

技術とトロッコ問題

—自動運転車の技術倫理—

Technology and Trolley Problem: Ethics in Technology of Self-driving Cars

南 雲 功*

Isao NAGUMO

Abstract : With the development of artificial intelligence, ethical thought experiments, which have been virtually discussed, are emerging as a problem of technology in the real world. For example, in the case of The Trolley Problem, it is a problem of who the self-driving car should sacrifice in a situation where sacrifice is unavoidable. This issue has been discussed in various fields such as ethics, law, psychology, and brain science. The trolley problem is an ethical problem for the field. If a situation such as the trolley problem is assumed at the design stage, it is the principle of design to take measures to prevent even one victim from occurring.

キーワード : トロッコ問題, 設計基準, 技術倫理, 人工知能 (AI), 自動運転車

1. トロッコ問題とは

多数を助けるために少数を犠牲にすることの是非については、「カルネアデスの板ⁱ」など古代ギリシアより論じられている。トロッコ問題の嚆矢は、1967年にイギリスの哲学者Foot, Philippaによって、提起された思考実験である (Foot, P., 1978)。Foot, P. は妊娠中絶に関し、母体の生命が危険な状況にあっても、宗教上に理由で妊娠中絶が禁止されているカトリック教徒の実情に対し、様々な事例を挙げて、倫理的考察を行っている。トロッコ問題も彼女の提起した事例のひとつである。

1.1 Foot, P. が提起した問題

Foot, P. が提起した、多数を助けるため少数者を犠牲にすることの是非を問う、主な事例を表1に示す。

Foot, P. は、中世以来の倫理判断基準である「二重結果論ⁱⁱ (The Double Effect)」に対して、

* なぐも いさお 准研究員・放送大学在学

表 1 Foot,P. の事例

事例	概 要 #
【事例 F1】	洞窟探検隊が、出口で太った男が岩に挟まり動けなくなった。彼がどかなければ全員が洞窟から出ることができない。探検隊がダイナマイトで太った男を爆殺することの是非。
【事例 F2】	裁判官が無罪と判断している人間を死刑にしなければ、暴徒により住民の多くが犠牲となる。無罪の被告を死刑にすることの是非。
【事例 F3】	住民の多い地域に墜落しかけている飛行機のパイロットが、人口の少ない地域に舵をきることの是非。
【事例 F4】	暴走している路面電車が直進すれば線路作業員 5 人が死亡し、右の引き込み線にポイントを切り替えれば作業員 1 人が死ぬ。運転士のポイントの切り替えに対する是非。
【事例 F5】	重症患者が治癒する分の薬がある。後から現れた軽症患者 5 人は重傷患者 1 人分の薬で 5 人が救われる。薬の投与先の選択。
【事例 F6】	健康な一人の青年の犠牲で、臓器移植を待つ 5 人が救われる。青年から臓器を取り出すことが許されるか。

概要 (出典)

概要は Foot,P. の提起事例の著者による要約 (Foot,P.,1978)

新たな判断基準として、助ける行動をとる「積極的な義務 (positive duty)」と行動を控えることにより助ける「消極的な義務 (negative duty)」を対比する判断基準を提案した (Foot,P., 1978)。Foot,P. の説を岡本裕一郎は「二重義務論」と名付けている (岡本裕一郎, 2017)。**【事例 F4】** ではどちらでも「消極的な義務」であるため、犠牲の少ないほうを是とする。ところが、**【事例 F6】** では、「積極的な義務」と「消極的な義務」の衝突となるため、「消極的な義務」を優先して青年を犠牲にすることを非とする。

1.2 Thomson, Judith Jarvis のトロッコ事例

【事例 F4】 は路面電車の運転士がポイントを操作ⁱⁱⁱする事例であり、この問題をトロリー問題 (The Trolley Problem) として定式化したのが Thomson,J.J. である。路面電車の問題に関する事例をいくつか比較し、その是非について議論している (Thomson,J.J.,1985)。これらの問題を日本ではトロッコ問題と呼ばれることが多い。

Thomson,J.J. が提起したトロッコ問題の事例を表 2 に示す。

他にも、乗客が当事者となる場合や、ループ線の場合などを挙げているが、問題が複雑になるので、本論においては下記 3 例を代表的事例として取り上げた。

表 2 Thomson,J.J. が提起したトロッコ問題

事例	概 要 #
【事例 T1：運転士】	Foot,P. の 【事例 F1】 と同じ状況で暴走した路面電車の運転士がポイントを操作する。
【事例 T2：通行人】	たまたま通りかかった通行人が、 【事例 T1：運転士】 と同じ状況でポイントを操作する。
【事例 T3：突落し】	レールの分岐はなく、線路上に 5 人の作業員がいる。橋上で暴走電車を見た通行人が隣の太った男を突き落として電車を止める。

概要 (出典)

概要は Thomson,J.J. の提起事例の著者による要約 (Thomson,J.J.,1985)

2. トロッコ問題に対する諸学説

トロッコ問題に対して、倫理学をはじめ、心理学、脳科学、法学など様々な学問分野で、それぞれの目的に沿って議論されている。

2.1 倫理学上の議論

Foot,P.は、【事例 T1：運転士】について、二重義務論から、一人を犠牲にすることを是としている（Foot, P.,1978）。Thomson,J.J.は生存の権利が功利性に優先するという立場から以下の論考を試みている。【事例 T2：通行人】では通行人に殺す意図がなく、5人を救う行ための結果として1人が犠牲になるのではとし、【事例 T3：突落し】の太った男を突き落とす事例では、殺す意図があることから、生きる権利が侵害されるとして非とした（Thomson,J.J.,1985）。ただし、岡本裕一郎は、Thomson,J.J.が【事例 T2：通行人】について2008年の考察で通行人が自己犠牲となる選択肢も考慮し、一人の他者を犠牲にすることは、自分の身代わりとなることから結論を変えていると述べている（岡本裕一郎. 2017）。また、Sandel, Michaelは、【事例 T1：運転士】を是とし、【事例 T3：突落し】を非とする一般通念に対して、倫理的判断の難しさを議論している（Sandel,M.,2009.p.21-24）。この議論は、規範倫理学の義務論と功利主義の対立ともなる^{iv}。

2.2 心理学的調査

心理学の立場からトロッコ問題の倫理感に対する無作為抽出によるアンケート調査も行われている。Cushman,Fiery;Young.Liane;Hauser,Marcらは、①行動、放置②手段としての犠牲、結果としての犠牲③身体接触の有無についてアンケート調査を行っている（Cushman,F.;Young.L.;Hauser,M.,2006）。また遠藤薫は、【事例 T2：通行人】と【事例 T3：突落し】の比較において、条件変更も含めて日本、米国、中国の比較を行っている（遠藤薫. 2019）。さらに、山本翔子らによる国別の文化の違いによる倫理観の調査も行われている（山本翔子, 結城正樹. 2019）。以上の報告では、様々な状況でどのような倫理的判断が行われるかの心理的分析を目的としているが、本論では、以上の報告の中から、Thomson,J.J.の事例のみについて比較し、表3にまとめた。

表3 トロッコ事例ごとの国別アンケート調査比較のまとめ

	事例 T1：運転士	事例 T2：通行人	事例 T3：突落し
日本	89.6% ⁽¹⁾	50.8% ⁽⁴⁾	20.0% ⁽¹⁾
アメリカ	84.4% ⁽²⁾	68.2% ⁽⁵⁾	17.1% ⁽²⁾
中国	78.6% ⁽³⁾		21.2% ⁽³⁾

#各事例において多数を助けるために少数を犠牲にする行動をとるとする人の比率
#調査条件および（出典）

(1)日本全国の20代～70代の男女：サンプル数 5000：インターネットモニター調査：2019.3（遠藤薫. 2019.）

(2)アメリカの20代～70代の男女：サンプル数 500：インターネットモニター調査：2019.3（遠藤薫. 2019.）

(3)中国の20代～70代の男女：サンプル数 500：インターネットモニター調査：2019.3（遠藤薫. 2019.）

(4)募集広告の応募者：サンプル数 124：インターネットクラウドソーシングサイト（Lancers）（山本翔子, 結城正樹. 2019.）

(5)募集広告の応募者：サンプル数 132：インターネットクラウドソーシングサイト（Amazon Mechanical Turk）（山本翔子, 結城正樹. 2019.）

その結果【事例 T1：運転士】【事例 T2：通行人】【事例 T3：突落し】の順に一人を犠牲にすることに消極的となる。このことから、倫理に対する人間の判断の多様性が示唆される。

2.3 脳科学、心理学からの論考

【事例 T2：通行人】では是としながら【事例 T3：突落し】では非とする人間の判断基準の矛盾について、Green, Joshua は、心理学および脳科学の立場から、二重過程脳説として、理性と情動のせめぎあいにより、異なる倫理的判断を説明している (Green,J.,2015)。【事例 T3：突落し】では、情動により人を殺すことに罪悪感が働くのに対し、【事例 T2：通行人】では理性に基づき功利主義的な計算により多数を助けるため、ポイントを切り替えることができるというものである。Singer,Peter も Green,J. の説を擁護し、義務論と功利主義に対立する規範的倫理の再考を提唱している。(Singer,P.,2005)

2.4 法学的諸説

法学の分野では、トロッコ問題よりも、現実の自動運転車事故に対する法学的解釈が行われている。自動運転車には、法的人格がないとする説が、通説となっているが、有責性について、議論が分かれている。また、緊急避難がどこまで許されるかの問題もある。この点について後述する。

3. 自動運転車におけるトロッコ問題

3.1 事例

自動運転車がトロッコ問題のような状況に遭遇した場合について、様々な分野で論じられている。現時点での議論の多くは AI 自ら倫理的判断ができる「強い AI」^vではなく、あくまで人間がプログラミングした通りに作動する「弱い AI」に対するものである。すなわち、事故に対して人間があらかじめ、倫理的判断をおこない、設定しておくべきものとしての AI である。自動運転車に関わるトロッコ問題には表 4 のようなものがある。

トロッコ問題として様々な状況が考えられることから、自動運転車の緊急時の設定をどのようにプログラミングするかが問題となる。小林正啓は、法学の立場から「優先順位をあらかじめ決めておく (小林正啓, 2017)」必要があるとしている。法律の立場では被害者への賠償や刑事責任など現実に起こる問題に対応しなければならない。また設計者に過度の負担を掛けないための配慮でもある。富川雅満は、プログラミング設計者の刑事責任などについて、ドイツの法制度を紹介し日本の法制度と比較している (富川雅満, 2017)。そこでは、ドイツでは生命の損失を許

表 4 自動運転車に関わるトロッコ問題

事例	概 要 #
【事例 A1】	自動運転車の前に突然暴走車 (トラック) が進路変更して現れた。直進すれば、乗員全員の死亡とトラック運転手の負傷、左は断崖で乗員全員の死亡、右にハンドルを切れば乗員が助かりが、1 人の歩行者をひき殺す。
【事例 A2】	事故が避けられない状況で路上の 10 人か、それとも歩道の 1 人のどちらを犠牲にするか。

概要 (出典)

【事例 A1】の概要は (小林正啓, 2017) 事例の著者による要約

【事例 A2】の概要は (河島 茂生, 北村 智, 柴内康文, 2017) 事例の著者による要約

容するハードルは高く、日本では低いと報告している。このように、法学では、被害者の保護と並んで、法的責任の所在が問題となる。すなわち、車の所有者、同乗者、製造者、AIの責任分担が議論されている。

3.2 事例に対する考察

倫理的判断は、多様な状況により様々に変わってくる。1.2項における Thomson,J.J. の事例【事例 T1：運転士】～【事例 T3：突落し】と 3.1 項の自動運転車事例【事例 A1】、【事例 A2】について、その原因、犠牲者の所属、人数について比較する。原因については、【事例 T1：運転士】～【事例 T3：突落し】は、電車の整備不良が原因であり、加えて5人作業に監視員を置かなかったという教育の不徹底、待避所を作らなかったことなど、鉄道会社に責任がある。自動運転車の【事例 A1】では暴走車が事故原因であるのに対し、【事例 A2】では歩行者の注意不足が原因となる。この場合、自動運転車は、法令を遵守し周囲に対し十分な注意を払っているものとする。法令を遵守していなければ、自動車の製造者が有責となる。所属に関しては、【事例 T1：運転士】、【事例 T2：通行人】の作業者は鉄道会社の関係者であり、【事例 T3：突落し】の太った男はまったく無関係である。【事例 A1】、【事例 A2】の歩行者は、自動運転車とは無関係であり、同乗者は、自動運転車より便益をうけている関係者となろう。このように、事故に対する責任などの関わりにおいて、様々な状況が発生する。すべての事故に同じ状況はないであろう。以上より優先させることは、人数だけとは限らないといえる。

4. 論理的考察

4.1 条件

トロッコ問題の前提条件を検討し、数学的論考を行った。いずれの事例でも以下の条件を前提としている。

「前提条件：すべての人間に「有限」で「平等」な価値がある。」

人間1人の価値を x ($0 < x < \infty$) とすると、トロッコ問題の価値の不等式は

$$x < 5x \quad (1)$$

となる。式1が成り立つのであれば、トロッコ問題は解決される。すなわち1人の犠牲により五人が救われることが、あらゆる場合において数学的に正しいことになる。前提条件だけで判断するのであれば、Foot,P. や Thomson,J.J. の事例は、すべての事例において少数の犠牲にして多数を助けることは、正しい。しかし、この「有限」「平等」とする前提の是非が問題となる。次に、この条件を除外した場合について検討した。

4.2 論考

まず人の価値は「平等」であるが、「有限」とは限らないとすると、1人の価値と5人の価値が等しい場合があり得る。

$$x = 5x \quad (2)$$

方程式(2)は実数において3個の解をもつ。すなわち

$$x = 0 \quad (3)$$

$$x = \infty \quad (4)$$

$$x = -\infty \quad (5)$$

となる。解(5) $x = -\infty$ は、人間がすべて滅亡することが、地球の最大幸福に最も貢献することであり、例え真実であったとしても、倫理的考察を無意味にすることになるので、本論では採用しない。

解(3) $x = 0$ は、自然主義の立場である。人は、多様な生物の中の一つの種であり、進化の過程で出現した偶然的存在者である。そのような存在者に意味や価値を認めない。倫理や道徳は人間の脳活動がつくりだした幻想にしかすぎないとする立場である。トロッコ問題においては犠牲が1人であろうが5人であろうが同等である。

解(4) $x = \infty$ は、人間主義の立場であり、一人の生命は、何物にも代えることができない価値がある。人の無限の価値には、個人の価値が平等であることも超越し個々人の差に意味がない。この場合もどちらを犠牲にしても同等であるが、(3)と(4)とではその重みが異なる。解(3)では失うものはなにもないが、解(4)では無限の価値を失う。

つぎに、人間の価値は「有限」であるが、「平等」とは限らないとする。個人の価値が、普遍的標準により定められるのであれば、功利主義において最大幸福量^{vi}を計算でき、だれもが納得できるであろうが、人の価値は個人や集団の立場により相対的に価値が異なり、計算の根拠が定められない。

$x_n = 0$ と $\sum_{n=1}^5 x_n$ の比較となる。変数 x_n の n は全世界の人口数を取りえるので、解が (人口の6乗)/720 通り存在する^{vi}。1人の価値と5人の価値の総和の比較は、組合せにより異なった解となる。Thomson, J.J. が5人の作業員がマフィアである事例を提起している。すなわち、5より1が価値をもつ場合がある (Thomson, J.J., 1985)。この場合、人間の価値は平等ではなく負の場合もあるということである。

法律や経済の実務の場では賠償責任の算定、保険料算定などで、人間の価値を有限で個々に異なる数値として算定されている。自動運転車が事故直前に、犠牲になる可能性のある人を特定して、その人の価値を、即座に世界中のデータから判断して、進路を決めることが可能であろうか。また遺族や社会通念として容認されるであろうか。さらに、法哲学の立場で、無辜の歩行者を、自己救済のために他人を殺すことを意図した設計条件が、許容されるであろうか。

5. 設計条件

5.1 設計対応と現場対応

緊急時の技術対応として、設計時に緊急時を想定して設計条件に織り込む設計対応と、現場の判断により対応する現場対応がある。設計で十分検討されていないことに対し、現場が補っているのが現実である。設計の制御が及ばない事態に対し現場の担当者が、自分の技術的倫理的判断により対応している。

自動運転車の場合、現場対応する運転手が不在のため、すべての事態を設計時に織り込む必要がある。すべてが、設計者の意図された通り制御されていないなければならない。

5.2 トロッコ問題は現場対応である

トロッコ問題が発生するのは、想定された状況が制御不能となる場合である。意図された管理状態を逸脱した状態で現場の担当者が倫理的判断が要求されるのである。衝突時に犠牲となる人

数により優先順位をつけることを予め設定することは、制御不能にも関わらず、いかにも制御されているかのように設計に織り込むことであり、適切な倫理判断とはいえない。

6. 設計段階でトロッコ問題が想定される場合の設計条件

技術の設計段階で、想定される危険はその頻度に応じてすべて解決策がもとめられる。もし、トロッコ問題のように、倫理的判断を必要する事例が頻繁に起こりうるとするならば、現行車も含めて自動車の使用を禁止しなければならないであろう。しかし、トロッコ問題が発生するのは、現実には非常に稀なケースである。頻度の少ない事例は多様な状況が現出する。このような場合、個々の事態を想定し対応することは不可能である。そこで原則に基づいた対処をとることが要請される。設計時点での原則について著者が提案するのは、1. 人命尊重、2. 法令遵守、3. 緊急時の機械停止である。

6.1 人命尊重

歩行者、運転者のいずれも人命を最優先にしなければならない。命に優先順位をつけてはならない。しかし、現実には死亡災害が多発している。特に医療の現場では、命の選択を避けることができない。しかし、少なくとも設計条件としてまったく無関係の少数者を意図して殺害することを織り込むことはあってはならない。

6.2 法令遵守

現在の事故の多くの原因は、人間の不注意による法令軽視と運転ミスによるものである。さらにスピード超過、一時停止無視、駐車違反車による見通し不良などの意識的違反が常習化している。これらは、法令が遵守されれば、多くの事故が防げることが推測される。人の判断に依拠しない自動運転車は、予め設定したこと以外の行動基準がないため、いっそう法令および規則を遵守しなければならない。法令を順守できない自動運転車は設計ミスである。他車が法令無視する事故もあるが、自車が法令順守することで、事故の被害を最小にすることができる。現在の道路システムが人間の随意性（南雲功、2020）に深く依存しているという問題があり、すべての車が法令順守した場合、スムーズな交通を阻害するような矛盾が露呈されるかもしれない。その場合には法令を改定しなければならないであろう。事故責任は法令違反者が引き受けることになる。今後、自動運転車を利用する上で、法律もそれに合わせて改正されていなければならぬだろう。その上で法令遵守により、自動運転車起因の事故はほとんど起こらないはずである。

6.3 緊急時の原理

緊急時といえども他者を犠牲にして自車の同乗者を助けること許されるのであろうか。自動車に対して歩行者は弱者である。すべての自動運転車が緊急避難として自車の同乗者を守り、弱者の歩行者や他車を犠牲にする挙動をとるならば、ホップスの、自然状態ⁱⁱⁱとなる。緊急時の自動運転車はどのような行動をとるべきであろうか。「機械安全規格を活用して災害防止を進めるためのガイドブック」（中央労働災害防止協会、2015）によれば、異常時の機械の原則は、以下である。

- (1) 本質安全の原則 ⇒ 危険源を除去する、又は人に危害を与えない程度にする
- (2) 隔離の原則 ⇒ 人と機械の危険源が接近・接触出来ないようにする

(3)停止の原則 ⇒一般的に機械は止まっていれば危険でなくなる

(1)、(2)が設計段階および交通システムについて充分検討されなければならない。その上、道路交通法を遵守したとしても、判断不能な状況が発生した場合、まず停車することが原則である。ただし飛行中の航空機や、トンネル内の列車火災のようにあきらかに停止が大きな事故に結びつく場合はこのかぎりでない。緊急時の自動運転車が判断停止状態になったときは、速やかに停車することが原則である。

7. 安全に対する交通機関の対応

7.1 鉄道におけるトロック問題

現実の技術が、トロック問題をどのように解決しているのだろうか。トロック問題の現場事情だけに捉われて全体を見なければ、ジレンマから抜け出せない。俯瞰的で総合的な解決を探るため、Thomson,J.J. のトロック問題の発生原因および異常行動を列挙する。

【原因1】列車の暴走

【原因2】安全対策のない線路作業

【原因3】運転士が意識をうしない乗客の運転代行。(本論で取り上げなかった運転士が気絶した場合)

【原因4】通行人がポイントを切り替えることができる。

【原因5】橋上から人を突き落とすことができる。

などである。そして、現代の日本の鉄道において、トロック問題が技術的には解決されている。以上の不安全状態に対する現実の鉄道における対策には以下ようになる。

【対策1】列車には2種以上の緊急停止装置があり、運転ができなくなる列車の暴走が起きる確率は非常に低い。

【対策2】線路作業時に列車見張り員の配置等が義務付けられている。もちろん一人作業は禁止されている。線路を横断するときにも指差呼称が義務付けられている。さらに、鉄橋、トンネル内にも退避設備が設けられている。

【対策3】新幹線では制限速度の超過や、停止駅で運転士が意識を失った場合、列車を停止するシステムが働く^{ix}。在来線でも赤信号を無視して進行すれば停止する。この場合乗客は運転室に入ることはできない。2019年12月にドイツのボンで路面電車の運転士が気絶したため、乗客が停止させた事故が発生させしている。日本でもバスの運転手の気絶に対し、乗客が停止させたことがある。しかし、乗客が公共交通機関を運転することは危険がある。ホーム上にある緊急停止ボタンによる緊急停止などの対策が適切であろう。本来はシステムで停止すべきであろう。

【対策4】路面電車をふくめた日本の鉄道で、通行人がポイントの操作をできる個所は皆無であろう。大きな事故の原因にもなるし、避けるべきである。

【対策5】日本の多くの線路上の橋には転落防止柵が設けられているので、突然人を橋上から突き落とすことはできない。

航空機では、さらに厳しい安全対策が取られている。このような安全対策は、長期にわたる多くの犠牲者の上に確立されてきたものであり、結果的には、トロック問題のような現場における倫理問題を、より包括的に解決するものである。その背景には一人の犠牲者をも出すことが許されないという技術における設計段階での根本原則があるからであり、いかなる場合であっても、

犠牲者を許容する設計条件を認めるべきではないのである。このような努力の積み重ねにより、京急踏切事故や、JAL 機同士のニアミスなど一つ間違えば大事故になった事例はあるもの、鉄道では羽越線脱線事故（2005年12月25日）、航空機ではJAL 御巣鷹山墜落（1985年8月12日）以降日本では事故起因による乗客の死亡事故は起きていない。

7.2 道路交通の問題

自動車の安全は、運転者の技量にかなりの部分が依存しており、人命を最優先に考えられてこなかったのではないだろうか。例えば、アクセルとブレーキが同じ右足で踏み込むという動作が、ブレーキの踏み間違いの原因となる。自転車、バイク、鉄道、航空機など加速と停止は異なる操作となっている。また、自動車のブレーキ系がフェールセーフ^xになっていない。さらにオートマチック変速車のギアが、前からP, R, Dでいいのだろうか。これはバックとドライブを間違える原因ではないのか。また、死角の存在など改善点が多々ある。

事故の原因が運転手の不注意によるものとされ、真の原因まで追求することなく、従って、設計条件にまで、落とし込んでこなかったのではないだろうか。現在、様々な技術が開発されてきたことから、もう一度人命最優先の立場で、個々の自動車から総合交通システムまで見直す時ではないだろうか。

たとえば安全体系のための既存技術としては、すでに以下のような道具がある。

GPS、緊急停止装置、無線ネットワーク（アクター間の情報交換）、眠気センサー

これらの組み合わせにより、自動運転車だけでなく、現行車にも緊急ブレーキの設置ご義務付けることは可能であろう。すると、現代の鉄道の緊急停止ボタンと同様に、危険暴走運転車を強制停車させることができよう。道具を有効に利用し、交通システムとして安全対策が取られることにより、トロッコ問題を回避できるのではないだろう。

8. 結語 — それでも事故は起きる

事故を完全になくすことはできない。現在の安全は多くの犠牲の上に講じられている。安全に対する技術は日々の小さな異常現象を観察しながら改良されている。小さな異変に気付き原因を究明し、対策を立てることにより安全が維持されていく。このような技術者倫理に加え、基本的人権の一つである生存権の保証の立場からも、無辜の歩行者の生命を奪うことを前提にすることは、法の理念にも反する。従って、二者択一的な状況に至らない総合的システム設計が要請されるのであり、安易に犠牲を容認する設計条件を設計時点で設定すべきではないであろう。トロッコ問題で一人の犠牲者もださないことが設計の原則となる。

文献

- 中央労働災害防止協会. 2015. 機械安全規格を活用して災害防止を進めるためのガイドブック
- Cushman, Fiery; Young, Liane; Hauser, Marc. 2006. The role of conscious reasoning and intuition in moral judgment :Testing three principles of harm . Psychological Science, 17, p.1082-1089.
- 遠藤薫. 2019. 「論説：AI化する社会と倫理的ジレンマ」学習院法務研究所, 14, p.1-13
- Foot, Philippa. 1978. The Problem of Abortion and the Doctrine of the Double Effect in Virtues and Vices (Oxford: Basil Blackwell, 1978 : 1967. Oxford Review, No. 5. Included in Foot, 1977/2002 Virtues and Vices

- and Other Essays in Moral Philosophy. (Minor stylistic amendments have been made.).
- Green, Joshua . 2015. 竹田円訳. モラル・トライブズ. 岩波書店
- 河島 茂生, 北村 智, 柴内康文. 2017. 自動運転車の「トロッコ問題」などに関する意識：日本に居住する人に対する質問紙調査を通じて. 社会情報学会 (SSI) 学会大会 プログラム.
<http://gmshattori.komazawa-u.ac.jp/ssi2017/wp-content/uploads/2017/07/10.pdf>. (2020.10.22 閲覧)
- 小林正啓. 2017. 自動運転車の実現に向けた法制度の課題. 情報管理, 60 (4), p.240-250
- 南雲功. 2020. AI と観念的諸概念の関係. 文教大学生生活科学研究, 42, p.107-112
- 岡本裕一郎. 2017. 一冊の本. 朝日新聞出版, 22 (8), p.22-26
- Sandel, Michael . 2009. Justice: What's the Right Thing to Do?. Farrar Straus & Giroux (邦訳) マイケル・サンデル. 2011. 鬼澤忍訳. 『これからの「正義」の話しよう』早川書房
- Searle, J. R.. 1980. Minds, brains, and programs. Behavioral and Brain , Sciences, 3, p.417-457
- Singer, Peter. 2005. Ethics and Intuition. The Journal of Ethics , p.331-352
- Thomson, Judith Jarvis . 1985. The Trolley Problem. Yale Law Journal, 94, p.1395-1415
- 富川雅満. 2017. 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」. 千葉大学法学論集, 32, (1・2). p.157-185
- 山本翔子, 結城正樹. 2019. トロッコ問題への反応の文化差はどこから来るのか?. 社会心理学研究, 35 (2), p.61-71

註

- ⁱ カルネアデスの板とは、古代ギリシアの哲学者カルネアデスが提示した問題。難船の後、一枚の板に一人がさがりついている時、助けを求める者がきたらどう行動するかを問う。(広辞苑第七版. 岩波書店)
- ⁱⁱ 二重結果説とは、他人を死に至らしめる意図には、殺意を持った「直接意図 (direct intention)」と別の目的を達成する行為の結果として他人を犠牲にすることが明らかな「間接意図 (oblique intention)」に分け、間接意図により死に至る行為は状況により許容しようとするもの (Foot,P.1978)。二重結果説は、トマス・アクィナス (1225-74) までさかのぼる哲学上の問題 (Green,J.,2015.289)。
- ⁱⁱⁱ ポイント切り替えを運転士が行うことについて奇異に感じることもあろうが、実際の路面電車のポイントの切り替えには様々な形式があり、日本でも長崎や広島などの路線が複雑な都市では運転士がポイントを切り替える方式である。但し、暴走している場合のポイント切り替えは困難である。
- ^{iv} 規範倫理は事前の意志を問題にするカント, イマヌエル (1724-1804) の義務論と、事後の結果の予測を重視する功利主義に大きく分けられる。
- ^v (社)人工知能学会では「人間の知能そのものをもつ機械を作ろうとする立場、もう一つは、人間が知能を使っていることを機械にさせようとする立場」(人工知能学会. 2019) という二つの立場を挙げている。Searle,J. R.は前者を「自由意志」を持つ「強い AI (strong AI)」、後者を人間の知的活動を補佐する「弱い AI (weak AI)」と称している (Searle,J.R.1980)。(南雲功. 2020)
- ^{vi} 最大多数の最大幸福とは、ベンサム, ジェレミ (1778-1832) が導入した、功利主義のスローガン。倫理的判断は全体の幸福を最大にする選択をすべしとする。
- ^{vii} 厳密には人口 n から 6 を選ぶ組み合わせは $\frac{n!}{6!(n-6)!}$ となるが、6 に比べ n は極端に大きいことから $\frac{n^6}{6!}$ に近似できる。
- ^{viii} ホブズ, トマス (1588-1679) の自然状態とは、自然状態に生きる人間はいかなる法の支配下にもなく、専ら自己保存の基本的権利を追求し、そのために最善手段と判断されるあらゆることに自由 (権利) を有している (思想・哲学事典. 岩波書店)
- ^{ix} JR 西日本山陽新幹線が 2003 年 2 月 26 日岡山駅で運転士が約 8 分間も居眠り運転をしたにもかかわらず、安全装置により停車した。
- ^x フェールセーフ (fail safe) とは操作しない状態や、電気系、油圧系に支障が生じた場合、停止する設計になっている。例えば、機械式バネなどによりブレーキが利いている状態を通常とし、走行時に油圧などでブレーキを解除するシステム。