FOUNDATIONS OF NON-COOPERATIVE GAMES

Ron Holzman
Technion-Israel Institute of Technology, Israel

Keywords: Bayesian game, Bayesian Nash equilibrium, chess-like game, consistency, determinacy, extensive form, information, Nash equilibrium, non-cooperative game, non-zero-sum game, payoff function, rationality, strategic form, strategy, subgame perfect equilibrium, value, zero-sum game.

Contents

Summary

This article introduces the basic models and ideas of the theory of non-cooperative games. We begin by treating games in the usual sense of the word, such as chess. We show that for a certain class of games, the outcome is completely determined if the players play optimally. Then we indicate how the descriptive framework of game theory, including the extensive and strategic form representations, can serve to model interactions between agents which do not qualify as games in the usual sense. For zero-sum games, where one player's gain is the other's loss, we introduce the concept of value, which is the expected outcome when the game is played optimally. If players are allowed to use mixed (i.e., randomized) strategies, the minimax theorem asserts that the value exists. Non-zero-sum games are more complex, and we cannot hope to pinpoint their expected outcome as in the zero-sum case. The central concept for these games is

that of a Nash equilibrium, which is a choice of strategies for the players having the property that every player does best by playing his strategy if the others do the same. Nash's theorem guarantees the existence of a Nash equilibrium in mixed strategies. Finally, we turn to the modeling of incomplete information, which occurs when the players lack information about the game they are facing. We present the concepts of a "state of the world" and the "type" of a player, and show how they are incorporated in the Bayesian game model.

1. Introduction

Non-cooperative game theory studies situations in which a number of agents are involved in an interactive process. The outcome of this process is determined by the agents' individual decisions (sometimes in conjunction with chance) and affects the well being of each agent in a possibly different way. The most obvious examples of such situations are parlor games. The terminology we use has its roots in this area: the entire situation is called a game, the agents are called players, their acts are called moves, their overall plans of action are called strategies, and their evaluations of the outcome are called payoffs. However, the range of situations that we have in mind is much wider, and includes interactions in areas such as economics, politics, biology, computing, etc. Thus, the significance of non-cooperative game theory to the understanding of social and natural phenomena is far bigger than its name may suggest.

The basic premise of our analysis is that players act rationally, meaning first and foremost, that they strive to maximize their payoffs. However, since their payoffs are affected not only by their own decisions but also by the other players' decisions, they must reason about the other players' reasoning. In doing so they take into account that the other players also act rationally.

The qualification "non-cooperative" refers to the assumption that players make their decisions individually, and are not allowed to forge binding agreements with other players that stipulate the actions to be taken by the parties to the agreement. The players may be allowed to communicate with each other prior to the play of the game and discuss joint plans of action. However, during the play of the game they act as autonomous decision makers, and as such they will follow previously made joint plans only if doing so is rational for them.

The theory of non-cooperative games comprises three main ingredients. The first of these is the development of formal models of non-cooperative games that create unified frameworks for representing games in a manner that lends itself to formal mathematical analysis. The second ingredient is the formulation of concepts that capture the idea of rational behavior in those models. The main such concept is that of equilibrium. The third ingredient is the use of mathematical tools in order to prove meaningful statements about those concepts, such as existence and characterizations of equilibrium.

In any concrete application of the theory, the first step is to represent the situation at hand by one of the available formal models. Because real-life situations are typically very complex and not totally structured, it is often impossible and/or unhelpful to incorporate all the elements of the situation in the formal model. Therefore, this step requires judicious decisions identifying those important features that must be modeled. Once the model is constructed, its analysis is carried out based on the appropriate concept of equilibrium, drawing on general results of non-cooperative game theory or, as the case may be, exploiting attributes of the specific application. This analysis yields conclusions, which may then be reformulated in terms of the real-life situation, providing insights into, or predictions about, the behavior of the agents and the long-term steady states of the system being investigated.

Another type of application is sometimes called game theoretic engineering. It involves making recommendations to organizations on how to set up the "rules of the game" so that the rational behavior of the agents will lead to results that are desirable from the organization's point of view. Examples include the revision of electoral systems, the design of auctions, the creation of markets for emission permits as a means to efficiently control pollution, etc. (see *Mechanism Theory*). This sort of application seems to be gaining more and more recognition lately.

This article is organized according to several standard criteria for classifying non-cooperative games: how the payoffs of the players are related, the presence or the absence of chance, the nature of the information that the players have.

We do not attempt to present here a comprehensive survey of non-cooperative game theory. The omission of certain parts of the theory, even important ones, is unavoidable in such an article. Some of these areas are covered in other articles within this topic. What we try to do here is offer a gentle introduction to a small number of basic models, ideas and concepts.

2. Chess-Like Games

2.1 The Description of the Game

The game of chess is a prime example of a class of games that we call *chess-like games*. We first give a verbal description of what we mean by a chess-like game.

In such a game, two players take turns making their moves. One of the players is designated as the one who starts, we call this player White and the other player Black. Whenever a player chooses a move, he is perfectly informed of all moves made prior to that stage. The play of the game is fully determined by the players' choices, that is, it does not involve any chance. For every initial sequence of moves made alternatingly by the two players, the rules of the game determine whether the player whose turn it is to play should choose a move, in which case they also determine what his legal moves are, or whether the play has ended, in which case they also determine the outcome: a win for White, a win for Black, or a draw. It may be the case that the rules allow for a play consisting of an infinite sequence of moves, but such infinite plays must also be classified as resulting in one of the three possible outcomes mentioned above.

Examples of chess-like games are chess, checkers, and tic-tac-toe. Note that Kriegspiel (a version of chess in which a player does not observe his opponent's moves) is not a chess-like game, due to the lack of perfect information. Backgammon is not a chess-like game because it involves dice.

We formally describe a chess-like game by means of a rooted tree (T, r), called the *game-tree*. Here T is a tree (that is, a connected acyclic graph which may be finite or infinite), and r, the root, is a designated node of T. We think of T as being oriented away from r. The edges that fan out from r correspond to the legal moves of White at his first turn. Each of these edges leads to a new node, and the edges that fan out from this new node correspond to the legal moves of Black after White chose the specified edge from r as his first move. It goes on like this: non-terminal nodes whose distance from r is even (respectively odd) are decision nodes of White (respectively Black), and the edges that fan out correspond to the legal moves of the respective player given the moves made so far. A maximal branch of (T, r) is a path that starts at the root r and either ends at a terminal node or is infinite. Every maximal branch corresponds to a play of the game. (Note the distinction we make between the usage of "game" and "play". A game is the totality of rules that define it, whereas a play of the game is a complete account of what happens a particular time when the game is played.) To complete the formal description of the game, we specify for each maximal branch whether it is a win for White, a win for Black, or a draw.

The realization that every chess-like game can in principle be represented by a game-tree as above, even though for most games constructing the actual game-tree is impractical, is an important conceptual step towards the analysis of such games.

2.2 The Determinacy of Chess-Like Games

The next important concept is that of a *strategy*. By a strategy we mean a complete set of instructions that tell a player what to do in every situation that may arise in the play of the game in which he is called upon to make a move. In terms of the formal description of a chess-like game by a game-tree, a strategy of a player is a function σ from the set of his decision nodes into the set of edges, such that $\sigma(x)$ is one of the edges that fan out of x.

It is clear from the above that a pair of strategies, σ for White and τ for Black, fully determines a play of the game. In particular any such pair (σ, τ) results in a win for White, a win for Black, or a draw. In effect, the comprehensive definition of the concept of strategy renders the actual play of the game unnecessary. We may imagine the players announcing (simultaneously) their respective strategies to a referee or a machine, who can determine right away the outcome based on the announced strategies. A strategy $\bar{\sigma}$ of White is a *winning* strategy, if for every strategy τ of Black, the pair $(\bar{\sigma}, \tau)$ results in a win for White. In other words, $\bar{\sigma}$ guarantees a win for White. A winning strategy for Black is defined similarly. A *drawing* strategy for a player is a strategy that guarantees at least a draw for that player, that is, using that strategy he will win or draw against any strategy of his opponent.

The following theorem was discovered by John Von Neumann (but is often referred to, wrongly, as Zermelo's theorem).

Theorem 1

Let G be a chess-like game in which the length of a play is finitely bounded (i.e., there exists $M < \infty$ such that in every play of the game there are at most M moves). Then one of the following statements is true:

a. White has a winning strategy in G.
b. Black has a winning strategy in G.
c. Both White and Black have drawing strategies in G.

It is important to understand that Theorem 1 does not merely state the tautological fact that every play of the game results in one of the three possible outcomes. Rather, the theorem asserts that every *game* (as opposed to every play of a game) satisfying the theorem's assumptions may be classified as a win for White, a win for Black, or a draw, in the sense that it will always end that way when played optimally. We refer to this property of a game G as *determinacy*.

The proof of Theorem 1 proceeds by "backward induction". Namely, one classifies every subgame starting at a terminal node of the game-tree (this is trivial). Then one classifies every subgame starting at a node from which all moves lead to terminal nodes, and so on, working one's way to the root of the tree. For every subgame encountered in this process, it is classified as satisfying (a), (b), or (c), and at the end of the process, G itself is classified. Thus, the proof of Theorem 1 is constructive, leading to an algorithm for classifying a game and finding a winning strategy for the player who has one (or drawing strategies for both players).

Nevertheless, for most real-life games, the computational complexity of constructing the game-tree, let alone running this algorithm, is much too high. Thus, although Theorem 1 renders the games it applies to uninteresting, in principle, to players with unbounded computational abilities, in practice the games still hold interest to humans. The issues related to discovering winning strategies and the complexity of this task are studied within the field of "combinatorial games", which grew quite independently of the rest of game theory.

Incidentally, the question of whether or not the game of chess itself satisfies the assumptions of Theorem 1 depends on a careful scrutiny of the rules of chess regarding, e.g., the repetition of positions on the board. Assuming an interpretation of the rules that makes the length of a play of chess finitely bounded, we can conclude from Theorem 1 that chess is either a win for White, a win for Black, or a draw. However, nobody knows which of these it is!

There are results extending Theorem 1 to certain classes of chess-like games with infinite plays. These results, starting from Gale and Stewart's theorem on "open games", depend on topological assumptions on the set of maximal branches of the game-tree which constitute a win for a given player.

Using the Axiom of Choice, Gale and Stewart also showed that there exist infinite chess-like games, which are undetermined, that is, for which the conclusion of Theorem 1 is false. This direction of research has revealed strong connections with the foundations of mathematics, and has grown quite independently of game theory, in the field called "descriptive set theory".

3. Representations of Non-Cooperative Games

3.1 An Informal Description of the Class of Games

The games that we consider here are more general than the chess-like games considered above in several respects:

- There is a finite number n of players, possibly $n > 2$.
- The order in which the players are called upon to make a move is arbitrary, and the identity of the player who has to move at a given stage may depend on the moves made up to that stage.
- Information may be imperfect, meaning that at the time when a player is called upon to make a move, he may have only partial information on the moves made prior to that time. This includes also the possibility of simultaneous moves, which may be represented as sequential moves with the provision that the player who moves later must do so without being informed of the choice of the player who moved earlier.
- There may be chance moves, that is, moves not controlled by any player but rather selected from a set of possible moves according to some probability distribution.
- The outcome associated with any play of the game, rather than being a win for some player or a draw, is represented by an n-tuple of real numbers $(u_1, ..., u_n)$, where u_i measures the utility that player i derives from the outcome. This permits to represent the outcomes of chess, for instance as (1, 0), (0, 1), or $(\frac{1}{2}, \frac{1}{2})$, and allows for much more general situations, as we will see below.

This more flexible framework includes a variety of parlor games like Kriegspiel, backgammon, bridge, poker, and monopoly. More importantly, many real-life situations, which are not normally thought of as "games", may usefully be modeled as such. Examples include: competition between firms in an oligopolistic market, campaigns of opposing candidates running for election, struggle between genes as part of evolution, interaction between processors involved in a parallel computation, and more.

There is, however, one important implicit assumption about the games we consider here, the validity of which must be assessed for any real-life application. This is the assumption of *complete information*, meaning that a player knows the entire description of the game. In addition, the player also knows that every other player knows the entire description of the game, and furthermore he knows that every other player knows that every other player knows the entire description of the game, and so forth. This condition is expressed concisely by saying that the description of the game is common knowledge. This concept should be distinguished from the concept of perfect information. While perfect information pertains to knowledge of what happened in the current play of the game, complete information pertains to knowledge of the game itself (its rules, the relevant probability distributions, who is informed of what and when, the utilities of the various outcomes to the various players). For example, bridge players who have mastered the rules of the game are engaged in a game of complete information which has, however, imperfect information: a player is not informed of the cards dealt to the other players. In modeling real-life situations, the assumption of complete information is more problematic.

-
-
-

Bibliography

Aumann R.J. and Hart S. (Eds.). (1992,1994,2002). *Handbook of Game Theory with Economic Applications*. Volumes I II and III, Amsterdam: North-Holland. [This handbook contains surveys of various areas in game theory. Chapter 2, Games in Extensive and Strategic Forms, by S. Hart (pp. 19-40) introduces basic concepts of non-cooperative game theory. Chapter 3, Games with Perfect Information, by J. Mycielski (pp. 41-70) surveys work on infinite games with perfect information. Chapter 20, Zero-Sum Two-Person Games, by T.E.S. Raghavan (pp. 735-759) presents zero-sum games, the minimax theorem and some of its extensions. Chapter 24, Communication, Correlated Equilibria and Incentive Compatibility, by R.B. Myerson (pp. 827-847) includes, among other things, a discussion of correlated equilibrium, an important concept that is not covered in this article.]

Berlekamp E.R., Conway J.H. and Guy R.K. (1982). *Winning Ways for Your Mathematical Plays*. London: Academic Press. [This is the main reference to the area of combinatorial games.]

Fudenberg D. and Tirole J. (1991). *Game Theory*, 579 pp. Cambridge, MA: MIT Press. [This textbook presents non-cooperative game theory with special attention to its economic applications.]

Luce R.D. and Raiffa H. (1957). *Games and Decisions: Introduction and Critical Survey*, 509 pp. New York: John Wiley & Sons. [This classical book presents the main ideas and concepts of game theory with a minimum of technicalities.]

Osborne M.J. and Rubinstein A. (1994). *A Course in Game Theory*, 352 pp. Cambridge, MA: MIT Press. [This textbook presents game theory and in particular treats the concept of rationality and its bounded versions.]

Owen G. (1995). *Game Theory*, 3rd Edition, 447 pp. San Diego: Academic Press. [This textbook is oriented mainly to the mathematics student. It includes chapters on linear programming and its connections to zero-sum games.]

van Damme E. (1991). *Stability and Perfection of Nash Equilibria*, 2nd Edition, 339 pp. Berlin: Springer-Verlag. [This is a comprehensive reference on refinements of Nash equilibrium.]

Von Neumann J. and Morgenstern O. (1944). *Theory of Games and Economic Behavior*. Princeton: Princeton University Press. [This classical book laid the foundations of game theory. The second and third editions include an appendix on utility theory.]

Biographical Sketch

Ron Holzman is Professor of Mathematics at the Technion–Israel Institute of Technology. He studied at the Hebrew University of Jerusalem, where he obtained his Ph.D. in 1985 with a thesis on game theoretic aspects of voting theory. After occupying a number of visiting positions in Belgium and the U.S., he spent a few years at the Weizmann Institute of Science. He has been at the Technion since 1993. He serves on the editorial board of *Mathematical Social Sciences*. His research interests include combinatorics and game theory.