

人工智能美学如何可能

陶 锋

2017年6月,人工智能(Artificial Intelligence)系统AlphaGo在围棋比赛中全胜世界冠军柯洁,AlphaGo所展现出的人工智能技术方面的巨大进步以及向人类智能挑战的野心,引发了人们对人工智能这一领域的又一次讨论热潮。百度推出的人工智能“小冰”写作的诗集、索尼利用人工智能摹仿披头士(Beatles)乐队创作音乐,还有人工智能在视觉和情感识别上面的突破,不由得引发人们的想象,人工智能会不会最终有一天能像人那样创造出艺术品?人工智能是否能够真正地拥有人类的情感以及意识呢?

机器具有智能的设想,最早源于计算机之父阿兰·图灵(Alan Turing),他在1950年的论文“计算机器和智能”中提出了“机器思考吗”的问题,他设想,可以设计一种对话实验,来测试机器是否能像人一样思考。他的结论是,机器终有一天可以实现智能⁽¹⁾。“人工智能”这一概念则是在1956年于美国达特茅斯学院举行的第一次人工智能大会上提出来的,当时,人们预测,机器人将在二三十年来实现智能化。可是,人们过高估计了人工智能技术的发展水平,而低估了人类智能研究的困难,人工智能在随后的几十年里,进展缓慢,甚至被人怀疑是科学界的一大骗局,直至近二十年来,人工智能在神经网络、专家系统、深度学习等理念和技术的发展,才使得人工智能获得了突破性的进步。

但是,人工智能在取得巨大进步的同时,却也面临着很多困难,例如对人类智能中的常识性推理、视觉等感觉、情感、想象力、创造力等方面,显得束手无策,简单的计算机逻辑符号与推理模式难以去描述这些人类习以为常的思维和行为。而在这些计算机难以形式化的地方,正是人类艺术所需要运用到能力,是美学的研究对象。人工智能哲学早已吸引了如哲学家塞尔(J.Seale)以及人工智能之父麦卡锡(J.McCarthy)等人的注意,他们纷纷发表相关文章讨论人工智能与哲学的关系,这些讨论从一定程度上促进了人工智能研究方向的转变。但是人工智能与美学的关联研究却似乎乏人问津,笔者认为,这并非是因为美学研究在

人工智能领域内研究不重要,而是因为,人工智能技术长期未能染指美学研究的情感和艺术领域,如今,人工智能在这些方面已经取得了初步进展,学者们也试图写出一些有着人类情感和创造力的算法,所以,人工智能的美学研究也应该像哲学研究一样,提上日程。另外,由于人工智能和美学的研究方法、知识积累和术语各方面相差甚远,一般的人文学者很难理解抽象的逻辑、计算机符号,这也加大了该领域研究的难度。笔者希望能够通过对人工智能美学的初步探讨,吸引更多的有识之士加入到这一领域的研究中来。笔者相信,人工智能美学研究不仅能为人工智能的情感和艺术方面工作提供参考,还能触及人类身心问题以及艺术本质问题。

一、人工智能美学是什么

1. 人工智能与哲学

关于人工智能,许多专家从不同角度给出了定义,图灵将机器与思考联系起来,认为机器智能化,就是能像人一样的思考(think)⁽²⁾。博登(M.Boden)认为:“AI就是研究如何制造或者程序化计算机,使得它们能够做那些心灵可以做的事。”⁽³⁾此定义将计算机与心灵联系起来,指出了人工智能应对人类心灵进行摹仿。而在《智能研究》一书中,作者强调了计算机与世界的理解和互动关系,并提出了计算机有自己独特的世界和思维方式,计算机“是智能的,在于它能够通过一种分层的记忆模式与它的世界进行理解和互动,并且能够以一种类似你我思考我们的世界一样的方式来思考它的世界”⁽⁴⁾。《人工智能:一种现代方法》的作者发展了图灵的界定,扩展出了人工智能的四种模式,即像人一样思考、理性地思考、像人一样行动、理性的行动⁽⁵⁾,区分了思考和行动。这几个定义虽然各有侧重点,但是其基础都在于将机器与人的心灵、思考和行为联系起来,因此,我们可以简单地概括,人工智能就是人造机器对人类的思维和行为的模拟活动。

另外,专家们根据机器智能化的程度不同,将之

分为弱人工智能和强人工智能,弱人工智能指的是“机器能够智能地行动吗”,而强人工智能则是“机器能真正地思考吗”⁽⁶⁾。从这个定义来看,弱人工智能是从机器的行为和效果来看机器是否是智能的,现在已经有了许多智能机器人能从某种程度上实现智能行为,如 AlphaGo 下棋,小冰写诗,等等。而强人工智能所说的“真正地思考”,要求则高得多,它意味着人工智能的思维模式要接近人类的思维模式,并且有可能需要“意识到”自己在思考,如塞尔所说的那样。这种程度的机器,目前只能存在于设想之中,而本文所探讨的人工智能与美学、艺术的问题,就要涉及这种理想的强人工智能。

人工智能研究和哲学的关系异常密切,这是因为,现代哲学从根本而言,就是人类思维和存在方式的方法论研究,无论是一般的计算机还是更高级的人工智能,其对人类思维进行模拟时,都要涉及人类思维的方法论,学者普遍认为,早期的计算机工作机制——即主张以符号逻辑模拟人类思维,这主要是受到了莱布尼兹的数理逻辑、弗雷格与前期维特根斯坦的语言逻辑学派的影响。而在 20 世纪 80 年代,人工智能哲学的研究就已经拉开序幕,著名心灵哲学家塞尔针对人工智能和图灵测试问题进行了批判。英国哲学家博登认为人工智能哲学与心灵哲学、语言哲学、认知科学哲学、计算心理哲学以及认识论具有紧密的关系⁽⁷⁾。人工智能之父麦卡锡则认为人工智能与哲学的关系比科学与哲学的关系要更加密切,这是因为人工智能需要哲学、特别是认识论的介入⁽⁸⁾。美国哲学家德雷福斯(H.Dreyfus)则批评了早期的人工智能方向——符号主义,认为其背后的哲学基础是西方哲学的理性主义和原子论传统,他建议用海德格尔的哲学来解释连接主义,并将之视作人工智能的新方向。在中国近二十年来,也有一些学者介绍了西方人工智能哲学的研究情况。

但是,为什么人工智能美学领域却少有人研究呢?人工智能和美学真的能够结合起来吗?

2. 人工智能美学何以可能

我们先来看看美学的定义和研究范围。“美学”(Aesthetics)一词最早由德国哲学家鲍姆嘉通提出,这个词源于希腊语“Aesthetica”,原意为“感性、感觉”。鲍姆嘉通将美学界定为研究那种低于理性能力的一种类理性能力,即感性,而感性的完善就是美⁽⁹⁾。黑格尔在其著作《美学》中提出,美学的研究范围包括了“美的广大领域,更进一步说就是艺术,也即是说它的范围是美的艺术”,他将美学等同于“艺术哲学”⁽¹⁰⁾。康德则重点研究了审美判断力,认为审美判断是一种人的内心能力的和谐一致。康德之后的审美心理学派则从审美距离、内感官等方面进行研究。综上,我们可以将美学的研究主要分成感性(包括情感、审美心理等)和艺术哲学研究两部分,而这部分又是相互联系的,如黑格尔所说的美是“理念的感性的显现”⁽¹¹⁾。

那么人工智能领域,可不可以成为美学研究的对象呢?笔者认为,这是可能的。首先从人工智能的定义来看,人工智能就必然和哲学、美学相联系,因为所谓智能,除了包括“需求和应用知识、推理和思考的能力”⁽¹²⁾,还包括“通过五种感觉来感知”“情感体验”等能力⁽¹³⁾。智能就是人类的思维能力,科学家将人类思维分成三种,“抽象(逻辑)思维,形象(直感思维)和灵感(顿悟)思维”⁽¹⁴⁾。人工智能如果想要真正成为智能,就不能仅仅模拟人类的抽象思维能力,即推理和逻辑能力,还要模拟人类的情感、感性、创造等与形象、灵感有关的能力,而后者正是美学所研究的对象。

其次,从人工智能在情感、感性方面的应用来看,随着人类与计算机交互关系越来越紧密,人们已经越来越重视人工智能情感的重要性。曾经有学者认为,智能机器只不过是帮助人类工作的工具,因此让它具有情感系统和体验既困难又没有必要⁽¹⁵⁾,但是现在人们已经认识到机器具有某种程度上的情感是有助于其工作的,“情感计算机不仅能够在协助人们提供更好表现,还能提升计算机做决策的能力”⁽¹⁶⁾。例如智能系统在与顾客交流时,可以通过顾客的服装打扮、面部表情、肢体动

作来分析顾客的情感,如果智能系统也要具有相应的情感传达,就能使得双方交流更加平等顺畅。因此,如何设计情感程序成了如今人工智能的一大难题。

再次,从人工智能在艺术方面的应用来看,人类借助人工智能计算机创作艺术、甚至编写智能程序,让智能机器在一定程度上独立地进行艺术创作,都已经成为现实。智能机器已经不只是艺术创作的工具,而是核心了⁽¹⁷⁾。而使用计算机进行创作或者设计程序让计算机某种程度上独立创作的艺术品,当然也可以成为美学的研究对象。而艺术风格、计算机艺术的独特性、人与计算机在艺术活动中的关系(创作者、工具、欣赏者)都属于美学研究范畴。

最后,在人工智能模拟人类的感性和艺术创作的时候,也需要对人类情感和艺术本质进行研究。人工智能在模拟人类的推理和逻辑时,探索出了非经典逻辑、不确定性推理等,为人类的常识推理机制进行了有益的研究。同理,智能体在模拟人类情感时,也需要对之进行系统的分析,例如情感产生的机制是什么?它与人体机能、与环境的关系、与性格和文化的关系如何?情感是如何在艺术中表达和传递的?审美愉快感是什么?在艺术本质方面,人工智能是否能真正地像人一样创造艺术品?人工智能创造的艺术品与人所创造的艺术品是否有本质的不同?相比起机器而言,艺术的人类本质在那里?

综上,我们可以初步得出,人工智能美学这一概念是可以成立的,它有其研究的合法性、对象和范围。在国外,有学者从人工智能计算机与美学的关系方面给出了一些相关定义,如费希维克(P.Fishwick)提出了“审美计算”(aesthetic computing)这一术语,“我们将审美计算定义为艺术对于计算机领域的理论和实践”⁽¹⁸⁾。不过审美计算重视的是如何从计算机层面来研究艺术和情感的程序化,如“1. 使用定制的表述程序和数据结构,有文化意义的特殊符号;2. 将艺术方法整合进典型的计算机强化行为中,如科学的可视化;3. 提升与计算机的情感和文化层面的互动”⁽¹⁹⁾。此外,有关计算机艺术(computer art)的研究也都只是关于计算机所创作的艺术本身的鉴赏和风格的探讨。⁽²⁰⁾

笔者试图从美学角度出发,结合其他专家在人工智能哲学、审美计算、计算机艺术等相近领域所做出的成果,为“人工智能美学”(the Aesthetics of Artificial Intelligence)做一个初步的定义:人工智能美学研究的是在人工智能技术发展过程中所出现的与美学有关的一些问题,其主要内容包括人工智能对人类感性(包括

情感)和艺术的模拟、人工智能艺术的风格与鉴赏、人工智能视野下人类情感和艺术本质问题等,其方法主要是哲学美学的,并结合诸多跨学科如脑科学、神经科学、生物进化等理论以及人工智能领域最新进展来进行研究。⁽²¹⁾

以上定义显然是不全面的,随着人工智能领域的研究发展推进,一些现在研究问题以后可能并没有研究的必要,例如图灵在驳斥那些反对人工智能的观点时,曾经认为一些所谓的“超感官能力”(如心灵感应、透视、预测未来、远距离致动)是人工智能无法实现的⁽²²⁾。而一些目前所无法预测的内容可能会成为重点,例如模拟人脑神经系统的连接主义和模拟人类进化的进化主义,在符号主义盛行时代都是闻所未闻的。所以,这样一种概念的内涵和外沿也必定随着时间的推移而不断发生改变。但是,其基本的两大内容,即对人类感性特别是情感表达,以及对人类艺术行为的模拟,应该在相当长时间之内是人工智能美学研究的对象,人工智能到底能不能实现对人类情感的最终模拟,人工智能最终能不能像人一样创造艺术品?这种人工智能要彻底地模拟人类智能,需要解决什么问题呢?鉴于技术在不断地发展,生物克隆技术已经能够完全创造出生命体来,所以,在此笔者将人工智能载体限制为无生命的机器,以此为前提探讨。下面笔者就准备从高级人工智能机器是否具有情感和意向性来进行初步探讨。

二、人工智能美学如何研究

1. 人工智能艺术现状

人们以计算机为工具、甚至编制智能程序进行艺术创作,已经不是新闻。在计算机视觉艺术方面,早在1952年,拉珀斯基(B.Laposky)就用示波器创作了《电子抽象》(Electronic Abstractions)的作品;随后在德国诞生了第一个计算机绘画程序;在1965年,第一幅真正意义的计算机绘画作品完成;而美国艺术家科恩(H. Cohen)设计了艺术创作软件“Aaron”,该软件所创作的绘画已能具有自己的风格。在国内,天津大学的孙济州教授通过“中国水墨画效果的计算机模拟与绘制系统”,用计算机模拟生成了水墨画效果,而浙江大学的徐颂华通过对中国书法和绘画进行元素分解,并设计出相应的程序,使计算机能够模拟生成书画图像。⁽²³⁾在文学写作方面,网上早就有了生成诗歌的程序,通过对词语的联想搜索和拼接,能创作出看似人写的诗歌,而百度于去年推出了其研发的智能体小冰的诗集,据

说在诗歌论坛发布时已能以假乱真。而音乐因其形式化最强,早就被人工智能成功模拟,80年代美国音乐教授库柏(D.Cope)设计了名为“音乐智能试验”的程序,通过分析和提取音乐大师作品中的旋律特征并加以重组,使得智能体创作的音乐几能与大师作品媲美,而最近索尼公司设计的程序则可以创作出酷似披头士乐队的作品。其实分析这些创作艺术品的智能程序可知,它们是利用了专家系统(Expert Database)和机器学习(Machine Learning)技术,存储并分析大量艺术家作品,寻找其中有规律性的特征,然后再加以重组,而基于神经网络的深度学习(Deep Learning)技术能够让计算机程序实现不断累积和更新,从而掌握更好的解决问题的方法,围棋智能机器AlphaGo的原理也与此类似,只不过在此基础上利用神经网络基础,优化了搜索技术和决策机制,其基础还是存储的大量棋谱和自我训练学习⁽²⁴⁾。

很明显,这种方式仍然只是一种初级的模拟,仅仅将已有的人类材料进行重新组装变形而已,而真正的人类艺术要求有情感的参与和传达、需要想象力和创造性,仅仅利用存储的艺术家作品再加以分解和重组,未免显得生硬而且缺乏创造性。如今一些智能专家,已经开始试图为情感、创造性提供了可以形式化的算法,那么我们就来分析一下,人类艺术创作中的情感真的能被形式化和符号化吗?

2. 艺术中的情感(emotion): 感情计算与情感语法

神经科学理论认为,“情感是一种大脑的高级功能,而且并不是某个‘中枢’可以独立完成的,参与情感的产生和表达可能是一个结构和机能相互联系的回路”。⁽²⁵⁾ 大脑中的Papez环和杏仁核团都可以影响和调节情感,情感是大脑皮层中的边缘系统比较发达之后的产物,所以只有进化到一定程度的动物才出现情感表达⁽²⁶⁾。虽然从解剖学、脑科学方面对情感的物质基础可以进行一定的解释,但是情感到底如何产生的,不同情感是由脑区、还是由化学分子决定的,现在科学暂时都无法解释。

艺术与科学、哲学等不同的地方在于,艺术品的创作和欣赏都有人的情感参与其中,艺术中所传达的并不是简单的信息,而主要是艺术家的情感、文化和思考。但是,情感毕竟是通过艺术形象来传递的,两者之间并没有绝对的对应关系。一个图形,仅从色彩、形状、线条等来分析,很难肯定地将之与某种情感对应起来。这样的话,就为机器创造艺术提供了可能,因为机器只需要储存和模拟现有艺术家的艺术形象再加以重新组

合,如果程序设计得足够合理、能够遵循一定的艺术逻辑的话,计算机似乎是能够创造出以假乱真的作品的。

因此,如果仅从行为主义来解释人工智能艺术的话,智能体创作的艺术品似乎和人类艺术品没有太大的区别。但是如果从艺术创作主体的感觉来考察,恐怕就会出现問題。因为智能机器人并没有任何感觉和情感,它的创作是一种不涉及到任何生物学反应的数据的输出。因此,我们是不可能从艺术品中感受到真正的艺术创作主体情感的。人与其他生物不同在于,人是社会动物,而艺术是一种“社会的表现”⁽²⁷⁾,正是艺术这种社会活动,使得人与人之间的情感和文化的交流成了可能,艺术品是艺术家与观众、观众与观众之间的情感交流纽带。诗人艾略特认为,诗歌是“感情的载体”,诗人会与读者分享情感,如果写的诗歌无法为其他人所欣赏、所分享,那就是无用的⁽²⁸⁾。古希腊哲学家柏拉图认为诗人的情感会迎合观众的情感,所以他批评这种悲剧艺术⁽²⁹⁾。与之相反,亚里士多德则认为悲剧可以帮助人们实现情感宣泄⁽³⁰⁾。康德认为,审美活动中的愉快感是人类普遍情感,这种情感就是一种“生命的情感”⁽³¹⁾。接受美学代表耀斯(H.Jauss)则认为,审美活动中以快感为主的主体经验可以“开辟了通向互为主体经验的道路”⁽³²⁾。因此,真正要想像人类那样创造艺术品,人工智能也必须能够具有情感、能够表达情感才行。

目前来看,智能专家解决计算机情感问题的方式主要有内外两种方法,一种是经验归纳法,即通过计算人类情感所呈现出来的表面特征,来建立一个情感的可数据化的标准,这是从外部来设计情感的方法,其代表就是“感情计算”。另一种方式是演绎法,即从人的内在感官反应来简化、形式化情感,可以视作从内部出发模拟人类情感。

皮卡德(R.Picard)提出的“感情计算”(affective computing)试图利用计算机强大的储存、搜索和运算能力,来计算、分析与情感相关的外在表现,如面部表情、心跳速率、皮肤温度等生理特征,以及艺术家在作品中留下能体现情感的相关痕迹,如色彩、形状、线条、文本等⁽³³⁾。笔者认为,这种对情感的计算发挥了机器的计算优势,不失为一种可行的方法。但是,计算机虽然能够得到情感与艺术表现手法的一些公式,却没有真正地将情感与表现真正地联系起来,因此,它无法去创造性地运用情感来表现,也难以去理解新的情感表现手法,最重要的是,计算机仍然不具备情感。

有些智能专家认为,可以用形式化符号语言来

表达情感,如斯洛曼(A. Sloman)相信,情感“是由动机激发因素产生的状态,同时包含着新动机激发因素的产生”⁽³⁴⁾。这种动机的激发、生成、妨碍、识别等可以用符号语言来表示。他认为,身体的感受并非情感的最重要决定因素,情感是源于一种整体性的认知结构,比如行为和外部原因比起身体原因更能引发情感,因此,斯洛曼得出结论:“所以我们用像‘害怕’‘失望’‘大喜’‘狂怒’或‘极度悲伤’这样一些术语,来描述一个异己者或没有生理反应的、高度精密的机器人心态,是合理的。”⁽³⁵⁾但是斯洛曼也承认,人类情感是非常复杂的,对于愉快和痛苦以及幽默都是目前难以解释的。斯洛曼的情感是建立在动机论基础上的,但是,我们知道,有些情感可能并非产生于某种特定动机,比如康德所说的审美愉快感,与有功利性的快适和善中的愉快相比,这是一种无功利性的、无目的的纯粹形式化情感⁽³⁶⁾。

脑科学认为快乐情感与脑内分泌的多巴胺物质有关,但是到底是因为快乐而分泌多巴胺,还是因为分泌了多巴胺才产生的快乐,具体生成机制还不太清楚。

康德在《实用人类学》中将高兴(痛苦)定义为:“高兴是生命提升的感觉,痛苦是生命阻滞的感觉。”⁽³⁷⁾审美愉快(不愉快)也是如此,由此,康德将这两种人类最基本的情感与生命力(状态)的提升与阻滞关联起来。可以设想,智能专家会说,我们可以将这种提升和阻滞转化为二元运算形式。我们可以先不去管人类情感产生机制的根本原因,而只是按照行为和结果来设置智能体,让智能体在面对相同的条件之下,产生相同的结果,这也就是当前人工智能常用的方法,这种方法或许在实际操作上是有可能性的,但是仍然没能从根本上解决智能情感问题,既不能让人工智能像人那样真正地产生情感,也很难照顾到情感的多样性和复杂性。

除此之外,艺术的创作和欣赏中,也并非仅仅只有情感在起作用,还有创作者和欣赏者的性格、体验、文化等方面的原因,或许还有黑格尔所说的“精神”(Geist)或者是哲学家德勒兹(G. Deleuz)所说的“在场感”⁽³⁸⁾等等。

综上,即使人工智能或许能够有一天模拟生成情感,甚至也能够艺术品中表现出一定的情感来,但这是否就意味着人工智能真正有了情感呢?还是说只是一种简单地执行指令?是否意味着人工智能是真正地在进行艺术创作呢?真正地创作,至少要意识到自己在创作,而不是像猩猩那样在人的训练下无意识地涂鸦,

或者是蜜蜂在空中摆出的轨迹,或者仅仅是一朵植物开出的花。那么人工智能是否能够有意识地创作呢?

3. 艺术创作的意向性:图灵测试 VS 塞尔的心灵哲学

哲学家、心理学家用“意向性”(Intentionality)来指人类有意识地做某事,是意识主体对对象的关涉或者指向。

哲学家塞尔用一种思想实验“中文屋”对图灵测试提出了反驳。他设想有一个人在屋里,完全不懂中文,但是可以根据一本标有简单对应中英文符号的手册,写下中文,以回答屋外人的中文问题,从而让人误以为他懂中文。塞尔的结论是,虽然从行为上,也许数字计算机能够做出和人一致的活动,但是就其过程而言,机器并不理解为什么要这么做,它的一切行为都只是遵照指令和程序而已。塞尔认为,只有机器真正地意识到了自己的行为以及价值,才能被认为具有智能了,也就是说智能的标准就在于是否具有意向性。如果我们将图灵测试与塞尔中文屋的争论应用于艺术上,即问答结果用艺术形象等方式表现出来,例如让被试者画一幅具有悲伤表情的画,那么以现在人工智能的水平,已经有一定概率可以蒙骗过测试者了。但是能够画出这种图像的人工智能真的知道它在做什么吗?如果说没有意识地去绘画(没有人能够真正没有意识地画画,即使达达主义也无法完全做到),那么这种绘画和自然现象、动物无意识所作出的图像有什么区别呢?所以人工智能要真正地创造艺术品,它就必须意识到自己在创作。

塞尔认为:“意向性是某些心理状态和事件的特征,它是心理状态和事件(在以下这些词的特殊含义上)指向、关于、涉及或表现某些其他客体和事物的特征。”⁽³⁹⁾意向性包括两个方面,一个是思想如何指向外在对象的,一个是我们如何确信正在发生的事情⁽⁴⁰⁾。比如说我在画画,一方面是我的行为指向了画画这件事,另一方面则是我确信我正在画画这个行为。正是由于意向性,使得我们能够与外界发生联系,而且能够确信这种联系,从而形成系统性的意识和知识。塞尔的意向性理论是建立在其身心理论之上的,他持有一种“生物自然主义”(Biological naturalism)思想⁽⁴¹⁾,实际上是将所有心灵(mind)活动归结为身体(大脑)的生物活动,因此,意识的产生也就是大脑神经元活动的结果,“意识和意向性的基本形式是由神经元的行为引起的,并由脑部系统所认识到,而该系统也是由神经元构成的”⁽⁴²⁾。心灵的意向性是语言的意向性的基础,

“意义是一种意向性”⁽⁴³⁾,因此语言所具有的丰富的意义实际上是建立在意向性基础之上的,最终也就是建立在一种神经生物反应的基础之上。正是这种生物学构造的不同,使得塞尔能够质疑由硅、金属这些物质构造的数字计算机,能否真正地产生意识,同时,机器也难以理解具有丰富语义的语言。“任何计算机程序自身不足以使一个系统具有一个心灵。简言之,程序不是心灵,它们自身不足以构成心灵。”⁽⁴⁴⁾ 塞尔可以说抓住了人工智能计算机与人类最大的区别,就是计算机没有生物性,而塞尔以及生物学家认为,人类所有的意识和思维都是以生物性为基础的。

在塞尔之前,图灵在提出测试的时候就已经反驳过这种挑战了。当时有人批评到,机器如果只是通过“符号的偶然坠落”,而不是因为“思想和情感”,不知道它做了何事,它就不能算作有情感。图灵反驳道,如果说计算机要像人一样有意识地去完成某件事才能算是有智能,那么同样,也没有人知道做此事的主体是否是有意识的,除了他自己。也就是说,如果不从行为上判断,而从动机上判断,他人是没法判断有没有意识,这种意识只能导致一种“唯我论”⁽⁴⁵⁾。

图灵的想法代表了大部分智能学者的观点,博登就认为,虽然塞尔说人类大脑的神经元活动产生了意向性,但是到底如何产生的,科学家也说不清楚,既然如此,我们也不能证明神经元蛋白就比硅、金属之类的物质更有可能产生意向性。我们也不清楚意向性到底怎么形成的,或者说意向性是什么,甚至有没有塞尔所谓的意向性也是一个问题。而且,人工智能所要实现的就是智能行为和活动结果本身,至于它本身是如何运作的,并不是决定性的,“从心理学的观点来看,重要的并不是生物化学性本身,而是建立在它之上的担负信息的功能”⁽⁴⁶⁾。虽然智能学者从现实角度出发,认为达到目的、解决问题就可以了,但是人类意识的内容和产生机制,我们仍然需要弄清楚,因为人工智能对意识的模拟,不仅仅是一种实际应用,它也是人类探索思维科学的重要途径。

在塞尔写作《心灵、大脑与程序》一文时,人工智能研究仍然以符号主义为主,即主张人类思维和认知可以用符号运算来实现,塞尔也明确了其批判对象是数字计算机,他也设想过,如果能够建造一台与人类的神经系统充分相像的计算机,并且能够用某种化学机制产生意识,那么也可能会产生出智能来。而在上世纪 80 年代后期基于神经系统研究的连接主义的兴起,使得模拟人类神经系统的工作又迈进了一步。1943 年

麦卡洛克(W. McCulloch)和皮茨(W. Pitts)提出了一种神经网络的模拟机制,他们认为神经活动的生理关系与命题的关系是对应的,因此,“对任一神经元的每个反应,都存在一个对应的简单命题陈述”。他们将神经网络的工作机制简化为神经元的“刺激”与“抑制”,对应于计算机的二进制工作机制。⁽⁴⁷⁾ 这种方式,将人类神经网络的复杂性大大降低了,使得计算机模拟人脑在一定程度上成为可能。而后来的专家系统和深度学习,都是得益于这种神经网络的广泛运用。在连接主义之后,出现了持机器进化思想的行为主义学派,1991 年麻省理工学院的布鲁克教授成功研制了能在未知的动态环境中漫游的仿生机器虫。但是不论是连接主义还是行为主义,都只是从外部(如反应、行动)来进行模拟,而内部具体如何实现的,仍然解释不清楚,内部成了“黑盒”,因此也难以分析人类认识和意识的真正生成原因和工作机制了。如果按照塞尔的标准而言,任何机器,无法自己生成意识和意向性,那么也就无法意识到自己是否在创作艺术,也就不是真正地实现智能了。

其实无论是图灵等智能学者坚持的行为主义,还是塞尔所坚持的自然生物主义,都无法真正地解决一个哲学问题,那就是主客关系以及由此引发的身心关系问题,即主客如何统一、身心如何统一。塞尔坚持心灵源于人脑、也即源于神经生物学系统的身心统一论,而智能学者则自相矛盾,一方面说机器和程序本来就是统一的,另一方面,又说程序可以模拟人脑,并不需要考虑到机器。对于哲学家而言,无论是生物体、机器还是程序的存在,都是无法证明的。唯一具有明证性的,就是笛卡尔(R. Descartes)所说的“我思”,即我可以怀疑一切,但是却无法怀疑我正在怀疑这一行为。“我思”发展到胡塞尔(E. Husserl)就是现象学意义上的“意向性”,这种“意向性”并非塞尔意义上的心理学活动,而是一种悬隔了一切未经证明的事物所剩余下来的“纯粹意识”⁽⁴⁸⁾,它既是对象的直接给予方式,也是主体的主动关涉,主客观在这里实现了统一。因此,在很多哲学家看来,图灵所批评的唯我论,恰恰是人类认识以及事物存在的起点。而身心关系,也是一种主客体关系,我们的心灵、精神甚至灵魂是如何产生的? 单纯由物质组成的身体为何能产生非物质的心灵实体? 除了心灵之外的所有事物有没有可能只是我们心灵的假象? 就像电影《黑客帝国》里所设想的,我们可能有一天被机器人所控制,我们所感知的世界其实只不过是一种幻象。

虽然目前仍然没能完全解释人类意识的产生以及运动规律,但是,科学表明,人类思维包括意识都是在生命有机体经过数亿年的进化中产生的,生命是由细胞构成的,而细胞的生殖和变化是一切生命活动的基础⁽⁴⁹⁾,这种由细胞构成的有机体和由硅、金属等构成的无机体之间是有根本区别的,例如运动、生长和衰亡。正是生命体的运动,使之在与环境的适应下,产生了进化。人的进化既有自然进化和也有社会进化,两者互相影响和促进⁽⁵⁰⁾,正是在这两种进化的合力下,人类产生出了意识,从而最终产生了包括艺术在内的诸多文化形式,“‘文化’是被印刻在人类有机体之上的,并且最终无法与这种生物区分开来,或者说人类的现象类型(phenotype)——自然性区分开来”⁽⁵¹⁾,文化是身体在环境影响下产生的,又反过来通过对人类行为的塑造和对人类的生活方式(饮食、生产、疾病)等的改变,影响到人体成分的生长状态和人类增长情况,从而影响到人类基因。⁽⁵²⁾因此,我们很难将意识、艺术以及文化的产生与人类有机体这个载体区分开来。如果真是如此,那么以无机物质为载体的计算机,如何能够真正地产生意识并且有意识地创造艺术呢?黑格尔将艺术美和自然美区分开来,认为艺术美是理念发展的产物,人们在艺术品中灌注了生气,使得艺术品有了自己的“精神”,这是艺术品和自然根本的不同,艺术品“绝对不是一种自然产品,也绝不会按照其自然方面而具有自然生命性”⁽⁵³⁾。

三、结论与展望

目前看来,情感和意向性是哲学家和智能专家关于人工智能是否能思维的争论焦点。人工智能强大的计算能力,是否能够完全计算模拟出人类大脑的功能,是否可以让机器自己产生如人一样的意识,现在恐怕仍然难以下定论。笔者倾向于认为人类智能有赖于其有机生命体不断进化发展而产生的,而人工智能的非生命体难以产生出情感、意向性等人类思维,因此也就不能说真正地创造艺术品,目前人类设计智能程序所创造的艺术品,仍然是属于人的作品,即使有一天,人工智能真的脱离人类操控创作,它也是无意识的,和动植物的无意识展示没有区别。

不过,目前科学家也开始了“人工生命”(artificial life)的研究并获得了一定的进展。1987年克里斯·兰顿(Chris Langton)提出人工生命概念,“人工生命是研究显示自然生命系统的行为特征的人造系统”⁽⁵⁴⁾。如果说传统的符号主义人工智能模式是自上而下的,即

是由统一的中央处理系统去发布命令,然后执行指令,那么人工生命模式则是一种从下而上的,它摹仿的是自然生命个体在实际环境中所遇到的情况,以及生命体与环境互动的行为,这种方式被称作“以个体为基础”(individual-based)⁽⁵⁵⁾。人工生命研究也试图解释生命和思维之间的关系,与认知科学着重于探究人类思维如何从神经元的生物学行为产生不同⁽⁵⁶⁾,人工生命则想揭示的是生命如何从无机物中产生的。在2014年,依照人工生命设想而研制出来的仿生机器虫,已经可以模仿某种低级爬虫的行为了。说不定有一天,具有和生命一样功能的机器智能也可能会出现,而到那一天,我们如今的人工智能美学又将重新改写了。

可以想见,在未来的日子里,人工智能美学研究的对象和范围肯定会一再发生变化,因为人工智能本身就在不断地改变,人工智能从符号计算到神经网络、再到仿生进化,不同的智能设计理念必然会影响到美学研究;而关于常识推理、混沌逻辑、不确定性逻辑,则会让人工智能可能模拟出人类日常语言甚至是艺术语言;按照不同方法设计出来的人工智能,也可能具有某种意义上的个性,而网络智能的兴起,或许是人工智能社会性的体现;人工智能的材质也可能发生改变,有机体与无机体可能结合起来,甚至人与机器结合起来;而如果真正实现了人工机器生命,实现了从无机物生产出有机生命,那人类历史都终将改写。

但是,无论技术如何改变,人工智能美学的基本出发点和关注重点不会变,它始终是以人为本,从人与人造物的关系出发,来研究人工智能对人类思维和行为的模拟,其最终目的是更好地认识人类情感和思维本身,是为了给人类创造更美好的生活。

[基金项目:本文系国家社科基金一般项目“阿多诺哲学中的语言思想研究”(项目编号:16BZX118)的阶段性成果。]

注释: -----

- (1)(2)(22)(45) Turing, A. "Computing Machinery and Intelligence", *Mind*. 1950, pp.433-460. p.433, p.433, p.453, p.446.
- (3) Boden, M. *Artificial intelligence*, San Diego: Academic Press, Foreword.XV.
- (4)(15) Hawkins, J. and S. Blakeslee. *On Intelligence*. St. Martin's Griffin, Reprint edition, 2005, p.142, p.140.
- (5)(6) Russell, S. and P. Norvig, *Artificial Intelligence: A*

- Modern Approach*, Second Edition, New Jersey: Pearson Education, Inc., 2003. p.2, p.1020.
- (7) (34) (35) (46) (47) [英]博登编著:《人工智能哲学》,刘西瑞、王汉琦译,上海译文出版社,2001年版,第2页,第325页,第330页,第127页,第31页。
- (8) J.McCarthy, “What has AI in Common with Philosophy?”, 1996, <http://www-formal.stanford.edu/jmc/>.
- (9) [英]鲍桑葵:《美学史》,张今译,商务印书馆,1985年版,第241-242页。
- (10) (11) (53) Hegel, G.: *Werke 13: Vorlesungen über die Ästhetik I*, Frankfurt am Main: Suhrkamp, 1986, p.13, p.151, p.57.
- (12) Goertzel, B., C. Pennachin, *Artificial General Intelligence*, Berlin Heidelberg: Springer-Verlag, 2007, p.6.
- (13) (17) (20) S.Wilson, “Computer Art: Artificial Intelligence and the Arts”, *Leonardo*, Vol. 16, No. 1 (Winter, 1983), pp.15-20, p.17, p.16, p.15.
- (14) 钱学森:《关于思维科学》,上海人民出版社,1986年版,第129页。
- (16) (33) Picard, R. *Affective Computing, M.I.T Media Laboratory Perceptual Computing Section Technical Report*, No. 321. <http://www.media.mit.edu/~picard/>, p.1, pp.12-13.
- (18) (19) Fishwick, P. *Aesthetic Computing*, Cambridge: The MIT Press, 2006, p.6, p.6.
- (21) 关于人工智能美学的定义,笔者曾在2017年6月28日于德国波恩大学(Bonn University)哲学系以及计算机系举办的两次讲座“Can AlphaGo Create Authentic Artworks?—The Preliminary Study on the Aesthetics of Artificial Intelligence”中提出,与参会哲学和计算机专家进行了讨论。<https://www.ioa.uni-bonn.de/de/abteilungen/sinologie/bonner-sinologisches-kolloquium>
- (23) 徐颂华:《中国书画艺术电子化创作的初步算法性探索——美、智能与计算》,浙江大学博士论文,2007年。
- (24) Silver, D. and etc., “Mastering the game of Go with deep neural networks and tree search”, *Nature*, vol.529, 2016, p.484.
- (25) (26) 孙久荣:《脑科学导论》,北京大学出版社,2001年版,第231页,第234页。
- (27) [德]格罗塞:《艺术的起源》,蔡慕晖译,商务印书馆,1996年版,第39页。
- (28) Eliot, T. “The Social Function of Poetry”, *On Poetry and Poets*[M], London: Faber and Faber Limited, 1957, p.19.
- (29) [古希腊]柏拉图:《理想国》,郭斌和、张竹明译,1986年版,商务印书馆,第405页。
- (30) [古希腊]亚里士多德:《诗学》,陈中梅译注,商务印书馆,1996年版,第62页。
- (31) (36) (37) Kant, I. *Werkausgabe X: Kritik der Urteilskraft*[M], Hg.W. Weischedel. Frankfurt am Main: Suhrkamp, 1974, p.115 (A4), p.116 (A5), p.551 (A171).
- (32) [德]耀斯:《审美经验与文学解释学》,顾建光、顾静宇等译,上海译文出版社,1997年版,第49页。
- (38) Deleuze, G. *the Logic of Sensations*. Trans., Daniel W. Smith, London: Continuum, 2003, p.42.
- (39) (44) [美]塞尔:《心、脑与科学》,杨音莱译,上海译文出版社,1991年版,第110页,第30页。
- (40) (41) (42) Searle, J. *Mind: A Brief Introduction*, New York: Oxford University Press, 2004, p.159, p.113, p.164.
- (43) [美]塞尔:《意向性:论心灵哲学》,刘叶涛译,上海世纪集团,2007年版,第164页。
- (48) [德]胡塞尔:《纯粹现象学通论:纯粹现象学和现象学哲学的观念》第一卷,李幼蒸译,商务印书馆,1992年版,第99页。
- (49) 翟中和:《细胞生物学》,高等教育出版社,1995年版,第3页。
- (50) 李难:《进化生物学基础》,高等教育出版社,2005年版,第135页。
- (51) Marks, J. “Ten Facts about Human Variation”, *Human Evolutionary Biology*, ed., Michael Muehlenbein, Cambridge: Cambridge University Press, 2010, pp.265-277, p.273.
- (52) Ulijaszek, S. “Variation in Human Growth Patterns due to Environmental Factors”, *Human Evolutionary Biology*, ed., Michael Muehlenbein, Cambridge: Cambridge University Press, 2010, pp.396-404, p.397.
- (54) Langton, C. “A New Definition of Artificial Life”, <http://www.chairemetal.com/cm03/intro.htm>
- (55) Bedau, M. “Artificial life: organization, adaptation and complexity from the bottom up”, *TRENDS in Cognitive Sciences*[J], Vol.7 No.11, November, 2003, p.506.
- (56) O’ Callaghan, J. “Are we on the brink of creating artificial life? Scientists digitise the brain of a WORM and place it inside a robot”, *Mail Online*, 2014, Nov.27, 2014.
- (作者单位:南开大学哲学院、德国波恩大学)
(责任编辑:李明彦)