

机器真的能思维吗？

WILL MACHINES EVER THINK?

彼得·J·丹宁

本文为美籍华裔教授王羽先生推荐,原载《美国科学家(American Scientist)》杂志,1986年7~8月号。

机器能思维吗？这个问题迷住发明家、哲学家和逻辑学家们已几百年。1938年英国数学家艾伦·图灵(Alan Turing)证明,一个简单的计算模型,即现在所说的图灵机,可以进行任何计算:它可以模拟任何其他图灵机的计算。这种模拟计算,就是存储程序式计算机的基础。这样,人们似乎可以认为,存储程序式计算机最终将能模拟人的任意行为。如果再掌握了每一神经元状态的信息以及神经兴奋的机制,似乎也就可以计算出某人每个神经的下一状态,从而可预示他的下一步行动。诚然,实际上不可能获得进行准确预测所需的足够信息。但是,从原理上讲,似乎有可能由程序控制机器,使其象人一样动作。

1950年,图灵提出一种方法,用来确定一台机器和一个真人之间的相似程度。这种方法现在叫做图灵测试法¹。按照这种方法,设想一个查询者,通过电传打字机,与两个对象交换信息,其中一个对象是人,而另一个是一台计算机。这个查询者需要花多长时间,才能确定出两个对象谁是谁呢?这种判断的可靠性又有多大?但图灵断言:50年内,计算机就可以很好地从事这种模仿游戏,以致70%的查询者在5分钟对话后,仍不能正确辨别它们。这个判断准则,已成为当今评价专家系统完善程度的方法。一个专家在区分一个程序和一个专家的行为所花时间越长,则该程序越成功。根据这一判断准则,人工智能程序员们就可容忍程序出一些差错,只要其出错率不比专家本

人为高即可。

图灵巧妙地把“机器能思维吗”这一个问题,替换成“一个数字计算机程序能在什么程度上模仿人”。对前者从哲学上对其用词就可提出许许多多的问题。如什么是机器?什么是思维?图灵提出一个可以通过实验科学地解决的相关但不等价的问题。要用实验来回答最初那个问题,无异于用实验来确定我是否有灵魂一样。

这个替代问题的恰当程度如何?计算机程序真的能如实模拟人的行为吗?虽则逻辑学和科学有助于阐明这个问题,但单纯靠推理和实验是不能解决问题的,认识到这一点十分重要。

人工智能的研究,在这个似乎越来越可能解决的问题上争论了许多年。1967年,麻省理工学院的马文·明斯基(Marvin Minsky)推测,只要拥有100万个适当组织的、有关日常生活的一组数据的一台计算机,就可以表现出很高的智能。如今,已可轻而易举地存储几倍于这个数量的信息,但却没有谁能成功地构造出一个智能程序。1970年左右,象棋大师戴维·利维(David Levy)对象棋程序不久将达到象棋大师水平的预言提出了挑战。他还为能在十年内设计出可以击败他的计算机程序的人设了奖。但直到今天,还没有任何人能构造出一个程序,可以在和利维的十盘比赛中胜一盘或两盘以上。1980年前后,日本人断言,他们将有能力构造出许多计算机程序,这些程序能在特定的领域中具有专家的技术水平。他们真能

彼得·J·丹宁 (Peter J. Denning) 美国国家航空和航天局 (NASA) 艾姆斯研究中心高级计算机科学研究所主任。

成功吗?

智能已成为模拟装置追求的目标。绝大多数人都可以把复杂的机构和创造这种机构的思维过程区分开来(专利法也是这样划分的)。自动驾驶仪是很好的飞行控制仪,但不是飞行员。一个能够超前探测四步的计算机程序可能是令人生畏的棋手,但却不是大师。1966年,麻省理工学院的约瑟夫·威森鲍姆(Joseph Weizenbaum)设计了一个叫伊莱扎(Eliza)的程序。它的一个模型曾模拟了心理学家的谈话²。伊莱扎可以对输入文件中一些事先定义好的关键字进行处理。对“我感到X”的输入,伊莱扎可能打印出“你为什么感到X?”威森鲍姆证实,借助伊莱扎有可能不需要给程序赋予任何理解能力,就可以令人信服地表现出智能。一旦什么人找到了使人类某一级行为机械化的方法时,人们也就再也不会把这种行为看成是智能了。事实上,人们总是让机器承担那些繁琐的日常工作。

对日益增多的人工智能研究人员来说,这个领域的根本问题却总是那样难以捉摸,而问题的解又总是变幻不定,从而引起了不少疑团。戴维·帕纳斯(David Parnas)和约翰·肖尔(John Shore)对构造专家系统的这种试凑(trial-and-error)法提出了批评。休伯特(Hubert)和斯图尔特·德赖弗斯(Stuart Dreyfus)则断言,经过25年的研究以后,人工智能并没有实现其许诺,也没有迹象表明能够实现这个许诺。道格拉斯·霍夫施塔特(Douglas Hofstadter)也说:许多人工智能问题的提出方式是肤浅的。越来越多的人工智能研究人员都疑惑不解:为什么从机器中产生智能行为如此困难?是工程学无能?是目标过于奢求?还是有什么根本性的东西使我们达不到目标?这种无休止的争论,可能会导致一种对机器在人类社会中的地位的新看法。

最近出版的三本书给这个老问题投射上了一丝引人兴趣的亮光。第一本是休伯特和斯图尔特·德赖弗斯合著的《机器的思维(Mind over Machine)》,它分析了机器可能获得人类技能的能力³。第二本是特里·威诺格拉德(Terry Winograd)和弗尔南多·弗洛里斯(Fernando Flores)合著的《判断计算机和认知过程(Understanding Computers and Cognition)》,它分析了人和机器间可能的关系,并提出了新的计算机基本设计原则⁴。第三本是詹姆斯·阿尔伯斯(James Albus)著的《大脑、行为和机器人学(Brains, Behavior and Robotics)》,它分析了智能行为的特征及模拟这些特征的计算结构⁵。你将会发现,这三本书彼此

之间是相互关联的。

德赖弗斯派指出,人类具有机器所不能匹敌的智能。为了支持这种论点,他们分五个层次来分析智能这个人的技能的重要方面。这五个层次是:新手、已有一定经验的新手、合格的工作者、熟练工作者和专家。新手只了解某项工作的基本事实以及使用这些事实的孤立规则。已有一定经验的新手则可以根据一些实例制订出行为规则,并能考虑到事件的前因后果。合格工作者是高水平工作人员,他行为的目的性很强,他不用思索各项规则就能采取行动,并能审时度势。熟练工作者能全局在胸,对形势无需仔细研究就能直接用于现实工作中。专家则把全部精力集中在整体布局上,很少注意技巧问题,也很少有意识地去使用分析推理法。他总是从他所能理解和控制的整个对象和形势的宏观上采取行动。

我认为,写作技能是说明这种观点的最好例证。一个新手总是要依赖有关语法、标点和拼写规则的书籍,要花较长时间,而且要反复查考规则书籍才能写出正确无误的句子和段落。一个已有一定经验的新手就掌握了多数主要规则,而且从一些佳作范例中又学到其他一些规则。一个合格的写作者更多注意的是情景的沟通,而不是具体细节的写作,他已掌握了某种文体的写作原则,并能自觉地运用;他还很善于构思和利用写作大纲。一个熟练的写作者会感到写作是件轻而易举的事,在语法或文体上他极少出差错,他能分析和编辑自己的作品,并能模仿他人的写作风格,他极善修辞和运用各种修辞技巧,具有构思和完成大型著作的能力。一个专业作家则生活在一种意念的世界中,总怀着一种需要交流的紧迫感,他不断探索出新事物的描述方法和旧事物的新描述途径,他能把神韵表达出来,并能模仿许多写作风格,然而他又具有自己的风格,而且非常有特色,以至于编辑人员改动了他作品中的一个字时,他也可以一眼看出。一个专业作家可以在头脑中勾画出大型著作的轮廓,并能自如地把握总体思想和具体表现手法。

在这种分层的描述中,运用规则和符号是最低水准的特征,而调用高度抽象的模式则是最高水准的特征。最低水准的规则,一个新手可以不太费力地写出,并直接加以使用,而最高水准的技能则只有通过多次反复才能掌握。德赖弗斯派把专门技巧视为高水准能力,它与使用符号和规则不同。他们认为,存储程序式计算机在使用符号和规则方面是很好的,但不能对所有的完整模式进行快速计算,因为这大大超出了它的能力。基于这种思想,他们对专门技巧就要求提出全新的

计算方式。全息摄影就是一种现有的,具有某些专门技巧特征的装置。它把整个图象在每个区域都进行编码,用一束光线照射两个全息摄影图象就可确定其相似性。实际上,至今尚不知道这种模式计算机的结构或功能;至少不知道这些计算机能否具备人所具有的从其他图象中抽象出图象的能力。

德赖弗斯派的结论是:用规则来处理事务,存储程序式计算机就不可能实现高水准技能。这种计算机的内部结构没有能力用与真正专家相同的方法处理信息。他们警告说,如果确信事务处理计算机很能干,尽管要花费大量资金,但人们仍将尝试利用专家系统,建立在这种想法基础上的研究规划,就潜伏着危险。极少有合格的程序能用在需要具有真正专家技能的地方。他们指出,用于教学和培训目的的专家系统,只能使学生在规则和符号使用方面得到提高。但真正的专家活动却远远超出这些,达到专门技巧的程度。

和德赖弗斯派一样,威诺格拉德和弗洛里斯对当代人工智能研究没有能力实现其预期目标而感到困惑。他们希望通过对正在进行研究的结构、要求解答的问题的性质以及解决这些问题的方法等几个方面的分析,来打开新的局面。

根据他们的观点,计算中令人感兴趣的基本问题是设计问题:在人类实践中计算机服务的目的是什么?设计工作不但要考虑计算机是如何操作的,而且还要考虑它对人的影响如何。例如,一台文字处理器,不同的人对它的看法不同:对制造商来说,它是一个电子器件箱;对程序员来说,它是一个软件包;对作家来说它是一种有效的表达工具。如果把文字处理器分解开来,那么对上述三种人来说它都是一种可见的客观存在。

威诺格拉德和弗洛里斯发现,传统观念,如过去的经验、观察世界的方法、公认的不言而喻的假设等,对我们的设计方法有很大影响。而传统观念又被其非理性掩盖。西方科学的传统(科学的方法)坚持认为,知识是客观的,它重视能够系统地应用于预测部分客观世界的明确的理论。威诺格拉德和弗洛里斯称之为理性化传统观念。起源于理性化传统观念的计算机传统观念重视信息、表达和决策,也强调对人脑、智能以及其他拟人行为进行模拟。

西方科学家极力推崇的理性化问题求解过程由三部分组成:给出以客观事实及其相互关系表达的状态的准确描述;枚举出问题所有可供选择的解;评价每一个可供选择的解,直到得出一个满意的低费用的解为止。存储程序式计算机非常适合于这一求解过程:客观事实、相互关系和规则都

可以明确地表示为程序语言结构;演绎和查询可以用来枚举可供选择的解;计算和比较可以用来选出一个低费用解。当今的人工智能研究都是基于这样一种研究倾向,即用启发式过程来使理性化解题方法等价于智能。威诺格拉德和弗洛里斯提出这样的问题:如果这些是人类认知的唯一要素,那么为什么人们认为计算机具有类似于思维的性质,而提供和处理信息的其他设备却没有呢?

威诺格拉德和弗洛里斯用鲜明的事例指出:认知不只是基于表达的系统手法。德赖弗斯关于技能水准和技巧的论点切中了要害。理性化的问题求解方法至多只是部分地描述了人们是如何解决问题的:人们是在一定处境中做出反应的。世界总是通过它内部的有机体表现出来。一给定的有机体,其结构在确定这个有机体具有哪种表现能力方面起着主要作用;相似的有机体使这种表现更趋于完善,更具有普遍性,使有机体之间得到相互交流。存储程序式计算机和人的关键差别就在于:计算机处理符号时并不关心其意义;人则是在一种解释框架(*framework of interpretations*)范围内处理每一件事物。因此,存储程序式计算机要模仿人类进行计算,其结构是错误的。

威诺格拉德和弗洛里斯提出一个新的研究方向:不要追求怎样使计算机的行为更象人类,而是要寻求设计出能帮助人类更有效地工作的计算机。有些是计算机很擅长的工作,如执行运算、检索信息、信号处理和滤波、辅助交流、过程监控等。机器应用于这些领域,对社会能产生相当积极的影响。强调用创造性的方法使机器有助于人类,这样一个设计过程是有可能成功的。

威诺格拉德和弗洛里斯的书给人一种印象,即寻求可以思维和理解问题的机器是徒劳的。德赖弗斯派则没有这么肯定,尽管他们也认为存储程序式计算机不能具有类似思维的性能,但仍不排除其他结构的计算机取得成功的可能性。

阿尔伯斯则更相信,机器中可以产生出类似于思维的行为。他最感兴趣的是结构类似于人的神经系统的机器。(这使人联想到威诺格拉德和弗洛里斯的“结构决定功能”的主张。)1950年,图灵探讨了构造能在智能方面和人媲美的机器的两种方法。第一种是要进行很抽象的活动,如逻辑推理或下棋。第二种是给机器装上复杂的感应元件和一个用来存储表示经验要素的复杂编码装置,经过反复接触感应样本,机器渐渐可以识别相似的样本,并进行相应的动作。阿尔伯斯说,大多数人工智能研究都是沿着第一个方向前进的。这

大概是因为许多先驱者们都是善于抽象推理的数学家的缘故。实际上,人工智能的全部工作就是集中在对思维的这种抽象推理能力的模型化上。

阿尔伯斯主张遵循图灵的第二种方法。他要从生物功能分层结构的底层,而不是从顶层开始研究,再造昆虫、鸟、动物和灵长类动物的控制功能和行为模式,再把这些功能和模式当做构成元件,组成人类的高级功能,最终推出达到人脑智能和思维能力的装置。作为朝这个方向迈出的一步,阿尔伯斯分析了人的神经系统的组成,提出了一个“小脑模型算术计算机(Cerebellar Model Arithmetic Computer)”,作为能再现许多神经系统行为的一种结构(这是他的专利)。阿尔伯斯著作的很大篇幅都用于介绍如何教会这台计算机,和经过学习后,它又是如何实现多种神经系统功能以及如何对复杂的刺激进行反应的。阿尔伯斯甚至还介绍了如何教会这台计算机以存储程序式计算机的方式工作。

我认为,这几位作者的论点新奇而令人鼓舞。他们并不反对人工智能的研究,相反而为其打开了新的大门。他们的某些探索性工作将导致具有更强的、类似专家信息处理能力的计算机的

诞生。其他的一些研究也将促进设计的改进。但我不相信,关于计算机能否模仿思维的争论会就此终止。也许存储程序式计算机是真的不能思维的,而其他结构的计算机也许可能……。

参考书:

1. D.Hofstadter. 1985. A coffeehouse conversation on Turing's test. In *Metamagical Themas*. Basic Books.
2. J.Weizenbaum. 1976. *Computer Power and Human Reason*. Freeman.
3. H.Dreyfus and S. Dreyfus. 1986. *Mind over Machine*. Free Press. Excerpts: Why computers may never think like people. *Technol. Rev.*, January.
4. T.Winograd and F. Flores. 1986. *Understanding Computers and Cognition: A New Foundation for Design*. Ablex.
5. J. Albus. 1981. *Brains, Behavior, and Robotics*. McGraw-Hill.

(何京翔译 吴永智校)

简讯

第五届国际材料力学行为会议在北京召开

第五届国际材料力学行为会议于1987年6月3日至6日在北京召开。这次会议是由中国金属学会、航空学会、力学学会共同组织召开的。颜鸣皋和师昌绪教授任会议组织委员会主任,郑哲敏和章守华教授任学术委员会主任。参加会议的有来自24个国家共328名中外代表,其中国外代表118名。在全体会议上世界著名材料科学家如英国的米勒教授,美国的阿根教授,法国的比努教授,瑞士的斯贝特教授等15位专家作了专题学术报告。在四个分组会上就材料在工程实践中的断裂、疲劳、环境影响、高温形变与失效、复合材料与非金属材料力学性能五个专题举行了论文报告会,共宣读了138篇学术论文,其中我国57篇。这次会议是一次材料科学领域的空前盛会,也是一次难得的国际间学术交流。美国的曼森、阿根、纽曼、雷切、蔡维伦、魏伯英等著名学者以及我国材料科学界肖纪美、李敏华、林栋梁、蔡其巩、雷廷权、赖祖涵、孙训芳、康沫狂等专家、教授都参加了会议。会议在高温材料引入断裂力学,复合材料结

构设计与应用,高分子材料强化,小裂纹屏蔽效应,宏观力学与微观组织结构相结合,工程法解决应力腐蚀问题,腐蚀疲劳机理与定量分析方法等方面,均反映了当前国际上的研究动态与先进水平,以及近几年世界各国的较新研究成果。美国加州大学雷切教授关于小裂纹屏蔽效应,改变微观结构提高疲劳裂纹起始和小裂纹扩展寿命的研究,美国国家航空和航天局纽曼提出的定量预测小裂纹寿命的方法,瑞士斯贝特教授提出的解决应力腐蚀失效的一系列有效的工程方法等引起与会代表的普遍注意。我国固体物理所孔庆平研究员用内耗法研究材料蠕变强度变化。钢铁学院谢锡善教授研究缺口影响下蠕变、疲劳强度及其应用等也均引起国外许多学者的浓厚兴趣。会议期间,国际材料力学行为会议理事会召开了有19个国家理事参加的全体会议,讨论通过了会议章程,决定第六届国际材料力学行为会议1991年在日本京都召开。我国北京航空材料研究所颜鸣皋教授被推选为本届理事会主席,吴学仁博士任秘书。