

# AI 技术与人类主体想象

——基于人工直觉在线讨论的研究

雅克布·约翰森 王 鑫

**摘 要:**人工直觉和神经形态硬件是人工智能发展的两大趋势,这种现象在热门网站上被广泛讨论。人工直觉是指能够凭直觉行动的人工智能(AI),而神经形态硬件的灵感则来自人脑的神经元功能。通过对两者的阐释,意图揭示两种技术叙事中所揭示的人类主体观念。人类被视为 AI 的发展蓝图,使技术向更加灵活、自主和高效的方向发展。然而,AI 和人类之间的关系是矛盾的。关于神经形态硬件研究的一些数据表明,AI 可以从人类大脑中学习;而另一些数据则表明,AI 在未来将会超越人类。直觉被认为是使 AI 更像人类的必要步骤,虽然由神经形态硬件支撑的人工直觉可以使技术进步,但也存在一种危险,即直觉被误解和误用,从而遮蔽了 AI 的透明性、可靠性和可解释性。

**关键词:**人工智能(AI);人工直觉;神经形态硬件;机器认知;人类主体;媒体

**中图分类号:**G206;TP18 **文献标识码:**A **文章编号:**2096-5443(2019)06-0035-10

**项目基金:**国家社会科学基金一般项目(16BXW037)

## 一、引言:AI、神经形态硬件、人工直觉

AI 在媒体上的讨论越来越多,很多广为流行的专业网站上的相关探讨涵盖了 AI 发展的最新趋势和观点。AI 被认为是正在进行的一项具有重大意义的技术开发,人类正身处其中<sup>[1-3]</sup>。但是,正如 Goode 指出的那样,这种讨论的特点往往被“神奇的思维”左右,夸大和夸张的宣称其实与事实并非相符。<sup>[2]</sup>正是 AI 的“情感效力”<sup>[2]</sup>激发人们对其的普遍看法,显示出作者对这项技术的惊奇、热情和恐惧。本研究并不把重点放在 AI 的准确性、作者的情感投入或受众对 AI 的潜在反应上,而是遵循另一目标:通过考察当代 AI 的两个特定趋势或方面(而非 AI 整体)时,旨在分析神经形态硬件和人工直觉的热门文章阐述了人类主体的哪些观点。这种关注促使人们能够对人工智能的具体发展和讨论进行深入审视,并从根本上关注 AI 与人类、化身、意识和直觉等之间的关系。这一目标被视为关于 AI 和人类之间不断变化边界的更广泛的讨论,特别是在逻辑和直觉方面,人类大脑的功能也激发了 AI 硬件的开发。正如研究者对人工直觉和神经形态硬件的讨论所认为的那样,AI 试图模仿和采用人脑的特征,并且两者都让 AI 看起来比现在更高效、更自主、更人性化也更灵活。本研究讨论的数据表明,AI 可以从人类的大脑、人类的情感和人类的直觉中学习,建立在这些功能之上的 AI 可以比人类的大脑更好、更先进。虽然这些观点体现了 AI 社区中许多人的意愿,他们希望人工智能可以像人类一样,比人类自身更高效地执行任务,但是这些观点并没有对 AI 的透明性、可解释性和可问责性等问题给予足够的重视。

关于 AI,有许多定义<sup>[4]</sup>。比如,AI 是指“对从环境接收感知并执行动作的智能体的研究”。<sup>[4]</sup>通常是指机器,比如数字计算机。Turner 将 AI 定义为非自然实体通过评估过程做出选择的能力。<sup>[5]</sup>对 Turner 来说,非自然指的是人造的和机器制造(例如,当 AI 系统创建其他 AI 时)。而《大英百科全

书》将 AI 定义为“数字计算机或计算机控制机器人执行与智能人相关任务的能力”。<sup>[6]</sup>在这个意义上,智能生命就是指人类,他们有能力“推理、发现意义、概括或从过去的经验中学习”。<sup>[6]</sup>虽然 AI 技术自 20 世纪 50 年代以来一直在发展中,但近年来,在模拟人类认知特征方面,AI 的复杂性有所提高。<sup>[7]</sup>总的来说,AI 在过去 10—15 年里成长迅速,是得益于物联网<sup>[8]</sup>的发展,以及数据处理和基于传感器的技术在不同领域的广泛进步。AI 的普遍存在,促使人文和社会科学领域对此的学术讨论不断增加,也得到了新闻传播学界、业界对 AI 的广泛关注。鉴于人工智能技术试图模拟或模仿人类行为(如解决问题、认知能力或模式识别),并提升人类的这些特征,特别是 AI 被视为具有复制和提高人类大脑的核心能力,它理所应当会吸引学者和记者的大量关注。不过,“即使有一台几乎拥有无限能力的电脑,人类仍然不知道它如何达到大脑的智力水平。”<sup>[4]</sup>然而,在开发 AI 技术时,一方面,AI 的某些部分试图寻求建立在人类大脑的结构和能力之上,例如,人工神经网络建立了大脑活动的数学模型;与此相反,思考大脑有助于扩展技术概念的范围。<sup>[4]</sup>本研究基于 AI 当前发展的趋势和状态,重点关注两个以“人”作为发展蓝图的方面,即:神经形态硬件和人工直觉。两者都被视为当前讨论 AI 趋势的范例,并且这些趋势在一些关注技术发展的网站上被广泛热议。

人工神经网络是以人脑为模型。这些软件,通常是由能够“通过从数据中学习解决方案来解决问题的自我设计”的算法组成<sup>[9]</sup>。这些系统“学习”执行任务,不需要监督,也没有特定的规则。比如,在图像识别中,通过人工分类的图像训练,可以对图像进行分类。与此相关的是神经形态硬件的最新发展,也是模拟大脑结构和生物神经网络,这被认为是 AI 的重大发展。<sup>[10-11]</sup>

传统的中央处理器(CPU)处理指令基于“时钟时间”——信息以固定的间隔传输,就像由节拍器管理一样。通过将神经元的数字等高线封装起来,神经形态学就可以利用“尖峰”(需要时就可以发送的突发电流)进行并行通信(而且不像计时的时间那么死板)。就像我们的大脑一样,芯片的神经元通过处理传入的电流来进行交流——每个神经元都能从传入的脉冲决定是否向下一个神经元发送电流。让这成为一件大事的是,这些芯片处理 AI 算法所需的能量要小得多。<sup>[12]</sup>

与神经形态硬件和神经网络相比,人工直觉还处在起步阶段<sup>[13-15]</sup>,尚未实现。一些观点认为它将“取代”AI<sup>[16]</sup>。虽然有对人工直觉的炒作嫌疑,但这使分析变得更加有趣,因为它既试图被创造,也被设想如何成为现实。人工直觉是指 AI 系统做出直观选择或对问题做出直观反应的能力<sup>[14]</sup>。AI 在计算机科学中的探讨和研究,也引起人文学者的兴趣。这可以被看作更广泛的跨人类/后人类<sup>[17-19]</sup>和“非人类转向”<sup>[20]</sup>研究的一部分,从不同视角研究了主体和客体、人类和机器之间的关系<sup>[7-8,21-24]</sup>。N. Katherine Hayles 在最近的一本书中提出了这样的观点:人类和计算技术,如 AI 驱动系统,都具有认知能力。Hayles 特别关注认知与她所声称的无意识认知之间的关系。无意识认知是指在神经元层面上的过程,这些过程对意识是不可接近的,但对意识却是至关重要的。机器本身并没有意识,尽管如此,它还是能够进行无意识认知。无意识认知和直觉两者之间有一些相似之处<sup>[7]</sup>。根据这一观点,AI 的未来可能是这样一种情况,其特征在于机器能够比人类更好地思考,这是由于无意识认知的直觉方面(另见<sup>[25]</sup>)——正如看到的那样,AI 行业和大众评论同样对这一前景有了新的预期。

事实上,只有少数学者关注和研究 AI 是如何在流行网站被构建,尤其是通过这种叙述所展示的人类形象<sup>[1-3]</sup>。来自人文学者们对计算技术和 AI 的批评更为普遍,但是他们关注的只是理性,忽视了人的本质和模仿人类特征的其他复杂性<sup>[26-31]</sup>。目前还没有研究关注神经形态硬件和人工直觉在更广泛的科技界是如何被讨论的。

本研究选择了神经形态硬件和人工直觉这两种现象,前者是硬件的开发,后者则是基于代码的开发,作为 AI 技术的两个重要方面,二者也是相互关联的。笔者重点考察这个问题在面向非专业受众的公共网站(如 Wired 或 Ars Technica)上是如何被讨论的。贯穿和引导数据收集和分析的主要问

题是:在关于神经形态硬件和人工直觉的热门文章中,对人类主体的看法是什么?为了对数据进行定性、详细的内容分析,本研究采取小样本的分析方法,即选择谷歌搜索前两页的文章。这些文章是2019年5月通过谷歌搜索关键词“神经形态硬件”和“人工直觉”获得的。虽然大样本可能是有益的,但本研究以实证分析为主,辅以相关理论阐释,着重于详细的讨论,因此数据量有限。本研究依据内容并按主题划分共讨论了七个网站,这些主题涵盖在接下来的分析中。

## 二、AI 对人脑的模拟与超越:关于大脑内外神经形态硬件的讨论

毫不奇怪,本研究参考的网站几乎都提供了神经形态硬件现象背后的功能和技术解释。比如前面提及的在 Wired<sup>[12]</sup> 和 Ars Technica 的引文中,都是参照人类大脑来解释这一现象:

该芯片包含 100 万个“神经元”,每个神经元由一组数字晶体管表示,以及 2.56 亿个“突触”,也就是它们之间的有线连接。两个特性使芯片比传统芯片更高效。首先,像大脑一样,它通过“尖峰”(spike)进行沟通,也就是从一个神经元向下游神经元发送一个大小的信息包。因为它可以这样通信,所以信号很简单(尖峰或没有尖峰),而且只有在神经元传输数据包时才偶尔传输。第二,和大脑一样,处理和记忆是同时定位的——在神经元和它们的突触中。在传统的计算机中,处理单元不断地从不同的内存区域获取信息,对其执行操作,然后将新信息返回到内存中,这导致了大量缓慢而浪费的通信。<sup>[32]</sup>

上述引文的惊人之处在于,尽管人类大脑和神经形态芯片之间存在着一种隐含的区别,而这种区别部分是由大脑的某些神经功能所激发的,但通过将两者等同起来,这种区别就被打破了,“神经元的等价物”使用“尖峰”。突触、神经元和尖峰都是人类大脑功能的基本方面,尽管它们在描述神经形态硬件时都被用引号括起来,但这些术语仍然被使用。“神经元的数字等价物”“就像我们自己的大脑一样”和“就像大脑一样”的表达方式表明,这种技术已经直接模仿了人类大脑。它不仅受到神经元如何交流的启发,还有效地重现了芯片中大脑功能某些方面的技术性。这是该技术的一个重要框架,因为它等同于人类大脑和用于 AI 的硬件。在现实中,一些外部的东西(人类大脑)已经通过技术被模拟出来,并且广泛地用于展示神经形态硬件是如何像人类大脑一样<sup>[1]</sup>。这一框架暗示人类大脑的复杂性已经在硬件上被成功地再现和模拟了。一个硬件(神经元、突触、尖峰)被映射到另一个硬件(芯片、电子、瓦特)上。这些公式已经揭示了科技行业内外许多人的固有信念,即计算机技术可以成功地采用和调整大脑的高效神经元动力学,以实现其自身的目的。这是否属实,仍有待商榷。从上面的引文中,人们会认为大脑实际上比 AI 在信息传输的效率和经济使用方面更好。神经形态芯片比传统芯片更高效、更节能,因为这种芯片消除了一些活跃的或者不活跃的“无用”通信。人类大脑和 AI 硬件的这个等式在下面的引文中得到了进一步的证明:

神经形态计算的灵感来自人类大脑的功能,目前人类大脑在所有需要创造力或将知识转移到其他问题上的表现仍然优于机器。此外,人脑的能量效率令人难以置信,仅使用约 20 瓦<sup>[2]</sup>。大脑中主要的计算单元是神经元,成年人脑中大约有 1000 亿个神经元。这些神经元由超过 15 万公里神经纤维和 150 万亿突触<sup>[4]</sup>连接,使大脑结构大规模并行。作为对比:截至 2017 年 11 月,世界上最快的超级计算机拥有 10 649 600 个处理器内核,使用 15 731kW 的能量<sup>[4]</sup>。单个 CPU 核心可以执行的计算量不能与单个神经元执行的计算量直接比较。然而,特别是在人工智能的应用中,人类的大脑可以作为一个非常强大和高效的计算系统的蓝图。<sup>[33]</sup>

Sherry Turkle 早于 2005 年就指出,自从个人电脑在人类日常生活中获得如此基本的功能之后,个体就开始用电脑和数字技术的功能来描述他们自己和他们的思想<sup>[34]</sup>。类似的情况发生在上述引自 Wired, Ars Technica 和 MI Garage 的文章描述中。大脑(据说)像电脑一样有效地运作,所以这些功能可以转移到硬件上。它的“主要计算单元”是神经元。然而,在这一点上,大脑仍然是一台比任何以它为模型的硬件更好的计算机,或者在技术上更先进。当涉及像解决问题这样的复杂任务时,它

的表现“优于”神经形态计算;它也是节能的,它使用瓦特就像计算硬件一样。从上面的引文中可以清楚地看出,大脑可作为能量和信息传输效率的蓝图。考虑到 AI 的最终目标是在本体论和认识论,或者说硬件和软件方面复制人类智能,这或许并不令人意外。但这样的方程式是否能同时反映人脑的复杂性,以及更广泛意义上作为研究对象——人的复杂性,仍然是个问题。稍后我们将更详细地讨论这一点。前面引用 Wired 同一篇文章的另一段内容讨论了在计算机上模拟的复杂大脑模型 Spaun:

Spaun 证明,计算机可以与环境进行流畅的互动,并执行类似人类的认知任务,比如识别图像和控制机器人手臂记录它所看到的東西。这台机器并不完美,但它是一个惊人的证明:有一天,电脑可能会模糊人类与机器认知之间的界线。<sup>[12]</sup>

另一个评论者写道:

随着 AI 的成熟,在日益沉浸和复杂的场景中开发和训练这些系统的需求成为进展的主要关注点。我们知道, AI 领域的最终目标是生产出一种通用人工智能 (AGI),它在人类努力的所有领域都能超越人类,我们需要建设一个足够富裕、足够大、有足够多复杂实体的世界,以实现这一发展的最大化。<sup>[35]</sup>

这些引述表明, AI 的未来指日可待。在未来,计算机将对人的大脑进行充分的模拟。计算机能“执行类似人类的认知任务”,哪怕是相当简单的任务。然而,目标是明确的: AI 必将超越人类的认知能力<sup>[7]</sup>,这样的叙述以一种必然面向进步的方式构建了 AI。AGI 只是一个用足够的的数据创建足够复杂的环境来训练系统的问题,更好和更复杂技术的因果关系已然建立。(实际上, Artificial Mind 的研发者并不认为这是一个需要足够多数据和复杂的环境训练就能够解决的问题,而是使用不同于现有 AI 的新的计算机语言才有可能实现)人类的大脑,无论多么复杂和聪明,最终都会输给 AI。然而,我们还没有做到这一点:神经形态硬件(以及软件)目前只能完成特定的任务。例如,曼彻斯特大学的 SpiNNaker (Spiking Neural Network Architecture) 项目被命名为能够找到“数独 (sudoku) 的解决方案……或者用四种颜色给地图上色,而相邻国家的颜色不相同……”<sup>[36]</sup>。或如下述:

传统上,编写计算机程序需要编写循序渐进的指令。例如,教计算机识别狗,可能需要列出一组规则来指导它的判断:看看是不是动物;检查它是否有四条腿;检查它是否比猫大、比马小;检查它是否吠叫等。但良好的判断需要灵活性,如果电脑遇到一只不会叫、只有三条腿的小狗怎么办?那么,也许需要更多的规则,但是为计算机做出的每种类型的决策必须列出无穷无尽的规则并重复该过程是低效且不切实际的。<sup>[32]</sup>

AI 需要更多的灵活性,更多的跳出框框的思考,而不是将特定的问题与规定的解决方案或属性进行比较。有一些人提倡,值得探索的解决方案可能是人工直觉。

### 三、人工直觉和 AI 的人性化:直觉理论与人工直觉的技术逻辑想象

#### (一)关于直觉的理论叙述

关于“直觉”这个术语有很多角度的讨论。本研究主要借鉴文化研究和情感理论中对直觉的定义(关于这个术语的广泛讨论,请参阅 Chudnoff, 2013<sup>[37]</sup>)。从历史上看,直觉作为一种较低或相对没有价值的认知形式,常常被置于理性的对立面。性别歧视的话语将直觉与女性联系在一起,女性的逻辑思维能力会更弱,因为她们被认为比男性更情绪化<sup>[38]</sup>。直觉依赖于个人的主体性,因此每个人的直觉都是不同的。直觉通常被理解为一种经验和感觉的认知形式,而不是完全有意识的认知。它通常与“直觉”“预感”或“有某种感觉”等术语联系在一起。直觉指的是一种身体状态或经验,在这种状态下,个体的行为不必有(经验的)证据或证据;它是一种转瞬即逝的东西,被跟随或倾听,然后可能在并不能充分理解原因的情况下付诸行动;直觉还指一种即时的经验模式,在个体完全没有意识到和认知到之前,就已经发生。这与某些情感理论是一致的,这些理论强调的是具身体验,这些

体验主要是感觉到的,而不是人类身体完全了解的。情感指的是身体在有意识地了解或反思这些经历之前,在感官上受到影响的能力<sup>[22,39-40]</sup>。然而,直觉并不反对理性,也不与理性相矛盾,它指的是本能的感觉。Greg Seigworth 曾写道,“通过直觉能力的经验和实验……超越或溢出智力”<sup>[41]</sup>。对于 Pedwell 来说,直觉是在将其转化为表征性和分析性思维范畴之前的具体经验<sup>[42]</sup>。当然,关于直觉的理论叙述无论是在哲学家或者文学理论家那里还有很多,根据本研究的需要,在此不再赘述。目前,AI 仍然是基于逻辑通过算法来实现认知和理解的,但人类不仅具有逻辑认知,还具有超越逻辑和语言限制的直觉。

## (二)关于人工直觉发展的技术想象

直觉最近引起了 AI 领域计算机科学家和评论者的极大兴趣,在某种程度上,这也与神经形态硬件有关。直觉和它被普遍理解的方式似乎是 AI 一个重要的逻辑组成部分:

我们希望机器能够凭直觉思考的一个原因是为了我们的安全。应用人工直觉的一个很好的例子就是自动驾驶汽车或自动驾驶设备。搭载 AI 软件的机载传感器一直是这些系统的工作方式,但在某些情况下,它们仍然容易发生事故。通过人工直觉,自动驾驶汽车可以预测道路上可能发生的不可预测的事情。<sup>[43]</sup>

考虑到最近发生的自动驾驶汽车事故,人们可能会同意上述说法,以及 AI 驱动系统需要更直观地发挥作用,而不是像目前这样死板、容易出错。在上面的引文中,直觉被设定为具有灵活性和动态性的特征,这是至关重要的。但机器凭直觉思考到底意味着什么呢?基于这些评论,直觉是如何被理解和构建的?

在游戏中,我们已经看到许多早期成功的深层强化学习算法的初步部署,比如 DeepMind 的 AlphaGo Zero 已经掌握了超越任何人类玩家能力的围棋和国际象棋棋艺,显示出人类从未考虑过的战略洞见。例如 AlphaGo Zero 并没有遵循通常控制人类游戏的 Reinfield 值(国际象棋棋子的具体值为:棋子=1,骑士=3,皇后=9等),而是优化棋盘位置,几乎完全忽略棋子值。有趣的是,这源于它自学而不是观察人类游戏的能力,从而避免了人类偏见的污染。然而,在这个过程中,AlphaGo Zero 产生的性能比任何其他国际象棋引擎都更人性化,它更依赖于启发式或直觉,以达成一个聪明的战略,而较少考虑棋局的位置。<sup>[35]</sup>

虽然对 AI 直觉系统的需求是可以理解的,即使不存在术语上的矛盾,但问题仍然存在。虽然人类的直觉是基于经验和积累足够多的“数据”,在给定的情况下凭直觉行事;但当一个人凭直觉行事时,数据或经验是如何以及为什么被利用,远比之前描述的要模糊和复杂得多。AlphaGo Zero(主要基于神经网络)或许能够更直观地行动,并自学新的策略,但这与人类的直觉并不相同。事实上,正如 Paolo Bory 所讨论的那样,最初的 AlphaGo(AlphaGo Zero 的前身)击败世界冠军的举动被认为是创造性的,甚至是漂亮的。实际上,它与人类并不像。“(比赛)第二天,DeepMind 透露,AlphaGo 之所以决定下这一步棋,是因为人类棋手下这一步棋的可能性是 1:10 000。”<sup>[3]</sup>。对于 AI“决定”执行某一特定动作的解释,根植于逻辑而非直觉。对概率的计算导致了一个特定的举动,导致人类玩家的失败。这是一种创造性的形式,但不是直觉,它看起来是实用的,更重要的是,对其他人来说是可以解释的。AlphaGo 的这种非人类特征在它的下一个版本中得到了进一步的发展:

经过仅 3 天的自我训练,AlphaGo Zero 就强势打败了此前战胜李世石的旧版 AlphaGo,战绩是 100:0。40 天后,它以 90% 的胜率击败了最先进的 AlphaGo 原始软件。通过不使用人类数据——不以任何方式使用人类专业知识——我们实际上消除了人类知识的限制。因此,它能够根据基本原理创造知识;从零开始。这使得它比以前的版本功能强大得多。<sup>[43]</sup>

无论 AlphaGo Zero 的动作多么接近直觉,它用来自我学习的数据仍然可以获得。它“有意识”地进行自我数据学习。也许“灵活性”比“直觉”在这里更合适。虽然目前的 AI 可能(仍旧)无法实现接近人类直觉复杂性的直觉能力,但上述讨论提出了有关意识和对 AI 的无意识认知的相关问题。

在此引入一些其他数据来更详细地讨论这个问题:

直觉更多地与直感有关,而不是经过计算的决策过程。直觉和智力不一样的。它们实际上是两个不同的认知过程。智力是建立在已知的基础上的,而直觉是处理未知的。直觉更多的是基于感觉,而智力则是依靠逻辑。人类可以根据他们的感觉做出决定,而不一定是合乎逻辑的。计算机不像人类那样有情感,对一台机器而言,由于只是二进制的,在做决定时使用“直觉”是相当了不起的。<sup>[43]</sup>

但是,直觉确实是一种认知的形式,也是一种智力的形式,尽管对文化理论家来说,直觉不是逻辑,也是完全不需要计算的。计算机是否能够基于直觉做出决策,这个问题很重要,因为它可能会为自动化系统提供更多的动力。例如,一个预测性的警务软件标记了一个通过闭路电视摄像机接受调查的人,不是因为该人的脸基于现有数据库更容易被识别;而是因为该软件的直觉判断,不是基于现有的知识/数据,而是基于一些未知的内容。这显然引发了重大的伦理问题,因为无法提出正式证据来解释为什么这个人会被软件标记出来,然后被警方逮捕。直觉几乎不能成为一个有效的理由。另一方面,如果一辆自动驾驶汽车突然停了下来,因为它能感觉到一个行人可能会跑到汽车前面,直觉会非常有效。下面的讨论呈现了类似的困境:

首先,直觉仅仅是不依靠完整的知识而做出的正确决定的标签。如果人们是直觉的,我们是可以接受的,事实上,我们也提升其价值并钦佩它们。但是我们对机器做出涉及直觉、模糊或不完整数据的决定的前景感到不安。如果直觉只是一堆我们没有意识到的变量呢?<sup>[44]</sup>

答案可能是让 AI 更加富有情感和人性化。

科学家们把传感器放在人们的手指上,记录他们在驾驶模拟器中的脉搏,以此作为衡量清醒程度的标准。算法利用这些记录——四个人共用 80 分钟——来学习预测一个人在驾驶途中每一刻的脉搏。然后,它利用这些“可怕”的信号作为向导,学习在虚拟世界中的驾驶。<sup>[45]</sup>

例如,科学家教 AI 自动驾驶汽车如何在撞车发生前“害怕”这一过程,可能会帮助科学家得到一个解决方案。然而,与我们之前讨论的神经形态硬件相似,一个基本的人类特征(恐惧和或者因为害怕产生的本能行为)被映射到技术上。通过这样的尝试,AI 被拟人了,无论是在实践中还是在话语建构中。人工直觉被认为是可以使 AI 更有效的东西,最终需要更少人的监督或控制。

实际上,AI 非常适合做出这些高度直观的决定。因为它不是有意识的,所以它对观察到的东西没有偏见,所以它能意识到影响特殊决定的每件事。这有点像机器版的夏洛克·福尔摩斯(Sherlock Holmes),他会注意到每个微小的细节,然后通过判断这些细节中哪些是重要的来得出观察结果。<sup>[44]</sup>

上面这段引述最能说明一些人是如何看待 AI 的。事实上,AI 远非没有偏见。AI 驱动的系统最初是由人类构建的,带着固有的偏见,而这些偏见往往(非)有意识地嵌入到技术中<sup>[46-48]</sup>。这个问题引发关于 AI 可解释性更广泛的争论。AI 系统记录并存储它们自己的决策和过程,以便人类能够仔细检查和排除故障。从这个意义上说,AI 总是有意识的。然而,通常难以向外行人解释 AI 是如何做出特定的决定,因为它们太复杂了。直觉会更加强化它,因为某些部分无法通过话语解释,需要 AI (或人类)解释清楚,才能让它负责。因此,提倡 AI 的直觉是危险的,因为从这个意义上讲,直觉即使可解释,也是很难解释的。如果 AI 凭直觉行事,它就有理由不透明,也无法解释特定的直觉行为。没有责任感的直觉是危险的。基于本文开头讨论的直觉的定义,AI 社区错误地将直觉视为一种可以添加到 AI 以使其更加灵活、动态和自治的技术性。对于文化理论家来说,直觉是一种与世界直接的情感和感官上的接触<sup>[40-42]</sup>。对 AI 的评论者来说,这是一个应该构建的技术性问题。AI 的认知和“意识”(这是有争议的)是基于代码的。AI 无法像人类那样凭直觉感知,人工直觉也是被编程到 AI 中。它通过话语(代码)来定义 AI 可以直觉做什么以及如何做,这显示了真正直觉的局限性,因为对

于文化研究学者来说,这个术语指的是与话语相冲突或超越话语的感性、情感过程。人类直觉的获得是不能通过教导来实现的,为了完善直觉而遵循一套预先定义的规则也是行不通的,对于 AI 亦是如此。同时,直觉作为一种思维方式,并不是通过逻辑推理来实现的,但这并不意味着直觉只是一种依赖于主观判断、缺乏客观依据的感性直觉。直觉作为一种思维方式,是人类与生俱来的一种能力。直觉思维具有超越感觉和理性的特征。从这个意义上说,AI 的类人直觉是否能够实现,至少目前还不确定。

超越人工直觉的技术方面,在热门网站和科学家中如何使用(误用)该术语表明了对人类主观性的特定理解。AI 不仅通过倡导直觉而被拟人化并变得更人性化,同时,人类比我们在文中研究的数据更具有机械性、算法性和技术性。这是通过“人类知识的约束”(Hadean)这样的表述来实现的,“直觉仅仅是不依靠完整的知识而做出的正确决定的标识”(INC),我们的直觉是“没有意识到”的“一堆变量”(INC)。与我们对神经形态硬件的讨论类似,人类主体被看作是一维的和功能主义的。芯片和大脑功能相同,几乎可以互换。AI 的创始人之一约翰·麦卡锡(John McCarthy)在接受采访时也表达了类似的观点:

米什洛夫:现在我们来谈谈这个问题——我知道,很多人对你很有敌意,你怎么可以把一个人描述成一个系统?人有更多的东西——我们有直觉,我们有灵性,我们有超越自身机械化存在的方面。

麦卡锡:这个观点几百年来一直在退步,随着对人类生理和心理的了解越来越多。我想,也许可以用“拳击”这个比喻:它能跑,但不能藏。

米什洛夫:你能详细说明一下吗?我不太清楚你在说什么。

麦卡锡:嗯,人类意识的这些方面在机器和计算机程序中还没有被实现,它们的实现还存在一些困难,但我们对 AI 持乐观态度,希望能够解决这些问题。<sup>[49]</sup>

这样的等式并没有公正地对待人类主体性的复杂。人类的思维比计算机更加矛盾和混乱<sup>[34]</sup>。尤其是人的主体性不仅建立在语言和话语的基础上,还有其他与情感和感官相关的人类经验的记录。直觉只是其中之一。

#### 四、结论

本研究旨在分析人工智能的两个趋势——神经形态硬件和人工直觉是如何在热门网站上被讨论的。我们旨在探寻这些叙述如何揭示了人类和 AI 之间的关系,以及对人类主体的理解如何在数据中被暗示或明示。虽然无法充分评估这两种现象的适应性,但笔者认为,它们表明了关于 AI 技术应该如何进步的特殊愿望。从某种意义上说,AI 试图同时在本体论和认识论上采用人类的特征。人类大脑的生物化学功能以及它们如何映射到神经形态芯片上,人类认识的直觉以及它如何映射到 AI 特征上,这两者都让 AI 看起来比现在更高效、更自主、更人性化也更灵活。然而,人与 AI 是矛盾的。讨论的数据表明,AI 既可以从人类大脑、人类情感和人类直觉中学习,将它们整合,使其可以在这些方面表现得更好、更聪明。这样的叙述显示了 AI 社区中许多人的愿望,即让 AI 尽可能的像人类,以便在总体上实现对人类的超越。他们承认这尚未发生,但相信这种状态一定能实现。这种关于人工直觉的炒作可能因为它不仅使 AI 能看起来更人性化,而且更有同情心,更讨人喜欢<sup>[3,23]</sup>,为了掩盖 AI 的透明性、可解释性和可问责性,或许可以故意这样做。正如已经讨论过的,这种情况引发了 AI 的自主性、问责性和透明度的关键伦理问题。与实现通用人工智能(AGI)的目标相关的是 AM 技术<sup>[50]</sup>,在中国正得到特别的发展,其特征是完全的人工意识、感知和情感<sup>[51]</sup>。《从 AI 到 AM:人工智能的知识观》一文认为,“基于 AM 的类人思维将是解决人机交互中真实语义理解和语用一致性问题的关键。”AM(Artificial Mind)技术或许将超越模拟神经网络的深度学习过程,成为实现人机交互的关键。<sup>[50]</sup>

虽然只讨论了极少数的数据,但它们仍然指出了 AI 流行叙述中的一个偏见。AI 仍然过于男性化,而关于其未来的论述严重偏向于某种特定的声音——主要是男性科幻作家和以技术为中心的科学家、未来学家和企业家——AI 技术很容易被描述成一种不可避免的崇高景象<sup>[2]</sup>。我们还要补充一点,即对所研究的数据中存在对人类主体特殊的理解或世界观,认为人类主体是理性的,是随时被人工智能技术利用的技术蓝图,如果仅用生化术语解释,像直觉这种模糊、复杂和难以定义的东西,为了以特定的方式建构和发展人工智能,被转化为一个机械性的概念。另一种理解认为直觉并不是对 AI 当前的僵化性进行有效论证的最佳术语。AI 需要变得更灵活——而不是直觉——这应该以一种合乎道德和透明的方式发生。

#### 参考文献:

- [1] S. Natale and A. Ballatore. *Imagining the Thinking Machine: Technological Myths and the Rise of Artificial Intelligence*. *Convergence*, 2017, <https://doi.org/10.1177%2F1354856517715164>.
- [2] L. Goode. *Life, But Not as We Know It; AI and the Popular Imagination*. *Culture Unbound: Journal of Current Cultural Research*, 2018, 10(2): 185-207; 186; 187; 204.
- [3] P. Bory. *Deep New: The Shifting Narratives of Artificial Intelligence from Deep Blue to AlphaGo*. *Convergence*, 2019, <https://doi.org/10.1177%2F1354856519829679>.
- [4] S. Russell, P. Norvig. *Artificial Intelligence: A Modern Approach*. Third Edition. Saddle River, NJ: Pearson; viii, 2010: 12; 727.
- [5] J. Turner. *Robot Rules: Regulating Artificial Intelligence*. Basingstoke: Palgrave Macmillan, 2019.
- [6] B. J. Copeland. *Artificial Intelligence*. *Encyclopedia Britannica Entry*, 2019-05-09. [2019-05-16] <https://www.britannica.com/technology/artificial-intelligence>.
- [7] N. K. Hayles. *Unthought: The Power of the Cognitive Nonconscious*. Chicago: University of Chicago Press, 2017.
- [8] M. Bunz, G. Meikle. *The Internet of Things*. 2017. Cambridge: Polity.
- [9] R. Muthic. *Adaptive Neural Algorithms for PCA and ICA*. J. R. R. Dopico, J. Dorado de la Calle, A. P. Sierra (Eds). *Encyclopedia of Artificial Intelligence*. New York: Information Science Reference, 22-30; 2.
- [10] J. Hasler, H. B. Marr. *Finding A Roadmap to Achieve Large Neuromorphic Hardware Systems*. *Frontiers in Neuroscience*, 2013, 7, 1-29.
- [11] T. Bohnstingl, F. Scherr, C. Pehle, K. Meier, W. Maass. *Neuromorphic hardware learns to learn*. arXiv preprint, 2019. arXiv: 1903. 06493. <https://arxiv.org/abs/1903.06493>.
- [12] A. Frank. *The Future of AI is Neuromorphic*. *Meet the Scientists Building Digital 'Brains' for Your Phone*. *Wired*, 2017-03-08. [2019-04-10] <https://www.wired.co.uk/article/ai-neuromorphic-chips-brains>, online.
- [13] W. Tao, P. He. *Intuitive Learning and Artificial Intuition Networks*. 2009 Second International Conference on Education Technology and Training. IEEE. <https://ieeexplore.ieee.org/abstract/document/5381516>.
- [14] J. Crowder, S. Friess. *Artificial Psychology: The Psychology of AI*. *Proceedings of the 3rd Annual International Multi-Conference on Informatics and Cybernetics*. Orlando, FL, 2012.
- [15] A. Srdanov, N. R. Kovacevic, S. Vasic, D. Milovanovic, D. Emulation of Artificial Intuition Using Random Choice and Logic. 2016 13th Symposium on Neural Networks and Applications (NEUREL). IEEE, 1-6.
- [16] P. Nelson. *Artificial Intuition Will Supersede Artificial Intelligence, Experts Say*. *Network World*, 2017-02-10. [2019-04-10] <https://www.networkworld.com/article/3168320/artificial-intuition-will-supersede-artificial-intelligence-experts-say.html>.
- [17] N. K. Hayles. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago, IL: University of Chicago Press, 1999.
- [18] N. Badmington. *Posthumanism*. Ed, Cambridge: Macmillan Press, 2000.
- [19] P. K. Nayar. *Posthumanism*. London: Polity, 2014.
- [20] R. Grusin. *The Nonhuman Turn*. Minneapolis: University of Minnesota Press, 2015.

- [ 21 ] W. H. K. Chun. *Updating to Remain the Same: Habitual New Media*. Minneapolis: University of Minnesota Press, 2017.
- [ 22 ] P. T. Clough. *The User Unconscious: On Affect, Media, and Measure*. Minneapolis: University of Minnesota Press, 2018.
- [ 23 ] A. McStay. *Emotional AI: The Rise of Empathic Media*. London: Sage, 2018.
- [ 24 ] L. Blackman. *Haunted Data: Affect, Transmedia, Weird Science*. London: Bloomsbury, 2019.
- [ 25 ] L. Parisi. *Critical computation: Digital Automata and General Artificial Thinking*. *Theory, Culture & Society*, 2019, 36 (2): 3-16.
- [ 26 ] L. Méro. *Ways of Thinking: The Limits of Rational Thought and Artificial Intelligence*. London: World Scientific, 1990.
- [ 27 ] J. Rodado, M. Rendon. Can Artificial Intelligence be of Help to Psychoanalysis... or... Vice Versa? *The American Journal of Psychoanalysis*, 1996, 56(4), 395-413.
- [ 28 ] J. Johnston. *The Allure of Machinic Life: Cybernetics, Artificial Life, and the New AI*. Massachusetts: MIT Press, 2008.
- [ 29 ] E. A. Wilson. Affect, Artificial Intelligence, and Internal Space. *Emotion, Space and Society*, 2008, 1(1): 22-27.
- [ 30 ] D. Columbia. *The Cultural Logic of Computation*. Cambridge, MA: Harvard University Press, 2009.
- [ 31 ] T. Essig. Psychoanalysis Lost-and Found-in Our Culture of Simulation and Enhancement. *Psychoanalytic Inquiry*, 2012, 32(5): 438-453.
- [ 32 ] M. Hutson. Pocket Brains: Neuromorphic Hardware Arrives for our Brain-inspired Algorithms. *Ars Technica*, 2017-07-16. [ 2019-04-04 ] <https://arstechnica.com/science/2017/07/pocket-brains-neuromorphic-hardware-arrives-for-our-brain-inspired-algorithms>, online.
- [ 33 ] D. Justus. Neuromorphic Computing: The Tech Behind the Hype. *MI Garage*, 2018-07-09. [ 2019-04-10 ] <https://www.migarage.ai/intelligence/neuromorphic-computing-tech-behind-hype>, online.
- [ 34 ] S. Turkle. *The Second Self: Computers and the Human Spirit*. Twentieth Anniversary Edition. Massachusetts: MIT Press, 2005: 22.
- [ 35 ] R. Mansoor. Building Worlds to Grow the Artificial Mind: The AI Petri-dish. *Hadean*, 2018-01-29. [ 2019-04-10 ] <https://www.hadean.com/blog/building-worlds-to-grow-the-artificial-mind>, online.
- [ 36 ] Justus, D. Neuromorphic Computing: The Tech Behind the Hype. *MI Garage*, 2017-07-16. [ 2019-04-04 ] <https://www.migarage.ai/intelligence/neuromorphic-computing-tech-behind-hype>, online.
- [ 37 ] E. Chudnoff. *Intuition*. Oxford: Oxford University Press, 2013.
- [ 38 ] L. H. Schwartzman. Intuition, Thought Experiments, and Philosophical Method: Feminism and Experimental Philosophy. *Journal of Social Philosophy*, 2012, 43(3): 307-316.
- [ 39 ] B. Massumi. *Parables for the Virtual: Movement, Affect, Sensation*. Durham: Duke University Press, 2002.
- [ 40 ] L. Berlant. Intuitionists: History and the Affective Event. *American Literary History*, 2008, 20(4): 845-860.
- [ 41 ] G. J. Seigworth. Cultural Studies and Gilles Deleuze. G. Hall, C. Birchall (Eds). *New Cultural Studies: Adventures in Theory*. Edinburgh, Edinburgh University Press, 2006: 107-126; 118.
- [ 42 ] C. Pedwell. Digital Tendencies: Intuition, Algorithmic Thought and New Social Movements. *Culture, Theory and Critique*, 2019: 1-16; 127.
- [ 43 ] V. Tabora. Artificial Intuition and Reinforcement Learning, the Next Steps in Machine Learning. *Becoming Human*, 2018-11-06. [ 2019-04-10 ] <https://becominghuman.ai/artificial-intuition-and-reinforcement-learning-the-next-steps-in-machine-learning-6f2abeb9926b>.
- [ 44 ] T. Koulopoulos. Did A. I. Just Make the Leap to Being Intuitive? *Inc. com*, 2017-10-31. [ 2019-04-10 ] <https://www.inc.com/thomas-koulopoulos/did-ai-just-make-leap-to-being-intuitive.html>, online.
- [ 45 ] M. Hutson. Scientists Teach Computers Fear-to Make Them better Drivers. *Science Mag*, 2019-05-10. [ 2019-05-16 ] <https://www.sciencemag.org/news/2019/05/scientists-teach-computers-fear-make-them-better-drivers>, online.
- [ 46 ] J. Cheney-Lippold. *We are Data: Algorithms and the Making of our Digital Selves*. New York: NYU Press, 2017.
- [ 47 ] S. U. Noble. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press, 2018.
- [ 48 ] S. M. West, M. Whittaker, K. Crawford. Discriminating Systems: Gender, Race and Power in AI. *AI Now Institute*, 2019-04-10. [ 2019-04-10 ] <https://ainowinstitute.org/discriminatingsystems.html>.

- [49] L. Mishlove, J. McCarthy. Artificial Intelligence with John McCarthy, PhD. Interview, 2011. [2019-04-10] <https://www.youtube.com/watch?v=Ozipf13jRr4>. Transcript available at: <http://www.intuition.org/txt/mccarthy.htm>.
- [50] 王鑫, 沙永锋. 从 AI 到 AM: 人工智能的知识观. 新闻与传播评论, 2018, 6: 41-50; 49.
- [51] Y. Deng, H. Jiang. The Development Overview of Artificial mind. In Proceedings of the 2nd International Conference on E-Education, E-Business and E-Technology, 2018. ACM: 111-116.

## AI Technology and Human Subject Imagination: Online Discussions Based on Artificial Intuition

*Jacob Johanssen* (St Mary's University)

*Wang Xin* (Liaoning University)

**Abstract:** Artificial intuition and neuromorphic hardware are two recent trends in artificial intelligence (AI). Those phenomena are discussed on popular websites. Artificial intuition refers to AI that is able to act intuitively, whereas neuromorphic hardware takes its inspiration from neuronal functions of the human brain. Following an explanation of the two, we pay attention to what kind of idea of the human subject is revealed in the narratives. The central arguments of this article are that the human is understood as a blueprint for AI that can make the technology more flexible, autonomous, and efficient. Yet, the relationship between AI and humans is ambivalent. Some data about neuromorphic hardware suggest that AI can learn from the human brain, whereas other argue that AI will outperform humans in the future. Intuition is framed as a necessary step to make AI more human-like. While an intuitive AI, backed by neuromorphic hardware, can contribute to technical advancements, there is a danger that intuition is misunderstood and misused to mask transparency, accountability and explainability of AI.

**Key Words:** Artificial Intelligence (AI); artificial intuition; neuromorphic hardware; machine cognition; human subject; media

---

■收稿日期: 2019-05-17

■作者单位: 雅克布·约翰森, 英国圣玛丽大学教育、人文与社会学院商务、法律与社会学系; 英国伦敦 TW14SX  
王鑫, 辽宁大学新闻与传播学院; 辽宁沈阳 110136

■责任编辑: 刘金波