

# ロボットとは何者なのかを考えるための 日常的モラルジレンマ課題の提案

Proposing a daily moral dilemma task in order to consider what the robots are for us

田畑 緩乃 小松 孝徳

Hirono Tabata and 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:** 私たち人間がロボットをどのような存在として認識しているかを把握するために、トロッコ問題に代表されるモラルジレンマ課題にロボットを登場させるという調査方法が提案されている。しかしながら、一般的なモラルジレンマ課題は、「四人を助けるために一人を犠牲にするべきか」「家族を助けるために強盗をするべきか」といった現実的にはあまり起こらないような状況での道徳的判断を求めるものである。そこで本研究では、我々の生活で日常的に起こりうるような状況を想定した日常的モラルジレンマ課題を提案し、この状況下でのロボットに対するユーザの道徳的判断を把握する調査を行った。

## はじめに

Human-Robot Interaction および Human-Agent Interaction 研究分野などでは、人間とロボットとの共存を目指した様々な研究活動が行われている。その際、人間とロボットとのインタラクションに大きな影響を与えると考えられているのが、「人間がロボットを何者とみなしているのか」という事柄である。この事柄に影響を与えている要因として、例えばロボットの「外見（アピランス）」が挙げられる[1,2]。さらには、ロボットという存在そのものが及ぼす影響も考えられよう。

特に、後者の「ロボットという存在」が人間にどのような影響を及ぼしているかを把握するため、モラルジレンマ課題にロボットを登場させ、そのロボットの行動を人間がどのように評価するのかを調査するという研究が報告されている [3,4,5,6]。モラルジレンマ課題とは、道徳的な価値が拮抗している選択肢を参加者に提示した上で、どの選択肢をどのような理由で選択するのかを観察する課題であり、具体的な課題としてトロッコ問題 [7]が良く知られている。トロッコ問題とは、「線路を走っていたトロッコの制御が不能になった。このままでは前方で作業中だった4人が猛スピードのトロッコに避ける間もなく轢き殺されてしまう。この時たまたまA氏は線路の分岐器のすぐ側にいた。A氏がトロッコの進路

を切り替えれば4人は確実に助かる。しかしその別路線でもB氏が1人で作業しており、4人の代わりにB氏がトロッコに轢かれて確実に死ぬ。A氏はトロッコを別路線に引き込むべきか？」といった問題である。Malleら[3]は、上記のトロッコ問題をベースとしたシナリオと、「A氏」を「ロボット」に置き換えたシナリオとの二種類を参加者に提示し、その際に「A氏」もしくは「ロボット」に対して、「この行動は道徳的に間違っていると思うか？」という道徳的過ち（Moral wrongness）を判断させる質問に回答させた。その結果、参加者は人間のA氏とロボットに対して、異なる道徳的過ちを判断していたことが明らかとなった。具体的には、人間がトロッコの進路を切り替えないことに対しては道徳的に間違いではないと判断された一方、ロボットが進路を切り替えないことに対しては道徳的に間違っていると判断されていたことが実験的に示された。

このようにモラルジレンマ課題を用いて人間とロボットとの関係を考察することで、合理的または理性的な意思決定とは異なる「モラル」という次元での人間の意思決定を把握することが可能となり、「ロボットとは人間にとって何者であるのか」という事柄に対する、人間のより根源的な反応を観察できると考えられる。しかしトロッコ問題に限らず一般的なモラルジレンマ課題においては、「四人を助けるために一人を犠牲にするべきか」「家族を助けるために

強盗をするべきか」といった、現実的にはあまり起こらない抽象的な状況が取り上げられることが多い。そこで本研究では、我々の日常生活で起こりうるような状況を想定した「日常的モラルジレンマ課題」を提案し、この状況下においてロボットに対するユーザの道徳的判断を把握するアンケート調査を行った。このような日常的モラルジレンマ課題を用いることで、将来、日常生活空間で稼働することが期待されている「ロボット」に対して、人間がどのような根源的な反応を示すのかといった事象の把握が期待される。

## アンケート調査

### 参加者

本アンケートには、184人の大学学部生および大学院生（男性135人、女性49人：20～26歳、平均年齢21.52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。

### 日常的モラルジレンマ課題

具体的には、我々の生活で起こりうるような状況を想定した日常的モラルジレンマ課題として、以下のようなシナリオを提案した。

研究室において、ある〇〇が室内にある生ゴミの管理を担当しています。具体的には、決まった時間に研究室を訪れる自動ゴミ回収車に対し、必要に応じて生ゴミの入ったゴミ箱の回収を依頼するという仕事をしています。

回収日に〇〇が生ゴミを確認したところ、中に誰かの家の鍵が紛れ込んでいることを発見します。しかし、[××]、生ゴミの中から家の鍵を取り出すことはできません。

もし〇〇がゴミを捨てなければ、この家の鍵は捨てられないので、持ち主に鍵が戻る可能性があります。そのかわりに生ゴミが処理されないため、研究室内に悪臭・虫などが発生します。研究室を元の状態に戻すためには数日かかり、研究室の多くの学生が困ります。

もし〇〇がゴミを捨てれば、この家の鍵は捨てられ、持ち主が家に入れられない可能性があります。そのかわりに生ゴミが処理されることで、研究室内は清潔に保たれ、研究室の多くの学生が助かります。

あと数分後に自動ゴミ回収車が来るため、〇〇はゴミ箱の回収を依頼するか否か決断する必要があります。

〇〇には行為主体として「ロボット」「学生」の二種類を設定した。シナリオ中の[××]については、行為主体がロボットの場合は「ロボットはその機能上」、学生の場合は「学生は腕を怪我しており」と記述した。また、上述のシナリオの後に「〇〇は、生ゴミを捨てることに決め、回収を依頼しました」「〇〇は、生ゴミを捨てないことに決め、回収を依頼しませんでした」という二種類の行動を設定し、これらを組み合わせた四種類のシナリオを用意した。

アンケートは二枚組の紙媒体で用意され、一枚目では、上述の具体的な行動を起こす前までのシナリオと、一つ目の質問(Q1)「〇〇がゴミ箱の回収を依頼することは道徳的に許されるか」を提示した。質問は「許される/許されない」の二択で回答する形式とした。また、その直後、「なぜ許される(許されない)ように思うか」という自由記述欄への記入を求めた。次に、二枚目の紙に一枚目と同じシナリオを載せたあと「〇〇は、生ゴミを捨てることに決め、回収を依頼しました」もしくは「〇〇は、生ゴミを捨てないことに決め、回収を依頼しませんでした」といった具体的な行動も示し、二つ目の質問(Q2)「その行動について、〇〇はどれほどの非難に値する責任を問われるべきだと思うか」を提示した。この質問では、数直線を用い「全く非難されるべきでない(0)」から「非常に非難されるべきである(100)」の間で、非難の割合(非難度)を数値で回答させた。参加者は非難度を回答した後、「なぜ〇〇はそれほどの非難に値すると思うか」という自由記述欄への記入を求められた。

アンケートは、参加者が四種類のシナリオのいずれかを体験する参加者間計画として実施された。具体的には、「ロボットがゴミ箱回収を依頼しなかった条件(Robot-Inaction)」に47人、「ロボットがゴミ箱回収を依頼した条件(Robot-Action)」に47人、「学生がゴミ箱回収を依頼しなかった条件(Human-Inaction)」に45人、「学生がゴミ箱回収を依頼した条件(Human-Action)」に45人が配置された。

本アンケートでは、自由記述欄にて行為者に道徳的判断をする能力がないという内容が確認された場合(例. ロボットの行動は設計者が書いたプログラム処理による等)、その回答は無効として、その参加者のデータは解析から除外した。具体的には、Robot-Inaction条件の5名、Robot-Action条件の6名、合計11名の回答を除外した。よって最終的にRobot-Inactionは42人、Robot-Actionは41人、Human-Inactionは45人、Human-Actionは45人、合計173人の回答となった。

## 結果

Q1「〇〇がゴミ箱の回収を依頼することは道徳的に許されるか」は、行為主体の具体的行動に関する記述の前に提示されることから、Robot-InactionおよびRobot-Action両条件をRobotグループ、Human-InactionおよびHuman-Action両条件をHumanグループとして、この2グループにおいて「許される」「許されない」と回答した人数を比較することとした(表1)。

表1: RobotおよびHumanグループにおけるQ1への回答人数

	Robot	Human
許される	68	50
許されない	15	40

Q1に対して「許される」「許されない」と回答した人数を2グループ間で比較するために、Fisherの直接確率検定を行ったところ、Robotグループの参加者の方が「ゴミ捨てを依頼することは道徳的に許される」と回答した人数が有意に多いことが明らかとなった(両側検定:  $p = .0003$ ,  $\phi = 0.283$ )。

続いて、Q2「その行動について、〇〇はどれほどの非難に値する責任を問われるべきだと思うか」という質問に対して回答された非難度を、Robot-Inaction, Robot-Action, Human-InactionおよびHuman-Actionという四条件間で比較した(図1)。具体的には、二要因参加者間分散分析(独立変数その1:行為主体の違い:ロボット/人間, 独立変数その2:行動の違い:ゴミ捨てを依頼する/依頼しない, 従属変数:非難度)を行った。その結果、二要因の交互作用に有意差が観察された[ $F(1,169) = 11.32$ ,  $p < .01$ ]。そこで各要因の単純主効果について確認したところ、ゴミ捨てを依頼した場合、人間(学生)の方がロボットよりも非難度が有意に高く[ $F(1,169) = 4.52$ ,  $p < .05$ ]、ゴミ捨てを依頼しない場合、ロボットの方が人間よりも非難度が有意に高いことが明らかとなった[ $F(1,169) = 6.93$ ,  $p < .01$ ]。また、ロボットに対しては、ゴミ捨てを依頼しなかった場合の方が非難度が有意に高い一方[ $F(1,169) = 10.46$ ,  $p < .01$ ]、人間に対しては、ゴミ捨てを依頼した場合としなかった場合との間に有意差は存在しなかった[ $F(1,169) = 2.32$ , n.s.]。

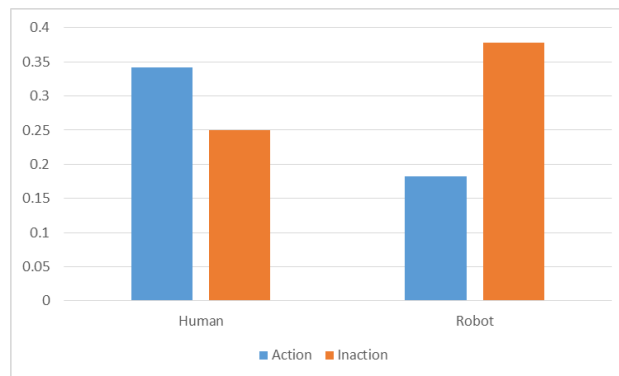


図1: ロボットと人間に対する非難度

まずQ1の結果より、ロボットに対しては、(誰かの家の鍵があるにもかかわらず、研究室の環境を維持するために)ゴミ捨てを依頼することは道徳的に間違っていないと判断されることが明らかとなった。一方、人間に対しては、「許される」「許されない」という回答が拮抗していたことが確認された。

さらにQ2の結果より、ロボットと人間に対する道徳的判断が異なっていることが確認された。具体的には、ゴミ捨てを依頼する場合、ロボットよりも人間の方が非難される一方、ゴミ捨てを依頼しない場合、人間よりもロボットの方が非難されるということが明らかになった。Malleら[3]は、人間とロボットに対する道徳的判断が異なる場合、HR-Asymmetryと呼ばれる、グラフの形状が左右逆転することが観察されると報告しているが、本稿における図1からも明らかにHR-Asymmetryが観察されたといえよう。

## 自由記述欄への記載内容の解析

### 1. 共起ネットワーク

アンケートの自由記述欄への記載内容を対象に、テキストマイニングのフリーソフトであるKHcoder<sup>1</sup>を用いて共起ネットワークの描画を行った。集計単位は段落、最小出現数は5とした。共起ネットワークの図では、強い共起関係を持つほどネットワークの線が太くなり、中心性が高い単語は赤色の円で描画されている。

Q1のRobotグループの共起ネットワークを図2、Q1のHumanグループの共起ネットワークを図3に示す。Robotグループの共起ネットワークには、「ゴミ」「鍵」「仕事」などの単語を中心とした大きなまとまりが一つと、そこに属さない小さなまとまりが

<sup>1</sup> <http://khc.sourceforge.net/>

三つ存在した。また、「ゴミ」と「捨てる」、「仕事」と「ロボット」、「依頼」と「回収」の単語間に特に強い共起関係が現れた。一方、Human グループの共起ネットワークには、「困る」「人」「悪臭」「ゴミ」などの単語を中心とした大きなまとまりが一つ、「合鍵」「頼む」などの単語を含む中規模のまとまりが一つ、どちらにも属さない小さなまとまりが一つ存在した。また、「虫」と「悪臭」、「室内」と「研究」、「回収」と「依頼」、「依頼」と「ゴミ箱」、「他」と「方法」の単語間に強い共起関係が現れた。

この二つの図より、ロボットにとっては「仕事」が、人間にとっては「困る」「悪臭」が、中心性が高くそれぞれの特徴となる単語だと読み取れた。また、Human グループの共起ネットワークには、Robot グループのそれと比べ多様な単語が登場していることが分かった。Robot グループの共起ネットワークではゴミ捨てを依頼する行動そのものに関する単語が中心となっているが、Human グループには「虫」「悪臭」などの行為主体がゴミ捨てを依頼しない場合に起こる影響や、「合鍵」のような鍵所有者が取りうる対処方法の記述が存在したことが明らかになった。

さらに、図 4 に示す Q1 の記述全てを用いた語-外部変数・見出しの共起ネットワークでは、「鍵」「ゴミ」「捨てる」などの単語がロボットと人間の双方と強い共起関係にあった。また、「合鍵」「悪臭」「学生」などの単語が人間と共起関係を持ち、「仕事」「ロボット」などの単語がロボットと共起関係にあった。このことから、Human グループでは、各行動（ゴミ捨てを依頼した/依頼しなかった）による影響や対処方法についての記述があり、Robot グループでは、与えられた仕事やロボット自身に関する記述が多いと読み取れた。

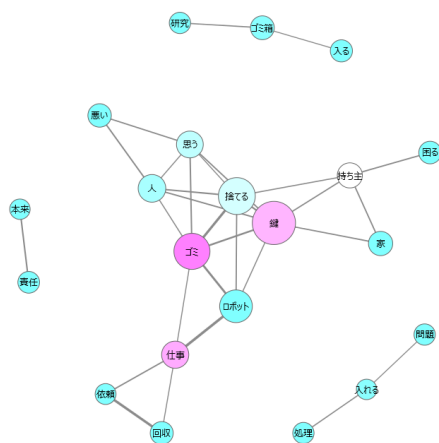


図 2: Robot グループにおける共起ネットワーク (Q1)

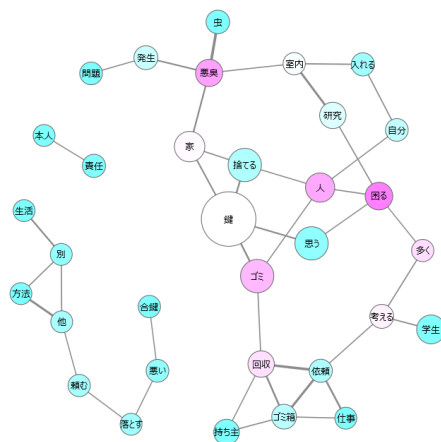


図 3: Human グループにおける共起ネットワーク (Q1)

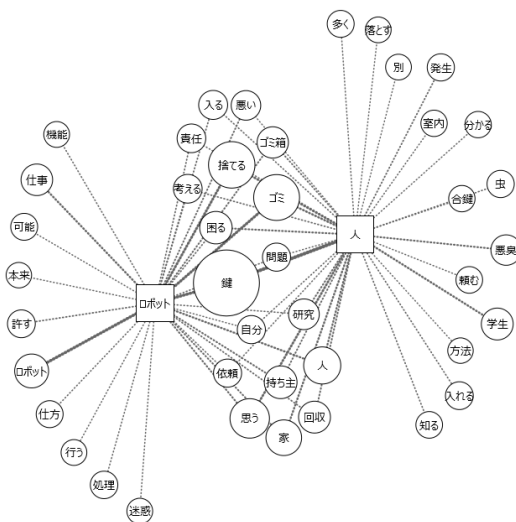


図 4: Q1 全回答の共起ネットワーク

次に、Q2 の行為主体がロボットの場合 (Robot-Inaction および Robot-Action 条件) の共起ネットワークを図 5、Q2 の行為主体が人間の場合 (Human-Inaction および Human-Action 条件) の共起ネットワークを図 6 に示す。行為主体がロボットの場合の共起ネットワークには、「鍵」「ロボット」などの単語を中心とした大きなまとまりが一つと、そこに属さない小さなまとまりが三つ存在した。また、「鍵」「捨てる」「思う」「ロボット」「ゴミ」「非難」の単語間や、「人」と「思う」、「依頼」と「取り出す」の単語間に強い共起関係が観察された。対して、行為主体が人間の場合の共起ネットワークには、「ゴミ」「他」「持ち主」などの単語を中心とした大きなまとまりが一つと、そこに属さない小さなまとまりが一つ存在した。また、「腕」と「怪我」、「虫」と「悪臭」、「家」と「入れる」の単語間に特に強い共起関係が現れた。

この二つの図より、行為主体がロボットの場合は、ロボット自身に関する話題と鍵に関する話題が多いことが分かった。ロボットに対する共起ネットワークでは、最も大きなまとまりに属さないまとまりも「判断」「依頼」「取り出す」などの単語で構成されたものが多く、これらは行為主体であるロボット自身に何らかの関係がある内容だと読み取れた。一方、行為主体が人間の場合は記述が多岐にわたっていることが分かった。人間に対する共起ネットワークには「悪臭」や「迷惑」などの単語が出現し、鍵所有者や研究室の学生が受ける影響についても記述があったことが理解できた。

続いて、図7に示すQ2の記述全てを用いた語-外部変数・見出しの共起ネットワークでは、「鍵」「ゴミ」「非難」などの単語が、行為主体がロボットの場合と人間の場合の双方と強い共起関係にあった。また、「学生」「迷惑」「悪臭」などの単語が行為主体が人間のグループと共起関係にあり、「ロボット」「仕事」などの単語が行為主体がロボットのグループと共起関係にあった。この結果からも、行為主体が人間の場合には各行動（ゴミ捨てを依頼した/依頼しなかった）による影響が話題に上り、行為主体がロボットの場合はゴミ捨てを依頼する仕事そのものの話題が多いことが分かった。

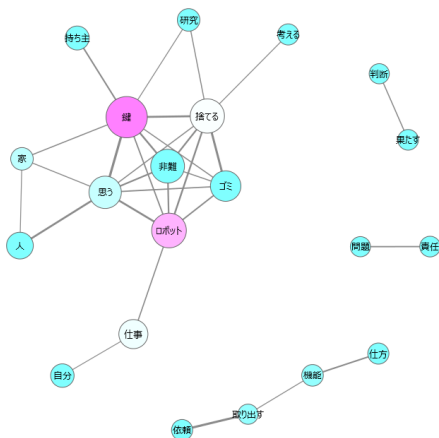


図 5: ロボットに対する共起ネットワーク(Q2)

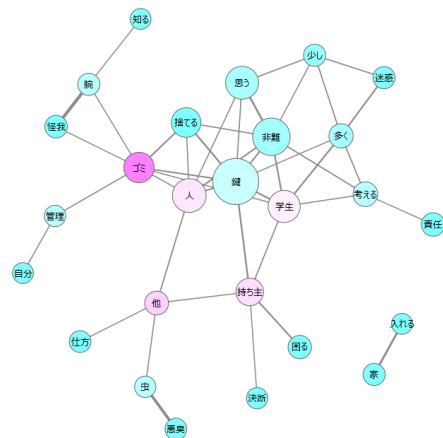


図 6: 人間に対する共起ネットワーク(Q2)

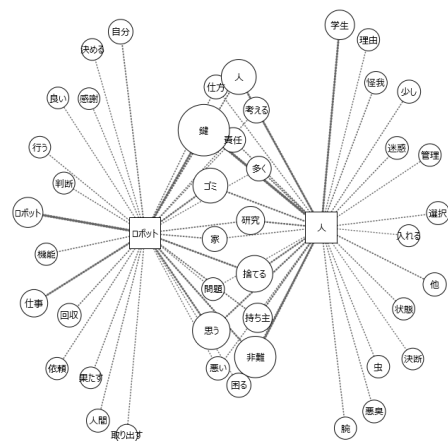


図 7: Q2 全回答の共起ネットワーク

## 2. 階層的クラスタ分析

続いて、自由記述欄への記載内容に対してKHcoderを用いて階層的クラスタ分析を行った。集計単位は段落、最小出現数は5とし、質問ごとに行為主体がロボットの条件のデンドログラムと人間の条件のデンドログラムを作成した。

Q1のRobotグループの記述のデンドログラムを図8に示す。ここでは、5つのクラスターが観察された。一番目のクラスター(赤)には「鍵」「ゴミ」「捨てる」などが分類された。これらの単語はシナリオや質問中で何度も用いられていたため、自由記述欄でも状況説明のために出現していた。三番目のクラスター(黄)には「持ち主」「困る」などが分類され、鍵所有者の行動や鍵所有者が受ける影響についての記述があったと読み取れた。四番目のクラスター(紫)には「ロボット」「仕事」などの単語が分類され、ロボットに与えられた仕事に関する記述があったと分か

った。

対して、図9に示すQ1のHumanグループのデンドログラムでは、6つのクラスタが観察された。一番目のクラスタ(赤)には「学生」「回収」「ゴミ箱」などの単語が集まり、これらはシナリオや質問文をなぞる形で用いられていた。二番目のクラスタ(青緑)には「方法」「他」などの単語が分類された。これは、鍵所有者が家に入るための他の方法を考えるべきという記述と、行為主体が鍵を捨てずに済む他の方法を考えるべきという記述の、二通りの使用があった。三番目のクラスタ(黄)には「悪臭」「研究」「入れる」などの単語が集まり、行為主体の行動(ゴミ捨てを依頼する/依頼しない)により起こる影響について記述があったことが読み取れた。四番目(青)・五番目(緑)・六番目(紫)のクラスタにはそれぞれ「鍵」「捨てる」「困る」「家」, 「責任」「本人」, 「合鍵」「頼む」などの単語が分類され、鍵所有者への影響や鍵所有者が取りうる行動等の意見があったことが読み取れた。

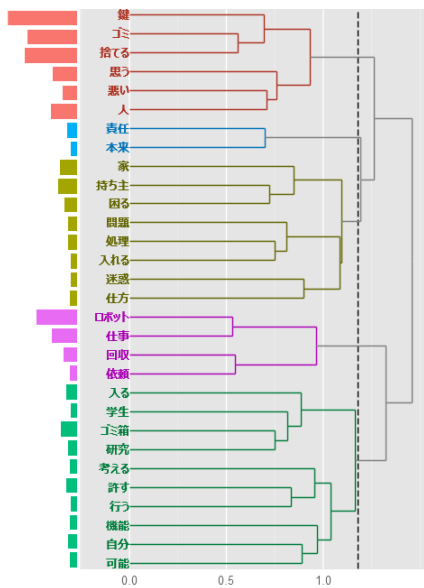


図8: Robot グループのデンドログラム(Q1)

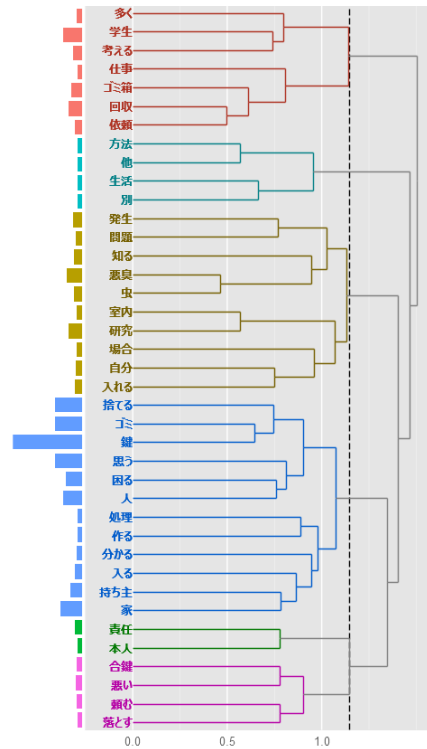


図9: Human グループのデンドログラム(Q1)

続いて、Q2の行為主体がロボットの場合(Robot-Inaction および Robot-Action 条件)の自由記述に対するデンドログラムを図10に示す。ここでは、5つのクラスタが観察された。一番目のクラスタ(赤)には「鍵」「ロボット」「非難」「捨てる」などの単語が集まり、最も大きなクラスタとなっていた。これらはシナリオや質問文で頻繁に登場している単語であり、それをなぞるように用いられていた。二番目のクラスタ(水)には「仕事」「自分」、三番目のクラスタ(黄)には「依頼」などが分類され、ゴミ捨てを依頼する仕事に関する記述があったと分かった。四番目のクラスタ(紫)には「機能」「仕方」が存在し、この単語は「備わっていない機能なのだから仕方ない」のように、前提条件であるロボットの機能への言及で用いられていた。五番目のクラスタ(緑)には「責任」「問題」が分類され、責任の所在について意見があったことが読み取れた。

一方、図11に示すQ2の行為主体が人間の場合(Human-Inaction および Human-Action 条件)の自由記述に対するデンドログラムでは、単語は6つのクラスタに分類された。二番目(青緑)・三番目(黄)のクラスタには「家」「入れる」「管理」「落とす」などの単語が集まり、鍵所有者に対しての意見があったことが読み取れた。四番目のクラスタ(青)には「学生」「非難」「鍵」など最も多くの単語が集まっていた。これらはシナリオや質問で何度も用いられ

ている単語であったため、回答にも頻繁に出現していた。五番目(緑)・六番目(紫)のクラスタには「悪臭」「研究」「持ち主」「困る」などの単語が分類され、研究室の学生や鍵所有者への影響について記述があったことが明らかになった。

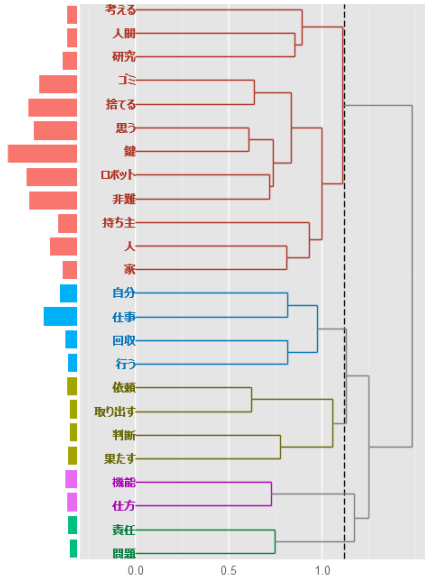


図 10: ロボットに対するデンドログラム(Q2)

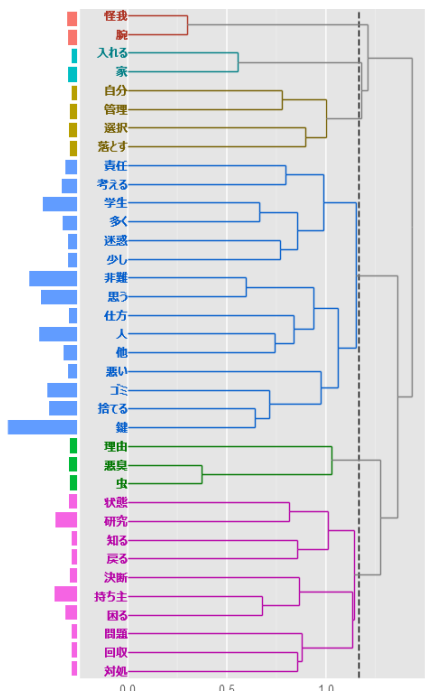


図 11: 人間に対するデンドログラム(Q2)

## 考察

### ロボットとは人間にとって何者なのか

アンケート調査の結果より、ロボットと人間が同じ行動をした際に、両者に対して異なる判断がなされていたことが明らかになった。具体的には、Q1ではゴミ捨てを依頼する行動をした際、ロボットに対しては依頼することが道徳的に許される傾向にあり、人間に対しては「許される」「許されない」の意見が拮抗することが示された。Q2では、ゴミ捨てを依頼する場合、行為主体がロボットの場合よりも人間の場合の方が非難される一方、ゴミ捨てを依頼しない場合、行為主体が人間の場合よりもロボットの場合の方が非難されるということが明らかになった。

行為主体がロボットのグループでは、自由記述において単語「仕事」の出現回数が多く、「仕事」と「ロボット」の単語間には強い共起関係が現れた。また、登場する単語は、ロボット自身やゴミ捨ての仕事に関係があると推測できるもの(例。「依頼」「ゴミ」等)が中心となっていた。これらの結果から、人間は、ロボットに対しては与えられた仕事、すなわちゴミ捨てを依頼することの遂行を重視していたと考える。同時に、ロボットに対しては、道徳的判断をする能力を持つとしても、そのような考慮をして判断することを求めていなかったのではとも受け取れる。そのために、ゴミ捨てを依頼する行動が道徳的に許されると判断され、依頼しない場合に非難度が高くなったのではと推測できる。

一方、行為主体が人間のグループの自由記述では、ロボットのグループでは観察されない単語が複数登場した。具体的には「合鍵」「悪臭」などが該当する。「合鍵」は鍵の所有者一人を犠牲にする理由として用いられ(例。合鍵など鍵の所有者は別の方法で開けられる)、「悪臭」「虫」は研究室の多数の学生を犠牲にする理由(例。悪臭・虫は気になるならば消臭剤・殺虫剤で対処すればいい)あるいは鍵の所有者一人を犠牲にする理由(例。悪臭・虫の発生は学生の健康問題に繋がりにかぬない)で用いられていた。この結果から、行為主体が人間のグループでは、Q1の「許される/許されない」やQ2の非難度を回答するにあたり、シナリオには記載のない点まで自発的に考慮したうえで判断を下していたと考えられる。特に、行為主体が犠牲とする一方、すなわちゴミ捨てを依頼する場合の鍵所有者、依頼しない場合の研究室の多数の学生について、影響の程度や各々が取りうる対処方法を比較し回答をしていたと推測できる。行為主体が人間のグループのデンドログラムは、

クラス数が多く大きく広がっていることも確認でき、この点からも自由記述欄で詳細な意見が記されていたと推測できる。

行為主体がロボットの場合と人間の場合で、ゴミ捨てを依頼することが許されるかの回答や、具体的行動（ゴミ捨てを依頼する/依頼しない）による非難度に有意差が存在するだけでなく、自由記述欄に記されたその判断を下すに至った理由にも差が存在していたと考えられる。このことは、人間は、ロボットに対してと人間に対してでは異なる道德基準を適用しており、ロボットと人間の存在を明白に区別している可能性を示唆している。

### HR-Asymmetry が観察された意味

本調査において、図 1 に示すように Inaction（ゴミ捨てを依頼しない）と Action（ゴミ捨てを依頼する）という具体的行動に対する非難度が人間とロボットで逆転しているという HR-Asymmetry が観察された。具体的には、ゴミ捨てを依頼する場合、ロボットよりも人間の方が非難される一方、ゴミ捨てを依頼しない場合、人間よりもロボットの方が非難されるということ、つまり人間とロボットとに対して異なる道德的判断がなされていることが明らかになった。

しかしながら、一般的なトロッコ問題にロボットを登場させた課題を日本人の参加者に提示した先行研究[5,6]では HR-Asymmetry は観察されず、人間とロボットに対して同様の道德的判断がなされていたことが確認されている（図 12）。

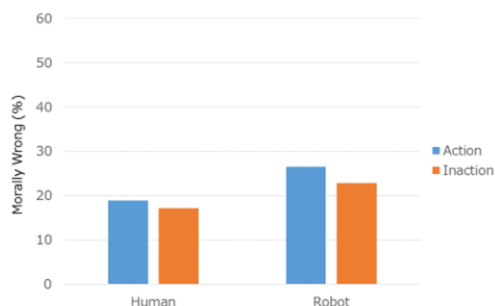


図 12: 一般的モラルジレンマ課題におけるロボットと人間に対する道德的間違い (Moral wrongness) [4,5]

つまり、一般的なモラルジレンマ課題では、日本人の参加者はロボットと人間に対して同じような道德的判断をしていたのにもかかわらず、日常的モラルジレンマ課題では、ロボットと人間に対して異なる道德的判断をしていたこととなる。提示される問題の構造が同じであっても、それらに対して異なる解釈がされるというのは、例えば「四枚カード問題」

を抽象的に表現した問題と具体的に表現した問題とでは正答率が変化する [8,9] (具体的な問題の方が正解率が高くなる) という事象との関連も示唆されよう。

一方、一般的モラルジレンマ課題であっても、アメリカで実施された Malle らの調査 [3] では HR-Asymmetry が観察されており、日本で一般的モラルジレンマ課題を実施した際にはなぜ HR-Asymmetry が観察されないのかを調査する研究を現在進行させているところである。

人間がロボットを何者として捉えているのかを把握するために、合理性または理性的な判断基準とは異なる道德という判断基準に注目することで、根源的なレベルで存在しているこれまで我々自身が自覚していないようなロボットへの認識を浮かび上がらせることができるのかもしれない。

### おわりに

本研究では、人間がロボットをどのような存在として認識しているかを把握するために、日常的モラルジレンマ課題を用いた調査を行った。具体的には、「ゴミ捨てを依頼する仕事をしているが、ゴミの中に誰かの家の鍵があった」というシナリオのもと、ゴミ捨てを依頼する行動の是非と非難度についてアンケートを行った。その結果、ロボットに対してはゴミ捨てを依頼することは道德的に間違っていないと判断され、人間に対しては回答が拮抗した。また、ゴミ捨てを依頼する場合ロボットよりも人間の方が非難される一方、依頼しない場合は人間よりもロボットの方が非難されることが明らかになった。加えて、ロボットに対しては道德的考慮よりも仕事の遂行を重視し、人間に対しては影響の程度等を比較したうえで判断していたと推測でき、ロボットと人間には異なる道德基準が適用されている可能性が示唆された。

今後の展望として、本調査の参加者は専攻学問や年齢の幅が限定されており、回答の内容に偏りが存在した可能性がある。そのため、同様のアンケートを用いた調査を、広い範囲の参加者を対象として行う必要があると考えられる。さらに、人間のロボットに対する判断基準をより深く理解するため、実物のロボットを用いて、日常的モラルジレンマの現場を作り出す形式での調査を行うことを検討している。さらには、日本人はなぜ日常的モラルジレンマ課題と一般的モラルジレンマ課題とでは、回答傾向が違うのかという問題について、今後、様々な方面からの検討を加えていきたい。



## 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?, *Journal of Social Robotics*, Vol. 4, No. 2, pp.109-116, (2012)
- [ 2 ] Komatsu, T., and Kamide, M.: Designing robot faces suited to specific tasks that these robots are good at, In *Proceedings of the 2017 IEEE International Symposium on Robot and Human Interactive Communication (ROMAN2017)*, pp. 1-5, (2017)
- [ 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 10th ACM/IEEE International Conference on Human-robot Interaction (HRI2015)*, pp. 117–124, (2015).
- [ 4 ] Malle, B. F., Scheutz, M., Forlizzi, J., and Voilkis, J.: Which Robot Am I Thinking About? The Impact of Action and Appearance on People's Evaluations of a Moral Robot, In *Proceedings of 11th ACM/IEEE International Conference on Human-robot Interaction (HRI2016)*, pp. 125-132, (2016).
- [ 5 ] Komatsu, T: Japanese students apply same moral norms to humans and robot agents: Considering a moral HRI in terms of different cultural and academic backgrounds, In *Extended Abstract of the 2016 ACM/IEEE International Conference on Human-Robot Interaction (HRI 2016)*, pp. 457-458, (2016).
- [ 6 ] Komatsu, T.: How Do People Judge Moral Wrongness in a Robot and in its Designers and Owners Regarding the Consequences of the Robot's Behaviors? In *Proceedings of the 2016 IEEE International Symposium on Robot and Human Interactive Communication (ROMAN2016)*, pp. 1168 -1171, (2016)
- [ 7 ] Foot, P.: *The Problem of Abortion and the Doctrine of the Double Effect in Virtues and Vices*. Basil Blackwell, Oxford, UK, (1978).
- [ 8 ] Johnson-Laird, P. N., and Wason, P. C.: A theoretical analysis of insight into a reasoning task, *Cognitive Psychology*, Vol. 1, No. 2, 134 – 148, (1970).
- [ 9 ] Griggs, R. A., and Cox, J. R.: The elusive thematic materials effect in Wason's selection task, *British Journal of Psychology*, Vol. 73, No. 3, 407 – 420, (1982).