

人工智能与 知识工程

陈世福 陈兆乾 等编著



南京大学出版社

内 容 简 介

本书系统地讲述了人工智能的基本概念、基本原理及其发展趋势,并列举了具有参考价值的人工智能程序和算法。本书总结了我們多年来的科研和教学成果及实践经验,内容涉及人工智能的大部分分支,取材新颖,内容丰富,由浅入深,层次分明。

全书内容包括人工智能的逻辑基础、知识表示技术、搜索策略、非精确性推理、人工智能语言、专家系统、专家系统开发工具、机器学习和人工神经网络等内容。书中各章均附有一定数量的习题,并列出了相关的参考文献,以便引导读者进行更深入的学习和研究。

本书用作高等院校有关专业的高年级学生或研究生教材,亦可供从事计算机科学、人工智能等有关方面工作的科技人员参考。

人工智能与知识工程

陈世福 陈兆乾 等编著

*

南京大学出版社出版

(南京大学校内,邮政编码:210093)

江苏省新华书店发行 常熟文化印刷厂印刷

*

开本 787×1092 1/16 印张 27.75 字数 684 千
1997 年 12 月第 1 版 1997 年 12 月第 1 次印刷

印数:1—4000

ISBN 7-305-03134-8/TP·174

定价:30.00 元

出版说明

人工智能是50年代兴起的一门新兴学科,40多年来,研究的进展和取得的成就都十分惊人。因此,它被誉为本世纪的重大科学技术成就之一。

人工智能是计算机科学的一个重要分支,它的研究领域十分广泛,涉及模式识别、物景分析、自然语言理解、数据库的智能检索、博弈、自动定理证明、自动程序设计、专家系统、机器翻译、机器学习、机器人等领域。随着人工智能的基本理论和技术的深入研究和广泛应用,它不仅正在深刻地影响着计算机科学的其它领域,而且对信息科学、控制科学、系统科学、数学、心理学、电子学、生物学、医学、语言学和哲学等学科产生愈来愈深刻的影响,并带来巨大的社会效益。

应广大科技人员、计算机工作者、研究生、大学生系统地学习人工智能的需要,我们陆续出版了下列一套具有特色的人工智能丛书:

- 1.《人工智能与知识工程》,1997年
- 2.《TURBO PROLOG 程序设计》,1989年
- 3.《TURBO PROLOG 工具库》,1988年
- 4.《知识工程语言与应用》,1989年

前 言

本书是针对高等学校“人工智能与知识工程”课程编写的一本通用教材。

40年来,人工智能又称机器智能取得了长足的进步与巨大的成就,一股研究人工智能和智能计算机的热潮正席卷全球,它已经成为举世瞩目的高新技术。为了配合我国人工智能的科研、教学和应用开发等工作的需要,我们根据多年来从事人工智能和知识工程研究和教学工作的一些经验,综合有关资料编写了这本书。

全书共分十二章,第一章绪论,第二章谓词演算与消解原理,第三章知识表示,第四章状态空间搜索,第五章启发式搜索,第六章 Prolog 语言,第七章 LISP 语言,第八章专家系统,第九章非精确性推理,第十章专家系统开发工具,第十一章机器学习,第十二章人工神经网络。

本书在编写过程中力求做到概念清楚,取材新颖,通俗易懂,选材考虑了广度、深度和先进性。为了帮助读者掌握人工智能原理和技术,书中列举了大量的例子和习题,其中有些例子已在 IBM PC 机上经过调试和运行,当然,它们的算法未必最佳,程序也未必十分完善,但对于读者深入理解书中的基本原理以及开发应用,也许会起到抛砖引玉的作用。

本书由陈世福(第一、八、九章)、陈兆乾(第二至五章、第七章、第十一章)、谢俊元(第六章)、潘金贵(第十章)、周志华(第十二章)等合作完成,并由陈世福和陈兆乾进行了修改和统编。

在本书的编写过程中,得到了谢琪、陆庆文、李红兵、尹胜等同志很多帮助。

由于编写时间仓促以及水平有限,书中错误和不妥之处在所难免,恳请读者批评指正。书中引用的文献可能会有遗漏,请有关专家谅解。

目 录

前言

第一章 绪论	(1)
1.1 人工智能	(1)
1.1.1 什么是人工智能	(1)
1.1.2 什么是智能	(2)
1.2 人工智能的发展史	(6)
1.2.1 第一阶段——孕育期	(6)
1.2.2 第二阶段——人工智能基础技术的研究与形成	(8)
1.2.3 第三阶段——发展和实用化阶段	(9)
1.2.4 第四阶段——知识工程与专家系统	(11)
1.3 人工智能的研究领域	(13)
1.3.1 专家系统	(13)
1.3.2 自然语言处理	(17)
1.3.3 机器学习	(20)
1.3.4 定理证明	(21)
1.3.5 分布式人工智能	(21)
1.3.6 机器人	(22)
1.3.7 模式识别	(23)
1.3.8 博弈	(24)
1.3.9 计算机视觉	(24)
1.3.10 人工神经网络	(25)
习题	(26)
第二章 谓词演算与消解原理	(27)
2.1 命题演算	(27)
2.1.1 符号和命题	(27)
2.1.2 命题演算的语义	(28)
2.2 谓词演算	(30)
2.2.1 谓词的语法和命题	(30)
2.2.2 谓词演算的语义	(34)
2.3 使用推理规则产生谓词演算表达式	(38)
2.3.1 推理规则	(38)
2.3.2 合一	(40)
2.3.3 合一的一个例子	(43)
2.4 应用：一个基于逻辑的金融投资辅助决策程序	(46)

2.5	消解原理	(48)
2.5.1	引言	(48)
2.5.2	为消解否定产生子句形式	(49)
2.5.3	消解证明过程	(52)
2.5.4	消解的策略和简化技术	(55)
2.5.5	从消解否定中提取解答	(60)
	习题	(62)
第三章	知识表示	(64)
3.1	引言	(64)
3.1.1	知识	(64)
3.1.2	使用知识	(65)
3.1.3	不确定性和语义原语	(66)
3.1.4	模块性和理解力	(67)
3.1.5	明确知识和灵活性	(68)
3.1.6	陈述性与过程性表示	(68)
3.2	知识表示技术综述	(68)
3.2.1	逻辑表示模式	(69)
3.2.2	语义网络	(69)
3.2.3	过程表示和产生式系统	(70)
3.2.4	模拟或直接表示法	(71)
3.2.5	特性表	(73)
3.2.6	框架和剧本	(73)
3.3	逻辑表示模式	(74)
3.3.1	命题演算	(75)
3.3.2	谓词演算	(76)
3.3.3	逻辑表示在人工智能中的应用	(77)
3.3.4	小结	(82)
3.4	基于规则的产生式系统	(82)
3.4.1	引言	(82)
3.4.2	问题求解的方法	(84)
3.4.3	问题求解中的搜索策略	(90)
3.5	语义网络	(93)
3.5.1	什么是语义网络	(93)
3.5.2	扩展的语义网络	(94)
3.5.3	部分语义网络	(97)
3.6	框架表示法	(98)
3.6.1	基本概念	(99)
3.6.2	框架的实例化以及用不同观点表示实体	(102)
3.6.3	类属特性,默认值和槽条件	(102)
3.6.4	框架结构	(104)
3.6.5	匹配和识别	(105)
3.6.6	推理	(107)

3.6.7	类比推理	(107)
3.6.8	错误和遗漏检查	(108)
3.6.9	过程附加	(108)
3.6.10	框架表示法的特点与不足	(109)
3.7	概念从属	(109)
3.7.1	引言	(109)
3.7.2	概念从属基本构成	(110)
3.7.3	概念从属示例	(114)
	习题	(115)
第四章	状态空间搜索	(117)
4.1	图论	(119)
4.1.1	状态空间搜索的结构	(119)
4.1.2	问题的状态空间表示法	(120)
4.2	状态空间搜索策略	(125)
4.2.1	数据驱动和目标驱动的搜索	(125)
4.2.2	图搜索的实现	(127)
4.2.3	深度和广度优先搜索	(129)
4.2.4	有界深度优先搜索	(135)
4.3	谓词演算推理的状态空间表示法	(135)
4.3.1	逻辑的状态空间描述	(135)
4.3.2	与/或图	(136)
4.3.3	讨论	(138)
4.4	基于递归的搜索	(145)
4.4.1	递归	(145)
4.4.2	递归搜索	(146)
4.5	模式驱动搜索	(147)
4.6	产生式系统	(153)
4.6.1	定义与历史	(153)
4.6.2	产生式系统示例	(154)
4.6.3	产生式系统搜索的控制	(159)
4.6.4	人工智能产生式系统的优点	(162)
	习题	(163)
第五章	启发式搜索	(165)
5.1	启发信息和估价函数	(167)
5.2	启发式搜索算法	(169)
5.2.1	局部择优搜索法(瞎子爬山法)	(169)
5.2.2	最好优先搜索法(有序搜索法)	(170)
5.2.3	启发估价函数的实现	(172)
5.2.4	启发式搜索与专家系统	(177)
5.3	启发式搜索过程的可采纳性、单调性、信息性	(178)
5.3.1	可采纳性	(178)

5.3.2	单调性	(179)
5.3.3	信息性	(180)
5.4	启发式搜索在游戏中的应用	(180)
5.4.1	强力搜索图上的极大极小化过程	(180)
5.4.2	固定层数的极大极小法	(183)
5.4.3	启发式搜索在一字棋游戏中的应用	(184)
5.4.4	α - β 剪枝技术	(187)
5.5	复杂性问题	(188)
	习题	(190)
第六章	Prolog 语言	(192)
6.1	概述	(192)
6.2	Prolog 语言基本概念和语法	(192)
6.2.1	数据结构	(192)
6.2.2	程序组成	(194)
6.2.3	Prolog 语法形式描述	(195)
6.3	Prolog 系统的搜索策略	(196)
6.3.1	例化, 匹配	(196)
6.3.2	回溯	(197)
6.3.3	搜索策略	(198)
6.4	Prolog 的内部谓词	(201)
6.4.1	比较	(201)
6.4.2	算术表达式求值	(201)
6.4.3	输入输出	(201)
6.4.4	文件操作	(202)
6.4.5	控制谓词	(202)
6.4.6	复杂目标的建立	(204)
6.4.7	项的分类	(204)
6.4.8	结构分量的建立与获得	(204)
6.4.9	子句的增、删、改	(205)
6.5	Prolog 程序设计	(206)
6.5.1	表处理和集合处理	(206)
6.5.2	数学函数	(209)
6.5.3	梵塔问题	(211)
6.5.4	八皇后问题	(212)
	习题	(214)
第七章	LISP 语言	(216)
7.1	引言	(216)
7.1.1	LISP 语言及其 PC SCHEME 版本	(216)
7.1.2	语言分类	(216)
7.1.3	LISP 语言的特性	(217)
7.2	LISP 语言简介	(217)

7.2.1	LISP 的数据结构——S-表达式	(217)
7.2.2	基本函数	(219)
7.2.3	常用的 LISP 系统函数	(222)
7.2.4	条件函数,定义函数	(226)
7.2.5	递归与迭代	(229)
7.3	LISP 程序设计及在人工智能中的应用实例	(233)
7.3.1	LISP 程序结构和用 LISP 语言解题	(234)
7.3.2	梵塔问题	(234)
7.3.3	水壶量水问题	(236)
7.3.4	皇后问题	(237)
7.3.5	向量程序设计	(239)
7.3.6	博弈	(242)
7.4	基于 LISP 的专家系统	(245)
7.4.1	用函数表示动物分类专家系统	(246)
7.4.2	用表表示动物分类专家系统	(248)
7.4.3	一个小型医疗诊断专家系统	(250)
	习题	(252)
第八章	专家系统	(254)
8.1	什么是专家系统	(254)
8.1.1	专家系统的历史	(254)
8.1.2	专家系统的特性	(256)
8.1.3	专家系统的构造者及其工具	(257)
8.1.4	专家系统的优越性	(258)
8.2	如何组织专家系统	(260)
8.2.1	知识的组织	(260)
8.2.2	知识表示	(262)
8.3	专家系统与传统程序的区别	(264)
8.3.1	专家系统的基本特点	(264)
8.3.2	专家系统存在的问题	(267)
8.4	专家系统的应用领域	(267)
8.4.1	专家系统的分类和基本活动	(267)
8.4.2	专家系统求解的类型	(272)
8.5	开发专家系统的需求分析	(280)
8.5.1	什么情况下开发专家系统是可能的	(280)
8.5.2	什么情况下开发专家系统是合理的	(281)
8.5.3	什么情况下开发专家系统是合适的	(282)
8.6	建造专家系统	(285)
8.6.1	建造专家系统的任务	(285)
8.6.2	专家系统开发的阶段	(287)
8.7	知识获取	(288)
8.7.1	知识获取过程	(288)
8.7.2	访问专家	(290)

8.8	专家系统设计示例	(293)
8.8.1	识别食物的简单产生式专家系统	(293)
8.8.2	动物分类专家系统	(294)
	习题	(300)
第九章	非精确性推理	(301)
9.1	确定性理论	(301)
9.1.1	医疗专家系统 MYCIN 概述	(301)
9.1.2	MYCIN 系统的推理策略	(302)
9.1.3	MYCIN 系统的知识表示	(306)
9.1.4	MYCIN 系统的非精确推理算法	(307)
9.2	主观概率论	(313)
9.2.1	PROSPECTOR 系统概述	(314)
9.2.2	推理网络	(315)
9.2.3	似然推理模型	(315)
	习题	(326)
第十章	专家系统开发工具	(327)
10.1	专家系统开发工具的特征	(327)
10.1.1	应用于开发专家系统的程序设计语言	(327)
10.1.2	知识工程语言	(328)
10.2	知识工程语言 M.1	(329)
10.2.1	M.1 概述	(329)
10.2.2	M.1 的体系结构	(330)
10.2.3	操作原理	(331)
10.2.4	M.1 的特征	(332)
10.2.5	M.1 应用领域	(333)
10.3	M.1 的知识表示	(334)
10.3.1	知识库的语法	(334)
10.3.2	项的语法	(336)
10.4	M.1 的推理机	(340)
10.4.1	概述	(340)
10.4.2	推理机工作过程	(341)
10.4.3	可信度计算	(344)
10.4.4	调节推理过程	(346)
10.5	M.1 的应用实例	(348)
10.5.1	微型照相顾问	(348)
10.5.2	交通工具顾问	(352)
	习题	(355)
第十一章	机器学习	(356)
11.1	引言	(356)
11.2	机器学习概述	(356)
11.2.1	什么是学习	(356)

11.2.2	机器学习的意义	(357)
11.2.3	机器学习研究的发展	(357)
11.2.4	机器学习的目标	(358)
11.3	机器学习的模型、策略和方法	(358)
11.3.1	简单的学习模型	(358)
11.3.2	主要的学习策略	(359)
11.3.3	基本的学习方法	(360)
11.4	机械学习、示教学习和类比学习	(361)
11.4.1	机械学习	(361)
11.4.2	示教学习	(362)
11.4.3	类比学习	(364)
11.5	基于示例的归纳学习	(365)
11.5.1	归纳学习分类	(365)
11.5.2	示例学习基本概念	(366)
11.5.3	示例学习的完备性条件与一致性条件	(369)
11.5.4	示例学习中的噪音问题	(370)
11.5.5	建立示例学习系统要考虑的几个问题	(370)
11.6	ID 系列算法	(371)
11.6.1	判定树	(371)
11.6.2	ID3 算法	(371)
11.6.3	ID4 算法	(374)
11.6.4	ID5R 算法	(375)
11.6.5	ID 系列算法的比较分析	(379)
11.7	一个基于示例的归纳学习系统	(380)
11.7.1	系统概述	(380)
11.7.2	系统的总体结构	(382)
11.7.3	系统的核心算法	(383)
11.7.4	系统运行实例	(387)
习题	(390)
第十二章	人工神经网络	(391)
12.1	引言	(391)
12.2	人工神经网络概述	(391)
12.2.1	什么是神经网络	(391)
12.2.2	脑神经信息活动的特征	(392)
12.2.3	神经网络研究的目的和意义	(392)
12.2.4	神经网络研究的发展	(392)
12.2.5	人工神经网络研究的局限性	(394)
12.3	神经网络基本模型	(395)
12.3.1	生物神经元的结构	(395)
12.3.2	M-P 模型	(396)
12.3.3	感知机模型	(397)
12.3.4	Hopfield 网络模型	(399)

12.4	神经网络的学习	(400)
12.4.1	“突触修正”假说	(400)
12.4.2	学习算法的分类	(400)
12.4.3	误差修正型学习	(401)
12.4.4	竞争型学习	(402)
12.4.5	赫布型学习	(403)
12.4.6	随机型学习	(403)
12.5	误差逆传播网络	(404)
12.5.1	基本理论	(404)
12.5.2	学习算法	(406)
12.5.3	标准误差逆传播和累积误差逆传播	(407)
12.5.4	全局最小和局部极小	(408)
12.6	自适应谐振理论	(408)
12.6.1	基本理论	(408)
12.6.2	学习算法	(410)
12.6.3	与 ART 有关的定理	(411)
12.6.4	快速学习与慢速学习	(411)
12.6.5	基于 ART 的其他网络模型	(411)
12.7	基于域理论的自适应谐振网络	(412)
12.7.1	概述	(412)
12.7.2	域理论	(412)
12.7.3	FTART 网络基本理论	(413)
12.7.4	FTART 学习算法	(414)
12.7.5	FTART 网络的判别方法	(415)
12.7.6	FTART 算法测试	(416)
12.7.7	运行实例	(420)
	习题	(421)
	参考文献	(422)

第一章 绪 论

1.1 人工智能

1.1.1 什么是人工智能

人工智能(Artificial Intelligence)又称机器智能(Machine Intelligence)是计算机科学中一门新兴的边缘学科。和计算机科学的许多分支一样,一个普通的读者如果不了解什么是人工智能及人工智能有什么用途,他就不大可能很好地理解该书的内容,就像一个人如果没有某一领域的知识,他就不能读懂某些杂志和报纸一样。所以读者了解一些人工智能的概貌对理解和阅读全书是很有用的。

1956年夏,麦卡锡(J. McCarthy)、明斯基(M. L. Minsky)、洛切斯特(N. Lochester)和香农(C. E. Shannon)等邀请了西蒙(H. A. Simon)、纽厄尔(A. Newell)、塞缪尔(A. L. Samuel)、莫尔(T. More)、塞尔夫利奇(O. Selfridge)、卡纳奇(Carnagic)和索罗孟夫(R. Solomonff)等10多位数学、心理学和信息论等方面的学者在美国达特茅斯(Dartmouth)大学举行了历时两月之久的有关如何利用计算机在数学、物理学、心理学和电子工程学等方面模拟人类智能行为的学术研讨会。会上达特茅斯大学的年轻助理教授麦卡锡提出,将所讨论的领域作为计算机科学中的一个新的研究分支,并称之为人工智能,即在会上正式决定使用人工智能这一术语。这是一次具有重大历史意义的会议,它标志着人工智能这门新兴的边缘学科的正式诞生。

自那时起,世界许多国家的科学工作者都先后开展了程度不等的人工智能研究,各方面的研究成果层出不穷。例如,1965年美国斯坦福大学费根鲍姆(E. A. Feigenbaum)等人研制出第一个成功的专家系统 DENDRAL,它能根据质谱仪的数据推知物质的分子结构。1976年,E. Shortliffe 等人又研制成功用于细菌感染疾病诊治的系统 MYCIN。1977年,费根鲍姆教授在第五届国际人工智能会议上(IJCAI-77)提出了知识工程(Knowledge Engineering),它为人工智能的应用奠定了基础,大大促进了专家系统的发展。

至今,人工智能的研究已有40多年的历史,虽然AI科学家从不同角度对人工智能进行了各种描述,但至今尚无严格的定义。1981年,费根鲍姆在“人工智能手册”中指出:“人工智能是计算机科学中的一个分支,涉及到智能计算机系统的设计,该系统显示人类行为中与智能有关的某些特征”。绍特里夫(Shortliffe)于1984年在“规则基专家系统”一文中指出:“人工智能是计算机科学中的一个分支,它研究问题求解的符号方法和非算法方法”。M. Boden(1977)认为人工智能是“利用计算机程序和程序设计技术来认识普通的智能原理和具体的人类思想”。他这个定义强调了计算机软件,在他看来只要写出一段合适的指令序列代码就建立了某种类型的人工智能机器。但并非所有的人工智能研究者都认为计算这么重要。例如,一个程序能够模仿一个机器人的行为,但如果没有合适的运行它的物理系统,这个程序就没有实际价值。M. Minsky 的一个经常被引用的定义指出了人工智能的实际特征:

人工智能就是让机器完成那些如果由人来做则需要智能的事情的科学。

但这个定义没有说出智能的本质,我们知道,对人来说做那些复杂的计算通常认为需要智能,但一般的人工智能研究者很少把这些计算列入人工智能的领域。常常以问题求解、推理和博弈等活动来说明人工智能的一些特征。

另一种方法就是用更通用的术语列出人工智能的目标,这样就可给出人工智能的一个实用的定义。例如,A. Sloman(1978)认为人工智能有三个主要目标:

- 对智能行为有效解释的理论分析
- 解释人类智能
- 构造智能的人工制品

他还给出了下列各种考虑,智能行为与“构造、描述、解释、比较、修改和使用复杂的结构,包括像句子、图画、地图和行动计划等符号结构”的能力有关,更进一步地说:即人工智能的研究与其他科目的研究相重叠,这些科目都要用到推理,如心理学、教育学、人类学和生理学等等。很显然,如果不研究词法和语义学就无法让计算机理解自然语言。

当然,还有其他许多种定义,这里不再赘述。众所周知,传统程序就是告诉机器干些什么,而如何干则是靠人设计出方法给出算法并写出程序而实现的。AI 程序要求机器知道怎么干,只要把问题描述清楚,机器就能自动实现求解。包括由“数值计算”过渡到“符号处理”。表 1.1 列出了人工智能程序和通常计算机程序的不同。

表 1.1 人工智能和通常程序的比较

人工智能	通常计算机程序
主要是符号处理	主要是数字处理
启发式搜索	依靠算法
控制结构和知识域相分离	信息和控制联结在一起
易于修改、更新和改变	难以修改
允许不正确的答案	要求正确的回答

总之,人工智能是一个非常广义的词语,基于人工智能技术编写的程序,能使计算机具有人类在处理问题时需要的智能。当机器像一个知识渊博的专家那样工作时,它能工作得很出色甚至有时会超过它们的导师。例如,人工智能研究者已创造出各种实用专家系统,它们具有高级工程师的专门知识和权威,可在某种程度上模仿人的思维和记忆,因此能够诊断疾病,勘探地下矿产和地下水,能够理解和翻译书面材料等。普遍的观点认为

人工智能的研究是困难的,这种观点在一定程度上是因为对智能这个概念产生的畏惧,所以讨论一下智能是十分必要的。

1.1.2 什么是智能

定义人工智能的主要困难就是“智能”本身不易被人们理解,至今对智能是什么还没有统一的看法,“智能”一词源于拉丁文的“Legere”,意思是收集、汇集,智能通常用来表示从中进行选择、理解和感觉。Feigenbaum 和 McCorduck(1983)曾解释:如果我们能够研制一个人工制品,它具有收集、汇集、选择、理解和感觉的功能,那么我们就认为该人工制品具有“智能”。

本书主要是讨论建立机器智能,但我们对自然智能也应当有所了解。人工智能工作者不可避免地要仔细地研究自然智能,以识别自然智能的关键特征和决定性的属性,对于生物系统中的智能行为有了深入的理解和清晰的概念,就可以在人工制品中研究如何加入智能因素了。

1. 自然智能

所谓自然智能就是人类和一些动物所具有的智力和行为能力。事实上,人们关于自然智能的观点是随着生物学的发展和人工制品的新性能的出现而不断地变化,即智能有多方面的含义,而人工智能的发展将有助于阐明一个聪明的生物种类的内部秘密。

在日常应用中,“智能”和它的衍生词“智力”是同义词,说这个人或那个人聪明一般是不会被弄错含义的,也就是说“智力”是人类的一个特征。但是当我们试图定义“智能”时,我们就会发现这个词变得难以捉摸了。它是用来处理数字、单词或其他符号的吗?它是否与世界上的实际动作有关?它对创造和发明是否有影响?是在精神活动中还是在实际行动中才能发现智力?智力是如何与意志、学习、记忆和情绪等现象相关联的?这一系列问题出现在我们面前。后来人们发现可以通过人工制品来建立智力的度量而无须完全了解自然智能的实质。但了解人类智能,可以知道对机器来说哪些是可能实现的,哪些是现在或将来不可能实现的。

关于自然智能的类型是多方面的,我们通常认为智力包括解决问题、进行多种运算、学习和处理各种新情况的能力。智力是针对具体情况的,根据不同的情况有不同的含义。例如我们说教授、经理、医生等处于责任重大岗位的人是聪明的,但有时我们又会陷入以下悖论:“某某人既然是那么聪明,他的行为或他说的话为何有时却非常的愚蠢?”这就告诉我们要识别不同类型的智能,例如卓越的数学家也许没有政治远见,一个著名的医生可能口才很差等等现象。智能可以是“行为上的”、“口头上的”和“哲学上的”等等,这些区别对于理解人工智能是很重要的。现在的计算机看来像是个很熟练的数学家,但对于需要判断和处理常识的任务则不太胜任。正如,生物智能已经演化了好几百万年,在从幼稚到成熟的过程中扩展了智能的种类,类似地,计算机智能也有一个演化的过程。

有时需要区别智力和具体的能力,“智力”是指学会某种技能的能力,而不是指技能本身,因此一个要用一个月时间来掌握一个简单的微分方程的人就不如一个只用几天就学会的人聪明;一种技能可以在一段漫长的时间内通过艰苦的努力获得,而该种技能的存在也对应着某种智力,当然有些技能是无法传授给某些人的。

获取技能的能力是一种综合的能力,而技能则又有多种,以往对人工智能的一种批评就是不同的人工智能程序只是针对不同的孤立的任务,如游戏、定理证明、故事描述等等,而缺乏真实的智能的综合性特点,很显然,这是一个很能说明问题的批评,人工智能工作者正在积极地攻克该问题。实际上计算机可以很容易地扩展人工智能程序的能力范围,例如纽厄尔(Newell)等研制的通用问题求解器 GPS(General Problem Solving)已经具有所需的综合能力。

有些学者试图把智能这一概念具体化,错误地把一些智力特征说成是灵魂或精神,事实上,智力应认为是对一定行为的抽象,只有通过行为我们才能了解智力,这个方法对评估计算机系统的智能性很重要,可以根据计算机系统能干些什么来评价它是否聪明,即看它能否解答问题,作出决定和判断。

对智能的各种定义和描述说明了研究者所强调的重点不同。下面是一些著名的研究者对智能给出的定义,并且可以看到大部分内容都和推理能力有关。

- Binet: 善于判断、理解和推理
- Spearman: 综合智力主要是指“相互关系的训练”
- Terman: 形成概念和掌握含义的能力

- Vernon: 全面考虑的能力或思维的效率
- Burt: 先天的、综合的和认知的能力
- Heim: 智力活动包括在某一情况下抓住本质并对它们作出适当的反应
- Wechsler: 适当地行动、理智地思考、有效地适应环境的总体能力
- Diaget: 身体和社会环境的适应性

以上定义中的各种能力:判断、理解、推理、形成概念、适当的反应和适应性等说明了智力具有多面性的本质。例如人工智能研究者 D. Hofstadter 认为“智能中最重要的能力包括适应环境、适应偶然性事件、能分辨模糊的或矛盾的信息、在孤立的情况中找出相似性、产生新概念和思想”。智力的多重特征再次明显地显露出来。

Hofstadter 定义中最重要的能力是指诸如判断、洞察力和创造力等与计算机没有关系的元素。事实上,人类的常识对于计算机来说是无法想象的。因为计算机只会盲从于规则,就像铁轨上的列车失去了独立创造性一样,但人工智能工作的乐趣在于试图把不同的规则加入到严格的形式体系中以使僵死的计算机变得更灵活些。

智力的多重性暗示着在人工制品中构造智能的某些功能比另一些功能更容易。例如,确定和测量计算能力比确定和测量判断、创造力等功能容易得多,如果不能从某种意义上对某个功能进行确认,就很难了解在一个计算机程序中是如何模拟它的。计算机科学的发展开辟了一条人类心理和智力的认识途径,这条心理学途径依次对人类的各种精神活动过程进行计算机系统的模型化。

人类大脑的认识模型是基于信息处理的思想,假设每个人都装备如图 1.1 所示的基本信息处理系统,那么所有的精神活动原则上都可根据信息处理系统的操作进行解释。

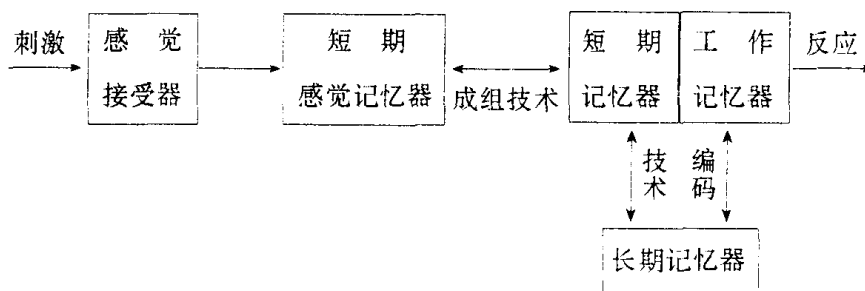


图1.1 人类大脑的信息处理模型

感觉接收器接收外部世界的信息,依赖于自然现象如气压、温度和电磁辐射等变化。这些信息先存放在一个短期感觉记忆器中,又称感觉缓冲器。这些记忆器有很大的容量而能装入由感觉接收器所提供的所有信息,但只能存储很短的一段时间,例如,视觉信息只能存 0.5 秒。根据不同需要,可以把注意力投向或不投向短期感觉记忆器中保存的信息,一般情况下,大量信息由于没有受到注意而逐渐消失,当注意力转向短期感觉记忆器时,有些信息就可能被传送到短期记忆器中,在这个记忆器中,可将一些信息保存 15 秒之久,直到另外一些新来的信息将它们“推出”。也可以使意识里的信息保持活跃状态,这个工作叫作“复述”。短期记忆器的容量比短期感觉记忆器小得多,但“成组技术”可用来聚合信息,这样就能增加当前感觉的范围。

工作记忆器有时被认为是短期记忆器的附属物,它容量有限,具有存储中间项的能力。工作记忆器作为一个高速暂存器促使有意识的智力活动能够完成。和短期记忆器一样,如果不采

取复述策略,工作记忆器中的信息很快就会丢失。

短期记忆器中的信息可以转到长期记忆器中,长期记忆器可保存大量的信息且不会随着时间的推移而丢失,由于种种原因,有时检索通路被堵塞,在堵塞畅通之前信息项将会丢失,在没有阻塞时,可使用查找策略来进行信息项的检索。

上述的存储和操作因素表示了人类的所有特征,人与人之间仍有很大的差别,某个人可能比别的人记得多,并且能更好地迅速地搜索长期记忆。另外,成组技术和信息编码(将短期记忆变为长期记忆)等操作对每个人来说效率也是不同的。这些差异说明了为什么每个人的智力水平是不同的。

在信息处理模型的框架中可以对不同类型的智能进行研究。例如,阅读需要口头上的智能,包括在记忆中查找每个词素,看它是否已出现过,确定其含义。由于一段文章中大量词素,因此对不同的人来说在搜索时的微小差别也会变得十分显著。同样地,一个人短期记忆器的容量也部分地决定了他能否很容易地阅读一篇文章,一个能一次读许多词装入短期记忆器的人容易进行阅读。

上述的认知方法对人工智能的研究是很重要的,因为它给出了一个能表示已存在的计算机系统的特征模型,这并非巧合,计算机科学的进展已经成为推动研究人员开发关于心理现象的信息处理模型的关键因素,因此,弄清楚人类的心理现象将能极大地扩展人工智能的研究领域。例如,如果我们能够给出情感的信息处理解释,就有可能将情感加入人工制品。人们已经建造了能够学习、记忆和推理的系统,在不远的将来构造具有想象、直觉和情感的人工制品已不再是一个幻想。

2. 智能的测量

智能这个词往往被人误解,例如当谈及某人做什么事时,习惯用该词表示不寻常的创造能力和聪明。结果,人们对人工智能期望过高,似乎是能产生一个非凡的聪明与洞察力的技术,事实上,人工智能最感兴趣的问题是试图复制“普通人的智能”,如视觉和自然语言能力等等。目前所研制的很多人工智能系统已初步达到该目标。但怎样判定一个系统是否具有智能,或者说机器是否具有智能呢?这是一个十分复杂而有用的问题,智能模拟的先驱者图灵(A. M. Turing)从理论上解决了这个问题。

1950年英国数学家图灵(Turing)在思想(mind)杂志上发表了“计算机与智力”一文,提出了一种人工智能的实用测试方法,后来以他的名字命名为Turing测试。Turing的目的是为计算机是否具有智能功能建立一种测试,他将其测试建立在当时的客厅游戏的基础之上,其客厅游戏玩法如下:有两个客人,即男女客人各一个,藏在两个分离的房间中,第三个人,作为裁判,在第三个房间中,看不见并不熟悉这两个客人,裁判的任务是决定哪个客人是女的,完全依靠裁判与两个客人之间传递的信息,这两个人都是尽量使裁判相信她是个女的,问题是决定谁讲的是真话。

Turing提出了该游戏的一个变种,即一个是人,一个是机器人,它们都通过电传打字机与“裁判”相联,如图1.2所示。其中测试者A可以是任何一个人,他可以通过终端设备与受试者B和C交谈。A从未见过B和C,只知道B和C中一个是人,一个是机器,测试时A可以通过终端设备向B、C提出各种智能方面的问题,但不能问及它们的物理特征;而人和机器都将尽量使“测试者”相信它们自己是“人”,而另一个是“机器”。在该前提下要求A区分B、C中哪个是机器,哪个是人,如果不管如何更换测试者和受试者中的人,测试者成功的概率都小于