

第 1 章

人工智能简史

人工智能的研究源于电子计算机的出现。

电子计算机出现了，这给人们带来很大影响。计算快和记忆力强一向被认为是人类高智商的表现，而像计算机这样的机器居然可以在计算和记忆上有如此好的表现，人们不禁思考：这样一台机器，除了计算和记忆，是否还可以完成其他的“智能”任务？艾伦·图灵对此有过深入思考。他在 1950 年发表了论文《计算机器与智能》和《机器能思考吗？》。在此之前有一些关于人工智能思想的论文，但图灵是第一个从数字计算机角度讨论智能问题的，因此《计算机器与智能》被认为是人工智能方面的第一篇论文。

艾伦·图灵（Alan Turing，1912—1954），英国数学家、逻辑学家、计算机科学家，被称为计算机科学之父、人工智能之父。

他提出的图灵机模型为现代计算机的研究奠定了基础；他是第一个从数字计算机角度讨论智能问题的人；他提出的图灵测试引导了人工智能的发展。

在学术界，计算机领域最高奖以他的名字命名：图灵奖。



艾伦·图灵

1.1 里程碑式的事件：达特茅斯会议

计算机出现之后，一些科学家就开始研究如何让计算机完成智能任务，例如：下棋、证明定理、识别图形。1956 年夏天，约翰·麦卡锡等人在美国

达特茅斯学院组织了一个为期两个月的学术讨论班。来自不同领域的十位科学家在这里讨论了一些各自关心和正在研究的问题。

组织这样的研讨会，是学术界的一种传统。这样，位于不同地区但具有相同兴趣的研究人员可以在一起讨论、交流。直至现在，人们仍然沿用这种做法，组织各类学术会议，希望通过交流和讨论，获得灵感，学习知识，促进合作。

这个学术研讨会把当时做人工智能相关研究的一些人召集在一起进行研讨。会议组织者认为应该用充足的时间来进行深入的讨论，因此就把时间安排在暑期，这样可以利用两个月的时间专心开会，而不被学校的工作打扰。在这期间，人们非常自由、充分、深入地开展了报告、交流、研讨。

参加讨论班的 10 位科学家是：约翰·麦卡锡（John McCarthy）、马文·明斯基（Marvin Minsky）、克劳德·香农（Claude Shannon）、艾伦·纽厄尔（Allen Newell）、赫伯特·西蒙（Herbert Simon）、雷·索洛莫诺夫（Ray Solomonoff）、纳撒尼尔·罗切斯特（Nathaniel Rochester）、特伦查德·莫尔（Trenchard More）、奥利弗·塞尔弗里奇（Oliver Selfridge）、阿瑟·塞缪尔（Arthur Samuel）。其中，马文·明斯基于 1969 年获得图灵奖；约翰·麦卡锡于 1971 年获得图灵奖；艾伦·纽厄尔和赫伯特·西蒙于 1975 年获得图灵奖。图 1-1 是参会的 10 位科学家的照片。

他们认为，他们正在从事的是一项全新的研究工作，并给这些研究起了一个响亮的名字“人工智能”（artificial intelligence, AI）。

达特茅斯会议被认为是人工智能的起始点，1956 年就是人工智能元年。这个时期，产生了一些人工智能方面的成果。例如：

约翰·麦卡锡发明了 LISP（list processing）语言，这是非常适合人工智能研究的计算机编程语言。LISP 语言曾经发挥过很大的作用，现在仍然有研究人员在使用这种语言。

IBM 的阿瑟·塞缪尔开发了跳棋程序。人们认为下棋是一个很典型的智能任务。他们研发跳棋程序时，使用了很多技术，其中包括在人工智能研究中有重要影响的启发式搜索技术（详见第 2 章）。注意到棋盘上不同位置的重要性不同，他们给棋盘的不同位置赋予不同的重要性权重，再利用启发式搜



约翰·麦卡锡



马文·明斯基



克劳德·香农



雷·索洛莫诺夫



艾伦·纽厄尔



赫伯特·西蒙



阿瑟·塞缪尔



奥利弗·塞尔弗里奇



纳撒尼尔·罗切斯特



特伦查德·莫尔

图 1-1 达特茅斯会议参会人员（1956 年）

索方法确定走棋路径。他们的跳棋程序在 1962 年曾经和美国的州跳棋冠军比赛并获胜。

艾伦·纽厄尔、赫伯特·西蒙和克里夫·肖（Cliff Shaw）设计和编写了证明数学定理的逻辑理论机（在中文中，“机”通常指物理的机械设备。但英文不一样，英文中一个计算机程序也被称为一台机器（machine））。这个程序可以证明罗素的《数学原理》第二章中的 52 个定理中的 38 个。后来经过改进，其程序可以证明全部的 52 个定理。



“

完全没想到，1956 年程序就可以下棋和证明定理了。

”

“

是。可以想象当时人们激动的心情。当然，人工智能只是在近 10 年才被大众广泛了解，大家不知道这些事也是自然的。

”



1.2 繁荣时期：1956 年到 20 世纪 70 年代初

这是人工智能的第一个繁荣时期。这个时期研究人员取得了一系列研究成果，下面列举几个。

(1) 问题求解与搜索。人们希望研发一些技术来解决各种智能问题。搜索技术就是这些研究的一系列成果。

(2) 计算复杂性理论。这是关于算法计算复杂性方面的理论。这些研究成果告诉我们，有些问题的解决非常困难，而有些问题相对容易。

(3) SHRDLU 系统。该系统意为积木世界操纵程序，即让机器人（机械臂）完成一些简单任务。其考虑的任务环境是，桌上有一些大小不同、形状不同的积木。机器人（机械臂）能把一块积木（其上面没有别的积木）抓起来，放到另一个地方（桌子上或者别的积木上面）。这个系统要完成的任务就是根据用户的要求确定机械臂的先后动作。例如，当前桌上有一块圆柱形积木在一块大长方体积木的上方，旁边摆放着一块小长方体积木，见图 1-2（a）。希望积木的最终状况是：小长方体积木在圆柱体上，圆柱形积木在桌子上，见图 1-2（b）。这个系统要完成的任务就是要自动确定先抓取什么，放到什么位置，然后再抓取什么，放到什么位置。在这个问题中，就是要先抓起圆柱形积木放到桌子上，然后再抓起小长方体放到圆柱形积木上。这是对于真实世界问题简化和抽象后的一个场景。在很长一段时间里，积木世界这样一个简单的几何体环境成为人工智能研究采用的场景。这也是研究人员通常采用的方法：先研究简单情况，做初步的尝试和探索。

SHRDLU 系统使用的是模拟的场景，而不是要解决一个物理世界的桌上的积木的摆放问题。这样做是因为不需要考虑构造一个实际的机器人抓取物体等问题，降低了系统实现的难度。

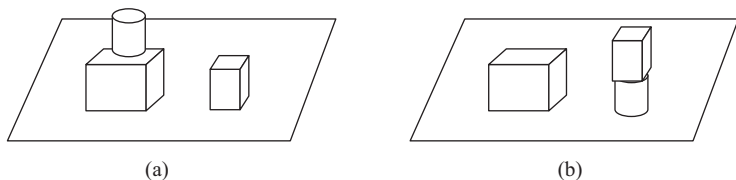


图 1-2 SHRDLU 系统要知道如何移动积木把 (a) 变为 (b)

“这个任务对于人来说易如反掌。但是，让系统自己决定先做什么，后做什么是一大类智能系统要完成的基本任务。这被称作规划（planning）。”



“如果让一个机器人“给我倒一杯水端过来”，就需要机器人自己做好规划：先倒水，再端过来。”

“规划有不同的细致程度。以倒水为例：打开柜子，拿出一个杯子，放到桌子上，倒水。这是一个比较粗略的规划。而对于机器人研究人员来说，就有更细致的规划。以拿起一个杯子为例：确定杯子的位置（坐标），确定杯子的抓取位置，移动机械臂到杯子的抓取位置，抓紧杯子，向上拿起杯子。”



（4）机器人 SHAKEY。和 SHRDLU 系统不同，这是一个可以移动的、真实的机器人。为了让机器人知道周围都有什么物体，它装备有一台摄像机。此外，它还使用激光测距仪来确定与周围物体之间的距离。这个机器人能够自动制定一系列需要执行的步骤从而完成一个给定的任务（这也被称作规划）。为此，研究团队开发了一个用于机器人规划的系统 STRIPS（Stanford Research Institute problem solver，斯坦福研究院问题解决器）。这个系统相当复杂，研究难度很大，需要把感知系统、规划系统、执行系统完整地结合起来，此外还要考虑各种技术和工程问题，是一项了不起的工作。

这个时期，研究人员致力于研发出一套能解决各种智能任务的理论和方法，也就是现在人们所说的通用人工智能研究。毫无疑问，这是非常宏大的理想和目标。但是，当时的计算机硬件水平、软件水平都还比较低，人们对于人工智能的研究深度和广度都还很不够，使得曾经对这一领域盲目乐观的

研究人员的很多承诺和预测没能兑现。

后来，对于人工智能研究的资助逐渐被削减。20 世纪 70 年代初，人工智能没有太多有用的核心研究进展，与此同时，人们对于人工智能的批评很多，这些导致了长达十年的人工智能严冬。不仅如此，甚至在其后的几十年里，人工智能都被一些人视为一门“伪科学”。



“

这是人工智能历史上最难熬的一段时间。
研究人员得不到经费支持，论文无处发表，毕业生找不到工作。

”

“

兑现不了承诺的确会让人失望，更别提一而再，再而三地食言了。

”



1.3 知识就是力量：20 世纪 70 年代末到 80 年代末

这个时期的主要工作是关于知识的研究。研究人员指出，以前研究对于知识的关注不够。获取知识和使用知识应该是人工智能发展的关键。这些思想和观点引导了这个时期的研究，取得了很多成果，繁荣和发展了人工智能的研究。

实际上，对于知识的表示、知识的获取在 20 世纪六七十年代就一直有研究成果不断出现。这些研究成果在这个时期得到重视。这个时期的很多工作是建立由大规模的知识库构成的专家系统。

在人们的生活中，很多专家的工作是智能的集中表现，例如，医生看病、法律咨询等。这些专家掌握了大量系统、深入的专业知识，因而能够解决实际问题。学习这样的知识需要人类花费很长的时间，而具备这样的知识的专家又非常少。因此，研究人员希望计算机系统也能掌握这些知识，从而完成一些智能任务，由此开始了专家系统的研发。



“

听起来挺合理的。
学医一学就是八年，太烧脑了。

”

“

这些专业需要学习大量的专业知识，比较适合作为研究的切入点。

”



专家系统是一个比较大型的软件系统。这个系统中包含大量的人类专业知识。系统利用这些知识解决特定的问题，例如：医学专家系统、法律咨询专家系统、设备维修专家系统等。这个时期，研究人员构建了一些专家系统，并取得了成功。下面介绍其中的几个。

专家系统 MYCIN，这是当时一个著名的医学专家系统。该系统对于人类的血液病的诊断可以提供建议。美国斯坦福大学人工智能专家组和医学院专家组用了大约 5 年时间共同开发完成了该系统。这个系统具有很多重要特点，例如：它可以向用户提问，以得到更多的信息；它可以给出其诊断的过程解释，这样人类专家可以知道其为什么给出这样的诊断结果。1979 年这个系统经过了 10 个实际病例的测试。测试结果表明，在血液病诊断方面，它与人类专家相当。

专家系统 DENDRAL，该系统可帮助化学家根据质谱仪的数据确定化合物的成分和结构。20 世纪 80 年代中期，每天有成百上千人使用这个系统。这表明专家系统可以落地使用。

专家系统 R1/XCON，该系统用于帮助人们配置虚拟地址扩展（VAX）系列计算机。使用该系统的美国数字设备公司（DEC）称，该系统为其公司节省了超过 4000 万美元。这表明专家系统可以商用。



“ 4000 万美元是不小的一笔钱。 ”

“

那还是在 80 年代，而不是几十年后的今天。

”



除了上述这些系统，还有一些其他的成功案例。这些研发工作表明，专家系统可以在解决一些具体的问题时达到人类的水平。这是人工智能取得的成就。

这样，知识以软件系统为载体，让知识从抽象和无形变为具体和有形，可以解决实际问题，可以盈利。这些都导致了相关技术、系统研究和开发的繁荣。直到现在，人们有时还希望构建一个小的专家系统解决一些具体、特殊的实际问题。

由于之前“人工智能”的负面影响，这个时期的工作被称为“基于知识的智能系统”，或“知识工程”，而不是人工智能。



“ 是不想被“人工智能”拖累吧？ ”

“

研究人员知道，以前人工智能的研究失败事出有因。人工智能并不是伪科学。现在探索出了一条新路，这株幼苗不应该被扼杀。

”



专家系统看起来更多的是工程技术。因此，一些研究人员思考，基于知识的专家系统的数学基础是什么？这些思考和研究导致了基于逻辑的人工智能的研究范式的形成，逻辑成为了知识表示的基础。在这样的基础上，基于逻辑的人工智能研究得到了快速发展。到了 20 世纪 80 年代初，它成为了人工智能的研究主流。其影响不仅限于人工智能的研究，还影响了整个计算机领域。由此

出现了逻辑编程 (logic programming) 和 PROLOG (programming in logic) 语言。

PROLOG 语言由鲍勃·科瓦爾斯基 (Bob Kowalski)、阿兰·科尔默劳尔 (Alain Colmerauer) 发明。这是一个以一阶谓词逻辑为基础的程序设计语言。PROLOG 语言功能强大、代码简洁, 虽然最终没有成为一种通用的计算机语言, 但是目前仍然被广泛使用。

这个时期取得的这些成就使得人们再次确认, 知识是不可缺少的。这让道格·莱纳特 (Doug Lenat) 给自己的团队进一步设立了一个宏大的目标: 建立一个包罗万象的知识库 Cyc, 这样就可以实现通用人工智能, 解决各种各样的实际问题。遗憾的是, 在研究中人们发现, 人类的知识非常的丰富、复杂, 特别是常识的获取和表示非常的困难 (见第 6 章)。经过十年的努力, 系统完成了, 但是没有达到预期的效果。人们测试系统, 发现系统中虽然存在大量知识, 但是有些方面的知识非常缺乏, 此外有大量的知识碎片。这一次专家又没有兑现之前的承诺。

20 世纪 80 年代末, 专家系统的繁荣结束了。专家系统的研究以失败而结束。

虽然 Cyc 工程失败了, 但是其观点“通用人工智能的本质是知识体系问



“为什么知识库非要包罗万象、无所不有呢? 研究一个一个的专家系统不是挺好吗?”

“

一个一个地构建专家系统不够有效, 如果能建设一个包罗万象的知识库, 岂不是一劳永逸? 科学家总是希望不断地攀登新高峰。因为不知道技术的极限在哪里, 所以就会不断探索, 当然也会失败。对于科研来说, 这是正常的。



”



“嗯, 更高、更快、更强!”

”

题”并没有被推翻。人们意识到，这个系统使用的人工输入知识的方法无法成功。而在十几年后，谷歌利用互联网上的大量数据建立了规模很大的知识图谱，并利用知识图谱为人们的搜索提供答案。这也是“知识工程”的成功。当前以 ChatGPT 为代表的预训练语言大模型中包含了更多的知识，特别是常识，并能回答各类问题。回顾当时的研究，虽然 Cyc 工程失败了，但是其思想很超前。

1.4 快速发展、繁荣：20 世纪 80 年代末及之后的 20 年

这个时期出现了一系列重要的研究成果。

在专家系统的研究逐渐走下坡路的时候，一些研究人员开始反思，并对基于知识表示和推理的人工智能提出了批评，其代表人物是罗德尼·布鲁克斯（Rodney Brooks）。他们认为，一个人工智能系统必须要感知环境，并且和环境交互。感知环境是非常困难的，也是非常重要的，而以前的研究都回避和绕开了这些问题。知识和推理并不是智能系统的必备条件，而只是智能的具象和表现，以知识和推理作为人工智能核心是错误的。这些研究被称为基于行为的人工智能。他们采用了新的技术并且在一些简单的机器人任务中取得了成功，如智能扫地机器人。

从当前的研究看，ChatGPT 等大模型通过“阅读”大量文本而获得了大量知识，并具备了一些推理能力。从某种意义上讲，这支持了几十年前的研究思想和观点。

目前，经过几十年的发展，人们对于自然智能和人工智能有了更多的了解和认识。人们发现人工智能包含很多方面内容，如同本书的目录中所列的章节题目一样。人们也发现，很多智能任务也只涉及其中的一个或者几个方面，完成一些简单任务甚至不需要对环境有深入的理解。这在后面第 5 章关于自然语言处理的词袋模型部分有所解释。基于行为的人工智能也是研究人员不断探索和研究的一个成果，丰富了人工智能的研究，加深了人们对于智能的认识。

基于行为的人工智能所采用的技术存在很大局限。当要设计的系统的基础功能组件很多时，厘清这些组件之间的关系，并让这些组件能很好地协调和工作就变得非常复杂。此外，构建某个系统的方法很难用于其他系统的构建。人们只能一个系统一个系统地重新设计、构建、测试。