计 算 机 科 学 丛 书

Artificial Intelligence
A New Synthesis

人工智能

(英文版)

(美) Nils J. Nilsson 著

机械工业出版社
China Machine Press

MORGAN KAUFMANN

计算机科学丛书

人工智能

(英文版)

Artificial Intelligence
A New Synthesis

(美) Nils J. Nilsson 著

机械工业出版社
China Machine Press

本书版权登记号：图字：01-1999-2024

图书在版编目（CIP）数据

人工智能：英文/(美)尼尔松(Nilsson, N. J.)著.-北京：机械工业出版社，1999.9
(计算机科学丛书)
ISBN 7-111-07438-6

Ⅰ.人… Ⅱ.尼… Ⅲ.人工智能－英文 Ⅳ.TP18

中国版本图书馆CIP数据核字(1999)第33788号

机械工业出版社（北京市西城区百万庄大街22号 邮政编码 100037）
北京牛山世兴印刷厂印刷·新华书店北京发行所发行
1999年9月第1版·2001年5月第2次印刷
787mm×1092mm1/16·33.5印张
印数：2 001-3 000册
定价：45.00元

For Scott and Ryan

About the author

Nils J. Nilsson's long and rich research career has contributed much to Al. His previous books, considered classics in the field, include Learning Machines, Problem-Solving Methods in Artificial Intelligence, Logical Foundations of Artificial Intelligence, and Principles of Artificial Intelligence. Dr. Nilsson is Kumagai Professor of Engineering, Emeritus, at Stanford University. He has served on the editorial boards of Artificial Intelligence and Machine Learning and as an area editor for the Journal of the Association for Computing Machinery. Former chairman of the Department of Computer Science at Stanford and Former Director of the SRI Artificial Intelligence Center, he is also a past president and fellow of the American Association for Artificial Intelligence.

Preface

This introductory textbook employs a novel perspective from which to view topics in artificial intelligence (AI). I will consider a progression of AI systems or "agents," each slightly more complex than its predecessor. I begin with elementary agents that react to sensed properties of their environments. Even such simple machines allow me to treat topics in machine vision, machine learning, and machine evolution. Then, by stages, I introduce techniques that allow agents to exploit information about the task environment that cannot be immediately sensed. Such knowledge can take the form of descriptive information about the state of the environment, iconic models of the environment, state-space graphs, and logical representations. Because the progression follows what plausibly might have been milestones in the evolution of animals, I have called this approach *evolutionary artificial intelligence*. I intend the book to be as much a proposal about how to think about AI as it is a description of AI techniques. Examples will be used to provide motivation and grounding.

Although I use agents to motivate and illustrate AI techniques, the techniques themselves have much broader application. Many ideas invented by AI researchers have been assimilated into computer science generally for applications in expert systems, natural language processing, human-machine interaction, information retrieval, graphics and image processing, data mining, and robotics (to name some examples). The agents theme serves to unify what might otherwise seem to be a collection of disparate topics.

Regarding coverage, my intention is to treat the middle ground between theory and applications. This middle ground is rich in important AI *ideas*, and in this book I try to motivate and explain the ideas that I think have lasting value in AI. (Being subject to the usual human frailties, I admit to possible errors of omission

and commission in selecting topics for inclusion.) Also, some subjects are treated in more depth than others—both because I thought some subjects more important and because I wanted to provide at least some examples of greater depth of exposition. Although some pseudocode algorithms are presented, the book is not an AI programming and implementation book. (Some "AI techniques" books are [Shoham 1994, Norvig 1992, Tracy & Bouthoorn 1997].) I do not give proofs of all of the important theoretical results, but I try to give intuitive arguments and citations to formal proofs. My goals are to present a modest-sized textbook for a one-semester introductory college course, to give the student and reader sufficient motivation and preparation to go on to more advanced AI courses, and to make the extensive literature on AI accessible.

A somewhat unconventional feature of the book is that machine learning is not treated as a separate topic; instead, various aspects of learning arise throughout the book. Neural nets and fundamental ideas about supervised learning are presented early; techniques for learning search heuristics and action policies are discussed in the chapters on search; rule learning, inductive logic programming, and explanation-based learning are treated toward the end of the chapters on logic; and learning plans is presented after discussing logic-based planning.

In my previous books, I included a "bibliographic and historical remarks" section at the end of each chapter. (Some readers may find those sections of some interest still.) I have not done so in the present book, both because AI history has now accumulated to such a great extent and because the longer text by [Russell & Norvig 1995] has already done such a thorough job in that regard. Instead, I include remarks and citations as appropriate throughout the text and provide some additional ones in discussion sections at the end of most chapters. The serious student who intends to specialize in AI research will want to consult many of the references. I hope the casual reader is not bothered by the many citations.

Sample exercises are included at the end of each chapter. They vary in difficulty from routine application of ideas presented in the book to mildly challenging. I expect that instructors will want to augment these problems with favorite ones of their own, including computer exercises and projects. (In keeping with my decision to concentrate on ideas instead of programs, I have not included any computer exercises or projects. Several good programming and project ideas can be found in texts devoted to AI programming techniques.)

The following typographical conventions are used in this book. Sans serif font is used for the names of actions and for "proto-English" sentences communicated among agents. SANS SERIF capitals are used for the names of computer languages, algorithms, and AI systems. Boldface capital letters, such as **W** and **X**, are used for vectors, matrices, and modal operators. `Typewriter` font is used for genetic programs, for expressions and subexpressions in the predicate calculus, and for

STRIPS rules and operators. Lowercase Greek letters are used for metavariables ranging over predicate-calculus expressions, subexpressions, and occasionally for substitutions. Uppercase Greek letters are used to denote sets of predicate-calculus formulas. Lowercase *p*'s are used to denote probabilities.

Students and researchers will find much helpful material about AI on the World Wide Web. I do not provide URLs here; any list written today would be incomplete and inaccurate within months. Use of one of the web search engines will quickly steer the reader to sites with sample applications, frequently asked questions, extensive bibliographies, research papers, programs, interactive demonstrations, announcements of workshops and conferences, homepages of researchers, and much more.

Material specifically in support of this book is provided on a Web page on the publisher's Web site at *www.mkp.com/nils*. If you discover any errors, please email them to the publisher at *aibugs@mkp.com*. Errata and clarifications can be found at *http://www.mkp.com/nils/clarified*.

My previous AI textbook, *Principles of Artificial Intelligence* (Morgan Kaufmann, 1980), is by now quite out of date, but some of the material in that book is still useful, and I have borrowed freely from it in preparing the present volume. Cross-checking against other AI textbooks (particularly [Russell & Norvig 1995, Rich & Knight 1991, Stefik 1995]) was also very helpful.

Students and teaching assistants in my Stanford courses on artificial intelligence and machine learning have already made several useful suggestions. I hope the following list includes most of them: Eyal Amir, David Andre, Scott Benson, George John, Steve Ketchpel, Ron Kohavi, Andrew Kosoresow, Ofer Matan, Karl Pfleger, and Charles Richards. Colleagues and reviewers at Stanford and elsewhere helped me learn what they already knew. Thanks to Helder Coelho, Oscar Firschein, Carolyn Hayes, Giorgio Ingargiola, Leslie Kaelbling, Daphne Koller, John Koza, Richard Korf, Pat Langley, John McCarthy, Bart Selman, Yoav Shoham, Devika Subramanian, Gheorghe Tecuci, and Michael Wellman. Special thanks go to Cheri Palmer, my production editor at Morgan Kaufmann, who kept me on schedule, cheerfully accepted my endless changes, and worked extra hard to meet a difficult publication date. Work on this book was carried on in the Robotics Laboratory of Stanford's Department of Computer Science and at the Santa Fe Institute. Continuing research support by the National Science Foundation is gratefully acknowledged.

Contents