

【法学研究】

无人驾驶汽车碰撞程序的刑法正当性

——基于行为功利主义理论*

杨 昆 鹏

(吉林大学 司法文明协同创新中心, 吉林 长春 130012)

摘 要:“无人驾驶汽车的时代”已经来临,“电车难题”伦理困境也已经变成无人驾驶汽车碰撞程序的现实选择。无人驾驶汽车碰撞行人或牺牲乘客必然侵害法益构成刑法违法,而无人驾驶汽车正常进入交通系统又使其碰撞程序存在违法阻却事由。借助于行为功利主义理论的规范性、刑法适用性、平等性的优势,通过人数比较将抽象化的“法益”转化成可计算的利益衡量,相关利益主体民主商议达成“被牺牲者同意”,从而构建“补强的紧急避险”理论,为无人驾驶汽车碰撞程序在行人之间、行人与乘客之间的选择提供违法阻却事由。

关键词:无人驾驶;行为功利主义;法益;补强的紧急避险;违法阻却

DOI: 10.13568/j.cnki.issn1000-2820.2021.05.003

中图分类号: D924.1

文献标识码: A

文章编号: 1000-2820(2021)05-0019-09

一、问题的提出

近年来人工智能、机器人、物联网等现代高新技术的迅猛发展,给人类社会带来了前所未有的变革与挑战,既改善了人们的生活品质,又给社会增添了新的“问题”。而无人驾驶汽车^①便是新技术所引发问题的最集中之处。当前我国法律实践与相关学术研究中对该技术的称谓还比较混乱,主要用语有“智能汽车”“自动驾驶汽车”“无人驾驶汽车”^②。本文只关注于“无人驾驶汽车”,其属于“智能汽车”“自动驾驶汽车”的最高技术等级,而该类汽车才是完全不同于人为操作汽车。

无人驾驶汽车就是为方便人们出行、大幅度地减少人为因素导致的交通事故而研发的,谷歌

自2009年开始研制无人驾驶汽车,就是在“最大程度上保证乘客的安全”^③;前瞻产业研究院预测未来数年,世界范围内的无人驾驶汽车数量将突破千百万辆,^④“无人驾驶汽车的时代”已经来临。现在无人驾驶汽车技术正日趋进步,借助于高分辨率摄像头、传感器、云计算、个人信息库等,无人驾驶汽车未来可能会识别行人信息(其中人数最确定),但无人驾驶汽车在现实中还是会发生概率极低的意外碰撞,这是任何技术都无法完全克服的难题。因此,无人驾驶汽车在生产过程中就要设计碰撞选择程序,如遇碰撞情形时如何在行人之间、行人与乘客之间选择,这最早引起了世界伦理学界^⑤的广泛争论。“法律的主要功能在于社会控制”,“人们以各种各样的方式用法律来控制、引导

* 收稿日期: 2019-12-05

基金项目: 国家社会科学基金重大项目“马克思主义法学方法论研究”(11&ZD077);重庆市教委人文社科基地重点研究项目“智能司法伦理风险法律控制机制研究”(20JD011)。

作者简介: 杨昆鹏(1990—),男,河南登封人,吉林大学司法文明协同创新中心博士研究生,从事法学理论、司法哲学研究。

①美国交通部采用SAE五类分级,其中第四、五等级已达到无人参与的驾驶水平,而本文探讨的便是此类完全无人驾驶汽车。参见《为何美国交通部选用SAE的自动驾驶分级,而弃NHTSA》,网址: http://www.sohu.com/a/116000253_115873,访问日期2019年8月14日。“SAE Levels of Driving Automation”,网址: <http://cyberlaw.stanford.edu/blog/2013/12/sae-levels-driving-automation>,访问日期2019年8月14日。

②分别以“智能汽车”“自动驾驶汽车”“无人驾驶汽车”为“标题”在“北大法宝司法案例”数据库中检索,有16个案例以“智能汽车”为标题,有1个案例的标题含有“无人汽车”,检索日期2019年8月14日;分别以“智能汽车”“自动驾驶汽车”“无人驾驶汽车”为“题名”在“中国知网期刊数据库”中检索,分别有940篇、1291篇、330篇论文,检索日期2019年8月14日。

③参见江溯《自动驾驶汽车对法律的挑战》,《中国法律评论》,2018年第2期,第180页。

④See Jean-François Bonnefon, Azim Shariff, Iyad Rahwan. The Social Dilemma of Autonomous Vehicles. Science, 2016, Vol. 352, Iss. 6293, pp. 1573-1576; M. Mitchell Waldrop. No Drivers Required. Nature, 2015, Vol. 518 (Feb 5), pp. 20-21.

和计划法庭外的生活”^[1],面对这个现实难题,我国法学界进行了广泛的规制控制研究;这集中于无人驾驶汽车碰撞民事责任分配、对交通肇事罪的挑战、立法方面,^①这些研究主要关注法律规则的制定,有助于我国将来相关的立法或修法。其中在本文所讨论的无人驾驶汽车碰撞的刑法方面,交通肇事罪的主体认定研究尤为突出:一种观点认为无人驾驶汽车承担交通肇事罪责任。这是视“无人驾驶汽车”具有法律人格,要承担刑事责任,该观点突破了人才有“自由意志”的刑法观念,笔者不能认同。一种观点认为无人驾驶汽车制造商(包括系统设计者)不承担交通肇事罪责任。理由是只要他们按相关的标准制造和操作,对具体的肇事情景就没有注意义务,这是基于实证法立场所做出的论证,但并没有进一步论证程序设计的正当性问题,而这是无人驾驶汽车正式进入交通系统的首要条件,交通肇事罪的责任认定是事后利益权衡的结果。而本文就是基于操作性强的行为功利主义理论,结合利益衡量来进行人数比较,在刑法上为无人驾驶汽车碰撞程序的选择提供一种辩护。这其中不涉及车与财物之间的碰撞选择,只考虑人,至于碰撞致人受伤或死亡的后果涉及事后追责,那是另一个问题,本文将不涉及,而碰撞程序的设计只是基于事实模型选择路线,这需要刑法上的正当性。

二、无人驾驶汽车碰撞程序中的 行为功利主义选择

关于无人驾驶汽车如何选择碰撞,伦理学界

主要有两种相互对立的理论,即义务论(deontology)与后果论(consequentialism),功利主义^②便属于后者,而伦理学界实质上所说的是行为功利主义^③理论。因此,我们由“电车难题”的实验困境到无人驾驶汽车碰撞程序选择的现实困境,可以运用行为功利主义理论作出一种选择。

(一)行为功利主义理论

行为功利主义又称为古典功利主义^④(classical utilitarianism),认为“行为的正确或错误仅仅取决于其结果总体上的善与恶,也就是行为对所有人类(或许是所有具有感知之生物)福利的效果”^[3]。行为功利主义思想主要来自于边沁与密尔。

边沁的功利主义就是通过利益有关者幸福的增减来赞成或非难行动(应当与否、对与错),个人评判行为以是否符合功利的法规或命令为基准;边沁的“利益有关者”包括个人和共同体;“幸福”等同于“实惠”“好处”“快乐”“利益”,这都是功利这种客体的一种性质。^⑤在边沁的语境中,“功利主义为直觉、公众伦理准则或个人道德提供了唯一可能的终极检验”^[4]。

密尔在《功利主义》(Utilitarianism)一书中详细阐明了功利主义的含义:第一,行为的对错与幸福的增进或不幸的倾向成正比。而幸福就是快乐和免除痛苦,不幸就是痛苦和丧失快乐。第二,快乐有数量和质量之别。快乐的质量也有高等和低等之别,快乐质量的评定与衡量来自于人的经验偏好、自我意识与自我观察、比较的方法。第三,功利主义的行为标准,不是行为人为本人的最大幸福,而是全体相关人员的最大幸福,行为者要作为“公正

①参见郑志峰《自动驾驶汽车的交通事故侵权责任》,《法学》,2018年第4期,第16-29页;冯珏《自动驾驶汽车致损的民事侵权责任》,《中国法学》,2018年第6期,第109-132页;杨立新《自动驾驶机动车交通事故责任规则的规则设计》,《福建师范大学学报》(哲学社会科学版),2019年第3期,第75-88页;龙敏《自动驾驶交通肇事刑事责任的认定与分配》,《华东政法大学学报》,2018年第6期,第77-82页;周铭川《论自动驾驶汽车交通肇事的刑事责任》,《上海交通大学学报》(哲学社会科学版),2019年第1期,第36-43页;崔俊杰《自动驾驶汽车准入制度:正当性、要求及策略》,《行政法学研究》,2019年第2期,第90-103页;李烁《自动驾驶汽车立法问题研究》,《行政法学研究》,2019年第2期,第104-113页。

②当代功利主义内部又分为“行为功利主义(act utilitarianism)”“规则功利主义(rule utilitarianism)”“制度功利主义(institution utilitarianism)”。参见晋运锋《当代西方功利主义研究述评》,《哲学动态》,2010年第10期,第57-61页;姚大志《当代功利主义哲学》,《世界哲学》,2012年第2期,第50-61页;[澳]J.J.C.斯玛特、[英]伯纳德·威廉斯《功利主义:赞成与反对》,劳东燕、刘涛译,北京:北京大学出版社,2018年,第15-21页。

③参见和鸿鹏《无人驾驶汽车的伦理困境、成因及对策分析》,《自然辩证法研究》,2017年第11期,第60页;谢惠媛《民用无人驾驶技术的伦理反思——以无人驾驶汽车为例》,《自然辩证法研究》,2017年第9期,第42页;孙保学《自动驾驶汽车的伦理困境——危机时刻到底该救谁》,《道德与文明》,2018年第4期,第36页。

④古典功利主义可追溯到伊壁鸠鲁学派,“功利”一词最早出现在休谟《道德原则研究》一书中。参见[英]弗雷德里克·罗森《古典功利主义:从休谟到密尔》,曹海军译,南京:译林出版社,2018年,第15-67页。

⑤参见[英]边沁《道德与立法原理导论》,时殷弘译,北京:商务印书馆,2000年,第58-59页。

无私的旁观者”来衡量自己与他人的幸福。^①在密尔的语境中,需要关注两点:首先,幸福不是一个抽象概念,而是一个整体的概念。幸福是唯一值得欲求的目的,其他欲求是达到幸福的手段,但这些手段(金钱、权力、名声、美德等)是为达到目的而值得欲求的,其也就成为了目的的一部分。其次,“最大幸福原理”是全体相关人员幸福的质量和数量的最大化。它不是关注一个人的幸福,而是多数人或所有人的幸福,“只考虑幸福的大小”^②。若依据人类的经验来衡量幸福的质量和数量,功利原则就有外在约束力(动机)和内在约束力(道德感情)。

行为功利主义由于有强大的预测性和“感受性”^③,从19世纪便开始主导西方伦理学,但最近几十年主要受到的批评之一来自于义务论:第一,人是目的而不是手段,行为功利主义是将人视为手段。“你的行动,要把你自己人身中的人性,和其他人身中的人性,在任何时候都同样看作是目的,永远不能只看作是手段。”^④其实在康德看来,人可为主观目的和客观目的,主观目的就是手段,人作为我们行为的结果而存在,客观目的是人的自在目的,即人作为一个人本身而存在。康德并未否认人在实际生活中会成为“手段”,他反对的是不应该仅仅将人视为手段,重视“人是目的”的理想状态,而手段与目的实则总是交织在一起。^⑤但“人是目的”的实践命令并未否认在必须牺牲少数人与多数人之间不作出选择。第二,行为功利主义会侵害个人自由权利。批评者认为行为功利主义视“自由和权利本身没有独立的价值”^⑥,功利最大化会侵害人的权利和自由。罗尔斯认为“功利主义并不在人与人之间做出严格的区分”^⑦,自由或权利比社会福利重要,行为功利主义就是追求“最大净余额”(即福利 welfare)^⑧,而福利是指生活上的利益^⑨。功利主义者也承认个人的权利或自由,只不过认为它们是达到福利或幸福的一种手段,有时可以牺牲个人

的权利或自由来达到福利的最大值。权利本身就是一种利益,而福利就是利益,福利的含义其实包含了权利与自由,行为功利主义与个人自由权利也不是绝对对立的。所以义务论并未完全战胜行为功利主义,行为功利主义在当代还有其存在价值,而这便体现在无人驾驶汽车碰撞程序的设计中。

(二)从“电车难题”到无人驾驶汽车碰撞程序选择

“电车难题”^⑩(Trolley Problem)由菲利普·富特(Philippa Foot)1967年提出,而弗朗西丝·默纳·卡姆(F.M.KAMM)构建了18类“电车难题”案例^⑪,但通过个案分析使该问题更加复杂,诚如谢利·卡根(Shelly Kagan)所言,“我们或许将会被引向这样一项原则,即它虽然只符合甚至更少的具体案例直觉,但它却具备一种极其重要且更加令人信服的根本理由”^⑫。而行为功利主义就是一种选择的理由,如上文所述,行为功利主义就是利益相关者的利益最大化,在“电车难题”中若选择撞向一人,这种人数的比较就是奉行行为功利主义。^⑬

但“电车难题”的伦理困境是否适用于无人驾驶汽车碰撞的现实困境,存在两种对立的观点:一种是直接将“电车难题”设定为无人驾驶汽车碰撞的一种伦理困境。无人驾驶汽车在行驶中需要其程序自主决定,对两难情形会陷入伦理困境,其中一个困境就是“电车难题”。^⑭另一种认为无人驾驶汽车碰撞所面对的伦理困境不同于“电车难题”。有研究者反对将自动驾驶汽车事故算法类比为“电车难题”,并且提出两者在三个方面存在不同:第一,全部的决定情形及其特征。前者是提前作出选择以应对瞬间情景,而后者是瞬间决定;前者的决策是由多个利益相关者协商一致达成,而后者是一个个体作出的道德选择;前者考虑的因素不受限制,而后者只考虑非常有限的几个因素。第二,道德和法律责任的作用。前者在编程

①参见[英]约翰·穆勒《功利主义》,徐大建译,北京:商务印书馆,2014年,第7-31页。

②参见甘绍平《功利主义的当代价值》,《中国社会科学院研究生院学报》,2010年第3期,第43-44页。

③参见俞吾金《如何理解康德关于“人是目的”的观念》,《哲学动态》,2011年第5期,第26-27页。

④参见[美]约翰·罗尔斯《正义论》,何怀宏、何包钢、廖申白译,北京:中国社会科学出版社,1988年,第22页。

⑤参见《汉典:福利》,网址:<https://www.zdic.net/hans/福利>,访问日期2020年1月15日。

⑥See Philippa Foot. Virtues and Vices and Other Essays in Moral Philosophy. Oxford University Press,2002, pp.19-31.

⑦参见[美]弗朗西丝·默纳·卡姆《电车难题之谜》,埃里克·拉科夫斯基编,常云译,北京:北京大学出版社,2018年,第341-346页。此为案例示意图,具体内容参考正文。

⑧在Cathcart设定的全民意法庭上,辩方就是用行为功利主义理论为扳动电车轨道而拯救五人的被告辩护。参见[美]托马斯·卡思卡特《电车难题:该不该把胖子推下桥?》,朱沅之译,北京:北京大学出版社,2014年,第28-53页。

⑨参见和鸿鹏《无人驾驶汽车的伦理困境、成因及对策分析》,《自然辩证法研究》,2017年第11期,第58-59页。

时必须要考虑道德和法律责任,而后者完全忽视了该问题。第三,决定者的认知情况。前者需要对不确定和风险进行推理,而后者只是对已知事实和特定结果进行推理。^①并且“电车难题”是旁观者的决策,而无人驾驶汽车的碰撞选择还有自我(乘客)利益与他人(行人)利益的现实冲突。^②虽然“电车难题”不能直接适用于无人驾驶汽车碰撞选择,但“电车难题”思想实验为无人驾驶汽车碰撞的现实解决提供了理论来源,“电车难题”适用于无人驾驶汽车碰撞的理由有:第一,“电车难题”遭遇到的选择,无人驾驶汽车也会在现实中遇到。第二,两者都要进行计算与衡量。“电车难题”是在五人与一人之间进行比较选择,而无人驾驶汽车要处理更为多元、复杂的信息,但最终也要利用计算机程序中的认知模型进行计算,作出确定、有效的指令。^③第三,“旁人”向“参与者”的转化。“电车难题”只是从旁人的角度作出决策,轨道上的人并没有表达自己的意愿,对此可用行为功利主义理论作出选择,而无人驾驶汽车碰撞程序的设计是多方利益主体协商一致的结果,在此“电车难题”中的“旁人”可能成为无人驾驶汽车碰撞程序设计的“参与者”。若“旁人”都成为“参与者”,行为功利主义可以成为无人驾驶汽车碰撞程序设计的原则,为碰撞提供一个“内在的理由”^④。因此,“电车难题”可作为无人驾驶汽车碰撞选择的一种“模型”,“电车难题”的行为功利主义选择也可适用于无人驾驶汽车碰撞程序设计。

此外,无人驾驶汽车碰撞程序本身的特点也决定了行为功利主义适用的可能性。无人驾驶汽车碰撞程序需有如下特点:第一,简单化。虽然无人驾驶汽车行驶中要收集庞大而复杂的信息,但最终都要通过计算机程序的算法作出决策,所以这个碰撞程序必须简单,快速对无人驾驶汽车发

出指令。第二,普遍化。无人驾驶汽车的批量生产,其在行驶中遇到的各种具体情形,都要求无人驾驶汽车碰撞程序能普遍适用,决非每遇到一个情形就作出一种新的选择。第三,民主化。由于无人驾驶汽车涉及到多方利益主体,所以无人驾驶汽车碰撞程序的设计需要民主决策而定。而行为功利主义就是一种简单的计算公式(对人数增减进行计算),也具有普遍适用性(利益覆盖范围广),并且通过每个参与者的民主商议可以作出行为功利主义选择。因此,行为功利主义有充足理由适用于无人驾驶汽车碰撞程序的选择。行为功利主义在无人驾驶汽车碰撞选择中的适用,需要“无人驾驶汽车的碰撞使全体相关人员的幸福(利益)最大化”,其中“全体相关人员”除了被撞者、乘客两类切实利益相关者之外,还包括无人驾驶汽车生产者(包含程序设计,除质量问题之外不再是相关者)、其他普通公民(利益在于参与程序设计),而“幸福最大化”就要整体考虑全体社会人员的利益,具体到碰撞就是人数的比较,因为这是社会整体利益分配和风险分担的结果。而在刑法层面上,行为功利主义适用于无人驾驶汽车碰撞程序设计还需要进一步论证其正当性。

三、无人驾驶汽车碰撞^⑤中的刑法违法类型

对于刑法违法的评判标准,当前我国刑法学界存在两种对立的观点,即结果无价值论与行为无价值论。^⑥而笔者认可结果无价值论的立场,评判违法的标准就是看“是否侵害法益”。本文的“违法”指不法(包括违法阻却事由)中的违法,就是“什么样的行为在刑法上是不正当的、是被刑法禁止的”^{⑦⑧}。违法是由刑法目的或任务来理解的。刑法的目的就是保护法益,刑法禁止的对象就是侵害或危险法益。^⑨我国现行刑法第2条明确规定其目的就是保

① See Sven Nyholm, Jilles Smids. The Ethics of Accident-Algorithms for Self-Driving Cars: an Applied Trolley Problem?. Ethical Theory and Moral Practice, 2016, 19(5), pp. 1275-1289.

② 参见孙保学《自动驾驶汽车的伦理困境——危机时刻到底该救谁》,《道德与文明》,2018年第4期,第35页。

③ 参见蒋舸《作为算法的法律》,《清华法学》,2019年第1期,第65-66页。

④ 哈特提出了规则的内在面向(internal aspect)和外在面向(external aspect),以此来证成义务规则,把规则或规范视为一个行动的理由。See H.L.A. Hart. The Concept of Law. Oxford University Press, 2012, pp. 88-91. 参见陈景辉《什么是“内在观点”》,《法制与社会发展》,2007年第5期,第3-13页。

⑤ 在此无人驾驶汽车碰撞是在其程序指令下的碰撞,下文“刑法违法类型和违法阻却”部分皆是在“无人驾驶汽车碰撞程序实施”的语义中进行论述,都是为了论证无人驾驶汽车碰撞程序的刑法正当性问题。

⑥ 参见张明楷《结果无价值论的法益观——与周光权教授商榷》,《中外法学》,2012年第1期,第24-43页;周光权《行为无价值论的法益观》,《中外法学》,2011年第5期,第944-957页。

⑦ 参见张明楷《行为无价值论的疑问——兼与周光权教授商榷》,《中国社会科学》,2009年第1期,第102页。

护法益,惩罚犯罪、与犯罪行为斗争、预防犯罪都是保护法益。^①关于“法益”(Rechtsgüter)概念,李斯特(Liszt)认为法益就是“由法律所保护的利益”^[1],罗克信(Roxin)认为“对于安全、自由的、保障所有个人人权和公民权的社会生活所必要的,或者对于建立在此目标上的国家制度的运转所必要的现实存在或者目的设定就是法益”^[12]。因此,法益实质上就是法律认可的实现个人自由发展的生活利益。而在无人驾驶汽车行驶中,碰撞行人或乘客的自我牺牲在刑法层面上都是对法益(利益)的侵害,都构成刑法违法,但这种违法行为是否存在阻却事由是另一个问题(下文详述)。通过上文分析无人驾驶汽车碰撞中行为功利主义的人数比较,同时也基于我们的驾车经验,笔者认为无人驾驶汽车碰撞中的刑法违法存在两种类型。

(一)在行人之间的选择

无人驾驶汽车在行人之间的选择属于碰撞最常见的情景类型之一,这类似于“电车难题”情形,该类选择具体可描述为:道路前方出现一群行人穿越,无人驾驶汽车若前行或向左行都会撞到多个行人,而路右侧有一个行人,如果汽车向右行驶则刚好会撞到这个行人,此时若选择撞向一个行人,结果会造成其死亡。无人驾驶汽车的出现是将“电车难题”思想实验搬进了现实,将轨道的选择变成程序设计路线的选择。而“电车难题”遭遇到的选择困境,无人驾驶汽车同样会遇到,但其必须作出一种普遍适用的正当选择。因此,无人驾驶汽车碰撞在行人之间的选择就是无人驾驶汽车在道路行驶中,前方紧急出现行人,而改变行驶方向将会撞到其他行人的情形,该处可选择的路线至少有两个以上,其中行人数有一个对多个(两个以上)等多种两方人数对比有差别或相等的情形。在此那个(些)被撞的“无辜”行人只是恰好站在那个位置,不管其是否违反交通规则,他(们)正常情况下都是没有放弃其生命健康权的意愿,并且他(们)的生命健康权受刑法保护,所以撞击该行人就是侵害他(们)的法益,在此无人驾驶汽车碰撞行人构成刑法违法。

(二)在乘客与行人之间的选择

无人驾驶汽车在乘客与行人之间的选择是碰撞的又一常见情景类型,这类似于“隧道难题”

(Tunnel Problem)情形,该类情形具体可描述为:“某人坐在一辆无人驾驶汽车上,汽车行驶在一条单车道山路上,正快速驶进一条狭窄的隧道。但在进入隧道之前,一名儿童误入道路,绊倒在车道中央,处于隧道的入口。但汽车已经无法及时刹车以避免碰撞,现在只有两种选择,即撞死这个孩子,或者突然转向,撞到隧道两边的墙上,乘客因此丧命。”^[13]该情形可以再转化成一群行人出现在隧道入口,若选择撞向隧道的两面墙可挽救多数行人。因此,无人驾驶汽车碰撞在乘客与行人之间的选择就是无人驾驶汽车在道路行驶中,有行人挡住行车路线,在此要么继续按原路线行驶而撞死行人,要么转向路边而造成乘客死亡,其中行人与乘客之间的人数对比存在多种两方人数有差别或相等的情形。刑法中的法益理论是为实现个人的自由发展,防止国家不必要的干预而产生的。以赛亚·伯林区分了两种自由,消极自由(被允许不受他人干预)和积极自由(决定某人做这个或成为这样),^②因此刑法中“侵害法益”存在两个向度:一个是国家对个人自由的干预而侵害法益,一个是因个人自戕(自我牺牲)而侵害法益。自我牺牲虽然使法益受损,但自我牺牲者会受到社会舆论的称赞,这种牺牲属于一种道德层面的约束,而刑法层面不会归罪追责自我牺牲者。在此无人驾驶汽车碰撞行人或碰撞墙壁而致乘客受伤或死亡都是刑法上的侵害法益,无人驾驶汽车在乘客与行人之间的碰撞选择构成刑法违法。

(三)两种违法类型中的“人数”比较

无人驾驶汽车在行人与行人、行人与乘客之间通过人数对比来选择碰撞是最直观快捷的。2018年Nature刊登了一篇《道德机器Moral Machine测试》,收集了233个国家和地区数百万人在面对自动驾驶汽车碰撞问题时的4000万个决定,进而总结了全球的道德偏好、个人偏好差异、跨文化的伦理差异、该三类差异与现代制度和深层文化特征的关系。^③这说明了无人驾驶汽车碰撞选择的复杂性,但计算机程序最终量化为人数进行选择是最具有可操作性的,这也需要我们进一步来论证其刑法正当性。依经验来看,无人驾驶汽车碰撞的现实违法情形都可归到这两种类型中来,但关于“人”存在两种不同的观点:一种是应当

①参见张明楷《刑法学》(第五版),北京:法律出版社,2016年,第22页。

②参见[英]以赛亚·伯林《自由论》,胡传胜译,南京:译林出版社,2003年,第186-246页。

③See Edmond Award, Sohan Dsouza, Richard Kim et al. The Moral Machine experiment. Nature, 2018, Vol. 563 (Nov1), pp. 59-63.

考虑人的差别。无人驾驶汽车技术将来可能会识别行人的身份信息,“基于全信息的道德决策就只能完全交给汽车的人工智能系统自主决定”^[14],这样就把复杂多样的现实情形交给无人驾驶汽车的计算机系统进行选择,但这又必然涉及到更具颠覆性的“机器人格化”问题。因为违反当前伦理观念,笔者反对这种基于“机器人格化”而在无人驾驶汽车碰撞中考虑人的差别。另一种是应当摒弃人的差别,将碰撞中的人抽象化。“事物越具体,可比较的不等式就越多。”^[15]如面对不同的人,人们的情感强度都是不同的,人为操作汽车的驾驶人在面对自己的亲人或朋友时,情感会促使其优先保护他们,而无人驾驶汽车需要预先设定碰撞程序,不存在现实的瞬间决断。因此,笔者赞同将无人驾驶汽车碰撞中的人抽象化,因为越具体反而越复杂,而只基于人数比较(包括人数相等)。所以,这两种违法类型构成了无人驾驶汽车碰撞选择在刑法层面上的“元模型”,一切的现实碰撞选择皆来自于此,而来源于行为功利主义的“人数比较”就是碰撞选择程序的“源代码”。

四、无人驾驶汽车碰撞的刑法违法阻却

无人驾驶汽车碰撞构成刑法违法,若对其刑事归罪还需要不存在违法阻却事由,而无人驾驶汽车碰撞程序选择又必须存在违法阻却,否则汽车制造商就构成故意犯罪。碰撞程序的设计是由国家、生产商、消费者等多方利益主体商定,并且无人驾驶汽车技术应该最大程度地保护乘客和行人,因此无人驾驶汽车碰撞程序应该存在刑法违法阻却事由。接下来笔者基于行为功利主义理论,借助利益衡量和民主决策的“紧急避险”来论证无人驾驶汽车碰撞的两种刑法违法类型存在违法阻却,但首先需要阐明行为功利主义理论在此论证中的优势。

第一,行为功利主义具有规范性。作为一个精神正常的人,我们做出的每一个行为都至少需要一个正当的理由,否则就可能因违背法律或习惯而给我们带来不利后果。拉兹(Raz)认为规则可以成为行动的正当理由,但并非所有的规则都是行动的理由,只有“构成性规则”(constitutive

rules,即人造规则)才能成为行动的理由。^①同样无人驾驶汽车碰撞程序也需要一个正当理由来作出选择,而这种计算机程序的选择就是一种人造规则的执行。行为功利主义“是一种规范性的理论”^{[16]97},“它认为正当行为是指其实际后果(与预期后果相对)对内在善的事态有最大贡献行为”^{[16]98}。因此行为功利主义具有规范性,可以用于阻却碰撞行为的刑法违法。

第二,行为功利主义具有刑法适用性。有研究者利用 Amazon Mechanical Turk 进行自动驾驶汽车碰撞测试,参与者都赞成功利型自动驾驶汽车,为了更大的利益可以牺牲乘客,但他们自己更愿意乘坐不惜一切代价保护乘客的自动驾驶汽车。^②这就产生了自相矛盾的结果,从“旁观者”角度赞成行为功利主义原则,从“参与者”角度则奉行自保原则,这实则制造了“两个身份”;只要乘坐无人驾驶汽车或认可(明示或默示)无人驾驶汽车在道路上行驶,就只能遵循一种程序选择的原则。根据刑法两阶层说,违法阻却属于不法阶层,而违法阻却实质上就是利益衡量,“结果无价值论使刑法与伦理相区别”^{[17]47},“行为的反社会伦理性,并不直接成为刑罚处罚的根据”^{[17]48}。刑法上的行为功利主义摒弃了以违背社会伦理作为归罪处罚的依据,这样行为功利主义在伦理学中的“幸福”就转化成刑法中的“利益”,而刑法上的行为功利主义又保留了其伦理学中的“计算”形式。因此,在无人驾驶汽车碰撞中,刑法上的行为功利主义可以消除其伦理性,使“旁观者”与“参与者”采用统一的行为功利主义原则。

第三,行为功利主义体现平等性。在无人驾驶汽车碰撞中,刑法不会关注被撞死的人是谁。2017年德国自动驾驶汽车全国伦理委员会规定:“在不可避免的事故中,严禁基于个人特征进行任何区分。禁止相互抵消受害者。减少受伤者人数的一般程序是合理的。”^[18]这样是为了防止机器程序根据个人特征来选择碰撞的目标。我国宪法规定“公民在法律面前一律平等”同样体现了个人的平等性。无人驾驶汽车碰撞程序设计应该不区分被撞的人,只计算人数,这也是当前技术提供的最具确定性和实用性的方法。因此,无人驾驶汽车碰撞

①See Joseph Raz. Between Authority and Interpretation: On the Theory of Law and Practical Reason. Oxford University Press, 2009, pp. 203-219.

②See Jean-François Bonnefon, Azim Shariff, Iyad Rahwan. The Social Dilemma of Autonomous Vehicles. Science, 2016, Vol. 352, Iss. 6293, pp. 1573-1576.

程序采取行为功利主义是个人平等性的体现。

(一)行人之间选择的违法阻却

根据我国刑法规定,在行人之间选择,无人驾驶汽车碰撞程序的违法阻却事由只能适用“紧急避险”^①。紧急避险的本质是不具有社会危害性,阻却违法是法益衡量的结果。^②紧急避险来自于社会连带说,“以立场的互换可能性为前提,负有相互地成为他人的‘牺牲品’”^[19]。通常紧急避险是针对两种正当法益的整体比较(如伤害与死亡、人身与财物),但若涉及到生命法益的比较就是“超法规的紧急避险,也仅限于被牺牲者特定化的场合”^[20]。因此,当前刑法理论中的紧急避险是一种基于人类自保的“弱紧急避险”,不足以为侵害“生命法益”提供正当性。通过行为功利主义的人数比较和整体利益衡量来补强“紧急避险”,无人驾驶汽车碰撞程序可适用“补强的紧急避险”理论来阻却在行人之间选择的违法。一方面,将来无人驾驶汽车正式进入交通道路是一种在行为功利主义考量下对社会整体利益和整体风险合理分配的结果,因为每个人都可能成为这个“被牺牲者”,这就是利益与风险的分配;另一方面无人驾驶汽车碰撞程序是数字化计算,行为功利主义的人数比较既符合计算程序,又可化解“法益衡量”的抽象化,以人数来计算利益的大小。所以,“补强的紧急避险”就是基于行为功利主义的整体利益衡量,适用于侵害“生命法益”的特殊场合,运用人数所代表的利益大小来采取紧急避险行为。这样无人驾驶汽车在行人之间选择碰撞的违法阻却就可以适用“补强的紧急避险”理论。

在此选择中又具体存在两种情形:第一,两方行人人数有明确的差异(用五个行人与一个行人做对比)。无人驾驶汽车碰撞程序进行同等利益衡量,计算得出五个人的利益大于一个人的,因此选择碰撞一个人。因为无人驾驶汽车在撞向五个人还是一个人之间必须有一个选择,这是两个作为义务之间的冲突,侵害一个人的法益应当优于侵害五个人的法益。如有学者所言,紧急避险对此可以阻却违法性,只不过汽车的生产者需要根据预见的情形在编程时决定牺牲哪个(些)人。^③

无人驾驶汽车的预见性要远远强于人为操作下的汽车驾驶员,所以只要无人驾驶汽车能判定属于碰撞的紧急情形,就可以通过紧急避险来阻却违法。在该情形中,对于选择碰撞一个人而避开五个人就是适用“补强的紧急避险”阻却违法。第二,两方行人人数相等。在此视受侵害的法益相等,无人驾驶汽车就要“随机选择碰撞”^④。因为在人为操作汽车情形中,驾驶人都是瞬间作出的随机选择,都有可能撞向其中的一方,这也是两个同等的作为义务之间的冲突,即使不作为而继续按原路线行驶也是一种“作为”,并且没有更高等级的义务要求驾驶员必须选择碰撞某一方。因此,在无人驾驶汽车碰撞中,人数相同也就是利益衡量相等,程序可以设定成随机选择,在此适用“补强的紧急避险”也可以阻却违法。

(二)乘客与行人之间选择的违法阻却

对于无人驾驶汽车在乘客与行人之间选择的情形,首先需要在刑法上消除一个伦理困境,那就是乘客与行人孰为优先保护?笔者反对优先保护乘客,认为应同等对待乘客与行人,如前文论述“人”的抽象化问题,在此还可具体解释为:第一,乘客与行人的换位。作为乘客,乘坐无人驾驶汽车,优先考虑的就是其安全性(适用于产品责任),但遇到碰撞情形,就是一种外部施加的“场域”迫使无人驾驶汽车作出选择。如果任何时候都优先保护乘客,最后的结果可能危害到换位成“行人”的乘客,因为任何人都可能成为无人驾驶汽车碰撞中的“行人”。第二,优先保护乘客违反平等原则。“每个人对与其他人所拥有的最广泛的基本自由体系相容的类似自由体系都应有一种平等的权利”^{[8]60-61},这是“确定与保障公民的平等自由”^{[8]61}。第三,优先保护乘客的原则可能会阻挡无人驾驶汽车行业的发展。这存在消费者安全与无人驾驶汽车行业效益两种利益的冲突,无人驾驶汽车相较于传统汽车,大幅度地减少了交通事故的发生,而碰撞只是一个发生概率极低的意外,无人驾驶汽车行业带来的整体利益远远大于碰撞中优先保护乘客的利益。因此过度强调优先保护乘客可能

①《中华人民共和国刑法》第21条第1款规定:“为了使国家、公共利益、本人或者他人的人身、财产和其他权利免受正在发生的危险,不得已采取的紧急避险行为,造成损害的,不负刑事责任。”

②参见黎宏《紧急避险法律性质研究》,《清华法学》,2007年第1期,第42页;张明楷《违法阻却事由与犯罪构成体系》,《法学家》,2010年第1期,第35页。

③参见[德]托马斯·魏根特《自动驾驶汽车的紧急避险权》,樊文译,《中国应用法学》,2019年第4期,第180页。

④参见[德]托马斯·魏根特《自动驾驶汽车的紧急避险权》,樊文译,《中国应用法学》,2019年第4期,第181-182页。

会阻碍无人驾驶汽车行业的发展,反而损害作为“消费者”的乘客利益。这样乘客与行人的法益同等重要,无人驾驶汽车在乘客与行人之间选择碰撞的违法阻却就只基于人数比较的利益衡量。

在此选择中有三种具体情形(人数为五人与一人):第一,在五个行人与一个乘客之间的选择。基于行为功利主义的利益衡量可知一个乘客受损的法益小于侵害五个行人的法益。并且即使以行人违反交通规则作为理由,我们也很难认为乘坐一人的无人驾驶汽车直接撞击五个行人在刑法上正当。这里就涉及到乘客的“自我牺牲”,还需进一步的阐释来适用“补强的紧急避险”。支持紧急避险“社会团结(社会连带)义务说”的学者否定自我牺牲,只是赞同侵害较小的法益来保护较大的法益。^①如果紧急避险制度允许本人为了保存自己的利益而侵害无辜他人的利益,那么为何本人不能为了他人的利益而牺牲自己的利益呢?这就存在一种矛盾,而刑法学界更多关注于“牺牲第三人利益而保护自我利益”的正当性^②,因此必须从刑法上论证紧急避险中“自我牺牲”的可能性。而无人驾驶汽车的生产者“在只有自毁才可避免与车外的人相撞的情况下,必须对车外要撞上的人数与陷入生命危险的乘客人数进行权衡”^[21]。这就需要在无人驾驶汽车同样保护乘客与行人利益的前提下,国家、消费者和汽车制造商等相关利益主体基于行为功利主义通过民主商议来达成“被牺牲者同意”。在此如果达不成“自我牺牲”,那无人驾驶汽车就要选择碰撞五个行人,如此就会发生“只要无人驾驶汽车不存在故障或违反交通规则,就可以直接撞击多于乘客人数的行人”的奇怪现象,并且也冲击我们“优先救多数人”的道德情感。因此,乘客与行人的法益同等,每个人都可能是“行人”或“乘客”,这还是两种作为义务的选择,只有达成“被牺牲者同意”而补强“紧急避险”才可统一碰撞选择原则。所以适用“补强的紧急避险”可以阻却此情形中侵害乘客利益的违法。第二,在一个行人与五个乘客之间的选择。基于行为功利主义理论,保护五个人的利益大于撞向一个人的利益,在此适用“补强的紧急避险”选择撞向一个行人阻却违法。第三,在一个行人

与一个乘客之间的选择。因为人数相同,利益衡量相等,可以由无人驾驶汽车程序随机选择。这里涉及到乘客的“自我牺牲”同上文所述。适用“补强的紧急避险”可以阻却无人驾驶汽车选择碰撞一个行人或致一个乘客牺牲的刑法违法。

五、结 语

对于无人驾驶汽车碰撞程序问题,伦理学界借助于思想实验、电脑模型、平台测试等方式已经争论了许久,而无人驾驶汽车终究要在现实道路上行驶,碰撞问题是无法完全避免的,这必然会侵害法益构成刑法违法。但无人驾驶汽车正式进入交通系统又必然使碰撞程序的设计存在违法阻却事由。基于行为功利主义理论,以人数比较进行利益衡量,进而通过“补强的紧急避险”可以阻却无人驾驶汽车碰撞程序实施选择的刑法违法。对于行人之间的选择,碰撞较少人数,当人数相等时,由程序随机选择碰撞;对于行人与乘客之间的选择,撞击人数较少的行人或乘客,若行人和乘客人数相同,就由程序随机选择碰撞。如此,无人驾驶汽车碰撞程序在刑法层面便正当化,这为无人驾驶汽车的现实化提供了有力的法律支撑。

参考文献:

- [1] 何佳,戎辉,王文扬,等.百度谷歌无人驾驶汽车发展综述[J].汽车电器,2017(12):21.
- [2] HART H L A. The Concept of Law[M]. Oxford: Oxford University Press,2012:40.
- [3] J.J.C.斯玛特,伯纳德·威廉斯.功利主义:赞成与反对[M].劳东燕,刘涛,译.北京:北京大学出版社,2018:5.
- [4] 蒂姆·莫尔根.理解功利主义[M].谭志福,译.唐冉,校.济南:山东人民出版社,2012:61.
- [5] 姚大志.当代功利主义哲学[J].世界哲学,2012(2):52.
- [6] 伊曼努尔·康德.道德形而上学原理[M].苗力田,译.上海:上海人民出版社,2012:37.
- [7] 晋运锋.权利莫基于功利——当代功利主义对权利之基础的思考[J].法制与社会发展,2012(6):131.
- [8] 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.
- [9] 弗朗西丝·默纳·卡姆.电车难题之谜[M].埃里克·拉

①参见王钢《对生命的紧急避险新论——生命数量权衡之否定》,《政治与法律》,2016年第10期,第100-108页;方军《紧急避险的体系再定位研究》,《现代法学》,2018年第2期,第137-138页。

②参见陈璇《生命冲突、紧急避险与责任阻却》,《法学研究》,2016年第5期,第135-139页;王钢《紧急避险中无辜第三人的容忍义务及其限度——兼论紧急避险的正当化根据》,《中外法学》,2011年第3期,第609-625页。

- 科夫斯基,编.常云云,译.北京:北京大学出版社,2018:214-215.
- [10] 张明楷.行为功利主义违法观[J].中国法学,2011(5):113.
- [11] 李斯特.德国刑法教科书[M].徐久生,译.何秉松,校.北京:法律出版社,2006:6.
- [12] 克劳斯·罗克信.刑法的任务不是法益保护吗?[A]//樊文,译.陈兴良.刑事法评论:第19卷.北京:北京大学出版社,2007:152.
- [13] MILLAR J. An Ethics Evaluation Tool for Automating Ethical Decision-Making in Robots and Self-Driving Cars[J].Applied Artificial Intelligence,2016,30(8):787.
- [14] 白慧仁.自动驾驶汽车的“道德算法”困境[J].科学学研究,2019(1):21.
- [15] 赵汀阳.有轨电车的道德分叉[J].哲学研究,2015(5):99.
- [16] 尤金·贝尔斯.行为功利主义:是对行为正当性特征的解释还是决策程序?[A]//陈江进,译.徐向东.后果主义与义务论.杭州:浙江大学出版社,2011.
- [17] 张明楷.行为无价值论与结果无价值论[M].北京:北京大学出版社,2012.
- [18] LUETGE C. The German Ethics Code for Automated and Connected Driving[J].Philos.Technol,2017(30):552.
- [19] 储陈城.自动汽车程序设计中解决“电车难题”的刑法正当性[J].环球法律评论,2018(3):88.
- [20] 张明楷.刑法学[M].第5版.北京:法律出版社,2016:222.
- [21] 托马斯·魏根特.自动驾驶汽车的紧急避险权[J].樊文,译.中国应用法学,2019(4):183.

[责任编辑:王正良]

Legitimacy of the Criminal Law on the Procedure of Driverless Car Crashes ——Based on the Theory of Act Utilitarianism

YANG Kun-peng

(Collaborative Innovation Center of Judicial Civilization, Jilin University, Changchun, Jilin,130012)

Abstract: The "era of driverless cars" has arrived, and the ethical dilemma of the "electric vehicle problem" has also become a realistic option for the procedure of driverless car crashes. The driverless cars colliding with pedestrians or sacrificing passengers inevitably infringe on the legal interests of the criminal law, while the driverless cars normally included in the traffic system make the procedure of the collision the cause of illegality deterrence. Based on the normativeness, applicability of criminal law, and equality of the theory of act utilitarianism, this paper, with the comparison of the number of persons, converted the abstract "legal interests" into a calculable measure of benefits, and, with the relevant parties consulting to reach the "agreement of the victims", constructed the theory of "emergency risk avoidance of reinforcement" in order to provide the cause of illegality deterrence to the choice between pedestrians, pedestrians and passengers, in the procedure of driverless car crashes.

Keywords: Driverless Cars, Act Utilitarianism, Legal Interests, Emergency Risk Avoidance of Reinforcement, Illegality Deterrence