

93, 20(3)

1-3

人工智能

计算机

计算机科学1993 Vol.20 期.3

人工智能研究的历史与展望

黄林鹏 编译

孙永强 审校

TP18

图灵奖获得者Maurice V. Wilkes教授最近应CACM杂志社之邀,撰文评述了人工智能研究的历史及展望,该文对机器学习、专家系统、神经网络、数字/模拟计算机等做了极其精辟的评价,读来颇受启发。

我知道我获得了名誉是因为我批判了对人工智能(AI)所提出的一些主张,但令我反感的是,我听到了某些夸大其词的宣传和缺乏应有的自我批评。然而,人工智能的哲学基础问题一直深深地吸引着我。我甚至能宣称我是第一个AI专业人员,我说的专业人员是指为(人工智能)效力并得到了回报的人。1950年当图灵的文章“Calculating Machinery and Intelligence”在《Mind》上发表的时候,我读过并深受影响。那时,没打算和图灵的博识与智慧竞争,我写了一篇相同主题的短文,投送《The Spectator》杂志——一家注重当代科学发展动向的刊物。编辑采用了,我也很快收到了稿酬。

除了数学论文外,我一直认为在《Mind》上发表的这篇文章是图灵写下的最好的文章。图灵以“机器能思维吗?”问题开始,揭示这个难以满意阐述的问题,他尝试抽取有意义的本质,提出了著名的图灵测试。图灵建议,不是问一个特别的机器是否能思维,而是问它能否通过这个测试。图灵测试就是让机器装做人并使一个询问者不能辨别它是男士还是女士。

尽管图灵满怀信心,但他承认他还没有强有力的证据来证明有一天数字计算机能通过他的测试。图灵提出一个有趣的建议是,机器可以装备一个学习程序,然后像儿童一

样进行教育。他指出的途径是,机器被程序设计后,可以模拟儿童而不是成年人的大脑。这说明思维和学习有着密切的联系。图灵深知机器如同儿童一样,其教育是一个漫长的过程。他也看到了实际的困难:计算机没有腿,不可能把它像儿童一样送去上学。即使灵巧的机械工程师能够克服这一困难,但图灵怕机器在学校里会成为其他儿童取笑的对象。

书写一个程序实现一种简单的学习,如学习识别人名的缩写,没什么困难。程序可以包含一张它能理解的人名缩写表。当给出一个不熟悉的缩写时,程序将做出猜测。当把猜测的正确性告诉程序后,它能修改其缩写表。尽管这没有什么深奥的东西,但也理应称之为学习。

在图灵论文的推动下,我和我的同事写了一些上述的学习程序。这些程序的局限性很快就表现了出来,它们只能做给它们编写过的事情,不会更多。由于这个原因,一旦第一次运行过这些程序后,就没什么意义了。我很快意识到,在我称之为广义的学习程序(它能继续学习新东西)研究方向上需要一个突破,也许称它为不受限制的学习程序更好。

假如在十七世纪末就有了计算机,而且人类已懂得了如何书写不受限制的学习程

序，那么一台具备这种程序的机器就能吸收牛顿的研究成果，然后吸收法拉第与爱因斯坦的成果，于是现在它就有可能在黑洞研究上做出贡献。假如它读过获更斯的小说，它将能参加由图灵的丰富想象力所创造的半开玩笑式对话：

图灵的梦想——和计算机对话(询问者是人类，智者是一台计算机)

询问者：你的十四行诗的首行为“你如同夏日”，你不觉得“春日”更好吗？

智者：它不合韵。

询问者：“冬日”如何？它可是完全合韵的。

智者：它是合韵，但没有人愿被比为“冬日”。

询问者：你不是说过匹克威克先生让你能想起圣诞节吗？

智者：是的。

询问者：圣诞节是冬天的一个日子，我想匹克威克对这个比喻不会介意吧。

智者：我认为你不严谨，“冬日”指的是一般的冬天的日子，而不是某个特别的日子，如圣诞节。

开始时，我希望在编制了充分多的简单学习程序之后，我们能辨明它们的共同点，并由此编写“不受限制”的学习程序。然而事情并非如此简单，我也很快认识到除非有一个天才能把这个主题了解得非常透彻，不受限制的学习程序的编写是不可能的。

出乎我的意料，其他人并没有很快认识到这一点。我所书写的程序仅仅用了两个星期，而其他人花了两年或更长的时间，所得到的程序和我差不多。

Arthur Samuel在国际象棋程序方面的开创性工作可以说明这个问题。Samuel的主要兴趣是探索可被象棋程序设计者成功采用的棋盘递归搜索算法，而且他也有兴趣看一看他能否使程序从经验中学习并进一步改进其性能。

在Samuel的程序中，经常出现的一个基本步骤是棋盘局势评价。Samuel设计了一些不同的评价函数，他不是选择其中之一，而是对每个评价函数赋予权重并取其线性组合。Samuel发现他能设计一些通过自动修改

权值来优化程序性能的方法，但他也发现他的程序不能自行发明一种评价函数并把它加入已有的评价函数集中。换言之，他能书写一个以数学形式而不是以智能概念进行处理的程序。

人们已编制了一些非常复杂的学习程序，它们的演示给非技术人员留下了深刻的印象。但结果证明，这些程序的性能优化都借助了调整参数或修改数据结构来影响程序的内部状态。

一个引人注目的高水平AI程序是1971年由Terry Winograd编写的，该程序旨在控制一台能堆积木的机器人。机器人已被大大简化并由计算机图示模拟。由于那时计算机图示学的使用尚未普及，这一特点十分突出。控制机器人的命令用英语给出，程序尝试从英语句子中抽取信息。然而，正是图示学和自然语言界面的结合魅力使其被誉为机器智能的骄傲。

专家系统与图灵梦想

最初，AI这一术语专门用在图灵梦想意义下，即一台计算机可被编程使其表现人类的智能行为。然而，近年来，AI多半用作为程序的称号，如果这些程序不是在AI界中出现的话，可被看成是使用诸如COMIT、SNOBOL语言编程或E.T.Irons研究开创性的语法制导编译器的自然产物。我管叫专家系统。

在简单的专家系统中，所有的知识都被程序设计者并入程序，由专家系统的另一别名——“基于知识的系统”可更确切地说明这一点。这就宛如通过大脑外科手术而使儿童掌握乘法表。在更精致的专家系统中，在系统生存期内可以修改其内部数据库，这些系统显示了如我们上述的程序所具备的类似学习能力，当然也有同样的局限性。专家系统是AI界奉献给人类社会的一份珍贵礼物，但它和图灵的梦想毫无关系。

图灵在1950年曾预言他的梦想将在50年内成为现实，具体地说在一台有128MB内存

的计算机上实现。这50年使我们到达了2000年，但他的梦想显然不可能成为现实。我们所面临的事实是，自1950伊始已过去了40年，但我们在实现图灵所定义的机器智能方面没有任何明显的进展。也许过去的时间使我们认识到图灵的梦想不可能在一台数字计算机上实现。

我们使用术语“数字计算机”来描述我们所建造的计算机，尽管“真正的数字计算机”事实上只是一个抽象概念。现实中的任何数字计算机都必须由模拟电路构成，图灵也认识到这一点，那时凡目睹了工程师为了使真空管以数字方式工作而做出艰苦努力的人都不难明白这个事实。工程师努力的结果使早期的数字计算机偶尔出错，但在基本电路一级，工程师们仍需认识到他们所设计的电路基本上是一些模拟器件，他们所使用的仍是仿真模拟器和其它的模拟设计工具。

数字计算机只好在逻辑系统中工作，众所周知在这样的系统里有许多事无法完成。我们无需引用Goedel定理来证明这点，Desargues关于三角形的定理将同样可以证明。然而，对这些层次上的证据间关系我没有把握。

在一个特别层次上，关于数字计算机局限性的例子之一是，它不能直接求解微分方程，通常用差分方程来代替微分方程以绕过这一困难。然而，微分方程和差分方程是两类十分不同的方程类型。例如，带有两点边界条件的二阶线性微分方程有无限多的独立解，而相应的差分方程仅有有限多的解。数学家可以通过极限把微分方程和差分方程统一起来，但对于数值领域中没有接触过极限的工作者就没有办法。

人脑——数字或模拟？

如果我们把人脑看成一台机器，那么我们应能证明这种机器能表现智能。然而，如果我们不能假定人脑是以数字方式工作的，那么我们就无法解决数字计算机能否表现智能的问题。一旦我们作出这样的假定，我们

将面临这样一个问题：人类神经元的速度比现代计算机的门电路速度大约慢五个数量级，如果人脑是以数字形式工作的，那么它是如何达到足够快的速度呢？那些把人脑想像为数字式的人通常会说，人脑具有高度的并行性，可弥补神经元速度上的不足。然而，一般高度并行计算机很难在速度上得到100甚至10的加速因子。即使那些热衷于并行计算的硬汉所持的加速因子也都停留在100,000这个数值上。

不过，上述论据是纯学术性的，我们没有理由认为人脑是数字设备。事实上，以数字和模拟划分，对于刻画人脑的功能完全是不适合的。就像我上面所指出的那样，数字计算机是一个抽象概念——一种人类设计者用以组织其思想的有用手段。没有理由认为一个非人类设计者应以相同的方式工作。根据进化论假说，甚至错误地认为人脑天生就满足用符号表示的要求。“盲目的进化偶尔发现……看！男人和女人能思考。”

我建议AI界在与人类活动竞争时，应在其机器中包括某些模拟元件，可能工作得更好。我对神经网络不太热心，但我发现神经网络也包含了一些模拟判别电路。

我不想给人留下我认为图灵的梦想会由于模拟机器代替数字机器而成为现实的印象。我不作出这样的预言。事实上，我们所建造的模拟机器也有其自身的局限性，这些局限性可能和数字机器一样，也可能不一样。然而，我确实提出了一个这样的现实假设：图灵意义下的智能行为超出了数字计算机所能处理的范围。

一个否定的原理可能对引导研究具有重大意义。如果没有热力学第一定律，那么许多机械工程系的优秀学生可能会埋头研究永动机。认识到图灵的梦想不可能借助数字计算机来实现，将有助于学生避开渺茫的研究方向。

参考文献：CACM Aug.1992