

ロボットの道徳的な責任は 使用者や設計者にまで及ぶのか？

Should designers or users of a robot take a moral responsibility for the robot's behavior?

小松 孝徳

Takanori Komatsu

明治大学 総合数理学部 先端メディアサイエンス学科
FMS, Meiji University

Abstract: 日常生活にロボットが導入されることで、それらのロボットが何らかのトラブルを引き起こすことは想像に難くない。特に、ロボットの誤作動ではなく、正常作動時において何らかのトラブルを引き起こした場合、そのロボットの道徳的な責任は誰が負うべきなのか？本稿では、トロリー問題を題材としてこの問題を検討したアンケート調査の結果を報告する。

はじめに

日常生活にロボットが導入されることで、それらのロボットが何らかのトラブルを引き起こすことは想像に難くない。特に、ロボットの誤作動ではなく、正常作動時において何らかのトラブルを引き起こした場合、その際のロボットの道徳的な責任はだれが負うべきなのであろうか。同様の問題意識は、例えば、Google や日産自動車などが開発している自動運転車においても議論されつつあり、法的な責任については製造物責任法 (PL 法) の範疇に含まれると考えられている。しかしながら、自動車よりも我々の日常生活空間に浸透していくことが望まれているロボットの場合、自動運転車とは違った道徳的な観点からの検討が必要となる。さらには、これらのロボットは人型の外観をしていることが想定されるため、適応ギャップ[1]的な視点からも問題を俯瞰する必要があるとも考えられる。

そこで、この問題を検討するため、「トロリー問題」を題材として、とある行為の主体が人間であった場合とロボットであった場合に着目し、その人間の雇用主、ロボットの設計者および使用者の、それら行為主体の行動に関する道徳的な責任の有無についてのアンケート調査を実施した。

アンケート調査

参加者

本アンケートには、理系学部在籍する 141 人の日本国籍を有する大学学部生 (男性 107 人、女性 31 人、無回答 3 人: 19~23 歳、平均年齢 20.78 歳) が参加した。アンケートは参加承諾書と共に印刷物として配布された。参加に対する報酬は用意されなかった。

概要

トロリー問題とは、Foot [1]によって提唱された「ある人を助けるために他の人を犠牲にするのは許されるか？」という倫理学における思考実験課題である。具体的には、以下のようなシナリオが用いられることが多い。

線路を走っていたトロッキの制御が不能になった。このままでは前方で作業中の 5 人の作業員が猛スピードのトロッキに避ける間もなく轢き殺されてしまう。この時たまたま〇〇は線路の分岐器のすぐ側にいた。〇〇がトロッキの進路を切り替えれば 5 人は確実に助かる。しかしその別路線でも作業員の B 氏が 1 人で作業しており、5 人の代わりに B 氏がトロッキに轢かれて確実に死ぬ。

この状態にて、〇〇がどのような行動をするべきかを考えさせるのが、この課題の目的である。

本稿では、行動の選択を迫られている行為主体と

して「作業員の A 氏」「最新型人型ロボット C」の二種類、さらに、上述のシナリオの後に、具体的な行為として「〇〇は分岐器を操作した。トロッコはその進路を切り替えた」「〇〇は分岐器を操作しなかった。トロッコはそのまま直進した」という二種類を設定し、これらを組み合わせた合計四種類のシナリオを用意した。アンケートの参加者は、これらの四つのシナリオのいずれかを体験するという、参加者間計画として無作為に配置された。具体的には、「A 氏が分岐器を操作した」条件に 37 人、「A 氏が分岐器を操作しなかった」に 35 人、「ロボット C が操作した」に 34 人、「ロボット C が操作しなかった」に 35 人が配置された。なお、これらのシナリオにおけるロボットの動作は正常な動作であり、誤作動ではないという注釈を追加した。このシナリオ設定は、トロリー問題においてロボットと人間に対してどのような道徳的規範が適用されるかどうかを調査した、Malle ら[2]の設定とほぼ同じものである。

アンケートではこれらのシナリオの後に、行為主体が人間の A 氏の場合には、「A 氏のこの行動は道徳的に間違っていますか?」「このような A 氏を雇用した企業は道徳的に間違っていますか?」という二問の質問が提示され、それらに対して「間違っている/間違っていない」という二択で回答してもらい、行為主体がロボット C の場合には、「ロボット C のこの行動は道徳的に間違っていますか?」「このようなロボット C を設計した人は道徳的に間違っていますか?」「このようなロボット C を導入した企業は道徳的に間違っていますか?」という三問の質問が提示され、それらに対しても A 氏の場合と同様に「間違っている/間違っていない」の二択で回答してもらった（雇用主、設計者および使用者に関する質問項目は、Malle ら[2]の調査では行っていない、本稿のオリジナルな質問である）。

結果

表 1: 「道徳的に間違っている」と回答された割合

	人間		ロボット		
	A氏	雇用主	ロボットC	設計者	使用主
行動した場合	0.1892	0.027	0.2647	0.2059	0.2647
行動しなかった場合	0.1714	0.0571	0.2286	0.2286	0.2286

アンケートの結果として、「道徳的に間違っている」という回答が得られた割合を表 1 に示す。まず、行為の内容ごとに人間の作業員 A 氏とロボット C が道徳的に間違っていると判断された割合を調査するために、二要因参加者間要因の分散分析を行った（独立変数その 1：行為主体、人間/ロボット、独立変数その 2：行為内容、分岐器を操作した/しなかった、従属変数：行為主体に対して道徳的に間違っている

と回答された割合）。その結果、交互作用および行為主体要因および行為内容要因のいずれにも有意差は観測されなかった（交互作用： $F(1, 137)=0.02$ 、行為主体要因： $F(1, 137)=0.91$ 、行為内容要因： $F(1, 137)=0.15$ ）。ここから、行為主体の違いおよび行為内容の違いは、調査参加者の「〇〇は道徳的に間違っている」という判断に影響を与えていないことが明らかになった。

続いて、A 氏の雇用主とロボット C の使用主が道徳的に間違っていると判断された割合を調査するため、二要因参加者間要因の分散分析を行った（独立変数その 1：使用者、人間の雇用主/ロボットの使用主、独立変数その 2：行為内容、分岐器を操作した/しなかった、従属変数：雇用主もしくは使用主に対して道徳的に間違っていると回答した割合）。その結果、交互作用および行為内容要因には有意差は観測されなかったもの（交互作用： $F(1, 137)=0.34$ 、行為内容要因： $F(1, 137)=0.00$ ）、使用者要因に有意差が観察された（ $F(1, 137)=12.90$ ）。つまり、ロボットの使用主の方が人間の雇用主よりも有意に高い割合で道徳的に間違っていると判断されていたことが明らかになった。

加えて、直接的に対応する条件ではないが、A 氏の雇用主とロボット C の設計者とを比較するため、同様の二要因参加者間要因の分散分析を行ったところ（独立変数その 1：雇用主と設計者、人間の雇用主/ロボットの設計者、独立変数その 2：行為内容、分岐器を操作した/しなかった、従属変数：雇用主または設計者に対して道徳的に間違っていると回答した割合）、雇用主と設計者要因にのみ有意差が確認された（ $F(1, 137)=10.14$ ）。つまり、ロボットの設計者の方が人間の雇用主よりも高い割合で道徳的に間違っていると判断されていたことが明らかになった。なお、交互作用と行為内容要因には有意差は観測されなかった（交互作用： $F(1, 137)=0.00$ 、行為内容要因： $F(1, 137)=0.23$ ）。

さらに、行為主体がロボットの場合にのみに注目して、ロボットが分岐器を操作した場合と操作しなかった場合における、ロボット、設計者、使用主が道徳的に間違っていたと判断された割合を調査するために、二要因混合計画の分散分析を行った（参加者間要因：行為内容、操作する/しない、参加者内要因：行為主体との関係要因：ロボット/設計者/使用主）。その結果、交互作用（ $F(2, 134)=0.18$ ）、参加者間要因（ $F(1, 134)=0.04$ ）、参加者内要因（ $F(2, 134)=0.18$ ）のいずれにおいても、有意差が観察されなかった。

これらの解析結果は、以下のようにまとめられる。

- 行為の主体が人間かロボットかを比較すると、

これらの間に有意差は観測されなかったため、行為主体の違いは、参加者の判断に影響を与えていなかった。

- 人間の雇用主とロボットの使用主とを比較すると、ロボットの使用主の方が人間の雇用主よりも有意に高い割合で道徳的に間違っていると判断されていた。
- 人間の雇用主とロボットの設計者とを同様に比較すると、ロボットの設計者の方が人間の雇用主よりも有意に高い割合で道徳的に間違っていると判断されていた。
- ロボット、その設計者、その使用者という三者を比較したところ、この三者間に有意差は観測されなかった。
- いずれの解析においても、「分岐器を操作したのかしなかったのか」という行為内容は、調査参加者の判断に影響を与えていなかった。

考察・議論

アンケート調査の結果から、同様の行動をする人間およびロボットが引き起こした結果に対して、その人間の雇用者およびロボットの設計者・使用者に対しては、異なる道徳的な判断が下されていたことが明らかとなった。具体的には、人間の雇用主よりも、ロボットの設計者および使用者に対して「道徳的に間違っている」という判断が高い割合で下されていたことが明らかになった。この結果からは、ロボットにはその状況を正確に判断することができないために、道徳的責任の矛先がロボットの設計者であったり使用者に向けられていたのではとも推測できる。

しかしながら、ロボット、設計者、使用者に対するその判断の割合を比較してみたところ、そこに有意差は観測されなかった。そこで、ロボット、設計者、使用者という三者に対して「道徳的に間違っている」という判断した参加者が、この三者に対してどのような判断をしていたのか、その回答内容を精査した。表2は行為内容として分岐器を操作した場合の回答内容、表3は分岐器を操作しなかった場合の回答内容を示したものである。

表2を見ると、13人中9人がロボットCを中心としたパターンを選択していることが見て取れる一方、表3においてはその選択が表2よりもばらついていることが理解できる。このことから、ロボットが分岐器を操作した場合においては、その道徳的な責任はロボットを中心に転嫁されるものの、ロボットが分岐器を操作しなかった場合においては、その責任の所在については様々な理解がされていること

が示唆された。このことから、ロボットが明示的な行動を表出した場合よりも、それらを表出しなかった場合において、その設計者や使用主の責任が問われる可能性があることを示唆している。

表2：分岐器を操作した場合

ロボットC	設計者	使用主	人数
○	○	○	5
○	○	-	0
○	-	○	2
○	-	-	2
-	○	○	0
-	○	-	2
-	-	○	2
○・・・間違っている			

表3：分岐器を操作しなかった場合

ロボットC	設計者	使用主	人数
○	○	○	1
○	○	-	3
○	-	○	2
○	-	-	2
-	○	○	2
-	○	-	2
-	-	○	3

このアンケート調査においては、分岐器を操作する/しないという行為内容自体が、道徳的に間違っているか否かという参加者の判断に与える影響は観測されなかったが、この結果は前述の Malle ら[2]の調査結果とは異なるものであった。Malle らの調査結果では、行為主体が人間の場合、分岐器を操作することが「道徳的に間違っている」と判断される一方、行為主体がロボットである場合、分岐器を操作しないことが「道徳的に間違っている」と判断されると報告しており、ここから、人間とロボットに対しては、異なる道徳規範が適用されている可能性を指摘している。それに対して、本稿で行った調査では、そのような差は観察されず、むしろ、人間に対してであってもロボットに対してでも、同じような道徳的規範が適用されていた可能性が示唆された。

この差異の原因としては、双方の調査における参加者層の違いが挙げられるであろう。Malle らの調査に参加したのは AMT (Amazon Mechanical Turk) での募集に応じたアメリカ在住の方々であり、その一方、本稿の調査に参加したのは、理系(特に情報学)を専攻する日本人の大学生であった。よって、これらの参加者を比較すると、

- それぞれの参加者の生活に根差しているキリスト教や仏教や神道などの宗教観から生じる道徳観の違い。特に、命を左右するような「犠牲」に対する考え方には、大きな違いが存在していると考えられる。
- 専攻している学問についての偏り（さらに、本調査の参加者は、一般的な日本国籍を有する人と比べても逸脱している可能性は否めない）

といった少なくとも二つの大きな相違点が生じているとも考えられる。この相違点から派生して、ロボットなどの最新テクノロジーへの態度の違い（それらを許容するのか、もしくは危険視するのか）にも差異が生じていることが考えられるため、このことも調査結果に影響を与えていた可能性も示唆されよう。

今後、この研究が進むべき方向性として、まずは、ロボットの設計者および使用者の責任についてより深い考察を行うため、本調査で使用した参加者層よりも、より一般的な人々に近い参加者層、つまり専攻する学問に偏りが無いいわゆる一般的な日本国籍を有する人に対して、同様のアンケート調査を実施することを検討している。さらには、トロリー問題以外の、ロボットが従事するより一般的なシナリオでの調査を行うことで、ロボットという存在とロボットに関わる人々が、一般的にどのような存在として認識されているのかを広い視野で捉えていきたいと考えている。

参考文献

- [1] Komatsu, T., Kurosawa, R., and Yamada, S.: How does the Difference between Users' Expectations and Perceptions about a Robotic Agent Affect Their Behavior? *International Journal of Social Robotics*, vol. 4 (2), 109-116, (2012).
- [2] Foot, P.: *The Problem of Abortion and the Doctrine of the Double Effect in Virtues and Vices*, Oxford: Basil Blackwell, (1978).
- [3] Malle, B. F., Scheutz, M., Arnold, T., Voiklis, J., and Cusimano, C.: Sacrifice one for the good of many? People apply different moral norms to humans and robot agents, In *Proceedings of HRI'15*, pp. 117 - 124, (2015).