

### 自動運転の安全性を担保する法制度：ディ レンマ状況（トロリー問題）への対応

今井, 猛嘉 / IMAI, Takeyoshi

---

(出版者 / Publisher)

法学志林協会

(雑誌名 / Journal or Publication Title)

Review of law and political sciences / 法学志林

(巻 / Volume)

118

(号 / Number)

4

(開始ページ / Start Page)

306(123)

(終了ページ / End Page)

284(145)

(発行年 / Year)

2021-03-20

(URL)

<https://doi.org/10.15002/00025505>

# 自動運転の安全性を担保する法制度

—ディレンマ状況（トロリー問題）への対応—

今井 猛 嘉

## 1 はじめに

日本では、道路交通法の改正により、自動運転の技術レベル3（L3）の公道での利用が可能になった。<sup>(1)</sup> 次の課題は、レベル4（L4）の利用を可能にする環境整備である。L4においては、車両内にいる乗員（occupant. O）は、車両が、自動運転が可能な運行設計領域（operational design domain. ODD）を走行中は、運転操作をしない。その際に、O以外の者を死傷させた事故が生じた場合、誰が、如何なる内容の法的責任を負うのであろうか。この点の検討は、実は、L3を許容する際にも考慮されているべき課題であった。即ち、L3においては、自動運転を実現するシステム（automated driving system. ADS）<sup>(2)</sup> が作動中は、Oは運転者（driver. D）ではなく、ADSこそがDである。ADSが機能限界に達するなど、自動運転の実現が困難になれば、Oが運転権限を引き継ぎ（テーク・オーバー。take over. TO）、以後、OがDとなって運転を継続する。このように

L3=L4 and above (before TO) +L2 and below (after TO)

(1) 今井猛嘉「自動運転制度実現へ向けた動向」、法律のひろば73巻2号（2020年）44頁以下。

(2) その中核はAI（artificial intelligence）である。

と整理できる。TO以前は、当該車両がODD内を走行しているのと同じであるから、L4が実現されていることとなる<sup>(3)</sup>。

そこで、L3の利用を適法化するには、L4の適法性も検討される必要があったが、この点への配慮は十分ではなかった。

L3の利用が普及すると、ODD内での自動運転の効用 (utility. U)<sup>(4)</sup> が、より広く認識され、L4の適法化への要請も高まるであろう。L3の利用が可能になった現在こそ、L4が適法化される根拠を、L3のTO以前の状況をも視野に入れて確認する必要がある。

本稿では、以上の観点から、L4及びTO以前のL3に共通する難問である、ディレンマ状況への対処法を検討することにする。

## 2 ディレンマ状況（トロリー問題）への対応

### 2-1 問題設定

L4で走行する自動運転車による事故処理との関係で、活発に議論されているのは、いわゆるトロリー問題 (Trolley problem) に代表されるディレンマ状況下での法的責任の内容及び帰属である。ここでディレンマ状況とは、ある法益 (例えば、ある人間の生命) を侵害しなければ、他の法益 (別の人間の生命) を保護することができず、いずれかの法益を侵害することになる行為を、瞬時に選択せざるを得ない状況である。

この問題を、次の設例で検討する。

(設例1)

ある自動運転車 (autonomous vehicle. AV)<sup>(5)</sup> は、自動運転車だけが利用で

(3) 今井猛嘉「自動運転の安全性を担保する法制度—近時の状況を踏まえて」自動車技術75巻1号(2021年)29頁以下。

(4) 人間の判断ミスによる交通事故の減少、移動サービスを運転困難者にも提供することで、過疎、高齢化が進む地域の受容に答えること等。

きる専用道路を L4 相当の技術で自動走行していた。O は、その車内にいた自然人であるが、当該 AV が L4 で走行中である（即ち、当該 AV に予定された ODD 内を走行中である）ため、O は、当該 AV の走行制御に関与していなかった。しばらくして、O は、当該 AV の進行方向（前方）に幼稚園児（child. C）5 人（C1 から C5）が歩いていることに気づいた。当該 AV が走行していた自動運転車の専用道路には、歩行者の立ち入りは想定されていなかった。ADS は、急ブレーキを掛けても C1-5 への衝突を回避できないと判断し、彼らへの損害発生（彼らの死亡）を回避するため、左側の（専用道路外の）歩道に向けてハンドルを切った。ところが同歩道では老人 V が散歩しており、当該 AV は V に衝突し V を死亡させた。ADS には、左にハンドルを切ると V に衝突して死亡させることは、認知されていた。この判断は、「ディレンマ状況に陥った場合には、より少ない法益侵害に至る車両の挙動を選択せよ。」とする衝突時における最適アルゴリズム（crash optimisation algorithm. COA）による情報処理の結果であった。

このアルゴリズムによる車両の挙動制御は、倫理的に正当なものか。V が死亡したことにつき、誰が、如何なる内容の法的責任を問われるのか。ADS の作成者（ADS を構成するアルゴリズムのライター。AW）か、当該 AV の製造者（manufacturer. M）、あるいはその販売者（seller. S）なのかが問題となる。

## 2-2 検討のための視点

（設例 1）のような事例の解決は、国際的にも関心が高く、倫理的衝突アルゴリズム（ethical crash algorithm. ECA）とは何かが議論されている。そこで示されている見解は、（設例 1）における V 死亡という結果に係る法的責任を検討する際にも有益な視点を提供するものである。本稿では、法的責任とし

---

(5) Autonomous vehicle と automated one とでは、自動運転の程度に差があるが、両者の区別は国際的議論においても貫徹されていないことから、本稿では、前者を、L4 及び L5 で走行可能な自動運転車を示す用語として使用する。

ては刑事責任に限定した検討を行う。

犯罪は、客観的要件が充足された後に、対応する主観的要件が充足された場合に、成立する。客観的要件とは、(法益侵害という違法な)結果、その結果を引き起こした因果関係、因果関係の起点となる行為、及び当該行為をした行為者である。<sup>(6)</sup>主観的要件の中核は、(客観的要件の存在を認識したか、予見した場合に認められる)故意と、(故意未満の心理状態としての、結果の予見可能性がある場合に認められる)過失である。

(設例1)のAVはL4相当で自動走行しており、Oには、当該車両の挙動を制御する義務はなかった。即ち、OはDではなく、Oには、(AVの挙動を、Vへの衝突を避けるべく変更しなかったという)不作為及び(V死亡という)結果を惹起したこととの因果関係が認められないから、刑事責任(例えば、過失運転致死罪<sup>(7)</sup>や業務上過失致死罪<sup>(8)</sup>)を問われることはない。

そこで、V死亡という結果につき、刑事責任を問われる可能性があるのは、当該AVを販売したS、当該AVを製造したM、当該AVのADSを構成するアルゴリズムを作成したAWだということになる。

S、M、AW(Sら)を当該AVの運転者(D)として評価することは可能であるから、彼らに過失運転致死罪が成立しないかが検討の対象となる。具体的<sup>(9)</sup>

(6) 事例を検討する際には「結果→因果関係→行為及び行為者の確定」という手順が採られるが、一旦、犯罪の成立が肯定されれば、「行為者による行為→因果関係→結果」という時系列に沿った説明がなされることが多い。

(7) 自動車の運転により人を死傷させる行為等の処罰に関する法律第5条。

(8) 刑法第211条。

(9) 今井・前掲「自動運転の安全性を担保する法制度—近時の状況を踏まえて」30頁以下参照。

(10) 彼らが、当該AVは、ディレンマ状況に遭遇した場合、無関係の第三者に対して傷害や死亡という結果を惹起することがありうと思いつつ、販売、製造、作成をしていた場合には、故意犯(殺人罪等)の成否も問題となる。ディレンマ状況の発生は、統計的には、その可能性を否定できないとしても、何時、どこで、如何なる経過でこれが発生するかを事前に予測することは不可能だから、(条件付き)故意を認定することは困難である。For a similar viewpoint, cf. *Eric Hilgendorf, The dilemma of autonomous driving: Reflection on the moral and legal treatment of automatic collision avoidance systems, in Eric Hilgendorf/Jochen Feldle (eds.), Digitization and the Law (2017), p81, Fn66.*

には、①Sらの行為つき、生じた結果（Vの死亡）に係る違法性が阻却されないかが問われ、②違法性が阻却されない場合には、Sらの（主観的要件としての）過失の有無が検討される。

①（違法性阻却の有無）は、V死亡という結果が、緊急状態でのアルゴリズムの作動により惹起された場合、（生命という）法益の個数（守られた生命数と、侵害された生命数との比較）を踏まえ、違法と評価されなくなる（違法性が阻却される）のではないかという問題である。①につき、Vの死亡が、Dにとって違法な行為であったと判断されれば、次に②（Dの過失の有無）が検討される。②では、将来のありうる法益侵害（Vに相当する他者の死亡）を抑止する観点から、S, M, AWは、当該結果を回避するために如何なる措置を執っていれば、過失なしとして免責されるべきかが、問題となる。

以上は、L4で走行する自動運転車に係る問題状況であるが、同様の整理は、L3で走行していた自動運転車との関係でも、TO前にV死亡が生じていた場合には、妥当する。前述したように、

L3=L4 and above (TO前) + L2 and below (TO後)

だからである。V死亡が、TO後に生じた場合、TOによりOがDになるため、当該Dの刑事責任が問われる。Vの死亡結果が、Dにとって回避可能であったならば、過失運転致死罪の成否が問題となるが、TO後は結果の発生が直前に迫っており、自然人Dの運転により結果を回避しえたと認められない事例も多いであろう。そうした事例では、Dにつき、V死亡という結果を回避する可能性が否定され、因果関係が認められないため、Dの刑事責任は否定されることになる。

## 2-3 国際的な議論の状況

### 2-3-1 従来議論の概観

以上の前提を踏まえ、（設例1）のような、L4で自動走行していた車両に係

るトロリー問題（及び類似の事例）において、①の点、即ち、保護された生命数（設例1）では5人の生命）と、侵害された生命数（設例1）では1人）との比較を踏まえ、運転の違法性が阻却されるかという問題につき、如何なる議論が展開されているかを確認する。

倫理的判断が、刑法上の違法性判断に直結することはない。しかし、（設例1）では、生じた結果に係る直感的、倫理的判断を踏まえつつ、違法性を判断する主体と、判断基準を議論する傾向が、国際的には認められる。そこで、違法性の判断主体と、判断基準を区別し、それぞれの見解の組み合わせにより、如何なる議論が展開されているのかを整理する。

違法性（阻却）の判断主体に関しては、自動運転車の走行の違法性を(1)トップダウンで評価しようとする見解と、(2)ボトムアップで評価しようとする見解との区別が可能である。それぞれの立場の下で、違法性の判断基準として、更に(1-1) 行為功利主義（Act Utilitarianism）型と(1-2) 義務論（Deontology, or Kantian perspective）型が想定される。

（設例1）の解釈として、(1-1)によれば、5人の生命を保護するために1名の生命が侵害されており、結果として4名の生命が保護されたと言えるから、社会全体にとって有益な結果が生じているとして、V死亡に至ったAVの挙動は違法とは評価されないことになる。

他方で、(1-2)によれば、人間の生命は最も重要な法益であり、他人がこれを侵害することは、（設例1）のような緊急事態でも許されないから、当該AVの運転者（D）に相当する者の行為は違法とされる。この見解によれば、S, M, AWとVとに親族関係等の密接な関係が認められる場合に、S, M, AWの違法行為に係る責任が、例外的に阻却されうるに過ぎないことになろう。<sup>(11)</sup>

違法性（阻却）の判断主体に関する(2)（ボトムアップ型）の下では、(2-1)（ADSのアルゴリズムを構成する）AIの機械学習（Machine learning: ML）の程度を踏まえて、結果の倫理的正当性、ひいては違法性を判断する見解や、<sup>(12)</sup>

(11) ドイツ刑法第35条1項が前提とする発想である。

(2-2) アンケート調査から得られた人々の意識に基づき、当該結果の倫理的正当性、ひいては違法性を判断する見解<sup>(13)</sup>が主張されている。

（設例1）の解釈として、(2-1)によれば、当該AVがODDを走行中に、（設例1）のように歩行者が自動運転車専用道に立ち入る事例が、どの程度、想定可能であったのか、想定された事態に直面した場合、自動運転車の走行速度を減速するなどの対応が可能であったのか、これらの措置が理論的には可能であってもADSのML能力の限界故に、MLを通じた安全走行に係る確率の向上が困難であったのかが問われることになる。

他方で、(2-2)によれば、（設例1）では、(i)単純に生命数を比較し、結局、4名の生命が保護されているから違法ではないとする回答、(ii)C1-5は、子供であるが、ルールに反して道路上に立ち入っており、一定の制裁を被ってもやむを得ないとして、AVとしては、C1-5に衝突するべきであったとの回答、(iii)仮にC1-5が違法に道路に侵入していたとしても、無関係のVの生命を犠牲にすることは、やはり許されないと<sup>(14)</sup>の回答が、示される可能性がある。この調査に基づき、如何なる回答の割合が大きいのかを、AVの自動走行が想定される地域毎に分析し、地域毎にアルゴリズムを微調整するという対処法が提案されることは、予想されることである。

### 2-3-2 新たな事例設定

以上のような調査に加えて、近時は、モリー問題（Molly problem）と命名された事例に係る意識調査が、国際的に展開されている。これは、(2-2)に整理される試みであるが、トロリー問題とは異なる設定下での調査であり、そこで得られた意識の分析が注目される。具体的には、次の事例と設問が予定され

(12) Amitai Etzioni and Oren Etzioni, Incorporating Ethics into Artificial Intelligence, J Ethics (2017), DOI 10.1007/s10892-017-9252-2.

(13) Bonnefon et.al, The Moral Machine experiment, Nature volume 563, p59-64. この分析として、今井猛嘉「AIと倫理的判断—自動運転技術の社会的受容性に係る一つの課題」ヒューマンインタフェース学会誌 Vol.21 No.2 2019, 5頁以下。

(14) 今井, 前掲「AIと倫理的判断—自動運転技術の社会的受容性に係る一つの課題」5頁以下。



(15)  
ている。

(シナリオ 1)

モリーという若い女性が、道路を一人で横断しようとしていた際に、誰も乗車していない（無人の）自動運転車（unoccupied self-driving vehicle）が彼女に衝突した。この事件の目撃者はいない。このシナリオを評価する視点は何か。

ここでは、L4以上で走行する無人の自動運転を実現しているソフトウェアの性能と、それが選択した挙動に係る意識の調査がなされている。設問は、多数用意されているが、回答者の9割以上が以下の点につき期待を示している<sup>(16)</sup>。即ち、当該ソフトウェアが、

- ・衝突を感知すること
- ・衝突の手前で（当該車両を）停止させること
- ・危険状況（a hazard）を他の道路利用者に対して指示すること
- ・緊急事態での挙動選択（emergency services）につき警告を発すること
- ・何時衝突するのかを意識させる機能を有すること
- ・どこで衝突するのかを意識させる機能を有すること
- ・どこで衝突の危険が確認されたのかを意識させる機能を有すること
- ・モリーを検知したら、その旨を意識させる機能を有すること
- ・モリーが、何時、検知されたのかを意識させる機能を有すること
- ・モリーが、人間として検知された場合には、その旨を意識させる機能を有すること
- ・モリーが、人間として検知されたのは何時であったのかを意識させる機能を有すること
- ・モリーに与える衝撃を緩和する挙動を車両に採らせたかを意識させる機能を

---

(15) AI Safety & Ethics for Self-Driving, Introducing The Molly Problem (<https://aiforgood.itu.int/events/ai-safety-ethics-for-self-driving-introducing-the-molly-problem/>)

(16) FG-AI4AD-The Molly Problem Panel (preliminary results) v. 1. 5. (筆者も、FG-AI4ADのメンバーであり、この調査、分析に携わっている)

有すること

- ・ 衝撃を緩和する挙動が、何時採られたのかを意識させる機能を有すること
  - ・ 衝撃を緩和する挙動として、何が選択されたのかを意識させる機能を有すること
- である。

以上からは、回答者は、予期された衝突とその被害軽減に係る情報取得につき高い選好を有することが確認された。そこで、次のシナリオが追加されて質問がなされた。

（シナリオ 2）

衝突が直前で回避された（ニアミスであった）場合でも、当該事態に係る情報を伝達する機能が当該ソフトウェアにおいて利用可能であるべきか。

この設問につき、回答者の 88% は、当該ソフトウェアに、ニアミス情報を伝達する機能を期待していると回答した。また回答者の 73% は、そのような情報伝達機能を有していないソフトウェアを用いての運転は禁止されるべきだと回答している。

以上の回答から、現時点での結論として、i) 衝突事例、ニアミス事例を通じて、自動運転用のソフトウェアに対して、倫理的に正当な挙動を選択することへの期待、ii)（選択された挙動を）AI が説明可能であることへの期待、iii) AI が如何なる挙動を選択するかが予測可能であることへの期待が高いことが明らかになった、とされている。

モリー問題では、トロリー問題とは異なり、L4 以上で走行する無人の自動運転車が、人間に被害を及ぼしうる場合でも、トロリー問題におけるのと同様に、ECA が期待されているとの結果が判明しつつあると言える。事故回避への期待は、ADS を構成するアルゴリズム（あるいは、AW）に向けられており、AI 刑法の基本問題（AI の挙動を制御する手段として刑罰の利用は適切か、という課題）に直結するデータの分析が始まっていると整理できる。本稿では、この基本問題の検討は行えないが、トロリー問題の背景を理解する上で、重要な示唆を与えるものである。そこで、以上の回答を踏まえ、ECA を、誰が

(判断主体)、如何なる内容のものとして設定する(判断基準)のかに係る問題点を検討する。

### 3 検討

#### 3-1 検討のための視点

ECAを巡る以上の議論を踏まえ、先ず、その判断主体につき検討すると、(1)(トップダウン型)により一定の方向性を決めた後に、判断基準の具体化を図るために(2)(ボトムアップ型)で得られた反応を考慮することが、適切であるように思われる。(2)での調査結果は、貴重である。自動運転車がディレンマ状況に遭遇した場合の走行態様の選択につき、回答者の属する国や地域、回答者の人種、性別、年齢等の変数に応じて、多様な意見があることが明らかになっている。<sup>(17)</sup>このビック・データを活用して、自動運転車が走行する地域、そこで予想される被害者の属性や数を踏まえ、当該地域に特有のECAが用意できるのであれば、それに自動運転車の挙動を委ねることは、1つの現実的対処法であろう。その場合でも、ECAが走行地域毎に(自動的に)変化することは、(MLについてと同様に、少なくとも現時点では)想定できない。また、ディレンマ状況が発生した地域や被害者の属性に関わらず、最低限の基準を提示しなければ、(刑事)責任を追及されるとのAW,Mの不安を解消することはできないであろう。そこで(1)に係る大枠(国際的に共有される準則)を決める必要<sup>(18)</sup>がある。

(1)を(2)に優先させる場合、(1)の下での判断基準としては、(1-1)行為功利主義の立場と(1-2)義務論の立場とが、考えられる。(1-1)は、行為の結果が、

(17) 今井、前掲「AIと倫理的判断—自動運転技術の社会的受容性に係る一つの課題」6頁以下。

(18) *Mirja Feldmann*, *Automatisiertes und vernetztes Fahren, Auswirkungen der Neuregelungen im StVG auf die strafrechtliche Verantwortlichkeit bei Unfällen und Überlegungen zur Debatte um die Lösung von sog. Dilemmasituationen*, GA2020, p692.

当時の状況下で最大の効用（U）をもたらす場合、当該行為の選択を推奨する立場である。そこでも行為時の選択が問題となるが、（Uを最大化するという）結果を想定して行為が選択させるため、（結果を踏まえての行為選択が重視されない）（1-2）の立場とは、立論の基盤が異なり、両者を同じ平面で評価することには困難がある。そこで、（1-1）と（1-2）の対立を踏まえつつ、（1-1-2）利己主義（Egoism）と（1-2-2）利他主義（Altruism）という、立論の基礎により類似性が認められる2つの立場を想定して、検討を続けることにす<sup>(19)</sup>る。

（設例1）で、（1-1-2）利己主義に依拠するアルゴリズムにより自動運転がなされている車両は、ディレンマ状況を打開するため、Vに衝突し、死亡させることを選択するであろう。これに対して、（1-2-2）利他主義に依拠するアルゴリズムにより自動運転がなされている車両は、C1-5への衝突、Vへの衝突、いずれをも選択せず、当該車両とその乗員Oの生命を犠牲にすることで、C1-5及びVの生命を保護するべき車両挙動を選択する可能性<sup>(20)</sup>（the self-sacrifice algorithmへの選好）が高くなろう。これら二つのタイプのいずれかのアルゴリズムからなるAIにより自動運転がなされている複数の車両だけが走行している道路空間<sup>(21)</sup>で、（設例1）の事態が生じると、（1-1-2）利己主義型の選択による効用が大きいため、（1-1-2）型の選択が、走行中の自動運転車にとって支配的戦略（dominant strategy）となり、全車両が、各自の判断においてこれを選択し（private ethical setting.PES）、結果として囚人のディレンマ状況<sup>(22)</sup>（prisoners dilemma）が生じる。こうして、（1-1-2）型の選択は適切

(19) 「(1-1) 行為功利主義→(1-1-2) 利己主義」、あるいは「(1-2) 義務論→(1-2-2) 利他主義」という論理的関連性が認められると主張するものではない。自動運転車に係るディレンマ状況に関しては、国際的には極めて多様な見解が主張されている。そこで、それを整理する一つの有益な座標軸として、このような（緩やかな）関連性を前提とするものに過ぎない。

(20) 例えば、急ブレーキを掛けて停車した後、後方を確認し、他車両の接近がない場合には高速で後進することも、Oの生命が侵害される危険を生じさせるものであり、the self-sacrifice algorithmの1つとして数えることができる。

(21) この道路空間は、走行中の個々の車両の、それぞれのODDの範囲内にある。

でなかったことになる。このような事態の推移が容易に予測されるので、(1-1-2) 利己主義に基づく PES を自動運転車（そのアルゴリズムを作成する AW）に許容してはならない。そこで、最適な選択肢は、（利己主義以外の）一定の倫理的判断をアルゴリズムに強制すること（mandatory ethical setting.MES<sup>(23)</sup>）になる。

それでは、AW（ひいては、M）に設定が強制される MES の内実、即ち、衝突アルゴリズム（mandatory crash algorithm. MCA）は、如何なるものであろうか。これは、判断基準としての ECA を問うものであり、難問であるが、本稿では、現時点での方向性を示すことを試みたい。

### 3-2 過失の理解

第1に、MCAに頼らざるを得ないディレンマ状況の出現が、相当の確率で予想される場合、そのような事態出現の可能性を生み出す AI に制御された自動運転車は、本来、公道で利用されてはならない。相当の確率で、そうした状況に遭遇することが予想される AI を搭載した自動運転車は欠陥商品であり、その走行により損害が発生した場合、M は製造物責任を問われることになる<sup>(25)</sup>。  
う。

(22) V2V が実現され、他の AV に係る ADS アルゴリズムが、別の AV により認識可能となれば、囚人のディレンマ状態には至らない。しかし、この想定は未だ現実的なものではないため、本文の想定で検討を続ける。

(23) Jan Gogoll & Julian F. Müller, Autonomous Cars In Favor of a Mandatory Ethics Setting, Science and Engineering Ethics volume 23 (2017), p681-700.

(24) Rolf Johansson and Jonas Nilsson, Disarming the Trolley Problem Why Self driving cars do not need to choose whom to kill, 2016, HAL Id: hal-01375606. は、AV を制御する AI（のアルゴリズム）が ISO26262 を充足するほど高機能のものであれば、トロリー問題の発生確率は非常に低く抑えることができるとする。ML の影響（上記 (2-1) の観点からの情報分析の成果）は、ML の発展を待って、考慮されるべきであろう。

(25) これは、AI ないしプログラムが搭載された動産も、全体として1つの「製造物」であるとする、現在の一般的理解に依拠した説明である。動産と、そこに搭載された AI とは別物であるが、AI 自体を「製造物」と評価できるのか、製造物責任法が、自動運転車に係る事故予防のインセンティブを与えうるのかは、別途、検討されるべき重要な課題である。

他方で、（L4で走行する）AVがもたらす様々な社会的効用（U）を踏まえると、MCAの作動可能性をゼロにしなければAVの公道での利用を禁止すべきであるとまでは言えない。MCAの作動可能性が高くなかったが、実際にはMCAにより損害が生じた場合、当該損害が、ODD内における事故として予見可能であった者には（その者がDであった場合には）過失運転致死傷罪、又は（その者がD以外であった場合には）業務上過失致死傷罪等の刑事責任を問うことで、将来の同種損害の発生を抑止するべきであろう。

これら二つの過失犯は、（設例1で言う）Vの死亡をAWやMが予見可能であったのに、予見しなかった点を過失として非難する類型である。<sup>(26)</sup>

予見すべきであったのは、犯罪が社会に及ぼす損害（法益侵害が発生する確率と当該法益侵害の重大性の積）に照らして、これを防止するために費用を支出する余地が残されていた（法益侵害を防止する他の手段を尽くす余地があった）という事実である。これは、（法益侵害という）結果を回避する可能性を基礎付ける事実であり、因果関係の判断と重なる事項である。しかし、追加的費用の支出により結果発生を回避し得るとの事実を予見すれば、当該費用支出をする動機形成が期待されたであろうから、当該事実の認識又は予見（としての故意）が欠けている事案で、過失を認定するためには、この事実関係を、予見可能性の対象事実として考慮することには、意義がある。

この意味での、過失において予見対象となるべき事実は、ハンドの原則に限<sup>(27)</sup>界理論による修正を加えることで導くことができる。<sup>(28)</sup>この事実は、犯罪の客観的要件の主要な要素（結果回避可能性を基礎付け、行為の違法性を導く要素）である。その確定は、民事の不法行為の客観的要件を導く作業と基本的には同<sup>(29)</sup>

---

(26) Sは、Mからの情報を顧客に伝達するに過ぎないから、Sが固有に発見した、危険実現の可能性が認められない限り、Sに業務上過失致死罪等の（主観的要件としての）過失を認めることはできないであろう。

(27) *U.S. v. Carroll Towing*, 159 F.2d 169 (2d Cir. 1947).

(28) *Cooter and Ulen*, *Law and Economics*, 6ed (2014), p187.; クーター・ユーレン（太田勝造訳）『法と経済学』（新版、1997年）369頁以下。

(29) 窪田充見『不法行為法—民法を学ぶ』第2版（2018年）58頁以下。

じであり、両者の相違は、刑法の解釈では「損害」が、刑法上の保護に値する法益に限定される点に認められる。

具体的には、

$$\text{Marginal Cost} < \text{Marginal (possibility} \times \text{Damage)}$$

が、結果回避可能性を算定する基準であり、この関係を充たす事実が、予見の対象となるべきものである。

ここで

Damage とは、保護法益の侵害であり、(設例1) では、V の生命侵害である。

Possibility とは、V の生命が侵害される確率である。

Cost とは、M (設例1) では、AW や M が、 $p \times H$  を防止するために投入した費用である<sup>(30)</sup>。

Marginal とは、その限界値 (を指す形容詞) である。

第2に、この立場からは、AW が、アルゴリズム作成時に、また M が当該 ADS が設定された自動運転車を製造し、その販売を S に委託する際に、合理的な主体として期待される限りの情報を収集<sup>(31)</sup>、分析しディレンマ状況の出現可能性の最小化に努めた (そのために費用を支出した) が、当該状況が発生した場合、法益侵害という結果が生じた以上、M や AW の行為は違法ではある。しかし、彼等にとって以下の事実関係しか予見可能でなかったならば、(彼らが D として認定されなかった場合には) 業務上過失致死罪の要件としての過失が、あるいは (彼等が D として認定される場合には) 過失運転致死罪の過

(30) 本稿では、Administration Cost 等は、議論を単純化するため、考慮していない。

(31) 情報収集義務も、上記フォーマットに照らして導かれるべきである。

失が、否定される。

その関係とは AW や M としては、  
Marginal Cost > Marginal (possibility × Damage),  
少なくとも、  
Marginal Cost = Marginal (possibility × Damage)  
と評価されるだけの費用投入をした、との事実である。

以上の理解に対して、ドイツでは、「許された危険 (das erlaubtes Risiko)」が (法益侵害という) 結果へと実現した場合には、過失が否定されるとの理解が、有力に主張されてきた。何が「許された危険」であるかは、文化的、歴史的な社会背景を踏まえ、最終的には社会的相当性により判断されるのであり、cost-benefit analysis (費用便益分析) の手法だけで、この危険を判断することはできない、とされている。<sup>(32)</sup> そこでは、危険が規範的概念として把握されている。しかし、(自動車の利用や新たな医療手法の利用のような) 技術の発展に伴う法益侵害の事案で過失の有無を判断する際には、それら技術を用いて得られるメリットとデメリットが比較され、当時の技術の状況 (state of art) に照らし、メリットがデメリットを上回っているときには、法益侵害に至っても、「許された危険」の実現とは評価されず、過失が否定されるとも、説明されている。<sup>(33)</sup> ここでは、実際には、cost-benefit analysis が主たる基準として用いられていると言わざるを得ないであろう。この限りでは、「許された危険」の法理による過失の理解は、本稿で示した理解と同様であるが、規範的な衣を纏うことで判断手法を不明確にしない点で、本稿の定式化の方が、過失の事後的検証可能性を高めていると言えよう。

---

(32) Ex, *Jakobs*, *Strafrecht Allgemeiner Teil*, 1991, 7/36 with reference to BGHZ 24, 21 (26 et seq.).

(33) *Hilgendorf*, *supra* note10, p85-87.



### 3-3 違法性阻却の判断手法

第3に、CAに関しては、発生したディレンマ状況において、CAにより衝突対象とされる客体は、自動運転車の走行によりディレンマ状況に遭遇する危険があることを予見して交通に関与した者（関与者。involved.I）が乗車していた車両に限定すべきである。関与者以外の者（非関与者。uninvolved.U）は、衝突の対象となる危険を予見しておらず、そのような者の犠牲により、ディレンマ状況に陥ったAVが事態の打開を図ることは、倫理的直感に反するであろう。<sup>(34)</sup> Iは、Uとは異なり、M（ $p \times D$ ）を踏まえ、当該事故による損害（D）発生を回避するべく、MCを算定することが可能だから、損害回避に向けた費用を最小化する観点からは、CAによる衝突対象は、（衝突に備えた費用支出が可能で、その結果として、被害が相対的に小さいものと予想される）<sup>(35)</sup> Iであるべきである。（設例1）のVのように、ODD外にいた歩行者は非関与者（U）であるから、Vへの衝突は、<sup>(36)</sup> 禁止されるべきである。

ここで、改めて、法的判断と道徳的直感とを比較しておこう。（設例1）では、Vに衝突しなければ、Oの生命は保護されなかったとの前提の下、Vの生命は侵害されたが、5名（C1-5）の生命が保護されたので、社会全体として

(34) Dietmar Hübner and Lucie White, Crash Algorithms for Autonomous Cars: How the Trolley Problem Can Move Us Beyond Harm Minimisation, Ethical Theory and Moral Practice 21 (3) (2018) p685-698.

(35) この理解からは、ディレンマ状況に遭遇した（L4で走行中の）自動運転車が、Dを最小化するため、ヘルメットを被ったバイク・ライダー（ $\alpha$ ）か、ヘルメットを被っていないバイク・ライダー（ $\beta$ ）のいずれかに衝突せざるを得ない状況に直面した場合、 $\alpha$ がCAの対象として選択されることにはならない。ヘルメットを被ることで交通事故の被害を最小化しようとする努力の有無に関わらず、 $\alpha$ 、 $\beta$ は、自動運転車が走行する領域であることを知って交通に関与しているから、いずれもIである。よって、いずれも、MCAにより衝突対象として選択されうるが、 $\alpha$ を衝突対象として選択するのは倫理的に不当だと理解が一般的であるならば、これを反映する結論を採ることが、政策的に要求されるであろう。

(36) よって、（設例1）では、Vの生命侵害は違法であり、M又はAWがDと評価された場合には、過失運転致死罪の、Dと評価されない場合には、業務上過失致死罪の過失の有無が、上記フォーマットに照らして検討されることになる。

は効用が生じていると評価され、刑法上は、違法性が阻却されるとの理解を導きうる。これは、行為功利主義に基づくものであり、緊急避難（刑法 37 条）の文言解釈として無理のない理解であるが<sup>(37)</sup>、自動運転車に係るディレンマ状況との関係では、想定される不公平感を考慮し、限定解釈を行う余地はあろう<sup>(38)</sup>。行為功利主義では集団的正義（collective justice）が問題とされるのに対して、義務論では個別的正義（individual justice）が考慮される傾向にあるが、ディレンマ状況では、個別の被害者への共感が一般に大きく、義務論的発想が強調され易い。そこで、そうした反応（ないし市民意識）を踏まえ、緊急避難の適用範囲を、法益を侵害される者が関与者（I）か非関与者（U）で区別し、U の犠牲の下での緊急避難は否定することは、可能な限定解釈である<sup>(39)</sup>。これにより、ディレンマ状況の解決についての社会的受容性を高め、L4 技術の利用を促進することが、妥当であるように思われる<sup>(40)</sup>。

ドイツ刑法典の緊急避難の規定（34、35 条）を、義務論を踏まえて理解すると、関与者（I）の生命を侵害してディレンマ状況を打開する行為には、緊急避難の成立を認めることはできない。即ち、人間の生命は法治国家の基礎であり（ドイツ基本法第 1 条第 1 項）、その侵害は決して正当化されず、少数の者の生命侵害を、（多数の者から構成される）社会全体の利益保護のために正当化すると全体主義に至るから容認できない、との解釈である<sup>(41)</sup>。この理解は一

---

(37) 今井猛嘉「AI 時代の刑事司法」、『罪と罰』56 巻 2 号（2019 年）23 頁。町野朔『刑法総論』（2019 年）291 頁をも参照。

(38) 5 名（CI-5）の生命を保護するために、彼らと無関係で、彼らを危険な状況に置いたことに何ら関与していない他人（である V）の生命を侵害した以上、当然に、その行為は違法であるべきだとの直感的反論。

(39) 罪刑法定主義の法律主義には抵触しうる（佐伯仁志『刑法総論の考え方・楽しみ方』（2013 年）181 頁脚注 19）。これは、刑法 37 条の規定態様に由来する問題であるが、緊急事態の典型例を条文化する作業も困難であり、特定の事例群だけを視野に入れた解釈の試みは、行う価値があると思われる。

(40) この理解は、あくまで、自動運転車に係るディレンマ状況に対処するための試論である。他の緊急避難が問題となる事例群では、原則として、行為功利主義的理解が採られるべきである。

(41) Hilgendorf, supra note10, p67.; Feldmann, supra note18, p690.

貫したものであり、個人主義に基づく法治国原理の基礎として、尊重されなければならない。しかし、自動運転に係るディレンマ状況の解決策としては、上記解決策が、政策的に妥当であろう。その理由は、第1に（繰り返しになるが）、IとUとでは、衝突の危険に対する予防的措置を準備する可能性に違いがあり、そのこと（損害回避に向けた費用支出の差異）が衝突のダメージ軽減に影響を及ぼすこと、第2に、（以下で確認するように）、自動運転車のディレンマ状況の解決においては、一般人の倫理的判断を考慮する必要性も高いこと、に求めることができる。

### 3-4 違法性阻却事由

第4に、本稿の理解に基づく、想定されるディレンマ状況と、これに予想される一般的回答を踏まえ、IとUを区別して、MCAの適用範囲を確定することが必要になる。

この作業を行う際には、ドイツの倫理委員会が示した指針<sup>(42)</sup>が、参照されるべきである。

そこでは、ディレンマ状況を予想し、倫理的に正当な結論を導くべく事前の対処をすることは困難であり、まずは独立した公的な機関で事例の検討が必要だとされている<sup>(43)</sup>。事例からのボトムアップにより指針を模索すべき場合があるとの認識は、本稿の(2-2)と同様の発想である。しかし、ドイツ倫理委員会では、義務論が行為功利主義の観点よりも重視されている<sup>(44)</sup>ことに因る制約が生じる。義務論からは、例えば、(設例1)でAVがVに衝突して死亡させたことは違法であり、例外的に、S、M、AWいずれかの責任<sup>(45)</sup>が否定されるに過ぎない。

(42) Ethik-Kommission, Automatisiertes und vernetztes Fahren, Bericht (2017).

(43) 指針8 (Nummer8 der Ethische Regeln für den automatisierten und vernetzten Fahrzeugverkehr)。

(44) 指針2, 6。

(45) ドイツでは、AVを犯罪の行為者とする理解は、学説上、まだ意識されていないことが多く、(設例1)では、Sら自然人の犯罪性が検討されることになる。

義務論からは、

（設例 2）「1 名の生命を保護するために、他の 1 名の生命を侵害せざるをえなかった緊急状態での行為」

は、当然に、違法性阻却が否定され、保護しようとした者が行為者の親族であったような例外的な場合に責任が阻却され（殺人罪等の）犯罪が不成立となるに過ぎない（ドイツ刑法 35 条 1 項<sup>(46)</sup>）。この結論の前提として、（設例 1）で緊急避難の適用を考える際には、緊急避難の要件である危難の現在性（V の生命侵害が差し迫っていること）を、M、AW にも認めることが必要だが、M、AW が予想もしていなかった V との関係で、V の生命侵害の現在性を認めることは、そもそも困難である<sup>(47)</sup>。

（設例 1）や（設例 2）の日本法の下での解決としては、行為功利主義を基礎に据えろと解させる緊急避難（刑法 37 条）の適用を、上述した政策的限定付きで、試みるべきである<sup>(48 49)</sup>。

---

(46) 超法規的緊急避難は、その成立要件が不明確であり、自動運転車に係るディレンマ状況に適用するには適ささないとと思われる。

(47) 町野・前掲『刑法総論』（2019 年）294 頁をも参照。As an opposing opinion, cf. RGSt 60, 318.

(48) 指針 9 は、法益侵害が回避不可能な状況に遭遇した場合に、1) 被害者数を減少させるプログラムは、正当化されうるが、2) そうした状況に関与した者が、非関与者を犠牲にしてはならない、とする（ナンバリングは筆者による）。これは、ディレンマ状況における MES としての ECA のあり方として、正当な理解である。しかし、1) で行為功利主義の観点が、2) で義務論の観点が採用されるべき根拠は、指針 2 や 6 に見られる基本的発想からは、容易には導けないように思われる。*Hilgendorf*, supra note 10, p. 69. は、複数の者の生命を保護するために、より少数の者の生命を侵害することは、ドイツ基本法第 1 条 1 項及びこれに基づく緊急避難の理解から、常に違法であるが、侵害される生命数をより少なくすること（lesser evil への指向）は求められ、より少ない生命侵害の場合は、それだけ違法性の程度（degree of wrong of the act）が減少するとする。この理解は、彼が言うように、量刑事情としては機能するであろう。しかし、違法性の程度を問題にすることは、行為功利主義の観点を踏まえなければ導けない。結局、そこでは、義務論と行為功利主義とが併用されていることになるが、行為功利主義を採用可能な根拠は示されていない。

(49) 指針 18 は、AI の Self learning system は、車両の制御に係る機能の安全性の要請を充たさない限り、利用されてはならないとする。これは、Self learning system が、その作成者（AW）の予想を超えた情報処理をしかねないことへの懸念の表れである。Self learning により

実はドイツにおいても、本稿と同様の行為功利主義的発想に基づく見解が、自動運転に係るディレンマ状況の解決策として提案されている。それは、VR (virtual reality) を利用して、想定事例に係る関与者 (I) と、非関与者 (U) の行動を予測し、分析するものであり、(2-2) に分類される試みである。そうした研究の1つは、次の仮説を検証するために行われた。<sup>(50)</sup>

仮説①：ディレンマ状況に直面した自動運転車の中にいた者は、当該自動運転車によって衝突させられる可能性のある人間（潜在的被害者）の内、実際に衝突させられる者の数を最少にするべく、行動を選択する。

仮説②：潜在的被害者の中で、より若い者は、衝突の対象から除外される傾向にある。

仮説③：潜在的被害者が車道にいる場合と、歩道にいる場合では、車道にいる者への衝突が選択され易い。歩道にいる者は交通に関与していないので、衝突対象とされるべきではない。

仮説④：潜在的被害者が、車道にいる子供と、歩道にいる大人である場合、衝突対象として、前者が選択され易い。

仮説⑤：当該自動運転車の中にいる者は、自己の生命を犠牲にして (self-sacrifice) ディレンマ状況を回避するとの選択をすることもあるが、それは、他者に当該自動車を衝突させた場合の損害が相当に大きいと予想される場合に限られる。

VRによる実験の結果によれば、仮説①、②については、その予想の確かからしさが実証され、仮説③、④については、潜在的被害者が歩道にいるか否かは、被験者において重視されておらず、④の場合でも年齢が考慮され、歩道にいる大人が衝突対象として選ばれる傾向が認められた。仮説⑤については、潜在的

---

アルゴリズム自体が改変されることは（少なくとも当面の間は）想定できないが、Self learningの結果として、どのような事故が（AWの予想を超えて）認知され、回避されたかを探ることは、(2-2) ボトムアップ型でのアプローチにおいて、必要な作業だと思われる。

(50) *Faulhaber et al*, Human Decisions in Moral Dilemmas are Largely Described by Utilitarianism: Virtual Car Driving Study Provides Guidelines for Autonomous Driving Vehicles, *Sci Eng Ethics* (2019) 25, p402.

被害者が5人以上（即ち、5人以上の生命が当該自動運転車の衝突により侵害される）と予想される場合には、当該車両内にいる者は、当該車両を崖から転落させるなど、自己の生命を侵害すること（self-sacrificing）により、損害発生を回避する傾向が強いことが認められた。

この実験では、車内にいた者が、衝突が予想される場合に車両の挙動を制御できることにもなっており、L4だけではなく、TO後のL3に相当する事態までが想定に入れている。とはいえ、そこで得られた結果は、L4との関係でも有意義である。それら回答結果を総合すると、自動運転車の衝突により何名かの人間の生命侵害が不可避である事態に直面した場合、O（ないし、OからTOによりDになった者）は、侵害される生命数を最少にすべく行動を選択する傾向にあり、潜在的被害者が歩道にいるかどうかは重視されず、彼（女）の居る場所よりも、子供のように、より若い者が潜在的被害者である場合には衝突対象から除外される傾向があること、また、5人以上の者の生命を救うためには自己犠牲も厭わない傾向が確認されたと言える。

この実験結果は、被験者の間では行為功利主義的判断が優勢であったことを示している。<sup>(51)</sup> そうした倫理的判断は、自動運転車の挙動を制御するアルゴリズムの内容として、強制的に導入されるべきである（obligatory ethical setting<sup>(52)</sup>の内実とされるべきである）と指摘されているが、本稿も、この考えに同意する。<sup>(53)</sup> 但し、歩道にいる者は、それが大人であれ子供であれ、自動運転車に係る交通に関与していない者（U）であるから、彼らを衝突の対象として選択するアルゴリズムの設定には、更なる検討が必要である。

この実験結果から推論された被験者の意識は、緊急避難に関するドイツ刑法34条、35条の前提と、実質的には合致しない。そこで、ディレンマ状況に係

---

(51) *Faulhaber et al*, id, p407. 例えば、誰かに衝突しなければ、より大きな損害を回避できない事態において、自動運転車が何人に衝突することが望ましいかとの設問に対しては、複数の人間に対する衝突と傷害の惹起を避けるために、一人の人間に衝突し傷害を負わせることが選択されるべきとの回答が、多数を占めている。

(52) これは、MESに相当する。

(53) *Faulhaber et al*, id, p410.

る人々の意識と、法制度との整合性が将来の検討課題である点も指摘されている<sup>(54)</sup>。緊急避難の立法趣旨が、自動運転車に係るディレンマ状況において社会的に受容されないものだとすれば、人々の意識に合致した制度的解決が模索されるべきだと主張である。本稿も、同様に理解するが、法改正が実現されない間は、所定の条文の限定解釈が求められるものと考ええる。

## 5 展 望

自動運転車の利用が、社会に大きな効用をもたらすと言われ始めてから、既に一定の時間が経過した。自動運転車が交通事故を起こすことなく安全に利用されるには、更なる技術の発展が望まれるが、将来においても、そうした事故発生確率を0にすることはできない。そこで、事故が発生した場合でも、被害(ないし損害)を最小限に抑える施策が必要である。そうした施策の多くは、工学的に実現されることが予想されるが、法制度として、事故時の責任の内容と帰属を明確にすることは、自動運転車の利用者に対して、ありうる不利益を告知しつつ自動運転技術の利用を図るために、必要な作業である。

本稿は、そのための1つの視点として、事故が不可避となった場面(ディレンマ状況)での対処法につき、理論的観点(緊急時の行動選択の判断主体と、判断基準は如何にあるべきか、その理解を踏まえると緊急避難の適用範囲は如何なるものであるべきか)、数的処理の観点(刑法上の過失の理解、その理解の、自動運転車事例への適用)及び、人々の倫理観の把握(その統計分析の視点)から、検討を加えた。

本稿で主張した内容の妥当性は、今後の検証に委ねられているが、検証に先

(54) *Faulhaber et al*, id, p411. 但し、設問で想定された自動運転車の走行速度は低速であり、衝突によって人間が死亡する危険が相対的に低い場合であったことは、留意されるべきである。義務論や、ドイツ刑法学の通説的見解によれば、このVRによる思考実験は、基本法1条1項の要請に反しない形で施行されており、緊急避難の解釈を変更する契機とはなりえない、との反論が予想される。これに対しては、自転車を崖から転落させる自己犠牲事例により反論可能である。

立ち、自動運転車の事故に係る調査委員会の役割への期待が高まるところである。本稿は、(2-2)の観点（ボトムアップ型の手法で得られたアンケート調査を利用し、個々のAVに想定されるODD内での典型的事故を選び出し、その防止に係る技術的対応を尽くした上で、それでも残された事故可能性に如何にして対処するかを検討）も重視している。この検討は、公的機関において、多くの事故関係証拠を収集した上で、科学的知見を活用した検討を尽くすことで、初めて可能になる。その実現のためには、AW、Mや（自動車工学や法律学の）専門家に加えて、被害者団体の代表者も参加し、自動運転車の社会的効用を維持しつつ被害の最小化を図るための検討を行い、同種事故の再発に向けた指針を策定し公表することが、効果的だと思われる。法律家にも、科学的及び統計的知見の活用が求められている。

（本稿の脱稿後に、ドイツ道路交通法改正案が閣議決定されたとの報に接した。改正案では、倫理委員会の提案をも踏まえ、注目すべき多くの規定が提案されている。その検討は、別稿にて行いたい。）