

Foundations of Empirical Beliefs

J.W. Lloyd

School of Computing
College of Systems and Society
The Australian National University

Incomplete Draft – May 12, 2026

© J.W. Lloyd, 2026

Preface

EMPIRICAL beliefs are beliefs that are acquired from observations by an agent situated in an environment. Agents use empirical beliefs to maintain a model of their environment and select actions to achieve their goals. The empirical belief base of an agent is a set of empirical beliefs. An empirical belief is a function from some space into the space of probability measures on another space; this is a conditional empirical belief. A special case is where an empirical belief is just a probability measure on a space; this is a nonconditional empirical belief. In a common case, the belief base of an agent is a single probability measure on a state space. More generally, an agent can have a large number of conditional and nonconditional empirical beliefs. This book provides a mathematical theory of empirical beliefs. In particular, it examines in detail the structure of empirical beliefs, and how to acquire, utilize, and logicize them.

Probability theory is used to model uncertainty about beliefs and provides a form of ‘degree of belief’. Belief acquisition is viewed here as stochastic filtering, but in a very general setting. Utilizing beliefs means performing computation involving the definitions of empirical beliefs. Logicizing beliefs involves constructing a theory in a modal higher-order logic that has the belief base as a model; reasoning about beliefs can then be performed by theorem proving and equational reasoning in this logic.

The main contributions are the concept of an empirical schema from which empirical beliefs are obtained; the structure that empirical schemas may have such as being a product or sum; a framework for studying the space of filters that can be nonconditional or conditional, particle, or variational, or a combination of these; the presentation of the syntax and semantics of a suitable modal higher-logic for logicizing beliefs; and computation and proof mechanisms to reason about such beliefs.

The book could be of interest to researchers in computer science, engineering, logic, or philosophy. In computer science, artificial intelligence and machine learning researchers are concerned about the problem of acquiring and utilizing a sophisticated and detailed model of the environment and other agents. This model is expressed as a collection of empirical beliefs that the agent acquires from observations, and uses to act and communicate. Thus the theoretical results of this book are directly applicable in agent applications. In addition, AI safety research could benefit from the precision and expressiveness of the theoretical formalism. In engineering, researchers in signal processing and control theory study stochastic filtering. Here, acquisition of empirical beliefs is also via stochastic filtering; however, the setting here generalizes that setting because the concept of a state distribution in stochastic filtering is a special kind of empirical belief as defined here. In logic, researchers are concerned about the use of logic for knowledge representation. Here, reasoning is carried out in an unusually expressive logic, namely, modal higher-order logic,

which admits the direct modelling of probabilistic, doxastic, and temporal aspects of empirical beliefs. Usually, modal logic is used to *analyze* agent systems; in contrast, here, modal higher-order logic is used as the language in which beliefs are represented. In philosophy, the ideas in the book could be useful to epistemologists in that they provide a precise definition of the concept of an empirical belief that has considerable generality and naturalness, and hence could be used to concretize epistemological theories. Also, the approach of stochastic filtering, used here to acquire empirical beliefs, takes a particular philosophical position on belief acquisition that would be interesting to investigate. Furthermore, the highly expressive logic in which beliefs are expressed provides opportunities for investigations in formal epistemology.

A primary goal of the theory is to use it in practical applications, and a number of novel results show promise for that. Potential applications include robotics, autonomous vehicles, home automation, smart grids, and virtual personal assistants.

The book consists of five chapters and four appendices. The chapters contain the core material. To avoid interrupting the flow of the core material, two extensive appendices contain the necessary mathematical background on measure-theoretic probability theory and logic to support the key results. The theoretical results are presented in a technically precise style. Wherever appropriate, examples and diagrams help provide the intuition behind the theoretical results.

The first chapter provides an overview of the contents of the book. The second chapter is concerned with state distributions, the prototypical kind of empirical belief. The third chapter studies the structure of empirical beliefs. The fourth chapter shows how to acquire empirical beliefs. The fifth chapter presents the logical representation of beliefs and shows how to reason with beliefs.

The first appendix gives background material on the relevant aspects of probability theory, especially regular conditional distributions and regular probability kernels, at the level of a senior undergraduate course in discrete stochastic processes. The second appendix gives the syntax and semantics of the logic, and describes how computation and proof are carried out. It also presents structural induction. The third appendix collects together all the filtering algorithms for easy reference. The fourth appendix provides some technical details about particle families.

The two main concepts of the book are those of empirical schema and empirical belief, where empirical beliefs are obtained from empirical schemas by instantiating them with the current history. Empirical schemas are sequences of regular probability kernels, the definition of which provides a criterion for the correctness of empirical schemas. Since regular probability kernels are fundamentally dependent on the concept of conditional expectation, the latter concept appears prominently throughout the theory.

Here is a summary of the main contributions of the book. Overall, the book provides a mathematical theory of empirical beliefs. Its theme is doxastic rationality, that is, the ability to acquire beliefs that capture aspects of the environment as accurately as possible given the available observations. The definitions of empirical schema and empirical belief are given that emphasize the important correctness property that the concept of a regular probability kernel provides. The practical importance of dealing with probability measures over structured spaces in the codomains of empirical beliefs is explained. Based on these definitions and the need to handle structured spaces, a theoretical account of the construction and deconstruction of empirical schemas and empirical beliefs is provided.

The recurrence equations for stochastic filtering of empirical schemas and empirical beliefs are established. Stochastic filtering is a natural method for acquiring empirical beliefs.

A suitable logic for logicizing empirical beliefs and reasoning about them is introduced. This logic is highly expressive and supports the reasoning needed for an agent to use empirical beliefs for the selection of actions. The basic theoretical results concerning the computation and proof aspects of the logic are established. Reasoning systems for the logic (in various forms) have had prototype implementations over the last 20 years. The evidence from these experiments suggests that the reasoning system presented here does seem to be feasible and practical.

The theoretical results in this book are presented with what is intended to be a high level of mathematical precision. One reason is that many of these results rely heavily on the probabilistic concepts of conditional expectation and regular probability kernel. In particular, regular probability kernels link the practical and theoretical aspects of probability. One side of the equation defining a regular probability kernel is a conditional expectation, a key theoretical concept which includes conditional probability as a special case. The other side concerns a probability kernel (or, as is more usual in applications, a conditional probability density), a key practical concept. Equating these two sides allows, for example, the derivation of recurrence equations for stochastic filters that relate empirical schemas, transition schemas, and observation schemas; these derivations are obtained by reasoning about conditional expectations. Requiring that certain probability kernels be regular formalizes their intuitive meaning and is the key to developing a rich and useful theory of empirical beliefs.

This book is a contribution to the mathematical theory of an aspect of artificial intelligence. There is now rapidly expanding coverage of all aspects of this theory. For just a few examples, at the 2022 International Congress of Mathematicians, there was a paper on the mathematics of deep neural networks [80] and a paper on the mathematics of causality [139]. Also [144] is a graduate level textbook on the mathematical foundations of deep learning models. This book can be regarded as a companion to these three and others, being about the mathematics of belief representation, acquisition, and reasoning.

One motivation for this book is the widely-quoted maxim that ‘nothing is as practical as a good theory’. Developing the theory of empirical beliefs has led already to many insights. For example, the observation that empirical schemas, transition schemas, and observation schemas are (sequences of) regular probability kernels led directly to finding the exact forms of the recurrence equations for the various cases of filtering by greatly restricting the forms that could be possible. It also led directly to the precise form of several versions of the environment synthesis and observation schema synthesis theorems. These insights have led to practical uses of the theory in this book.

I thank Kee Siong Ng who contributed significantly to the material of Chapter 5 and Appendix B.3 through a series of papers we wrote about computational logic for agents. I also thank Dawei Chen and Kee Siong Ng for many enlightening discussions about the material in the book on filtering during a collaboration on this topic that we have had over a number of years.

In its present form, this manuscript is a snapshot of an on-going research endeavour. For some sections there is still much work to be done. The biggest gaps are in the sections on logicization of empirical beliefs, point processes, intensity measures, intensity kernels, intensity filtering, and causality which are also still in early draft form. There are

also numerous minor gaps throughout the manuscript; all these gaps are flagged in blue font. My intention is to post regular updates over the next couple of years. Comments, suggestions, and corrections are greatly appreciated.

Sydney, May 2026

John Lloyd

Contents

1 Introduction	1
1.1 Rationality	1
1.2 Probability and Logic	2
1.3 Agent-environment Systems	3
1.4 Empirical Belief Structure	6
1.5 Empirical Belief Acquisition	10
1.6 Empirical Belief Reasoning	14
Bibliographical Notes	17
2 State Distributions	19
2.1 Action and Observation Processes	19
2.2 Agents and Environments	21
2.3 State Schemas	26
Bibliographical Notes	44
Exercises	45
3 Structure of Empirical Beliefs	47
3.1 Empirical Schemas and Empirical Beliefs	47
3.2 Construction of Empirical Schemas	52
3.2.1 Finite Products	52
3.2.2 Infinite Products	53
3.2.3 Sums	55
3.3 Deconstruction of Empirical Schemas	57
3.3.1 Finite Products	58
3.3.2 Infinite Products	61
3.3.3 Sums	63
3.3.4 Quotients	65
3.3.5 Deconstruction Examples	67
3.4 Representation Issues	71
3.4.1 Sets	73
3.4.2 Multisets	77
3.4.3 Sets and Multisets as Point Measures	80
3.4.4 Lists	81
3.4.5 Graphs	82
3.4.6 Quotients	86

3.4.7	Function Spaces	88
	Bibliographical Notes	93
	Exercises	93
4	Acquisition of Empirical Beliefs	95
4.1	Stochastic Filtering	95
4.2	Nonconditional Setting	100
4.2.1	Filter Recurrence Equations for the Nonconditional Setting	100
4.2.2	Environment Synthesis for the Nonconditional Setting	112
4.2.3	Sequential Bayesian Inference for the Nonconditional Setting	118
4.2.4	Simulation for the Nonconditional Setting	122
4.2.5	Codomain of Empirical Schema is Probability Measures on a Function Space	125
4.3	Constant-valued Conditional Setting	126
4.3.1	Filter Recurrence Equations for the Constant-valued Conditional Setting	127
4.3.2	Environment Synthesis for the Constant-valued Conditional Setting	137
4.3.3	Observation Schema Synthesis for the Constant-valued Conditional Setting	144
4.3.4	Simulation for the Constant-valued Conditional Setting	151
4.4	Functional Conditional Setting	158
4.4.1	Filter Recurrence Equations for the Functional Conditional Setting	159
4.4.2	Environment Synthesis for the Functional Conditional Setting	163
4.4.3	Observation Schema Synthesis for the Functional Conditional Setting	167
4.5	Conditional Setting	170
4.5.1	Filter Recurrence Equations for the Conditional Setting	171
4.5.2	Environment Synthesis for the Conditional Setting	181
4.5.3	Observation Schema Synthesis for the Conditional Setting	185
4.5.4	Sequential Bayesian Inference for the Conditional Setting	191
4.5.5	Simulation for the Conditional Setting	193
4.6	Intensity Setting	198
4.7	Distribution Filtering Algorithms	204
4.7.1	Standard Filter	207
4.7.2	Standard Particle Filter	208
4.7.3	Standard Variational Filter	214
4.7.4	Conditional Filter (Constant-valued Case)	217
4.7.5	Conditional Particle Filter (Constant-valued Case)	217
4.7.6	Conditional Variational Filter (Constant-valued Case)	223
4.7.7	Conditional Filter (General Case)	224
4.7.8	Conditional Particle Filter (General Case)	224
4.7.9	Conditional Variational Filter (General Case)	225
4.7.10	Factored Filter	225
4.7.11	Factored Particle Filter	229
4.7.12	Factored Variational Filter	239
4.7.13	Factored Conditional Filter (Constant-valued Case)	241
4.7.14	Factored Conditional Particle Filter (Constant-valued Case)	242
4.7.15	Factored Conditional Variational Filter (Constant-valued Case)	250

4.7.16 Factored Conditional Filter (General Case)	251
4.7.17 Factored Conditional Particle Filter (General Case)	251
4.7.18 Factored Conditional Variational Filter (General Case)	251
4.7.19 Interface between Nonconditional and Conditional Filters	251
4.8 Intensity Filtering Algorithms	257
4.8.1 Standard Intensity Filter	258
4.8.2 Conditional Intensity Filter (Constant-valued Case)	259
4.8.3 Conditional Intensity Filter (General Case)	260
4.8.4 Factored Intensity Filter	261
4.8.5 Factored Conditional Intensity Filter (Constant-valued Case)	262
4.8.6 Factored Conditional Intensity Filter (General Case)	264
4.8.7 Derivation of the Standard Intensity Filter	265
Bibliographical Notes	265
Exercises	268
5 Reasoning with Empirical Beliefs	271
5.1 Modal Higher-order Logic	271
5.2 Logicization of Empirical Beliefs	282
5.3 Computation Examples	284
5.4 Proof Examples	293
5.5 Computation and Proof Examples	303
5.6 Reasoning about Beliefs	307
5.7 Reasoning about Empirical Beliefs	327
5.8 Reasoning for Choosing Actions	327
Bibliographical Notes	330
Exercises	333
A Probability	335
A.1 Notation	335
A.2 Measurable Spaces	337
A.3 Measurable Functions	338
A.4 Measures	343
A.5 Kernels	347
A.6 Densities	362
A.7 Conditional Densities	367
A.8 Topological Properties	374
A.9 Random Elements	377
A.10 Conditional Expectation	382
A.11 Regular Conditional Distributions	390
A.12 Conditional Independence	413
A.13 Point Processes	422
A.14 Intensity Measures	424
A.15 Intensity Kernels	438
A.16 Finite Products of Probability Kernels	448
A.17 Infinite Products of Probability Kernels	473
A.18 Sums of Probability Kernels	486

A.19 Quotients of Probability Kernels	500
A.20 Restrictions of Probability Kernels	508
A.21 Products of Conditional Probability Densities	508
A.22 Sums of Conditional Probability Densities	519
A.23 Quotients of Conditional Probability Densities	522
A.24 Restrictions of Conditional Probability Densities	522
A.25 Causality	522
A.26 Computing Integrals	533
Bibliographical Notes	535
Exercises	535
B Logic	539
B.1 Syntax	539
B.1.1 Types	539
B.1.2 Terms	542
B.1.3 Occurrences	547
B.1.4 Substitutions	548
B.1.5 Term Replacement	552
B.1.6 α -Conversion	553
B.1.7 Composition of Substitutions	554
B.1.8 Matching	556
B.1.9 Representation of Individuals	559
B.1.10 Polymorphism	562
B.1.11 Standard Predicates	571
B.1.12 Predicate Rewrite Systems	580
B.2 Semantics	582
B.2.1 Interpretations	583
B.2.2 Denotations	585
B.2.3 Admissible Substitutions	605
B.2.4 Term Replacement and Denotations	608
B.2.5 Denotations of α -equivalent Terms	609
B.2.6 β -Reduction	611
B.2.7 Validity and Consequence	612
B.3 Reasoning	618
B.3.1 Computation	618
B.3.2 Proof	629
B.3.3 Computation and Proof Combined	642
B.3.4 Decidability and Termination	646
B.4 Structural Induction	647
B.4.1 Principle of Structural Induction	647
Bibliographical Notes	649
Exercises	651
C Filtering Algorithms	653
C.1 Algorithms for Distribution Filtering	653
C.2 Algorithms for Intensity Filtering	663

Contents xi

D Particle Families	667
D.1 Particle Families with Equal Weights	667
D.2 Particle Families in General	671
References	675
Index	685