Uncertain Inference

Coping with uncertainty is a necessary part of ordinary life and is crucial to an understanding of how the mind works. For example, it is a vital element in developing artificial intelligence that will not be undermined by its own rigidities. There have been many approaches to the problem of uncertain inference, ranging from probability to inductive logic to nonmonotonic logic. This book seeks to provide a clear exposition of these approaches within a unified framework.

The principal market for the book will be students and professionals in philosophy, computer science, and artificial intelligence. Among the special features of the book are a chapter on evidential probability, an interpretation of probability specifically developed with an eye to inductive and uncertain inference, which has not received a basic exposition before; chapters on nonmonotonic reasoning and theory replacement that concern matters rarely addressed in standard philosophical texts; and chapters on Mill’s methods and statistical inference that cover material sorely lacking in the usual treatments of artificial intelligence and computer science.

Henry E. Kyburg, Jr. is the Burbank Professor of Philosophy and Computer Science, at the University of Rochester and a research scientist at the Institute for Human and Machine Cognition at the University of West Florida. He is the author or editor of a number of books on induction, probability, and statistical inference, as well as an editor (with Ron Loui and Greg Carlson) of Knowledge Representation and Defeasible Reasoning. He is a Fellow of the American Academy of Arts and Sciences and a Fellow of the American Association for the Advancement of Science.

Choh Man Teng is a research scientist at the Institute for Human and Machine Cognition at the University of West Florida. She is interested in uncertain reasoning and in machine learning.
Uncertain Inference

HENRY E. KYBURG, JR.
University of Rochester and
Institute for Human and Machine Cognition,
University of West Florida

and

CHOH MAN TENG
Institute for Human and Machine Cognition,
University of West Florida
## Contents

*Preface*  
Preface xi  

1 Historical Background  
1.1 Introduction 1  
1.2 Inference 1  
1.3 Roots in the Past 6  
1.4 Francis Bacon 7  
1.5 The Development of Probability 10  
1.6 John Stuart Mill 13  
1.7 G. H. von Wright 17  
1.8 Bibliographical Notes 19  
1.9 Exercises 19  
Bibliography 20  

2 First Order Logic  
2.1 Introduction 21  
2.2 Syntax 24  
2.3 Semantics 28  
2.4 W. V. O. Quine’s Mathematical Logic 30  
2.5 Arguments from Premises 33  
2.6 Limitations 34  
2.7 Summary 39  
2.8 Bibliographical Notes 39  
2.9 Exercises 39  
Bibliography 41  

3 The Probability Calculus  
3.1 Introduction 42  
3.2 Elementary Probability 44  
3.2.1 Combinations and Permutations 44  
3.2.2 The Probability Calculus 47  
3.2.3 Elementary Theorems 49
vi CONTENTS

3.3 Conditional Probability 50
   3.3.1 The Axiom of Conditional Probability 51
   3.3.2 Bayes’ Theorem 53
3.4 Probability Distributions 54
   3.4.1 Frequency Functions and Distribution Functions 55
   3.4.2 Properties of Distributions 59
3.5 Sampling Distributions 61
3.6 Useful Distributions 62
3.7 Summary 64
3.8 Bibliographical Notes 65
3.9 Exercises 65
   Bibliography 67
4 Interpretations of Probability 68
   4.1 Introduction 68
   4.2 The Classical View 68
   4.3 Empirical Interpretations of Probability 71
      4.3.1 The Limiting Frequency Interpretation 72
      4.3.2 The Propensity Interpretation 78
   4.4 Logical Interpretations of Probability 80
   4.5 Subjective Interpretations of Probability 87
      4.5.1 Dutch Book 87
      4.5.2 Conditionalization 89
   4.6 Summary 93
   4.7 Bibliographical Notes 95
   4.8 Exercises 95
   Bibliography 96
5 Nonstandard Measures of Support 98
   5.1 Support 98
   5.2 Karl Popper 99
      5.2.1 Corroboration 100
      5.2.2 Levi’s Criticism 102
   5.3 Other Measures 103
   5.4 Dempster–Shafer Belief Functions 104
      5.4.1 Belief Functions and Mass Functions 105
      5.4.2 Reduction to Sets of Probabilities 106
      5.4.3 Combining Evidence 108
      5.4.4 Special Cases 110
      5.4.5 Assessment of Belief Functions 111
   5.5 Sets of Probability Functions 113
   5.6 Summary 114
   5.7 Bibliographical Notes 114
   5.8 Exercises 115
   Bibliography 116
6 Nonmonotonic Reasoning 117
   6.1 Introduction 117
   6.2 Logic and (Non)monotonicity 117
CONTENTS

6.3 Default Logic 121
   6.3.1 Preliminaries 121
   6.3.2 Transformation of Open Default Theories 123
   6.3.3 Extensions 124
   6.3.4 Need for a Fixed Point 126
   6.3.5 Number of Extensions 127
   6.3.6 Representation 128
   6.3.7 Variants of Default Logic 131

6.4 Autoepistemic Logic 134
   6.4.1 Modal Logic 134
   6.4.2 Autoepistemic Reasoning vs Default Reasoning 136
   6.4.3 Stable Expansions 138
   6.4.4 Alternative Fixed-Point Formulation 140
   6.4.5 Groundedness 142

6.5 Circumscription 143

6.6 Unresolved Issues 146
   6.6.1 “Intuition”: Basis of Defaults 146
   6.6.2 Computational Complexity 147
   6.6.3 Multiple Extensions 147

6.7 Summary 148

6.8 Bibliographical Notes 148

6.9 Exercises 149

6.10 Bibliography 150

7 Theory Replacement 152
   7.1 Introduction 152
   7.2 Theory Change 153
      7.2.1 Expansion 153
      7.2.2 Contraction 154
      7.2.3 Revision 155
   7.3 Rationality Considerations 156
   7.4 The AGM Postulates 158
      7.4.1 Expansion 158
      7.4.2 Contraction 159
      7.4.3 Revision 161
   7.5 Connections 163
   7.6 Selecting a Contraction Function 164
   7.7 Epistemic Entrenchment 166
   7.8 Must It Be? 168
      7.8.1 Belief Bases 168
      7.8.2 Updates 169
      7.8.3 Rationality Revisited 170
      7.8.4 Iterated Change 171
   7.9 Summary 171
   7.10 Bibliographical Notes 172
   7.11 Exercises 172

Bibliography 173
viii CONTENTS

8 Statistical Inference 175
  8.1 Introduction 175
  8.2 Classical Statistics 178
    8.2.1 Significance Tests 179
    8.2.2 Hypothesis Testing 182
    8.2.3 Confidence Intervals 186
  8.3 Bayesian Statistics 192
  8.4 Summary 195
  8.5 Bibliographical Notes 197
  8.6 Exercises 197
    Bibliography 198

9 Evidential Probability 200
  9.1 Introduction 200
  9.2 Background Issues and Assumptions 201
  9.3 The Syntax of Statistical Knowledge 203
  9.4 Reference Classes and Target Classes 205
    9.4.1 Reference Formulas 205
    9.4.2 Target Formulas 208
  9.5 Prima Facie Support 209
    9.5.1 Indefinite Probabilities 210
    9.5.2 Definite Probabilities 210
  9.6 Sharpening 212
    9.6.1 Precision 213
    9.6.2 Specificity 213
    9.6.3 Richness 216
    9.6.4 Sharpens 217
  9.7 Partial Proof 219
  9.8 Extended Example 220
  9.9 A Useful Algorithm 225
  9.10 Relations to Other Interpretations 225
  9.11 Summary 226
  9.12 Bibliographical Notes 227
  9.13 Exercises 227
    Bibliography 228

10 Semantics 230
  10.1 Introduction 230
  10.2 Models and Truth 232
  10.3 Model Ratios 235
  10.4 Relevant Models 236
  10.5 Partial Validity 240
  10.6 Remarks 242
  10.7 Summary 244
  10.8 Bibliographical Notes 245
  10.9 Exercises 245
    Bibliography 245
# CONTENTS

<table>
<thead>
<tr>
<th>Number</th>
<th>Sections</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Applications</td>
<td></td>
</tr>
<tr>
<td>11.1</td>
<td>Introduction</td>
<td>247</td>
</tr>
<tr>
<td>11.2</td>
<td>Elementary Results</td>
<td>248</td>
</tr>
<tr>
<td>11.3</td>
<td>Inference from Samples</td>
<td>254</td>
</tr>
<tr>
<td>11.4</td>
<td>Example</td>
<td>259</td>
</tr>
<tr>
<td>11.5</td>
<td>Statistical Induction</td>
<td>261</td>
</tr>
<tr>
<td>11.6</td>
<td>Bayesian Induction</td>
<td>262</td>
</tr>
<tr>
<td>11.7</td>
<td>Sequences of Draws</td>
<td>264</td>
</tr>
<tr>
<td>11.8</td>
<td>Summary</td>
<td>266</td>
</tr>
<tr>
<td>11.9</td>
<td>Bibliographical Notes</td>
<td>267</td>
</tr>
<tr>
<td>11.10</td>
<td>Exercises</td>
<td>268</td>
</tr>
<tr>
<td></td>
<td>Bibliography</td>
<td>268</td>
</tr>
<tr>
<td>12</td>
<td>Scientific Inference</td>
<td></td>
</tr>
<tr>
<td>12.1</td>
<td>Introduction</td>
<td>270</td>
</tr>
<tr>
<td>12.1.1</td>
<td>Objectivity</td>
<td>270</td>
</tr>
<tr>
<td>12.1.2</td>
<td>Evidential and Practical Certainty</td>
<td>272</td>
</tr>
<tr>
<td>12.1.3</td>
<td>Statistical Inference</td>
<td>273</td>
</tr>
<tr>
<td>12.2</td>
<td>Demonstrative Induction</td>
<td>274</td>
</tr>
<tr>
<td>12.3</td>
<td>Direct Measurement</td>
<td>275</td>
</tr>
<tr>
<td>12.4</td>
<td>Indirect Measurement</td>
<td>280</td>
</tr>
<tr>
<td>12.5</td>
<td>Theory, Language, and Error</td>
<td>285</td>
</tr>
<tr>
<td>12.6</td>
<td>Summary</td>
<td>286</td>
</tr>
<tr>
<td>12.7</td>
<td>Bibliographical Notes</td>
<td>287</td>
</tr>
<tr>
<td>12.7.1</td>
<td>Measurement</td>
<td>287</td>
</tr>
<tr>
<td>12.7.2</td>
<td>Theories</td>
<td>287</td>
</tr>
<tr>
<td>12.7.3</td>
<td>Datamining</td>
<td>287</td>
</tr>
<tr>
<td>12.8</td>
<td>Exercises</td>
<td>288</td>
</tr>
<tr>
<td></td>
<td>Bibliography</td>
<td>288</td>
</tr>
<tr>
<td>11.10</td>
<td>Exercises</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Names Index</td>
<td>291</td>
</tr>
<tr>
<td></td>
<td>Index</td>
<td>293</td>
</tr>
</tbody>
</table>

© in this web service Cambridge University Press www.cambridge.org
Preface

This book is the outgrowth of an effort to provide a course covering the general topic of uncertain inference. Philosophy students have long lacked a treatment of inductive logic that was acceptable; in fact, many professional philosophers would deny that there was any such thing and would replace it with a study of probability. Yet, there seems to many to be something more traditional than the shifting sands of subjective probabilities that is worth studying. Students of computer science may encounter a wide variety of ways of treating uncertainty and uncertain inference, ranging from nonmonotonic logic to probability to belief functions to fuzzy logic. All of these approaches are discussed in their own terms, but it is rare for their relations and interconnections to be explored. Cognitive science students learn early that the processes by which people make inferences are not quite like the formal logic processes that they study in philosophy, but they often have little exposure to the variety of ideas developed in philosophy and computer science. Much of the uncertain inference of science is statistical inference, but statistics rarely enter directly into the treatment of uncertainty to which any of these three groups of students are exposed.

At what level should such a course be taught? Because a broad and interdisciplinary understanding of uncertainty seemed to be just as lacking among graduate students as among undergraduates, and because without assuming some formal background all that could be accomplished would be rather superficial, the course was developed for upper-level undergraduates and beginning graduate students in these three disciplines. The original goal was to develop a course that would serve all of these groups. It could make significant demands on ability and perseverance, and yet it would have to demand relatively little in the way of background—in part precisely because relatively little could be assumed about the common elements of the backgrounds of these diverse students. In this event, the only formal prerequisite for the course was a course in first order logic. At the University of Rochester, this is the kind of course designed to produce a certain amount of “mathematical maturity.”

The course has been taught for two years from earlier versions of this book. It is a difficult course, and students work hard at it. There are weekly exercises and a final research paper. All the material is covered, but some students find the pace very demanding. It might be that a more leisurely approach that concentrated, say, on probability and nonmonotonic acceptance, would be better for some groups of students.
xii

PREFACE

The most common suggestion is that probability theory, as well as logic, should be required as a prerequisite. At Rochester, this would make it difficult for many students to fit the course into their schedules. Furthermore, although the axiomatic foundations of probability and a few elementary theorems seem crucial to understanding uncertainty, the need for an extensive background in probability is questionable. On the other hand, because the whole topic, as we view it, is a kind of logic, a strong background in logic seems crucial.

Support for the research that has allowed us to produce this volume has come from various sources. The University of Rochester, the University of New South Wales, and the Institute for Human and Machine Cognition have each contributed time and facilities. Direct financial support has come from the National Science Foundation through award ISI-941267.

We also wish to acknowledge a debt of a different kind. Mark Wheeler was a graduate student in Philosophy at the University of Rochester many years ago when this volume was first contemplated. At that time he drafted the first chapter. Although the book has been through many revisions and even major overhauls, and has even changed authors more than once, Mark’s chapter has remained an eminently sensible introduction. He has allowed us to include it in this, our final version. We are very grateful, and herewith express our thanks.

Henry E. Kyburg, Jr.
University of Rochester and the Institute for Human and Machine Cognition,
University of West Florida

Choh Man Teng
Institute for Human and Machine Cognition,
University of West Florida