

## いま改めて考える「自動運転車のトロッコ問題」

藤居 学（AIG 総合研究所 主任研究員）

自動車による痛ましい事故が続いています。池袋では 2019 年 4 月、高齢者の運転する自動車の暴走により 12 人が死傷し、滋賀県・大津では同 5 月に、T 字交差点において直進車に対向右折車が衝突する「右直事故」により、保育園児および保育士 16 人が死傷しました。事故に逢われた被害者の方々に、心より哀悼の意を表します。

最近では、こういった自動車事故のニュースのたび、自動運転の早期実現を希望する声もあがります。実際のところ、池袋の事故・大津の事故、いずれも、当事車両がすべて自動運転車であったと仮定するならば、事故は回避されていたと考えられます。

ただ、私たちが自動運転の普及やそれに伴う自動車事故の防止といった問題を議論する場合、そのような「路上に存在するすべての自動車が自動運転車になっている**未来**」をいきなり想定すると、重要な論点を見落としてしまうこととなります。そうではなく、私たちがまず考えなければならないのは、「自動運転車と従来型のマニュアル運転車が混在する**近未来**」についてです。

本コラムでは、大津の事故に見られるような「交差点右直事故」を例にとり、自動運転という革新的テクノロジーが浸透していく「**過渡期**」により強く現れると想定される、自動運転車の起こす事故に関するひとつの論点について考えていきたいと思います。

交差点右直事故とは、交差点を直進する自動車に、対向車線から走行してくる右折車が衝突する事故をいいますが、大津事故のように、右折車が直進車の前方進路を妨害するのではなく、直進車の側面に衝突するケースでは、直進車側の操作による事故の事前回避は特に難しいと考えられます。さらに、衝突地点から歩道までの距離も短く、衝突により進路が変わった後に急ブレーキをかけたとしても、交差点内で安全に停止することは困難である可能性が高いでしょう。

このような事故を想定した場合、直進車が完全自動運転だったとしても、対向右折車がマニュアル運転でこのような形で側面から突っ込んできた場合には衝突は避けられず、自動運転車の AI（人工知能）には、衝突により曲げられた進行方向に存在する歩行者への二次事故に対応するための緊急判断が求められる事態となります。

例えば、ここで「急ブレーキのみ」の場合には制動距離が足りず歩行者に被害を与えてしまい、「急ブレーキ+急ハンドル」の場合には歩行者を巻き込まずに済む一方、車両横転によって車内の搭乗者が被害を受ける、という状況だと AI が認識した場合、AI はどちらの緊急操作をとるべきなのでしょうか？

このように、何らかの緊急操作が必要な状況に陥った自動運転車が、どのような操作を選択しても被害が避けられない場合に、どう判断すべきかという議論は、有名な倫理学の思考実験との類似性から、自動運転車の「トロッコ問題」と呼ばれます。

自動運転車のトロッコ問題については唯一無二の正解はなく、自動運転車の AI には、想定される個別のケースごとに、もっとも社会的合意の得られる選択を行うことが期待されますが、その「社

会的合意」が具体的にどのようなものであるのかを世界的な規模で調査しているのが、マサチューセッツ工科大学の「モラル・マシーン」というサービス（日本語版は <http://moralmachine.mit.edu/hl/ja>）です。

このサービスでは、自動運転車が直面するさまざまな緊急的ジレンマ状況が提示され、回答者は自動運転車がどのような判断をすべきかを選択します。ネイチャー誌（<https://www.nature.com/articles/d41586-018-07135-0>）に掲載されたモラル・マシーンへの回答の分析や、サイエンス誌（<http://science.sciencemag.org/content/352/6293/1573>）に掲載された関連調査からは、自動運転車に期待される「倫理観」について、いくつもの興味深い事実が浮かび上がってきます。

まず、人々が自動運転車に期待する倫理観は、概ね「功利主義」に近い、ということが示されています。功利主義とは、選択の「結果」を重視する倫理観で、総合的な損害が最小となる選択をすべきという倫理観です。簡単に言うなら、被害者が1人となる選択と2人となる選択があれば、1人となるほうを選ぶべき、というのが、世界の多くの人の意見だということになります。

ところが、さらに続けて、車内外の人を区別せず「被害者数」で判断する自動運転車があったとして、あなたはそれを買いたいですか？ という質問をすると、多くの人が「買いたくない」と答えたのです。つまり、自動運転車を販売する自動車メーカーの立場からは、車外の人よりも車内の人の安全を優先するように AI を調整しないと自動運転車が売れない、しかし自動運転車のオーナー以外の人々からはそのような AI は反倫理的であると批判される、という社会的ジレンマが生じる可能性が高いことを、この調査は示しています。

さらに、国によってジレンマ状況に対する倫理観が大きく異なることも示されています。国ごとの比較（<http://moralmachineresults.scalablecoop.org/>）を見ると、日本は世界平均に対し大幅に「反功利主義的」（被害者の「数」よりも属性を世界一重視し、「操作によって被害者を切り替える」ことを避ける傾向も強い）であり、また「歩行者」と「法令遵守者」を優先して守るべきという意見が極めて強い（それぞれ世界第1位、第4位）など、欧米諸国などともまったく異なる独自の倫理観をもっていることがわかります。このような事実は、自動運転車の AI の倫理についての国際基準を作ることが極めて難しいことを示唆しています。

そもそも、個々の自動運転車の AI の倫理観は、誰が設定すべきなのでしょう。国が法令により定めるのか、自動車メーカーが決めるのか、それともオーナーがカスタマイズできるようにするのがいいのでしょうか。そして、事故が起きた場合に、被害者が AI の設定により「選ばれた」と判明した場合、誰がどのような責任を負うべきなのでしょう。

さらにいえば、もしも自動運転車ごとに AI の「倫理観」が異なっているという状況が生じた場合、そのような様々な倫理観が混在する自動運転車が走る路上というのは、歩行者や自転車、さらには従来型のマニュアル運転の自動車にとって、不気味で安心できない空間になってしまいかねません。

自動運転車は、これからの新しい社会を実現していく技術革新、AI の社会への応用事例の筆頭ともいえますが、その実現にあたっては、単なる AI の技術論にとどまらず、今回ご紹介した「トロッコ問題」のように、法・制度論（責任論を含む）や哲学・倫理論まで、幅広い領域の専門家・関係者を巻き込んだ深く横断的な議論が不可欠です。近年、「AI の倫理」に関する議論が注目を集め

ようになっていますが、その背景には、いよいよ AI が社会に浸透しつつある新しい時代のなか、技術論や法・制度論の枠に留まらない、より複雑な問題が次々と顕在化していることがあげられるでしょう。

自動運転車におけるトロッコ問題も決して新しい論点ではなく、これまでも継続的に論じられてきたものではあります。令和という新しい時代を迎え、また日々繰り返される痛ましい自動車事故を心に銘じ、改めて議論を深めていくべき時期がきているのではないのでしょうか。

※本ドキュメントは保険もしくはその他一切の金融商品の販売、勧誘を意図したものではありません。また、本ドキュメントは具体的な特定の取引をご提案するものではなく、その実現性を保証するものでもありません。

※AIG 総合研究所（以下「AIG」と呼びます。）は、本ドキュメントの利用あるいは利用の結果に関して、その正確性、精度、信頼性などについていかなる表明および保証も行わないものではなく、その利用の結果については責任を負いません。AIG は、本ドキュメントがいかなる場所においても適切であり利用可能であることを表明するものではありません。AIG は、正確かつ最新の情報を本ドキュメントで提供しよう合理的な努力をしていますが、誤差・脱漏が生じる場合があります。

※AIG あるいは本ドキュメントの企画、作成または提供に関わるいかなる当事者も、お客様が本ドキュメントを利用したことあるいは利用できなかったことに起因する直接的、偶発的、結果的、間接的損害あるいは懲罰的賠償の責任を負うものではありません。

※本ドキュメントに掲載されている内容に関する権利は、AIG および AIG が利用許諾を得た著作権者に帰属します。無断で転用・複製・改変をすることはできません。