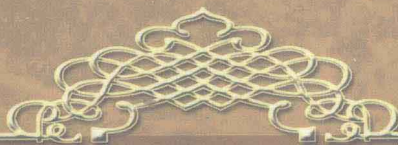
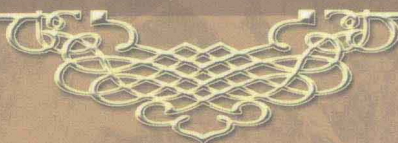


面向计算机科学与技术专业规范系列教材



人工智能

刘凤岐 编著

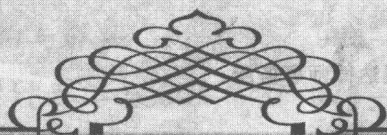


Artificial Intelligence



机械工业出版社
China Machine Press

面向计算机科学与技术专业规范系列教材



人工智能

刘凤岐 编著

Artificial Intelligence



机械工业出版社
China Machine Press

本书较全面地介绍了人工智能领域的各种课题，其中包括使用神经网络和遗传算法的感知和适应性、附有本体论的智能 agent、自动推理、自然语言分析，以及机器学习的随机途径。论述的许多技术和策略可用于应对现今计算机科学面临的众多挑战。此外，本书进一步从哲学、心理学和神经生理学等角度讨论了人工智能研究工作。

本书可作为大学本科生和研究生的教材，还可供从事人工智能研究和应用的专业人员参考。

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

图书在版编目 (CIP) 数据

人工智能 / 刘凤岐编著. — 北京: 机械工业出版社, 2011.6
(面向计算机科学与技术专业规范系列教材)

ISBN 978-7-111-34970-9

I. 人… II. 刘… III. 人工智能—教材 IV. TP18

中国版本图书馆 CIP 数据核字 (2011) 第 105467 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: 刘立卿

北京诚信伟业印刷有限公司印刷

2011 年 7 月第 1 版第 1 次印刷

185mm×260mm • 22.5 印张

标准书号: ISBN 978-7-111-34970-9

定价: 38.00 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: hzsj@hzbook.com

出版者的话

机械工业出版社华章公司是国内重要的教育出版公司，培生教育集团（拥有 Addison Wesley、Prentice Hall 等品牌）是全球知名的教育出版集团，双方在过去长达十余年的合作中秉承“全球采集内容，服务教育事业”的理念，遴选、移译了国外大量的在计算机科学界享誉盛名的专家名著与名校教材，其中包括 Donald E. Knuth、Alfred V. Aho、Jeffrey D. Ullman、John E. Hopcroft、Dennis Ritchie 等大师名家的经典作品（收录在大理石封面的“计算机科学丛书”中），这些作品对国内计算机教育及科研事业的发展起到了积极的促进作用。

随着国内计算机科学与技术专业学科建设的不断完善、教学研究的蓬勃发展，以及教材改革的逐渐深化，计算机科学与技术专业的优秀课程及教材不仅仅是“引进来”（版权引进），而且需要“走出去”（版权输出）了。

近几年以来，教育部计算机科学与技术专业教学指导分委员会根据我国计算机专业教育的现状以及社会对人才的需求，发布了《高等学校计算机科学与技术专业发展战略研究报告暨专业规范（试行）》（以下简称《规范》）。为配合《规范》的实施推广，同时为落实中央“提高高等教育质量”的最新指导思想，在教育部计算机科学与技术专业教学指导分委员会的指导下，在国内知名高校众多教授的帮助下，我们出版了这套“面向计算机科学与技术专业规范系列教材”。

本套教材的作者在长达数十年的科研和教学经历中积累了大量的知识和经验，也奠定了他们在学术和教学领域的地位，教材的内容体现了他们的教学思想和教学理念，本套教材也是传承他们优秀教学成果的最好载体，是中国版的专家名著和名校教材，相信它们的出版对提高计算机科学与技术专业的教育水平和教学质量能够起到积极的作用。

华章与培生作为专业的出版团队，愿与高等院校的老师共同携手，在这套教材的出版上引进国际先进教材出版经验，在教学配套资源的建设上做出新的尝试，为促进中国计算机科学与技术专业教育事业的发展，为增进中国与世界文化的交流而努力。



华章教育



培生教育集团

序 言

近 20 年里, 计算机学科有了很大的发展, 人们普遍认为, “计算机科学”这个名字已经难以涵盖该学科的内容, 因此, 改称其为计算学科 (Computing Discipline)。在我国本科教育中, 1996 年以前曾经有计算机软件专业和计算机及应用专业, 之后被合并为计算机科学与技术专业。2004 年以来, 教育部计算机科学与技术专业教学指导分委员会根据我国计算机专业教育和计算学科的现状, 为更好地满足社会对计算机专业人才的需求, 发布了《高等学校计算机科学与技术专业发展战略研究报告暨专业规范 (试行)》(以下简称《规范》), 提出在计算机科学与技术专业名称之下, 构建计算机科学、计算机工程、软件工程和信息技术四大专业方向。《规范》中四大专业方向的分类, 在于鼓励办学单位根据自己的情况设定不同的培养方案, 以培养更具针对性和特色的计算机专业人才。

为配合《规范》的实施, 落实中央“提高高等教育质量”的精神, 我们规划了“面向计算机科学与技术专业规范系列教材”。本系列教材面向全新的计算学科, 针对我国高等院校逐步向新的计算机科学与技术专业课程体系过渡的趋势编写, 在知识选择、内容组织和教学方法等方面满足《规范》的要求, 并与国际接轨。本套教材具有以下几个特点:

(1) **体现《规范》的基本思想, 满足其课程要求。**为使教材符合我国高等院校的教学实际, 编委会根据《规范》的要求规划本套教材, 广泛征集在国内知名高校中从事一线教学和科研工作、经验丰富的优秀教师承担编写任务。

(2) **围绕“提高教育质量”的宗旨开发教材。**为了确保“精品”, 本系列教材的出版不走盲目扩大的路子, 每本教材的选题都将由编委会集体论证, 并由一名编委担任责任编委, 最大程度地保证这套教材的编写水准和出版质量。

(3) **教材内容的组织科学、合理, 体系得当。**本套教材的编写注重研究学科的新发展和新成果, 能够根据不同类型人才培养需求, 合理地进行内容取舍、组织和叙述, 还精心设计了配套的实验体系和练习体系。

(4) **教材风格鲜明。**本套教材按 4 个专业方向统一规划, 分批组织, 陆续出版。教材的编写体现了现代教育理念, 探讨先进的教学方法。

(5) **开展教材立体化建设。**根据需要配合主教材的建设适时开发实验教材、教师参考书、学生参考书、电子参考资料等教辅资源, 为教学实现多方位服务。

我们衷心希望本系列教材能够为我国高等院校计算机科学与技术等专业的教学作出贡献, 欢迎广大读者广为选用。

“面向计算机科学与技术专业规范系列教材”编委会

面向计算机科学与技术专业规范系列教材

编委会

主任委员： 蒋宗礼

副主任委员： 王志英 钱乐秋

委员：（以姓氏拼音为序）

陈道蓄	陈 明
傅育熙	何炎祥
黄刘生	贾云得
姜守旭	李仁发
李晓明	刘 辰
马殿富	齐 勇
孙吉贵	孙茂松
吴功宜	吴 跃
谢长生	于 戈
张 钢	周兴社

秘书组： 温莉芳 刘立卿 姚 蕾

本书责任编辑： 钱乐秋



前言

尽管人工智能(AI)与大部分工程学科相同,必须对实际问题提出具体解决方法,在商业世界中证明自己,但我们希望探讨和理解使智力思维与活动成为可能的智能的机理。国学大师辜鸿铭先生(1857—1928)在上世纪之初就指出了“……人与动物的区别何在?就在于智能”^[26],对智能给出了界定,但直到电子数字计算机诞生之后才有可能精确探讨智能的属性。若对智能有了深入开发和探讨,将使我们关于世界的知识走向更高水准。我们相信由设计与评价智能人造物能有效研究可能智能的空间。本书将拒绝“智能是人类独有的能力”这一稍显狭隘的观念。

上世纪末期, AI 界在“干净派”与“邋遢派”间曾发生过激烈争论,前者强调智能形式理论的详细阐述与分析;后者相信智能本身是某类大构架,可很好地按应用驱动的特定方式逼近。尽管两派间的争论已引申出许多其他争论,包括物理符号系统的支持者与神经网络的研究者,逻辑学家与多半按非逻辑方式发展出的人工生命形式的设计者,专家系统的设计者与基于事例的推理器的研究者,最后是相信 AI 已达到的人与相信 AI 绝不会发生的人,但已证明这种两分法过分简单。应该客观地看到,这些争论对 AI 的发展起到了积极推动作用,促进了 AI 研究人员对智能属性的深思,扩展了 AI 的研究范围。

智能用任何单一理论描述都太复杂,研究人员们已构造一理论层次,在多层抽象上表征智能。在该层次的最低层上,神经网络、遗传算法和其他形式的浮现计算使我们能理解适应、感知、体现,以及与物理世界交互的过程,这些必须是任何形式的智力活动的基础。工作在更高层上的逻辑学家们在亚里士多德开创的更强调理性的逻辑推断模式基础上,已探索出演绎、反绎、归纳、真值维护和无数其他推理模式与方式的轮廓。在进一步的更高抽象层上,诊断系统、智能 agent 和自然语言理解程序的设计者们已逐渐认识到社会进程在建立、传播和延续知识中的作用。

人工智能是一门边缘科学,要涉及哲学、心理学、社会学、神经生理学等领域。可以认为形式主义(或唯理主义)和经验主义支配了 AI 的先前发展,但近年来的研究已浮现出智能体系的激发兴趣的图像。

在 AI 事业中,唯理主义和经验主义极端似乎只导致了有限结果,两极端经受了有限的适用性和普遍性。本书将兼顾经验主义的修整(语义网、脚本、包容体系)与唯理主义的清晰、明确的观念(谓词演算、非单调逻辑、自动推理),并且给出基于贝叶斯的另一观点。关系不变性(规律性)的体验制约智能 agent 的期望,并且学习这些规律性,这本身又有偏向未来的期望。哲学家们有责任评论 AI 事业认识论的有效性。在本书第 15 章将讨论唯理主义的设想、经验主义的困境,并且介绍基于贝叶斯的构造主义和谐状态。本书中将涉及当前 AI 事业中的所有这些层面。

本书的一宗旨是,在 AI 方法学中高级表示形式体系和搜索技术的中心地位。这恐怕是本书和 AI 的许多早期工作的最有争议的方面,在浮现计算中的许多研究人员质疑是否符号推理

和指称语义在智能中全然有任何作用。虽然表示作为对事物给出名称的思想已受到由神经网络或人工生命的浮现模式提供的隐式表示的挑战，但我们相信理解表示和搜索对任何严肃的人工智能实践者仍是实质的。我们还认为第 1 章中对 AI 历史传统和先驱者们的综述是 AI 教育的关键部分。此外，这些对分析像神经网络的表达能力或穿过遗传算法的适合度情况的候选问题解的进展之类的非符号 AI 方面是无法估价的工具。在第 15 章中将给出对现代 AI 的比较和评论。

本书的另一宗旨是把人工智能安排在实验科学、经验学科范围中。按 Newell 和 Simon 图灵奖演说的精神^[16]，我们引用：

……AI 不是对科学传统的某种离奇的偏离，而是关于智能本身的知识，以及理解智能的一般探讨的一部分。此外，我们的 AI 程序设计工具，与探索式程序设计方法学一起……对探索一环境是理想的。我们的工具为理解与质疑给出了一种手段。我们构造性地达到评价与认识若干现象，即由逐步逼近。

于是我们把每种设计与程序都视为与大自然的一种实验：我们提出一种表示，产生一种搜索算法，并且询问我们对部分智力现象的解释的刻画的适宜性。而自然界对我们的询问给出响应。我们的实验可以解构、修正、扩展，并且再运行。我们的模型可以精炼，我们的理解可以扩展。

下面给出本书各章及附录的内容概要。

第 1 章由试图按哲学、心理学及其他研究领域理解智能的简史开始。按一种重要意义，AI 是一门老科学，其根源至少可追溯到亚里士多德。对该背景的正确评价对理解现代研究致力的问题是实质的。本章还给出 AI 中某些重要应用领域的综述。本章的目标是为以后的理论和应用提供背景和动机。

第 2 章中简要介绍作为表示与搜索的 AI，试图使读者对 AI 的该重要研究途径有某些直觉，并且介绍 AI 研究的初衷。第 3 章介绍既作为数学系统也作为描述问题实质特征的表示语言的谓词演算。第 4 章介绍搜索，以及用于实现搜索的算法和数据结构，组织对问题状况的探索。第 5 章讨论启发式在聚焦和约束基于搜索的问题求解中的实质作用。第 6 章介绍实现这些搜索算法的若干软件体系，包括黑板和产生式系统。

第 7 章介绍 AI 表示模式的发展历史。将由讨论基于联想的网络开始，并且将该模型扩展到包括概念相依理论、框架与脚本。而后深入审查一具体形式体系，即概念图，强调在表示知识中涉及的认识论问题，并且说明在现代表示语言中如何致力于这些问题。在第 14 章中将重述该形式体系，说明概念图如何用于实现自然语言数据库前端。第 7 章最后介绍表示的更现代途径，包括面向 agent 的体系。

第 8 章介绍基于规则的专家系统，以及基于事例和基于模型的推理，包括 NASA 空间计划的例子。问题求解的这些途径可描述为第 2、3、4、5、6 章内容的自然发展：使用对谓词演算表达式的产生式系统与图搜索适当结合。本章将由分析知识密集型问题求解的各途径的力量和弱点结束。

第 9 章介绍不确定与使用不可靠信息的推理模型。将介绍贝叶斯模型、信念网络、Dempster-Shafer 证据论、因果模型以及不确定情况中推理的肯定因数代数。本章还包含真值维护

算法，由最小模型的推理，基于逻辑的反绎，以及贝叶斯信念网络的团树算法。

第 10 章对基于符号的学习算法给出详细审查，这是引出若干不同问题和求解方法的富有成效的研究领域。这些学习算法按它们的目标、考虑的训练数据、它们的学习策略以及它们使用的知识表示变化。基于符号的学习包括归纳、概念学习、版本空间搜索和 ID3。研究了归纳偏向的作用，由数据模式的普遍化，以及在基于解释的学习中有效使用知识由一个实例学习。范畴学习或概念聚类是按无监督的学习介绍。最后介绍增强式学习，或者把来自环境的反馈集成到做新决策的策略中。

第 11 章介绍神经网络，它常称为学习的亚符号或连接主义模型。在神经网络中信息隐含在对一组连接的处理器组织和加权中，而学习涉及反复排列与修改全部加权的节点和系统的结构。将介绍若干连接主义体系，包括感知器学习、反向传播和逆传播；将说明 Kohonen、Grossberg 和 Hebb 模型；最后介绍结合学习与吸引区模型，包括 Hopfield 网络的例子。

遗传算法与学习的进化途径在第 12 章中介绍。按此观点可把学习视为浮现与适应过程。在根据遗传算法求解问题的几个例子之后，将介绍遗传技术在更一般问题求解器中的应用，包括分类器系统和遗传程序设计。而后将由人工生命研究的例子描述基于社会的学习。最后介绍浮现计算专题研究的例子。

定理证明常认为是自动推理，是 AI 研究的最老领域之一。第 13 章将讨论该领域中最早的程序，包括逻辑理论家和通用问题求解器。本章的主要焦点是二元归结证明程序，特别是归结反驳。最后，把 PROLOG 解释器描述为 Horn 短句与基于归结的推理系统，并且视 PROLOG 的计算为逻辑程序设计范型的一实例。

第 14 章介绍自然语言理解。语言理解的传统途径已由第 7 章介绍的许多语义结构的例子说明，本章将用随机途径补充。这些包括使用 Markov 模型、相互信息聚类 and 基于统计的语法分析。最后介绍应用自然语言技术产生数据库查询与原文摘要系统的例子。

第 15 章致力于智能系统科学的可能性问题，研究对 AI 的当代挑战，讨论 AI 的当前限制，以及激发未来的研究课题。

在附录 A 中介绍随机方法学，这是在不确定情况中推理的重要技术。在附录 B 中描述机器学习的随机途径。我们由定义隐式 Markov 模型(HMM)开始，而后介绍几种重要变形，包括自回归和层次 HMM。其次介绍动态贝叶斯网络，这是 HMM 的推广，能跨越时间周期追踪系统。这些技术对模拟复杂环境中的变化是有益的，对诊断和预测推理是需要的。最后在增强式学习中加入概率成分，其中包括描述 Markov 决策过程和部分可观察 Markov 决策过程。

请读者不要忽视两附录的内容，随机技术已日益增大了对 AI 的影响，特别是在诊断和预测推理、自然语言分析、机器人学和机器学习之类的领域中。为支持这些出现的技术，我们较详细地介绍了贝叶斯定理、Markov 模型、贝叶斯信念网络以及相关的图模型。我们还介绍了极多使用的概率有限状态机和隐式 Markov 模型。

我们假定读者们已学习过“离散数学”课程，其中包括谓词演算、集合论、计数和图论。还假定读者有“数据结构”基础，其中包括使用栈、队列和优先排队的树、图和基于递归的搜索。

编写本书的基本参考资料是 G. F. Luger 的著作，特别是反映人工智能近期进展的内容。本书中的参考文献并不完善，更详细的参考文献可查阅基本参考资料。由于本书追求的目标是尽量向读者展示人工智能的最新进展，对许多新内容笔者理解得不一定深刻，所以错误之处难免，恳请读者们批评指正。

使用本书

人工智能是一门大前沿科学领域，因此是一门大课程。本书可分为两部分：第一部分，第 1~9 章，可称为“人工智能导论”；第二部分，第 10~15 章，可称为“高级人工智能”。

本书应作为计算机专业本科生两学期的教材，共需 120 学时。学时安排如下。

第 1 章：10 学时。第 2 章：2 学时。第 3 章：4 学时。

第 4 章：6 学时。第 5 章：6 学时。第 6 章：4 学时。

第 7 章：8 学时。第 8 章：8 学时。第 9 章：8 学时。

第 10 章：10 学时。第 11 章：8 学时。第 12 章：8 学时。

第 13 章：8 学时。第 14 章：10 学时。第 15 章：10 学时。

此外，可根据学生的概率论基础情况，在第 9 章前适当插入附录 A，4 学时。在第 12 章之后可补充附录 B，6 学时，以介绍机器学习的概率途径。若学生缺乏谓词演算、图论和数据结构的基础知识，教师应对第 3 章、第 4 章和第 5 章花费更多时间。

刘凤岐

2011 年 4 月



目 录

出版者的话

序 言

前 言

第 1 章 导论 1

1.1 人工智能的定义与基础 1

1.1.1 何谓人工智能? 1

1.1.2 AI 基础简史 3

1.1.3 AI 与唯理主义和经验主义 传统 5

1.1.4 形式逻辑的发展 6

1.1.5 图灵测试 8

1.1.6 智能的生物与社会模型: agent 理论 10

1.2 AI 应用领域概述 12

1.2.1 博弈 12

1.2.2 自动推理与定理证明 13

1.2.3 专家系统 13

1.2.4 自然语言理解 15

1.2.5 模拟人的性能 15

1.2.6 规划与机器人 15

1.2.7 AI 语言和环境 16

1.2.8 机器学习 16

1.2.9 神经网络与遗传算法 17

1.2.10 AI 与哲学 18

1.3 人工智能概要 18

练习 19

第 2 章 表示与搜索导引 20

2.1 表示系统 21

2.2 搜索 22

第 3 章 谓词演算 26

3.1 命题演算 26

3.1.1 语法 26

3.1.2 语义 27

3.2 谓词演算 28

3.2.1 语法 28

3.2.2 语义 29

3.3 谓词演算的推断规则 32

3.3.1 推断规则 32

3.3.2 一致化 33

3.4 应用: 基于逻辑的家庭财务咨询 系统 36

练习 38

第 4 章 状态空间搜索 40

4.1 状态空间搜索的结构 41

4.1.1 图论 41

4.1.2 有限状态机 42

4.1.3 问题的状态空间表示 44

4.2 状态空间搜索策略 47

4.2.1 数据驱动和目标驱动的 搜索 47

4.2.2 广度优先和深度优先搜索 48

4.2.3 逐步加深的深度优先搜索 52

4.3 与/或图和逻辑推理 53

4.4 应用: 家庭财务咨询系统 54

练习 56

第 5 章 启发式搜索 57

5.1 爬山法和动态规划 59

5.1.1 爬山法 59

5.1.2 动态规划 60

5.2 最好优先搜索算法 62

5.2.1 实现最好优先搜索 62

5.2.2 实现启发估价函数 64

5.2.3 启发式搜索与专家系统 66

5.3 可采纳性、单调性及信息度 67

5.3.1 可采纳性 67

5.3.2 单调性	68	7.4 基于 agent 的分布式问题求解 ...	113
5.3.3 A* 算法的比较	69	7.4.1 面向 agent 的问题求解： 一种定义	113
5.4 搜索博弈图	69	7.4.2 agent 范型的示例及其存在 问题	115
5.4.1 极小极大程序	69	练习	116
5.4.2 固定深度的 minimax	71	第 8 章 知识系统	117
5.4.3 α - β 剪枝	74	8.1 专家系统技术概述	118
5.5 计算复杂度问题	75	8.1.1 设计基于规则的专家系统 ...	118
练习	77	8.1.2 问题领域的选择与知识 工程过程	119
第 6 章 状态空间搜索的控制算法	78	8.1.3 概念模型及其在知识获取 中的作用	121
6.1 基于递归的搜索	78	8.2 基于规则的专家系统	122
6.1.1 递归搜索	78	8.2.1 产生式系统与目标驱动的 问题求解	122
6.1.2 递归搜索示例：模式驱动的 推理	80	8.2.2 目标驱动的推理的解释与 透明性	124
6.2 产生式系统	82	8.2.3 产生式系统与数据驱动的 推理	125
6.2.1 定义及简史	82	8.2.4 专家系统的启发性与控制 ...	127
6.2.2 产生式系统示例	84	8.3 基于模型、基于事例和混合 系统	128
6.2.3 搜索的控制	86	8.3.1 基于模型的推理导引	128
6.2.4 产生式系统的优点	89	8.3.2 基于模型的推理：NASA 的 例子	130
6.3 问题求解的黑板体系	90	8.3.3 基于事例的推理导引	132
练习	91	8.3.4 混合设计	134
第 7 章 知识表示	93	8.4 规划	137
7.1 AI 表示研究简史	93	8.4.1 STRIPS	139
7.1.1 含义的联想主义理论	93	8.4.2 目的-反应式规划	142
7.1.2 语义网络的早期工作	96	8.4.3 规划：NASA 的例子	143
7.1.3 网络关系的标准化	97	练习	145
7.1.4 脚本	101	第 9 章 不确定推理	146
7.1.5 框架	103	9.1 基于逻辑的反绎推理	147
7.2 概念图	105	9.1.1 非单调推理逻辑	147
7.2.1 概念图导引	105	9.1.2 真值维护系统	149
7.2.2 类型、个体和名称	106	9.1.3 基于最小模型的逻辑	153
7.2.3 类型的层次结构	107		
7.2.4 概念图的操作	107		
7.2.5 命题节点	109		
7.2.6 概念图与逻辑	110		
7.3 替代显式表示	110		
7.3.1 Brooks 的包容体系	111		
7.3.2 多种表示、本体论与知识 服务	112		

9.1.4 集合覆盖与基于逻辑的 反绎	154	10.6.2 概念聚类	200
9.2 反绎：替代逻辑	156	10.6.3 COBWEB 与分类知识的 结构	201
9.2.1 肯定因数代数	156	10.7 增强式学习	205
9.2.2 模糊推理	158	10.7.1 增强式学习的成分	205
9.2.3 Dempster 和 Shafer 的 证据论	160	10.7.2 示例：井字棋博弈	206
9.3 不确定性的随机途径	164	10.7.3 增强式学习的推理算法	208
9.3.1 有向图模型：贝叶斯信念 网络	164	练习	209
9.3.2 有向图模型：d-分隔	165	第 11 章 神经网络	210
9.3.3 有向图模型：推理算法	166	11.1 神经网络基础	211
9.3.4 有向图模型：动态贝叶斯 网络	168	11.2 感知器学习	212
9.3.5 Markov 模型：离散 Markov 过程	169	11.2.1 感知器训练算法	212
9.3.6 Markov 模型：改型	171	11.2.2 感知器学习用于分类	213
练习	171	11.2.3 梯度下降法与 δ -规则	215
第 10 章 基于符号的机器学习	173	11.3 反向传播学习	217
10.1 基于符号的学习框架	174	11.3.1 反向传播算法	217
10.2 版本空间搜索	178	11.3.2 例 1：NETtalk	218
10.2.1 一般化操作与概念空间	178	11.3.3 例 2：异或函数	219
10.2.2 候选排除算法	179	11.4 竞争学习	220
10.2.3 LEX：归纳搜索启发式	182	11.4.1 分类的 WTA 学习算法	220
10.2.4 候选排除算法的评价	185	11.4.2 Kohonen 的学习原型 网络	221
10.3 ID3 决策树归纳算法	185	11.4.3 重复传播网络	222
10.3.1 自顶向下归纳决策树	187	11.5 Hebb 的叠合学习	224
10.3.2 信息论的测试选择	188	11.5.1 无监督 Hebb 学习示例	225
10.3.3 评价 ID3	190	11.5.2 有监督 Hebb 学习	226
10.3.4 决策树的数据问题	190	11.5.3 结合存储与线性结合器	227
10.4 归纳偏向与学习能力	191	11.6 吸引状态网络	229
10.4.1 归纳偏向	191	11.6.1 双向结合存储	230
10.4.2 学习能力理论	192	11.6.2 BAM 处理示例	231
10.5 知识与学习	193	11.6.3 自结合存储与 Hopfield 网络	233
10.5.1 基于解释的学习	194	练习	235
10.5.2 类比推理	197	第 12 章 学习的遗传与浮现模型	236
10.6 无监督的学习	199	12.1 遗传算法	237
10.6.1 发现与无监督的学习	199	12.1.1 遗传算法示例	238
		12.1.2 遗传算法的评价	240
		12.2 分类器系统与遗传程序设计	242

12.2.1 分类器系统	242	14.3.3 语义: ATN 分析器	285
12.2.2 遗传程序设计	245	14.3.4 用 ATN 组合语法和语义 知识	288
12.3 人工生命和基于社会的学习 ...	249	14.4 语言理解的随机工具	291
12.3.1 “生命博弈”	249	14.4.1 语法分析的概率途径	291
12.3.2 进化规划	251	14.4.2 概率上下文无关分析器 ...	292
12.3.3 浮现的专题研究	252	14.5 自然语言的应用	293
练习	255	14.5.1 故事理解与回答问题	293
第 13 章 自动推理	256	14.5.2 数据库前端	294
13.1 通用问题求解器	256	14.5.3 对 Web 的信息抽取和摘要 系统	296
13.2 归结定理证明器	260	练习	298
13.2.1 谓词演算表达式化为短句 集合	261	第 15 章 结束语: 评述与展望	300
13.2.2 归结证明程序	263	15.1 对 AI 几种途径的评述	301
13.2.3 归结策略和简化技术	265	15.1.1 智能与物理符号系统 假设	301
13.2.4 由归结反驳抽取答案	268	15.1.2 连接主义计算	303
13.3 PROLOG 与自动推理	269	15.1.3 agent、浮现与智能	305
练习	273	15.1.4 概率模型和随机技术	307
第 14 章 自然语言理解	275	15.2 现代认知科学	308
14.1 理解语言的符号途径	277	15.2.1 心理学的约束	308
14.2 语法	279	15.2.2 认识论问题	309
14.2.1 上下文无关文法的规范与 分析	279	15.3 AI: 当前的挑战与未来方向 ...	314
14.3 跃迁网分析器与语义	280	附录 A 随机方法导论	317
14.3.1 跃迁网分析器	280	附录 B 随机方法的应用	330
14.3.2 Chomsky 层次与上下文有 关文法	283	参考文献	345

导 论

1.1 人工智能的定义与基础

1.1.1 何谓人工智能?

1956 年夏季由美国数学家、计算机科学家和人工智能早期研究者 John McCarthy 等人倡导,在达德茅斯大学召开了首次人工智能研讨会,在此次会议上正式将人工智能作为一门独立学科。“人工智能”(Artificial Intelligence, AI)一词由 McCarthy 提出,但迄今尚难以给出确切定义。定义人工智能的主要困难是,目前我们还无法对“智能”给出确切定义,原因是通常认为智能是人类有别于其他动物的独有属性,但我们对智能的本质所知不多而难以理解。智能是一种本领还是像学习、推理、语言、感知、理解等看似不相关的各种能力的总称?在何程度上可以讲智能是先天的还是后天学到的?在学习时究竟发生了什么?什么是创造力?什么是直觉?智能可由可观察的行为来推断,还是需有具体内部机制的依据?在生物神经组织中如何表示知识,这对设计智能机器有何启迪?什么是“觉悟”,它在智能中起何作用?此外,按人类智能已知的什么模造智能计算机程序是否必要,或者对问题的严格“工程”途径是否充分?计算机是否可达到智能,或者智能体是否需要只在生物生存中才能找到的丰富感觉和体验?这些都是难以回答的问题,但它们有助于成型构成现代 AI 核心的问题和解决方法学。定义 AI 的另一问题是“A”。若世界上真的出现了克隆人,他们的智能也并非不是人造的。当然这不是 AI 研究的初衷。

尽管有这些困难,我们还是试图对 AI 给出一些定义,以便对该学科的性质、研究的问题、研究方法学、研究方向给出一些理解。Marvin Minsky 的定义是:

人工智能是研究使机器做人能做的需要智能的工作的一门学科。

该定义是面向机器的,并且“需要智能的工作”似乎容易理解,如下棋、写程序、解数学题、驾车等都是需要智能的工作,难以有争议。但该定义隐含着智能是人类独有的属性,本书不做该假设。另一进一步的定义由 G. F. Luger 给出:

人工智能是计算机科学中关于智力行为自动化的分支。

该定义确信 AI 是计算机科学的一部分,因此必须基于计算机科学的合理理论和应用原理。这些原理包括用于表示知识的数据结构、应用知识所需的算法,以及用于实现的语言和程序设计技术。值得注意的是,上述二定义都不能回避定义或理解“智能”的困难,因为目前我们不可能有规律地枚举出所有已见到或预见到的“需要智能的工作”和“智力行为”。虽然多数人在观察中都肯定知道智力行为或需要智能的工作,但按足够具体的方法帮助评价可能的智能计算机程序,又能捕捉人类智能的活性和复杂性来定义智能是可疑的。

实际上,人工智能的部分魅力是为精确探讨智能的属性提供独特而强有力的工具。AI 为智能理论的研究给出手段和试验平台:这些理论可按计算机程序语言叙述,并且通过在计算机上运行这些程序进行测试和验证。因此,AI 的上述二定义并未达到对领域的清晰定义。人

工智能仍是年青学科，它的结构、关注点和方法学相比物理学之类的成熟学科不太容易明确定义。如果物理学以百年为单位度量它的进展，人工智能只能以十年为单位度量自己的进展。人工智能是发展中的学科，何时能达到相对稳定，恐怕再过 50 年后才可能会有些眉目。但要强调的是，相比确定计算机科学的边界，人工智能总是更加关注扩展计算机科学的能力。使这类探讨有合理的理论原理是 AI 界面临的挑战之一。

关于人工智能学科的性质，A. Newell 和 H. A. Simon 在 ACM 图灵奖演说中指出^[16]：

计算机科学是一门经验学科。我们可以把它称为实验科学，但类似天文学、经济学和地质学，它的某些独特观察和实验形式并不适合狭隘的老一套实验方法。尽管如此，它还是实验的。构造的每台新机器都是一种实验。实际上，构造的机器提出自然界的问题；我们由观察机器的运行以及采用所有可能的解析和测量手段对它们进行分析，倾听答案。构造的每道新程序都是一种实验。它们提出自然界的问题，并且它们的行为对答案给出线索。机器和程序都不是黑箱子；已设计的硬件和软件都是人造物，我们可以把它们打开，观察它们的内部。我们可以将它们的结构和行为相联系，并且由单一实验提取许多经验和教训。

对许多人而言，AI 最惊奇的方面之一是与许多计算机科学相同，在范围上原来竟是经验学科。这种惊奇是由于大部分人开始时是按他们的数学或工程基础设想该领域。由数学观点或称为“干净”观点，有将证明和解析标准带到设计智能计算设备中的唯理主义愿望。由工程或“邋遢”观点，常将任务视为只是制作社会希望称为“智能”的成功人工产品。遗憾或幸运的是，取决于每个人的哲学，智能软件的复杂性及其在与人活动的世界的交互中固有的含糊性，阻碍了纯数学或纯工程观点的分析。

此外，若人工智能达到了某种科学水准并且成为“智能系统科学”的关键部分，解析和经验方法的混合必然包括在它的人工产品的设计、运行和分析中。按此观点，每道 AI 程序都可视为一种实验：它提出自然界的问题，而结果是大自然的响应。大自然对我们的设计和程序行为的响应可形成我们对形式主义和机械论的理解，最终形成对智能本质的理解^[16]。与许多对人的认知更为传统的研究不同，我们作为智能计算机产品的设计者，可以检查我们“主体”的内部工作，可以适当停止程序的运行，审查内部状态并且修改结构。另一方面，如文献 [16] 中指出的，计算机的结构及其程序预示了它们的可能行为：它们可以被审查，理解它们的知识表示和搜索算法。计算机作为理解智能的工具的力量是这种双向性的产物。适当程序控制的计算机有能力达到请求按心理学术语表征的语义水平和行为的复杂性，并且给出机会检查它们的内部状态，在很大程度上可拒绝科学家们对许多其他智能生命形式的研究。

关于 AI 学科的意义，J. Hopcroft 在 ACM 图灵奖演说中更明确地指出：

计算机科学的潜能若充分开发和探讨，将使我们走向关于世界知识的更高水准。计算机科学将帮助我们获得对智力过程的深入理解。这将增进我们关于学习、推理和推理过程的知识。计算机科学将为认知科学提供各种模型和概念工具。完全与物理科学在 20 世纪人们探讨物质的本质和宇宙起源期间统治了许多人智力上的努力相同，今天我们开始探讨思想、知识结构和语言的智力世界。我预见到不断做出的重要进展将使我们的生活有重大改变。……可预见到我们可以理解知识如何组织和运用……

由于 AI 界的不懈努力，迄今人工智能研究已取得了重要进展。由基于符号推理和对各种事物都给出名称的指称语义的物理符号系统途径诞生了专家系统和知识系统，由神经生物学研究诞生了人工神经网络或连接主义网络，由达尔文进化论研究诞生了遗传算法，由社会学研究诞生了基于 agent 的分布式问题求解。这些成就扩展了计算机科学的能力，使我们对智能

有了进一步理解。例如，今天我们知道智力功能不是一律的巨石一块，而是模块化和分布的；同样，人的学习也不是一律的同类能力，学习是对多种环境和不同社会结构的一种功能，但都适于达到具体目的。此外，这些成就也使人工智能研究由实验室进入了商业市场，影响着我们的社会。

按照 AI 研究已取得的各种成就的经验，我们扩展由 G. F. Luger 对 AI 给出的适于当前人工智能发展趋势的进一步定义为：

人工智能研究构成智力行为基础或呈现智力行为的机制，途径是构造与评价按制定的这些机制设计的人造物。

按该定义，人工智能不太是关于构成智能基础或呈现出智力行为的各种机制的理论，更多的是构造与测试支持这种理论的各种可能模型的实验方法学。它承诺以模型求精和进一步实验为目标，设计、运行和评价实验的科学方法。但更重要的是，该定义类似 AI 自身的领域，直接抨击了关于智能本质数世纪的哲学上的蒙昧（愚民）主义。它给人们想要理解的恐怕是我们定义为人的特征要替代宗教信仰、迷信行为、上帝的存在，或者在某种科学尚待揭示的怪癖中搜寻智能。如果支撑人工智能的科学对人类知识做出了任何贡献，可以确信智能不是神秘、不可触及的，而是可以理解并且适用于智能机器设计的一组原理和机制的效能。必须指出，AI 的该定义并不是定义智能，而是提出“人工”智能在探讨各种智力现象的表现和本质中的合理作用。

由于构造一般智能是极困难的任务，AI 研究人员常设想设计、制作具体智能人造物的作用。这些常按诊断、预测或形象化工具形式出现，使它们的用户能完成复杂任务。这些工具的例子包括语言理解的隐式 Markov 模型，证明数学中新定理的自动推理系统，跟踪穿过脑皮层网络信号的动态贝叶斯网络，以及基因表示数据模式的形象化。

1.1.2 AI 基础简史

现代人工智能研究的哲学基础已发展了数千年。但合理的历史起点应源于古希腊思想家和哲学家亚里士多德，他以精细分析和严谨思维将古希腊传统文化的各种见识、不可思议和恐惧的事物编织在一起，成为许多现代科学的标准。对亚里士多德来讲，自然界最有魅力的方面是变化。在他的《物理学》中，将“大自然的哲学”定义为“研究变化的事物”。他区分出事物的“物质”和“形式”：雕像由青铜“材料”和人“形”制成。青铜浇铸为新形式时出现了变化。物质和形式的区分为符号计算和数据抽象之类的现代概念提供了哲学基础。在计算中（即使使用数字）我们是操纵电磁材料形式的模式，这些材料形式的变化表示求解过程的各种样式。由它的表现手段抽象出形式，不只允许这些形式在计算上进行操作，也为现代计算机科学的核心——数据结构理论——提供了前提。这同样支持建立“人工”智能。

在亚里士多德的《形而上学》中，由句子“任何人生性都有求知欲望”开头，开发了永恒事物科学，包括他的宇宙论和神学。但对人工智能最中肯的是，在他的《逻辑学》中讨论的认识论或者分析人如何“认识”他们的世界。亚里士多德称逻辑为“工具”（工具论），因为他感到思维本身的研究是所有知识的基础。在他的《逻辑学》中分析了某些命题是否可称为“真”，取决于其他已知为“真”的事物。因此，若知道“任何人都是要死的”并且“孔夫子是人”，则可断定“孔夫子是要死的”。该论证是亚里士多德称为使用演绎形式的“假言推理”三段论的例子。虽然推理所需的形式公理化在 2000 多年后完全注入到 Gottlob Frege、Bertrand Russell、Kurt Gödel、Alan Turing、Alfred Tarski 等人的工作中，但其根源可追溯到亚里士多德。

建立在希腊传统上的文艺复兴思想，激发了思考人类及其与自然界关系的一种不同且强