

### EVOLUTION OF THE SOCIAL CONTRACT

In this new edition of Evolution of the Social Contract, Brian Skyrms uses evolutionary game theory to analyze the genesis of social contracts and investigates social phenomena including justice, communication, altruism, and bargaining. Featuring new material on evolution and information transfer, and including recent developments in game theory and evolution literature, his book introduces and applies appropriate concepts of equilibrium and evolutionary dynamics, showing how key issues can be modeled as games and considering the ways in which evolution sometimes supports, and sometimes does not support, rational choice. He discusses topics including how bargaining with neighbors promotes sharing of resources, the diversity of behavior in ultimatum bargaining in small societies, the Prisoner's Dilemma, and an investigation into signaling games and the spontaneous emergence of meaningful communication. His book will be of great interest to readers in philosophy of science, social science, evolutionary biology, game and decision theory, and political theory.

**Brian Skyrms** is Distinguished Professor of Logic and Philosophy of Science and of Economics at the University of California, Irvine and Professor of Philosophy at Stanford University. His publications include *The Stag Hunt and the Evolution of Social Structure* (Cambridge, 2004), *Signals: Evolution, Learning, and Information* (2010), *From Zeno to Arbitrage: Essays on Quantity, Coherence, and Induction* (2012), and *Evolution of the Social Contract* (first edition, Cambridge, 1996), which won the 1999 Lakatos Award in Philosophy of Science.





# Evolution of the Social Contract

Second Edition

BRIAN SKYRMS
University of California,
Irvine





# CAMBRIDGE UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org Information on this title: www.cambridge.org/9781107434288

© Brian Skyrms 2014

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 1996 12<sup>th</sup> printing 2011

Second edition first published 2014

A catalogue record for this publication is available from the British Library

Library of Congress Cataloging-in-Publication Data Skyrms, Brian.

Evolution of the social contract / Brian Skyrms. – Second edition.

pages cm

ISBN 978-1-107-07728-7 (Hardback) – ISBN 978-1-107-43428-8 (Paperback)

1. Social contract. 2. Game theory. I. Title.

JC336.S49 2014 320.1'1-dc23 2014025188

ISBN 978-1-107-07728-7 Hardback ISBN 978-1-107-43428-8 Paperback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.



For Pauline, Michael, and Gabriel





Two men who pull the oars of a boat, do it by an agreement or convention, tho' they have never given promises to each other. Nor is the rule concerning the stability of possession the less derived from human conventions, that it arises gradually, and acquires force by a slow progression ... In like manner are languages established by human conventions without any promise.

David Hume, A Treatise of Human Nature





# **CONTENTS**

| Preface to the second edition Preface to the original edition Acknowledgments |                          | page xi |
|---|--------------------------|---------|
|   |                          | xiii    |
|   |                          | xvi     |
| 1   | Sex and justice          | 1       |
| 2   | Fairness and commitment  | 23      |
| 3   | Mutual aid               | 43      |
| 4   | Correlated convention    | 64      |
| 5   | The evolution of meaning | 81      |
|   | Postscript               | 106     |
| Notes   |                          | 111     |
| References  |                          | 127     |
| Index   |                          | 145     |

ix





# PREFACE TO THE SECOND EDITION

**S** OME things stay the same, and some things change. We know more about some of the issues addressed in this book than we did when it was written. I have made changes, some small and some larger, to bring the discussion up to date. There are new footnotes and lots of new references, of course. Some larger modifications are made in the substantive discussions.

In Chapter 1 I say a little more to preview the importance of correlation that is to be the main theme in Chapter 3. I refer to correlation set up by local interaction in the form of bargaining with neighbors. This is work done with Jason McKenzie Alexander that was produced subsequent to the publication of the first edition. I also make a historical correction: Darwin himself really cracked the sex ratio puzzle.

I am more circumspect in Chapter 2 about the connection between rationality and backward induction, perhaps one might say evasive, since I do not want to plunge into the counterfactual reasoning involved. That is not what this book is about. I change terminology from "modular rationality" to "sequential rationality" to bring it into alignment with the standard terminology in the game theory literature. The empirical discussion of behavior in ultimatum bargaining now includes the work of anthropologists, which reveals great variability across small-scale societies. This fits well with a



### Preface to the second edition

discussion of framing of social norms that was only hinted at in the first edition. There is a reference to new work with Kevin Zollman on a model of such framing.

Chapter 3 now has a new section on inclusive fitness, showing quite simply how to get a version of Hamilton's rule from just correlated encounters. There is also a new section discussing important correlation mechanisms: the family, partner choice, reciprocal altruism implemented in various ways, local interaction, and Maynard Smith's haystack model of a kind of group selection.

The discussion of the formation of correlated conventions by symmetry-breaking in Chapter 4 is largely unchanged. But Chapter 5 had to be modified to take into account the fact that a great deal has been learned about the dynamics of signaling games. What was conjectured on the basis of simulations combined with partial analysis of the simplest signaling games has now been proved. But one has to be careful about generalizing to more complicated signaling games. The basic idea, that meaning – in the form of signaling system equilibria – can arise spontaneously, continues to hold good. But this is guaranteed to happen only in some circumstances. Chapter 5 is now corrected, but there is a lot more to say. I wrote another book about this.

The postscript was supposed to point to direction for further investigations. I tried to follow my own directions, and I could now add a pointer to work by myself and others on dynamics of social network formation.



# PREFACE TO THE ORIGINAL EDITION

THE best-known tradition approaches the social contract in terms of rational decision. It asks what sort of contract rational decision makers would agree to in a preexisting "state of nature." This is the tradition of Thomas Hobbes and – in our own time – of John Harsanyi and John Rawls. There is another tradition – exemplified by David Hume and Jean-Jacques Rousseau – which asks different questions. How can the existing implicit social contract have evolved? How may it continue to evolve? This book is intended as a contribution to the second tradition.

Hegel and Marx are, in a way, on the periphery of the second tradition. Lacking any real evolutionary dynamics, they resorted to the fantasy of the dialectical logic of history. It was Darwin who recognized that the natural dynamics of evolution is based on differential reproduction. Something like differential reproduction operates on the level of cultural as well as biological evolution. Successful strategies are communicated and imitated more often than unsuccessful ones. In the apt language of Richard Dawkins, we may say that both cultural and biological evolution are processes driven by differential replication. There is a simple dynamical model of differential replication now commonly called the *replicator dynamics*. Although this dynamics is surely oversimplified from both

xiii



### Preface to the original edition

biological and cultural perspectives, it provides a tractable model that captures the main qualitative features of differential replication. The model can be generalized to take account of mutation and recombination. These biological concepts also have qualitative analogues in the realm of cultural evolution. Mutation corresponds to spontaneous trial of new behaviors. Recombination of complex thoughts and strategies is a source of novelty in culture. Using these tools of evolutionary dynamics, we can now study aspects of the social contract from a fresh perspective.

Some might argue that, in the end, both traditions should reach the same conclusion because natural selection will weed out irrationality. This argument is not quite right, and one way of reading the book is to concentrate on how it is not right. Chapter 1 juxtaposes the biological evolution of the sex ratio with cultural evolution of distributive justice. It shows how evolution imposes a "Darwinian veil of ignorance" that often (but not always) leads to selection of fair division in a simple bargaining game. In contrast, rational decision theory leads to an infinite number of equilibria in informed rational selfinterest. Chapter 2 shows that evolution may not eliminate behavior that punishes unfair offers at some cost to the punisher. Such strategies can survive even though they are "weakly dominated" by alternatives that could do better and could not do worse. Chapter 3 widens the gap between rational decision and evolution. If evolutionary game theory is generalized to allow for correlation of encounters between players and like-minded players, then strongly dominated strategies - at variance with both rational decision and game theory - can take over the population. Correlation implements a "Darwinian categorical imperative" that provides a general unifying account of the conditions for the evolution of altruism and mutual aid. Chapter 4 deals in general with situations in which rational choice cannot decide between symmetric optimal

xiv



### Preface to the original edition

options. Evolutionary dynamics can break the "curse of symmetry" and lead to the formation of correlated conventions. The genesis of "ownership" behavior and thus the rudiments of the formation of the concept of property are a case in point. Chapter 5 shows how meaning is spontaneously attached to tokens in a signaling game. Here rational choice theory allows "babbling equilibria" where tokens do not acquire meaning, but consideration of the evolutionary dynamics shows that the evolution of meaning is almost inevitable. Throughout a range of problems associated with the social contract, the shift from the perspective of rational choice theory to that of evolutionary dynamics makes a radical difference. In many cases, anomalies are explained and supposed paradoxes disappear.

The two traditions, then, do not come to the same conclusions. There are points of correspondence, but there are also striking differences. In pursuing the tradition of Hume, my aims are explanatory rather than normative. Sometimes, I am happy explaining how something could have evolved. Sometimes I think I can say why something must have evolved, given any plausible evolutionary dynamics. In intermediate cases, we can perhaps say something about the range of initial conditions that would lead to a given result. When I contrast the results of the evolutionary account with those of rational decision theory, I am not criticizing the normative force of the latter. I am just emphasizing the fact that the different questions asked by the two traditions may have different answers.

Although there is real game theory and real dynamics behind the discussions in this book, I have reserved the technical details for scholarly journals. No special background is presupposed. Useful concepts are introduced along the way. I hope and believe that this book should be generally accessible to readers who wish to pursue the fascinating issues of a naturalistic approach to the social contract.



# **ACKNOWLEDGMENTS**

have many friends to thank for useful comments on some  $oldsymbol{1}$  or all of the contents of this book. In particular, I would like to thank Justin d'Arms, Francisco Ayala, Cristina Bicchieri, Ken Binmore, Vincent Crawford, Branden Fitelson, Steven Frank, Alan Gibbard, Clark Glymour, Peter Godfrey-Smith, Bill Harms, Bill Harper, John Harsanyi, Jack Hirshleifer, Richard Jeffrey, Ehud Kalai, Philip Kitcher, Karel Lambert, David Lewis, Barbara Mellers, Jordan Howard Sobel, Patrick Suppes, Peter Vanderschraaf, Bas van Fraassen, and Bill Wimsatt. I owe much to discussions with my late friend and colleague Greg Kavka. The first draft of this book was completed at the Center for Advanced Study in the Behavioral Sciences in 1993–4. I am grateful for financial support provided by the National Science Foundation, the Andrew Mellon Foundation, and the University of California President's Fellowship in the Humanities. The University of California at Irvine provided a grant of computer time at the San Diego Supercomputer Center, which made some of the larger simulations reported here possible. Substantial portions of "Sex and Justice" and "Darwin Meets 'The Logic of Decision' " are reprinted here with permission of The Journal of Philosophy and Philosophy of Science, respectively.

xvi