※ 本书被全世界89个国家的900多所大学用作教材

大学计算机教育国外著名教材系列（影印版）

PEARSON Prentice Hall

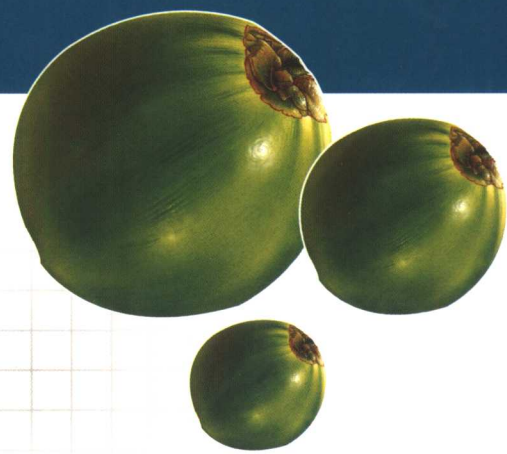
ARTIFICIAL INTELLIGENCE

A MODERN APPROACH

SECOND EDITION

人工智能

—种现代的方法（第2版）



Stuart J. Russell
Peter Norvig 著

PEARSON Education

清华大学出版社

English reprint edition copyright © 2006 by **PEARSON EDUCATION ASIA LIMITED and TSINGHUA UNIVERSITY PRESS.**
Original English language title from Proprietor's edition of the Work.

Original English language title: Artificial Intelligence: A Modern Approach, Second Edition **by** Stuart J. Russell, Peter Norvig, Copyright © 2003
All Rights Reserved.

Published by arrangement with the original publisher, Pearson Education, Inc., publishing as Prentice Hall, Inc.

This edition is authorized for sale and distribution only in the People's Republic of China (excluding the Special Administrative Region of Hong Kong, Macao SAR and Taiwan).
本书影印版由 Pearson Education（培生教育出版集团）授权给清华大学出版社出版发行。

For sale and distribution in the People's Republic of China exclusively (except Taiwan, Hong Kong SAR and Macao SAR).
仅限于中华人民共和国境内(不包括中国香港、澳门特别行政区和中国台湾地区)销售发行。

北京市版权局著作权合同登记号　图字 01-2003-7188

版权所有，翻印必究。举报电话：010-62782989　13501256678　13801310933
本书封面贴有 Pearson Education(培生教育出版集团)激光防伪标签，无标签者不得销售。

图书在版编目(CIP)数据

人工智能：一种现代的方法 = Artificial Intelligence: A Modern Approach: 第 2 版 / （ ）罗素（Russell, S.），（ ）诺维格（Norvig, P.）著. —影印本. —北京：清华大学出版社，2006.5
（大学计算机教育国外著名教材系列）
ISBN 7-302-12829-4

Ⅰ. 人…　Ⅱ. ①罗… ②诺…　Ⅲ. 人工智能－高等学校－教材－英文　Ⅳ. TP18

中国版本图书馆 CIP 数据核字（2006）第 033483 号

出 版 者：清华大学出版社　　地　　址：北京清华大学学研大厦
　　http://www.tup.com.cn　　邮　　编：100084
　　社总机：010-6277 0175　　客户服务：010-6277 6969
印 刷 者：清华大学印刷厂
装 订 者：三河市化甲屯小学装订二厂
发 行 者：新华书店总店北京发行所
开　　本：185×230　印张：69.75
版　　次：2006 年 5 月第 2 版　2006 年 5 月第 1 次印刷
书　　号：ISBN 7-302-12829-4/TP・8162
印　　数：1~2500
定　　价：128.00 元

出 版 说 明

进入 21 世纪，世界各国的经济、科技以及综合国力的竞争将更加激烈。竞争的中心无疑是对人才的竞争。谁拥有大量高素质的人才，谁就能在竞争中取得优势。高等教育，作为培养高素质人才的事业，必然受到高度重视。目前我国高等教育的教材更新较慢，为了加快教材的更新频率，教育部正在大力促进我国高校采用国外原版教材。

清华大学出版社从 1996 年开始，与国外著名出版公司合作，影印出版了“大学计算机教育丛书（影印版）”等一系列引进图书，受到国内读者的欢迎和支持。跨入 21 世纪，我们本着为我国高等教育教材建设服务的初衷，在已有的基础上，进一步扩大选题内容，改变图书开本尺寸，一如既往地请有关专家挑选适用于我国高校本科及研究生计算机教育的国外经典教材或著名教材，组成本套“大学计算机教育国外著名教材系列（影印版）”，以飨读者。深切期盼读者及时将使用本系列教材的效果和意见反馈给我们。更希望国内专家、教授积极向我们推荐国外计算机教育的优秀教材，以利我们把“大学计算机教育国外著名教材系列（影印版）”做得更好，更适合高校师生的需要。

清华大学出版社

序 言

《人工智能：一种现代方法》（英文，第 1 版）是一部人工智能（AI）的优秀教材，1995 年问世后在世界各地出版、发行，很快成为一部畅销书。在我国，英文影印本于 2002 年出版，也很受大众的欢迎。

事隔 8 年(2003 年)，该书的第 2 版（作者 Stuart Russell, Peter Norvig， Pearson Education 出版集团出版）又出现在我们面前。作者是这样解释出版新书的原因，他说，自 1995 年该书第 1 版发行以来，AI 有了很大的变化，它的技术更趋实用，因此新书的每一章都经过重写以反映该领域的最新成就，同时重新解释了原有的结果，使之更加符合新的发现。这充分反映了作者对新书的负责精神与严肃态度。

这部书的主要特点如下。

（1）在智能 Agent（自主体，代理，行为者）的概念下，将 AI 中相互分离的领域统一起来，克服了以往 AI 教材中难以避免的内容零散且互不相关的现象，从而使 AI 变得更加理论化、系统化。

（2）理论与实际并重。作者在论述各个领域的原理与方法时，尽量运用数学（形式化）的语言，力图让它们建立在严格的理论基础之上。同时又介绍最新的实用算法，特别是能够解决现实世界问题的方法，尽量使 AI 从“玩具世界”中走向实用。

本书所具有的以上特点正好反映了 AI 当前发展的大趋势。大家知道，人工智能从上个世纪中叶诞生以来，一直未能形成系统的理论体系。因此，有的人把 AI 看成是一门“工程”，有的则认为是一门“技术”，也有的甚至认为只是一门“艺术”。大家也许记得，上个世纪 80 年代，以斯坦福大学的 N. J. Nilsson 为代表与以耶鲁大学 R. C. Schank 为代表，曾经展开过一场关于 AI 究竟是一门“工程技术”，还是一门“艺术”的争论。当时存在这种争论是很自然的。在 AI 发展的初期，大多数研究者采取的研究方法是，首先凭借直觉或者启发式建立起 AI 的相关假设，然后在“玩具世界”中论证假设的合理性，由此建立起一套 AI 的理论与方法。为了克服数学方法的“局限性”，他们总是避免使用数学工具，尽量与传统的严格科学保持距离。但是，随着 AI 走向成熟，AI 的“传统”发生了变化，它们逐步向科学靠拢，向实用靠近。一方面，尽量使用现代科学工具，使 AI 逐步变成一门科学。一方面，尽量面向现实世界，提出可行的算法，使 AI 走向实际应用。该书作者将这两大趋势及时地反映在教材中，从而形成自己的特色。

作者在前言中特别说明了本书与第 1 版的区别。我认为，第 2 版加强并进一步突出了以上两个特点。作者在重写过程中，对于理论部分尽量采用已有的、成熟的科学方法，如数学、心理学、计算机工程、神经科学等。在第 2 版中，还进一步强调了经济学中决策理论、运筹学，以及控制理论、控制论等与 AI 的关系，将 AI 与其它科学领域联系起来。在各个章节中都补充了新的内容，将这期间所取得的研究成果尽量纳入到新书中。

尽管本书已由清华大学计算机系姜哲等老师译为中文，并于 2004 年出版，为了使读者能够读到英文原书，我以为出版第 2 版的影印本还是有必要的。

中国科学院院士 清华大学教授

张钹

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 in presenting results, although we strive to cover only the most central ideas in the main part of each chapter. Pointers are given to further results in the bibliographical notes at the end of each chapter.

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 than they might otherwise have been.

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 production systems, reactive agents, real-time conditional planners, neural networks, 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 millenia of related work. We have tried to avoid excessive formality in the presentation of these ideas while retaining precision. Wherever appropriate, we have included pseudocode algorithms to make the ideas concrete; our pseudocode is described briefly in Appendix B. Implementations in several programming languages are available on the book's Web site, **aima.cs.berkeley.edu**.

This book is primarily intended for use in an undergraduate course or course sequence. It can also be used in a graduate-level course (perhaps with the addition of some of the primary sources suggested in the bibliographical notes). Because of its comprehensive coverage and large number of detailed algorithms, it is useful as a primary reference volume for AI graduate students and professionals wishing to branch out beyond their own subfield. The only prerequisite is familiarity with basic concepts of computer science (algorithms, data structures, complexity) at a sophomore level. Freshman calculus is useful for understanding neural networks and statistical learning in detail. Some of the required mathematical background is supplied in Appendix A.

Overview of the book

The book is divided into eight parts. Part I, **Artificial Intelligence**, offers a view of the AI enterprise based around the idea of intelligent agents—systems that can decide what to do and then do it. Part II, **Problem Solving**, concentrates on methods for deciding what to do when one needs to think ahead several steps—for example in navigating across a country or playing chess. Part III, **Knowledge and Reasoning**, discusses ways to represent knowledge about the world—how it works, what it is currently like, and what one's actions might do—and how to reason logically with that knowledge. Part IV, **Planning**, then discusses how to use these reasoning methods to decide what to do, particularly by constructing *plans*. Part V, **Uncertain Knowledge and Reasoning**, is analogous to Parts III and IV, but it concentrates on reasoning and decision making in the presence of *uncertainty* about the world, as might be faced, for example, by a system for medical diagnosis and treatment.

Together, Parts II–V describe that part of the intelligent agent responsible for reaching decisions. Part VI, **Learning**, describes methods for generating the knowledge required by these decision-making

试读结束：需要全本请在线购买： www.ertongbook.com

components. Part VII, **Communicating, Perceiving, and Acting**, describes ways in which an intelligent agent can perceive its environment so as to know what is going on, whether by vision, touch, hearing, or understanding language, and ways in which it can turn its plans into real actions, either as robot motion or as natural language utterances. Finally, Part VIII, **Conclusions**, analyzes the past and future of AI and the philosophical and ethical implications of artificial intelligence.

Changes from the first edition

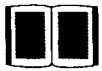
Much has changed in AI since the publication of the first edition in 1995, and much has changed in this book. Every chapter has been significantly rewritten to reflect the latest work in the field, to reinterpret old work in a way that is more cohesive with new findings, and to improve the pedagogical flow of ideas. Followers of AI should be encouraged that current techniques are much more practical than those of 1995; for example the planning algorithms in the first edition could generate plans of only dozens of steps, while the algorithms in this edition scale up to tens of thousands of steps. Similar orders-of-magnitude improvements are seen in probabilistic inference, language processing, and other subfields. The following are the most notable changes in the book:

- In Part I, we acknowledge the historical contributions of control theory, game theory, economics, and neuroscience. This helps set the tone for a more integrated coverage of these ideas in subsequent chapters.
- In Part II, online search algorithms are covered and a new chapter on constraint satisfaction has been added. The latter provides a natural connection to the material on logic.
- In Part III, propositional logic, which was presented as a stepping-stone to first-order logic in the first edition, is now presented as a useful representation language in its own right, with fast inference algorithms and circuit-based agent designs. The chapters on first-order logic have been reorganized to present the material more clearly and we have added the Internet shopping domain as an example.
- In Part IV, we include newer planning methods such as GRAPHPLAN and satisfiability-based planning, and we increase coverage of scheduling, conditional planning, hierarchical planning, and multiagent planning.
- In Part V, we have augmented the material on Bayesian networks with new algorithms, such as variable elimination and Markov Chain Monte Carlo, and we have created a new chapter on uncertain temporal reasoning, covering hidden Markov models, Kalman filters, and dynamic Bayesian networks. The coverage of Markov decision processes is deepened, and we add sections on game theory and mechanism design.
- In Part VI, we tie together work in statistical, symbolic, and neural learning and add sections on boosting algorithms, the EM algorithm, instance-based learning, and kernel methods (support vector machines).
- In Part VII, coverage of language processing adds sections on discourse processing and grammar induction, as well as a chapter on probabilistic language models, with applications to information retrieval and machine translation. The coverage of robotics stresses the integration of uncertain sensor data, and the chapter on vision has updated material on object recognition.
- In Part VIII, we introduce a section on the ethical implications of AI.

Using this book

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. Alternatively, a course can be tailored to suit the interests of the instructor and student. Through its broad coverage, the book can be used to support such

courses, whether they are short, introductory undergraduate courses or specialized graduate courses on advanced topics. Sample syllabi from the more than 600 universities and colleges that have adopted the first edition are shown on the Web at aima.cs.berkeley.edu, along with suggestions to help you find a sequence appropriate to your needs.



The book includes 385 exercises. 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. 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 10,000 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.

NEW TERM

Using the Web site

At the **aima.cs.berkeley.edu** Web site you will find:

- implementations of the algorithms in the book in several programming languages,
- a list of over 600 schools that have used the book, many with links to online course materials,
- 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
- copies of the figures in the book, along with slides and other material for instructors.

Acknowledgments

Jitendra Malik wrote most of Chapter 24 (on vision). Most of Chapter 25 (on robotics) was written by Sebastian Thrun in this edition and by John Canny in the first edition. Doug Edwards researched the historical notes for the first edition. Tim Huang, Mark Paskin, and Cynthia Bruyns helped with formatting of the diagrams and algorithms. Alan Apt, Sondra Chavez, Toni Holm, Jake Warde, Irwin Zucker, and Camille Trentacoste at Prentice Hall tried their best to keep us on schedule and made many helpful suggestions on the book's design and content.

Stuart would like to thank his parents for their continued support and encouragement and his wife, Loy Sheflott, for her endless patience and boundless wisdom. He hopes that Gordon and Lucy will soon be reading this. RUGS (Russell's Unusual Group of Students) have been unusually helpful.

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

We are indebted to the librarians at Berkeley, Stanford, MIT, and NASA, and to the developers of CiteSeer and Google, who have revolutionized the way we do research.

We can't thank all the people who have used the book and made suggestions, but we would like to acknowledge the especially helpful comments of Eyal Amir, Krzysztof Apt, Ellery Aziel, Jeff Van Baalen, Brian Baker, Don Barker, Tony Barrett, James Newton Bass, Don Beal, Howard Beck, Wolfgang Bibel, John Binder, Larry Bookman, David R. Boxall, Gerhard Brewka, Selmer Bringsjord, Carla Brodley, Chris Brown, Wilhelm Burger, Lauren Burka, Joao Cachopo, Murray Campbell, Norman Carver, Emmanuel Castro, Anil Chakravarthy, Dan Chisarick, Roberto Cipolla, David Cohen, James Coleman, Julie Ann Comparini, Gary Cottrell, Ernest Davis, Rina Dechter, Tom Dietterich, Chuck Dyer, Barbara Engelhardt, Doug Edwards, Kutluhan Erol, Oren Etzioni, Hana Filip, Douglas

Fisher, Jeffrey Forbes, Ken Ford, John Fosler, 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, Vasant Honavar, Tim Huang, Seth Hutchinson, Joost Jacob, Magnus Johansson, Dan Jurafsky, Leslie Kaelbling, Keiji Kanazawa, Surekha Kasibhatla, Simon Kasif, Henry Kautz, Gernot Kerschbaumer, Richard Kirby, Kevin Knight, Sven Koenig, Daphne Koller, Rich Korf, James Kurien, John Lafferty, Gus Larsson, John Lazzaro, Jon LeBlanc, Jason Leatherman, Frank Lee, Edward Lim, Pierre Louveaux, Don Loveland, Sridhar Mahadevan, Jim Martin, Andy Mayer, David McGrane, Jay Mendelsohn, Brian Milch, Steve Minton, Vibhu Mittal, Leora Morgenstern, Stephen Muggleton, Kevin Murphy, Ron Musick, Sung Myaeng, Lee Naish, Pandu Nayak, Bernhard Nebel, Stuart Nelson, XuanLong Nguyen, Illah Nourbakhsh, Steve Omohundro, David Page, David Palmer, David Parkes, Ron Parr, Mark Paskin, Tony Passera, Michael Pazzani, Wim Pijls, Ira Pohl, Martha Pollack, David Poole, Bruce Porter, Malcolm Pradhan, Bill Pringle, Lorraine Prior, Greg Provan, William Rapaport, Philip Resnik, Francesca Rossi, Jonathan Schaeffer, Richard Scherl, Lars Schuster, Soheil Shams, Stuart Shapiro, Jude Shavlik, Satinder Singh, Daniel Sleator, David Smith, Bryan So, Robert Sproull, Lynn Stein, Larry Stephens, Andreas Stolcke, Paul Stradling, Devika Subramanian, Rich Sutton, Jonathan Tash, Austin Tate, Michael Thielscher, William Thompson, Sebastian Thrun, Eric Tiedemann, Mark Torrance, Randall Upham, Paul Utgoff, Peter van Beek, Hal Varian, Sunil Vemuri, Jim Waldo, Bonnie Webber, Dan Weld, Michael Wellman, Michael Dean White, Kamin Whitehouse, Brian Williams, David Wolfe, Bill Woods, Alden Wright, Richard Yen, Weixiong Zhang, Shlomo Zilberstein, and the anonymous reviewers provided by Prentice Hall.

About the Cover

The cover image was designed by the authors and executed by Lisa Marie Sardegna and Maryann Simmons using SGI Inventor™ and Adobe Photoshop™. The cover depicts the following items from the history of AI:

1. Aristotle's planning algorithm from *De Motu Animalium* (c. 400 B.C.).
2. Ramon Lull's concept generator from *Ars Magna* (c. 1300 A.D.).
3. Charles Babbage's Difference Engine, a prototype for the first universal computer (1848).
4. Gottlob Frege's notation for first-order logic (1789).
5. Lewis Carroll's diagrams for logical reasoning (1886).
6. Sewall Wright's probabilistic network notation (1921).
7. Alan Turing (1912–1954).
8. Shakey the Robot (1969–1973).
9. A modern diagnostic expert system (1993).

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 director of Search Quality at Google, Inc. He is a Fellow and Executive Council member of the American Association for Artificial Intelligence. 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. Before that he served as chief scientist at Junglee, where he helped develop one of the first Internet information extraction services, and as a senior scientist at Sun Microsystems Laboratories working on intelligent information retrieval. 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 has been a professor at the University of Southern California and a research faculty member at Berkeley. He has over 50 publications in computer science including the books *Paradigms of AI Programming: Case Studies in Common Lisp*, *Verbmobil: A Translation System for Face-to-Face Dialog*, and *Intelligent Help Systems for UNIX.*

Summary of Contents

I Artificial Intelligence

1 Introduction 1
2 Intelligent Agents 32

II Problem-solving

3 Solving Problems by Searching 59
4 Informed Search and Exploration 94
5 Constraint Satisfaction Problems 137
6 Adversarial Search 161

III Knowledge and reasoning

7 Logical Agents 194
8 First-Order Logic 240
9 Inference in First-Order Logic 272
10 Knowledge Representation 320

IV Planning

11 Planning 375
12 Planning and Acting in the Real World 417

V Uncertain knowledge and reasoning

13 Uncertainty 462
14 Probabilistic Reasoning 492
15 Probabilistic Reasoning over Time 537
16 Making Simple Decisions 584
17 Making Complex Decisions 613

VI Learning

18 Learning from Observations 649
19 Knowledge in Learning 678
20 Statistical Learning Methods 712
21 Reinforcement Learning 763

VII Communicating, perceiving, and acting

22 Communication 790
23 Probabilistic Language Processing 834
24 Perception 863
25 Robotics 901

VIII Conclusions

26 Philosophical Foundations 947
27 AI: Present and Future 968

A Mathematical background 977
B Notes on Languages and Algorithms 984

Bibliography 987

Index 1045

Contents

I Artificial Intelligence

1 Introduction 1

1.1 What is AI? . . . 1
Acting humanly: The Turing Test approach . . . 2
Thinking humanly: The cognitive modeling approach . . . 3
Thinking rationally: The "laws of thought" approach . . . 4
Acting rationally: The rational agent approach . . . 4

1.2 The Foundations of Artificial Intelligence . . . 5
Philosophy (428 B.C.–present) . . . 5
Mathematics (c. 800–present) . . . 7
Economics (1776–present) . . . 9
Neuroscience (1861–present) . . . 10
Psychology (1879–present) . . . 12
Computer engineering (1940–present) . . . 14
Control theory and Cybernetics (1948–present) . . . 15
Linguistics (1957–present) . . . 16

1.3 The History of Artificial Intelligence . . . 16
The gestation of artificial intelligence (1943–1955) . . . 16
The birth of artificial intelligence (1956) . . . 17
Early enthusiasm, great expectations (1952–1969) . . . 18
A dose of reality (1966–1973) . . . 21
Knowledge-based systems: The key to power? (1969–1979) . . . 22
AI becomes an industry (1980–present) . . . 24
The return of neural networks (1986–present) . . . 25
AI becomes a science (1987–present) . . . 25
The emergence of intelligent agents (1995–present) . . . 27

1.4 The State of the Art . . . 27

1.5 Summary . . . 28

Bibliographical and Historical Notes . . . 29

Exercises . . . 30

2 Intelligent Agents 32

2.1 Agents and Environments . . . 32

2.2 Good Behavior: The Concept of Rationality . . . 34
Performance measures . . . 35
Rationality . . . 35
Omniscience, learning, and autonomy . . . 36

2.3 The Nature of Environments . . . 38
Specifying the task environment . . . 38
Properties of task environments . . . 40

2.4 The Structure of Agents . . . 44
Agent programs . . . 44
Simple reflex agents . . . 46
Model-based reflex agents . . . 48

试读结束：需要全本请在线购买： www.ertongbook.com

Goal-based agents . . . 49
Utility-based agents . . . 51
Learning agents . . . 51
2.5 Summary . . . 54
Bibliographical and Historical Notes . . . 55
Exercises . . . 56

II Problem-solving

3 Solving Problems by Searching **59**
3.1 Problem-Solving Agents . . . 59
Well-defined problems and solutions . . . 62
Formulating problems . . . 62
3.2 Example Problems . . . 64
Toy problems . . . 64
Real-world problems . . . 67
3.3 Searching for Solutions . . . 69
Measuring problem-solving performance . . . 71
3.4 Uninformed Search Strategies . . . 73
Breadth-first search . . . 73
Depth-first search . . . 75
Depth-limited search . . . 77
Iterative deepening depth-first search . . . 78
Bidirectional search . . . 79
Comparing uninformed search strategies . . . 81
3.5 Avoiding Repeated States . . . 81
3.6 Searching with Partial Information . . . 83
Sensorless problems . . . 84
Contingency problems . . . 86
3.7 Summary . . . 87
Bibliographical and Historical Notes . . . 88
Exercises . . . 89

4 Informed Search and Exploration **94**
4.1 Informed (Heuristic) Search Strategies . . . 94
Greedy best-first search . . . 95
A* search: Minimizing the total estimated solution cost . . . 97
Memory-bounded heuristic search . . . 101
Learning to search better . . . 104
4.2 Heuristic Functions . . . 105
The effect of heuristic accuracy on performance . . . 106
Inventing admissible heuristic functions . . . 107
Learning heuristics from experience . . . 109
4.3 Local Search Algorithms and Optimization Problems . . . 110
Hill-climbing search . . . 111
Simulated annealing search . . . 115
Local beam search . . . 115
Genetic algorithms . . . 116
4.4 Local Search in Continuous Spaces . . . 119

4.5 Online Search Agents and Unknown Environments . . . 122
Online search problems . . . 123
Online search agents . . . 125
Online local search . . . 126
Learning in online search . . . 127
4.6 Summary . . . 129
Bibliographical and Historical Notes . . . 130
Exercises . . . 134

5 Constraint Satisfaction Problems **137**
5.1 Constraint Satisfaction Problems . . . 137
5.2 Backtracking Search for CSPs . . . 141
Variable and value ordering . . . 143
Propagating information through constraints . . . 144
Intelligent backtracking: looking backward . . . 148
5.3 Local Search for Constraint Satisfaction Problems . . . 150
5.4 The Structure of Problems . . . 151
5.5 Summary . . . 155
Bibliographical and Historical Notes . . . 156
Exercises . . . 158

6 Adversarial Search **161**
6.1 Games . . . 161
6.2 Optimal Decisions in Games . . . 162
Optimal strategies . . . 163
The minimax algorithm . . . 165
Optimal decisions in multiplayer games . . . 165
6.3 Alpha–Beta Pruning . . . 167
6.4 Imperfect, Real-Time Decisions . . . 171
Evaluation functions . . . 171
Cutting off search . . . 173
6.5 Games That Include an Element of Chance . . . 175
Position evaluation in games with chance nodes . . . 177
Complexity of expectiminimax . . . 177
Card games . . . 179
6.6 State-of-the-Art Game Programs . . . 180
6.7 Discussion . . . 183
6.8 Summary . . . 185
Bibliographical and Historical Notes . . . 186
Exercises . . . 189

III Knowledge and reasoning

7 Logical Agents **194**
7.1 Knowledge-Based Agents . . . 195
7.2 The Wumpus World . . . 197
7.3 Logic . . . 200
7.4 Propositional Logic: A Very Simple Logic . . . 204
Syntax . . . 204

Semantics . . . 206
A simple knowledge base . . . 208
Inference . . . 208
Equivalence, validity, and satisfiability . . . 210
7.5 Reasoning Patterns in Propositional Logic . . . 211
Resolution . . . 213
Forward and backward chaining . . . 217
7.6 Effective propositional inference . . . 220
A complete backtracking algorithm . . . 221
Local-search algorithms . . . 222
Hard satisfiability problems . . . 224
7.7 Agents Based on Propositional Logic . . . 225
Finding pits and wumpuses using logical inference . . . 225
Keeping track of location and orientation . . . 227
Circuit-based agents . . . 227
A comparison . . . 231
7.8 Summary . . . 232
Bibliographical and Historical Notes . . . 233
Exercises . . . 236

8 First-Order Logic **240**
8.1 Representation Revisited . . . 240
8.2 Syntax and Semantics of First-Order Logic . . . 245
Models for first-order logic . . . 245
Symbols and interpretations . . . 246
Terms . . . 248
Atomic sentences . . . 248
Complex sentences . . . 249
Quantifiers . . . 249
Equality . . . 253
8.3 Using First-Order Logic . . . 253
Assertions and queries in first-order logic . . . 253
The kinship domain . . . 254
Numbers, sets, and lists . . . 256
The wumpus world . . . 258
8.4 Knowledge Engineering in First-Order Logic . . . 260
The knowledge engineering process . . . 261
The electronic circuits domain . . . 262
8.5 Summary . . . 266
Bibliographical and Historical Notes . . . 267
Exercises . . . 268

9 Inference in First-Order Logic **272**
9.1 Propositional vs. First-Order Inference . . . 272
Inference rules for quantifiers . . . 273
Reduction to propositional inference . . . 274
9.2 Unification and Lifting . . . 275
A first-order inference rule . . . 275
Unification . . . 276

Storage and retrieval . . . 278
9.3 Forward Chaining . . . 280
First-order definite clauses . . . 280
A simple forward-chaining algorithm . . . 281
Efficient forward chaining . . . 283
9.4 Backward Chaining . . . 287
A backward chaining algorithm . . . 287
Logic programming . . . 289
Efficient implementation of logic programs . . . 290
Redundant inference and infinite loops . . . 292
Constraint logic programming . . . 294
9.5 Resolution . . . 295
Conjunctive normal form for first-order logic . . . 295
The resolution inference rule . . . 297
Example proofs . . . 297
Completeness of resolution . . . 300
Dealing with equality . . . 303
Resolution strategies . . . 304
Theorem provers . . . 306
9.6 Summary . . . 310
Bibliographical and Historical Notes . . . 310
Exercises . . . 315

10 Knowledge Representation **320**
10.1 Ontological Engineering . . . 320
10.2 Categories and Objects . . . 322
Physical composition . . . 324
Measurements . . . 325
Substances and objects . . . 327
10.3 Actions, Situations, and Events . . . 328
The ontology of situation calculus . . . 329
Describing actions in situation calculus . . . 330
Solving the representational frame problem . . . 332
Solving the inferential frame problem . . . 333
Time and event calculus . . . 334
Generalized events . . . 335
Processes . . . 337
Intervals . . . 338
Fluents and objects . . . 339
10.4 Mental Events and Mental Objects . . . 341
A formal theory of beliefs . . . 341
Knowledge and belief . . . 343
Knowledge, time, and action . . . 344
10.5 The Internet Shopping World . . . 344
Comparing offers . . . 348
10.6 Reasoning Systems for Categories . . . 349
Semantic networks . . . 350
Description logics . . . 353
10.7 Reasoning with Default Information . . . 354

Open and closed worlds . . . 354
Negation as failure and stable model semantics . . . 356
Circumscription and default logic . . . 358
10.8 Truth Maintenance Systems . . . 360
10.9 Summary . . . 362
Bibliographical and Historical Notes . . . 363
Exercises . . . 369

IV Planning

11 Planning **375**
11.1 The Planning Problem . . . 375
The language of planning problems . . . 377
Expressiveness and extensions . . . 378
Example: Air cargo transport . . . 380
Example: The spare tire problem . . . 381
Example: The blocks world . . . 381
11.2 Planning with State-Space Search . . . 382
Forward state-space search . . . 382
Backward state-space search . . . 384
Heuristics for state-space search . . . 386
11.3 Partial-Order Planning . . . 387
A partial-order planning example . . . 391
Partial-order planning with unbound variables . . . 393
Heuristics for partial-order planning . . . 394
11.4 Planning Graphs . . . 395
Planning graphs for heuristic estimation . . . 397
The GRAPHPLAN algorithm . . . 398
Termination of GRAPHPLAN . . . 401
11.5 Planning with Propositional Logic . . . 402
Describing planning problems in propositional logic . . . 402
Complexity of propositional encodings . . . 405
11.6 Analysis of Planning Approaches . . . 407
11.7 Summary . . . 408
Bibliographical and Historical Notes . . . 409
Exercises . . . 412

12 Planning and Acting in the Real World **417**
12.1 Time, Schedules, and Resources . . . 417
Scheduling with resource constraints . . . 420
12.2 Hierarchical Task Network Planning . . . 422
Representing action decompositions . . . 423
Modifying the planner for decompositions . . . 425
Discussion . . . 427
12.3 Planning and Acting in Nondeterministic Domains . . . 430
12.4 Conditional Planning . . . 433
Conditional planning in fully observable environments . . . 433
Conditional planning in partially observable environments . . . 437
12.5 Execution Monitoring and Replanning . . . 441

12.6 Continuous Planning . . . 445
12.7 MultiAgent Planning . . . 449
Cooperation: Joint goals and plans . . . 450
Multibody planning . . . 451
Coordination mechanisms . . . 452
Competition . . . 454
12.8 Summary . . . 454
Bibliographical and Historical Notes . . . 455
Exercises . . . 459

V Uncertain knowledge and reasoning

13 Uncertainty **462**
13.1 Acting under Uncertainty . . . 462
Handling uncertain knowledge . . . 463
Uncertainty and rational decisions . . . 465
Design for a decision-theoretic agent . . . 466
13.2 Basic Probability Notation . . . 466
Propositions . . . 467
Atomic events . . . 468
Prior probability . . . 468
Conditional probability . . . 470
13.3 The Axioms of Probability . . . 471
Using the axioms of probability . . . 473
Why the axioms of probability are reasonable . . . 473
13.4 Inference Using Full Joint Distributions . . . 475
13.5 Independence . . . 477
13.6 Bayes' Rule and Its Use . . . 479
Applying Bayes' rule: The simple case . . . 480
Using Bayes' rule: Combining evidence . . . 481
13.7 The Wumpus World Revisited . . . 483
13.8 Summary . . . 486
Bibliographical and Historical Notes . . . 487
Exercises . . . 489

14 Probabilistic Reasoning **492**
14.1 Representing Knowledge in an Uncertain Domain . . . 492
14.2 The Semantics of Bayesian Networks . . . 495
Representing the full joint distribution . . . 495
Conditional independence relations in Bayesian networks . . . 499
14.3 Efficient Representation of Conditional Distributions . . . 500
14.4 Exact Inference in Bayesian Networks . . . 504
Inference by enumeration . . . 504
The variable elimination algorithm . . . 507
The complexity of exact inference . . . 509
Clustering algorithms . . . 510
14.5 Approximate Inference in Bayesian Networks . . . 511
Direct sampling methods . . . 511
Inference by Markov chain simulation . . . 516

试读结束：需要全本请在线购买： www.ertongbook.com