

# 人工智能与人类的原初问题

江 怡

(山西大学 哲学社会学学院、科学技术哲学研究中心, 山西 太原 030006)

摘要: 人工智能是 20 世纪人类科技取得的最为重要的成果之一。如何从源头上考察人工智能的性质, 是当代哲学研究的重要问题。追问“‘人工智能’是什么”的问题, 就是对人工智能的一种形而上学的追问, 也是对人类存在本身的追问。通过分析可以看到, 人工智能是人类探索自身存在意义的一种现代形式, 是人类重新认识自我的一次重要尝试, 也是对人类原初问题的一次重新回归。在其起源上, 人工智能概念来自近代哲学家们最初提出的模拟大脑思维的设想, 而在其未来发展上, 人工智能问题又是与人类的原初问题密切相关, 最终都要回答诸如“我是谁”这样的问题。唯有在尝试回答类似问题的努力方向上, 人工智能问题才能得到真正的理解。哲学的起点是对人类存在本质的思考, 人工智能以揭示智能奥秘的方式完成着哲学家们对人类存在的追问。如何回应人工智能的这种追问, 应当是当下哲学研究必须面对的紧迫问题。

关键词: 人工智能; 原初问题; 人类存在; 哲学的挑战

中图分类号: B01; B80-0 文献标识码: A 文章编号: 0257-0246 (2020) 01-0207-07

自从 1956 年达特茅斯会议以来, “人工智能”概念就已经成为人们用于指称能够模拟、扩展和开发人类智能的机器智能的专有名词。从历史上看, 这种机器模拟人类智能的想法早已存在于古代神话、中世纪传说和近代哲学家的论述之中, 在当代则主要来源于数理逻辑的巨大进步, 人工计算方式的革命使得理智推理的形式化成为可能。然而, 人们至今似乎对“人工智能”概念并没有形成统一的清晰认识。<sup>①</sup> 仅从智能模拟的角度看人工智能, 主要是一种功能主义的解释方式, 但这已经遭到许多哲学家和科学家的质疑。从实际应用范围看, 人工智能技术的确取得了显著成就, 但依然缺少对这种技术性质的理解。由于不同领域的研究者仅从自己领域出发去认识人工智能, 这就使得人工智能成为一个“百宝箱”, 无论如何解释似乎都可以得到我们所需要的结果。而在当代人工智能热潮中潜伏的最大危机, 则是对人工智能本身性质的模糊认识以及由此引发的对人类未来的担忧。这就迫使我们从源头上思考人工智能的性质问题。

## 一、“人工智能”是什么?

从字面上理解, 人工智能似乎是模仿人类智能的一种人工计算方式, 或者说, 是让计算机完成类似人类一样的理性思维。图灵在 1950 年的文章中设想, 计算机可以使得人类相信它不再是一台机器。

基金项目: 国家社会科学基金重点项目 (17AZX002)。

作者简介: 江怡, 山西大学哲学社会学学院、科学技术哲学研究中心教授, 研究方向: 西方哲学、分析哲学史、人工智能哲学。

<sup>①</sup> Haugeland, John, *Artificial Intelligence, The Very Idea*, Cambridge, Mass.: MIT Press, 1985; Charniak, E. & D. McDermott, *Introduction to Artificial Intelligence*, Reading MA: Addison-Wesley, 1985; Schalkoff, Robert J., *Artificial Intelligence: An Engineering Approach*, New York: McGraw-Hill, Inc., 1990.

霍格兰则提出,当机器可以像人类一样思考时,机器就有了人工智能。<sup>①</sup>查耐克认为,计算模式的运用就可以让机器像人类一样从事理性的思考。<sup>②</sup>然而,无论这些科学家们在何种意义上理解人工智能,他们似乎都没有触及一个最为关键的问题——“‘人工智能’是什么?”这个问题并不是简单地给出对人工智能特征描述或功能解释,而是对人工智能的性质给出说明,是对人工智能的一种形而上学追问。

在形而上学意义上,追问“是什么”的问题,意味着一种对规定性的要求。由于当代哲学已经放弃了柏拉图主义对本质确定性的要求,因此,对人工智能提出这种规定性要求似乎已经不合时宜了。但是,正如海德格尔所言,人类的存在本身始终处于被追问的过程之中,无论我们以何种方式展现着这样的存在。在今天,人工智能已经成为人类存在的一种独特方式,对人工智能技术的高度依赖成为当人类存在方式的一种重要标志。按照马克思的说法,存在决定意识,人类存在本身就是一个值得不断发问的对象。这样,我们对人工智能性质的追问,应当是一种对人类存在本身的追问。

从源头上看,关于模仿人类思维的设想最初出现在古希腊神话之中。例如,传说中的古希腊火神赫菲斯托斯(Hephaestus)的小金人和塞浦路斯国王皮格马利翁(Pygmalion)的美女加拉提亚(Galatea)。<sup>③</sup>这些事物都被赋予了人类的生命和智慧。在中世纪,赋予某些事物以人类思维能力则被看做制造神秘力量的重要方法,例如,炼金术就声称可以把灵魂置于某些事物之中而使其具有某种神秘力量。在历史上,从古代埃及、希腊和东方,直到近代的西方科学,人类从来没有停止过对自动机器(automatons)的想象。<sup>④</sup>但这些都不过是毫无灵性的机械产物,用于满足人类的某些特殊需要,虽然也有一些宗教信仰者试图赋予它们与人类一样的智慧和情感。

而人类真正提出人工智能的观念,是基于这样的信念,即人类的思维活动是可以被机器化的,或者说,机器可以模仿人类的思维活动。这种具有革命性的观念出自近代哲学家和科学家莱布尼茨,他明确提出,人类理性可以被还原为机器演算。代数和几何学的系统化使得这种还原成为可能。人类只需要按照推理规则做出形式推导,就可以完成逻辑的推理活动。霍布斯曾在《利维坦》中提出,理性不过就是计算。莱布尼茨则设计出一种“普遍语言”(characteristica universalis),试图以此把所有的论证都还原为演算。所以,在他看来,两个哲学家之间的争论如同两个会计之间的争论一样,是完全没有必要的。因为只要让他们拿起笔来,坐在石凳旁,相互说一句,让我们来算一下吧,这就足够了。这种观念的重要意义在于,它使得理性的计算变成了物理符号之间的交换,并最终使得形式演算的外在化和系统化成为可能。当然,人工智能概念被真正确立,则完全是由于20世纪的数理逻辑的建立。这被看做使得人工智能成为可能的革命性突破。这一革命首先来自于布尔的《思想的规律》和弗雷格的《概念文字》。在弗雷格思想的影响下,罗素和怀特海于1913年出版了《数学原理》,首次尝试以形式化的方式奠定数学的基础。受到他们的启发,希尔伯特提出了这样的问题——“一切数学推理都可以被形式化吗?”哥德尔的不完全性定理、图灵的机器思维和丘奇的 $\lambda$ 演算(Lambda calculus,即最小通用程序设计语言),都以不同的方式回答了这个问题。<sup>⑤</sup>

从以上历史描述中可以看出,人类思考智能模拟的基本出发点是为了尝试理解人类理智本身,把理智活动作为人类存在的一个主要标志,并以人类理智和情感作为衡量一切事物具有灵性的标准。这种“万物有灵论”的观念曾支配着古代和近代哲学家和科学家。虽然“万物有灵论”在近代哲学中被解释为一种“泛神论”并最后演变成为一种宗教信仰,但相信可以用人类的理智精神去解释万物运动的根本原因,这始终是人类对自身理性自信的充分表现,也是人类得以作为万物之灵长的主要依

① Haugeland, John, *Artificial Intelligence, The Very Idea*, Cambridge, Mass.: MIT Press, 1985.

② Charniak, E. & D. McDermott, *Introduction to Artificial Intelligence*, Reading MA: Addison-Wesley, 1985.

③ McCorduck, Pamela, *Machines Who Think (2nd ed.)*, Natick, MA: A. K. Peters, Ltd., 2004, p. 5.

④ McCorduck, Pamela, *Machines Who Think (2nd ed.)*, Natick, MA: A. K. Peters, Ltd., 2004, pp. 6-9.

⑤ Crevier, Daniel, *AI: The Tumultuous Search for Artificial Intelligence*, New York: Basic Books, 1993, p. 190.

据。当代人工智能思想来自“机器思维”的观念，从性质上看，依然是把人类思维作为衡量机器工作能否达到人类水平的主要标准。图灵的“模仿游戏”就是把“机器能否思维”的问题转换为一种推理方式，通过回答某个更为具体的问题（如“性别问题”），来判断一个机器是否可以得到与人类相同的结论。虽然塞尔曾以“中文屋”的思想实验反对图灵式的功能主义观点，但以行为方式作为判断机器思维能力的设想，无论是对哲学家还是对科学家，都依然具有极大的吸引力。例如，《人工智能哲学》一书的编者博登就在反驳塞尔实验的《逃出中文屋》一文中指出“重要的问题不是‘一台机器何时理解了某件事情？’（这个问题暗示存在着某个明确的断点，理解在那里终止了，这是一种误导），而是‘为了能够作出理解，一台机器（无论是不是生物的）必须能够做到哪些事情？’这个问题不仅关系到计算心理学是否可能，而且关系到它的内涵。”<sup>①</sup>

显然，“能够做到”是用于判断“理解”的主要标准。当代计算机科学正是以这种标准建立了模拟人类智能的机器演算，用各种算法代替人类推理，实现了人类智能所无法完成的许多工作理想。

然而，这种对人类智能的模仿首先是建立在对人类存在方式的理解之上的。按照这种观念，人类与其他动物的重要区别就在于人类具有的理智能力，即通过计算和推理完成人类自身对外部世界的理性认识和行为反应。这种理智能力，一方面是人类的天然自然能力，另一方面也是人类在与自然的联系中优于其他动物的能力。这构成了人类存在的一种特殊方式，即休谟所说的“有理性的存在物”，或者莱布尼茨所说的“具有天赋观念的存在”。正是凭借这种能力，人类可以用语言、社会和行动等不同方式展现自身作为一类物种在世界上存在的特殊性。以这种存在方式作为人工智能观念的最初起点，恰好说明人工智能并非一种单纯的技术上的革命，更是一种人类存在方式的革命，即把人类存在的理智方式外化为非生命的机器（计算机或学习机）的计算方式，用可计算性的形式化方法诠释人类存在的另一种意义。在这种意义上，我们可以说，人工智能是人类探索自身存在意义的一种现代形式，是人类重新认识自我的一次重要尝试，也是对人类原初问题的一次重新回归。

## 二、什么是人类的原初问题？

在美国科幻电影《黑客帝国》(Matrix)中，人类在与人工智能机器人的战争中失利，导致人类必须学会适应机器人，并以“先知”（研究人类心智的程序）作为导师而在“矩阵”或“母体” (Matrix) 中拥有选择的权利。但最终，人类的第六代“救世主”尼奥由于人类特有的爱情而选择自我牺牲，拯救了人类。在这里，爱情和选择被赋予为人类最初的存在方式：爱情是人类的情感本能，而选择则是人类的自由意志。这些似乎是在暗示，无论人类如何发展变化，人类最初的存在方式依然是人类的本能，但这种本能并非是生物学或生理学意义上的动物本能，而是人类作为一种情感存在的反思本能。当机器人无法完全控制人类时，机器人的重要工作就是试图去理解人类。而理解人类的首要任务，则是要回答这样一些对人类来说最为根本的问题“我是谁”“我从哪里来”“我为何存在”。当机器人可以理解并能够回答这些问题，就意味着机器人具有了人类同样的自我意识。机器人的自我意识觉醒既是机器人真正实现了对人类的理解，又再次证明自我意识对人类存在的根本意义。这些似乎表明，自我意识应当被看做人类存在的原初问题。

然而，当今科学实验证据证明，我们在其他动物身上也可以看到类似人类的自我意识的存在。例如，在与人类最为接近的猿人那里，自我意识的存在是显而易见的。<sup>②</sup>即使是在家养的猫狗等宠物身上，我们也可以很容易看到某些自我意识的表现。虽然这些自我意识是以现象意识的形式出现的，但

<sup>①</sup> 博登《人工智能哲学》，刘西瑞、王汉琦译，上海：上海译文出版社，2001年，第140页。

<sup>②</sup> Gallup, G. G. Jr., "Chimpanzees: Self-Recognition," *Science*, 167 (3914), 1970, pp. 86-87.

它们却构成了我们理解这些动物活动特征的重要成分。<sup>①</sup>可见,仅有自我意识的存在并不能完全区分人类与非人类的动物以及具有人工智能的机器人。真正构成人类存在原初问题的一定是其他的东西。我们还是从历史中寻找一下这个问题的踪迹。

应当说,西方哲学的历史就是人类自我认识的历史。从古希腊人面对世界变化而提出世界本源学说开始,人类就在不断追问本质或始基问题。古代哲学家们对这个问题的回答基本上是围绕人与世界的关系,无论是从人类视角出发还是从诸神视角出发。世界把人类置于思考的中心,但人类始终无法摆脱面向世界的命运安排。因而,哲学家们对人类存在原初问题的追问,就只能是借助于世界的存在而展开。柏拉图的“彼岸世界”、亚里士多德的实体观念以及阿奎那的神圣世界,都是以高高在上的客观世界的存在为前提。近代哲学家们开始自觉地把人类抽离于世界,让人类站在世界的主宰地位,用人类的眼光去审视这个世界。笛卡尔的心身二元论就是这种审视方式的最初尝试。无论是经验论者还是唯理论者,他们都把人类的经验或观念看作理解世界存在的入口。在这种意义上,西方近代哲学才在真正意义上开始了人类自我认识之旅。这种自我认识过程就是从康德批判哲学开始的对人类理性能力的全面考察,通过对绝对自我意识(费希特哲学)和绝对本身(谢林哲学)的批判,达到了黑格尔的绝对精神。而在这一过程中的核心问题则是,“我能够知道什么”“我应当做什么”和“我可以期望什么”,最后归结为“人是什么”这个问题。正是围绕“人是什么”这个问题,哲学家们的思考才真正进入了自我反思的阶段。

虽然人类达到自我反思阶段走过了这样一个漫长的路程,但自我反思本身却使人类真正回到自身。这就是人们常说的人类的“返乡”过程。“返乡”就是返回自身,也是回到人类存在的初始状态,回到人类存在的原初问题。这个问题就是“我是谁?”显然,这里的“我”不是具体的个人,它既不是具有身体的物质个体,也不是作为精神存在的自我主体。这里的“我”是指人类存在的总体,是对人类存在的自我发问。人们常说,几千年来的哲学都是为了回答这个根本的问题,但至今依然没有最终的答案。然而,这是一种错觉。事实上,“我是谁”并不是一个可以回答的真正问题,而是一种自我反思的提问方式,是一类具有本质要求的问题句法形式。历史上的哲学家们所有试图回答这个问题的努力,都不是为了得到这个问题的答案,而只是为了寻求认识人类自身的不同途径。

从句法形式上看,“我是谁”具有明确的主谓结构,对这个问题的回答似乎就必须满足这个结构的要求,即给出“是什么”的句子形式。但如果为了满足这个要求而给出这个问题的答案,就没有真正领会这个问题的真正含义。在句法上,这个句子结构是具有误导性的,似乎所有的回答(如果有答案的话)都必须满足这个问题的预设前提,即一定存在某个东西作为可以满足这个问题的答案。然而,如上所见,这个东西实际上并不存在。相反,“我是谁”是对人类自我意识的反思,是以一种完全开放的态度面对人类的存在本身。这里的“反思”与“开放”恰好表明了人类存在的基本状态:人类是以反思的方式区别于非人类的动物和具有人工智能的机器人,是以开放的态度面对一切进入反思中的对象和事物。从具有自我意识到进入自我反思,人类完成了认识自我的重大飞跃。由于自我意识并非人类与非人类的主要区别,因此,以确立自我意识作为人工智能研究方向就变得不具有重要的价值。换言之,制造出具有自我意识的人工智能机器人并不意味着人类被完全复制或对人类存在构成致命挑战。事实上,如何使这样的机器人具有自我意识,目前仅仅属于技术上的难题,而不是观念上的鸿沟。但要使这样的机器人具有自我反思的能力,则是一个难以实现的任务。“自我反思”在这里不仅意味着具有自我意识的主体对自身的认识,而且意味着这个主体对自我的反思能力。然而,由于人工智能技术主要是通过数据计算或通过建立人工神经网络而实现自主学习或深度学习的目的,其本

<sup>①</sup> 这里的“现象意识”(Phenomenal Consciousness)是与“取用意识”(Access Consciousness)相对立的,前者表征的是作为意识活动内容的知觉经验,后者表征作为判断意识活动存在的心理状态。这个区分来自当代美国哲学家布洛克,参见Block Ned, "On a confusion about a Function of Consciousness," *Behavioral and Brain Sciences*, Vol. 18, 1995, pp. 227-247.

身并不具有这样的反思能力。即使是人工智能计算机的无监督学习技术,也不过是对已有程序重新做出整理分配,使用不同算法(如聚类、自动编码等)或特征分离技术等,但无法实现对作为整体的机器本身的自我反思,即回答“我为何存在”或“我为何这样做”的问题。这也就意味着,我们无法把人工智能机器人看作与人类相同的“行动者”(agent),因为它们并不具有与人类相同的“能动性”(agency)。<sup>①</sup>在这种意义上,无论未来的人工智能如何发展,在性质上,它似乎永远无法具有人类的自我反思能力。

### 三、从人工智能到人类的原初问题

上文的最后一个论断,并不意味着人工智能由此无法完全模拟人类的智能活动,而只是为我们讨论人工智能与人类智能的关系,提供了另一个可选择的方向:由于人工智能无法完全实现人类智能的目标,因而,人工智能研究就会另辟蹊径,以人类智能的名义,实现着人类智能无法完成的任务。事实上,在完成智能任务方面,人工智能早已在许多领域超出了人类所能达到的极限,或者是为人类实现了许多其不可能完成的目标。例如,无人机的定点目标工作,帮助人类实现了远距离空中侦察和实施行动的任务;智能机器人在许多危险领域完成的任务也是人类望尘莫及的。在这些方面,人工智能已经很好地完成了人类智能的工作。或许,正是由于人工智能在这些方面取得的巨大成绩,使得人类产生了对人工智能最终取代人类智能的担忧。但从以上分析中我们可以清楚地看到,这种担忧完全是多余的。因为人工智能所做的事情,正是人类智能无需或无法完成的工作,因此,两者的关系应当是相互补充,而不是相互取代。然而,人工智能相比于人类智能所具有的某些优势,让我们不得不重新反思人类智能的根本性质,由此探讨人工智能的未来发展的突破口。

的确,当今对人类智能的反思主要来自于人工智能的出现。按照图灵的理解,人类智能的主要表征形式是计算。这里的计算不仅是指形式化的数字演算,而且是指人类本能性的推理活动。所谓“本能性的”是指人们不需要后天的教育或训练就可以完成的行为。由于作为人类智能特征的计算是可以机器模仿的,因此,人们就认为,人类智能也是可以机器模仿的。从计算主义到功能主义再到联想主义,人工智能领域的科学家们已经在这方面做出了许多努力,试图为实现人类智能的机器模拟找到更为恰当的方式。这个努力过程也是人类对自身智能的反思过程。目前,人们已经形成了对人类智能的基本理解,把这种智能归结为人类复杂的认知能力和高层次的动机(motivation)与自我觉察(self-awareness)。人类的认知能力是指学习、形成概念、理解、运用逻辑和进行推理等方面的能力,其中包括了识别形状、理解观念、做出计划、解决问题、做出决定、获得信息、语言交流等。<sup>②</sup>所有这些似乎都指向了机器模拟人类智能的可能性。然而,在人类智能中,更为重要的是获得经验和进行思考的能力。自我意识与自我反思都依赖于经验的获得和思考的结果,因而,人类智能的首要工作不是认知能力,而是心理能力(mental capacity)。这与人类的思维过程有关,也与人类应对外部对象的反应能力有关,更与人类的自我认识能力有关。这是一种内在于人类心灵的思维活动,也是人类获得外部经验时所感受到的最初的心理体验。但是,由于这种智能活动无法直接模仿,因此,人工智能的最初设想者就把智能表征(intelligent representation)作为模拟智能的主要方式,<sup>③</sup>这就把内在的

<sup>①</sup> Barandiaran, X. E., E. Di Paolo & M. Rohde, "Defining Agency: Individuality, Normativity, Asymmetry, and Spatio-Temporality in Action," *Adaptive Behavior*, Vol. 17, No. 5, 2009, pp. 367-386, 382.

<sup>②</sup> Hunt, Earl, *Human Intelligence*, Cambridge: Cambridge University Press, 2011; Sternberg, Robert J. & Kaufman, Scott Barry, ed., *The Cambridge Handbook of Intelligence*, Cambridge: Cambridge University Press, 2011.

<sup>③</sup> Nilsson, N., *Principles of Artificial Intelligence*, New York, Springer-Verlag, 1987; Ginsberg, M., *Essentials of Artificial Intelligence*, New York: Morgan Kaufmann, 1993; Luger, G. & Stubblefield, W., *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*, Redwood, CA: Benjamin Cummings, 1993.

心理活动表征为外在的命题演算和逻辑推理，因为前者是不可见的，而后者则是可观察的。人工智能正是建立在对内在心理活动的外在表征基础之上，通过计算和推理，完成了对人类智能的模拟任务。在这种意义上，人工智能就是一种对人类智能的外在表征形式。<sup>①</sup>

根据这种理解，人工智能对人类智能的表征，事实上是倒逼我们对人类智能性质的深入讨论。人类智能所包含的认知能力和目标动机等，在解决问题的能力上属于人类智能中的高级层次，需要复杂的计算才能完成。人工智能在解决这些问题方面不仅可以很好地模仿人类智能，而且在某些方面还可以超越人类智能。但人工智能在解决较低层次的人类智能问题时，却往往表现得无能为力。例如，面对经验中的现象意识所呈现的感觉性质，即“感受质”，由于个人的体验不同而使得机器无法处理这种性质。这就是人工智能中的“莫拉维克悖论”（Moravec's Paradox）。这个悖论是说，在人工智能领域，最困难的问题是最容易得到解决的，而最简单的问题却是最难解决的。莫拉维克说“要让电脑如成人般地下棋是相对容易的，但是要让电脑有如一岁小孩般的感知和行动能力，却是相当困难，甚至是不可能的。”<sup>②</sup> 同样，对于人工智能而言，解决有关演算和推理的问题是最为简单的，因为人工智能的设计目的就是为了解决这些问题；然而，要解决自我意识问题却是最为困难的，因为意识活动已经超出了人工智能的设计范围。这样，人类智能就有了人工智能在性质上无法取代的部分。

必须指出，我们这样说的目的，并不是要表明人工智能无法取代人类智能，而是要表明，通过对人工智能与人类智能关系的这种理解，我们看到了人类智能所面对的根本问题，也就是人类存在的原初问题。而这个原初问题的凸显，正是由于我们看到了目前人工智能发展的限度。只有从人类存在问题出发，我们才能真正理解人工智能对当代人类的重要意义。

#### 四、人工智能与人类的存在

从人类存在看人工智能的意义，首先意味着把人工智能作为人类存在的一种特殊形态。这是从当代哲学和科学发展的视角出发对人工智能的哲学定位。这一定位有两个主要理由。

其一，人工智能作为对人类智能的表征性模仿，试图揭示人类智能的心灵奥秘，用外在的计算和推理，度量内在的心灵活动过程。传统哲学的二元论在当代哲学中早已被消解，因而，当代哲学家们在处理心身关系时总是持有各种不同的一元论立场。早期的科学主义哲学家们（如马赫、罗素、维也纳学派等）倡导一种中立一元论，希望能够用独立于身心的中立要素去解释身心关系。欧洲大陆现象学传统则坚持一种意识一元论，用具有独立于身体存在的意识内容去说明身心关系的变化。当代心灵哲学家们则普遍持有一种外在论的一元论，坚持用物理主义和现象主义去描述身一心状态。当代人工智能哲学家们对身心问题的处理采用的是心灵哲学家们的路径：人工智能就是用外在的表征方式说明存在于人类内心的理智活动。但这种表征方式与心灵哲学家的理论主张不同，它以可观察的计算和推理（即计算输出和推理结果）表征（模仿）了人类的理智过程。人工智能的目的不是为了简单地模仿，而是为了使人类智能得以重新呈现，是人类智能的外在表征形式。因此，在这种意义上，人工智能就是人类智能存在的另一种形式。

其二，当代科学发展已经越来越清楚地表明，一切科学理论和技术进步都是为人类的存在提供更加便利的条件，也是为人类的未来提供更加可靠的保障。无论是宏观的黑洞理论，还是微观的量子学说，或者是中观的人工智能，这些都是对人类存在的意义提供说明。它们的意义与外部对象的存在无

<sup>①</sup> 围绕这种表征，在心灵哲学家和认知科学哲学家们中引发了激烈争论，其中的主要观点是“内在主义”与“外在主义”的对立。由于该问题涉及范围极广，本文无法在这里详细讨论。参见 Silvers, Stuart, ed., *Rerepresentations: Readings in the Philosophy of Mental Representation*, Dordrecht: Kluwer Academic Publishers, 1989; McLaughlin, Brian and R. Cohen, eds., *Contemporary Debates in the Philosophy of Mind*, London: Blackwell Press, 2008.

<sup>②</sup> Moravec, H., *Mind Children*, Cambridge, Mass.: Harvard University Press, 1988, p. 15.

关,甚至与它们所研究的具体问题无关。它们都把人类存在本身设定为理论前提和问题讨论的出发点。黑洞理论是要揭示在宇宙空间中存在的黑洞现象对人类所在的地球天体的长远影响;量子理论说明了微观世界中量子如何构成了我们所认识的具体物质形态;人工智能则是以模拟人类智能的方式为人类存在提供更为广阔的发展前景。在这种意义上,人工智能对人类存在本身所具有的重要意义,无论如何都不能低估。对此,人工智能的科学家和哲学家都给予了充分的说明。<sup>①</sup>这恰好表明,人工智能表征了人类存在的一种特殊方式。

从人类存在看人工智能的意义,还意味着人类理智对自我限度的挑战。从人类智能的发展看人工智能的限度,即人工智能能够做和不能做的事情,这已经在前面得到了说明。但从人类存在的视角看待人工智能的意义,则更是对人类理智本身的挑战。这也可以从两个方面加以说明。

其一,人工智能是以模拟人类智能的方式表征了人类存在的基本形式,即以计算和推理作为人类理智的重要特征,通过对人类计算和推理能力的形式化,把人类作为万物之灵长的鲜明特征凸显出来,并由此构造了接近人类理智的人工智能。这既是对人类智能的完美模仿,也是对人类智能范围的准确刻画。正是由于这种刻画,人类智能表现出了自身的限度,即在运算速度和解题方式上逊于人工智能,这在“深蓝大战”和“阿尔法狗”获胜的案例中已有体现。不仅如此,人工智能在未来城市建设、人机对话、无人驾驶等方面所展现的引人注目的成就,使得人类智能也不得不依赖于人工智能的作用。这些表明,人类智能的限度已经在人工智能面前显露无疑,人工智能则以许多方面的优势弥补了人类智能的不足。这不是人工智能对人类智能的胜利,而是人工智能帮助人类更好地展现自己的理智能力,并以此推动人类对自我的理解。

其二,哲学的起点是对人类存在本质的思考,人工智能就是以揭示智能奥秘的方式完成着哲学家们对人类存在的追问。自古以来,哲学关注的都是人类本身的问题,围绕着人类的理智和情感发问。但这样的追问与其说是根据事实和证据,不如说是根据猜想和假设。一切关于人类自身的问题,在哲学家看来都是没有最终答案的,因而也就不需要清楚的事实和确凿的证据加以说明。而且,哲学的发问方式总是从已然存在的现象入手,而不是从事物的发生说起,事物的发生和发展反倒是各种哲学理论主张竭力阐明的内容。两千多年来的西方哲学主要是依靠哲学家的想象和推理实现着对人类存在本身的追问。虽然这样的追问带来许多影响深远的哲学理论,但后人对这些理论的解释则更多地是为了说明自己的哲学观念,形成自己的哲学理论,而不是为了让这些追问本身得到更为准确的回答。当然,我们也可以说,得到这样的回答并不是哲学家的工作,相反,哲学正是以无法回答的问题不断追问。但是,如果我们能够让这样的追问变得更具有可操作性,更具有成为真正问题的可能性,那么,这样的追问就会取得更为丰硕的成果,也会推进人类智能的发展,人工智能的重要努力正体现于此。事实证明,人工智能在揭示人类智能奥秘上所完成的工作,已经为我们理解人类智能提供了有力的帮助,例如,类脑智能(Brain-inspired Intelligence)、卷积神经网络(Convolutional Neural Networks, CNN)、深度学习引发的通用人工智能(Artificial General Intelligence, AGI)等。

综上所述,从人类存在的原初问题看,人工智能不仅是人类存在的一种特殊形态,而且构成了对人类智能的一种挑战。如何应对这个挑战,特别是如何回应人工智能对人类存在的追问,这是当下哲学研究必须面对的紧迫问题。

责任编辑:马妮

<sup>①</sup> Penrose, R., *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*, Oxford: Oxford University Press, 1999; Goldberg, E., *The New Executive Brain: Frontal Lobes in a Complex World*, Oxford: Oxford University Press, 2009; Bostrom, N., *Superintelligence: Paths, Dangers, Strategies*, Oxford: Oxford University Press, 2014; Cominelli, L., Mazzei, D., & De Rossi, D. E., "SEAI: Social Emotional Artificial Intelligence Based on Damasio's Theory of Mind," *Frontiers in Robotics and AI*, Vol. 5, No. 6, 2018.

Chinese aesthetics , although it is not a new topic. In 1725 , Giovanni Battista Vico , an Italian philosopher of history , raised “Poetic Geography” in his *New Science*. This concept implies that Geography itself is poetic , literary and aesthetical , so it is not an objective measure in shaping the form of Literature and Arts. In other words , Geography is not a material reality in the space of nature , but an artificial creation. In terms of history , this idea of Geography derives from the worship of circle in both the East and the West , which especially presupposes the cognition of the Center between the Heaven and the Earth , Nine States , Four Directions , Four Oceans , and the Universe for Chinese ancestors. Later on , with this idea the Chinese carry out their lives. From the ancient three generations to Wei and Jin Dynasties , Chinese culture and aesthetics divide into East and West , two parts separated by Tai-hang Mountains , and Yellow River becomes the cultural and geographical axis across these parts. From the Eastern Jin Dynasty to the late Qing Dynasty and the early Republic of China , the East and West division transforms into the South and North , divided by mountains and rivers. This sense leads to such spatial expressions as the Schools , the Paintings , Literatures in the South and the North , and so forth. However , in general , it does not break away from the ancient Chinese aesthetical geographical recognition as Chinese culture , aesthetics , literature and arts have been carried out in four independent styles. Conversely , the divisions correspond to the ancestors’ ideas. With this perspective , this archetype of Geography gives us an indispensable presupposition for discussing the regional issues in Chinese Literature and Arts. The so-called “the beauty of China” or “the beautiful China” in traditional sense of Geography comes from the joint creation between the presupposed national geographical ideas and the writers and artists in later generations.

### **A Study on Chinese Aesthetics and Traditional National Geography**

*Jiang Yi* ( 207)

Artificial Intelligence is one of the most significant achievements of science and technology in the 20th century. Thus , it is necessary for contemporary philosophy to expound the nature of AI from its origin , which in other words means to seek a metaphysical answer to AI and human beings. This study holds that AI is a current form of human being’s life , an attempt of rediscovering human beings and a return to the original problems of human beings. In terms of its origin , the concept of AI is derived from the idea of simulating brain thinking proposed firstly by modern philosophers. In terms of its future development , the problems of AI are closely related to the original problems of human beings , as to answer such questions as “who am I” , which ultimately amounts to key to the problems of AI. The starting point of philosophy is the reflection on the nature of human existence. AI completes philosophers’ questioning of human existence by revealing the mysteries of intelligence. Thus , this paper concludes that how philosophy responds to the inquiry of AI is the urgent problem that current philosophy research faces.