

〈研究論文〉

刑法解釈学と法倫理学の観点から見た
完全自動運転車とトロリージレンマの問題*

安 秀 吉

(訳：中村真利子／監訳：裴相均)

Fully Automated Vehicle and Trolley Dilemma

—From the Perspective of Criminal Law Dogmatics and Legal Ethics—

Sugil AN

(Translator: Mariko NAKAMURA/ Translation Supervisor: Sangkyun BAE)

Abstract

The era of fully self-driving cars is just around the corner. Accordingly, the trolley dilemma is drawing renewed attention: Is it allowed to program a fully self-driving car to sacrifice a small number of people for saving the great number in cases where a fully self-driving car cannot avoid human casualties by any means while driving? At present, there are several opinions on this issue. In this article, these opinions were classified into seven types, and they were criticized from the perspective of criminal law dogmatics and legal ethics. And from the point of view of liberal criminal law theories and deontological ethics, this article tried to present the conditions that could justify programming autonomous vehicles in a way that sacrifices a small number to save the majority.

Key Words

Deontological Ethics, Fully Automated Vehicle, Trolley Dilemma, Utilitarian Ethics

目 次

- I 自動運転車が刑法学に投げかける問い
- II 刑法学が出すことのできる答え
- III 一緒に議論すべき事例
- IV ま と め

* 本稿は、2023年8月4日(金)に中央大学後楽園キャンパスで日本比較法研究所、中央大学 ELSI センター、新潟大学 ELSI センターが共同で主催した日韓比較法シンポジウムで、筆者が報告した原稿を若干修正したものです。筆者にシンポジウムで報告する機会をくださった日本比較法研究所、中央大学 ELSI センター、新潟大学 ELSI センターの関係者の皆さまに感謝申し上げます。特に、今回の報告に関して、筆者を物心両

面で後援してくださった中央大学の中村真利子先生、韓国刑事・法務政策研究院の裴相均先生に深い感謝の言葉を申し上げます。また、拙稿に対して貴重なご意見をくださった中央大学の平野晋先生と新潟大学の根津洗希先生にも感謝します。最後に、シンポジウムで筆者の報告を聞いていただき、筆者を歓待してくださったシンポジウムの参加者の皆さまにも、感謝の気持ちをお伝えします。

I 自動運転車が刑法学に投げかける問い

「完全自動運転車」の時代を目前にして¹⁾、「トロリージレンマ (trolley dilemma)」が注目されている。トロリージレンマを解決できなければ、完全自動運転車を公道に解き放つことができないという認識が広がっているのである。トロリージレンマとは、次のような事例をいう。

[事例 1] 疾走するトロリーの制動装置が言うことを聞かない。トロリーが走る線路の前方では、甲・乙 2 名が作業をしている。このまま行けば、両者ともトロリーにひかれて死亡する。ところが、トロリーと甲・乙の間の線路には分岐がある。そのため、トロリー運転手はトロリーを甲・乙のいる線路ではなく、他の線路に進入させることができる。しかし、その線路上でも、やはり丙 1 名が作業をしている。したがって、もし運転手が線路を変更すると、丙が死亡することになる²⁾。

この場合、トロリー運転手は、トロリーの線路を維持して 2 名が死ぬのを待つか、線路を変更し、1 名を犠牲にして 2 名を生かすことになる。このようなジレンマに自動運転車が直面し得ることは言うまでもない。したがって、次のような状況にどのように対応するかを議論する必要が生じる。

[事例 2] 自動運転車が高速度で走行中である。

ところが、自動運転車が疾走している車道の前方に、甲・乙 2 名が突然飛び出した。自動運転車は高速で走行していたので、急停止しても 2 名と衝突する前に車を止めることはできない。ただし、車の進行方向を左に変更することができるが、左側には丙 1 名が立っている。もし車のハンドルを左に切ると、丙が死亡してしまう。

この事例で提起される問いの重点は、自動運転車のレベルによって異なる³⁾。自動運転車に搭載された自動運転システムと、自動運転車に搭乗した運転者が運転を分担する場合（「部分的自動運転車」）には、発生した結果の責任を、自動運転車の設計者、販売者、運転者のうち誰にどの程度帰属させるかがポイントとなる⁴⁾。一方、自動運転車に搭載された自動運転システムが、自動運転車の運転を専担する場合（「完全自動運転車」）には、自動運転システムのアルゴリズムをどのように設定するかがポイントとなる。多数を助けるために少数を犠牲にするようにプログラムしても良いかという問いが前面に出てくるのである⁵⁾。

このうち部分的自動運転車に関しては、既に〔訳者注：討論者の〕根津洗希先生が詳述している⁶⁾。また、自動運転車とトロリージレンマの問題は、完全自動運転車を中心に議論されるのが一般的である。そこで、以下では、完全自動運転車の場合に限って議論を進めることとする。

1) 대한민국 정책브리핑, “2027년 ‘완전자율주행’ 시대 열린다 … 3년 뒤 도심 항공 교통 상용화”, 2022. 9. 19.; Nezu, “Strafrechtlicher Problemaufriss von (teil) autonomen Fahrzeugen in der Gegenwart und Zukunft — Darstellung möglicher Lösungsansätze —”, Hannover Law Review 4/2019, S. 268 ff.

2) 特に英米圏では、今日、我々が「トロリージレンマ」と呼ぶこのような問題を定式化した学者として、フィリッパ・フット (Philippa Foot) を挙げるのが一般的である。Foot, “The Problem of Abortion and the Doctrine of the Double Effect”, The Oxford Review 5, 1967, p. 8.

3) 自動運転車のレベルについては、SAE International, Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, 2021. 4. 30.; National Highway Traffic Safety Administration, Preliminary Statement of Policy Concerning Automated Vehicles, 2013を参照。

4) Nezu, 前掲論文 (注 1), S. 268 ff; Gless/Janal, “Hochautomatisiertes und autonomes Autofahren — Risiko und rechtliche Verantwortung”, Juristische Rundschau 2016 (10), S. 561 ff.

5) Weigend, “Notstandsrecht für selbstfahrende Autos?”, Zeitschrift für Internationale Strafrechtsdogmatik 10/2017, S. 599 ff.

6) Nezu, 前掲論文 (注 1), S. 268 ff.

II 刑法学が出すことのできる答え

運転者が事例2のような状況に置かれた場合について、刑法解釈学はどのように処理するかということから確認してみよう。もし運転者が2名を救うためにハンドルを左に切って1名を犠牲にする事件が発生した場合、刑法解釈学は、この事件を「緊急避難」という制度に当てはめる⁷⁾。ここでは、保護された法益と侵害された法益の重さを比較する。保護された法益が侵害された法益よりも本質的に優越する場合—例えば、人を救うために物を壊した場合—運転者の行為は違法性が阻却される⁸⁾。「優越的利益の原則」を適用するのである。一方、ハンドルを切ることで侵害されたのが人の生命である場合、運転者の行為は違法性が阻却されない。人の生命は同等で、量的にも質的にも衡量できないので、「優越的利益の原則」を適用することができないのである⁹⁾。

このように、刑法解釈学は、多数を救おうとハンドルを切って少数を犠牲にする行為を違法と評価する¹⁰⁾。それでは、運転者には違法にならないように行動できる道はないのであろうか。見解が分かれる。まず、一つあるという学説がある。2名の犠牲を甘受することが唯一合法的な選択であるというものである。車の進路を変更する行為（作為）とそのままする行為（不作為）のうち、いずれかを選択しなければならない場合、現状（status quo）尊重の法理によって、ハンドルをそのままにしなければならないということである¹¹⁾。しかし、この場合も違法性が阻却されないという学説もあ

る。運転者が方向を変更するのが違法であるからといって、方向を変更しないことが合法になるわけではないということである¹²⁾。これによると、いったんトロリージレンマの状況に置かれると、合法的な行為をすることができる道は塞がれてしまうことになる。

このような刑法解釈学の理論から出発すると、自動運転車がトロリージレンマの状況に置かれた場合に、多数を生かすために少数を犠牲にする方向へ進むようプログラムすることは、違法であるから許容されないという結論に至る。それでは、どうすれば良いのだろうか。車の進路をそのままにすることだけが唯一適法な行為であるという学説によれば、自動運転車が進路を維持して多数の死亡を甘受するようにプログラムされなければならない（方案3）¹³⁾。一方、車の進路をそのままにしても左に変更しても違法な行為になるという学説によれば、トロリージレンマに対処するプログラムを別途設けてはならない（方案2）¹⁴⁾。一部の学者は、トロリージレンマを解決する方法がないとして、完全自動運転車を使用すること自体を阻止することまで要求する（方案1）¹⁵⁾。

それでは、我々は、この3つの方案のうちどれを選ぶべきだろうか。他の選択肢はないのだろうか。

1 方案1

方案1は、社会にとって大きな利益になると予

7) Engländer, “Das selbstfahrende Kraftfahrzeug und die Bewältigung dilemmatischer Situationen”, Zeitschrift für Internationale Strafrechtsdogmatik 9/2016, S. 608.

8) 韓国刑法第22条.

9) 이재상/장영민/강동범, 형법총론, 제10판, 박영사, 2019, 266면; Roxin/Greco, Strafrecht Allgemeiner Teil I, 5. Aufl., 2020, § 16 Rn. 33.

10) 김준호, “폭주하는 전차 사례에 대한 형사법적 결론을 생각해보기”, 연세법학 제29권 제2호, 연세대학교 법학연구원, 2019, 12면 이하.

11) Roxin/Greco, 前掲書籍 (注9), § 16 Rn. 117; Wörner, “Der Weichensteller 40 — Zur strafrechtlichen Verantwortlichkeit des Programmierers im Notstand für Vorgaben an autonome Fahrzeuge”, Zeitschrift für Internationale Strafrechtsdogmatik 1/2019, S. 44.

12) 김준호, 前掲論文 (注10), 24면 이하; Hilgendorf, “Dilemma-Probleme beim automatisierten Fahren”, Zeitschrift für die gesamte Strafrechtswissenschaft 2018, S. 691.

13) Wörner, 前掲論文 (注11), S. 44. 本稿に対する討論文を執筆した根津先生は、方案3を「トロッコ問題不介入説」と称した。適切な命名であると思われる。

14) Hilgendorf, 前掲論文 (注12), S. 691. 根津先生は、方案2を「トロッコ問題絶対回避説」と称する。

想される新技術（自動運転車）について、あまり発生しないと思われる難問（トロリージレンマ）のために全て放棄するのが正しいのかという反論を受ける¹⁶⁾。今まで人間は、利益とリスクを伴う新技術が出てくれば、利益は最大化しながら、リスクは最小化・社会化する方法を求めて文明を発展させてきた。新技術の利用を許可しつつ、それによって発生する損害は、一不法 (Unrecht) でも不幸 (Unglück) でもない一事故 (Unfall) として、事故の被害者には、社会が連帯して積み立てた保険金で補償する方法を考案したのである¹⁷⁾。完全自動運転車を商用化するにあたって、この方法をとることができるだろう。交通事故を減らし、交通の安全性を高め、特に交通弱者の移動権を拡充すると予想される完全自動運転車の導入に反対するには、トロリージレンマの回避不可能性に加えて、さらに大きな理由を提示しなければならない¹⁸⁾。

現在、このような反対論拠として、トロリージレンマにどのように対応するかという問いは、個々人が自身の良心と道徳によって判断するしかない、正解のない問題であるという議論がある。トロリージレンマに一律に対応するアルゴリズムを搭載した完全自動運転車を許容するということは、必ず

人々にそのアルゴリズムの論理を一自身の道徳に反しても一受け入れることを強要するわけであるが、これは自由民主主義社会では許容されないという後見主義的な発想である¹⁹⁾。

2 方案2

方案2は、トロリージレンマを無視するものであるという指摘を受ける。自動運転システムを設計するには、最善の努力を尽くしてトロリージレンマと遭遇しないようにしなければならないが、それでもトロリージレンマと遭遇することになるだろう²⁰⁾。方案2を選択したドイツ倫理委員会 (Ethik-Kommission) は、トロリージレンマが絶対に発生しないように自動運転システムを設計しなければならないと強調したが、トロリージレンマが実際に発生する可能性を否定することはしていない。ところが、倫理委員会は、トロリージレンマに遭遇する可能性があるからといって、これに対応するアルゴリズムを用意してはならず、トロリージレンマを前にしてどのような決定をするかは、道徳的で責任感のある運転者に任せなければならないという意見を表明した²¹⁾。しかし、完全自動運転車においては、トロリージレンマの状況に搭乗者が介入できる可能性は全くないと思われる。それでは、道徳的で責任感のある運転者に決定を任せるという言葉は、何を意味するのだろうか。以下の2つのいずれかのようなものである。

一つは、自動運転システムが、通常のアルゴリズムに従って行った決定を受け入れようとしている可能性である。しかし、トロリージレンマに対応するアルゴリズムのない自動運転システムが、トロリージレンマの状況にどのように反応するかを予測することは難しい。進路を維持するという

15) 고인석, “자율주행자동차를 어떻게 규율할 것인가?”, 철학논총 제 96 권 제 2 권, 새한철학회, 2019, 81 면 이하; Sander/Hollering, “Strafrechtliche Verantwortlichkeit im Zusammenhang mit automatisiertem Fahren”, Neue Zeitschrift für Strafrecht 2017, S. 204. 根津先生は、方案1を「全面禁止説」と呼ぶことができるという。

16) 이종기/오병두, “자율주행자동차와 로봇윤리: 그 법적 시사점”, 홍익법학 제 17 권 제 2 호, 홍익대학교 법학연구소, 2018, 7 면. Roxin/Greco, 前掲書籍 (注9), § 16 Rn. 125h.

17) Neumann, Recht als Struktur und Argumentation. Beiträge zur Theorie des Rechts und zur Wissenschaftstheorie der Rechtswissenschaft, 2008, S. 190 f.

18) Hevelke/Nida-Rümelin, “Selbstfahrende Autos und Trolley-Probleme: Zum Aufrechnen von Menschenleben im Falle unausweichlicher Unfälle I”, Jahrbuch für Wissenschaft und Ethik 19 (1), 2015, S. 6.

19) 고인석, 前掲論文 (注15), 97 면 이하.

20) Gless/Janal, 前掲論文 (注4), S. 561; Lin, “Why Ethics Matters for Autonomous Cars”, Maurer u. a. (Hrsg.), Autonomes Fahren, 2015, S. 71.

21) Bundesministerium für Verkehr und digitale Infrastruktur, Ethik-Kommission. Automatisiertes und vernetztes Fahren, 2017, S. 10 f.

保証はない。むしろ、進路を変更する可能性の方が高いかもしれない。自動運転車は、被害を最小化する方向に進行するようプログラムされているからである。そのため、完全自動運転車のアルゴリズムを設定するにあたって、トロリージレンマのような特定の状況では「決定を下さない」ことを選択するようにプログラムすることは不可能であるとの指摘が出てきている²²⁾。何らかの方法で決定を下すしかないということである²³⁾。もう一つは、完全自動運転車の搭乗者に、もしトロリージレンマが発生した場合、これにどのように対応するかをその倫理観に応じて予め選択しておくようにするという点かもしれない。これは、規範が正解を提示することができない状況では、個人が良心に従って下した決断を尊重しなければならないという点では納得できるが、人の生死を決する問題を個人の恣意に委ねることになるのではないかとこの反論に直面することになる²⁴⁾。

一方で、方案2によれば、一旦トロリージレンマの状況に陥ると、合法的な行為ができる道は閉ざされ、違法行為を犯すしかない。そうであれば、この場合、結果の不法を引き起こした自動運転車の設計者を、刑法でどのように処理するか悩むところである。一部では、この問題を「許された危険」の法理で打開しようとする。自動運転車を道路交通に投入すれば、社会全体に大きな利益となる。ただし、注意義務を果たして自動運転システ

ムを設計しても、自動運転車がトロリージレンマの状況に置かれるリスクは残る。しかし、このリスクによって発生する損害は、自動運転システムが社会全体に与える利益に比べるとわずかである²⁵⁾。したがって、トロリージレンマに直面するリスクは、社会的に相当なリスク、まさに許された危険であると認めることができるということである²⁶⁾。トロリージレンマに対処するプログラムまで用意することは、自動運転システムの設計者に課される注意義務の範囲に含まれないと見るのである²⁷⁾。しかし、これに反対する声もある。方向を転換して1名を死に至らしめるのは違法な行為であると評価しておきながら、この違法行為が、許された危険の法理によって適法行為に変わると説明するのは、一貫しない解釈であるということである²⁸⁾。

3 方案3

多数の死亡を甘受することが唯一の正解であるとする方案3は、一般人の直感と常識に合致しないという欠点がある²⁹⁾。トロリージレンマのような状況では、人命被害を最小限に抑えるという選択が最も多く支持されていることは、いくつかの研究で重ねて確認されている厳しい事実である³⁰⁾。

また、方案3は、人の生命は絶対に衡量できないという原則は重視するが、人の生命を最大限保

22) Hörnle/Wohlers, "The Trolley Problem Reloaded. Wie sind autonome Fahrzeuge für Leben-gegen-Leben-Dilemmata zu programmieren?", *Goldammer's Archiv für Strafrecht* 2018, S. 20.

23) Schuster, "Coping with Moral Dilemmas in German Criminal Law Theory and Justice: Classical Cases and Modern Variants", *Criminal Law Forum* 2023, p. 26.

24) 本稿に対する討論をされた平野先生は、ドイツ倫理委員会の方案は多数の「同伴自殺」〔訳者注：無理心中のことで、討論文では「心中」とされている〕を強要するような立場であり、賛成できないという見解を表明した。このような平野先生の見解は、ドイツ倫理委員会が提示した方案の盲点を鋭く刺すものである。

25) Bayerischer Landtag, Ausschuss für Wirtschaft und Medien, Infrastruktur, Bau und Verkehr, Energie und Technologie. 38. Sitzung, 2015. 10. 29., S. 50.

26) 根津先生も、社会が自動運転車を利用することで得られる利益が、生じる可能性のある損害よりもはるかに大きいと判断されれば、自動運転車について、許された危険の法理を適用できると考える。Nezu, 前掲論文(注1), S. 272 f.

27) 이승준, "자율주행자동차 사고시 형사책임에 따른 자율주행자동차의 운행과 책임에 관한 법률안" 시도, *법제연구* 제53호, 한국법제연구원, 2017, 616 면.

28) Engländer, 前掲論文(注7), S. 612 f.; Wörner, 前掲論文(注11), S. 46 f.

29) Hörnle/Wohlers, 前掲論文(注22), S. 20 f.; Weigend, 前掲論文(注5), S. 600.

護しなければならないという原則は軽視するものであるとの批判も受ける³¹⁾。方案3は、人間の尊厳という理念を死守するためには、犠牲になる人間の数は無視しても良いという立場であるということが出来るが、この立場が、人間の尊厳という思想に真に一致するものなのかということも併せて考える必要がある³²⁾。

4 方案4

上記の3つの方案とは異なる道を探すとすれば、まず功利主義を前面に打ち出す方法を検討してみることができる。ある法益を保護するためには、他の法益を侵害するしかないという場合にも、その侵害される法益よりも保護される法益が大きい場合、小さな法益を侵害して大きな法益を助ける行為を正当と考えるのである。功利主義の発祥の地である英米では、人の生命も量的に衡量できるという考えをタブー視しない³³⁾。アメリカの模範刑法典 (Model Penal Code)³⁴⁾にも、このような考えの痕跡がある³⁵⁾。これによると、2名を救う方が1名を救うよりも良いことになる。

大陸法系でも、緊急避難の正当化の是非を考える際には「優越的利益の原則」を用いる。しかし、

緊急避難の正当性は、このような功利主義の物差しによってのみ判断するわけではない。例えば、それぞれ異なる臓器に問題がある5名が、急いで移植を受けなければ死亡するという場合、この5名を救うために、死刑囚である青年1名を殺害して、その臓器を5名に移植して生かしたとすれば、社会全体の利益は増加したと判断することもできるだろう。しかし、刑法はこれを正当な緊急避難として認めない。これは、ある行為が緊急避難として正当化されるためには、「優越的利益の原則」とは別の独自の基準に適用ものでなければならないことを意味するが、それは「相当性」ないし「連帯性」の原則である³⁶⁾。ある個人に対して、他人

30) 이기쁨 외, “자율주행 자동차 딜레마 시나리오에서 운전자 관점의 도덕코드, 한국인을 대상으로”, 한국디자인리서치 제5권 제1호, 한국디자인리서치학회, 2020, 131면 이하; Bonnefon/Shari/Rahwan, Autonomous Vehicles Need Experimental Ethics: Are We Ready for Utilitarian Cars?, 2015.

31) Hörnle, “Töten, um viele Leben zu retten. Schwierige Notstandsfälle aus moralphilosophischer und strafrechtlicher Sicht”, Putzke u. a. (Hrsg.), Strafrecht zwischen System und Telos. Festschrift für Rolf Dietrich Herzberg zum siebzigsten Geburtstag am 14. Februar 2008, 2008, S. 562 ff.

32) 人間の尊厳という思想が具体的な人間の生活に及ぼす結果を度外視すれば、人間を利するために存在する人間の尊厳という思想 (Menschenwürde) は、むしろ人間を害する負担 (Menschenbürde) となる可能性がある。Neumann, 前掲書籍 (注17), S. 35 ff.

33) Hevelke/Nida-Rümelin, 前掲論文 (注18), S. 7; Schuster, 前掲論文 (注23), p. 10.

34) Section 3.02. Justification Generally: Choice of Evils.

(1) Conduct which the actor believes to be necessary to avoid a harm or evil to himself or to another is justifiable, provided that:

(a) the harm or evil sought to be avoided by such conduct is greater than that sought to be prevented by the law defining the offense charged; and

(b) neither the Code nor other law defining the offense provides exceptions or defenses dealing with the specific situation involved; and

(c) a legislative purpose to exclude the justification claimed does not otherwise plainly appear.

(2) When the actor was reckless or negligent in bringing about the situation requiring a choice of harms or evils or in appraising the necessity for his conduct, the justification afforded by this Section is unavailable in a prosecution for any offense for which recklessness or negligence, as the case may be, suffices to establish culpability.

35) Bourget/Chalmers, “What do philosophers believe?”, Philosophical Studies 170, 2014, pp. 465-500によると、英米圏の哲学者を対象としてアンケートを行った結果、トロリージレンマの状況では、トロリーの進路を変更し、少数を犠牲にして多数を救わなければならないと回答した者が絶対多数であり、進路を変更してはならないという回答は、8%にすぎなかったという。

36) 이재상 / 장영민 / 강동범, 前掲書籍 (注9), 265면 이하; Neumann, “§ 34”, Kindhäuser/Neumann/Paeffgen (Hrsg.), Nomos Kommentar Strafgesetzbuch, Band 1, 5. Aufl., 2017, Rn. 9 ff.

の大きな利益を保護するために自身の利益が侵害されることを甘受するよう要求できる根拠は、そのような甘受を個人が社会の構成員として負担すべき義務、すなわち、緊急状況にある他人と連帯する義務であると解釈することで見出すことができる。しかし、個人をただ（他人に代替することができる）利益の「管理者」としてではなく、（他人に代替することができない）利益の「権利者」と認める国家では、ある個人に要求できる連帯義務に限界がある。その生命まで犠牲にするよう要求することは、たとえこれによって得られる利益が莫大なものであったとしても、決して正当化することができない³⁷⁾。

功利主義の物差しだけでトロリージレンマを見ると、社会全体の利益のために個人の生命を犠牲にしても良いという結論に至りやすい。このような結論は、個人を手段化するものであり、個人の自由を最高の価値と認め、個人の権利を保障する自由主義的な法治国家³⁸⁾では容認できない。韓国の国土交通部が、韓国交通安全公団、韓国交通研究院とともに2020年12月に発行した「自動運転自動車倫理ガイドライン」でも、自動運転車に関する倫理を定めるにあたって、「功利主義原則を表面化しようという主張」に難色を示したことがある³⁹⁾。

5 方案5

完全自動運転車のアルゴリズムでは、「決定を下さない」という選択肢はあり得ないとの指摘があることは前述した。同様の文脈で、自動運転車の場合には、そのまま受け入れなければならない運命というものはないとの主張が提起される。自動運転車は、事前に入力されたアルゴリズムに起因して、リスク要因を絶えず分析して走行するため、元々決まった自然な進路はないということであ

る⁴⁰⁾。このように見ると、自動運転車と衝突するリスクに直面しているように見える多数と、そのようなリスクから離れているように見える少数が、実は全て同じように自動運転車と衝突するリスクの前に置かれていると見ることができる。

そうすると、ここで核心となる問いは、多数が置かれている宿命的なリスクを少数に人為的に転嫁しても良いかということではなく、全員リスクにさらされた状態で誰を救うのかということになる。トロリージレンマの前では、自動運転システムは必ず「甲と乙を殺さない義務」と「丙を殺さない義務」を同時に負担することになるのである。ところが、両義務をいずれも履行する道はないため、結局、義務衝突に陥る。生命を助けるべき義務が衝突する場合には、「解決できない義務衝突」となるが⁴¹⁾、ここで生命の数を勘案すべきかどうかをめぐっては争いがある。衝突する義務の中で、より多くの生命を生かすことができる義務を履行すれば違法性が阻却されるという見解と、少数を生かす義務を履行しても違法性が阻却されるという見解が対立している。前者によれば、トロリージレンマの状況では、自動運転車が丙1名だけに衝突するようにプログラムしなければならないという結論に至るだろう。後者による場合には、自動運転車の進路を任意に定めることができるようになるが、これはランダムなアルゴリズムを入力する方法で具現化すれば良いとの提案がある⁴²⁾。この提案の内容については後述する。

6 方案6

人の生命を衡量することが禁止されており、このことから多数を生かしても少数を生かしても違法行為にならざるを得ないとすると、自動運転車の進路を運に任せる方が良いと考えることもできる。トロリージレンマの状況で誰を救うかを抽選で決めようというものである。多数でも少数でも

37) Neumann, 前掲論文（注36）, Rn. 77d.

38) von der Pfordten, "Normativer Individualismus und das Recht", Juristenzeitung 2005, S. 1069.

39) 국토교통부, 자율주행자동차 윤리가이드라인, 2020, 12면 이하.

40) Weigend, 前掲論文（注5）, S. 602.

41) Weigend, 前掲論文（注5）, S. 603 f.

42) Roxin/Greco, 前掲書籍（注9）, § 16 Rn. 125q.

いずれかの死亡を避けられないときは、ランダムにいずれかが選択されるようにする「ランダム生成器 (random generator)」を自動運転システムに入力し、そこから出る結果を受け入れようということである⁴³⁾。歴史を見ると、人間の恣意や偏見を排除し、公正を確保する方法として抽選を活用してきた事例は少なくない⁴⁴⁾。最近でも、コロナ19で病床不足の事態が生じ、「トリアージ (Triage)」問題が懸念として浮上した際に、その解決策としてくじ引きが提案されたこともある⁴⁵⁾。

ただし、多数が死亡するか少数が死亡するか不確かな状況で、人の生死を直ちに運に任せることは、性急で無責任であるという印象を残す点は否定しにくい⁴⁶⁾。

7 方案7

本稿で取り上げる最後の選択肢は、トロリー運転手の場合 (事例1) と完全自動運転車の場合 (事例2) を比較すると、人が決定を下す時点が異なるという事実に注目する。

トロリー運転手の場合には、人間 (トロリー運転手) が決定を下す瞬間に、その決定に従って保護

され、又は犠牲になる者が誰であるかが既に具体化 (個別化) されている。トロリーの線路をそのまましておけば甲・乙が死亡し、変更すれば丙が死亡する。アイデンティティのある個人のいずれを犠牲にするかという問題が発生するのである。トロリー運転手が決定を下す時点で、個人である甲・乙の利益と個人である丙の利益は鋭く対立する。トロリーが走行する線路上にいる甲・乙にとっては、トロリーの進路が変更されることが利益であるが、丙にとっては線路が維持されることが利益である。この場合、片方の生存は、もう片方の犠牲の対価としてのみ保証することができる。

これとは異なり、自動運転車がトロリージレンマの状況にどのように対応するかというプログラムを人間が決定する時点は、自動運転車がトロリージレンマの状況に置かれる時点に先行する。決定を下す時点では、その決定によって後に保護され、又は犠牲になる者が誰であるかが全く具体化 (個別化) されていない。トロリー運転手の事例とは異なり、個人の利益は対立しない。むしろトロリージレンマの犠牲者となる確率を減らすという同じ利益を共有する。したがって、自動運転車がトロリージレンマの状況に置かれた場合には、死者の数を最小化する方向へ進めるようにすれば、後で生存者になる者であれ犠牲者になる者であれ、誰にとっても利益となる⁴⁷⁾。

このような発想に基づいて死亡者数を最小限に抑えるようにプログラムすることは、一功利主義はもちろんのこと一義務論によっても正当化することができる。このようなプログラミングは、イマヌエル・カント (Immanuel Kant) 的に言えば、人間の価値を衡量し、人間を手段として考えるのではなく、むしろ人間を目的として考え、人間の生存可能性を最大化するものであるからである。ジョン・ロールズ (John Rawls) 風に表現すると、

43) Roxin/Greco, 前掲書籍 (注9), § 16 Rn. 125q; Weigend, 前掲論文 (注5), S. 603 f. 根津先生は、この方案を「ランダム義務履行説」と命名する。

44) Steiner/Frey, "In Extremsituationen sollten selbstfahrende Autos per Zufall entscheiden", Welt, 2023, 3. 7.

45) Lossau, "Ist Zufall besser als Triage?", Welt, 2021, 3. 25.; Poser, "Plädoyer für den Zufall", Welt, 2022, 5. 14.

46) Stahnke, "Das autonome Fahrzeug im Dilemma: Strafrechtliche Risiken des Softwareentwicklers und gesetzgeberische Möglichkeiten", Bonner Rechtsjournal 02/2019, S. 84. 平野先生は、トロリージレンマの解決を運に委ねるという方案6には、不断の努力によって発展させた自動運転車の技術を無益なものとするリスクがあることを見抜いた。したがって、ランダム生成器は、本当に回避できない場合に最後の手段としてのみ最小限度で使用し、偶然に左右される領域を最小化する技術の発展に尽力しなければならないだろう。

47) Hevelke/Nida-Rümelin, 前掲論文 (注18), S. 10 ff.; Neumann, "Schadensmindernde Umsteuerung autonomer Fahrzeuge als Problem des strafrechtlichen Notstandes", 2015, S. 4 ff.

上記のようなプログラムは、無知のヴェール (veil of ignorance) の下で行った正義の選択でもある。ロールズによれば、社会を構成する個人々が、後に自身が社会で得る地位と引き受ける役割が何であるかがわからない状態で、社会の利益と損害を分配するルールに合意すれば、その合意されたルールは正義であるとみなされる⁴⁸⁾。自動運転車がトロリージレンマの状況に対応するプログラムを組んだ時点で、個人々が、後で自身がそのプログラムの受益者になるのか被害者になるのかわからない場合は、自身の死亡リスクを最小限に抑えるようにプログラミングしようと合意することが、合理的で正義に適う選択となり得る⁴⁹⁾。

筆者は、この方案が最も無難であると考えている。もちろん、筆者と異なる立場からは次のような反論がある。まず、この方案も結果的には人 (少数) を他人 (多数) の安全のために犠牲にすべき標的とする「死のアルゴリズム (Algorithmen des Todes)]⁵⁰⁾を採用することを意味し、これは、人の生命はいかなる場合でも衡量できないという根本的な価値観を弱めるおそれがあるということである⁵¹⁾。また、プログラムを設計する時点では、誰が受益者になって誰が犠牲者になるかわからないという点についても、プログラムが実行される時点では、結局、具体的な個人が恩恵を受け、あるいは犠牲になることになるので、この時点では自動運転車の場合とトロリー運転手の場合とで本質的な差異がないという反論もある⁵²⁾。

48) Rawls, A Theory of Justice, Revised Edition, 1971, pp. 10-18.

49) Neumann, 前掲論文 (注 47), 2015, S. 4 ff.

50) Hilgendorf, "Recht und autonome Maschinen — ein Problemaufriß", Hilgendorf/Hötitzsch (Hrsg.), Das Recht vor den Herausforderungen der modernen Technik. Beiträge der 1. Würzburger Tagung zum Technikrecht im November 2013, 2015, S. 20.

51) Engländer, 前掲論文 (注 7), S. 613; Hilgendorf, 前掲論文 (注 12), S. 694; Roxin/Greco, 前掲書籍 (注 9), § 16 Rn. 125k.

52) Engländer, 前掲論文 (注 7), S. 613; Hilgendorf, 前掲論文 (注 12), S. 694.

Ⅲ 一緒に議論すべき事例

完全自動運転車のプログラムを設計するにあたっては、トロリージレンマ (事例 2) のほか、次のような事例にも備えなければならない。

1 対称的危険共同体 (生存者と犠牲者の数が異なる場合)

[事例 3] 自動運転車が一方通行の道路を走行中である。ところが、突然、児童 3 名が自動運転車の進路の前方に飛び出した。自動運転車が児童 3 名と衝突する前に車を止めることはできない。なお、児童甲は、自動運転車の左側に立っており、児童乙と丙は自動運転車の右側に立っている。したがって、自動運転車が進路を変更せずにそのまま走行すれば、甲、乙、丙が全て死亡するが、車のハンドルを左に切ると甲が死亡し、右に切ると乙と丙が死亡するという状況に置かれた。

事例 2 では、多数でも少数でも一部だけ死亡し、一部は生き残る。しかし、この事例では、一部だけではなく、全員死亡することになる。まさに危険共同体 (Gefahrengemeinschaft) の事例である。

方案 2 によれば、ここでもやはり自動運転車の進路を決めるアルゴリズムを設定してはならない。方案 3 の論理は、誰もが危険にさらされている事例 3 には全く適用されない。方案 2 や方案 3 によると、ここでは、3 名とも死亡する結果が生じても、これを甘受しなければならない。もちろん、方案 2 や方案 3 を選択する学者も、ここでは被害を最小化する必要があると認めるが⁵³⁾、そうすると、人の生命は絶対に衡量してはならないという方案 2 や方案 3 の前提を修正するしかない。このことは、方案 2 や方案 3 が、完全自動運転車のプログラムについて、一貫性をもって導くことのできる原理となる難しさを示唆している。

53) Hilgendorf, 前掲論文 (注 12), S. 696.

方案4や方案7によれば、事例3では方向を左に変更するようにプログラムしなければならない。方案5や方案6によれば、この場合には、左を向くようにプログラムしたり、あるいはランダム生成器に運命を委ねることになるだろう。

2 対称的危険共同体（生存者と犠牲者の数が同じ場合）

[事例4] 事例3と状況は同じである。ただ、自動運転車の前方に突然飛び出した児童は、3名ではなく甲と乙2名で、甲は自動運転車の左側に、乙は右側に立っている。

ジレンマの対象となる生命の数が同じこの事例では、自動運転車にどのように反応させるかを定めることは、事実上不可能である。少なくとも両方を犠牲にする方案は排除すべきである。ランダム生成器に運命を委ねる道のほかに、代案が見えてこない⁵⁴⁾。

3 交通ルールの違反者

[事例5] 自動運転車が高速度で走行中である。ところが、甲と乙2名が、自動運転車が疾走中の車道の前方に、交通ルールに違反して突然飛び出した。自動運転車は高速で走行していたので、急停止しても2名と衝突する前に車を止めることはできない。ただし、車の進行方向を左に変更することができるが、左側には交通ルールを遵守している丙1名が立っている。もし車のハンドルを左に切ると、丙が死亡してしまう。

交通ルール違反の有無は問わず、生存可能性の最大化というアルゴリズムだけを選択するとすれば、事例5では、交通ルールを遵守している丙が死亡してしまう。これは、法を守った者が法を破った者によって犠牲になるということであり、正義

ではない。また、法を守った者が損害を被れば、人々は交通ルールを遵守せず、その後、道路交通は危険なものとなる。そのため、交通ルールを遵守した者が違反した者よりも優先的に保護されなければならないとの見解が、幅広い支持を受ける⁵⁵⁾。これによると、甲と乙2名が犠牲にならないといけない。

ところが、この見解にも盲点がある。注意力がなくルールに違反しやすい児童等が、非常に不利な立場に置かれる可能性があるのである。また、自動運転車が、交通ルールに違反する背景を正確に把握することは極めて困難であるため、交通ルールの遵守いかんのみを強調すると、むしろ不当な結果が生じることもある。例えば、ある者が突然道路上に飛び出す行為は、意図的なルール違反かもしれないが、強盗を避けるためのものかもしれないし、突然の心臓発作でバランスを失った結果であるかもしれない⁵⁶⁾。このような複雑な背景を無視したまま、表面的な交通ルールの遵守いかんを考慮することは望ましいことではない。

したがって、交通ルールを守る者が優先されなければならないという原則が採用されたとしても、この原則をいつどの程度実現するかは、交通ルールに違反する状況を把握する技術のレベルを勘案して決するしかない⁵⁷⁾。

4 搭乗者と歩行者

[事例6] 自動運転車が高速度で走行中である。ところが突然、甲と乙2名が、自動運転車が疾走中の車道の前方に飛び出した。自動運転車は高速で走行していたので、急停止しても2名と衝突する前に車を止めることはできない。2名と衝突しないようにするためには、車の進行方

55) Hevelke/Nida-Rümelin, 前掲論文（注18）, S. 19 f.; Hörnle/Wohlers, 前掲論文（注22）, S. 24 f.

56) Mitsch, "Die Probleme der Kollisionsfälle beim autonomen Fahren", *Kriminalpolitische Zeitschrift* 2/2018, S. 72.

57) Roxin/Greco, 前掲書籍（注9）, § 16 Rn. 125p.

54) Roxin/Greco, 前掲書籍（注9）, § 16 Rn. 125q.

向を左に変更するしかないが、左側には大きなコンクリートの電信柱がある。もし車のハンドルを左に切ると車は電信柱と衝突し、これにより車の搭乗者である丙が死亡することになる。

自動運転車の搭乗者の生命と、歩行者の生命が衝突するこの事例では、搭乗者が歩行者に比べて特別に保護されなければならないかが問題となる。自動運転システムのアルゴリズムは、「自己保護 (self-preservation)」と「自己犠牲 (self-sacrifice)」のうち、いずれの原則に基づくべきかという問題である⁵⁸⁾。

自己保護の原則を支持する立場は、搭乗者が優先的に保護されなければならないと考える。その理由としては、倫理的にも法的にも、搭乗者が他人のために犠牲になる義務はないことが挙げられる⁵⁹⁾。また、搭乗者も犠牲になるようにプログラムすることは、自動運転車の購入者と生産者の双方に対して過度の要求であるとも言われる。搭乗者も犠牲になり得るアルゴリズムを搭載した自動運転車は、消費者にそっぽを向かれるであろうし、そうすれば自動運転車の発展が阻害される⁶⁰⁾。この立場によると、事例6では甲と乙の死亡を甘受すべきということになる。

これに対して、自己犠牲の原則を支持する立場は、自動運転車の利益を享受する搭乗者は、自動運転車のリスクをも負担することが当然であると反論する。他人の生命を犠牲にしてまで搭乗者の生命を保護すれば、搭乗者に過度の利益を与えることになり、不公正であるということである⁶¹⁾。

58) Mayer/Bell/Buchner, "Self-protective and self-sacrificing preferences of pedestrians and passengers in moral dilemmas involving autonomous vehicles", *PLoS ONE* 16 (12), 2021, pp. 1-25.

59) Weber, "Dilemmasituationen beim autonomen Fahren", *Neue Zeitschrift für Verkehrsrecht* 2016, S. 253.

60) Gurney, "Crashing into the Unknown: An Examination of Crush-optimization Algorithms through the Two Lanes of Ethics and Law," *Albany Law Review* 79, 2015/2016, pp.183-266.

搭乗者が歩行者のために生命を犠牲にする義務を負わないように、歩行者も搭乗者のために自己の生命を犠牲にする義務を負わない⁶²⁾。したがって、事例6では、搭乗者である丙が犠牲にならないといけないということになる。

自動運転車を導入する趣旨について、自動運転車の搭乗者だけではなく、道路交通に関わる全ての者の生存可能性を高めることに見出すと、搭乗者と歩行者のいずれかを優先するのではなく、双方とも均等に保護するようにプログラムすることが公正であるように思われる⁶³⁾。

IV ま と め

完全自動運転車の商用化に反対する立場(方案1)を選択しないとすると、完全自動運転車がトロリージレンマの状況(事例2)に置かれた場合にどのように反応するプログラムにするかという問題に、何らかの答えを出すほかない。その答えは、前述した方案(方案2~7)のうち一つを選択したり、二つ以上を組み合わせて見出すことができる。一例として、トロリージレンマの状況には、許された危険(方案2)と義務衝突(方案5)の法理を適用しつつ、ランダム生成器(方案6)を活用する方法が提示されている⁶⁴⁾。

この提案は説得力があるように見受けられるが、

61) Gless/Janal, 前掲論文(注4), S. 575.

62) 歩行者のために搭乗者を犠牲にするアルゴリズムを搭載した自動運転車は、消費者にそっぽを向かれるとの主張の基礎には、搭乗者のために歩行者を犠牲にするアルゴリズムを搭載した自動運転車が、消費者の支持を得るという予測がある。しかし、この予測が妥当であるかは疑問である。道路交通に関わる市民(消費者)は、搭乗者であるか歩行者であるかいずれかというのではなく、搭乗者でも歩行者でもあるからである。そうすると、歩行者を犠牲にするアルゴリズムも、消費者にそっぽを向かれる公算が大きい。これは、一平野先生が議論したように一ダイムラー(Daimler)社が自動運転車において搭乗者保護を最優先すると明らかにしたところ、市民の激しい反発を買ったという事実からよくわかる。

63) Roxin/Greco, 前掲書籍(注9), § 16 Rn. 125p.

64) Roxin/Greco, 前掲書籍(注9), § 16 Rn. 125d ff.

ただし、この提案を一貫した形で体系的に確立できるようにするためには、方案7を大原則とする必要がある。方案7は、義務論の視点から見れば矛盾なく正当化することができる。まず、自動運転車のアルゴリズムを組むにあたって、「人命被害が発生しないように、交通ルールを遵守して防衛的に運転せよ」という命令を基本原理とするならば、これはカントが定言命法で要請した「普遍的な法則になることができる準則」を定立するものと言える。また、トロリージレンマのように人の死亡を回避することができない場合には、犠牲を最小限に抑えるようにプログラムすることは、人間を手段ではなく目的とし（カント）、無知のヴェールの下で各個人を同等の道徳的人格として扱うこと（ロールズ）の結果である。

もちろん、方案7が「死のアルゴリズム」を正当化するという批判は無視できない。しかし、「死のアルゴリズム」を選択する方案7が、これを否定する方案2よりも、人間の生存可能性を高めるのに有利であるという事実を見落とすべきではな

い。また、方案7も結果的には人の生命を衡量して少数を犠牲にするものであるとの反論も一理ある。しかし、方案7によると、トロリージレンマの状況で少数の死亡に繋がるのは、個人の生命を計り、少数の生命を手段化する功利主義的（集団主義的）決定による結果ではない。個人の生存可能性を最大化するという正当な目的を達成するために下した合理的な決定（プログラミング）が、自動運転システムのアルゴリズムを通じて具現化された結果である⁶⁵⁾。

危険共同体という形で発生するジレンマ状況（事例3、4）でも、方案7に基づき多数を救う方向でプログラミングを行うことになるが、どうしても生存者と犠牲者の数が同じである場合には、ランダム生成器を利用することは避けられないように思われる。交通ルールを遵守した者が、違反した者よりも優先されるようにするプログラムは、自動運転車のレベルに応じて段階的に実現していき（事例5）、歩行者と搭乗者は均等に保護しなければならない（事例6）。

65) この問題では、義務論と功利主義は、論拠と論証の方法は異なるものの、同じ結論に至る。出発点と経路は異なるが、同じ地点に到着するのである。これは、自動運転車のトロリージレンマ問題においては、義務論と功利主義が調和できる端緒があることを示唆する。本稿の議論を受けて、根津先生は、今後重要なのは、「義務論か功利主義かといった二分法」〔訳者注：討論文では「義務論か功利主義かといった二項対立的な次元」とされている〕を浮き彫りにするのではなく、「互いを尊重する」〔訳者注：討論文では「互いに敬意を表する」とされている〕「価値相対主義」に基づいて、合理的な解決策を探ることにあるという意見を開陳した。完全に同意する。