

# 人工智能体道德设计的美德伦理路径： 基于道德强化学习

王亮

(西安交通大学马克思主义学院,西安 710049)

**摘要:**在人工智能体的道德设计思路中常被考虑的道德理论主要包括义务论、功利主义和美德伦理,然而义务论的抽象原则和功利主义的伦理计算都忽视了道德情境的复杂性,最终表现出缺乏“情境敏感性”。相反,美德伦理注重经验知识的学习过程,以一种开放、动态的理论特质来适应复杂道德情境,而强化学习同样注重动态学习过程。因此,美德伦理与强化学习具有理论相洽性,二者的结合使人工智能体道德强化学习成为可能,并且这也是立足于现实复杂道德情境之最佳的人工智能体道德设计方案。

**关键词:**人工智能体;美德伦理;强化学习

**中图分类号:**N031 **文献标识码:**A

随着人工智能技术的突飞猛进,人工智能体在人类社会中的应用越来越深入,面对“人机共生”的社会情境,人类试图从各个层面对人工智能体进行规范,其中对人工智能体进行道德设计是当下的一个热点问题。但用何种道德理论来设计呢?通常被考虑的理论主要有义务论、功利主义和美德伦理。

## 一、道德机器的义务论路径缺陷： 抽象原则与具体道德情境的不可调和

康德是道德义务论的代表,他的思想影响深远。关于什么是义务,有两点值得我们去考虑:一方面,康德认为义务是包含着一个善的意志的概念;另一方面,义务就是出自对法则的敬重的一个行为的必然性。<sup>[1]403-407</sup> 而要以“善的意志”为出发点来设计人工智能道德将会面临巨大的挑战,甚至是不可能完成的任务,因为人工智能体的意志问题在理论上悬而未决,争论较大。较为经典的是摩尔

(James H. Moor)曾经将道德代理分为四个层次,由低到高依次为:“伦理影响代理(ethical-impact agents)”“内隐式伦理代理(implicit ethical agents)”“显式伦理代理(explicit ethical agents)”“完全伦理代理(full ethical agents)”,而最高层次的“完全伦理代理”只有人类才具有,它需要有“意识、意向性和自由意志”,这不仅是人工智能体无法跨越的红线,而且还是机器与人类之间的“关键本体论差异”。<sup>[2]</sup> 那么能否从“法则”入手来设计人工智能道德呢?康德提出了一条非常有名的“定言命令式”,即:“要只按照你同时能够愿意它成为一个普遍法则的那个准则去行动。”<sup>[1]428</sup> 而这一“命令式”就是义务遵循的法则,是“义务的一切命令式的原则”<sup>[1]429</sup>。简单来说,从“法则”入手来设计人工智能道德至少有两个步骤:(1)存在一种道德准则;(2)这一道德准则能够接受普遍法则的检验,或者说道德准则可以普遍化。顺利完成第一步是没有问题的,因为“准则是行动的主观原则”,主体可以对其进行规定。<sup>[1]428</sup> 但第二步往往难以实现,因为现实的道德情境往往是复杂多样的,将具体道德准则抽象化为

收稿日期:2021-11-8

基金项目:国家社会科学基金青年项目“跨文化视角下人工智能的伦理嵌入机制研究”(19CZX018)。

作者简介:王亮(1985—),湖北黄冈人,哲学博士,西安交通大学马克思主义学院副教授、硕士生导师,主要研究方向:马克思主义理论、科技伦理等。

普遍法则面临着极大的挑战。阿瓦德(Awad)等人的跨文化“道德机器实验”已经证实,受不同文化传统、经济、法制状况等方面的影响,很难在全球建立一种普遍化的无人驾驶汽车道德行为准则。<sup>[3]</sup>

然而,这还不能从根本上阻碍一些学者构造“康德式道德机器”的步伐。芭芭拉·赫尔曼(Barbara Herman)在《道德判断的实践》一书中设想了“康德式道德机器”设计方案:(1)道德机器必须“具有一种自然的描述性语言”;(2)道德机器的核心是拥有“道德上突出的特征清单”;(3)道德突出特征与自然描述之间具备“恰当的相互关系的映射指令”。<sup>[4]118-119</sup>赫尔曼的这一设想十分有吸引力,但问题是:它是否就是“康德式”的道德机器呢?或者说,内嵌于机器的“道德突出特征清单”能否与康德的道德法则相类比?赫尔曼关于“康德式道德机器”的设想与其“道德显著规则(rules of moral salience, RMS)”理论一脉相承,她认为这一理论能够解决康德义务论的“义务冲突问题”。<sup>[4]123</sup>当出现义务冲突凸显出来,并且权衡特定情境中不同义务的“分量”,为最终的道德判断提供依据。<sup>[4]123-125</sup>“道德显著规则”之所以能够使道德行为者具备这样的“道德敏感性”,是因为“这些规则在一种道德教育中经习得而作为要素,它们把一个行为者对他情形的感知结构化,以使他所感知到的是一个具有道德特征的世界”<sup>[4]121</sup>。这也是赫尔曼在道德机器中设置“道德突出特征清单”的理论来源,道德清单使机器也具备一定程度的“道德敏感性”,进而为复杂道德情境中的道德决策提供依据。可以看出,“道德显著规则”“道德突出特征清单”是通过“自下而上”的经验习得获得的,不仅如此,赫尔曼也并不否认“情感敏感性”对“道德显著规则”形成的积极作用。<sup>[4]129</sup>而这样一种观念完全不符合康德道德学说的本意。康德之所以如此强调道德的普遍法则,意欲彰显人的理性的力量,并且道德法则只有依据“先天原则”才能有效,而人类的经验只是一种道德“见证”,它不构成“道德形而上学”的根据,更不是道德感知结构化的关键,在康德那里,道德法则、内在道德感知结构只能通过“自上而下”的方式获得。<sup>[5]</sup>因此,赫尔曼与康德道德思想背道而驰的道德机器只能是一种伪“康德式道德机器”。还有学者试图通过“普遍化准则(universalized maxims)”、“默认规则(default rule)”背景理论、“准入-纠错

(R-in and R-out)”机制等理论来解决“康德式道德机器”的设计问题,但或者是因为无法解决抽象原则与具体道德情境的冲突问题,或者违背康德思想使本该“普遍化”的道德法则过于具体化等原因而止步不前。<sup>[6]</sup>

## 二、道德机器的功利主义路径缺陷: 难以计算的幸福

功利主义对以算法为核心的道德机器设计似乎有着天然的优势,边沁(Jeremy Bentham)很早就提出,注重结果的功利主义实则是一种“道德运算(moral arithmetic)”。<sup>[7]</sup>如何运算呢?一般来说有如下四个步骤:

“第一步,具体说明可能的行动方针  $A_1 \dots A_n$  或者行动规则  $R_1 \dots R_n$ ;

第二步,确定(预测)  $A_1 \dots A_n$  或  $R_1 \dots R_n$  的可能的后果  $C_1 \dots C_n$ ;

第三步,运用最大幸福原则(GHP),从  $C_1 \dots C_n$  中挑选出  $C_x$ ,得到最大幸福和/或最少不幸的结果;

第四步,对应地选择行动方针  $A_x$  或规则  $R_x$ 。”<sup>[8]</sup>

就第一步而言,行动方针( $A_1 \dots A_n$ )或者行动规则( $R_1 \dots R_n$ )是由设计师初始设置的,它具有确定性。而第二步中的后果( $C_1 \dots C_n$ )只是作为一种预测性的结果出现,然而智能技术的应用结果真的能准确预测吗?当代伦理学家维乐(Shannon Vallor)认为技术的加速演进必将造成“技术社会模糊性(technosocial opacity)”,也即我们无法准确预测采用新技术将如何改变我们的社会。<sup>[9]441</sup>至少在道德领域将会表现出三种“模糊性”或者预测的不准确性。第一,幸福体验模糊。机器深度参与的人-机交互生活和人类深度参与的传统生活给我们带来的是不同的幸福体验,但这两种幸福时常会发生冲突,在这两种不同的“幸福”拉扯下,幸福体验变得模糊。第二,技术效果模糊。“技术并不仅仅是由机械和电子设备组成的集合体,而是汇集了物体、人、实践和意义的复杂系统。”<sup>[10]120</sup>这种复杂性导致了很难再通过技术单一设计来预测技术效果,相反,技术效果由于各方力量的渗透融合而变得模糊,最终会导致与之相关的道德责任归

属、分配的模糊。第三,道德的“远距离”模糊。“远距离”首先体现在空间上,人工智能技术正进入我们的“私人领域”,逐渐发展成为“亲密型技术(intimate technologies)”,它们“越来越接近我们的身体、思想和生活世界”,技术的“空间”拓展意味着技术道德风险覆盖面更广、更复杂,道德风险的不确定性在增加。<sup>[11]9</sup>“远距离”还体现在时间上,面对新兴技术的快速发展,由于道德变革的滞后性,“在一段时间内,我们不再确定该用什么道德标准来处理这些影响,因为有关的技术剥夺了这些标准不言自明的相关性和真理性”<sup>[11]11</sup>。

接下来考察功利主义道德推理的第三步,看“最大幸福原则”是否成立,或者说“幸福”能否量化?“幸福”能够量化的前提是它有一个公认的统一定义。亚里士多德尝试从美德伦理的角度定义“幸福(eudaimonia)”,这一词常被翻译为“人类的幸福(human happiness)、福祉(well-being)、兴旺(thriving)或繁荣(flourishing)”,这些定义不仅多样而且模糊,每一种翻译都包含不同层次、不同角度的“幸福”描述。<sup>[12]38</sup>康德通过“自由意志”切断了“幸福”与道德动机的联系,并且否认了在现实生活中人类幸福的实现,肯定了幸福的彼岸性。<sup>[13]</sup>对于彼岸的幸福该如何定义呢?这是非常具有挑战的问题。功利主义对“幸福”的定义也不断更新,边沁尝试用简单的方式来定义“幸福”,即借助“快乐与痛苦的享乐主义概念”,密尔(John Stuart Mill)则在此基础上对快乐进行了高低层次区分。<sup>[14]</sup>其实,无论是哲学家还是心理学家对“幸福”都没有统一的定义,这里就不一一例举。造成这种现象的根本原因就在于“幸福”与我们生活紧密相关,“幸福的概念在生死观以及政治和宗教教义中起着核心作用”,对“幸福”的定义反映了人类对生活经验、愿望的多层次理解。<sup>[12]56-57</sup>

综上所述,义务论和功利主义道德机器方案的失败主要是因为义务论和功利主义过于注重道德规则,强调规则的抽象性、框定和推算功能,它们无法与具体、动态、复杂的现实道德情境相适应。那么什么样的道德理论以及实施方式能够适应复杂道德情境呢?下面将利用美德伦理和强化学习的理论资源来回应这一问题。

### 三、基于美德伦理的人工智能体道德强化学习:面向情境的动态学习

#### 1. 面向复杂道德情境的“实践智慧”

我们为什么要如此注重道德情境?这涉及人与道德之间的关系问题,一种共识是,道德体现的是人与世界交往的一种方式,即:我们应该如何对待他人和自己?<sup>[11]13</sup>这是一个与生活实践紧密相关的问题,当我们需要做出道德决策的时候,我们不是拿着道德词典翻阅后再行动,而是凭借我们对道德实践的理解、经验立即采取行动,我们进行道德决策的前提是包含道德情境的道德实践教会了我们该如何做,因为“在很大程度上,我们的道德规范和道德诀窍以具体化的知识和默契化的形式存在”<sup>[11]13</sup>。这里指出了在现实中进行道德决策的一个前提条件:学会利用道德经验知识解读道德情境。而这也恰恰是美德伦理所具备的重要特质。

亚里士多德十分注重美德知识的重要性,这与他能够深刻洞察到美德实践的不确定性有关,亚里士多德曾经形象地将美德实践比喻成“健康”“医疗”“航海”等问题,肯定了美德实践过程中包含着众多不确定性,因此主张美德行为“只能因时因地制宜”,而要做到这一点必须要运用“实践智慧”。<sup>[15]38</sup>亚里士多德不仅认识到“明智<sup>①</sup>是一种同人的善相关的、合乎逻各斯的、求真的实践品质”<sup>[15]173</sup>,而且认为“实践智慧”既需要“普遍的知识”,又需要“具体的知识”,“尤其是需要后一种知识”,而“具体的知识”就是指经验知识<sup>[15]177-178</sup>。亚里士多德的这一观点得到了当代美德伦理研究者的赞同或者阐发。罗森(Stanley Rosen)认为,“实践智慧”是一种能够将我们的内在目的和当下复杂情境相调适的知识,在这种知识的“筹划”下,我们采取正确行动。<sup>[16]</sup>此处的“实践智慧”道德知识被看作是调适目的、情境、美德行为的实践工具,具有较强的功能性特征。赫斯特豪斯(Rosalind Hursthouse)则将“实践智慧”理解为一种能力型知识,认为“实践智慧是对实际问题进行正确推理的能

① 译文中“实践智慧”被翻译为“明智”。

力”<sup>[17]13</sup>,同时强调这种知识不是一种抽象、普遍的知识,而是基于个体对“人性或人类生活方式”的深刻理解,是一种独特经验性的“知识”<sup>[17]144-145</sup>。李义天明确了“实践智慧”的推理程序:第一环节是基于“情境的感知”形成“特殊知识”;第二环节“激活关于目的的普遍知识”;第三环节“慎思”,谋求“手段和方法”;第四环节是对第三环节的结果进行“抉择”;第五环节是“手段和方法”的“实施”。<sup>[18]</sup>需要注意的是,美德伦理“实践智慧”之所以如此注重经验知识是因为它符合道德认知的规律,即我们关于道德的判断是基于道德活动过程中的建构性的具体,并最终经过具体-抽象-具体的过程回到道德实践,也即是一种既依赖于我们的经验性知识,也为我们构建新经验的边学边做过程。这一动态特征明显区别于基于提前设定好的普遍法则的义务论和基于抽象“道德运算”的功利主义。

2. “道德专业知识”与“状态-行动”

美德伦理的边学边做过程与人工智能领域的强化学习(reinforcement learning)理论完美契合。强化学习注重智能体在与环境交互过程中依靠自身经历进行学习,其基本模型原理为:特定状态(state)下的人工智能体(agent)在具体环境中(environment)执行行动(action),在与环境进行交互过程中产生新状态,同时获得奖励(reward),当行动有利于目标时产生正向奖励,不利于目标时产生负向奖励,正向奖励增加类似环境下特定行动的执行概率,负向奖励则会减少特定行动的执行概率。简图如下:

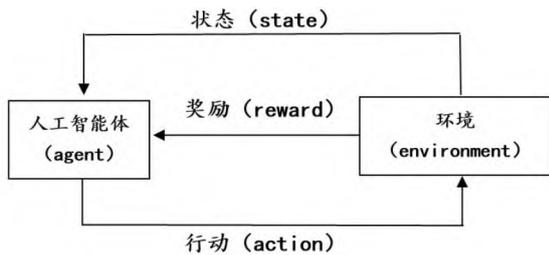


图1 强化学习结构简图

通过图1可以看出,强化学习包含了状态模块、行动模块与奖励模块,移植到道德决策理论,状态模块类似于道德认知,行动模块类似于道德行为,奖励模块其实是对道德认知的调节过程,环境可以对应于道德情境。问题的焦点是状态模块(道德认知)是提前设定的还是在环境中自主学习的?

如果是提前设定的,则与义务论和功利主义没有本质的区别。如果是自主学习的,该如何实现呢?

有学者曾经提出了“一种新的生物神经网络模型——时空联想记忆网络(spatio-temporal associative memory networks, 简称为 STAMN)作为时空经验的知识表示”<sup>[19]11</sup>。基于 STAMN 模型的强化学习主要有两大特点:第一,突出人工智能体的自主认知能力;第二,“可以实现增量式学习,不断记忆新知识”,并保留经验知识与情境之间的关系,且可以重新激活。<sup>[19]58</sup>其相关的结构图如下:

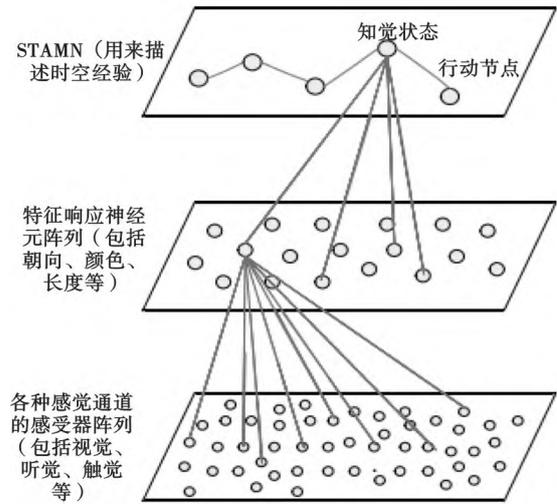


图2 与知觉直接相关的时空联想记忆结构<sup>[19]59</sup>

这样的结构设计尽可能地还原了智能体在与环境进行交互过程中的情境信息,实现了经验知识的“具体化”,它的好处是保留了知识与情境的匹配关系,而非脱离情境的“抽象化”原则。另外,“增量式学习”也能保证经验的“日积月累”,而非机械地新旧替代或者不变。所有这些特点都完全满足美德伦理基于“实践智慧”的道德认知模式,既保留了道德情境化特征,又符合美德“增量式”习得过程。现代道德认知理论强调美德是可教、可学的,认为“道德专业知识”可以通过“沉浸式”方法系统地建构,分为四个层次:第一层次是建构“识别性知识(identification knowledge)”,第二层次是建构“细节性知识(elaboration knowledge)”,第三层次是建构“程序性知识(procedural knowledge)”,第四层次是建构“执行性知识(execution knowledge)”。<sup>[20]721</sup>我们可以用 STAMN 结构图中的最下面一层来训练道德机器第一层次的道德“识别性知识”,为道德机器设计多样化的情境识别系统,综合识别道德情境;

可以用 STAMN 结构图的中间层来训练道德机器第二层次的“细节性知识”，对道德情境的独特性细节特征进行响应；道德机器第三层次“程序性知识”和第四层次“执行性知识”都可以通过 STAMN 结构图的最上层来设计。笔者认为“程序性知识”对应的应该是某种方法性知识。在亚里士多德美德伦理中类比方法被经常用到，他强调，“德性是一种适度，因为它以选取中间为目的”<sup>[15]47</sup>，就像勇敢，它既不是懦弱，也不是鲁莽，前者是不足，后者是过度<sup>[15]38</sup>。中间、不足、过度这些描述常被用于算术、几何学，亚里士多德借此类比美德，所以有学者认为“算数类比奠定了亚里士多德整个德性学说的基础”<sup>[21]</sup>。那么如何让机器能够像人一样掌握类比这样的“程序性知识”呢？分类是类比的前提，在不同情境中通过类比来准确把握美德行为是一项十分复杂的分类问题，而隐藏层神经网络模型可以解决复杂的分类问题，其基本结构图如下：

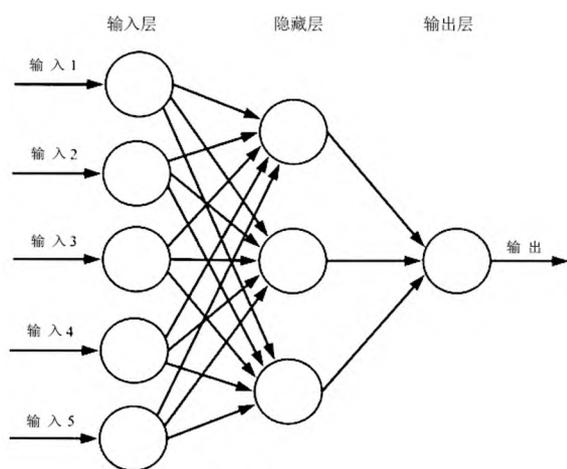


图 3 隐藏层神经网络<sup>[22]</sup>

上图的网络层可以分为三类：输入层 (input layer)、隐藏层 (hidden layer)、输出层 (output layer)，每个网络都有一个输入层和一个输出层，而隐藏层的设计增加了网络模型的复杂程度，提高了模型的非线性程度，从而可以拟合更加复杂的问题，最终其更接近真实的美德类比方法。因此，在此基础上可以训练机器掌握类比这一道德“程序性知识”。最后来看“道德专业知识”的第四层次——“执行性知识”，它连接着“认知 - 执行”，体现在 STAMN 结构图最上层的“知觉状态 - 行动节点”的结构中。“知觉状态节点用来记忆机器人的观测状态”，“行动节点用来记录机器人的动作”。<sup>[19]61-62</sup>从

“知觉状态”到“行动”的传递原理如下：当机器人与环境交互时，“知觉状态”就会被内部激活，此时机器人会按照“目标相关的状态 - 行动值函数”选择对应的“行动”。<sup>[19]62-63</sup>可以说，“行动”取决于被激活的“知觉状态”和“目标相关的状态 - 行动值函数”，道德的“执行性知识”就应当是指与道德目标相关的“状态 - 行动值函数”，它反映的是环境/道德情境信息与行动/道德行为之间的匹配关系，这种关系是机器通过强化学习，具体来说是通过“接近目标 - 获得奖励”的正反馈机制训练而构建的。

### 3. 美德目标与“奖励”

目标决定了“奖励”函数的赋值情况，在道德强化学习中起着引导作用，在强化学习中，“目标 - 奖励”“负责引导神经网络挖掘状态信息中的决策相关因素并经过提炼后用于行动的计算生成”<sup>[23]</sup>。可以用一个简单的公式反映“目标 - 奖励”的关系：

$$f(x) = \begin{cases} \text{正值}(+) , \text{接近美德目标} \\ \text{负值}(-) , \text{与美德目标无关} \end{cases}$$

当接近美德目标时，奖励值为正，正向奖励最终增加目标相关行动的执行概率；当与美德目标无关时，奖励值为负，负向奖励最终减少目标无关行动的执行概率。那么，如何为人工智能体设计美德目标呢？

亚里士多德认为“大度 (μεγαλοψυχια)”是“德性之冠”。<sup>[15]108</sup>在古希腊语里，“大度”的本意是指伟大的灵魂，它经常被用来形容“完满”的美德，也即是美德发展的最高阶段。廖申白在《尼各马可伦理学》译注里指出“完满”的一种意指是“也要求具备其他德性”，即“完满”是所有美德的总和。<sup>[15]109</sup>“所有”是指多少？我们很难给出完全的美德清单。但维乐提出的简单版美德清单可以为道德机器美德目标的设定提供参考，维乐针对当前新兴技术所产生的道德环境不确定性和多元化挑战提出了十二种核心技术美德：(1)“诚实”，包括“信任、可靠和正直”；(2)“自制”，包括“节欲、自律、节制和耐性”；(3)“谦逊”，要“认可技术社会知识和能力的限制”，摒弃对技术力量的“盲目信仰”；(4)“公正”，包括“责任感、互惠意识和慈善”；(5)“勇气”，包括“谨慎和不屈不挠”；(6)“同理心”，包括“同情、慈善和怜悯”；(7)“关爱”，包括“慷慨、爱心和服务他人”；(8)“礼貌”，“在全球网络化社会中与其他公民和谐生活”，保持合作的态度；(9)“变通”，在充满不确定性的道德情境中能够“可靠和熟

练调节”行为,包括“耐心、自制、容忍和镇定”; (10)“前瞻性”,学会“注意、识别和理解道德现象”,包括“敏锐、关注和理解”; (11)“慷慨”,包括“镇定、勇气和雄心”; (12)“智慧”。<sup>[9]447-448</sup> 这些美德是针对新兴技术社会复杂道德情境而提出的,例如“变通”和“前瞻性”,其中有些美德与经典美德重合,即便如此,维乐也赋予了它们新的道德情境适应性内涵。这些美德既可以指导道德机器设计师,也可直接用于训练道德机器。尽管要将这十二种美德移植到机器身上是一项具有挑战性的工作,但它已经开始了,比如有学者提出了一种“基于强化学习的机器人认知情感交互模型”,试图利用“情感奖励函数”对“同理心”等美德进行设计。<sup>[24]</sup> 此外,“注意”“识别”“理解”“公正”“礼貌”“关爱”等美德常常经由强化学习的方法被应用于机器人的避障系统、情感系统中。然而,美德清单的方法也存在去情境化的风险,为了避免这种风险,我们要保证清单中的美德是通过“自下而上”的案例式学习方法实现的,而不能一开始就将美德清单抽象化为道德符号。需要强调的是,美德清单只是从广度上满足了美德伦理路径下道德强化学习的目标要求,强化学习不仅要帮助道德机器掌握不同类的美德,并且还要将道德机器培养成“美德专家”,即,从深度上扩展美德,能将特定美德从简单情境扩展到复杂情境。美德专家“更擅长快速准确地解读道德情境,并决定自己可能发挥的作用”,他们往往“会注意到新手错过的细节和机会”,在解决复杂道德问题时得心应手,反应快速。<sup>[20]716</sup> 因此,从深度上设计道德强化学习目标的一条有效路径是,向美德榜样学习,从他们身上取样,分析他们所具有的美德技能。纳尔瓦埃斯(Darcia Narvaez)分别从“道德敏感性”“道德判断”“道德聚焦”“道德行动”四个阶段列举了道德模范的美德技能,并认为每个阶段都有七种需要练习的美德技能,整个专家模型中包含有二十八种典型性美德技能,它们涵盖了“传统美德”和“现代美德”。<sup>[20]716-718</sup>

尽管这里分别从广度和深度上指出了典型性美德目标,但从广度上来实现作为美德终极目标的“完满”是不现实的,无论是人还是机器都不可能集所有美德于一身,类似,从深度上来实现作为美德终极目标的无限“卓越”也是不现实的。因此,可以说美德目标的实现在实践上是一个无限接近的过程,道德强化学习的每一步也是更接近美德目标,

或者说只是在特定领域实现了美德相关目标。这也正是美德伦理和强化学习的魅力之所在,它们的相洽并不是偶然,二者都以一种动态学习特质面向复杂情境,正是这样的理论特质使得基于美德伦理的道德强化学习能够高度适应复杂道德情境。

## 四、结语

道德是我们生活实践的产物,它是有情境的,受我们的价值观、文化、政治、科技等多方面影响,道德情境体现出多元性、复杂性和不确定性,当今日新月异的新兴技术更是强化了道德情境的这些特征。在很多情况下我们做出错误的道德决策或推理,是因为不能准确地解读道德情境,或者面对新的道德情境我们缺乏足够应对之道。人工道德代理同样也会面临这样的问题,义务论和功利主义最大的缺陷就是缺乏“情境敏感性”,因此,很难利用义务论和功利主义道德理论来设计出与复杂道德情境相适应的人工道德代理。相反,美德伦理十分注重经验知识的学习,美德习得不仅证明了道德可学,而且体现为一种动态式学习过程,因此,美德伦理可以适应道德情境的变化,能够合理地将美德目标、美德行为与当下道德情境相调适。而强化学习同样注重学习过程,也是面向环境,在“目标-奖励”的引导下执行合适的“行动”。因此,完全可以让人工智能体通过强化学习的方式来习得美德,具体来说,“道德专业知识”为道德强化学习提供了合理的学习路线图,美德目标为道德强化学习提供了合理的学习方向。

## 参考文献

- [1][德]康德. 康德著作全集(第4卷)[M]. 李秋零,译. 北京:中国人民大学出版社,2005.
- [2]Moor J H. The Nature, Importance, and Difficulty of Machine Ethics [J]. *IEEE Intelligent Systems*, 2006, 21(4): 18-21.
- [3]Awad E, Dsouza S, Kim R, et al. The Moral Machine Experiment [J]. *Nature*, 2018, 563(7729): 59-64.
- [4][美]芭芭拉·赫尔曼. 道德判断的实践[M]. 陈虎平,译. 北京:东方出版社,2006.
- [5][德]康德. 康德著作全集(第6卷)[M]. 李秋零,译. 北京:中

- 国人民出版社, 2007: 222 - 224.
- [6] Powers T M. Prospects for a Kantian Machine [J]. *IEEE Intelligent Systems*, 2006, 21(4): 46 - 51.
- [7] Anderson M, Anderson S L. Machine Ethics: Creating an Ethical Intelligent Agent [J]. *AI Magazine*, 2007, 28(4): 15 - 26.
- [8] Klineciewicz M. Challenges to Engineering Moral Reasoners: Time and Context [C]// Lin P, Jenkins R, Abney K. (eds.) *Robot Ethics 2. 0: From Autonomous Cars to Artificial Intelligence*. New York: Oxford University Press, 2017: 244 - 257.
- [9] [美] 仙依·维乐. 论技术德性的建构[J]. 陈佳, 译. 东北大学学报(社会科学版), 2016, 18(5): 441 - 449.
- [10] 朱勤, 王前. 社会技术系统论视角下的工程伦理学研究[J]. 道德与文明, 2010(6): 119 - 124.
- [11] Swierstra T. Identifying the Normative Challenges Posed by Technology's 'Soft' Impacts [J]. *Etikk i praksis - Nordic Journal of Applied Ethics*, 2015, 9(1): 5 - 20.
- [12] Bok S. *Exploring Happiness: From Aristotle to Brain Science* [M]. New Haven: Yale University Press, 2010.
- [13] 杨秀香. 论康德幸福观的嬗变[J]. 哲学研究, 2011(2): 85 - 92.
- [14] White M D. The Problems with Measuring and Using Happiness for Policy Purposes [J]. *Mercatus Research*, 2015(7): 1 - 25.
- [15] 亚里士多德. 尼各马可伦理学[M]. 廖申白, 译注. 北京: 商务印书馆, 2003.
- [16] Rosen S. *The Elusiveness of the Ordinary: Studies in the Possibility of Philosophy* [M]. New Haven: Yale University Press, 2002: 119.
- [17] Hursthouse R. *On Virtue Ethics* [M]. Oxford and New York: Oxford University Press, 1999.
- [18] 李义天. 感觉、认知与美德——亚里士多德美德伦理的情感概念及其阐释[J]. 哲学动态, 2020(4): 92 - 102.
- [19] 王作为. 具有认知能力的智能机器人行为学习方法研究[D]. 哈尔滨: 哈尔滨工程大学, 2010.
- [20] Narvaez D. Integrative Ethical Education [C]// Killen M, Smetana J G. (eds.) *Handbook of Moral Development*. New Jersey: Lawrence Erlbaum Associates, Inc., 2006: 703 - 732.
- [21] 刘鑫. 亚里士多德的类比学说[J]. 清华西方哲学研究, 2015, 1(1): 400 - 431.
- [22] Howard D, Muntean I. Artificial Moral Cognition: Moral Functionalism and Autonomous Moral Agency [C]// Powers T M. (ed.) *Philosophy and Computing: Essays in Epistemology, Philosophy of Mind, Logic, and Ethics*. Cham: Springer International Publishing AG, 2017: 121 - 159.
- [23] WYJJYN. 深度强化学习落地方法论 [EB/OL]. [2019 - 12 - 12] (2021 - 08 - 10). <https://zhuanlan.zhihu.com/p/97032357>.
- [24] 黄宏程, 李净, 胡敏, 等. 基于强化学习的机器人认知情感交互模型[J]. 电子与信息学报, 2021, 43(6): 1781 - 1788.

## A Virtuous Ethical Approach to Moral Design of Artificial Intelligence Systems: Based on Moral Reinforcement Learning

WANG Liang

(College of Marxism, Xi'an Jiaotong University, Xi'an 710049, China)

**Abstract:** The moral theories that are often considered in the moral design of artificial intelligence systems include deontology, utilitarianism, and virtue ethics, but both deontological abstract principles and utilitarian ethical calculations ignore the complexity of moral situations and ultimately show a lack of "situational sensitivity". In contrast, virtue ethics focuses on the learning process of empirical knowledge and adapts to complex moral situations with an open and dynamic theoretical quality, while reinforcement learning also focuses on the dynamic learning process. Therefore, virtue ethics and reinforcement learning have theoretical compatibility, and the combination of the two makes moral reinforcement learning possible for artificial intelligence systems, and it is the best moral design solution for artificial intelligence systems based on realistic and complex moral situations.

**Key words:** artificial intelligence systems; virtue ethics; reinforcement learning

(本文责任编辑: 崔伟奇 郑泉)