natural element of chance.

Secondly, the tribe making its choice also knew the other tribe's objective, namely to win. And Charlie Brown should have known that Lucy enjoyed seeing him fall flat on his back. Players have such perfect knowledge of the other player's or players' objectives in many simple games and sports, but that is not necessarily the case in games people play in business, politics, and social interactions. Motives in such games are complex combinations of selfishness and altruism, concern for justice or fairness, short-run and long-run considerations, and so on. To figure out what the other players will choose at future points in the game, you need to know what their objectives are and, in the case of multiple objectives, how they will trade one off against the other. You can almost never know this for sure and must make educated guesses. You must not assume that other people will have the same preferences as you do, or as a hypothetical "rational person" does, but must genuinely think about their situation. Putting yourself in the other person's shoes is a difficult task, often made more complicated by your emotional involvement in your own aims and pursuits. We will have more to say about this kind of uncertainty later in this chapter and at various points throughout the book. Here we merely point out that the uncertainty about other players' motives is an issue for which it may be useful to seek advice from an objective third party—a strategic consultant.

Finally, players in many games must face uncertainty about other players' choices; this is sometimes called strategic uncertainty to distinguish it from the natural aspects of chance, such as a distribution of cards or the bounce of a ball from an uneven surface. In 21-flats there was no strategic uncertainty, because each tribe saw and knew exactly what the other had done previously. But in many games, players take their actions simultaneously or in such rapid sequence that one cannot see what the other has done and react to it. A soccer goalie facing a penalty kick must decide

whether to move to his/her own right or left without knowing which direction the shooter will aim for; a good shooter will conceal his/her own intentions up to the last microsecond, by which time it is too late for the goalie to react. The same is true for serves and passing shots in tennis and many other sports. Each participant in a sealed-bid auction must make his/her own choice without knowing what the other bidders are choosing. In other words, in many games the players make their moves simultaneously, and not in a preassigned sequence. The kind of thinking that is needed for choosing one's action in such games is different from, and in some respects harder than, the pure backward reasoning of sequential-move games like 21-flags; each player must be aware of the fact that others are making conscious choices and are in turn thinking about what he himself is thinking, and so on. The games we consider in the next several chapters will elucidate the reasoning and solution tools for simultaneous-move games. In this tools chapter, however, we focus solely on sequential-move games, as exemplified by 21-flags and, at a much higher level of complexity, chess.

Do People Actually Solve Games by Backward Reasoning?

Backward reasoning along a tree is the correct way to analyze and solve games where the players move sequentially. Those who fail to do so either explicitly or intuitively are harming their own objectives; they should read our book or hire a strategic consultant. But that is an advisory or normative use of the theory of backward reasoning. Does the theory have the usual explanatory or positive value that most scientific theories do? In other words, do we observe the correct outcomes from the play of actual games? Researchers in the new and exciting fields of behavioral economics and behavioral game theory have conducted experiments that yield mixed evidence.

What seems to be the most damaging criticism comes from the ultimatum game. This is the simplest possible negotiation game: there is just one take-it-or-leave-it offer. The ultimatum game has two players, a "proposer," say A, a "responder," say B, and a sum

A QUICK TRIP TO THE GYM: REVERSE ULTIMATUM GAME

In this variant of the ultimatum game, A makes an offer to B about how to divide up the 100 dollars. If B says yes, the money is divided up and the game is over. But if B says no, then A must decide whether to make another offer or not. Each subsequent offer from A must be more generous to B. The game ends when either B says yes or A stops making offers. How do you predict this game will end up?

In this case, we can suppose that A will keep on making offers until he has proposed 99 to B and 1 for himself. Thus, according to tree logic, B should get almost all of the pie. If you were B, would you hold out for 99:1? We'd advise against it.

of money, say 100 dollars. Player A begins the game by proposing a division of the 100 dollars between the two. Then B decides whether to agree to A's proposal. If B agrees, the proposal is implemented; each player gets what A proposed and the game ends. If B refuses, then neither player gets anything, and the game ends.

Pause a minute and think. If you were playing this game in the A role, what division would you propose?

Now think how this game would be played by two people who are "rational" from the point of view of conventional economic theory—

that is, each is concerned only with his or her self-interest and can calculate perfectly the optimal strategies to pursue that interest. The proposer (A) would think as follows. "No matter what split I propose, B is left with the choice between that and nothing. (The game is played only once, so B has no reason to develop a reputation for toughness, or to engage in any tit-for-tat response to A's actions.) So B will accept whatever I offer. I can do best for myself by offering B as little as possible—for example, just one cent, if that is the minimum permissible under the rules of the game." Therefore A would offer this minimum and B would accept.*

Pause and think again. If you were playing this game in the B role, would you accept one cent?

Numerous experiments have been conducted on this game.⁶ Typically, two dozen or so subjects are brought together and are matched randomly in pairs. In each pair, the roles of proposer and

* This argument is another example of tree logic without drawing a tree.

responder are assigned, and the game is played once. New pairs are formed at random, and the game played again. Usually the players do not know with whom they are matched in any one play of the game. Thus the experimenter gets several observations from the same pool in the same session, but there is no possibility of forming ongoing relationships that can affect behavior. Within this general framework, many variations of conditions are attempted, to study their effects on the outcomes.

Your own introspection of how you would act as proposer and as responder has probably led you to believe that the results of actual play of this game should differ from the theoretical prediction above. And indeed they differ, often dramatically so. The amounts offered to the responder differ across proposers, but one cent or one dollar, or in fact anything below 10 percent of the total sum at stake, is very rare. The median offer (half of the proposers offer less than that and half offer more) is in the 40–50 percent range; in many experiments a 50:50 split is the single most frequent proposal. Proposals that would give the responder less than 20 percent are rejected about half the time.

IRRATIONALITY VERSUS OTHER-REGARDING RATIONALITY

Why do proposers offer substantial shares to the responders? Three reasons suggest themselves. First, the proposers may be unable to do the correct backward reasoning. Second, the proposers may have motives other than the pure selfish desire to get as much as they can; they act altruistically or care about fairness. Third, they may fear that responders would reject low offers.

The first is unlikely because the logic of backward reasoning is so simple in this game. In more complex situations, players may fail to do the necessary calculations fully or correctly, especially if they are novices to the game being played, as we saw in 21-flags. But the ultimatum game is surely simple enough, even for novices. The explanation must be the second, the third, or a combination thereof.

responder are assigned, and the game is played once. New pairs are formed at random, and the game played again. Usually the players do not know with whom they are matched in any one play of the game. Thus the experimenter gets several observations from the same pool in the same session, but there is no possibility of forming ongoing relationships that can affect behavior. Within this general framework, many variations of conditions are attempted, to study their effects on the outcomes.

Your own introspection of how you would act as proposer and as responder has probably led you to believe that the results of actual play of this game should differ from the theoretical prediction above. And indeed they differ, often dramatically so. The amounts offered to the responder differ across proposers, but one cent or one dollar, or in fact anything below 10 percent of the total sum at stake, is very rare. The median offer (half of the proposers offer less than that and half offer more) is in the 40–50 percent range; in many experiments a 50:50 split is the single most frequent proposal. Proposals that would give the responder less than 20 percent are rejected about half the time.

IRRATIONALITY VERSUS OTHER-REGARDING RATIONALITY

Why do proposers offer substantial shares to the responders? Three reasons suggest themselves. First, the proposers may be unable to do the correct backward reasoning. Second, the proposers may have motives other than the pure selfish desire to get as much as they can; they act altruistically or care about fairness. Third, they may fear that responders would reject low offers.

The first is unlikely because the logic of backward reasoning is so simple in this game. In more complex situations, players may fail to do the necessary calculations fully or correctly, especially if they are novices to the game being played, as we saw in 21-flags. But the ultimatum game is surely simple enough, even for novices. The explanation must be the second, the third, or a combination thereof.

Early results from ultimatum experiments favored the third. In fact, Harvard's Al Roth and his coauthors found that, given the pattern of rejection thresholds that prevailed in their subject pool, the proposers were choosing their offers to achieve an optimal balance between the prospect of obtaining a greater share for themselves against the risk of rejection. This suggests a remarkable conventional rationality on part of the proposers.

However, later work to distinguish the second and the third possibilities led to a different idea. To distinguish between altruism and strategy, experiments were done using a variant called the dictator game. Here the proposer dictates how the available total is to be split; the other player has no say in the matter at all. Proposers in the dictator game give away significantly smaller sums on average than they offer in the ultimatum game, but they give away substantially more than zero. Thus there is something to both of those explanations; proposers' behavior in the ultimatum game has both generous and strategic aspects.

Is the generosity driven by altruism or by a concern for fairness? Both explanations are different aspects of what might be called a regard for others in people's preferences. Another variation of the experiment helps tell these two possibilities apart. In the usual setup, after the pairs are formed, the roles of proposer and responder are assigned by a random mechanism like a coin toss. This may build in a notion of equality or fairness in the players' minds. To remove this, a variant assigns the roles by holding a preliminary contest, such as a test of general knowledge, and making its winner the proposer. This creates some sense of entitlement to the proposer, and indeed leads to offers that are on average about 10 percent smaller. However, the offers remain substantially above zero, indicating that proposers have an element of altruism in their thinking. Remember that they do not know the identity of the responders, so this must be a generalized sense of altruism, not concern for the well-being of a particular person.

A third variation of individual preferences is also possible: contributions may be driven by a sense of shame. Jason Dana of the

University of Illinois, Daylian Cain of Yale School of Management, and Robyn Dawes of Carnegie-Mellon University performed an experiment with the following variation of the dictator game.⁷ The dictator is asked to allocate \$10. After the allocation is made, but before it is delivered to the other party, the dictator is given the following offer: You can have \$9, the other party will get nothing, and they will never know that they were part of this experiment. Most dictators accept this offer. Thus they would rather give up a dollar to ensure that the other person never knows how greedy they were. (An altruistic person would prefer keeping \$9 and giving away \$1 to keeping \$9 while the other person gets nothing.) Even when a dictator had offered \$3, he would rather take that away to keep the other person in the dark. This is much like incurring a large cost to cross the street to avoid making a small donation to a beggar.

Observe two things about these experiments. First, they follow the standard methodology of science: hypotheses are tested by designing appropriate variations of controls in the experiment. We mention a few prominent variations of this kind here. (Many more are discussed in Colin Camerer's book cited in chapter 2, note 6.) Second, in the social sciences, multiple causes often coexist, each contributing part of the explanation for the same phenomenon. Hypotheses don't have to be either fully correct or totally wrong; accepting one need not mean rejecting all others.

Now consider the behavior of the responders. Why do they reject an offer when they know that the alternative is to get even less? The reason cannot be to establish a reputation for being a tough negotiator that may bear fruit in future plays of this game or other games of division. The same pair does not play repeatedly, and no track record of one player's past behavior is made available to future partners. Even if a reputational motive is implicitly present, it must take a deeper form: a general rule for action that the responder follows without doing any explicit thinking or calculation in each instance. It must be an instinctive action or an emotion-driven response. And that is indeed the case. In a new emerging line of experimental research called neuroeconomics,

the subjects' brain activity is scanned using functional magnetic resonance imaging (fMRI) or positron emission tomography (PET) while they make various economic decisions. When ultimatum games are played under such conditions, it is found that the responders' anterior insula shows more activity as the proposers' offers become more unequal. Since the anterior insula is active for emotions, such as anger and disgust, this result helps explain why second movers reject unequal offers. Conversely, the left-side prefrontal cortex is more active when an unequal offer is accepted, indicating that conscious control is being exercised to balance between acting on one's disgust and getting more money.⁸

Many people (especially economists) argue that while responders may reject small shares of the small sums that are typically on offer in laboratory experiments, in the real world, where stakes are often much larger, rejection must be very unlikely. To test this, ultimatum game experiments have been conducted in poorer countries where the amounts were worth several months' income for the participants. Rejection does become somewhat less likely, but offers do not become significantly less generous. The consequences of rejection become more serious for the proposers just as they do for the responders, so proposers fearing rejection are likely to behave more cautiously.

Although behavior can be explained in part by instincts, hormones, or emotions hardwired into the brain, part of it varies from one culture to another. In experiments conducted across many countries, it was found that the perception of what constitutes a reasonable offer varied by up to 10 percent across cultures, but properties like aggressiveness or toughness varied less. Only one group differed substantially from the rest: among the Machiguenga of the Peruvian Amazon, the offers were much smaller (average 26 percent) and only one offer was rejected. Anthropologists explain that the Machiguenga live in small family units, are socially disconnected, and have no norms of sharing. Conversely, in two cultures the offers exceeded 50 percent; these have the custom of lavish giving when one has a stroke of good luck, which places an obligation on the recipients to return the favor even

more generously in the future. This norm or habit seems to carry over to the experiment even though the players do not know whom they are giving to or receiving from.⁹

Evolution of Altruism and Fairness

What should we learn from the findings of these experiments on the ultimatum game, and others like them? Many of the outcomes do differ significantly from what we would expect based on the theory of backward reasoning with the assumption that each player cares only about his or her own reward. Which of the two—correct backward calculation or selfishness—is the wrong assumption, or is it a combination? And what are the implications?

Consider backward reasoning first. We saw the players on *Survivor* fail to do this correctly or fully in 21-flags. But they were playing the game for the first time, and even then, their discussion revealed glimpses of the correct reasoning. Our classroom experience shows that students learn the full strategy after playing the game, or watching it played, just three or four times. Many experiments inevitably and almost deliberately work with novice subjects, whose actions in the game are often steps in the process of learning the game. In the real world of business, politics, and professional sports, where people are experienced at playing the games they are involved in, we should expect that the players have accumulated much more learning and that they play generally good strategies either by calculation or by trained instinct. For somewhat more complex games, strategically aware players can use computers or consultants to do the calculations; this practice is still somewhat rare but is sure to spread. Therefore, we believe that backward reasoning should remain our starting point for analysis of such games and for predicting their outcomes. This first pass at the analysis can then be modified as necessary in a particular context, to recognize that beginners may make mistakes and that some games may become too complex to be solved unaided.

We believe that the more important lesson from the experimental research is that people bring many considerations and

preferences into their choices besides their own rewards. This takes us beyond the scope of conventional economic theory. Game theorists should include in their analysis of games the players' concerns for fairness or altruism. "Behavioral game theory *extends* rationality rather than abandoning it."¹⁰

This is all to the good; a better understanding of people's motives enriches our understanding of economic decision making and strategic interactions alike. And that is already happening; frontier research in game theory increasingly includes in the players' objectives their concerns for equity, altruism, and similar concerns (and even a "second-round" concern to reward or punish others whose behavior reflects or violates these precepts).¹¹

But we should not stop there; we should go one step further and think about why concerns for altruism and fairness, and anger or disgust when someone else violates these precepts, have such a strong hold on people. This takes us into the realm of speculation, but one plausible explanation can be found in evolutionary psychology. Groups that instill norms of fairness and altruism into their members will have less internal conflict than groups consisting of purely selfish individuals. Therefore they will be more successful in taking collective action, such as provision of goods that benefit the whole group and conservation of common resources, and they will spend less effort and resources in internal conflict. As a result, they will do better, both in absolute terms and in competition with groups that do not have similar norms. In other words, some measure of fairness and altruism may have evolutionary survival value.

Some biological evidence for rejecting unfair offers comes from an experiment run by Terry Burnham.¹² In his version of the ultimatum game, the amount at stake was \$40 and the subjects were male Harvard graduate students. The divider was given only two choices: offer \$25 and keep \$15 or offer \$5 and keep \$35. Among those offered only \$5, twenty students accepted and six rejected, giving themselves and the divider both zero. Now for the punch line. It turns out that the six who rejected the offer had testosterone levels 50 percent higher than those who accepted the offer.

To the extent that testosterone is connected with status and aggression, this could provide a genetic link that might explain an evolutionary advantage of what evolutionary biologist Robert Trivers has called "moralistic aggression."

In addition to a potential genetic link, societies have nongenetic ways of passing on norms, namely the processes of education and socialization of infants and children in families and schools. We see parents and teachers telling impressionable children the importance of caring for others, sharing, and being nice; some of this undoubtedly remains imprinted in their minds and influences their behavior throughout their lives.

Finally, we should point out that fairness and altruism have their limit. Long-run progress and success of a society need innovation and change. These in turn require individualism and a willingness to defy social norms and conventional wisdom; selfishness often accompanies these characteristics. We need the right balance between self-regarding and other-regarding behaviors.

VERY COMPLEX TREES

When you have acquired a little experience with backward reasoning, you will find that many strategic situations in everyday life or work lend themselves to "tree logic" without the need to draw and analyze trees explicitly. Many other games at an intermediate level of complexity can be solved using computer software packages that are increasingly available for this purpose. But for complex games such as chess, a complete solution by backward reasoning is simply not feasible.

In principle, chess is the ideal game of sequential moves amenable to solution by backward reasoning.¹³ The players alternate moves; all previous moves are observable and irrevocable; there is no uncertainty about the position or the players' motives. The rule that the game is a draw if the same position is repeated ensures that the game ends within a finite total number of moves. We can start with the terminal nodes (or endpoints) and work backward. However, practice and principle are two different

things. It has been estimated that the total number of nodes in chess is about 10^{120} , that is, 1 with 120 zeroes after it. A supercomputer 1,000 times as fast as the typical PC would need 10^{103} years to examine them all. Waiting for that is futile; foreseeable progress in computers is not likely to improve matters significantly. In the meantime, what have chess players and programmers of chess-playing computers done?

Chess experts have been successful at characterizing optimal strategies near the end of the game. Once the chessboard has only a small number of pieces on it, experts are able to look ahead to the end of the game and determine by backward reasoning whether one side has a guaranteed win or whether the other side can obtain a draw. But the middle of the game, when several pieces remain on the board, is far harder. Looking ahead five pairs of moves, which is about as much as can be done by experts in a reasonable amount of time, is not going to simplify the situation to a point where the endgame can be solved completely from there on.

The pragmatic solution is a combination of forward-looking analysis and value judgment. The former is the science of game theory—looking ahead and reasoning backward. The latter is the art of the practitioner—being able to judge the value of a position from the number and interconnections of the pieces without finding an explicit solution of the game from that point onward. Chess players often speak of this as “knowledge,” but you can call it experience or instinct or art. The best chess players are usually distinguished by the depth and subtlety of their knowledge.

Knowledge can be distilled from the observation of many games and many players and then codified into rules. This has been done most extensively with regard to openings, that is, the first ten or even fifteen moves of a game. There are hundreds and hundreds of books that analyze different openings and discuss their relative merits and drawbacks.

How do computers fit into this picture? At one time, the project of programming computers to play chess was seen as an integral part of the emerging science of artificial intelligence; the aim was

to design computers that would think as humans do. This did not succeed for many years. Then the attention shifted to using computers to do what they do best, crunch numbers. Computers can look ahead to more moves and do this more quickly than humans can.* Using pure number crunching, by the late 1990s dedicated chess computers like Fritz and Deep Blue could compete with the top human players. More recently, computers have been programmed with some knowledge of midgame positions, imparted by some of the best human players.

Human players have ratings determined by their performances; the best-ranked computers are already achieving ratings comparable to the 2800 enjoyed by the world's strongest human player, Garry Kasparov. In November 2003, Kasparov played a four-game match against the latest version of the Fritz computer, X3D. The result was one victory each and two draws. In July 2005, the Hydra chess computer demolished Michael Adams, ranked number 13 in the world, winning five games and drawing one in a six-game match. It may not be long before the rival computers rank at the top and play each other for world championships.

What should you take away from this account of chess? It shows the method for thinking about any highly complex games you may face. You should combine the rule of look ahead and reason back with your experience, which guides you in evaluating the intermediate positions reached at the end of your span of forward calculation. Success will come from such synthesis of the science of game theory and the art of playing a specific game, not from either alone.

BEING OF TWO MINDS

Chess strategy illustrates another important practical feature of looking forward and reasoning backward: you have to play the

* But good chess players can use their knowledge to disregard immediately those moves that are likely to be bad without pursuing their consequences four or five moves ahead, thereby saving their calculation time and effort for the moves that are more likely to be good ones.

game from the perspective of both players. While it is hard to calculate your best move in a complicated tree, it is even harder to predict what the other side will do.

If you really could analyze all possible moves and countermoves, and the other player could as well, then the two of you would agree up front as to how the entire game would play out. But once the analysis is limited to looking down only some branches of the tree, the other player may see something you didn't or miss something you've seen. Either way, the other side may then make a move you didn't anticipate.

To really look forward and reason backward, you have to predict what the other players will actually do, not what you would have done in their shoes. The problem is that when you try to put yourself in the other players' shoes, it is hard if not impossible to leave your own shoes behind. You know too much about what you are planning to do in your next move and it is hard to erase that knowledge when you are looking at the game from the other player's perspective. Indeed, that explains why people don't play chess (or poker) against themselves. You certainly can't bluff against yourself or make a surprise attack.

There is no perfect solution to this problem. When you try to put yourself in the other players' shoes, you have to know what they know and not know what they don't know. Your objectives have to be their objectives, not what you wish they had as an objective. In practice, firms trying to simulate the moves and countermoves of a potential business scenario will hire outsiders to play the role of the other players. That way, they can ensure that their game partners don't know too much. Often the biggest learning comes from seeing the moves that were not anticipated and then understanding what led to that outcome, so that it can be either avoided or promoted.

To end this chapter, we return to Charlie Brown's problem of whether or not to kick the football. This question became a real issue for football coach Tom Osborne in the final minutes of his championship game. We think he too got it wrong. Backward reasoning will reveal the mistake.

**CASE STUDY: THE TALE OF TOM OSBORNE
AND THE '84 ORANGE BOWL**

In the 1984 Orange Bowl the undefeated Nebraska Cornhuskers and the once-beaten Miami Hurricanes faced off. Because Nebraska came into the Bowl with the better record, it needed only a tie in order to finish the season with the number-one ranking.

Coming into the fourth quarter, Nebraska was behind 31-17. Then the Cornhuskers began a comeback. They scored a touchdown to make the score 31-23. Nebraska coach Tom Osborne had an important strategic decision to make.

In college football, a team that scores a touchdown then runs one play from a hash mark 2 1/2 yards from the goal line. The team has a choice between trying to run (or pass) the ball into the end zone, which scores two additional points, or trying the less risky strategy of kicking the ball through the goalposts, which scores one extra point.

Coach Osborne chose to play it safe, and Nebraska successfully kicked for the one extra point. Now the score was 31-24. The Cornhuskers continued their comeback. In the waning minutes of the game they scored a final touchdown, bringing the score to 31-30. A one-point conversion would have tied the game and landed them the title. But that would have been an unsatisfying victory. To win the championship with style, Osborne recognized that he had to go for the win.

The Cornhuskers went for the win with a two-point conversion attempt. Irving Fryar got the ball but failed to score. Miami and Nebraska ended the year with equal records. Since Miami beat Nebraska, it was Miami that was awarded the top place in the standings.

Put yourself in the cleats of Coach Osborne. Could you have done better?

Case Discussion

Many Monday morning quarterbacks fault Osborne for going for the win rather than the tie. But that is not our bone of

contention. Given that Osborne was willing to take the additional risk for the win, he did it the wrong way. He would have done better to first try the two-point conversion. If it succeeded, then go for the one-point; if it failed, attempt a second two-pointer.

Let us look at this more carefully. When down by 14 points, he knew that he needed two touchdowns plus three extra points. He chose to go for the one-point and then the two. If both attempts succeeded, the order in which they were made becomes irrelevant. If the one-point conversion was missed but the two-point was successful, here too the order is irrelevant and the game ends up tied with Nebraska getting the championship. The only difference occurs if Nebraska misses the two-point conversion. Under Osborne's plan, that results in the loss of the game and the championship. If, instead, they had tried the two-point conversion first, then if it failed they would not necessarily have lost the game. They would have been behind 31-23. When they scored their next touchdown this would have brought them to 31-29. A successful two-point attempt would tie the game and win the number-one ranking!*

We have heard the counterargument that if Osborne went for the two-pointer first and missed, his team would have been playing for the tie. This would have provided less inspiration and perhaps they might not have scored the second touchdown. Moreover, by waiting until the end and going for the desperation win-lose two-pointer his team would rise to the occasion knowing everything was on the line. This argument is wrong for several reasons. Remember that if Nebraska waits until the second touchdown and then misses the two-point attempt, they lose. If they miss the two-point attempt on their first try, there is still a chance for a tie. Even though the chance may be diminished, something is better than nothing. The momentum argument is also flawed. While Nebraska's offense may rise to the occasion in a single play for the championship, we expect the Hurricanes' defense to rise as well.

* Furthermore, this would be a tie that resulted from the failed attempt to win, so no one would criticize Osborne for playing to tie.

The play is important for both sides. To the extent that there is a momentum effect, if Osborne makes the two-point conversion on the first touchdown, this should increase the chance of scoring another touchdown. It also allows him to tie the game with two field goals.

One of the general morals of this story is that if you have to take some risks, it is often better to do so as quickly as possible. This is obvious to those who play tennis: everyone knows to take more risk on the first serve and hit the second serve more cautiously. That way, if you fail on your first attempt, the game won't be over. You may still have time to take some other options that can bring you back to or even ahead of where you were. The wisdom of taking risks early applies to most aspects of life, whether it be career choices, investments, or dating.

For more practice using the principle of look forward, reason backward, have a look at the following case studies in chapter 14: "Here's Mud in Your Eye"; "Red I Win, Black You Lose"; "The Shark Repellent That Backfired"; "Tough Guy, Tender Offer"; "The Three-Way Duel"; and "Winning without Knowing How."

Prisoners' Dilemmas and How to Resolve Them

MANY CONTEXTS, ONE CONCEPT

What do the following situations have in common?

- Two gas stations at the same corner, or two supermarkets in the same neighborhood, sometimes get into fierce price wars with each other.
- In general election campaigns, both the Democratic and the Republican parties in the United States often adopt centrist policies to attract the swing voters in the middle of the political spectrum, ignoring their core supporters who hold more extreme views to the left and the right, respectively.
- "The diversity and productivity of New England fisheries was once unequalled. A continuing trend over the past century has been the overexploitation and eventual collapse of species after species. Atlantic halibut, ocean perch, Haddock and Yellowtail Flounder . . . [have joined] the ranks of species written-off as commercially extinct."¹
- Near the end of Joseph Heller's celebrated novel *Catch-22*, the Second World War is almost won. Yossarian does not want to be among the last to die; it won't make any difference to the outcome. He explains this to Major Danby, his superior officer.

When Danby asks, "But, Yossarian, suppose everyone felt that way?" Yossarian replies, "Then I'd certainly be a damned fool to feel any other way, wouldn't I?"²

Answer: They are all instances of the prisoners' dilemma.* As in the interrogation of Dick Hickock and Perry Smith from *In Cold Blood* recounted in chapter 1, each has a personal incentive to do something that ultimately leads to a result that is bad for everyone when everyone similarly does what his or her personal interest dictates. If one confesses, the other had better confess to avoid the really harsh sentence reserved for recalcitrants; if one holds out, the other can cut himself a much better deal by confessing. Indeed, the force is so strong that each prisoner's temptation to confess exists regardless of whether the two are guilty (as was the case in *In Cold Blood*) or innocent and being framed by the police (as in the movie *L.A. Confidential*).

Price wars are no different. If the Nexon gas station charges a low price, the Lunaco station had better set its own price low to avoid losing too many customers; if Nexon prices its gas high, Lunaco can divert many customers its way by pricing low. But when both stations price low, neither makes money (though customers are better off).

If the Democrats adopt a platform that appeals to the middle, the Republicans may stand to lose all these voters and therefore the election if they cater only to their core supporters in the economic and social right wings; if the Democrats cater to their core supporters in the minorities and the unions, then the Republicans can capture the middle and therefore win a large majority by being more centrist. If all others fish conservatively, one fisherman going

* No prizes for correct answers—after all, the prisoners' dilemma is the subject of this chapter. But we take this opportunity to point out, as we did in chapter 2, that the common conceptual framework of game theory can help us understand a great variety of diverse and seemingly unrelated phenomena. We should also point out that neighboring stores do not *constantly* engage in price wars, and political parties do not *always* gravitate to the center. In fact, analyses and illustrations of how the participants in such games can avoid or resolve the dilemma is an important part of this chapter.

for a bigger catch is not going to deplete the fishery to any significant extent; if all others are fishing aggressively, then any single fisherman would be a fool to try single-handed conservation.³ The result is overfishing and extinction. Yossarian's logic is what makes it so difficult to continue to support a failed war.

A LITTLE HISTORY

How did theorists devise and name this game that captures so many economic, political, and social interactions? It happened very early in the history of the subject. Harold Kuhn, himself one of the pioneers of game theory, recounted the story in a symposium held in conjunction with the 1994 Nobel Prize award ceremonies:

Al Tucker was on leave at Stanford in the Spring of 1950 and, because of the shortage of offices, he was housed in the Psychology Department. One day a psychologist knocked on his door and asked what he was doing. Tucker replied: "I'm working on game theory," and the psychologist asked if he would give a seminar on his work. For that seminar, Al Tucker invented prisoner's dilemma as an example of game theory, Nash equilibria, and the attendant paradoxes of non-socially-desirable equilibria. A truly seminal example, it inspired dozens of research papers and several entire books.⁴

Others tell a slightly different story. According to them, the mathematical structure of the game predates Tucker and can be attributed to two mathematicians, Merrill Flood and Melvin Dresher, at the Rand Corporation (a cold war think tank).⁵ Tucker's genius was to invent the story illustrating the mathematics. And genius it was, because presentation can make or break an idea; a memorable presentation spreads and is assimilated in the community of thinkers far better and faster, whereas a dull and dry presentation may be overlooked or forgotten.

A Visual Representation

We will develop the method for displaying and solving the game using a business example. Rainbow's End and B. B. Lean are rival

mail-order firms that sell clothes. Every fall they print and mail their winter catalogs. Each firm must honor the prices printed in its catalog for the whole winter season. The preparation time for the catalogs is much longer than the mailing window, so the two firms must make their pricing decisions simultaneously and without knowing the other firm's choices. They know that the catalogs go to a common pool of potential customers, who are smart shoppers and are looking for low prices.

Both catalogs usually feature an almost identical item, say a chambray deluxe shirt. The cost of each shirt to each firm is \$20.* The firms have estimated that if they each charge \$80 for this item, each will sell 1,200 shirts, so each will make a profit of $(80 - 20) \times 1,200 = 72,000$ dollars. Moreover, it turns out that this price serves their joint interests best: if the firms can collude and charge a common price, \$80 is the price that will maximize their combined profits.

The firms have estimated that if one of them cuts its price by \$1 while the other holds its price unchanged, then the price cutter gains 100 customers, 80 of whom shift to it from the other firm, and 20 who are new—for example, they might decide to buy the shirt when they would not have at the higher price or might switch from a store in their local mall. Therefore each firm has the temptation to undercut the other to gain more customers; the whole purpose of this story is to figure out how these temptations play out.

We begin by supposing that each firm chooses between just two prices, \$80 and \$70.† If one firm cuts its price to \$70 while the

* This includes not only the cost of buying the shirt from the supplier in China but also the cost of transporting it to the United States, any import duties, and the costs of stocking it and of order fulfillment. In other words, it includes all costs specifically attributable to this item. The intention is to have a comprehensive measure of what economists would call marginal cost.

† This specification, and in particular the assumption that there are just two possible choices for the price, is just to build the analytical method for such games in the simplest possible way. In the following chapter we will allow the firms much greater freedom in choosing their prices.

other is still charging \$80, the price cutter gains 1,000 customers and the other loses 800. So the price cutter sells 2,200 shirts while the other's sales drop to 400; the profits are $(70 - 20) \times 2,200 = \$110,000$ for the price cutter, and $(80 - 20) \times 400 = \$24,000$ for the other firm.

What happens if both firms cut their price to \$70 at the same time? If both firms reduce their price by \$1, existing customers stay put, but each gains the 20 new customers. So when both reduce their price by \$10, each gains $10 \times 20 = 200$ net sales above the previous 1,200. Each sells 1,400 and makes a profit of $(70 - 20) \times 1,400 = \$70,000$.

We want to display the profit consequences (the firms' payoffs in their game) visually. However, we cannot do this using a game tree like the ones in chapter 2. Here the two players act simultaneously. Neither can make his move knowing what the other has done or anticipating how the other will respond. Instead, each must think about what the other is thinking at the same time. A starting point for this thinking about thinking is to lay out all the consequences of all the combinations of the simultaneous choices the two could make. Since each has two alternatives, \$80 or \$70, there are four such combinations. We can display them most easily in a spreadsheet-like format of rows and columns, which we will generally refer to as a game table or payoff table. The choices of Rainbow's End (RE for short) are arrayed along the rows, and those of B. B. Lean (BB) along the columns. In each of the four cells corresponding to each choice of a row by RE and of a column by BB, we show two numbers—the profit, in thousands of dollars, from selling this shirt. In each cell, the number in the southwest corner belongs to the row player, and the number in the northeast corner belongs to the column player.* In the jargon of game theory, these

* Thomas Schelling invented this way of representing both players' payoffs in the same table while making clear which payoff belongs to which player. With excessive modesty, he writes: "If I am ever asked whether I ever made a contribution to game theory, I shall answer yes . . . the invention of staggered payoffs in a matrix." Actually Schelling developed many of the most important concepts of game theory—focal points, credibility, commitment, threats and promises, tip-

numbers are called payoffs.* To make it abundantly clear which payoffs belong to which player, we have also put the numbers in two different shades of gray for this example.

		B. B. Lean (BB)	
		80	70
Rainbow's End (RE)	80	72,000	110,000
	70	24,000	70,000

Before we "solve" the game, let us observe and emphasize one feature of it. Compare the payoff pairs across the four cells. A better outcome for RE does not always imply a worse outcome for BB, or vice versa. Specifically, both of them are better off in the top left cell than in the bottom right cell. This game need not end with a winner and a loser; it is not zero-sum. We similarly pointed out in chapter 2 that the Charlie Brown investment game was not zero-sum, and neither are most games we meet in reality. In many games, as in the prisoners' dilemma, the issue will be how to avoid a lose-lose outcome or to achieve a win-win outcome.

The Dilemma

Now consider the reasoning of RE's manager. "If BB chooses \$80, I can get \$110,000 instead of \$72,000 by cutting my price to \$70. If BB chooses \$70, then my payoff is \$70,000 if I also charge

ping, and much more. We will cite him and his work frequently in the chapters to come.

* Generally, higher payoff numbers are better for each player. Sometimes, as with prisoners under interrogation, the payoff numbers are years in jail, so each player prefers a smaller number for himself. The same can happen if the payoff numbers are rankings where 1 is best. When looking at a game table, you should check the interpretation of the payoff numbers for that game.

\$70, but only \$24,000 if I charge \$80. So, in both cases, choosing \$70 is better than choosing \$80. My better choice (in fact my best choice, since I have only two alternatives) is the same no matter what BB chooses. I don't need to think through their thinking at all; I should just go ahead and set my price at \$70."

When a simultaneous-move game has this special feature, namely that for a player the best choice is the same regardless of what the other player or players choose, it greatly simplifies the players' thinking and the game theorists' analysis. Therefore it is worth making a big deal of it, and looking for it to simplify the solution of the game. The name given by game theorists for this property is *dominant strategy*. A player is said to have a dominant strategy if that same strategy is better for him than all of his other available strategies no matter what strategy or strategy combination the other player or players choose. And we have a simple rule for behavior in simultaneous-move games:*

RULE 2: If you have a dominant strategy, use it.

The prisoners' dilemma is an even more special game—not just one player but both (or all) players have dominant strategies. The reasoning of BB's manager is exactly analogous to that of RE's manager, and you should fix the idea by going through it on your own. You will see that \$70 is also the dominant strategy for BB.

The result is the outcome shown in the bottom right cell of the game table; both charge \$70 and make a profit of \$70,000 each. And here is the feature that makes the prisoners' dilemma such an important game. When both players use their dominant strategies, both do worse than they would have if somehow they could have jointly and credibly agreed that each would choose the other, dom-

* In chapter 2, we could offer a single, unifying principle to devise the best strategies for games with sequential moves. This was our Rule 1: Look forward and reason backward. It won't be so simple for simultaneous-move games. But the thinking about thinking required for simultaneous moves can be summarized in three simple rules for action. These rules in turn rest on two simple ideas—dominant strategies and equilibrium. Rule 2 is given here; Rules 3 and 4 will follow in the next chapter.

inated strategy. In this game, that would have meant charging \$80 each to obtain the outcome in the top left cell of the game table, namely \$72,000 each.*

It would not be enough for just one of them to price at \$80; then that firm would do very badly. Somehow they must both be induced to price high, and this is hard to achieve given the temptation each of them has to try to undercut the other. Each firm pursuing its own self-interest does not lead to an outcome that is best for them all, in stark contrast to what conventional theories of economics from Adam Smith onward have taught us.†

This opens up a host of questions, some of which pertain to more general aspects of game theory. What happens if only one player has a dominant strategy? What if none of the players has a dominant strategy? When the best choice for each varies depending on what the other is choosing simultaneously, can they see through each other's choices and arrive at a solution to the game? We will take up these questions in the next chapter, where we develop a more general concept of solution for simultaneous-move games, namely Nash equilibrium. In this chapter we focus on questions about the prisoners' dilemma game per se.

In the general context, the two strategies available to each player are labeled "Cooperate" and "Defect" (or sometimes "Cheat"), and we will follow this usage. Defect is the dominant strategy for each, and the combination where both choose Defect yields a worse outcome for both than if both choose Cooperate.

* Actually, \$80 is the common price that yields the two the highest possible joint profit; that is the price they would choose if they could get together and cartelize the industry. Rigorous proof of this statement requires some math, so just take our word for it. For readers who want to follow the calculation, it is on the book's web site.

† The beneficiaries from this price cutting by the firms are of course the consumers, who are not active players in this game. Therefore it is often in the larger society's interest to prevent the two firms from resolving their pricing dilemma. That is the role of antitrust policies in the United States and other countries.

Some Preliminary Ideas for Resolving the Dilemma

The players caught on the horns of this dilemma have strong incentives to make joint agreements to avoid it. For example, the fishermen in New England might agree to limit their catch to preserve the fish stocks for the future. The difficulty is to make such agreements stick, when each faces the temptation to cheat, for example, to take more than one's allotted quota of fish. What does game theory have to say on this issue? And what happens in the actual play of such games?

In the fifty years since the prisoners' dilemma game was invented, its theory has advanced a great deal, and much evidence has accumulated, both from observations about the real world and from controlled experiments in laboratory settings. Let us look at all this material and see what we can learn from it.

The flip side of achieving cooperation is avoiding defection. A player can be given the incentive to choose cooperation rather than the originally dominant strategy of defection by giving him a suitable reward, or deterred from defecting by creating the prospect of a suitable punishment.

The reward approach is problematic for several reasons. Rewards can be internal—one player pays the other for taking the cooperative action. Sometimes they can be external; some third party that also benefits from the two players' cooperation pays them for cooperating. In either case, the reward cannot be given before the choice is made; otherwise the player will simply pocket the reward and then defect. If the reward is merely promised, the promise may not be credible: after the promisee has chosen cooperation, the promisor may renege.

These difficulties notwithstanding, rewards are sometimes feasible and useful. At an extreme of creativity and imagination, the players could make simultaneous and mutual promises and make these credible by depositing the promised rewards in an escrow account controlled by a third party.⁶ More realistically, sometimes the players interact in several dimensions, and cooperation in one can be rewarded with reciprocation in another. For example,

among groups of female chimpanzees, help with grooming is reciprocated by sharing food or help with child minding. Sometimes third parties may have sufficiently strong interests in bringing about cooperation in a game. For example, in the interest of bringing to an end various conflicts around the world, the United States and the European Union have from time to time promised economic assistance to combatants as a reward for peaceful resolutions of their disputes. The United States rewarded Israel and Egypt in this way for cooperating to strike the Camp David Accords in 1978.

Punishment is the more usual method of resolving prisoners' dilemmas. This could be immediate. In a scene from the movie *L.A. Confidential*, Sergeant Ed Exley promises Leroy Fontaine, one of the suspects he is interrogating, that if he turns state's witness, he will get a shorter sentence than the other two, Sugar Ray Coates and Tyrone Jones. But Leroy knows that, when he emerges from jail, he may find friends of the other two waiting for him!

But the punishment that comes to mind most naturally in this context arises from the fact that most such games are parts of an ongoing relationship. Cheating may gain one player a short-term advantage, but this can harm the relationship and create a longer-run cost. If this cost is sufficiently large, that can act as a deterrent against cheating in the first place.*

A striking example comes from baseball. Batters in the American League are hit by pitches 11 to 17 percent more often than their colleagues in the National League. According to Sewanee professors Doug Drinen and John-Charles Bradbury, most of this difference is explained by the designated hitter rule.⁷ In the American League, the pitchers don't bat. Thus an American League pitcher who plunks a batter doesn't have to fear direct retaliation from the opposing team's pitcher. Although pitchers are unlikely to get hit, the chance goes up by a factor of four if they have just plunked someone in the previous half inning. The fear of retaliation is clear.

* Robert Aumann was awarded the 2005 Nobel Prize in Economics for his instrumental work in the development of the general theory of tacit cooperation in repeated games.

As ace pitcher Curt Schilling explained: "Are you seriously going to throw at someone when you are facing Randy Johnson?"⁸

When most people think about one player punishing the other for past cheating, they think of some version of tit for tat. And that was indeed the finding of what is perhaps the most famous experiment on the prisoners' dilemma. Let us recount what happened and what it teaches.

TIT FOR TAT

In the early 1980s, University of Michigan political scientist Robert Axelrod invited game theorists from around the world to submit their strategies for playing the prisoners' dilemma in the form of computer programs. The programs were matched against each other in pairs to play a prisoners' dilemma game repeated 150 times. Contestants were then ranked by the sum of their scores.

The winner was Anatol Rapoport, a mathematics professor at the University of Toronto. His winning strategy was among the simplest: tit for tat. Axelrod was surprised by this. He repeated the tournament with an enlarged set of contestants. Once again Rapoport submitted tit for tat and beat the competition.

Tit for tat is a variation of the eye for an eye rule of behavior: Do unto others as they have done onto you.* More precisely, the strat-

* In Exodus (21:22-25), we are told: "If men who are fighting hit a pregnant woman and she gives birth prematurely but there is no serious injury, the offender must be fined whatever the woman's husband demands and the court allows. But if there is serious injury, you are to take life for life, eye for eye, tooth for tooth, hand for hand, foot for foot, burn for burn, wound for wound, bruise for bruise." The New Testament suggests more cooperative behavior. In Matthew (5:38-39) we find: "You have heard that it was said, 'Eye for eye, and tooth for tooth.' But I tell you, Do not resist an evil person. If someone strikes you on the right cheek, turn to him the other also." We move from "Do unto others as they have done onto you" to the golden rule: "Do to others as you would have them do to you" (Luke 6:31). If people were to follow the golden rule, there would be no prisoners' dilemma. And if we think in the larger perspective, although cooperation might lower your payoffs in any particular game, the potential reward in an afterlife may make this a rational strategy even for a selfish

egy cooperates in the first period and from then on mimics the rival's action from the previous period.

Axelrod argues that tit for tat embodies four principles that should be present in any effective strategy for the repeated prisoners' dilemma: clarity, niceness, provocability, and forgivingness. Tit for tat is as *clear* and simple as you can get; the opponent does not have to do much thinking or calculation about what you are up to. It is *nice* in that it never initiates cheating. It is *provocable*—that is, it never lets cheating go unpunished. And it is *forgiving*, because it does not hold a grudge for too long and is willing to restore cooperation.

One of the impressive features about tit for tat is that it did so well overall even though it did not (nor could it) beat any one of its rivals in a head-on competition. At best, tit for tat ties its rival. Hence if Axelrod had scored each competition as a winner-take-all contest, tit for tat would have only losses and ties and therefore could not have had the best track record.*

But Axelrod did not score the pairwise plays as winner-take-all: close counted. The big advantage of tit for tat is that it always comes close. At worst, tit for tat ends up getting beaten by one defection—that is, it gets taken advantage of once and then ties from then on.

The reason tit for tat won the tournament is that it usually managed to encourage cooperation whenever possible while avoiding exploitation. The other entries either were too trusting and open to exploitation or were too aggressive and knocked one another out.

In spite of all this, we believe that tit for tat is a flawed strategy. The slightest possibility of a mistake or a misperception results in a complete breakdown in the success of tit for tat. This flaw was not apparent in the artificial setting of a computer tournament,

individual. You don't think there is an afterlife? Pascal's Wager says that the consequences of acting on that assumption can be quite drastic, so why take the chance.

* Since every loser must be paired with a winner, it must be the case that some contestant will have more wins than losses, else there will be more losses than wins overall. (The only exception is when every single match is a tie.)

because mistakes and misperceptions did not arise. But when tit for tat is applied to real-world problems, mistakes and misperceptions cannot be avoided, and the result can be disastrous.

The problem with tit for tat is that any mistake "echoes" back and forth. One side punishes the other for a defection, and this sets off a chain reaction. The rival responds to the punishment by hitting back. This response calls for a second punishment. At no point does the strategy accept a punishment without hitting back.

Suppose, for example, that both Flood and Dresher start out playing tit for tat. No one initiates a defection, and all goes well for a while. Then, in round 11, say, suppose Flood chooses Defect by mistake, or Flood chooses Cooperate but Dresher mistakenly thinks Flood chose Defect. In either case, Dresher will play Defect in round 12, but Flood will play Cooperate because Dresher played Cooperate in round 11. In round 13 the roles will be switched. The pattern of one playing Cooperate and the other playing Defect will continue back and forth, until another mistake or misperception restores cooperation or leads both to defect.

Such cycles or reprisals are often observed in real-life feuds between Israelis and Arabs in the Middle East, or Catholics and Protestants in Northern Ireland, or Hindus and Muslims in India. Along the West Virginia-Kentucky border, we had the memorable feud between the Hatfields and the McCoys. And in fiction, Mark Twain's Grangerfords and Shepherdsons offer another vivid example of how tit for tat behavior can end in a cycle of reprisals. When Huck Finn tries to understand the origins of the Grangerford-Shepherdson feud, he runs into the chicken-or-egg problem:

"What was the trouble about, Buck?—land?"

"I reckon maybe—I don't know."

"Well, who done the shooting? Was it a Grangerford or a Shepherdson?"

"Laws, how do I know? It was so long ago."

"Don't anybody know?"

"Oh, yes, pa knows, I reckon, and some of the other old people; but they don't know now what the row was about in the first place."

What tit for tat lacks is a way of saying "Enough is enough." It is too provocative, and not forgiving enough. And indeed, subsequent versions of Axelrod's tournament, which allowed possibilities of mistakes and misperceptions, showed other, more generous strategies to be superior to tit for tat.*

Here we might even learn something from monkeys. Cotton-top tamarin monkeys were placed in a game where each had the opportunity to pull a lever that would give the other food. But pulling the lever required effort. The ideal for each monkey would be to shirk while his partner pulled the lever. But the monkeys learned to cooperate in order to avoid retaliation. Tamarin cooperation remained stable as long as there were no more than two consecutive defections by one player, a strategy that resembles tit for two tats.⁹

MORE RECENT EXPERIMENTS

Thousands of experiments on the prisoners' dilemma have been performed in classrooms and laboratories, involving different numbers of players, repetitions, and other treatments. Here are some important findings.¹⁰

First and foremost is that cooperation occurs significantly often, even when each pair of players meets only once. On average, almost half of the players choose the cooperative action. Indeed, the most striking demonstration of this was on the Game Show Network's production of *Friend or Foe*. In this show, two-person

* In 2004, Graham Kendall at Nottingham ran a contest to celebrate the twentieth anniversary of Axelrod's original tournament. It was "won" by a group from England's Southampton University. The Southampton group submitted multiple entries, sixty in all. There were 59 drones and 1 queen. All their entries started with an unusual pattern so they would recognize each other. Then the drone programs sacrificed themselves so that the queen would do well. The drone programs also refused to cooperate with any rival program so as to knock down the opponents' scores. While having an army of drones prepared to sacrifice themselves on your behalf is one way to increase your payoff, it doesn't tell us much about how to play a prisoners' dilemma.

teams were asked trivia questions. The money earned from correct answers went into a "trust fund," which over the 105 episodes ranged from \$200 to \$16,400. To divide the trust fund, the two contestants played a one-shot dilemma.

Each privately wrote down "friend" or "foe." When both wrote down friend, the pot was split evenly. If one wrote down foe while the other wrote friend, the person writing foe would get the whole pot. But if both wrote foe, then neither would get anything. Whatever the other side does, you get at least as much, and possibly more, by writing down foe than if you wrote friend. Yet almost half the contestants wrote down friend. Even as the pot grew larger there was no change in the likelihood of cooperation. People were as likely to cooperate when the fund was below \$3,000 as they were when it was above \$5,000. These were some of the findings in a pair of studies by Professors Felix Oberholzer-Gee, Joel Waldfogel, Matthew White, and John List.¹¹

If you are wondering how watching television counts as academic research, it turns out that more than \$700,000 was paid out to contestants. This was the best-funded experiment on the prisoners' dilemma, ever. There was much to learn. It turns out that women were much more likely to cooperate than men, 53.7 percent versus 47.5 percent in season 1. The contestants in season 1 didn't have the advantage of seeing the results from the other matches before making their decision. But in season 2, the results of the first 40 episodes had been aired and this pattern became apparent. The contestants had learned from the experience of others. When the team consisted of two women, the cooperation rate rose to 55 percent. But when a woman was paired with a guy, her cooperation rate fell to 34.2 percent. And the guy's rate fell, too, down to 42.3 percent. Overall, cooperation dropped by ten points.

When a group of subjects is assembled and matched pairwise a number of times, with different pairings at different times, the proportion choosing cooperation generally declines over time. However, it does not go to zero, settling instead on a small set of persistent cooperators.

If the same pair plays the basic dilemma game repeatedly, they

often build up to a significant sequence of mutual cooperation, until one player defects near the end of the sequence of repetitions. This happened in the very first experiment conducted on the dilemma. Almost immediately after they had thought up the game, Flood and Dresher recruited two of their colleagues to play the dilemma game a hundred times.¹² On 60 of these rounds, both players chose Cooperate. A long stretch of mutual cooperation lasted from round 83 to round 98, until one player sneaked in a defection in round 99.

Actually, according to the strict logic of game theory, this should not have happened. When the game is repeated exactly 100 times, it is a sequence of simultaneous-move games, and we can apply the logic of backward reasoning to it. Look ahead to what will happen on the 100th play. There are no more games to come, so defection cannot be punished in any future rounds. Dominant strategy calculations dictate that both players should choose Defect on the last round. But once that is a given, the 99th round becomes effectively the last round. Although there is one more round to come, defection on the 99th round is not going to be selectively punished by the other player in the 100th round because his choice in that round is foreordained. Therefore the logic of dominant strategies applies to the 99th round. And one can work back this sequential logic all the way to round 1. But in actual play, both in the laboratory and the real world, players seem to ignore this logic and achieve the benefits of mutual cooperation. What may seem at first sight to be irrational behavior—departing from one's dominant strategy—turns out to be a good choice, so long as everyone else is similarly "irrational."

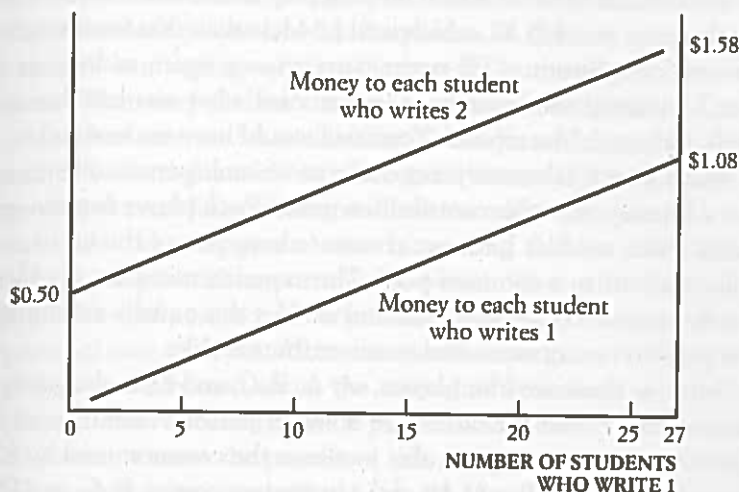
Game theorists suggest an explanation for this phenomenon. The world contains some "reciprocators," people who will cooperate so long as the other does likewise. Suppose you are not one of these relatively nice people. If you behaved true to your type in a finitely repeated game of prisoners' dilemma, you would start cheating right away. That would reveal your nature to the other player. To hide the truth (at least for a while), you have to behave nicely. Why would you want to do that? Suppose you started by

acting nicely. Then the other player, even if he is not a reciprocator, would think it possible that you are one of the few nice people around. There are real gains to be had by cooperating for a while, and the other player would plan to reciprocate your niceness to achieve these gains. That helps you, too. Of course you are planning to sneak in a defection near the end of the game, just as the other player is. But you two can still have an initial phase of mutually beneficial cooperation. While each side is waiting to take advantage of the other, both are benefiting from this mutual deception.

In some experiments, instead of pairing each subject in the group with another person and playing several two-person dilemmas, the whole group is engaged in one large multiperson dilemma. We mention a particularly entertaining and instructive instance from the classroom. Professor Raymond Battalio of Texas A&M University had his class of 27 students play the following game.¹³ Each student owned a hypothetical firm and had to decide (simultaneously and independently, by writing on a slip of paper) whether to produce 1 and help keep the total supply low and the price high or produce 2 and gain at the expense of others. Depending on the total number of students producing 1, money would be paid to students according to the following table:

Number of students who write 1	Payoff to each student who writes 1	Payoff to each student who writes 2
0		\$0.50
1	\$0.04	\$0.54
2	\$0.08	\$0.58
3	\$0.12	\$0.62
...
25	\$1.00	\$1.50
26	\$1.04	\$1.54
27	\$1.08	

This is easier to see and more striking in a chart:



The game is "rigged" so that students who write 2 (Defect) always get 50 cents more than those who write 1 (Cooperate), but the more of them that write 2, the less their collective gain. Suppose all 27 start planning to write 1, so each would get \$1.08. Now one thinks of sneaking a switch to 2. There would be 26 1s, and each would get \$1.04 (4 cents less than in the original plan), but the switcher would get \$1.54 (46 cents more). The same is true irrespective of the initial number of students thinking of writing 1 versus 2. Writing 2 is a dominant strategy. Each student who switches from writing 1 to writing 2 increases his own payout by 46 cents but decreases that of each of his 26 colleagues by 4 cents—the group as a whole loses 58 cents. By the time everyone acts selfishly, each maximizing his own payoff, they each get 50 cents. If they could have successfully conspired and acted so as to minimize their individual payoff, they would each receive \$1.08. How would you play?

In some practice plays of this game, first without classroom

discussion and then with some discussion to achieve a "conspiracy," the number of cooperative students writing 1 ranged from 3 to a maximum of 14. In a final binding play, the number was 4. The total payout was \$15.82, which is \$13.34 less than that from totally successful collusion. "I'll never trust anyone again as long as I live," muttered the conspiracy leader. And what was his choice? "Oh, I wrote 2," he replied. Yossarian would have understood.

More recent laboratory experiments of multiperson dilemmas use a format called the contribution game. Each player is given an initial stake, say \$10. Each can choose to keep part of this and contribute a part to a common pool. The experimenter then doubles the accumulated common pool and divides this equally among all the players, contributors and noncontributors alike.

Suppose there are four players, say A, B, C, and D, in the group. Regardless of what the others are doing, if person A contributes a dollar to the common pool, this increases the common pool by \$2 after the doubling. But \$1.50 of the increment goes to B, C, and D; A gets only 50 cents. Therefore A loses out by raising his contribution; conversely he would gain by lowering it. And that is true no matter how much, if anything, the others are contributing. In other words, contributing nothing is the dominant strategy for A. The same goes for B, C, and D. This logic says that each should hope to become a "free rider" on the efforts of the others. If all four play their dominant strategy, the common pool is empty and each simply keeps the initial stake of \$10. When everyone tries to be a free rider, the bus stays in the garage. If everyone had put all of their initial stakes in the common pool, the pool after doubling would be \$80 and the share of each would be \$20. But each has the personal incentive to cheat on such an arrangement. This is their dilemma.

The contribution game is not a mere curiosity of the laboratory or theory; it is played in the real world in social interactions where some communal benefit can be achieved by voluntary contributions from members of the group, but the benefit cannot be withheld from those who did not contribute. Flood control in a village, or conservation of natural resources, are cases in point: it is not possible to build levees or dams so that flood waters will selectively go to

the fields of those who did not help in the construction, and it is not practicable to withhold gas or fish in the future from someone who consumed too much in the past. This creates a multiperson dilemma: each player has the temptation to shirk or withhold his contribution, hoping to enjoy the benefits of the others' contributions. When they all think this way, the total of contributions is meager or even zero, and they all suffer. These situations are ubiquitous, and of such magnitude that all of social theory and policy needs a good understanding of how the dilemmas might be resolved.

In what is perhaps the most interesting variant of the game, players are given an opportunity to punish those who cheat on an implicit social contract of cooperation. However, they must bear a personal cost to do so. After the contribution game is played, the players are informed about the individual contributions of other players. Then a second phase is played, where each player can take an action to lower the payoffs of other players at a cost to himself of so many cents (typically 33) per dollar reduction chosen. In other words, if player A chooses to reduce B's payoff by three dollars, then A's payoff is reduced by one dollar. These reductions are not reallocated to anyone else; they simply return to the general funds of the experimenter.

The results of the experiment show that people engage in a significant amount of punishment of "social cheaters," and that the prospect of the punishment increases the contributions in the first phase of the game dramatically. Such punishments seem to be an effective mechanism for achieving cooperation that benefits the whole group. But the fact that individuals carry them out is surprising at first. The act of punishing others at a personal cost is itself a contribution for the general benefit, and it is a dominated strategy; if it succeeds in eliciting better behavior from the cheater in the future, its benefits will be for the group as a whole, and the punisher will get only his small share of this benefit. Therefore the punishment has to be the result of something other than a selfish calculation. That is indeed the case. Experiments on this game have been conducted while the players' brains were being imaged by PET scan.¹⁴ These revealed that the act of imposing the penalty

activated a brain region called the dorsal striatum, which is involved in experiencing pleasure or satisfaction. In other words, people actually derive a psychological benefit or pleasure from punishing social cheaters. Such an instinct must have deep biological roots and may have been selected for an evolutionary advantage.¹⁵

HOW TO ACHIEVE COOPERATION

These examples and experiments have suggested several preconditions and strategies for successful cooperation. Let us develop the concepts more systematically and apply them to some more examples from the real world.

Successful punishment regimes must satisfy several requirements. Let us examine these one by one.

Detection of cheating: Before cheating can be punished, it must be detected. If detection is fast and accurate, the punishment can be immediate and accurate. That reduces the gain from cheating while increasing its cost, and thus increases the prospects for successful cooperation. For example, airlines constantly monitor each other's fares; if American were to lower its fare from New York to Chicago, United can respond in under five minutes. But in other contexts, firms that want to cut their prices can do so in secret deals with the customers, or hide their price cuts in a complicated deal involving many dimensions of delivery time, quality, warranties, and so on. In extreme situations, each firm can only observe its own sales and profits, which can depend on some chance elements as well as on other firms' actions. For example, how much one firm sells can depend on the vagaries of demand, not just on other firms' secret price cuts. Then detection and punishment become not only slow but also inaccurate, raising the temptation to cheat.

Finally, when three or more firms are simultaneously in the market, they must find out not only whether cheating has occurred but who has cheated. Otherwise any punishments cannot be targeted to hurt the miscreant but have to be blunt, perhaps unleashing a price war that hurts all.

Nature of punishment: Next, there is the choice of punishment. Sometimes the players have available to them actions that hurt others, and these can be invoked after an instance of cheating even in a one-time interaction. As we pointed out in the dilemma in *L.A. Confidential*, the friends of Sugar and Tyrone will punish Leroy when he emerges from jail after his light sentence for turning state's witness. In the Texas A&M classroom experiment, if the students could detect who had reneged on the conspiracy for all of them to write 1, they could inflict social sanctions such as ostracism on the cheaters. Few students would risk that for the sake of an extra 50 cents.

Other kinds of punishments arise within the structure of the game. Usually this happens because the game is repeated, and the gain from cheating in one play leads to a loss in future plays. Whether this is enough to deter a player who is contemplating cheating depends on the sizes of the gains and losses and on the importance of the future relative to the present. We will return to this aspect soon.

Clarity: The boundaries of acceptable behavior, and the consequences of cheating, should be clear to a prospective cheater. If these things are complex or confusing, the player may cheat by mistake or fail to make a rational calculation and play by some hunch. For example, suppose Rainbow's End and B. B. Lean are playing their price-setting game repeatedly, and RE decides that it will infer that BB has cheated if RE's discounted mean of profits from the last seventeen months is 10 percent less than the average real rate of return to industrial capital over the same period. BB does not know this rule directly; it must infer what rule RE is using by observing RE's actions. But the rule stated here is too complicated for BB to figure out. Therefore it is not a good deterrent against BB's cheating. Something like tit for tat is abundantly clear: if BB cheats, it will see RE cutting its price the very next time.

Certainty: Players should have confidence that defection will be punished and cooperation rewarded. This is a major problem in some international agreements like trade liberalization in the

World Trade Organization (WTO). When one country complains that another has cheated on the trade agreement, the WTO initiates an administrative process that drags on for months or years. The facts of the case have little bearing on the judgment, which usually depends more on dictates of international politics and diplomacy. Such enforcement procedures are unlikely to be effective.

Size: How harsh should such punishments be? It might seem that there is no limit. If the punishment is strong enough to deter cheating, it need never actually be inflicted. Therefore it may as well be set at a sufficiently high level to ensure deterrence. For example, the WTO could have a provision to nuke any nation that breaks its undertakings to keep its protective tariffs at the agreed low levels. Of course you recoil in horror at the suggestion, but that is at least in part because you think it possible that some error may cause this to happen. When errors are possible, as they always are in practice, the size of the punishment should be kept as low as is compatible with successful deterrence in most circumstances. It may even be optimal to forgive occasional defection in extreme situations—for example, a firm that is evidently fighting for its survival may be allowed some price cuts without triggering reactions from rivals.

Repetition: Look at the pricing game between Rainbow's End and B. B. Lean. Suppose they are going merrily along from one year to the next, holding prices at their joint best, \$80. One year the management of RE considers the possibility of cutting the price to \$70. They reckon that this will yield them an extra profit of $\$110,000 - \$72,000 = \$38,000$. But that can lead to a collapse of trust. RE should expect that in future years BB will also choose \$70, and each will make only \$70,000 each year. If RE had kept to the original arrangement, each would have kept on making \$72,000. Thus RE's price cutting will cost it $\$72,000 - \$70,000 = \$2,000$ every year in the future. Is a one-time gain of \$38,000 worth the loss of \$2,000 every year thereafter?

One key variable that determines the balance of present and

future considerations is the interest rate. Suppose the interest rate is 10% per year. Then RE can stash away its extra \$38,000 and earn \$3,800 every year. That comfortably exceeds the loss of \$2,000 in each of those years. Therefore cheating is in RE's interest. But if the interest rate is only 5% per year, then the \$38,000 earns only \$1,900 in each subsequent year, less than the loss of \$2,000 due to the collapse of the arrangement; so RE does not cheat. The interest rate at which the two magnitudes just balance is $2/38 = 0.0526$, or 5.26% per year.

The key idea here is that when interest rates are low, the future is relatively more valuable. For example, if the interest rate is 100%, then the future has low value relative to the present—a dollar in a year's time is worth only 50 cents right now because you can turn the 50 cents into a dollar in a year by earning another 50 cents as interest during the year. But if the interest rate is zero, then a dollar in a year's time is worth the same as a dollar right away.*

In our example, for realistic interest rates a little above 5%, the temptation for each firm to cut the price by \$10 below their joint best price of \$80 is quite finely balanced, and collusion in a repeated game may or may not be possible. In chapter 4 we will see how low the price can fall if there is no shadow of the future and the temptation to cheat is irresistible.

Another relevant consideration is the likelihood of continuation of the relationship. If the shirt is a transient fashion item that may not sell at all next year, then the temptation to cheat this year is not offset by any prospect of future losses.

But Rainbow's End and B. B. Lean sell many items besides this shirt. Won't cheating on the shirt price bring about retaliation on all those other items in the future? And isn't the prospect of this huge retaliation enough to deter the defection? Alas, the usefulness of multiproduct interactions for sustaining cooperation is not

* If you read the financial press, you have often seen the statement: "Interest rates and bond prices move in opposite directions." The lower the interest rate, the higher the prices of bonds. And bonds, being promises of future income, reflect the importance of the future. This is another way to remember the role of interest rates.

so simple. The prospect of multiproduct retaliation goes hand in hand with that of immediate gains from simultaneous cheating in all of those dimensions, not just one. If all the products had identical payoff tables, then the gains and losses would both increase by a factor equal to the number of products, and so whether the balance is positive or negative would not change. Therefore successful punishments in multiproduct dilemmas must depend in a more subtle way on differences among the products.

A third relevant consideration is the expected variation in the size of the business over time. This has two aspects—steady growth or decline, and fluctuations. If the business is expected to grow, then a firm considering defection now will recognize that it stands to lose more in the future due to the collapse of the cooperation and will be more hesitant to defect. Conversely, if the business is on a path of decline, then firms will be more tempted to defect and take what they can now, knowing that there is less at stake in the future. As for fluctuations, firms will be more tempted to cheat when a temporary boom arrives; cheating will bring them larger immediate profits, whereas the downside from the collapse of the cooperation will hit them in the future, when the volume of business will be only the average, by definition of the average. Therefore we should expect that price wars will break out during times of high demand. But this is not always the case. If a period of low demand is caused by a general economic downturn, then the customers will have lower incomes and may become sharper shoppers as a result—their loyalties to one firm or the other may break down, and they may respond more quickly to price differences. In that case, a firm cutting its price can expect to attract more customers away from its rival, and thereby reap a larger immediate gain from defection.

Finally, the composition of the group of players is important. If this is stable and expected to remain so, that is conducive to the maintenance of cooperation. New players who do not have a stake or a history of participation in the cooperative arrangement are less likely to abide by it. And if the current group of players expects new ones to enter and shake up the tacit cooperation in

the future, that increases their own incentive to cheat and take some extra benefit right now.

SOLUTION BY KANTIAN CATEGORICAL IMPERATIVE?

It is sometimes said that the reason some people cooperate in the prisoners' dilemma is that they are making the decision not only for themselves but for the other player. That is wrong in point of fact, but the person is acting as if this is the case.

The person truly wants the other side to cooperate and reasons to himself that the other side is going through the same logical decision process that he is. Thus the other side must come to the same logical conclusion that he has. Hence if the player cooperates, he reasons that the other side will do so as well, and if he defects, he reasons that it will cause the other side to defect. This is similar to the categorical imperative of the German philosopher Immanuel Kant: "Take only such actions as you would like to see become a universal law."

Of course, nothing could be further from the truth. The actions of one player have no effect whatsoever on the other player in the game. Still people think that somehow their actions can influence the choice of others, even when their actions are invisible.

The power of this thinking was revealed in an experiment done with Princeton undergraduates by Eldar Shafir and Amos Tversky.¹⁶ In their experiment, they put students in a prisoners' dilemma game. But unlike the usual dilemma, in some treatments they told one side what the other had done. When students were told that the other side had defected on them, only 3 percent responded with cooperation. When told that the other side had cooperated, this increased cooperation levels up to 16 percent. It was still the case that the large majority of students were willing to act selfishly. But many were willing to reciprocate the cooperative behavior exhibited by the other side, even at their own expense.

What do you think would happen when the students were not told anything about the other player's choice at all? Would the percentage of cooperators be between 3 and 16 percent? No; it rose to

37 percent. At one level, this makes no sense. If you wouldn't cooperate when you learned that the other side had defected and you wouldn't cooperate when you learned that the other side had cooperated, why would you then cooperate when you don't know what the other side had done?

Shafir and Tversky call this "quasi-magical" thinking—the idea that by taking some action, you can influence what the other side will do. People realize they can't change what the other side has done once they've been told what the other side has done. But if it remains open or undisclosed, then they imagine that their actions might have some influence—or that the other side will somehow be employing the same reasoning chain and reach the same outcome they do. Since Cooperate, Cooperate is preferred to Defect, Defect, the person chooses Cooperate.

We want to be clear that such logic is completely illogical. What you do and how you get there has no impact at all on what the other side thinks and acts. They have to make up their mind without reading your mind or seeing your move. However, the fact remains that if the people in a society engage in such quasi-magical thinking, they will not fall victim to many prisoners' dilemmas and all will reap higher payoffs from their mutual interactions. Could it be that human societies deliberately instill such thinking into their members for just such an ultimate purpose?

DILEMMAS IN BUSINESS

Armed with the tool kit of experimental findings and theoretical ideas in the previous sections, let us step outside the laboratory and look at some instances of prisoners' dilemmas in the real world and attempts at resolving them.

Let us begin with the dilemma of rival firms in an industry. Their joint interests are best served by monopolizing or cartelizing the industry and keeping prices high. But each firm can do better for itself by cheating on such an agreement and sneaking in price cuts to steal business from its rivals. What can the firms do? Some factors conducive to successful collusion, such as growing demand or

lack of disruptive entry, may be at least partially outside their control. But they can try to facilitate the detection of cheating and devise effective punishment strategies.

Collusion is easier to achieve if the firms meet regularly and communicate. Then they can negotiate and compromise on what are acceptable practices and what constitutes cheating. The process of negotiation and its memory contributes to clarity. If something occurs that looks *prima facie* like cheating, another meeting can help clarify whether it is something extraneous, an innocent error by a participant, or deliberate cheating. Therefore unnecessary punishments can be avoided. And the meeting can also help the group implement the appropriate punishment actions.

The problem is that the group's success in resolving their dilemma harms the general public's interest. Consumers must pay higher prices, and the firms withhold some supply from the market to keep the price high. As Adam Smith said, "People of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices."¹⁷ Governments that want to protect the general public interest get into the game and enact antitrust laws that make it illegal for firms to collude in this way.* In the United States, the Sherman Antitrust Act prohibits conspiracies "in restraint of trade or commerce," of which price fixing or market-share fixing conspiracies are the prime instance and the ones most frequently attempted. In fact the Supreme Court has ruled that not only are explicit agreements of this kind forbidden, but also any explicit or tacit arrangement among firms that has the effect of price fixing is a violation of the Sherman Act, regardless of its primary intent. Violation of these laws can lead to jail terms for the firms' executives, not just fines for the corporations that are impersonal entities.

Not that firms don't try to get away with the illegal practices. In

* Not all governments care enough about the general interest. Some are beholden to the producers' special interests and ignore or even facilitate cartels. We won't name any, lest they ban our book in their countries!

1996 Archer Daniels Midland (ADM), a leading American processor of agricultural products, and their Japanese counterpart, Ajinomoto were caught in just such a conspiracy. They had arranged market sharing and pricing agreements for various products such as lysine (which is produced from corn and used for fattening up chickens and pigs). The aim was to keep the prices high at the expense of their customers. Their philosophy was: "The competitors are our friends, and the customers are our enemies." The companies' misdeeds came to light because one of the ADM negotiators became an informant for the FBI and arranged for many of the meetings to be recorded for audio and sometimes also video.¹⁸

An instance famous in antitrust history and business school case studies concerns the large turbines that generate electricity. In the 1950s, the U.S. market for these turbines consisted of three firms: GE was the largest, with a market share of around 60 percent, Westinghouse was the next, with approximately 30 percent, and Allied-Chalmers had about 10 percent. They kept these shares, and obtained high prices, using a clever coordination device. Here's how it worked. Electric utilities invited bids for the turbines they intended to buy. If the bid was issued during days 1–17 of a lunar month, Westinghouse and Allied-Chalmers had to put in very high bids that would be sure losers, and GE was the conspiracy's chosen winner by making the lowest bid (but still at a monopolist's price allowing big profits). Similarly, Westinghouse was the designated winner in the conspiracy if the bid was issued during days 18–25, and Allied-Chalmers for days 26–28. Since the utilities did not issue their solicitations for bids according to the lunar calendar, over time each of the three producers got the agreed market share. Any cheating on the agreement would have been immediately visible to the rivals. But, so long as the Department of Justice did not think of linking the winners to the lunar cycles, it was safe from detection by the law. Eventually the authorities did figure it out, some executives of the three firms went to jail, and the profitable conspiracy collapsed. Different schemes were tried later.¹⁹

A variant of the turbine scheme later appeared in the bidding at the airwave spectrum auctions in 1996–1997. A firm that wanted

the right for the licenses in a particular location would signal to the other firms its determination to fight for that right by using the telephone area code for that location as the last three digits of its bid. Then the other firms would let it win. So long as the same set of firms interacts in a large number of such auctions over time and so long as the antitrust authorities do not figure it out, the scheme may be sustainable.²⁰

More commonly, the firms in an industry try to attain and sustain implicit or tacit agreements without explicit communication. This eliminates the risk of criminal antitrust action, although the antitrust authorities can take other measures to break up even implicit collusion. The downside is that the arrangement is less clear and cheating is harder to detect, but firms can devise methods to improve both.

Instead of agreeing on the prices to be charged, the firms can agree on a division of the market, by geography, product line, or some similar measure. Cheating is then more visible—your salespeople will quickly come to know if another company has stolen some of your assigned market.

Detection of price cuts, especially in the case of retail sales, can be simplified, and retaliation made quick and automatic, by the use of devices like "matching or beating competition" policies and most-favored-customer clauses. Many companies selling household and electronic goods loudly proclaim that they will beat any competitor's price. Some even guarantee that if you find a better price for the same product within a month after your purchase, they will refund the difference, or in some cases even double the difference. At first sight, these strategies seem to promote competition by guaranteeing low prices. But a little game-theoretic thinking shows that in reality they can have exactly the opposite effect. Suppose Rainbow's End and B. B. Lean had such policies, and their tacit agreement was to price the shirt at \$80. Now each firm knows that if it sneaks a cut to \$70, the rival will find out about it quickly; in fact the strategy is especially clever in that it puts the customers, who have the best natural incentive to locate low prices, in charge of detecting cheating. And the prospective defector also knows that

the rival can retaliate instantaneously by cutting its own price; it does not have to wait until next year's catalog is printed. Therefore the cheater is more effectively deterred.

Promises to meet or beat the competition can be clever and indirect. In the competition between Pratt & Whitney (P&W) and Rolls-Royce (RR) for jet aircraft engines to power Boeing 757 and 767 planes, P&W promised all prospective purchasers that its engines would be 8 percent more fuel-efficient than those of RR, otherwise P&W would pay the difference in fuel costs.²¹

A most-favored-customer clause says that the seller will offer to all customers the best price they offer to the most favored ones. Taken at face value, it seems that the manufacturers are guaranteeing low prices. But let's look deeper. The clause means that the manufacturer cannot compete by offering selective discounts to attract new customers away from its rival, while charging the old higher price to its established clientele. They must make general price cuts, which are more costly, because they reduce the profit margin on all sales. You can see the advantage of this clause to a cartel: the gain from cheating is less, and the cartel is more likely to hold.

A branch of the U.S. antitrust enforcement system, the Federal Trade Commission, considered such a clause that was being used by DuPont, Ethyl, and other manufacturers of antiknock additive compounds in gasoline. The commission ruled that there was an anticompetitive effect and forbade the companies from using such clauses in their contracts with customers.*

TRAGEDIES OF THE COMMONS

Among the examples at the start of this chapter, we mentioned problems like overfishing that arise because each person stands to benefit by taking more, while the costs of his action are visited upon

* This ruling was not without some controversy. The commission's chairman, James Miller, dissented. He wrote that the clauses "arguably reduce buyers' search costs and facilitate their ability to find the best price-value among buyers." For more information, see "In the matter of Ethyl Corporation et al.," FTC Docket 9128, *FTC Decisions* 101 (January-June 1983): 425-686.

numerous others or on future generations. University of California biologist Garrett Harding called this the "tragedy of the commons," using among his examples the overgrazing of commonly owned land in fifteenth- and sixteenth-century England.²² The problem has become well known under this name. Today the problem of global warming is an even more serious example; no one gets enough private benefit from reducing carbon emissions, but all stand to suffer serious consequences when each follows his self-interest.

This is just a multiperson prisoners' dilemma, like the one Yossarian faced in *Catch-22* about risking his life in wartime. Of course societies recognize the costs of letting such dilemmas go unresolved and make attempts to achieve better outcomes. What determines whether these attempts succeed?

Indiana University political scientist Elinor Ostrom and her collaborators and students have conducted an impressive array of case studies of attempts to resolve dilemmas of the tragedy of the commons—that is, to use and conserve common property resources in their general interest and avoid overexploitation and rapid depletion. They studied some successful and some unsuccessful attempts of this kind and derived some prerequisites for cooperation.²³

First, there must be clear rules that identify who is a member of the group of players in the game—those who have the right to use the resource. The criterion is often geography or residence but can also be based on ethnicity or skills, or membership may be sold by auction or for an entry fee.*

* The establishment of property rights is what actually happened in England. In two waves of "enclosures," first by local aristocrats during the Tudor period and later by acts of Parliament in the eighteenth and nineteenth centuries, previously common land was given to private owners. When land is private property, the invisible hand will shut the gate to just the right extent. The owner will charge grazing fees to maximize his rental income, and this will cut back on use. This will enhance overall economic efficiency but alter the distribution of income; the grazing fees will make the owner richer and the herdsmen poorer. Even absent concern for the distributional consequences, this approach is not always feasible. Property rights over the high seas or SO₂ and CO₂ emissions are hard to define and enforce in the absence of an international government: fish

Second, there must be clear rules defining permissible and forbidden actions. These include restrictions on time of use (open or closed seasons for hunting or fishing, or what kinds of crops can be planted and any requirements to keep the land fallow in certain years), location (a fixed position or a specified rotation for inshore fishing), the technology (size of fishing nets), and, finally, the quantity or fraction of the resource (amount of wood from a forest that each person is allowed to gather and take away).

Third, a system of penalties for violation of the above rules must be clear and understood by all parties. This need not be an elaborate written code; shared norms in stable communities can be just as clear and effective. The sanctions used against rule breakers range from verbal chastisement or social ostracism to fines, the loss of future rights, and, in some extreme cases, incarceration. The severity of each type of sanction can also be adjusted. An important principle is graduation. The first instance of suspected cheating is most commonly met simply by a direct approach to the violator and a request to resolve the problem. The fines for a first or second offense are low and are ratcheted up only if the infractions persist or get more blatant and serious.

Fourth, a good system to detect cheating must be in place. The best method is to make detection automatic in the course of the players' normal routine. For example, a fishery that has good and bad areas may arrange a rotation of the rights to the good areas. Anyone assigned to a good spot will automatically notice if a violator is using it and has the best incentive to report the violator to others and get the group to invoke the appropriate sanctions. Another example is the requirement that harvesting from forests or similar common areas must be done in teams; this facilitates mutual monitoring and eliminates the need to hire guards.

Sometimes the rules on what is permissible must be designed in

and pollutants move from one ocean to another, SO_2 is carried by the wind across borders, and CO_2 from any country rises to the same atmosphere. For this reason, whaling, acid rain, or global warming must be handled by more direct controls, but securing the necessary international agreements is no easy matter.

the light of feasible methods of detection. For example, the size of a fisherman's catch is often difficult to monitor exactly and difficult even for a well-intentioned fisherman to control exactly. Therefore rules based on fish quantity quotas are rarely used. Quantity quotas perform better when quantities are more easily and accurately observable, as in the case of water supplied from storage and harvesting of forest products.

Fifth, when the above categories of rules and enforcement systems are being designed, information that is easily available to the prospective users proves particularly valuable. Although each may have the temptation after the fact to cheat, they all have a common prior interest to design a good system. They can make the best use of their knowledge of the resource and of the technologies for exploiting it, the feasibility of detecting various infractions, and the credibility of various kinds of sanctions in their group. Centralized or top-down management has been demonstrated to get many of these things wrong and therefore perform poorly.

While Ostrom and her collaborators are generally optimistic about finding good solutions to many problems of collective action using local information and systems of norms, she gives a salutary warning against perfection: "The dilemma never fully disappears, even in the best operating systems. . . . No amount of monitoring or sanctioning reduces the temptation to zero. Instead of thinking of overcoming or conquering tragedies of the commons, effective governance systems cope better than others."

NATURE RED IN TOOTH AND CLAW

As you might expect, prisoners' dilemmas arise in species other than humans. In matters like building shelter, gathering food, and avoiding predators, an animal can act either selfishly in the interest of itself or its immediate kin, or in the interest of a larger group. What circumstances favor good collective outcomes? Evolutionary biologists have studied this question and found some fascinating examples and ideas. Here is a brief sample.²⁴

The British biologist J. B. S. Haldane was once asked whether

he would risk his life to save a fellow human being and replied: "For more than two brothers, or more than eight cousins, yes." You share half of your genes with a brother (other than an identical twin), and one-eighth of your genes with a cousin; therefore such action increases the expected number of copies of your genes that propagate to the next generation. Such behavior makes excellent biological sense; the process of evolution would favor it. This purely genetic basis for cooperative behavior among close kin explains the amazing and complex cooperative behavior observed in ant colonies and beehives.

Among animals, altruism without such genetic ties is rare. But reciprocal altruism can arise and persist among members of a group of animals with much less genetic identity, if their interaction is sufficiently stable and long-lasting. Hunting packs of wolves and other animals are examples of this. Here is an instance that is a bit gruesome but fascinating: Vampire bats in Costa Rica live in colonies of a dozen or so but hunt individually. On any day, some may be lucky and others unlucky. The lucky ones return to the hollow trees where the whole group lives and can share their luck by regurgitating the blood they have brought from their hunt. A bat that does not get a blood meal for three days is at risk of death. The colonies develop effective practices of mutual "insurance" against this risk by such sharing.²⁵

University of Maryland biologist Gerald Wilkinson explored the basis of this behavior by collecting bats from different locations and putting them together. Then he systematically withheld blood from some of them and saw whether others shared with them. He found that sharing occurred only when the bat was on the verge of death, and not earlier. Bats seem to be able to distinguish real need from mere temporary bad luck. More interesting, he found that sharing occurred only among bats that already knew each other from their previous group, and that a bat was much more likely to share with another bat that had come to its aid in the past. In other words, the bats are able to recognize other individual bats and keep score of their past behavior in order to develop an effective system of reciprocal altruism.

CASE STUDY: THE EARLY BIRD KILLS THE GOLDEN GOOSE

The Galápagos Islands are the home of Darwin's finches. Life on these volcanic islands is difficult and so evolutionary pressures are high. Even a millimeter change in the beak of a finch can make all the difference in the competition for survival.*

Each island differs in its food sources, and finches' beaks reflect those differences. On Daphne Major, the primary food source is a cactus. Here the aptly named cactus finch has evolved so that its beak is ideally suited to gather the pollen and nectar of the cactus blossom.

The birds are not consciously playing a game against each other. Yet each adaptation of a bird's beak can be seen as its strategy in life. Strategies that provide an advantage in gathering food will lead to survival, a choice of mating partners, and more offspring. The beak of the finch is a result of this combination of natural and sexual selection.

Even when things seem to be working, genetics throws a few curveballs into the mix. There is the old saying that the early bird gets the worm. On Daphne Major, it was the early finch that got the nectar. Rather than wait until nine in the morning when the cactus blossoms naturally open for business, a dozen finches were trying something new. They were prying open the cactus blossom to get a head start.

At first glance, this would seem to give these birds an edge over their late-coming rivals. The only problem is that in the process of prying open the blossom, the birds would often snip the stigma. As Weiner explains:

[The stigma] is the top of the hollow tube that pokes out like a tall straight straw from the center of each blossom. When the stigma is cut, the flower is sterilized. The male sex cells, in the

* This example is motivated by Jonathan Weiner's wonderful book, *The Beak of the Finch: A Story of Evolution in Our Time* (New York: Knopf, 1994). See especially chapter 20: "The Metaphysical Crossbeak."

pollen cannot reach the female sex cells in the flower. The cactus flower withers without bearing fruit.²⁶

When the cactus flowers wither, the main source of food disappears for the cactus finch. You can predict the end result of this strategy: no nectar, no pollen, no seeds, no fruit, and then no more cactus finch. Does that mean that evolution has led the finches into a prisoners' dilemma where the eventual outcome is extinction?

Case Discussion

Not quite, on two counts. Finches are territorial and so the finches (and their offspring) whose local cactus shut down may end up as losers. Killing next year's neighborhood food supply is not worth today's extra sip of pollen. Therefore these deviant finches would not appear to have a fitness advantage over the others. But that conclusion changes if this strategy ever becomes pervasive. The deviant finches will expand their search for food and even those finches that wait will not save their cactus's stigma. Given the famine that is sure to follow, the birds most likely to survive are those who started in the strongest position. The extra sip of nectar could make the difference.

What we have here is a cancerous adaptation. If it stays small, it can die out. But if it ever grows too large, it will become the fittest strategy on a sinking ship. Once it ever becomes advantageous even on a relative scale, the only way to get rid of it is to eliminate the entire population and start again. With no finches left on Daphne Major, there will be no one left to snip the stigmas and the cacti will bloom again. When two lucky finches alight on this island, they will have an opportunity to start the process from scratch.

The game we have here is a cousin to the prisoners' dilemma, a life and death case of the "stag hunt" game analyzed by the philosopher Jean-Jacques Rousseau.* In the stag hunt, if everyone works together to capture the stag, they succeed and all eat well. A

* There are other interpretations of Rousseau's stag hunt, to which we return in the history section of the next chapter.

problem arises if some hunters come across a hare along the way. If too many hunters are sidetracked chasing after hares, there won't be enough hunters left to capture the stag. In that case, everyone will do better chasing after rabbits. The best strategy is to go after the stag *if and only if* you can be confident that most everyone is doing the same thing. You have no reason not to chase after the stag, except if you lack confidence in what others will do.

The result is a confidence game. There are two ways it can be played. Everyone works together and life is good. Or everyone looks out for themselves and life is nasty, brutish, and short. This is not the classic prisoners' dilemma in which each person has an incentive to cheat no matter what others do. Here, there is no incentive to cheat, so long as you can trust others to do the same. But can you trust them? And even if you do, can you trust them to trust you? Or can you trust them to trust you to trust them? As FDR famously said (in a different context), we have nothing to fear but fear itself.

For more practice with prisoners' dilemmas, have a look at the following case studies in chapter 14: "What Price a Dollar?" and "The King Lear Problem."

A Beautiful Equilibrium

BIG GAME OF COORDINATION

Fred and Barney are Stone Age rabbit hunters. One evening, while carousing, they happen to engage in some shop talk. As they exchange information and ideas, they realize that by cooperating they could hunt much bigger game, such as stag or bison. One person on his own cannot expect any success hunting either stag or bison. But done jointly, each day's stag or bison hunting is expected to yield six times as much meat as a day's rabbit hunting by one person. Cooperation promises great advantage: each hunter's share of meat from a big-game hunt is three times what he can get hunting rabbits on his own.

The two agree to go big-game hunting together the following day and return to their respective caves. Unfortunately, they caroused too well, and both have forgotten whether they decided to go after stag or bison. The hunting grounds for the two species are in opposite directions. There were no cell phones in those days, and this was before the two became neighbors, so one could not quickly visit the other's cave to ascertain where to go. Each would have to make the decision the next morning in isolation.

Therefore the two end up playing a simultaneous-move game of

deciding where to go. If we call each hunter's quantity of meat from a day's rabbit hunting 1, then the share of each from successful coordination in hunting either stag or bison is 3. So the payoff table of the game is as shown here:

		Barney's choice		
		Stag	Bison	Rabbit
Fred's choice	Stag	3 3	0 0	1 0
	Bison	0 0	3 3	1 0
	Rabbit	0 1	0 1	1 1

This game differs from the prisoners' dilemma of the previous chapter in many ways. Let us focus on one crucial difference. Fred's best choice depends on what Barney does, and vice versa. For neither player is there a strategy that is best regardless of what the other does; unlike in the prisoners' dilemma, this game has no dominant strategies. So each player has to think about the other's choice and figure out his own best choice in light of that.

Fred's thinking goes as follows: "If Barney goes to the grounds where the stags are, then I will get my share of the large catch if I go there too, but nothing if I go to the bison grounds. If Barney goes to the bison grounds, things are the other way around. Rather than take the risk of going to one of these areas and finding that Barney has gone to the other, should I go by myself after rabbits and make sure of my usual, albeit small, quantity of meat? In other words, should I take 1 for sure instead of risking either 3 or nothing? It depends on what I think Barney is likely to do, so let me put myself in his shoes (bare feet?) and think what he is thinking. Oh, he is wondering what I am likely to do and is trying to put himself in my shoes! Is there any end to this circular thinking about thinking?"

SQUARING THE CIRCLE

John Nash's beautiful equilibrium was designed as a theoretical way to square just such circles of thinking about thinking about other people's choices in games of strategy.* The idea is to look for an outcome where each player in the game chooses the strategy that best serves his or her own interest, in response to the other's strategy. If such a configuration of strategies arises, neither player has any reason to change his choice unilaterally. Therefore, this is a potentially stable outcome of a game where the players make individual and simultaneous choices of strategies. We begin by illustrating the idea with some examples of it in action. Later in this chapter we discuss how well it predicts outcomes in various games; we find reasons for cautious optimism and for making Nash equilibrium a starting point of the analysis of almost all games.

Let us develop the concept by considering a more general version of the pricing game between Rainbow's End and B. B. Lean. In chapter 3 we allowed each company the choice of just two prices for the shirt, namely \$80 and \$70. We also recognized the strength of the temptation for each to cut the price. Let us therefore allow more choices in a lower range, going in \$1 steps from \$42 to \$38.† In the earlier example, when both charge \$80, each sells 1,200 shirts. If one of them cuts its price by \$1 while the other holds its price unchanged, then the price cutter gains 100 cus-

* For those readers who have not seen the movie *A Beautiful Mind*, starring Russell Crowe as Nash, or read Sylvia Nasar's best-selling book of the same name, we should add that John Nash developed his fundamental concept of equilibrium in games around 1950 and went on to make contributions of equal or greater importance in mathematics. After several decades of severe mental illness, he recovered and was awarded the 1994 Nobel Prize in Economics. This was the first Nobel Prize for game theory.

† The \$1 increment and the restricted range of prices are chosen merely to simplify our entrée into this game by keeping the number of strategies available to each player finite. Later in the chapter we will consider briefly the case where each firm can choose its price from a continuous range of values.

tomers, 80 of whom shift from the other firm and 20 of whom shift from some other firm that is not a part of this game or decide to buy a shirt when they would otherwise not have done so. If both firms reduce their price by \$1, existing customers stay put, but each gains 20 new ones. So when both firms charge \$42 instead of \$80, each gains $38 \times 20 = 760$ customers above the original 1,200. Then each sells 1,960 shirts and makes a profit of $(42 - 20) \times 1,960 = 43,120$ dollars. Doing similar calculations for the other price combinations, we have the game table below.

		B. B. Lean's price				
		42	41	40	39	38
Rainbow's End's price	42	43,120	43,260	43,200	42,940	42,480
	41	41,360	41,580	41,600	41,420	41,040
	40	39,600	39,900	40,000	39,900	39,600
	39	37,840	38,220	38,400	38,380	38,160
	38	36,080	36,540	36,800	36,860	36,720

The table may seem daunting but is in fact easy to construct using Microsoft Excel or any other spreadsheet program.

TRIP TO THE GYM NO. 2

Try your hand at constructing this table in Excel.

Best Responses

Consider the thinking of RE's executives in charge of setting prices. (From now on, we will simply say "RE's thinking," and similarly for BB.) If RE believes that BB is choosing \$42, then RE's profits from choosing various possible prices are given by the numbers in the southwest corners of the first column of profits in the above table. Of those five numbers, the highest is \$43,260, corresponding to RE's price \$41. Therefore this is RE's "best response"

to BB's choice of \$42. Similarly, RE's best response is \$40 if it believes that BB is choosing \$41, \$40, or \$39, and \$39 if it believes BB is choosing \$38. We show these best-response profit numbers in bold italics for clarity. We also show BB's best responses to the various possible prices of RE, using bold, italicized numbers in the northeast corners of the appropriate cells.

Before proceeding, we must make two remarks about best responses. First, the term itself requires clarification. The two firms' choices are simultaneous. Therefore, unlike the situation in chapter 2, each firm is not observing the other's choice and then "responding" with its own best choice given the other firm's actual choice. Rather, each firm is formulating a belief (which may be based on thinking or experience or educated guesswork) about what the other firm is choosing, and responding to this belief.

Second, note that it is not always best for one firm to undercut the other's price. If RE believes that BB is choosing \$42, RE should choose a lower price, namely \$41; but if RE believes that BB is choosing \$39, RE's best response is higher, namely \$40. In choosing its best price, RE has to balance two opposing considerations: undercutting will increase the quantity it sells, but will leave it a lower profit margin per unit sold. If RE believes that BB is setting a very low price, then the reduction in RE's profit margin from undercutting BB may be too big, and RE's best choice may be to accept a lower sales volume to get a higher profit margin on each shirt. In the extreme case where RE thinks BB is pricing at cost, namely \$20, matching this price will yield RE zero profit. RE does better to choose a higher price, keeping some loyal customers and extracting some profit from them.

Nash Equilibrium

Now return to the table and inspect the best responses. One fact immediately stands out: one cell, namely the one where each firm charges \$40, has both of its numbers in bold italics, yielding a profit of \$40,000 to each firm. If RE believes that BB is choosing the price of \$40, then its own best price is \$40, and vice versa. If the two firms choose to price their shirts at \$40 each, the beliefs of

each about the other's price are confirmed by the actual outcome. Then there would be no reason for one firm to change its price if the truth about the other firm's choice were somehow revealed. Therefore these choices constitute a stable configuration in the game.

Such an outcome in a game, where the action of each player is best for him given his beliefs about the other's action, and the action of each is consistent with the other's beliefs about it, neatly squares the circle of thinking about thinking. Therefore it has a good claim to be called a resting point of the players' thought processes, or an equilibrium of the game. Indeed, this is just a definition of Nash equilibrium.

To highlight the Nash equilibrium, we shade its cell in gray and will do the same in all the game tables that follow.

The price-setting game in chapter 3, with just two price choices of \$80 and \$70, was a prisoners' dilemma. The more general game with several price choices shares this feature. If both firms could make a credible, enforceable agreement to collude, they could both charge prices considerably higher than the Nash equilibrium price of \$40, and this would yield larger profits to both. As we saw in chapter 3, a common price of \$80 gives each of them \$72,000, as opposed to only \$40,000 in the Nash equilibrium. The result should impress upon you how consumers can suffer if an industry is a monopoly or a producers' cartel.

In the above example, the two firms were symmetrically situated in all relevant matters of costs and in the quantity sold for each combination of own and rival prices. In general this need not be so, and in the resulting Nash equilibrium the two firms' prices can be different.

For those of you who want to acquire a better grasp of the methods and the concepts, we offer this as an "exercise"; casual readers should feel free to peek at the answer in the workouts.

The pricing game has many other features, but they are more

TRIP TO THE GYM NO. 3

Suppose Rainbow's End locates a cheaper source for its shirts, so its cost per shirt goes down from \$20 to \$11.60, while B. B. Lean's cost remains at \$20. Recalculate the payoff table and find the new Nash equilibrium.

complex than the material so far. Therefore we postpone them to a position later in this chapter. To conclude this section, we make a few general remarks about Nash equilibria.

Does every game have a Nash equilibrium? The answer is essentially yes, provided we generalize the concept of actions or strategies to allow mixing of moves. This was Nash's famous theorem. We will develop the idea of mixing moves in the next chapter. Games that have no Nash equilibrium, even when mixing is allowed, are so complex or esoteric that we can safely leave them to very advanced treatments of game theory.

Is Nash equilibrium a good solution for simultaneous-move games? We will discuss some arguments and evidence bearing on this issue later in this chapter, and our answer will be a guarded yes.

Does every game have a unique Nash equilibrium? No. In the rest of this chapter we will look at some important examples of games with multiple Nash equilibria and discuss the new issues they raise.

Which Equilibrium?

Let us try Nash's theory on the hunting game. Finding best responses in the hunting game is easy. Fred should simply make the same choice that he believes Barney is choosing. Here is the result.

		Barney's choice		
		Stag	Bison	Rabbit
Fred's choice	Stag	3, 3	0, 0	1, 0
	Bison	0, 0	3, 3	1, 0
	Rabbit	0, 1	0, 1	1, 1

So the game has three Nash equilibria.* Which of these will emerge as the outcome? Or will the two fail to reach any of the equilibria at all? The idea of Nash equilibrium does not by itself give the answers. Some additional and different consideration is needed.

If Fred and Barney had met at the stag party of a mutual friend, that might make the choice of Stag more prominent in their minds. If the ritual in their society is that as the head of the family sets out for the day's hunting he calls out in farewell, "Bye, son," the choice of Bison might be prominent. But if the ritual is for the family to call out in farewell "Be safe," the prominence might attach to the safer choice that guarantees some meat regardless of what the other chooses, namely rabbit hunting.

But what, precisely, constitutes "prominence"? One strategy, say Stag, may be prominent in Fred's mind, but that is not enough for him to make that choice. He must ask himself whether the same strategy is also prominent for Barney. And that in turn involves asking whether Barney will think it prominent to Fred. Selecting among multiple Nash equilibria requires resolution of a similar problem of thinking about thinking as does the concept of Nash equilibrium itself.

To square the circle, the "prominence" must be a multilevel back-and-forth concept. For the equilibrium to be selected successfully when the two are thinking and acting in isolation, it must be obvious to Fred that it is obvious to Barney that it is obvious to Fred . . . that is the right choice. If an equilibrium is obvious ad infinitum in this way, that is, if the players' expectations converge upon it, we call it a *focal point*. The development of this concept was just one of Thomas Schelling's many pioneering contributions to game theory.

Whether a game has a focal point can depend on many circumstances, including most notably the players' common experiences, which may be historical, cultural, linguistic, or purely accidental. Here are some examples.

* If mixing moves is allowed, there are other Nash equilibria as well. But they are somewhat strange and mostly of academic interest. We discuss them briefly in chapter 5.

We begin with one of Schelling's classics. Suppose you are told to meet someone in New York City on a specific day but without being told where or at what time. You don't even know who the other person is, so you cannot contact him/her in advance (but you are told how you would identify each other if and when you do meet). You are also told that the other person has been given identical instructions.

Your chances of success might seem slim; New York City is huge, and the day is long. But in fact people in this situation succeed surprisingly often. The time is simple: noon is the obvious focal point; expectations converge on it almost instinctively. The location is harder, but there are just a few landmark locations on which expectations can converge. This at least narrows down the choices considerably and improves the chances of a successful meeting.

Schelling conducted experiments in which the subjects were from the Boston or New Haven areas. In those days they traveled to New York by train and arrived at Grand Central Station; for them the clock in that station was focal. Nowadays, many people would think the Empire State Building is a focal point because of the movie *Sleepless in Seattle* (or *An Affair to Remember*); others would think Times Square the obvious "crossroads of the world."

One of us (Nalebuff) performed this experiment in an ABC *Primetime* program titled *Life: The Game*.¹ Six pairs of mutual strangers were taken to different parts of New York and told to find others about whom they had no information except that the other pair would be looking for them under similar conditions. The discussions within each pair followed Schelling's reasoning remarkably well. Each thought about what they believed would be the obvious places to meet and about what others would think they were thinking: each team, say team A, in its thinking recognized the fact that another team, say B, was simultaneously thinking about what was obvious to A. Eventually, three of the pairs went to the Empire State Building and the other three to Times Square. All chose noon for the time. There remained some further issues to be sorted out: the Empire State Building has observation decks on two different levels, and Times Square is a big place. But

with a little ingenuity, including a display of signs, all six pairs were successful in meeting.*

What is essential for success is not that the place is obvious to you, or obvious to the other team, but that it is obvious to each that it is obvious to the others that . . . And, if the Empire State Building has this property, then each team has to go there even though it may be inconvenient for them to get there, because it is the only place each can expect the other team to be. If there were just two teams, one of them might think the Empire State Building the obvious focal point and the other might think Times Square equally obvious; then the two would fail to meet.

Professor David Kreps of Stanford Business School conducted the following experiment in his class. Two students were chosen to play the game, and each had to make his/her choice without any possibility of communication with the other. Their job was to divide up a list of cities between them. One student was assigned Boston, and the other was assigned San Francisco (and these assignments were public so that each knew the other's city). Each was then given a list of nine other U.S. cities—Atlanta, Chicago, Dallas, Denver, Houston, Los Angeles, New York, Philadelphia, and Seattle—and asked to choose a subset of these cities. If their choices resulted in a complete and nonoverlapping division, both got a prize. But if their combined list missed a city or had any duplicates, then they both got nothing.

How many Nash equilibria does this game have? If the student assigned Boston chooses, say, Atlanta and Chicago, while the student assigned San Francisco chooses the rest (Dallas, Denver, Houston, Los Angeles, New York, Philadelphia, and Seattle), that

* One of the pairs sat outside the Empire State Building for almost an hour, waiting for noon. If they had decided to wait inside, they would have done much better. It was also instructive that the teams of men went running from one site to another (Port Authority, Penn Station, Times Square, Grand Central, Empire State Building) without any sign that would help them be found by another team. As might be expected, the male teams even crossed paths without recognizing each other. In contrast, the all-women teams made signs and hats. They picked a single spot and waited to be found.

is a Nash equilibrium: given the choice of one, any change in the choice of the other will create either an omission or an overlap and would lower the payoff to the deviator. The same argument applies if, say, one chooses Dallas, Los Angeles, and Seattle while the other chooses the other six. In other words, there are as many Nash equilibria as there are ways of dividing up the list of nine numbers into two distinct subsets. There are 2^9 , or 512, such ways; therefore the game has a huge number of Nash equilibria.

Can the players' expectations converge to create a focal point? When both players were Americans or long-time U.S. residents, over 80 percent of the time they chose the division geographically; the student assigned Boston chose all the cities east of the Mississippi and the student assigned San Francisco chose those west of the Mississippi.* Such coordination was much less likely when one or both students were non-U.S. residents. Thus nationality or culture can help create a focal point. When Kreps's pairs lacked such common experience, choices were sometimes made alphabetically, but even then there was no clear dividing point. If the total number of cities was even, an equal split might be focal, but with nine cities, that is not possible. Thus one should not assume that players will always find a way to select one of multiple Nash equilibria by a convergence of expectations; failure to find a focal point is a distinct possibility.†

Next, suppose each of two players is asked to choose a positive integer. If both choose the same number, both get a prize. If the two choose different numbers, neither gets anything. The overwhelmingly frequent choice is 1: it is the first among the whole

* Perhaps in a few years' time this will no longer work, if the news stories about the deterioration of geographic knowledge among American schoolchildren are true.

† The game of dividing cities might seem uninteresting or irrelevant, but think of two firms that are trying to divide up the U.S. market between them to allow each to enjoy an uncontested monopoly in its assigned territory. U.S. antitrust laws forbid explicit collusion. To arrive at a tacit understanding requires a convergence of expectations. Kreps's experiment suggests that two American firms may achieve this better than could an American firm and a foreign firm.

numbers (positive integers), it is the smallest, and so on; therefore, it is focal. Here the reason for its salience is basically mathematical.

Schelling gives the example of two or more people who have gone to a crowded place together and get separated. Where should each go in the expectation of finding the other? If the place, say a department store or a railway station, has a Lost and Found window, it has a good claim to be focal. Here the reason for its salience is linguistic. Sometimes meeting places are deliberately created to guarantee a convergence of expectations; for example, many railway stations in Germany and Switzerland have a well-signposted *Treffpunkt* (meeting point).

What is neat about the game of meeting is not just that the two players find each other but that the focal point ends up being relevant to so many strategic interactions. Probably the most important is the stock market. John Maynard Keynes, arguably the twentieth century's most famous economist, explained its behavior by analogy with a newspaper contest that was common in his time, where a number of photographs of faces were presented, and readers had to guess which face the majority of other voters would judge the most beautiful.² When everyone thinks along these lines, the question becomes which face most people think that most others will think that most others will think . . . is the most beautiful. If one contestant was significantly more beautiful than all the others, this could provide the necessary focal point. But the reader's job was rarely that easy. Imagine instead that the hundred finalists were practically indistinguishable except for the color of their hair. Of the hundred, only one is a redhead. Would you pick the redhead?

The aim becomes not to make any absolute judgment of beauty but to find a focal point of this process of thinking. How do we agree on that? The reader must figure out the realized convention without the benefit of communication. "Pick the most beautiful" might be the stated rule, but that could be significantly more difficult than picking the redhead, or the one with an interesting gap between her two front teeth (Lauren Hutton) or the mole (Cindy Crawford). Anything that distinguishes becomes a focal point and allows people's expectations to converge. For this reason, we

should not be surprised that many of the world's top models do not have perfect features; rather, they are almost perfect but have some interesting flaw that gives their look a personality and a focal point.

Keynes used the beauty contest as a metaphor for the stock market, where each investor wants to buy the stocks that will rise in price, which means the stocks that investors, in general, think will appreciate. The hot stock is the one that everyone thinks that everyone else thinks . . . is the hot stock. There can be different reasons why different sectors or stocks become hot at different times—a well-publicized initial public offering, a famous analyst's recommendation, and so on. The focal point concept also explains the attention paid to round numbers: 10,000 for the Dow, or 2,500 for the Nasdaq. These indexes are just values of a specified portfolio of stocks. A number like 10,000 does not have any intrinsic meaning; it serves as a focal point only because expectations can converge more easily on round numbers.

The point of all this is that equilibrium can easily be determined by whim or fad. There is nothing fundamental that guarantees the most beautiful contestant will be chosen or the best stock will appreciate the fastest. There are some forces that work in the right direction. High forecast earnings are similar to the beauty contestant's complexion—one of the many necessary but by no means sufficient conditions needed to anchor otherwise arbitrary whims and fads.

Many mathematical game theorists dislike the dependence of an outcome on historical, cultural, or linguistic aspects of the game or on purely arbitrary devices like round numbers; they would prefer the solution be determined purely by the abstract mathematical facts about the game—the number of players, the strategies available to each, and the payoffs to each in relation to the strategy choices of all. We disagree. We think it entirely appropriate that the outcome of a game played by humans interacting in a society should depend on the social and psychological aspects of the game.

Think of the example of bargaining. Here the players' interests seem to be totally conflicting; a larger share for one means a smaller share for the other. But in many negotiations, if the two

parties fail to agree, neither will get anything and both may suffer serious damage, as happens when wage bargaining breaks down and a strike or a lockout ensues. The two parties' interests are aligned to the extent that both want to avoid such disagreement. They can do so if they can find a focal point, with the common expectation that neither will concede anything beyond that point. That is why a 50:50 split is so often observed. It is simple and clear, it has the advantage of appearing fair, and, once such considerations get a foothold, it serves for the convergence of expectations.

Consider the problem of excessive compensation of CEOs. Often a CEO really cares about prestige. Whether the person gets paid \$5 million or \$10 million won't really have a big impact on the person's life. (That's easy for us to say from where we sit, where both numbers are quite abstract.) What's the meeting place that the CEOs care about? It is being better than average. Everyone wants to be in the top half. They all want to meet there. The problem is that this meeting spot only allows in half of the folks. But the way they get around this is via escalating pay. Every firm pays its CEO above last year's average, so everyone can think they have an above-average CEO. The end result is wildly escalating CEO salaries. To solve the problem, we need to find some other focal meeting point. For example, historically CEOs got prestige in their community via public service. Competing in that dimension was good all around. The current focal point on pay was created by *Business Week* surveys and compensation consultants. Changing it won't be easy.

The issue of fairness is also one of choosing a focal point. The Millennium Development Goals and Jeff Sachs's book *The End of Poverty* emphasize that contributing 1 percent of gross domestic product (GDP) to development will end poverty by 2025. The key point here is that the focal point of contributions is based on a percentage of income, not an absolute amount. Thus rich countries have a bigger obligation to contribute than the less rich. The apparent fairness of this can contribute to the convergence of expectations. Whether the promised funds will actually materialize remains to be seen.

BATTLES AND CHICKENS

In the hunting game, the two players' interests are perfectly aligned; both prefer one of the big-game equilibria, and the only question is how they can coordinate their beliefs on a focal point. We now turn to two other games, which also have non-unique Nash equilibria, but have an element of conflicting interests. Each leads to different ideas about strategy.

Both of these games date from the 1950s and have stories that fit those times. We will illustrate them using variants of the game between our Stone Age hunters, Fred and Barney. But we will relate the original sexist stories too, partly because they explain the names that have come to be attached to these games and partly for the amusement value of looking back on the quaint thoughts and norms of old times.

The first game is generically called battle of the sexes. The idea is that a husband and wife have different preferences in movies, and the two available choices are very different. The husband likes lots of action and fighting; he wants to see *300*. The wife likes three-handkerchief weepies; her choice is *Pride & Prejudice* (or *A Beautiful Mind*). But both prefer watching either movie in the other's company to watching any movie on their own.

In the hunting version, remove the Rabbit choice and keep only Stag and Bison. But suppose Fred prefers stag meat and rates the outcome of a jointly conducted stag hunt 4 instead of 3, while Barney has the opposite preference. The revised game payoff table is as shown below.

		Barney's choice	
		Stag	Bison
Fred's choice	Stag	4, 3	0, 0
	Bison	0, 0	3, 4

As usual, best responses are shown in bold italics. We see at once that the game has two Nash equilibria, one where both choose Stag, and the other where both choose Bison. Both players prefer to have either equilibrium outcome than to hunt alone in one of the two nonequilibrium outcomes. But they have conflicting preferences over the two equilibria: Fred would rather be in the Stag equilibrium and Barney in the Bison equilibrium.

How might one or the other outcome be sustained? If Fred can somehow convey to Barney that he, Fred, is credibly and unyieldingly determined to choose Stag, then Barney must make the best of the situation by complying. However, Fred faces two problems in using such a strategy.

First, it requires some method of communication before the actual choices are made. Of course, communication is usually a two-way process, so Barney might try the same strategy. Fred would ideally like to have a device that will let him send messages but not receive them. But that is not without its own problems; how can Fred be sure that Barney has received and understood the message?

Second, and more important, is the problem of credibly conveying an unyielding determination. This can be faked, and Barney might put it to the test by defying Fred and choosing Bison, which would leave Fred with a pair of bad choices: give in and choose Bison, which leads to humiliation and destruction of reputation, or go ahead with the original choice of Stag, which means missing the opportunity of the joint hunt, getting zero meat, and ending up with a hungry family.

In chapter 7 we will examine some ways that Fred could make his determination credible and achieve his preferred outcome. But we will also examine some ways that Barney could undermine Fred's commitment.

If they have two-way communication before the game is played, this is essentially a game of negotiation. The two prefer different outcomes, but both prefer some agreement to complete disagreement. If the game is repeated, they may be able to agree to a compromise—for example, alternate between the two grounds on alternate days. Even in a single play, they may be able to achieve a compromise in the sense of a statistical average by tossing a coin

and choosing one equilibrium if it comes up heads and the other equilibrium if it comes up tails. We will devote an entire chapter to the important subject of negotiation.

The second classic game is called chicken. In the standard telling of this story, two teenagers drive toward each other on a straight road, and the first one to swerve to avoid a collision is the loser, or chicken. If both keep straight, however, they crash, and that is the worst outcome for both. To create a game of chicken out of the hunting situation, remove the Stag and Bison choices, but suppose there are two areas for rabbit hunting. One, located to the south, is large but sparse; both can go there and each will get 1 of meat. The other, located to the north, is plentiful but small. If just one hunter goes there, he can get 2 of meat. If both go there, they will merely interfere and start fighting with each other and get nothing. If one goes north and the other goes south, the one who goes north will enjoy his 2 of meat. The one going south will get his 1. But his and his family's feeling of envy for the other who comes back at the end of the day with 2 will reduce his enjoyment, so we will give him a payoff of only $1/2$ instead of 1. This yields the game payoff table shown below.

		Barney's choice	
		North	South
Fred's choice	North	0 0	<i>2</i> <i>$1/2$</i>
	South	<i>$1/2$</i> <i>2</i>	1 1

As usual, best responses are shown in bold italics. We see at once that the game has two Nash equilibria, with one player going north and the other going south. The latter is then the chicken; he has made the best of a bad situation in responding to the other's choice of North.

Both games, the battle of the sexes and chicken, have a mixture of

common and conflicting interests: in both, the two players agree in preferring an equilibrium outcome to a nonequilibrium outcome, but they disagree as to which equilibrium is better. This conflict is sharper in chicken, in the sense that if each player tries to achieve his preferred equilibrium, both end up in their worst outcome.

Methods for selecting one of the equilibria in chicken are similar to those in the battle of the sexes. One of the players, say Fred, may make a commitment to choosing his preferred strategy, namely going north. Once again, it is important to make this commitment credible and to ensure that the other player knows it. We will consider commitments and their credibility more fully in chapters 6 and 7.

There is also the possibility of compromise in chicken. In a repeated interaction Fred and Barney may agree to alternate between North and South; in a single play, they may use a coin toss or other randomizing method to decide who gets North.

Finally, chicken shows a general point about games: even though the players are perfectly symmetric as regards their strategies and payoffs, the Nash equilibria of the game can be asymmetric, that is, the players choose different actions.

A LITTLE HISTORY

In the course of developing examples in this chapter and the one before it, we have introduced several games that have become classics. The prisoners' dilemma, of course, everyone knows. But the game of the two Stone Age hunters trying to meet is almost equally well known. Jean-Jacques Rousseau introduced it in an almost identical setting—of course he did not have *Flintstones* characters to add color to the story.

The hunters' meeting game differs from the prisoners' dilemma because Fred's best response is to take the same action as Barney does (and vice versa), whereas in a prisoners' dilemma game Fred would have a dominant strategy (just one action—for example, Rabbit—would be his best choice regardless of what Barney does) and so would Barney. Another way to express the difference is to say

that in the meeting game, Fred would go stag hunting if he had the assurance, whether by direct communication or because of the existence of a focal point, that Barney would also go stag hunting, and vice versa. For this reason, the game is often called the *assurance game*.

Rousseau did not put his idea in precise game-theoretic language, and his phrasing leaves his meaning open to different interpretations. In Maurice Cranston's translation, the large animal is a deer, and the statement of the problem is as follows: "If it was a matter of hunting a deer, everyone well realized that he must remain faithfully at his post; but if a hare happened to pass within the reach of one of them, we cannot doubt that he would have gone off in pursuit of it without scruple and, having caught his own prey, he would have cared very little about having caused his companions to lose theirs."³ Of course if the others were going for the hare, then there would be no point in any one hunter's attempting the deer. So the statement seems to imply that each hunter's dominant strategy is to go after a hare, which makes the game a prisoners' dilemma. However, the game is more commonly interpreted as an assurance game, where each hunter prefers to join the stag hunt if all the others are doing likewise.

In the version of chicken made famous by the movie *Rebel Without a Cause*, two teenagers drive their cars in parallel toward a cliff; the one who first jumps out of his car is the chicken. The metaphor of this game was used for nuclear brinkmanship by Bertrand Russell and others. The game was discussed in detail by Thomas Schelling in his pioneering game-theoretic analysis of strategic moves, and we will pick this back up in chapter 6.

To the best of our knowledge, the battle of the sexes game does not have such roots in philosophy or popular culture. It appears in the book *Games and Decisions* by R. Duncan Luce and Howard Raiffa, an early classic on formal game theory.⁴

FINDING NASH EQUILIBRIA

How can we find Nash equilibrium for a game? In a table, the worst-case method is cell-by-cell inspection. If both of the pay-

off entries in a cell are best responses, the strategies and payoffs for that cell constitute a Nash equilibrium. If the table is large, this procedure can get tedious. But God made computers precisely to rescue humans from the tedium of inspection and calculation. Software packages to find Nash equilibria are readily available.⁵

Sometimes there are shortcuts; we now describe one that is often useful.

Successive Elimination

Return to the pricing game between Rainbow's End and B. B. Lean. Here again is the table of payoffs:

		B. B. Lean's price				
		42	41	40	39	38
Rainbow's End's price	42	43,120	41,360	43,200	42,940	42,480
	41	41,360	41,580	41,600	41,420	41,040
	40	39,600	39,900	40,000	39,900	39,600
	39	37,840	38,220	38,400	38,380	38,160
	38	36,080	36,540	36,800	36,860	36,720

RE does not know what price BB is choosing. But it can figure out what price or prices BB is not choosing: BB will never set its price at \$42 or \$38. There are two reasons (both of which apply in our example, but in other situations only one may apply).⁶

First, each of these strategies is uniformly worse for BB than another available strategy. No matter what it thinks RE is choosing, \$41 is better for BB than \$42, and \$39 is better than \$38. To see this, consider the \$41 versus \$42 comparison; the other is similar. Look at the five numbers for BB's profits from choosing \$41 (shaded in dark gray) versus those from \$42 (shaded in light gray).

For each of RE's five possible choices, BB's profit from choosing \$42 is smaller than that from choosing \$41:

$$43,120 < 43,260,$$

$$41,360 < 41,580,$$

$$39,600 < 39,900,$$

$$37,840 < 38,220,$$

$$36,080 < 36,540.$$

So no matter what BB expects RE to do, BB will never choose \$42, and RE can confidently expect BB to rule out the \$42 strategy, and, likewise, \$38.

When one strategy, say A, is uniformly worse for a player than another, say B, we say that A is *dominated* by B. If such is the case, that player will never use A, although whether he uses B remains to be seen. The other player can confidently proceed in thinking on this basis; in particular, he need not consider playing a strategy that is the best response only to A. When solving the game, we can remove dominated strategies from consideration. This reduces the size of the game table and simplifies the analysis.*

The second avenue for elimination and simplification is to look for strategies that are *never best responses* to anything the other player might be choosing. In this example, \$42 is never BB's best response to anything RE might be choosing within the range we are considering. So, RE can confidently think, "No matter what BB is thinking about my choice, it will never choose \$42."

Of course, anything that is dominated is a never best response. It is more instructive to look at BB's option to price at \$39. This can *almost* be eliminated for being a never best response. A price of \$39 is only a best response to an RE price of \$38. Once we know that \$38 is dominated, then we can conclude that a BB price of \$39 will never be a best response to anything RE will ever play. The

* If A is dominated by B, then conversely, B dominates A. So if A and B were the only two strategies available to that player, B would be a dominant strategy. With more than two strategies available, it is possible that A is dominated by B, but B is not dominant, because it does not dominate some third strategy C. In general, elimination of dominated strategies may be possible even in games that do not have any dominant strategies.

advantage, then, of looking for never best responses is that you are able to eliminate strategies that are not dominated but would still never be chosen.

We can perform a similar analysis for the other player. RE's \$42 and \$38 strategies are eliminated, leaving us with a 3-by-3 game table:

		B. B. Lean's price		
		41	40	39
Rainbow's End's price	41	41,580 41,580	41,600 39,900	41,420 38,220
	40	39,900 41,600	40,000 40,000	39,900 38,400
	39	38,220 41,420	38,400 39,900	38,380 38,380

In this simplified game, each firm has a dominant strategy, namely \$40. Therefore our Rule 2 (from chapter 3) indicates that as a solution for the game.

The \$40 strategy was not dominant in the original larger game; for example, if RE thought that BB would charge \$42, then its profits from setting its own price at \$41, namely \$43,260, would be more than its profits from choosing \$40, namely \$43,200. The elimination of some strategies can open up the way to eliminate more in a second round. Here just two rounds sufficed to pin down the outcome. In other examples it may take more rounds, and even then the range of outcomes may be narrowed somewhat but not all the way to uniqueness.

If successive elimination of dominated strategies (or never-best-response strategies) and choice of dominant strategies does lead to a unique outcome, that is a Nash equilibrium. When this works, it is an easy way to find Nash equilibria. Therefore we summarize our discussion of finding Nash equilibria into two rules:

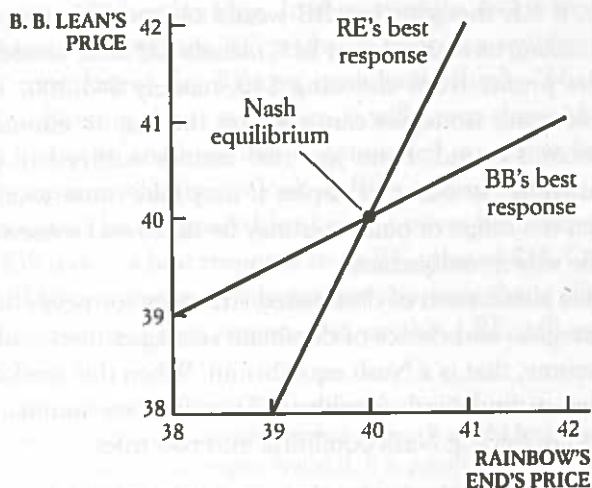
RULE 3: Eliminate from consideration any dominated strategies and strategies that are never best responses, and go on doing so successively.

RULE 4: Having exhausted the simple avenues of looking for dominant strategies or ruling out dominated ones, next search all the cells of the game table for a pair of mutual best responses in the same cell, which is a Nash equilibrium of the game.

GAMES WITH INFINITELY MANY STRATEGIES

In each of the versions of the pricing game we discussed so far, we allowed each firm only a small number of price points: only \$80 and \$70 in chapter 3, and only between \$42 and \$38 in \$1 steps in this chapter. Our purpose was only to convey the concepts of the prisoners' dilemma and Nash equilibrium in the simplest possible context. In reality, prices can be any number of dollars and cents, and for all intents and purposes it is as if they can be chosen over a continuous range of numbers.

Our theory can cope with this further extension quite easily, using nothing more than basic high-school algebra and geometry. We can show the prices of the two firms in a two-dimensional graph, measuring RE's price along the horizontal or X axis and BB's price along the vertical or Y axis. We can show the best



responses in this graph instead of showing bold italic profit numbers in a game table of discrete price points.

We do this for the original example where the cost of each shirt to each store was \$20. We omit the details of the mathematics and merely tell you the result.⁷ The formula for BB's best response in terms of RE's price (or BB's belief about the price RE is setting) is

$$\text{BB's best response price} = 24 + 0.4 \times \text{RE's price (or BB's belief about it)}.$$

This is shown as the flatter of the two lines in the above graph. We see that for each \$1 cut in RE's price, BB's best response should be to cut its own price but by less, namely 40 cents. This is the result of BB's calculation, striking the best balance between losing customers to RE and accepting a lower profit margin.

The steeper of the two curves in the figure is RE's best response to its belief about BB's price. Where the two curves intersect, the best response of each is consistent with the other's beliefs; we have a Nash equilibrium. The figure shows that this occurs when each firm charges \$40. Moreover, it shows that this particular game has exactly one Nash equilibrium. Our finding a unique Nash equilibrium in the table where prices had to be multiples of \$1 was not an artificial consequence of that restriction.

Such graphs or tables that allow much more detail than we could in the simple examples are a standard method for computing Nash equilibria. The calculation or graphing can quickly get too complicated for paper-and-pencil methods, and too boring besides, but that's what computers are for. The simple examples give us a basic understanding of the concept, and we should reserve our human thinking skills for the higher-level activity of assessing its usefulness. Indeed, that is our very next topic.

A BEAUTIFUL EQUILIBRIUM?

John Nash's equilibrium has a lot of conceptual claim to be the solution of a game where each player has the freedom of choice. Perhaps the strongest argument in its favor takes the form of a

counterargument to any other proposed solution. A Nash equilibrium is a configuration of strategies where each player's choice is his best response to the other player's choice (or the other players' choices when there are more than two players in the game). If some outcome is not a Nash equilibrium, at least one player must be choosing an action that is not his best response. Such a player has a clear incentive to deviate from that action, which would destroy the proposed solution.

If there are multiple Nash equilibria, we do need some additional method for figuring out which one will emerge as the outcome. But that just says we need Nash plus something else; it does not contradict Nash.

So we have a beautiful theory. But does it work in practice? One answers this question by looking for instances where such games are played in the real world, or by creating them in a laboratory setting and then comparing the actual outcomes against the predictions of the theory. If the agreement is sufficiently good, that supports the theory; if not, the theory should be rejected. Simple, right? In fact the process turns complicated very quickly, both in implementation and in interpretation. The results are mixed, with some reasons for optimism for the theory but also some ways in which the theory must be augmented or altered.

The two methods—observation and experiment—have different merits and flaws. Laboratory experiments allow proper scientific “control.” The experimenters can specify the rules of the game and the objectives of the participants quite precisely. For example, in pricing games where the subjects play the roles of the managers of the firms, we can specify the costs of the two firms and the equations for the quantities each would sell in relation to the prices both charge, and give the players the appropriate motivation by paying them in proportion to the profits they achieve for their firm in the game. We can study the effects of a particular factor, keeping all other things constant. By contrast, games that occur in real life have too many other things going on that we cannot control and too many things about the players—their true motivations, the firms' costs of production, and so on—that we

do not know. That makes it hard to make inferences about the underlying conditions and causes by observing the outcomes.

On the other hand, real-world observations do have some advantages. They lack the artificiality of laboratory experiments, in which the subjects are usually students, who have no previous experience in business or the similar applications that motivate the games. Many are novices even to the setting of the laboratory where the games are staged. They have to understand the rules of the game and then play it, all in a matter of an hour or two. Think how long it took you to figure out how to play even simple board games or computer games; that will tell you how naïve the play in such settings can be. We already discussed some examples of this problem in chapter 2. A second issue concerns incentives. Although the experimenter can give the students the correct incentives by designing the structure of their monetary payments to fit their performance in the game, the sizes of the payments are usually small, and even college students may not take them sufficiently seriously. By contrast, business games and even professional sports in the real world are played by experienced players for large stakes.

For these reasons, one should not rely solely on any one form of evidence, whether it supports or rejects a theory, but should use both kinds and learn from each. With these cautions in mind, let us see how the two types of empirical approaches do.

The field of industrial organization in economics provides the largest body of empirical testing of game-theoretic competition among firms. Industries like auto manufacturing have been studied in depth. These empirical investigators start with several handicaps. They do not know the firms' costs and demands from any independent source, and must estimate these things from the same data that they want to use for testing the pricing equilibrium. They do not know precisely how the quantities sold by each firm depend on the prices charged by all. In the examples in this chapter, we simply assumed a linear relationship, but the real-world counterparts (demand functions, in the jargon of economics) can be nonlinear in quite complicated ways. The investigator

must assume some specific form of the nonlinearity. Real-life competition among firms is not just about prices; it has many other dimensions—advertising, investment, research and development. Real-life managers may not have the pure and simple aims of profit (or shareholder value) maximization that economic theory usually assumes. And competition among firms in real life extends over several years, so an appropriate combination of backward reasoning and Nash equilibrium concepts must be specified. And many other conditions, such as income and costs, change from one year to the next, and firms enter or exit the industry. The investigator must think about what all these other things might be and make proper allowance for (control for, in statistical jargon) their effects on quantities and prices. Real-world outcomes are also affected by many random factors and so, uncertainty must be allowed for.

A researcher must make a choice in each of these matters and then derive equations that capture and quantify all the relevant effects. These equations are then fitted to the data, and statistical tests performed to see how well they do. Then comes an equally difficult problem: What does one conclude from the findings? For example, suppose the data do not fit your equations very well. Something in your specification that led to the equations was not correct, but what was it? It could be the nonlinear form of the equations you chose; it could be the exclusion of some relevant variable, like income, or of some relevant dimension of competition, like advertising; or it could be that the Nash equilibrium concept used in your derivations is invalid. Or, it could be a combination of all these things. You cannot conclude that Nash equilibrium is incorrect when something else might be wrong. (But you would be right to raise your level of doubt about the equilibrium concept.)

Different researchers have made different choices in all these matters and, predictably, have found different results. After a thorough survey of this research, Peter Reiss and Frank Wolak of Stanford University give a mixed verdict: "The bad news is that the underlying economics can make the empirical models extremely

complex. The good news is that the attempts so far have begun to define the issues that need to be addressed."⁸ In other words, more research is needed.

Another active area for empirical estimation concerns auctions where a small number of strategically aware firms interact in bidding for things like bandwidths in the airwave spectrum. In these auctions, asymmetry of information is a key issue for the bidders and also for the auctioneer. Therefore we postpone the discussion of auctions to chapter 10, after we have examined the general issues of information in games in chapter 8. Here we merely mention that empirical estimation of auction games is already having considerable success.⁹

What do laboratory experiments have to say about the predictive power of game theory? Here the record is also mixed. Among the earliest experiments were the markets set up by Vernon Smith. He found surprisingly good results for game theory as well as for economic theory: small numbers of traders, each with no direct knowledge of the others' costs or values, could achieve equilibrium exchanges very quickly.

Other experiments with different kinds of games yielded outcomes that seemed contradictory to theoretical predictions. For example, in the ultimatum game, where one player makes a take-it-or-leave-it offer to the other for dividing a given sum between the two, the offers were surprisingly generous. And in prisoners' dilemmas, good behavior occurred far more frequently than theory might lead people to believe. We discussed some of these findings in chapters 2 and 3. Our general conclusion was that the participants in these games had different preferences or valuations than the purely selfish ones that used to be the natural assumption in economics. This is an interesting and important finding on its own; however, once the realistic "social" or "other-regarding" preferences are allowed for, the theoretical concepts of equilibrium—backward reasoning in sequential-move games and Nash in simultaneous-move games—yield generally good explanations of the observed outcomes.

When a game does not have a unique Nash equilibrium, the

players have the additional problem of locating a focal point or some other method of selection among the possible equilibria. How well they succeed depends on the context, in just the way that theory suggests. If the players have sufficiently common understanding for their expectations to converge, they will succeed in settling on a good outcome; otherwise disequilibrium may persist.

Most experiments work with subjects who have no prior experience playing the particular game. The behavior of these novices does not initially conform to equilibrium theory, but it often converges to equilibrium as they gain experience. But some uncertainty about what the other player will do persists, and a good concept of equilibrium should allow players to recognize such uncertainty and respond to it. One such extension of the Nash equilibrium concept has become increasingly popular; this is the *quantal response equilibrium*, developed by professors Richard McKelvey and Thomas Palfrey of Caltech. This is too technical for a book like ours, but some readers may be inspired to read and study it.¹⁰

After a detailed review of the relevant work, two of the top researchers in the field of experimental economics, Charles Holt of the University of Virginia and Alvin Roth of Harvard University, offer a guardedly optimistic prognosis: "In the last 20 years, the notion of Nash equilibrium has become a required part of the tool kit for economists and other social and behavioral scientists. . . . There have been modifications, generalizations, and refinements, but the basic equilibrium analysis is the place to begin (and sometimes end) the analysis of strategic interactions."¹¹ We think that to be exactly the right attitude and recommend this approach to our readers. When studying or playing a game, begin with the Nash equilibrium, and then think of reasons why, and the manner in which, the outcome may differ from the Nash predictions. This dual approach is more likely to give you a good understanding or success in actual play than either a totally nihilistic—anything goes—attitude or a slavishly naïve adherence to the Nash equilibrium with additional assumptions, such as selfishness.

CASE STUDY: HALF WAY

A Nash equilibrium is a combination of two conditions:

- i. Each player is choosing a best response to what he believes the other players will do in the game.
- ii. Each player's beliefs are correct. The other players are doing just what everyone else thinks they are doing.

It is easier to describe this outcome in a two-player game. Our two players, Abe and Bea, each have beliefs about what the other will do. Based on those beliefs, Abe and Bea each choose to take an action that maximizes their payoffs. The beliefs prove right: Abe's best response to what he thinks Bea is doing is just what Bea thought Abe would do, and Bea's best response to what she thought Abe would do is indeed just what Abe expected her to do.

Let's look at these two conditions separately. The first condition is quite natural. If otherwise, then you'd have to argue that someone is not taking the best action given what he or she believes. If he or she had something better, why not do it?

Mostly, the rub comes in the second condition—that everyone is correct in what they believe. For Sherlock Holmes and Professor Moriarty this was not a problem:

"'All that I have to say has already crossed your mind,' said he.

'Then possibly my answer has crossed yours,' I replied.

'You stand fast?'

'Absolutely.'"

For the rest of us, correctly anticipating what the other side will do is often a challenge.

The following simple game will help illustrate the interplay between these two conditions and why you might or might not want to accept them.

Abe and Bea are playing a game with the following rules: Each player is to pick a number between 0 and 100, inclusive. There is a \$100 prize to the player whose number is closest to half the other person's number.

We'll be Abe and you can play Bea. Any questions?

What if there's a tie?

Okay, in that case we split the prize. Any other questions?

No.

Great, then let's play. We've picked our number. Time for you to pick yours. What is your number? To help keep yourself honest, write it down.

Case Discussion

We picked 50. No, we didn't. To see what we actually picked, you'll have to read on.

Let's start by taking a step back and use the two-step approach to finding a Nash equilibrium. In step 1, we believe that your strategy had to be an optimal response to something we might have done. Since our number has to be something between 0 and 100, we figure that you couldn't have picked any number bigger than 50. For example, the number 60 is only an optimal response if you thought we would pick 120, something we couldn't do under the rules.

What that tells us is that if your choice was truly a best response to something we might have done, you had to pick a number between 0 and 50. By the same token, if we picked a number based on something that you might have done, we would have picked something between 0 and 50.

Believe it or not, many folks stop right there. When this game is played among people who haven't read this book, the most common response is 50. Frankly, we think that is a pretty lame answer (with apologies if that's what you picked). Remember that 50 is only the best choice if you think that the other side was going to pick 100. But, in order for the other side to pick 100, they would have to have misunderstood the game. They would have had to pick a number that had (almost) no chance of winning. Any number less than 100 will beat 100.

We will assume that your strategy was a best response to something we might have done and so it is between 0 and 50. That means our best choice should be something between 0 and 25.

Note that at this juncture, we have taken a critical step. It may seem so natural that you didn't even notice. We are no longer relying on our first condition that our strategy is a best response. We have taken the next step and proposed that our strategy should be a best response to something that is a best response from you.

If you are going to do something that is a best response, we should be doing something that is a best response to a best response.

At this point, we are beginning to form some beliefs about your actions. Instead of imagining that you can do anything allowed by the rules, we are going to assume that you will actually have picked a move that is a best response. Given the quite sensible belief that you are not going to do something that doesn't make sense, it then follows that we should only pick a number between 0 and 25.

Of course, by the same token, you should be realizing that we won't be picking a number bigger than 50. If you think that way, then you won't pick a number bigger than 25.

As you might have guessed, the experimental evidence shows that after 50, 25 is the most common guess in this game. Frankly, 25 is a much better guess than 50. At least it has a chance of winning if the other player was foolish enough to pick 50.

If we take the view that you are only going to pick a number between 0 and 25, then our best response is now limited to numbers between 0 and 12.5. In fact, 12.5 is our guess. We'll win if our guess is closer to half your number than your number is to half ours. That means we win if you picked anything higher than 12.5.

Did we win?

Why did we pick 12.5? We thought you would pick a number between 0 and 25, and that's because we thought you'd think we'd pick a number between 0 and 50. We could of course go on with our reasoning and conclude that you'd figure we'd pick a number

between 0 and 25, leading you to choose something between 0 and 12.5. If you had thought that, then you'd be one step ahead of us and would have won. Our experience suggests that most people don't think more than two or three levels, at least on their first go-around.

Now that you've had some practice and better understand the game, you might want a rematch. That's fair. So write down your number again—we promise not to peek.

We are pretty confident that you expect us to pick something less than 12.5. That means you'll pick something less than 6.25. And if we think you'll pick something less than 6.25, we should pick a number less than 3.125.

Now if this were the first go-around, we might stop there. But we just explained that most folks stop after two levels of reasoning, and this time we expect that you are determined to beat us, so you'll engage in at least one more level of thinking ahead. If you expect us to pick 3.125, then you'll pick 1.5625, which leads us to think of 0.78125.

At this point, we are guessing that you can see where this is all heading. If you think we are going to pick a number between 0 and X , then you should pick something between 0 and $X/2$. And if we think you are going to pick something between 0 and $X/2$, then we should pick something between 0 and $X/4$.

The only way that we can both be right is if we both pick 0. That's what we've done. This is the Nash equilibrium. If you pick 0, we want to pick 0; if we pick 0, you want to pick 0. Thus if we both correctly anticipate what the other will do, we both do best picking 0, just what we expected the other to do.

We should have picked 0 the first time around as well. If you pick X and we pick 0, then we win. That is because 0 is closer to $X/2$ than X is to $0/2 = 0$. We knew this all along but didn't want to give it away the first time we played.

As it turned out, we didn't actually need to know anything about what you might be doing to pick 0. But that is a highly unusual case and an artifact of having only two players in the game.

Let's modify the game to add more players. Now the person

whose number is closest to half the average number wins. Under these rules, it is no longer the case that 0 always wins.* But it is still the case that the best responses converge to zero. In the first round of reasoning, all players will pick something between 0 and 50. (The average number picked can't be above 100, so half the average is bounded by 50.) In the second iteration of logic, if everyone thinks others will play a best response, then in response everyone should pick something between 0 and 25. In the third iteration of logic, they'll all pick something between 0 and 12.5.

How far people are able to go in this reasoning is a judgment call. Again, our experience suggests that most people stop at two or three levels of reasoning. The case of a Nash equilibrium requires that the players follow the logic all the way. Each player picks a best response to what he or she believes that the other players are doing. The logic of Nash equilibrium leads us to the conclusion that all players will pick 0. Everyone picking 0 is the only strategy where each of the players is choosing a best response to what they believe other players are doing and each is right about what they believe the others will be doing.

When people play this game, they rarely pick zero on the first go-around. This is convincing evidence against the predictive power of Nash equilibrium. On the other hand, when they play the game even two or three times, they get very close to the Nash result. That is convincing evidence in favor of Nash.

Our view is that both perspectives are correct. To get to a Nash equilibrium, all players have to choose best responses—which is relatively straightforward. They also all have to have correct beliefs about what the other players will be doing in the game. This is much harder. It is theoretically possible to develop a set of internally consistent beliefs without playing the game, but it is often easier to play the game. To the extent that players learn that their beliefs were wrong by playing the game and then learn how to do a

* If there are three players and the other two have picked 1 and 5, then the average of the three numbers (0, 1, and 5) is 2, half the average is 1, and the person picking 1 will win.

better job predicting what others will do, they will converge to a Nash equilibrium.

While experience is helpful, it is no guarantee of success. One problem arises when there are multiple Nash equilibria. Consider the annoying problem of what to do when a mobile phone call gets dropped. Should you wait for the other person to call you, or should you call? Waiting is a best response if you think the other person will call, and calling is a best response if you think the other person will wait. The problem here is that there are two equally attractive Nash equilibria: You call and the other person waits; or you wait and the other person calls.

Experience doesn't always help get you there. If you both wait, then you might decide to call, but if you both happen to call at the same time, then you get busy signals (or at least you did in the days before call waiting). To resolve this dilemma, we often turn to social conventions, such as having the person who first made the call do the callback. At least that way you know the person has the number.

EPILOGUE TO PART I

In the previous four chapters, we introduced several concepts and methods, using examples from business, sports, politics, and so forth as vehicles. In the chapters to follow, we will put the ideas and techniques to work. Here we recapitulate and summarize them for ready reference.

A *game* is a situation of strategic interdependence: the outcome of your choices (strategies) depends upon the choices of one or more other persons acting purposely. The decision makers involved in a game are called players, and their choices are called *moves*. The interests of the players in a game may be in strict conflict; one person's gain is always another's loss. Such games are called *zero-sum*. More typically, there are zones of commonality of interests as well as of conflict and so, there can be combinations of mutually gainful or mutually harmful strategies. Nevertheless, we usually refer to the other players in a game as one's rivals.

The moves in a game may be *sequential* or *simultaneous*. In a game of sequential moves, there is a linear chain of thinking: If I do this, my rival can do that, and in turn I can respond in the following way. Such a game is studied by drawing a *game tree*. The best choices of moves can be found by applying *Rule 1: Look forward and reason backward*.

In a game with simultaneous moves, there is a logical circle of reasoning: I think that he thinks that I think that . . . and so on. This circle must be squared; one must see through the rival's action even though one cannot see it when making one's own move. To tackle such a game, construct a *table* that shows the outcomes corresponding to all conceivable combinations of choices. Then proceed in the following steps.

Begin by seeing if either side has a *dominant strategy*—one that outperforms all of that side's other strategies, irrespective of the rival's choice. This leads to *Rule 2: If you have a dominant strategy, use it*. If you don't have a dominant strategy, but your rival does, then count on his using it, and choose your best response accordingly.

Next, if neither side has a dominant strategy, see if either has a *dominated strategy*—one that is uniformly worse for the side playing it than all the rest of its strategies. If so, apply *Rule 3: Eliminate dominated strategies from consideration*. Go on doing so successively. If during the process any dominant strategies emerge in the smaller games, they should be chosen. If this procedure ends in a unique solution, you have found the prescriptions of action for the players and the outcome of the game. Even if the procedure does not lead to a unique outcome, it can reduce the size of the game to a more manageable level. Finally, if there are neither dominant nor dominated strategies, or after the game has been simplified as far as possible using the second step, apply *Rule 4: Look for an equilibrium, a pair of strategies in which each player's action is the best response to the other's*. If there is a unique equilibrium of this kind, there are good arguments why all players should choose it. If there are many such equilibria, one needs a commonly understood rule or convention for choosing one over the others. If there is no such equilibrium, that usually means that any systematic behavior can be exploited by one's rivals, which indicates the need for *mixing one's plays*, the subject of the next chapter.

In practice, games can have some sequential moves and some simultaneous moves; in that case a combination of these techniques must be employed to think about and determine one's best choice of actions.