

免费提供
电子教案

高等院校规划教材
计算机科学与技术系列

人工智能导论

鲍军鹏 张选平 等编著



机械工业出版社
CHINA MACHINE PRESS

高等院校规划教材·计算机科学与技术系列

人工智能导论

鲍军鹏 张选平 等编著



机械工业出版社

本书系统地阐述了人工智能的基本理论、基本技术、研究方法和应用领域等内容,比较全面地反映了近20年来人工智能研究领域的进展,并根据人工智能的发展动向对一些传统内容作了取舍,详细介绍了机器学习方面的内容。全书分为8章,内容涉及人工智能的基本概念、知识工程、确定性推理和不确定性推理、搜索策略、机器学习、人工神经网络,以及模式识别、自然语言处理和智能体等方面。每章后面都附有习题,以供读者练习。

本书可作为计算机专业本科生和其他相关学科本科生、研究生的教材,也可作为有关科技人员的参考用书。

图书在版编目(CIP)数据

人工智能导论/鲍军鹏,张选平,吕园园编著. —北京:机械工业出版社, 2009.11
(高等院校规划教材·计算机科学与技术系列)
ISBN 978-7-111-28837-4

I. 人… II. ①鲍…②张…③吕… 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(2009)第207453号

机械工业出版社(北京市百万庄大街22号 邮政编码100037)
责任编辑:陈皓 常建丽
责任印制:杨曦

北京蓝海印刷有限公司印刷

2010年1月第1版·第1次印刷
184mm×260mm·20.25印张·501千字
0001—3000册
标准书号:ISBN 978-7-111-28837-4
定价:35.00元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

电话服务

网络服务

社服务中心:(010)88361066

门户网:<http://www.cmpbook.com>

销售一部:(010)68326294

教材网:<http://www.cmpedu.com>

销售二部:(010)88379649

封面无防伪标均为盗版

读者服务部:(010)68993821

出版说明

计算机技术的发展极大地促进了现代科学技术的发展，明显地加快了社会发展的进程。因此，各国都非常重视计算机教育。

近年来，随着我国信息化建设的全面推进和高等教育的蓬勃发展，高等院校的计算机教育模式也在不断改革，计算机学科的课程体系和教学内容趋于更加科学和合理，计算机教材建设逐渐成熟。在“十五”期间，机械工业出版社组织出版的大量计算机教材，包括“21世纪高等院校计算机教材系列”、“21世纪重点大学规划教材”、“高等院校计算机科学与技术‘十五’规划教材”、“21世纪高等院校应用型规划教材”等，均取得了可喜成果，其中多个品种的教材被评为国家级、省部级的精品教材。

为了进一步满足计算机教育的需求，机械工业出版社策划开发了“高等院校规划教材”。这套教材是在总结我社以往计算机教材出版经验的基础上策划的，同时借鉴了其他出版社同类教材的优点，对我社已有的计算机教材资源进行整合，旨在大幅提高教材质量。我们邀请多所高校的计算机专家、教师及教务部门针对此次计算机教材建设进行了充分的研讨，达成了许多共识，并由此形成了“高等院校规划教材”的体系架构与编写原则，以保证本套教材与各高等院校的办学层次、学科设置和人才培养模式等相匹配，满足其计算机教学的需要。

本套教材包括计算机科学与技术、软件工程、网络工程、信息管理与信息系统、计算机应用技术以及计算机基础教育等系列。其中，计算机科学与技术系列、软件工程系列、网络工程系列和信息管理与信息系统系列是针对高校相应专业方向的课程设置而组织编写的，体系完整，讲解透彻；计算机应用技术系列是针对计算机应用类课程而组织编写的，着重培养学生利用计算机技术解决实际问题的能力；计算机基础教育系列是为大学公共基础课层面的计算机基础教学而设计的，采用通俗易懂的方法讲解计算机的基础理论、常用技术及应用。

本套教材的内容源自致力于教学与科研一线的骨干教师与资深专家的实践经验和研究成果，融合了先进的教学理念，涵盖了计算机领域的核心理论和最新的应用技术，真正在教材体系、内容和方法上做到了创新。另外，本套教材根据实际需要配有电子教案、实验指导或多媒体光盘等教学资源，实现了教材的“立体化”建设。本套教材将随着计算机技术的进步和计算机应用领域的扩展而及时改版，并及时吸纳新兴课程和特色课程的教材。我们将努力把这套教材打造成为国家级或省部级精品教材，为高等院校的计算机教育提供更好的服务。

对于本套教材的组织出版工作，希望计算机教育界的专家和老师们能提出宝贵的意见和建议。衷心感谢计算机教育工作者和广大读者的支持与帮助！

机械工业出版社

前 言

人工智能是计算机科学的一个核心研究领域。计算机科学发展的终极目标就是不但让计算机具有超强的计算能力，还让计算机具有像人一样的智能，最终使计算机能够做人类所做的一切事情，并且可能有比人类更强的能力。这是人类的一个梦想。人工智能作为一门学科从正式提出到现在，已经走过了半个多世纪。在过去的半个多世纪里，人工智能学科的发展经历了风风雨雨、起起落落。现在，人工智能正处于一个蓬勃发展、更加深入的阶段。

人工智能涉及的子领域以及从事研究的学者范围非常广泛。这是一个典型的交叉科学。因为智能本身就是一个极其复杂的现象。不同的人从不同角度和不同观点都可以获得对智能的认识和模拟。进入 21 世纪以来，人工智能的发展一日千里。其研究成果和各种文献更是不计其数。为了让大家能够了解人工智能发展的基本现状，掌握人工智能研究的大致热点和基本原理与方法，我们编写了本书。本书并不试图详细、深入地介绍人工智能中每一方面的最新成果。因为本书的每一章都涵盖了人工智能的一个或者多个研究领域。如果要深入、全面地介绍这个领域，那么每一章至少都是一部专著，所以本书的意图就是让读者了解人工智能研究和发展的基本轮廓，对人工智能有一个基本的认识，知道目前人工智能研究中的一些热点，掌握人工智能研究和应用中的一些基本的、普遍的、比较广泛的原理和方法。归结成一句话，就是把读者引进人工智能的大门。入门之后，读者就要根据自己的兴趣选择各自的研究方向，然后再阅读该方向的专著。因此本书定名为“导论”。

全书共分 8 章。第 1 章是绪论，介绍一些关于人工智能的基本观点、基本途径和人工智能中的主要研究内容。第 2 章是知识工程，主要介绍关于知识的表示和知识组织的问题。第 3 章介绍的确定性推理与第 4 章介绍的不确定性推理，都是关于如何运用知识（即推理）的问题。第 5 章介绍了搜索策略。这些都是传统人工智能教材的内容。

知识和推理是智能的核心。有些学者所称的人工智能仅指这一部分；而把人工神经网络、模式识别、机器学习等问题与人工智能并列起来，甚至对立起来。笔者认为，传统的知识和推理部分主要运用物理符号系统进行研究，应该只是人工智能研究途径中的一种，并非全部。以人工神经网络为代表的连接主义途径也是研究人工智能的基本途径。学习问题是扩展知识、发展智能的必要手段，当然也是人工智能必不可少的一部分。模式识别正是人工智能在人类社会中的重要应用。这几部分也是当前人工智能研究中的一些热点和重点。

本书根据人工智能近 20 年来的发展，引入了一些新内容。本书的第 6 章介绍了当前机器学习的一些基本问题、基本方法和基本思路。第 7 章介绍了人工神经网络的基本模型和算法。由于模式识别的内容和机器学习的内容有很大一部分是重叠的，所以本书没有把模式识别单独作为一章进行介绍。第 8 章把模式识别、自然语言处理和智能体等当前人工智能研究和发展的热点集中在一起，进行了简单介绍。

本书的第 1、6、7 章由鲍军鹏编写，第 2~5 章由张选平编写，第 8 章由吕园园编写。由于作者水平有限，书中难免有不妥之处，恳请广大读者批评指正。

编 者

目 录

出版说明

前言

第 1 章 绪论	1
1.1 什么是人工智能	1
1.1.1 关于智能	1
1.1.2 人工智能的研究目标	2
1.2 人工智能发展简史	3
1.3 人工智能研究方法	8
1.3.1 人工智能研究的特点	8
1.3.2 人工智能研究的途径	9
1.3.3 人工智能研究的资源	11
1.4 人工智能研究及应用领域	12
1.4.1 问题求解与博弈	12
1.4.2 专家系统	13
1.4.3 自动定理证明	14
1.4.4 机器学习	15
1.4.5 人工神经网络	16
1.4.6 模式识别	17
1.4.7 计算机视觉	17
1.4.8 自然语言处理	18
1.4.9 智能体	19
1.4.10 智能控制	20
1.4.11 机器人学	21
1.4.12 人工生命	22
1.5 本章小结	23
1.6 习题	23
第 2 章 知识工程	24
2.1 概述	24
2.2 知识表示方法	26
2.2.1 经典逻辑表示法	26
2.2.2 产生式表示法	27
2.2.3 层次结构表示法	32
2.2.4 网络结构表示法	34

2.2.5	其他表示法	36
2.3	知识获取与管理	39
2.3.1	知识获取的任务	39
2.3.2	知识获取的方式	41
2.3.3	知识管理	42
2.3.4	本体论	44
2.4	基于知识的系统	46
2.4.1	什么是知识系统	46
2.4.2	专家系统	47
2.4.3	知识系统举例	48
2.5	本章小结	51
2.6	习题	51
第3章	确定性推理	53
3.1	概述	53
3.1.1	推理方式与分类	53
3.1.2	推理控制策略	54
3.1.3	知识匹配	55
3.2	自然演绎推理	57
3.3	归结演绎推理	59
3.3.1	海伯伦理论	59
3.3.2	鲁宾逊归结原理	62
3.3.3	归结反演	65
3.3.4	归结策略	66
3.3.5	应用归结原理求解问题	73
3.4	与或形演绎推理	76
3.4.1	与或形正向演绎推理	76
3.4.2	与或形逆向演绎推理	79
3.4.3	与或形双向演绎推理	81
3.5	本章小结	82
3.6	习题	83
第4章	不确定性推理	85
4.1	概述	85
4.2	基本概率方法	88
4.3	主观贝叶斯方法	90
4.3.1	不确定性的表示	91
4.3.2	不确定性的传递算法	92
4.3.3	结论不确定性的合成算法	95

4.4	可信度方法	96
4.4.1	基本可信度模型	96
4.4.2	带阈值限度的可信度模型	98
4.4.3	加权的可信度模型	99
4.4.4	前件带不确定性的可信度模型	100
4.5	模糊推理	102
4.5.1	模糊理论	102
4.5.2	简单模糊推理	110
4.5.3	模糊三段论推理	116
4.5.4	多维模糊推理	118
4.5.5	多重模糊推理	121
4.5.6	带有可信度因子的模糊推理	121
4.6	证据理论	123
4.6.1	D-S 理论	123
4.6.2	基于证据理论的不确定性推理	125
4.7	粗糙集理论	128
4.7.1	粗糙集理论的基本概念	128
4.7.2	粗糙集在知识发现中的应用	132
4.8	本章小结	133
4.9	习题	134
第5章	搜索策略	137
5.1	概述	137
5.1.1	什么是搜索	137
5.1.2	状态空间表示法	137
5.1.3	与或树表示法	139
5.2	状态空间搜索	142
5.2.1	状态空间的一般搜索过程	142
5.2.2	广度优先搜索	143
5.2.3	深度优先搜索	145
5.2.4	有界深度优先搜索	146
5.2.5	启发式搜索	148
5.2.6	A* 算法	151
5.3	与或树搜索	154
5.3.1	与或树的一般搜索过程	154
5.3.2	与或树的广度优先搜索	155
5.3.3	与或树的深度优先搜索	156
5.3.4	与或树的有序搜索	157

5.3.5	博弈树的启发式搜索	161
5.3.6	剪枝技术	164
5.4	本章小结	164
5.5	习题	165
第6章	机器学习	167
6.1	概述	167
6.1.1	什么是机器学习	167
6.1.2	机器学习方法的分类	169
6.1.3	机器学习的基本问题	171
6.1.4	评估学习结果	173
6.2	决策树学习	176
6.2.1	决策树表示法	177
6.2.2	ID3 算法	178
6.2.3	决策树学习的常见问题	180
6.2.4	用决策树学习客户分类	182
6.3	贝叶斯学习	186
6.3.1	贝叶斯法则	186
6.3.2	朴素贝叶斯方法	189
6.3.3	贝叶斯网络	189
6.3.4	EM 算法	191
6.3.5	用贝叶斯方法过滤垃圾邮件	193
6.4	统计学习	194
6.4.1	统计学习理论	195
6.4.2	支持向量机	198
6.4.3	核函数	201
6.4.4	基于支持向量机的车牌识别	202
6.5	遗传算法	205
6.5.1	进化计算	205
6.5.2	遗传算法原理	207
6.5.3	问题编码策略	210
6.5.4	遗传算子	213
6.5.5	遗传算法的理论分析	217
6.5.6	用遗传算法解决 TSP 问题	219
6.6	聚类	221
6.6.1	聚类问题	222
6.6.2	分层聚类方法	223
6.6.3	划分聚类方法	226

6.6.4	基于密度的聚类方法	227
6.6.5	基于网格的聚类方法	229
6.6.6	股票信息的聚类分析	230
6.7	特征选择与提取	232
6.7.1	特征选择	232
6.7.2	常用的特征函数	233
6.7.3	主成分分析	234
6.8	其他学习方法	235
6.8.1	强化学习	235
6.8.2	隐马尔可夫模型	237
6.9	本章小结	239
6.10	习题	239
第7章	人工神经网络	241
7.1	概述	241
7.1.1	人脑神经系统	241
7.1.2	人工神经网络的研究内容与特点	244
7.1.3	人工神经网络的基本形态	245
7.2	感知器	249
7.2.1	简单感知器	249
7.2.2	多层感知器	252
7.3	前馈神经网络	253
7.3.1	反向传播算法	253
7.3.2	反向传播算法中的问题	258
7.3.3	径向基函数网络	259
7.4	反馈神经网络	262
7.4.1	Hopfield 网络	263
7.4.2	离散型 Hopfield 网络	263
7.4.3	连续型 Hopfield 网络	265
7.4.4	Hopfield 网络中的问题	266
7.4.5	用 Hopfield 网络解决 TSP 问题	267
7.5	随机神经网络	268
7.5.1	模拟退火算法	268
7.5.2	波尔兹曼机	270
7.6	自组织神经网络	272
7.6.1	竞争学习	273
7.6.2	自组织特征映射网络	274
7.7	本章小结	275

7.8	习题	276
第8章	人工智能的其他领域	277
8.1	模式识别	277
8.1.1	模式识别的基本问题	277
8.1.2	统计模式识别	278
8.1.3	句法模式识别	280
8.1.4	模糊模式识别	282
8.1.5	神经网络模式识别	282
8.1.6	模式识别的应用	283
8.2	自然语言处理	284
8.2.1	自然语言处理的基本问题	284
8.2.2	语法分析	287
8.2.3	语义分析	289
8.2.4	大规模文本处理	291
8.2.5	自然语言处理的应用	294
8.3	智能体	296
8.3.1	智能体模型	297
8.3.2	多智能体系统的模型	300
8.3.3	多智能体系统的协作、协商与协调	302
8.3.4	多智能体系统的学习与规划	304
8.3.5	智能体间的通信	305
8.3.6	智能体的应用	307
8.4	本章小结	308
8.5	习题	309
	参考文献	310

第1章 绪 论

科学发展的道路从来都不是一条坦途。人工智能发展至今也经历了很多风风雨雨。本章主要介绍关于人工智能的一般观点和研究途径；同时也简单介绍了人工智能的发展简史；最后概括地列举了目前人工智能的一些比较热门的研究方向和应用领域。

1.1 什么是人工智能

人工智能(Artificial Intelligence, AI)从20世纪50年代明确提出以来,已经有了迅猛的发展。如同钢铁机械延展了人的四肢能力一样,人工智能是对人类大脑能力的延展。人工智能研究和发展的本质就是延长和扩展人的智能。如果读者能够理解工业社会中机器如何替代和减轻人的体力劳动,就同样可以想象人工智能将会替代和减轻人的脑力劳动。所以人工智能的研究和发展将会改变人类社会。

事实上,人工智能的应用成果已经给我们的现实生活带来了许多方便。例如,农业专家系统可以帮助农业人员对农作物进行最优栽培和管理;字符识别系统使得计算机可以“看懂”文字,帮助邮政人员自动投递或者分拣信件;语音识别系统使得计算机能够“听懂”人类的语言,有些操作就不必由人亲自动手。本书主要阐述人工智能领域中的基本理论、原理和方法,以及重要应用。读者通过本书可以鸟瞰人工智能的主要研究,对人工智能有一个基本的认识。

1.1.1 关于智能

人工智能,顾名思义,就是用人工的方法实现智能。“人工”是指人可以控制每一个步骤,并且能够达到预期结果的一个物理过程。目前,这个物理过程一般是通过电子线路实现的。现在正在研究的光子、量子 and 分子计算机,则试图用新型材料代替电子半导体,实现逻辑运算。未来的新型计算机将具有可并行计算、运算速度快、能耗低等特性。无论采用什么类型的计算机,在理论上任何可计算问题都可以用计算机软件来实现(未必解决该问题)。所以绝大多数的人工智能研究都是在计算机上进行的。人工智能有时候也被称为机器智能,一般都指使计算机表现出智能或者具有智能行为。

但什么是“智能”这个问题目前并没有一个统一的结论。《现代汉语词典》对“智能”的解释是:智慧和才能;或者具有人的某些智慧和才能。《牛津高阶英语词典》(OXFORD ADVANCED LEARNERS DICTIONARY)对“智能(Intelligence)”的解释是:以逻辑的方式学习、理解、思考事物的能力(The ability to learn understand and think in a logical way about things)。

从唯物主义哲学观点来说,智能是大脑特别是人脑运动的结果或者产物。由于人类对自身以及脑的功能原理还没有认识清楚,所以很难对智能给出确切的定义。知识理论的观点认为智能的基础是知识。因为一个系统之所以有智能是因为它具有可运用的知识,没有知识显然就不可能有智能。思维理论的观点认为智能的关键是思维。因为人的一切智能都来自大脑的思维活动,人类的一切知识都是人类思维的产物。进化理论的观点认为智能取决于感知和

行为，智能就是在系统与周围环境不断“感知—动作”的交互中发展和进化的。

瑟斯蒂姆(Thurstone)认为智能由语言理解、用词流畅、数、空间、联系性记忆、感知速度及一般思维 7 种因子组成。一般认为，智能是知识与智力的总和。其中知识是一切智能行为的基础，而智力是获取知识并运用知识求解问题的能力，是头脑中思维活动的具体体现。

由于对智能有不同理解，所以人工智能现在没有统一的定义。麦卡锡(John McCarthy)认为，人工智能就是要让机器的行为看起来就像是人所表现出的智能行为一样。尼尔逊(Nilsen)认为，人工智能是关于人造物的智能行为，包括知觉、推理、学习、交流和在复杂环境中的行为。巴尔(A. Barr)和费根鲍姆(E. A. Feigenbaum)认为，人工智能属于计算机科学的一个分支，旨在设计智能的计算机系统。也就是说，对照人类在自然语言理解、学习、推理问题求解等方面的智能行为，它所设计的系统应呈现出与之类似的特征。本书认为人工智能就是研究如何使一个计算机系统具有像人一样的智能特征，使其能模拟、延伸、扩展人类智能。通俗地讲，人工智能就是研究如何使得计算机能听、说、读、写、学习、推理，能够适应环境变化，能够模拟出人脑思维活动。总之，人工智能就是要使计算机能够像人一样去思考和行动，完成人类能够完成的工作，甚至在某些方面比人更强。

1.1.2 人工智能的研究目标

人工智能的定义实际上就很明确地指出了人工智能研究的最终目标，即造出一个像人一样具有智能，会思维和行动的计算机系统。对于这个目标可以有两种理解：一种观点认为人工智能的最终目标就是制造出真正能推理和解决问题的智能机器，并且这样的机器能将被认为是具有知觉的，有自我意识的。这种观点被称为“强人工智能”。强人工智能又有两类：一类是类人的人工智能，即机器的思考和推理方式和人的思维方式一模一样。另一类是非类人的人工智能，即机器可以有与人不一样的知觉和意识，使用与人不一样的推理方式。另一种关于人工智能终极目标的观点被称为“弱人工智能”。这种观点认为不可能制造出能真正地推理和解决问题的智能机器，智能机器只不过看起来像是智能的，但是并不真正拥有智能，也不会有自主意识。

目前，人们还没有完全认识清楚智能的本质，但人们对智能的理解在不断变化。在古代如果有一台机器能够自动进行算术运算，恐怕就被认为有智能了。在 19 世纪如果有一台机器能够和人类下棋，那必定会被认为有智能。现代计算机能轻而易举地完成上述功能，还可以超越人类。然而，现在人们对智能行为有了更高的要求。虽然人工智能研究已经取得了很多可观的成就，但是研究人员并不认为目前的智能机器具有智能。

判断一台机器是否具有智能，或者说如何检测一台机器是否会思维，目前还是一个有争议的问题。对于这个问题英国数学家阿兰·图灵(Alan M. Turing)在 1950 年发表了题为《计算机与智能》(Computing Machinery and Intelligence)的论文。文章以“机器能思维吗?”开始论述并提出了著名的“图灵测试”，作为衡量机器智能的准则。图灵测试的基本思想是，让受试者和计算机分别回答一定数量的问题，如果审查答案的人在多数情况下不能正确地判断哪边回答问题的是计算机，那么就说明这台计算机的智能达到了与人接近的水平。

图灵设计了一段对话作为测试内容。在这段对话中，Q 代表人，A 代表机器，并且假设他们都阅读过狄更斯(C. Dickens)的一本小说《匹克威克外传》。对话内容如下：

Q: 你的 14 行诗的首行为“你如同夏日”。你不觉得“春日”更好吗?

A: 它不合韵。

Q: “冬日”如何? 它可是完全合韵的。

A: 它确是合韵, 但没有人愿被比为“冬日”。

Q: 你不是说过匹克威克先生让你能想起圣诞节吗?

A: 是的。

Q: 圣诞节是冬天的一个日子。我想匹克威克先生对这个比喻不会介意吧。

A: 我认为你不够严谨。“冬日”指的是一般的冬天的日子, 而不是某个特别的日子, 如圣诞节。

图灵测试第一次给出了检验计算机是否具有智能的哲学说法。它用一种“模仿游戏”的方式来回避难以严格定义的“智能”。由于判断答案和人回答的相似程度本身很难客观界定, 而且还有很多判断者的主观因素在里面, 因此对图灵测试的质疑有很多。

例如, 加利福尼亚大学伯克利分校的瑟利(John R. Searle)教授提出: 假设一个不懂汉语的人(或者机器)A 有一个充分详细的汉语问答手册。我们不考虑查手册的时间代价。然后对 A 提出一个汉语问题, A 通过汉语符号查阅使用手册, 给出回答。那么, 如果 A 通过查手册做出的回答与懂汉语的人一样, A 懂汉语吗?

图灵测试在实际中使用还有很多困难。实现人工智能的终极目标仍将是一个漫长的过程, 是一个任重道远的艰巨任务。在近期, 人工智能研究的目标是使现有的计算机系统更聪明、更有用, 使它不仅能做一般的数值计算及非数值信息处理, 而且能运用知识处理问题, 能模拟人类的部分智能行为, 成为人类的智能化辅助工具。

人工智能研究的终极目标与近期目标是相辅相成的, 二者之间并无严格的界限。远期目标为近期目标指明了方向, 而近期目标的研究则是实现远期目标的基础。近期目标的研究成果不仅可以造福于当代社会, 还可以进一步增强人们对实现远期目标的信心。科学研究正是通过实现一个又一个的近期目标而逐步接近和达到最终目标的。

1.2 人工智能发展简史

人工智能的发展历史就是人类思索自身的历史。人类能够区别于其他生物, 能够成为万物主宰, 最重要的一点就是人有智能, 并且能够通过学习向下一代稳定地传递智能和不断地扩展智能。而动物只有本能。本能是与生俱来的, 不需要学习获得。动物在后天的活动中也会获得一些经验。但是动物无法像人类一样能把有用的经验积累起来并一代一代稳定地传授下去, 更不会像人类一样能在旧经验的基础上扩展出新经验。“举一反三”、“温故而知新”这些词语描述了人类的思维能力。思维是智能行为的一种体现。那么, 人是如何进行思维的? 这个问题自始至终地贯穿于人工智能的发展过程。对这个问题的探索, 一方面推动着人类对自身的认识, 另一方面也推动着人工智能的进步。

人类从很早的时候就开始思考自身。但是人工智能作为一门科学正式诞生于 1956 年在美国达特茅斯大学(Dartmouth University)召开的一次学术会议上。到目前为止, 人工智能的发展经历了 3 个阶段: 第 1 阶段为孕育阶段(1956 年之前), 第 2 阶段为形成阶段(1956 ~ 1969), 第 3 阶段为发展阶段(1970 年至今)。

1. 孕育阶段(1956 年之前)

自古以来,人类就根据当时的技术条件和认识水平,一直力图创造出某种机器来代替人的部分劳动(包括体力劳动和脑力劳动),以提高征服自然的能力。但是真正对思维和智能进行理性的探索,并抽象成理论体系,则经历了相当漫长的一段时期。在这期间的一些重大研究有:

1) 早在公元前,希腊哲学家亚里士多德(Aristotle)在其著作《工具论》中提出了形式逻辑的一些主要定律。他提出的三段论至今仍是演绎推理的基本依据。

2) 英国哲学家培根(F. Bacon)在《新工具》中系统地提出了归纳法,还提出了“知识就是力量”的警句。

3) 1642 年,法国数学家帕斯卡(Blaise Pascal)发明了第一台机械计算器——加法器(Pascaline),开创了计算机械的时代。

4) 德国数学家莱布尼兹(G. Leibniz)在帕斯卡加法器的基础上发展并制成了可进行全部四则运算的计算器。他还提出了“通用符号”和“推理计算”的概念,使形式逻辑符号化。他认为,可以建立一种通用的符号语言以及在此符号语言上进行推理的演算。这一思想不仅为数理逻辑的产生和发展奠定了基础,而且还是现代机器思维设计思想的萌芽。

5) 英国逻辑学家布尔(G. Boole)创立了布尔代数。他在《思维法则》一书中,首次用符号语言描述了思维活动的基本推理法则。

6) 英国数学家图灵在 1936 年提出了一种理想计算机的数学模型,即图灵机。这为后来电子数字计算机的问世奠定了理论基础。

7) 美国神经生理学家麦库仑奇(W. McCulloch)和佩兹(W. Pitts)在 1943 年提出了第一个神经网络模型——M-P 模型,开创了微观人工智能的研究工作,奠定了人工神经网络发展的基础。

8) 美国数学家马士利(J. W. Mauchly)和埃克特(J. P. Eckert)在 1946 年研制出了世界上第一台电子数字计算机 ENIAC(Electronic Numerical Integrator and Calculator)。

9) 1950 年,图灵发表论文《Computing Machinery and Intelligence》,提出了图灵测试。

19 世纪以来,数理逻辑、自动机理论、控制论、信息论、仿生学、计算机、心理学等科学技术的进展,为人工智能的诞生奠定了思想、理论和物质基础。在 20 世纪 50 年代,计算机应用还局限在数值处理方面,如计算弹道等。但是,1950 年 C. E. Shannon 完成了第一个下棋程序,开创了非数值计算的先河。Newell、Simon、MaCarthy 和 Minsky 等均提出以符号为基础的计算。这一切使得人工智能作为一门独立的学科呼之欲出。

2. 形成阶段(1956 ~ 1969)

1956 年夏季, J. McCarthy、M. L. Minsky、N. Lochester 和 C. E. Shannon 邀请 T. More、A. L. Samuel、O. Selfridge、R. Solomonoff、A. Newell、H. A. Simon 等 10 人在美国的 Dartmouth 大学召开了一次学术研讨会,讨论关于机器智能的有关问题。在会上, McCarthy 提议正式采用“人工智能”这一术语。这次历史性的会议被认为是人工智能学科正式诞生的标志。从此在美国开始形成了以人工智能为研究目标的几个研究组,如 Newell 和 Simon 的 Carnegie-RAND 协作组、Samuel 和 Gelernter 的 IBM 公司工程课题研究组、Minsky 和 McCarthy 的 MIT 研究组等,这一时期人工智能的研究工作有:

1) 1956 年, Samuel 研究出了具有自学习能力的西洋跳棋程序。这个程序能从棋谱中学

习,也能从下棋实践中提高棋艺。这是机器模拟人类学习过程卓有成就的探索。1959年这个程序曾战胜设计者本人,1962年还击败了美国 Connecticut 州的跳棋冠军。

2) 1957年,A. Newell、J. Shaw 和 H. Simon 等人的心理学小组编制出一个称为逻辑理论机(The Logic Theory Machine)的数学定理证明程序。当时该程序证明了 B. A. W. Russell 和 A. N. Whitehead 的《数学原理》一书第2章中的38个定理。1963年修订的程序在大机器上证明了该章中的52个定理。1958年,美籍华人王浩在 IBM-740 机器上用3~5分钟证明了《数学原理》中有关命题演算的全部定理220个,还证明了谓词演算中150个定理的85%。1965年,J. A. Robinson 提出了归结原理,为定理的机器证明作出了突破性贡献。

3) A. Newell、J. Shaw 和 H. Simon 等人揭示了人在解题时的思维过程大致可归结为3个阶段:先想出大致的解题计划;根据记忆中的公理、定理和推理规则组织解题过程;进行方法和目的分析,修正解题计划。这种思维活动不仅解数学题时如此,解决其他问题时也大致如此。基于这一思想,他们于1960年又编制了能解10种不同类型课题的通用问题求解程序(General Problem Solving, GPS)。

4) A. Newell、J. Shaw 和 H. Simon 等人还发明了编程的表处理技术和 NSS 国际象棋机。后来,他们的学生还做了许多工作,如人的口语学习和记忆的 EPAM 模型(1959年)、早期自然语言理解程序 SAD-SAM 等。此外,他们还对启发式求解方法进行了探讨。

5) 1959年,Selfridge 推出了一个模式识别程序。1965年,Roberts 编制了可分辨积木构造的程序。

6) 1960年,McCarthy 在 MIT 研制出了人工智能语言 LISP。

7) 1965年,Stanford 大学的 E. A. Feigenbaum 开展了专家系统 DENDRAL 的研究,并于1968年投入使用。这个专家系统能根据质谱仪的试验,通过分析推理决定化合物的分子结构。其分析能力接近甚至部分超过有关化学专家的水平。该专家系统的成功不仅为人们提供了一个实用的智能系统,而且对知识表示、存储、获取、推理及利用等技术是一次非常有益的探索,为以后的专家系统树立了一个榜样,对人工智能的发展产生了深刻的影响。

这些早期的成果充分表明了人工智能作为一门新兴学科正在蓬勃发展,也使当时的研究者们对人工智能产生了盲目乐观的看法。例如,1958年 Newell 和 Simon 提出了以下4个预测:

- 10年内,计算机将成为世界象棋冠军。
- 10年内,计算机将发现或证明有意义的数学定理。
- 10年内,计算机将能谱写优美的乐曲。
- 10年内,计算机将能实现大多数的心理学理论。

而实际上却是:

- 1997年5月12日,IBM公司的“深蓝”超级电脑才第一次击败国际象棋世界冠军卡斯帕罗夫(Kasparov)。
- 1976年,美国数学家 Kenneth Appel 等人在3台大型电子计算机上用了1200小时 CPU 时间完成了四色定理证明。1977年,我国数学家吴文俊在《中国科学》上发表论文《初等几何判定问题与机械化问题》,提出了一种几何定理机械化证明方法,被称之为“吴氏方法”。1980年,吴文俊在 HP9835A 机上成功证明了勾股定理、西姆逊线定理、帕普斯定理、帕斯卡定理、费尔巴哈定理,并在45个帕斯卡点中发现了20条帕斯卡圆

锥曲线。

- 至今仍未有公认的由计算机谱写的名曲。
- 至今计算机仍然缺乏对常识的推理能力。

3. 发展阶段(1970 年至今)

从 20 世纪 70 年代开始, 人工智能的研究已经逐渐在世界各国开展起来。此时召开并创办了多个人工智能国际会议和国际期刊, 对推动人工智能的发展, 促进学术交流起到了重要作用。1969 年成立了国际人工智能联合会议(International Joint Conference on Artificial Intelligence, IJCAI)。1970 年创刊了著名的国际期刊《人工智能》。1974 年成立了欧洲人工智能会议(European Conference on Artificial Intelligence, ECAI)。此外, 许多国家也都有本国的人工智能学术团体。英国爱丁堡大学那时候就成立了“人工智能”系。日本和西欧一些国家虽起步较晚, 但发展都较快。前苏联对人工智能研究也开始予以重视。我国是从 1978 年开始人工智能课题研究的, 那时候主要集中在定理证明、汉语自然语言理解、机器人及专家系统等方面的研究。我国也先后成立了中国人工智能学会、中国计算机学会、人工智能和模式识别专业委员会和中国自动化学会模式识别与机器智能专业委员会等学术团体, 开展这方面的学术交流。

此时也涌现出了一批重要的研究成果。例如, 1972 年法国马赛大学的 A. Comerauer 提出并实现了逻辑程序设计语言 PROLOG。Stanford 大学的 E. H. Shortliffe 等人从 1972 年开始研制用于诊断和治疗感染性疾病的专家系统 MYCIN。

但是, 在 20 世纪 60 年代末至 70 年代末, 人工智能研究遭遇了一些重大挫折。例如, Samuel 的下棋程序与世界冠军对弈时, 五局中败了四局。机器翻译研究中碰到了不少问题。例如, “果蝇喜欢香蕉”的英语句子“Fruit flies like a banana.”会被翻译成“水果像香蕉一样飞行”。“心有余而力不足”的英语句子“The spirit is willing but the flesh is week.”被翻译成俄语, 然后再由俄语翻译回英语后, 竟变成了“The vodka is strong but meat is rotten.”, 即“伏特加酒虽然很浓, 但肉是腐烂的”。这种错误都是由多义词造成的。这说明仅仅依赖一部双向词典和简单的语法、词法知识还不足以实现准确的机器翻译。在问题求解方面, 即便是对于良结构问题, 当时的人工智能程序也无法面对巨大的搜索空间。更何况现实世界中的问题绝大部分是非良结构的或者是不确定的。在人工神经网络方面, 感知机模型无法通过学习解决异或(XOR)等非线性问题。

这些问题使人们对人工智能研究产生了质疑。1966 年, ALPAC 的负面报告导致美国政府取消了对机器翻译的资助。1973 年, 英国剑桥大学应用数学家赖特黑尔(James Lighthill)爵士的报告认为, “人工智能研究即使不是骗局, 至少也是庸人自扰。”当时英国政府接受了该报告的观点, 取消了对人工智能研究的资助。这些指责中最著名的是明斯基(Minsky)的批评。1969 年, 明斯基和贝波特(Papert)出版了《感知机》(Perceptron)一书。在书中他批评感知机无法解决非线性问题, 如异或(XOR)问题等。而复杂性信息处理应该以解决非线性问题为主。而且他还认为, 几何方法应该代替分析方法作为主要数学手段。明斯基的批评导致美国政府取消了对人工神经网络研究的资助。人工神经网络的研究此后被冷落了 20 年。人工智能研究出现了一个暂时的低潮。

面对困难和挫折, 人工智能研究者们对以前的思想和方法进行了反思和检讨。1977 年, Edward Feigenbaum 在第五届国际人工智能联合会议上作了题为《人工智能的艺术: 知识工程