method (p. 111, 122);
omething similar. Smith
upon us in this man-
ar in Adams (2015). The
40–1); Seneca (2007
p. 181) claims life is
discusses "man's forecast

# PART 5 Strategic Interaction

CHAPTER 10

# Analytical Game Theory

## 10.1 Introduction

So far, we have assumed that the outcomes that you enjoy or suffer are determined jointly by your choices and the state of the world, which may be unknown to you at the time of the decision. For many real-time decisions, however, this is far from the whole story. Instead, many of the decision problems that you face in real life have an *interactive* or *strategic* nature. This means that whatever happens depends not just on what you do but also on what other people do. If you play chess, whether you win or lose is determined not just by what you do but also by what your opponent does. If you invest in stock, whether or not you make money depends not only on your choice of stock but also on whether the stock goes up or down. And that is a function of supply and demand, which is a function of whether other people buy or sell the stock.

The presence of strategic interaction adds a whole new layer of complexity to the analysis. If you are a defense attorney, the outcome of your case depends not only on your actions but on the actions of the prosecutor. Since you are well aware of this, you do your best to anticipate her decisions. Thus, your decisions will depend on what you think her decisions will be. Her decisions will reflect, in part, what she thinks you will do. So your decisions will depend on what you think she thinks you will do. But what she thinks you will do depends in part on what she thinks you think she will do ... and so on. It is obvious that the correct analysis of strategic interaction is not obvious at all.

The analysis of strategic interaction is the province of **game theory**. In this chapter, I offer a brief overview of the standard theory, sometimes referred to as **analytical game theory**.

## 10.2 Nash equilibrium in pure strategies

The following story is a true internet legend.

**Example 10.1 The makeup exam** One year there were two students taking Chemistry. They both did so well on quizzes, midterms, and labs that they decided to leave town and go partying the weekend before the exam. They mightily enjoyed themselves. However, much like a scene in *The Hangover: Part III*, they overslept and did not make it back to campus in time for the exam.

So they called their professor to say that they had got a flat tire on
to the exam, did not have a spare, and had to wait for a long time. Th
sor thought about it for a moment, and then said that he was glad to gi
a makeup exam the next day. The two friends studied all night.

At the assigned time, the professor placed them in separate rooms,
them the exams, and asked them to begin. The two friends looked at
problem, which was a simple one about molarity and solutions and w
5 points. "Easy!" they thought to themselves. Then, they turned the p
saw the second question: "(95 points) Which tire?"

This example illustrates the interactive or strategic nature of many
problems. Here, the final grade of either friend will depend not jus
answer to the question, but on the other friend's answer too. The two
As whenever they give the same answer to the question and Fs whene
do not.

More formally speaking, you are playing a **game** whenever you fac
sion problem in which the final outcome depends not just on your actio
whatever state of the world obtains, but also on the actions of at least c
agent. According to this definition, the two friends are in fact playing
against each other. And this is true whether or not they think of it as
Notice that you can play a game in this sense without competing *aga*
other. Here, the name of the game is cooperation – and coordination. Th
involved in games are called **players**. A **strategy** is a complete plan of ac
describes what a player will do under all possible circumstances. In th
the makeup exam, each friend has four strategies to choose from: he c
"Front Left (FL)," "Front Right (FR)," "Rear Left (RL)," or "Rear Right (

Given a number of players, a set of strategies available to each pla
a set of payoffs (rewards or punishments) corresponding to each
combination of strategies, a game can be represented using a **payof**
A payoff matrix is a table representing the payoffs of the players for e
sible combination of strategies. The payoff matrix of the game playe
two friends can be represented as in Table 10.1. A **strategy profile** is
of strategies, one for each player. ⟨FL, RR⟩ is a strategy profile; so is ⟨
Thus, the payoff matrix shows the payoffs resulting from each strat
file. Of course, the payoff matrix looks much like the tables represent
strategic decision problems, except for the fact that each column rep
choice by the other player rather than a state of the world.

**Table 10.1** The makeup exam

| | **FL** | **FR** | **RL** | **RR** |
|---|---|---|---|---|
| FL | A | F | F | F |
| FR | F | A | F | F |
| RL | F | F | A | F |
| RR | F | F | F | A |

Analytical game theory is built around the concept of an equilibrium. The most prominent equilibrium concept is that of **Nash equilibrium**.

---

**Definition 10.2 Nash equilibrium** *A Nash equilibrium is a strategy profile such that each strategy in the profile is a best response to the other strategies in the profile.*

---

In the makeup-exam game from Example 10.1, ⟨FL, FL⟩ is a Nash equilibrium: given that Player I plays FL, FL is a best response for Player II, and given that Player II plays FL, FL is a best response for Player I. In equilibrium, given the other players' strategies, no one player can improve his or her payoff by unilaterally changing to another strategy. By contrast, ⟨FL, RR⟩ is not a Nash equilibrium: given that Player I plays FL, Player II can do better than playing RR, and given that Player II plays RR, Player I can do better than playing FL. In this section, we will limit our analysis to Nash equilibria in pure strategies: Nash equilibria in which each player simply plays one of the individual strategies available to him or her (compare Section 10.3). In all, there are four Nash equilibria in pure strategies, one for each tire.

**Example 10.3 Coffee shops** You and your study partner are planning to meet at noon at one of two coffee shops, Lucy's Coffee and Crestwood Coffee. Unfortunately, you failed to specify which one, and you have no way of getting in touch with each other before noon. If you manage to meet, you get a utility of 1; otherwise, you get a utility of 0. Draw the payoff matrix and find the Nash equilibria in pure strategies.

The payoff matrix is Table 10.2. The convention is for the first number in each cell to represent the payoff of Player I, whose strategies are listed in the left-most column; the second number in each cell represents the payoff of Player II, whose strategies are listed in the top row. The Nash equilibria in pure strategies are ⟨Lucy's, Lucy's⟩ and ⟨Crestwood, Crestwood⟩.

**Table 10.2** A pure coordination game

| | **Lucy's** | **Crestwood** |
|---|---|---|
| Lucy's | 1,1 | 0,0 |
| Crestwood | 0,0 | 1,1 |

The coffee-shop game is an example of a **pure coordination game**: a game in which the players' interests are perfectly aligned. The makeup-exam game, obviously, is also a pure coordination game. In some coordination games, however, interests fail to align perfectly. The point is typically made by means of the politically-incorrectly named **battle of the sexes**.

**Example 10.4 Battle of the sexes** A husband and wife must decide whether to have dinner at the steak house or at the crab shack. All things equal, both would rather dine together than alone, but the man (Player I) prefers the steak house and the woman (Player II) prefers the crab shack. The man gets 2 units of utility if both dine at the steak house, 1 if both dine at the crab shack, and 0 if they dine apart; the woman gets 2 units of utility if both dine at the crab shack, 1 if both dine at the steak house, and 0 if they dine apart. Draw the payoff matrix and find the Nash equilibria in pure strategies.

The payoff matrix is Table 10.3. There are two Nash equilibria in pure strategies. ⟨Steak House, Steak House⟩ is one. Because this is Player I's best outcome, he cannot improve his payoff by changing strategies. Although Player II would prefer it if *both* switched to Crab Shack, she cannot improve her payoff by *unilaterally* deviating: if she plays Crab Shack when Player I plays Steak House, she will end up with a payoff of 0 rather than 1. Of course ⟨Crab Shack, Crab Shack⟩ is the other Nash equilibrium in pure strategies.

**Table 10.3** An impure coordination game

| | **Steak House** | **Crab Shack** |
|---|---|---|
| Steak House | 2,1 | 0,0 |
| Crab Shack | 0,0 | 1,2 |

Because Player I prefers the one equilibrium and Player II prefers the other, the battle of the sexes, sometimes euphemistically called "Bach or Stravinsky," is an example of an **impure coordination game**. Here are some exercises.

**Exercise 10.5 Nash equilibrium in pure strategies** Find all Nash equilibria in the games in Table 10.4, where Player I chooses between Up (U), (Middle (M)), and Down (D) and Player II chooses between Left (L), (Middle (M)), and Right (R).

**Table 10.4** Nash equilibrium exercises

| | **L** | **R** |
|---|---|---|
| U | 2,2 | 0,0 |
| D | 0,0 | 1,1 |

(a)

| | **L** | **R** |
|---|---|---|
| U | 5,1 | 2,0 |
| D | 5,1 | 1,2 |

(b)

| | **L** | **M** | **R** |
|---|---|---|---|
| U | 6,2 | 5,1 | 4,3 |
| M | 3,6 | 8,4 | 2,1 |
| D | 2,8 | 9,6 | 3,0 |

(c)

Notice that in Exercise 10.5(a), there are two Nash equilibria in pure strategies, though one is clearly inferior to the other from the point of view of both players. In Exercise 10.5(b), ⟨U, L⟩ and ⟨D, L⟩ are not both Nash equilibria although they are "just as good" in the sense that they lead to the same payoffs. And in Exercise 10.5(c), there are outcomes that are better for both players than the Nash equilibrium.

As these games illustrate, there is no straightforward connection between Nash equilibria and "best" outcomes for the players. As a result, it would be a mistake to try to identify the former by searching for the latter. An even more striking example of the general phenomenon is the **prisoners' dilemma** (Figure 10.1).

**Example 10.6 Prisoners' dilemma** Two criminals are arrested on suspicion of two separate crimes. The prosecutor has sufficient evidence to convict the two

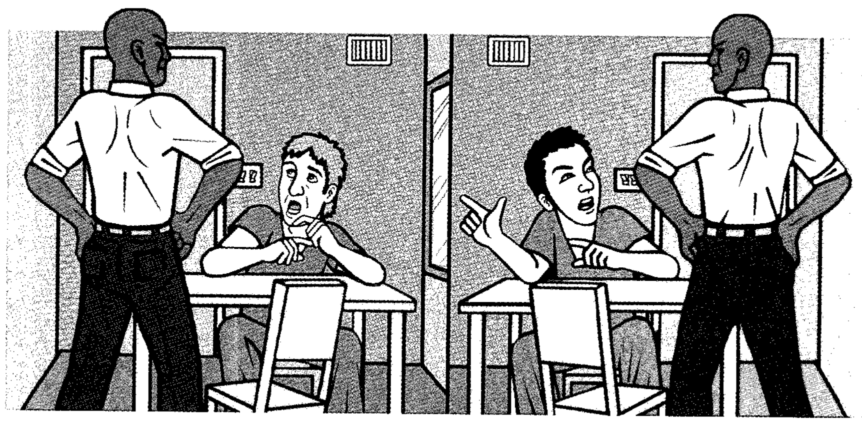

**Figure 10.1** The suspects. Illustration by Cody Taylor

on the minor charge, but not on the major one. If the two criminals *cooperate* (C) with each other and stay mum, they will be convicted on the minor charge and serve two years in jail. After separating the prisoners, the prosecutor offers each of them a reduced sentence if they *defect* (D), that is, testify against each other. If one prisoner defects but the other one cooperates, the defector goes free whereas the cooperator serves 20 years in jail. If both defect, both get convicted on the major charge but (as a reward for testifying) only serve ten years. Assume that each prisoner cares about nothing but the number of years he himself spends in jail. What is the payoff matrix? What is the Nash equilibrium?

The payoff matrix in terms of jail sentences is Table 10.5(a); in terms of utilities, the payoff matrix can be represented as Table 10.5(b). Let us consider Player I first. If Player II cooperates, Player I has the choice between cooperating and defecting; by defecting, he can go free instead of serving two years in jail. If Player II defects, Player I still has the choice between cooperating and defecting; by defecting, he can serve 10 instead of 20 years in jail. In brief, Player I is better off defecting no matter what Player II does. But the same thing is true for Player II. Thus, there is only one Nash equilibrium. Both defect and serve 10 years in jail.

**Table 10.5** The prisoners' dilemma

| | C | D |
|---|---|---|
| C | 2 years, 2 years | 20 years, 0 years |
| D | 0 years, 20 years | 10 years, 10 years |

(a)

| | C | D |
|---|---|---|
| C | 3,3 | 0,5 |
| D | 5,0 | 1,1 |

(b)

One way to identify the unique Nash equilibrium in the prisoners' dilemma is to eliminate all strictly dominated strategies. A strategy $X$ is said to strictly dominate another strategy $Y$ if choosing $X$ is better than choosing $Y$ no matter what the other player does. Because no rational agent will play a strictly ( inated strategy, such strategies can be eliminated from consideration v searching for Nash equilibria. In the prisoners' dilemma, defection st dominates cooperation, so cooperation can be eliminated. No rational p will cooperate, and both will defect.

Notice that the result holds even though both prisoners agree th would have been better if they had both cooperated. An outcome $X$ is sa **Pareto dominate** another $Y$ if all players weakly prefer $X$ to $Y$ and at leas player strictly prefers $X$ to $Y$. An outcome is **Pareto optimal** if it is not P dominated by any other outcome. In the prisoners' dilemma, the cooper outcome ⟨C, C⟩ Pareto dominates the Nash equilibrium ⟨D, D⟩ . In fact, players strictly prefer the former to the latter. Still, rational players will jo choose an outcome which is not Pareto optimal. For this reason, the pri ers' dilemma is sometimes presented – for example, in the film *A Bea Mind*, about game theory inventor and Nobel laureate John Nash – as refu Adam Smith's insight that the rational pursuit of individual self-interest l to socially desirable outcomes.

Many real-world interactions have features that are reminiscent of prisoners' dilemma. Arms races are classic examples. Consider the nu buildup in India and Pakistan. Whether or not India has nuclear a Pakistan wants them. If India has them, Pakistan needs them to preserve balance of power; if India does not have them, Pakistan wants them to ge upper hand. For the same reason, India wants nuclear arms whether or Pakistan has them. Thus, both countries acquire nuclear arms, neither cou has the upper hand, and both countries are worse off than if neither had th Overfishing, deforestation, pollution, and many other phenomena are o classic examples. The idea is that no matter what other players do, each pl has an incentive to fish, cut down forests, and pollute, but, if everyone d everyone is worse off than if nobody had acted.

A number of different solutions might occur to you. What if the two p oners, before committing the crime, got together and promised to coope in the event that they are caught? Surely, you might think, a gentlem agreement and a handshake would do the trick. The solution fails, h ever, because whatever verbal agreement the prisoners might have ent into before getting caught will not be binding. At the end of the day, e has to make a choice in isolation, defection strictly dominates cooperat and a rational agent has no choice but to defect. "Talk is cheap," the say goes, which is why game theorists refer to non-binding verbal agreement **cheap talk**.

What if the game could be repeated? Repetition, you might think, sho afford a prisoner the opportunity to punish defection by defecting. But s pose the two prisoners play ten prisoners' dilemma games against each ot To find the equilibrium in the repeated game, we start at the end and u procedure called **backward induction**. In the last round, no rational prisc will cooperate, because his opponent has no way to retaliate against de tion; so in round ten, both prisoners will defect. In the next to last roun rational prisoner already knows that his opponent will defect in round which means that it does not matter whether he cooperates or defects; s round nine, both prisoners will defect. The same thing is true for round eight, round seven, and so on. In this way, the prospect of rational cooperation in the repeated prisoners' dilemma unravels from the end. Repetition does not necessarily solve the problem.

Cooperation can be sustained if there is no last round, that is, if the game is repeated indefinitely. In the indefinitely repeated prisoners' dilemma, there is a Nash equilibrium in which both prisoners cooperate throughout but are prepared to punish defection by defecting. The cooperative solution presupposes that the players do not discount the future too much: if they do, no amount of repetition will save the prisoners. And there is no guarantee that rational individuals will play that particular equilibrium. In fact, there is an *infinite* number of equilibria in the indefinitely repeated prisoners' dilemma, and in one of those equilibria prisoners always defect. In brief, indefinite repetition upholds the prospect of rational cooperation in the prisoners' dilemma, but cooperation is far from guaranteed.

There is only one sure-fire way for rational agents to avoid defection, and it is to make sure they are not playing a prisoners' dilemma at all. Suppose that the two criminals, before committing their crimes, go to the local contract killer and instruct him to kill anyone who defects in the event that he is caught. If death at the hands of a contract killer is associated with a utility of $-\infty$, the payoff matrix of the two prisoners will now look like Table 10.6. Here, cooperation strictly dominates defection for both players and ⟨C, C⟩ is the unique Nash equilibrium. You might think that it would never be in a person's interest to ask to be killed by a contract killer, no matter what the conditions; yet, by proceeding in this way, the prisoners can guarantee themselves a much better payoff than if they had not. But notice that cooperation is the uniquely rational strategy only because the two prisoners are no longer playing a prisoners' dilemma.

**Table 10.6** The modified prisoners' dilemma

| | C | D |
|---|---|---|
| C | 3,3 | $0,-\infty$ |
| D | $-\infty,0$ | $-\infty,-\infty$ |

**Example 10.7 The Leviathan** The seventeenth-century political philosopher Thomas Hobbes offered a justification of political authority by imagining what life would be like without it. In one of the most famous lines in the history of Western philosophy, Hobbes described this "state of nature" as follows:

> [During] the time men live without a common power to keep them all in awe, they are in that condition which is called war, and such a war as is of every man against every man … In such condition there is … continual fear and danger of violent death, and the life of man, solitary, poor, nasty, brutish, and short.

The solution, according to Hobbes, is a covenant according to which people give up their right to kill and maim other people in exchange for the right


**Figure 10.2** The Leviathan. Detail of the frontispiece from the 1651 edition

not to be killed and maimed, and which at the same time establishes an ov whelming power – a *Leviathan* – to ensure that people adhere to the term the covenant (see Figure 10.2).

Game theory offers a new way to interpret the nature of this "war of against all." These days, many people think of Hobbes's story as a vi description of a scenario in which people are forced to play prisoners' dile mas against each other, and in which the pursuit of rational self-interest the fore leads to the worst possible outcome for all involved. The Leviathan Hobbes's story serves the same function as the contract killer in the scena above: by holding people to their promises, he ensures that rational s interest coincides with social desirability.

## 10.3 Nash equilibrium in mixed strategies

Some games have no Nash equilibria in pure strategies. But that does mean that they do not have Nash equilibria.

**Example 10.8 Coffee shops, cont.** Suppose that you still have to go to of the two coffee shops in Example 10.3 and that your ex has to also. You not want to run into your ex, but your ex wants to run into you. What kind game would you be playing against each other?

If a player gets a utility of 1 whenever his or her goal is attained and 0 otherwise, the payoff matrix is Table 10.7.

**Table 10.7** A pure coordination game

| | Lucy's | Crestwood |
|---|---|---|
| Lucy's | 1,0 | 0,1 |
| Crestwood | 0,1 | 1,0 |

This game has no Nash equilibria in pure strategies. If you go to Lucy's, your ex will want to go there too, but then you want to go to Crestwood, in which case your ex wants to do so too. This coffee shop game, by the way, has the same payoff structure as a game called **matching pennies**. When two people play matching pennies, each flips a penny. If both coins come up heads, or if both coins come up tails, Player I wins; otherwise, Player II wins. This also happens to be an example of a **zero-sum game**, a game in which whenever one player wins, another player loses.

The game does, however, have a **Nash equilibrium in mixed strategies**. Suppose that you figure out where to go by flipping a coin, and that your ex does the same. Given that you have a 50 percent chance of ending up at Lucy's and a 50 percent chance of ending up at Crestwood, your ex is indifferent between Lucy's and Crestwood and can do no better than flipping a coin. And given that your ex has a 50 percent chance of ending up at Lucy's and a 50 percent chance of ending up at Crestwood, you are indifferent between Lucy's and Crestwood and can do no better than flipping a coin. Hence, the two of you are in a Nash equilibrium, though you are playing mixed rather than pure strategies.

In a game like this, the mixed-strategy equilibrium is easy to find. In other games it can be more tricky. Consider the battle of the sexes (Example 10.4). In order to find a mixed-strategy equilibrium in a game like this, there is one critical insight: in order for players to rationally play a mixed strategy, they must be indifferent between the pure strategies they are mixing. Why? If a player strictly preferred one over the other, the only rational thing to do would be to play the preferred strategy with probability one. Thus, you can find the mixed-strategy equilibrium in a game by setting up equations and solving for the probabilities with which the players play different strategies.

**Example 10.9 Battle of the sexes, cont.** In order to find the mixed-strategy equilibrium in the battle of the sexes (Table 10.8), let us assume that Player I plays U with probability $p$ and D with probability $(1 - p)$ and that Player II plays L with probability $q$ and R with probability $(1 - q)$.

Consider Player I first. In order to play a mixed strategy, he must be indifferent between U and D, meaning that $u(\text{U}) = u(\text{D})$. The utility of playing U will depend on what Player II does, that is, on what $q$ is. When playing U, Player I has a probability of $q$ of getting 2 utiles and a probability of $(1 - q)$ of getting 0. Consequently, $u(\text{U}) = q * 2 + (1 - q) * 0 = 2q$. When playing D, Player I has a probability of $q$ of getting 0 utiles and a probability of $(1 - q)$ of getting 1. Thus, $u(D) = q * 0 + (1 - q) * 1 = 1 - q$. So $u(U) = u(D)$ entails th $2q = 1 - q$, meaning that $q = 1/3$.

Next, consider Player II. In order to play a mixed strategy, she needs to b indifferent between L and R, meaning that $u(L) = u(R)$. Now $u(L) = p * 1$ $(1 - p) * 0 = p$ and $u(R) = p * 0 + (1 - p) * 2 = 2 - 2p$. So $u(L) = u(R)$ entai that $p = 2 - 2p$, meaning that $p = 2/3$.

Hence, there is a Nash equilibrium in mixed strategies in which Player plays U with probability 2/3 and Player II plays L with probability 1/3. In th mixed-strategy equilibrium, Player I gets payoff $u(U) = u(D) = 2q = 2/3$ an Player II gets payoff $u(L) = u(R) = p = 2/3$.

**Table 10.8** An impure coordination game

| | L | R |
|---|---|---|
| U | 2,1 | 0,0 |
| D | 0,0 | 1,2 |

As this example shows, games with pure-strategy equilibria may also have mixed equilibria.

**Exercise 10.10 Mixed-strategy equilibrium** Find the mixed-strategy Nash equilibria in Tables 10.4(a) and (b).

In the mixed-strategy equilibrium in (a), notice that Player I is more likely to play D than U and that Player II is more likely to play R than L. This might seem strange, since you would perhaps expect the players to be more likely to play the strategy associated with the more desirable equilibrium ⟨U, L⟩. But assume for a proof by contradiction that the two players are in a mixed-strategy equilibrium in which Player I plays U and Player II plays L with some high probability. If so, Player I would strictly prefer U to D and Player II would strictly prefer L to R. Thus, the two players would not be in equilibrium at all, contrary to the initial assumption. For the two players to want to mix, they must be indifferent between the two pure strategies, and this can happen only when Player I is more likely to play D than U and when Player II is more likely to play R than L.

Notice also that the probability $p$ with which Player I plays U is a function not of Player I's own payoffs, but of Player II's payoffs. This might seem equally counterintuitive. Yet, it follows from the fact that $p$ must be selected in such a manner as to make Player II indifferent between her pure strategies. Similarly, the probability $q$ with which Player II plays L is determined not by her payoffs, but by her opponent Player I's payoffs. This is a fascinating feature of Nash equilibria in mixed strategies.

**Exercise 10.11 Pure vs. mixed equilibria** Find all Nash equilibria (in pure and mixed strategies) in the games depicted in Table 10.9.

Although a mixed-strategy equilibrium may at first blush seem like an artificial construct of mainly academic interest, mixed strategies are important and common in a wide variety of strategic interactions. Even if you are a tennis player with a killer cross-court shot, it would be unwise to hit the cross-court shot every time, or your opponent will learn to expect it. Every so often, you must hit the ball down the line. In games like these, in order to keep your opponent guessing, you must mix it up a bit. This analysis shows that it is not a mistake, but *necessary*, every so often to hit your weaker shot.

**Table 10.9** Mixed Nash equilibrium exercises

| | L | R |
|---|---|---|
| U | 5,2 | 1,1 |
| D | 1,1 | 2,5 |

(a)

| | L | R |
|---|---|---|
| U | 4,1 | 2,0 |
| D | 5,1 | 1,2 |

(b)

| | L | R |
|---|---|---|
| U | 1,1 | 0,0 |
| D | 0,0 | 0,0 |

(c)

**Example 10.12 Spousonomics** According to the authors of the book *Spousonomics*: "Economics is the surest route to marital bliss" because "it offers dispassionate, logical solutions to what can often seem like thorny, illogical, and highly emotional domestic disputes." Suppose that you are stuck in an equilibrium where you do the dishes, make the bed, and empty the cat litter while your spouse sits back and relaxes. Spousonomics, apparently, teaches that you can turn your spouse into an acceptable (if not ideal) partner by playing a mixed strategy, by sometimes doing the laundry, sometimes not, sometimes making the bed, sometimes not, and so on.

**Exercise 10.13 Rock-paper-scissors**

(a) Draw the payoff matrix for the game rock-paper-scissors. Suppose that a win gives you 1 utile, a tie 0, and a loss −1.

(b) What is the unique Nash equilibrium in this game?

We already know that not all games have Nash equilibria in pure strategies. But now that we have access to the concept of a Nash equilibrium in mixed strategies it is possible to prove a famous theorem originally due to John Nash. Simplified and expressed in words:

***Theorem 10.14 Nash's theorem*** *Every finite game – that is, every game in which all players have a finite number of pure strategies – has a Nash equilibrium.*

*Proof.* Omitted. □

Given this theorem, the search for Nash equilibria is not futile. As long as the number of pure strategies available to each player is finite – and whether this condition is satisfied is fairly easy to determine – we know that the game has at least one Nash equilibrium in pure or mixed strategies. This is neat.

**Example 10.15 Chess** Chess is a finite game. We know this because every player has a finite number of moves to choose from at any point in the game and because every game ends after a finite number of moves. Because it finite game, Nash's theorem establishes that it has an equilibrium.

This suggests that chess should be uninteresting, at least when played experienced players. Assuming Player I plays the equilibrium strategy, Pla II can do no better than playing the equilibrium strategy, and vice versa. Th we should expect experienced players to implement the equilibrium stra gies every time, and the outcome to be familiar and predictable.

Yet Nash's theorem only establishes the existence of an equilibrium; it d not reveal what the equilibrium strategies are. As of yet, no computer is po erful enough to figure out what they are. And even if we knew what the stra gies were, they might be too complex for human beings to implement. Th chess is likely to remain interesting for a good long time.

Before we move on, two more exercises.

**Exercise 10.16 Chicken** The game of **chicken** was popularized in the 19 film *Rebel Without a Cause*, starring James Dean. The game is played by two pe ple who drive cars straight at each other at high speed; the person who swer first is called "chicken" and becomes an object of contempt. In this game, t British philosopher Bertrand Russell saw an analogy with Cold War policy:

> Since the nuclear stalemate became apparent, the Governments of E and West have adopted the policy which [US Secretary of State] Mr Dul calls "brinkmanship". This is a policy adapted from a sport which am told, is practised by some youthful degenerates. This sport is call "Chicken!" ... As played by irresponsible boys, this game is consider decadent and immoral, though only the lives of the players are risked. B when the game is played by eminent statesmen, who risk not only th own lives but those of many hundreds of millions of human beings, it thought on both sides that the statesmen on one side are displaying a hig degree of wisdom and courage, and only the statesmen on the other si are reprehensible. This, of course, is absurd.

Imagine that each player has the choice between swerving (S) and not swer ing (¬S), and that the payoff structure is that of Table 10.10. Find all Na equilibria in this game.

**Table 10.10** Chicken

| | S | ¬S |
|---|---|---|
| S | 3,3 | 2,5 |
| ¬S | 5,2 | 1,1 |

In a branch of game theory called **evolutionary game theory**, this game figure prominently under the heading of **hawk & dove**. Hawks are willing to fight t the death whereas doves easily give up. The best possible outcome for you when you are a hawk and your opponent is a dove, the second best outcom is when both of you are doves, the third best is when you are a dove and your opponent is a hawk, and the worst possible outcome is when both of you are hawks. If doves "swerve" and hawks do not, the payoff structure of hawk & dove is the same as that of chicken. In evolutionary game theory, the mixed-strategy equilibrium is interpreted as describing a population in which hawks and doves coexist in given proportions – just as they do in the real world.

**Exercise 10.17 The stag hunt** This game is due to Jean-Jacques Rousseau, the eighteenth-century French philosopher. Rousseau describes a scenario in which two individuals go hunting. The two can hunt hare or deer but not both. Anyone can catch a hare by himself, but the only way to bag a deer is for both hunters to pursue the deer. A deer is much more valuable than a hare. The **stag hunt**, which is thought to provide an important parable for social cooperation, is usually represented as in Table 10.11. What are the Nash equilibria (in pure and mixed strategies) of this game?

**Table 10.11** The stag hunt

| | D | H |
|---|---|---|
| D | 3,3 | 0,1 |
| H | 1,0 | 1,1 |

Notice how superficially subtle differences in the payoff structure between the prisoners' dilemma (Table 10.5(b)), chicken (Table 10.10), and the stag hunt (Table 10.11) lead to radically different results.

## 10.4 Equilibrium refinements

The concept of a Nash equilibrium is associated with a number of controversial results. In this section, we consider two alternative equilibrium concepts, designed to deal with supposedly problematic cases.

**Example 10.18 Trembling-hand perfection** Let us return to Table 10.9(c). As you know, ⟨U, L⟩ is a Nash equilibrium. ⟨D, L⟩ is *not* an equilibrium, since Player I can improve his payoff by playing U instead of D, and neither is ⟨U, R⟩. But consider ⟨D, R⟩. If Player II plays R, Player I can do no better than playing D; if Player I plays D, Player II can do no better than playing R. Thus, ⟨D, R⟩ is a Nash equilibrium. There are no mixed equilibria. No matter what Player II does, Player I will never be indifferent between U and D, and no matter what Player I does, Player II will never be indifferent between L and R.

There is nothing wrong with the analysis here, but there is something odd about the second equilibrium ⟨D, R⟩. A strategy $X$ is said to weakly dominate another strategy $Y$ if choosing $X$ is no worse than choosing $Y$ no matter what the other player does, and choosing $X$ is better than choosing $Y$ for at least one strategy available to the other player. In Example 10.18, U weakly dominates

D and L weakly dominates R. Thus, there seems to be no reason why ratic individuals would ever play the second equilibrium ⟨D, R⟩. And the prob is not that (1,1) Pareto dominates (0,0) (see Section 10.2).

The concept of a **trembling-hand-perfect equilibrium** was designec handle this kind of situation.

---

***Definition 10.19 Trembling-hand-perfect equilibrium*** *A trembling-hanc perfect equilibrium is a Nash equilibrium that remains a best response for eac player even when others have some minuscule probability of* ***trembling****, that i accidentally playing an out-of-equilibrium strategy.*

---

In Table 10.9(c), ⟨U, L⟩ is a trembling-hand-perfect equilibrium: even if th is a minuscule probability $\epsilon > 0$ that Player II plays R, she still plays L w probability $(1 - \epsilon)$ and U remains a best response for Player I. (And simila for the other player.) By contrast, ⟨D, R⟩ is not trembling-hand perfect. If th is a minuscule probability $\epsilon > 0$ that Player II plays L, no matter how small is a strictly preferred strategy for Player I.

**Exercise 10.20 Battle of the sexes, cont.** Are the two pure-strategy equi ria in the battle of the sexes (Table 10.8) trembling-hand perfect?

Trembling-hand-perfect equilibrium is a **refinement** of Nash equilibrium. T means that every trembling-hand-perfect equilibrium is a Nash equilibriu but not every Nash equilibrium is trembling-hand perfect.

**Exercise 10.21 Trembling-hand perfection** Find (**a**) all Nash equilibria pure strategies in Table 10.12 and (**b**) identify which of them are tremblir hand perfect.

**Table 10.12** Trembling-hand perfection, cont.

| | L | M | R |
|---|---|---|---|
| U | 1,4 | 0,0 | 0,0 |
| M | 0,0 | 4,1 | 0,0 |
| D | 0,0 | 0,0 | 0,0 |

Substituting the concept of trembling-hand-perfect equilibrium for the co cept of Nash equilibrium would eliminate some problematic implications the Nash-equilibrium concept. The concept of trembling-hand equilibrium however, insufficient to deal with all problematic cases.

**Example 10.22 Credible versus non-credible threats** Consider a game wi two stages. In the first stage, Player I plays U or D. If Player I plays D, bo players get a payoff of 2. If Player I plays U, it is Player II's turn. In the secon stage, Player II plays L or R; if Player II plays L, Player I gets 5 and Player gets 1. If Player II plays R, both get 0. What are the Nash equilibria of this gam

The game can be represented as in Table 10.13. There are two Nash equilibria: ⟨U, L⟩ and ⟨D, R⟩.

**Table 10.13** Subgame perfection

| | L | R |
|---|---|---|
| U | 5,1 | 0,0 |
| D | 2,2 | 2,2 |

Yet there is something odd about the second of the two equilibria. The only thing preventing Player I from playing U is the threat of Player II playing R. But suppose that Player I did play U. Then, Player II would have the choice of playing L (for a payoff of 1) and R (for a payoff of 0). In the second stage, it is not in Player II's interest to play R. So while it is perfectly possible for Player II to threaten to play R if Player I plays U, she would have no interest in carrying out the threat. Knowing this, it appears that Player I should just go ahead and play U. Game theorists say that the problem is that Player II's threat is not **credible**. Many people think that it is problematic that a Nash equilibrium might involve a non-credible threat. And the problem is not that the Nash equilibrium is not trembling-hand perfect.

Games with multiple stages are called **sequential**. To analyze such games, it is often useful to use a tree-like representation called the **extensive form**. The game from Example 10.22 can, for example, be represented as in Figure 10.3. This representation affords another way to spell out the problem. Consider that part of the game which starts at the node where Player II moves (see the shaded area in the figure). We refer to it as a **subgame** of the original game. In the subgame, Player II has two strategies (L and R) and there is only one Nash equilibrium: to play L (for a payoff of 1) rather than R (for a payoff of 0). Yet the Nash equilibrium in the game requires Player II to play R in the subgame. One way to spell out the problem, then, is to say that

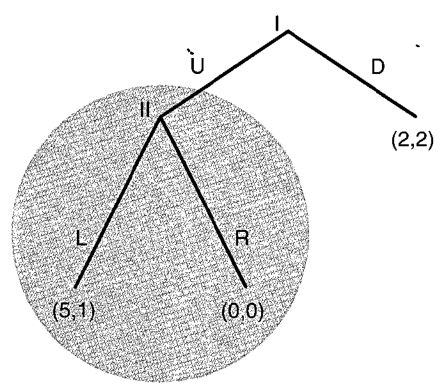


**Figure 10.3** Subgame perfection

the Nash equilibrium of the *game* requires Player II to play a strategy that is not a Nash equilibrium in the *subgame*.

Consistent with this analysis, game theorists have proposed another equilibrium concept: **subgame-perfect equilibrium**. As suggested in the previous paragraph, a subgame of a game is any part of that game which in itself constitutes a game. A game is always its own subgame, but in this case there is a proper subgame starting at the node where Player II moves.

***Definition 10.23 Subgame-perfect equilibrium*** *A subgame-perfect equilibrium is a strategy profile that constitutes a Nash equilibrium in each subgame.*

Like trembling-hand-perfect equilibrium, subgame-perfect equilibrium is a refinement of Nash equilibrium: all subgame-perfect equilibria are Nash equilibria, but not all Nash equilibria are subgame perfect.

One way to find subgame-perfect equilibria is to start at the end and use backward induction. Backward induction would tell you to start with the last subgame, that is, at the node where Player II moves (the shaded area of the figure). Since L would lead to a payoff of 1 and R would lead to a payoff of 0, L is the unique Nash equilibrium strategy. So, in subgame-perfect equilibrium, Player II will play L. Given that Player II will play L, what will Player I do at the first node? Player I has the choice between playing U for a payoff of 3, and playing D for a payoff of 2. Thus, Player I will play U. In brief, there is only one subgame-perfect equilibrium in this game, and it is ⟨U, L⟩.

**Example 10.24 MAD** Mutually assured destruction (MAD) is a military doctrine according to which two superpowers (such as the US and the USSR) can maintain peace by threatening to annihilate the human race in the event of an enemy attack. Suppose that the US moves first in a game like that in Figure 10.3. The US can launch an attack (U) or not launch an attack (D). If it launches an attack, the USSR can refrain from retaliating (L) or annihilate the human race (R). Given the payoff structure of the game in the figure, ⟨D, R⟩ is a Nash equilibrium. The doctrine is flawed, however, in that the threat is not credible: the MAD Nash equilibrium presupposes that USSR forces are willing to annihilate the human race in the event of a US attack, which would obviously not be in their interest. Thus, the MAD Nash equilibrium is not subgame perfect.

In Stanley Kubrick's 1963 film *Dr Strangelove*, the USSR tries to circumvent the problem by building a **doomsday machine**: a machine that in the event of an enemy attack (or when tampered with) automatically launches an attack powerful enough to annihilate the human race. Such a machine would solve the strategic problem, because it guarantees retaliation to enemy attack and therefore makes the threat credible. As the film illustrates, however, such machines are associated with other problems. To begin with, you must not forget to tell your enemy that you have one.

**Exercise 10.25 Subgame perfection** Use backward induction to find the unique subgame-perfect equilibrium in the game in Figure 10.4. Recall that a

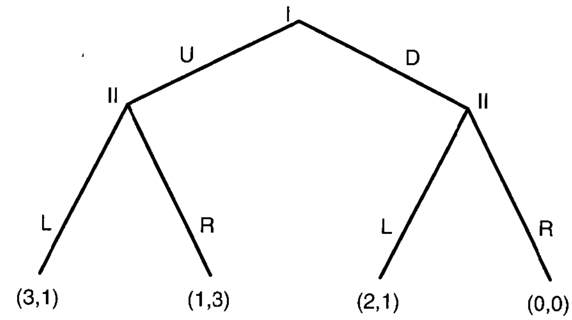


**Figure 10.4** Subgame perfection exercise

strategy is a complete plan of action, which means that a strategy for Player II will have the form "L at the first node and L at the second (LL)," "R at the first node and L at the second (RL)," and the like. In this game, then, whereas Player I only has two strategies to choose from, Player II has four.

Finally, one more exercise:

**Exercise 10.26 The centipede game** The centipede game has four stages (see Figure 10.5). At each stage, a player can Take, thereby ending the game, or Pass, thereby increasing the total payoff and allowing the other player to move.

(a) Use backward induction to find the unique subgame-perfect equilibrium.
(b) Would the outcome of the game differ if it had 1000 stages instead of four?


**Figure 10.5** The centipede game

## 10.5 Discussion

Like the theories we came across earlier in this book, analytical game theory admits of descriptive and normative interpretations. According to the descriptive interpretation, game theory captures the manner in which people behave when they engage in strategic interactions. In this view, game theory predicts that people will jointly choose an equilibrium strategy profile. Specific predictions, of course, will depend not only on the game played, but on the equilibrium concept that is employed. According to the normative interpr tion, game theory describes how rational agents should behave when t engage in strategic interaction. In this view, game theory says that play should jointly choose an equilibrium strategy profile. Again, the spec advice offered by the theory will depend on the game played and the equi rium concept employed.

One thing to notice is that games do not necessarily have a unique eq librium. And analytical game theory in itself does not contain the resour required to identify *which* equilibrium people will or should play. While theory can be interpreted as predicting that the outcome of strategic int action will or should be a Nash equilibrium, this is only to say that so Nash equilibrium will or should obtain. In this sense, then, the theory is in terminate. And because some games (like the indefinitely repeated pris ers' dilemma) have an infinite number of equilibria, the theory is radica indeterminate.

If we want determinate predictions, we must augment the theory w additional resources. The most famous such effort is the theory of **fo points**, due to 2005 Nobel laureate Thomas C. Schelling. According to t theory, some equilibria tend to stand out in the minds of the players. Schelli predicts that people will frequently succeed in selecting such equilibria. T precise feature of an equilibrium that makes it stand out in the minds of t players is far from obvious.

> Finding the key ... may depend on imagination more than on logic; it m depend on analogy, precedent, accidental arrangement, symmetry, ae thetic or geometric configuration, casuistic reasoning, and who the parti are and what they know about each other.

This theory can explain why people favor ⟨U, L⟩ over ⟨D, R⟩ in Table 10.9( When there is a unique Pareto-optimal outcome that also happens to be Nash equilibrium, it seems plausible to assume that people will use Pare optimality as a focal point. If so, we might be able to explain observed beha ior without making the transition to trembling-hand-perfect equilibrium.

In the next chapter, we will explore behavioral economists' challenge analytical game theory.

## ADDITIONAL EXERCISES

**Exercise 10.27 Paradoxes of rationality** Experimental economists have invited students with different majors to play prisoners' dilemma games against each other. In an experiment pitching economics majors against economics majors, and non-majors against non-majors, who would you expect to do better?

Chapter 11 contains more game-theoretic exercises.

## FURTHER READING

There are many fine introductions to game theory, including Binmore (2007), Dixit et al. (2009), and Osborne and Rubinstein (1994). *Spousonomics* is Szuchman and Anderson (2011, pp. xii–xv, 294–98). Life in the state of nature is described in Hobbes (1994 [1651], xiii, 8–9, p. 76). Skyrms (1996) discusses the doctrine of mutually assured destruction (pp. 22–25) and the games of chicken and hawk & dove (pp. 65–67); Russell (1959, p. 30) examines the game of chicken. The theory of focal points is due to Schelling (1960, p. 57). Evidence about economics majors' performance in prisoner dilemma games can be found in Frank et al. (1993).