



# Artificial Intelligence

A Modern Approach

*Third Edition*



**PRENTICE HALL SERIES  
IN ARTIFICIAL INTELLIGENCE**  
*Stuart Russell and Peter Norvig, Editors*

FORSYTH & PONCE  
GRAHAM  
JURAFSKY & MARTIN  
NEAPOLITAN  
RUSSELL & NORVIG

*Computer Vision: A Modern Approach*  
*ANSI Common Lisp*  
*Speech and Language Processing, 2nd ed.*  
*Learning Bayesian Networks*  
*Artificial Intelligence: A Modern Approach, 3rd ed.*

# Artificial Intelligence

## A Modern Approach

### *Third Edition*

Stuart J. Russell and Peter Norvig

*Contributing writers:*

Ernest Davis  
Douglas D. Edwards  
David Forsyth  
Nicholas J. Hay  
Jitendra M. Malik  
Vibhu Mittal  
Mehran Sahami  
Sebastian Thrun

**Prentice Hall**

Upper Saddle River   Boston   Columbus   San Francisco   New York  
Indianapolis   London   Toronto   Sydney   Singapore   Tokyo   Montreal  
Dubai   Madrid   Hong Kong   Mexico City   Munich   Paris   Amsterdam   Cape Town

Vice President and Editorial Director, ECS: Marcia J. Horton  
Editor-in-Chief: Michael Hirsch  
Executive Editor: Tracy Dunkelberger  
Assistant Editor: Melinda Haggerty  
Editorial Assistant: Allison Michael  
Vice President, Production: Vince O'Brien  
Senior Managing Editor: Scott Disanno  
Production Editor: Jane Bonnell  
Senior Operations Supervisor: Alan Fischer  
Operations Specialist: Lisa McDowell  
Marketing Manager: Erin Davis  
Marketing Assistant: Mack Patterson  
Cover Designers: Kirsten Sims and Geoffrey Cassar  
Cover Images: Stan Honda/Getty, Library of Congress, NASA, National Museum of Rome,  
Peter Norvig, Ian Parker, Shutterstock, Time Life/Getty  
Interior Designers: Stuart Russell and Peter Norvig  
Copy Editor: Mary Lou Nohr  
Art Editor: Greg Dulles  
Media Editor: Daniel Sandin  
Media Project Manager: Danielle Leone

**Copyright © 2010, 2003, 1995 by Pearson Education, Inc.,  
Upper Saddle River, New Jersey 07458.**

All rights reserved. Manufactured in the United States of America. This publication is protected by Copyright and permissions should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. To obtain permission(s) to use materials from this work, please submit a written request to Pearson Higher Education, Permissions Department, 1 Lake Street, Upper Saddle River, NJ 07458.

The author and publisher of this book have used their best efforts in preparing this book. These efforts include the development, research, and testing of the theories and programs to determine their effectiveness. The author and publisher make no warranty of any kind, expressed or implied, with regard to these programs or the documentation contained in this book. The author and publisher shall not be liable in any event for incidental or consequential damages in connection with, or arising out of, the furnishing, performance, or use of these programs.

**Library of Congress Cataloging-in-Publication Data on File**

**Prentice Hall**  
is an imprint of



[www.pearsonhighered.com](http://www.pearsonhighered.com)

10 9 8 7 6 5 4 3 2 1  
ISBN-13: 978-0-13-604259-4  
ISBN-10: 0-13-604259-7

*For Loy, Gordon, Lucy, George, and Isaac — S.J.R.*

*For Kris, Isabella, and Juliet — P.N.*

*This page intentionally left blank*

# Preface

**Artificial Intelligence** (AI) is a big field, and this is a big book. We have tried to explore the full breadth of the field, which encompasses logic, probability, and continuous mathematics; perception, reasoning, learning, and action; and everything from microelectronic devices to robotic planetary explorers. The book is also big because we go into some depth.

The subtitle of this book is “A Modern Approach.” The intended meaning of this rather empty phrase is that we have tried to synthesize what is now known into a common framework, rather than trying to explain each subfield of AI in its own historical context. We apologize to those whose subfields are, as a result, less recognizable.

## New to this edition

This edition captures the changes in AI that have taken place since the last edition in 2003. There have been important applications of AI technology, such as the widespread deployment of practical speech recognition, machine translation, autonomous vehicles, and household robotics. There have been algorithmic landmarks, such as the solution of the game of checkers. And there has been a great deal of theoretical progress, particularly in areas such as probabilistic reasoning, machine learning, and computer vision. Most important from our point of view is the continued evolution in how we think about the field, and thus how we organize the book. The major changes are as follows:

- We place more emphasis on partially observable and nondeterministic environments, especially in the nonprobabilistic settings of search and planning. The concepts of *belief state* (a set of possible worlds) and *state estimation* (maintaining the belief state) are introduced in these settings; later in the book, we add probabilities.
- In addition to discussing the types of environments and types of agents, we now cover in more depth the types of *representations* that an agent can use. We distinguish among *atomic* representations (in which each state of the world is treated as a black box), *factored* representations (in which a state is a set of attribute/value pairs), and *structured* representations (in which the world consists of objects and relations between them).
- Our coverage of planning goes into more depth on contingent planning in partially observable environments and includes a new approach to hierarchical planning.
- We have added new material on first-order probabilistic models, including *open-universe* models for cases where there is uncertainty as to what objects exist.
- We have completely rewritten the introductory machine-learning chapter, stressing a wider variety of more modern learning algorithms and placing them on a firmer theoretical footing.
- We have expanded coverage of Web search and information extraction, and of techniques for learning from very large data sets.
- 20% of the citations in this edition are to works published after 2003.
- We estimate that about 20% of the material is brand new. The remaining 80% reflects older work but has been largely rewritten to present a more unified picture of the field.

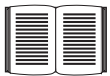
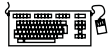


## Overview of the book

The main unifying theme is the idea of an **intelligent agent**. We define AI as the study of agents that receive percepts from the environment and perform actions. Each such agent implements a function that maps percept sequences to actions, and we cover different ways to represent these functions, such as reactive agents, real-time planners, and decision-theoretic systems. We explain the role of learning as extending the reach of the designer into unknown environments, and we show how that role constrains agent design, favoring explicit knowledge representation and reasoning. We treat robotics and vision not as independently defined problems, but as occurring in the service of achieving goals. We stress the importance of the task environment in determining the appropriate agent design.

Our primary aim is to convey the *ideas* that have emerged over the past fifty years of AI research and the past two millennia of related work. We have tried to avoid excessive formality in the presentation of these ideas while retaining precision. We have included pseudocode algorithms to make the key ideas concrete; our pseudocode is described in Appendix B.

This book is primarily intended for use in an undergraduate course or course sequence. The book has 27 chapters, each requiring about a week's worth of lectures, so working through the whole book requires a two-semester sequence. A one-semester course can use selected chapters to suit the interests of the instructor and students. The book can also be used in a graduate-level course (perhaps with the addition of some of the primary sources suggested in the bibliographical notes). Sample syllabi are available at the book's Web site, [aima.cs.berkeley.edu](http://aima.cs.berkeley.edu). The only prerequisite is familiarity with basic concepts of computer science (algorithms, data structures, complexity) at a sophomore level. Freshman calculus and linear algebra are useful for some of the topics; the required mathematical background is supplied in Appendix A.



NEW TERM

Exercises are given at the end of each chapter. Exercises requiring significant programming are marked with a **keyboard** icon. These exercises can best be solved by taking advantage of the code repository at [aima.cs.berkeley.edu](http://aima.cs.berkeley.edu). Some of them are large enough to be considered term projects. A number of exercises require some investigation of the literature; these are marked with a **book** icon.

Throughout the book, important points are marked with a *pointing* icon. We have included an extensive index of around 6,000 items to make it easy to find things in the book. Wherever a **new term** is first defined, it is also marked in the margin.

## About the Web site

[aima.cs.berkeley.edu](http://aima.cs.berkeley.edu), the Web site for the book, contains

- implementations of the algorithms in the book in several programming languages,
- a list of over 1000 schools that have used the book, many with links to online course materials and syllabi,
- an annotated list of over 800 links to sites around the Web with useful AI content,
- a chapter-by-chapter list of supplementary material and links,
- instructions on how to join a discussion group for the book,

- instructions on how to contact the authors with questions or comments,
- instructions on how to report errors in the book, in the likely event that some exist, and
- slides and other materials for instructors.

## About the cover

The cover depicts the final position from the decisive game 6 of the 1997 match between chess champion Garry Kasparov and program DEEP BLUE. Kasparov, playing Black, was forced to resign, making this the first time a computer had beaten a world champion in a chess match. Kasparov is shown at the top. To his left is the Asimo humanoid robot and to his right is Thomas Bayes (1702–1761), whose ideas about probability as a measure of belief underlie much of modern AI technology. Below that we see a Mars Exploration Rover, a robot that landed on Mars in 2004 and has been exploring the planet ever since. To the right is Alan Turing (1912–1954), whose fundamental work defined the fields of computer science in general and artificial intelligence in particular. At the bottom is Shakey (1966–1972), the first robot to combine perception, world-modeling, planning, and learning. With Shakey is project leader Charles Rosen (1917–2002). At the bottom right is Aristotle (384 B.C.–322 B.C.), who pioneered the study of logic; his work was state of the art until the 19th century (copy of a bust by Lysippos). At the bottom left, lightly screened behind the authors' names, is a planning algorithm by Aristotle from *De Motu Animalium* in the original Greek. Behind the title is a portion of the CPSC Bayesian network for medical diagnosis (Pradhan *et al.*, 1994). Behind the chess board is part of a Bayesian logic model for detecting nuclear explosions from seismic signals.

Credits: Stan Honda/Getty (Kasparaov), Library of Congress (Bayes), NASA (Mars rover), National Museum of Rome (Aristotle), Peter Norvig (book), Ian Parker (Berkeley skyline), Shutterstock (Asimo, Chess pieces), Time Life/Getty (Shakey, Turing).

## Acknowledgments

This book would not have been possible without the many contributors whose names did not make it to the cover. Jitendra Malik and David Forsyth wrote Chapter 24 (computer vision) and Sebastian Thrun wrote Chapter 25 (robotics). Vibhu Mittal wrote part of Chapter 22 (natural language). Nick Hay, Mehran Sahami, and Ernest Davis wrote some of the exercises. Zoran Duric (George Mason), Thomas C. Henderson (Utah), Leon Reznik (RIT), Michael Gourley (Central Oklahoma) and Ernest Davis (NYU) reviewed the manuscript and made helpful suggestions. We thank Ernie Davis in particular for his tireless ability to read multiple drafts and help improve the book. Nick Hay whipped the bibliography into shape and on deadline stayed up to 5:30 AM writing code to make the book better. Jon Barron formatted and improved the diagrams in this edition, while Tim Huang, Mark Paskin, and Cynthia Bruyns helped with diagrams and algorithms in previous editions. Ravi Mohan and Ciaran O'Reilly wrote and maintain the Java code examples on the Web site. John Canny wrote the robotics chapter for the first edition and Douglas Edwards researched the historical notes. Tracy Dunkelberger, Allison Michael, Scott Disanno, and Jane Bonnell at Pearson tried their best to keep us on schedule and made many helpful suggestions. Most helpful of all has

been Julie Sussman, P.P.A., who read every chapter and provided extensive improvements. In previous editions we had proofreaders who would tell us when we left out a comma and said *which* when we meant *that*; Julie told us when we left out a minus sign and said  $x_i$  when we meant  $x_j$ . For every typo or confusing explanation that remains in the book, rest assured that Julie has fixed at least five. She persevered even when a power failure forced her to work by lantern light rather than LCD glow.

**Stuart would like to thank** his parents for their support and encouragement and his wife, Loy Sheflott, for her endless patience and boundless wisdom. He hopes that Gordon, Lucy, George, and Isaac will soon be reading this book after they have forgiven him for working so long on it. RUGS (Russell's Unusual Group of Students) have been unusually helpful, as always.

**Peter would like to thank** his parents (Torsten and Gerda) for getting him started, and his wife (Kris), children (Bella and Juliet), colleagues, and friends for encouraging and tolerating him through the long hours of writing and longer hours of rewriting.

**We both thank** the librarians at Berkeley, Stanford, and NASA and the developers of CiteSeer, Wikipedia, and Google, who have revolutionized the way we do research. We can't acknowledge all the people who have used the book and made suggestions, but we would like to note the especially helpful comments of Gagan Aggarwal, Eyal Amir, Ion Androutsopoulos, Krzysztof Apt, Warren Haley Armstrong, Ellery Aziel, Jeff Van Baalen, Darius Bacon, Brian Baker, Shumeet Baluja, Don Barker, Tony Barrett, James Newton Bass, Don Beal, Howard Beck, Wolfgang Bibel, John Binder, Larry Bookman, David R. Boxall, Ronen Brafman, John Bresina, Gerhard Brewka, Selmer Bringsjord, Carla Brodley, Chris Brown, Emma Brunskill, Wilhelm Burger, Lauren Burka, Carlos Bustamante, Joao Cachopo, Murray Campbell, Norman Carver, Emmanuel Castro, Anil Chakravarthy, Dan Chisarick, Berthe Choueiry, Roberto Cipolla, David Cohen, James Coleman, Julie Ann Comparini, Corinna Cortes, Gary Cottrell, Ernest Davis, Tom Dean, Rina Dechter, Tom Dietterich, Peter Drake, Chuck Dyer, Doug Edwards, Robert Egginton, Asma'a El-Budrawy, Barbara Engelhardt, Kutluhan Erol, Oren Etzioni, Hana Filip, Douglas Fisher, Jeffrey Forbes, Ken Ford, Eric Fosler-Lussier, John Fosler, Jeremy Frank, Alex Franz, Bob Futrelle, Marek Galecki, Stefan Gerberding, Stuart Gill, Sabine Glesner, Seth Golub, Gosta Grahne, Russ Greiner, Eric Grimson, Barbara Grosz, Larry Hall, Steve Hanks, Othar Hansson, Ernst Heinz, Jim Hendler, Christoph Herrmann, Paul Hilfinger, Robert Holte, Vasant Honavar, Tim Huang, Seth Hutchinson, Joost Jacob, Mark Jelasity, Magnus Johansson, Istvan Jonyer, Dan Jurafsky, Leslie Kaelbling, Keiji Kanazawa, Surekha Kasibhatla, Simon Kasif, Henry Kautz, Gernot Kerschbaumer, Max Khesin, Richard Kirby, Dan Klein, Kevin Knight, Roland Koenig, Sven Koenig, Daphne Koller, Rich Korf, Benjamin Kuipers, James Kurien, John Lafferty, John Laird, Gus Larsen, John Lazzaro, Jon LeBlanc, Jason Leatherman, Frank Lee, Jon Lehto, Edward Lim, Phil Long, Pierre Louveaux, Don Loveland, Sridhar Mahadevan, Tony Mancill, Jim Martin, Andy Mayer, John McCarthy, David McGrane, Jay Mendelsohn, Risto Miikkulanien, Brian Milch, Steve Minton, Vibhu Mittal, Mehryar Mohri, Leora Morgenstern, Stephen Muggleton, Kevin Murphy, Ron Musick, Sung Myaeng, Eric Nadeau, Lee Naish, Pandu Nayak, Bernhard Nebel, Stuart Nelson, XuanLong Nguyen, Nils Nilsson, Illah Nourbakhsh, Ali Nouri, Arthur Nunes-Harwitt, Steve Omohundro, David Page, David Palmer, David Parkes, Ron Parr, Mark

Paskin, Tony Passera, Amit Patel, Michael Pazzani, Fernando Pereira, Joseph Perla, Wim Pijls, Ira Pohl, Martha Pollack, David Poole, Bruce Porter, Malcolm Pradhan, Bill Pringle, Lorraine Prior, Greg Provan, William Rapaport, Deepak Ravichandran, Ioannis Refanidis, Philip Resnik, Francesca Rossi, Sam Roweis, Richard Russell, Jonathan Schaeffer, Richard Scherl, Hinrich Schuetze, Lars Schuster, Bart Selman, Soheil Shams, Stuart Shapiro, Jude Shavlik, Yoram Singer, Satinder Singh, Daniel Sleator, David Smith, Bryan So, Robert Sproull, Lynn Stein, Larry Stephens, Andreas Stolcke, Paul Stradling, Devika Subramanian, Marek Suchenek, Rich Sutton, Jonathan Tash, Austin Tate, Bas Terwijn, Olivier Teytaud, Michael Thielscher, William Thompson, Sebastian Thrun, Eric Tiedemann, Mark Torrance, Randall Upham, Paul Utgoff, Peter van Beek, Hal Varian, Paulina Varshavskaya, Sunil Vemuri, Vandiver Verma, Ubbo Visser, Jim Waldo, Toby Walsh, Bonnie Webber, Dan Weld, Michael Wellman, Kamin Whitehouse, Michael Dean White, Brian Williams, David Wolfe, Jason Wolfe, Bill Woods, Alden Wright, Jay Yagnik, Mark Yasuda, Richard Yen, Eliezer Yudkowsky, Weixiong Zhang, Ming Zhao, Shlomo Zilberstein, and our esteemed colleague Anonymous Reviewer.

# About the Authors

**Stuart Russell** was born in 1962 in Portsmouth, England. He received his B.A. with first-class honours in physics from Oxford University in 1982, and his Ph.D. in computer science from Stanford in 1986. He then joined the faculty of the University of California at Berkeley, where he is a professor of computer science, director of the Center for Intelligent Systems, and holder of the Smith–Zadeh Chair in Engineering. In 1990, he received the Presidential Young Investigator Award of the National Science Foundation, and in 1995 he was cowinner of the Computers and Thought Award. He was a 1996 Miller Professor of the University of California and was appointed to a Chancellor’s Professorship in 2000. In 1998, he gave the Forsythe Memorial Lectures at Stanford University. He is a Fellow and former Executive Council member of the American Association for Artificial Intelligence. He has published over 100 papers on a wide range of topics in artificial intelligence. His other books include *The Use of Knowledge in Analogy and Induction* and (with Eric Wefald) *Do the Right Thing: Studies in Limited Rationality*.

**Peter Norvig** is currently Director of Research at Google, Inc., and was the director responsible for the core Web search algorithms from 2002 to 2005. He is a Fellow of the American Association for Artificial Intelligence and the Association for Computing Machinery. Previously, he was head of the Computational Sciences Division at NASA Ames Research Center, where he oversaw NASA’s research and development in artificial intelligence and robotics, and chief scientist at Junglee, where he helped develop one of the first Internet information extraction services. He received a B.S. in applied mathematics from Brown University and a Ph.D. in computer science from the University of California at Berkeley. He received the Distinguished Alumni and Engineering Innovation awards from Berkeley and the Exceptional Achievement Medal from NASA. He has been a professor at the University of Southern California and a research faculty member at Berkeley. His other books are *Paradigms of AI Programming: Case Studies in Common Lisp* and *VerbMobil: A Translation System for Face-to-Face Dialog* and *Intelligent Help Systems for UNIX*.

# Contents

## I Artificial Intelligence

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	What Is AI? . . . . .	1
1.2	The Foundations of Artificial Intelligence . . . . .	5
1.3	The History of Artificial Intelligence . . . . .	16
1.4	The State of the Art . . . . .	28
1.5	Summary, Bibliographical and Historical Notes, Exercises . . . . .	29
<b>2</b>	<b>Intelligent Agents</b>	<b>34</b>
2.1	Agents and Environments . . . . .	34
2.2	Good Behavior: The Concept of Rationality . . . . .	36
2.3	The Nature of Environments . . . . .	40
2.4	The Structure of Agents . . . . .	46
2.5	Summary, Bibliographical and Historical Notes, Exercises . . . . .	59

## II Problem-solving

<b>3</b>	<b>Solving Problems by Searching</b>	<b>64</b>
3.1	Problem-Solving Agents . . . . .	64
3.2	Example Problems . . . . .	69
3.3	Searching for Solutions . . . . .	75
3.4	Uninformed Search Strategies . . . . .	81
3.5	Informed (Heuristic) Search Strategies . . . . .	92
3.6	Heuristic Functions . . . . .	102
3.7	Summary, Bibliographical and Historical Notes, Exercises . . . . .	108
<b>4</b>	<b>Beyond Classical Search</b>	<b>120</b>
4.1	Local Search Algorithms and Optimization Problems . . . . .	120
4.2	Local Search in Continuous Spaces . . . . .	129
4.3	Searching with Nondeterministic Actions . . . . .	133
4.4	Searching with Partial Observations . . . . .	138
4.5	Online Search Agents and Unknown Environments . . . . .	147
4.6	Summary, Bibliographical and Historical Notes, Exercises . . . . .	153
<b>5</b>	<b>Adversarial Search</b>	<b>161</b>
5.1	Games . . . . .	161
5.2	Optimal Decisions in Games . . . . .	163
5.3	Alpha–Beta Pruning . . . . .	167
5.4	Imperfect Real-Time Decisions . . . . .	171
5.5	Stochastic Games . . . . .	177

5.6	Partially Observable Games . . . . .	180
5.7	State-of-the-Art Game Programs . . . . .	185
5.8	Alternative Approaches . . . . .	187
5.9	Summary, Bibliographical and Historical Notes, Exercises . . . . .	189
<b>6</b>	<b>Constraint Satisfaction Problems</b>	<b>202</b>
6.1	Defining Constraint Satisfaction Problems . . . . .	202
6.2	Constraint Propagation: Inference in CSPs . . . . .	208
6.3	Backtracking Search for CSPs . . . . .	214
6.4	Local Search for CSPs . . . . .	220
6.5	The Structure of Problems . . . . .	222
6.6	Summary, Bibliographical and Historical Notes, Exercises . . . . .	227
<b>III</b>	<b>Knowledge, reasoning, and planning</b>	
<b>7</b>	<b>Logical Agents</b>	<b>234</b>
7.1	Knowledge-Based Agents . . . . .	235
7.2	The Wumpus World . . . . .	236
7.3	Logic . . . . .	240
7.4	Propositional Logic: A Very Simple Logic . . . . .	243
7.5	Propositional Theorem Proving . . . . .	249
7.6	Effective Propositional Model Checking . . . . .	259
7.7	Agents Based on Propositional Logic . . . . .	265
7.8	Summary, Bibliographical and Historical Notes, Exercises . . . . .	274
<b>8</b>	<b>First-Order Logic</b>	<b>285</b>
8.1	Representation Revisited . . . . .	285
8.2	Syntax and Semantics of First-Order Logic . . . . .	290
8.3	Using First-Order Logic . . . . .	300
8.4	Knowledge Engineering in First-Order Logic . . . . .	307
8.5	Summary, Bibliographical and Historical Notes, Exercises . . . . .	313
<b>9</b>	<b>Inference in First-Order Logic</b>	<b>322</b>
9.1	Propositional vs. First-Order Inference . . . . .	322
9.2	Unification and Lifting . . . . .	325
9.3	Forward Chaining . . . . .	330
9.4	Backward Chaining . . . . .	337
9.5	Resolution . . . . .	345
9.6	Summary, Bibliographical and Historical Notes, Exercises . . . . .	357
<b>10</b>	<b>Classical Planning</b>	<b>366</b>
10.1	Definition of Classical Planning . . . . .	366
10.2	Algorithms for Planning as State-Space Search . . . . .	373
10.3	Planning Graphs . . . . .	379

10.4	Other Classical Planning Approaches . . . . .	387
10.5	Analysis of Planning Approaches . . . . .	392
10.6	Summary, Bibliographical and Historical Notes, Exercises . . . . .	393
<b>11</b>	<b>Planning and Acting in the Real World</b>	<b>401</b>
11.1	Time, Schedules, and Resources . . . . .	401
11.2	Hierarchical Planning . . . . .	406
11.3	Planning and Acting in Nondeterministic Domains . . . . .	415
11.4	Multiagent Planning . . . . .	425
11.5	Summary, Bibliographical and Historical Notes, Exercises . . . . .	430
<b>12</b>	<b>Knowledge Representation</b>	<b>437</b>
12.1	Ontological Engineering . . . . .	437
12.2	Categories and Objects . . . . .	440
12.3	Events . . . . .	446
12.4	Mental Events and Mental Objects . . . . .	450
12.5	Reasoning Systems for Categories . . . . .	453
12.6	Reasoning with Default Information . . . . .	458
12.7	The Internet Shopping World . . . . .	462
12.8	Summary, Bibliographical and Historical Notes, Exercises . . . . .	467
<b>IV</b>	<b>Uncertain knowledge and reasoning</b>	
<b>13</b>	<b>Quantifying Uncertainty</b>	<b>480</b>
13.1	Acting under Uncertainty . . . . .	480
13.2	Basic Probability Notation . . . . .	483
13.3	Inference Using Full Joint Distributions . . . . .	490
13.4	Independence . . . . .	494
13.5	Bayes' Rule and Its Use . . . . .	495
13.6	The Wumpus World Revisited . . . . .	499
13.7	Summary, Bibliographical and Historical Notes, Exercises . . . . .	503
<b>14</b>	<b>Probabilistic Reasoning</b>	<b>510</b>
14.1	Representing Knowledge in an Uncertain Domain . . . . .	510
14.2	The Semantics of Bayesian Networks . . . . .	513
14.3	Efficient Representation of Conditional Distributions . . . . .	518
14.4	Exact Inference in Bayesian Networks . . . . .	522
14.5	Approximate Inference in Bayesian Networks . . . . .	530
14.6	Relational and First-Order Probability Models . . . . .	539
14.7	Other Approaches to Uncertain Reasoning . . . . .	546
14.8	Summary, Bibliographical and Historical Notes, Exercises . . . . .	551
<b>15</b>	<b>Probabilistic Reasoning over Time</b>	<b>566</b>
15.1	Time and Uncertainty . . . . .	566



15.2	Inference in Temporal Models . . . . .	570
15.3	Hidden Markov Models . . . . .	578
15.4	Kalman Filters . . . . .	584
15.5	Dynamic Bayesian Networks . . . . .	590
15.6	Keeping Track of Many Objects . . . . .	599
15.7	Summary, Bibliographical and Historical Notes, Exercises . . . . .	603
<b>16</b>	<b>Making Simple Decisions</b>	<b>610</b>
16.1	Combining Beliefs and Desires under Uncertainty . . . . .	610
16.2	The Basis of Utility Theory . . . . .	611
16.3	Utility Functions . . . . .	615
16.4	Multiattribute Utility Functions . . . . .	622
16.5	Decision Networks . . . . .	626
16.6	The Value of Information . . . . .	628
16.7	Decision-Theoretic Expert Systems . . . . .	633
16.8	Summary, Bibliographical and Historical Notes, Exercises . . . . .	636
<b>17</b>	<b>Making Complex Decisions</b>	<b>645</b>
17.1	Sequential Decision Problems . . . . .	645
17.2	Value Iteration . . . . .	652
17.3	Policy Iteration . . . . .	656
17.4	Partially Observable MDPs . . . . .	658
17.5	Decisions with Multiple Agents: Game Theory . . . . .	666
17.6	Mechanism Design . . . . .	679
17.7	Summary, Bibliographical and Historical Notes, Exercises . . . . .	684
<b>V</b>	<b>Learning</b>	
<b>18</b>	<b>Learning from Examples</b>	<b>693</b>
18.1	Forms of Learning . . . . .	693
18.2	Supervised Learning . . . . .	695
18.3	Learning Decision Trees . . . . .	697
18.4	Evaluating and Choosing the Best Hypothesis . . . . .	708
18.5	The Theory of Learning . . . . .	713
18.6	Regression and Classification with Linear Models . . . . .	717
18.7	Artificial Neural Networks . . . . .	727
18.8	Nonparametric Models . . . . .	737
18.9	Support Vector Machines . . . . .	744
18.10	Ensemble Learning . . . . .	748
18.11	Practical Machine Learning . . . . .	753
18.12	Summary, Bibliographical and Historical Notes, Exercises . . . . .	757
<b>19</b>	<b>Knowledge in Learning</b>	<b>768</b>
19.1	A Logical Formulation of Learning . . . . .	768

19.2	Knowledge in Learning . . . . .	777
19.3	Explanation-Based Learning . . . . .	780
19.4	Learning Using Relevance Information . . . . .	784
19.5	Inductive Logic Programming . . . . .	788
19.6	Summary, Bibliographical and Historical Notes, Exercises . . . . .	797
<b>20</b>	<b>Learning Probabilistic Models</b>	<b>802</b>
20.1	Statistical Learning . . . . .	802
20.2	Learning with Complete Data . . . . .	806
20.3	Learning with Hidden Variables: The EM Algorithm . . . . .	816
20.4	Summary, Bibliographical and Historical Notes, Exercises . . . . .	825
<b>21</b>	<b>Reinforcement Learning</b>	<b>830</b>
21.1	Introduction . . . . .	830
21.2	Passive Reinforcement Learning . . . . .	832
21.3	Active Reinforcement Learning . . . . .	839
21.4	Generalization in Reinforcement Learning . . . . .	845
21.5	Policy Search . . . . .	848
21.6	Applications of Reinforcement Learning . . . . .	850
21.7	Summary, Bibliographical and Historical Notes, Exercises . . . . .	853
<b>VI</b>	<b>Communicating, perceiving, and acting</b>	
<b>22</b>	<b>Natural Language Processing</b>	<b>860</b>
22.1	Language Models . . . . .	860
22.2	Text Classification . . . . .	865
22.3	Information Retrieval . . . . .	867
22.4	Information Extraction . . . . .	873
22.5	Summary, Bibliographical and Historical Notes, Exercises . . . . .	882
<b>23</b>	<b>Natural Language for Communication</b>	<b>888</b>
23.1	Phrase Structure Grammars . . . . .	888
23.2	Syntactic Analysis (Parsing) . . . . .	892
23.3	Augmented Grammars and Semantic Interpretation . . . . .	897
23.4	Machine Translation . . . . .	907
23.5	Speech Recognition . . . . .	912
23.6	Summary, Bibliographical and Historical Notes, Exercises . . . . .	918
<b>24</b>	<b>Perception</b>	<b>928</b>
24.1	Image Formation . . . . .	929
24.2	Early Image-Processing Operations . . . . .	935
24.3	Object Recognition by Appearance . . . . .	942
24.4	Reconstructing the 3D World . . . . .	947
24.5	Object Recognition from Structural Information . . . . .	957

24.6	Using Vision . . . . .	961
24.7	Summary, Bibliographical and Historical Notes, Exercises . . . . .	965
<b>25</b>	<b>Robotics</b>	<b>971</b>
25.1	Introduction . . . . .	971
25.2	Robot Hardware . . . . .	973
25.3	Robotic Perception . . . . .	978
25.4	Planning to Move . . . . .	986
25.5	Planning Uncertain Movements . . . . .	993
25.6	Moving . . . . .	997
25.7	Robotic Software Architectures . . . . .	1003
25.8	Application Domains . . . . .	1006
25.9	Summary, Bibliographical and Historical Notes, Exercises . . . . .	1010
<b>VII</b>	<b>Conclusions</b>	
<b>26</b>	<b>Philosophical Foundations</b>	<b>1020</b>
26.1	Weak AI: Can Machines Act Intelligently? . . . . .	1020
26.2	Strong AI: Can Machines Really Think? . . . . .	1026
26.3	The Ethics and Risks of Developing Artificial Intelligence . . . . .	1034
26.4	Summary, Bibliographical and Historical Notes, Exercises . . . . .	1040
<b>27</b>	<b>AI: The Present and Future</b>	<b>1044</b>
27.1	Agent Components . . . . .	1044
27.2	Agent Architectures . . . . .	1047
27.3	Are We Going in the Right Direction? . . . . .	1049
27.4	What If AI Does Succeed? . . . . .	1051
<b>A</b>	<b>Mathematical background</b>	<b>1053</b>
A.1	Complexity Analysis and $O()$ Notation . . . . .	1053
A.2	Vectors, Matrices, and Linear Algebra . . . . .	1055
A.3	Probability Distributions . . . . .	1057
<b>B</b>	<b>Notes on Languages and Algorithms</b>	<b>1060</b>
B.1	Defining Languages with Backus–Naur Form (BNF) . . . . .	1060
B.2	Describing Algorithms with Pseudocode . . . . .	1061
B.3	Online Help . . . . .	1062
	<b>Bibliography</b>	<b>1063</b>
	<b>Index</b>	<b>1095</b>

# 1

## INTRODUCTION

*In which we try to explain why we consider artificial intelligence to be a subject most worthy of study, and in which we try to decide what exactly it is, this being a good thing to decide before embarking.*

INTELLIGENCE

We call ourselves *Homo sapiens*—man the wise—because our **intelligence** is so important to us. For thousands of years, we have tried to understand *how we think*; that is, how a mere handful of matter can perceive, understand, predict, and manipulate a world far larger and more complicated than itself. The field of **artificial intelligence**, or AI, goes further still: it attempts not just to understand but also to *build* intelligent entities.

ARTIFICIAL  
INTELLIGENCE

AI is one of the newest fields in science and engineering. Work started in earnest soon after World War II, and the name itself was coined in 1956. Along with molecular biology, AI is regularly cited as the “field I would most like to be in” by scientists in other disciplines. A student in physics might reasonably feel that all the good ideas have already been taken by Galileo, Newton, Einstein, and the rest. AI, on the other hand, still has openings for several full-time Einsteins and Edisons.

AI currently encompasses a huge variety of subfields, ranging from the general (learning and perception) to the specific, such as playing chess, proving mathematical theorems, writing poetry, driving a car on a crowded street, and diagnosing diseases. AI is relevant to any intellectual task; it is truly a universal field.

### 1.1 WHAT IS AI?

---

RATIONALITY

We have claimed that AI is exciting, but we have not said what it *is*. In Figure 1.1 we see eight definitions of AI, laid out along two dimensions. The definitions on top are concerned with *thought processes* and *reasoning*, whereas the ones on the bottom address *behavior*. The definitions on the left measure success in terms of fidelity to *human* performance, whereas the ones on the right measure against an *ideal* performance measure, called **rationality**. A system is rational if it does the “right thing,” given what it knows.

Historically, all four approaches to AI have been followed, each by different people with different methods. A human-centered approach must be in part an empirical science, in-