Jörgen W. Weibull

Evolutionary Game Theory

Evolutionary Game Theory

Jörgen W. Weibull

The MIT Press
Cambridge, Massachusetts
London, England

Second printing, 1996

This book was set in Times Roman by Windfall Software using ZzTEX and was printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Weibull, Jörgen W., 1948–
Evolutionary game theory / Jörgen W. Weibull.
p. cm.
Includes bibliographical references and index.
ISBN 0-262-23181-6 (alk. paper)
1. Game theory. 2. Evolution–Mathematical models. I. Title
QA269.W45 1995
519.3–dc20

94-44425
CIP

for Sofia and Anna

Foreword

Ken Binmore

When von Neumann and Morgenstern's *Theory of Games and Economic Behavior* appeared in 1944, it was greeted with great enthusiasm. It was thought that a complete theory of strategic behavior had sprung into existence from nowhere, as Athena burst fully armed from the forehead of Zeus. However, it soon became apparent that von Neumann and Morgenstern had provided only the beginnings of a theory, and those seeking quick applications became disillusioned. Game theory then spent a long period in the doldrums. The mathematics of the theory of two-person, zero-sum games continued to be studied. Much effort was also devoted to developing cooperative game theory. But the problems of noncooperative game theory in general were left largely untouched.

Von Neumann and Morgenstern being no more, the Nobel Prize for Economics was recently awarded to three game theorists, John Nash, John Harsanyi, and Reinhard Selten. Nash's work was published in the early 1950s, but it was not until the early 1970s that it was fully realized what a powerful tool Nash had provided in formulating the equilibrium concept that bears his name. Game theory then enjoyed a renaissance as economists applied the idea to a wide range of problems. However, a fly in the ointment was awaiting discovery. Games typically have many Nash equilibria. In two-person, zero-sum games, this creates no problem because all equilibria are then interchangeable and payoff-equivalent. But the equilibrium selection problem for more general games has no such easy solution.

At first it was thought that the problem could be tackled by refining the Nash equilibrium concept. Despite Nash's remarks in his thesis about a possible evolutionary interpretation of the idea of a Nash equilibrium, attention at that time was focused almost entirely on its interpretation as the only viable outcome of careful reasoning by ideally rational players. Various bells and whistles were therefore appended to the definition of rationality. These allowed some Nash equilibria to be discarded as inadequately rational according to whatever new definition of rationality was being proposed. However, different game theorists proposed so many different rationality definitions that the available set of refinements of Nash equilibrium became embarrassingly large. Eventually, almost any Nash equilibrium could be justified in terms of someone or other's refinement. As a consequence a new period of disillusionment with game theory seemed inevitable by the late 1980s.

Fortunately the 1980s saw a new development. Maynard Smith's book *Evolution and the Theory of Games* directed game theorists' attention away from their increasingly elaborate definitions of rationality. After all, insects can hardly be said to think at all, and so rationality cannot be so crucial if game theory somehow manages to predict their behavior under appropriate conditions. Simultaneously the advent of experimental economics brought home the fact that human subjects are no great shakes at thinking either. When they find their way to an equilibrium of a game, they typically do so using trial-and-error methods.

As the appearance of this book indicates, the 1990s have therefore seen a turning away from attempts to model people as hyperrational players. The new approach to the equilibrium selection problem emphasizes the almost tautological assertion that the equilibrium selected will be a function of the equilibriating process by means of which it is achieved. The process may be slow, as in biological evolution. It may be fast, as in social evolution, when the mechanism for the transmission of superior strategies from one head to another is imitation. It may be almost instantaneous, as when the price adjusts to equate supply and demand in the Chicago wheat market. However, we have learned that all these different processes have features in common that make it worthwhile considering evolutionary processes in the abstract.

Such studies teach us some painful lessons. We learn that there is nearly always evolutionary pressure against the various types of behavior labeled as "irrational" in the refinements' literature, but these pressures can vary enormously in their relative strengths. If the pressure against one type of irrationality is weak, the pressures against other types of irrationality may rush the system to an equilibrium before the pressure against the first type of irrationality has a chance to have much effect. For example, weakly dominated strategies need not be eliminated. Even strongly dominated strategies can survive in certain special cases.

We also learn that historical and institutional factors cannot be ignored. This is not a hard lesson for biologists, for whom the realities of genetic inheritance and the accidents of geography are brute facts that cannot be overlooked. But economists remain resistant to the idea that the same game might receive a different analysis if the players have a different history of experience, or live in different societies, or operate in different industries. One sometimes even reads that theories that ignore such considerations are "superior" to those that do because they are able to generate predictions with less data! However, if there is one fact that work on evolutionary games has established beyond doubt, it is that some details of the equilibriating process can have a major impact on the equilibrium selected. One of the major tasks awaiting us is to identify such significant details so that applied workers know what to look for in the environments within which the games they care about are played.

However, such a program is for the future. Jörgen Weibull's book is a compendium of progress so far in the area in which biology and economics overlap. Much of the material is his own work and that of his collaborators. It is distinguished by the clarity of the exposition and the elegance of the mathematics. He does not pretend to cover the whole field. One must look elsewhere for the nitty-gritty of population genetics or the properties of evolutionary processes with a strong stochastic component. But within his chosen area, his coverage is satisfyingly comprehensive.

Evolutionary game theory is here to stay, and I suspect this book will be a staple of its literature for many years to come. Its author is to be congratulated on having done such a fine job.

Introduction

The standard interpretation of noncooperative game theory is that the analyzed game is played exactly once by fully rational players who know all the details of the game, including each other's preferences over outcomes. Evolutionary game theory, instead, imagines that the game is played over and over again by biologically or socially conditioned players who are randomly drawn from large populations.[1] More specifically, each player is "pre-programmed" to some behavior—formally a strategy in the game—and one assumes that some evolutionary selection process operates over time on the population distribution of behaviors. What, if any, are the connections between the long-run aggregate behavior in such an evolutionary process and solution concepts in noncooperative game theory? More specifically: Are dominated strategies wiped out in the long run? Will aggregate behavior tend toward a Nash equilibrium of the game? Are some Nash equilibria more likely to emerge in this fashion than others? What is the nature of long-run aggregate behavior if it does not settle down on some equilibrium? These are the kinds of questions addressed in this book.

Similar questions have, of course, been raised in the domains of economics and biology. Market competition is usually thought to weed out firms that are not profit maximizers and to bring about the equilibrium outcomes predicted by economic theory. This is the basis for the so-called "as if" defense of economic theory, which claims that it is not important that managers think the way microeconomic theory says they do; what counts is whether they behave as if they did (Friedman 1953). Likewise natural selection is usually thought to result in animal behavior that is well adapted to the environment. In the simplest cases this environment is exogenously fixed, while in other cases the environment of an individual is itself composed of other individuals who are subject to the same forces of natural selection (this is also true for market selection). What is optimal for an individual or firm in such an interactive setting is endogenous in the sense of depending on the distribution of behaviors in the population with which the individual or firm interacts. Evolutionary game theory is designed to enable analysis of evolutionary selection in precisely such interactive environments.

1. In his unpublished Ph.D. dissertation (Nash 1950a) John Nash suggests a population-statistical interpretation of his equilibrium concept in which he imagines that players are randomly drawn from large populations, one for each player position in the game. These players were not assumed to "have full knowledge of the total structure of the game, or the ability and inclination to go through any complex reasoning process" (op. cit., p. 21); see Leonard (1994), Weibull (1994), and Björnerstedt and Weibull (1993).

Plan of the Book

Evolutionary game theory provides a tool kit of wide applicability. Its potential domain ranges from evolutionary biology to the social sciences in general and economics in particular. This book does not try to cover all the developments in the field, not even all the most important ones. Instead, it strives to give a self-contained treatment of a selected set of core elements, focused on conceptual and technical connections between evolutionary and noncooperative game theory.

Chapter 1 gives a concise introduction to noncooperative game theory. Notation, definitions, and results of relevance to the subsequent discussion are introduced, along with a number of examples that are used throughout the book. Chapters 2 through 4 deal with single-population evolutionary models of pairwise interactions represented as a symmetric two-player game. Chapter 2 considers a few static models, centered around the key concept of an evolutionarily stable stragegy. Chapter 3 focuses on a particular dynamic model of evolutionary selection in continuous time, the so-called replicator dynamics. Chapter 4 develops a few variations on the theme in chapter 3, including dynamic models of social evolution. Chapter 5 develops both static and dynamic models of multipopulation interactions represented as an n-player game. The dynamic models developed in chapters 3 through 5 use systems of ordinary differential equations to describe the evolution of aggregate behavior over time. Chapter 6 provides a concise introduction to the theory of ordinary differential equations. All chapters contain examples that illustrate the workings of the discussed methods.

The presentation of the material in many instances proceeds from the special to the general. Several themes first appear in simple examples, thereafter in specific but broader contexts, and finally in more general and abstract settings. It may annoy some mathematically well-versed readers to first see a claim proved in a special case and later in a more general case. However, it is hoped that this procedure will facilitate an operational "hands on," and not only abstract, understanding of the methods used.

The reader is assumed to have some familiarity with standard notions in mathematics (basic set theory, topology, and calculus) at about the level achieved after the first year of graduate studies in economics. Although chapter 1 provides the tools needed from noncooperative game theory, this treatment will most likely appear terse for a reader who is not acquainted with the basic ideas in noncooperative game theory. Also here the reader is presumed to have a knowledge at about the level achieved after first-year graduate studies in economics.

How to read the book, and how to use it in class? One obvious way is to read chapter 1, selected parts of chapter 2, make a short excursion into selected parts of chapter 6, and finally read selected parts of chapters 3 through 5. A shorter course could focus on parts of chapters 1, 2, 3, and 5 (e.g., sections 1.1–1.3, 1.5, 2.1–2.3, 3.1–3.3, 3.5, and 5.2).

To enable a self-contained and yet concise treatment, only deterministic models of games in normal form are discussed in this book, despite the fact that there now are a few promising evolutionary stochastic models and evolutionary models of extensive-form games. Each of these two extensions of the scope would require additional technical machinery. The reader who is interested in these and other developments in evolutionary game theory not covered here may consult the bibliography at the end of the book. For example, stochastic models are discussed in Foster and Young (1990), Kandori, Mailath, and Rob (1993), and Young (1993). Models of games in extensive form may be found in Selten (1983), van Damme (1987), and Nöldeke and Samuelson (1993). A number of other important contributions can be found in recent issues of economics and biology journals.

Acknowledgments

During the writing of this book I benefited greatly from conversations with colleagues and students. I am particularly grateful for comments on various drafts of the manuscript from Jonas Björnerstedt, Anette Bjørsted, Immanuel Bomze, Ross Cressman, Eric van Damme, Martin Dufwenberg, Josef Hofbauer, Oliver Kirchkamp, Johan Lindén, Oleg Malafeyev, Per Molander, Peter Norman, Jörg Oesschler, Klaus Ritzberger, Larry Samuelson, Karl Schlag, Björn Segendorff, Karl Sigmund, Birgitte Sloth, Tony E. Smith, Joakim Sonnegård, Johan Stennek, David Strömberg, and Karl Wärneryd.

I would also like to thank Assar Lindbeck, Janos Kornai, and Lars E. O. Svensson for their moral and intellectual support during my transition from applied mathematics to economics, and for generously sharing their profound insights. Especially stimulating for the present project was a course I took many years back in dynamic modeling of ecological systems, taught by Ingemar Nåsell, as well as conversations I had with Janos Kornai, also a long time ago, about the evolution of firm behavior in markets with varying degrees of state intervention.

Terry Vaughn of The MIT Press was a great support during the preparation of the manuscript. Special thanks go to Dana Andrus of The MIT Press for her fine editing, and in Stockholm, Sweden, to Jonas Björnerstedt for all computer illustrations in the book, to Molly Åkerlund for her excellent help with page proofs, and to Karl Wärneryd for his contribution to the cover illustration.

I am grateful to the Industrial Institute for Economic and Social Research (IUI), Stockholm, for its sponsoring of part of the research behind this text.

Mathematical Notation

Lowercase letters are mostly used for real numbers, vectors of real numbers, and for functions, while capital letters usually signify matrices and sets. Euclidean spaces are typically denoted R^n, where n is a positive integer—the dimension of the space. The subset of vectors x in R^n that have all coordinates x_i nonnegative is denoted R^n_+, and the subset of vectors that have all coordinates positive is written R^n_{++}. The *inner* (or *scalar*) *product* of two vectors x and y in R^n is a real number (scalar) written $x \cdot y = \sum_{i=1}^n x_i y_i$. The euclidean *norm* (or length) of a vector $x \in R^n$ is denoted $\|x\| = \sqrt{x \cdot x}$, and the *distance* between two points (vectors) x and y in R^n is written $d(x, y) = \|x - y\|$. The transpose of an $n \times n$ matrix A is denoted A^T.

In this book $\subset$ denotes *weak* set inclusion. Hence $X \subset Y$ signifies that all elements of X are also elements of Y. The *complement* of a set $X \subset R^n$ is written $\sim X$. By a *neighborhood* of a point (vector) x in R^n is meant an open set $U \subset R^n$ containing x. The *interior* of a set $X \subset R^n$ is written $\text{int}(X)$; this is the subset of points x in X such that X also contains some neighborhood of x. The *boundary* of a set $X \subset R^n$ is written $\text{bd}(X)$; this is the set of points $y \in R^n$ such that every neighborhood of y contains some point from X and some point from $\sim X$. The *closure* of a set $X \subset R^n$ is denoted $\overline{X}$; this is the union of X and its boundary. A *function* f from a set X to a set Y is viewed as a rule that to each element x of X assigns precisely one element, $f(x)$, of Y. Likewise a *correspondence* φ from a set X to a set Y is a rule that to each element x of X assigns precisely one nonempty subset, $\varphi(x)$, of Y.

Evolutionary Game Theory

Contents

1 Elements of Noncooperative Game Theory

This chapter provides an introduction to the concepts and results in noncooperative game theory that will be used in the subsequent evolutionary analysis. The material in this chapter is organized as follows: In section 1.1 the structure of finite normal-form games is outlined. In particular, the geometry of strategy spaces and multilinearity of payoff functions is emphasized. Section 1.2 discusses dominance orderings of a player's strategy space and formalizes the notion of "best replies." Section 1.3 considers Nash equilibria as fixed points of the best-reply correspondence, and studies some properties of the set of Nash equilibria. Section 1.4 gives a brief account of some point- and setwise refinements of the Nash equilibrium concept. Section 1.5 introduces some special notation for, and properties of, symmetric two-player games; the basic setting in chapters 2 through 4. Many of the examples introduced in the chapter will be used later to illustrate evolutionary concepts.

The reader who wishes to have a fuller treatment of noncooperative game theory is advised to consult Fudenberg and Tirole (1991) or, for a more concise and technical treatment, van Damme (1987).

1.1 Strategies and Payoff Functions

The analysis in this book is restricted to *finite games in normal form*. More precisely, let $I = \{1, 2, \ldots, n\}$ be the set of *players*, where n is a positive integer. For each player $i \in I$, let S_i be her finite set of *pure strategies*. For notational convenience, we will label every player's pure strategies by positive integers. Hence the pure-strategy set of any player $i \in I$ is written $S_i = \{1, 2, \ldots, m_i\}$, for some integer $m_i \geq 2$. A vector s of pure strategies, $s = (s_1, s_2, \ldots, s_n)$, where s_i is a pure strategy for player i, is called a *pure-strategy profile*. The set of pure strategy profiles in the game is thus the cartesian product $S = \times_i S_i$ of the players' pure strategy sets, sometimes to be called the *pure-strategy space* of the game.

For any strategy profile $s \in S$ and player $i \in I$, let $\pi_i(s) \in R$ be the associated payoff to player i. In economics the payoffs are usually firms' profits or consumers' (von Neumann-Morgenstern) utility, while in biology payoffs usually represent individual fitness (expected number of surviving offspring). The finite collection of real numbers $\pi_i(s)$ defines the ith player's *(pure-strategy) payoff function* $\pi_i : S \to R$, for each player $i \in I$. The combined pure-strategy payoff function of the game, $\pi : S \to R^n$, assigns to each pure-strategy profile s the full vector $\pi(s) = (\pi_1(s), \ldots, \pi_n(s))$ of payoffs.

In terms of pure strategies, a game in normal form may be summarized as a triplet $G = (I, S, \pi)$, where I is its player set, S its pure-strategy space, and π its combined payoff function. In the special case when there are only two players, one may conveniently write each of the two payoff functions π_1 and π_2 in tabular form as an $m_1 \times m_2$ matrix. We will usually denote the first player's payoff matrix $A = (a_{hk})$, where $a_{hk} = \pi_1(h, k)$ for each $h \in S_1$ and $k \in S_2$, and will likewise denote the second player's payoff matrix $B = (b_{hk})$, where $b_{hk} = \pi_2(h, k)$. Each row in both matrices thus corresponds to a pure strategy for player 1, and each column to a pure strategy for player 2. Any two-player game can be fully represented by the associated payoff matrix pair (A, B), where player 1 is understood to be the "row player" and player 2 the "column player."

Example 1.1 The most widely known game is probably the Prisoner's Dilemma Game, a two-player game in which each player has only two pure strategies. A typical configuration of payoffs is given in the matrix pair

$$A = \begin{pmatrix} 4 & 0 \\ 5 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & 5 \\ 0 & 3 \end{pmatrix}. \tag{1.1}$$

Evidently player 1's second pure strategy ("defect") gives a higher payoff than her first pure strategy ("cooperate"), irrespective of which strategy is used by player 2; each entry in the second row of matrix A exceeds the corresponding entry in the first row. Likewise player 2's second pure strategy always gives her a higher payoff than her first pure strategy; each entry in B's second column exceeds the corresponding entry of its first column. Hence individual rationality leads each player to select her second pure strategy (defect). The dilemma consists in the fact that both players would earn higher payoffs if they were to select their first pure strategy (cooperate).

1.1.1 The Geometry of Mixed-Strategy Spaces

A *mixed strategy* for player i is a probability distribution over her set S_i of pure strategies. Since for each player $i \in I$ the set S_i is finite, we can represent any mixed strategy x_i for player i as a *vector* x_i in m_i-dimensional euclidean space R^{m_i}, its hth coordinate $x_{ih} \in R$ being the probability assigned by x_i to the player's hth pure strategy.