PREFACE

---

*What we have to learn to do
we learn by doing. . .*

—ARISTOTLE, *Ethics*

Welcome to the Sixth Edition!

I was very pleased to be asked to produce the sixth edition of my artificial intelligence book. It is a compliment to the earlier editions, started over twenty years ago, that our approach to AI has been so highly valued. It is also exciting that, as new development in the field emerges, we are able to present much of it in each new edition. We thank our many readers, colleagues, and students for keeping our topics relevant and our presentation up to date.

Many sections of the earlier editions have endured remarkably well, including the presentation of logic, search algorithms, knowledge representation, production systems, machine learning, and, in the supplementary materials, the programming techniques developed in Lisp, Prolog, and with this edition, Java. These remain central to the practice of artificial intelligence, and a constant in this new edition.

This book remains accessible. We introduce key representation techniques including logic, semantic and connectionist networks, graphical models, and many more. Our search algorithms are presented clearly, first in pseudocode, and then in the supplementary materials, many of them are implemented in Prolog, Lisp, and/or Java. It is expected that the motivated students can take our core implementations and extend them to new exciting applications.

We created, for the sixth edition, a new machine learning chapter based on stochastic methods (Chapter 13). We feel that the stochastic technology is having an increasingly larger impact on AI, especially in areas such as diagnostic and prognostic reasoning, natural language analysis, robotics, and machine learning. To support these emerging technologies we have expanded the presentation of Bayes' theorem, Markov models, Bayesian

---

belief networks, and related graphical models. Our expansion includes greater use of probabilistic finite state machines, hidden Markov models, and dynamic programming with the Earley parser and implementing the Viterbi algorithm. Other topics, such as emergent computation, ontologies, stochastic parsing algorithms, that were treated cursorily in earlier editions, have grown sufficiently in importance to merit a more complete discussion. The changes for the sixth edition reflect emerging artificial intelligence research questions and are evidence of the continued vitality of our field.

As the scope of our AI project grew, we have been sustained by the support of our publisher, editors, friends, colleagues, and, most of all, by our readers, who have given our work such a long and productive life. We remain excited at the writing opportunity we are afforded: Scientists are rarely encouraged to look up from their own, narrow research interests and chart the larger trajectories of their chosen field. Our readers have asked us to do just that. We are grateful to them for this opportunity. We are also encouraged that our earlier editions have been used in AI communities worldwide and translated into a number of languages including German, Polish, Portuguese, Russian, and two dialects of Chinese!

Although artificial intelligence, like most engineering disciplines, must justify itself to the world of commerce by providing solutions to practical problems, we entered the field of AI for the same reasons as many of our colleagues and students: we want to understand and explore the mechanisms of mind that enable intelligent thought and action. We reject the rather provincial notion that intelligence is an exclusive ability of humans, and believe that we can effectively investigate the space of possible intelligences by designing and evaluating intelligent artifacts. Although the course of our careers has given us no cause to change these commitments, we have arrived at a greater appreciation for the scope, complexity, and audacity of this undertaking. In the preface to our earlier editions, we outlined three assertions that we believed distinguished our approach to teaching artificial intelligence. It is reasonable, in writing a preface to the present edition, to return to these themes and see how they have endured as our field has grown.

The first of these goals was *to unify the diverse branches of AI through a detailed discussion of its theoretical foundations*. At the time we first adopted that goal, it seemed that the main problem was in reconciling researchers who emphasized the careful statement and analysis of formal theories of intelligence (the *neats*) with those who believed that intelligence itself was some sort of grand hack that could be best approached in an application-driven, *ad hoc* manner (the *scruffies*). That dichotomy has proven far too simple.

In contemporary AI, debates between neats and scruffies have given way to dozens of other debates between proponents of physical symbol systems and students of neural networks, between logicians and designers of artificial life forms that evolve in a most illogical manner, between architects of expert systems and case-based reasoners, and finally, between those who believe artificial intelligence has already been achieved and those who believe it will never happen. Our original image of AI as frontier science where outlaws, prospectors, wild-eyed prairie prophets and other dreamers were being slowly tamed by the disciplines of formalism and empiricism has given way to a different metaphor: that of a large, chaotic but mostly peaceful city, where orderly bourgeois neighborhoods draw their vitality from diverse, chaotic, bohemian districts. Over the years that we have devoted to the different editions of this book, a compelling picture of the architecture of intelligence has started to emerge from this city's structure, art, and industry.

Intelligence is too complex to be described by any single theory; instead, researchers are constructing a hierarchy of theories that characterize it at multiple levels of abstraction. At the lowest levels of this hierarchy, neural networks, genetic algorithms and other forms of emergent computation have enabled us to understand the processes of adaptation, perception, embodiment, and interaction with the physical world that must underlie any form of intelligent activity. Through some still partially understood resolution, this chaotic population of blind and primitive actors gives rise to the cooler patterns of logical inference. Working at this higher level, logicians have built on Aristotle's gift, tracing the outlines of deduction, abduction, induction, truth-maintenance, and countless other modes and manners of reason. At even higher levels of abstraction, designers of diagnostic systems, intelligent agents, and natural language understanding programs have come to recognize the role of social processes in creating, transmitting, and sustaining knowledge.

At this point in the AI enterprise it looks as though the extremes of rationalism and empiricism have only led to limited results. Both extremes suffer from limited applicability and generalization. The author takes a third view, that the empiricist's conditioning: semantic nets, scripts, subsumption architectures and the rationalist's clear and distinct ideas: predicate calculus, non-monotonic logics, automated reasoning - suggest a third viewpoint, the Bayesian. The experience of relational invariances conditions intelligent agents's expectations, and learning these invariances, in turn, bias future expectations. As philosophers we are charged to critique the epistemological validity of the AI enterprise. For this task, in Chapter 16 we discuss the rationalist project, the empiricists dilemma, and propose a Bayesian based constructivist rapprochement. In this sixth edition, we touch on all these levels in the presenting the AI enterprise.

The second commitment we made in earlier editions was to *the central position of advanced representational formalisms and search techniques in AI methodology*. This is, perhaps, the most controversial aspect of our previous editions and of much early work in AI, with many researchers in emergent computation questioning whether symbolic reasoning and referential semantics have any role at all in intelligence. Although the idea of representation as giving names to things has been challenged by the implicit representation provided by the emerging patterns of a neural network or an artificial life, we believe that an understanding of representation and search remains essential to any serious practitioner of artificial intelligence. We also feel that our Chapter 1 overview of the historical traditions and precursors of AI are critical components of AI education. Furthermore, these are invaluable tools for analyzing such aspects of non-symbolic AI as the expressive power of a neural network or the progression of candidate problem solutions through the fitness landscape of a genetic algorithm. Comparisons, contrasts, and a critique of modern AI are offered in Chapter 16.

Our third commitment was made at the beginning of this book's life cycle: to place artificial intelligence within the context of empirical science. In the spirit of the Newell and Simon (1976) Turing award lecture we quote from an earlier edition:

> ... AI is not some strange aberration from the scientific tradition, but . . . part of a general quest for knowledge about, and the understanding of, intelligence itself. Furthermore, our AI programming tools, along with the exploratory programming methodology . . . are ideal for exploring an environment. Our tools give us a medium for both understanding

and questions. We come to appreciate and know phenomena constructively, that is, by progressive approximation.

Thus we see each design and program as an experiment with nature: we propose a representation, we generate a search algorithm, and we question the adequacy of our characterization to account for part of the phenomenon of intelligence. And the natural world gives a response to our query. Our experiment can be deconstructed, revised, extended, and run again. Our model can be refined, our understanding extended.

New with The Sixth Edition

The biggest change for the sixth edition is the extension of the stochastic approaches to AI. To accomplish this we revised Section 9.3 and added a new chapter (13) introducing probability-based machine learning. Our presentation of stochastic AI tools and their application to learning and natural language is now more comprehensive.

From probability theory's foundations in set theory and counting we develop the notions of probabilities, random variables, and independence. We present and use Bayes' theorem first with one symptom and one disease and then in its full general form. We examine the hypotheses that underlie the use of Bayes and then present the *argmax* and *naive Bayes* approaches. We present examples of stochastic reasoning, including the analysis of language phenomena and the Vierbi algorithm. We also introduce the idea of conditional independence that leads to Bayesian belief networks, the BBN, in Chapter 9.

In Chapter 13 we introduce hidden Markov models, HMMs, and show their use in several examples. We also present several HMM variants, including the auto-regressive and hierarchical HMMs. We present dynamic Bayesian networks, DBNs, and demonstrate their use. We discuss parameter and structure learning and present the expectation maximization algorithm and demonstrate its use with loopy belief propagation. Finally, we present Markov decision processes, the MDP, and partially observable Markov decision process, the POMDP, in the context of an extension to the earlier presentation of reinforcement learning.

We include several more examples of *probabilistic finite state machines* and *probabilistic acceptors*, as well as the use of *dynamic programming*, especially with stochastic measures (the *Viterbi algorithm*). We added a stochastic English language parser (based on the work of Mark Steedman at the University of Edinburgh) as well as the use of dynamic programming with the Earley parser.

We made a major decision to remove the Prolog and Lisp chapters from the book. Part of the reason for this is that these were getting too large. We have also accumulated a number of AI algorithms written in Java. When we added the new Chapter 13 on stochastic approaches to machine learning, we determined that the book was getting too large/cumbersome. Thus the sixth edition is more than 150 pages smaller than the fifth and the AI algorithms in Prolog, Lisp, and Java are being released as supplementary materials. From our earliest days in AI we have always felt that the way to understand the power (and limitations) of AI algorithms is constructively - that is, by building them! We encourage our present generation of readers to do exactly this: to visit the supplementary materials: to build and experiment directly with the algorithms we present.

Finally, we have done the usual updating of references and materials that a new edition warrants. In a revised Chapter 16, we return to the deeper questions on the nature of intelligence and the possibility of creating intelligent machines.

Sixth Edition: The Contents

Chapter 1 introduces artificial intelligence, beginning with a brief history of attempts to understand mind and intelligence in philosophy, psychology, and other areas of research. In an important sense, AI is an old science, tracing its roots back at least to Aristotle. An appreciation of this background is essential for an understanding of the issues addressed in modern research. We also present an overview of some of the important application areas in AI. Our goal in Chapter 1 is to provide both background and a motivation for the theory and applications that follow.

Chapters 2, 3, 4, 5, and 6 (Part II) introduce the research tools for AI problem solving. These include, in Chapter 2, the predicate calculus presented both as a mathematical system as well as a representation language to describe the essential features of a problem. Search, and the algorithms and data structures used to implement search, are introduced in Chapter 3, to organize the exploration of problem situations. In Chapter 4, we discuss the essential role of heuristics in focusing and constraining search-based problem solving. In Chapter 5, we introduce the stochastic methodology, important technology for reasoning in situations of uncertainty. In Chapter 6, we present a number of software architectures, including the blackboard and production system, for implementing these search algorithms.

Chapters 7, 8, and 9 make up Part III: representations for AI, knowledge-intensive problem solving, and reasoning in changing and ambiguous situations. In Chapter 7 we present the evolving story of AI representational schemes. We begin with a discussion of association-based networks and extend this model to include conceptual dependency theory, frames, and scripts. We then present an in-depth examination of a particular formalism, conceptual graphs, emphasizing the epistemological issues involved in representing knowledge and showing how these issues are addressed in a modern representation language. Expanding on this formalism in Chapter 14, we show how conceptual graphs can be used to implement a natural language database front end. We conclude Chapter 7 with more modern approaches to representation, including Copycat and agent-oriented architectures.

Chapter 8 presents the rule-based expert system along with case-based and model-based reasoning, including examples from the NASA space program. These approaches to problem solving are presented as a natural evolution of the material in Part II: using a production system of predicate calculus expressions to orchestrate a graph search. We end with an analysis of the strengths and weaknesses of each of these approaches to knowledge-intensive problem solving.

Chapter 9 presents models for reasoning with uncertainty as well as the use of unreliable information. We introduce Bayesian models, belief networks, Dempster-Shafer, causal models, and the Stanford certainty algebra for reasoning in uncertain situations.

Chapter 9 also contains algorithms for truth maintenance, reasoning with minimum models, logic-based abduction, and the clique-tree algorithm for Bayesian belief networks.

Part IV, Chapters 10 through 13, is an extensive presentation of issues in machine learning. In Chapter 10 we offer a detailed look at algorithms for symbol-based learning, a fruitful area of research spawning a number of different problems and solution approaches. These learning algorithms vary in their goals, the training data considered, their learning strategies, and the knowledge representations they employ. Symbol-based learning includes induction, concept learning, version-space search, and ID3. The role of inductive bias is considered, generalizations from patterns of data, as well as the effective use of knowledge to learn from a single example in explanation-based learning. Category learning, or conceptual clustering, is presented with unsupervised learning. Reinforcement learning, or the ability to integrate feedback from the environment into a policy for making new decisions concludes the chapter.

In Chapter 11 we present neural networks, often referred to as sub-symbolic or connectionist models of learning. In a neural net, information is implicit in the organization and weights on a set of connected processors, and learning involves a re-arrangement and modification of the overall weighting of nodes and structure of the system. We present a number of connectionist architectures, including perceptron learning, backpropagation, and counterpropagation. We demonstrate Kohonen, Grossberg, and Hebbian models. We present associative learning as well as attractor models, including examples of Hopfield networks.

Genetic algorithms and evolutionary approaches to learning are introduced in Chapter 12. On this viewpoint, learning is cast as an emerging and adaptive process. After several examples of problem solutions based on genetic algorithms, we introduce the application of genetic techniques to more general problem solvers. These include classifier systems and genetic programming. We then describe society-based learning with examples from artificial life, called *a-life*, research. We conclude the chapter with an example of emergent computation from research at the Santa Fe Institute.

Chapter 13 presents stochastic approaches to machine learning. We begin with a definition of hidden markov models and then present several important variations including the auto-regressive and hierarchical HMM. We then present dynamic Bayesian networks, a generalization of the HMM, and also able to track systems across periods of time. These techniques are useful for modeling the changes in complex environments as is required for diagnostic and prognostic reasoning. Finally, we add a probabilistic component to reinforcement learning first introduced in Chapter 10. This includes presentation of the Markov decision process (or MDP) and the partially observed Markov decision process (or POMDP).

Part V, Chapters 14 and 15, presents automated reasoning and natural language understanding. Theorem proving, often referred to as automated reasoning, is one of the oldest areas of AI research. In Chapter 14, we discuss the first programs in this area, including the Logic Theorist and the General Problem Solver. The primary focus of the chapter is binary resolution proof procedures, especially resolution refutations. More advanced inferencing with hyper-resolution and paramodulation is also presented. Finally, we describe the Prolog interpreter as a Horn clause and resolution-based inferencing system, and see Prolog computing as an instance of the logic programming paradigm.

Chapter 15 presents natural language understanding. Our traditional approach to language understanding, exemplified by many of the semantic structures presented in Chapter 7, is complemented with the stochastic approach. These include using Markov models, CART trees, CHART parsing (the Earley algorithm), mutual information clustering, and statistics-based parsing. The chapter concludes with examples applying natural language techniques to database query generation, a text summarization systems well as.the use of machine learning to generalize extracted results from the WWW.

Finally, Chapter 16 serves as an epilogue for the book. It addresses the issue of the possibility of a science of intelligent systems, and considers contemporary challenges to AI; it discusses AI's current limitations, and projects its exciting future.

Using This Book

Artificial intelligence is a big field, and consequently, this is a large book. Although it would require more than a single semester to cover all of the material offered, we have designed our book so that a number of paths may be taken through the material. By selecting subsets of the material, we have used this text for single semester and full year (two semester) courses.

We assume that most students will have had introductory courses in discrete mathematics, including predicate calculus, set theory, counting, and graph theory. If this is not true, the instructor should spend more time on these concepts in the "optional" sections at the beginning of the introductory chapters (2.1, 3.1, and 5.1). We also assume that students have had courses in data structures including trees, graphs, and recursion-based search, using stacks, queues, and priority queues. If they have not, then spend more time on the beginning sections of Chapters 3, 4, and 6.

In a one quarter or one semester course, we go quickly through the first two parts of the book. With this preparation, students are able to appreciate the material in Part III. We then consider the Prolog, Lisp, or the Java code in the supplementary materials for the book and require students to build many of the representation and search techniques of the second part of the book. Alternatively, one of the languages, Prolog, for example, can be introduced early in the course and be used to test out the data structures and search techniques as that are encountered. We feel the meta-interpreters presented in the language materials are very helpful for building rule-based and other knowledge-intensive problem solvers. Prolog, Lisp, and Java are excellent tools for building natural language understanding and learning systems; these architectures are presented in Parts II and III and there are examples of them in the supplementary course materials.

In a two-semester or three-quarter course, we are able to cover the application areas of Parts IV and V, especially the machine learning chapters, in appropriate detail. We also expect a much more detailed programming project from students. We also think that it is very important in the second semester for students to revisit many of the primary sources of the AI literature. It is crucial for students to see both where we are in the evolution of the AI enterprise, as well as how we got here, and to have an appreciation of the future promises of artificial intelligence. We use materials from the WWW for this purpose or select a collected set of readings, such as, *Computation and Intelligence* (Luger 1995).

The algorithms of our book are described using a Pascal-like pseudo-code. This notation uses the control structures of Pascal along with English descriptions of the tests and operations. We have added two useful constructs to the Pascal control structures. The first is a modified case statement that, rather than comparing the value of a variable with constant case labels, as in standard Pascal, lets each item be labeled with an arbitrary boolean test. The case evaluates these tests in order until one of them is true and then performs the associated action; all other actions are ignored. Those familiar with Lisp will note that this has the same semantics as the Lisp cond statement.

The other addition to our pseudo code language is a return statement which takes one argument and can appear anywhere within a procedure or function. When the return is encountered, it causes the program to immediately exit the function, returning its argument as a result. Other than these modifications we used Pascal structure, with a reliance on the English descriptions, to make the algorithms clear.

Supplemental Material Available

The sixth edition has an ***attached web site*** maintained by my graduate students. This site, built originally by two UNM students, Alejandro CdeBaca and Cheng Liu, includes supplementary ideas for most chapters, some sample problems with their solutions, and ideas for student projects. Besides the ***Prolog, Lisp, and Java programs*** in the supplementary materials for this book, we have included many other AI algorithms in Java and C++ on the web site. Students are welcome to use these and supplement them with their own comments, code, and critiques. The web url is www.cs.unm.edu/~luger/ai-final/.

The ***Prolog, Lisp, and Java programs*** implementing many of the AI data structures and search algorithms of the book are available through your Addison-Wesley Pearson Education representative. There is also an ***Instructor's Guide*** available which has many of the book's exercises worked out, several practice tests with solutions, a sample syllabus, and ideas supporting teaching the material. There are also a full set of ***PowerPoint presentation materials*** for use by instructors adopting this book. Again, consult your local A-W Pearson representative for access and visit www.aw.com/luger.

My e-mail address is luger@cs.unm.edu, and I enjoy hearing from my readers.

Acknowledgements

Although I am the sole author of the sixth edition, this book has always been the product of my efforts as Professor of Computer Science, Psychology, and Linguistics at the University of New Mexico along with my fellow faculty, professional colleagues, graduate students, and friends. The sixth edition is also the product of the many readers that have e-mailed comments, corrections, and suggestions. The book will continue this way, reflecting a *community* effort; consequently, I will continue using the prepositions *we, our,* and *us* when presenting material.

I thank Bill Stubblefield, the co-author for the first three editions, for more than twenty years of contributions, but even more importantly, for his friendship. I also thank

Artificial intelligence is an exciting and rewarding discipline; may you enjoy your study as you come to appreciate its power and challenges.

George Luger
1 January 2008
Albuquerque