

21世纪高等学校计算机规划教材

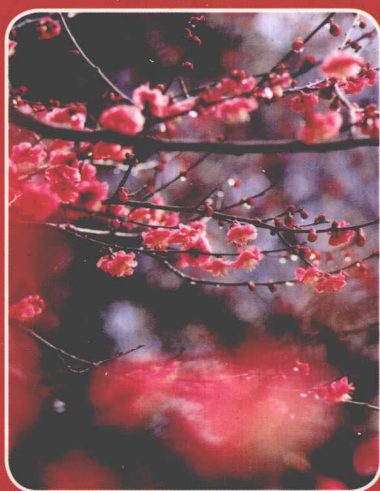
21st Century University Planned Textbooks of Computer Science

人工智能

Artificial Intelligence

王万森 编著

- 理论、方法、技术一体
- 认知、动手、创新启迪
- 学习、研究、教学皆宜



精品系列



人民邮电出版社
POSTS & TELECOM PRESS

21世纪高等学校计算机规划教材

21st Century University Planned Textbooks of Computer Science

人工智能

Artificial Intelligence

王万森 编著



精品系列

人民邮电出版社

北京

图书在版编目 (C I P) 数据

人工智能 / 王万森编著. — 北京 : 人民邮电出版社, 2011. 3
21世纪高等学校计算机规划教材
ISBN 978-7-115-23979-2

I. ①人… II. ①王… III. ①人工智能—高等学校—教材 IV. ①TP18

中国版本图书馆CIP数据核字(2011)第002764号

内 容 提 要

本书是作者在多年人工智能教学实践和多本人工智能教材编写的基础上,结合国内外人工智能领域的最新进展和我国人工智能教学改革的现状,编写的一本教材。全书共10章,除第1章人工智能概述外,其余内容可划分为四大部分。第一部分由第2章~第4章组成,包括知识表示、确定性推理和搜索策略;第二部分由第5章、第6章组成,包括计算智能和不确定性推理;第三部分由第7章、第8章组成,包括机器学习和自然语言理解;第四部分由第9章、第10章组成,包括分布智能和新型专家系统。此外,在附录中还给出了人工智能实验说明。

本书可作为高等院校计算机、自动化、智能、通信、电子信息、信息管理及其他相关学科专业的本科生或研究生教材,也可供从事相关领域研究、开发和应用的科技工作者参考。

21 世纪高等学校计算机规划教材

人 工 智 能

-
- ◆ 编 著 王万森
责任编辑 武恩玉
 - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街14号
邮编 100061 电子函件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
大厂聚鑫印刷有限责任公司印刷
 - ◆ 开本: 787×1092 1/16
印张: 16.5 2011年3月第1版
字数: 433千字 2011年3月河北第1次印刷

ISBN 978-7-115-23979-2

定价: 29.50 元

读者服务热线: (010)67170985 印装质量热线: (010)67129223
反盗版热线: (010)67171154

前言

智能是人类生命体的精髓，用机器模拟、延伸和扩展人的智能，是人类数千年来的一种梦想。今天，人工智能用它 50 多年的坎坷与辉煌，使人们看到了实现这一梦想的希望。

回顾人工智能的发展历程，从游戏世界到现实世界，从理论探讨到实际应用，从符号模拟到结构模拟和行为模拟，从个体智能到群体智能和社会智能，从机器智能到自然智能与机器智能的集成，从孤立的智能研究到智能与情感的交互研究，从智能产品到智能产业和智能社会，人工智能在飞速发展，其新理论、新方法、新技术、新系统、新应用层出不穷。

作为一本教材，如何在有限的篇幅内全面、系统、科学、合理地讨论人工智能，是一个很值得深思的问题。从第一次参加王永庆主编的《人工智能原理·方法·应用》的编写算起，这已是笔者主编和参编的第 4 本人工智能教材。这本教材与笔者三年前出版的《人工智能原理及其应用》（第 2 版）相比，此次编写的基本思路是：减少理论叙述及相关证明，增加方法讨论及相关应用，压缩传统方法及陈旧技术，加强现代方法及新兴技术，使得本教材能更好地体现其基础性和应用性。基于这一思想，笔者采取了如下措施。

（1）在不确定性推理中，删除了证据理论，增加了基于贝叶斯网络的概率推理，并对其他不确定性推理方法的讨论过程做了进一步精炼。

（2）在计算智能中，新增了粗糙集的概念、理论和方法，丰富了计算智能方面的相关内容。

（3）在搜索策略中，删除了一般图搜索过程，对各种搜索方法的讨论均基于搜索树进行，并把讨论的重点放在启发式搜索方面。至于盲目搜索，删掉了深度优先策略，仅以宽度优先策略给予简单说明。

（4）在确定性推理中，删掉了基于规则的演绎推理，并将原确定性推理中关于推理方向等方面的内容与知识表示中关于产生式系统的内容合并，作为一种确定性推理方法在第 3 章单独讨论。

（5）在知识表示中，删掉了过程表示方法和语义网络知识表示法中逻辑关系的表示。

（6）在机器学习中，删掉了解释学习，增加了学习的心理学观点。

（7）对人工智能的概述，更准确地反映了人工智能的概貌和最新进展。

（8）对附录中的人工智能实验，在笔者实际开设过的两届人工智能实验课的基础上，给出了更加详细、准确的实验说明。

（9）对其他方面的内容和结构，也做了一定的修改、更新和重构。

全书共 10 章，各章内容如下。

第 1 章为人工智能概述，包括人工智能的基本概念、基本内容、历史回顾、学派观点、研究应用领域及现状与思考等，可以给出一个更加清晰的人工智能概貌。

第2章~第4章为确定性人工智能的三大基本技术,包括第2章的知识表示方法、第3章的确定性推理和第4章的搜索策略。本教材以兼顾基础和强调应用并重的原则,对这部分内容进行了精心优化和取舍。

第5章为计算智能,包括神经计算、进化计算、模糊计算和粗糙集,这些内容都是当今人工智能研究和应用的热点。

第6章为不确定性推理,这些内容也都是当今不确定性推理方面的主要方法和技术。

第7章为机器学习,包括符号学习和连接学习,这些学习方法都是机器学习中最基本的学习方法。

第8章为自然语言理解,概括性地讨论了书面语言的理解方法。

第9章和第10章分别为分布智能和新型专家系统。重点讨论了基于多 Agent 技术的分布智能和各种新型专家系统,这些都是人工智能面向实际应用的成功典范。

本书还为任课教师提供电子课件和部分习题解答及实验指导,有需要者请登录人民邮电出版社教学服务与资源网(<http://www.ptpedu.com.cn>)免费下载。

本书的编写,也吸取了众多国内外同行在其报告、论文、专著和教材中的精华。在此,谨向这些专家和作者表示感谢。

人工智能是一门正在快速发展的新兴学科,其研究和应用领域十分广阔,对其不少方面,作者还缺乏深入研究,再加上水平有限、时间仓促,书中难免存在一些缺点和错误,恳请各位专家和读者不吝指教。

王万森

目 录

第 1 章 人工智能概述1	2.2 谓词逻辑表示法25
1.1 人工智能的基本概念.....1	2.2.1 谓词逻辑表示的逻辑学基础.....25
1.1.1 智能的概念.....1	2.2.2 谓词逻辑表示的方法.....27
1.1.2 人工智能的概念.....2	2.2.3 谓词逻辑表示的经典例子.....28
1.1.3 人工智能的研究目标.....4	2.2.4 谓词逻辑表示的特性.....30
1.2 人工智能的基本内容.....5	2.3 产生式表示法31
1.2.1 与脑科学和认知科学的交叉研究.....5	2.3.1 产生式表示的基本方法.....31
1.2.2 智能模拟的方法和技术研究.....6	2.3.2 产生式表示的简单例子.....32
1.3 人工智能的历史回顾.....7	2.3.3 产生式表示的特性.....32
1.3.1 孕育期.....7	2.4 语义网络表示法33
1.3.2 形成期.....7	2.4.1 语义网络的基本概念.....33
1.3.3 知识应用期.....8	2.4.2 事物和概念的表示.....35
1.3.4 从学派分立走向综合.....9	2.4.3 情况和动作的表示.....37
1.3.5 智能科学技术学科的兴起.....9	2.4.4 语义网络的基本过程.....38
1.4 人工智能研究中的不同学派.....10	2.4.5 语义网络表示法的特征.....38
1.4.1 符号主义学派.....10	2.5 框架表示法39
1.4.2 联结主义学派.....10	2.5.1 框架理论.....39
1.4.3 行为主义.....11	2.5.2 框架结构和框架表示.....39
1.5 人工智能的研究应用领域.....11	2.5.3 框架系统.....43
1.5.1 机器思维.....12	2.5.4 框架系统的基本过程.....43
1.5.2 机器感知.....13	2.5.5 框架表示法的特性.....44
1.5.3 机器行为.....14	2.6 面向对象表示法45
1.5.4 机器学习.....15	2.6.1 面向对象的基本概念.....45
1.5.5 计算智能.....16	2.6.2 知识的面向对象表示.....46
1.5.6 分布智能.....17	2.6.3 面向对象表示与框架表示的 区别.....46
1.5.7 智能系统.....18	习题.....46
1.5.8 人工心理与人工情感.....18	第 3 章 确定性推理48
1.5.9 人工智能的典型应用.....19	3.1 推理概述.....48
1.6 人工智能的现状与思考.....20	3.1.1 推理的概念.....48
习题.....22	3.1.2 推理方法及其分类.....49
第 2 章 确定性知识表示23	3.1.3 推理控制策略及其分类.....50
2.1 知识表示的基本概念.....23	3.2 产生式系统51
2.1.1 知识的概念.....23	3.2.1 产生式系统的基本结构.....51
2.1.2 知识表示和知识表示方法的概念.....24	3.2.2 产生式系统的推理过程.....51

3.2.3 产生式系统的示例	55	第 5 章 计算智能	106
3.3 自然演绎推理	56	5.1 概述	106
3.3.1 自然演绎推理的逻辑基础	57	5.1.1 什么是计算智能	106
3.3.2 自然演绎推理方法	59	5.1.2 计算智能的产生与发展	107
3.4 归结演绎推理	60	5.1.3 计算智能与人工智能的关系	107
3.4.1 归结演绎推理的逻辑基础	61	5.2 神经计算	108
3.4.2 子句集及其应用	61	5.2.1 神经计算基础	108
3.4.3 鲁宾逊归结原理	65	5.2.2 人工神经网络的互连结构	111
3.4.4 归结演绎推理的方法	68	5.2.3 人工神经网络的典型模型	113
3.4.5 归结演绎推理的归结策略	71	5.3 进化计算	117
3.4.6 用归结反演求取问题的答案	75	5.3.1 进化计算概述	118
习题	76	5.3.2 遗传算法	121
第 4 章 搜索策略	79	5.3.3 遗传算法应用简例	129
4.1 搜索概述	79	5.4 模糊计算	132
4.1.1 搜索的含义	79	5.4.1 模糊集及其运算	132
4.1.2 状态空间法	79	5.4.2 模糊关系及其运算	135
4.1.3 问题归约法	83	5.5 粗糙集	137
4.2 状态空间的盲目搜索	85	5.5.1 粗糙集概述	137
4.2.1 广度优先和深度优先搜索	86	5.5.2 粗糙集的基本理论	137
4.2.2 代价树搜索	87	5.5.3 决策表的约简	139
4.3 状态空间的启发式搜索	89	习题	144
4.3.1 启发性信息和估价函数	89	第 6 章 不确定性推理	146
4.3.2 A 算法	89	6.1 不确定性推理的基本概念	146
4.3.3 A* 算法	91	6.1.1 不确定性推理的含义	146
4.3.4 A* 算法的特性	91	6.1.2 不确定性推理的基本问题	147
4.3.5 A* 算法应用举例	94	6.1.3 不确定性推理的类型	148
4.4 与/或树的盲目搜索	95	6.2 可信度方法	148
4.4.1 与/或树的一般搜索	95	6.2.1 可信度的概念	149
4.4.2 与/或树的广度优先和深度优先 搜索	96	6.2.2 可信度推理模型	149
4.5 与/或树的启发式搜索	97	6.2.3 可信度推理的例子	153
4.5.1 解树的代价与希望树	97	6.3 主观 Bayes 方法	154
4.5.2 与/或树的启发式搜索过程	98	6.3.1 主观 Bayes 方法的概率论基础	154
4.6 博弈树的启发式搜索	100	6.3.2 主观 Bayes 方法的推理模型	155
4.6.1 概述	100	6.3.3 主观 Bayes 推理的例子	159
4.6.2 极大极小过程	101	6.4 模糊推理	161
4.6.3 α - β 剪枝	102	6.4.1 模糊知识表示	161
习题	103	6.4.2 模糊概念的匹配	163

6.4.3 模糊推理方法	164	8.4.1 语义文法	207
6.5 概率推理	168	8.4.2 格文法	208
6.5.1 贝叶斯网络的概念及理论	168	8.5 自然语言理解系统的层次模型	209
6.5.2 贝叶斯网络推理的概念和类型	171	习题	210
6.5.3 贝叶斯网络的精确推理	171		
习题	173	第 9 章 分布智能	212
第 7 章 机器学习	175	9.1 分布智能概述	212
7.1 机器学习概述	175	9.1.1 分布智能的概念	212
7.1.1 机器学习的概念及其发展过程	175	9.1.2 分布式问题求解	213
7.1.2 学习系统的概念及模型	177	9.1.3 多 Agent 系统	213
7.1.3 机器学习的类型	179	9.2 Agent 的结构	216
7.2 记忆学习	179	9.2.1 Agent 的机理	216
7.3 示例学习	180	9.2.2 反应 Agent 的结构	217
7.3.1 示例学习的类型	180	9.2.3 认知 Agent 的结构	217
7.3.2 示例学习的模型	181	9.2.4 混合 Agent 的结构	218
7.3.3 示例学习的归纳方法	182	9.3 Agent 通信	218
7.4 决策树学习	184	9.3.1 Agent 通信的基本问题	218
7.4.1 决策树的概念	184	9.3.2 Agent 通信方式	219
7.4.2 ID3 算法	185	9.3.3 Agent 通信语言 KQML	220
7.5 联结学习	188	9.4 多 Agent 合作	223
7.5.1 联结学习的心理学基础	188	9.4.1 Agent 的协调	223
7.5.2 联结学习的学习规则	188	9.4.2 Agent 的协作	224
7.5.3 感知器学习	190	9.4.3 Agent 的协商	226
7.5.4 BP 网络学习	192	9.4.4 多 Agent 应用示例	227
7.5.5 Hopfield 网络学习	195	9.5 移动 Agent	228
习题	197	9.5.1 移动 Agent 系统的一般结构	228
		9.5.2 移动 Agent 的实现技术及应用	229
第 8 章 自然语言理解	198	习题	231
8.1 自然语言理解的基本概念	198	第 10 章 新型专家系统	232
8.1.1 自然语言的含义与组成	198	10.1 专家系统概述	232
8.1.2 自然语言理解的含义及任务	199	10.1.1 专家系统的产生与发展	232
8.1.3 自然语言理解的发展过程	199	10.1.2 新型专家系统的特征与类型	233
8.1.4 自然语言理解的层次	200	10.2 专家系统基础	234
8.2 词法分析	201	10.2.1 专家系统的基本结构	235
8.3 句法分析	201	10.2.2 基于规则的专家系统	236
8.3.1 句法规则的表示方法	202	10.2.3 基于框架的专家系统	237
8.3.2 自顶向下与自底向上分析	204	10.3 典型的新型专家系统	238
8.3.3 扩充转移网络分析	205	10.3.1 模糊专家系统	238
8.4 语义分析	207	10.3.2 神经网络专家系统	239

10.3.3 基于 Web 的专家系统	241	10.4.3 开发工具与环境	246
10.3.4 分布式专家系统	242	习题	248
10.3.5 协同式专家系统	243	附录 A 人工智能实验	250
10.4 专家系统的开发	244	参考文献	255
10.4.1 开发步骤	244		
10.4.2 知识获取	245		

第 1 章

人工智能概述

人工智能（Artificial Intelligence, AI）作为一门引领未来的科学技术，目前已渗透到人类社会的各个方面，并逐步改变着人们的学习、工作和生活方式。“智能”这个具有无限潜力的学科，正在以其无穷的魅力推动着现代科学技术的发展和人类文明的进步。人工智能前景诱人，同时也任重而道远。本章主要是先就人工智能做一些简要介绍，包括基本概念、基本内容、发展历史、学派观点、研究应用领域等，以建立起对人工智能的初步认识。

1.1 人工智能的基本概念

认识智能机理，建造智能实体，用人工的方法去模拟和实现人类智能，这应该是人工智能的一个最基本的解释。从这一意义上讲，人工智能的定义应依赖于对智能的定义。因此，本节拟从智能的概念入手，讨论人工智能的一些基本概念。

1.1.1 智能的概念

智能是对人类智能或自然智能的简称，其确切定义还有待于脑科学和认知科学对人脑奥秘的彻底揭示。下面，我们就从脑科学和认知科学的角度出发，谈几点对智能的认识。

1. 认识智能的不同观点

人类在认识智能的过程中提出了许多不同的观点，其中最具代表性的观点有以下三种。

（1）智能来源于思维活动。这种观点被称为思维理论。它强调思维的重要性，认为智能的核心是思维，人的一切智慧或智力都来自于大脑的思维活动，人的一切知识都是思维的产物，因而通过对思维规律与思维方法的研究，可望揭示智能的本质。

（2）智能取决于可运用的知识。这种观点被称为知识阈值理论。它把智能定义为：智能就是在巨大的搜索空间中迅速找到一个满意解的能力。知识阈值理论着重强调知识对智能的重要意义和作用，认为智能行为取决于知识的数量及其可运用的程度，一个系统所具有的可运用知识越多，其智能就会越高。

（3）智能可由逐步进化来实现。这种观点被称为进化理论。它是美国麻省理工学院（MIT）的布鲁克斯（R. A. Brooks）教授在对人造机器虫进行研究的基础上提出来的。他认为智能取决于感知和行为，取决于对外界复杂环境的适应，智能不需要知识、不需要表示、不需要推理，智能可以由逐步进化来实现。

上述三种观点对智能的认识角度不同，有些看起来好像是相互对立的，但如果把它们放到智

能的层次结构中去考虑，又是统一的。

2. 智能的层次结构

根据脑科学的研究，人类智能总体上可分为高、中、低三个层次，不同层次智能的活动由不同的神经系统来完成。其中，高层智能以大脑皮层为主，大脑皮层又称为抑制中枢，主要完成记忆、思维等活动；中层智能以丘脑为主，也称为感觉中枢，主要完成感知活动；低层智能以小脑、脊髓为主，主要完成动作反应活动。

可见，把智能的不同观点和智能的层次结构联系起来看，思维理论和知识阈值理论对应于高层智能，进化理论对应于中层智能和低层智能。

3. 智能所包含的能力

认知科学的研究认为，智能是一种综合能力。具体地说，它包含的主要能力如下。

(1) 感知能力。感知能力是指人们通过感觉器官感知外部世界的能力。它是人类最基本的生理、心理现象，也是人类获取外界信息的基本途径。人类对感知到的外界信息，通常有两种不同的处理方式。一种是对简单或紧急情况，可不经大脑思索，直接由低层智能做出反应。另一种是对复杂情况，一定要经过大脑的思维，然后才能做出反应。

(2) 记忆与思维能力。记忆与思维是人脑最重要的功能，也是人类智能最主要的表现形式。记忆是对感知到的外界信息或由思维产生的内部知识的存储过程。思维是对所存储的信息或知识的本质属性、内部规律等的认识过程。人类基本的思维方式有抽象思维、形象思维和灵感思维。

(3) 学习和自适应能力。学习是一个具有特定目的的知识获取过程。学习和自适应是人类的一种本能，一个人只有通过学习，才能增加知识、提高能力、适应环境。尽管不同的人在学习方法、学习效果等方面有较大差异，但学习却是每个人都具有的一种基本能力。

(4) 决策和行为能力。行为能力是指人们对感知到的外界信息做出动作反应的能力。引起动作反应的信息可以由感知直接获得的外部信息，也可以是经思维加工后的内部信息。完成动作反应的过程，一般通过脊髓来控制，并由语言、表情、体姿等来实现。

综上所述，我们可以给出智能的一个一般解释：智能是人类在认识客观世界的过程中，由思维过程和脑力活动所表现出来的综合能力。

1.1.2 人工智能的概念

什么是人工智能？顾名思义是人造的智能。如果智能能够严格定义，那么人工智能也就容易定义了。但从以上分析可知，其前提并不成立，因此人工智能还无法形式化定义。尽管如此，在人工智能诞生的五十多年里，人们还是从不同方面给出了一些不同的解释。其中，较具代表性的是从类人、理性、思维、行为这四个方面给出的定义方法。

1. 类人行为方法

类人行为方法也称为图灵测试方法，它是一种基于人类自身的智能去定义一个机器或系统是否具有智能的方法。其典型代表是库兹韦勒（Kurzweil）于1990年提出的定义：人工智能是一种创建机器的技艺，这种机器能够执行那些需要人的智能才能完成的功能。

早在1950年，人工智能还没有作为一门学科正式出现之前，英国数学家图灵（A. M. Turing, 1912~1954年）就在他发表的一篇文章《Computing Machinery and Intelligence（计算机器与智能）》中提出了“机器能思维”的观点，并设计了一个很著名的测试机器智能的实验，称为“图灵测试”或“图灵实验”。

图灵实验可描述如下：该实验的参加者由一位测试主持人和两个被测试对象组成。其中，两个被测试对象中的一个是人，另一个是机器。测试规则为：测试主持人和每个被测试对象分别位于彼此看不见的房间中，相互之间只能通过计算机终端进行会话。测试开始后，由测试主持人向被测试对象提出各种具有智能性的问题，但不能询问测试者的物理特征。被测试对象在回答问题时，都应尽量使测试者相信自己是“人”，而另一位是“机器”。在这个前提下，要求测试主持人区分这两个被测试对象中哪个是人，哪个是机器。如果测试主持人不能分辨出被测试对象中哪个是人，哪个是机器，则认为该机器具有了智能。

按照以上实验要求，一台计算机要能够通过测试，至少应该具有以下能力。

- (1) 自然语言处理。实现用自然语言与计算机的交流。
- (2) 知识表示。存储它所知道的或听到的知识或信息。
- (3) 自动推理。运用存储的知识或信息来回答问题，并提取新的结论。
- (4) 机器学习。能适应新的环境，并能自我获取新的知识。

如果还需要测试被测对象利用视频信号的感知能力和传递接受物体的行为能力，即所谓的完全图灵测试，则计算机还应该具有如下能力。

- (5) 计算机视觉。可以感知物体。
- (6) 机器人技术。可以操纵和移动物体。

这6个领域构成了人工智能的大部分内容。图灵测试的主要贡献在于它给出了一个客观的智能概念。对图灵测试，也有人提出了质疑：认为该测试仅反映了结果的比较，既没有涉及思维的过程，也没有明确参加实验的人是小孩还是具有良好素质的成年人。并且，也很少有人花很大精力去尝试通过图灵测试。尽管如此，图灵测试对人工智能学科发展所产生的影响却是十分深远的。

2. 类人思维方法

类人思维方法也称为认知模型方法，它是一种基于人类思维工作原理的可检测理论定义智能的方法。其典型代表是贝尔曼（Bellman）于1978年提出的定义：人工智能是那些与人的思维、决策、问题求解和学习等有关的活动的自动化。

认知科学是一门研究人类智能如何由物质产生和人脑怎样进行思维信息处理的科学。它把来自人工智能的计算机模型和来自心理学的实验技术结合起来，试图创立一种精确而且可检测的人脑思维过程的工作模型。要确定人脑思维的真实过程，可以有两种方法，一种是通过内省去捕捉自身的思维过程，另一种是通过心理实验。一旦我们能够得到关于思维的足够精确的理论，就可以把这种理论表示为计算机程序。如果该程序的输入/输出和实时行为能够与人类一致，那就证明该程序的某些机制可能是按照人脑思维模式运行的。

这方面的典型例子是艾伦·纽厄尔（Allen Newell）和赫伯特·西蒙（Herbert Simon）等人于1960年研制的通用问题求解（General Problem Solving, GPS）程序。

3. 理性思维方法

理性思维方法也称为思维法则方法，它是一种基于逻辑推理定义智能的方法。其典型代表是查尼艾克（E.Charniak）和麦克德莫特（D.McDermott）于1985年提出的定义：人工智能是通过计算模型的使用来进行心智能力研究的。

这里的计算模型主要是指能“正确思维”的逻辑学模型。古希腊哲学家亚里士多德（Aristotle）是首先严格定义“正确思维”的人之一，他将其定义为“不能辩驳的推理过程”。他的三段论推理方法，就给出了一种能正确思维的推理模式，即当已知前提正确时，总能推

出正确的结论。

理性思维方法正是人工智能领域中所谓的逻辑主义观点，他们希望通过编制逻辑程序来建造智能系统。但是，这种方法存在两个主要问题：第一，非形式的知识用形式的逻辑符号表示不易实现，尤其是对不确定的知识；第二，原则上可以解决的问题与实际解决问题之间存在较大差异，需要考虑推理过程的控制。

4. 理性行为方法

理性行为方法也称为理性智能体方法，它是一种基于智能体定义智能的方法。其典型代表是尼尔森（N.J.Nilsson）于1998年提出的定义：人工智能关心的是人工制品中的智能行为。这里的人工制品主要是指能够感知环境、适应变化、自主操作、执行动作的理性智能体（Agent）。按照这种方法，可以认为人工智能就是研究和建造理性智能体。

理性行为方法与理性思维方法相比，二者既有区别，又有联系。首先，理性行为和理性思维强调的重点不同。理性思维方法强调的是正确思维，而理性行为方法强调的则是理性行动。所谓行动上的理性，指的是已知某些信念，执行某些动作以达到某个目标。其次，理性行为可以依据理性思维进行。例如，对一些能够通过理性思维做出正确结论的事情，实现理性行为的途径往往是先通过逻辑推理得出该行为能达到的目标和结论，然后再付诸实施。再其次，理性行为不一定要依据理性思维进行。例如，对有些事情，即使理性思维无法证明哪些行为是正确的，而其他行为是错误的，理性行为也必须有所行动。

对以上4种定义方法，1和2属于类人方法，3和4属于理性方法；2和3强调思维，1和4强调行为。从历史上看，这4种方法都有人做了很多工作，当然在以人为中心的方法和以理性为中心的方法之间也存在着一定的争议。例如，以人为中心的方法是一种经验科学，它需要涉及很多假设和实验证实；而以理性为中心的方法则涉及到把数学与工程相结合。对此，我们不做过多评论，原因是人工智能还处在快速发展中。

综上所述，我们可以从“能力”和“学科”两个方面给出人工智能的如下解释。从能力的角度看，人工智能是指用人工的方法在机器（计算机）上实现的智能。从学科的角度看，人工智能是一门研究如何构造智能机器或智能系统，使它能够模拟、延伸和扩展人类智能的学科。

1.1.3 人工智能的研究目标

关于人工智能的研究目标，目前也还没有一个统一的说法。斯洛曼（A. Sloman）于1978年给出了人工智能的以下三个主要目标。

- （1）对智能行为有效解释的理论分析。
- （2）解释人类智能。
- （3）构造智能的人工制品。

要实现斯洛曼的这些目标，需要同时开展对智能机理和智能构造技术的研究。对图灵所期望的那种智能机器，尽管它没有提到思维过程，但要真正实现这种智能机器，却同样离不开对智能机理的研究。因此，揭示人类智能的根本机理，用智能机器去模拟、延伸和扩展人类智能应该是人工智能研究的根本目标，或者叫远期目标。

人工智能的远期目标涉及脑科学、认知科学、计算机科学、系统科学、控制论及微电子等多种学科，并有赖于这些学科的共同发展。但从目前这些学科的现状来看，实现人工智能的远期目标还需要一个较长的时期。

在这种情况下,人工智能的近期目标是研究如何使现有的计算机更聪明,即使它能够运用知识去处理问题,能够模拟人类的智能行为,如推理、思考、分析、决策、预测、理解、规划、设计和学习等。为了实现这一目标,人们需要根据现有计算机的特点,研究实现智能的有关理论、方法和技术,建立相应的智能系统。

实际上,人工智能的远期目标与近期目标是相互依存的。远期目标为近期目标指明了方向,而近期目标为远期目标奠定了理论和技术基础。同时,近期目标和远期目标之间并无严格界限,近期目标会随人工智能研究的发展而变化,并最终达到远期目标。

1.2 人工智能的基本内容

关于人工智能的研究内容,不同学派、不同研究领域及人工智能发展的不同时期,对其都有着一些不同的看法。下面主要从智能机理和对智能的模拟这两个方面进行讨论。

1.2.1 与脑科学和认知科学的交叉研究

脑科学和认知科学是人工智能的重要理论基础,对人工智能的研究具有重要的指导和启迪作用,因此人工智能应该重视与脑科学和认知科学的交叉研究。

1. 脑科学

脑科学亦称神经科学,是一门研究脑与心智现象及规律的科学,其主要目标就是要揭示脑功能的本质,认识脑与智能的规律,保护脑和创造脑。关于脑科学或神经科学的定义,美国神经科学学会给出的解释是:神经科学是为了了解神经系统内分子水平、细胞水平及细胞间的变化过程,以及这些过程在中枢的功能、控制系统内的整合作用所进行的研究。

在脑科学中,脑的含义可从狭义和广义两个方面来理解:从狭义方面,脑是指中枢神经系统,有时特指大脑;从广义方面,脑可泛指整个神经系统。由于人们一般是从最广泛的交叉学科的角度来理解脑科学的,因此其学科范畴涵盖了所有与认识脑和神经系统有关的研究。又由于人工智能的主要任务是要用机器来模拟人脑,因此脑科学的研究应该是人工智能研究的必要前提。

人脑被认为是自然界中最复杂、最高级的智能系统。这种复杂性主要表现在人脑是由巨量神经元相互之间其突触的广泛并行互联所形成的一个巨复杂的系统。现代脑科学的基本问题主要包括:揭示神经元之间的联结形式,奠定行为的脑机制的结构基础;阐明神经活动的基本过程,说明在分子、细胞到行为等不同层次上神经信号的产生、传递、调制等基本过程;鉴别神经元的特殊细胞生物学特性;认识实现各种功能的神经回路基础;解释脑的高级功能机制等。

脑科学研究的任何进展,都将会对人工智能的研究起到积极的推动作用,因此脑科学是人工智能的一个重要基础,人工智能应该加强与脑科学的交叉研究,以及人类智能与机器智能的集成研究。

2. 认知科学

认知科学(或称思维科学),是一门研究人类智能如何由物质产生和人脑怎样进行思维信息处理的科学,其主要目的是要说明和解释人类在完成认知活动时是如何进行信息加工的。所谓认知(cognition)可一般地认为是和情感、动机、意志相对应的理智或认识过程,或者说是为了一定目

的，在一定的心理结构中进行的信息加工过程。美国心理学家霍斯顿（Houston）等人曾把对认知的看法归纳为以下 5 种主要类型。

（1）认知是信息的处理过程。

（2）认知是心理上的符号运算。

（3）认知是问题求解。

（4）认知是思维。

（5）认知是一组相关的活动，如知觉、记忆、思维、判断、推理、问题求解、学习、想象、概念形成及语言使用等。

人类的认知过程是一种非常复杂的行为，人们对其研究形成了认知科学。认知科学也称思维科学，是一门研究人类感知和思维信息处理过程的学科，它包括从感觉的输入到复杂问题的求解，从人类个体智能到人类社会智能的活动，以及人类智能和机器智能的性质。其主要研究目的就是说明和解释人类在完成认知活动时是如何进行信息加工的。

认知科学也是人工智能的重要理论基础，对人工智能发展起着根本性的作用。认知科学涉及的问题非常广泛，除了像霍斯顿提出的知觉、语言、学习、记忆、思维、问题求解、创造、注意、想象等相关联的活动外，还会受到环境、社会、文化背景等方面的影响。从认知观点看，人工智能不能仅限于逻辑思维的研究，还必须深入开展对形象思维和灵感思维的研究。只有这样，才能使人工智能具有更坚实的理论基础，才能为智能计算机系统的研制提供更新的思想，创造更新的途径。

1.2.2 智能模拟的方法和技术研究

要用机器模拟人类智能，就必须开展对机器感知、思维、学习、行为的研究，以及智能系统和智能机器建造技术的研究。

所谓机器感知，就是要让计算机具有类似于人的感知能力，如视觉、听觉、触觉等。在机器感知中，目前研究较多和较为成功的是机器视觉（或叫计算机视觉）和机器听觉（或叫计算机听觉）。这些研究目前已在人工智能中形成了一些专门的研究领域，如计算机视觉、模式识别、自然语言理解等。

所谓机器思维，就是让计算机能够对感知到的外界信息和自己产生的内部信息进行思维性加工。由于人类智能主要来自于大脑的思维活动，因此机器思维也应该是机器智能的重要组成部分。为了实现机器的思维功能，需要在知识的表示、组织及推理方法，各种启发式搜索及控制策略，神经网络、思维机理等方面进行研究。

所谓机器学习就是让计算机能够像人那样自动地获取新知识，并在实践中不断地完善自我和增强能力。机器学习是机器具有智能的重要标志，也是人工智能研究的核心问题之一。目前，人们根据对人类自身学习的已有认识，已经研究出了不少机器学习方法，如记忆学习、归纳学习、解释学习、发现学习、神经学习和遗传学习等。

所谓机器行为就是让计算机能够具有像人那样的行动和表达能力，如走、跑、拿、说、唱、写、画等。如果把机器感知看成智能系统的输入部分，那么，机器行为可看成智能系统的输出部分。例如，智能控制、智能制造、智能调度、智能机器人等。

建立智能系统或构造智能机器既是人工智能面向实际应用的关键，也是人工智能近期目标和远期目标的一种必然要求，因此需要开展对智能系统和智能机器建造技术的研究。它包括系统模型、构造技术、构造工具及语言环境等方面的研究。

1.3 人工智能的历史回顾

人工智能诞生 50 多年来,走过了一条坎坷和曲折的发展道路。回顾历史,可以按照人工智能在不同时期的主要特征,将其产生与发展过程分为以下五个阶段。

1.3.1 孕育期

自远古以来,人类就有着用机器代替人进行脑力劳动的幻想。早在公元前 900 多年,我国就有歌舞机器人流传的记载。到公元前 850 年,古希腊也有了制造机器人帮助人们劳动的神话传说。此后,在世界上的许多国家和地区也都出现了类似的民间传说或神话故事。为追求和实现人类的这一美好愿望,很多科学家都为之付出了艰辛的劳动和不懈的努力。人工智能可以在顷刻间诞生,而孕育这个学科却需要经历一个相当漫长的历史过程。

从古希腊伟大的哲学家亚里士多德(Aristotle, 公元前 384 年~公元前 322 年)创立的演绎法,到德国数学和哲学家莱布尼茨(G.W.Leibnitz, 1646 年~1716 年)奠定的数理逻辑的基础;再从英国数学家图灵(A.M.Turing, 1912 年~1954 年)1936 年创立图灵机模型,到美国数学家、电子数字计算机的先驱莫克利(J.W.Mauchly, 1907 年~1980 年)等人 1946 年研制成功的世界上第一台通用电子计算机等,这些都为人工智能的诞生奠定了重要的思想理论和物质技术基础。

此外,1943 年,美国神经生理学家麦卡洛克(W.McCulloch)和皮茨(W.Pitts)一起研制出了世界上第一个人工神经网络模型(MP 模型),开创了以仿生学观点和结构化方法模拟人类智能的途径;1948 年,美国著名数学家威纳(N.Wiener, 1874 年~1956 年)创立了控制论,为以行为模拟观点研究人工智能奠定了理论和技术基础;1950 年,图灵又发表题为《计算机能思维吗?》的著名论文,明确提出了“机器能思维”的观点。至此,人工智能的雏形已初步形成,人工智能的诞生条件也已基本具备。通常,人们把这一时期称为人工智能的孕育期。

1.3.2 形成期

人工智能诞生于一次历史性的聚会。为使计算机变得更“聪明”,或者说使计算机具有智能,1956 年夏季,当时在达特茅斯(Dartmouth)大学的年轻数学家、计算机专家麦卡锡(J.McCarthy, 后为 MIT 教授)和他的三位朋友,哈佛大学数学家、神经学家明斯基(M.L.Minsky, 后为 MIT 教授),IBM 公司信息中心负责人洛切斯特(N.Lochester),贝尔实验室信息部数学研究员香农(C.E.Shannon)共同发起,并邀请 IBM 公司的莫尔(T.More)和塞缪尔(A.L.Samuel),MIT 的塞尔弗里奇(O.Selfridge)和索罗蒙夫(R.Solomonff),以及兰德(RAND)公司和卡内基(Carnegie)工科大学的纽厄尔(A.Newell)和西蒙(H.A.Simon)共 10 人,在美国达特茅斯大学举行了一个为期两个月的夏季学术研讨会。这 10 位来自美国数学、神经学、心理学、信息科学和计算机科学方面的杰出年轻科学家,在一起共同学习和探讨了用机器模拟人类智能的有关问题,并由麦卡锡提议正式采用了“人工智能(Artificial Intelligence AI)”这一术语。从而,一个以研究如何用机器来模拟人类智能的新兴学科——人工智能诞生了。

在人工智能诞生以后的十多年中,很快就在定理证明、问题求解、博弈等领域取得了重大突破。通常,人们把 1956 年~1970 年这段时间称为人工智能的形成期,也有人称为高潮时期。在这一

时期所取得的主要研究成果如下。

1956 年, 塞缪尔研制成功了具有自学习、自组织和自适应能力的西洋跳棋程序, 该程序于 1959 年击败了塞缪尔本人, 1962 年又击败了一个州的冠军。1957 年, 纽厄尔、肖 (J.Shaw) 和西蒙等人的心理学小组研制了一个称为逻辑理论机 (Logic Theory machine, LT) 的数学定理证明程序, 开创了用计算机研究人类思维活动规律的工作。1965 年, 鲁滨逊 (J.A.Robinson) 提出了归结 (消解) 原理。1968 年, 美国斯坦福大学费根鲍姆 (E.A.Feigenbaum) 领导的研究小组研制成功了化学专家系统 DENDRAL。此外, 在人工神经网络方面, 1957 年, 罗森布拉特 (F.Rosenblatt) 等人研制了感知器 (perceptron), 利用感知器可进行简单的文字、图像、声音识别。

1.3.3 知识应用期

正当人们在为人工智能所取得的成就而高兴的时候, 人工智能却遇到了许多困难, 遭受了很大的挫折。然而, 在困难和挫折面前, 人工智能的先驱者们并没有退缩, 他们在反思中认真总结了人工智能发展过程中的经验教训, 从而又开创了一条以知识为中心、面向应用开发的研究道路, 使人工智能又进入了一条新的发展道路。通常, 人们把从 1971 年到 20 世纪 80 年代末这段时间称为人工智能的知识应用期, 也有人称为低潮时期。

1. 挫折和教训

人工智能在经过形成时期的快速发展之后, 很快就遇到了许多麻烦。例如:

(1) 在博弈方面, 塞缪尔的下棋程序在与世界冠军对弈时, 五局中败了四局。

(2) 在定理证明方面, 发现鲁滨逊归结法的能力有限。当用归结原理证明两个连续函数之和还是连续函数时, 推了 10 万步也没证出结果。

(3) 在机器翻译方面, 原来人们以为只要有一本双解字典和一些语法知识就可以实现两种语言的互译, 但后来发现并不那么简单, 甚至会闹出笑话。例如, 把“心有余而力不足”的英语句子 “The spirit is willing but the flesh is weak” 翻译成俄语, 然后再翻译回来时竟变成了“酒是好的, 肉变质了”, 即英语句子为 “The wine is good but the meat is spoiled”。

(4) 在神经网络方面, 神经生理学研究发现人脑由 $10^{11} \sim 10^{12}$ 个神经元所组成, 在现有技术条件下用机器从结构上模拟人脑是根本不可能的。并且, 明斯基于 1969 年出版的专著《Perceptrons》, 指出了感知器模型存在的严重缺陷, 致使人工神经网络的研究落入低潮。

在其他方面, 人工智能也遇到了这样那样的问题。一些西方国家的人工智能研究经费被削减, 研究机构被解散, 全世界范围内的人工智能研究陷入困境、跌入低谷。

值得庆幸的是, 在这种极其困难的环境下, 仍有一大批人工智能学者不畏艰辛、潜心研究。他们在认真总结前一阶段研究工作的经验教训的同时, 从费根鲍姆以知识为中心开展人工智能研究的观点中找到了新的出路。

2. 以知识为中心的研究

科学的真谛总是先由少数人创造出来的。早在 20 世纪 60 年代中期, 当大多数人工智能学者正热衷于对博弈、定理证明、问题求解等进行研究时, 专家系统这一个重要研究领域也开始悄悄地孕育。正是由于专家系统这棵幼小萌芽的存在, 才使得人工智能能够在后来出现的困难和挫折中很快找到前进的方向, 又迅速地再度兴起。

专家系统 (Expert System, ES) 是一个具有大量专门知识, 并能够利用这些知识去解决特定领域中需要由专家才能解决的那些问题的计算机程序。专家系统实现了人工智能从理论研究走向