1 Introduction

One thing I learned from Pop was to try to think
as people around you think.
And on that basis, anything's possible.
Al Pacino alias Michael Corleone
in *The Godfather – Part II*

What is this book about? In our lives we are continuously asked to make choices – small choices about daily issues, but also big choices with important consequences. Often, the final outcome of a choice does not only depend on our own decision, but also on decisions made by other people surrounding us. For instance, if you are about to meet a couple of friends in a pub this evening, then whether you have to wait for your friends or not depends on your own arrival time, but also on the times your friends arrive. And if you are at a flea market negotiating about the price of a rare Beatles record, then the final price will not only depend on your own bargaining strategy, but also on that of the seller.

Such situations, in which the final outcome does not only depend on your own choice but also on the choices of others, are called *games*. The discipline that studies such situations is called *game theory*. The name *game theory* is perhaps a bit misleading as it suggests that its main application is to recreational games – such as chess or poker – but this is not true. In fact, game theory can be applied to *any* situation where you must make a choice, and in which the final outcome also depends on the decisions of others. The people whose choices directly influence the final outcome are called *players* – so you are one of the players in the two real-life situations sketched above – and we usually refer to the *other* players as your *opponents*, even if these are your friends.

In order to evaluate the possible consequences of your own choice, it is important to form some *belief* about the likely choices of your opponents, as these will affect the final result. Moreover, to make a *good* choice it is necessary to form a *reasonable* belief about your opponents. But in general not every belief about your opponents will be reasonable: your opponents will have in general some choices that seem more plausible than others. But to determine which choices of your opponent are plausible and which are not, you must put yourself in the shoes of the opponent, and think about the possible beliefs *he* can have about *his* opponents. That is, you must *reason about your opponents*

before you can form a reasonable belief about them. And this reasoning process will precisely be the main topic of this book.

More precisely, this book studies several plausible ways in which you can reason about your opponents before you make your final choice in a game. As different people reason differently, we do not believe in a *unique* way of reasoning. Instead, we offer the reader a *spectrum* of plausible reasoning patterns in this book. We also investigate how your eventual decision will be affected by the type of reasoning you use. The discipline that studies these patterns of reasoning, and how they influence the eventual choices of the players, is called *epistemic game theory*. This explains the title of this book, *Epistemic Game Theory: Reasoning and Choice*.

Why this book? We have just seen that reasoning about your opponents is a crucial step towards making a good choice in a game. In fact, Oskar Morgenstern – one of the early founders of game theory – had stressed the importance of this reasoning process, in particular to form beliefs about the beliefs of others, in his paper Morgenstern (1935). But strangely enough it is exactly this reasoning step that has largely been overlooked by the game theory literature – including the game theory textbooks – during the last sixty years! This immediately raises the question: Why? In our opinion, this phenomenon is largely due to the concept of *Nash equilibrium* and its various refinements, which have dominated the game theory literature for so many decades. Nash equilibrium describes just one possible way – and not even a very plausible one – in which the players in a game may choose, or form a belief about the opponents' choices. Yet many game theory textbooks and articles *assume* that the players' reasoning process will eventually lead them to choose in accordance with Nash equilibrium, without explicitly describing this reasoning process.

We find this approach unsatisfactory for two reasons. First, we believe that the reasoning is an essential part of the decision-making process for a player in a game, and hence deserves to be discussed explicitly. Second, we will see in this book that Nash equilibrium is based on some rather implausible assumptions about the way players reason, which makes Nash equilibrium a rather unnatural concept to use when analyzing the reasoning of players in a game. That is also the reason why Nash equilibrium plays only a minor role in this book – it is only discussed in Chapter 4.

Things started to change with the rise of *epistemic game theory*, some twenty-five years ago. This relatively young discipline attempts to bring game theory back to its basic elements – namely the reasoning by players about their opponents. In recent years it has developed a whole spectrum of concepts that are based on more plausible assumptions than Nash equilibrium. But to date there is no textbook on epistemic game theory, nor is there any other game theory textbook that focuses on the reasoning process of the players. The aim of this book is to fill this gap, by providing a text that concentrates on the way people can reason about their opponents before making a choice in a game. This book will thus be the first textbook on epistemic game theory. We feel there is a need for such a textbook, because reasoning about your opponents is such an important and natural ingredient of the decision-making process in games.

Moreover, for researchers it will also be valuable to have a book that discusses the various concepts from epistemic game theory in a systematic and unified way.

While this book was being written, Pierpaolo Battigalli, Adam Brandenburger, Amanda Friedenberg and Marciano Siniscalchi were also working on a book on *epistemic game theory*, and so were Eric Pacuit and Olivier Roy but their books were not finished at the time this introduction was written. So, shortly there will be two new books on epistemic game theory on the market, which is very good news for the field. The first book mentioned above will be rather complementary to ours, as it will treat topics like games with incomplete information, infinite games and psychological games, which are not covered by our book.

Intended audience. This textbook is primarily written for advanced bachelor students, master students and Ph.D. students taking a course in game theory. This course could either be an introductory course or a more specialized follow-up course. In fact, the book does not presuppose any knowledge about game theory, and should thus be accessible also for students who have not studied game theory before. Moreover, the mathematics that we use in this book is very elementary, and the book can therefore be used within any program that has a game theory course in its curriculum. But the book can also be used for self-study by researchers in the field, or people who want to learn about epistemic game theory.

Structure of the book. The book has been divided into three parts, according to the type of game and type of beliefs we consider.

In Part I we assume that the game is *static* – that is, all players choose only once, and in complete ignorance of the opponents' choices. In this part we assume moreover that the belief of a player about the opponents is a *standard belief*, represented by a *single* probability distribution. This is the type of belief that is most commonly used for static games. Part I includes Chapters 2–4.

In Part II, which contains Chapters 5–7, we still assume that the game is static, but now we model the belief of a player about the opponents by a *lexicographic belief*, consisting of *various* probability distributions instead of only one. Lexicographic beliefs are particularly useful if we want to model *cautious reasoning* – that is, a state of mind in which you do not completely rule out *any* opponent's choice from consideration. This type of belief is not as well known as standard beliefs, but we believe it is a very natural way to describe cautious reasoning about your opponents.

Part III, which contains Chapters 8 and 9, is dedicated to *dynamic games*, where players may choose one after the other, and may fully or partially observe what the opponents have done in the past before making a choice themselves. For such games we model the belief of a player about the opponents by a sequence of *conditional beliefs* – one at every point in time where the player must make a choice. So, instead of holding just one belief once and for all, the player holds a separate belief at every instance when a choice has to be made, and the player's belief about the opponents may *change* as the game proceeds.

The first two chapters of Part I – Chapter 2 and Chapter 3 – form the basis for this book, as they introduce the central idea of *common belief in rationality*. This concept

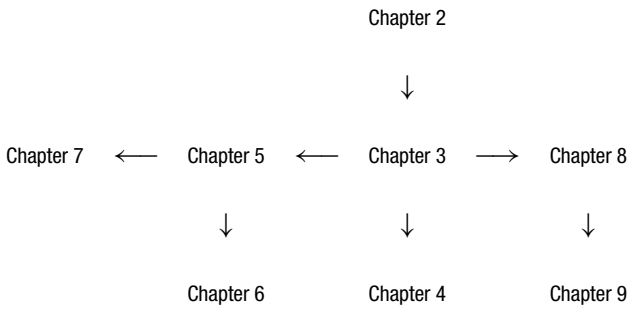

Figure 1.1 Logical connection between the chapters

is at the heart of epistemic game theory, and all other concepts in this book may be viewed as variations on the idea of common belief in rationality. Every other chapter is about exactly one such variation, describing one particular way of reasoning about your opponents. Some of these chapters can be read independently of each other, whereas others cannot. Figure 1.1 gives an overview of the logical connection between the chapters. Here, the arrow from Chapter 3 to Chapter 4 means that Chapter 4 builds upon Chapter 3. As there is no arrow between Chapter 4 and Chapter 5, you should be able to understand Chapter 5 without having read Chapter 4. The other arrows are to be interpreted in a similar fashion.

Players can be male or female, but for simplicity we will often refer to a player as "he" or "him."

Structure of the chapters. Every chapter in this book is on one particular concept, describing one possible way of reasoning about your opponents. The main structure of each chapter is as follows: We begin the chapter with one or more examples, illustrating the way of reasoning we will discuss in that chapter. Subsequently, we show how this particular way of reasoning can be formally described within an *epistemic model* – that is, a model that describes the belief a player holds about the opponents' choices, the belief held about the opponents' beliefs about the other players' choices, and so on. In most of the chapters we then provide a simple *existence proof* for this concept, showing that this particular way of reasoning is always *possible* within any game, so never leads to logical contradictions. For every concept – except Nash equilibrium – we design an *algorithm* that yields for every game precisely those choices you can rationally make if you reason in accordance with that concept. Such an algorithm typically proceeds by iteratively removing choices or strategies from the game. An exception is the algorithm we discuss in Chapter 6. These algorithms are relatively easy to use, and greatly facilitate the task of finding those choices that are selected by the concept. All formal proofs are given in a separate section at the end of the chapter, and these proofs are mainly for Ph.D. students and researchers. In every chapter we also provide seven practical problems and three theoretical problems that the student or reader can work on. Every chapter concludes with a literature section in which

we discuss the relevant literature for that chapter, and provide references to relevant articles, books and book chapters. We have decided not to include these references in the main text of the chapter, as we believe this would distract the student's attention too much from the main ideas in that chapter. After all, this is a textbook and not a monograph or survey.

One-person perspective. Throughout this book we take a *one-person perspective* to analyze game-theoretic situations. That is, we always view the game from the perspective of *one single player*, and put restrictions only on the beliefs of this particular player – including beliefs about the opponents' beliefs – without imposing restrictions on the *actual* beliefs of the opponents. We believe this approach to be plausible, as we cannot look inside the minds of our opponents at the time we make a choice. So, you can only base your choice on your *own* beliefs about the opponents, and not on the actual beliefs and choices of your opponents, since these are not known to you. But then, if we want to analyze the reasonable choices a player can make in a game, it is sufficient to concentrate only on the beliefs of this particular player, as they encompass everything that can be used to make a decision. Although we believe the one-person perspective to be very natural, it crucially differs from the usual approach to games in books and articles, which typically proceed by imposing restrictions on the beliefs of *all* players, and not only one player.

Descriptive versus normative approach. In this book we do not tell people what they should do or believe in a certain game. We only explore different intuitive ways of reasoning that you could use in a game, and see how these various ways of reasoning would affect the final choice – or choices – you could rationally make *if* you were to use this particular way of reasoning. But it is eventually up to the reader, or the modeler, to choose the preferred way of reasoning. That is, we take a purely *descriptive* approach in this book, and *not* a *normative* one.

In Chapters 6 and 7 of this book we even introduce two ways of reasoning that in some games lead to completely opposing choices! We do not believe that one of these two concepts is better – or more intuitive – than the other, they are just different. In fact, we believe that both ways of reasoning are very plausible, so we see no problem in presenting both concepts next to each other, even if in some games they lead to completely different choices. The same actually holds for the two concepts we discuss in Chapters 8 and 9.

As a consequence, we do *not* believe in a *unique* concept for game theory. In my opinion we must accept that different people tend to reason differently in the same game-theoretic situation, and to me there is simply no best way of reasoning in a game – only *different* ways of reasoning.

Rational and reasonable choices. The word *rational* has often led to confusion in game theory. What do we mean precisely when we say that a player chooses *rationally*? In this book, we say that a player chooses *rationally* if some belief is formed about the opponents' choices, and then the player makes a choice that is optimal under this belief.

However, the belief held about the opponents may be completely unreasonable, and hence – within our terminology – a rational choice is not necessarily a *reasonable* choice. Of course, the meaning of a *reasonable* choice is very subjective, as it depends on the way of reasoning one has in mind. As we have already argued above, there are intuitive concepts in this book that in some games lead to completely opposed choices. Hence, what is a "reasonable" choice under one concept may be "unreasonable" under another concept. Since we do not believe in a unique concept for game theory, we also do not believe in a unique definition of what is a reasonable choice in a game. To a large extent, it is up to the reader to decide. The book is only there to help the reader make this choice.

Examples and problems. The book is full of examples and practical problems, and in my opinion they constitute the most important ingredient of this book. Each of the examples and practical problems is based on a story – mostly inspired by everyday life situations – in which *you* (the reader) are the main character. The reason we choose scenarios from everyday life is that it makes it easier for the reader to identify with such situations. We could also have chosen scenarios from professional decision making, such as business, politics, economics or managerial decision making, but such situations would probably be more distant for some readers.

All of the stories in the examples and practical problems take place in an imaginary world in which you experience some adventures together with your imaginary friends Barbara and Chris. These stories often have a humorous side, and thereby also serve as points of relaxation in the book. The one-person perspective we discussed above is very strongly present in these examples and practical problems, as *you* always play the main role in these stories, and the situation is always analyzed completely from *your* perspective. The crucial question is always: "Which choice would *you* make in this situation, and why?"

The examples are there to illustrate the main ideas behind the concepts, and to show how the various concepts and algorithms can be applied to concrete situations. As every example is based on a particular story, it will make it easier for the reader to remember the various examples, and to keep these examples in the back of the mind as a benchmark.

The theoretical problems in the book are of a completely different type compared to the practical problems. They do not refer to any story, but rather discuss some general theoretical issues related to the concept of that chapter. Usually, these problems require the reader to formally prove some statement. These theoretical problems are primarily meant for Ph.D. students and researchers who wish to deepen their theoretical insights.

Beliefs diagrams. In Chapters 2, 3 and 4 of this book we use a *beliefs diagram* to graphically represent the belief hierarchy of a player – that is, the belief about the opponents' choices, the belief about the opponents' beliefs about the other players' choices, and so on. We invented this beliefs diagram because a belief hierarchy may seem rather abstract and complicated when stated formally – certainly for readers that are new to epistemic game theory. However, belief hierarchies are crucial for

developing the various concepts in this book. By visualizing the various levels of a belief hierarchy – by means of a *beliefs diagram* – we hope the reader will find it easier to work with such belief hierarchies, and to understand what a belief hierarchy really represents. Moreover, the beliefs diagrams also play a crucial role in the examples and practical problems of Chapters 2, 3 and 4.

Choices and beliefs. In this book we always make a very clear distinction between the *choices* and the *beliefs* of a player. The reason we raise this issue here is that in some books and articles this distinction is not very clear. Some books, when they introduce the concept of Nash equilibrium for instance, state that the *mixed strategy* of a player can either be interpreted as the actual choice by this player, or as the belief his opponents hold about the player's choice. But what is then the real interpretation of a mixed strategy? This often remains unclear. But if the meaning of a mixed strategy remains ambiguous, it is likely to cause confusion, which is of course very undesirable and unnecessary. Such confusion can easily be avoided by always being clear about the interpretation of the various objects that are being introduced. In particular, we believe we must always make a very clear distinction between the choices and the beliefs of a player, as these are completely different objects. And that is precisely what we do in this book.

Randomized choices. The concept of a *mixed strategy* – or *randomized choice* – is still used as an object of choice in many books and articles in game theory. Strictly speaking, a *randomized choice* means that a player, before making a choice, uses a randomization device and bases the actual choice on the outcome of the randomization device. For instance, the player rolls a dice, and chooses based on the outcome of the dice roll. We believe, however, that decision makers do not randomize when making serious decisions! The reason is that there is nothing to gain for a player by randomizing. Namely, randomizing between two choices a and b can only be optimal for the player if there is the belief that a and b yield the same maximal utility. But in that case, the player could just as well choose a or b – without randomizing – and save the trouble of having to roll a dice.

In this book we take seriously the fact that people typically do not randomize when making choices. Throughout the book, we assume that players do not use randomized choices, but always go for one particular choice (with probability one). Randomized choices are only used in this book as *artificial auxiliary* objects, used to characterize *rational* and *irrational choices* – see Theorem 2.5.3 in Chapter 2. Within that theorem, randomized choices are not interpreted as real objects of choice, but are rather used as an abstract mathematical tool to verify whether a given choice is rational or not.

Using the book for a course. This book is well suited for a course in game theory at any university. Moreover, it can be used at different levels – for advanced bachelor students, master students and Ph.D. students. Depending on the type, the level and the length of the game theory course, one can use all chapters or selected chapters from this book – of course respecting the arrows in Figure 1.1. As we mentioned above, Chapters 2 and 3 present the central idea in epistemic game theory – *common belief*

in rationality – and these two chapters should thus be part of any course based on this book. Chapter 2 is a preparatory chapter, meant to make the reader familiar with the main ingredients of a game, and which informally introduces the first reasoning steps that will eventually lead to the concept of *common belief in rationality*. Chapter 3 shows how these reasoning steps can be modeled formally, and how these first few steps can be completed to arrive at the concept of *common belief in rationality*. When designing a course in game theory one could decide to discuss Chapter 2 only briefly, or to merge it with Chapter 3, depending on the type of course. For instance, for a trimester course of seven weeks, one could teach Chapters 2 and 3 in week 1, and dedicate every subsequent week to one of the remaining six chapters.

Irrespective of the level of the course, we think that the examples should play a prominent role in class. From my own experience I know that these examples are a powerful tool for revealing the main ideas behind the concepts. Moreover, the examples are likely to stick in the students' heads as they are all based on some particular story. If time allows, we would also strongly advise dedicating at least one session per week to discussing some of the practical problems at the end of each chapter. By working on these problems, the student will be trained in applying the concepts and algorithms to concrete situations, and we believe this is the best way for a student to master the various concepts. Besides, we really hope – and do believe – that the practical problems are a lot of fun. I, at least, had a lot of fun inventing them.

As an example of how to use the book for a course, we will briefly outline the course in epistemic game theory that Christian Bach and I will give at Maastricht University during the academic year 2011/2012. The course is designed for master students and Ph.D. students, and lasts seven weeks. Every week there are two theory lectures of two hours, and one problem session of two hours. In week 1 we cover Chapters 2 and 3 of the book, whereas every subsequent week covers one of the remaining Chapters 4–9. Every week the students must work on three practical problems from the associated chapter, and the solutions to these problems are discussed during the problem session of that week. But this is just an example – every teacher can design the course depending on the time available and the teaching method used.

Limitations of this book. As with any book, this book has its limitations. First, we only consider *finite* games, which means that for a static game, every player only has finitely many possible choices, whereas in the context of dynamic games we additionally assume that the game will stop after finitely many moves by the players. By doing so, we exclude dynamic games of possible infinite duration such as infinitely repeated games, evolutionary games or stochastic games. But the concepts we develop in this book should be applicable to such infinite games.

We also assume throughout the book that the players' utility functions are completely transparent to all the players involved. That is, players do not have any uncertainty about their opponents' utility functions. We thereby exclude games with *incomplete information*, in which some or all of the players do not completely know their opponents' utility functions. Some of the concepts in this book have been extended to games with incomplete information, but we do not discuss these extensions here.

Recently, game theorists have started to study situations in which some of the players are *unaware* of certain elements in the game. For instance, you may be unaware of some of the opponent's choices. In this book we do not study such situations, but this is certainly a very interesting line of research. In particular, it would be interesting to see whether – and if so how – the various concepts in this book can be extended to such games with unawareness.

Finally, we have taken the concept of *common belief in rationality* as the basis for this book. Every other concept in this book can be seen as a *refinement* of common belief in rationality, since they are obtained by taking the concept of common belief in rationality and imposing additional restrictions. There are also intuitive concepts in the literature that *violate* some of the conditions in common belief in rationality – especially within the bounded rationality literature – but such concepts are not discussed in this book.

Surveys on epistemic game theory. As we mentioned above, this is the first ever textbook on epistemic game theory. In the literature there are, however, survey papers that give an overview of some of the most important ideas in epistemic game theory. The interested reader may consult the surveys by Brandenburger (1992a), Geanakoplos (1992), Dekel and Gul (1997), Battigalli and Bonanno (1999), Board (2002) and Brandenburger (2007).