Empirical Beliefs

J.W. Lloyd
School of Computing
College of Engineering and Computer Science
The Australian National University

Incomplete Draft – May 19, 2022

Preface

Empirical beliefs are beliefs that are acquired from observations by an agent situated in an environment. Agents use empirical beliefs to track reality. This manuscript is a draft of a mathematical theory of empirical beliefs. In particular, it examines in detail the structure of empirical beliefs, and how to acquire and utilize them.

The account here of empirical beliefs is probabilistic and modal. Probability theory is used to model uncertainty about beliefs and provides a form of 'degree of belief'. Modal operators provide doxastic and temporal aspects of beliefs. The main contributions are the introduction of the concept of a schema from which empirical beliefs are obtained, the ability to acquire from observations beliefs that are conditional distributions, the sophistication of the representation language for empirical beliefs, and the ability 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 two 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 probability 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 probability kernels and regular conditional distributions. 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 two main concepts of the book are those of schema and empirical belief, where empirical beliefs are obtained from schemas by instantiating them with the current history. Schemas are sequences of regular conditional distributions, the definition of which provides a criterion for the correctness of schemas. Since regular conditional distributions are primarily 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 schema and empirical belief are given that emphasize the important correctness property that the concept of a regular conditional distribution 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 schemas and empirical beliefs is provided. The recurrence equations for stochastic filtering of 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.

In addition to a systematic theoretical account of empirical beliefs, contributions of note are the following:

1. The definitions of schema and empirical belief. (Definitions 3.1.1 and 3.1.3).

2. The results concerning the construction and deconstruction of schemas and empirical beliefs. (Propositions 3.2.1 to 3.3.5 and 3.3.1 to 3.3.4.)

3. The filter recurrence equations for schemas and empirical beliefs in the conditional case. (Propositions 4.2.3 and 4.2.11.)

4. The results showing that the environment can be synthesized from a schema and the transition and observation models for the schema in the conditional case. (Propositions 4.2.4 and 4.2.13.)

5. The observation model synthesis results. (Propositions 4.2.5 and 4.2.12.)

6. Bayesian inference is a special case of filtering. (Propositions 4.1.7 and 4.2.9.)

7. The algorithm for a conditional particle filter. (Figures 4.20 and 4.21.)

8. The algorithm for a factored conditional particle filter. (Figures 4.31 and 4.32.)

9. The definition of the denotation of a modal term. (Part 6 of Definition B.2.10.)

10. The definitions of computation of rank 0 (Definition B.3.1), proof of rank 0 (Definition B.3.4), computation of rank k (Definition B.3.6), and proof of rank k (Definition B.3.7).

Kee Siong Ng contributed significantly to this book through a series of papers that we wrote on the material of Chapter 5 and Appendix B.3. The results of a collaboration with Dawei Chen, Samuel Yang-Zhao, and Kee Siong Ng on a more extensive account of filtering algorithms than is currently presented in Chapter 4 and their application to modelling epidemic processes will appear elsewhere.

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. My intention is to post regular updates over the next couple of years. Comments, suggestions, and corrections are greatly appreciated. Finally, notwithstanding the earlier experimental work on reasoning in the logic and the more recent filtering experiments mentioned above, there remains much more experimental work that needs to be done to demonstrate that the machinery for acquiring and utilizing empirical beliefs proposed here pays sufficient dividends in terms of intelligent behaviour of agents. I would be pleased to hear from anyone interested in pursuing such experimental work.

Sydney, May 2022 John Lloyd

Contents