

## 自动汽车程序设计中解决“电车难题”的刑法正当性

储陈城

**内容提要:**自动汽车的普及指日可待。随着自动汽车时代的到来,人为因素导致的交通事故数量虽然能够大幅减少,但是囿于技术发展的局限性,因技术瑕疵引发事故的可能性仍然存在。例如自动汽车在程序设计过程中所面临的“电车难题”,直至目前各国尚没有得出合理的解决方案。通过程序设计优先保护车内人的生命法益,侵害行人生命法益,因不符合具体的利益优越性和期待可能性,所以难以通过紧急避险阻却违法及责任;而优先保护行人生命法益,侵害车内人的生命法益,既不符合汽车的市场需求,也无法通过被害人承诺予以正当化。客观归责中被允许的危險法理或许是合理的解决路径。如果自动汽车的程序设计行为,满足被允许的危險理论的要素——概括的利益优越性和为立法所确立的最善义务,就可能在构成要件阶段被排除归责,从而在合理的限度内,保障自动汽车技术的发展。

**关键词:**紧急避险 被允许的危險 电车难题 人工智能 正当化事由

储陈城,安徽大学法学院副教授。

### 一 问题的提出

人类驾驶汽车发生的事故,有90%是因为人的失误导致,根据日本警察厅《平成28年交通事故发生状况》显示,交通事故的原因多为“未确认安全状态”、“不注意周围情况”、“粗心驾驶”等人为过失。未来随着自动汽车的普及,这些人为过失会得到有效地控制。<sup>[1]</sup> 自动汽车已经成为人工智能产业的排头兵,各国均觊觎未来巨大的市场规模,美国、日本、德国等科技与汽车产业巨头已经开始进行研究、开发和生产。<sup>[2]</sup> 据估计,到2035年自动汽车将占到汽车总数的25%,其中有10%为全自动汽车。如果自动汽车实

[1] 参见 <https://www.npa.go.jp/news/release/2017/20170322002.html>,最近访问时间[2017-04-10]。

[2] 参见 <https://www.dmv.ca.gov/portal/dmv/detail/vr/autonomous/testing>,最近访问时间[2017-04-10]。

现市场化普及,交通事故的发生量将会骤然减少。围绕自动汽车的法律问题,也自然是备受各国法律界关注的重要话题。比如以汽车产业著称的德国拜恩州的地方议会——“经济、媒体、基础建设、交通、能源及科技委员会”,于2015年10月29日举行的听证会的主题就是自动汽车,其中最为重要的议题之一就是自动汽车程序设计的适法性问题。<sup>[3]</sup>

根据美国道路交通安全局的定义,大致可将自动汽车的实现分为四个阶段。第一阶段:加速、操控方向盘、刹车等任意一个是由系统来操作,其余驾驶任务仍然由驾驶员来进行。第二阶段:加速、操控方向盘、刹车等当中有数个同时由系统掌控,则驾驶员具有监视义务。第三阶段:加速、操控方向盘、刹车全部由系统进行操作,只有在系统请求的情况下,驾驶员才需要做出相应的处理。第四阶段:加速、操控方向盘、刹车全部由系统操控,驾驶人员完全无需进行任何操作。其中,第一阶段是“安全驾驶支援系统”,第二和第三阶段则是“准自动驾驶系统”,第四阶段则能够被称之为“完全自动驾驶系统”。<sup>[4]</sup>目前国内外自动驾驶系统技术,如特斯拉汽车公司,已经达到第二阶段的水平。<sup>[5]</sup>在自动汽车技术的发展过程中,虽然可以大量减少交通惨剧,但是仍然无法避免很多突发因素导致的事故,尤其是在自动汽车技术商业化的早期阶段。虽然汽车厂商都在竭尽全力进行自动汽车相关技术的研发,但在脱离人为控制后,汽车事故还是会不时发生。对于优步自动汽车致路人死亡事件,有专家认为这次事故暴露了自动驾驶领域的技术难题,“任何一个传感器都无法确保100%的准确率”。“这起事故主要涉及准确及时地感知、识别到障碍物(本案例中为行人)并采取刹车等措施从而避免碰撞。那么,如果是某个传感器或摄像头出了故障(即存在缺陷),导致没有感知到行人,则优步可能需要承担责任。或者传感器都正常工作,但优步的软件系统没有识别到行人,导致没有采取刹车,亦可能需要承担责任”。<sup>[6]</sup>“美国优步自动驾驶汽车致人死亡事件,让开发企业和社会公众意识到,目前仍然处于发展期的自动驾驶技术,还有很多不完善的地方”。<sup>[7]</sup>总而言之,自动汽车虽然能够大幅提升驾驶的安全性,但是并不能完全消除因技术瑕疵或其他外因(如受到黑客攻击、病毒侵袭)导致的交通事故的发生。正因如此,美国非营利性组织FLI(Future of Life Institute)在就人工智能问题所发表的报告中预计,“自动驾驶的引入虽然可以减少一半以上的死亡案件,但是并不会因此收到感谢信,反倒有可能会让诉讼纠纷接踵而至”。<sup>[8]</sup>其原因是,当自动汽车进入第三阶段时,针对“电车难题”所进行的程序设计选择会导致相关主

[3] Vgl. <http://www.sueddeutsche.de/muenchen/landkreismuenchen/mobilitaet-teststrecke-in-die-zukunft-1.3050171>, 最近访问时间[2017-04-10]。

[4] 参见 Brian A. Browne, *Self-Driving Cars: On the Road to a New Regulatory Era*, 8 *Case W. Res. J. L. Tech. & Internet* 2-4 (2017)。

[5] 参见[日]中西孝樹著:《自動車業界のいまと未来がわかる本》,洋泉社2015年版,第225页;Brodsky, Jessica S. *Autonomous Vehicle Regulation: How an Uncertain Legal Landscape May Hit the Brakes on Self-Driving Cars*, 31 *Berkeley Tech. L. J.* 851 (2016)。

[6] 建峰、罗杰:《无人车“闯祸”事故谁“买单”》,《文汇报》2018年3月31日第6版。

[7] 参见袁曾:《人工智能有限法律人格审视》,《东方法学》2017年第5期,第50-57页;蒲晓磊:《自动驾驶立法步伐要跟上》,《法制日报》2018年3月27日第9版。

[8] Future of Life Institute Research, “Priorities for Robust and Beneficial AI”, 2015. [http://futureoflife.org/static/data/documents/research\\_priorities.pdf](http://futureoflife.org/static/data/documents/research_priorities.pdf), 最近访问时间[2017-04-10]。

体承担法律责任,尤其是刑事责任。<sup>[9]</sup>

“电车难题”,是指一方的重大法益只有通过侵害其他人的重大法益才可能得以保全,此时该如何选择的困境问题。<sup>[10]</sup> 这一难题在自动汽车的场合可以具体化为:车内人员重大身体法益或者生命法益,只有通过侵害车外行人的重大身体法益或者生命法益,才能够得以保护,自动汽车中的程序该如何设计的问题。比如第三阶段或者第四阶段的自动汽车,在十字路口正常行驶需要直行的时候,突遇紧急情况(包括传统的紧急情况:如由于初级阶段,自动驾驶汽车无法完备所有技术细节,导致在部分情况下遇到高速行驶的大货车自右向左驶来,难以躲闪;还包括新型的紧急状况:由于自动汽车较为依赖网联系统,而当网联系统受到黑客攻击或病毒侵袭,会导致相关的传感器系统失效等)。<sup>[11]</sup> 第三阶段的自动汽车来不及将操纵权转交给驾驶员来选择,或者第四阶段的自动汽车来不及进行刹车避免冲撞。此时,自动汽车如果保护车内人的安全,只能让汽车向左转弯,直接撞上步行道上的行人,导致行人死亡;如果要保护道路行人的安全,则将致使车内人员的死亡。<sup>[12]</sup> 那么此时,面临两难选择的自动汽车的程序系统应该如何设计才具有适法性?<sup>[13]</sup> 为了能够让自动驾驶技术得以良性发展,德国交通运输部正在专门针对自动驾驶技术的道德准则制定相关法律规范。我国发布的《人工智能标准化白皮书(2018)》中,就重点提到了人工智能的安全及标准问题,因此,自动汽车中程序设计问题的研究正当其时。

## 二 问题的哲学源头:生命权衡量应否禁止

自动汽车程序设计中“电车难题”的原型,最早始于古希腊学者卡涅阿德斯(Carneades)构想的思想实验,船舶遇难而在海上漂泊的甲乙两人,同时游到临近的一块木板上,但是由于木板的浮力只能支撑一个人,甲将乙撞开致其溺亡,自己独占该木板并得救。<sup>[14]</sup> 而在现实中,也真实发生过类似的案件,如英国的 Mignonette 号客船案以及美国的

[9] 参见[日]河島茂生、北村智、柴内康文:《自動運転車の「トロッコ問題」などに関する意識》,2017年社会情報学会(SSI)学会大会2017年9月17日講演。

[10] Armin Engländer, Das selbstfahrende Kraftfahrzeug und die Bewältigung dilemmatischer Situationen, ZIS (2016), S. 608 f.

[11] 参见吴汉东:《人工智能时代的制度安排与法律规制》,《法律科学》2017年第5期,第128-130页;Sergio Ferraz; Victor Del Nero:《人工智能伦理与法律风险的探析》,《科技与法律》2018年第1期,第21页。

[12] 这一问题已经成为各国学者的研究热点,德国、日本和美国学者认为,应对这种情形,必须在自动驾驶汽车设计时考虑到,并提前通过程序设计做出选择。Philip Weber, Dilemmasituationen beim autonomen Fahren, NZV 2016, S. 249 f.; Armin Engländer, Das selbstfahrende Kraftfahrzeug und die Bewältigung dilemmatischer Situationen, ZIS 2016, S. 608 f.; 平野晋:《「ロボット法」と自動運転の「派生型トロッコ問題」》,《NBL》2016年第1083号;金井猛嘉:《自動化運転を巡る法的諸問題》,《国際交通安全学会誌》2016年第2期;小林正啓:《自動運転車の実現に向けた法制度上の課題》,《情報管理》2017年第4号;Belay, Nick, Robot Ethics and Self-Driving Cars: How Ethical Determinations in Software Will Require a New legal Framework, 40 J. Legal Prof. 119 (2015); Huang, Bert I., Law and Moral Dilemmas, 130 Harv. L. Rev. 659 (2016)。

[13] 参见 Jean-François Bonnefon, Azim Shariff, Iyad Rahwan, The Social Dilemma of Autonomous Vehicles, Science 24 Jun 2016: Vol. 352, pp. 1573-1576。

[14] Vgl. Alexander Aichele, Was ist und wozu taugt das Brett des Carneades? JRE 11 (2003), 245 f.

Essex 号捕鲸船案。<sup>[15]</sup> 近年来,对于这一问题的讨论因为“电车难题”案的提出,又引起了学术界的关注<sup>[16]</sup>。

由此,理论界形成了两派对立的观点,一方以康德的“人是目的”学说为主,主张在任何时间、地点都禁止生命衡量;另一方则是以边沁的“最大多数人的最大幸福”学说为主,主张通过功利主义思想进行利益衡量,承认生命衡量的可能性。<sup>[17]</sup>

### (一) 康德和边沁的基本理论

近代西方社会经过文艺复兴、宗教革命和启蒙运动三个重要时期的洗礼,才逐渐形成欧洲的现代文明。由此形成了当时西方社会对于人、人性以及人的尊严的极度尊重。在这样的历史背景下,康德高扬人的地位和价值,提升人的境界和操守,提出人和物的本质区别,认为人是理性的存在者,而物是非物理性的存在者。作为理性的存在者的人,具有自我的意志;非理性存在者的物,则是依照自然意志存在。因此,人是具有绝对价值的,而物则没有,只具有相对价值。这就导出这样的论断:所有的理性存在者们都服从的法则是,不仅不能将自己和其他人视为手段,而且不论何时,都要同时将自己和他人作为目的对待,绝不能仅当作手段来使用。<sup>[18]</sup> 简言之,物可以作为手段,而人只能作为目的。“人是目的”的命题是康德在《道德形而上学原理》中提出的“目的公式”所蕴含的核心思想,是启蒙精神的集中体现,具有不可磨灭的理论意义。<sup>[19]</sup>

根据康德的“人是目的”的基本观点,我们应该尊重每个理性存在者所任意设定目的的能力,也就是其主观目的和意图,简言之,尊重每个理性主体的选择。这意味着,“即便一个人自愿选择通过损人利己的手段来获得自己的利益,我们也应该尊重他,将目的公式做此种解读显然违背了康德的道德律令”。<sup>[20]</sup> 从这一原则出发,将他人的生命仅仅作为手段来衡量是被禁止的。在“电车难题”案中,也不能以救自己或者多数人为由,而牺牲他人或者少数人的生命。基于对人的高度重视之立场,康德的“人是目的”学说一直备受关注与支持。

而 19 世纪初始,功利主义成为了西方道德和哲学主流观点之一,并且逐渐形成了其他学派对其无法构成威胁的局面。功利主义作为一种道德哲学,而非利益哲学,它要回答的是“我们应该做什么”的道德命题。功利主义的正统类型是行为功利主义,<sup>[21]</sup> 行为功利主义的主要观点是,以结果作为判断行为是否道德正确的标准。如果行为结果能够让最大多数人产生最大幸福,则该行为就是道德的。“最大多数人分享最大幸福”一语源于贝卡利亚《论犯罪与刑罚》引言中关于法律制定目的的表述。<sup>[22]</sup> 功利主义的代表人物边沁

[15] 参见 Nathaniel Philbrick, *In the Heart of the Sea: The Tragedy of the Whaleship Essex*, Penguin Books 2001, p. 1。

[16] 参见 Guerra-Pujol, F. E., Trolley Problems, 63 *Drake L. Rev. Discourse* 101 (2014)。

[17] 参见[日]川口浩一:《生命の衡量禁止原則とその例外》,《例外状態と法に関する諸問題》関西大学法学研究所 2016 年版,第 49 页。

[18] 参见《康德著作全集》(第 4 卷),李秋零译,中国人民大学出版社 2005 年版,第 436-437 页。

[19] 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 页。

[20] 王福玲:《“人是目的”的限度——生命伦理学视域的考察》,《中国人民大学学报》2017 年第 6 期,第 110 页。

[21] 其他还包括规则功利主义、双层功利主义等,参见晋运锋:《权利奠基于功利——当代功利主义对权利之基础的思考》,《法制与社会发展》2012 年第 6 期,第 130 页。

[22] 参见[意]贝卡利亚著:《论犯罪与刑罚》,黄风译,中国法制出版社 2002 年版,第 6 页。

受其影响,将休谟的幸福哲学与贝卡利亚的立法目的理论,形成行为功利主义的典型表达公式。<sup>[23]</sup>

边沁主张,行为是否在道德上正确与行为结果是否促进幸福相关联。幸福就是趋利避害的正面结果。功利主义“把追求快乐和避免痛苦看作行为的唯一目的,其他一切东西都是达到这一目的的手段”。<sup>[24]</sup>“最大多数人的最大幸福”是行为之后所达到的功利水平中的一种,即最大程度实现了最大功利的产生。将功利主义的判断公式推而广之,即“一切举措都应该通过功利原则进行最大限度的检验。具体说来,对于政府颁行的法令、政策乃至整个社会制度的评价标准,就在于看其是否指向了社会整体功利的最大化”。<sup>[25]</sup>

因此,从社会功利主义的观点来看,要防止出现导致社会利益变成负值的制度。在具体的个案的解决中也彻底贯彻这一立场,肯定生命可以进行衡量。在为了救助自己,将无关系的第三人杀害的案件中,只要满足补充性要件的话,就具有被正当化的可能性。<sup>[26]</sup>

## (二) 康德和边沁理论的盲点与对峙

当然,无论是康德的“人是目的”理论亦或边沁的“最大多数人的最大幸福”学说,在解决生命能否衡量这一问题上,都具有难以自圆其说的灰色地带。

康德的思想在启蒙时代,具有唤醒人类主体意识的伟大价值,促进了人类文明的进步。但是在后启蒙时代,尤其是全球一体化的今天,康德的“人是目的”理论,将人作为绝对的目的,把物当作绝对的手段,这种观点只能停留在纯粹理论的层面,在现实生活中不具有可操作性。因为,在现实生活中,人与物、目的与手段之间并不存在绝对的界限。在亚当·斯密看来,在实际的生产和生活当中,人与人的交往,往往在目的和手段之间不断变换身份。当人的生产不进行任何交换,只是自己自足,则每个人都是将自己视为目的。但是,一旦出现了物物交换,出于“利己”的考虑,每个人都需要进行特定的生产,以满足自己和家庭的需要,由此产生的结果是“利他”的。在进行产品交换的时候,自己的产品提供给他人,是将自己视为他人的手段,而将他人物品收入囊中的时候,则是将他人视为自己的手段。“在某种意义上,市场经济乃至整个市民社会的运作机制就是对人同时把自己和他人视为手段的普遍观念的认可。”<sup>[27]</sup>

而功利主义最大的疑问点在于:首先,其无法契合我们的伦理感官。即便假设其立场是正确的,只将生命的数量作为衡量的基准,就会出现“医生将一个健康的人杀害,将脏器取出、移植来救助 5 个人”的案件也是正当的,但是这显然和我们的伦理感觉相违背,难以被接受。有学者指出,即便在“电车难题”中,为了救助五个人而杀死一个人这样“极度现实的选择”是“现实中确定唯一的方法”,但是,多少也要考虑那个被杀的人的个体需要被尊重这样问题。<sup>[28]</sup>其次,功利主义为纳粹人种优劣论提供了理论支撑。“边沁思想的

[23] 参见湛洪果:《法律实证主义的功利主义自由观:从边沁到哈特》,《法律科学》2006 年第 4 期,第 18-19 页。

[24] 姚大志:《当代功利主义哲学》,《世界哲学》2012 年第 2 期,第 51 页。

[25] 刘雪梅、顾肃:《功利主义的理论优势及其在当代的新发展》,《学术月刊》2007 年第 8 期,第 45 页。

[26] 参见[日]平野龍一:《刑法の基礎正当防衛と緊急避難》,《法学セミナー》1969 年第 133 号,第 17 页。

[27] 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 25-28 页。

[28] 参见[日]一ノ瀬正樹著:《功利主義と分析哲学——経験論哲学入門》,放送大学教育振興会 2010 年版,第 105 页。

最大问题在于,将人类的本性视为快乐的函数。”<sup>[29]</sup>而在上述的讨论中,只涉及快乐的量的问题,而没有谈到质的问题。但是,边沁曾认为一种儿童玩的弹球游戏,“如果除去偏见,这种游戏拥有与艺术、音乐、诗学等同样的价值”。<sup>[30]</sup>密尔将边沁的这句话予以放大,认为对于边沁来说,“这种游戏和诗歌一样伟大”<sup>[31]</sup>对于密尔来说,具有决定性的是快乐的质。与不思进取的蠢货相比,不满足的苏格拉底更伟大。但是,一旦引入这种质的观点,功利计算就会变得非常复杂,而且一旦考虑到生命的质,则会为纳粹的“无生存价值的生命”这样的恐怖思想提供理论注脚。

由此,哲学上的生命衡量问题就此陷入了理论僵局。显然,不论是“人是目的”还是“最大多数人分享最大幸福”的观点都无法很好地回应“电车难题”,对于自动汽车程序设计中涉及的“电车难题”也无法进行有效的指导。那么我们将视角转换到作为具体部门法的刑法,来探讨是否存在解决问题的可能性。

### 三 审视刑法传统解决之道:陷入紧急避险的函矢相攻

一直以来,“电车难题”虽引起了学术界的研究兴趣,但是,由于这种情景发生的概率较低,所以对司法实践的意义也很有限。但是随着自动汽车的普及,“电车难题”发生的概率将变大,讨论该问题的重要性也大幅增加,域外学者已经就此展开先行研究。<sup>[32]</sup>这是因为,在“电车难题”中应该采取何种行动而不承担法律责任,在自动汽车的开发阶段,已不再是单纯的思想性实验。自动汽车的设计制造者,必须要清晰地确定事故发生时汽车的行动模式。根据情况的不同,处于“电车难题”状况的自动汽车,必须要作出行动选择。因此,该选择是否适当、是否具有刑法理论支撑,就有了研究的意义。

#### (一)人类驾驶汽车选择侵害第三者生命法益可能阻却责任

在研究自动汽车的设计制造者如何进行程序设计,才能免受刑罚处罚之前,让我们先来讨论,如果是驾驶员在驾驶车辆的时候遭遇“电车难题”时,选择侵害第三人的生命法益,挽救自己的生命法益该如何评价的问题。在“电车难题”中,为了救助一方的法益,不得不牺牲他人法益的时候,在刑法中能够想到的违法或者责任阻却事由就是紧急避险。

首先,关于紧急避险正当化的根据,支配性的观点是采取优越的利益说,即紧急避险“之所以阻却违法,是法益衡量的结果”。<sup>[33]</sup>而优越利益说的前提是社会连带性说。<sup>[34]</sup>根据这一学说之观点,法益的牺牲是附加给第三人的义务,这种附加给他人义务的正当化

[29] [日]小野纪明著:《西洋政治思想史講義——精神史的考察》,岩波書店2015年版,第236页。

[30] Jeremy Bentham, *The Rationale of Reward* (1830), John and H. L. Hunt, p. 206.

[31] John Stuart Mill, “Bentham”, in *Dissertations and Discussions*, Vol. I (1859), p. 389.

[32] 参见[日]平野晋:《『ロボット法』と自動運転の『派生型トロッコ問題』》,《NBL》2016年第1083号,第37页; [日]久保木水生等:《ロボットからの倫理学入門》,名古屋大学出版会2017年版,第37页。Vgl. Volker M. Jänich, Paul T. Schrader, Vivian Reck, *Rechtsprobleme des autonomen Fahrens*, NZV 2015, S. 313f.; Jan-Philipp Günther, *Roboter und rechtliche Verantwortung*, 2016, S. 45 f.

[33] 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。

[34] 参见[日]橋田久:《緊急避難に対する緊急避難》,《法政論集》2014年第256号,第65页。

理由是,基于相互性保证同一性规则的构成要素。社会的所有人,必须以自己都会陷入紧急避险为前提。陷入这种状况中以牺牲他人来避免危险的妥当性,是以承认自己在紧急避险状况中,也会为他人而牺牲为基础。换言之,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”,这样一种连带性的社会义务。正因如此,紧急避险中对他人附加牺牲的义务才具有可能性。当然,并不是任何场合都会产生连带义务,归根结底,只限于保全之利益要显著优越于侵害的利益,才能附加给他人甘受侵害的义务。驾驶员在驾驶机动车过程中,为了避免对自己或者他人遭受的现实的危险,而侵害第三人法益的时候,只要是所保全的法益显著优越于被侵害的法益,就可以认定构成正当化的紧急避险。比如,在驾驶机动车中突然出现障碍物,如果直接撞上该障碍物可能导致驾驶员死亡。此时,为了避免这样的危险而转动方向盘,撞上停在路边第三人的汽车,造成该车损毁的案件。因为所保全的利益是人的生命,远远高于损毁的汽车这一财产性利益,所以,可以认为成立正当化紧急避险。

相反,如果是因回避措施而导致第三者死亡或者重大伤害的情形,则无法认定保全利益具有显著优越性。根据支配性的观点,生命法益是不容许进行比较衡量的。<sup>[35]</sup> 即便是为了营救多数人而牺牲一个人,也不能认定具有正当性。<sup>[36]</sup> 那么,遇到这种情形只能考虑免责事由。

由于主流观点认为,我国刑法中的“紧急避险的本质,应当和正当防卫一样,从该行为不具有客观的社会危害性即不具有违法性的角度来理解,而不是从行为人主观上‘不足深责’的角度来考虑”。<sup>[37]</sup> 所以,探讨免责的紧急避险只能参考德国的刑法规定。<sup>[38]</sup> 免责的紧急避险可以阻却责任的理由是期待可能性理论,<sup>[39]</sup> 在正当化紧急避险要件不充分的情形中,每个人保持自己现存利益,只剩最后一个方法的时候,承认这可以作为免责的根据。人人必须相互地承认各自所持有的现存利益,对于每个个体来说,正是这种相互承认保障了重要法益。为了保持这种法益而侵害他人法益的行为,因为不能被认定为是正当化的紧急避险,即属于违法的。但是,保持自己的现存利益是大家相互承认的,以此为基础,这一利益只要不劣于被侵害者的利益,即可免责。<sup>[40]</sup> 基于这一点,保护现存法益的侵害行为和一般侵害行为之间具有本质的不同,这也是免责的根据所在。免责的紧急避险限定于自身、亲属或者其他关系亲密者,原因在于具有亲密关系的人的法益保护被涵盖在了自己现存利益的保护范围之内,而不是所谓的代理的利益保护问题。当现实存在威胁驾驶员生命的危险,且只能通过撞死行人才能予以回避。即便考量比例原则,行为者

[35] Perron, in: Schönke/Schröder, Strafgesetzbuch. Kommentar, 29. Aufl. 2014, § 34 Rn. 23.

[36] Vgl. BVerfGE 115, 118 (151 ff.); BGHSt 35, 347 (350); Neumann, in: Kindhäuser/Neumann/Paeffgen (Hrsg.), Nomos Kommentar, Strafgesetzbuch, Bd. 1, 4. Aufl. 2013, § 34 Rn. 74; Wessels/Beulke/Satzger, Strafrecht, Allgemeiner Teil, 45. Aufl. 2015, Rn. 467.

[37] 黎宏:《紧急避险法律性质研究》,《清华法学》2007年第1期,第42页;刘明祥:《论紧急避险的性质》,《法学研究》1997年第4期,第93页;张明楷著:《刑法学》(第5版),法律出版社2016年版,第217-218页。

[38] 日本刑法中的紧急避险也只是作为违法阻却事由,没有规定责任阻却的紧急避险,因此本文对于免责的紧急避险的探讨,是以《德国刑法》第35条规定为参考。

[39] 参见[日]高桥则夫著:《刑法总论》(第三版),成文堂2016年版,第309页。

[40] Vgl. Engländer, in: Matt/Renzikowski (Hrsg.), Strafgesetzbuch, Kommentar, 2013, § 35 Rn. 19 f.

确定会出现死亡结果的情况下,那么致死性紧急避险行为就有必要。

《德国刑法》第35条第1项第2款规定,对于自招危险的情况,限制免责的紧急避险的成立。那么,驾驶人员选择驾驶汽车是否属于自招危险就是需要回答的先决问题。要想认定驾驶员引起了危险,则必须要说明该危险的惹起有非难的必要性。在遵守交通规则的前提下,驾驶员驾驶汽车的行为,不能被视为是具有非难可能性的自招危险行为。交通的危险是社会所能够接受的,对于道路交通的参与者而言,这种危险不应该成为刑法的责任要素。因此,驾驶员选择驾驶汽车的行为不应属于免责紧急避险的限制适用范围。那么,根据《德国刑法》第35条第1项第1款的规定,对生命以及身体的危险,对于行为者自身或者亲属、关系亲密者来说迫在眉睫,就可以考虑通过紧急避险来免责。

因此,人驾驶车辆遭遇“电车难题”时,选择侵害第三人的生命法益,在理论上,可以根据免责的紧急避险阻却责任。

## (二) 程序设计选择侵害第三者生命法益无法阻却责任

那么,在自动汽车的情形中,应如何评价设计制造者的程序设计行为呢?汽车设计制造者提前通过程序设计决定当自动汽车遭遇“电车难题”时,选择牺牲第三人(其他交通参与者)的生命法益,这种情况是否也可能成立紧急避险呢?

如上所述,驾驶员驾驶汽车时,为了保护自己的生命而导致第三人死亡的情况,无法认定为正当化的紧急避险。那么,这就意味着刑法在违法性阶段附加给设计制造者一项义务,即不能为了自动汽车的使用者,而在车辆上设计牺牲第三人生命的程序。然而,有学者批判这种观点,认为这种附加的义务和自然法则上的人类形象不一致。<sup>[41]</sup>这是因为,附加这一义务意味着间接强制车辆利用者在遭遇“电车难题”时要牺牲自己的现实性利益。但是,这一批判性观点又忽略了不得牺牲他人生命来保全自己生命的合法性秩序要求。因此,设计制造者如果为了使处于“电车难题”状况中的自动汽车,因保护使用者的生命法益,而侵害第三人的生命法益,只能被宽恕而不能免除处罚。因为其行为同样是违法的,只能继续考察是否可以通过免责的紧急避险予以免责。

此时,首先需要讨论,自动汽车的设计制造者提前进行程序设计选择侵害第三者的生命法益,是否满足危险的现实性的要求?要满足免责的紧急避险的要求,首先需要符合危险的现实性。在“电车难题”发生时,生命、身体及其他法益存在危险,这一点无可争议。而和本要件相关的问题是,危险的现实性可以被认定吗?在程序设计的时候,还没有产生具体状况中的法益危险。这种情况下,免责的紧急避险的成立就遇到了先决性问题。

关于现实性要件,很多观点认为可以和正当防卫中的急迫性采取同样的理解,这种解释要求法益侵害危险迫在眉睫。<sup>[42]</sup>危险的现实性中,是否要求时间的迫切性,在学理上是存在争议的。<sup>[43]</sup>围绕时间迫切性的争议焦点在于:危险尚不存在,但如果

[41] Philip Weber, Dilemmasituationen beim autonomen Fahren, NZV 2016, S. 253 f.

[42] 参见[日]大塚仁ほか編:《大コンメンタール刑法2巻》(第3版),青林書院2016年版,第689页;[日]前田雅英ほか編:《条解刑法》(第3版),弘文堂2013年版,第122页。

[43] 参见[日]遠藤聡太:《緊急避難論の再検討(2)》,《法学協会雑誌》第131卷第2号,第450-491页;[日]小林憲太郎:《刑法判例と実務11回緊急避難》,《判例時報》第2305号,第8页。

即开始侵害他人法益的避险行为,结果是否无法避免。对此,即便是将来可能发生的危险,只要满足现阶段不介入就会错过时机这一条件,还是有很多判例<sup>[44]</sup>和学说<sup>[45]</sup>肯定危险的“现实性”的。即便是采取形式时间迫切性的观点,需要限定的只是危险发生以前,实施对他人法益产生实际侵害的行为。但是,“电车难题”状况中的避险行为是程序设计行为,这种程序设计行为时,并没有对他人的法益产生现实的侵害。归根到底,在陷入“电车难题”状况时,只不过是牺牲他人的可能性,将他人法益置于抽象危险化的问题。

即便是在正当防卫中,诸如事前在墙头安装防小偷的碎玻璃片等防卫装置的行为,该行为的防卫效果只是在实际侵害迫切发生的时点才产生,但是也不能因此否定其迫切性。<sup>[46]</sup> 如果将现实性和迫切性做同等平行考量的话,紧急避险也可以这样来对待。因此,为了对应所生的危险,实施程序设计的行为,不能否定危险的现实性。

但是,车辆利用者对于自动汽车的设计制造者来说,并不是亲属或者其他关系密切者,因为不满足免责的紧急避险中人的关系性要件,所以我们可以期待设计制造者不通过程序设计来侵害第三人的生命法益,因此设计制造者也就无法通过免责的紧急避险来阻却责任。

### (三) 程序设计选择侵害车内人员生命法益无法阻却违法

自动汽车的设计制造者,既然无法通过程序设计来侵害第三人生命法益,保护驾驶人员生命法益来免责,反之,剩下一种选择则是,通过程序设计让自动汽车在遭遇“电车难题”时,选择侵害车内人员(包括驾驶员和同乘人员)的生命法益。暂且不论这样的自动汽车是否有销售市场,<sup>[47]</sup> 先从刑法的角度分析是否可行。那么,将程序设计为牺牲车辆利用者及车内其他乘客的利益该如何评价? 比如,自动汽车驾驶过程中,突然出现障碍物,回避可能会碾压附近的行人,所以程序设计为直接撞上障碍物,导致车内人员死亡。这里就可能要考察设计制造者是否成立作为犯<sup>[48]</sup>的可罚性问题。

有学者认为,只要自动汽车的设计制造者已经向消费者履行了告知义务——将车辆遭遇“电车难题”状况会进行何种选择的信息告知给汽车的使用者,那么自动汽车的使用者在真的遭遇“电车难题”状况之前,一直掌握着关于自动汽车程序选择的信息,或者说,已经认识到了自动汽车选择的方向性。那么,当出现“电车难题”状况,发生的构成要件结果就不能归属到设计制造者身上。这是因为自动汽车的使用者可以被认定为自我归责中的自陷危险。不仅认识到了风险状况,而且启动车辆的判断也归属于利用者的话,危险化支配也能够被肯定。而对于车内其他乘客的侵害结果,基本上可根据同样的原理推导,结论也是妥当的。车辆利用者在认识到风险状况的情况下,就是有合意的其他人的

[44] RGSt 60, 318.

[45] Schönke/Schröder/Perron, Strafgesetzbuch Kommentar 29. Aufl. (2014). § 34 Rn. 17.

[46] 参见[日]川端博(ほか)編:《裁判例コンメンタール刑法1卷》,立花書房2006年版,第240页。

[47] Vgl. Eric Hilgendorf, Können Roboter schuldhaft handeln?: Zur Übertragbarkeit unseres normativen Grundvokabulars auf Maschinen, in: Susanne Beck (Hrsg.), Jenseits von Mensch und Maschine, 2012, S. 119 f.

[48] Freund, in: Joecks/Miebach (Hrsg.), Münchener Kommentar zum Strafgesetzbuch, Bd. 1, 2. Aufl. 2011, § 13 Rn. 8 f.

自陷风险的问题,可以通过被害人同意来否定违法性。<sup>[49]</sup>

“明知自己实施或参与他人某种行为存在风险,但仍然自愿实施或参与,对于最终出现的结果如何归责,这个问题在刑法理论上称之为被害人自陷风险(或者危险接受)。”在被害人自陷风险的场合中,如何排除行为人行为的违法性?被害人同意理论是最具针对性的,可作为被害人自陷风险案件中违法阻却事由。当然,在国内外刑法学界,既有很多学者支持将被害人同意理论适用于被害人自陷风险的场合,也有不少学者反对。本文限于篇幅,就两者的争议暂且不再赘述。国外刑法理论的主流观点和司法判例对此持肯定的态度,<sup>[50]</sup>本文暂以主流观点作为后续讨论的前提。

那么在自动汽车领域,设计制造者已经告知使用者,自动汽车在遭遇“电车难题”时,会造成车内人员重大身体法益甚至生命法益的损害,使用者表示同意接受该风险,并现实发生了法益侵害的结果。这种情况下,可以认定使用者属于自陷风险。那么,对于设计制造者来说,可以通过被害人同意理论排除违法性吗?

一方面,尽管被害人同意可以作为被害人自陷风险的出罪理由,但是并非没有限制。被害人同意需要在同意范围、同意能力、同意对象、同意的现实性等方面同时满足相应的条件方可得作为违法阻却事由。在同意的范围中,即使是同意侵害自己的法益,也必须受一定范围的限制。比如,重大身体法益和生命法益受他人故意侵害,原则上无法通过被害人同意阻却违法。<sup>[51]</sup>比如相约自杀中,由其中一方杀死另一方,继而自杀未得逞的。尽管未死亡的一方是在死者一方同意的前提下实施的杀害行为,但仍然要对杀死对方的行为承担故意杀人罪的刑事责任。<sup>[52]</sup>

另一方面,需要讨论设计制造者对于使用者最终的死亡是持故意还是过失的态度?是否该当故意杀人罪的构成要件?

在故意和过失之间,间接故意和过于自信的过失密切关联,两者主要的区别在于行为人对法益的不同态度,前者是一种积极的蔑视,后者则是消极的不保护。其判断的依据是,行为人是否有避免法益侵害结果发生的能力。<sup>[53]</sup>“过于自信的过失是一种有认识的过失,即应当避免而没有避免。应当避免是避免义务与避免能力的统一。虽有避免义务,但没有避免能力,仍属于缺乏应当避免这一要件。在过于自信的过失中,行为人认为凭借自己熟练的技术、敏捷的动作、高超的技能、丰富的经验、有效的防范,完全可以避免发生危害结果,但实际上过高地估计了自己的力量,因而未能防止危害结果的发生。”<sup>[54]</sup>如果行为人既没有经过专业培训,也没有通过实践积累长期丰富的经验,则行为不仅完全丧失

[49] BGHSt 53, 55(60 f.); Rengier, Strafrecht, Allgemeiner Teil, 8. Aufl. 2016, § 13 Rn. 81 f.; Armin Engländer, Das selbstfahrende Kraftfahrzeug und die Bewältigung dilemmatischer Situationen, ZIS 2016, S. 608 f.

[50] 参见车浩:《过失犯中的被害人同意与被害人自陷风险》,《政治与法律》2014年第5期,第27-36页;江溯:《过失犯中被害人自陷风险的体系性位置——以德国刑法判例为线索的考察》,《北大法律评论》2013年第1期,第115-142页。

[51] 参见[日]前田雅英著:《刑法总论讲义》,东京大学出版会2015年版,第244-245页。

[52] 参见张明楷著:《刑法学》(第5版),法律出版社2016年版,第849-850页。

[53] 参见付玉明:《犯罪故意的事实认识与内容解读》,《中国刑事法杂志》2016年第6期,第81-95页。

[54] 孙伟铭以危险方法危害公共安全案二审判决书,四川省高级人民法院(2009)川刑终字第690号。

对危害的有效防范,而且大大降低其驾驭风险的能力,造成重大损害结果的发生是必然的,主观心理状态上不是过于自信的过失而是间接故意。比如,德国的一起典型案例中,行为人打赌如果能够用手枪击中被害人手中所持有的玻璃球,即可获得赌金,行为人因射术不精,射伤了被害人,最终行为人被认定为故意伤害罪。<sup>[55]</sup>

那么,在设计制造者设定的程序是当车辆遭遇“电车难题”状况时,选择让驾驶人员死亡是确定的结果,而且设计制造者对于驾驶人员一方死亡的结果是有认识的。虽然其对于该结果并不积极追求,但是没有能力予以避免,不属于过于自信的过失,对于驾驶人员的死亡应该是间接故意的态度。因此,设计制造者的行为在外观上满足了杀人罪的构成要件。既然如此,则无法通过被害人同意理论排除违法。

由此可见,作为自动汽车的设计制造者,对汽车陷入“电车难题”时,通过程序设定牺牲第三人生命法益来保全驾驶人员的生命法益,无法通过紧急避险予以正当化,即便是引入了免责的紧急避险,也难以满足免责的紧急避险中人的关系性这一要素。而通过程序设定直接牺牲驾驶人员的生命法益,由于故意让他人侵害自己的生命法益不属于被害人同意的范围,故无法脱逸违法性。因此,在当前的规范体系下,无论设计制造者如何进行程序设计,都无法通过阻却违法和责任来排除犯罪的成立。另外,如果考虑到自动汽车的市场需求,自动汽车的程序设定应当趋向于保护驾驶人员一方的法益。<sup>[56]</sup>那么,后续的讨论是以处于“电车难题”状况中自动汽车的程序设定,选择侵害第三方行人的生命法益为前提。即讨论此种程序设计是否存在让设计制造者出罪的可能。

#### 四 绕开生命权衡的理论路径:被允许的危險理论排除犯罪性

从上述讨论可知,无论是在违法还是在责任判断阶段,都难以绕开具体的生命权被衡量这一制约。虽然在哲学层面,康德的“人是目的”学说与边沁的“最大多数人分享最大幸福”学说针锋相对、势均力敌,但是在实际的规范层面,无论是宪法还是部门法,都重视保障具体个体的生命权益,通过具体的生命法益衡量上的优越或者平衡,达到阻却行为违法或者责任,难以跨越规范性障碍。因此,跳出违法和责任判断的传统思维,提前从构成要件该当性的判断阶段进行思考,摆脱具体的生命法益衡量与平衡的困扰,或许能够为自动汽车的程序设计寻找到适法的路径。

##### (一) 只有被允许的危險可能排除对设计行为的归责

在客观归责理论中,以条件因果关系为前提,将没有通过客观归责检证的行为排除出归责的范围,从而能够将行为在构成要件该当性阶段予以出罪。那么,在自动汽车的程序设计下,“电车难题”状况中,为了保护车内人员的生命法益,自动汽车选择侵害第三方行人的生命法益,是否能够通过客观归责的判断,直接摆脱入罪的可能性呢?

[55] 参见[日]井田良著:《刑法總論の理論構造》,成文堂2008年版,第75页。

[56] 参见[日]小林正啓:《自動運転車の実現に向けた法制度上の課題》,《情報管理》2017年第4号,第240-250页;Eric Hilgendorf, Recht und autonome Maschinen-ein Problemaufriß, in: Eric Hilgendorf/Sven Hötitzsch (Hrsg.), Das Recht vor den Herausforderungen der modernen Technik, 2015, S. 11 f.

客观归责遵循行为是否制造了不被允许的危险、行为是否实现了不被允许的危险以及结果是否超出构成要件的保护范围三个判断要素,最终得出结果是否可以归责于行为人。其中,行为是否制造了不被允许的危险是首要且基础性判断要素。要回答行为是否制造了不被允许的危险,基于客观归责的属性,学界无法给出正面的回答,而主要是从反面回答什么不是不被允许的危险。<sup>[57]</sup> 罗克辛教授将其概括为四类:一是,使得危险降低的行为;二是,没有创设出危险的行为;三是,假定因果关系的行为不排除归责;四是,创设被允许危险的行为。<sup>[58]</sup> 其中第一、第二和第四点是从反面来判断,即不属于创设不被允许的危险,而第三则是从正面判断,认为如果是基于假定因果关系,行为对造成的结果仍然应该被归责。下面将从上述四个要素出发,分别检视它们能否让设计制造者的行为被排除出不被允许的危险的范畴。

如前所述,自动汽车的普及能够大幅度地减少交通事故,从整体社会效果来看,将给社会带来巨大的利益,降低交通事故的死亡人数,而其中程序设定所导致的行人生命法益的侵害只是个案。那么,是否通过法益衡量就可以得出自动汽车中的程序设计行为属于使危险降低的行为的结论呢?显然答案是否定的。首先,这是将自动汽车整体和程序设计的局部之间予以混淆,降低交通事故的死亡人数是自动汽车整体所致,而个别行人生命的受损是汽车遭遇“电车难题”的程序设计所致,二者不可同时评价。其次,所谓使危险降低应该是在具体个案中进行判断,而不是从整体、抽象的层面进行利益对比。从个案中进行判断,自动汽车在具体“电车难题”中所做的选择,并没有降低周边行人法益的危险,而是升高了危险。所以,“使得危险降低的行为”这一判断要素并不能帮助设计制造者出罪。

典型的没有创设危险的行为是,如不想支付抚养费的父亲,想让自己四岁的孩子被流星击中死亡,于是将其带到草原,假如真的被流星击中,也不存在杀害行为。之所以将这种行为称之为没有创设危险,是因为行为人的劝说属于日常生活行为,按照人们的通常认识,不存在对法益造成危险的可能性。但是,如果行为人对此有特别认知的话,如父亲是一个天文学家,对于流星的运行了如指掌则另当别论。<sup>[59]</sup> 此时,刑法就需要对具体行为人提出高于一般人的行为要求,以使其尽量控制行为的危险。作为自动汽车的设计制造者,对于汽车的程序设计、驾驶过程、最终的结果有着清晰且专业的认知,知道一旦汽车处于“电车难题”状况,将一定会侵害到周围第三方行人的生命法益,因此,设计制造者的行为并非是没有创设危险的行为,而是创设了法益侵害的危险。

那么,设计制造者是否能够凭借以下理由来脱逸出不被允许的危险范围呢?即便设计制造者不通过程序设计,让处于“电车难题”状况中的自动汽车,选择侵害第三方行人的生命法益,在驾驶员驾驶并操作汽车,遭遇“电车难题”状况时,一般也会选择保护自己

[57] 参见周光权:《行为无价值论与客观归责理论》,《清华法学》2015年第1期,第130-144页。

[58] 参见[德]ロクシン著:《刑法総論第一卷・第一分册「犯罪論の構造」》,山中敬一监译,信山社2003年版,第401-412页。

[59] Jan C. Joerden, Strafrechtliche Perspektiven der Robotik, in: Eric Hilgendorf/Jan-Philipp Günther (Hrsg.), Robotik und Gesetzgebung (Robotik und Recht Bd. 2), 2013, S. 195 f.

的生命,侵害行人的生命。因此,即便没有自动汽车的设计制造人的程序设计,行人生命受到侵害的结果也仍然会由其他人或事件引发。当然,这一观点是无法成立的,这正是假定因果关系不排除归责的典型。这是因为“刑法需要处理也有能力处理的,只能是合法或不法,而不是幸运或不幸”。并且,“归责的判断对象,是行为与特定的不法构成要件结果之间的关系,而不是行为与某个抽象性结果之间的关系”。<sup>[60]</sup>

综上,无论是通过使得危险降低的行为、没有创设出危险的行为,还是假定因果关系的行为不排除归责等要素,都无法使得设计制造者的上述程序设计行为,被认定为不属于“制造不被允许的危险”的行为。以下本文将讨论设计制造者的程序设计行为是否可以被认定为“创设被允许危险的行为”?

## (二)从责任论、违法论到构成要件论中被允许的危险法理

所谓被允许的危险是指,如果只要某种行为具有危险,就以此为由一律禁止,那么对社会有用性的行为或者国民的行动自由就会遭受严重的桎梏。因此,在一定范围内,也即危险行为所产生的社会有用性或者行动自由等利益明显超过负面恶害,可以容许该危险的存在。<sup>[61]</sup> 比如汽车驾驶会给行人的生命或者身体带来一定的危险,如果因此就一律禁止,会显著阻碍人们的出行自由,而且会导致物流产业受到严重冲击。为此我们制定了交通规则,让汽车驾驶者遵守该规则,这使得将危险抑制到最低限度和多少因此受到限制的出行自由和物流产业实现一种平衡。也即,最终达成一个结论——只要遵守交通规则,就可以容许驾驶汽车。

最早提出被允许的危险概念的是德国学者巴尔,随着 19 世纪后半叶工业技术的迅猛发展,“德国的矿山、煤气、铁道企业发展迅速”,“社会生活中不可避免地出现一定的危险行为。如果为了完全防止实害发生而要求采取过度的预防措施,就排除了所有企业活动的可能性。因此,即使行为人懈怠了过度的预防措施,也不成立犯罪”。<sup>[62]</sup> 巴尔并没有将被允许的危险概念予以理论化,仅仅处于构思阶段,虽然提出了利益衡量后的行为和日常行为视为被允许的危险,但是并没有进一步细化。需要说明的是,在当时构成要件论、违法论和责任论还没有完全被区分的时期,巴尔的观点相当于将被允许的危险放在了违法性判断阶段。<sup>[63]</sup>

随后有学者将过失犯中的责任和被允许的危险的法理进行联系展开。其认为“犯罪论当中,只有在过失犯才有的,遵守社会相当性的危险而阻却刑罚的体系地位,既不是构成要件该当性也不是违法性,而是责任”。因此,“和被要求的注意义务相一致的,遵守社会相当性的危险,在发生违法结果的时候,并不是什么违法阻却事由,而是必要的责任阻却事由”。而且认为这种结论和我们的既有认识是一致的,因为惹起违法结果的危险行

[60] 车浩:《假定因果关系、结果避免可能性与客观归责》,《法学研究》2009 年第 5 期,第 147 页。

[61] 参见[日]中山研一(ほか)編:《現代刑法講座》(第三卷),成文堂 1979 年版,第 25 页;[日]前田雅英著:《可罰的違法性論の研究》,東京大学出版会 1982 年版,第 219 页。

[62] 张明楷:《论被允许的危险的法理》,《中国社会科学》2012 年第 2 期,第 113 页。

[63] Vgl. Roeder, Die Einhaltung des sozialadäquaten Risikos und ihr systematischer Standort im Verbrechenaufbau, 1969, S. 28 f.

为在事前就是违法的,这是我们所拥有的存在论的认识中,就已经确立的。<sup>[64]</sup>

将被允许的危险和责任论联系的观点并没有成为支配性学说,经过之后一些代表性学者的推进,将被允许的危险从责任问题又重新推回到违法性阶段的问题。<sup>[65]</sup>

原捷克布拉格查理大学教授米利克将巴尔提出的被允许的危险略加精致化展开,在讨论危险范围的时候,认为超过一定基准的——不被允许的、非一般的、异常性的危险才能被认为是对刑法而言“重要的危险”。并且,将危险和有用性的衡量方法设定了一定精度的判断基准,包括危险的程度、被置于危险中法益价值的高度、应生侵害范围的广度以及侵害的可能性等。在此基础上,进一步判断危险是否超过了行为目的和社会适合性的程度。<sup>[66]</sup>

而奥地利学者艾克斯勒主张对被允许的危险持高度亲密性态度,他对“感情责任论”展开讨论,认为虽然过失责任的本质是行为人因注意力不集中而实施侵害行为,追究的是对法益义务违反的轻视态度。但是,即便能够预见结果,行为人只要遵守了“社会生活上必要的注意”,就可以免除可罚性。<sup>[67]</sup> 其中包括:因为结果在法律当中是偶然的,所以欠缺责任的情形,以及行为所伴随的一定危险被所允许,故欠缺违法的情形。而后者则包括合规地进行危险的手术却失败的案件、进行危险操作的工厂的所有者,尽管履行了所有必要的安全措施,仍然发生事故的案件等。之所以允许这些危险,并让社会承受这些危险,其根据是行为所产生的社会有用性。

德国著名法学家宾丁则在艾克斯勒的基础上,提倡“适度的危险”的概念,并通过这一概念来限制过失犯的成立。也即,危险的被允许性要通过从法的角度考察伴随着危险的行为的避免可能性和避免不可能性来确定。而这又需要通过危险和不可或缺的行为的法律价值之间的比例关系进行具体判断:第一,如果从法律的角度来看,行为达到了不可或缺的程度,即便是有较大的危险,法律也需要予以允许。这包括:(1)负有行为义务的人,要冒着义务履行的必要危险;(2)在实施法律所允许的行为时,不得不冒不可避免的危险,原则上也是正当的;(3)行为者在权利行使受到妨碍时,被允许的危险就越大;(4)让法益危险化是法益保全的唯一手段时,一般是被允许的危险。第二,如果从法律的角度来看,行为并不是必须的话,此时需要承受的被允许的危险就会小。其中包括:(1)行为故意违反法律;(2)完全无益的行为。第三,所有的危险只有对行为来说具有必要性,才有可能被正当化。<sup>[68]</sup>

德国学者恩吉施虽立足于消极构成要件要素论,但是将被允许的危险理解为违法阻却事由,讨论其与“外部注意”之间的关系。其针对将过失仅仅作为责任问题的观点——“行为者缺乏对于犯罪事实的认识是基于不注意”(内部注意),提出应该是作为

[64] Vgl. Roeder, Die Einhaltung des sozialadäquaten Risikos und ihr systematischer Standort im Verbrechenaufbau, 1969, S. 94 f.

[65] 参见[日]花井哲也著:《过失犯の基本構造》,信山社1993年版,第264-265页。

[66] Vgl. Miricka, Die Formen der Strafschuld und ihre gesetzliche Regelung, 1903, S. 140 f.

[67] Vgl. Franz Exner, Das Wesen der Fahrlässigkeit. Eine strafrechtliche Untersuchung, 1910, S. 192 f.

[68] Vgl. Karl Binding, Die Normen und ihre Übertretung. Eine Untersuchung über die rechtmässige Handlung und die Arten des Delikts, 4. Bd. Die Fahrlässigkeit, 1919, S. 432 f.

结果回避所必要的外部态度(外部注意)问题。所以将被允许的危险作为违法性的要素来理解。并且,将外部注意的内容区分为“不实施危险行为的注意”、“作为危险状况中外部行为的注意”和“作为履行法遵守义务的注意”三种,<sup>[69]</sup>其中主要就前两者中的被允许的危险的理论进行展开。

综上可知,从巴尔到恩吉施,都是在工业技术发展的背景下,为防止刑法的介入限制经济技术的发展,而将被允许的危险作为限定过失犯成立的重要理论手段。并且被允许的危险是以社会整体价值衡量为基础,被作为和正当化事由一样的违法阻却事由,即放在违法性阶段进行判断。

而“持法规范违反说的行为无价值论者认为,行为人实施被允许的危险行为,或者行为实现的是被允许的危险时,该行为就不符合构成要件;而行为是否被允许,则取决于行为人是否遵守了行为规则”。<sup>[70]</sup>罗克辛教授认为,所谓被允许的危险,虽然是创设了法律中重要危险的行为,但是,这在一般场合说(并在非具体场合)是被允许的。因此,其不同于正当化事由,而应该被作为阻却客观构成要件归责的行为来理解。被允许的危险的典型例子是所有的遵守道路交通规则的汽车。不可否认,汽车对于生命、健康以及财物来说都具有重大的危险,通过事故数据的统计,可以得出正面的支持。尽管如此,由于存在压倒性的公共福祉、利益的需求,立法者通过制定一定的注意性规定,来允许汽车驾驶行为。这显然是不同于诸如紧急避险那样的正当化事由。在此,并不需要进行具体的利益衡量,而是进行概括性衡量,目的就是为了排除犯罪类型的成立。因此,即便是在具体的场合中,(驾驶汽车)并不是为了追求更高的利益,甚至是为实施不值得鼓励的目的(比如犯罪预备),都不妨碍允许汽车驾驶行为的合法性,显然,这不符合基于正当化事由的具体利益衡量。

另外,考虑到即便遵守了所有的交通规则仍然会产生惹起法益侵害的行为,而这是不符合构成要件的行为。所以,被允许的危险理论既可以用于过失犯罪,也可以用于故意犯罪。比如“B 尽管遵守了所有交通注意义务,仍然(意识到但不得不)撞到 A,导致 A 当场死亡。在这起确实是由 B(故意)引起了 A 的死亡”<sup>[71]</sup>的案件中,就不能认为 B 构成了杀人罪。也即,只有超出了被允许的危险,该实现行为才能被视为符合构成要件行为。<sup>[72]</sup>

### (三)程序设计行为可以作为被允许的危险

至此,被允许的危险已经成为可以同时用于故意犯罪和过失犯罪,并且在构成要件该当性阶段排除归责的理论,而这一理论的前提则是:(1)概括性的利益衡量;(2)有立法设定的注意义务。只有符合概括性的利益衡量,立法才有必要为之设置一定的注意义务,而立法设置注意义务是为了防止以概括性利益优越之名,滥用被允许的危险。因此,满足概括性利益衡量且未超过法定的注意义务,才能被认为是被允许的危险。

那么什么是概括性的利益衡量,以汽车驾驶行为为例,其之所以能够成为被允许的危

[69] Engisch, Untersuchungen über Vorsatz und Fahrlässigkeit im Strafrecht, 1930, S. 286 f.

[70] 张明楷:《论被允许的危险的法理》,《中国社会科学》2012 年第 2 期,第 118 页。

[71] 在这种情况下,B 虽然不具有交通肇事罪的外观,但却表面符合故意杀人罪的外观。

[72] Vgl. Roxin, Strafrecht Allgemeiner Teil Band I, 4. Aufl., 2006, S. 381 - 382 f.

险,是因为其具备了普遍性、社会利益性和社会接受性三个特点。首先,被允许的危險的前提是具有社会普遍性,如果汽车驾驶行为只是个别人群的个别行为,则不能为了部分人的方便,而牺牲大部分人的安全利益。只有汽车驾驶行为成为社会的主流行为,才有将其纳入到社会整体利益中进行考量的必要。其次,被允许的危險的核心是满足社会利益性,如果汽车驾驶行为只具有普遍性,但是并没有给人类带来实质的利益,或者造成的侵害巨大,超过了所带来的利益,这就无法满足被允许的危險成立的核心要素。最后被允许的危險的试金石是社会接受度。由于概括性利益衡量是社会整体性衡量,利益接受和损害承担的主体都是社会全体,因此社会对于具有一定危險的行为是否能够接纳,是考察前两者是否成立的试金石。汽车驾驶行为仅仅被一部分人认为具有概括性的利益优越性,则不能使其成为被允许的危險,因此也就无法促使立法者为其进行立法,设定相应的注意义务。

接下来,我们将上述理论用于检视自动汽车中设计制造者的程序设计行为,看其是否能够成为被允许的危險,并在构成要件该当性中被排除归责。

基于进行概括性的利益衡量层面的判断。第一,要进行普遍性检测。当前,驾驶自动汽车还没有成为人类生活的主流行为,尚不具备普遍性。但是,目前日本、美国、德国以及其他欧洲诸国都在开展自动汽车的研制、生产,在我国,北京、上海、深圳、重庆等城市都在积极探索自动驾驶实践。<sup>[73]</sup> 在可预期的将来,自动汽车将会进入人类社会并普及。因此,普遍性要素并不会成为障碍。第二,要进行社会利益性检测。如前所述,作为人工智能的典型产业,自动汽车将人工智能快速与人类生产生活联系起来,并且将大量减少人为驾驶汽车可能带来的伤亡事故。而且,人工智能正在我们的日常生活中扮演着越来越重要的角色。<sup>[74]</sup> 有人断言,在不久的将来,人工智能将取代或加强各种人类职业。<sup>[75]</sup> 据估计,包括自动汽车在内的人工智能将为我们的公共领域、生活领域和产业领域带来巨大的便利。比如在公共领域,人工智能会给我们的公共基础设施建设带来便捷,帮助我们更好地进行数据收集、分析和预测,应对异常的气候和灾害。更好地帮助人类建设舒适、安全和效率的智慧城市系统。在生活领域,能够通过人工智能的使用,实现人类日常生活的辅助化,甚至与人类的自然对话。在产业领域,人工智能将会通过自动栽培、农业无人机等,大幅度提高农业生产效率和产量。与此同时,人工智能在制造业、建设业、金融业、物流运输业、医疗护理行业、教育研究以及服务业等方面的潜力,不一而足。<sup>[76]</sup> 2017年7月,国务院印发的《新一代人工智能发展规划》,人工智能技术的发展已经为党中央、国务院高度重视,获得了国家战略层面的大力支持。通过人工智能技术的深度参与和研究,我国将实

[73] 参见李洪兴:《自动驾驶如何“弯道超车”》,《人民日报》2018年1月8日第5版。

[74] 参见 Monika Hengstler et al., Applied Artificial Intelligence and Trust-The Case of Autonomous Vehicles and Medical Assistance Devices, 105 *Technological Forecasting & Social Change* 105, 114 (2016)。

[75] 参见 Karamjit S. Gill, Artificial Super Intelligence; Beyond Rhetoric, 31 *AI & SOCIETY* 137, 137 (2016)。

[76] 参见[日]西田豊明:《選択肢としての人工知能》,《情報管理》2017年第9号,第655-661页;日本経済産業政策局:《ビッグデータ?人工知能がもたらす経済社会の変革》,2015年4月21日;日本総務省:《平成28年版情報通信白書のポイント》,第232-241页;Sabine Gleß, Thomas Weigend, Intelligente Agenten und das Strafrecht, *ZStW* 126 (2014), S. 561 f。

现高新技术领域的弯道超车,领跑世界。<sup>[77]</sup>因此,自动汽车满足社会利益性要素。第三,要进行社会接受性的检测。由于自动汽车处于“电车难题”状况时,程序设定是保护车内人员而牺牲车外人员。这种对第三人生命的牺牲必须要得到社会的多数性认同。当然,这需要以社会实证调查作为数据支撑,当自动汽车正式进入产业化生产之前,必须要以社会接受性的调查作为前提。2016年6月24日的《科学》杂志刊发了法国图卢兹大学、美国俄勒冈大学和麻省理工学院的学者的联合研究成果,就自动汽车时代,当自动汽车遭遇电车难题时,是救乘客还是路人进行了实证调查。在接受调查的451人当中,其中有75%的人认为当车上只有一名乘客并且只有一名行人时,应该优先保护乘客。但是当行人的数量增加时,如有5名行人和1名乘客时,则只有50%的人认为应该优先救助乘客。但是当行人人数达到100人,乘客是1人时,有80%左右的受试者表示要救行人。当被受试者被询问是否会购买牺牲驾驶员的自动汽车时,只有不足30%的人表示愿意。<sup>[78]</sup>我国立法在规制自动汽车程序设计时,宜采纳这种调查方法,查实社会接受程度,并作为立法的数据基础。

另外,还需进行立法设定的注意义务判断。“我们既要对所谓风险刑法观保持警惕,也要谨慎对待被允许的危险的法理;动辄以进入危险社会为由扩大处罚范围的观点与做法,或者动辄以被允许的的危险为由为行为人开脱罪责的观点与做法,都是相当危险的。”<sup>[79]</sup>而防止被允许的的危险法理被滥用,就必须要通过立法为制造被允许的的危险的行为主体,设定合理的注意义务。在自动汽车领域,由于前一层次的判断尚未完成,因此我国当前并没有关于自动汽车设计、生产的安全规范。但是,在人工智能标准化论坛发布的《人工智能标准化白皮书(2018)》中,提到了人工智能的安全问题。在自动汽车进入并普及之前,有必要对自动汽车的设计、生产安全规范做出预案。关于自动汽车遭遇“电车难题”的程序设计问题,作为原则,需要设计、生产者尽到最善的义务,将自动汽车遭遇“电车难题”概率以及对第三人的法益侵害程度都降到最低。这就要求自动汽车应尽量避免陷入电车难题。从车辆的设计和编程开始,应考虑以一种具备防御性和可预期的方式驾驶,自动驾驶应重在完善技术,并通过可控交通环境的应用、车辆传感器、刹车性能、危险情况中给受威胁者发出信号提示、通过“智能”道路基础设施等来预防危险,最大程度地增强道路安全性,尽量避免“电车难题”的发生。未来相关的安全规范应该以当时的平均技术水准为依据,设定合理的概率和程度标准。只有当自动汽车出现“电车难题”的可能性低于该概率,法益损害程度低于该标准,才可以准予设计、生产,最终出现法益侵害结果也才可以通过被允许的的危险这一法理排除归责。

要让对于自动汽车中程序设计行为的刑法评价,跳出纠缠不清的、具体的生命衡量的哲学困境,必须要跨越违法性和责任性判断阶层,提前在构成要件该当性阶层寻找可行性路径。而客观归责理论中,程序设计行为如果不属于制造不被允许的的危险,就可以直接排

[77] 参见卢泽华:《人工智能:全球竞赛中国领跑》,《人民日报(海外版)》2017年3月11日第10版。

[78] 参见 Jean-François Bonnefon, Azim Shariff, Iyad Rahwan, The Social Dilemma of Autonomous Vehicles, *Science* 24 Jun 2016: Vol. 352, pp. 1573 - 1576。

[79] 张明楷:《论被允许的危险的法理》,《中国社会科学》2012年第2期,第115页。

除归责。自动汽车中的程序设计行为与被允许的危险这一法理相契合,只要满足概括性的利益衡量和有立法设定的注意义务,就可以成为被允许的危险。这就为自动汽车的设计生产者在程序设计时考虑让遭遇“电车难题”的汽车牺牲第三人生命法益而保护车内人的生命安全提供了出罪的可能性。

人工智能业已成为国家战略,作为人工智能前端产业的自动汽车普及时代即将到来。自动汽车面临“电车难题”时,是优先保护车内人的生命还是保护车外第三人的生命,取决于设计制造阶段的程序设计。通过紧急避险理论无法解决这样的难题,被害人承诺也具有局限性。在符合车辆设计目的的前提下,只要满足被允许的危险的构成要件,通过程序优先保护车内人生命的设计行为,可以直接在构成要件该当性阶段被排除归责。这将为我国自动汽车的技术发展,提供理论支撑,注入刑法保障的强心剂。

[本文为作者主持的中国法学会 2017 年度部级法学研究课题立项“网络犯罪司法适用的平衡性研究——以《刑法修正案(九)》为样本的分析”[CLS(2017)D63]的研究成果。]

---

---

[ **Abstract** ] With the arrival of the age of self-driving cars, the number of automobile accident caused by human factors will decrease drastically. However, because of the limitation of technological development, there remains the possibility of accidents resulting from technical flaws. For example, currently no reasonable solution has been found to the “trolley problem” in the programming of self-driving cars. Programming that gives priority to the legal interest of life of those in the car over that of pedestrians can not avoid the violation of law and the corresponding liability on ground of necessity because it does not meet the requirements of the principle of concrete superiority of interest and the principle of probability of expectation, whereas programming that give priority to pedestrians’ legal interest of life over that of those in the car is neither compatible with market demand of self-driving cars, nor justifiable on ground of the consent of the victim. The doctrine of permissible danger in the theory of objective imputation may offer a reasonable solution to this problem. If programming can meet the requirements of the doctrine of permissible danger, namely general superiority of interest and the obligation of best intention established by legislation, it can eliminate imputation at the stage of constitutive elements, thereby promoting the development of self-driving car technology within reasonable limits.

---

---

(责任编辑:郑佳)