

人工智能 基础与应用

樊重俊 主编

清华大学出版社

北 京

内 容 简 介

人工智能(Artificial Intelligence, AI)是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门技术科学。本书以人工智能为主要研究对象,较全面地介绍人工智能的基本原理、常见算法和应用技术。全书共分为12章,主要内容包括绪论、知识与知识表示、自动推理与专家系统、搜索算法与智能计算、机器学习、深度学习、计算机视觉、自然语言与语音处理、智能机器人以及大数据与区块链、Python编程基础等。同时,为了便于读者自学,每章结尾附有小结与习题,便于读者进一步巩固所学知识。

本书全面系统地介绍了人工智能的理论体系,在内容编排上不仅注重基本理论的讲述,而且以发展的眼光设计各章知识点和习题,引导学生去思考人工智能理论知识的应用与实践,使得理论学习更加生动直观,便于培养此类学生对人工智能理论的理解和应用能力。本书面向高等院校管理学院的诸多专业,可作为高等院校经济管理类专业本科生人工智能课程要求重点掌握人工智能的基本理论知识、相关算法的初步实践操作以及Python编程基础的教材,也可作为成人教育和自学考试同名课程的参考教材,还可供从事人工智能领域研究、开发和应用的工程技术人员和工程技术人员阅读参考。

本书配套的电子课件和课后习题参考答案可以到<http://www.tupwk.com.cn/downpage>网站下载,也可以通过扫描前言中的二维码来下载。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

人工智能基础与应用 / 樊重俊 主编. —北京:清华大学出版社, 2020.9

ISBN 978-7-302-55947-4

I. ①人… II. ①樊… III. ①人工智能 IV. ①TP18

中国版本图书馆CIP数据核字(2020)第120429号

责任编辑:胡辰浩

封面设计:高娟妮

版式设计:孔祥峰

责任校对:成凤进

责任印制:丛怀宇

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦A座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者:三河市少明印务有限公司

经 销:全国新华书店

开 本:185mm×260mm 印 张:23.25 字 数:580千字

版 次:2020年8月第1版 印 次:2020年8月第1次印刷

定 价:78.00元

产品编号:087378-01

编写组 ◀

主 编 樊重俊

副主编 施柏州 尹 裴 杨云鹏 熊红林

编 写 刘 臣 关晓飞 鞠晓玲 安艾芝

臧悦悦 朱 玥 樊鸿飞 王雅琼

李璟轩 余 莹 黄 耐 蒋雨桥

徐 佩

序

1956年，一次历史性的聚会标志着人工智能的诞生。来自美国 Dartmouth 学院与 IBM 公司、RAND 公司、Carnegie Mellon 大学的十位学者在会上第一次正式使用了人工智能(Artificial Intelligence, AI)这一术语，开创了人工智能的研究方向。自诞生以来，人工智能的理论和技術日益成熟，应用领域也不断扩大，对人类社会的发展产生了方方面面的影响。人工智能理论中的知识表示和推理机制、问题求解和搜索算法以及计算智能技术等，可以解决从前难以解决的问题；综合应用语法、语义和人工智能的形式知识表示方法，有可能改善知识的自然语言表达形式；潜在的知识、灵感等也能阐述为适用的人工智能形式；专家系统更是深入各行各业，带来巨大的宏观效益……这些无不代表着人工智能正在通过对科学知识的扩展与升华，对人类的思维方式、社会经济、文化生活带来实质性的改变。

人工智能是国家政府、社会公众普遍关注的创新领域，也是创新文化的综合集成体和国计民生发展的重要应用方向。中国在全球人工智能创新竞争中居于前列，人工智能与相关产业的融合应用发展推动人民生活越来越便捷。中国在人工智能领域已有许多独角兽企业峥嵘突起，产业蓬勃。随着国家一系列相关政策的发布，众多重大项目启动实施，创新平台布局日趋完善，学科建设、人才培养、创新创业等各方面蓬勃开展，全社会支持人工智能发展的良好氛围正在形成。面对激烈的科技竞争态势，必须以学科建设为引领，多措并举、齐抓共管，不断提升人工智能教育水平，不断增强对人工智能发展的支撑能力、推动能力、引领能力，从而培养出真正具有问题解决能力的创新型人工智能人才。在大力发展人工智能新技术的同时，必须高度重视相关人才培养与学术研究。人工智能学科是规范性的，也是建设性的，该学科的研究旨在揭示人工智能学科发展面临的难题，同时为人工智能技术的发展提供价值指引，确保人工智能在正确的轨道上前行。

由于人工智能学科发展的多样性、交叉性、综合性、知识跨越性、复杂性、专业性等特点，人工智能教育所需要的知识不可能由个人或少数人全部掌握，如何合理组织学科内容，针对不同人群进行精准施教是高校需要着力研究的重点。人工智能课程知识点较多，涵盖模式识别、机器学习、数据挖掘等众多内容，概念抽象，不易学习。对于管理类专业的人才，大部分都会走向更具体化的管理岗位，具有多学科素养，但这也导致很多学生所学知识杂而不精。学生在基础未夯实的情况下去学习更高层面的知识，给学生学习与老师教学造成了很大困扰。《人

《人工智能基础与应用》一书主要面向经管类专业人才，由相关业内人士编写，内容涉及绪论、知识表示与推理、常用算法、机器学习、神经网络、区块链、Python 编程基础等，内容丰富、层次分明、重点突出，关注学科最新动态，理清了管理类人才人工智能基础的学习思路，帮助读者掌握人工智能知识基础及专业技能，同时培养自主学习和知识探索能力，为管理类人才的人工智能教学培养提供有力借鉴。

陈宏民
于上海交通大学
2020年5月

前言 ◀

人工智能已经成为推动经济社会发展的新引擎。数字经济下,人工智能成为人类认知世界、改造世界的新切入点,人工智能技术逐步成为科技未来发展的趋势,对各行各业产生巨大影响。当前人工智能已经广泛应用于教育、医疗、交通、零售、物流、安防等领域,成为人们生活和工作中不可或缺的应用技术,在未来世界的发展中,人工智能技术已经成为不可或缺的发展要素之一。2018年4月,教育部制定了《高等学校人工智能创新行动计划》,其核心目标之一是推动高校人工智能领域人才培养体系的完善。越来越多的高校参与到人工智能课程建设中来,不断地进行探索与革新,对于培养掌握人工智能技术的管理类人才愈发重视。

作为一门深刻改变世界、有远大发展前途的前沿学科,人工智能覆盖面广、包容性强、应用需求空间巨大的学科特点,有利于更好地培养学生的技术创新思维与能力。《中国新一代人工智能发展报告 2019》于2019年5月在上海发布,报告显示我国在人工智能领域的多个方面取得了快速进步,但随之而来的是市场对人工智能领域专业人才需求的迫切性仍在持续。对于管理类人才来说,通过对知识表示、搜索算法与智能计算、机器学习、深度学习、计算机视觉、自然语言与语音处理、大数据与区块链等核心内容的深入学习,不断将人工智能技术与管理知识结合起来,提高解决实际问题的能力,对学生未来职业生涯的发展有着重要作用。

本书将理论知识与实践应用联系起来,每章除了对人工智能基础理论、算法进行讲解以外,还包括各类算法的实践应用与习题巩固环节。此外,随着大数据与人工智能的兴起,我国不少大学以及世界知名大学均已意识到编程能力对于未来的重要性,因此,本书最后增设对 Python 编程基础的讲解,培养读者采用 Python 解决人工智能相关问题的能力。

全书共 12 章,包括以下内容:第 1 章是绪论,主要介绍人工智能的概念以及发展简史、当前人工智能的发展现状及未来趋势等内容。第 2 章主要介绍知识与知识表示,包含知识与知识表示的相关概念特征、表示方法、知识图谱及其应用等内容。第 3 章主要介绍自动推理与专家系统,包含自动推理的基本知识以及专家系统的概念、结构、设计与实现、应用与发展等内容。第 4 章主要介绍搜索算法与智能计算,包含搜索算法、遗传算法、蚁群算法以及粒子群优化算法的相关内容。第 5 章主要介绍机器学习,包含机器学习的概念与类型、机器学习的流程、模型选择、常见分类方法、常见聚类方法以及集成学习等内容。第 6 和 7 章主要介绍深度学习,其中第 6 章包含神经元与神经网络、BP 神经网络、卷积神经网络、循环神经网络、贝叶斯深度学习等内容,第 7 章包含注意力与记忆机制、自编码器、强化学习、对抗学习等内容。第 8 章主要介绍计算机视觉,包含计算机视觉概述、图像分析和理解、计算机视觉应用等内容。第 9

章主要介绍自然语言与语音处理，包含情感分类、机器翻译、自然语言人机交互，还包含语音识别、合成、转换等内容。第 10 章主要介绍智能机器人，包含机器人概述、机器人组织结构、工作原理以及机器人的应用等内容。第 11 章主要介绍大数据与区块链，包含大数据的基本概念、应用、关键技术以及区块链的技术基础与应用、区块链与人工智能等相关内容。第 12 章主要介绍 Python 编程基础知识，包含 Python 的安装、版本、使用方法、代码编写、IPython、Python 的各种库等内容。本书作为高等院校人工智能课程的教材以 32~48 学时为宜。

本书的编写分工为：樊重俊、杨云鹏、熊红林、刘臣、朱玥、樊鸿飞、王雅琼负责编写全书的大纲和框架，并负责全书的组织、审校，他们还对各章分别进行了修改。李璟轩、施柏州、樊重俊撰写第 1 章，鞠晓玲、施柏州、樊重俊撰写第 2 章，臧悦悦、施柏州、樊重俊撰写第 3 章，安艾芝、施柏州、樊重俊撰写第 4 章，余莹、尹裴、樊重俊撰写第 5 章，黄耐、尹裴、樊重俊撰写第 6 章，安艾芝、尹裴、樊重俊撰写第 7 章，鞠晓玲、施柏州、樊重俊撰写第 8 章，鞠晓玲、尹裴、樊重俊撰写第 9 章，臧悦悦、施柏州、樊重俊撰写第 10 章，蒋雨桥、尹裴、樊重俊撰写第 11 章，李璟轩、尹裴、樊重俊撰写第 12 章。关晓飞、杨云鹏参与第 4~7 章、第 11 和 12 章的撰写。熊红林、刘臣、施柏州、尹裴对全部内容进行了修改。徐佩参与全部章节的协调工作与部分章节的撰写。杨云鹏现为上海交通大学博士后，关晓飞现为同济大学副教授，樊鸿飞博士为金山云架构师，其他编写人员工作单位均为上海理工大学。

平台经济专家、中国管理科学与工程学会副理事长、上海市人民政府参事、《系统管理学报》主编、上海交通大学行业研究院副院长陈宏民教授对本书给予了大量支持，提出了很多修改建议，并在百忙之中为本书撰写了序，特此致谢！

笔者近年来专注于人工智能、大数据、电子商务、“互联网+”等领域的研究、教学与咨询。本书有些内容是我们团队在为企业做咨询服务时的一些思考与知识积累。上海财经大学常务副校长徐飞教授，信息安全专家、全国高等学校计算机教育研究会常务理事、复旦大学计算机科学技术学院原副院长赵一鸣，上海机场(集团)有限公司技术中心总经理冉祥来博士，原中国电子商务协会副理事长、中国出入境检验检疫协会唐生副会长，中国出入境检验检疫协会段小红秘书长，数字经济专家、国家创新与发展战略研究会副理事长吕本富教授，国家创新与发展战略研究会副会长兼秘书长王博永博士，产业互联网 CIP 模式创始人张勇军博士，著名管理咨询与数字经济专家、中驰车福董事长兼 CEO、联想集团原全球副总裁张后启博士，东方钢铁电子商务有限公司张春前总经理，上海市民政局信息研究中心黄爱国主任，北京大学信息科学技术学院数字媒体研究所贾惠柱副所长，同济大学博士生导师张建同教授、王洪伟教授，华东理工大学博士生导师李英教授，均在本书写作过程中不同程度地给予了一些有益的建议，在此一并感谢。

上海理工大学党委副书记、上海市高等学校信息技术水平考试委员会副主任、教育部大学计算机课程教学指导委员会委员顾春华教授；上海理工大学管理学院院长赵来军教授、管理科学与工程博士后流动站站长马良教授、朱小栋副教授、刘勇副教授、张宝明副教授、倪静副教

授、张惠珍副教授、刘宇熹博士、刘雅雅博士、赵敬华博士对本书提出了很多有益的建议。本书获得上海高校课程思政领航计划支持，在此一并感谢。

在本书的编写过程中，力求跟踪人工智能学科最新的技术水平和发展方向，引入新的技术和方法。由于笔者的水平有限和人工智能快速发展的特性，书中难免有不尽如人意之处，甚至是错漏，敬请诸位专家、读者批评指正。我们的电话是010-62796045，信箱是huchenhao@263.net。

本书配套的电子课件和课后习题参考答案可以到 <http://www.tupwk.com.cn/downpage> 网站下载，也可以通过扫描下方二维码来下载。



樊重俊
于上海理工大学
2020年5月

目 录 ◀

第 1 章 绪论	1
1.1 人工智能的概念	1
1.1.1 智能	1
1.1.2 人工智能	4
1.1.3 图灵测试	4
1.2 人工智能发展简史	5
1.2.1 孕育阶段	5
1.2.2 形成阶段	5
1.2.3 发展阶段	6
1.2.4 人工智能的学派	8
1.3 人工智能的研究及应用领域	12
1.4 人工智能的发展现状和趋势	18
1.5 本章小结	20
1.6 习题	20
参考文献	20
第 2 章 知识与知识表示	22
2.1 知识与知识表示的内涵	22
2.1.1 知识的概念	22
2.1.2 知识的特征	23
2.1.3 知识表示的概念	24
2.2 知识表示方法	24
2.2.1 一阶谓词逻辑表示法	24
2.2.2 产生式表示法	33
2.2.3 状态空间表示法	36
2.2.4 框架表示法	40
2.3 知识图谱及其应用	44
2.3.1 知识图谱概述	44
2.3.2 知识图谱应用示例	45
2.4 本章小结	47
2.5 习题	48
参考文献	48

第 3 章 自动推理与专家系统	50
3.1 自动推理基本知识	50
3.1.1 确定性推理	50
3.1.2 不确定性推理	59
3.2 专家系统	63
3.2.1 专家系统的内涵	63
3.2.2 专家系统的结构	63
3.2.3 专家系统的设计与实现	66
3.2.4 专家系统的应用与发展	71
3.3 本章小结	77
3.4 习题	78
参考文献	78
第 4 章 搜索算法与智能计算	80
4.1 搜索算法	80
4.1.1 盲目搜索策略	81
4.1.2 启发式搜索	83
4.1.3 博弈搜索策略	89
4.2 遗传算法	91
4.2.1 基本遗传算法	91
4.2.2 遗传算法的基本操作	93
4.2.3 遗传算法的一般步骤	99
4.2.4 遗传算法的应用	100
4.3 蚁群算法	100
4.3.1 蚁群算法的基本思想	101
4.3.2 基本的蚁群算法模型	101
4.3.3 蚁群算法的参数选择	103
4.3.4 蚁群算法的应用	103
4.4 粒子群优化算法	105
4.4.1 粒子群优化算法的基本原理	105
4.4.2 粒子群优化算法的流程	106
4.4.3 粒子群优化算法的参数分析	107

4.4.4 粒子群优化算法的应用	108	6.2.3 BP 学习算法	153
4.5 本章小结	110	6.2.4 BP 神经网络在模式识别中 的应用	155
4.6 习题	110	6.3 卷积神经网络	156
参考文献	112	6.3.1 卷积神经网络的结构	157
第5章 机器学习	113	6.3.2 卷积神经网络的模型实例	162
5.1 机器学习的概念与类型	113	6.3.3 几种典型的卷积神经网络	163
5.1.1 机器学习的概念	113	6.4 循环神经网络	164
5.1.2 机器学习的类型	114	6.4.1 循环神经网络的结构	165
5.2 机器学习的流程	119	6.4.2 BPTT 学习算法	166
5.2.1 模型	119	6.4.3 梯度消失和梯度爆炸	167
5.2.2 训练	120	6.4.4 基于门控机制的循环神经网络	167
5.3 模型性能度量	122	6.5 贝叶斯深度学习	170
5.3.1 模型精度	122	6.5.1 贝叶斯公式	170
5.3.2 查准率、查全率与 F 值	122	6.5.2 贝叶斯深度学习	171
5.3.3 ROC 与 AUC	123	6.5.3 基于贝叶斯深度学习的预测 和训练	172
5.4 常见分类方法	124	6.5.4 贝叶斯深度学习框架	173
5.4.1 逻辑回归与 Softmax 回归	124	6.6 本章小结	174
5.4.2 KNN	126	6.7 习题	174
5.4.3 朴素贝叶斯	129	参考文献	176
5.4.4 决策树	130		
5.4.5 支持向量机	132		
5.5 常见聚类方法	137	第7章 深度学习(2)	177
5.5.1 原型聚类	137	7.1 注意力与记忆机制	177
5.5.2 密度聚类	138	7.1.1 注意力	177
5.5.3 层次聚类	141	7.1.2 注意力机制	178
5.6 集成学习	142	7.1.3 常见的记忆方式	180
5.7 本章小结	143	7.1.4 典型的记忆网络	183
5.8 习题	143	7.1.5 典型场景应用	184
参考文献	144	7.2 自编码器	185
第6章 深度学习(1)	145	7.2.1 传统自编码器	185
6.1 神经元与神经网络	145	7.2.2 改进的自编码器	186
6.1.1 神经元模型	146	7.2.3 自编码器的应用	192
6.1.2 神经网络的结构	147	7.3 强化学习	194
6.1.3 神经网络的优化算法	148	7.3.1 强化学习的基本原理	194
6.1.4 神经网络的发展历程	151	7.3.2 强化学习的分类及任务	195
6.2 BP 神经网络	152	7.3.3 强化学习算法	195
6.2.1 BP 神经网络的结构	152	7.3.4 强化学习的主要应用	197
6.2.2 BP 神经网络的基本原理	153	7.4 对抗学习	198

7.4.1 GAN 的提出背景	198
7.4.2 GAN 的核心原理	200
7.4.3 GAN 的学习方法	201
7.4.4 GAN 的衍生模型	202
7.4.5 GAN 的应用领域	203
7.4.6 GAN 的思考与展望	204
7.5 联邦学习	207
7.5.1 联邦学习的提出背景	207
7.5.2 联邦学习的基本内涵	207
7.5.3 联邦学习的应用探索	209
7.6 本章小结	210
7.7 习题	211
参考文献	211
第 8 章 计算机视觉	213
8.1 计算机视觉概述	213
8.2 图像的分析 and 理解	215
8.2.1 基于浅层模型的方法	216
8.2.2 基于深度模型的方法	219
8.3 计算机视觉的典型应用	222
8.3.1 模式识别	222
8.3.2 动态行为分析	225
8.3.3 机器视觉	226
8.4 本章小结	227
8.5 习题	228
参考文献	228
第 9 章 自然语言处理与语音处理	229
9.1 自然语言处理	229
9.1.1 自然语言处理概述	229
9.1.2 情感分类	233
9.1.3 机器翻译	244
9.1.4 自然语言人机交互	246
9.2 语音处理	256
9.2.1 语音识别	256
9.2.2 语音合成	261
9.2.3 语音转换	265
9.3 本章小结	266
9.4 习题	267
参考文献	267
第 10 章 机器人	269
10.1 机器人概述	269
10.1.1 机器人的定义	269
10.1.2 机器人的分类	270
10.1.3 机器人的特性	271
10.1.4 机器人三原则	271
10.1.5 智能机器人的发展历程	271
10.2 机器人的基本构成	272
10.3 机器人的工作原理	273
10.4 人工智能技术在机器人中的应用	274
10.4.1 智能感知	275
10.4.2 智能导航与规划	277
10.4.3 智能控制与操作	279
10.4.4 智能交互	280
10.5 机器人的应用	282
10.5.1 机器人的典型应用场景	282
10.5.2 智能机器人的发展展望	287
10.6 本章小结	289
10.7 习题	289
参考文献	289
第 11 章 大数据与区块链	291
11.1 大数据	291
11.1.1 大数据的基本概念与特征	291
11.1.2 大数据的应用	293
11.1.3 大数据的关键技术	297
11.1.4 大数据与云计算、物联网的关系	303
11.1.5 大数据与人工智能应用探讨	303
11.2 区块链	306
11.2.1 区块链概述	306
11.2.2 区块链的技术基础	308
11.2.3 区块链与人工智能	310
11.2.4 区块链的应用探讨与展望	314
11.3 大数据与区块链的关系	316
11.4 本章小结	317

11.5 习题	318	12.7.1 NumPy 简史	330
参考文献	318	12.7.2 安装 NumPy	330
第 12 章 Python 编程基础	319	12.7.3 ndarray 对象	331
12.1 Python——编程语言	319	12.7.4 基本操作	335
12.2 Python——解释器	320	12.8 Pandas 库	338
12.3 安装 Python	321	12.8.1 Pandas: Python 数据 分析库	338
12.4 使用 Python	321	12.8.2 安装 Pandas	338
12.4.1 Python shell	321	12.8.3 测试 Pandas 是否安装成功	339
12.4.2 运行完整的 Python 程序	322	12.8.4 Pandas 数据结构简介	339
12.4.3 使用 IDE 编写代码	322	12.9 matplotlib 库	347
12.4.4 与 Python 交互	322	12.9.1 安装 matplotlib	348
12.5 编写 Python 代码	322	12.9.2 pyplot 模块	348
12.5.1 Python 基础语法	322	12.9.3 为图表添加更多元素	353
12.5.2 数学运算	325	12.10 本章小结	357
12.5.3 导入新的库和函数	326	12.11 习题	357
12.5.4 函数式编程	328	参考文献	358
12.6 PyPI 软件仓库——Python 包索引	329		
12.7 NumPy 库	330		



第 1 章

绪 论

人工智能(Artificial Intelligence, AI)是在计算机科学、控制论、信息论、神经心理学、哲学、语言学等多学科研究的基础上发展起来的综合性很强的交叉学科,是一门新思想、新观念、新理论、新技术不断出现的新兴学科,也是正在迅速发展的前沿学科。自 1956 年正式提出“人工智能”这个术语并把它作为一门新兴学科的名称以来,尽管该学科的发展经历了曲折的过程,但它在知识表示、自动推理、认知建模、机器学习、神经计算、自然语言理解、专家系统、智能机器人等方向开展了大量的研究工作,获得了迅速发展,并取得了惊人成就。人工智能与空间技术、原子能技术一起被誉为 20 世纪三大科学技术成就。有人称它为继三次工业革命后的又一次革命,认为前三次工业革命主要扩展了人手的功能,把人类从繁重的体力劳动中解放出来,而人工智能则扩展了人脑的功能,实现脑力劳动的自动化。

本章将首先介绍人工智能的概念以及发展简史,然后介绍当前人工智能的发展现状以及未来的发展趋势,最后简要描述本课程的基本内容,使读者对人工智能广阔的研究及应用领域有总体的了解。

1.1 人工智能的概念

人工智能的目标是用机器实现人类的部分智能。因此,下面首先讨论人类的智能行为。

1.1.1 智能

智能及智能的本质是古今中外许多哲学家、脑科学家一直在努力探索和研究的课题,但至今仍然没有完全了解。智能的发生与物质的本质、宇宙的起源、生命的本质一起被列为自然界四大奥秘。近年来,随着脑科学、神经心理学等研究的进展,人们对人脑的结构和功能有了初步认识,但对整个神经系统的内部结构和作用机制,特别是脑的功能原理还没有认识清楚,有待进一步探索。目前还不可能对智能给出精确的、公认的定义,这就导致对于智能的多种说法。

目前,根据对人脑已有的认识,结合智能的外在表现,从不同的角度、不同的侧面,用不同的方法对智能进行研究,人们提出了几种不同的观点。其中影响较大的观点有思维理论、知识阈值理论及进化理论等。

1. 思维理论

思维理论认为智能的核心是思维,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,因而通过对思维规律与方法的研究渴望揭示智能的本质。

不同的划分观点认为,思维科学体系的基础科学包括两大类:一类是总结人类思维经验、揭示思维对象的普遍规律和思维本身普遍规律的各种思维科学,包括哲学世界观、哲学史、认识论和逻辑学,是理论的思维科学。另一类思维科学包括研究思维主体——人脑的生理结构和功能,揭示思维过程生理机制的神经生理学和神经解剖学等。这种观点将认识论归在思维科学的基础科学范围内。其实这两种观点,都不否认人工智能和哲学是通过认识论相联系的。

2. 知识阈值理论

知识阈值理论着重强调知识对于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一般化程度。一个系统之所以智能,是因为它具有可运用的知识。因此,知识阈值理论把智能定义为:智能就是在巨大的搜索空间中迅速找到一个令人满意的解的能力。这一理论在人工智能的发展史中有着重要影响,知识工程、专家系统等都是在这一理论的影响下发展起来的。

3. 进化理论

进化理论着重强调控制,该理论认为人的本质能力是动态环境中的行走能力、对外界事物的感知能力、维持生命和繁衍生息的能力。正是这些能力对智能的发展提供了基础,因此智能是某种复杂系统所浮现的性质,是由许多部件交互作用产生的。智能仅取决于感知和行为,它可以在没有明显的可操作的内部表达的情况下产生,也可以在没有明显的推理系统出现的情况下产生,因而智能是在系统与周围环境不断“刺激-反应”的交互中发展和进化的。该理论的核心是用控制取代表示,从而取消概念、模型及显式表示的知识,否定抽象对于智能及智能模拟的必要性,强调分层结构对于智能进化的可能性与必要性。目前这些观点尚未形成完整的理论体系,有待进一步研究。

目前,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。具体地说,智能具有以下特征。

1) 具有感知能力

感知能力是指通过视觉、听觉、触觉、嗅觉等感觉器官感知外部世界的的能力。感知是人类获取外部信息的基本途径,人类的大部分知识都是通过感知获取有关信息,然后经过大脑加工获得的。如果没有感知,人们就不可能获得知识,也不可能引发各种智能活动。因此,感知是产生智能活动的前提与必要条件。根据有关研究,视觉与听觉在人类感知中占有主导地位,80%以上的外界信息是通过视觉得到的,10%是通过听觉得到的。因此,在人工智能的机器感知研究方面,主要研究机器视觉及机器听觉。

2) 具有记忆与思维能力

记忆与思维是人脑最重要的功能，是人有智能的根本原因。记忆用于存储由感知器官感知到的外部信息以及由思维产生的知识；思维用于对记忆的信息进行处理，即利用已有的知识对信息进行分析、计算、比较、判断、推理、联想及决策等。思维是一个动态过程，是获取知识以及运用知识求解问题的根本途径。思维可分为逻辑思维、形象思维以及顿悟思维等。

逻辑思维

逻辑思维又称为抽象思维，是一种根据逻辑规则对信息进行处理理性思维方式。人们首先通过感觉器官获得外部事物的感性认识，将它们存储于大脑中，然后通过匹配选出相应的逻辑规则，并且作用于已经表示成一定形式的已知信息，进行相应的逻辑推理。这种推理一般比较复杂，一般不是用一条规则进行一次推理就能解决问题，而是要对第一次推出的结果再运用新的规则进行新一轮的推理。推理是否成功取决于两个因素，一是用于推理的规则是否完备，二是已知的信息是否完善、可靠。如果推理规则是完备的，由感性认识获得的初始信息是完善、可靠的，那么通过逻辑思维可得到合理、可靠的结论。

形象思维

形象思维又称为直感思维，是一种以客观现象为思维对象、以感性形象认识为思维材料、以意象为主要思维工具、以指导创造物化形象的实践为主要目的的思维活动。思维过程有两次飞跃：第一次飞跃是从感性形象认识到理性形象认识的飞跃，即把对事物的感觉组合起来，形成反映事物多方面属性的整体性认识(即知觉)，再在知觉的基础上形成具有一定概括性的感觉反映形式(即表象)，然后经形象分析、形象比较、形象概括及组合形成对事物的理性形象认识；第二次飞跃是从理性形象认识到实践的飞跃，即对理性形象认识进行联想、想象等加工，在大脑中形成新的意象，然后回到实践中，接受实践的检验。这个过程不断循环，就构成了形象思维从低级到高级的运动发展。

顿悟思维

顿悟思维又称为灵感思维，是一种显意识与潜意识相互作用的思维方式。当人们遇到一个无法解决的问题时，会“苦思冥想”。这时，大脑处于一种极为活跃的思维状态，会从不同的角度用不同方法去寻求解决问题的方法。有时一个“想法”突然从脑中涌现出来，使人“茅塞顿开”，问题便迎刃而解。像这样用于沟通有关知识或信息的“想法”通常被称为灵感。灵感也是一种信息，可能是与问题直接有关的重要信息，也可能是与问题并不直接相关并且不起眼的信息，只是由于它的到来使解决问题的智慧被启动起来了。顿悟思维比形象思维更复杂，至今人们还不能确切地描述灵感的机理。1830年奥斯特在指导学生实验时，看见电流能使磁针偏转，从而发现了电磁关系。虽然很偶然，但也是在他10年探索的基础上发现的。

3) 具有学习能力

学习是人的本能。每个人都在通过与环境的相互作用，不断地学习，从而积累知识，适应环境的变化。学习既可能是自觉的、有意识的，也可能是不自觉的、无意识的。只是由于个人所处的环境不同，条件不同，学习的效果亦不相同，体现出不同的智能差异。

4) 具有行为能力

人们通常用语言或者某个表情、眼神及形体动作来对外界的刺激做出反应，传达某个信息，

这些称为行为能力或表达能力。如果把人们的感知能力看作信息的输入，则行为能力就可以看作信息的输出，它们都受到神经系统的控制。

▼ 1.1.2 人工智能

所谓人工智能，就是人工的方法在机器(计算机)上实现的智，或者说是人类智能在机器上的模拟，因此又可称为机器智能。

现在，“人工智能”这个术语已被用于“研究如何在机器上实现人类智能”这门学科的名称。从这个意义上说，可把它定义为：人工智能是一门研究如何构造智能机器(智能计算机)或智能系统，使它能模拟、延伸、扩展人类智能的学科。通俗地说，人工智能就是要研究如何使机器能听、会说、能看、会写、能思维、会学习，能适应环境变化，能解决各种面临的实际问题。

▼ 1.1.3 图灵测试

关于“人工智能”的含义，早在它正式提出之前，就由英国数学家图灵(A. M. Turing)提出了。1950，图灵发表了题为《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”，形象地指出了什么是人工智能以及机器应该达到的智能标准。图灵在这篇论文中指出不要问机器是否具有思维，而是要看它能否通过如下测试：分别让人与机器位于两个房间里，二者之间可以通话，但彼此都看不到对方，如果通过对话，人的一方不能分辨对方是人还是机器，那么就可以认为对方的那台机器达到了人类智能的水平。为了进行这个测试，图灵还设计了很有趣且智能性很强的对话内容，称为“图灵的梦想”。

现在许多人仍把图灵测试作为衡量机器智能的准则。但也有许多人认为图灵测试仅仅反映了结果，没有涉及思维过程。即使机器通过了图灵测试，也不能认为机器就有智能。针对图灵测试，哲学家约翰·塞尔勒在1980年设计了“中文屋思想实验”以说明这一观点。在该实验中，一个完全不懂中文的人在一间密闭的屋子里，有一本中文处理规则的书。他不必理解中文就可以使用这些规则。屋外的测试者不断通过门缝给他写一些有中文语句的纸条。他在书中查找处理这些中文语句的规则，根据规则将一些中文字符抄在纸条上作为对相应语句的回答，并将纸条递出房间。这样，在屋外的测试者看来，仿佛屋里的人是一个以中文为母语的人，但实际上并不理解他所处理的中文，也不会在此过程中提高自己对中文的理解。用计算机模拟这个系统，可以通过图灵测试。这说明按照规则执行的计算机程序不能真正理解其输入输出的意义。许多人对约翰·塞尔勒的“中文屋思想实验”进行了反驳，但还没有人能够彻底将其驳倒。

实际上，要使机器达到人类智能的水平，是非常困难的。但是，人工智能的研究正朝着这个方向前进，图灵的梦想总有一天会变成现实。特别是在专业领域内，人工智能能够充分利用计算机的特点，具有显著的优越性。

1.2 人工智能发展简史

人工智能是在 1956 年作为一门新兴学科的名称正式提出的，自此之后，它已经取得惊人的成就，获得迅速的发展，它的发展历史，可归结为孕育、形成、发展三个阶段。

1.2.1 孕育阶段

这个阶段主要是指 1956 年以前。自古以来，人们就一直试图用各种机器来代替人的部分脑力劳动，以提高人们征服自然的能力，其中对人工智能的产生、发展有重大影响的主要研究成果包括：

- 早在公元前384年至公元前322年，伟大的哲学家亚里士多德(Aristotle)就在他的名著《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。
- 英国哲学家培根(F. Bacon)曾系统地提出了归纳法，还提出了“知识就是力量”的警句。这对于研究人类的思维过程，以及自20世纪70年代人工智能转向以知识为中心的研究产生了重要影响。
- 德国数学家和哲学家莱布尼茨(G. W. Leibniz)提出了万能符号和推理计算的思想，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，而且是现代机器思维设计思想的萌芽。
- 英国逻辑学家布尔(C. Boole)致力于使思维规律形式化和实现机械化，并创立了布尔代数。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。
- 英国数学家图灵(A. M. Turing)在1936年提出了一种理想计算机的数学模型，即图灵机，为后来电子数字计算机的问世奠定了理论基础。
- 美国神经生理学家麦克洛奇(W. McCulloch)与匹兹(W. Pitts)在1943年建成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领域，为后来神经网络的研究奠定了基础。
- 美国爱荷华州立大学的阿塔纳索夫(Atanasoff)教授和他的研究生贝瑞(Berry)在1937年至1941年间开发的世界上第一台电子计算机“阿塔纳索夫-贝瑞计算机(Atanasoff-Berry Computer, ABC)”为人工智能的研究奠定了物质基础。需要说明的是，世界上第一台计算机不是许多书上所说的由美国的莫克利和埃柯特在1946年发明的。这是美国历史上一桩著名的公案。

由上面的发展过程可以看出，人工智能的产生和发展绝不是偶然的，而是科学技术发展的必然产物。

1.2.2 形成阶段

这个阶段主要是指 1956 年至 1969 年。1956 年夏季，由当时达特茅斯大学的年轻数学助教、现任斯坦福大学教授麦卡锡(J. McCarthy)联合哈佛大学年轻数学和神经学家、麻省理工学院教授明斯基(M. L. Minsky)，IBM 公司信息研究中心负责人洛切斯特(N. Rochester)，贝尔实验

室信息部数学研究员香农(C. E. Shannon)共同发起,邀请普林斯顿大学的莫尔(T. Moore)和 IBM 公司的塞缪尔(A. L. Samuel)、麻省理工学院的塞尔夫里奇(O. Selfridge)和索罗莫夫(R. Solomonff)以及兰德(RAND)公司和卡内基梅隆大学的纽厄尔(A. Newell)、西蒙(H. A. Simon)等在美国达特茅斯大学召开了一次为时两个月的学术研讨会,讨论关于机器智能的问题。会上经麦卡锡提议正式采用了“人工智能”这一术语。麦卡锡因而被称为“人工智能之父”。这是一次具有历史意义的重要会议,它标志着人工智能作为一门新兴学科正式诞生了。此后,美国形成了多个人工智能研究组织,如纽厄尔和西蒙的 Carnegie-RAND 协作组、明斯基和麦卡锡的 MIT 研究组、塞缪尔的 IBM 工程研究组等。

自这次会议之后的 10 多年间,人工智能的研究在机器学习、定理证明、模式识别、问题求解、专家系统及人工智能语言等方面取得了许多引人注目的成就,例如:

- 在机器学习方面,1957年Rosenblatt研制成功了感知机。这是一种将神经元用于识别的系统,它的学习功能引起了广泛的兴趣,推动了连接机制的研究,但人们很快发现了感知机的局限性。
- 在定理证明方面,美籍华人数理逻辑学家王浩于1958年在IBM-704机器上用三到五分钟时间证明了《数学原理》中有关命题演算的全部定理(220条),并且还证明了谓词演算中150条定理的85%。1965年鲁滨逊(J. A. Robinson)提出了归结原理,为定理的机器证明做出突破性的贡献。
- 在模式识别方面,1959年塞尔夫里奇推出了一个模式识别程序,1965年罗伯特(Roberts)编制出可分辨积木构造的程序。
- 在问题求解方面,1960年纽厄尔等人通过心理学试验总结出人们求解问题的思维规律,编制了通用问题求解程序(General Problem Solver, GPS),可以用来求解11种不同类型的问题。
- 在专家系统方面,美国斯坦福大学的费根鲍姆(E. A. Feigenbaum)领导的研究小组自1965年开始专家系统DENDRAL的研究,1968年完成并投入使用。该专家系统能根据质谱仪的实验,通过分析推理决定化合物的分子结构,其分析能力已接近甚至超过有关化学专家的水平,在美、英等国得到了实际应用。该专家系统的研制成功不仅为人们提供了一个实用的专家系统,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,为以后专家系统的建造树立了榜样,对人工智能的发展产生了深刻影响,其意义远远超过系统本身在实用性方面创造的价值。
- 在人工智能语言方面,1960年麦卡锡研制出人工智能语言(LISP),成为建造专家系统的重要工具。
- 1969年成立的国际人工智能联合会议(International Joint Conferences On Artificial Intelligence, IJCAD)是人工智能发展史上一个重要的里程碑,它标志着人工智能这门新兴学科已经得到世界的肯定和认可。1970年创刊的国际性杂志《人工智能》(Artificial Intelligence)对推动人工智能的发展、促进研究者们的交流起到重要的作用。

1.2.3 发展阶段

这个阶段主要是指 1970 年以后。进入 20 世纪 70 年代,许多国家都开展了人工智能的研

究,涌现出大量的研究成果。例如,1972年法国马赛大学的科麦瑞尔(A. Comerauer)提出并实现了逻辑程序设计语言 PROLOG,斯坦福大学的肖特利夫(E. H. Shorliffe)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,和其他新兴学科的发展一样,人工智能的发展道路也不是平坦的。例如,机器翻译的研究没有像人们最初想象的那么容易。当时人们总以为只要一部双向词典及一些词法知识就可以实现两种语言文字间的互译。后来发现机器翻译远非这么简单。实际上,由机器翻译出来的文字有时会出现十分荒谬的错误。例如,当把“眼不见,心不烦”的英语句子“Out of sight, out of mind”翻译成俄语后变成“又瞎又疯”;当把“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,然后翻译回来时竟变成“The wine is good but the meat is spoiled”,翻译过来成了“酒是好的,但肉变质了”;当把“光阴似箭”的英语句子“Time flies like an arrow”翻译成日语,然后翻译回来的时候,竟成了“苍蝇喜欢箭”。由于机器翻译出现的这些问题,1960年美国政府顾问委员会的一份报告裁定:“还不存在通用的科学文本机器翻译,也没有很近的实现前景。”因此,英国、美国当时中断了对大部分机器翻译项目的资助。在其他方面,如问题求解、神经网络、机器学习等,也都遇到了困难,使人工智能的研究一时陷入困境。

人工智能研究的先驱者们认真反思,总结前一段研究的经验和教训。1977年费根鲍姆在第五届国际人工智能联合会议上提出了“知识工程”的概念,对以知识为基础的智能系统的研究与建造起到重要的作用。大多数人接受了费根鲍姆关于以知识为中心展开人工智能研究的观点。从此,人工智能的研究又迎来了蓬勃发展的以知识为中心的新时期。

在这个时期,专家系统的研究在多个领域取得了重大突破,各种不同功能、不同类型的专家系统雨后春笋般建立起来,产生了巨大的经济效益及社会效益。例如,地矿勘探专家系统 PROSPECTOR 拥有15种矿藏知识,能根据岩石标本及地质勘探数据对矿藏资源进行估计和预测,能对矿床分布、储藏量、品位及开采价值进行推断,制定合理的开采方案。人们应用该系统成功找到了超亿美元的钼矿。专家系统 MYCIN 能识别51种病菌,正确处理23种抗菌素,可协助医生诊断、治疗细菌感染性血液病,为患者提供最佳处方。该系统成功地治愈了数百个病例,并通过了严格的测试,显示出较高的医疗水平。美国 DEC 公司的专家系统 XCON 能根据用户要求确定计算机的配置。由专家做这项工作一般需要3小时,而该系统只需要0.5分钟,速度提高大约360倍。DEC 公司还建立了另外一些专家系统,由此产生的净收益每年超过4000万美元。信用卡认证辅助决策专家系统 American Express 能防止不应有的损失,据说每年可节省2700万美元左右。

专家系统的成功,使人们越来越清楚地认识到知识是智能的基础,对人工智能的研究必须以知识为中心来进行。对知识的表示、利用及获取等的研究取得较大的进展,特别是对不确定性知识的表示与推理取得了突破,建立了主观贝叶斯理论、确定性理论、证据理论等,对人工智能中模式识别、自然语言理解等领域的发展提供了支持,解决了许多理论及技术上的问题。

人工智能在博弈中的成功应用也举世瞩目。人们对博弈的研究一直抱有极大兴趣,早在1956年人工智能刚刚作为一门学科问世时,塞缪尔就研制出跳棋程序。这个程序能从棋谱中学习,也能从下棋实践中提高棋艺。1959年它击败了塞缪尔本人,1962年又击败了美国一个州的冠军。1991年8月在悉尼举行的第12届国际人工智能联合会议上,IBM公司研制的“深思”(Deep Thought)计算机系统就与澳大利亚象棋冠军约翰森(D. Johansen)举行了一场人机对抗赛,

结果以 1:1 平局告终。1957 年西蒙曾预测 10 年内计算机可以击败人类的世界冠军。虽然在 10 年内没有实现，但 40 年后深蓝计算机击败国际象棋棋王卡斯帕罗夫(Kasparov)，仅仅比预测迟了 30 年。

1996 年 2 月 10 日至 17 日，为了纪念世界上第一台电子计算机诞生 50 周年，美国 IBM 公司出巨资邀请国际象棋棋王卡斯帕罗夫与 IBM 公司的深蓝计算机系统进行了六局的“人机大战”。这场比赛被人们称为“人脑与电脑的世界决战”。参赛双方分别代表了人脑和电脑的世界最高水平。当时的深蓝是一台运算速度达每秒 1 亿次的超级计算机。第一盘，深蓝就给卡斯帕罗夫一个下马威，赢了这位世界冠军，给世界棋坛带来极大震动。但卡斯帕罗夫总结经验，稳扎稳打，在剩下的五盘中赢三盘，平两盘，最后以总比分 4:2 获胜。一年多后，1997 年 5 月 3 日至 11 日，深蓝再次挑战卡斯帕罗夫。这时，深蓝是一台拥有 32 个处理器和强大并行计算能力的 RS/6000SP/2 超级计算机，运算速度达每秒 2 亿次。计算机里存储了百余年来世界顶尖棋手的棋局，5 月 3 日棋王卡斯帕罗夫首战击败深蓝，5 月 4 日深蓝扳回一局，之后双方战平三局。双方的决胜局于 5 月 11 日拉开了帷幕，卡斯帕罗夫在这盘比赛中仅仅走了 19 步便放弃了抵抗，比赛用时只有 1 小时多一点。这样，深蓝最终以 3.5:2.5 的总比分赢得这场举世瞩目的“人机大战”的胜利。深蓝的胜利表明了人工智能达到的成就。尽管棋路还远非真正对人类思维方式的模拟，但它已经向世人说明，计算机能够以人类远远不能企及的速度和准确性，实现属于人类思维的大量任务。深蓝精湛的残局战略使观战的国际象棋专家们大为惊讶。卡斯帕罗夫也表示：“这场比赛中有许多新的发现，其中之一就是计算机有时也可以走出人性化的棋步。在一定程度上，我不能不赞扬这台机器，因为它对盘势因素有着深刻的理解，我认为这是一项杰出的科学成就。”因为这场胜利，IBM 的股票升值为 180 亿美元。

1.2.4 人工智能的学派

根据前面的论述，我们知道要理解人工智能，就要研究如何在一般的意义上定义知识，可惜的是，准确定义知识也是件十分复杂的事情。严格来说，人们最早使用的知识定义是柏拉图在《泰阿泰德篇》中给出的：“被证实的、真的和被相信的陈述”(Justified True Belief, 简称 JTB 条件)。

然而，这个延续了两千多年的定义在 1963 年被哲学家盖梯尔否定了。盖梯尔提出了一个著名的悖论(简称“盖梯尔悖论”)。该悖论说明柏拉图给出的知识定义存在严重缺陷。虽然后来人们给出了很多知识的替代定义，但直到现在仍然没有定论。

但关于知识，至少有一点是明确的，那就是知识的基本单位是概念。要想精通掌握任何一门知识，就必须从这门知识的基本概念开始学习。而知识自身也是概念。因此，如何定义一个概念，对于人工智能具有非常重要的意义。给出定义看似简单，实际上却非常难，因为经常会涉及自指的性质(自指：词性的转化——由谓词性转化为体词性，语义则保持不变)。一旦涉及自指，就会出现非常多的问题，很多的语义悖论都出于概念自指。

自指与转指这对概念最早出自朱德熙先生的《自指与转指》(《方言》1983 年第一期，《朱德熙文集》第三卷)。陆俭明先生在《八十年代中国语法研究》中说：“自指和转指的区别在于，自指单纯是词性的转化——由谓词性转化为体词性，语义则保持不变；转指则不仅词性发生转化，语义也发生变化，尤指行为动作或性质本身转化为与行为动作或性质相关的事物。”

举例：

(1) 教书的来了(“教书的”是转指，转指教书的“人”)，教书的时候要认真(“教书的”语义没变，是自指)。

(2) **unplug** 一词的原意为“不使用(电源)插座”，是自指，常用来转指为不使用电子乐器的清唱。

(3) **colored** 在表示 **having colour**(着色)时是自指，**colored** 在表示有色人种时就是转指。

(4) **rich**，富有的，是自指；**the rich**，富人，是转指。

知识本身也是一个概念。据此，人工智能的问题就变成了如下三个问题：一、如何定义(或者表示)一个概念、如何学习一个概念、如何应用一个概念。因此，对概念进行深入研究就非常必要了。

那么，如何定义一个概念呢？为简单起见，这里先讨论最为简单的经典概念。经典概念的定义由三部分组成：第一部分是概念的符号表示，即概念的名称，说明这个概念叫什么，简称概念名；第二部分是概念的内涵表示，由命题表示，命题就是能判断真假的陈述句；第三部分是概念的外延表示，由经典集合表示，用来说明与概念对应的实际对象是哪些。

举一个经典概念的常见例子——素数(**prime number**)，其内涵表示是一个命题，即只能被被1和自身整除的自然数。

概念有什么作用呢？或者说概念定义的各个组成部分有什么作用呢？经典概念定义三部分各有作用，且彼此不能互相代替。具体来说，概念有三个作用或功能，要掌握一个概念，就必须清楚其三个功能。

第一个功能是概念的指物功能，即指向客观世界的对象，表示客观世界的对象的可观测性。对象的可观测性是指对象对于人或者仪器的知觉感知特性，不依赖于人的主观感受。举一个《阿Q正传》里的例子：那赵家的狗，何以看我两眼呢？句中“赵家的狗”应该是指现实世界当中的一条真正的狗。但概念的指物功能有时不一定能够实现，有些概念设想存在的对象在现实世界里并不存在，例如“鬼”。

第二个功能是指心功能，即指向人的心智世界里的对象，代表心智世界里的对象表示。鲁迅有一篇著名的文章《论丧家的资本家的乏走狗》，显然，这个“狗”不是现实世界中的狗，只是他心智世界中的狗，即心里的狗。概念的指心功能一定存在。如果对于某个人，一个概念的指心功能没有实现，则该词对于该人不可见，简单地说，该人不理解该概念。

最后一个功能是指名功能，即指向认知世界或符号世界中表示对象的符号名称，这些符号名称可组成各种语言。最著名的例子是乔姆斯基的“**colorless green ideas sleep furiously**”，这句话翻译过来是“无色的绿色思想在狂怒地休息”。这句话没有什么意思，但是完全符合语法，纯粹在语义符号世界里，仅仅指向符号世界而已。当然也有另外，“鸳鸯两字怎生书”指的就是“鸳鸯”这两个字组成的名字。一般情形下，概念的指名功能依赖于不同的语言系统或符号系统，由人类创造，属于认知世界。同一个概念在不同的符号系统里，概念名不一定相同，如汉语称“雨”，英语称“**rain**”。

根据波普尔的三个世界理论，认知世界、物理世界与心理世界虽然相关，但各不相同。因此，一个概念的三个功能虽然彼此相关，但也各不相同。更重要的是，人类文明发展至今，这三个功能不断发展，彼此都越来越复杂，但概念的三个功能并没有改变。

在现实生活中，如果想要了解一个概念，就需要知道这个概念的三个功能：要知道概念的

名字，也要知道概念所指的对象(可能是物理世界)，更要在自己的心智世界里具有该概念的形象(或图像)。如果只有一个，那是不行的。

知道了概念的三个功能之后，就可以理解人工智能的三个学派以及各学派之间的关系。

人工智能也是一个概念，而要使一个概念成为现实，自然要实现概念的三个功能。人工智能的三个学派关注于如何才能让机器具有人工智能，并根据概念的不同功能给出不同的研究路线。专注于实现 AI 指名功能的人工智能学派称为符号主义，专注于实现 AI 指心功能的人工智能学派称为连接主义，专注于实现 AI 指物功能的人工智能学派称为行为主义。

1. 符号主义

符号主义的代表人物是 Simon 与 Newell，他们提出了物理符号系统假设，即只要在符号计算上实现了相应的功能，那么在现实世界中就实现了对应的功能，这是智能的充分必要条件。因此，符号主义认为，只要在机器上是正确的，现实世界就是正确的。说得更通俗一点，指名对了，指物自然正确。

在哲学上，关于物理符号系统假设也有一个著名的思想实验——前面 1.1.3 节中提到的图灵测试。图灵测试要解决的问题就是如何判断一台机器是否具有智能。

图灵测试将智能的表现完全限定在指名功能里。但马少平教授的故事已经说明，只在指名功能里实现概念的功能，并不能说明一定实现了概念的指物功能。实际上，根据指名与指物的不同，哲学家约翰·塞尔专门设计了思想实验用来批判图灵测试，这就是著名的“中文屋实验”。

中文屋实验明确说明，即使符号主义成功了，全是符号的计算跟现实世界也不一定搭界，即完全实现指名功能也不见得具有智能。这是哲学上对符号主义的正式批评，明确指出了按照符号主义实现的人工智能不等同于人的智能。

虽然如此，符号主义在人工智能研究中依然扮演了重要角色，早期工作的主要成就体现在机器证明和知识表示上。在机器证明方面，早期 Simon 与 Newell 做出了重要的贡献，王浩、吴文俊等华人也得出了很重要的结果。机器证明以后，符号主义最重要的成就是专家系统和知识工程，最著名的学者就是 Feigenbaum。如果认为沿着这条路就可以实现全部智能，显然存在问题。日本第五代智能机就是沿着知识工程这条路走的，后来的失败在现在看来是完全合乎逻辑的。

实现符号主义面临的观实挑成主要有三个。第一个是概念的组合爆炸问题。每个人掌握的基本概念大约有 5 万个，且形成的组合概念是无穷的。因为常识难以穷尽，推理步骤可以无穷。第二个是命题的组合悖论问题。两个都是合理的命题，合起来就成了没法判断真假的句子了，比如著名的柯里悖论(Curry's Paradox)。第三个也是最难的问题，就是经典概念在实际生活当中是很难得到的，知识也难以提取。上述三个问题成了符号主义发展的瓶颈。

2. 连接主义

连接主义认为大脑是一切智能的基础，主要关注于大脑神经元及其连接机制，试图发现大脑的结构及其处理信息的机制、揭示人类智能的本质机理，进而在机器上实现相应的模拟。前面已经指出知识是智能的基础，而概念是知识的基本单元，因此连接主义实际上主要关注于概念的心智表示以及如何计算机上实现其心智表示，这对应着概念的指心功能。2016 年发表在 *Nature* 上的一篇学术论文揭示了大脑语义地图的存在性，文章指出概念都可以在每个脑区找到

对应的表示区，确确实实概念的心智表示是存在的。因此，连接主义也有其坚实的物理基础。

连接主义学派的早期代表人物有麦克洛奇、匹兹、霍普菲尔德等。按照这条路，连接主义认为可以实现完全的人工智能。对此，哲学家普特南设计了著名的“缸中之脑实验”，可以看作对连接主义的一次哲学批判。

缸中之脑实验描述如下：一个人(可以假设是你自己)被邪恶科学家进行了手术，脑被切下来并放在存有营养液的缸中。脑的神经末梢被连接在计算机上，同时计算机按照程序向脑传递信息。对于这个人来说，人、物体、天空都存在，神经感觉等都可以输入，这个大脑还可以被输入、截取记忆，比如截取掉大脑手术的记忆，然后输入他可能经历的各种环境、日常生活，甚至可以输入代码，“感觉”到自己正在阅读一段有趣而荒唐的文字。

缸中之脑实验说明即使连接主义实现了，指心没有问题，但指物依然存在严重问题。因此，连接主义实现的人工智能也不等同于人的智能。

尽管如此，连接主义仍是目前最为大众所知的一条 AI 实现路线。在围棋上，采用了深度学习技术的 AlphaGo 战胜了李世石，之后又战胜了柯洁。在机器翻译上，深度学习技术已经超过人的翻译水平。在语音识别和图像识别上，深度学习也已经达到实用水准。客观地说，深度学习的研究成就已经取得工业级的进展。

但是，这并不意味着连接主义就可以实现人的智能。更重要的是，即使要实现完全的连接主义，也面临极大挑战。到现在为止，人们并不清楚人脑表示概念的机制，也不清楚人脑中概念的具体表示形式、表示方式和组合方式等。现在的神经网络与深度学习实际上与人脑的真正机制距离尚远。

3. 行为主义

行为主义假设智能取决于感知和行动，不需要知识、表示和推理，只需要将智能行为表现出来就好，只要能实现指物功能就可以认为具有智能了。这一学派的早期代表作是 Brooks 的六足爬行机器人。

对此，哲学家普特南也设计了一个思想实验，可以看作对行为主义的哲学批判，这就是“完美伪装者和斯巴达人”。完美伪装者可以根据外在的需求进行完美的表演，需要哭的时候可以哭得让人撕心裂肺，需要笑的时候可以笑得让人兴高采烈，但是其内心可能始终冷静如常。斯巴达人则相反，无论其内心是激动万分还是心冷似铁，其外在总是一副泰山崩于前而色不变的表情。完美伪装者和斯巴达人的外在表现都与内心没有联系，这样的智能如何从外在行为进行测试？因此，行为主义路线实现的人工智能也不等同于人的智能。

对于行为主义路线，其面临的实现最大困难可以用莫拉维克悖论来说明。所谓莫拉维克悖论，是指对计算机来说困难的问题是简单的、简单的问题是困难的，最难以复制的反而是人类技能中那些无意识的技能。目前，模拟人类的行动技能面临很大挑战。比如，在网上看到波士顿动力公司人形机器人可以做高难度的后空翻动作，大狗机器人可以在任何地形负重前行，其行动能力似乎非常强。但是这些机器人都有大的缺点——能耗过高、噪音过大。大狗机器人原是美国军方订购的产品，但因为大狗机器人开动时的声音在十里之外都能听到，大大提高了其成为一个活靶子的可能性，使其在战场上几乎没有实用价值，美国军方最终放弃了采购。

1.3 人工智能的研究及应用领域

在大多数学科中存在着几个不同的研究领域，每个领域都有其特有的感兴趣的研究课题、研究技术和术语。在人工智能中，这样的领域包括自然语言处理、自动定理证明、自动程序设计、智能检索、智能调度、机器学习、机器人学、专家系统、智能控制、模式识别、视觉系统、神经网络、智能体、计算智能、问题求解、人工生命、人工智能方法和程序设计语言等。在过去 50 多年中，已经建立了一些具有人工智能的计算机系统，例如，能够求解微分方程、下棋、设计分析集成电路、合成人类自然语言、检索情报、诊断疾病，以及能够控制太空飞行器、地面移动机器人和水下机器人。

本节首先介绍人工智能一些最基本的概念和基本原理，为后面的各种应用建立基础。值得指出的是，正如不同的人工智能子领域不是完全独立的一样，这里简介的各种智能特性也不是互不相关的。把它们分开介绍只是为了便于指出现有的人工智能程序能够做些什么和还不能做什么。大多数人工智能研究课题都涉及许多智能领域。

1. 问题求解与博弈

人工智能的第一大成就是发展了能够求解难题的下棋(如国际象棋)程序(已在 1.2.3 节做详细介绍)。在下棋程序中应用的某些技术，如向前看几步，并把困难的问题分成一些比较容易的子问题，发展为搜索和问题消解(归约)这样的人工智能基本技术。

通用人工智能是人工智能领域的一项重大挑战，目标是让机器能够像人一样完成各种各样的任务。具体到游戏领域，通用对弈游戏(General Game Playing, GGP)致力于开发一种能够以人类水准玩任意已知或未知游戏的人工智能系统。IBM 的超级计算机“深蓝”战胜国际象棋冠军卡斯帕罗夫标志着人工智能达到一个新的高度。然而，“深蓝”的智能至少在两个方面还存在局限性：其一是只能玩国际象棋这种游戏，不具有通用性；其二是依赖于大量的人类游戏经验，不具有完全自主学习能力。GGP 研究的目标就是突破这些局限，它设置的环境要求机器必须在没有人类游戏经验的指导下玩各种各样的游戏。因此，GGP 研究的进展反映了机器游戏智能在通用性和自主学习方面的发展，GGP 比赛则成为一种评价机器游戏智能的标准。

今天的计算机程序能够下锦标赛水平的各种方盘棋、十五子棋、中国象棋和国际象棋，并取得前面提到的计算机棋手战胜国际象棋和中国象棋冠军的成果。另一种问题求解程序把各种数学公式符号汇编在一起，其性能达到很高的水平，并正在为许多科学家和工程师所应用。有些程序甚至还能够用经验来改善性能。

如前所述，这个问题中未解决的问题包括人类棋手具有的但尚不能明确表达的能力，如国际象棋大师们洞察棋局的能力。另一个未解决的问题涉及问题的原概念，在人工智能中叫做问题表示的选择。人们常常能够找到某种思考问题的方法从而使求解变易而解决该问题。到目前为止，人工智能程序已经知道如何考虑它们要解决的问题，即搜索解答空间、寻找较优的解答。

2. 逻辑推理与定理证明

早期的逻辑演绎研究工作与问题和难题的求解相当密切。已经开发出的程序能够借助于对事实数据库的操作来“证明”断定；其中每个事实由分立的数据结构表示，就像数理逻辑中由

分立公式表示一样。与人工智能的其他技术的不同之处是，这些方法能够完整和一致地加以表示。也就是说，只要本原事实是正确的，程序就能够证明这些从事实得出的定理，而且也仅仅是证明这些定理。

逻辑推理是人工智能研究中最持久的子领域之一。特别重要的是要找到一些方法，只把注意力集中在一个大型数据库中的有关事实上，留意可信证明，并在出现新信息时适时修正这些证明。

对数学中臆测的定理寻找证明或反证，确实称得上是一项智能任务。为此不仅需要有能力根据假设进行演绎的能力，而且需要某些直觉技巧。但是至少在当前，人工智能在数学推理中的表现并不尽如人意。这种推理对人工系统具有很大的挑战性，因为它不仅仅涉及处理数字，还需要一套认知能力，包括学习基本公理以及以正确的顺序进行推理、计划和做事的能力。

3. 计算智能

计算智能(computational intelligence)涉及神经计算、模糊计算、进化计算、粒群计算、自然计算、免疫计算和人工生命等研究领域。

进化计算(evolutionary computation)是指一类以达尔文进化论为依据来设计、控制和优化人工系统的技术和方法的总称，包括遗传算法(genetic algorithm)、进化策略(evolutionary strategy)和进化规划(evolutionary programming)。自然选择的原则是适者生存，即物竞天择、优胜劣汰。

自然进化的这些特征早在 20 世纪 60 年代就引起霍兰(Holland)的极大兴趣。受达尔文进化论思想的影响，他逐渐认识到在机器学习中，为获得一个好的学习算法，仅靠单个策略的建立和改进是不够的，还要依赖于包含许多候选策略的群体的繁殖。他还认识到生物的自然遗传现象与人工自适应系统行为的相似性，因此提出在研究和设计人工自主系统时可以模仿生物自然遗传的基本方法。20 世纪 70 年代初，霍兰提出了“模式理论”，并于 1975 年出版了《自然系统与人工系统的自适应》专著，系统地阐述了遗传算法的基本原理，奠定了遗传算法研究的理论基础。

遗传算法、进化规划、进化策略具有共同的理论基础——生物进化论，因此，把这三种方法统称为进化计算，而把相应的算法称为进化算法。

人工生命是 1987 年提出的，旨在用计算机和精密机械等人工媒介生成或构造出能够表现自然生命系统行为特征的仿真系统或模型系统。自然生命系统行为具有自组织、自复制、自修复等特征以及形成这些特征的混沌动力学、进化和环境适应。

人工生命的理论和方法有别于传统人工智能和神经网络的理论和方法。人工生命把生命现象所体现的自适应机理通过计算机进行仿真，对相关非线性对象进行更真实的动态描述和动态特征研究。人工生命学科的研究内容包括生命现象的仿生系统、人工建模与仿真、进化动力学、人工生命的计算理论、进化与学习综合系统以及人工生命的应用等。

4. 分布式人工智能与智能体

分布式人工智能(Distributed AI, DAI)是分布式计算与人工智能结合的结果。DAI 系统以鲁棒性作为控制系统质量的标准，并具有互操作性，即不同的异构系统在快速变化的环境中具有交换信息和协同工作的能力。

分布式人工智能的研究目标是创建一种能够描述自然系统和社会系统的精确概念模型。

DAI 中的智能并非独立存在的概念,只能在团体协作中实现,因而主要研究智能体间的合作与对话,包括分布式问题求解和多智能体系统(Multi-Agent System, MAS)。在这两个领域, MAS 更能体现人类的社会智能,具有更大的灵活性和适应性,更适合开放和动态的世界环境,因而倍受重视,已成为人工智能乃至计算机科学和控制科学与工程的研究热点。MAS 解决实际问题的方式可以理解作为一种基于智能体的协作问题,而分布式约束则可以描述领域对象的性质、相互关系、任务要求、目标,因此可以作为一种有效的方法表示智能体间的协作关系。

5. 自动程序设计

自动程序设计能够以各种不同的目的描述来编写计算机程序。对自动程序设计的研究不仅可以促进半自动软件开发系统的发展,而且也使通过修正自身数码进行学习的人工智能系统得到发展。程序理论方面的有关研究工作对人工智能的所有研究工作都是很重要的。

自动编制一份程序来获得某种指定结果的任务与证明一份给定程序将获得某种指定结果的任务是紧密相关的,后者叫做程序验证。

自动程序设计研究的重大贡献之一是作为问题求解策略的调整概念。人们已经发现,对于程序设计或机器人控制问题,先产生一个不费事的有错误的解,再修改它,这种做法比坚持求第一个完全没有缺陷的解的做法有效得多。

6. 专家系统

一般来说,专家系统是智能计算机程序系统,其内部具有大量专家水平的某个领域知识与经验,能够利用人类专家的知识 and 解决问题的方法来解决该领域的问题。

发展专家系统的关键是表达和运用专家知识,即来自人类专家的并已被证明对解决有关领域内的典型问题有用的事实和过程。专家系统和传统的计算机程序的本质区别在于专家系统所要解决的问题一般没有算法解,并且经常要在不完全、不精确或不确定的信息基础上得出结论。

随着人工智能整体水平的提高,专家系统也获得发展。正在开发的新一代专家系统有分布式专家系统和协同式专家系统等。在新一代专家系统中,不但采用基于规则的方法,而且采用基于框架的技术和基于模型的原理。

7. 机器学习

学习是人类智能的主要标志和获得知识的基本手段。机器学习(自动获取新的事实及推理算法)是使计算机具有智能的根本途径。此外,机器学习还有助于发现人类学习的机理并揭示人脑的奥秘。

传统的机器学习倾向于使用符号表示而不是数值表示,使用启发式方法而不是算法。传统机器学习的另一倾向是使用归纳(induction)而不是演绎(deduction)。前一倾向使它有别于人工智能的模式识别等分支,后一倾向使它有别于定理证明等分支。

按系统对导师的依赖程度可将学习方法分类为机械式学习、讲授式学习、类比学习、归纳学习、观察发现式学习等。

近十多年来又发展了下列各种学习方法:基于解释的学习、基于事例的学习、基于概念的学习、基于神经网络的学习、遗传学习、增强学习以及数据挖掘和知识发现等。数据挖掘和知识发现是 20 世纪 90 年代初期崛起的一个活跃的研究领域。在数据库基础上实现的知识发现系

统,通过综合运用统计学、粗糙集、模糊数学、机器学习和专家系统等多种学习手段和方法,从大量的数据中提炼出抽象的知识,从而揭示出这些数据背后蕴涵的客观世界的内在联系和本质规律,实现知识的自动获取。大规模数据库和互联网的迅速发展,使人们对数据库的应用提出新的要求。数据库中包含的大量知识无法得到充分发掘与利用,会造成信息的浪费,并产生大量的数据垃圾。另一方面,知识获取仍是专家系统研究的瓶颈问题。从领域专家获取知识是非常复杂的个人之间的交互过程,具有很强的个性和随机性,没有统一的办法。因此,人们开始考虑以数据库作为新的知识源。数据挖掘和知识发现能自动处理数据库中大量的原始数据,抽取出具有必然性的、富有意义的模式,成为有助于人们实现其目标的知识,找出人们对所需问题的解答。这已经成为许多领域研究的前沿与热点。例如人工智能技术在电力系统和综合能源系统中的应用,将改变能源传统利用模式,促进系统进一步智能化;在农业领域中,基于机器学习方法开展农作物分类研究,以确保粮食安全和生态安全;在生物医学领域中,可以将影像学和以机器学习等为代表的人工智能技术相结合,为探寻疾病客观生物学标志物及其临床应用提供科学证据和参考依据。

8. 大数据与深度学习

大数据是人工智能的基础,人工智能的决策依赖于大数据的分析。大数据时代改变了基于数理统计的传统数据科学,促进了数学分析方法的创新,从机器学习和多层神经网络演化而来的深度学习是当前数据处理与分析的研究前沿。

大数据是数据分析的前沿技术,其核心价值在于对海量数据进行存储和分析。数据是资源,更可以说是战略资源,海量的数据运用“大数据+人工智能”技术可以在庞大、复杂的数据海洋中迅速得到需要的信息,从而对实践的指导更加具有现实意义。人工智能、机器算法、大数据的结合运用领域,可以说对人类的未来有着举足轻重的影响。例如以物联网、云计算、信息物理系统、社会信息物理系统、大数据和深度学习等新一代信息通信技术为基础的第二代智能制造对制造业产生了翻天覆地的影响,极大提高了生产的性能与效率;在生物医学领域中,深度学习方法基于患者的疾病相关数据,通过模型预测异常病变或发病风险,进行疾病的辅助诊断,为医师提供参考,使其判断不受主观因素的干扰,在减轻医师工作负担的同时提升效率和诊断准确率。

大数据与深度学习技术一方面可以解放劳动力,另一方面还可以促进生产力的发展和经济的进步。未来那些简单、重复且操作单一的工种毫无疑问会在人工智能和大数据时代被淘汰。

9. 自然语言理解

语言处理也是人工智能的早期研究领域之一,并引起进一步的重视。人们已经编写出能够从内部数据库回答问题的程序,这些程序通过阅读文本材料和建立内部数据库,能够把句子从一种语言翻译为另一种语言,执行给出的指令和获取知识等。有些程序甚至能够在一定程度上翻译从话筒输入的口头指令。

当人们用语言互通信息时,他们几乎可以不费力地进行极其复杂却又只需要一点点理解的过程。语言已经发展成为智能动物之间的一种通信媒介,可在某些环境条件下把一点“思维结构”从一个头脑传输到另一个头脑,而每个头脑都拥有庞大的、高度相似的周围思维结构作为公共的文本。这些相似的、前后有关的思维结构中的一部分允许每个参与者知道对方也拥有这

种共同结构，并能在通信“动作”中用来执行某些处理。语言的生成和理解是极为复杂的编码和解码问题。

10. 机器人学

人工智能研究中日益受到重视的另一个分支是机器人学。一些并不复杂的动作控制问题，如移动式机器人的机械动作控制问题，表面上看并不需要很多智能。然而人类几乎下意识就能完成的这些任务，要是由机器人实现，就要求机器人具备在求解需要较多智能的问题时可能用到的能力。

机器人和机器人学的研究促进了许多人工智能思想的发展，所带来的一些技术可用来模拟世界的状态，用来描述从一种世界状态转变为另一种世界状态的过程。

智能机器人的研究和应用体现出广泛的学科交叉，涉及众多的课题，如机器人体系结构、机构、控制、智能、视觉、触觉、力觉、听觉、机器人装配、在恶劣环境中工作的机器人以及机器人语言等。机器人已在各种工业、农业、商业、旅游业，在空中和海洋以及国防等领域获得越来越普遍的应用。

随着人工智能技术的发展，类生命机器人的研发也受到广泛的关注。类生命机器人由活体生物系统与传统机电系统深度有机融合而成，具有高能量效率、高本质安全性、高灵敏度以及可自修复等潜在优点。由于类生命机器人具有单独生命体或者以机电系统为主体的传统机器人系统所不具备的特性，因此对它的研究已成为当今的热点，并且在近些年的研究中取得了一定的重要成果。

11. 模式识别

计算机硬件的迅速发展，计算机应用领域的不断开拓，急切要求计算机能更有效地感知诸如声音、文字、图像、温度、振动等人类赖以发展自身并改造环境而运用的信息资料。

着眼于拓宽计算机的应用领域，提高感知外部信息能力的学科——模式识别得到迅速发展。人工智能所研究模式识别是指用计算机代替人类或帮助人类感知模式，是对人类感知外界功能的模拟，研究的是计算机模式识别系统，也就是使计算机系统具有模拟人类通过感官接收外界信息、识别和理解周围环境的感知能力。

实验表明，人类接收外界信息的80%以上来自视觉，10%左右来自听觉。所以，早期的模式识别研究工作集中在对视觉图像和语音的识别上。

模式识别是一个不断发展的新学科，它的理论基础和研究范围也在不断发展。随着生物医学对人类大脑的初步认识，模拟人脑构造的计算机实验——人工神经网络方法已经成功用于手写字符的识别、汽车牌照的识别、指纹识别、语音识别等方面。

12. 机器视觉

机器视觉或计算机视觉已从模式识别的一个研究领域发展为一门独立的学科。在视觉方面，人们已经给计算机系统装上电视输入装置以便能“看见”周围的东西。在人工智能中研究的感知过程通常包含一组操作。

整个感知问题的要点是形成精练的表示以取代难以处理的、极其庞大的未经加工的输入数据。最终表示的性质和质量取决于感知系统的目标。不同系统有不同的目标，但所有系统都必

须把来自输入的、多得惊人的感知数据简化为一种易于处理的且有意义的描述。

计算机视觉通常可分为低层视觉与高层视觉两类。低层视觉主要执行预处理功能，如边缘检测、运动目标检测、纹理分析以及通过阴影获得形状、立体造型、曲面色彩等。高层视觉主要是理解对象，需要掌握与对象相关的知识。

机器视觉的前沿研究领域包括实时并行处理、主动式定性视觉、动态和时变视觉、三维景物的建模与识别、实时图像的压缩传输和复原、多光谱和彩色图像的处理与解释等。在一些传统领域，往往离不开人的判断，机器视觉可以一定程度上替代或辅助这些工作。例如在土木工程领域，机器视觉已被公认为检查和监测基础设施寿命和损耗情况的重要方法；在产品生产和制造的过程中，产品品质感官审评往往存在主观判断失误的缺陷，机器视觉技术也可以对产品品质进行快速无损评价。

13. 神经网络

研究结果已经证明，用神经网络处理直觉和形象思维信息具有比传统处理方式好得多的效果。神经网络的发展有着非常广阔的科学背景，是众多学科研究的综合成果。神经生理学家、心理学家与计算机科学家的共同研究得出的结论是：人脑是功能特别强大、结构异常复杂的信息处理系统，其基础是神经元及其互联关系。因此，对人脑神经元和人工神经网络的研究可能创造出新一代人工智能机——神经计算机。

对神经网络的研究始于20世纪40年代初期，经历了一条十分曲折的道路，几起几落，自20世纪80年代初以来，对神经网络的研究再次出现高潮。

对神经网络模型、算法、理论分析和硬件实现的大量研究，为神经计算机走向应用提供了物质基础。人们期望神经计算机能重建人脑的形象，极大地提高信息处理能力，在更多方面取代传统的计算机。

14. 智能控制

人工智能的发展促进自动控制向智能控制发展。智能控制是一类无需(或需要尽可能少的)人的干预就能独立地驱动智能机器实现其目标的自动控制。或者说，智能控制是驱动智能机器自主地实现其目标的过程。许多复杂的系统，难以建立有效的数学模型和用常规控制理论进行定量计算与分析，而必须采用定量数学解析法与基于知识的定性方法的混合控制方式。随着人工智能和计算机技术的发展，已有可能把自动控制和人工智能以及系统科学的某些分支结合起来，建立一种适用于复杂系统的控制理论和技术。智能控制正是在这种条件下产生的，是自动控制的最新发展阶段，也是用计算机模拟人类智能的一个重要研究领域。

智能控制是同时具有以知识表示的非数学广义世界模型和以数学公式模型表示的混合控制过程，也往往是含有复杂性、不完全性、模糊性、不确定性以及不存在已知算法的非数学过程，并以知识进行推理，以启发来引导求解过程。智能控制的核心在高层控制，即组织级控制，其任务在于对实际环境或过程进行组织，即决策和规划，以实现广义问题求解。

15. 智能调度与指挥

确定最佳调度或组合的问题是人们感兴趣的又一类问题。一个古典的问题就是推销员旅行问题(Traveling Salesman Problem, TSP)。许多问题具有这类相同的特性。

在这些问题中,有几个问题(包括推销员旅行问题)属于理论计算机科学家所说的 NP 完全性问题。根据理论上的最佳方法计算出所耗时间(或所走步数)的最坏情况来排列不同问题的难度。所耗时间或所走步数是随着问题大小的某种量度增长的。

人工智能学家曾经研究过若干组合问题的求解方法。有关问题域的知识再次成为比较有效的求解方法的关键。智能组合调度与指挥方法已被应用于汽车运输调度、列车的编组与指挥、空中交通管制以及军事指挥等系统。

16. 智能检索

随着科学技术的迅速发展,出现了“知识爆炸”的情况。对国内外种类繁多且数量巨大的科技文献进行检索远非人力和传统检索系统所能胜任。研究智能检索系统已成为科技持续快速发展的重要保证。

数据库系统是存储某学科大量事实的计算机软件系统,它们可以回答用户提出的有关该学科的各种问题。数据库系统的设计也是计算机科学的一个活跃分支。为了有效地表示、存储和检索大量事实,已经发展出许多技术。

智能信息检索系统的设计者们将面临以下几个问题。首先,建立能够理解以自然语言陈述的询问系统本身就存在不少问题。其次,即使可以通过规定某些机器能够理解的形式化询问语句来回避语言理解问题,也仍然存在一个如何根据存储的事实演绎出答案的问题。最后,为理解询问和演绎答案所需要的知识可能超出该学科领域数据库所能表示的知识。

17. 系统与语言工具

除了直接瞄准实现智能的研究工作外,开发新的方法也往往是人工智能研究的一个重要方面。人工智能对计算机界的某些最大贡献已经以派生的形式表现出来。计算机系统的一些概念,如分时系统、编目处理系统和交互调试系统等,已经在人工智能研究中得到发展。一些能够简化演绎、机器人操作和认识模型的专用程序设计和系统常常是新思想的丰富源泉。几种知识表达语言(把编码知识和推理方法作为数据结构和过程计算机的语言)已在 20 世纪 70 年代后期开发出来,以探索各种建立推理程序的思想。20 世纪 80 年代以来,计算机系统,如分布式系统、并行处理系统、多机协作系统和各种计算机网络等,都有了长足发展。在人工智能程序设计语言方面,除了继续开发和改进通用和专用的编程语言新版本和新语种外,还研究出一些面向目标的编程语言和专用开发工具。对关系数据库研究取得的进展,无疑为人工智能程序设计提供了新的有效工具。



1.4 人工智能的发展现状和趋势



如 1.2.4 节所述,符号主义认为只要实现指名功能就能实现人工智能,连接主义认为只要实现指心功能就可以实现人工智能,行为主义认为只要实现指物功能就可以实现人工智能。人工智能的三大流派虽然取得了很大进展,但各自也面临巨大挑战。简单地说,人工智能的三大流派假设能够成立的前提是指名、指物、指心功能等价。

可是这个前提成立吗?早期的人工智能研究使用经典概念,而经典概念至少具有以下 5 个

假设：①概念的外延表示可以用经典集合表示；②概念的内涵表示存在命题表示；③指称对象的外延表示与内涵表示名称一致；④概念表示唯一，同一个概念表示与个体无关，对于同一个概念，每个人的表示都是一样的；⑤概念的内涵表示与外延表示在指称对象上功能等价。

可以明显看出，在上述5个假设之下，经典概念的指心、指物、指名功能是等价的，即指名意味着指物、指心。但是，日常生活中使用的概念一般并不满足经典概念的5个假设，因此也不能保证其指心、指物、指名功能等价。《周易·系辞上》中也说，“书不尽言，言不尽意”，明确指出了指名、指心与指物不一定等价。下面给出两个例子，以说明日常生活中概念上的指名、指物功能并不等价。

第一个例子，微信上流传过一个著名的段子：一个人说手头有一个亿，谁有项目通知一下，一起投资。不然，再晚一点，就洗手不干了。听的人以为指物，即真有一个亿的资金；而实际上，这个人只是在手上写了三个字“一个亿”而已。这里纯粹指名，“手头有一个亿”仅仅是符号“一个亿”而已。这个段子显然利用了概念上的指名与指物不一定等价的性质。

第二个例子，西方绘画史上有幅著名的画，画面上画了一个烟斗，题字却说这不是烟斗。其显然是想说明符号与实物不同，即指名与指物不等价。在现实生活中，也可以发现指名与指心不等价的例子。

综上所述，概念上的指名、指物与指心功能在生活中并不等价，单独实现概念上的一个功能并不能保证具有智能。因此，单独遵循一个学派并不足以实现人工智能，现在的人工智能研究已经不再强调遵循人工智能的单一学派，很多时候会综合各个流派的技术。比如，从专家系统发展起来的知识图谱已经不完全遵循符号主义的路线了。在围棋上战胜人类顶尖棋手的AlphaGo综合使用了三种学习算法——强化学习、蒙特卡罗树搜索、深度学习，而这三种学习算法分属于三个人工智能流派(强化学习属于行为主义，蒙特卡罗树搜索属于符号主义，深度学习属于连接主义)。无人驾驶技术同样是突破了人工智能三大流派限制的综合技术。虽然人工智能发展至今，各个流派依然在发展，也都取得了很好的进展，但是对各个流派进行融合已经是大势所趋，特别是在大数据和云计算的助力下，新一代人工智能将带来社会的第四次技术革命。

然而，目前的人工智能还有很大的缺陷，其使用的知识表示还是建立在经典概念的基础之上。图灵测试的文章发表于1950年，当时人们对于经典概念的普适性还没有提出质疑，因此，其使用的概念是基于经典概念的。在图灵测试中，最重要的概念之一是人，包括提问者和回答者。但是什么样的人才是合适的提问者和回答者，是中国人还是英国人、是圣人还是智力障碍者或装傻者，图灵测试并没有定义。如果人存在经典定义，在图灵测试里，就容易确定什么样的人作为提问者和回答者。可惜的是，人并不能用经典概念来定义。

经典概念的基本假设还是指心、指名与指物等价，这与人类的日常生活经验严重不符，过于简单化了。在人类的现实生活中，概念上的指名、指物、指心并不总是等价的。在基于经典概念的知识表示框架下，现在的机器表现有时显得极具智障，缺乏常识、缺乏理解能力，严重缺乏处理突发状况的能力。实际上，维特根斯坦在1953年出版的《哲学研究》中明确提出，日常生活中使用的概念是没有经典概念定义的。这实际上给图灵测试带来了很大的不确定性。因此，在经典概念表示不成立的情形下，如何进行概念表示是一个极具挑战性的问题。人工智能研究已取得了阶段性进展，但是目前仍然没有任何一个智能系统能接近人类水平，具备多模态协同感知、协同多种不同认知的能力，对复杂环境具备极强的自适应能力，对新事物、新环境具备人类水平的自主学习、自主决策能力等。人工智能研究离真正实现信息处理机制类脑、认

知能力全面类人的智能系统还有很长的路要走。

1.5 本章小结

本章一开始主要介绍了人工智能的概念。人工智能是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学。智能及智能的本质是古今中外许多哲学家、脑科学家一直在努力探索和研究的课题。目前，根据对人脑已有的认识，结合智能的外在表现，从不同角度、不同侧面，用不同方法对智能进行研究，人们提出了思维理论、知识阈值理论、进化理论等不同的观点。

在对人工智能的相关概念进行阐述之余，本章后面主要讲述了人工智能发展简史、人工智能的研究及应用领域、人工智能的发展现状和趋势。人工智能作为一门新兴学科，经历了孕育、形成、发展这三个阶段。孕育阶段主要是指 1956 年以前，人类用各种机器来代替人的部分脑力劳动，以提高人类征服自然的能力。形成阶段主要是指 1956—1969 年，以人工智能作为一门新兴学科正式诞生作为起点，之后的十多年间，人工智能的研究在机器学习、定理证明、模式识别、问题求解、专家系统及人工智能语言等方面获得了长足发展。发展阶段主要是指 1970 年以后，许多国家都开展人工智能的研究，涌现出大量的研究成果，人工智能领域的学者关注于如何才能让机器具有人工智能，并根据概念的不同功能给出了不同的研究路线。专注于实现 AI 指名功能的人工智能学派称为符号主义，专注于实现 AI 指心功能的人工智能学派称为连接主义，专注于实现 AI 指物功能的人工智能学派称为行为主义。

现今，人工智能领域诞生了多个特有的研究课题、研究技术和术语，包括自然语言处理、自动定理证明、自动程序设计、智能检索、智能调度、机器学习、大数据与深度学习、机器人学、专家系统、智能控制、模式识别、视觉系统、神经网络、智能体、计算智能、问题求解、人工生命、人工智能方法和程序设计语言等。

1.6 习题

1. 试论述人工智能的定义。
2. 试论述人工智能的三个流派。
3. 试论述知识与概念之间的关系。

参考文献

- [1] 王万良. 人工智能导论[M]. 4 版. 北京: 高等教育出版社, 2017.
- [2] 李陶深. 人工智能[M]. 重庆: 重庆大学出版社, 2001.
- [3] 蔡自兴, 徐光祐. 人工智能及其应用[M]. 4 版. 北京: 清华大学出版社, 2015.
- [4] 李德毅, 于剑. 人工智能导论[M]. 北京: 中国科学技术出版社, 2018.

- [5] <https://baike.baidu.com/item/%E8%87%AA%E6%8C%87/4379696?fr=aladdin>.
- [6] 张海峰, 刘当一, 李文新. 通用对弈游戏: 一个探索机器游戏智能的领域[J]. 软件学报, 2016, 27(11): 2814-2827.
- [7] Sean O'Neill. 数学推理挑战人工智能[J]. Engineering, 2019, 5(5): 16-19.
- [8] 段沛博, 张长胜, 张斌. 分布式约束优化方法研究进展[J]. 软件学报, 2016, 27(2): 264-279.
- [9] 杨挺, 赵黎媛, 王成山. 人工智能在电力系统及综合能源系统中的应用综述[J]. 电力系统自动化, 2019, 43(1): 2-14.
- [10] 黄双燕, 杨辽, 陈曦, 等. 机器学习法的干旱区典型农作物分类[J]. 光谱学与光谱分析, 2018, 38(10): 3169-3176.
- [11] 孙也婷, 陈桃林, 何度, 等. 基于精神影像和人工智能的抑郁症客观生物学标志物研究进展[J]. 生物化学与生物物理进展, 2019, 46(9): 879-899.
- [12] 姚锡凡, 刘敏, 张剑铭, 等. 人工智能视角下的智能制造前世今生与未来[J]. 计算机集成制造系统, 2019, 25(1): 19-34.
- [13] 李渊, 骆志刚, 管乃洋, 等. 生物医学数据分析中的深度学习应用[J]. 生物化学与生物物理进展, 2016, 43(5):472-483.
- [14] 张闯, 王文学, 席宁, 等. 类生命机器人发展与未来挑战[J]. Engineering, 2018, 4(4): 53-77.
- [15] Billie F.Spencer Jr., Vedhus Hoskere, Yasutaka Narazaki. 基于计算机视觉的民用基础设施的检查与监测研究进展[J]. Engineering, 2019, 5(2): 199-248.
- [16] 刘鹏, 吴瑞梅, 杨普香, 等. 基于计算机视觉技术的茶叶品质随机森林感官评价方法研究[J]. 光谱学与光谱分析, 2019, 39(1): 193-198.
- [17] 曾毅, 刘成林, 谭铁牛. 类脑智能研究的回顾与展望[J]. 计算机学报, 2016, 39(1): 212-222.