

Preface

Artificial Intelligence (AI) systems simulating human behavior are often called *intelligent agents*. By definition, these intelligent agents exhibit some form of human-like intelligence. For example, intelligent agents can learn from the environment, make plans and decisions, react to the environment with appropriate actions, express emotions, or revise beliefs. Intelligent agents typically represent human cognitive states using underlying beliefs and knowledge modeled in a knowledge representation language. We use the term *epistemic state* to refer to an actual or possible cognitive state that drives human behavior at a given point in time; the accurate determination (or estimation) of these epistemic states is crucial to an agent's ability to correctly simulate human behavior.

This book provides three fundamental and generic approaches (logical, probabilistic, and modal) to representing and reasoning with agent epistemic states, specifically in the context of decision making. In addition, the book introduces a formal integration of these three approaches into a single unified approach we call $\mathcal{P3}$ (Propositional, Probabilistic, and Possible World), that combines the advantages of the other approaches. Each of these approaches can be applied to create the foundation for intelligent, decision-making software agents. Each approach is symbolic in nature (unlike the sub-symbolic neural network approach, for example), yielding agent "thought processes" that are sufficiently transparent to provide users with understandable explanations of agent "reasoning." The classical logic and probability theory thrust of our approach naturally excludes from our consideration other potential symbolic approaches to decision-making such as fuzzy logics. The "symbolic argumentation" approach ($\mathcal{P3}$) to decision making combines logic and probability, and therefore offers several advantages over the traditional approach to decision-making based on simple rule-based expert systems or expected utility theory.

The generic logical, probabilistic, and modal reasoning techniques that we discuss in the book, such as logical deduction and evidence propagation, are applied to various decision-making problems that an agent will encounter in various circumstances. However, the book does not focus on particular applications of these reasoning techniques to such problems as planning,

learning, or belief revision. Neither does the book focus on communication and collaboration within agent societies. Such applications require extensive treatments in themselves and so are beyond the scope of the book; our purpose here is to provide the formal foundation necessary to support decision-making.

This book is divided into four parts by the four reasoning approaches we develop. The first part focuses on logic-based modeling of an epistemic state as a set of propositions expressed by sentences from classical propositional and first-order logics. The second part takes a probabilistic approach to modeling epistemic states. In this approach, states are represented by a probability measure defined over some space of events, which are represented as random variables. The third part focuses on extending logic-based modeling with modalities to yield various systems of modal logics, including epistemic logics. These logics are based on mental constructs that represent information and are suitable for reasoning with possible human cognitive states. Finally, the fourth part describes the combination of these three rather disparate approaches into the single integrated $\mathcal{P3}$ approach within an argumentation framework especially appropriate for simulating human decision making.

The purpose of this book is not to carry out an elaborate philosophical discussion on epistemic state modeling, but rather to provide readers with various practical ways of modeling and reasoning with agent epistemic states used in decision making. Therefore, the resolution-based logic programming paradigm is introduced under logic-based modeling; Bayesian belief networks are introduced as a way of modeling and reasoning with epistemic states where beliefs are represented by probability measures defined over some state space; and modal resolution schemes are introduced under modal logic-based modeling and reasoning with actual and possible worlds.

Frameworks specific to decision making, such as the logic-based production rule or belief-network-based influence diagrams that incorporate the notions of action and utility, are also detailed in the book. The use of logic-based reasoning (especially resolution theorem proving) and probabilistic reasoning (especially Bayesian belief networks) is widespread when implementing agents capable of making decisions, generating plans, revising beliefs, learning, and so on. But the book goes a step further by bridging the gap between these two approaches, augmenting them with informational mental constructs, such as beliefs and knowledge, and then producing a coherent approach that culminates in the penultimate chapter of the book in the form of a symbolic argumentation approach for weighing the pros and cons of decision-making options.

In the book I have tried to balance discussions on theory and practice. As for theory, I have stated results and developed proofs wherever necessary to provide sufficient theoretical foundations for the concepts and procedures introduced. For example, if I have introduced a specific resolution-based theorem-proving procedure, then I also provide the associated soundness and completeness of the procedure with respect to the stated semantics. On the practical side, I have provided detailed algorithms that can be encoded in computer programs in a straightforward manner to incorporate aspects of human behavior into intelligent decision-making agents. For example, an implementation of the evidence propagation algorithms for belief networks and influence diagrams can be incorporated into agents that make decisions within a probabilistic framework.

The decision-making problem is present in every industry; finance, medicine, and defense are a few of the big ones. Computer professionals and researchers in the decision science community within commerce and industry will find the book most useful for building intelligent and practical decision-aiding agents or, more generally, for embedding intelligence in various systems based on sound logical and probabilistic reasoning; the material is especially suited for building safety-critical decision aiding applications based on well-founded theories. This book may also be used as an AI textbook on logic and probability for undergraduate and graduate courses, and as a reference book for researchers in universities.

The chapters in the book and their dependencies are shown in the figure below. Readers are advised not to skip Chapter 1, which includes an informal overview of the rest of the book. A basic understanding of this chapter will prevent the reader from becoming lost later on. The beginner should not be dismayed by the use of unfamiliar terminology in this first chapter, as these terms will be explained in later chapters.

Readers who have been exposed to the concepts of basic mathematics and probability may omit Chapter 2, which provides a background on mathematical preliminaries. Chapters 3 and 4, on classical logics and logic programming, should be read in sequence. Chapter 6, on Bayesian belief networks, can be read independently of Chapters 3 and 4, but Chapter 8, on modal logics, requires an understanding of classical logic as presented in Chapter 3. Chapters 3, 4, 6, and 8 provide the semantic foundations of our three problem modeling approaches, namely, propositional, probabilistic, and possible world. Because production rules are a special form of rules in logic programming with embedded degrees of uncertainty I recommend to reading at least the logic programming section in Chapter 4 and the background material on probability theory before reading

Chapter 5. Chapter 7, on influence diagrams, requires understanding the basics of Bayesian belief networks as it extends this understanding with the concepts of action and utility. The final chapter (Chapter 9) on symbolic argumentation requires an overall understanding of the contents of Chapters 3-8. Ultimately, three different decision-making frameworks are provided, in Chapters 5, 7, and 9.

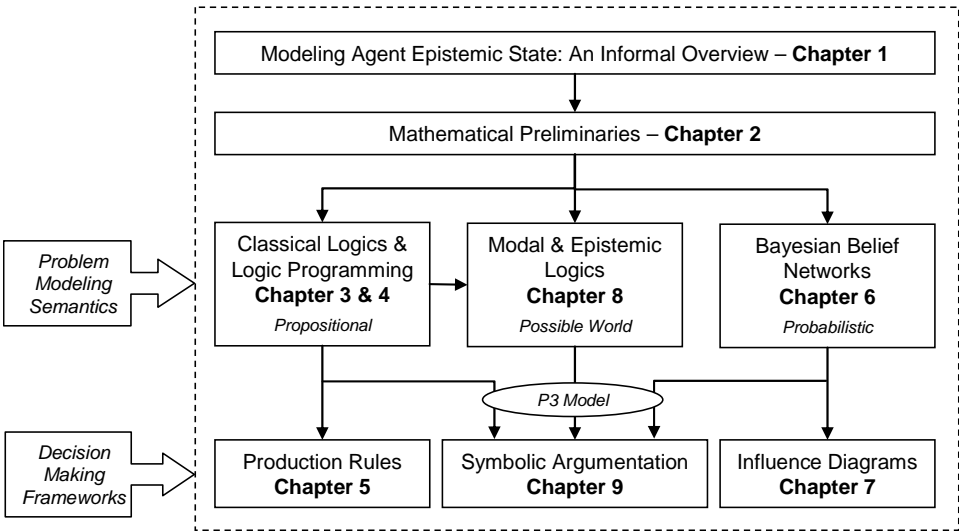


Figure 1-1: Chapters and their dependencies

There is another suggested way to follow this book for those readers who, with some exposure to mathematical logic and the theory of probability, are looking for some well-founded computational decision support technologies to build deployable systems. They can skip detailed theoretical materials as presented in some chapters (e.g. 3, 4, and 8), and focus more on the computational aspects (e.g. chapter 5, 6, 7, and 9).

My sincere thanks go to my wife Janique, my son Sébastien, and my daughter Kabita, for their love, patience and inspiration throughout the preparation of this book (My children’s perception of the book are very different from each other. Sébastien’s view is very philosophical, which he manifested in his drawing for the title page of a sage from the Museum of Fine Arts in Boston. However, Kabita is more pragmatic, expressed herself through a dog’s decision-making dilemma as depicted below.) Special thanks go to Karen DeSimone for her careful proofreading of the manuscript that enormously enhanced readability and enriched the content of the book. Thanks are due to my colleagues, Dave

Lawless and Josh Introne, with whom I have had many technical discussions on various aspects of this book. Thanks also to all of my colleagues at Charles River Analytics, especially Alex Feinman, Paul Gonsalves, Partha Kanjilal, Rich Reposa, and Greg Zacharias, and Avi Pfeffer at Harvard University, for their part in creating a stimulating intellectual environment which inspired me to write this book. Thanks to the World Scientific/Imperial College Press, especially Tjan Kwang Wei, for their help in producing the book from the beginning. Finally, I thank my parents, brothers, sisters and other family members back in one of several thousands of small villages in India for patiently accepting my absence and showing their encouragement and support through their many phone calls.

This book is based on the work of many researchers, and I express my gratitude to those who have directly or indirectly contributed so much to the fields of deductive logic, probability theory, and artificial intelligence. I have made my best effort to make this book informative, readable, and free from mistakes, and I would welcome any criticism or suggestion for improvement.



Subrata Das
Cambridge, MA
March 2007