Chapter 1
The name of the game

What is game theory about?

When my wife was away for the day at a pleasant little conference in Tuscany, three young women invited me to share their table for lunch. As I sat down, one of them said in a sultry voice, 'Teach us how to play the game of love', but it turned out that all they wanted was advice on how to manage Italian boyfriends. I still think they were wrong to reject my strategic recommendations, but they were right on the nail in taking for granted that courting is one of the many different kinds of game we play in real life.

Drivers manoeuvring in heavy traffic are playing a driving game. Bargain-hunters bidding on eBay are playing an auctioning game. A firm and a union negotiating next year's wage are playing a bargaining game. When opposing candidates choose their platform in an election, they are playing a political game. The owner of a grocery store deciding today's price for corn flakes is playing an economic game. In brief, a game is being played whenever human beings interact.

Antony and Cleopatra played the courting game on a grand scale. Bill Gates made himself immensely rich by playing the computer software game. Adolf Hitler and Josef Stalin played a game that killed off a substantial fraction of the world's population. Kruschev and Kennedy played a game during the Cuban missile crisis that might have wiped us out altogether.

With such a wide field of application, game theory would be a universal panacea if it could always predict how people will play the many games of which social life largely consists. But game theory isn't able to solve all of the world's problems, because it only works when people play games *rationally*. So it can't predict the behaviour of love-sick teenagers like Romeo or Juliet, or madmen like Hitler or Stalin. However, people don't always behave irrationally, and so it isn't a waste of time to study what happens when people put on their thinking caps. Most of us at least try to spend our money sensibly – and we don't do too badly much of the time or economic theory wouldn't work at all.

Even when people haven't thought everything out in advance, it doesn't follow that they are necessarily behaving irrationally. Game theory has had some notable successes in explaining the behaviour of spiders and fish, neither of which can be said to think at all. Such mindless animals end up behaving as though they were rational, because rivals whose genes programmed them to behave irrationally are now extinct. Similarly, companies aren't always run by great intellects, but the market is often just as ruthless as Nature in eliminating the unfit from the scene.

Does game theory work?

In spite of its theoretical successes, practical men of business used to dismiss game theory as just one more ineffectual branch of social science, but they changed their minds more or less overnight after the American government decided to auction off the right to use various radio frequencies for use with cellular telephones.

With no established experts to get in the way, the advice of game theorists proved decisive in determining the design of the rules of the auctioning games that were used. The result was that the

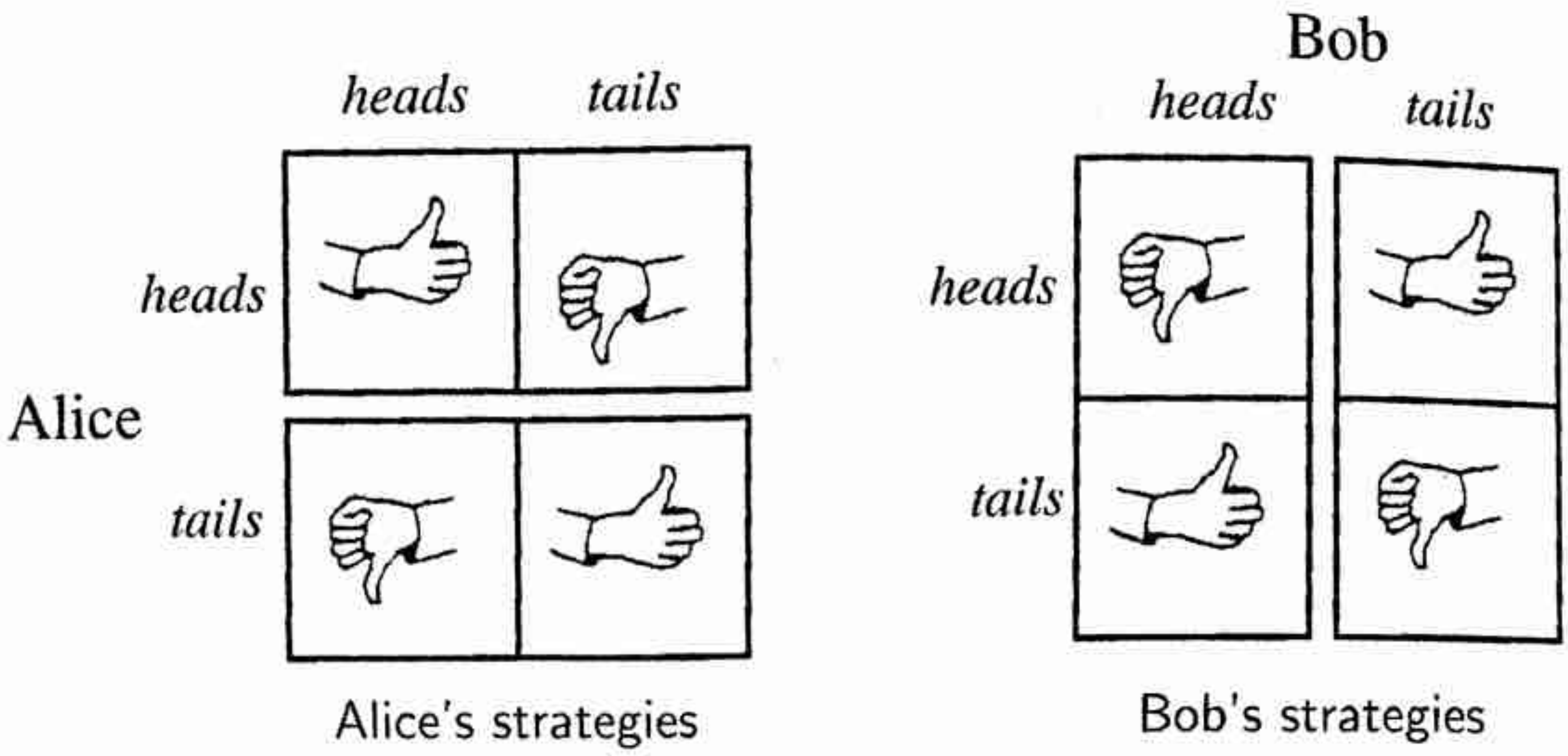
American taxpayer made a profit of $20 billion – more than twice the orthodox prediction. Even more was made in a later British telecom auction for which I was responsible. We made a total of $35 billion in just one auction. In consequence, *Newsweek* magazine described me as the ruthless, Poker-playing economist who destroyed the telecom industry!

As it turned out, the telecom industry wasn't destroyed. Nor is it at all ruthless to make the fat cats of the telecom industry pay for their licences what they think they are worth – especially when the money is spent on hospitals for those who can't afford private medical care. As for Poker, it is at least 20 years since I played for more than nickels and dimes. The only thing that *Newsweek* got right is that game theory really does work when applied by people who know what they are doing. It works not just in economics, but also in evolutionary biology and political science. In my recent book *Natural Justice*, I even outrage orthodox moral philosophers by using game theory when talking about ethics.

Toy games

Each new big-money telecom auction needs to be tailored to the circumstances in which it is going to be run. One can't just take a design off the shelf, as the American government found when it hired Sotheby's to auction off a bunch of satellite transponders. But nor can one capture all the complicated ins and outs of a new telecom market in a mathematical model. Designing a telecom auction is therefore as much an art as a science. One extrapolates from simple models chosen to mimic what seem to be the essential strategic features of a problem.

I try to do the same in this book, which therefore contains no algebra and a minimum of technical jargon. It looks only at toy games, leaving aside all the bells and whistles with which they are complicated in real life. However, most people find that even toy games give them plenty to think about.



1. Alice and Bob's decision problem in Matching Pennies

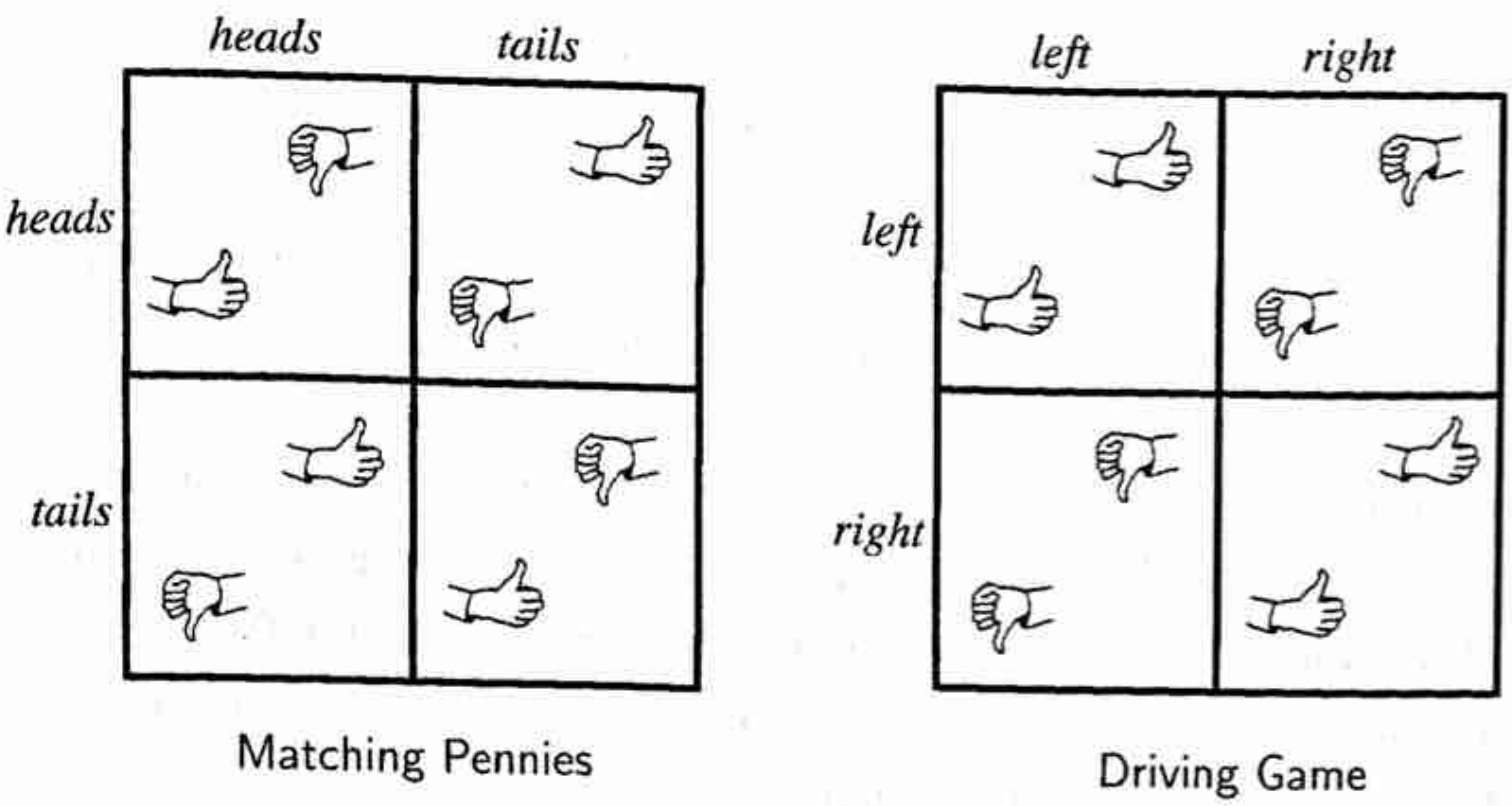
Conflict and cooperation

Most of the games in this book have only two players, called Alice and Bob. The first game they will play is Matching Pennies.

Sherlock Holmes and the evil Professor Moriarty played Matching Pennies on the way to their final confrontation at the Reichenbach Falls. Holmes had to decide at which station to get off a train. Moriarty had to decide at which station to lie in wait. A real-life counterpart is played by dishonest accountants and their auditors. The former decide when to cheat and the latter decide when to inspect the books.

In our toy version, Alice and Bob each show a coin. Alice wins if both coins show the same face. Bob wins if they show different faces. Alice and Bob therefore each have two strategies, *heads* and *tails*. Figure 1 shows who wins and loses for all possible strategy combinations. These outcomes are the players' *payoffs* in the game. The thumbs-up and thumbs-down icons have been used to emphasize that payoffs needn't be measured in money.

Figure 2 shows how all the information in Figure 1 can be assembled into a payoff table, with Alice's payoff in the southwest corner of each cell, and Bob's in the northeast corner. It also shows a two-player version of the very different Driving Game that we



2. Payoff tables. Alice chooses a row and Bob chooses a column

play every morning when we get into our cars to drive to work. Alice and Bob again have two pure strategies, *left* and *right*, but now the players' payoffs are totally aligned instead of being diametrically opposed. When journalists talk about a win-win situation, they have something like the Driving Game in mind.

Von Neumann

The first result in game theory was John Von Neumann's minimax theorem, which applies only to games like Matching Pennies in which the players are modelled as implacable enemies. One sometimes still reads dismissive commentaries on game theory in which Von Neumann is caricatured as the archetypal cold warrior – the original for Dr Strangelove in the well known movie. We are then told that only a crazed military strategist would think of applying game theory in real life, because only a madman or a cyborg would make the mistake of supposing that the world is a game of pure conflict.

Von Neumann was an all-round genius. Inventing game theory was just a sideline for him. It is true that he was a hawk in the Cold War, but far from being a mad cyborg, he was a genial soul, who liked to party and have a good time. Just like you and me, he preferred cooperation to conflict, but he also understood that the way to achieve cooperation isn't to pretend that people can't sometimes profit by causing trouble.

Cooperation and conflict are two sides of the same coin, neither of which can be understood properly without taking account of the other. To consider a game of pure conflict like Matching Pennies isn't to claim that all human interaction is competitive. Nor is one claiming that all human interaction is cooperative when one looks at a game of pure coordination like the Driving Game. One is simply distinguishing two different aspects of human behaviour so that they can be studied one at a time.

Revealed preference

To cope with cooperation and conflict together, we need a better way of describing the motivation of the players than simply saying that they like winning and dislike losing. For this purpose, economists have invented the idea of *utility*, which allows each player to assign a numerical value to each possible outcome of a game.

In business, the bottom line is commonly profit, but economists know that human beings often have more complex aims than simply making as much money as they can. So we can't identify utility with money. A naive response is to substitute happiness for money. But what is happiness? How do we measure it?

It is unfortunate that the word 'utility' is linked historically with Victorian utilitarians like Jeremy Bentham and John Stuart Mill, because modern economists don't follow them in identifying utility with how much pleasure or how little pain a person may feel. The modern theory abandons any attempt to explain how people behave in terms of what is going on inside their heads. On the contrary, it makes a virtue of making no psychological assumptions at all.

We don't try to explain *why* Alice or Bob behave as they do. Instead of an explanatory theory, we have to be content with a descriptive theory, which can do no more than say that Alice or Bob will be acting inconsistently if they did such-and-such in the past, but now plan to do so-and-so in the future. In game theory, the object is to observe the decisions that Alice and Bob make (or would make) when they aren't interacting with each other or anyone else, and to deduce how they will behave when interacting in a game.

We therefore don't argue that some preferences are more rational than others. We follow the great philosopher David Hume in regarding reason as the 'slave of the passions'. As he extravagantly remarked, there would be nothing *irrational* about his preferring the destruction of the entire universe to scratching his finger. However, we go even further down this road by regarding reason purely as an instrument for avoiding inconsistent behaviour. Any consistent behaviour therefore counts as rational.

With some mild assumptions, acting consistently can be shown to be the same as behaving as though seeking to maximize the value of something. Whatever this abstract something may be in a particular context, economists call it utility. It needn't correlate with money, but it sadly often does.

Taking risks

In acting consistently, Alice may not be aware that she is behaving as though maximizing something we choose to call her utility. But if we want to predict her behaviour, we need to be able to measure her utility on a utility scale, much as temperature is measured on a thermometer. Just as the units on a thermometer are called degrees, we can then say that a *util* is a unit on Alice's utility scale.

The orthodoxy in economics used to be that such cardinal utility scales are intrinsically nonsensical, but Von Neumann fortunately didn't know this when Oskar Morgenstern turned up at his house one day complaining that they didn't have a proper basis for the numerical payoffs in the book on game theory they were writing together. So Von Neumann invented a theory on the spot that measures how much Alice wants something by the size of the risk she is willing to take to get it. We can then figure out what choice she will make in risky situations by finding the option that will give her the highest utility on average.

It is easy to use Von Neumann's theory to find how much utility to assign to anything Alice may need to evaluate. For example, how many utils should Alice assign to getting a date with Bob?

We first need to decide what utility scale to use. For this purpose, pick two outcomes that are respectively better and worse than any other outcome Alice is likely to encounter. These outcomes will correspond to the boiling and freezing points of water used to calibrate a Celsius thermometer, in that the utility scale to be constructed will assign 0 utils to the worst outcome, and 100 utils to the best outcome. Next consider a bunch of (free) lottery tickets in which the only prizes are either the best outcome or the worst outcome.

When we offer Alice lottery tickets with higher and higher probabilities of getting the best outcome as an alternative to a date with Bob, she will eventually switch from saying *no* to saying *yes*. If the probability of the best outcome on the lottery ticket that makes her switch is 75%, then Von Neumann's theory says that a date with Bob is worth 75 utils to her. Each extra percentage point added to her indifference probability therefore corresponds to one extra util.

When some people evaluate sums of money using this method, they always assign the same number of utils to each extra dollar. We call such people risk neutral. Those who assign fewer utils to each extra dollar than the one that went before are called risk averse.

Insurance

Alice is thinking of accepting an offer from Bob to insure her Beverley Hills mansion against fire. If she refuses his offer, she faces a lottery in which she ends up with her house plus the insurance premium if her house doesn't burn down, and with only the premium if it does. This has to be compared with her ending up for sure with the value of the house less the premium if she accepts Bob's offer.

If it is rational for Bob to make the offer and for Alice to accept, he must think that the lottery is better than breaking even for sure, and she must have the opposing preference. The existence of the insurance industry therefore confirms not only that it can be rational to gamble – provided that the risks you take are calculated risks – but that rational people can have different attitudes to taking risks. In the insurance industry, the insurers are close to being risk neutral and the insurees are risk averse to varying degrees.

Notice that economists regard the degree of risk aversion that a person reveals as a matter of personal preference. Just as Alice may or may not prefer chocolate ice-cream to vanilla, so she may or may not prefer to spend $1,000 on insuring her house. Some philosophers – notably John Rawls – insist that it is *rational* to be risk averse when defending whatever alternative to maximizing average utility they prefer, but such appeals miss the point that the players' attitudes to taking risks have already been taken into account when using Von Neumann's method to assign utilities to each outcome.

Economists make a different mistake when they attribute risk aversion to a dislike of the act of gambling. Von Neumann's theory only makes sense when the players are entirely neutral to the actual act of gambling. Like a Presbyterian minister insuring his house, they don't gamble because they enjoy gambling – they gamble only when they judge that the odds are in their favour.

	heads	tails
heads	-1 +1	+1 -1
tails	+1 -1	-1 +1

Matching Pennies

	left	right
left	+1 +1	-1 -1
right	-1 -1	+1 +1

Driving Game

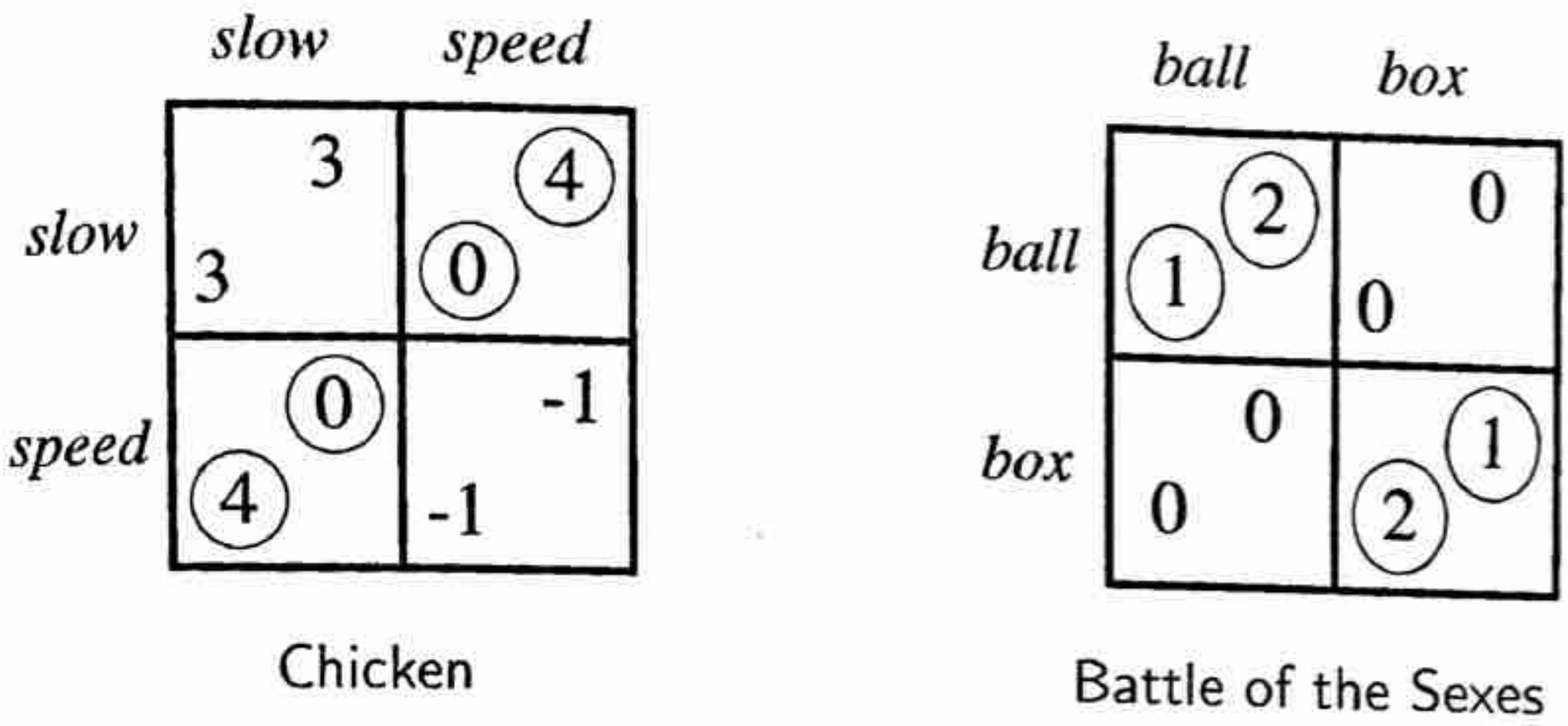
3. Numerical payoffs

Life isn't a zero-sum game

As with measuring temperature, we are free to choose the zero and the unit on Alice's utility scale however we like. We could, for example, have assigned 32 utils to the worst outcome, and 212 utils to the best outcome. The number of utils a date with Bob is worth on this new scale is found in the same way that one converts degrees Celsius into degrees Fahrenheit. So the date with Bob that was worth 75 utils on the old scale would be worth 167 utils on the new scale.

In the toy games we have considered so far, Alice and Bob have only the outcomes WIN and LOSE to evaluate. We are free to assign these two outcomes any number of utils we like, as long as we assign more utils to winning than to losing. If we assign plus one util to winning and minus one util to losing, we get the payoff tables of Figure 3.

The payoffs in each cell of Matching Pennies in Figure 3 always add up to zero. We can always fix things to make this true in a game of pure conflict. Such games are therefore said to be *zero sum*. When gurus tell us that life isn't a zero-sum game, they therefore aren't saying anything about the total sum of happiness in the world. They are just reminding us that the games we play in real life are seldom games of pure conflict.



4. Games with mixed motivations

Nash equilibrium

The old movie *Rebel without a Cause* still occasionally gets a showing because it stars the unforgettable James Dean as a sexy teenage rebel. The game of Chicken was invented to commemorate a scene in which he and another boy drive cars towards a cliff edge to see who will chicken out first. Bertrand Russell famously used the episode as a metaphor for the Cold War.

I prefer to illustrate Chicken with a more humdrum story in which Alice and Bob are two middle-aged drivers approaching each other in a street too narrow for them to pass safely without someone slowing down. The strategies in Figure 4 are therefore taken to be *slow* and *speed*.

The new setting downplays the competitive element of the original story. Chicken differs from zero-sum games like Matching Pennies because the players also have a joint interest in avoiding a mutual disaster.

The stereotypes embedded in the Battle of the Sexes pre-date the female liberation movement. Alice and Bob are a newly married couple honeymooning in New York. At breakfast, they discuss whether to go to a boxing match or the ballet in the evening, but

5. James Dean

fail to make a decision. They later get separated in the crowds and now each has to decide independently where to go in the evening.

The story that accompanies the Battle of the Sexes emphasizes the cooperative features of their problem, but there is also a conflictual element absent from the Driving Game, because each player prefers that they coordinate on a different outcome. Alice prefers the ballet and Bob the boxing match.

John Nash

Everybody has heard of John Nash now that his life has been featured in the movie *A Beautiful Mind.* As the movie documents, the highs and lows of his life are out of the range of experience of most human beings. He was still an undergraduate when he initiated the modern theory of rational bargaining. His graduate thesis formulated the concept of a Nash equilibrium, which is now regarded as the basic building block of the theory of games. He went on to solve major problems in pure mathematics, using methods of such originality that his reputation as a mathematical genius of the first rank became firmly established. But he fell prey

6. John Nash

to a schizophrenic illness that wrecked his career and finally left him to languish in obscurity for more than 40 years as an object of occasional mockery on the Princeton campus. His recovery in time to be awarded a Nobel Prize in 1994 seems almost miraculous in retrospect. But as Nash comments, without his 'madness', he would perhaps only have been another of the faceless multitudes who have lived and died on this planet without leaving any trace of their existence behind.

However, one doesn't need to be a wayward genius to understand the idea of a Nash equilibrium. We have seen that the payoffs in a game are chosen to make it tautological that rational players will seek to maximize their average payoff. This would be easy if players knew what strategies their opponents were going to choose. For example, if Alice knew that Bob were going to choose *ball* in the Battle of the Sexes, she would maximize her payoff by choosing *ball* as well. That is to say, *ball* is Alice's best reply to Bob's choice of *ball*, a fact indicated in Figure 4 by circling Alice's payoff in the cell that results if both players choose *ball*.

A Nash equilibrium is just a pair of strategies whose use results in a cell in which *both* payoffs are circled. More generally, a Nash equilibrium occurs when all the players are simultaneously making a best reply to the strategy choices of the others.

Both (*box*, *box*) and (*ball*, *ball*) are therefore Nash equilibria in the Battle of the Sexes. Similarly, (*slow*, *speed*) and (*speed*, *slow*) are Nash equilibria in Chicken.

Why should we care about Nash equilibria? There are two major reasons. The first supposes that ideally rational players reason their way to a solution of a game. The second supposes that people find their way to a solution by some evolutionary process of trial and error. Much of the predictive power of game theory arises from the possibility of passing back and forth between these alternative interpretations. We seldom know much about the details of evolutionary processes, but we can sometimes leap ahead to predict where they will eventually end up by asking what rational players would do in the situation under study.

Rational interpretation

Suppose that somebody even cleverer than Nash or Von Neumann had written a book that lists all possible games along with an authoritative recommendation on how each game should be played by rational players. Such a great book of game theory would necessarily have to pick a Nash equilibrium as the solution of each game. Otherwise it would be rational for at least one player to deviate from the book's advice, which would then fail to be authoritative.

Suppose, for example, that the book recommended that teenage boys playing Chicken should both choose *slow* as their mothers would wish. If the book were authoritative, each player would then know that the other was going to play *slow*. But a rational player in Chicken who knows that his opponent is going to choose *slow* will necessarily choose *speed*, thereby refuting the book's claim to be authoritative.

Notice that the reasoning in this defence of Nash equilibria is circular. Why does Alice play this way? Because Bob plays that way. Why does Bob play that way? Because Alice plays this way.

Various Latin tags are available to those who are unhappy with such circular arguments. When first accused of committing the fallacy of *circulus in probando* when talking about equilibria, I had to go and look it up. It turns out that I was lucky not to have been accused of the even more discreditable *petitio principii*. But all arguments must obviously either be circular or reduce to an infinite regression if one never stops asking *why*. Dictionary definitions are the most familiar example.

In games, we can either forever contemplate the infinite regression that begins:

> Alice thinks that Bob thinks that Alice thinks that Bob thinks . . .

or else take refuge in the circularity built into the idea of a Nash equilibrium. This short circuits the infinite regression by observing that any other strategy profile will eventually be destabilized when the players start thinking about what the other players are thinking. Or to say the same thing another way, if the players' beliefs about each other's plans are to be consistent, then they must be in equilibrium.

Evolutionary interpretation

The rational interpretation of Nash equilibrium had such a grip on early game theorists that the evolutionary interpretation was almost entirely neglected. The editors of the journal in which Nash published his paper on equilibria even threw out his remarks on this subject as being without interest! But game theory would never be able to predict the behaviour of ordinary people if the evolutionary interpretation were invalid. For example, the famous mathematician Emile Borel thought about game theory before Von Neumann but came to the conclusion that the minimax theorem was probably false. So what hope would there be for the rest of us, if even someone as clever as Borel couldn't reason his way to a solution of the simplest class of games!

There are many possible evolutionary interpretations of Nash equilibria, which differ in the adjustment process by means of which players may find their way to an equilibrium. In the simpler adjustment processes, the payoffs in a game are identified with how fit the players are. Processes that favour fitter strategies at the expense of their less successful brethren can then only stop working when we get to a Nash equilibrium, because only then will all the surviving strategies be as fit as it is possible to be in the circumstances. We therefore don't need our players to be mathematical whizzes for Nash equilibria to be relevant. They often predict the behaviour of animals quite well. Nor is the evolutionary significance of Nash equilibria confined to biology. They have a predictive role whenever an adjustment process tends to eliminate strategies that generate low payoffs.

For example, stockbrokers who do less well than their competitors go bust. The rules-of-thumb that stockbrokers use are therefore subject to the same kind of evolutionary pressures as the genes of fish or insects. It therefore makes sense to look at Nash equilibria in the games played by stockbrokers, even though we all know that some stockbrokers wouldn't be able to find their way around a goldfish bowl, let alone a game theory book.

Prisoner's Dilemma

The most famous toy game of all is the Prisoner's Dilemma. In the traditional story used to motivate the game, Alice and Bob are gangsters in the Chicago of the 1920s. The District Attorney knows that they are guilty of a major crime, but is unable to convict either unless one of them confesses. He orders their arrest, and separately offers each the following deal:

> If you confess and your accomplice fails to confess, then you go free. If you fail to confess but your accomplice confesses, then you will be convicted and sentenced to the maximum term in jail. If you both confess, then you will both be convicted, but the maximum sentence will not be imposed. If neither confesses, you will both be framed on a tax evasion charge for which a conviction is certain.

The story becomes more poignant if Alice and Bob have agreed to keep their mouths shut if ever put into such a situation. Holding out then corresponds to cooperating and confessing to defecting, as in the table on the left of Figure 7. The payoffs in the table correspond to notional years in jail (on the assumption that one util always corresponds to one extra year of freedom).

A less baroque story assumes that Alice and Bob each have access to a pot of money. Both are independently allowed either to give their opponent \$2 from the pot, or to put \$1 from the pot in their own pocket. On the assumption that Alice and Bob care only about money, we are led to the payoff table on the right of Figure 7 in which utils have been identified with dollars. In this case, the altruistic strategy of giving \$2 has been assigned the label *dove*,

	defect	*coop*
defect	−9 −9	−10 0
coop	0 −10	−1 −1

gangster version

	dove	*hawk*
dove	2 2	3 0
hawk	0 3	1 1

give-or-take version

7. Two versions of the Prisoner's Dilemma: in the version on the right, *dove* represents giving and *hawk* represents taking

and the selfish strategy of taking $1 has been assigned the label *hawk*.

Circling best replies reveals that the only Nash equilibrium in the give-or-take version of the Prisoner's Dilemma is for both Alice and Bob to play *hawk*, although each would get more if they both played *dove*. The gangster version is strategically identical. In the unique Nash equilibrium, each will defect, with the result that they will both spend a long time in jail, although each would get a much lighter sentence if they both cooperated.

Paradox of rationality?

A whole generation of scholars swallowed the line that the Prisoner's Dilemma embodies the essence of the problem of human cooperation. They therefore set themselves the hopeless task of giving reasons why game theory's resolution of this supposed 'paradox of rationality' is mistaken (See Fallacies of the Prisoner's Dilemma, Chapter 10). But game theorists think it just plain wrong that the Prisoner's Dilemma captures what matters about human cooperation. On the contrary, it represents a situation in which the dice are as loaded against the emergence of cooperation as they could possibly be.

If the great game of life played by the human species were adequately modelled by the Prisoner's Dilemma, we wouldn't have evolved as social animals! We therefore see no more need to solve an invented paradox of rationality than to explain why people drown when thrown into Lake Michigan with their feet encased in concrete. No paradox of rationality exists. Rational players don't cooperate in the Prisoner's Dilemma because the conditions necessary for rational cooperation are absent.

Fortunately the paradox-of-rationality phase in the history of game theory is just about over. Insofar as they are remembered, the many fallacies that were invented in hopeless attempts to show that it is rational to cooperate in the Prisoner's Dilemma are now mostly quoted as entertaining examples of what psychologists call magical reasoning, in which logic is twisted to secure some desired outcome. My favourite example is Immanuel Kant's claim that rationality demands obeying his categorical imperative. In the Prisoner's Dilemma, rational players would then all choose *dove*, because this is the strategy that would be best if everybody chose it.

Domination

The idea that it is necessarily irrational to do things that would be bad if everybody did them is very pervasive. Your mother was probably as fond of this argument as mine. The following knock-down refutation in the case of the Prisoner's Dilemma is therefore worth repeating.

So as not to beg any questions, we begin by asking where the payoffs that represent the players' preferences in the Prisoner's Dilemma come from. The theory of revealed preference tells us to find the answer by observing the choices that Alice and Bob make (or would make) when solving one-person decision problems.

Writing a larger payoff for Alice in the bottom-left cell of the payoff table of the Prisoner's Dilemma than in the top-left cell therefore means that Alice would choose *hawk* in the one-person decision problem that she would face if she knew in advance that Bob had chosen *dove*. Similarly, writing a larger payoff in the bottom-right cell means that Alice would choose *hawk* when faced with the one-person decision problem in which she knew in advance that Bob had chosen *hawk*.

The very definition of the game therefore says that *hawk* is Alice's best reply when she knows that Bob's choice is *dove*, and also when she knows his choice is *hawk*. So she doesn't need to know anything about Bob's actual choice to know her best reply to it. It is rational for her to play *hawk* whatever strategy he is planning to choose. In this unusual circumstance, we say that *hawk* dominates Alice's alternative strategies.

Objections?

Two objections to the preceding analysis are common. The first denies that Alice would choose to defect in the gangster version of the Prisoner's Dilemma if she knew that Bob had chosen to cooperate. Various reasons are offered that depend on what one believes about conditions in Al Capone's Chicago, but such objections miss the point. If Alice wouldn't defect if she knew that Bob had chosen to cooperate, then she wouldn't be playing the Prisoner's Dilemma. Here and elsewhere, it is important not to take the stories used to motivate games too seriously. It is the payoff tables of Figure 7 that define the Prisoner's Dilemma – not the silly stories that accompany them.

The second objection always puzzles me. It is said that appealing to the theory of revealed preference reduces the claim that it is rational to defect in the Prisoner's Dilemma to a tautology. Since tautologies have no substantive content, the claim can therefore be ignored! But who would say the same of $2 + 2 = 4$?

Experiments

An alternative response is to argue that it doesn't matter what is rational in the Prisoner's Dilemma, because laboratory experiments show that real people actually play *dove*. The payoffs in such experiments aren't usually determined using the theory of revealed preference. They are nearly always just money, but the results can nevertheless be very instructive.

Inexperienced subjects do indeed cooperate a little more than half the time on average, but the evidence is overwhelming in games like the Prisoner's Dilemma that the rate of defection increases steadily as the subjects gain experience, until only about 10% of subjects are still cooperating after ten trials or so.

Computer simulations are also mentioned which supposedly show that evolution will eventually generate cooperation in the Prisoner's Dilemma, but such critics have usually confused the Prisoner's Dilemma with its indefinitely repeated cousin in which cooperation is indeed a Nash equilibrium (See Tit-for-tat, Chapter 5).