眼科人工智能

杨卫华　吴茂念●主编

OPHTHALMOLOGY ARTIFICIAL INTELLIGENCE

湖北科学技术出版社
HUBEI SCIENCE & TECHNOLOGY PRESS

眼科人工智能

杨卫华　吴茂念●主编

OPHTHALMOLOGY ARTIFICIAL INTELLIGENCE

长江出版传媒 Changjiang Publishing & Media
湖北科学技术出版社 HUBEI SCIENCE & TECHNOLOGY PRESS

图书在版编目(CIP)数据

眼科人工智能 / 杨卫华，吴茂念主编. —武汉：湖北科学技术出版社，2018.2
ISBN 978-7-5706-0132-5

Ⅰ.①眼… Ⅱ.①杨… ②吴… Ⅲ.①人工智能—应用—眼科学 Ⅳ.①R77-39

中国版本图书馆 CIP 数据核字(2018)第 032848 号

责任编辑：冯友仁　程玉珊　　封面设计：喻　杨

出版发行：湖北科学技术出版社　　电话：027—87679447
地　　址：武汉市雄楚大街 268 号　　邮编：430070
　　（湖北出版文化城 B 座 13—14 层）
网　　址：http://www.hbstp.com.cn

印　　刷：武汉惜木印刷有限公司　　邮编：430070

787×1092　　1/16　　10 印张　　256 千字
2018 年 2 月第 1 版　　2018 年 2 月第 1 次印刷
定价：48.00 元

前言/PREFACE

人工智能概念诞生于1956年，在半个多世纪的发展历程中，由于受到智能算法、计算速度、存储水平等多方面因素的影响，人工智能技术和应用发展经历了多次高潮和低谷。2006年以来，以深度学习为代表的机器学习算法在机器视觉和语音识别等领域取得了极大的成功，识别准确性大幅提升，使人工智能再次受到学术界和产业界的广泛关注。云计算、大数据等技术在提升运算速度、降低计算成本的同时，也为人工智能发展提供了丰富的数据资源，协助训练出更加智能化的算法模型。而如今，各类智能化产品已经成为人类生活当中不可或缺的一部分。智能诊断、智能陪护、手术机器人、智能康复等，所有的这一切，都在表达着这样一个意念：人工智能的时代已经到来了！

以人工智能为代表的科技发展，在近几年呈现出一片欣欣向荣的傲人姿态。尤其是在医学领域，以深度学习为基础的人工智能的研究引起了广大人工智能和医学专家的关注和投入。医学人工智能的研发已经不再是欧美、日、韩等发达国家的专属，以中国为首的发展中国家也越来越关注医学人工智能的研发和实践。国内近几年新注册的近200家医学人工智能企业，越来越多的医学人工智能专题会议的举办，就是最好的说明。我国于2017年7月8日印发并实施的《新一代人工智能发展规划》，就是为抢抓人工智能发展的重大战略机遇，构筑我国人工智能发展的先发优势，加快建设创新型国家和世界科技强国而制订的。

深度学习（deep learning）是机器学习领域一个新的研究方向，近年来在语音识别、图像分析等多个类应用中取得突破性的进展。深度学习通过模拟人类上述的处理过程，组合图像低层特征形成更加抽象的高层表示、属性类别或特征，给出图像分类的识别标志，特别适合处理人工尚未发现的高度特异性识别标志图像的分类。深度学习技术与医学影像的融合和研究是未来医疗发展的重要方向之一。

眼科作为临床医学的一个重要学科，其临床诊疗中大量使用标准图像，如眼底彩色照相、眼底血管造影、眼前节照相、角膜地形图、眼位照相、眼部B超、眼部CT、视网膜光学断层扫描等。眼科影像特别是眼底照相图像非常适合训练深度学习算法，用于病灶部位特征提取和识别。如糖尿病视网膜病变在眼底照相图像上有明显的识别特征，单纯凭借

病变的颜色（血管瘤——红色、渗出——黄色或白色等）和其他特征就可以准确诊断。结合眼科影像的便捷优势的人工智能技术，有利于提高眼病的智能诊断水平，已成为医生的得力助手，有利于全民眼健康。

可以说，眼科人工智能正处在一个不错的历史际遇之中，无论是从国家角度，还是从社会需求、市场因素角度，眼科人工智能的研究和发展，都将是必不可少的。而在新一轮的时代热潮到来之前，探寻眼科人工智能发展的历程，学习其主要的应用技术，解读眼科人工智能技术的理论和原理，分享眼科人工智能研究与实践成果，同时展望这一门新兴科学在时代浪潮当中的发展前景，就是本书重点阐释的方面。

当然，由于眼科人工智能本身是一门专业知识非常强的跨界学科，其内容方面也必然会触及部分晦涩难懂的专业公式。作者在这里针对部分难点进行了整理筛选，务求通过深入浅出的语言文字，将庞杂人工智能系统之中的高深原理传达给读者。尽管我们团队做了很多努力，但由于编写时间仓促，加之经验水平有限，不足之处恳请各位专家同行多多指正。针对文章的任何意见和建议，都欢迎广大读者不吝笔墨，批评指正，以便我们及时进行更正和完善。

本书的出版，首先要感谢我们的团队和单位，感谢湖州师范学院医学人工智能重点实验室全体成员的努力，感谢湖州市第一人民医院的支持。眼科人工智能研究难度大，我们能取得今天的成绩，还要感谢在研究道路上给予支持和鼓励的老师、同行、朋友以及家人。

杨卫华　吴茂念

目录/CONTENTS

第一章 人工智能历史和现状

人工智能历史可以追溯到3 000多年前。世界各国（特别是四大文明古国的中国）不乏这方面的故事与史料。我们将简述人工智能从古至今的发展历史，从历史中获得人工智能发展的规律及特征。本书通过深度分析当前人工智能在世界各国（尤其是美国和中国）的发展现状，期待获得较好的启发，以期为眼科专家和眼科人工智能研究者较好地展示人工智能的发展趋势。

第一节 人工智能的定义

人工智能领域存在多种概念和定义，一些定义太过，而一些定义则不够，似乎就像瞎子摸象，很难全面地给出人工智能这个学科的准确定义。作为该领域创始人之一的Nilsson先生写道："人工智能缺乏通用的定义。"一本如今已经修订了三版的权威性人工智能教科书《人工智能——一种现代化方法》给出了8项定义，作者并没有透露其究竟倾向于哪种定义。实用的定义为：人工智能是对计算机系统能够如何履行那些只有依靠人类智慧才能完成的任务的理论研究。例如，视觉感知、语音识别、在不确定条件下做出决策、学习，还有语言翻译等。

比起研究人类如何进行思维活动，从人类能够完成的任务角度对人工智能进行定义，而非人类如何思考，在当今时代更加能够让我们绕开神经机制层面对智慧进行确切定义从而直接探讨它的实际应用。随着计算机为解决新任务挑战而升级换代并推而广之，人们对那些所谓需要依靠人类智慧才能解决的任务的定义门槛也越来越高。所以，人工智能的定义随着时间而演变，这一现象称之为"人工智能效应"，概括起来就是"人工智能就是要实现所有目前还无法不借助人类智慧才能实现的任务的集合"。

人工智能通常被认为是计算机学科中的一门重要分支，在控制论、心理学、神经生理学和计算机科学的影响下，诞生于20世纪二战后的40年代末50年代初，在经过多次激烈争论和起伏之后，20世纪80年代以来逐渐开始形成若干不仅相互竞争、而且相互补充的研究纲领，同时人工智能与认知科学、机器人等相近学科分化开来，成为新世纪以来的前瞻性重要交叉学科——智能科学的核心组成部分。当前人工智能已经通过学科认证，即将成为新的一级学科。北京航空航天大学已于2017年秋开始招收人工智能专业的本科生。

国家自然科学基金委信息部已经成立了人工智能处。

人工智能及其相关的研究成果，通过大众媒介于20世纪50年代末进入人们的视野，成为科幻作品和艺术中经久不衰的重要主题之一。它与生物工程、空间技术并列为21世纪的三大尖端技术。计算机的发明为人工智能的实现提供了强大的工具，反过来人工智能发展过程中的新思想和成果又促进计算机科学的发展，在计算机科学界的“诺贝尔奖”——图灵奖1966年到2001年的获奖者中，有将近六分之一是因为其人工智能方面的成就而获奖。

如果以1956年达特茅斯会议上确定“人工智能”这个名词作为学科的开端，那么这门学科60余年来的研究纲领与经典自然科学相比，其成熟度和统一性相对较低，不同学派之间或相互竞争或互为补充，而且由于是当下的科学或技术，涉及对在世科学家的评价，从而导致对其进行单独、严肃的科学史考察并不容易，即使是人工智能著名学者Nilsson自己编写人工智能史时也觉得为人工智能下一个准确的定义很困难，但为了编写学科史却不得不对其进行界定。

因此，不仅国内尚无专业的科学史学者对人工智能发展史进行深入研究，而且国外已有的人工智能史，也主要是以面向大众的非专业性普及读物出现，或者是零散的口述记录史，或者只是作为二级学科穿插在计算机科学、认知科学的历史描述中出现，与这门学科在目前大众传播中的热度相比并不相称。

但是，无论是国内还是国外，从认识论、哲学的角度研究人工智能的方法论基础，关注其哲学意义和社会影响，并考察其与认知心理学、神经心理学、逻辑学等其他学科的互动关系，都是非常活跃的。因此在这些研究文献中，经常会涉及人工智能发展历史中的问题，并与认知科学、逻辑学、数学等学科的历史描述存在大量交叉，从而成为研究人工智能历史发展的重要思想和资料来源，这为对其进行严肃的学科历史考察和研究提供了较好的外部基础。

目前在技术史的研究中，对计算机科学发展历程的研究已经出现了不少优秀的著作，但一般偏于硬件和程序设计语言的发展，对于人工智能这门分支学科，由于其脉络复杂，往往穿插在软件史中描述，偏重于实际的经验性问题，对其和心灵哲学、认识论、心理学的互动关系往往语焉不详。这带来了两个问题：首先是如果没有内容完整、严肃的人工智能史，也同样没有完整的计算机科学史，就好像如果没有好的代数学史、几何史、微积分史，就不能形成完整的数学史一样；其次是人们对计算机的认识早已突破其仅仅是一种计算工具的观念，而是将其作为研究人类智能行为的模拟工具，以及实现人类级别智能任务的技术手段，因此只是遵循将其作为计算机科学的二级学科的传统，局限在计算机学科或软件工程的历史发展来研究人工智能历史，把这门学科和数学、认知心理学、控制论、哲学等相关学科的理论渊源或联系割裂开来，则不足以解决在智能模拟、理解人类认知的任务背景下考察人工智能发展的脉络和动力机制问题。

第二节　人工智能发展简史

一、概述

人工智能在20世纪50年代就已经开始启动，这段探索的历史喧嚣与渴望、挫折和失望交替出现。

人工智能的历史源远流长。在古代的神话传说中，技艺高超的工匠可以制作人造人，并为其赋予智能或意识。现代意义上的AI始于古典哲学家用机械符号处理的观点解释人类思考过程的尝试。20世纪40年代基于抽象数学推理的可编程数字计算机的发明使一批科学家开始严肃地探讨构造一个电子大脑的可能性。

1956年夏天，美国达特茅斯大学（Dartmouth）召开了一次影响深远的历史性会议。这次聚会本来属于朋友间沙龙式的学术研讨，与会者也仅仅只有10个人。主要发起人是该校青年助教麦卡锡（John McCarthy），以及哈佛大学明斯基（Marvin Minsky）、贝尔实验室香农（E. Shannon）和IBM公司信息研究中心罗彻斯特（N. Lochester），他们邀请了卡内基梅隆大学纽厄尔（Newell）和赫伯特·西蒙（Herbert Simon）、麻省理工学院塞夫里奇（O. Selfridge）和索罗门夫（R. Solomamff），以及IBM公司塞缪尔（A. Samuel）和莫尔（T. More）。这些青年学者的研究专业包括数学、心理学、神经生理学、信息论和电脑科学，分别从不同的角度共同探讨人工智能的可能性。他们的名字人们并不陌生，例如，香农是《信息论》的创始人，塞缪尔编写了第一个电脑跳棋程序，麦卡锡、明斯基、纽厄尔和西蒙都是图灵奖的获奖者。

达特茅斯会议长达两个多月，学者们在充分讨论的基础上，首次提出了“人工智能(artificial intelligence)”这一术语，标志着人工智能（AI）作为一门新兴学科正式诞生。

数字电子计算机发明之后，计算机界的第一件伟大创举，为什么是迫不及待地向“人工智能”发起冲击？与今天的电脑相比，当年电子计算机的计算能力，甚至不如今天最简单的学生计算器。是什么在驱使这些青年学者，奋不顾身地投身于一个迄今为止都堪称“最硬的骨头”的领域？

20世纪50年代末期，西蒙和纽厄尔宣称，“在可以预见的未来，计算机的能力将和人类智力并驾齐驱”。1970年，明斯基甚至雄心勃勃地预测：“在3～8年的时间里，我们将研制出具有普通人一般智力的计算机。这样的机器能读懂莎士比亚的著作，会给汽车上润滑油，会玩弄政治权术，能讲笑话，会争吵。到了这个程度后，计算机将以惊人的速度进行自我教育。几个月之后，它将具有天才的智力，再过几个月，它的智力将无与伦比。”

虽然MIT人工智能实验室明斯基的同事们，都觉得这样的预测有点夸张，但他们还是一致同意，实现这个目标大约需要15年的时间，“终有一日，计算机将把人当宠物对待”。

最终研究人员发现自己大大低估了这一工程的难度。由于 James Lighthill 爵士的批评和国会方面的压力，美国和英国政府于 1973 年停止向没有明确目标的人工智能研究项目拨款。7 年之后受到日本政府研究规划的刺激，美国政府和企业再次在 AI 领域投入数十亿研究经费，但这些投资者在 20 世纪 80 年代末重新撤回了投资。AI 研究领域诸如此类的高潮和低谷不断交替出现，至今仍有人对 AI 的前景做出异常乐观的预测。

二、人工智能发展简史

在 20 世纪 40 年代和 50 年代，来自不同领域（数学、心理学、工程学、经济学和政治学）的一批科学家开始探讨制造人工大脑的可能性。1956 年，人工智能被确立为一门学科。

最初的人工智能研究是 20 世纪 30 年代末到 50 年代初的一系列科学进展交汇的产物。神经学研究发现大脑是由神经元组成的电子网络，其激励电平只存在“有”和“无”两种状态，不存在中间状态。维纳的控制论描述了电子网络的控制和稳定性。克劳德·香农提出的信息论则描述了数字信号（即高低电平代表的二进制信号）。图灵的计算理论证明数字信号足以描述任何形式的计算。这些密切相关的想法暗示了构建电子大脑的可能性。

这一阶段的工作包括一些机器人的研发，例如 W. Grey Walter 的“乌龟（turtles）”，还有“约翰霍普金斯兽（johns hopkins beast）”。这些机器并未使用计算机、数字电路和符号推理，控制它们的是纯粹的模拟电路。

Walter Pitts 和 Warren McCulloch 分析了理想化的人工神经元网络，并且指出了它们进行简单逻辑运算的机制。他们是最早描述所谓“神经网络”的学者。马文·闵斯基是他们的学生，当时是一名 24 岁的研究生。1951 年他与 Dean Edmonds 一道建造了第一台神经网络机，称为 SNARC。在接下来的 50 年中，闵斯基是 AI 领域最重要的领导者和创新者之一。

1951 年，Christopher Strachey 使用曼彻斯特大学的 Ferranti Mark 1 机器写出了一个西洋跳棋（checkers）程序；Dietrich Prinz 则写出了一个国际象棋程序。Arthur Samuel 在 20 世纪 50 年代中期和 60 年代初开发的国际象棋程序的棋力已经可以挑战具有相当水平的业余爱好者。游戏 AI 一直被认为是评价 AI 进展的一种标准。

1950 年，图灵发表了一篇划时代的论文，文中预言了创造出具有真正智能的机器的可能性。由于注意到“智能”这一概念难以确切定义，他提出了著名的图灵测试：如果一台机器能够与人类展开对话（通过电传设备）而不能被辨别出其机器身份，那么称这台机器具有智能。这一简化使得图灵能够令人信服地说明“思考的机器”是可能的。论文中还回答了对这一假说的各种常见质疑。图灵测试是人工智能哲学方面第一个严肃的提案。

20 世纪 50 年代中期，随着数字计算机的兴起，一些科学家直觉地感到可以进行数字操作的机器也应当可以进行符号操作，而符号操作可能是人类思维的本质。这是创造智能机器的一条新路。

1955 年，Newell 和（后来荣获诺贝尔奖的）Simon 在 J. C. Shaw 的协助下开发了

“逻辑理论家（logic theorist）”。这个程序能够证明《数学原理》中前 52 个定理中的 38 个，其中某些证明比原著更加新颖和精巧。Simon 认为他们已经“解决了神秘的心/身问题，解释了物质构成的系统如何获得心灵的性质”。（这一断言的哲学立场后来被 John Searle 称为“强人工智能”，即机器可以像人一样具有思想。）

1956 年达特茅斯会议的组织者是 Marvin Minsky、约翰·麦卡锡和另两位资深科学家 Claude Shannon 及 Nathan Rochester。会议提出“学习或者智能的任何其他特性的每一个方面都应能被精确地加以描述，使得机器可以对其进行模拟”。与会者包括 Ray Solomonoff、Oliver Selfridge、Trenchard More、Arthur Samuel、Newell 和 Simon，他们中的每一位都将在 AI 研究的第一个 10 年中做出重要贡献。会上纽厄尔和西蒙讨论了“逻辑理论家”，而麦卡锡则说服与会者接受“人工智能”一词作为本领域的名称。1956 年达特茅斯会议上 AI 的名称和任务得以确定，同时出现了最初的成就和最早的一批研究者，因此这一事件被广泛承认为 AI 诞生的标志。

达特茅斯会议之后的数年是大发现的时代。对许多人而言，这一阶段开发出的程序堪称神奇：计算机可以解决代数应用题，证明几何定理，学习和使用英语。当时大多数人几乎无法相信机器能够如此“智能”。研究者们在私下的交流和公开发表的论文中表达出相当乐观的情绪，认为具有完全智能的机器将在 20 年内出现。ARPA（国防高等研究计划署）等政府机构向这一新兴领域投入了大笔资金。

从 20 世纪 50 年代后期到 60 年代涌现了大批成功的 AI 程序和新的研究方向。下面列举其中最具影响的几个。

许多 AI 程序使用相同的基本算法。为实现一个目标（例如赢得游戏或证明定理），它们一步步地前进，就像在迷宫中寻找出路一般；如果遇到了死胡同则进行回溯。这就是“搜索式推理”。

这一思想遇到的主要困难是，在很多问题中，“迷宫”里可能的线路总数是一个天文数字（所谓“指数爆炸”）。研究者使用启发式算法去掉那些不太可能导出正确答案的支路，从而缩小搜索范围。

Newell 和 Simon 试图通过其“通用解题器（general problem solver）”程序，将这一算法推广到一般情形。另一些基于搜索算法证明几何与代数问题的程序也给人们留下了深刻印象。例如，Herbert Gelernter 的几何定理证明机（1958）和 Minsky 的学生 James Slagle 开发的 SAINT（1961）。还有一些程序通过搜索目标和子目标做出决策，如斯坦福大学为控制机器人 Shakey 而开发的 STRIPS 系统。

AI 研究的一个重要目标是使计算机能够通过自然语言（例如英语）进行交流。早期的一个成功范例是 Daniel Bobrow 的程序 STUDENT，它能够解决高中程度的代数应用题。

如果用节点表示语义概念（例如“房子”“门”），用节点间的连线表示语义关系（例如“有 — 一个”），就可以构造出“语义网（semantic net）”。第一个使用语义网的 AI 程序由 Ross Quillian 开发；而最为成功（也是最有争议）的一个则是 Roger Schank 的“概

念关联（conceptual dependency）”。

Joseph Weizenbaum 的 ELIZA 是第一个聊天机器人，可能也是最有趣的会说英语的程序。与 ELIZA “聊天” 的用户有时会误以为自己是在和人类，而不是和一个程序交谈。但是实际上 ELIZA 根本不知道自己在说什么。它只是按固定套路作答，或者用符合语法的方式将问题复述一遍。

20 世纪 60 年代后期，麻省理工学院 AI 实验室的 Marvin Minsky 和 Seymour Papert 建议 AI 研究者们专注于被称为 “微世界” 的简单场景。他们指出在成熟的学科中往往使用简化模型帮助基本原则的理解，例如物理学中的光滑平面和完美刚体。许多这类研究的场景是 “积木世界”，其中包括一个平面，上面摆放着一些不同形状、尺寸和颜色的积木。

在这一指导思想下，Gerald Sussman（研究组长）、Adolfo Guzman、David Waltz［“约束传播（constraint propagation）” 的提出者］，特别是 Patrick Winston 等人在机器视觉领域做出了创造性贡献。同时，Minsky 和 Papert 制作了一个会搭积木的机器臂，从而将 “积木世界” 变为现实。微世界程序的最高成就是 Terry Winograd 的 SHRDLU，它能用普通的英语句子与人交流，还能做出决策并执行操作。

第一代 AI 研究者们曾做出了如下预言：

1958 年，H. A. Simon：10 年之内，数字计算机将成为国际象棋世界冠军。

1965 年，H. A. Simon：20 年内，机器将能完成人能做到的一切工作。

1967 年，Marvin Minsky：一代之内……创造人工智能的问题将获得实质上的解决。

1970 年，Marvin Minsky：在 3～8 年的时间里我们将得到一台具有人类平均智能的机器。

1963 年 6 月，MIT 从新建立的 ARPA（即后来的 DARPA，国防高等研究计划局）获得了 220 万美元经费，用于资助 MAC 工程，其中包括 Minsky 和 McCarthy 5 年前建立的 AI 研究组。此后 ARPA 每年提供 300 万美元，直到 20 世纪 70 年代为止。ARPA 还对 Newell 和 Simon 在卡内基梅隆大学的工作组以及斯坦福大学 AI 项目（由 John McCarthy 于 1963 年创建）进行类似的资助。另一个重要的 AI 实验室于 1965 年由 Donald Michie 在爱丁堡大学建立。在接下来的许多年间，这 4 个研究机构一直是 AI 学术界的研究（和经费）中心。

经费几乎是无条件地提供的：时任 ARPA 主任的 J. C. R. Licklider 相信他的组织应该 “是资助人，而不是项目”，并且允许研究者去做任何感兴趣的方向。这导致了 MIT 无约无束的研究氛围及其黑客文化的形成，但是好景不长。

到了 20 世纪 70 年代，AI 开始遭遇批评，随之而来的还有资金上的困难。AI 研究者们对其课题的难度未能做出正确判断：此前的过于乐观使人们期望过高，当承诺无法兑现时，对 AI 的资助就缩减或取消了。同时，由于 Marvin Minsky 对感知器的激烈批评，联结主义（即神经网络）销声匿迹了 10 年。20 世纪 70 年代后期，尽管遭遇了公众的误解，AI 在逻辑编程、常识推理等一些领域还是有所进展。

20 世纪 70 年代初，AI 遭遇了瓶颈。即使是最杰出的 AI 程序也只能解决它们尝试解

决的问题中最简单的一部分，也就是说所有的 AI 程序都只是“玩具”。AI 研究者们遭遇了无法克服的基础性障碍。尽管某些局限后来被成功突破，但许多至今仍无法满意地解决。

当时的计算机有限的内存和处理速度不足以解决任何实际的 AI 问题。例如，Ross Quillian 在自然语言方面的研究结果只能用一个含 20 个单词的词汇表进行演示，因为内存只能容纳这么多。1976 年 Hans Moravec 指出，计算机离智能的要求还差上百万倍。他做了个类比：人工智能需要强大的计算能力，就像飞机需要大功率动力一样，低于一个门限时是无法实现的；但是随着能力的提升，问题逐渐会变得简单。

1972 年 Richard Karp 根据 Stephen Cook 于 1971 年提出的 Cook-Levin 理论证明，许多问题只可能在指数时间内获解（即计算时间与输入规模的幂成正比）。除了那些最简单的情况，这些问题的解决需要近乎无限长的时间。这就意味着 AI 中的许多玩具程序恐怕永远也不会发展为实用的系统。

许多重要的 AI 应用，如机器视觉和自然语言，都需要掌握大量的信息。程序应该知道它在看什么，或者在说些什么。这要求程序对这个世界具有儿童水平的认识。研究者们很快发现这个要求太高了：1970 年没人能够做出如此巨大的数据库，也没人知道一个程序怎样才能学到如此丰富的信息。

证明定理和解决几何问题对计算机而言相对容易，而一些看似简单的任务，如人脸识别或穿过屋子，实现起来却极端困难。这也是 20 世纪 70 年代中期机器视觉和机器人方面进展缓慢的原因。

采取逻辑观点的 AI 研究者们（例如 John McCarthy）发现，如果不对逻辑的结构进行调整，他们就无法对常见的涉及自动规划（planning or default reasoning）的推理进行表达。为解决这一问题，他们发展了新逻辑学［如非单调逻辑（non-monotonic logics）和模态逻辑（modal logics）］。

由于缺乏进展，对 AI 提供资助的机构（如英国政府、DARPA 和 NRC）对无方向的 AI 研究逐渐停止了资助。早在 1966 年 ALPAC（automatic language processing advisory committee，自动语言处理顾问委员会）的报告中就有批评机器翻译进展的意味，预示了这一局面的来临。NRC（national research council，美国国家科学委员会）在拨款 2 000 万美元后停止了资助。1973 年 Lighthill 针对英国 AI 研究状况的报告批评了 AI 在实现其“宏伟目标”上的完全失败，并导致了英国 AI 研究的低潮（该报告特别提到了指数爆炸问题，以此作为 AI 失败的一个原因）。DARPA 则对 CMU 的语音理解研究项目深感失望，从而取消了每年 300 万美元的资助。到了 1974 年已经很难再找到对 AI 项目的资助机构。

Hans Moravec 将批评归咎于他的同行们不切实际的预言：“许多研究者落进了一张日益浮夸的网中。”还有一点，自从 1969 年 Mansfield 修正案通过后，DARPA 被迫只资助“具有明确任务方向的研究，而不是无方向的基础研究”。20 世纪 60 年代那种对自由探索的资助一去不复返；此后资金只提供给目标明确的特定项目，比如自动坦克，或者战役管理系统。

一些哲学家强烈反对 AI 研究者的主张。其中最早的一个是 John Lucas，他认为哥德尔不完备定理已经证明形式系统（例如计算机程序）不可能判断某些陈述的真理性，但是人类可以。Hubert Dreyfus 讽刺 20 世纪 60 年代 AI 界那些未实现的预言，并且批评 AI 的基础假设，认为人类推理实际上仅涉及少量“符号处理”，而大多是具体的、直觉的、下意识的“窍门（know how）”。John Searle 于 1980 年提出“中文房间”实验，试图证明程序并不“理解”它所使用的符号，即所谓的“意向性（intentionality）”问题。Searle 认为，如果符号对于机器而言没有意义，那么就不能认为机器是在“思考”。

AI 研究者们并不太把这些批评当回事，因为它们似乎有些离题，而计算复杂性和“让程序具有常识”等问题则显得更加紧迫和严重。对于实际的计算机程序而言，“常识”和“意向性”的区别并不明显。Minsky 提到 Dreyfus 和 Searle 时说，“他们误解了，所以应该忽略”。在 MIT 任教的 Dreyfus 遭到了 AI 阵营的冷遇：他后来说，AI 研究者们“生怕被人看到在和我一起吃中饭”。ELIZA 程序的作者 Joseph Weizenbaum 感到他的同事们对待 Dreyfus 的态度不太专业，而且有些孩子气。虽然他直言不讳地反对 Dreyfus 的论点，但他“清楚地表明了他们待人的方式不对”。

Weizenbaum 后来开始思考 AI 相关的伦理问题，起因是 Kenneth Colby 开发了一个模仿医师的聊天机器人 DOCTOR，并用它当作真正的医疗工具。二人发生争执；虽然 Colby 认为 Weizenbaum 对他的程序没有贡献，但这于事无补。1976 年 Weizenbaum 出版著作《计算机的力量与人类的推理》，书中表示人工智能的滥用可能损害人类生命的价值。

感知器是神经网络的一种形式，由 Frank Rosenblatt 于 1958 年提出。与多数 AI 研究者一样，他对这一发明的潜力非常乐观，预言说“感知器最终将能够学习、作出决策和翻译语言”。整个 20 世纪 60 年代里这一方向的研究工作都很活跃。

1969 年 Minsky 和 Papert 出版了著作《感知器》，书中暗示感知器具有严重局限，而 Frank Rosenblatt 的预言过于夸张。这本书的影响是破坏性的：联结主义的研究因此停滞了 10 年。后来新一代研究者使这一领域获得重生，并使其成为人工智能中的重要部分；遗憾的是 Rosenblatt 没能看到这些，他在《感知器》问世后不久即因游船事故去世。

早在 1958 年，John McCarthy 就提出了名为“纳谏者（advice taker）”的一个程序构想，将逻辑学引入了 AI 研究界。1963 年，J. Alan Robinson 发现了在计算机上实现推理的简单方法：归结（resolution）与合一（unification）算法。然而，根据 20 世纪 60 年代末 McCarthy 和他的学生们的工作，这一想法的直接实现具有极高的计算复杂度：即使是证明很简单的定理也需要如天文数字般的步骤。20 世纪 70 年代 Robert Kowalsky 在爱丁堡大学的工作则更具成效：法国学者 Alain Colmerauer 和 Phillipe Roussel 在与他的合作中开发出成功的逻辑编程语言“Prolog”。

Dreyfus 等人针对逻辑方法的批评观点认为，人类在解决问题时并没有使用逻辑运算。心理学家 PeterWason、Eleanor Rosch、阿摩司·特沃斯基、Daniel Kahneman 等人的实验证明了这一点。McCarthy 则回应说，人类怎么思考是无关紧要的：我真正想要的是解题机器，而不是模仿人类进行思考的机器。

对 McCarthy 的做法持批评意见的还有他在 MIT 的同行们。Marvin Minsky、Seymour Papert 和 Roger Schank 等试图让机器像人一样思考，使之能够解决“理解故事”和“目标识别”一类问题。为了使用“椅子”“饭店”之类最基本的概念，他们需要让机器像人一样做出一些非逻辑的假设。不幸的是，这些不精确的概念难以用逻辑进行表达。Gerald Sussman 注意到，“使用精确的语言描述本质上不精确的概念，并不能使它们变得精确起来”。Schank 用“芜杂（scruffy）”一词描述他们这一“反逻辑”的方法，与 McCarthy、Kowalski、Feigenbaum、Newell 和 Simon 等人的“简约（neat）”方案相对。

在 1975 年的一篇开创性论文中，Minsky 注意到与他共事的“芜杂派”研究者在使用同一类型的工具，即用一个框架囊括所有相关的常识性假设。例如，当我们使用“鸟”这一概念时，脑中会立即浮现出一系列相关事实，如会飞、吃虫子等。我们知道这些假设并不一定正确，使用这些事实的推理也未必符合逻辑，但是这一系列假设组成的结构正是我们所想和所说的一部分。他把这个结构称为“框架（frames）”。Schank 使用了“框架”的一个变种，他称之为“脚本（scripts）”，基于这一想法他使程序能够回答关于一篇英语短文的提问。多年之后的面向对象编程即采纳了 AI“框架”研究中的“继承（inheritance）”概念。

在 20 世纪 80 年代，一类名为“专家系统”的 AI 程序开始为全世界的公司所采纳，而“知识处理”成了主流 AI 研究的焦点。日本政府在同一年代积极投资 AI 以促进其第五代计算机工程。20 世纪 80 年代早期另一个令人振奋的事件是 John Hopfield 和 David Rumelhart 使联结主义重获新生。AI 再一次获得了成功。

专家系统是一种程序，能够依据一组从专门知识中推演出的逻辑规则在某一特定领域回答或解决问题。最早的示例由 Edward Feigenbaum 和他的学生们开发。1965 年设计的 Dendral 能够根据分光计读数分辨混合物。1972 年设计的 MYCIN 能够诊断血液传染病。它们展示了这一方法的威力。

专家系统仅限于一个很小的知识领域，从而避免了常识问题；其简单的设计又使它能够较为容易地编程实现或修改。总之，实践证明了这类程序的实用性。直到此时 AI 才开始变得实用起来。

1980 年 CMU 为 DEC（digital equipment corporation，数字设备公司）设计了一个名为 XCON 的专家系统，这是一个巨大的成功。在 1986 年之前，它每年为公司省下 4 000 万美元。全世界的公司都开始研发和应用专家系统，到 1985 年它们已在 AI 上投入 10 亿美元以上，大部分用于公司内设的 AI 部门。为之提供支持的产业应运而生，其中包括 Symbolics、Lisp Machines 等硬件公司和 IntelliCorp、Aion 等软件公司。

专家系统的能力来自于它们存储的专业知识。这是 20 世纪 70 年代以来 AI 研究的一个新方向。Pamela McCorduck 在书中写道：“不情愿的 AI 研究者们开始怀疑，因为它违背了科学研究中对最简化的追求。智能可能需要建立在对分门别类的大量知识的多种处理方法之上。”“20 世纪 70 年代的教训是智能行为与知识处理关系非常密切。有时还需要在特定任务领域非常细致的知识。”知识库系统和知识工程成为了 20 世纪 80 年代 AI 研究的

主要方向。

第一个试图解决常识问题的程序 Cyc 也在 20 世纪 80 年代出现，其方法是建立一个容纳一个普通人知道的所有常识的巨型数据库。发起和领导这一项目的 Douglas Lenat 认为别无捷径，让机器理解人类概念的唯一方法是一个一个地教会它们。这一工程几十年也没有完成。

1981 年，日本经济产业省拨款 85 000 万美元支持第五代计算机项目。其目标是造出能够与人对话、翻译语言、解释图像，并且像人一样推理的机器。令“芜杂派”不满的是，他们选用 Prolog 作为该项目的主要编程语言。

其他国家纷纷做出响应。英国开始了耗资 35 000 万英镑的 Alvey 工程。美国一个企业协会组织了 MCC（microelectronics and computer technology corporation，微电子与计算机技术集团），向 AI 和信息技术的大规模项目提供资助。DARPA 也行动起来，组织了战略计算促进会（strategic computing initiative），其 1988 年向 AI 的投资是 1984 年的 3 倍。

1982 年，物理学家 John Hopfield 证明了一种新型的神经网络（现被称为“Hopfield 网络”）能够用一种全新的方式学习和处理信息。大约在同时（早于 Paul Werbos），David Rumelhart 推广了“反传法（backpropagation）”，一种神经网络训练方法。这些发现使 1970 年以来一直遭人遗弃的联结主义重获新生。

1986 年由 Rumelhart 和心理学家 James McClelland 主编的两卷本论文集《分布式并行处理》问世，这一新领域从此得到了统一和促进。20 世纪 90 年代神经网络获得了商业上的成功，它们被应用于光字符识别和语音识别软件。

20 世纪 80 年代中商业机构对 AI 的追捧与冷落符合经济泡沫的经典模式，泡沫的破裂也在政府机构和投资者对 AI 的观察之中。尽管遇到各种批评，这一领域仍在不断前进。来自机器人学这一相关研究领域的 Rodney Brooks 和 Hans Moravec 提出了一种全新的人工智能方案。

“AI 之冬（AI winter）”一词由经历过 1974 年经费削减的研究者们创造出来。他们注意到了人们对专家系统的狂热追捧，预计不久后人们将转向失望。事实被他们不幸言中：从 20 世纪 80 年代末到 90 年代初，AI 遭遇了一系列财政问题。

变天的最早征兆是 1987 年 AI 硬件市场需求的突然下跌。Apple 和 IBM 生产的台式机性能不断提升，到 1987 年时其性能已经超过了 Symbolics 和其他厂家生产的昂贵的 Lisp 机。老产品失去了存在的理由：一夜之间这个价值 5 亿美元的产业土崩瓦解。

XCON 等最初大获成功的专家系统维护费用居高不下。它们难以升级、难以使用、脆弱（当输入异常时会出现莫名其妙的错误），成了以前已经暴露的各种各样的问题［例如资格问题（qualification problem）］的牺牲品。专家系统的实用性仅仅局限于某些特定情景。

到了 20 世纪 80 年代晚期，战略计算促进会大幅削减对 AI 的资助。DARPA 的新任领导认为 AI 并非“下一个浪潮”，拨款将倾向于那些看起来更容易出成果的项目。

1991 年人们发现 10 年前日本人宏伟的“第五代工程”并没有实现。事实上其中一些目标，比如“与人展开交谈”，直到 2010 年也没有实现。与其他 AI 项目一样，期望比真