

论无人驾驶汽车的伦理设计原则

杜严勇

摘要: 随着无人驾驶技术的快速发展与各级政府的大力推进,无人驾驶汽车将会得到越来越广泛的应用。然而,国内外学术界不少学者还在为无人驾驶汽车的若干伦理困境争论不休。科技伦理的研究目标是为科技发展提供伦理规制与伦理支持,而不是阻碍技术的发展与应用。为了更好地推动无人驾驶技术的发展,我们需要更多地从用户的视角考虑伦理问题。交通规则应该成为无人驾驶汽车伦理设计的根本指导原则,由此可以消解目前学术界讨论的大多数所谓伦理困境,进而有力地促进无人驾驶汽车的推广应用,使其更好地造福社会。

关键词: 无人驾驶汽车;伦理设计;伦理困境;交通规则

中图分类号: B82 **文献标识码:** A **文章编号:** 1003-0751(2023)09-0108-06

近几年,很多著名的大型车企以及互联网、通信等非传统汽车公司都在积极研发无人驾驶汽车,并从多方面进行战略布局。不少工业园区、旅游景点和大学校园内已经可以看到无人驾驶汽车的忙碌身影。同时,很多城市也在积极推动无人驾驶汽车进入实际应用。据报道,截至2022年6月,我国已有北京、上海、广州、深圳、重庆、长沙等40多个城市积极出台相应管理办法或实施细则,从道路测试、载人测试、示范运营到无人化等多个阶段开展积极探索。对于已经来到我们身边的无人驾驶汽车,国内外不少学者还在为无人驾驶汽车的某些伦理问题纠结徘徊。针对目前学界关注的无人驾驶汽车的若干伦理困境,有必要对其进行批判性反思,并采取合理立场,把交通法规作为无人驾驶汽车的伦理设计原则,为实质性解决无人驾驶汽车的伦理困境提供一种较为可行的方案。

一、无人驾驶汽车的伦理困境

总体来看,目前国内外学术界关注的无人驾驶

汽车的伦理困境主要表现为以下几个方面。

1. 义务论与功利主义之争

义务论与功利主义之争是科技伦理研究中的代表性争论之一,在无人驾驶汽车领域亦颇受关注。功利主义看重行为的后果,强调人们的行为应该给最大多数人带来最多的益处。戈戈尔等人认为无人驾驶汽车伦理设计可以按照两种模式进行:一种是个人伦理设置,另一种是强制伦理设置。个人伦理设置充分尊重个人选择,但可能会导致囚徒困境,这是因为一个遵守道德规范的人不清楚其他人是有道德的还是自私的,因此他可能会这样想:我不能确定别人是否会像我这样遵守道德规范,那么我就不应该为了他们而牺牲我自己。为了避免囚徒困境,我们可以在无人驾驶汽车的伦理设计中采取强制伦理设置,政府为无人驾驶汽车提供伦理规制,使交通安全最大化,同时使所有可能受到影响的人的伤害最小化,也就是遵循功利主义原则^[1]。

与功利主义不同,义务论强调行为本身的原则性与规范性。参照阿西莫夫的机器人三定律,格迪斯等人提出了无人驾驶汽车的三定律:一是无人驾

收稿日期:2023-04-30

基金项目:国家社会科学基金重大项目“人工智能伦理风险防范研究”(20&ZD041)。

作者简介:杜严勇,男,同济大学人文学院特聘教授、博士生导师,国家社会科学基金重大项目首席专家(上海 200092)。

驶汽车不能碰撞行人和骑自行车的人;二是无人驾驶汽车不能碰撞其他汽车,除非与第一定律相冲突;三是无人驾驶汽车不能碰撞其他任何物体,除非与第一、第二定律相冲突^[2]。古柯-维拉认为无人驾驶汽车无法基于功利主义原则做出令人满意的、不会引起争议的抉择,对碰撞过程中的各种因素、各方利益等进行权衡比较也并非易事,无人驾驶汽车的碰撞算法应该基于刑法正当性理论的义务论原则来进行设计,由此可以提供一种法律上可以接受的解决方案^[3]。

2. 电车难题:义务论与功利主义之争的具体表现

电车难题可以看作是义务论与功利主义之争的具体表现,大多数讨论无人驾驶汽车伦理困境的论著都会涉及这个思想的实验。电车难题是由富特提出的著名思想实验,大致情况为:假如你是一名电车司机,正驾驶电车驶过弯道,突然发现前面有五位正在维修轨道的工人,如果你要避免撞倒他们,就必须停车,但不幸的是,刹车却失灵了。此时,你发现前面还有一条支线,可以将电车开到支线上从而避免撞倒五位工人,不过支线上有一名工人在施工。那么,你该如何选择?电车难题以后又有很多变化,也得到哲学家们的颇多关注^[4]。

帕特里克·林认为如果无人驾驶汽车得到应用的话,电车难题可能出现在现实世界中。后果论(功利主义)者会认为应该转换轨道来救五个人,尽管这样会牺牲一人。但是,非后果论者不仅仅是考虑事情的结果,他们可能认为转换轨道构成了谋杀(一人)行为而反对这样做,尽管不作为会听任某些人(五人)死亡,而谋杀在伦理和法律两方面都要比听任死亡更糟糕。谋杀意味着你对某人的死亡负有直接责任,而听任死亡对你而言只涉及较少的责任。在帕特里克·林看来电车难题中的任何一种选择都可以得到辩护,因此无人驾驶汽车的伦理设计应该保持透明,而不仅仅是公开最终的结果^[5]。不过,尼霍尔姆等人认为在无人驾驶汽车伦理问题的讨论中,把电车难题和事故算法进行简单类比是不恰当的。他们认为无人驾驶汽车事故算法与电车难题存在三方面的区别:一是在整体决策环境及其特征方面,前者涉及多方面的利益相关者,是一种预期决策,需要考虑的因素是无限多的,而后者只涉及单一个体在当下的决策,需要考虑的因素很有限;二是前者需要同时考虑道德与法律两方面的责任问题,而后者都没有进行考量;三是在决策者的认识情境方面,前者是在诸多不确定因素下的风险评估与决策的纠缠,而后者面临的因素是确

定的和已知的^[6]。

3. 车内人员还是车外人员优先问题

隧道难题可以看作是电车难题的变种,但隧道难题关注的重点是车内人员与车外人员的矛盾。隧道难题大致情况为:萨拉乘坐无人驾驶汽车沿着山上的单行道行驶,正在快速接近一条狭窄的隧道。忽然一名儿童误入道路中间,正好挡住了隧道入口,汽车已经来不及刹车以避免碰撞。现在萨拉面临两种选择:要么直行撞上儿童致其死亡,要么急转弯撞上隧道墙壁,撞死萨拉自己^[7]。为了解决车内人员与车外人员的矛盾,孔泰萨等人提出了伦理旋钮的设计思路,也就是给无人驾驶汽车安装一种装置,让乘客在三种设置中选择一种:一是优先保护车外人员的利他模式,二是同等对待车内与车外人员的中立模式,三是优先保护乘客的利己模式。当然,从利他模式到中立模式再到利己模式,还存在一些过渡状态^[8]。

4. 无人驾驶的责任问题

以自动驾驶汽车为典型,人工智能应用的责任鸿沟成了亟待解决的新问题。责任鸿沟实质上是机器智能体在自主性不断增强但又尚未获得责任主体地位阶段的一种责任无着落现象,它甚至可能构成人工智能应用的瓶颈^[9]。因此,国内外很多学者普遍关注无人驾驶的责任问题。也就是说,如果无人驾驶汽车发生了事故,究竟应该由谁来承担责任,是车内人员或车主,是汽车的设计者、制造者、销售商,还是无人驾驶汽车本身?有学者认为,无人驾驶汽车的伦理与法律责任在传统的伦理和法律框架中不可能得到合理的解决,我们只能从全新的视角来分析并解决这个问题。我们很容易发现,法学家在讨论无人驾驶汽车的责任问题时,主要是围绕产品开发商、生产商、销售商和使用者来展开的,无人驾驶汽车本身根本不是承担责任的主体。也有学者认为,无人驾驶汽车的责任困境产生的原因在于新技术对传统责任概念的重塑,同时,无人驾驶汽车改变了汽车与使用者的边界及二者的关系,使用者应该承担部分责任^[10]。

5. 是否需要考虑情感因素的困扰问题

情感在人类的道德活动中扮演着重要的角色,情感主义伦理学甚至认为人类的伦理判断主要是基于情感而不是理性做出的。现代认知神经科学与心理学等领域的研究都表明,情感在人类的道德判断中发挥着重要的基础性作用。同时,人类的道德判断通常以外在的情感反映出来,具体表现为人们的外部表情与行为举止等方面。因此,如果忽视情感的作用,那么理论上做出的道德推理与行为评价很可能与现实

相差甚远。而且,由于人类的内在情感是非常丰富的,不同个体持有不同的价值观,这些个体的差异性在无人驾驶汽车的设计中并没有得到充分的考虑,导致了无人驾驶汽车的同质化趋向^[11]。

二、关于无人驾驶汽车伦理设计的几点反思

对无人驾驶汽车伦理困境的理论探讨,目前而言,学者们难以达成较大程度的共识,从而在很大程度上使得无人驾驶汽车的伦理设计问题陷入僵局。或许我们应该转换视角、纠正预期,从无人驾驶汽车伦理设计的基本目标出发来反思这些伦理困境,进而为探讨可行的伦理设计原则提供有力支持。

1. 视角转换:从人类驾驶员转到无人驾驶汽车及用户

关于无人驾驶汽车伦理困境的讨论,目前大多数是从人类驾驶员的角度出发的。也就是说,研究者把无人驾驶汽车想象成一名司机,讨论司机在某种情境中应该如何进行伦理决策。但是,即使人们愿意让无人驾驶汽车做出某些伦理判断,但它们进行判断的机制毫无疑问与人类存在着重大差异。就情感因素而言,并不是所有的人类情感在伦理判断中所发挥的作用都是积极的。我们希望无人驾驶汽车比人类更理性、更客观、更冷静,从而做出比人类更合理的伦理判断。也就是说,我们不能从单个人类驾驶员的角度思考无人驾驶汽车的伦理设计问题,而应该从无人驾驶汽车的特点出发进行考量。

而且,无人驾驶汽车的本质是要代替人类司机,虽然有的无人驾驶汽车目前还保留着人类控制的模式,但随着技术的发展进步,人类将逐渐退出对汽车驾驶的干预。因此,我们应转换视角,更多地从用户的角度来讨论问题。通过对近两千名美国公众的调研表明,大多数人倾向于认为遵循功利主义的无人驾驶汽车是更道德的,但他们自己仍然倾向于乘坐优先保护车内人员的无人驾驶汽车。因此,如果采用两种不同伦理设计模式的汽车都投放市场的话,那么很少有人会愿意乘坐功利主义的无人驾驶汽车,尽管人们都希望别人去选择这样的汽车^[12]。用户的信任与接受是无人驾驶汽车能够广泛应用的基本前提,因而无人驾驶汽车的伦理设计必须满足公众的伦理预期。

2. 伦理设计的基本目标:为技术发展与接受提供伦理支持

人们普遍认为无人驾驶汽车可以使出行更加便

捷、安全,对于出行不便的人群更为有利,而且基本可以杜绝人为因素发生的交通事故,同时减轻交通拥堵、环境污染与停车难等问题,无人驾驶汽车的广泛应用是社会发展的趋势。因此,无人驾驶汽车伦理设计的基本目标是充分掌握技术现状与趋势,全面把握用户的认知与预期,为无人驾驶汽车提供更具有可操作性、更可接受的伦理设计原则,从而为技术发展提供有力的伦理支持,推动技术健康理性的发展,而不是试图阻碍技术发展,因为这是完全无法做到的。

通过电车难题等思想实验来论证无人驾驶汽车的某种伦理设计原则的合理性与局限性,当然具有重要的理论意义,但正如前文提及的尼霍尔姆的批评那样,电车难题确实是高度简化的思想实验,在现实情景中我们会面临更复杂的情况。还有学者讨论了电车难题中所涉及人员的年龄情况、健康状态、伤害程度、社会地位等因素,以及周围环境(如高山还是草地、高速公路还是城市道路等)的差异,由此得出的结果显然会有显著的不同。因此,无论是义务论与功利主义的权衡,还是电车难题与隧道难题的讨论,都无法为无人驾驶汽车的伦理设计提供一个较为可行的方案。不少学者都担心,如果我们不能就无人驾驶汽车的伦理与法律准则达成共识的话,很可能会影响日益成熟的无人驾驶汽车进入商用阶段,从而阻碍技术的进一步发展,影响技术带给我们的巨大红利。

3. 纠正预期:技术系统及其伦理判断不可能完美无缺

目前讨论无人驾驶汽车伦理设计原则的论著基本上都暗含了一个前提预设,就是无人驾驶汽车的某种伦理设计原则与现实的伦理判断、伦理抉择可以得到普遍认可与接受。无论是主张义务论的学者,还是赞同功利主义的学者,都希望自己的理论主张能够得到肯定。同时,大力宣传无人驾驶优越性的科技工作者虽尽力强调技术的可行性与先进性,但时常见诸报端的无人驾驶汽车事故又使得公众对其先进性持一定程度的怀疑态度。

事实上,无人驾驶汽车即使有诸多优点,也不可避免地会产生交通事故,甚至会引起新的社会问题。一是无人驾驶技术本身不可能是完美无缺的。由于交通环境的复杂性及人类认知的有限性,使得包括无人驾驶在内的任何技术都不可能是完美的。技术的先进性并不意味着不发生事故,只不过事故发生率与严重程度会显著降低而已。二是无人驾驶汽车的伦理判断与伦理抉择也不可能得到所有人的认同。从人类整体的角度看,当然存在某些人类可以普遍接受

的共同价值与伦理原则,但对包括无人驾驶汽车在内的具体技术产生的事故进行伦理评价,却很难仅仅以这些人类共同价值与原则作为标准。同时,几乎所有的伦理评价都是高度情境化、个性化的,更何况不同主体对同样的伦理原则还可能存在的理解。如果我们能做到让无人驾驶汽车的伦理设计原则得到大多数人的肯定,这就足够了。

4. 伦理设计的透明性:知情同意与知情选择的重要性

美国哈佛大学的心理学家格林认为,在生与死之间进行权衡令人不快,而且无论无人驾驶汽车采用哪种伦理原则,它们都会受到公开的强烈质疑,这使得制造商不愿意公开他们的伦理设计原则。因此,我们面临的关于无人驾驶汽车的伦理问题中最紧迫的可能是透明性问题^[13]。事实上,透明性原则是人工智能伦理原则中受到最多关注的伦理原则。虽然人工智能系统的透明性受到某种程度的局限,但人工智能决策过程与机制的可解释性与可说明性是至关重要的,也是用户能够普遍信任并接受智能产品的一个基本前提。也就是说,无人驾驶汽车的伦理设计原则必须是透明的,这是保障用户知情权的基本要求。

知情同意是科技伦理的一项基本原则,也是尊重人的自主性的基本表现。自主性是一个人按照她/他自己选择的计划决定她/他的行动方针的一种理性能力,人的自主性就可以有力地证明知情同意原则的必要性^{[14]236}。因此,无论是设计研发者还是制造商、销售商,都不能隐瞒无人驾驶汽车的伦理设计原则。如前文所述,为了使公众接受无人驾驶汽车,我们建议无人驾驶汽车的默认设计原则是优先保护车内人员,但是我们也建议增加设计的多样性供用户选择,比如,用户也可以选择优先保护其他人员的伦理设计模式,因为现实社会中确实有不少像雷锋那样道德高尚的人。我们鼓励无人驾驶汽车的用户学习雷锋,也提供相应的机会,但这显然不能成为一种硬性要求。同时,我们也不建议给这种利他型的无人驾驶汽车以特殊的标识,因为这很可能增加车内人员受伤的风险。

三、交通规则作为伦理设计原则的可行性

为了消解学术界讨论的关于无人驾驶汽车的伦理困境,根据前文所述的几点反思,笔者认为,交通规则作为无人驾驶汽车的伦理设计原则,具有很强的可行性。

1. 遵守交通规则是无人驾驶汽车基本的伦理与法律准则

人类驾驶员必须遵守交通法规,这是一条基本的伦理与法律准则。显然,无人驾驶汽车也应该遵守现行交通规则,这一点应该很容易得到人们的认可。有学者认为,关于无人驾驶汽车的伦理困境研究应该着力解决更加实际的问题,我们应该“改变规则思维,聚焦共识达成”^[15]。让无人驾驶汽车遵守交通规则,应该很容易成为人们的共识,在此基础上我们才能讨论其他的相关问题。

我们主张无人驾驶汽车在遵守交通规则的前提下,如果有不可避免的碰撞发生,应该优先保护车内人员,有人可能会认为这对车外人员不公平。但是,在无人驾驶汽车普遍应用的年代,乘客与行人的角色可能是随时变换的,很可能现在你是乘客,但没过多久你到达目的地之后就成了行人。因而对所有人来说机会是平等的,并不存在明显的歧视与不公平。

对于无人驾驶汽车的伦理设计模式,有的学者主张应该让无人驾驶汽车学习人类驾驶员的伦理判断模式,然后在学习的基础上做出与学习对象类似的伦理判断与抉择,并称之为“自我道德原则选择”^[16]。这种思路至少存在两方面的问题:一是人类驾驶员的伦理判断模式并不是最优的,甚至部分人类驾驶员还存在不良驾驶习惯,学习人类的驾驶模式并不是最优的选择;二是如果用户并不拥有驾驶资格与技能,那他该如何选择无人驾驶汽车的伦理设计模式呢?

当然,无人驾驶汽车确实需要人类驾驶员的灵活性。也就是说,无人驾驶汽车对交通规则的遵守应该保持一定程度的灵活性,而不是机械地无条件地遵守。特别是对某些特殊情况的处理,比如,假如道路中央是原则上不可跨越的实线,但如果无人驾驶汽车为了避免撞到前方的人或障碍物,就可以跨越实线。也就是说,为了避免碰撞事件的发生,可以在保障安全的情况下违反部分交通规则。

2. 把交通规则作为伦理设计原则兼有道义论、功利主义的优点

我们明确反对电车难题中简单地根据人数多少进行功利主义的选择,因为这样可能会产生更加不合理的后果,如法不责众。而且,如果简单地以人数多少来作为无人驾驶汽车的选择标准,那么会使少数严格遵守交通规则的人受到伤害,而多数不遵守交通规则的人反而受到保护,这其实违背了功利主义的基本原则。正如有学者指出的那样,正是由于

考虑人类利益的质量,所以最大幸福原则不允许用无辜者的生命来挽救事故当事人,即使用一人来挽救五人也不行,因为无辜者能够代表多数人的利益,因而他们的生命质量比有幸者的要高些^[17]。

交通法规作为无人驾驶汽车伦理设计原则,可以视为规则功利主义的伦理设计思路,它兼有义务论和功利主义的特点。康德的义务论要求:“要这样行动,使得你的意志的准则任何时候都能同时被看做一个普遍立法的原则。”^[18]⁴⁴无人驾驶汽车遵守交通规则,不用考虑各种复杂的场景,也不用考虑行为的后果,这是义务论的立场。绝大多数人都希望无人驾驶汽车比人类更严格遵守交通规则,这应该符合人们的伦理预期。同时,这也符合规则功利主义的特点,因为“每个人都始终应当确立和遵循会给一切相关者带来最大好处的规则”^[19]⁴⁴。无独有偶,有的法学研究者比较全面地对功利主义、规则功利主义以及义务论等不同伦理立场之后认为,无人驾驶汽车的伦理理念应当立足于康德义务论中的人本主义观,并具体采取规则功利主义的立场^[20]。

同样重要的是,从长时段的角度看,无人驾驶汽车的这种伦理设计原则会产生积极的长期影响。因为无论是在什么环境中,总会有一部分人的交通规则意识较为淡薄,而严格遵守交通规则的无人驾驶汽车优先保护遵守交通规则的人,不去简单比较可能受到伤害的人数,由此可以产生长期效应和示范效应,使人们更加自觉地严格遵守交通规则,这无论是对个人还是对社会都是有益的。

3. 遵守交通规则有利于无人驾驶技术的推广应用

虽然无人驾驶技术发展迅猛,但大范围推广应用可能还需要较长的时间。因此,在未来较长的一段时间内,应该是人类驾驶汽车与无人驾驶汽车混合存在的状态。从当前人类驾驶汽车主导的交通环境,到两种驾驶方式混合的交通环境,再到无人驾驶主导的交通环境,都遵循同样的交通规则,毫无疑问有利于交通环境与规则的平稳过渡。等到未来无人驾驶完全取代人类驾驶员之后,再按全新的交通环境转变交通规则就容易多了。

无人驾驶汽车产业的发展关系到全球汽车产业的未来发展。世界各国的交通规则虽然有一定差异,但总体上看大同小异,而且基本上趋于成熟和稳定,也易于无人驾驶汽车遵守。当无人驾驶技术在不同国家进行推广时,同样的伦理设计原则也易于

不同的国家接受。也就是说,让无人驾驶汽车遵守交通规则也有利于无人驾驶技术的跨国应用。不过,当前的交通规则主要是基于人类驾驶员而设计的,可能并不利于充分发挥无人驾驶技术的优点。但是,如前所述,在无人驾驶汽车逐渐推广应用的过程中,我们不能要求无人驾驶汽车能够拥有完美的技术表现,包括在伦理判断与伦理抉择方面。而且,随着相关技术的普及应用,我们可以积累大量的数据,在此基础上可以更加充分地发挥人工智能技术的优越性,进一步完善相关技术,使无人驾驶汽车在当前的交通规则下充分发挥技术的优越性。

4. 合理处理技术与交通环境的不确定性导致的事故

首先需要强调的是,目前关于无人驾驶汽车伦理困境的研究,主要集中于讨论那些不可避免发生碰撞的情况时的处理模式。但是,无人驾驶汽车可以更合理地控制车速与车距,更冷静、理性地驾驶,而且随着无人驾驶技术的日益完善与成熟,可以很大程度地减少碰撞事件,这一点也是学者们普遍认可的。当然,由于技术以及其他多种因素的不确定性,使得碰撞事件不可能完全避免,因而,虽然碰撞是小概率事件,但仍然有一定的研究意义。

无人驾驶汽车如果发生了包括碰撞在内的各种事故该如何分配责任呢?笔者认为至少应该做到两点:一是应澄清事故发生的主要原因。如果通过严格的检查程序认定事故是由技术缺陷导致的,那么技术研发者、制造商应该承担一定的责任。同时,我们也不应该将所有责任都交由研发者和制造商来承担。如果他们遵循了相应的技术标准和职业规范,不存在主观故意等过错,则可以主要由保险公司来承担赔偿责任。如果无人驾驶汽车得到普遍推广,到时所有的汽车也必须强制购买保险,那么由于事故率的大幅降低,保险公司可以大幅提高无人驾驶汽车事故的赔偿力度,这应该也是所有利益相关者比较容易接受的一种处理模式。二是如果事故发生的原因不是技术缺陷导致的,而是由于包括交通环境在内的各种不确定因素导致的,则可以把无人驾驶汽车发生此类事故归为一种不好的“道德运气”。有学者认为,凡在某人所做之事有某个重要方面取决于他所无法控制的因素,而我们仍然在那个方面把他作为道德判断对象之处,那就可以称之为道德上的运气^[21]²⁹。由于交通环境涉及人类、其他生命、自然环境、天气状况等很多方面的因素,每种因素都具有很大的不确定性、高度复杂性与多变性,因

而此类因素导致的交通事故是无法完全避免的。而且,此类事故也可以主要由保险公司来承担赔偿责任。总之,我们不建议由技术研发者与产品制造商、销售商来承担较多的责任,因为这对于技术的发展及应用是不利的。由此可见,无人驾驶汽车的责任困境可能并没有人们通常所论述的那样严重,而且也有较为合理的解决途径。

不少学者与公众对无人驾驶汽车的安全性表示担忧,我们非常理解这些忧虑,我们也确实应该对无人驾驶汽车的应用持谨慎态度。但是,就无人驾驶技术的发展现状与发展趋势来看,我们完全有理由对其持乐观态度。与其在理论层面对一些小概率事故的处理模式争论不休,还不如从更现实的角度分析如何解决无人驾驶汽车面临的实际问题,不应让科技伦理滞后于技术的发展速度,更不应让伦理观念的滞后成为影响技术发展的阻碍因素。本文的论证试图表明:让无人驾驶汽车严格地遵守现有的交通规则可以消解目前学术界讨论的所谓伦理困境,也显然有利于无人驾驶技术的推广应用,从而使其更好地造福社会。

参考文献

- [1] GOGOLL J, Müller J. Autonomous Cars: In Favor of a Mandatory Ethics Setting[J]. *Science and Engineering Ethics*, 2017(3): 681-700.
- [2] GERDES J, THORNTON S. Implementable Ethics for Autonomous Vehicles[C]// Maurer M, et al. (eds). *Autonomous Driving*, Berlin: Springer, 2016: 87-102.
- [3] COCA-VILA I. Self-Driving Cars in Dilemmatic Situations: An Approach Based on the Theory of Justification in Criminal Law[J]. *Criminal Law and Philosophy*, 2018(1): 59-82.
- [4] THOMSON J. The Trolley Problem[J]. *The Yale Law Journal*, 1985(6): 1395-1415.
- [5] LIN P. Why Ethics Matters for Autonomous Cars[C]// Maurer M, et al. (eds). *Autonomous Driving*, Berlin: Springer, 2016: 69-85.
- [6] NYHOLM S, SMIDS J. The Ethics of Accident-Algorithm for Self-Driving Cars: an Applied Trolley Problem? [J]. *Ethical Theory and Moral Practice*, 2016(5): 1275-1289.
- [7] MILLAR J. An Ethics Evaluation Tool for Automating Ethical Decision-Making in Robots and Self-Driving Cars[J]. *Applied Artificial Intelligence*, 2016(8): 787-809.
- [8] CONTISSA G, LAGIOIA F, SARTOR G. The Ethical Knob: Ethically-Customisable Automated Vehicles and the Law[J]. *Artificial Intelligence and Law*, 2017(3): 365-378.
- [9] 王天恩. 人工智能应用“责任鸿沟”的造世伦理跨越:以自动驾驶汽车为典型案例[J]. *哲学分析*, 2022(1): 15-30.
- [10] 白惠仁. 自动驾驶汽车的“道德责任”困境[J]. *大连理工大学学报(社科版)*, 2019(4): 13-19.
- [11] 谢惠媛. 民用无人驾驶技术的伦理反思[J]. *自然辩证法研究*, 2017(9): 39-43.
- [12] BONNEFON J, Shariff A, Rahwan I. The Social Dilemma of Autonomous Vehicles[J]. *Science*, 2016(6293): 1573-1576.
- [13] GREENE J. Our Driverless Dilemma[J]. *Science*, 2016(6293): 1514-1515.
- [14] 邱仁宗. *生命伦理学*[M]. 北京: 中国人民大学出版社, 2010.
- [15] 和鸿鹏. 无人驾驶汽车的伦理困境、成因及对策分析[J]. *自然辩证法研究*, 2017(11): 58-62.
- [16] 李伟, 华梦莲. 论自动驾驶汽车伦理难题与道德原则自我选择[J]. *科学学研究*, 2020(4): 588-594.
- [17] 陈晓平, 翟文静. 关于自动驾驶汽车的立法及伦理问题[J]. *山东科技大学学报(社科版)*, 2018(3): 1-7.
- [18] 康德. *康德三大批判合集: 下*[M]. 邓晓芒, 译. 北京: 人民出版社, 2009.
- [19] 蒂洛, 克拉斯曼. *伦理学与生活*[M]. 程立显, 刘建, 译. 成都: 四川人民出版社, 2020.
- [20] 李飞. 无人驾驶碰撞算法的伦理立场与法律治理[J]. *法制与社会发展*, 2019(5): 167-187.
- [21] 内格尔. *人的问题*[M]. 万以, 译. 上海: 上海译文出版社, 2004.

On the Ethical Design Principles of Driverless Cars

Du Yanyong

Abstract: It is obvious that driverless cars will be more and more widely used with the rapid development of driverless technology and the vigorous promotion of governments. However, domestic and foreign academic circles are still debating several ethical dilemmas for driverless cars. This essay argues that the research goal of science and technology ethics is to provide ethical regulation and ethical support for the development of science and technology, not to hinder the development and application of technology. In order to promote the development of driverless technology, we need to consider ethical issues from the perspective of users. Traffic rules should become the fundamental guiding principles for the ethical design of driverless cars, which can resolve most of the so-called ethical dilemmas currently discussed in academic circles, thus effectively promoting the application of driverless cars and making them better benefit society.

Key words: driverless cars; ethical design; ethical dilemma; traffic rules

责任编辑: 思 齐