The Probabilistic Foundations of Rational Learning

According to Bayesian epistemology, rational learning from experience is *consistent* learning, that is learning should incorporate new information consistently into one's old system of beliefs. Simon Huttegger argues that this core idea can be transferred to situations where the learner's informational inputs are much more limited than conventional Bayesianism assumes, thereby significantly expanding the reach of a Bayesian type of epistemology. What results from this is a unified account of probabilistic learning in the tradition of Richard Jeffrey's "radical probabilism". Along the way, Huttegger addresses a number of debates in epistemology and the philosophy of science, including the status of prior probabilities, whether Bayes' rule is the only legitimate form of learning from experience, and whether rational agents can have sustained disagreements. His book will be of interest to students and scholars of epistemology, of game and decision theory, and of cognitive, economic, and computer sciences.

SIMON M. HUTTEGGER is Professor of Logic and Philosophy of Science at the University of California, Irvine. His work focuses on game and decision theory, probability, and the philosophy of science, and has been published in numerous journals.

The Probabilistic Foundations of Rational Learning

SIMON M. HUTTEGGER
University of California, Irvine

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

4843/24, 2nd Floor, Ansari Road, Daryaganj, Delhi – 110002, India

79 Anson Road, #06–04/06, Singapore 079906

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/9781107115323
DOI: 10.1017/9781316335789

First published 2017

Printed in the United Kingdom by Clays, St Ives plc

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

ISBN 978-1-107-11532-3 Hardback

For my parents, Maria and Simon

Contents

Figures

Preface and Acknowledgments

The work presented here develops a comprehensive probabilistic approach to learning from experience. The central question I try to answer is: “What is a correct response to some new piece of information?” This question calls for an evaluative analysis of learning which tells us whether, or when, a learning procedure is rational. At its core, this book embraces a Bayesian approach to rational learning, which is prominent in economics, philosophy of science, statistics, and epistemology. Bayesian rational learning rests on two pillars: *consistency* and *symmetry*. Consistency requires that beliefs are probabilities and that new information is incorporated consistently into one’s old beliefs. Symmetry leads to tractable models of how to update probabilities. I will endorse this approach to rational learning, but my main objective is to extend it to models of learning that seem to fall outside the Bayesian purview – in particular, to models of so-called “bounded rationality.” While these models may often not be reconciled with Bayesian decision theory (maximization of expected utility), I hope to show that they are governed by consistency and symmetry; as it turns out, many bounded learning models can be derived from first principles in the same way as Bayesian learning models.

This project is a continuation of Richard Jeffrey’s epistemological program of *radical probabilism*. Radical probabilism holds that a proper Bayesian epistemology should be broad enough to encompass many different forms of learning from experience besides conditioning on factual evidence, the standard form of Bayesian updating. The fact that boundedly rational learning can be treated in a Bayesian manner, by using consistency and symmetry, allows us to bring them under the umbrella of radical probabilism; in a sense, a broadly conceived Bayesian approach provides us with “the one ring to rule them all” (copyright Jeff Barrett). As a consequence, the difference between high rationality models and bounded rationality models of learning is not as large as it is sometimes thought to be; rather than residing in the core principles of rational learning, it originates in the type of information used for updating.

Many friends and colleagues have helped with working out the ideas presented here. Jeff Barrett (who contributed much more than the ring

metaphor), Brian Skyrms, and Kevin Zollman have provided immensely helpful feedback prior to as well as throughout the process of writing this book. My late friend Werner Callebaut introduced me to Herbert Simon's ideas about bounded rationality. Hannah Rubin spotted a number of weaknesses in my arguments. Gregor Grehslehner, Sabine Kunrath, and Gerard Rothfus read the entire manuscript very carefully and gave detailed comments. Many others have provided important feedback: Johannes Brandl, Justin Bruner, Kenny Easwaran, Jim Joyce, Theo Kuipers, Louis Narens, Samir Okasha, Jan-Willem Romeijn, Teddy Seidenfeld, Bas van Fraassen, and Carl Wagner. I have also profited from presenting material at the University of Groningen, the University of Salzburg, the University of Munich, the University of Bielefeld, and the University of Michigan, and from conversations with Albert Anglberger, Brad Armendt, Cristina Bicchieri, Peter Brössel, Jake Chandler, Christian Feldbacher, Patrick Forber, Norbert Gratzl, Josef Hofbauer, Hannes Leitgeb, Arthur Merin, Cailin O'Connor, Richard Pettigrew, Gerhard Schurz, Reuben Stern, Peter Vanderschraaf, Kai Wehmeier, Paul Weingartner, Charlotte Werndl, Greg Wheeler, Sandy Zabell, and Francesca Zaffora Blando. I would, moreover, like to thank the team at Cambridge University Press and two anonymous referees.

UC Irvine provided time for a much needed sabbatical leave in 2013–14, which I spent writing the first third of this book by commuting between the Department of Philosophy at Salzburg and the Munich Center for Mathematical Philosophy. I'd like to thank these two institutions for their hospitality, as well as Laura Perna and Volker Springel for allowing me to live in their beautiful and wonderfully quiet Munich apartment.

Some parts of the book rely on previously published articles. Material from "Inductive Learning in Small and Large Worlds" (*Philosophy and Phenomenological Research*) is spread out over Chapters 2, 4, and 5; Chapter 6 is mostly based on "In Defense of Reflection" (*Philosophy of Science*) and "Learning Experiences and the Value of Knowledge" (*Philosophical Studies*); and Chapter 8 draws on my "Merging of Opinions and Probability Kinematics" (*The Review of Symbolic Logic*). I thank the publishers for permission to reproduce this material here.

My greatest personal thanks go to a number of people whose generosity and help have been essential for putting me in the position to write this book. Back in Salzburg, I'm particularly indebted to Hans Czermak and Georg Dorn; without Georg I would have left philosophy, and without Hans I wouldn't have learned any interesting mathematics. Since I first came to Irvine, the members of the Department of Logic and Philosophy of Science, and especially Brian Skyrms, have taken an interest in my intellectual development and my career that has gone far beyond the call of duty. The unwavering support of my parents, to whom I dedicate this book, has been invaluable; I learned from them that meaningfulness and deep engagement are more important than mere achievement or success. My sisters and my brother have been my earliest companions and friends, and they still are among my best. Finally, I thank Sabine, Teresa, and Benedikt for their love.