INTRODUCTION

What is game theory?

Game theory is the name given to the methodology of using mathematical tools to model and analyze situations of interactive decision making. These are situations involving several decision makers (called players) with different goals, in which the decision of each affects the outcome for all the decision makers. This interactivity distinguishes game theory from standard decision theory, which involves a single decision maker, and it is its main focus. Game theory tries to predict the behavior of the players and sometimes also provides decision makers with suggestions regarding ways in which they can achieve their goals.

The foundations of game theory were laid down in the book *The Theory of Games and Economic Behavior*, published in 1944 by the mathematician John von Neumann and the economist Oskar Morgenstern. The theory has been developed extensively since then and today it has applications in a wide range of fields. The applicability of game theory is due to the fact that it is a context-free mathematical toolbox that can be used in any situation of interactive decision making. A partial list of fields where the theory is applied, along with examples of some questions that are studied within each field using game theory, includes:

- **Theoretical economics.** A market in which vendors sell items to buyers is an example of a game. Each vendor sets the price of the items that he or she wishes to sell, and each buyer decides from which vendor he or she will buy items and in what quantities. In models of markets, game theory attempts to predict the prices that will be set for the items along with the demand for each item, and to study the relationships between prices and demand. Another example of a game is an auction. Each participant in an auction determines the price that he or she will bid, with the item being sold to the highest bidder. In models of auctions, game theory is used to predict the bids submitted by the participants, the expected revenue of the seller, and how the expected revenue will change if a different auction method is used.

- **Networks.** The contemporary world is full of networks; the Internet and mobile telephone networks are two prominent examples. Each network user wishes to obtain the best possible service (for example, to send and receive the maximal amount of information in the shortest span of time over the Internet, or to conduct the highest-quality calls using a mobile telephone) at the lowest possible cost. A user has to choose an Internet service provider or a mobile telephone provider, where those providers are also players in the game, since they set the prices of the service they provide. Game theory tries to predict the behavior of all the participants in these markets. This game is more complicated from the perspective of the service providers than from the perspective of the buyers, because the service providers can cooperate with each other
Introduction

(for example, mobile telephone providers can use each other’s network infrastructure to carry communications in order to reduce costs), and game theory is used to predict which cooperative coalitions will be formed and suggest ways to determine a “fair” division of the profit of such cooperation among the participants.

• **Political science.** Political parties forming a governing coalition after parliamentary elections are playing a game whose outcome is the formation of a coalition that includes some of the parties. This coalition then divides government ministries and other elected offices, such as parliamentary speaker and committee chairmanships, among the members of the coalition. Game theory has developed indices measuring the power of each political party. These indices can predict or explain the division of government ministries and other elected offices given the results of the elections. Another branch of game theory suggests various voting methods and studies their properties.

• **Military applications.** A classical military application of game theory models a missile pursuing a fighter plane. What is the best missile pursuit strategy? What is the best strategy that the pilot of the plane can use to avoid being struck by the missile? Game theory has contributed to the field of defense the insight that the study of such situations requires strategic thinking: when coming to decide what you should do, put yourself in the place of your rival and think about what he/she would do and why, while taking into account that he/she is doing the same and knows that you are thinking strategically and that you are putting yourself in his/her place.

• **Inspection.** A broad family of problems from different fields can be described as two-player games in which one player is an entity that can profit by breaking the law and the other player is an “inspector” who monitors the behavior of the first player. One example of such a game is the activities of the International Atomic Energy Agency, in its role of enforcing the Treaty on the Non-Proliferation of Nuclear Weapons by inspecting the nuclear facilities of signatory countries. Additional examples include the enforcement of laws prohibiting drug smuggling, auditing of tax declarations by the tax authorities, and ticket inspections on public trains and buses.

• **Biology.** Plants and animals also play games. Evolution “determines” strategies that flowers use to attract insects for pollination and it “determines” strategies that the insects use to choose which flowers they will visit. Darwin’s principle of the “survival of the fittest” states that only those organisms with the inherited properties that are best adapted to the environmental conditions in which they are located will survive. This principle can be explained by the notion of *Evolutionarily Stable Strategy*, which is a variant of the notion of *Nash equilibrium*, the most prominent game-theoretic concept. The introduction of game theory to biology in general and to evolutionary biology in particular explains, sometimes surprisingly well, various biological phenomena.

Game theory has applications to other fields as well. For example, to philosophy it contributes some insights into concepts related to morality and social justice, and it raises questions regarding human behavior in various situations that are of interest to psychology. Methodologically, game theory is intimately tied to mathematics: the study of game-theoretic models makes use of a variety of mathematical tools, from probability and
Introduction

combinatorics to differential equations and algebraic topology. Analyzing game-theoretic models sometimes requires developing new mathematical tools.

Traditionally, game theory is divided into two major subfields: strategic games, also called noncooperative games, and coalitional games, also called cooperative games. Broadly speaking, in strategic games the players act independently of each other, with each player trying to obtain the most desirable outcome given his or her preferences, while in coalitional games the same holds true with the stipulation that the players can agree on and sign binding contracts that enforce coordinated actions. Mechanisms enforcing such contracts include law courts and behavioral norms. Game theory does not deal with the quality or justification of these enforcement mechanisms; the cooperative game model simply assumes that such mechanisms exist and studies their consequences for the outcomes of the game.

The categories of strategic games and coalitional games are not well defined. In many cases interactive decision problems include aspects of both coalitional games and strategic games, and a complete theory of games should contain an amalgam of the elements of both types of models. Nevertheless, in a clear and focused introductory presentation of the main ideas of game theory it is convenient to stick to the traditional categorization. We will therefore present each of the two models, strategic games and coalitional games, separately. Chapters 1–15 are devoted to strategic games, and Chapters 16–21 are devoted to coalitional games. Chapters 22 and 23 are devoted to social choice and stable matching, which include aspects of both noncooperative and cooperative games.

How to use this book

The main objective of this book is to serve as an introductory textbook for the study of game theory at both the undergraduate and the graduate levels. A secondary goal is to serve as a reference book for students and scholars who are interested in an acquaintance with some basic or advanced topics of game theory. The number of introductory topics is large and different teachers may choose to teach different topics in introductory courses. We have therefore composed the book as a collection of chapters that are, to a large extent, independent of each other, enabling teachers to use any combination of the chapters as the basis for a course tailored to their individual taste. To help teachers plan a course, we have included an abstract at the beginning of each chapter that presents its content in a short and concise manner.

Each chapter begins with the basic concepts and eventually goes farther than what may be termed the “necessary minimum” in the subject that it covers. Most chapters include, in addition to introductory concepts, material that is appropriate for advanced courses. This gives teachers the option of teaching only the necessary minimum, presenting deeper material, or asking students to complement classroom lectures with independent readings or guided seminar presentations. We could not, of course, include all known results of game theory in one textbook, and therefore the end of each chapter contains references to other books and journal articles in which the interested reader can find more material for a deeper understanding of the subject. Each chapter also contains exercises, many of which are relatively easy, while some are more advanced and challenging.
Introduction

This book was composed by mathematicians; the writing is therefore mathematically oriented, and every theorem in the book is presented with a proof. Nevertheless, an effort has been made to make the material clear and transparent, and every concept is illustrated with examples intended to impart as much intuition and motivation as possible. The book is appropriate for teaching undergraduate and graduate students in mathematics, computer science and exact sciences, economics and social sciences, engineering, and life sciences. It can be used as a textbook for teaching different courses in game theory, depending on the level of the students, the time available to the teacher, and the specific subject of the course. For example, it could be used in introductory level or advanced level semester courses on coalitional games, strategic games, a general course in game theory, or a course on applications of game theory. It could also be used for advanced mini-courses on, e.g., incomplete information (Chapters 9, 10, and 11), auctions (Chapter 12), or repeated games (Chapters 13 and 14). As mentioned previously, the material in the chapters of the book will in many cases encompass more than a teacher would choose to teach in a single course. This requires teachers to choose carefully which chapters to teach and which parts to cover in each chapter. For example, the material on strategic games (Chapters 4 and 5) can be taught without covering extensive-form games (Chapter 3) or utility theory (Chapter 2). Similarly, the material on games with incomplete information (Chapter 9) can be taught without teaching the other two chapters on models of incomplete information (Chapters 10 and 11).

For the sake of completeness, we have included an appendix containing the proofs of some theorems used throughout the book, including Brouwer’s Fixed Point Theorem, Kakutani’s Fixed Point Theorem, the Knaster–Kuratowski–Mazurkiewicz (KKM) Theorem, and the Separating Hyperplane Theorem. The appendix also contains a brief survey of linear programming. A teacher can choose to prove each of these theorems in class, assign the proofs of the theorems as independent reading to the students, or state any of the theorems without proof based on the assumption that students will see the proofs in other courses.
Chapter summary
In the opening chapter of this book, we use the well-known game of chess to illustrate the notions of *strategy* and *winning strategy*. We then prove one of the first results in game theory, due to John von Neumann: in the game of chess either White (the first mover) has a winning strategy, or Black (the second mover) has a winning strategy, or each player has a strategy guaranteeing at least a draw. This is an important and nontrivial result, especially in view of the fact that to date, it is not known which of the above three alternatives holds, let alone what the winning strategy is, if one exists.

In later chapters of the book, this result takes a more general form and is applied to a large class of games.

We begin with an exposition of the elementary ideas in noncooperative game theory, by analyzing the game of chess. Although the theory that we will develop in this chapter relates to that specific game, in later chapters it will be developed to apply to much more general situations.

1.1 Schematic description of the game

The game of chess is played by two players, traditionally referred to as White and Black. At the start of a match, each player has sixteen pieces arranged on the chessboard. White is granted the opening move, following which each player in turn moves pieces on the board, according to a set of fixed rules. A match has three possible outcomes:

- Victory for White, if White captures the Black King.
- Victory for Black, if Black captures the White King.
- A draw, if:
  1. it is Black’s turn, but he has no possible legal moves available, and his King is not in check;
  2. it is White’s turn, but he has no possible legal moves available, and his King is not in check;
  3. both players agree to declare a draw;
  4. a board position precludes victory for both sides;
  5. 50 consecutive turns have been played without a pawn having been moved and without the capture of any piece on the board, and the player whose turn it is requests that a draw be declared;
The game of chess

6. or if the same board position appears three times, and the player whose turn it is
requests that a draw be declared.

Analysis and results

For the purposes of our analysis all we need to assume is that the game is finite, i.e.,
the number of possible turns is bounded (even if that bound is an astronomically large
number). This does not apply, strictly speaking, to the game of chess, but since our
lifetimes are finite, we can safely assume that every chess match is finite.

We will denote the set of all possible board positions in chess by \( X \). A board position
by definition includes the identity of each piece on the board, and the board square on
which it is located.

A board position, however, does not provide full details on the sequence of moves
that led to it: there may well be two or more sequences of moves leading to the same
board position. We therefore need to distinguish between a “board position” and a “game
situation,” which is defined as follows.

**Definition 1.1** A game situation (in the game of chess) is a finite sequence
\((x_0, x_1, x_2, \ldots, x_K)\) of board positions in \( X \) satisfying

1. \( x_0 \) is the opening board position.
2. For each even integer \( k \), \( 0 \leq k < K \), going from board position
   \( x_k \) to \( x_{k+1} \) can be
   accomplished by a single legal move on the part of White.
3. For each odd integer \( k \), \( 0 \leq k < K \), going from board position
   \( x_k \) to \( x_{k+1} \) can be
   accomplished by a single legal move on the part of Black.

We will denote the set of game situations by \( H \).

Suppose that a player wishes to program a computer to play chess. The computer would
need a plan of action that would tell it what to do in any given game situation that could
arise. A full plan of action for behavior in a game is called a strategy.

**Definition 1.2** A strategy for White is a function \( s_W \) that associates every game situation
\((x_0, x_1, x_2, \ldots, x_K) \in H \), where \( K \) is even, with a board position \( x_{K+1} \), such that going
from board position \( x_K \) to \( x_{K+1} \) can be accomplished by a single legal move on the part
of White.

Analogously, a strategy for Black is a function \( s_B \) that associates every game situation
\((x_0, x_1, x_2, \ldots, x_K) \in H \), where \( K \) is odd, with a board position \( x_{K+1} \) such that going from
board position \( x_K \) to \( x_{K+1} \) can be accomplished by a single legal move on the part
of Black.

Any pair of strategies \((s_W, s_B)\) determines an entire course of moves, as follows. In
the opening move, White plays the move that leads to board position \( x_1 = s_W(x_0) \).
Black then plays the move leading to board position \( x_2 = s_B(x_0, x_1) \), and so on.
The succeeding board positions are determined by \( x_{2K+1} = s_W(x_0, x_1, \ldots, x_{2K}) \) and
\( x_{2K+2} = s_B(x_0, x_1, \ldots, x_{2K+1}) \) for all \( K = 0, 1, 2, \ldots \).
1.2 Analysis and results

An entire course of moves (from the opening move to the closing one) is termed a play of the game.

Every play of the game of chess ends in either a victory for White, a victory for Black, or a draw. A strategy for White is termed a winning strategy if it guarantees that White will win, no matter what strategy Black chooses.

Definition 1.3 A strategy \( s_W \) is a winning strategy for White if for every strategy \( s_B \) of Black, the play of the game determined by the pair \( (s_W, s_B) \) ends in victory for White. A strategy \( s_W \) is a strategy guaranteeing at least a draw for White if for every strategy \( s_B \) of Black, the play of the game determined by the pair \( (s_W, s_B) \) ends in either a victory for White or a draw.

If \( s_W \) is a winning strategy for White, then any White player (or even computer program) adopting that strategy is guaranteed to win, even if he faces the world’s chess champion.

The concepts of “winning strategy” and “strategy guaranteeing at least a draw” for Black are defined analogously, in an obvious manner.

The next theorem follows from one of the earliest theorems ever published in game theory (see Theorem 3.13 on page 46).

Theorem 1.4 In chess, one and only one of the following must be true:

(i) White has a winning strategy.
(ii) Black has a winning strategy.
(iii) Each of the two players has a strategy guaranteeing at least a draw.

We emphasize that the theorem does not relate to a particular chess match, but to all chess matches. That is, suppose that alternative (i) is the true case, i.e., White has a winning strategy \( s_W \). Then any person who is the White player and follows the prescriptions of that strategy will always win every chess match he ever plays, no matter who the opponent is. If, however, alternative (ii) is the true case, then Black has a winning strategy \( s_B \) and any person who is the Black player and follows the prescriptions of that strategy will always win every chess match he ever plays, no matter who the opponent is. Finally, if alternative (iii) is the true case, then White has a strategy \( s_W \) guaranteeing at least a draw, and Black has a strategy \( s_B \) guaranteeing at least a draw. Any person who is the White player (or the Black player) and follows the prescriptions of \( s_W \) (or \( s_B \), respectively) will always get at least a draw in every chess match he ever plays, no matter who the opponent is. Note that if alternative (i) holds, there may be more than one winning strategy, and similar statements can be made with regard to the other two alternatives.

So, given that one of the three alternatives must be true, which one is it? We do not know. If the day ever dawns in which a winning strategy for one of the players is discovered, or strategies guaranteeing at least a draw for each player are discovered, the game of chess will cease to be of interest. In the meantime, we can continue to enjoy the challenge of playing (or watching) a good chess match.

Despite the fact that we do not know which alternative is the true one, the theorem is significant, because a priori it might have been the case that none of the alternatives...
The game of chess

White moves

Black moves

White moves

Figure 1.1 The game of chess presented in extensive form

was possible; one could have postulated that no player could ever have a strategy _always_ guaranteeing a victory, or at least a draw.

We present two proofs of the theorem. The first proof is the “classic” proof, which in principle shows how to find a winning strategy for one of the players (if such a strategy exists) or a strategy guaranteeing at least a draw (if such a strategy exists). The second proof is shorter, but it cannot be used to find a winning strategy for one of the players (if such a strategy exists) or a strategy guaranteeing at least a draw (if such a strategy exists).

We start with several definitions that are needed for the first proof of the theorem. The set of game situations can be depicted by a tree\(^1\) (see Figure 1.1). Such a tree is called a _game tree_. Each vertex of the game tree represents a possible game situation. Denote the set of vertices of the game tree by \(H\).

The _root vertex_ is the opening game situation \(x_0\), and for each vertex \(x\), the set of _children vertices_ of \(x\) is the set of game situations that can be reached from \(x\) in one legal move. For example, in his opening move, White can move one of his pawns one or two squares forward, or one of his two knights. So White has 20 possible opening moves, which means that the root vertex of the tree has 20 children vertices. Every vertex that can be reached from \(x\) by a sequence of moves is called a _descendant_ of \(x\). Every _leaf_ of the tree corresponds to a terminal game situation, in which either White has won, Black has won, or a draw has been declared.

Given a vertex \(x \in H\), we may consider the subtree beginning at \(x\), which is by definition the tree whose root is \(x\) that is obtained by removing all vertices that are not descendants of \(x\). This subtree of the game tree, which we will denote by \(\Gamma(x)\),

---

\(^1\) The mathematical definition of a _tree_ appears in the sequel (see Definition 3.5 on page 42).
1.2 Analysis and results

corresponds to a game that is called the subgame beginning at \( x \). We will denote by \( n_x \) the number of vertices in \( \Gamma(x) \). The game \( \Gamma(x_0) \) is by definition the game that starts with the opening situation of the game, and is therefore the standard chess game.

If \( y \) is a child vertex of \( x \), then \( \Gamma(y) \) is a subtree of \( \Gamma(x) \) that does not contain \( x \). In particular, \( n_x > n_y \). Moreover, \( n_x = 1 \) if and only if \( x \) is a terminal situation of the game, i.e., the players cannot implement any moves at this subgame. In such a case, the strategy of a player is denoted by \( \emptyset \).

Denote by
\[
\mathcal{F} = \{ \Gamma(x) : x \in H \}
\]
the collection of all subgames that are defined by subtrees of the game of chess.

Theorem 1.4 can be proved using the result of Theorem 1.5.

**Theorem 1.5** Every game in \( \mathcal{F} \) satisfies one and only one of the following alternatives:

(i) White has a winning strategy.

(ii) Black has a winning strategy.

(iii) Each of the players has a strategy guaranteeing at least a draw.

**Proof:** The proof proceeds by induction on \( n_x \), the number of vertices in the subgame \( \Gamma(x) \).

Suppose \( x \) is a vertex such that \( n_x = 1 \). As noted above, that means that \( x \) is a terminal vertex. If the White King has been removed from the board, Black has won, in which case \( \emptyset \) is a winning strategy for Black. If the Black King has been removed from the board, White has won, in which case \( \emptyset \) is a winning strategy for White. Alternatively, if both Kings are on the board at the end of play, the game has ended in a draw, in which case \( \emptyset \) is a strategy guaranteeing a draw for both Black and White.

Next, suppose that \( x \) is a vertex such that \( n_x > 1 \). Assume by induction that at all vertices \( y \) satisfying \( n_y < n_x \), one and only one of the alternatives (i), (ii), or (iii) is true in the subgame \( \Gamma(y) \).

Suppose, without loss of generality, that White has the first move in \( \Gamma(x) \). Any board position \( y \) that can be reached from \( x \) satisfies \( n_y < n_x \), and so the inductive assumption is true in the corresponding subgame \( \Gamma(y) \). Denote by \( C(x) \) the collection of vertices that can be reached from \( x \) in one of White’s moves.

1. If there is a vertex \( y_0 \in C(x) \) such that White has a winning strategy in \( \Gamma(y_0) \), then alternative (i) is true in \( \Gamma(x) \): the winning strategy for White in \( \Gamma(x) \) is to choose as his first move the move leading to vertex \( y_0 \), and to follow the winning strategy in \( \Gamma(y_0) \) at all subsequent moves.

2. If Black has a winning strategy in \( \Gamma(y) \) for every vertex \( y \in C(x) \), then alternative (ii) is true in \( \Gamma(x) \): Black can win by ascertaining what the vertex \( y \) is after White’s first move, and following his winning strategy in \( \Gamma(y) \) at all subsequent moves.

3. Otherwise:

   • (1) does not hold, i.e., White has no winning strategy in \( \Gamma(y) \) for any \( y \in C(x) \).

   Because the induction hypothesis holds for every vertex \( y \in C(x) \), either Black has
The game of chess

a winning strategy in \( \Gamma(y) \), or both players have a strategy guaranteeing at least a
draw in \( \Gamma(y) \).

- (2) does not hold, i.e., there is a vertex \( y_0 \in C(x) \) such that Black does not have a
winning strategy in \( \Gamma(y_0) \). But because (1) does not hold, White also does not have a
winning strategy in \( \Gamma(y_0) \). Therefore, by the induction hypothesis applied to \( \Gamma(y_0) \),
both players have a strategy guaranteeing at least a draw in \( \Gamma(y_0) \).

As we now show, in this case, in \( \Gamma(x) \) both players have a strategy guaranteeing at least a
draw. White can guarantee at least a draw by choosing a move leading to vertex \( y_0 \),
and from there by following the strategy that guarantees at least a draw in \( \Gamma(y_0) \). Black
can guarantee at least a draw by ascertaining what the board position \( y \) is after White’s
first move, and at all subsequent moves in \( \Gamma(y) \) either by following a winning strategy or
following a strategy that guarantees at least a draw in that subgame.

The proof just presented is a standard inductive proof over a tree: one assumes that the
theorem is true for every subtree starting from the root vertex, and then shows that it is true for
the entire tree. The proof can also be accomplished in the following way: select any vertex \( x \) that is neither a terminal vertex nor the root vertex. The subgame starting from
this vertex, \( \Gamma(x) \), contains at least two vertices, but fewer vertices than the original game
(because it does not include the root vertex), and the induction hypothesis can therefore be
applied to \( \Gamma(x) \). Now “fold up” the subgame and replace it with a terminal vertex whose
outcome is the outcome that is guaranteed by the induction hypothesis to be obtained in
\( \Gamma(x) \). This leads to a new game \( \Gamma' \). Since \( \Gamma(x) \) has at least two vertices, \( \Gamma' \) has fewer
vertices than the original game, and therefore by the induction hypothesis the theorem is
true for \( \Gamma' \). It is straightforward to ascertain that a player has a winning strategy in \( \Gamma' \) if
and only if he has a winning strategy in the original game.

In the proof of Theorem 1.5 we used the following properties of the game of chess:

(C1) The game is finite.
(C2) The strategies of the players determine the play of the game. In other words, there is
no element of chance in the game; neither dice nor card draws are involved.
(C3) Each player, at each turn, knows the moves that were made at all previous stages of
the game.

We will later see examples of games in which at least one of the above properties fails to
hold, for which the statement of Theorem 1.5 also fails to hold (see for example the game
“Matching Pennies,” Example 3.20 on page 52).

We next present a second proof of Theorem 1.4. We will need the following two facts
from formal logic for the proof. Let \( X \) be a finite set and let \( A(x) \) be an arbitrary logical
formula.\(^2\)

- If it is not the case that “for every \( x \in X \) the formula \( A(x) \) holds,” then there exists an
\( x \in X \) where the formula \( A(x) \) does not hold:

\[
\neg (\forall x(A)) \equiv \exists x(\neg A).
\] (1.2)

---

\(^2\) Recall that the logical statement “for every \( x \in X \) event \( A \) obtains” is written formally as \( \forall x(A) \), and the statement
“there exists an \( x \in X \) for which event \( A \) obtains” is written as \( \exists x(A) \), while “event \( A \) does not obtain” is written as
\( \neg A \). For ease of exposition, we will omit the set \( X \) from each of the formal statements in the proof.