

碰撞困境中自动驾驶的刑事正当化理论剖析

蔡孟兼

内容提要:对于自动驾驶碰撞困境的应对,法哲学上已有电车难题的讨论和因此引发的功利主义与自由主义之争,在此基础上,应将其价值选择的正当性延伸至碰撞法则的具体设计和相应法律适用上。通过紧急避险与社会连带性的观点可论证自动驾驶在碰撞困境中的危险承担义务。在对生命的攻击型紧急避险中,在自动驾驶负有危险承担义务的要求下,造成无辜的交通参与人的生命侵害不成立紧急避险,只能在侵害使用人或乘客的最低限度内采取必要的避让措施,最大限度地维护使用人和其他交通参与人的安全;在对生命的防卫型紧急避险中,基于引起事故危险之人有危险承担义务,自动驾驶对其造成生命侵害可成立紧急避险;自动驾驶与使用人或乘客基于信赖关系,可期待使用人或乘客履行承担危险的义务,故自动驾驶在碰撞困境中实施不具有正当性的避险行为,不应以欠缺期待可能性作为阻却责任的理由。

关键词:自动驾驶 碰撞困境 容许危险 紧急避险

蔡孟兼,广东财经大学法治与经济发展研究所副研究员。

一 问题的提出

2019年6月,我国国家新一代人工智能治理专业委员会于发布《新一代人工智能治理原则——发展负责任的人工智能》(下称“《治理原则》”),旨在更好地协调人工智能发展与治理的关系,确保人工智能安全可控可靠,推动经济、社会及生态可持续发展,共建人类命运共同体。2021年9月该委员会又发布《新一代人工智能伦理规范》(下称“《伦理规范》”),旨在将伦理道德融入人工智能全生命周期,为从事人工智能相关活动的自然人、法人和其他相关机构等提供伦理指引,增强全社会的人工智能伦理意识与行为自觉,积极引导负责任的人工智能研发与应用活动,促进人工智能健康发展。^[1]

[1] 《新一代人工智能伦理规范》提出了增进人类福祉、促进公平公正、保护隐私安全、确保可控可信、强化责任担当、提升伦理素养等六项基本伦理要求。

联合国欧洲经济委员会于2016年4月修正公布的《国际道路交通公约(维也纳)》(*Vienna Convention for Road Traffic*, 下称“《公约》”)将自动驾驶纳入其规范范畴。该《公约》规定,在全面符合联合国车辆管理条例或者驾驶人可以人工选择关闭该功能的情况下,将驾驶的职责交给车辆的自动驾驶技术可以明确地被应用到交通运输当中,这意味着自动驾驶在法律的层面上得到许可。不过,我国并非《公约》的签署国,在我国的法律层面上尚未承认自动驾驶的法律主体性。

在学说上虽然有基于意思自由的责任归属,从而否定人工智能的法律主体性,不过也有基于法人具有犯罪能力的观点说明人工智能的法律主体性,肯定人工智能的法律主体性。^[2] 本文认为,基于行为规范与制裁规范的关系,行为规范是由禁止、命令规范与容许规范所构成,一旦行为违反命令或禁止规范且又不符合容许规范时,应评价为不法而成立犯罪。换言之,任何能够履行行为规范之人在刑法当中即具有行为主体的资格。行为人故意实施符合犯罪构成要件的行为且具有实质违法性,如欠缺违法性认识而阻却责任,行为人因违反行为规范而成立犯罪,只是没有发动制裁规范的前提而不能处罚而已,并无碍于构成要件中行为主体的判断,在刑法中称为“无责任的行为主体”。^[3] 因此,当自动驾驶车辆行驶于公共交通运输环境中,承担了减少因人为因素引起交通事故造成重大的生命、财产的损失并保障公共交通运输环境安全的义务时,可以认为其是在遵守刑法基于保护法益的目的而设置的行为义务,进而也具有了其能否获得主体地位进行讨论的可能性,但这一主题并非本文关注的重点,故不再予以展开。

美国汽车工程师学会(Society of Automotive Engineers)于2021年4月发布最新的《自动驾驶分级与术语的定义》(*Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*)(下称“《J3016_202104》”)以驾驶人或是自动驾驶掌握车辆的驾驶权限区分自动驾驶的分级依据,据此可阐明自动驾驶在何种情况下会成为刑法上的行为主体。^[4] 详言之,在L0至L2时因驾驶人掌握驾驶权限而负有行为义务;^[5] 在L3时因驾驶人负有回应自动驾驶的要求而有接管驾驶权限的必要,因此自动驾驶只有在运作时负有行为义务,当系统发出要求驾驶人接管驾驶权限的指令或是自动驾驶未运作时,驾驶人负有行为义务;在L4与L5时,除了未运作或无法继续运作以外,自动驾驶负有完全的行为义务。因此,当L3至L5的自动驾驶掌握驾驶权限时,不仅负有遵守交通运输法规的行为义务,同时在刑法上也负有不得侵害交通运输安全的不作为义务与注意义务。本文所称面对碰撞困境的自动驾驶专指L3至L5掌握驾驶权限的自动驾驶,一旦自动驾驶陷入碰撞困境则有应对碰撞困境的行为义务,进而发生应如何决定优先保

[2] 佐久間修「AIによる自動運転と刑事責任」刑事法ジャーナル57号(2018年)12頁参照。

[3] 在刑法的三阶层理论中因严格区分故意与违法性认识的体系地位以及客观违法性论的立场,会存在“无责任的行为主体”,如无责任能力之人犯罪与间接正犯中的被利用人即属于此。参见余振华著:《刑法总论》,新学林出版社2015年版,第226、299页;川端博『刑法総論講義』(成文堂,2016年)424-427頁。

[4] 由于本次修正是美国汽车工程师学会和国际标准化组织共同合作修订,使得《J3016_202104》成为国际标准,对于我国自动驾驶的研发与法制化具有重大的意义,不可不重视。

[5] 根据《J3016_202104》的定义,L1与L2自动驾驶中具有紧急避让功能属于驾驶人辅助系统,该功能运作时会介入驾驶进行避让,纵使驾驶人短暂的时间无法操控车辆,驾驶人仍具有驾驶权限。

护对象的问题。

自动驾驶的碰撞困境是指自动驾驶在运行过程中,因突发紧急状况必须采取避让措施,但无论采取何种避让措施均无法完全避免侵害结果。当前自动驾驶的碰撞困境虽然基于交通运输环境的复杂性而出现各种各样的情形,但从导致自动驾驶陷入碰撞困境的原因观察,基本上能以是否因他人的行为导致自动驾驶陷入碰撞困境作为区分方式。一种是因他人的行为导致自动驾驶陷入碰撞困境的情形,如在自动驾驶的行进路线上突然出现从路旁窜出的行人或是行驶于高速公路时前车突然紧急刹车。另一种是非因他人的行为导致自动驾驶陷入碰撞困境的情形,如在自动驾驶的行进路线上突然出现障碍物,或是自动驾驶车辆出现无法预见的刹车失灵等机件故障,无论是自然因素或是自动驾驶车辆本身的因素均属之。由于自动驾驶的碰撞困境有一个很重要的特征,即无论采取何种避让措施一定会造成侵害,若是保全法益优越于被害法益,在必要限度内的侵害虽然具有正当性,但在生命不可比较衡量原则的要求下,若是以侵害他人生命保全自己生命的情形,如在碰撞困境中侵害其他交通参与人的生命以保护自动驾驶使用人的生命,是否具有正当性仍不无疑问。

本文拟从碰撞困境刑事正当性根据、以紧急避险为核心的碰撞法则以及碰撞困境中侵害生命的规则模式三个方面对自动驾驶如何应对碰撞困境带来的刑事责任问题展开分析。

二 自动驾驶应对碰撞困境中刑事正当性理论的缺陷

对于碰撞困境中牺牲他人生命保护自动驾驶使用人生命的正当性问题,在法哲学的讨论中已然有相当丰富的观点,如陈景辉认为无论是基于市场论证或乘客论证,并没有理由给予乘客优先保护的法律效力;^[6]朱振认为只有在付出合理代价并带来更大社会善的情形下,可对乘客施加可允许的伤害;^[7]翟小波主张以功利主义为指导,透过民主程序统一确定具有利他性和最小侵害性的选择方案;^[8]骆意中认为法律应允许自动驾驶在特定场景中优先保护乘客安全,^[9]但其实这些彼此争论的观点,本质上仍旧是为了解决生命之间价值无法取舍的根本伦理问题。为此,刑法学者尝试突破或缓和生命不可比较衡量的原则,分别从容许危险^[10]义务冲突^[11]与紧急避险等观点,试图由自动驾驶带来的利

[6] 参见陈景辉:《自动驾驶与乘客优先》,《华东政法大学学报》2020年第6期,第6页。

[7] 参见朱振:《自动驾驶汽车如何破解“电车难题”》,《华东政法大学学报》2020年第6期,第20页。

[8] 参见翟小波:《痛苦最小化与自动驾驶》,《华东政法大学学报》2020年第6期,第35页。

[9] 参见骆意中:《法理学如何应对自动驾驶的根本性挑战》,《华东政法大学学报》2020年第6期,第49页。

[10] Vgl. Hilgendorf, Autonomes Fahren im Dilemma, in: Hilgendorf (Hrsg.), Autonome Systeme und neue Mobilität, Ausgewählte Beiträge zur 3. und 4. Würzburger Tagung zum Technikrecht, 2017, S. 143; Freund, Strafrecht Allgemeiner Teil, 10. Aufl. 2018, § 2 Rn. 14, § 3 Rn. 44; Wörner, Der Weichensteller 4.0: Zur strafrechtlichen Verantwortlichkeit des Programmierers im Notstand für Vorgaben an autonome Fahrzeuge, In: ZIS 2019, 41, 41 ff.

[11] Vgl. Neumann, in: Kindhäuser/Neumann/Paeffgen Nomos StGB, 4. Aufl. 2013, § 34 Rn. 74; Perron, in: Schönke/Schröder StGB, 29. Aufl. 2014, Vorbem. § § 32 ff. Rn. 76; Weigend, Notstandsrecht für selbstfahrende Autos?, ZIS 2017, 600, 603.

益与不利益以及自动驾驶的保护义务价值高低说明生命取舍的正当性与自动驾驶的危险承担义务作为自动驾驶应对碰撞困境的正当性根据。

(一) 容许危险与自动驾驶的碰撞困境

1. 刑法教义学下容许危险的理论构造

所谓容许危险,是指具有法益侵害性质的采矿业、高速交通、医疗行为等社会活动,基于社会有用性而在一定范围内容许侵害发生。如高速交通的利用对人的生命或身体有侵害的危险,但若放弃高速交通则有碍于现代社会的各方面发展,因此在一定范围内对人的生命或身体造成侵害时仍具有正当性。^[12] 容许危险理论可以说是伴随德国的高度产业化发展而来。因产业活动带来的社会利益与由此所生的侵害危险进行比较衡量,若是利益大于危险,该活动则应被允许。不过,产业活动带来的危险未必仅以社会有用性作为容许的理由,从容许危险的概念被提出以后,其与规范违反之间的关系一直争论不休。

实际上,德国 19 世纪末的判例从区别违反营业命令的法规与过失,转变成将不遵守法规中安全措施的不作为直接认定为过失的根据。^[13] 在第二次世界大战之前虽然可以看出尝试着直接将行为的社会性格本身作为容许危险的理论根据,但这种观点是因目的行为论才开始全面展开。威尔泽尔(Hans Welzel)将过失结果犯作为惹起犯而处于与故意犯不同的体系地位,以社会相当性限定过失犯的处罚范围。在此所谓社会生活上必要的注意,也就是容许危险的社会相当性包含实现注意规定的行为在内,^[14] 认为容许危险是指行为人已履行法规范要求的注意义务,但在社会生活中所容许的危险范围内造成侵害的结果并非法所不容许的危险,因此欠缺构成要件的结果归责。^[15] 举例而言,餐厅在法规的允许范围内贩卖酒类饮料给顾客饮用属于容许危险的范围,顾客若是饮酒后因酒精作用不慎摔倒受伤,不可归责于餐厅贩卖酒类饮料的行为。因此,在判定容许危险之际,必须在各种安全规定当中慎重衡量行为的社会忍受可能性。

对于容许危险的体系定位问题,不同的学说立场可能会得不同的结论。基于行为无价值论的立场,有的学者认为容许危险可以作为独立的阻却违法事由。^[16] 出于新过失论的观点,只有在遵守社会生活中必要的注意仍发生法益侵害结果时可阻却违法性。然而,同样基于新过失论的立场,也有学者主张容许危险并非阻却违法事由而是认定注意义务的标准,若是已遵守注意义务却依然发生法益侵害的结果,会认定为容许危险而非过失而不具有构成要件该当性。^[17] 基于结果无价值论的立场,有观点认为没有必要特地将没有实质危险的情形称为容许危险,且有实质危险的行为可以被容许,必须是比较衡量保全法

[12] 只木誠『コンパクト刑法総論』(新世社,2018年)103頁参照。

[13] 松宮孝明『刑事過失論の研究(補正版)』(成文堂,2004年)10頁参照。

[14] 林美月子「許された危険」西田典之ほか編『刑法の争点』(有斐閣,2000年)38頁参照。

[15] クラウス・ロクシン著,平野龍一監修,町野朔=吉田宣之監訳『ロクシン刑法総論第一卷(第3版)』(信山社,2003年)313頁参照。

[16] 福田平『全訂刑法総論(第4版)』(有斐閣,2004年)145頁;大塚仁『刑法概説総論(第3版増補版)』(有斐閣,2005年)339頁;大谷実『刑法講義総論(新版第2版)』(成文堂,2007年)277頁参照。

[17] 井田良『刑法総論の理論構造』(成文堂,2005年)244頁参照。

益与被害法益后,保全法益优越于被害法益的情形。^[18] 基于这种观点,有鉴于生命与身体的优越性,在比较衡量生命、身体与遂行行为带来的利益时,如有侵害生命的危险程度,就拥护他人的生命而言,只有在极为限定的范围内才可认为是容许危险。^[19] 实际上,看似对立的两者并没有绝对的分歧。对履行结果回避义务而减少行为危险性的结果无价值而言,当危险减少后已失去过失犯的实行行为性,从一开始就没有称之为容许危险的必要,这种观点可以说与前述行为无价值论中认为容许危险不具有构成要件该当性的见解一致。对此,罗克辛也认为,若是容许危险不能放入正当性事由的框架内承认,则没有必要将其作为一种独立的正当性事由。对引起法规范关注的危险的容许,会导致这种行为无法归责,不会被评价为构成要件行为。如所有遵守交通法规的驾驶行为虽然无法否定对生命、健康以及财产可能会造成的重大危险,但因其具有压倒性的公共利益,立法者在一定的注意规定范围内容许驾驶行为。但是对容许危险的衡量并非像正当防卫或紧急避险一般具体衡量保全法益与被害法益,而是基于完全不会成立犯罪的概括性衡量,即使是在个别具体的事例中不存在更高的利益或是用于实现非法目的,也都并非基于正当性原则来衡量行为的容许危险。^[20]

本文认为,构成要件该当性旨在形式判断违法且有责的犯罪类型,违法性旨在具有构成要件该当性后实质评价其正当性,从这两种概念在犯罪论体系中的意义来观察,容许危险是阻却构成要件事由或是阻却违法事由的体系地位之争,是出于本应形式判断的构成要件该当性却以实质评价的方式为之。易言之,在构成要件该当性的判断中,如有对法益造成侵害结果或侵害危险即可满足构成要件中的结果要素,不问该侵害结果或侵害危险是否具有正当性,其正当性是在违法性当中透过社会相当性以及法益衡量的方式综合评价。容许危险同样具有法益侵害的性质,一旦造成法益的侵害结果或侵害危险依然会有构成要件该当性,没有考虑该侵害结果或侵害危险是否具有正当性的必要;但在违法性的评价中,法规范容许此等危险的存在是因为这是伴随社会上必要不可或缺的行为而来,带来的利益也显著大于可能造成的侵害,可以说同时通过了社会相当性与法益衡量的检验,纵使在一定的条件下该危险造成法益侵害而具有构成要件该当性,仍然可做为一种阻却违法事由而阻却违法。所以,如果要讨论容许危险能否作为独立的违法性阻却事由,可能还需要从社会相当性出发,考虑危险行为的有用性与必要性、危险的盖然性、受侵害的法益性质与程度、能够防止危险发生的手段等因素。

2. 自动驾驶在碰撞困境中的容许危险

确立容许危险在刑法理论中的意义与体系地位后,接下来要回答的问题是自动驾驶在碰撞困境中的容许危险所指为何? 容许危险的场景是实行行为已遵守法规范要求的行为义务,但仍然发生侵害的事实,所以容许危险理论应对当前自动驾驶碰撞困境的两种场景时,必须限定以自动驾驶遵守法规范的前提。据此,进一步分析自动驾驶在碰撞困境造

[18] 平野龍一『刑法総論』(有斐閣,1972年)198頁参照。

[19] 山口厚『刑法総論(第2版)』(成文堂,2007年)231頁参照。

[20] クラウス・ロクシン著,平野龍一監修,町野朔=吉田宣之監訳『ロクシン刑法総論第一卷(第3版)』(信山社,2003年)408頁参照。

成生命侵害的事实,学上有将容许危险区分为碰撞困境本身与碰撞困境中的生命侵害两种观点。

主张碰撞困境本身为容许危险的观点认为,自动驾驶对交通安全的提升效果显然优于在碰撞困境中带来的侵害,倘若为了避免在碰撞困境中造成任何人死亡或伤害而禁用自动驾驶,对于交通安全的发展而言显然是弊大于利,故根据容许危险建构自动驾驶在碰撞困境中的算法,并未违反基于最小侵害性原理的注意义务,^[21]由自动驾驶在碰撞困境中做出谁死亡或负伤的决定,并未充足过失致死罪或过失伤害罪的构成要件。纵使自动驾驶已尽其可能遵守注意义务的要求,仍然会出现碰撞困境导致无法避免的侵害,如要完全避免侵害发生,只有完全放弃自动驾驶才能避免,但自动驾驶的应用对社会而言具有重大的利益,故不可能赞同放弃自动驾驶的应用。以接种疫苗为例,如在接种疫苗时已采取所有保护措施,其副作用仍有可能危害一部分接种者的健康,虽然完全放弃接种疫苗可以完全避免副作用对健康的侵害,但会使所有人的健康处于相当高的感染风险,因此接种者必须忍受接种疫苗的副作用带来的容许危险,故因副作用造成的健康侵害不可归责于医师遵守医疗准则的施打疫苗行为。同此原理,当自动驾驶遵守交通法规而行驶,无论是陷入前述第一种或第二种碰撞困境而侵害他人生命,均不成立犯罪。

这种观点认为碰撞困境对自动驾驶的应用而言是可能发生的“副作用”,若为了避免“副作用”侵害交通参与人而放弃自动驾驶的应用,显然会提高整体的交通危险,故不应放弃自动驾驶的应用。不过,这种观点误解了碰撞困境中容许危险的内容。碰撞困境中的容许危险并非是碰撞困境本身,而是陷入碰撞困境后造成的生命侵害是否为刑法规范所容许。是否发生碰撞困境与自动驾驶的应用并没有任何关系,纵使放弃了自动驾驶的应用,碰撞困境也不会随之消失,只不过是驾驶人来面对而已,因此将碰撞困境本身理解为自动驾驶在碰撞困境中的容许危险显然并不妥当。

主张碰撞困境中的生命侵害为容许危险的观点认为,刑法学的价值判断中并未对生命的多寡建构价值序列,因此在设定自动驾驶应对碰撞困境的算法之际,“一人的生命与多数人的生命相比具有同等价值”的命题是否妥当,仍有检讨的必要。人的生命之所以不能成为计算衡量的对象,是因为刑法规范已揭示了个人生命具有最高价值,若是牺牲无辜的生命仍无法救助多数人的生命,则对牺牲的生命而言仍不具有正当性,故仍有必要维持生命的绝对价值。然而问题在于,自动驾驶的碰撞困境并非是对过去特定态度的法律评价,而是对将来如何解决问题的价值判断,自动驾驶本身造成侵害是否具有正当性的评价未必会成为问题,反而“比起多数人成为受害者,不如认为是牺牲更少人”这种在不法阶段的评价才会成为问题。^[22]据此,有论者认为要区分不法的基础与不法的阶段,在不法的基础中虽然不允许对生命多寡的比较衡量,但在不法的阶段中则可允许对生命多寡的比较衡量。因自然因素导致自动驾驶陷入碰撞困境造成的生命侵害,如自动驾驶为了

[21] 富川雅満「エリック・ヒルゲンドルフ法と自立型機械—問題概説」千葉大学法学論集 31 卷 2 号(2016 年) 128 頁参照。

[22] Vgl. Hilgendorf, Recht und autonome Maschinen - ein Problemaufriß, in: Hilgendorf, Hötitzsch (Hrsg.), Das Recht vor den Herausforderungen der modernen Technik, 2015, S. 26.

避免冲撞突然出现的障碍物以救助车上 3 名乘客,按照算法变换行车路线而碰撞 1 名路人,由于自动驾驶的车辆有相当程度减轻了发生一般事故的危险,因此对所有交通参与人而言也发挥了极大的作用,故在这种情形中造成的生命侵害可理解为容许危险,不可归责至任何罪名的客观构成要件。

由此可见,自动驾驶在碰撞困境中造成法益侵害的结果为容许危险,是以自动驾驶的运行必须遵守交通运输法规为前提。事实上自动驾驶的运行设计是以交通运输法规所要求的行为义务为基础的,所以在正常运作下不会发生违反交通运输法规的情形,故因他人的行为导致自动驾驶陷入碰撞困境造成生命侵害,如自动驾驶碰撞从路旁违规窜出的行人,或是在高速公路上碰撞因错过匝道而紧急刹车的违规车辆,造成他人伤害或死亡的结果均属于容许危险而不成立犯罪。但是,因自然因素或自动驾驶车辆本身的因素导致自动驾驶陷入碰撞困境造成生命侵害是否也能认为是容许危险,仍有讨论的空间。《J3016_202104》中有关驾驶人与自动驾驶在车辆行驶中的作用,说明驾驶人不仅可要求自动驾驶使事故风险最小化,自动驾驶本身也应具有事故风险最小化的作用,但在刑法规范的评价上并不表示自动驾驶在陷入碰撞困境时使用人必然享有优先保护的地位,而是应考虑是否有承担危险的义务。在应承担危险的情形中,自动驾驶必须对使用人采取最低限度侵害的避让措施,在不应承担危险的情形中,自动驾驶对应承担危险的交通参与者采取最低限度侵害的避让措施,才能认为是容许危险。据此,因自然因素或自动驾驶车辆本身的因素导致自动驾驶陷入碰撞困境,自动驾驶基于保护使用人生命而侵害其他交通参与人的生命不可认为是容许危险,无法证明自动驾驶在这种碰撞困境中造成生命侵害的正当性。

综上所述,认同容许危险的观点虽然是限于行为人不存在避免侵害发生的权限,且其不存在并非行为人的责任,但本文认为自动驾驶并不属于这种情形。事实上,纵使遵守所有的注意义务虽然无法完全避免碰撞困境的发生,不过仍有可能在有些情景中避免自动驾驶对其他交通参与人的侵害,故对交通参与人的侵害显然并非是不可回避的状况。碰撞困境的争点是自动驾驶的避让措施造成生命侵害是否具有正当性,但对刑法保护法益的目的而言并不乐见侵害生命的结果发生。自动驾驶与驾驶人在碰撞困境中的不同之处在于,自动驾驶必须按照避让算法应对碰撞困境,但驾驶人在碰撞困境中未必会按照刑法规范的要求进行避让,故基于预防侵害的目的,倘若要让自动驾驶的避让措施具有正当性,只要事前合理且适当地设定自动驾驶的避让算法,即有可能回避侵害结果的发生。^[23]

(二) 义务冲突与自动驾驶的碰撞困境

义务冲突是指两个以上对立的法律义务存在时,无论履行哪个义务势必无法履行其他义务的情形。交通运输法规的规范目的,旨在维持道路交通的顺畅,确保交通运输环境的平稳。所有交通参与人行驶车辆时应遵守交通运输法规,自动驾驶或非自动驾驶均无例外。据此,交通参与人在交通运输环境中应履行的义务,至少可被理解为遵守交通运输

[23] 富川雅满「アルミン・エングレーダー自動運転自動車とジレンマ状況の克服 Das Selbstfahrende Kraftfahrzeug und die Bewältigung dilemmatischer Situationen」千葉大学法学論集 32 卷 1・2 号(2017 年)167 頁参照。

法规的义务。然而,交通运输行为本来具有发生交通事故的危险性质,交通运输法规的规范手段并非出于提升交通运输安全,而是降低交通运输的危险,以此规范交通参与人的交通运输行为,制定一系列在驾驶状态下应遵守的行动准则,此即为交通运输法规赋予交通参与人的行为义务。在自动驾驶的情形,必然是透过算法对车辆的操控来体现交通运输法规中的行为义务。这种行为义务可区分为积极义务与消极义务,前者属于命令交通参与人应采取的一定作为,如按规定车道行驶、夜间行车应开启车灯等,后者属于禁止交通参与人的特定作为,如不可超速行车、不可闯红灯等,但无论是积极义务或是消极义务,均以降低交通运输的危险为规范目的,据此,自动驾驶陷入的碰撞困境实质上属于数个交通运输法规中行为义务之间的冲突。

从义务冲突的观点观察自动驾驶在碰撞困境中的避让措施是否具有刑事正当性,应确认自动驾驶在碰撞困境中的积极义务与消极义务。因自动驾驶在碰撞困境中的碰撞对象除了其他交通参与人以外,也包含自动驾驶的使用人,既然自动驾驶的应用旨在降低交通运输的危险,自动驾驶除了负有保护使用人安全的积极义务,同时也负有不可侵害其他交通参与人的消极义务。因此,碰撞困境中义务冲突的类型可分为积极义务与消极义务的冲突以及消极义务与消极义务的冲突。^[24]

所谓积极义务与消极义务的冲突,如自动驾驶面临紧急状态时采取的避让措施会侵害其他交通参与人,但不采取避让措施则会侵害使用人的情形。在这种义务冲突的碰撞困境中,本应按照自动驾驶的使用人与其他交通参与人的数量来决定碰撞对象,但自动驾驶主要的应用目的是保障使用人的安全,^[25]此时自动驾驶的积极义务显然高于消极义务,即使被撞的其他交通参与人多于自动驾驶的使用人,自动驾驶仍应优先选择碰撞其他交通参与人,^[26]才符合义务冲突的原理要求。所谓消极义务与消极义务的冲突,如自动驾驶在陷入无碍于使用人安全的紧急状态时,有选择碰撞5人与碰撞1人的情形。在这种义务冲突的碰撞困境中,不得碰撞5人的义务显然优于不得碰撞1人的义务,因此自动驾驶在选择碰撞1人时获得了一定程度上的刑事正当性。^[27]

基于义务冲突的观点,虽然可以从义务的优越性判断应优先履行的义务,但其判断的标准是否妥当仍不无疑问。按照《J3016_202104》的说明,使用人享有要求自动驾驶最小化事故风险的权利,且自动驾驶也有使事故风险最小化的义务,由此可知自动驾驶有优先保护使用人的义务。在积极义务与消极义务的冲突中,无论引起事故的危险是否可归属于自动驾驶,若是认为优先保护使用人而碰撞其他交通参与人具有正当性,对其他交通参与人而言反而会是一种歧视,也就是未使用自动驾驶的交通参与人,在刑法中的保护程度低于自动驾驶的使用人,对未使用自动驾驶的交通参与人反而保护不足,在法益保护上有

[24] 佐藤英明「自動運転車とトロッロ一問題」中央学院大学人間・自然論叢48号(2020年)37頁参照。

[25] 按照《J3016_202104》说明驾驶人于自动驾驶在车辆行驶中的作用,驾驶人不仅可要求自动驾驶使事故风险最小化,自动驾驶本身也应具有事故风险最小化的作用。由此可知,自动驾驶在陷入紧急事态时,驾驶人享有优先保护的地位。

[26] 参见王钰:《生命权冲突的紧急状态下自动驾驶汽车的编程法律问题》,《浙江社会科学》2019年第9期,第78页。

[27] 参见杨昆鹏:《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》,《新疆大学学报(哲学·人文社会科学版)》2021年第5期,第25页。

失均衡。义务冲突的观点着眼于实现义务后的价值,并未考虑到引起事故危险的原因以及危险承担义务,在消极义务与消极义务的冲突中,纵使引起事故危险的原因可归属于 5 人而仍应优先保护 5 人而选择碰撞 1 人,对被碰撞的 1 人而言,并没有义务承担 5 人引起的事故危险,故在这种情形下,基于义务冲突的观点而选择碰撞 1 人的理由仍不够充分。^[28]

三 探究以紧急避险为核心的自动驾驶碰撞法则

笔者认为,碰撞困境中对容许危险的认识出现了混淆,而义务冲突中衡量义务优越性的基准会导致刑法规范的保护歧视等问题,故这两种观点显然均难以充分说明自动驾驶的碰撞正当性。鉴于“容许危险”与“义务冲突”两种新兴理论的上述缺憾,对于自动驾驶的碰撞困境,我们仍应回到传统的“紧急避险”法理来寻找其基础。事实上,长期以来对于驾驶人碰撞困境,学说上均是以“紧急避险”作为其基础,通过衡量保全法益与被害法益来决定其采取的避让措施侵害他人生命是否具有刑事正当性。换言之,自动驾驶在面对突如其来的紧急状况时,基于保护使用人的要求,在可以同时保全使用人与其他交通参与人的情形当然不得采取侵害其他交通参与人的避让措施以保护使用人,但在无论采取何种避让措施都会侵害生命的碰撞困境,自动驾驶能否基于保护使用人的目的而对被撞的交通参与者主张紧急避险则有争议。

在陷入碰撞困境时,驾驶人大部分是出于无意识或是基于本能的判断而实行避让措施,自动驾驶则是受到操控车辆的软件算法所左右而实行避让措施。驾驶人与自动驾驶陷入碰撞困境时最根本的不同是,当驾驶人陷入碰撞困境时才会在自己与其他交通参与者之间做出避让的选择,但自动驾驶早已在事先的算法设定中做出避让的选择,因此在评价侵害生命的避让正当性时,对驾驶人而言是避让正当性的法则适用问题,而对自动驾驶而言不仅是法则适用问题,更是法则设定问题。再者,自动驾驶如何通过传感器掌握外部的行车环境是属于技术层面的问题,驾驶人与自动驾驶如何应对碰撞困境则属于刑法规范层面的问题,如果避让措施要获得刑事正当性,其选择侵害的对象需符合紧急避险衡量保全法益与被害法益的评价基准,且这种衡量法益的评价基准并不会因为是驾驶人或是自动驾驶实施的避让行为而改变。因此,当自动驾驶陷入碰撞困境时,为了避免使用人或是其他交通参与人的危难而造成侵害能否成立紧急避险,有两个问题必须探讨,其一为自动驾驶在碰撞困境中的避让措施与危难紧迫性的要求,其二为自动驾驶在碰撞困境中价值衡量的基准。

对于危难紧迫性问题,由于自动驾驶对碰撞困境的避让措施是通过算法预先设定,学说上有批评观点认为制造人在设计自动驾驶的避让措施时还不存在任何的危险,也就是欠缺紧急避险的危难紧迫性要件,因此设计时想象可能发生的危难而预先设定好自动驾驶

[28] 基于危险承担的观点批判义务冲突下的碰撞法则,认为在积极义务冲突与消极义务的冲突中,纵使是由自动驾驶引起事故的危险,仍应选择自撞以保护其他交通参与者方为妥当,但这种自动驾驶能否被大众接受,则是现实上应交由市场选择的问题。

驶的避让措施并无法成功保护使用人的安全。然而,制造人事先设定自动驾驶陷入紧急情况时的避让措施,性质上与预测到会有不法侵害而事先准备好高压电网等自动防卫装置的正当防卫并无不同,故制造人事先设定自动驾驶陷入紧急情况时的避让措施并无不妥,且就危难紧迫性的判断而言,并非以安装系统或是系统运作的时间点来决定,而是以自动驾驶采取避让措施的时间点来决定,也就是自动驾驶在紧急危难存在的一瞬间采取避让措施即可认为是紧急避险。

关于价值衡量基准问题,需要注意紧急避险中保全的法益必须具有显著优越性。基于多数说的优越利益说虽然是将法益的价值与危难的程度、结果回避的可能性、被害人的自律性、危险的答责性、危险忍受义务的存否、个人喜好之间的关联性作为显著优越性的衡量要素,但仍是以法益的价值高低为前提,若是保全法益的价值低于被害法益,纵使仅需造成被害法益的轻微伤害即可避免保全法益的危害,仍然不可主张紧急避险,更遑论保全法益与被害法益具有同等价值的情形。然而,在对生命的碰撞困境中能否成立紧急避险,有观点认为生命不可成为比较衡量的对象而不可成立紧急避险;^[29]有观点认为基于生命的法律价值平等,牺牲等价的生命来保全自己的生命是法律不可控制的原始本能,牺牲他人生命的紧急避险有利于实现社会的最大利益;^[30]有观点基于人的生命是等价的,主张应视不同情况综合判断,对生命的紧急避险有可能成立阻却违法事由,也有可能成立阻却责任事由。^[31]对生命的紧急避险之所以会成为争议问题,是因为各种评价基准的理论局限性所导致,如生命不可比较衡量的观点是为了避免产生法律纠葛,却导致法律只保护强者的局面;为了实现的最大利益而牺牲的生命若是没有为他人承担危难的义务,对其成立紧急避险会有欠公平;紧急避险观点无论是成立阻却违法事由或是阻却责任事由,最终的结果皆是不予制裁,如此一来似乎并没有区分的必要。

本文认为对生命的紧急避险会成为理论争议的原因,是将生命能否比较衡量的争议与对生命的衡量基准这两个问题混为一谈。因此,要彻底解决这个争议,必须要对这两个问题分别思考,先探讨生命能否比较衡量,后说明对生命的衡量基准。

(一) 生命能否比较衡量争议的解决思路

人的生命是基于对人性尊严的保障而具有最高价值,所以应严守对生命比较衡量的原则,为避免陷入法律纠葛放弃价值判断,任由当事人的实力高低来决定应保护的主体。若是严守生命不可比较衡量的原则,当自动驾驶面对突然出现的障碍物,为了避免碰撞以保护使用人的生命安全,系统变换行车路线而撞死路人不能成立紧急避险,无论使用人是年轻人还是只剩下数日或数周生命的老年癌症患者,均不具有比较衡量的重要性。因此,在牺牲少数人而可以拯救多数人的情形中,也必须维持生命不可比较衡量的原则,如载有40名学生的自动驾驶校车为了闪躲突然出现的障碍物,因而改变行进路线撞死少数的路人仍不具有故意杀人罪的刑事正当性,在刑法规范的价值判断上,1名路人的生命与40

[29] 参见马克昌主编:《刑法学》,高等教育出版社2003年版,第134页。

[30] 参见王政勋著:《正当行为论》,法律出版社2000年版,第269页。

[31] 参见张明楷著:《刑法学(上册)》,法律出版社2017年版,第221-222页。

名学生的生命同等重要,故自动驾驶不能将事故的危险转嫁给不须承担事故危险的路人。

不过,有论者主张在自动驾驶的碰撞困境中应改变不准对生命比较衡量的原则,^[32]认为驾驶人是在陷入碰撞困境时才开始决定牺牲的对象与救助的对象,其决定通常是尽可能挽救更多生命而不是为了驾驶人的利益,就事故的整体而言,是以牺牲生命的最低限度为考虑,而这种决定在具体的碰撞困境当中,通常与驾驶人本身的利益完全相反。然而,自动驾驶的碰撞困境则是完全不同的情形,这是由于自动驾驶在遭遇碰撞困境之前已预先设定自动驾驶的避让措施,但此时尚未分配牺牲的对象,因此只能在事前基于所有具体情况皆未知的状态下,通过风险最小化的算法来保障所有人的利益,而在实际的碰撞困境中,则根据个别的情况来回避危险并决定实际的牺牲者。

前述见解似乎合情合理,但进一步检讨则会发现这种主张显然是流于表面。自动驾驶是以“特定的碰撞困境下应采取特定的避让措施”为前提,设定具有适用普遍性的避让措施,但是在设定避让措施的阶段,事实上并未确定谁会按照自动驾驶的避让措施而受有利益与承担不利益。驾驶人在碰撞困境中有可能甚至是必然会优先选择保护自己,但这对自动驾驶而言只是选择之一。此外,在确定行动的适用普遍性之际,碰撞困境中的相关人完全不知道自己的地位与立场,也就是这些人完全不知道自动驾驶在碰撞困境中如何分配自己的风险,只有在发生现实上的碰撞困境而有适用避让措施时,才能确定如何分配风险。换言之,与驾驶人决意牺牲他人以阻止危险的情形完全相同,自动驾驶的避让措施必须要陷入碰撞困境时才会决定谁要承担不利益。

然而,有反对意见认为前述观点忽略了一个重要的问题,即决定时间点的不同。^[33]自动驾驶为了救助他人,在还没有确定谁是必须牺牲的时间点已做出决定,之后完全按照避让措施自动运作来实现救助与牺牲;驾驶人的决定反而与避让措施无关,而是在陷入紧急情况时才开始决定自己要怎么避让。不过,反对意见不具有说服力。驾驶人确实是在陷入现实上的紧急情况时才开始决定自己要如何避让,但做决定的时间点已经由刑法规范所确定,也就是驾驶人应如何避让,在碰撞困境发生前已经由法制化的紧急避险所决定,因此从刑法规范的角度观察,驾驶人已没有其他选择的余地。自动驾驶与驾驶人的不同之处是驾驶人的确有无视紧急避险的可能性,但无论是对自动驾驶或驾驶人,同样都是要求在已经有设定避让措施之际做出适当避让的决定,这一点并没有任何改变。因此,纵使驾驶人在碰撞困境中决意违反紧急避险的要求,在某种意义上,相当于自动驾驶因算法的前提条件有缺陷而无法实行避让措施。最终,若是坚持在自动驾驶的碰撞困境中应改变生命不可比较衡量的原则,必须要全盘考虑不能牺牲少数人生命以拯救多数人生命的原理,但从人性尊严的保障并无法充分说明改变这个原则的理由。

若是严守生命不可比较衡量的原则,很可能导致本来不应承担危险之人受到侵害,而本来有承担危险义务之人反而受到保护这种显失公平的情形发生。因此,有论者提出

[32] Vgl. Hevelke, Nida-Rümelin, Selbstfahrende Autos und Trolley-Probleme: Zum Aufrechnen von Menschenleben im Falle unausweichlicher Unfälle, in: Jahrbuch für Wissenschaft und Ethik, Bd. 19, 2015, S. 10 ff.

[33] Armin Engländer 著,田村翔译「シレンマ状況における自動走行車:トrolleyー問題 4.0」ノモス43号(2018年)121頁参照。

社会连带性的观点来缓和生命不可比较衡量的原则。所谓社会连带性,是指社会成员彼此之间负有参与社会共同生活的连带义务,每个成员均有义务牺牲少许自我利益来拯救陷入危难的其他社会成员,^[34]如自己陷入紧急危难时对他人的避险行为造成侵害可成立紧急避险,他人陷入相同的紧急危难时对自己的避险行为造成侵害也可成立紧急避险。换言之,基于一般人的立场,当双方的立场互换后仍有成立紧急避险的可能性,相互之间负有为对方牺牲的社会连带义务,因此在具有社会连带性的紧急避险中,他人负有牺牲的义务,但赋予他人牺牲的义务并非存在于所有的紧急避险中,而是只有在保全法益显著优越于被害法益时才成立。^[35]有别于优越利益说对于显著优越性的判断,社会连带性认为被害人的自律性与连带性的存否,是在设定衡量基准时的考虑事由而非具体的衡量要素,作为首要衡量要素仍然是法益的价值,但应考虑法益的损害程度与危险发生的可能性,如造成价值较高的法益轻微损害能避免价值较低的法益重大损害仍可成立紧急避险。至于如何判断法益价值的高低,则是以下对生命的衡量基准要说明的问题。

(二) 紧急避险中对生命的衡量基准

按照社会连带性的观点,决定各种利益的重要性需要进行整体的价值衡量,其中必须考虑到的是所有关于保全法益以及被害法益的保护必要性事由,不仅是生命、身体、财产等法益之间的法律位阶,法益的危险程度与侵害的规模、救助机会的高低与因自己的疏失而负有特别的危险承担义务等均属之。本文认为,从自动驾驶碰撞困境的成因来思考,虽然陷入碰撞困境的原因与形态不胜枚举,但大致上可区分为自然因素与人为因素。因自然因素导致自动驾驶陷入碰撞困境,如认为自动驾驶有承担危险的义务,其他处于安全状态的交通参与人不存在为自动驾驶的使用人或乘客牺牲的义务,自动驾驶将其应承担的危险转嫁给其他处于安全状态的交通参与人并造成生命侵害,不成立紧急避险;如自动驾驶与其他交通参与人均有承担危险的义务,可认为相互之间负有牺牲的义务,在这种情形下自动驾驶侵害应承担危险的交通参与人的生命,仍可成立紧急避险。因人为因素导致自动驾驶陷入碰撞困境,如因其他交通参与人违反交通运输法规而引起事故,违反交通运输法规的交通参与人不仅负有承担危险的义务,自动驾驶也不存在为其牺牲的义务,故自动驾驶为了保护使用人或乘客的生命而侵害违反交通运输法规的交通参与人生命,可成立紧急避险;如交通参与人的行为并未违反交通运输法规,但引起的事故危险导致自动驾驶陷入碰撞困境,自动驾驶同样不存在为其牺牲的义务,因此自动驾驶为了保护使用人或乘客的生命,对引起事故危险的其他交通参与人造成生命侵害仍可成立紧急避险。

四 自动驾驶在碰撞困境中成立紧急避险的类型与归责模式

前述紧急避险的原理通过社会连带性合理分配碰撞困境中的危险承担义务可支持自

[34] 参见余振华著:《刑法总论》,我国台湾地区新学林出版社2015年版,第256页。

[35] 富川雅满「アルミン・エングレーダー自動運転自動車とジレンマ状況の克服 Das Selbstfahrende Kraftfahrzeug und die Bewältigung dilemmatischer Situationen」千葉大学法学論集32卷1・2号(2017年)175頁参照。

自动驾驶的碰撞正当性,但自动驾驶在碰撞困境中的生命侵害成立紧急避险的范围,以及不成立紧急避险后应如何归责的问题,本文认为应从三种紧急避险的类型来分析检讨,也就是攻击型紧急避险、防卫型紧急避险与免责型紧急避险。按照三阶层的体系定位,攻击型紧急避险与防卫型紧急避险可归类为阻却违法型紧急避险,免责型紧急避险可归类为阻却责任型紧急避险。若是从碰撞困境的视角观察,攻击型紧急避险为前述非因他人的行为导致自动驾驶陷入碰撞困境的情形,防卫型紧急避险为前述因他人的行为导致自动驾驶陷入碰撞困境的情形,免责型紧急避险是在紧急危难中造成法益侵害的避险行为,虽然不成立阻却违法型紧急避险但可能阻却责任的情形。

(一) 阻却违法型紧急避险的成立范围

1. 攻击型紧急避险的界限

自动驾驶为了回避紧急危难而侵害他人法益,只有在保全利益显著优越于被害法益才能成立紧急避险而具有刑事正当性。举例而言,自动驾驶行驶中突然出现障碍物,若是撞上障碍物会导致使用人死亡,为了避免撞上障碍物而碰撞停靠在路边的车辆并造成损害,这种将自己面对的危险转嫁给无关第三人承担的紧急避险称为攻击型紧急避险,此时保全法益是使用人的生命,而被害法益是他人的车辆,基于生命比起车辆具有更高的价值的理由,自动驾驶的避让措施可成立紧急避险。相较于前述案例,如自动驾驶的避让措施并非造成他人车辆的损害,而是造成第三人死亡或重大伤害,如载有 40 名学童的自动驾驶校车为了回避出现在眼前的障碍物,撞死 1 名在人行道上的行人,按照社会连带性缓和生命不可比较衡量的原则,因自动驾驶校车面对事故危险负有承担危险的义务,人行道上的行人并没有承担自动驾驶校车事故危险的社会连带性,所以仍然不可主张紧急避险,最多只能免除其责任。^[36]

综上所述,自动驾驶在陷入这种碰撞困境时负有承担危险的义务,故自动驾驶不应将事故危险转嫁给不存在牺牲义务的交通参与者,而应严守交通运输法规有关避让行人、按规定路线行驶等规定,并在危害使用人的最低限度内采取必要措施,最大限度地维护使用人与其他交通参与人的安全。

2. 防卫型紧急避险的特殊性

所谓防卫型紧急避险,是指对引起事故危险之人实施避险行为而对其造成侵害的情形,如自动驾驶基于回避道路上的障碍物而变换行进方向撞伤摆设障碍物的行人。在这种情形中,只要保全法益不劣后于被害法益即为已足,也就是不要求保全法益显著优越于被害法益,^[37]此时事故的危险是由摆设障碍物的行人所引起而负有危险承担义务,因此在碰撞困境中的保护必要性已劣后于自动驾驶,故自动驾驶撞伤摆设障碍物的行人可成立紧急避险。换言之,在对生命的防卫型紧急避险中造成引起事故危险之人死亡,若是基于社会连带性的观点,因引起事故危险之人有承担危险的义务,自动驾驶为了保护使用人

[36] 富川雅満「アルミン・エングレーダー自動運転自動車とジレンマ状況の克服 Das Selbstfahrende Kraftfahrzeug und die Bewältigung dilemmatischer Situationen」千葉大学法学論集 32 卷 1・2 号(2017 年)163 頁参照。

[37] Armin Engländer 著,田村翔译「ジレンマ状況における自動走行車: トロリー問題 4.0」ノモス 43 号(2018 年) 121 頁参照。

或乘客的生命而造成其死亡应具有正当性。

危险应由其他交通参与者承担的情形,即前述因他人的行为导致自动驾驶陷入碰撞困境的情形。在公共运输交通环境中,所有交通参与者皆有遵守交通运输法规要求的义务与注意义务,当有交通参与者违反交通运输法规引起事故危险,或是虽未违反交通运输法规,但因他人行为引起事故危险,导致自动驾驶陷入碰撞困境,除了应承担事故危险以外,法律保护的必要性也随之降低,此时自动驾驶显然比引起事故危险的交通参与者更应受到法律保护,而引起事故危险的交通参与者负有承担危险的义务,故自动驾驶为了保护使用人或乘客的生命,造成引起事故危险的交通参与者生命侵害可成立紧急避险,其他未引起事故危险的交通参与者不存在为自动驾驶承担危险的义务,自动驾驶的避让措施对其造成的生命侵害不可成立紧急避险。

(二) 免责型紧急避险的可能性

免责型紧急避险是行为人的避险行为虽然不具有刑事正当性,但因欠缺责任要素而阻却责任。这种免责型紧急避险的典型立法例为《德国刑法典》(*Deutsches Strafgesetzbuch*)第35条,其背后的原理虽然为期待可能性,但并非所有欠缺期待可能性的避险行为均可阻却责任,而是限于没有任何避险方法可回避对生命、身体的紧急危难时,为了自己、亲人或具有密切身份关系之人采取违法的行为避免危难的情形(《德国刑法典》第35条第1项第1款),不过若是自招危难或是因特别法律关系而可期待行为人忍受危难时不阻却责任,但可减轻其刑罚(《德国刑法典》第35条第1项第2款),如驾驶人因超速行驶陷入碰撞困境而侵害他人法益,会基于驾驶人的义务违反性而不阻却责任。成立免责型紧急避险的前提虽然是行为人的避险行为不具有正当性,如在攻击型紧急避险中将危险转嫁给无辜的第三人,或是在防卫型紧急避险中侵害不需承担危险义务之人,但是通过社会连带性的观点限缩期待可能性在免责型紧急避险中的适用范围,限定在对自己、亲人或有密切身份关系之人面临的紧急危难,认为保护自己、亲人或有密切身份关系之人不仅是现存的利益,同时也是社会大众相互承认的利益,故在免责型紧急避险中,只要保全法益不劣于被害法益即可阻却责任。^[38]

我国刑法并没有明文规定免责型紧急避险的具体情形,若是自动驾驶在碰撞困境中做出不具有正当性的避险行为,理论上虽然有可能根据免责型紧急避险阻却责任,^[39]但本文认为不应以免责型紧急避险作为阻却自动驾驶责任的理由。详言之,自动驾驶与使用人或乘客之间不具有特定的身份关系,而是基于与信赖关系而成为共同承担行为义务的整体,信赖自动驾驶在碰撞困境中的避让措施能够合于法律要求,避免造成不当的侵害。因此,使用人或乘客与自动驾驶基于信赖关系,在攻击型紧急避险中同样负有承担危险的义务,不得将危险转嫁给无关的第三人,倘若自动驾驶违反承担危险的义务而对他人造成不当侵害,在责任评价上同样也可根据社会连带性的观点期待使用人或乘客履行承担危险的义务,故自动驾驶在碰撞困境中实施不具有正当性的避险行为,不应以欠缺期待

[38] Vgl. Engländer, in: Matt/Renzikowski StGB, 2013, § 35 Rn. 19 f.

[39] 参见储陈城:《自动驾驶汽车编程设计中解决“电车难题”的刑法正当性》,《环球法律评论》2018年第3期,第89页。

可能性作为阻却责任的理由。同此原理,引起事故危险之人在防卫型紧急避险中负有承担危险的义务,倘若自动驾驶对无关的第三人实施不具有正当性的避险行为,在有责任性的评价上,也不应以欠缺期待可能性作为阻却责任的理由。

(三) 自动驾驶欠缺碰撞正当性的归责分析

当自动驾驶未按照紧急避险的原理造成侵害,因不具有正当性而追究刑事责任时,由谁来承担刑事责任是碰撞困境的归责问题。学说上认为自动驾驶是一种人工智能的具体应用,可根据传感器感知外界环境而操控车辆作出反应,因此具备辨识能力与控制能力,倘若自动驾驶感知到碰撞困境却做出不具有正当性的避让措施造成侵害,能够以自动驾驶具备辨识能力与控制能力为理由承担刑事责任。^[40]

不过,这种观点混淆了责任能力中的辨识能力与控制能力,以及故意中的认识与意欲。故意中的认识与意欲是针对事实而言,也就是行为人是否有认识到自己的行为与后果符合某罪名的构成要件,如自动驾驶在碰撞困境中已感知到碰撞对象为行人,且预测到对行人的碰撞会造成行人死亡而仍决定碰撞,此为故意杀人罪的故意;但责任能力中的辨识能力与控制能力是针对规范而言,也就是能否辨识自己的行为是法所不容许行为的能力,与基于辨识而不为法所不容许的能力,如自动驾驶依其辨识判断碰撞行人为法所不容许但仍决意碰撞,此时自动驾驶即具有责任能力。

前述紧急避险的原理已明确自动驾驶的行为义务,制造人有义务按照紧急避险的原理编写自动驾驶陷入碰撞困境的避让算法,因此可认为自动驾驶陷入碰撞困境的避让措施反映了制造人的意志,而自动驾驶在碰撞困境中的避让措施是制造人意志的延伸,故有观点认为自动驾驶在碰撞困境中造成的不当侵害应归责于制造人,此时制造人应承担碰撞事故的刑事责任。不过,这种观点在牺牲使用人的情形下被批评与自然法则的人类形象不一致,认为对制造人附加编写具有牺牲使用人正当性的避让算法,是间接强制自动驾驶的使用人在碰撞困境中牺牲现实的利益。^[41]然而,这种批评忽略了不得在碰撞困境中牺牲无辜生命以保护自己生命的法规范要求。无论是具有驾驶权限的驾驶人,或是编写自动驾驶避让算法的制造人,甚至是自动驾驶的使用人均毫无例外地应遵守法规范要求,也就是在紧急危难发生之际,制造人设定的避让措施违反法规范要求,牺牲无辜生命而保全使用人生命仍然是一种违法态度的表现,除非是因为自动驾驶的避让算法发生无法发现的瑕疵可免责外,^[42]制造人仍负有刑事责任。

相较于牺牲使用人的归责,对于为了救助其他交通参与者而将自动驾驶设计成牺牲使用人或乘客的归责问题,如自动驾驶面对突然出现在行进路线上的障碍物,为了避免行人死亡而牺牲使用人或乘客。在这种情形中,虽然要考虑的是制造人成立过失致死罪的刑事责任,但若是使用人在碰撞困境中已经认识到自动驾驶会采取的避让措施,或是可能认识到自动驾驶会采取的避让措施,因此造成使用人或乘客的侵害不能归属于制造人的

[40] 参见刘宪权主编:《人工智能:刑法时代的挑战》,上海人民出版社 2018 年版,第 11 页。

[41] 参见储陈城:《自动驾驶汽车程序中解决“电车难题”的刑法正当性》,《环球法律评论》2018 年第 3 期,第 89 页。

[42] 参见刘宪权:《人工智能时代的刑事风险与刑法应对》,《法商研究》2018 年第 1 期,第 7-8 页。

算法设定,可认为是使用人的自我答责危险,同时包含对风险的认识在内,只有在使用人可判断自动驾驶的反应时,才能肯定使用人的危险支配性。^[43] 不过,使用人未必能获得自动驾驶在使用上的完整信息,若是误信自动驾驶在碰撞困境中会优先保护使用人的生命,则应根据容许危险排除归责。然而,改变避让措施的算法设定确实能避免对使用人的侵害,但刑法规范不允许牺牲无辜的生命,因此这种避让算法未必是刑法规范所容许,只能在设定避让算法之际排除其他不合理的碰撞选项,才能避免这种评价矛盾。

五 结 语

自动驾驶是当前人工智能的典型应用之一,其目的是旨在减少发生交通事故的人为因素,提升交通运输安全,因此自动驾驶的保护对象不仅包括自动驾驶的使用人,同时也包括利用交通环境的其他交通参与者,然而,面对交通环境的复杂性引发的紧急情况,自动驾驶经常被迫在使用人与其他交通参与者之间做出保护选择,此时应如何判断选择是否具有正当性而不构成犯罪?这正是自动驾驶人工智能伦理中面临的碰撞困境问题。

驾驶人碰撞困境的研究与自动驾驶不同之处在于驾驶人存在随机选择碰撞对象的可能性,而自动驾驶是在事前的算法中设定好陷入碰撞困境时应执行的避让措施,一旦发生碰撞,完全是按照算法设定来执行避让措施。因此,对自动驾驶碰撞困境的刑事正当性研究要透过法理分析探求自动驾驶的避让正当性基础,提出自动驾驶陷入碰撞困境时应采取的避让模式设定,以作为设定自动驾驶陷入碰撞困境时应执行的算法内涵,从而避免自动驾驶应对碰撞困境的法律纠葛。

自动驾驶的碰撞困境不仅是现在要面对的选择难题,更是日后扩大人工智能应用后可能面对的选择难题,虽然在法哲学上有电车难题的讨论,但仅限于事实上的价值选择,在具体的法律适用上,仍应赋予价值选择的法律意义以及正当性。本文从紧急避险的观点出发对自动驾驶碰撞困境的犯罪评价提出了具体的建议,但有以下几点需要说明:首先,在发展自动驾驶车辆之际,碰撞困境是自动驾驶汽车制造人无法规避的难题,从技术与法律结合的层面研究自动驾驶的碰撞困境也是非常有意义的;其次,自动驾驶的碰撞困境属于全新的研究领域,仍应进一步确认问题与阐明问题,但至少应接纳以紧急避险作为论证碰撞困境的出发点;最后,在自动驾驶的碰撞困境中,在对其他交通参与人的侵害进行犯罪评价之际,应以紧急避险的各种基准为基础,就结论而言,不应以最小侵害性原理为前提将生命作为计算衡量的标准。如此,即可要求自动驾驶汽车制造人在设计自动驾驶的算法时,负有将“不能牺牲其他交通参与人的生命来回避对自动驾驶使用人的生命危险”化为自动驾驶碰撞法则的行为义务。

[本文为作者参与的2022年度广东省普通高校创新团队项目(2022WCXTD008)“智慧法治研究团队”的研究成果。]

[43] 富川雅満「アルミン・エングレーダー自動運転自動車とジレンマ状況の克服 Das Selbstfahrende Kraftfahrzeug und die Bewältigung dilemmatischer Situationen」千葉大学法学論集 32 卷 1・2 号(2017 年)173 頁参照。

[**Abstract**] Being an application of artificial intelligence, autonomous driving sets in the prior algorithm the avoidance measures that should be implemented in a collision dilemma. In case of collision, it faithfully implements the avoidance measures set by the algorithm, unlike a driver who has the possibility of randomly selecting collision objects. With respect to the response to the collision dilemma of autonomous driving, there have been discussions on the “trolley dilemma” in the philosophy of law and the dispute between utilitarianism and liberalism. On this basis, the legitimacy of its value choice should be extended to the specific design of collision rules and their corresponding legal application. At present, collision dilemmas of autonomous driving can be divided into two categories: those caused by acts of other people and those that are not. The former include the sudden appearance of a pedestrian crossing the road in violation of traffic rules or someone doing an emergency brake in front of the autonomous car on the expressway, and the latter include natural factors or problems with autonomous driving vehicles. Some people hold that, the collision dilemma of autonomous driving belongs to the permitted danger under criminal law, but the permitted danger in collision dilemma refers to the infringement upon the right to life caused in collision dilemma allowed by criminal law norms, and it is obviously inappropriate to understand the collision dilemma itself as the permitted danger of autonomous driving; some others believe that the collision dilemma in autonomous driving belongs to the conflict among several behavioral obligations under transportation laws and regulations, and the legitimacy of infringement should be judged according to the level of obligation. However, the obligation conflict fails to take into account the causes of the risk of accident and the risk-taking obligation and will lead to disputes such as protection discrimination in criminal law norms. It is rational to demonstrate the risk-taking obligation of autonomous driving in a collision dilemma on the basis of necessity and social solidarity. In the case of aggressive life-threatening necessity, because of the risk-taking obligation of autonomous driving, the infringement upon the right to life of an innocent traffic participant does not constitute a necessity. Under such circumstance, avoidance measures must be able to maximize the protection of the safety of users and other traffic participants while minimizing the infringement upon the rights of the users or passengers of the autonomous vehicle. In the defensive life-threatening necessity, because the person who causes the risk of the accident has the risk-taking obligation, the infringement upon his right to life by autonomous driving constitutes necessity. Based on the trust relationship between autonomous driving and the users or passengers, users or passengers can be expected to fulfill their risk-taking obligation. Therefore, the lack of expectation possibility should not be taken as a ground for the elimination of liability when autonomous driving implements unjustifiable risk avoidance measures in a collision dilemma.

(责任编辑:贾元)