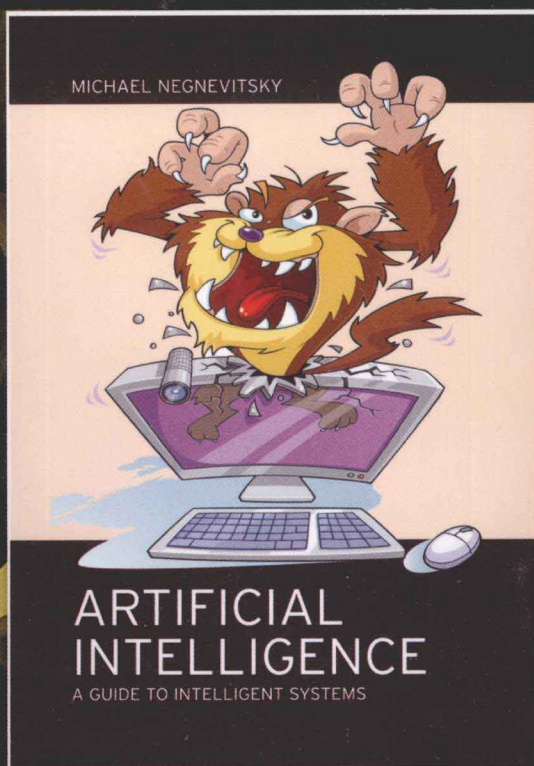


原书第3版

人工智能 智能系统指南

(澳) Michael Negnevitsky 著 陈薇 等译

Artificial Intelligence
A Guide to Intelligent Systems Third Edition



机械工业出版社
China Machine Press

计

算

丛

书

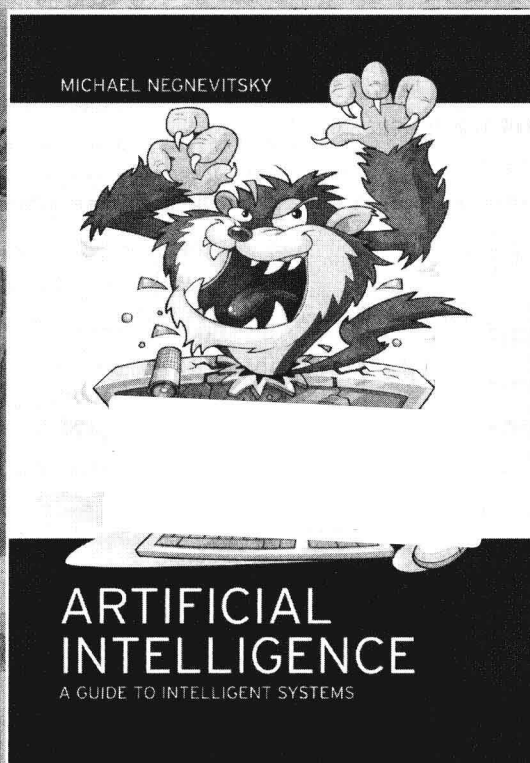
原书第3版

人工智能 智能系统指南

(澳) Michael Negnevitsky 著 陈薇 等译

Artificial Intelligence

A Guide to Intelligent Systems Third Edition



机械工业出版社
China Machine Press

本书是一本很好的人工智能入门书籍,内容丰富、浅显易懂。作者根据自己多年的教学、实践经验,并结合实际代码、图示、案例等讲解了人工智能的基本知识。

全书共分10章,主要内容包括:基于规则的专家系统、不确定性管理技术、模糊专家系统、基于框架的专家系统、人工神经网络、进化计算、混合智能系统、知识工程、数据挖掘等。另外,本书还提供了一个人工智能相关术语表和包含商业化的人工智能工具的附录。

本书既可以作为计算机科学相关专业本科生的入门教材,也可以作为非计算机专业读者的自学参考书。

Michael Negnevitsky: Artificial Intelligence: A Guide to Intelligent Systems, Third Edition (ISBN 978-1-4082-2574-5).

Copyright © 2011 by Pearson Education Limited.

This translation of Artificial Intelligence: A Guide to Intelligent Systems, Third Edition (ISBN 978-1-4082-2574-5) is published by arrangement with Pearson Education Limited.

All rights reserved.

本书中文简体字版由英国 Pearson Education 培生教育出版集团授权出版。

封底无防伪标均为盗版

版权所有,侵权必究

本书法律顾问 北京市展达律师事务所

本书版权登记号:图字:01-2011-4255

图书在版编目(CIP)数据

人工智能:智能系统指南(原书第3版)/(澳)尼格尼维斯基(Negnevitsky, M.)著;陈薇等译. —北京:机械工业出版社,2012.7

(计算机科学丛书)

书名原文:Artificial Intelligence: A Guide to Intelligent Systems, Third Edition

ISBN 978-7-111-38455-7

I. 人… II. ①尼… ②陈… III. 人工智能 IV. TP18

中国版本图书馆CIP数据核字(2012)第104490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)

责任编辑:高婧雅

北京市荣盛彩色印刷有限公司印刷

2012年8月第1版第1次印刷

185mm×260mm·20.75印张

标准书号:ISBN 978-7-111-38455-7

定价:49.00元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

客服热线:(010) 88378991; 88361066

购书热线:(010) 68326294; 88379649; 68995259

投稿热线:(010) 88379604

读者信箱:hzjsj@hzbook.com

文艺复兴以降,源远流长的科学精神和逐步形成的学术规范,使西方国家在自然科学的各个领域中取得了垄断性的优势;也正是这样的传统,使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中,美国的产业界与教育界越来越紧密地结合,计算机学科中的许多泰山北斗同时身处科研和教学的最前线,由此而产生的经典科学著作,不仅擘划了研究的范畴,还揭示了学术的源变,既遵循学术规范,又自有学者个性,其价值并不会因年月的流逝而减退。

近年,在全球信息化大潮的推动下,我国的计算机产业发展迅猛,对专业人才的需求日益迫切。这对计算机教育界和出版界都既是机遇,也是挑战;而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短的现状下,美国等发达国家在其计算机科学发展的几十年间积淀和发展的经典教材仍有许多值得借鉴之处。因此,引进一批国外优秀计算机教材将对我国计算机教育事业的发展起到积极的推动作用,也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章公司较早意识到“出版要为教育服务”。自1998年开始,我们就将工作重点放在了遴选、移译国外优秀教材上。经过多年的不懈努力,我们与Pearson, McGraw-Hill, Elsevier, MIT, John Wiley & Sons, Cengage等世界著名出版公司建立了良好的合作关系,从他们现有的数百种教材中甄选出Andrew S. Tanenbaum, Bjarne Stroustrup, Brian W. Kernighan, Dennis Ritchie, Jim Gray, Alfred V. Aho, John E. Hopcroft, Jeffrey D. Ullman, Abraham Silberschatz, William Stallings, Donald E. Knuth, John L. Hennessy, Larry L. Peterson等大师名家的一批经典作品,以“计算机科学丛书”为总称出版,供读者学习、研究及珍藏。大理石纹理的封面,也正体现了这套丛书的品位和格调。

“计算机科学丛书”的出版工作得到了国内外学者的鼎力襄助,国内的专家不仅提供了中肯的选题指导,还不辞劳苦地担任了翻译和审校的工作;而原书的作者也相当关注其作品在中国的传播,有的还专程为其书的中译本作序。迄今,“计算机科学丛书”已经出版了近两百个品种,这些书籍在读者中树立了良好的口碑,并被许多高校采用为正式教材和参考书籍。其影印版“经典原版书库”作为姊妹篇也被越来越多实施双语教学的学校所采用。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑,这些因素使我们的图书有了质量的保证。随着计算机科学与技术专业学科建设的不断完善和教材改革的逐渐深化,教育界对国外计算机教材的需求和应用都将步入一个新的阶段,我们的目标是尽善尽美,而反馈的意见正是我们达到这一终极目标的重要帮助。华章公司欢迎老师和读者对我们的工作提出建议或给予指正,我们的联系方式如下:

华章网站: www.hzbook.com

电子邮件: hzjsj@hzbook.com

联系电话: (010) 88379604

联系地址: 北京市西城区百万庄南街1号

邮政编码: 100037



华章科技图书出版中心

人工智能 (Artificial Intelligence, AI) 是计算机学科的一个分支, 被认为是 21 世纪三大尖端技术 (基因工程、纳米科学、人工智能) 之一。“人工智能”一词最初是在 1956 年达特茅斯学院研讨会上提出的。从那以后, 研究者们发展了众多理论和原理, 人工智能的概念也随之扩展。在经历了辉煌和低谷之后, 近 30 年来, 人工智能的研究取得了迅速的发展, 在很多领域 (例如工程、医疗、财经、商业和管理等) 都获得了广泛应用。人工智能的研究逐渐成熟, 已成为一个独立的分支, 无论在理论和实践上都已自成一个系统。

本书作者 Michael Negnevitsky 是澳大利亚塔斯马尼亚大学电气工程和计算机科学系教授。他的许多研究课题都涉及人工智能和软计算, 一直致力于电气工程、过程控制和环境工程中智能系统的开发和应用。他著有 200 多篇论文、两本书, 并获得了四项发明专利。

本书是一本很好的人工智能入门书籍, 内容丰富、浅显易懂、适应面广。美国、德国、日本等多所大学的计算机相关专业都采用本书作为教材或主要教学参考书。与第 2 版相比, 第 3 版引入了关于“数据挖掘”的新一章, 并展示了智能工具在解决复杂实际问题中的新应用。第 3 版还扩展了本书的参考文献和参考书目, 并更新了附录中的人工智能工具和厂商列表。通过本书的学习, 相信读者一定会对人工智能的完整知识体系有全面的了解。

陈薇组织并参与了本书的翻译和审校工作, 参加翻译工作的还有闫秋玲、黄威靖、欧高炎、刘璐。由于译者水平有限, 译文中疏漏和错误之处在所难免, 欢迎广大读者批评指正。

陈 薇

2012 年于北京大学

这本书的主要目的与第1版相同，即为读者提供一本能实际了解计算机智能领域相关知识的书。它适合作为一个学期课程的入门教程，学生需要具备一些微积分的知识，不要求具备编程的经验。

在涵盖内容方面，本书引入了关于数据挖掘的新的一章，并展示了智能工具在解决复杂实际问题中的新应用。主要的变化如下：

在新的“数据挖掘和知识发现”一章中，我们介绍了大型数据库中知识发现不可或缺的一部分——数据挖掘。涉及将数据转换为知识的主要技术和工具，包括统计方法、数据可视化工具、结构化查询语言、决策树和购物篮分析。同时还展示了几个数据挖掘应用的实例。

在第9章中，增加了采用自组织神经网络进行聚类分析的新的实例。

最后，我们还扩展了本书的参考文献和参考书目，更新了附录中的人工智能工具和厂商列表。

Michael Negnevitsky
2010年9月于澳大利亚
塔斯马尼亚州 霍巴特市

“The only way not to succeed is not to try.”

——Edward Teller

又是一本人工智能的书……我已经见过很多同类的书，为什么还要理会它？它有什么与众不同之处？

每年，有成百上千本书和博士论文拓展着计算机或人工智能的知识体系。专家系统、人工神经网络、模糊系统以及进化计算是应用于智能系统的主要技术，数百个工具支持着这些技术，数以千计的科学论文不断推进着该学科的发展。本书中的任何章节的内容都可以作为一本书的主题。然而，我想写一本能够阐述智能系统基础的书，更为重要的原因是，我想消除大家对人工智能理论的畏惧心理。

大多数人工智能文献是采用计算机科学的专业术语进行描述的，其中充斥着大量复杂的矩阵代数和微分方程，这当然给人工智能理论带来了令人敬佩的资本，但同时也令非计算机科学的科学家对其敬而远之。不过，这种情况已经有所改变！

个人电脑已经成为我们日常生活中不可或缺的部分，我们将它用于打字机和计算器、日历和通信系统、交互式数据库以及决策支持系统。并且我们还渴望更多。我们希望计算机智能化！我们发现智能系统正快速地走出实验室，而我们也想更好地利用它。

智能系统的原理是什么？它是如何构建的？智能系统的用处是什么？我们该如何选择适当的工具构建智能系统？这些问题都可以在本书中找到答案。

与许多介绍计算机智能的书不同，本书将智能系统背后的奥妙简单明了地展示给读者。它是基于作者多年来给没有多少微积分知识的学生授课时所用的讲义而编写的，读者甚至不用学习任何编程语言就可轻松理解！书中的素材已经经过笔者 15 年的教学实践的检验，写作中也考虑了学生们提出的典型问题和建议。

本书是一本计算机智能领域的入门书籍，内容包括基于规则的专家系统、模糊专家系统、基于框架的专家系统、人工神经网络、进化计算、混合智能系统、知识工程和数据挖掘。

总体来说，本书可作为计算机科学、计算机信息系统和工程专业的本科生的入门教材。我在教学过程中，会要求学生开发小型的基于规则和基于框架的专家系统，设计一个模糊系统，探究人工神经网络，采用遗传算法求解一个简单的优化问题，并开发混合的神经模糊系统。他们使用一些专家系统的核心程序（XpertRule、Exsys Corvid 和 Visual Rule Studio）、MATLAB 的模糊逻辑工具箱以及 MATLAB 的神经网络工具箱。我们选择这些工具的原因是它们能够便利地演示教学中的原理。然而，本书并不局限于任何特定的工具，书中给出的例子可以轻松地不同的工具中实现。

本书也适合非计算机专业的相关人士自学。对于他们来说，本书提供了进入基于知识

的系统和计算智能的前沿领域的钥匙。事实上，本书面向的专业读者群十分广泛：工程师和科学家、管理人员和商人、医生和律师，也就是所有面临挑战而无法用传统的方法解决问题的人，所有想了解计算机智能领域巨大成就的人。本书将帮助你实际了解智能系统的用途，发现与你的工作密切相关的工具，并最终学会如何使用这些工具。

希望读者能与我共同分享人工智能和软计算学科所带来的乐趣，并从本书中获益。

读者可以访问 <http://www.booksites.net/negnevitsky> 获得更多的信息。

Michael Negnevitsky
2001 年 2 月于澳大利亚
塔斯马尼亚州 霍巴特市

本书一共包含 10 章。

第 1 章简要介绍了人工智能的历史，从 20 世纪中期诞生人工智能的思想并在 60 年代设立了远大目标并积极实现到 20 世纪 70 年代早期理想破灭和资金投入大幅削减；从 20 世纪 70 年代第一代专家系统（DENDRAL、MYCIN、PROSPECTOR）的诞生到 20 世纪八九十年代专家系统技术的成熟和其在不同领域的广泛应用；从 20 世纪 40 年代简单的二元神经元模型的提出到 20 世纪 80 年代人工神经网络领域的复苏；从 20 世纪 60 年代模糊集理论的提出和它的存在被西方忽视到 20 世纪 80 年代日本生产出大量模糊用户产品，以及 20 世纪 90 年代软计算和文字计算在世界范围内的广泛接受。

第 2 章提供了基于规则的专家系统概述。作者简要介绍了知识的概念，以及专家用产生式规则表示知识的过程。作者介绍了专家系统开发团队的主要成员和基于规则系统的结构。分析了专家系统的基本特性，并指出专家系统不是万无一失的。然后回顾了前向链接和后向链接推理技术，并讨论了冲突消解策略。最后分析了基于规则的专家系统的优缺点。

第 3 章展示了专家系统使用的两种不确定性管理技术：贝叶斯推理和确信因子。作者分析了不确定知识的主要来源，并简要回顾了概率理论。作者考虑了可累积论据的贝叶斯方法并开发一个简单的基于贝叶斯方法的专家系统。然后讨论确信因子理论（贝叶斯推理的常用替代方法），并开发一个基于论据推理的专家系统。最后，比较贝叶斯推理和确信因子理论并分析它们的适用范围。

第 4 章介绍了模糊逻辑并讨论其背后的哲学思想。首先介绍模糊集的概念，考虑如何在计算机里表示一个模糊集并介绍了模糊集的操作。作者还定义了语言变量和模糊限制语（hedge）。然后作者叙述了模糊规则，并解释了经典规则和模糊规则的主要区别。该章主要研究两种模糊推理技术：Mamdani 法和 Sugeno 法，并就它们适宜的应用领域给出建议。最后介绍开发一个模糊专家系统的主要步骤，并通过构建和调试模糊系统的具体过程来阐明其理论。

第 5 章概述了基于框架的专家系统。介绍了框架的概念，并讨论如何将框架用于知识表达，以及阐明继承是基于框架系统的基本特征，还讨论了方法、守护程序和规则的应用。最后通过一个实例来介绍基于框架的专家系统的开发。

第 6 章介绍了人工神经网络，并讨论了机器学习的基本思想。叙述作为一个简单的计算单元的感知器的概念，并讨论了感知器的学习规则，还探索了多层神经网络，以及讨论了如何提高反向传播学习算法的计算效率。然后介绍循环神经网络，思考 Hopfield 网络训练算法和双向联想记忆（BAM）。最后介绍自组织神经网络并探讨 Hebbian 学习规则和竞争学习。

第 7 章概述了进化计算，其中包括遗传算法、进化策略和遗传编程。首先介绍了开发一个遗传算法的主要步骤，讨论遗传算法的工作机制，并通过具体的遗传算法应用阐明其理论。然后叙述一个进化策略的基本概念，比较了进化策略和遗传算法之间的不同。最后考虑遗传编程以及它的实际应用。

第 8 章讨论了结合不同智能技术的混合智能系统。首先介绍一种新型的专家系统——神经专

家系统，它将神经网络和基于规则的专家系统结合起来。然后考虑一个功能上等同于 Mamdani 模糊推理模型的神经 - 模糊系统，以及一个功能上等同于 Sugeno 模糊推理模型的自适应神经模糊推理系统 (ANFIS)。最后讨论进化神经网络和模糊进化系统。

第 9 章讨论了知识工程。首先讨论智能系统可以解决什么样的问题，并介绍知识工程过程的 6 个主要阶段。然后展示了专家系统、模糊系统、神经网络和遗传算法的典型应用。作者演示了如何构建智能系统来解决诊断、选择、预测、分类、聚类和优化问题。最后讨论了混合神经 - 模糊系统在决策支持和时序预测方面的应用。

第 10 章介绍了数据挖掘的概貌，讨论了将数据转换为知识的主要技术。首先，宽泛地定义了数据挖掘，并解释了在大型数据库中进行数据挖掘和知识发现的过程。介绍了一些统计学方法，包括主成分分析，并讨论了它们的局限性。随后展示了关系数据库中结构化查询语言的应用，介绍了数据仓库和多维数据分析。最后，介绍了最流行的数据挖掘工具——决策树和购物篮分析。

本书还包括一个术语表和一个附录。术语表包含了 300 余条用于专家系统、模糊逻辑、神经网络、进化计算、知识工程以及数据挖掘领域的定义。附录中提供了商业化的人工智能工具列表。

本书的网站地址为 <http://www.booksites.net/negnevitsky>。

在本书的出版过程中，我直接或间接地得到了很多人至诚的帮助。我首先要感谢 Vitaly Fay-bisovich 博士，他不仅对我在软计算领域的研究提出了许多建设性的意见，还在我过去 30 年间的奋斗生涯中与我建立了真诚的友谊并给予了我大力支持。

我还要感谢许许多多为本书提出评论和有益建议的检阅人，感谢 Pearson Education 的编辑们，特别是 Keith Mansfield、Owen Knight、Liz Johnson 和 Rufus Curnow，他们帮助我完成了本书的出版发行工作。

我要感谢我在澳大利亚塔斯马尼亚大学的本科生和研究生们，特别是我以前的博士生 Tan Loc Le、Quang Ha、Steven Carter 和 Mark Lim，他们对新知识的渴求对我来说既是挑战也是激励。

我想感谢美国波士顿大学的 Stephen Grossberg 教授、德国马德堡大学的 Frank Palis 教授、日本广岛大学的 Hiroshi Sasaki 教授、美国罗切斯特理工学院的 Walter Wolf 教授以及东京工业大学的 Kaoru Hirota 教授，他们为我提供了在学生中对书的内容进行测试的机会。

我还要衷心地感谢 Vivienne Mawson 和 Margaret Eldridge 博士对本书初稿的审读工作。

尽管本书的第 1 版出版了十多年，但是已经有很多人用过它并向我提出了宝贵的意见和建议，由于人数众多，在此不能一一致谢，但我至少要感谢那些提出了特别有益建议的人们：Martin Beck（英国普利茅斯大学）、Mike Brooks（澳大利亚阿德莱德大学）、Genard Catalano（美国哥伦比亚大学）、Warren du Plessis（南非比勒陀利亚大学）、Salah Amin Elewa（埃及美洲大学）、Michael Fang（中国浙江大学）、John Fronckowiak（美国曼达尔学院）、Patrick B. Gibson（加拿大温莎大学）、Lev Goldfarb（加拿大 New Brunswick 大学）、Susan Haller（美国威斯康星大学）、Evor Hines（英国沃里克大学）、Philip Hingston（澳大利亚埃迪斯科文大学）、Sam Hui（美国斯坦福大学）、Yong - Hyuk Kim（韩国光云大学）、David Lee（英国赫特福德郡大学）、Andrew Nunekpeku（加纳大学）、Vasile Palade（英国牛津大学）、Leon Reznik（美国罗切斯特理工学院）、Simon Shiu（中国香港理工大学）、Boris Stilman（美国科罗拉多大学）、Thomas Uthmann（德国美因茨约翰尼斯·古腾堡大学）、Anne Venables（澳大利亚维多利亚大学）、Brigitte Verdonk（比利时安特卫普大学）、Ken Vollmar（美国西南密苏里州立大学）、Kok Wai Wong（新加坡南洋理工大学）以及 Georgios N. Yannakakis（丹麦哥本哈根信息技术大学）。

出版者的话	
译者序	
第3版前言	
第1版前言	
本书概要	
致谢	
第1章 基于知识的智能系统概述	1
1.1 智能机	1
1.2 人工智能的发展历史, 从“黑暗时代”到基于知识的系统	3
1.2.1 “黑暗时代”, 人工智能的诞生 (1943~1956年)	3
1.2.2 人工智能的上升期, 远大目标积极实现的年代 (1956年~20世纪60年代晚期)	4
1.2.3 没有履行的诺言, 来自现实的冲击 (20世纪60年代晚期~20世纪70年代早期)	5
1.2.4 专家系统技术, 成功的关键因素 (20世纪70年代早期~20世纪80年代中期)	5
1.2.5 如何使机器学习, 神经网络的重生 (20世纪80年代中期至今)	8
1.2.6 进化计算, 在尝试中学习 (20世纪70年代早期至今)	9
1.2.7 知识工程的新纪元, 文字计算 (20世纪80年代后期至今)	9
1.3 小结	11
复习题	13
参考文献	13
第2章 基于规则的专家系统	16
2.1 知识概述	16
2.2 知识表达技术——规则	16
2.3 专家系统研发团队的主要参与者	18
2.4 基于规则的专家系统的结构	19
2.5 专家系统的基本特征	20
2.6 前向链接和后向链接推理技术	21
2.6.1 前向链接	22
2.6.2 后向链接	23
2.7 MEDIA ADVISOR: 基于规则的专家系统实例	25
2.8 冲突消解	29
2.9 基于规则的专家系统的优点和缺点	31
2.10 小结	32
复习题	33
参考文献	34
第3章 基于规则的专家系统中的不确定性管理	35
3.1 不确定性简介	35
3.2 概率论基本知识	36
3.3 贝叶斯推理	39
3.4 FORECAST: 论据累积的贝叶斯方法	41
3.5 贝叶斯方法的偏差	46
3.6 确信因子理论和基于论据的推理	47
3.7 FORECAST: 确信因子的应用	51
3.8 贝叶斯推理和确信因子的对比	52
3.9 小结	53
复习题	53
参考文献	54
第4章 模糊专家系统	56
4.1 概述	56
4.2 模糊集	57
4.3 语言变量和模糊限制语	60
4.4 模糊集的操作	63
4.5 模糊规则	66
4.6 模糊推理	68
4.6.1 Mamdani-style 推理	68

4.6.2 Sugeno-style 推理	73	参考书目	169
4.7 建立模糊专家系统	75	第8章 混合智能系统	170
4.8 小结	82	8.1 概述	170
复习题	83	8.2 神经专家系统	171
参考文献	83	8.3 神经-模糊系统	176
参考书目	84	8.4 ANFIS	182
第5章 基于框架的专家系统	86	8.5 进化神经网络	188
5.1 框架简介	86	8.6 模糊进化系统	192
5.2 知识表达技术——框架	87	8.7 小结	195
5.3 基于框架的系统中的继承	91	复习题	196
5.4 方法和守护程序	94	参考文献	197
5.5 框架和规则的交互	97	第9章 知识工程	198
5.6 基于框架的专家系统实例: Buy Smart ..	99	9.1 知识工程简介	198
5.7 小结	108	9.1.1 问题评估	198
复习题	109	9.1.2 数据和知识获取	199
参考文献	109	9.1.3 原型系统开发	200
参考书目	110	9.1.4 完整系统开发	201
第6章 人工神经网络	111	9.1.5 系统评价和修订	201
6.1 人脑工作机制简介	111	9.1.6 系统集成和维护	201
6.2 作为简单计算元素的神经元	113	9.2 专家系统可以解决的问题	202
6.3 感知器	114	9.3 模糊专家系统可以解决的问题	209
6.4 多层神经网络	117	9.4 神经网络可以解决的问题	214
6.5 多层神经网络的加速学习	123	9.5 遗传算法可以解决的问题	226
6.6 Hopfield 网络	126	9.6 混合智能系统可以解决的问题	229
6.7 双向联想记忆	131	9.7 小结	236
6.8 自组织神经网络	133	复习题	237
6.8.1 Hebbian 学习	133	参考文献	239
6.8.2 竞争学习	136	第10章 数据挖掘和知识发现	241
6.9 小结	141	10.1 数据挖掘简介	241
复习题	143	10.2 统计方法和数据可视化	243
参考文献	143	10.3 主成分分析	247
第7章 进化计算	145	10.4 关系数据库和数据库查询	255
7.1 进化是智能的吗	145	10.5 数据仓库和多维数据分析	258
7.2 模拟自然进化	145	10.6 决策树	265
7.3 遗传算法	146	10.7 关联规则和购物篮分析	271
7.4 遗传算法为什么可行	153	10.8 小结	277
7.5 案例研究: 用遗传算法来维护调度	154	复习题	278
7.6 进化策略	160	参考文献	279
7.7 遗传编程	161	术语表	281
7.8 小结	167	附录 人工智能工具和经销商	295
复习题	167	索引	310
参考文献	168		

基于知识的智能系统概述

本章讨论何为智能以及机器能否真正实现智能化。

1.1 智能机

哲学家们两千多年来一直致力于理解和解决宇宙的两大问题：人类是如何思考的？人类以外的物体是否也能够思考？至今，这两个问题依旧没有答案。

一部分哲学家接受由计算机领域的科学家提出的计算方法，认可人类所做的任何事情机器都能够做到的观点。另一部分哲学家则非常反对这一观点，认为一些高级复杂的行为（例如爱、创新、道德判断）远超出了任何机器的能力范围。

哲学的特质允许这些争论一直存在。但事实上，工程师和科学家们已经造出了具有“智能”的机器。何为“智能”？下面是来自《柯林斯基础英语词典》的定义：

- (1) 智能是人类理解和学习事情的能力。
- (2) 智能是思考和理解事情的能力，而非本能或机械地做事情。

——《柯林斯基础英语词典》2008 年版

按照第一条定义，智能是专属于人类的一个品质，但第二条定义更具弹性，指出只要具有思考和理解能力就具有智能，而不限主体是人还是物。而定义里的“思考”是什么意思呢？再次看词典中的定义：

思考是使用大脑考虑问题或创建新想法的活动。

——《柯林斯基础英语词典》2008 年版

为了能够思考，人或物需要拥有像大脑那样的一个器官，依赖它去学习、理解事物，去处理问题和做出决策。因此智能可以定义为“学习和理解事物、处理问题并做出决策的能力”。

计算机能否智能化，即机器能否思考，这一问题源于人工智能的“黑暗时代”（20 世纪 40 年代晚期开始）。作为一门科学，人工智能（Artificial Intelligence, AI）的目标是使机器像人那样具有智能去做事情（Boden, 1977）。因此，关于“机器能否思考”这一问题的答案对人工智能这一学科至关重要。答案本身不是“能”或“不能”之类的简单判断，而必定是模糊的。我们从常识中也能得出这个结论：比如一部分人在某些方面比其他聪明；日常生活中做出的抉择有些是明智的，有些则很愚蠢的；有人可能精通复杂的数学、工程问题，但对于哲学或历史却一窍不通；有人善于赚钱，而有人更善于花钱。虽然人类中的每一个成员都具有学习和理解事物、处理问题并做出决策的能力，但比较起来，这种能力在每一个领域都不均等。所以，如果机器能够思考，一部分机器在某些方面也许比另一部分机器聪明。

英国数学家阿兰·图灵（Alan Turing）在 50 多年前撰写的论文“计算机器和智能化”（Turing, 1950）是关于机器智能方面最早和最有影响的文章之一。图灵的思想经住了时间的考验，到现在为止仍是通用的。

阿兰·图灵在 20 世纪 30 年代早期重新发现了中心极限定理，他也由此开始了科学生涯。1937 年，他发表了可计算数字的论文，在这篇文章中提出了通用机器的概念。之后在二战期间，他在破解德国军用编密码机 Enigma 的工作中担任了重要角色。第二次世界大战结束后，他设计了“自动计算机器”，并设计了第一个国际象棋比赛程序，后来在曼彻斯特大学的计算机上得以实

现。图灵关于通用机的理论概念和译码的实践经验使他在人工智能的关键性基础问题方面取得进展。他曾经提出这样的问题：是否存在不需经验的思考？是否存在不需交流的思考？是否存在不使用的语言？是否存在生命之外的智能？显然，这些问题都是人工智能中的基础问题——“机器能否思考？”的不同表达。

图灵并没有对机器和思考进行定义，而是设计了一个游戏——图灵模拟游戏，用游戏阐述他对机器和思考的理解，从而避免了关于定义的异议。图灵认为，与其问“机器能否思考”，不如问“机器能否通过智能行为测试”。他曾经预测，到公元 2000 年时，计算机能够通过程序与人进行 5 分钟对话，并有 30% 的机会使对话的人认为对方不是计算机，而是人。图灵将计算机智能行为定义为在认知任务中达到人类行为水平的能力。换句话说，如果对话者依据所提问题的答案无法判断对方是人还是机器，计算机就通过了测试。

图灵模拟游戏包括两个阶段。在第一阶段，如图 1.1 所示，一个审讯员和一男一女分别被安排在单独的空间，他们只能通过远程终端之类的中介进行交流。审讯员的目标是通过对另外两个人提问，根据他们的回答来判断谁是男的、谁是女的。游戏规则是男士要尽力让审讯员认为他是女士，而女士则需向审讯员证明她是女士。

游戏的第二个阶段，如图 1.2 所示，用一台计算机代替男士，计算机的任务和第一阶段中男士的任务一样。对计算机编程时，为了更好地模拟人类，故意让计算机犯错或提供模糊的答案。如果计算机成功“骗过”审讯员的次数和男士成功“骗过”审讯员的次数一致，则认为计算机通过了智能行为测试。

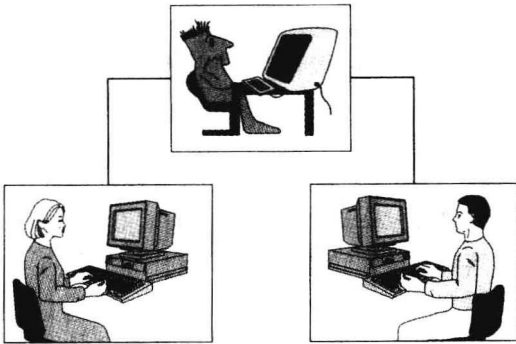


图 1.1 图灵模拟游戏：第一阶段

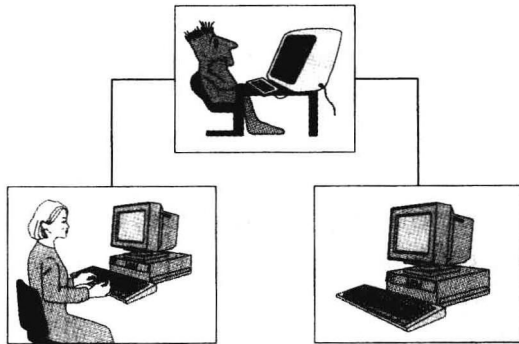


图 1.2 图灵模拟游戏：第二阶段

对于人的身体模拟在智能测试中并不重要。因此，图灵测试中，审讯员看不见、触不到、听不到计算机，因此不会受计算机的外表或声音的干扰。但是为了辨别出计算机，审讯员可以问任何问题，甚至是过激的问题。例如，审讯员可以让人和计算机做复杂的数学计算，如果一方提供了正确答案，并且比对方速度快，可以判断这一方就是计算机。因此，计算机必须知道何时故意出错、何时延迟回答。审讯员也可能会问双方关于短篇小说、诗歌、绘画方面的问题，由此可以推断出一方具有的明显的人类专属的情绪特征。当面对这一类问题时，计算机就必须模拟人类对事物的感性理解。

图灵测试很通用，原因是测试具备了两个显著特征：

- 由于使用终端进行人机交互，这一测试让我们对智能有客观的衡量标准，避免了对人类智能本质的争论，并排除了偏向人类的可能性。
- 测试本身独立于实验细节。既可以像前述的进行两阶段测试，也可以变成一个阶段的测试，从一开始就让审讯员对人与机器进行判断。审讯员既可以提任何领域的任何问题，也可以只关注答案本身。

图灵相信到20世纪末,能够通过对数字计算机编程来完成测试游戏。尽管现代计算机还没有通过图灵测试,但已经为我们提供了基于知识的系统的评估和验证基础。通过与人类专家行为对比,已经对一些狭窄专业领域的具有智能的程序进行了评估。

我们的大脑具有 10^{18} 位的容量,能够每秒处理 10^{15} 位的信息。到2020年,也许一块方糖大小的芯片就足以模拟人脑,也许会有计算机能够通过图灵测试游戏。但是,我们真的想让计算机在做数学计算时,像人一样又慢又会出错吗?实际上,智能机应该能够帮助人类进行决策、搜索信息、控制复杂对象,最终理解单词含义。也许没有必要片面追求使计算机像人一样聪明这一既抽象又艰涩的目标。为了构建一套智能机系统,我们必须去捕捉、组织和使用某个狭窄专业领域的人类专家知识。

1.2 人工智能的发展历史,从“黑暗时代”到基于知识的系统

人工智能成长为一门科学,历经三代研究者的努力。这一节就来介绍每一代研究者对人工智能学科所做的贡献和重大事件。

4

1.2.1 “黑暗时代”,人工智能的诞生(1943~1956年)

Warren McCulloch 和 Walter Pitts 在1943年发表了人工智能领域的开篇之作。McCulloch 拥有哥伦比亚大学的哲学和医学学位,后来在伊利诺伊大学的精神病学系担任基础研究实验室的主任。他在对中枢神经系统的研究中提出了大脑神经元模型,最早为人工智能做出了主要贡献。

McCulloch 和年轻的数学家 Walter Pitts 合作,提出了人工神经网络模型,这个模型假定每个神经元有两个状态:开、关(McCulloch and Pitts, 1943)。并证明他们的神经网络模型实际上等价于图灵机,并且任何可计算函数都可以由相连的神经网络进行计算。同时揭示了简单的网络结构具有学习功能。

神经网络模型从理论和实验两个方面激发了对大脑模型的研究。但是,实验结果清晰地表明将神经元限制为两个状态是不正确的。事实上,神经元具有明显的非线性特征,而非简单的二元状态。即使如此,McCulloch 是人工智能领域继阿兰·图灵之后的第二个创始人,为神经计算和人工智能网络(Artificial Neural Networks, ANN)奠定了基础。人工神经网络的研究在20世纪70年代经历了衰退,在20世纪80年代后期重新复苏。

人工智能的第三个奠基人是聪明的匈牙利籍数学家 John von Neumann。他在1930年加入了普林斯顿大学,在数学物理系任教。他是阿兰·图灵的同事和朋友。在第二次世界大战期间, von Neumann 在建造原子弹的“曼哈顿项目”中担任重要角色,并成为宾夕法尼亚大学的 ENIAC (Electronic Numerical Integrator and Calculator, 电子数字集成计算器)项目的顾问,参与设计了 EDVAC 计算机 (Electronic Discrete Variable Automatic Computer, 电子离散变量自动计算机)^①。McCulloch 和 Pitts 所提出的神经网络模型使 von Neumann 受到启发。因此,在1951年,当来自普林斯顿大学数学系的两个研究生 Marvin Minsky 和 Dean Edmonds 建造第一台神经网络计算机时, von Neumann 鼓励并资助了他们。

在第一代研究者中,不得不提 Claude Shannon。他毕业于麻省理工学院并于1941年加入贝尔实验室。Shannon 继承了阿兰·图灵的观点,认为可以使机器智能化。在1950年,他发表了一篇关于让机器下象棋的文章,文章指出一盘典型的棋赛大约有 10^{120} 次的移动(Shannon, 1950)。即

① EDVAC 是一个存储程序的机器。

使一台新的 von Neumann 式计算机每微秒就能测试一种走法，移动第一步也得花上 3×10^{106} 年。因此，Shannon 论证了在寻找解决方案时使用启发式策略的必要性。

另一位人工智能奠基人 John McCarthy 也来自普林斯顿大学。他从普林斯顿大学毕业后去达特茅斯学院工作。他说服了 Marvin Minsky 和 Claude Shannon 在达特茅斯学院组织一个暑期研讨会。在 1956 年，他们召集了对机器智能、人工神经网络和自动理论感兴趣的研究者，参加由 IBM 赞助的研讨会。尽管研讨会仅有 10 人参加，但这次会议却促成了人工智能学科的诞生。在之后的 20 年里，参加过达特茅斯学院研讨会的学者及他们的学生在人工智能领域一直起主导作用。

1.2.2 人工智能的上升期，远大目标积极实现的年代（1956 年 ~ 20 世纪 60 年代晚期）

在人工智能学科的发展早期，可以用热情巨大、想法大胆，但成果很少来形容。当时，计算机被用来做常规数学计算才几年时间，而人工智能领域的研究者却已在论证计算机的本领远超过这些。这真是一个有远大目标并积极付诸实现的年代。

John McCarthy 作为达特茅斯学院研讨会的组织者之一，也是“人工智能”这一术语的发明者，从达特茅斯学院转到了麻省理工学院。他设计了高级语言 LISP。LISP 是早期编程语言之一（FORTRAN 仅比它早了两年），到今天还有人在用。在 1958 年，McCarthy 发表了论文“基于常识的程序”，文中提出了一个名为“建议采纳者”（Active Taker）的程序，用于搜索一般性问题的解决方案（McCarthy, 1958）。McCarthy 论证了设计和生成程序的方法，他以驾车去机场为例，采用这个程序基于几个简单的规则设计了驾驶计划。更重要的是，程序无须变动，就可以接受来自不同专业领域的新的规则，即可以接受新知识。因此，“建议采纳者”是第一个真正集成了知识表达和推理的核心法则的基于知识的系统。

达特茅斯学院研讨会的另一个组织者 Marvin Minsky 也加入了麻省理工学院。他不像 McCarthy 那样专注于形式逻辑，而是建立了一套反逻辑的观点，以处理知识表达和推理。他的框架理论为知识工程做出了主要的贡献（Minsky, 1975）。

由 McCulloch 和 Pitts 开创的神经计算和人工神经网络的早期研究得以继续。随着学习方法的改进和提高，Frank Rosenblatt 证明了感知器收敛理论，并论证他的学习算法能够调整感知器的连接程度（Rosenblatt, 1962）。

GPS（General Problem Solver，通用问题解决方案）是在这个大胆假设年代最具野心的项目之一（Newell and Simon, 1961; 1972）。卡内基 - 梅隆大学的 Allen Newell 和 Herbert Simon 设计了一个通用程序，用于模拟人类在解决问题时的方法。GPS 也许是第一个尝试将解决问题的技术和数据分开的程序。它使用了现在已被称为手段 - 目的分析（means-ends analysis）的技术。Newell 和 Simon 假定可以用状态去定义问题。手段 - 目的的作用是判断问题的当前状态和期望状态或目标状态的差距，并选择操作算子去达到目标状态。如果从当前状态不能直接到达目标状态，就要建立一个更接近目标状态的新状态，重复这一过程直到达到目标状态。操作算子集合决定了解决方案。

但是，GPS 无法解决复杂问题。由于程序建立在形式逻辑的基础上，操作算子可能有无穷多个，导致效率低下。在解决实际问题时，GPS 需要大量运算时间和内存，这导致人们后来放弃 GPS 计划。

总之，在 20 世纪 60 年代，为了模拟复杂的思考过程，人工智能研究者试图研究通用方法来解决广泛的问题。他们使用了通用搜索策略以寻找解决方案。通用搜索策略现在被称为弱方法，由于使用的是问题域的弱信息，导致程序的性能低下。