

Jordan Howard Sobel has long been recognized as an important figure in philosophical discussions of rational decision. He has done much to help formulate the concept of causal decision theory.

In this volume of essays, Sobel explores the Bayesian idea that rational actions maximize expected values, where an action's expected value is a weighted average of its agent's values for its possible total outcomes. Newcomb problems and the Prisoners' Dilemma are discussed, and Allais-type puzzles are viewed from the perspective of causal world Bayesianism. The author establishes principles for distinguishing options in decision problems, and studies ways in which perfectly rational causal maximizers can be capable of resolute choices. Several of the essays concern games, with interacting ideally rational and well-informed maximizing rationality. Sobel also views critically David Gauthier's revisionist ideas about maximizing rationality.

This collection will be a desideratum for anyone working in the field of rational choice theory, whether in philosophy, economics, political science, psychology, or statistics.

"Howard Sobel's work in decision theory is certainly among the most important, interesting, and challenging that is being done by philosophers."

- David Gauthier



Taking chances



Cambridge Studies in Probability, Induction, and Decision Theory

General editor: Brian Skyrms

Advisory editors: Ernest W. Adams, Ken Binmore, Jeremy Butterfield, Persi Diaconis, William L. Harper, John Harsanyi, Richard C. Jeffrey, Wolfgang Spohn, Patrick Suppes, Amos Tversky, Sandy Zabell

This series is a forum for the most innovative and challenging work in the theory of rational decision. It focuses on contemporary developments at the interface between philosophy, psychology, economics, and statistics. The series addresses foundational theoretical issues, often quite technical ones, and therefore assumes a distinctly philosophical character.

Other titles in the series
Ellery Eells, Probabilistic Causality
Richard Jeffrey, Probability and the Art of Judgment
Robert Koons, Paradoxes of Strategic Rationality
Cristina Bicchieri and Maria Luisa Dalla Chiara (eds.), Knowledge, Belief, and Strategic Interaction
Patrick Maher, Betting on Theories
Cristina Bicchieri, Rationality and Coordination

Forthcoming

Patrick Suppes and Mario Zanotti, Foundations of Probability with Applications Clark Glymour and Kevin Kelly (eds.), Logic, Confirmation, and Discovery



Taking chances Essays on rational choice

Jordan Howard Sobel
University of Toronto





CAMBRIDGE UNIVERSITY PRESS Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521416351

© Cambridge University Press 1994

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 1994
This digitally printed version 2007

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

Library of Congress Cataloguing in Publication data

Sobel, Jordan Howard,

Taking chances: essays on rational choice / Jordan Howard Sobel.

 cm. - (Cambridge studies in probability, induction, and decision theory)

Includes bibliographical references and index. ISBN 0-521-41635-3 (hc)

1. Bayesian statistical decision theory. 2. Statistical decision.

I. Title. II. Series. QA279.5.S635 1994

003'.56 - dc20

93-4800 CIP

ISBN 978-0-521-41635-1 hardback ISBN 978-0-521-03898-0 paperback



Contents

Pr	eface	page ix
	PART I WORLD BAYESIANISM	
1	Utility theory and the Bayesian paradigm	3
	0. Introduction	3
	1. Utility theory	4
	2. The problem	5
	3. A reaction: Insist on complete basic alternatives	9
	4. Utility theory within a more general theory	13
	5. Summing up	26
	PART II PROBLEMS FOR EVIDENTIAL DECISION THEORY	
2	Newcomblike problems	31
	0. Introduction	31
A.	Analysis and variety	32
	1. An analysis of Newcomblike problems	32
	2. The variety of Newcomblike problems	33
В.	Objections	36
	1. These are problems for space cadets	36
	2. Deliberators cannot view their choices as signs of prior conditions	38
	3. Agents who believe in causes for their actions	41
	4. Agents who believe in causes for their choices	45
	5. Newcomblike problems are not possible for ideally	73
	rational agents	48
	6. Conclusions	58
	Appendix I: Statements of some Newcomblike problems	60
	Appendix II: Further structure for "Brian Skyrms's Uncle	
	Albert" and "How do you spell relief?"	63
	Appendix III: Dominance problems	65
	Appendix IV: Further structures for Newcomb's Problem	66
	Appendix V: Ratificationism	68



3	Not every prisoners' dilemma is a Newcomb problem	77
	0. Introduction	77
	1. Near-certainty Newcomb problems	78
	2. Near-certainty prisoners' dilemmas	80
	3. Near-certainty Newcomb problems and prisoners'	
	dilemmas compared	80
	4. Conclusions	86
4	Some versions of Newcomb's Problem are prisoners'	
	dilemmas	89
	0. Introduction	89
	1. Preliminaries	90
	2. Some, but not all, versions of Newcomb's Problem are	
	prisoners' dilemmas	92
	3. Further questions	94
	4. Concluding remarks	96
5	Infallible predictors	100
	0. Introduction	100
	1. Predictors who are certainly incapable of error	101
	2. Predictors who are certainly unerring	105
	3. Conclusions	111
6	Kent Bach on good arguments	119
7	Maximizing and prospering	126
	0. Introduction	126
	1. Gauthier's pragmatism	127
	2. Implications for Newcomb problems	128
	3. Possibly disturbing aspects of these implications	130
	4. Judgmental conclusions	132
	Appendix: Socializing sentiments	136
	PART III CAUSAL DECISION THEORY	
8	Notes on decision theory: Old wine in new bottles	141
	0. Introduction: Bayesian decision theory	141
	0. Introduction: Bayesian decision theory1. Jeffrey's logic of decision	141 143
	· · · · · · · · · · · · · · · · · · ·	
	1. Jeffrey's logic of decision	143
	 Jeffrey's logic of decision The popcorn problem 	143 145
	 Jeffrey's logic of decision The popcorn problem A causal decision theory 	143 145 152
9	 Jeffrey's logic of decision The popcorn problem A causal decision theory Conclusions Appendix: Demonstrations of partition theorems Partition theorems for causal decision theories	143 145 152 164
9	 Jeffrey's logic of decision The popcorn problem A causal decision theory Conclusions Appendix: Demonstrations of partition theorems 	143 145 152 164 165

vi



	2. Causal decision theories need partition theorems	178
	3. Two partition theorems	181
	4. Fishburn and conditional acts	185
	5. Armendt on conditional preferences	187
	6. Partitions for U and Exclusive Partitions compared	191
	7. Uses of partition theorems	192
	Appendix: A theorem for sufficiently fine partitions	195
10	Expected utilities and rational actions and choices	197
	0. Introduction	197
	1. Definitions, assumptions, and restrictions	198
	2. The ideal stability of rational decisions	200
	3. Principles that apply tests of expected utility to actions4. A principle that confines tests of expected utility to	202
	choices	212
	5. A principle for agents who are sure they can make	
	mixed choices	215
	6. Conclusion	215
11	Maximization, stability of decision, and actions in	
	accordance with reason	218
	0. Introduction	218
	1. Positions	218
	2. Arguments and perspectives	227
	3. The maximization and stability theory: Restatement	
	and elaboration	233
12	Useful intentions	237
	0. Introduction	237
	1. Senses in which the rational cannot intend irrational	
	actions	238
	2. Forming and adopting intentions	240
	3. Magical bootstrapping, and rational intentions and	
	preferences	243
	4. Rational adoption of intentions to do things that	
	would otherwise be irrational	246
	5. Conclusion	251
	6. Postscript	252
	PART IV INTERACTING CAUSAL MAXIMIZERS	
13	The need for coercion	257
	0. Introduction	257
	1. The hyperrational community	258

vii



	2. The community's need for coercion	262
	3. The individual's need for coercion	270
	Appendix: The Farmer's Dilemma and mutual trust	276
14	Hyperrational games	283
	0. Introduction	283
	1. The concept of a hyperrational normal-form game	283
	2. Resolutions of hyperrational games	294
	Necessary conditions for, or limitations on, resolutions	294
	Sufficient conditions for, and ways to, resolutions	303
	3. Problems for members of hyperrational communities	313
	Games in which they would not do well	313
	Games in which they could not do anything at all	315
	The significance of these problems	324
	Appendix: Lewis domination	325
15	Utility maximizers in iterated prisoners' dilemmas	330
15	Utility maximizers in iterated prisoners' dilemmas 0. Introduction	330 330
15	· · · · · · · · · · · · · · · · · · ·	
15	0. Introduction	
15	0. Introduction1. The defeat of utility maximizers in iterated prisoners'	330
15 16	0. Introduction1. The defeat of utility maximizers in iterated prisoners' dilemmas	330 331
	 Introduction The defeat of utility maximizers in iterated prisoners' dilemmas The significance of this defeat 	330 331 338
	 Introduction The defeat of utility maximizers in iterated prisoners' dilemmas The significance of this defeat Backward induction arguments: A paradox regained	330 331 338 345
	 Introduction The defeat of utility maximizers in iterated prisoners' dilemmas The significance of this defeat Backward induction arguments: A paradox regained Introduction 	330 331 338 345 345
	 Introduction The defeat of utility maximizers in iterated prisoners' dilemmas The significance of this defeat Backward induction arguments: A paradox regained Introduction Indicative rationality and belief premises 	330 331 338 345 345 346
	 Introduction The defeat of utility maximizers in iterated prisoners' dilemmas The significance of this defeat Backward induction arguments: A paradox regained Introduction Indicative rationality and belief premises A backward argument for defection 	330 331 338 345 345 346 348
	 Introduction The defeat of utility maximizers in iterated prisoners' dilemmas The significance of this defeat Backward induction arguments: A paradox regained Introduction Indicative rationality and belief premises A backward argument for defection Subjunctive rationality and belief premises 	330 331 338 345 345 346 348 351
16	 Introduction The defeat of utility maximizers in iterated prisoners' dilemmas The significance of this defeat Backward induction arguments: A paradox regained Introduction Indicative rationality and belief premises A backward argument for defection Subjunctive rationality and belief premises Interlude: Defense of common-knowledge premises 	330 331 338 345 345 346 348 351 358



Preface

The Bayesian idea with which the essays of this volume are concerned is that rational actions maximize agents' expected values. I understand an agent's expected value for an action to be a weighted average of his values for its possible outcome, where the weight for the value of a particular possible outcome is his probability for this outcome given the action.

The essay in Part I features an interpretation of "outcomes" according to which an action's possible outcomes detail all choice-relevant aspects of its possible worlds: "Outcomes" in expected values are "worlds for practical purposes." I explain how interpreting outcomes in this comprehensive manner leads to a theory that accommodates patterns of preference such as those framed by Maurice Allais that can embarrass less commodious Bayesianisms such as those of Leonard Savage and of R. Duncan Luce and Howard Raiffa.

The essays in Part II study problems that challenge so-called evidential decision theories, such as Richard Jeffrey's, that interpret "the probability for an outcome given an action" as a conditional probability that reflects the possible evidential bearing for the agent of this action on that outcome, a probability that would make an action's expected value a measure of the welcomeness to the agent of information that it will take place. I suggest in Chapter 2 that such theories – whether or not "metatickle enhanced," as Ellery Eells would have them be - yield wrong answers in some Newcomblike problems, and give wrong reasons even when they get right answers. Challenges to the coherence of Newcomblike problems are examined in that chapter; relations of Newcomb problems and prisoners' dilemmas are discussed in Chapters 3 and 4, and in Chapter 5 the two-box prescription of causal decision theories is defended even for infalliblepredictor Newcomb problems. Causal maximizing arguments are held to be good in a sense identified in Chapter 6, even when the agent is sure that the predictor has correctly predicted his choice; for even then, causal maximizing – which still prescribes the two-box choice – is certain to lead to a payoff as great as would any other kind of practical thinking. Chapter 7 responds to challenges to the intelligence of maximizers, especially causal



maximizers, that can be found in the writings of David Gauthier: I oppose the charge that maximizers, since they would in some circumstances not do as well as, are thereby less rational than, certain other thinkers.

Part III explores so-called causal decision theories already in play in Part II, theories that would interpret "the probability for an outcome given an action" as something like the unconditional probability of this action's causing that outcome. Such theories would make an action's expected value a measure of the welcomeness, in view of the differences it figures to make, of the fact that it will take place. A theory is described in Chapter 8 that features subjective probabilities for "chancy objective conditionals"; in calculating expected values, this theory uses not probabilities of causal action-outcome "would" conditionals but, rather, probabilityweighted averages of chances in action-outcome "chance" conditionals. An account is proposed in Chapter 10 of partitioning agents' options, and accounts are proposed in Chapters 8 and 9 of partitions of circumstances for applications of causal decision theories. Chapters 10 and 11 contain the suggestion that a rational action must not only maximize causal expected value; additionally, a decision for that action must, in a certain sense, be "ratifiable." These conditions are said to be individually necessary and jointly sufficient for an action's being, in a certain thin sense, rational. Chapter 12 explores the ways in which agents who are in this sense rational may be able to adopt useful intentions and make resolute choices, including intentions and choices to perform antecedently nonmaximizing irrational actions.

Part IV deals with interactions of causal maximizers. I maintain in Chapters 13 and 14 that not even ideally rational and well-informed causal maximizers would invariably do well when interacting: There are situations in which they would do less well than other agents. I contend, further, that situations exist in which they, in their perfect and pure causal maximizing transparency, could no nothing at all. I maintain in Chapters 15 and 16 that perfectly rational and well-informed causal maximizers not only do less well than other agents in prisoners' dilemmas but do so even when they realize they are in ongoing relationships and will meet in a sequence of such dilemmas.

Debts are acknowledged with appreciation in notes to the essays. Special thanks are due to Włodek Rabinowicz and Willa Freeman-Sobel for their comments and criticisms. Essays were revised and a manuscript for this volume assembled at St. Andrews during Whitsunday term in 1992, when I was a fellow at the Centre for Philosophy and Public Affairs, to which I am grateful. This book is dedicated to my teachers and friends Richard Cartwright and David Falk.