

新世纪高等学校教材

人工智能基础

蔡自兴 主编 蒙祖强 副主编



高等教育出版社

<http://www.hep.com.cn>
<http://www.hep.edu.cn>

新世纪高等学校教材

人工智能基础

蔡自兴 主编

蒙祖强 副主编

高等教育出版社

内容简介

本书是国家教育部立项建设的优秀网络课程“人工智能网络课程”的配套教材。全书共 10 章,主要内容:绪论、知识表示、搜索原理、推理技术、机器学习、规划系统、专家系统、自然语言理解、智能控制、人工智能程序设计。附录中给出了人工智能网络课程使用指南。

本书可作为本科学校和高职高专学校计算机相关专业的“人工智能”课程教材或教学参考书,还可供从事人工智能研究、开发与应用的科技工作者学习参考。建议读者尽可能结合“人工智能网络课程”进行学习和训练,充分利用该网络课程提供的丰富教学资源。

图书在版编目(CIP)数据

人工智能基础/蔡自兴主编. —北京:高等教育出版社,
2005.5

ISBN 7-04-016486-8

I. 人... II. 蔡... III. 人工智能-高等学校-教材 IV. TP18

中国版本图书馆 CIP 数据核字(2005)第 025195 号

策划编辑 董建波 责任编辑 萧 潇 封面设计 王凌波 责任绘图 朱 静
版式设计 张 岚 责任校对 俞声佳 责任印制 杨 明

出版发行 高等教育出版社
社 址 北京市西城区德外大街 4 号
邮政编码 100011
总 机 010-58581000
经 销 北京蓝色畅想图书发行有限公司
印 刷 北京市联华印刷厂

购书热线 010-58581118
免费咨询 800-810-0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>
网上订购 <http://www.landaco.com>
<http://www.landaco.com.cn>

开 本 787×1092 1/16
印 张 20
字 数 420 000
插 页 2

版 次 2005 年 5 月第 1 版
印 次 2005 年 5 月第 1 次印刷
定 价 25.60 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 16486-00

作者简介



联合国专家、纽约科学院院士、国际导航与运动控制科学院院士、全国政协委员、国家级教学名师奖和宝钢优秀教师特等奖获得者。从事计算机智能系统、人工智能、智能控制、智能机器人研究。1962年毕业于西安交通大学。1983—1985年留学美国普渡大学和内华达大学。1992—1993年留学美国伦塞勒工学院。2004年留学俄罗斯科学院圣彼得堡信息学与自动化研究所。

出版专著及教材 24 部，发表学术论文 560 多篇，主持并完成研究课题 30 多项。获国际奖励 1 项，国家级奖励 2 项，省部级奖励 12 项。

现任中南大学信息科学与工程学院博士生导师、教授，兼任中国人工智能学会副理事长、中国自动化学会理事、中国计算机学会模式识别与人工智能专业委员会委员及《控制理论与应用》、《机器人》、《控制与决策》等 5 种杂志的编委，并兼任国防科技大学、北京航空航天大学、北京邮电大学等 10 所高校兼职教授，为 IEEE 高级会员、国际科技发展联合会会员、国际小型和微型计算机学会会员等。曾任湖南省政协副主席、国际智能自动化联合会理事、国家自然科学基金委学科评议组成员、北京大学信息科学中心客座研究员、中国科学院自动化研究所客座研究员、美国伦塞勒工学院客座教授等职。

前 言

在过去十多年中，随着社会进步和科技发展，人工智能作为一门前沿和交叉学科，在自身取得很大进展的同时，也促进了其他学科的发展。进入 21 世纪以来，人类社会对人工智能表现出更好的认同，寄予更大的希望。人工智能正处在一个最好的时期，它的发展及其成果已经而且必将在更大的广度和深度上造福于人类。

为了落实教育部《面向 21 世纪教育振兴行动计划》，加快现代远程教育工程资源建设步伐，提高我国高等学校的教学现代化水平，教育部于 2000 年 5 月启动了国家级“新世纪网络课程建设工程”，以便采用更先进的教学手段，为提高教学质量和培养高素质人才服务。“人工智能网络课程”就是该工程首批支持的项目之一。在教育部及其现代远程教育资源建设委员会和中南大学的大力支持下，经过我们的深入研究和认真开发，本网络课程开发取得成功，于 2004 年先后通过教育部组织的专家验收和质量认证，并被评为了优秀网络课程。本书就是“人工智能网络课程”的配套教材。

本书共 10 章。第 1 章叙述人工智能的概况，列举出人工智能的研究与应用领域。第 2 章～第 4 章研究传统人工智能的知识表示方法、搜索技术和知识推理，并介绍了两种计算智能技术，即遗传算法和模拟退火算法。第 5 章～第 9 章比较详细地讨论了人工智能的主要应用，包括机器学习、自动规划、专家系统、自然语言理解和智能控制等。第 10 章简要介绍人工智能的程序设计工具，涉及 LISP 和 PROLOG 语言、专用开发工具和人工智能机等。书后附有人工智能网络课程使用指南，简要介绍该网络课程的使用方法，包括客户端的安装说明、网络课程使用说明、教师注册及使用说明等。

本书是在蔡自兴、徐光祐主编的《人工智能及其应用》（第二版）（清华大学出版社 1996 年出版）的基础上编写的。该书承蒙广大读者厚爱，已发行 12 多万册，被数百所院校用作教材或教学参考书。1992 年 3 月，台湾儒林图书出版公司出版了本书的繁体字版，向海外发行。该书曾获 1999 年度国家教育部科技进步一等奖和 2002 年国际优秀作品奖，并与其他著作和成果一起获得 2000 年度中国高校自然科学奖二等奖和省级优秀教学成果奖一等奖。蔡自兴主讲的“人工智能”课程于 2004 年 2 月被教育部评为国家精品课程。在此，特别感谢教育部和现代远程教育资源建设委员会和中南大学对本课程的支持，衷心感谢验收和质量认证专家组的指导与帮助。中南大学和高等教育出版社的有关领导和编辑对本书的出版给予高度重视，付出了辛勤劳动，如果没有他们的大力支持与合作，本书就不可能迅速与读者见面，在此，谨向他们致以诚挚的敬意。

在本书的编写过程中还引用了国内外人工智能领域的研究成果，使本书能够较全面地反

映人工智能学科的最新进展。在此，特对这些专家学者深表谢意，他们是：高济、何华灿、何新贵、何志均、陆汝钤、施鹏飞、史忠植、宋健、涂序彦、吴文俊、徐光祐、张钹、钟义信以及 J Durkin、E A Feigenbaum、C S G Lee、Z. Michalewics、N. J. Nilsson、P. Norvig、E. Rich、S. J. Russell、G. N. Saridis、A. B. Тимофеев 和 P H Winston 等。

本书由蔡自兴任主编，蒙祖强任副主编，陈白帆为主编助理。蔡竞峰、龚涛、李枚毅、孙国荣、魏世勇、肖晓明、文敦伟、邹小兵等参加了本书编写工作和“人工智能网络课程”的开发。此外，曹峰、曹科、陈海燕、杜文峰、凌群、刘娟、庞慧、秦鹏、石跃祥、田淑杭、王雷、王腾、文敦伟、许永华、杨志国、赵慧、朱光辉等也参加了网络课程的开发工作。他们中的许多人已经学成毕业，走向新的工作岗位。在本书出版之际，谨向他们表示由衷感谢，望他们再接再厉，与时俱进；祝他们事业成功，诸事如意。

由于编著者学识有限，编写时间又比较仓促，加之人工智能发展较快，书中难免存在不妥和错误之处，欢迎各位专家和本书读者不吝指教和帮助。

蔡自兴

2005年5月

目 录

第 1 章 绪论	1	1.4.17 人工生命	15
1.1 人工智能的定义与发展	1	1.4.18 系统与语言工具	15
1.1.1 人工智能的定义	1	1.5 人工智能对人类的影响	16
1.1.2 人工智能的起源与发展	2	1.5.1 人工智能对经济的影响	16
1.2 人类智能与人工智能	3	1.5.2 人工智能对社会的影响	16
1.2.1 研究认知过程的任务	3	1.5.3 人工智能对文化的影响	18
1.2.2 智能信息处理系统的假设	3	1.6 对人工智能的展望	19
1.2.3 人类智能的计算机模拟	5	1.6.1 更新的理论框架	19
1.3 人工智能的学派及其争论	6	1.6.2 更好的技术集成	19
1.3.1 人工智能的主要学派	6	1.6.3 更成熟的应用方法	20
1.3.2 对人工智能基本理论的争论	7	第 2 章 知识表示	21
1.3.3 对人工智能技术路线的争论	7	2.1 概述	21
1.4 人工智能的研究与应用领域	8	2.2 状态空间法	22
1.4.1 问题求解	8	2.2.1 问题状态描述	23
1.4.2 逻辑推理与定理证明	9	2.2.2 状态图示法	24
1.4.3 自然语言理解	9	2.2.3 状态空间表示举例	25
1.4.4 自动程序设计	10	2.3 问题归约法	28
1.4.5 专家系统	10	2.3.1 问题归约描述	28
1.4.6 机器学习	10	2.3.2 与或图表示	30
1.4.7 人工神经网络	11	2.3.3 问题归约机理	32
1.4.8 机器人学	11	2.4 谓词逻辑法	35
1.4.9 模式识别	11	2.4.1 谓词演算	35
1.4.10 机器视觉	12	2.4.2 谓词公式	37
1.4.11 智能控制	12	2.4.3 置换与合一	39
1.4.12 智能检索	13	2.5 语义网络法	40
1.4.13 智能调度与指挥	13	2.5.1 二元语义网络的表示	41
1.4.14 分布式人工智能与 Agent	13	2.5.2 多元语义网络的表示	42
1.4.15 计算智能与进化计算	14	2.5.3 连词和量化的表示	43
1.4.16 数据挖掘与知识发现	14	2.5.4 语义网络的推理过程	47

2.6 框架表示	51	3.5.1 模拟退火算法的模型	92
2.6.1 框架的构成	52	3.5.2 模拟退火算法的简单应用	93
2.6.2 框架的推理	55	3.5.3 模拟退火算法的参数控制问题	95
2.7 面向对象表示	57	第4章 推理技术	96
2.7.1 面向对象基础	57	4.1 消解原理	96
2.7.2 类与类继承	58	4.1.1 化为子句集	96
2.7.3 面向对象表示的实例	58	4.1.2 消解推理规则	99
2.8 剧本表示	62	4.1.3 含有变量的消解式	100
2.8.1 剧本的构成	63	4.1.4 消解反演求解过程	101
2.8.2 剧本的推理	64	4.2 规则演绎系统	105
2.9 过程式表示	65	4.2.1 规则正向演绎系统	105
第3章 搜索原理	68	4.2.2 规则逆向演绎系统	111
3.1 盲目搜索	68	4.2.3 规则双向演绎系统	113
3.1.1 图搜索策略	68	4.3 产生式系统	115
3.1.2 宽度优先搜索	70	4.3.1 产生式系统的组成及表示	116
3.1.3 深度优先搜索	72	4.3.2 正向与反向推理	122
3.1.4 等价价搜索	73	4.4 不确定性推理	124
3.2 启发式搜索	75	4.4.1 概率推理	124
3.2.1 启发式搜索策略	75	4.4.2 贝叶斯推理	128
3.2.2 估价函数	76	4.4.3 模糊逻辑推理与可能性理论	129
3.2.3 有序搜索	76	4.5 非单调推理	131
3.2.4 A*算法	79	4.5.1 缺省推理	131
3.3 博弈树搜索	81	4.5.2 非单调推理系统	134
3.3.1 博弈概述	81	第5章 机器学习	138
3.3.2 极小极大分析法	82	5.1 机器学习的研究意义与发展	
3.3.3 α - β 剪枝技术	83	历史	138
3.4 遗传算法	85	5.1.1 机器学习的定义和研究意义	138
3.4.1 遗传算法的结构	85	5.1.2 机器学习的发展史	139
3.4.2 遗传算法的基本原理	88	5.2 机器学习的主要策略与	
3.4.3 遗传算法的收敛性	88	基本结构	139
3.4.4 遗传算法的性能	89	5.3 常见的几种学习方法	140
3.4.5 进化算法	90	5.3.1 机械学习	140
3.4.6 遗传算法展望	91	5.3.2 基于解释的学习	143
3.5 模拟退火算法	92	5.3.3 基于事例的学习	145

5.3.4 基于概念的学习	149	7.3.2 基于框架的专家系统举例	200
5.3.5 基于类比的学习	151	7.4 基于模型的专家系统	201
5.3.6 归纳学习	151	7.4.1 基于模型的专家系统的概念	201
5.3.7 强化学习	154	7.4.2 基于模型的专家系统举例	202
5.4 基于神经网络的学习	157	7.5 专家系统的设计、评价与 开发	203
5.4.1 神经网络的组成与特性	157	7.5.1 专家系统的设计	203
5.4.2 基于反向传播网络的学习	160	7.5.2 专家系统的评价	206
5.4.3 基于 Hopfield 网络的学习	164	7.5.3 专家系统的开发工具	209
5.4.4 基于神经网络的推理	167	7.6 专家系统设计举例	214
第 6 章 规划系统	173	7.6.1 专家知识的描述	214
6.1 规划的作用与任务	173	7.6.2 知识的使用	218
6.1.1 规划的概念	173	7.6.3 决策的解释	221
6.1.2 规划的作用与主要问题	173	7.6.4 MYCIN 系统概述	222
6.2 基于谓词逻辑的规划	176	第 8 章 自然语言理解	225
6.2.1 规划世界模型的谓词 逻辑表示	176	8.1 语言及其理解的一般问题	225
6.2.2 基于谓词逻辑规划的基本 过程	177	8.1.1 语言和语言理解	225
6.3 STRIPS 规划系统	179	8.1.2 自然语言理解研究的进展	226
6.3.1 积木世界的机器人规划	179	8.1.3 自然语言理解过程的层次	227
6.3.2 STRIPS 规划系统	182	8.2 句法和语法的自动分析	228
6.4 分层规划	186	8.2.1 句法模式匹配和转移网络	228
6.4.1 长度优先搜索	186	8.2.2 扩充转移网络	229
6.4.2 NOAH 规划系统	187	8.2.3 词汇功能语法	231
第 7 章 专家系统	191	8.2.4 语义的解析	234
7.1 专家系统概述	191	8.3 语言理解	235
7.1.1 专家系统的一般特点	191	8.3.1 简单句的理解方法	235
7.1.2 专家系统的结构与类型	192	8.3.2 复合句的理解方法	238
7.2 基于规则的专家系统	197	8.4 机器翻译	240
7.2.1 基于规则的专家系统的基本 结构	197	8.5 语音识别	243
7.2.2 基于规则的专家系统举例	198	8.5.1 语音识别的发展历史	243
7.3 基于框架的专家系统	199	8.5.2 语音识别的基本原理	243
7.3.1 基于框架的专家系统的概念	199	8.5.3 语音识别中的难点	244
		8.5.4 语音识别的关键技术	245
		8.6 应用举例	246

8.6.1 自然语言自动理解系统	246	9.5.3 模糊控制系统	270
8.6.2 机器翻译系统 ARIANE	248	9.5.4 学习控制系统	273
8.6.3 自然语言问答系统	250	9.5.5 神经控制系统	276
第9章 智能控制	253	9.5.6 进化控制系统	280
9.1 智能控制概述	253	第10章 人工智能程序设计	284
9.1.1 智能控制的产生和发展	253	10.1 符号和逻辑处理编程语言	284
9.1.2 智能控制的定义	255	10.2 LISP 语言	285
9.2 智能控制的研究领域	255	10.2.1 LISP 的特点和数据结构	285
9.3 智能控制的学科结构理论	258	10.2.2 LISP 的基本函数	287
9.3.1 二元结构理论	259	10.2.3 递归和迭代	291
9.3.2 三元结构理论	259	10.2.4 LISP 编程举例	294
9.3.3 四元结构理论	260	10.3 PROLOG 语言	297
9.4 智能控制的特点与系统一般 结构	263	10.3.1 语法与数据结构	297
9.4.1 智能控制的特点	263	10.3.2 PROLOG 程序设计原理	298
9.4.2 智能控制系统的一般结构	263	10.3.3 PROLOG 编程举例	300
9.5 智能控制系统	264	10.4 专用开发工具与 人工智能机	301
9.5.1 递阶智能控制系统	264	附录 网络课程使用说明	302
9.5.2 专家控制系统	268	参考文献	306

第 1 章 绪 论

从 1956 年正式提出人工智能的概念起,近 50 年来,人工智能获得了长足的发展,成为一门应用广泛的交叉和前沿科学。通俗地说,人工智能的目的就是让计算机能够像人一样思维。人类已经制造出汽车、火车、飞机、收音机等,它们分别模仿人体某些器官的功能。那么,能否让机器模仿人脑的功能呢?目前,对于由数十亿神经细胞组成的人脑,人类还知之甚少,模仿它或许是天下最困难的事情了。

当计算机出现后,人类真正有了可以模拟自己思维的工具。在此后的 50 多年中,无数科学家和工程师为这个目标而不懈努力着。现在人工智能研究已不再是少数科学家的专利了,全世界几乎所有大学的计算机系都开设这门课程。经过长期的改进和发展,现在计算机已经变得十分聪明了。例如,1997 年 5 月,IBM 公司研制的深蓝(Deep Blue)计算机战胜了国际象棋大师卡斯帕洛夫(Kasparov)。人工智能始终是计算机科学的前沿学科,人工智能的发展带动着计算机学科许多领域的发展。

进入 21 世纪以来,人工智能理论正酝酿着新的突破。例如,人工生命的提出,意味着人类试图从传统的工程技术途径,开辟生物工程技术途径发展人工智能;同时,人工智能的发展,又将作为人工生命科学的重要支柱和推动力量。此外,计算智能已成为人工智能一个新的研究和应用领域,为人工智能提供新的理论基础和新的研究方向。可以预言:人工智能的研究成果必将能够创造出更多、更高级的智能“产品”,为发展国民经济和改善人类生活作出更大贡献。

1.1 人工智能的定义与发展

1.1.1 人工智能的定义

要给人工智能下个准确的定义是困难的。像许多新兴学科一样,人工智能至今尚无统一的定义。人类的许多活动,如解答算题、猜测谜语、进行讨论、编制计划、编写计算机程序、驾驶汽车和骑自行车等,都需要“智能”。如果机器能够执行这种任务,就可以认为机器已具有某种性质的“人工智能”。

逻辑学派(Logicism)、仿生学派(Bionicsism)、生理学派(Physiologism)、计算机学派、心理学派和语言学派等。不同学科背景的学者对人工智能有不同的理解,提出不同的观点。这些观点源于符号主义(Symbolism)、连接主义(Connectionism)和行为主义(Actionism)等。

在 1.3 节中将综述这些学派的主要观点。下面从狭义角度给出人工智能的一些相关定义。

定义 1.1 智能机器(intelligent machine) 能够在各类环境中自主地或交互地执行各种拟人任务(anthropomorphic tasks)的机器。

例 1.1 能够模拟人的思维,进行博弈的计算机,如由 IBM 公司研制的深蓝(Deep Blue)计算机。

例 1.2 能够进行深海探测的自主潜水机器人。

例 1.3 在星际探险中的移动机器人,如美国 2004 年 1 月发射的火星探测车“勇气号”(Spirit)和“机遇号”(Opportunity)以及 2004 年 8 月发射的水星探测车“信使号”(Messenger)。

定义 1.2 人工智能 斯坦福大学的尼尔逊(Nilsson)提出:人工智能是关于知识的科学(知识的表示、知识的获取以及知识的运用)。本书中,我们首先从学科的范围来定义。

定义 1.3 人工智能(学科)是计算机科学中涉及研究、设计和应用智能机器的一个分支。它的近期主要目标在于研究用机器来模仿和执行人脑的某些智能功能,并开发相关理论和技术。

定义 1.4 从人工智能所实现的功能来定义:人工智能(能力)是智能机器所执行的通常与人类智能有关的功能,如判断、推理、证明、识别、感知、理解、设计、思考、规划、学习和问题求解等思维活动。

1.1.2 人工智能的起源与发展

人工智能的发展是以硬件与软件的发展为基础的,它的发展经历了漫长的发展历程。人类很早就开始研究自身的思维形式。早在公元前 384—322 年,在亚里士多德(Aristotle)研究他称之为三段论的演绎推理时就迈出了向人工智能发展的早期步伐,可以把它看作原始的知识表达规范。三段论是以真言判断为其前提的一种演绎推理,它借助于一个共同项,把两个直言判断联系起来,从而得出结论。例如,一切金属都是能够熔解的;因为铁是金属,所以铁是能够熔解的。

人工智能的发展大致可分为以下 3 个阶段。

1. 孕育时期

- (1) 20 世纪 30 年代和 40 年代,两件最重要的事:数理逻辑和关于计算的新思想。
- (2) 20 世纪 40 年代,贝尔实验室 M 系列继电器计算机。

(3) 1946 年 2 月 15 日,世界上第一台通用电子数字计算机 ENIAC 研制成功。ENIAC 在通用性、简单性和可编程方面取得的成功,标志现代电子计算机成为现实。今天的计算机从袖珍计算器到巨型机,它们的出现都可追溯到 ENIAC。

2. 形成时期

- (1) 1956 年夏季,在美国的达特茅斯(Dartmouth)大学举行的首次人工智能研讨会,标

志着人工智能学科的诞生。

(2) 1969年,召开第一届国际人工智能联合会议(International Joint Conference on AI)。

(3) 1970年,《人工智能》国际杂志(*International Journal of AI*)创刊。

3. 发展时期

(1) 1968年,“专家系统和知识工程之父”费根鲍姆(Feigenbaum)领导的研究小组开发出第一个专家系统 DENDRAL,用于质谱仪分析有机化合物的分子结构。

(2) 1972—1976年,费根鲍姆小组又开发成功 MYCIN 医疗专家系统,用于抗生素药物治疗。

(3) 1977年,费根鲍姆进一步提出知识工程(knowledge engineering)的概念。

(4) 知识表示、知识利用和知识获取成为人工智能系统的3个基本问题。

自20世纪80年代以来,人工智能学术界出现了一些新的思想,引起热烈的讨论与争论,极大地促进了人工智能的发展。

1.2 人类智能与人工智能

1.2.1 研究认知过程的任务

这一节的开始首先讨论人类认知过程。人的心理活动具有不同的层次,它可以与计算机的层次相比较,见图1.1。

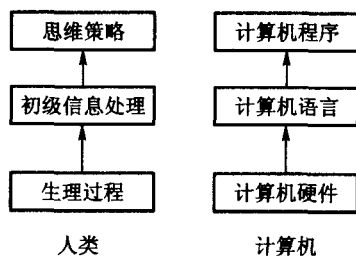


图 1.1 人类认知活动与计算机的比较

心理活动的最高层级是思维策略,中间一层是初级信息处理,最底层级是生理过程,即中枢神经系统、神经元和大脑的活动,与此相对应的是计算机程序、语言和硬件。

研究认知过程的主要任务是探求高层次思维决策与初级信息处理的关系,并用计算机程序来模拟人的思维策略水平,而用计算机语言模拟人的初级信息处理过程。

1.2.2 智能信息处理系统的假设

下面讨论智能信息处理系统的假设。

令 T 表示时间变量, x 表示认知操作(cognitive operation), x 的变化 Δx 为当时机体状态 S (机体的生理和心理状态以及脑子里的记忆等) 和外界刺激 R 的函数。当外界刺激作用到处于某一特定状态的机体时, 便发生变化, 即存在下列关系:

$$\begin{aligned} T &\rightarrow T + 1 \\ x &\rightarrow x + \Delta x \\ \Delta x &= f(S, R) \end{aligned} \quad (1.1)$$

计算机也以类似的原理进行工作。在规定时间内, 计算机存储的记忆相当于机体的状态; 计算机的输入相当于机体施加的某种刺激。在得到输入后, 计算机便进行操作, 使得其内部状态随时间发生变化。可以从不同的层次来研究这种计算机系统。这种系统以人的思维方式为模型进行智能信息处理(intelligent information processing)。显然, 这是一种智能计算机系统。设计适用于特定领域的这种高水平智能信息处理系统, 是研究认知过程的一个具体而又重要的目标。例如, 一个具有智能信息处理能力的自动控制系统就是一个智能控制系统, 它可以是专家控制系统, 也可以是智能决策系统等。人也是一种智能信息处理系统。

信息处理系统又叫符号操作系统(symbol operation system)或物理符号系统(physical symbol system)。所谓符号就是模式(pattern)。任一模式只要能与其他模式相区别, 它就是一个符号。例如, 不同的汉语拼音字母或英文字母就是不同的符号。对符号进行操作就是对符号进行比较, 从中找出相同的和不同的符号。物理符号系统的基本任务和功能就是辨认相同的符号和区别不同的符号。为此, 这种系统就必须能够辨别出不同符号之间的实质差别。符号既可以是物理符号, 也可以是头脑中的抽象符号, 或者是电子计算机中的电子运动模式, 还可以是头脑中神经元的某些运动方式。一个完善的符号系统应具有下列 6 种基本功能:

- (1) 输入符号(input)。
- (2) 输出符号(output)。
- (3) 存储符号(store)。
- (4) 复制符号(copy)。
- (5) 建立符号结构: 通过找出各符号间的关系, 在符号系统中形成符号结构。
- (6) 条件性迁移(conditional transfer): 根据已有符号, 继续完成活动过程。

如果一个物理符号系统具有上述全部 6 种功能, 能够完成这个全过程, 那么它就是一个完整的物理符号系统。人能够输入信号, 如用眼睛看, 用耳朵听, 用手触摸等。计算机也能通过卡片或纸带打孔、磁带或键盘打字等方式输入符号。人具有上述 6 种功能; 现代计算机也具备物理符号系统的这 6 种功能。

假设 任何一个系统, 如果它能表现出智能, 那么它就必定能够执行上述 6 种功能。反之, 任何系统如果具有这 6 种功能, 那么它就能够表现出智能; 这种智能指的是人类所具有的那种智能。把这个假设称为物理符号系统的假设。

物理符号系统的假设伴随有 3 个推论,或称为附带条件。

推论 1 既然人具有智能,那么他(她)就一定是一个物理符号系统。

推论 2 既然计算机是一个物理符号系统,它就一定能够表现出智能。

推论 3 既然人是一个物理符号系统,计算机也是一个物理符号系统,那么就能够用计算机来模拟人的活动。

现代控制论之父维纳(Wiener)于 1940 年提出现代计算机设计的五原则:

- (1) 不是模拟式,而是数字式;
- (2) 由电子元件构成,尽量减少机械部件;
- (3) 采用二进制,而不是十进制;
- (4) 内部存放计算表;
- (5) 在计算机内部存储数据。

1940 年,维纳开始考虑计算机如何能像大脑一样工作,他发现了二者的相似性。维纳认为计算机是一个进行信息处理和信息转换的系统,只要这个系统能得到数据,机器本身就应该几乎能做任何事情。维纳系统地创建了控制论,根据这一理论,一个机械系统完全能进行运算和记忆。

1.2.3 人类智能的计算机模拟

能否用机器模拟人类智能行为?帕梅拉·麦考达克(Pamela McCorduck)曾经在她的著名的人工智能历史著作《机器思维》(*Machine Who Think*, 1979)中指出:在复杂的机械装置与智能之间存在着长期的联系。从几世纪前出现的神话般的复杂巨钟和机械自动机开始,人们已对机器操作的复杂性与自身的智能活动进行直接联系。

被称为人工智能之父的著名英国科学家图灵(Turing)不仅创造了一个简单的、通用的非数字计算模型,而且直接证明了计算机可能以某种被理解为智能的方式工作。1950 年,图灵发表了题为《计算机能思考吗?》的论文,给人工智能下了一个定义,而且论证了人工智能的可能性。

图灵设计了一个用以判断机器是否具有智慧的“图灵测试”。测试由一男(A)、一女(B)和一名询问者(C)进行;C与A、B隔离,通过电传打字机与A、B对话。询问者只知道二人的称呼是X、Y,通过提问以及回答来判断,最终作出“X是A,Y是B”或者“X是B,Y是A”的结论。游戏中,A必须尽力使C判断错误,而B的任务是帮助C。

可用一个机器代替测试中的A,并且机器将试图使得C相信它是一个人。如果机器通过了图灵测试,就认为它是“智慧”的。

物理符号系统假设的推论 1 也告诉我们,人有智能,所以他是一个物理符号系统;推论 3 指出,可以编写出计算机程序去模拟人类的思维活动。这就是说,人和计算机这两个物理符号系统所使用的物理符号是相同的,因而计算机可以模拟人类的智能活动过程。

1.3 人工智能的学派及其争论

1.3.1 人工智能的主要学派

目前人工智能的主要学派有符号主义、连接主义和行为主义。

1. 符号主义(Symbolicism)

又称为逻辑主义(Logicism)、心理学派(Psychologism)或计算机学派(Computerism),其原理主要为物理符号系统(即符号操作系统)假设和有限合理性原理。

该学派认为人工智能源于数理逻辑。数理逻辑从19世纪末起就获迅速发展,到20世纪30年代开始用于描述智能行为。计算机出现后,又在计算机上实现了逻辑演绎系统。该学派早在1956年就首先采用“人工智能”这个术语,后来又发展了启发式算法、专家系统和知识工程理论与技术,并在20世纪80年代取得很大发展。符号主义曾长期一枝独秀,为人工智能的发展作出重要贡献,尤其是专家系统的成功开发与应用,为人工智能走向工程应用和实现理论联系实际具有特别重要意义。在人工智能的其他学派出现之后,符号主义仍然是人工智能的主流派。这个学派的代表有纽厄尔、肖、西蒙和尼尔逊等。

2. 连接主义(Connectionism)

又称为仿生学派(Bionicsism)或生理学派(Physiologism),其原理主要为神经网络及神经网络间的连接机制与学习算法。

该学派认为人工智能源于仿生学,特别是人脑模型的研究。它的代表性成果是1943年由生理学家麦卡洛克(McCulloch)和数理逻辑学家皮茨(Pitts)创立的脑模型,即MP模型。20世纪60—70年代,连接主义,尤其是对以感知机(perceptron)为代表的脑模型的研究曾出现过热潮。由于当时的理论模型、生物原型和技术条件的限制,脑模型研究在20世纪70年代后期至80年代初期落入低潮。直到霍普菲尔德(Hopfield)在1982年和1984年发表两篇重要论文,提出用硬件模拟神经网络时,连接主义又重获关注。1986年鲁梅尔哈特(Rumelhart)等人提出多层网络中的反向传播(BP)算法。此后,连接主义势头大振,从模型到算法,从理论分析到工程实现,为神经网络计算机走向市场打下基础。现在,人们对ANN的研究热情仍然不减。

3. 行为主义(Actionism)

又称进化主义(Evolutionism)或控制论学派(Cyberneticsism),其原理为控制论及“感知-动作”型控制系统。

该学派认为人工智能源于控制论。控制论思想早在20世纪40—50年代就成为时代思潮的重要部分,影响了早期的人工智能工作者。维纳和麦克洛克(McCulloch)等人提出的控制论和自组织系统以及钱学森等人提出的工程控制论和生物控制论,影响了许多领域。控

制论把神经系统的工作原理与信息理论、控制理论、逻辑以及计算机联系起来。早期的研究工作重点是模拟人在控制过程中的智能行为和作用,如对自寻优、自适应、自校正、自镇定、自组织和自学习等控制论系统的研究,并进行“控制论动物”的研制。到20世纪60—70年代,上述这些控制论系统的研究取得一定进展,播下智能控制和智能机器人的种子,并在20世纪80年代诞生了智能控制和智能机器人系统。行为主义是20世纪末才以人工智能新学派的面孔出现的,引起许多人的兴趣与研究。这一学派的代表作首推布鲁克斯(Brooks)的六足行走机器人,它被看做新一代的“控制论动物”,是一个基于“感知-动作”模式的模拟昆虫行为的控制系统。

1.3.2 对人工智能基本理论的争论

不同学派对人工智能的基本理论问题,诸如定义、基础、核心要素、认知过程、学科体系以及人工智能与人类智能的关系等,均有不同的看法。

1. 符号主义

该学派认为人的认知基元是符号,而且认知过程即符号操作过程。它认为人是一个物理符号系统,计算机也是一个物理符号系统,因此,人们就能够用计算机来模拟人的智能行为,即用计算机的符号操作来模拟人的认知过程。也就是说,人的思维是可操作的。它还认为,知识是信息的一种形式,是构成智能的基础。人工智能的核心问题是知识表示、知识推理和知识运用。知识可用符号表示,也可用符号进行推理,因而有可能建立起基于知识的人类智能和机器智能的统一理论体系。

2. 连接主义

该学派认为人的思维基元是神经元,而不是符号处理过程。它对物理符号系统假设持反对意见,认为人脑不同于电脑,并提出连接主义的大脑工作模式和用于取代符号操作的电脑工作模式。

3. 行为主义

该学派认为智能取决于感知和行动(所以被称为行为主义),提出智能行为的“感知-动作”模式。行为主义者认为智能不需要知识、不需要表示、不需要推理;人工智能可以像人类智能一样逐步进化(所以也称为进化主义);智能行为只能在现实世界中与周围环境交互作用而表现出来。行为主义还认为:符号主义(以及连接主义)对真实世界客观事物的描述及其智能行为工作模式是过于简化的抽象,因而是不能真实地反映客观存在的。

1.3.3 对人工智能技术路线的争论

如何在技术上实现人工智能系统、研制智能机器和开发智能产品,即沿着什么技术路线和策略来发展人工智能,也存在不同的派别,即不同的路线。