

# 弱人工智能向强人工智能转向中的生态伦理本质

姜子豪, 陈发俊

**摘 要:** 随着人工智能的发展, 其涉及的生态伦理问题愈发引人深思。人工智能在发展过程中可以进一步分为强人工智能和弱人工智能两个阶段。在弱人工智能阶段, 其生态伦理本质仍属于人与自然的关系范畴, 弱人工智能只是充当技术中介角色。而在强人工智能阶段, 从生态伦理学角度出发, 拥有自我意识的人工智能、人、自然都应作为道德关怀对象被合理对待, 此时人工智能亦由技术中介变为道德主体参与到人与自然关系的讨论中去。因此, 从根本上厘清强人工智能阶段与弱人工智能阶段的不同生态伦理本质, 是未来构建人、人工智能与自然和谐共生关系的必要前提。

**关键词:** 人工智能; 弱人工智能; 强人工智能; 生态伦理

**作者简介:** 姜子豪, 安徽大学哲学系硕士研究生; 陈发俊, 哲学博士, 安徽大学哲学系教授、博士研究生导师。

**基金项目:** 安徽省社会科学创新发展研究课题攻关研究项目“家庭与社会服务领域人工智能伦理规范问题研究”(2018CXF158)。

**中图分类号:** B82; TP18 **文献标志码:** A **Doi:** 10.3969/j.issn.2095-042X.2019.05.012

自第一台蒸汽机出现以来, 在工业革命各种新科技的“帮助”下, 人类对自然资源的超负荷掠夺引发了生态危机, 很大程度上破坏了自然界的生态平衡。与此同时, 人类自身也深陷科技异化的泥沼。进入21世纪后, 人工智能技术得到超高速发展, 它在给人们带来福利的同时, 也引发了更多前所未有的生态伦理问题。在《奇点临近》一书中, 雷·库兹韦尔提及奇点概念, 并预言, 2045年奇点来临时, 人工智能将超越人类智能, 最终达到第六纪元, 即“无智能”物质和宇宙机制将转变为精巧且具有高级形式的智能, 人机智能将胜于单纯的人类智能, 并成为人类文明的主要形式<sup>[1]</sup>。尽管目前的科技水平尚未使人工智能无限接近甚至超越人类, 但在如何妥善处理人、人工智能与自然三者的关系问题上, 我们仍需秉持未雨绸缪的态度。而在人工智能不同发展阶段中, 厘清生态环境所受的影响以及生态价值取向发生的变化, 便是思考这一问题的核心要义。

## 一、解构人工智能并审视其与生态伦理之联结

### (一) 人工智能定义

当某个事物作为讨论和研究的对象时, 人们总是试图先搞清楚它“是什么”, 对人工智能的研究也不例外。从这一概念的名称来看, 对它的定义至少应涉及“人工”和“智能”两个维度。

而作为核心维度的“智能”一词,人们目前对其并没有严格和统一的定义,所以对人工智能这个概念也就无法形成统一的结论。“智能”的原初意义主要基于“人类的”自然智能,要想对其实现揭秘最终还要靠对人脑认知机理的彻底研究。关于“智能”最流行的观点主要有“思维理论”“知识阈值理论”“进化理论”<sup>[2]1</sup>。思维理论主张智能源于思维活动。它认为思维是智能的核心,人的智力来自于脑部的思维活动,应通过对思维的研究去揭秘智能的本质。知识阈值理论将智能描述为在巨大搜索库快速找到满意的解的能力。进化理论认为智能取决于感知和行为,取决于对外部环境的适应,是一种综合能力,包括感知能力、记忆与思维能力、学习和自适应能力以及行为能力。

在对“智能”进行解构以后,我们再来探讨一下,何谓“人工智能”?目前,学术界在对人工智能定义的探索过程中,逐渐形成了三大学派,即“符号主义”“联结主义”“行为主义”。符号主义基于物理符号系统假设和有限合理性原理,认为人工智能来源于数理逻辑,计算机可以用来研究人类思维活动,模拟人的智能活动,主张通过研究人类认知系统结构,实现计算机对人的智能的模拟<sup>[3]9</sup>。艾伦·纽厄尔(Allen Newell)与赫伯特·西蒙(Herbert Simon)提出所谓的“物理符号系统”,认为计算机程序是可以用来控制物理机器操作的一系列形式化的符号,“物理符号系统对智能来说是必要的,也是充分的”<sup>[4]</sup>。联结主义基于神经网络及网络间的联结机制与学习算法,认为人工智能来源于仿生学,尤其是关于人脑模型的研究,主张通过采用模拟人的生理神经网络结构来解构人工智能<sup>[2]11</sup>。麦克洛奇(W. McCulloch)和皮兹(W. Pitts)于1943年提出神经元的数学模型(M-P模型)。M-P模型是人工神经网络最初模型,为人工智能创造了一条用电子装置模拟人脑结构和功能的新途径,亦为实现联结主义的智能模拟创造了条件<sup>[5]17</sup>。行为主义是基于控制论和“感知—动作”控制系统的人工智能学派。行为主义认为人工智能起源于控制论,提出“智能取决于感知和行为,取决于对外界复杂环境的适应,而不是表示和推理”<sup>[2]11</sup>。布鲁克斯(R. A. Brooks)提出智能不需要知识、不需要表示、不需要推理,并认为智能是在与环境交互作用中表现出来的,需要不同行为模块与环境交互,以此产生复杂行为<sup>[5]17</sup>。

由此可见,对“人工智能”的定义可谓仁者见仁,智者见智,在其发展进程中,不同学科背景的学者都提出过独特见解。英国计算机科学家阿兰·图灵(Alan Turing)提出图灵测试,即用人类表现来衡量机器智能,如果人类无法区分与之交流的是机器还是人,那么就认定该机器是智能的<sup>[5]2</sup>。蔡自兴等学者认为,就学科角度而言,人工智能是计算机科学中涉及研究、设计和应用智能机器的一个分支;就能力角度而言,人工智能(能力)是智能机器人所执行的通常与人类智能有关的智能行为<sup>[3]2</sup>。人工智能专家松尾丰认为,人工智能是采用人工方法制造的类人智能(具有发现和觉察功能的计算机)及其制造技术<sup>[6]</sup>。综合关于人工智能的种种观点,笔者从能力与学科两方面对人工智能进行阐述。从能力层面看,人工智能是采用人工手段使智能机器人能够模拟与人类智能有关的智能行为;从学科层面看,人工智能是一门涉及脑科学、计算机、心理学、哲学、认知科学等众多领域的学科。它致力于建构智能机器或智能系统,从而使得某些智能设备具备人类思维能力或实现脑力劳动自动化。

## (二) 强弱人工智能之分

20世纪80年代,哲学家塞尔(John Searle)曾提出“弱人工智能”与“强人工智能”两个概念。弱人工智能亦被称作“应用型人工智能”,它能够通过技术手段模拟人或者动物的智能,

其目的是专注于解决特定领域的问题,它本身不具有真正的智能或自主意识<sup>[7]</sup>。英国学者玛格丽特·A·博登(Margaret A. Boden)认为,“机器人根本没有意向状态,它只是受到电路和程序支配,简单地来回运动而已”<sup>[8]</sup>。也就是说,弱人工智能并没有学习思考的能力,无法拥有自己的判断思维能力。弱人工智能主要研究领域集中在智能机器人、语言识别、模式识别、图像识别、专家系统、自然语言处理以及智能搜索等方面。由此可见,目前我们所看到的人工智能算法与应用仍然属于弱人工智能范畴,在弱人工智能阶段,人类起到主导作用。随着弱人工智能在医疗、金融、军事等关键领域的应用,我们在享受其益处的同时,亦需要考虑其所引发的伦理问题,诸如隐私问题、安全与责任问题、歧视问题、失业问题等。这是我们在弱人工智能阶段必须深思的问题,亦是我們落实以人为本原则,实现人类美好生活愿望的本质要求。

与弱人工智能相对应的则是强人工智能。强人工智能又称完全人工智能,和其相似的概念还有很多,如高端“通用人工智能(Artificial General Intelligence)”、波斯特罗姆(Nick Bostrom)提出的“超级智能(Super intelligence)”、弗诺·文奇(Vernor Stefen Vinge)提出的“奇点(singularity)”等<sup>[9]</sup>。塞尔(John Searle)认为强人工智能与人类心灵等同,即具有在各方面都能与人类比肩的能力,能够进行思考、计划、解决问题、抽象思维、理解复杂理念、快速学习和从经验中学习等各种类人操作,与人类能力已无差别,甚至还能超过人类<sup>[10]</sup>。强人工智能的实质是达到人脑水平的机器智能,拥有自我意识、自我学习、自我决策的能力,不再只是一个简单工具,其本身就具有思维能力。拥有知觉和自我意识的强人工智能可以进行类人的思维活动,拥有自己的价值观和世界观体系,甚至有着和生物一样的生存需求。虽然我们距离强人工智能阶段还有很长的路要走,但随着技术指数级增长,在未来某个时刻,人工智能具有了自我意识以及人格以后,我们可能就需要考虑智能机器的道德地位、权利与责任、人机伦理关系等问题,只有这样才能在强人工智能阶段引导其向上、向善发展。

### (三) 人工智能与生态伦理关联

在当前的国际社会,不少国家仍以经济发展为主要价值导向,由此所带来的最大问题就是人类如何应对全球性的环境危机。生态环境问题已经超出了国家种族界限,从生存层面将人类共同的命运联结起来。与自然演化历史相比,人类很晚才出现在自然的舞台,但诞生不久,人类这个特殊的物种很快就适应了自然,并且极大程度地改变了地球的外貌,人类文明历史正逐渐取代自然的历史。在“顺从自然”还是“征服自然”的矛盾冲突中,人类逐渐抛弃对自然的敬畏之心,转而成为支配自然的主人。

人类逐渐强大,开始利用科学技术控制和改造自然以追求利益最大化,但是任何高新科技的出现都会对环境产生或多或少的影 响,人类过多干预自然的后果就是导致地球生态失衡。从旧石器时代起,人类就学会制造和使用工具,并且利用手中的工具去改造自己生存的环境。此后,在相当长一段时间内,人类对自然环境的破坏还很微小,尚在环境可承受的范围之内。到了工业革命时期,大机器生产的模式激发出人类前所未有的欲望,也向自然发出了前所未有的索求,各种废水、废气严重污染了地球环境,破坏了生态环境的自我调节功能。任何事物的出现都具有两面性,人工智能也不例外。它是人类社会发展到一定程度的科学技术产物,与以往的科技成果一样,无疑也会给自然的存续带来难以想象的困境。

目前,人工智能作为解放人类大脑的特殊工具,活跃在社会领域的方方面面。它一方面推动

社会更快、更好地发展;另一方面也无法逃脱科技异化的窠臼,逐渐远离科技发明的初衷,引发资源浪费、环境污染等全球性生态问题。人工智能引发的生态问题主要体现在以下三个方面:第一,自然资源的衰竭与退化。近代科技革命与产业革命以来,为了满足当前或未来的需要,人们对自然资源开发的欲望空前高涨。人工智能的发展,为人们开发与利用自然提供了更好的武器。举例而言,异构融合技术可用于自然资源数据和系统融合,图像识别技术可用于空间规划“多规合一”等。这些人工智能技术都可能导致人类对自然过度开发,引发森林锐减、水土流失、风沙加剧等生态环境问题,从而加剧资源枯竭与生物多样性减少。第二,自然环境的污染与破坏。人类不合理活动产生的物质和能量会污染自然环境,破坏生态系统。随着人工智能的发展,与其相关的智能产品得以大量研发与生产。人工智能产品不仅在使用后或淘汰后会产生污染,而且在研发过程中也会产生多种多样的污染,如电子垃圾、太空垃圾等。第三,自然生态的失衡与危机。由于人工智能产品更新迭代较快,引发的污染问题层出不穷,而人们则会因技术手段滞后而无法及时有效地应对与处理这些新发生的生态问题,从而加剧已有的生态失衡与生态危机。

生态伦理是人类道德进化的重要阶段,它是对传统伦理思想——将道德关系仅仅局限在人际间而将非人类存在排除于道德共同体之外——的批判;主张道德共同体应该包含非人类生命,乃至非生命存在物,或者说整个生态系统或自然<sup>[11]</sup>。生态伦理要求人们用一种全新的观念去对待自然,提倡人与自然和谐共生。正如学者卢风所说,“现代科技已渗透于文明(广义的文化)各个维度,科技的生态学转向不可能孤立地发生于科技领域。没有整个文明的革命,就不可能有科技的转向,反之亦然。现代文明是不可持续的,生态文明是人类文明的必由之路”<sup>[12]</sup>。当代社会发展的一个文明形态就是生态文明,以“征服、改造自然”为主题的发展模式已经不符合时代潮流。相应地,科学技术的目标亦开始由帮助人类改造自然,转变为促进人与自然和谐发展,进而更好地满足人类追求美好生活的需求。“人类社会历史的发展必然要求为科学技术的旨归注入新的时代内涵,这个内涵就要求科学技术实践要以生态维度为标尺,实现科学技术实践及其伦理规范的生态转向。”<sup>[13]</sup>在人工智能大发展时期,生态伦理的重心不仅是要探索人与自然的关系,更应该关注人、人工智能与自然三者之间的关系,必须厘清弱人工智能向强人工智能转向中的生态伦理本质,才能更好地构建人与自然、人工智能与自然以及人与人工智能之间的良性循环关系。

## 二、人与自然关系中弱人工智能体的中介地位

目前,人类所使用的人工智能都是以深度学习为基础的弱人工智能,人类用它们专注于完成某个特定的任务,例如语音识别、图像识别和翻译等。在此阶段,人工智能只被视作改造和利用自然的一种工具,抑或是一座联结人与自然的桥梁。所以,弱人工智能阶段的生态伦理本质表现为人工智能在人与自然关系中的媒介效应与工具属性。因此,在弱人工智能阶段,其引发的生态伦理问题就是人与自然的关系问题,即人在借助人工智能这一媒介时如何与自然和谐相处的问题。

人类诞生伊始就参与了自然的演化历程,同时,人与自然的关系随着科技文明的发展不断变化,人类对自然的态度也由崇拜、改造、征服逐渐转变为谋求人与自然共生共存、和谐发展。人与自然的关系问题是一个历久弥新的议题,更是人工智能时期无法回避的议题。



在原始社会,人只能被动适应自然。这一阶段人自身的认识能力较低,生产力水平低下,科学技术极端不发达。在这种原始的人与自然的关系中,自然界占据主导地位,人类只能被动接受自然的馈赠。同时,人类对自然界及人与自然关系的认识严重不足,深信天人合一和万物有灵,既崇拜自然的神秘力量,感谢自然的馈赠;又害怕改造自然的行为遭到神灵的惩罚。因此,纵然此时的人类有扩张的欲望,但是囿于人类极其有限的的能力以及对自然的高度敬畏,这种欲望也微乎其微;即使人类学会了利用火与石器工具对自然进行开采,也不会给自然带来明显的伤害,人与自然处于基本和谐状态。及至农耕文明阶段,人虽然没有改变依附自然的态势,但是生产力水平较原始社会有所提升,生产关系变得更加复杂,人逐步加深了对自然的认识,发挥了自身主观能动性对自然进行改造。虽然在这个阶段,人类开始想要摆脱自然的束缚,但是由于生产力水平还不够高,对自然的破坏仍然处在可控范围之内,自然尚可慢慢完成自我修复,因此,人与自然仍可以保持相对和谐的状态。

随着工业时代的到来,科技的快速发展加速了人们开发自然的进程。在人类中心主义的引导下,人类开始从崇拜自然转为改造和征服自然。哲学家笛卡尔(Rene Descartes)和康德(Immanuel Kant)为人类中心主义的合理性提供了论证。笛卡尔认为,只有具有理性的人才是自然的主人,而动物等其他自然存在没有这些品质,因此,人对动物等自然物没有义务<sup>[14]</sup>。康德则认为,只有理性的人才能被赋予道德关怀,而自然不具有理性,不是理性存在,它们只是人类这个目的的合适工具,故不应给予道德关怀<sup>[15]</sup>。这样的哲学思想为现代性的人与自然关系提供了合理化论证,加速了人对自然的征服与奴役。在现代社会中,人类坚持以牺牲自然环境为代价追求经济的快速发展,将人与自然置于征服与被征服的地位,打破了原本和谐的状态,人与自然的关系变得空前紧张。

在当代社会,弱人工智能的发展显然与可持续发展原则、人与自然共同进化原则以及人与自然平等原则之间存在一定的冲突。人工智能这种高科技手段,在其发展过程中伴随着资源过度消耗、环境污染以及生态破坏等问题,加剧了人与自然之间的矛盾。以智能化工厂为例,近年来很多企业为实现自动化生产开始大量购入机器人设备,在解放双手的同时也加剧了机器人原材料的快速消耗,导致了资源紧缺、社会环境不稳定等问题。而且,随着智能设备的更新换代周期越来越短,由此产生的智能产品垃圾也成为目前较为棘手的环境问题之一。更有甚者是太空垃圾,如今太空已经有超过4 500吨的太空垃圾,这些垃圾来自于废弃空间站、人造卫星、遗落的太空工具、火箭爆炸碎片等。太空垃圾极具危险性,一旦和地球相撞将产生很大危害。随着人类探索太空能力的不断增强,产生的太空垃圾也越来越多,而清理速度却远远滞后于制造速度。

因此,从人与自然关系的演变历程可以发现,弱人工智能阶段的生态伦理本质仍属于人与自然的关系。在弱人工智能阶段,生态伦理学科的年轻化导致生态理论不完善,技术增长速度过快使人工智能伦理设计出现偏颇,部分研发人员缺乏科学的生态伦理观……诸如此类的原因,导致自然开发过程中出现严重的科技异化现象,从而引发了环境伦理问题<sup>[16]</sup>。目前人工智能技术带来的生态环境问题主要有资源和能源危机以及智能机器更新换代导致的环境污染和生态破坏,人类在弱人工智能阶段无法回避这些问题。由于该阶段生态伦理的本质是基于人工智能媒介基础上的人与自然关系,所以我们在制定相关的伦理法则或法律法规时,仍然是以“人类应该如何……”这一思维模式来处理问题,但是,若要考虑到以后可能出现的强人工智能阶段,这样的

思维模式在处理问题时可能就会略显片面。

### 三、人、人工智能与自然关系中强人工智能体的主体地位

由于弱人工智能不具有自主意识,造就了其在人与自然之间只能扮演着联结桥梁的媒介角色。而在强人工智能阶段,该问题将发生质的变化。强人工智能是达到人类思维水平的机器智能,拥有知觉和意识。所以,在考量该阶段人工智能的生态伦理本质时,就需要将人工智能放在与人、自然同等的地位进行探究。因此,强人工智能阶段的生态伦理本质主要体现为一种与人类、自然平等的主体属性,而不再是一种工具属性。

#### (一) 强人工智能成为道德体的可能性

在弱人工智能阶段,我们只需要考虑人与自然两个维度;而在强人工智能阶段,由于有自主意识的人工智能的加入,使得人与自然的关系变得更加复杂。在环境伦理学的发展历程中,道德对象的范围该如何确定一直是人类中心主义与非人类中心主义争论的焦点。学者佩珀(David Pepper)认为,人类中心主义是把人类放在一切生物世界的中心,把人当作一切价值的来源,自然的价值由人赋予;美国环境伦理学家波斯勒(Richard G. Botzler)与阿姆斯特朗(Susan J. Armstrong)指出,作为哲学观点的人类中心主义把人的需要和利益当作最高宗旨,对非人实体的关注亦仅仅是因为它们自身的工具价值;埃克斯利(R. Eckersley)表示,人类中心主义认为人类是万物的主宰,是所有价值的来源和万物的尺度<sup>[17]</sup>。人类中心主义认为,人是自然界中唯一具有内在价值的存在,大自然对人类而言只具备工具价值,人是自然的主人和所有者,人类天生具有优越性,超越自然万物且与其他生物无伦理关系,故道德属于人的范畴,道德关怀的对象只能是人。各种非人类中心主义流派则提出了不同的见解。以彼得·辛格(Peter Singer)和汤姆·雷根(Tom Regan)为代表的动物解放/权利论者主张把道德关怀对象延伸至具有感知生命的动物身上。泰勒(Paul W. Taylor)的生物中心主义则提倡把道德关怀的对象扩大到一切有生命的存在。以利奥·波德(Aldo Leopold)的大地伦理学为代表的更加激进的环境主义者认为,自然界的所有事物都应是道德的关怀对象,这也被称为生态中心主义或生态整体主义。从非人类中心主义对道德对象的划分标准来看,虽然人工智能并非传统意义上的天然存在的东西,但是其具有和人类一样的独立思维和意识,甚至是行为能力,能对环境产生各种可能的影响,所以,在强人工智能阶段,人、人工智能与自然在道德上完全具有地位平等的可能性,不能片面地认为人永远是人工智能与自然的主人,也不能说人工智能是自然的主人。

#### (二) 强人工智能与人的道德关系

在强人工智能阶段,强人工智能具有成为道德主体的可能性,那么我们就需要考量拥有自我意识的人工智能的权利归属问题。当人工智能拥有自我意识后,其权利应如何归属,它是否需要被当作生命体公平对待?如何使人工智能更好地融入人类社会?显然,随着科技的发展,人工智能的权利问题是我们无法回避的。自然权利论将人权定义为从人所谓的固有属性或自然属性推演出的某种权利与规范,其强调人权主体是人而非有意识的非人存在物,本质上将人工智能排除于权利范围之外。康德则提出了一种理性权利论,认为理性的人才具有内在价值,而符号是理性的表征,人因为能使用符号表达普遍概念而受到道德关怀,非理性存在因为无法使用符号来表达普

遍概念,故而只具有工具价值<sup>[18]</sup>。其实,从这里可以看出,康德强调使用符号表达普遍概念的重要性。而强人工智能机器完全可能拥有自我思考能力,拥有使用符号表达普遍观点的理性能力。基于此,未来的强人工智能享受道德关怀、拥有权利,也具有一定合理性。契约论则将权利主体置入虚拟场景,并基于某种合乎逻辑的理由做出道德判断和选择。例如,罗尔斯提出的原始状态就是一片从未出现的空白,与洛克的白板一样,个体在这空白之中基于理性的自我利益与兴趣的考量,基于相互性需要的原则而达成一致的道德规则<sup>[19]</sup>。这从侧面说明权利是一种基于某种规则的构建活动,而具有独立意识的强人工智能当然可以参与到这种构建活动中。所以,在强人工智能阶段,人和人工智能应该被视为平等的道德构建主体。

### (三) 强人工智能与自然的道德关系

在分析了直接涉及人的伦理问题后,我们还需分析人工智能与自然之间的道德关系。在前文我们论述了强人工智能与人共同成为道德主体的可能性,强人工智能是有理性的非人存在,自然是无理性的非人存在,那么无理性的非人存在的自然是否有成为道德主体的可能性呢?如前文所述,拥有理性的强人工智能具有内在价值,因而具有成为伦理关照对象的可能。那么同样是非人类存在的无理性的自然是否拥有内在价值呢?如果有,则其具有成为道德关照的对象的可能性,反之则不能。关于内在价值,西方伦理学家一般有以下四种解释:一是可以让主体感受到愉悦的内在价值,也就是说某物或某种状态能直接让人或有意识的存在物具有愉快体验,该物或该状态就具有内在价值,这种意义的内在价值离不开有意识主体的评价;二是基于其内在属性而具有的价值,即一客体因为固有属性而拥有的价值;三是内在价值是一种客观价值,独立于评价者的主观评价;四是指一种非工具性的价值,如果客体存在自身就是一种目的,而不是为了他者实现其目的,那么它就拥有内在价值<sup>[20]303-304</sup>。西方环境伦理学通常采用第四种概念,并称之为天赋价值,当一个存在物将自身当作目的来维护时,它自存在的那刻起就具有了内在价值。从西方伦理学角度来看,自然物只要拥有生物学意义上的自我繁衍能力、生态学意义上的自我维持倾向、控制论意义上的自动平衡功能,其就具有内在价值<sup>[20]305</sup>。从这个角度看,无理性非人类存在的自然万物拥有内在价值,亦能成为道德关照的对象。虽然目前强人工智能与自然之间还没有实现事实上的地位平等,但是,在强人工智能面世的未来的道德构建中,二者之间平等的主体关系一定会得以体现。

综上所述,在未来的强人工智能阶段,无论是就强人工智能体自身而言,还是从它与人类和自然界的关系来看,它都具备了成为道德主体的资格。因此,在从弱人工智能向强人工智能转变的过程中,人工智能在生态伦理关系中的本质也必将从“媒介”转变为“主体”。

## 四、结语

由此可见,要厘清弱人工智能向强人工智能转向中的生态伦理本质,关键在于明确人工智能体本身属性的转变。通过对这一问题的探讨,我们可以发现,在弱人工智能阶段,人工智能被人类当作技术媒介用来处理人与自然的关系,其生态伦理本质主要是处理人与自然关系的一种媒介属性;而在强人工智能阶段,人工智能将成为与人类一样的道德主体参与到与自然相关的实践活动中去,其生态伦理本质由原来的技术媒介属性转变为与人类平等的主体属性。因此,只有提前

认清弱人工智能向强人工智能转向过程中的生态伦理本质,方有可能更好地应对强人工智能时代可能出现的生态伦理问题。尽管强人工智能时代尚未到来,我们也无法准确预测未来的强人工智能可以强到何种程度,但从反映未来人工智能的一些科幻电影(科幻电影在某种意义上就是对未来的一种预测)可以感知,未来的人工智能完全可能基于自己的利益诉求对生态伦理做出补益或者破坏。《黑客帝国》中,人工智能背叛了人类并与人类开战。它通过基因技术创造人类,使人类居住在其制造的虚拟世界中,以达到统治人类的目的。《罗萨姆的万能机器人》中,最早的机器人服从主人的指令,默默地为人类效力与服务,但在其意识觉醒后,便开始反抗人类的统治,最终导致人类灭亡。《银翼杀手》中,一群与人类具有完全相同智能和感觉的复制人,冒险乘坐太空船回到地球,想在其机械能量耗尽之前寻求长存的方法<sup>[21]</sup>。由此可以看出,当人工智能具有自我意识后便有可能异化,想要统治人类或者摧毁地球以构建自己的机器王国。“人们可能由于自动化而失业。人们可能拥有过多(或过少)的闲暇时间。人们可能会失去作为人的独一无二的感觉。人们可能会失去一些个人隐私权。人工智能系统的应用可能会导致责任感的丧失。人工智能的成功可能意味着人类种族的终结。”<sup>[22]</sup>基于这种预测,我们需要去考虑未来人工智能发展中,人工智能与人和自然的关系问题。“生态哲学通过对整体大全之存在以及人类理性之有限性的体认,而力主敬畏自然、保持理性的谦逊、超越物质主义,进而建设生态文明,谋求人与地球生物圈的和谐共生。”<sup>[23]</sup>当我们构建新的生态伦理法则时,应该将人工智能提升到与人类平等的地位去加以对待,让它从人类的“影子”里走出来。

#### 参考文献:

- [1] 雷·库兹韦尔. 奇点临近 [M]. 李庆诚, 董振华, 田源, 译. 北京: 机械工业出版社, 2011: 9-15.
- [2] 王万森. 人工智能原理及其应用 [M]. 北京: 电子工业出版社, 2007.
- [3] 蔡自兴, 徐光祐. 人工智能及其应用 [M]. 4版. 北京: 清华大学出版社, 2010.
- [4] 温德尔·瓦拉赫, 科林·艾伦. 道德机器——如何让机器人明辨是非 [M]. 王小红, 译. 北京: 北京大学出版社, 2017: 47.
- [5] 史忠植, 王文杰. 人工智能 [M]. 北京: 国防工业出版社, 2007.
- [6] 松尾丰. 人工智能狂潮 [M]. 赵函宏, 高华彬, 译. 北京: 机械工业出版社, 2015: 26.
- [7] 莫宏伟. 强人工智能与弱人工智能的伦理问题思考 [J]. 科学与社会, 2018 (1): 14-24
- [8] 玛格丽特·A·博登. 人工智能哲学 [M]. 刘西瑞, 王汉琦, 译. 上海: 上海译文出版社, 2006: 83.
- [9] 王彦雨. 基于历史视角分析的强人工智能论争 [J]. 山东科技大学学报(社会科学版), 2018, 20 (6): 16-27, 35.
- [10] 申锋, 何可人. 反思与超越: 人工智能的影响与应对 [J]. 常州大学学报(社会科学版), 2019, 20 (1): 108-116.
- [11] 周国文. 西方生态伦理学 [M]. 北京: 中国林业出版社, 2017: 9-13.
- [12] 卢风. 科技、自由与自然——科技伦理与环境伦理前沿问题研究 [M]. 北京: 中国环境科学出版社, 2011: 序 7.
- [13] 陈彬. 科技伦理问题研究——一种论域划界的多维审视 [M]. 北京: 中国社会科学出版社, 2014: 79.
- [14] 任雪萍, 陈浴新. 新型发展伦理价值的逻辑性 [J]. 学术界, 2015 (6): 84-92.
- [15] 霍尔姆斯·罗尔斯顿. 环境伦理学: 大自然的价值以及人对大自然的义务 [M]. 杨通进, 译. 许广明, 校. 北京: 中国社会科学出版社, 2000: 84.
- [16] 王云, 刘亚平. 人工智能技术发展中的生态伦理困境和启示 [J]. 长沙理工大学学报(社会科学版), 2018, 33 (4): 7-14.
- [17] 杨冠政. 环境伦理学概论 [M]. 北京: 清华大学出版社, 2013: 44.
- [18] 王珀. 康德式的动物伦理 [J]. 苏州科技学院学报(社会科学版), 2015, 32 (6): 33-41.
- [19] 约翰·罗尔斯. 正义论 [M]. 何怀宏, 何包钢, 廖申白, 译. 北京: 中国社会科学出版社, 1988: 120.



- [20] 何怀宏. 生态伦理: 精神资源与哲学基础 [M]. 保定: 河北大学出版社, 2002.
- [21] 王宁, 崔扬舟. 科幻影视表现人工智能的回顾与反思 [J]. 常州大学学报(社会科学版), 2019, 20 (1): 99-107.
- [22] 拉塞尔, 诺文. 人工智能——一种现代方法 [M]. 姜哲, 金奕江, 张敏, 等译. 2 版. 北京: 人民邮电出版社, 2010: 739-740.
- [23] 卢风, 余怀龙. 生态文明新时代的新哲学 [J]. 社会科学论坛, 2018 (6): 4-25.

## The Essence of Ecological Ethics in the Transition from Weak AI to Strong AI

Jiang Zihao, Chen Fajun

**Abstract:** With the development of artificial intelligence (AI), the ecological ethical issues involved in it are more and more thought-provoking. In the process of development, AI can be further divided into two stages: strong AI and weak AI. In the weak AI stage, the essence of ecological ethics still belongs to the category of the relationship between human and nature, AI just serves as a technical intermediary, while in the stage of strong AI, from the perspective of ecological ethics, AI with self-consciousness, man and natural should all be treated reasonably as objects of ethical care, then AI changes from a technology intermediary to a moral subject to participate in the discussion about the relationship between man and nature. Therefore, clarifying the difference in the essences of ecological ethics between strong AI and weak AI is a necessary prerequisite for building a harmonious symbiotic relationship among human, AI and nature in the future.

**Keywords:** artificial intelligence; weak AI; strong AI; ecological ethics

(收稿日期: 2019-04-04; 责任编辑: 陈鸿)