**人工智能的历史**

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，在达特茅斯学院举行的一次会议上正式确立了人工智能的研究领域。会议的参加者在接下来的数十年间是AI研究的领军人物。他们中有许多人预言，经过一代人的努力，与人类具有同等智能水平的机器将会出现。同时，上千万美元被投入到AI研究中，以期实现这一目标。

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。七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的前景作出异常乐观的预测。

尽管在政府官僚和风投资本家那里经历了大起大落，AI领域仍在取得进展。某些在20世纪70年代被认为不可能解决的问题今天已经获得了圆满解决并已成功应用在商业产品上。与第一代AI研究人员的乐观估计不同，具有与人类同等智能水平的机器至今仍未出现。图灵在1950年发表的一篇催生现代智能机器研究的著名论文中称，“我们只能看到眼前的一小段距离……但是，我们可以看到仍有许多工作要做”。

**目录**

**1 先驱**

1.1 神话，幻想和预言中的AI  
1.2 自动人偶  
1.3 形式推理  
1.4 计算机科学

**2 人工智能的诞生：1943 – 1956**

2.1 控制论与早期神经网络  
2.2 游戏AI  
2.3 图灵测试  
2.4 符号推理与“逻辑理论家”程序  
2.5 1956年达特茅斯会议：AI的诞生

**3 黄金年代：1956 – 1974**

3.1 研究工作  
3.1.1 搜索式推理  
3.1.2 自然语言  
3.1.3 微世界  
3.2 乐观思潮  
3.3 经费

**4 第一次AI低谷：1974 – 1980**

4.1 问题  
4.2 停止拨款  
4.3 来自大学的批评  
4.4 感知器与联结主义遭到冷落  
4.5 “简约派（the neats）”：逻辑，Prolog语言和专家系统  
4.6 “芜杂派（the scruffies）”：框架和脚本

**5 繁荣：1980 – 1987**

5.1 专家系统获得赏识  
5.2 知识革命  
5.3 重获拨款：第五代工程  
5.4 联结主义的重生

**6 第二次AI低谷：1987 – 1993**

6.1 AI之冬  
6.2 躯体的重要性：Nouvelle AI与嵌入式推理

**7 AI：1993 – 现在**

7.1 里程碑和摩尔定律  
7.2 智能代理  
7.3 “简约派”的胜利  
7.4 幕后的AI  
7.5 HAL 9000在哪里?

**先驱**

McCorduck写道：“某种形式上的人工智能是一个遍布于西方知识分子历史的观点，是一个急需被实现的梦想，”先民对人工智能的追求表现在诸多神话，传说，故事，预言以及制作机器人偶（automaton）的实践之中。

**神话，幻想和预言中的AI**

希腊神话中已经出现了机械人和人造人,如赫淮斯托斯的黄金机器人和皮格马利翁的伽拉忒亚。中世纪出现了使用巫术或炼金术将意识赋予无生命物质的传说,如贾比尔的Takwin,帕拉塞尔苏斯的何蒙库鲁兹和Judah Loew的魔像。19世纪的幻想小说中出现了人造人和会思考的机器之类题材，例如玛丽·雪莱的《弗兰肯斯坦》和卡雷尔·恰佩克的《罗素姆的万能机器人》。Samuel Butler的《机器中的达尔文（Darwin among the Machines）》一文（1863）探讨了机器通过自然选择进化出智能的可能性。至今人工智能仍然是科幻小说的重要元素。

**自动人偶**



加扎利的可编程自动人偶(1206年)

许多文明中都有创造自动人偶的杰出工匠，例如偃师（中国西周），希罗（希腊），加扎利和Wolfgang von Kempelen 等等。已知最古老的“机器人”是古埃及和古希腊的圣像，忠实的信徒认为工匠为这些神像赋予了思想，使它们具有智慧和激情。赫耳墨斯·特里斯墨吉斯忒斯（Hermes Trismegistus）写道“当发现神的本性时，人就能够重现他”。

**形式推理**

人工智能的基本假设是人类的思考过程可以机械化。对于机械化推理（即所谓“形式推理（formal reasoning）”）的研究已有很长历史。中国，印度和希腊哲学家均已在公元前的第一个千年里提出了形式推理的结构化方法。他们的想法为后世的哲学家所继承和发展，其中著名的有亚里士多德（对三段论逻辑进行了形式分析），欧几里得（其著作《几何原本》是形式推理的典范），花剌子密（代数学的先驱，“algorithm”一词由他的名字演变而来）以及一些欧洲经院哲学家，如奥卡姆的威廉和邓斯·司各脱。

马略卡哲学家拉蒙·柳利（1232-1315）开发了一些“逻辑机”，试图通过逻辑方法获取知识。 柳利的机器能够将基本的，无可否认的真理通过机械手段用简单的逻辑操作进行组合，以求生成所有可能的知识。Llull的工作对莱布尼兹产生了很大影响，后者进一步发展了他的思想。



莱布尼兹猜测人类的思想可以简化为机械计算

在17世纪中，莱布尼兹，托马斯·霍布斯和笛卡儿尝试将理性的思考系统化为代数学或几何学那样的体系。霍布斯在其著作《利维坦》中有一句名言：“推理就是计算（reason is nothing but reckoning）。” 莱布尼兹设想了一种用于推理的普适语言（他的通用表意文字），能将推理规约为计算，从而使“哲学家之间，就像会计师之间一样，不再需要争辩。他们只需拿出铅笔放在石板上，然后向对方说（如果想要的话，可以请一位朋友作为证人）：‘我们开始算吧。’” 这些哲学家已经开始明确提出形式符号系统的假设，而这一假设将成为AI研究的指导思想。

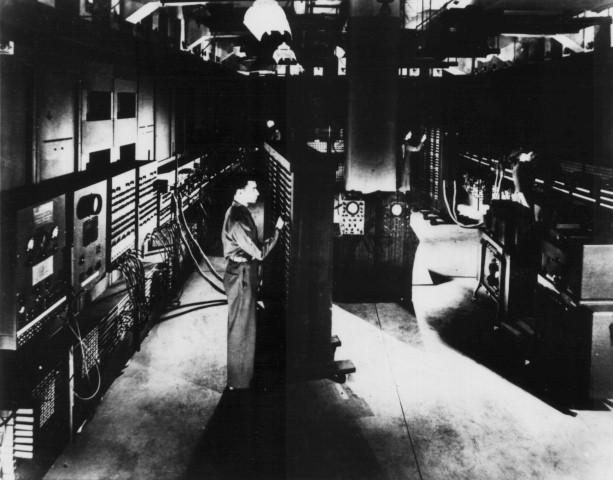
在20世纪，数理逻辑研究上的突破使得人工智能好像呼之欲出。这方面的基础著作包括布尔的《思维的定律》与弗雷格的《概念文字》。基于弗雷格的系统，罗素和怀特海在他们于1913年出版的巨著《数学原理》中对数学的基础给出了形式化描述。这一成就激励了希尔伯特，后者向20世纪20年代和30年代的数学家提出了一个基础性的难题：“能否将所有的数学推理形式化?” 这个问题的最终回答由哥德尔不完备定理，图灵机和Alonzo Church的λ演算给出。他们的答案令人震惊：首先，他们证明了数理逻辑的局限性；其次（这一点对AI更重要），他们的工作隐含了任何形式的数学推理都能在这些限制之下机械化的可能性。

邱奇-图灵论题暗示，一台仅能处理0和1这样简单二元符号的机械设备能够模拟任意数学推理过程。这里最关键的灵感是图灵机：这一看似简单的理论构造抓住了抽象符号处理的本质。这一创造激发科学家们探讨让机器思考的可能。

**计算机科学**

用于计算的机器古已有之；历史上许多数学家对其作出了改进。19世纪初，查尔斯·巴贝奇设计了一台可编程计算机（“分析机”），但未能建造出来。爱达·勒芙蕾丝预言，这台机器“将创作出无限复杂，无限宽广的精妙的科学乐章”。（她常被认为是第一个程序员，因为她留下的一些笔记完整地描述了使用这一机器计算伯努利数的方法。）

第一批现代计算机是二战期间建造的大型译码机（包括Z3，ENIAC和Colossus等）。后两个机器的理论基础是图灵和约翰·冯·诺伊曼提出和发展的学说。



在摩尔学校的电气工程的ENIAC计算机.

**人工智能的诞生：1943 – 1956**

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。

**控制论与早期神经网络**

最初的人工智能研究是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。



IBM 702:第一代AI研究者使用的电脑.

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。在接下来的五十年中，闵斯基是AI领域最重要的领导者和创新者之一。

**游戏AI**

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在五十年代中期和六十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。

**图灵测试**

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。

**符号推理与“逻辑理论家”程序**

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：AI的诞生**

1956年达特矛斯会议的组织者是Marvin Minsky，约翰·麦卡锡和另两位资深科学家Claude Shannon以及Nathan Rochester，后者来自IBM。会议提出的断言之一是“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研究的第一个十年中作出重要贡献。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年达特矛斯会议上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。

**黄金年代：1956 – 1974**

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二十年内出现。 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。

**研究工作**

从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。

**搜索式推理**

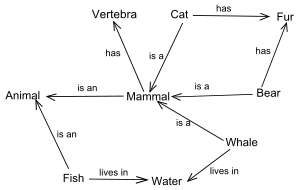
许多AI程序使用相同的基本算法。为实现一个目标（例如赢得游戏或证明定理），它们一步步地前进，就像在迷宫中寻找出路一般；如果遇到了死胡同则进行回溯。这就是“搜索式推理”。

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作出决策，如斯坦福大学为控制机器人Shakey而开发的STRIPS系统。

**自然语言**

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。



一个语义网的例子

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；[54] 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。

**微世界**

60年代后期，麻省理工大学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作出了创造性贡献。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，还能作出决策并执行操作。

**乐观思潮**

第一代AI研究者们曾作出了如下预言:

1958年，H. A. Simon，Allen Newell：“十年之内，数字计算机将成为国际象棋世界冠军。” “十年之内，数字计算机将发现并证明一个重要的数学定理。”

1965年，H. A. Simon：“二十年内，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：“在三到八年的时间里我们将得到一台具有人类平均智能的机器。”

**经费**

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）获得了二百二十万美元经费，用于资助MAC工程，其中包括Minsky和McCarthy五年前建立的AI研究组。此后ARPA每年提供三百万美元，直到七十年代为止。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。[65]在接下来的许多年间，这四个研究机构一直是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“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hacker文化的形成，但是好景不长。

**第一次AI低谷：1974 – 1980**

到了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作出正确判断：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）销声匿迹了十年。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。

**问题**

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。

**计算机的运算能力。**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二十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。

**计算复杂性和指数爆炸。**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。

**常识与推理**。许多重要的AI应用，例如机器视觉和自然语言，都需要大量对世界的认识信息。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。

**莫拉维克悖论**。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70年代中期机器视觉和机器人方面进展缓慢的原因。

**框架和资格问题**。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））。

**停止拨款**

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）在拨款二千万美元后停止资助。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，从而取消了每年三百万美元的资助。到了1974年已经很难再找到对AI项目的资助。

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”。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。

**来自大学的批评**

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讽刺六十年代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。

**感知器与联结主义遭到冷落**

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”。整个六十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十年。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。

**“简约派（the neats）”：逻辑，Prolog语言和专家系统**

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60年代末McCarthy和他的学生们的工作，对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天文数字的步骤。70年代Robert Kowalsky在Edinburgh大学的工作则更具成效：法国学者Alain Colmerauer和Phillipe Roussel在他的合作下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Peter Wason，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：真正想要的是解题机器，而不是模仿人类进行思考的机器。

**“芜杂派（the scruffies）”：框架和脚本**

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作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采纳了AI“框架”研究中的“继承（inheritance）”概念。

**繁荣：1980 – 1987**

在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”成为了主流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。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。

**专家系统获得赏识**

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年起设计的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。直到现在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，它每年为公司省下四千万美元。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上投入十亿美元以上，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。

**知识革命**

专家系统的能力来自于它们存储的专业知识。这是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。” “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。

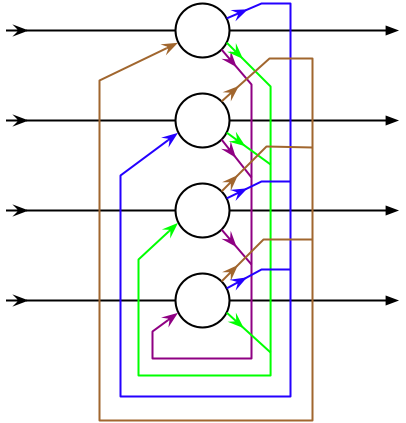
**重获拨款：第五代工程**

1981年，日本经济产业省拨款八亿五千万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。

其他国家纷纷作出响应。英国开始了耗资三亿五千万英镑的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年的三倍。

**联结主义的重生**

1982年，物理学家John Hopfield证明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（en: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。



一个四节点的Hopfield网络.

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。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。

**第二次AI低谷：1987 – 1993**

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之冬（en: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：一夜之间这个价值五亿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（en: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年人们发现十年前日本人宏伟的“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，期望比真正可能实现的要高得多。

**躯体的重要性：Nouvelle AI与嵌入式推理**

80年代后期，一些研究者根据机器人学的成就提出了一种全新的人工智能方案。他们相信，为了获得真正的智能，机器必须具有躯体 – 它需要感知，移动，生存，与这个世界交互。他们认为这些感知运动技能对于常识推理等高层次技能是至关重要的，而抽象推理不过是人类最不重要，也最无趣的技能（参见Moravec悖论）。他们号召“自底向上”地创造智能，这一主张复兴了从60年代就沉寂下来的控制论。

另一位先驱是在理论神经科学上造诣深厚的David Marr，他于70年代来到MIT指导视觉研究组的工作。他排斥所有符号化方法（不论是McCarthy的逻辑学还是Minsky的框架），认为实现AI需要自底向上地理解视觉的物理机制，而符号处理应在此之后进行。

在发表于1990年的论文“大象不玩象棋（Elephants Don’t Play Chess）”中，机器人研究者Rodney Brooks提出了“物理符号系统假设”，认为符号是可有可无的，因为“这个世界就是描述它自己最好的模型。它总是最新的。它总是包括了需要研究的所有细节。诀窍在于正确地，足够频繁地感知它。” 在80年代和90年代也有许多认知科学家反对基于符号处理的智能模型，认为身体是推理的必要条件，这一理论被称为“具身的心灵/理性/ 认知（embodied mind/reason/cognition）”论题。

**AI：1993 – 现在**

现已年过半百的AI终于实现了它最初的一些目标。它已被成功地用在技术产业中，不过有时是在幕后。这些成就有的归功于计算机性能的提升，有的则是在高尚的科学责任感驱使下对特定的课题不断追求而获得的。不过，至少在商业领域里AI的声誉已经不如往昔了。“实现人类水平的智能”这一最初的梦想曾在60年代令全世界的想象力为之着迷，其失败的原因至今仍众说纷纭。各种因素的合力将AI拆分为各自为战的几个子领域，有时候它们甚至会用新名词来掩饰“人工智能”这块被玷污的金字招牌。AI比以往的任何时候都更加谨慎，却也更加成功。

**里程碑和摩尔定律**

1997年5月11日，深蓝成为战胜国际象棋世界冠军卡斯帕罗夫的第一个计算机系统。2005年，Stanford开发的一台机器人在一条沙漠小径上成功地自动行驶了131英里，赢得了DARPA挑战大赛头奖。2009年，蓝脑计划声称已经成功地模拟了部分鼠脑。

这些成就的取得并不是因为范式上的革命。它们仍然是工程技术的复杂应用，但是计算机性能已经今非昔比了。事实上，深蓝计算机比Christopher Strachey在1951年用来下棋的Ferranti Mark 1快一千万倍。这种剧烈增长可以用摩尔定律描述：计算速度和内存容量每两年翻一番。计算性能上的基础性障碍已被逐渐克服。

**智能代理**

90年代，被称为“智能代理（en:intelligent agents）”的新范式被广泛接受。[132]尽管早期研究者提出了模块化的分治策略， 但是直到Judea Pearl，Alan Newell等人将一些概念从决策理论和经济学中引入AI之后现代智能代理范式才逐渐形成。当经济学中的“理性代理（rational agent）”与计算机科学中的“对象”或“模块”相结合，“智能代理”范式就完善了。

智能代理是一个系统，它感知周围环境，然后采取措施使成功的几率最大化。最简单的智能代理是解决特定问题的程序。已知的最复杂的智能代理是理性的，会思考的人类。智能代理范式将AI研究定义为“对智能代理的学习”。这是对早期一些定义的推广：它超越了研究人类智能的范畴，涵盖了对所有种类的智能的研究。

这一范式让研究者们通过学习孤立的问题找到可证的并且有用的解答。它为AI各领域乃至经济学，控制论等使用抽象代理概念的领域提供了描述问题和共享解答的一种通用语言。人们希望能找到一种完整的代理架构（像Newell的en:SOAR那样），允许研究者们应用交互的智能代理建立起通用的智能系统。

**“简约派”的胜利**

越来越多的AI研究者们开始开发和使用复杂的数学工具。人们广泛地认识到，许多AI需要解决的问题已经成为数学，经济学和运筹学领域的研究课题。数学语言的共享不仅使AI可以与其他学科展开更高层次的合作，而且使研究结果更易于评估和证明。AI已成为一门更严格的科学分支。 Russell和Norvig（2003）将这些变化视为一场“革命”和“简约派的胜利”。

Judea Pearl发表于1988年的名著将概率论和决策理论引入AI。现已投入应用的新工具包括贝叶斯网络，隐马尔可夫模型，信息论，随机模型和经典优化理论。针对神经网络和进化算法等“计算智能”范式的精确数学描述也被发展出来。

**幕后的AI**

AI研究者们开发的算法开始变为较大的系统的一部分。AI曾经解决了大量的难题，这些解决方案在产业界起到了重要作用。应用了AI技术的有数据挖掘，工业机器人，物流，语音识别，银行业软件，医疗诊断和Google搜索引擎等。

AI领域并未从这些成就之中获得多少益处。AI的许多伟大创新仅被看作计算机科学工具箱中的一件工具。Nick Bostrom解释说，“很多AI的前沿成就已被应用在一般的程序中，不过通常没有被称为AI。这是因为，一旦变得足够有用和普遍，它就不再被称为AI了。”

90年代的许多AI研究者故意用其他一些名字称呼他们的工作，例如信息学，知识系统，认知系统或计算智能。部分原因是他们认为他们的领域与AI存在根本的不同，不过新名字也有利于获取经费。至少在商业领域，导致AI之冬的那些未能兑现的承诺仍然困扰着AI研究，正如New York Times在2005年的一篇报道所说：“计算机科学家和软件工程师们避免使用人工智能一词，因为怕被认为是在说梦话。”

**HAL 9000在哪里?**

1968年亚瑟·克拉克和史丹利·库柏力克创作的《“2001太空漫游”》中设想2001年将会出现达到或超过人类智能的机器。他们创造的这一名为HAL-9000的角色是以科学事实为依据的：当时许多顶极AI研究者相信到2001年这样的机器会出现。

“那么问题是，为什么在2001年我们并未拥有HAL呢?” Marvin Minsky问道。 Minsky认为，问题的答案是绝大多数研究者醉心于钻研神经网络和遗传算法之类商业应用，而忽略了常识推理等核心问题。另一方面，John McCarthy则归咎于资格问题（en:qualification problem）。Ray Kurzweil相信问题在于计算机性能，根据摩尔定律，他预测具有人类智能水平的机器将在2029年出现。Jeff Hawkins认为神经网络研究忽略了人类大脑皮质的关键特性，而简单的模型只能用于解决简单的问题。还有许多别的解释，每一个都对应着一个正在进行的研究计划。目前以自然语言理解问题为突破口，以本源语义为对象，通过对“理解”与“智能”的界定研究，人类级别的人工智能研究已经取得进展。

End.