



第 1 章 神奇的人工智能

制造智能机器自古以来就是人类一直追求的梦想,蕴藏在浩如烟海的民间故事、轶事和传说之中。《列子·汤问》中记载了一位名叫偃师的巧匠,他制作的机械人能歌善舞,与真人无异。一天,偃师带着它拜见周穆王,却因机器人朝向穆王的嫔妃们抛媚眼惹得穆王大怒,将它剖开后才确信是机器人,可见其惟妙惟肖。三国时期,传说蜀汉丞相诸葛亮发明了可以自动行走的木牛流马,走山路如履平地,北伐时常用以搬运军粮,立了大功。在中世纪的阿拉伯,博学者加扎利发明了众多玩偶机器人,包括送饮料的女侍者、可演奏音乐的机器人乐团等,受到时人的普遍欢迎。

随着机械制造工艺的进步和电子电气技术的发展,人类发明了很多自动化机器,如各种加工机器、电器设备等。这些设备虽然具有一定程度的自主性,但远未达到“智能”的程度。计算机的出现极大地推动了智能机器的发展,由此催生了一门全新的学科——“人工智能”。自 20 世纪 50 年代诞生以来,人工智能技术飞速发展,从最初的简单搜索和推理,到自动定理证明和专家系统,再到如今各种能听会说、能思会想的智能机器,人工智能技术已经悄无声息地走入我们的日常生活,极大地改变了社会生产方式和人们的生活方式,并为我们提供了无限的想象空间。可以预期的是,未来人工智能技术必然会在各个领域完成一个又一个的创举,成为

人类探索自然、改变自然的强大工具和亲密战友。

1.1 什么是人工智能

首先,让我们来看看“什么是人工智能”。要回答这一问题,需要先了解“什么是智能”。依维基百科的定义,**智能**(Intelligence)是指生物一般性的精神能力。这个能力包括:推理、理解、计划、解决问题、抽象思维、表达意念以及语言和学习的能力。可以将“智能”通俗地理解为“思考的能力”。而**人工智能**(Artificial Intelligence, AI),就是让机器具有这种能力的科学,也就是说让机器像人一样能思会想。如果机器真的具备了这种能力,就可以称为**智能机器**(Intelligence Machine)。

从这种意义上讲,计算器是智能的,因为它可以推理出 $1+1=2$;汽车不是智能的,因为它只会在人的操纵下行走。然而,众所周知汽车不论在功能、复杂程度上都远高于计算器。因此,智能机器未必比非智能机器先进,非智能机器里也有很多智能的元素。例如,对于汽车来说,ABS 系统在汽车制动时,可自动调整制动力的的大小,使车轮不被抱死,这是智能的,但却不容易被大家察觉。因此,智能更多的是功能上的概念,而非对某一设备的评价。事实上,几乎所有的机器都具有一定的智能元素,只不过有些智能元素是局部的、辅助的、不易察觉的,而有些智能元素却是全局的、主体的、表现出明显的智能性。此外,人们对智能的感知具有相对性。自动洗衣机在问世之初无疑是让人新奇的智能机器,但当它普及之后,人们就习以为常了,智能的标签就不存在了。这也说明了“智能”和“智能机器”本身的模糊性,人们总是把新颖的、具有超常能力的机器看作是智能的。值得注意的是,目前学术界和产业界并没有一个关于人工智能的确切定义。人工智能的先驱 John McCarthy(约翰·麦卡锡)在 1955 年曾给出这样的定义:“人工智能是制造智能机器的科学与工程”。维基百科的定义是:“人工智能又称机器智能,是指由人制造出来的机器所表现出来的智能。通常人工智能是指通过普通计算机程序的手段实现的类人智能技术。”虽然没有一个公认的定义,但研究者对人工智能的研究目标是明确的,即为机器赋予人的思想和行为能力。

1.2 人工智能简史

1.2.1 数理逻辑:人工智能的前期积累

人工智能的发展初期以如何刻画人类的智能行为作为研究目标,特别是对知识的表达和推理过程的形式化。换句话说,就是如何将人类的智能行为用计算机模拟出来。事实上,对人类知识结构和推理方法的研究最早可上溯到古希腊哲学

家亚里士多德(Aristotle)的三段论逻辑以及欧几里得(Euclid)的形式推理方法。13世纪,加泰罗尼亚数学家和逻辑学家拉蒙·柳利(Raymundus Lullus)用机械手段模拟简单的逻辑操作,通过演绎运算从旧知识中推理出新知识。17世纪,英国哲学家霍布斯(Hobbes)和数学家莱布尼茨(Leibniz)等进一步提出“推理就是计算”的思路,将逻辑变得可计算化。到了20世纪,在布尔(Boole)、弗雷格(Frege)、希尔伯特(Hilbert)、罗素(Russell)等人的努力下,数理逻辑(Mathematical logic)成为一门独立的学科,标志着逻辑推理形式化的数学理论最终形成。

什么叫逻辑推理形式化呢?可以通过一个例子来简单理解。假设 p 、 q 、 r 分别表示“今天下雨”“我们今天不野餐”“我们明天野餐”,那么“如果今天下雨,那么我们今天将不野餐。”可表示为 $p \rightarrow q$;“如果我们今天不野餐,那么我们明天将野餐。”可以表示为 $q \rightarrow r$ 。通过连续运用推理规则,即可由 $p \rightarrow q$ 和 $q \rightarrow r$ 推理出 $p \rightarrow r$ 。这意味着“如果今天下雨,那么则明天野餐。”在上述过程中,我们将事实表示为符号,将推理表示成符号间的蕴含关系($\rightarrow$),如果再加上一系列限制条件和演算规则,即可得到一套逻辑系统。在这套系统中, p 、 q 、 r 是独立于事实本身的变量,因此该系统描述的不是某一个具体的推理任务,而是一类基于相同逻辑元素和统一推理规则的任务的抽象表示。基于此,推理过程被转化成符号演算,这是数理逻辑的基本思路。

数理逻辑的发展为未来的人工智能大厦奠定了第一块基石。数理逻辑的先驱们认为一切智能活动都可以转化为逻辑过程,因此逻辑过程的可计算意味着人类智能的可计算。希尔伯特(图1-1)甚至曾经设想一个一致完备的逻辑体系,只要基本假设是合理的,就可以通过运算推导出领域内的一切知识。这个大一统的梦想



图1-1 大卫·希尔伯特(德国人,1862—1943年,伟大的数学家)

注:1900年,希尔伯特在巴黎的国际数学家大会上提出了23个问题。这些问题为20世纪的数学研究指明了方向,被称为“希尔伯特问题”。希尔伯特的第2个问题为“算术公理之相容性”。在这一问题中,希尔伯特猜想一个公理系统可以一致、完备地生成所有真值命题。这一猜想在1930年被奥匈帝国数学家库尔特·哥德尔证明为伪。

最终被哥德尔(Godel)著名的不完备定理打破,但数理逻辑的强大描述能力已经深入人心,大大增加了人们制造智能机器的勇气。人们相信,只要逻辑系统设计得足够好,就有望将人类的智能过程通过计算完美地复现,尽管当时计算机还没有出现。

延伸阅读:哥德尔不完备定理

哥德尔于1930年证明,任意一个足够强大的逻辑系统都是不完备的,总有一些定理在该逻辑系统中无法被证明为真,也无法被证明为伪。哥德尔的证明类似“说谎者悖论”:如果有个人说“我说的是假话”,我们是无法判断这句话的真假的。如果这个人说的是真话,那么由“我说的是假话”这句话的意义可推知他(她)实际在说假话,与前提“他在说真话”相互矛盾;反之,如果这个人说的是假话,则“我说的是假话”这句话就不是真的,因此这个人事实上说的是真话,又与前提“他在说假话”相互矛盾。哥德尔证明类似的悖论在任何一个足够强的逻辑系统中都存在,因此任何一个逻辑系统总有它无法理解的命题存在。这说明任何一个系统都有其固有局限性,不同层次的系统局限性各不相同。计算机无法突破其固有局限性,因此模拟人类智能的方法有可能永远无法超过人类。

1.2.2 图灵:人工智能的真正创始人

1936年,年仅24岁的英国科学家图灵(Turing)在他的论文《论可计算数及其在判定问题上的应用》中提出**图灵机**(Turing Machine)模型,证明基于简单的读写操作,图灵机有能力处理非常复杂的计算,包括逻辑演算。1945年6月,美国著名数学家和物理学家约翰·冯·诺伊曼(John von Neumann)等人联名发表了著名的“101页报告”,阐述了计算机设计的基本原则,即著名的**冯·诺伊曼结构**。1946年2月14日,世界上第一台计算机ENIAC在美国宾夕法尼亚大学诞生。1951年,ENIAC的发明者电气工程师约翰·莫奇利(John William Mauchly)和普雷斯波·艾克特(J. Presper Eckert)依据冯·诺伊曼结构对ENIAC进行了升级,即著名的EDVAC计算机。计算机的出现为快速逻辑演算准备好了工具,奠定了人工智能大厦的第二块基石。

在美国人设计ENIAC的同时,图灵也在曼彻斯特大学负责曼彻斯特一号的软件开发工作,并开始关注让计算机执行更多智能性的工作。例如,他主张智能机器不该只复制成人的思维过程,还应该像孩子一样成长学习,这正是机器学习的早期思路;他认为可以通过模仿动物进化的方式获得智能;他还自己编写了一个下棋程序,这可能是最早的机器博弈程序了。为了对人工智能有个明确的评价标准,图灵于1950年提出了著名的**图灵测试**(Turing Test)。在这一测试中,图灵设想将一个人和一台计算机隔离开,通过打字进行交流。如果在测试结束后,机器有30%以上的可能性骗过测试者,让他(她)误以为自己是人,则说明计算机具有智能。这一测试标

准一直延续至今,可惜还没有一台计算机可以确定无疑地通过这一看似简单的测试。图灵的这些工作使他成为人工智能当之无愧的创始人(图 1-2)。



图 1-2 图灵和他的图灵测试

注:测试者通过键盘和机器及真人以自然语言对话,如果机器可以骗过测试者,让测试者以为它是真人,则认为该机器具有了智能。

1.2.3 达特茅斯会议:AI的开端

就在图灵开始他的人工智能研究不久,当时很多年轻人也开始关注这一崭新的领域,其中就包括美国达特茅斯学院数学助理教授约翰·麦卡锡(John McCarthy)、美国哈佛大学数学与神经学初级研究员马文·明斯基(Marvin Minsky)、贝尔电话实验室数学家克劳德·香农(Claude Shannon)、IBM 公司信息研究经理纳撒尼尔·罗切斯特(Nathaniel Rochester)。1956年,这些年轻人聚会在达特茅斯学院,讨论如何让机器拥有智能,这次会议被称为“达特茅斯会议”(图 1-3 和图 1-4)。正是在这次会议上,研究者们正式提出“人工智能”这一概念,AI 从此走上历史舞台。当时讨论的研究方向包括以下几个方面:

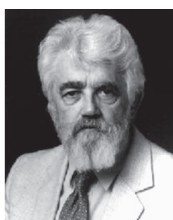
- 可编程计算机;
- 编程语言;
- 神经网络;
- 计算复杂性;
- 自我学习;
- 抽象表示方法;
- 随机性和创见性。

可见,当时人工智能的研究非常宽泛,像编程语言、计算复杂性这些现在看来并不算 AI 的范畴也需要人工智能的学者们考虑。这是因为当时计算机刚刚诞生不久,很多事情还没有头绪,AI 研究者们不得不从基础做起。尽管如此,现代人工智能的主要研究内容在这次会议上已经基本确定了。

达特茅斯会议被公认为是人工智能研究的开始,会议的参加者们在接下来的



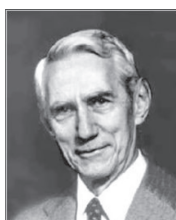
图 1-3 达特茅斯会议原址



John McCarthy
约翰·麦卡锡



Marvin Minsky
马文·明斯基



Claude Shannon
克劳德·香农



Ray Solomonoff
雷·索洛莫洛夫



Allen Newell
艾伦·纽厄尔



Herbert Simon
希尔伯特·西蒙



Arthur Samuel
亚瑟·塞缪尔



Nathaniel Rochester
纳撒尼尔·罗切斯特

图 1-4 达特茅斯会议的几位参加者

数十年里都是这个方向的领军人物,完成了一次又一次的创举和突破。

1.2.4 人工智能的三起两落

历史总是曲折的,同时也是螺旋式前进的,人工智能的发展也是如此。我们可

以将人工智能的发展分为以下几个阶段。

黄金十年(1956—1974 年) 达特茅斯会议后的十年被称为黄金十年,这是人工智能的第一次高潮。当时很多人持有乐观情绪,认为经过一代人的努力,创造出与人类具有同等智能水平的机器并不是个难题。1965 年,希尔伯特·西蒙(Herbert Simon)就曾乐观预言:“二十年内,机器人将完成人能做到的一切工作。”在这近二十年里,包括 ARPA 在内的资助机构投入大笔资金支持 AI 研究,希望制造出具有通用智能的机器。这一时期的典型方法是**符号方法**(Symbolic Method),该方法基于人为定义的知识,利用符号的逻辑演算解决推理问题。**启发式搜索**(Heuristic Search)是这一时期的典型算法,这一算法通过引入问题相关的领域知识(称为启发信息)对搜索空间进行限制,从而极大地提高了符号演算的效率。这一时期的典型成果包括定理证明、基于模板的对话机器人(ELIZA、SHRDLU)等。

AI 严冬(1974—1980 年) 到了 20 世纪 70 年代,人们发现 AI 并不像预想的那么无所不能,只能解决比较简单的问题。这其中有计算资源和数据量的问题,也有方法论上的问题。当时的 AI 以逻辑演算为基础,试图将人的智能方式复制给机器。这种方法在处理确定性问题(如定理证明)时表现很好,但在处理包含大量不确定性的实际问题时则具有极大的局限性。一些研究者开始怀疑用逻辑演算模拟智能过程的合理性。如休伯特·德莱弗斯(Hubert Dreyfus)就认为人类在解决问题时并不依赖逻辑运算^①,然而,不依赖逻辑运算的感知器模型被证明具有严重局限性^②,这使得研究者更加心灰意冷。AI 研究在整个 20 世纪 70 年代进入严冬。

短暂回暖(1980—1987 年) 到了 20 世纪 80 年代,人们渐渐意识到通用 AI 过于遥远,人工智能首先应该关注受限任务。这一时期发生了两件重要的事情,一是**专家系统**(Expert System)的兴起;二是**神经网络**(Neural Net)的复苏。前者通过积累大量领域知识,构造了一批可应用于特定场景下的专家系统,受到普遍欢迎;后者通过学习通用的非线性模型,可以得到更复杂的模型。这两件事事实上都脱离了传统 AI 的标准方法,从抽象的符号转向更具体的数据,从人为设计的推理规则转向基于数据的自我学习。

二次低潮(1987—1993 年) 20 世纪 80 年代后期到 20 世纪 90 年代初期,人们发现专家系统依然有很大的问题,知识的维护相当困难,新知识难以加入,老知识互相冲突。同时,日本雄心勃勃的“第五代计算机”也没能贡献有价值的成果。人们对 AI 的投资再次削减,AI 再次进入低谷。在这一时期,人们进一步反思传统人工智能中的符号逻辑方法,意识到推理、决策等任务也许并不是人工智能的当务之急,实现感知、移动、交互等基础能力也许是更现实、更迫切的事,而这些任务与

① Dreyfus H, Dreyfus SE, Athanasiou T. Mind over machine. Simon and Schuster. 2000.

② Minsky M, Papert S. Perceptrons: An essay in computational geometry. Cambridge, MA: MIT Press. 1969.

符号逻辑并没有必然联系。

务实与复苏(1993—2010 年) 经过 20 世纪 80 年代末和 20 世纪 90 年代初的反思,一大批脚踏实地的研究者脱去 AI 鲜亮的外衣,开始认真研究特定领域内特定问题的解决方法,如语音识别、图像识别、自然语言处理等。这些研究者并不在意自己是不是在做 AI,也不在意自己从事的研究与人工智能的关系。他们努力将自己的研究建立在牢固的数学模型基础上,从概率论、控制论、信息论、数值优化等各个领域汲取营养,一步步提高系统的性能。在这一过程中,研究者越来越意识到数据的重要性和统计模型的价值,贝叶斯模型(Bayes Model)和神经网络越来越受到重视,机器学习成为 AI 的主流方法。

迅猛发展(2011 至今) 人工智能再次进入大众的视野是在 2011 年。这一年苹果发布了 iPhone 4S,其中一款称为 Siri 的语音对话软件引起了公众的关注,重新燃起了人们对人工智能技术的热情。从技术上讲,这次人工智能浪潮既源于过去十年研究者在相关领域的踏实积累,同时也具有崭新的元素,特别是大数据的持续积累、以深度神经网络(Deep Neural Net, DNN)为代表的新一代机器学习方法的成熟,以及大规模计算集群的出现。这些新元素组合在一起,形成了聚合效应,使得一大批过去无法解决的问题得以解决,实现了真正的成熟落地。可以说,当前的人工智能技术比历史上任何一个时代都踏实和自信。

1.3 机器学习：现代人工智能的灵魂

从半个多世纪的发展历程可以看出,人工智能技术的进步走的是一条“反逻辑”的路。人类用一千多年的时间得到了可计算的逻辑,即数理逻辑。虽然绝大多数逻辑系统并不完备(可能存在不可证明真伪的命题),但在很多时候已经足以描述在数学和物理学上的很多知识(如概念、关系等)。这些知识是如此简洁美好,如果被计算机掌握,则有望实现理解、决策等智能行为,这也是最初的人工智能研究者所持有的基本思路。然而,人们在研究过程中一步步发现,人为设计的知识以及基于这些知识的推理过程在实际应用中非常困难。这不仅因为对知识进行形式化本身就很烦琐,即使完成了这一形式化,依然会有各种冲突和不确定性存在,使得推理很难完成。相反,从数据中学习得到的知识虽然可能是不精确、不全面的,但在很多时候更适合实际应用。因此,人工智能的研究者们不得不用数据学习逐渐取代人为设计。在这一过程中,我们失去了传统数理逻辑的简洁和清晰,越来越依赖从数据中得到统计规律,而这些规律天然具有模糊性和近似性。

这意味着当前人工智能技术与传统 AI 在方法论上已经有很大的不同了。当代人工智能的本质是让机器从数据中学习知识,而不再是对人类知识的复制,这一方法称为“机器学习”。基于这样的思路,人工智能已经不再是人的附庸,它将和人

类在平等的起跑线上汲取和总结知识,因而可能创造出比人类更巧妙的方法、生成比人类更高效的决策、探索人类从未发现过的知识空间。数据越丰富,计算能力越强,这种学习方法带来的效果越好,超越人的可能性越高。当前 AI 的很多成就很大程度是由庞大的数据资源和计算资源支撑的,典型的领域包括语音识别、图像识别、自然语言处理、生物信息处理等。21 世纪的 AI 是数据的 AI,是机器学习的 AI,“人工智能”里的“人工”更多的是设计学习原则,而非设计智能过程本身。基于此,本书将重点介绍基于机器学习的现代 AI 技术。关于传统 AI 方法,读者可参考朱福喜老师编著的《人工智能基础教程》一书^①。

1.3.1 什么是机器学习

1959 年,亚瑟·塞缪尔(Arthur Samuel)发表了一篇名为 *Some Studies in Machine Learning Using the Game of Checkers* 的文章。该文章描述了一种会学习的西洋棋计算机程序,只需告诉该程序游戏规则和一些常用知识,经过 8~10 小时的学习后,即可学到足以战胜程序作者的棋艺。这款西洋棋游戏是世界上第一个会自主学习的计算机程序,宣告了机器学习的诞生。

什么是机器学习?塞缪尔认为机器学习是“让计算机拥有自主学习的能力,而无须对其进行事无巨细的编程”的方法^②。尼尔斯·约翰·尼尔森(Nils J. Nilsson)则认为机器学习是“机器在结构、程序、数据等方面发生了基于外部信息的某种改变,而这种改变可以提高该机器在未来工作中的预期性能”^③。上述这些定义本质上是一致的,即认为机器学习是通过接收外界信息(包括观察样例、外来监督、交互反馈等),获得一系列知识、规则、方法和技能的过程。和传统算法相比,机器学习的一个巨大优势在于程序设计者不必定义具体的流程,只需告诉机器一些通用知识,定义一个足够灵活的学习结构,机器即可通过观察和体验积累实际经验,对所定义的学习结构进行调整、改进,从而获得面向特定任务的处理能力。

1.3.2 机器学习发展史

图 1-5 给出了机器学习发展历史上的一些重要人物和标志性事件。总体来说,20 世纪 50 年代以前是技术积累阶段,研究者在统计学习和优化方法上提出了一系列模型和准则。1950 年图灵提出图灵测试准则,开创了人工智能的广阔领域。机器学习伴随着人工智能的研究开始萌芽。1959 年亚瑟·塞缪尔的划时代

① 朱福喜,朱三元,伍春香. 人工智能基础教程[M]. 北京:清华大学出版社,2006.

② Samuel AL. Some studies in machine learning using the game of checkers. IBM Journal of Research and Development, 1959, 3(3): 210-229.

③ Nilsson NJ. Introduction to machine learning. URL <http://ai.stanford.edu/nilsson/mlbook.html>, lecture notes, 1998.

论文将“机器学习”这一重要概念引入人工智能,并开始独立解决实际问题。整个 20 世纪 60 年代,以符号逻辑为研究对象的符号学派(Symbolism)是人工智能研究的主流,人工神经网络、概率模型、遗传算法等更侧重“学习”的方法开始萌芽。进入 20 世纪 70 年代,人工智能的冬天来临,机器学习研究也走入困境,特别是在马文·明斯基发表《感知器》一书后,被寄予厚望的人工神经网络的研究几乎停滞。

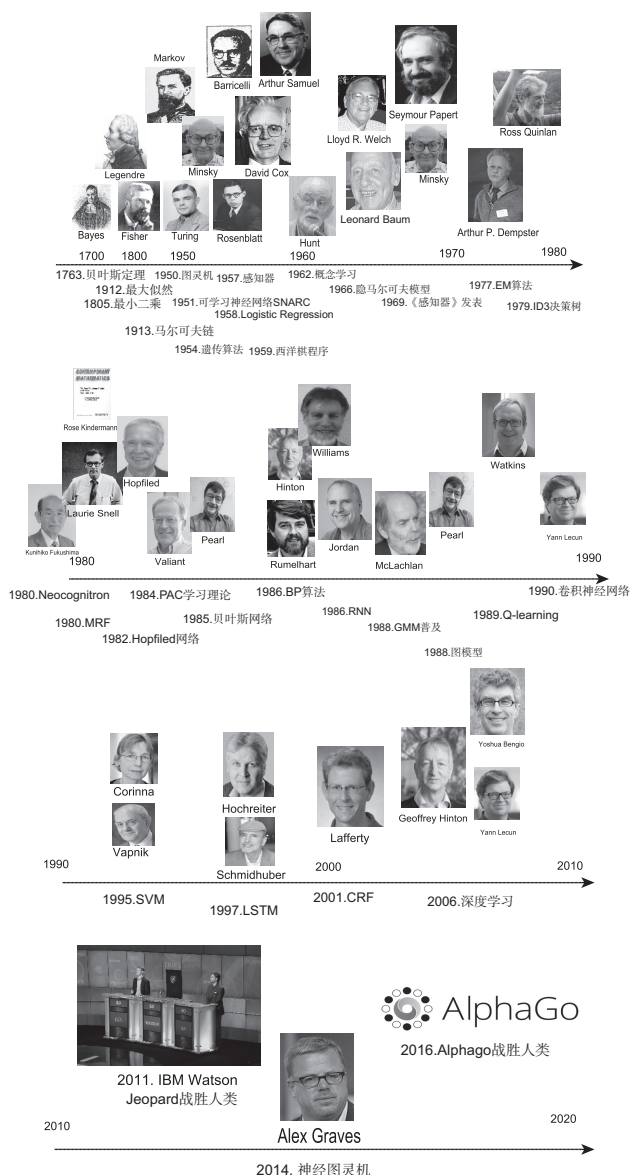


图 1-5 机器学习发展史上的若干重要人物和重要事件

注：坐标轴上方为人物,下方为对应人物的对应事件。