

科学研究費助成事業 研究成果報告書

平成 27 年 5 月 25 日現在

機関番号：13301

研究種目：基盤研究(B)

研究期間：2011～2014

課題番号：23320002

研究課題名(和文)意図的主体性のロボットの構築に向けて

研究課題名(英文)Toward constructing robots with intentional agency

研究代表者

柴田 正良 (SHIBATA, Masayoshi)

金沢大学・その他部局等・理事

研究者番号：20201543

交付決定額(研究期間全体)：(直接経費) 14,200,000円

研究成果の概要(和文)：本研究の成果は、ロボットと人間の間でプリミティブなレベルの共同注意現象を生じさせることによって、(1)共同注意を構成する相互の意図の読み込み合いが高次の「言語的理解」ではなく、情動や欲求によって可能となること、また(2)意図の「他者性」は、ロボットが人間との「協調性」よりむしろ「抵抗性」を示す場合の方がよりリアルに感じられることを、共同注意ロボットと被験者とのインタラクション実験によって明らかにしたことである。また、その際、(3)ロボットの内部構造を単純な記憶-反応型から情動-駆動型へと複雑化させたことによって、人間の共同注意機能のメカニズムにも有望な示唆を与えたことである。

研究成果の概要(英文)：The outcome of this research is, by producing phenomena of primitive joint attention between robots and humans and conducting some interaction experiments between them, to made it clear (1) that the mutual reading of each other's intentions constructing joint attention is made possible by desires or emotions rather than by "understanding at a higher level of using language", and (2) that "otherness" of intentions is felt more real when robots show "resistant character" rather than "cooperative character". And our research also gave (3) a useful suggestion to the elucidation of the mechanisms of human joint attention through making the internal structures of robots increasingly complex from simple memory-reflective type to emotion-driving type.

研究分野：哲学

 キーワード：共同注意 意図的主体性 自律型ロボット インタラクション 意図の重層性 注視対象 記憶反射型
 情動

1. 研究開始当初の背景

意図に関する哲学的な議論の一つは、現代ではもっぱら行為論の文脈において、アンスコムやデイヴィッドソンに代表されるように意図的行為の分析においてなされてきた。もう一つの文脈は、グライスに端を発する、話者の意図の階層性から言語の意味を説明しようとする意味論の流れである。しかし、いずれも成熟した個人の意図の論理構造的な分析が主眼であって、意図の相互共有や共同体による継承はほとんど問題にされてこなかった。しかしトマセロらが言うように、他者との意図の共有という現象は、人間の社会制度の基盤となる役割を果たすがゆえに、ヒトの進化を類人猿たちの進化から決定的に分かつキーポイントである。

他方、現代社会においてロボットがますます身近な存在として人間と相互交渉の密度を高めてくるとき、最初はロボットに託された「他人の意図」に、そして次には、自律性の度合いを高めていく「ロボット自身の意図」に、われわれは向き合わざるをえなくなる。いわば、人工物がそれ自身の意図を持った存在として人間と<相互交渉>をする社会的な現実が迫っているのであり、われわれは、この新しい社会的な現象が持つ哲学的な意味を十分に理解しておかなければならない。その意味で、本研究がターゲットに絞る探究対象は、従来の哲学的分析において軽視されてきた意図の発生論的・社会制度的な要素であり、進化論的含意をもった社会的現象としての「心の機能」である。

2. 研究の目的

本研究の目的は、意図的主体性をロボット上において人工的に構築することによって、人工物の「自律的」な意図と人間の意図の<相互交渉>という新しい社会現象の哲学的意味を探究し、同時に、意図の重層構造に関して「社会的な意図共有」という視点からの哲学的分析を深化させることである。

トマセロの言うように、ヒトの進化の独自性が知識基盤としての社会的な意図の共有・継承にあるとすれば、これらの研究目的の射程は、エンハンスメントによってサイボーグ化する人間とロボットとの未来の共生可能性を占うという重要性をもつ。その意味で本研究は、自律的ロボットと共生する社会のあり方、及び意図の重層的な形成・理解・共有を、人間の基礎的な心的機能の適応としてトータルに解明しようとする哲学・認知科学・ロボット工学の学際融合的研究である。

3. 研究の方法

本研究計画は、平成 23 年度当時にほぼ完成していた第一レベルのロボットを出発点に、第二、第三レベルのロボットを逐次制作することが基本線であった。それぞれのレベルのロボットは以下のようなロボットである。

(1) 第一レベル：単なる反射行動として親の視線を追跡するのではなく、親の視線から自分の見ようと意図する対象を記憶によって同定し、以前にそれがあった方向にそれを探すロボット

(2) 第二レベル：自分の見ようと意図する対象を親の意図する対象と一致させるための推論を行なうロボット

(3) 第三レベル：親の諾否の態度を理解することによって、自分の推論の誤りを正し、それによって親の意図する対象を見出そうとするロボット

ロボット制作には全員が参加し、その中心を北陸先端大の橋本・金野グループが担い、一貫してすべての情報をこの中心に集めるという体制を取った。他は4つのグループに分かれ、高次ロボット制作の前提となる基礎的研究として、共同注視現象の分析、意図の重層構造の分析、情動要因の分析、推論・評価メカニズムの開発をそれぞれの年度配置の下に行った。このように、各々の課題担当グループの成果を逐次、ロボット制作に取り込みつつ、25年度に第二レベル、26年度に第三レベルのロボットの制作を行った。各グループは相対的に独立して研究を展開しつつも、年に数回の打ち合わせにおいて全体の知識と情報の共有を図った。また26年度は、意図共有の社会性に関する哲学的分析を行った。

4. 研究成果

(1) ロボット製作と実験デザイン

まずわれわれは第二レベルのロボット制作に着手した。それまでの人とロボットのインタラクション実験では、テーブルに配置された4つのカラーボールを互いに見るという課題を設定していたが、注視対象を正確に分析しつつその数を増やすために、1から12までの数字が描かれたカードを配置する実験環境を設計した(図1)。また、それぞれのカードは近づけた状態にして、視線による判別が難しくなる配置にした。これは、相手の顔と対象の間で視線を行き来させる交互凝視行動を通じて、注視対象を確認するプロセスを観察できるようにするためである。



図1 新たに設計した実験環境

この環境ではロボットが子どもであり、人がその親に対応する。このロボットが、自ら見ようと意図する対象を、親の意図する対象と一致させる推論を行なうことを考えた場合（第二レベル）われわれの議論を通じた最初の重要な帰結は、ロボットの内部に、単なる注視経験の蓄積だけではなく、そこに付随する情動的内部状態の生起体験の蓄積が必要だということである。つまり、われわれが製作した第一段階のロボットは、人との視線インタラクションを通じて得る対象の注視体験の蓄積から、人の視線方向に応じた想起という形で、内部に注視対象を事前に保持することで意図性を持つという状態が実現されているのだが、他者の意図の推論には、その事前の想起に単純な頻度情報だけではなく、その経験を特徴付ける「何か」が必要不可欠であり、その「何か」に充当する1つの妥当な候補として考えられるのが情動だということである。

情動要因をロボットが蓄積する経験の中にも含ませることを考えた場合、当然のことながらロボットには生得的な情動生起のメカニズムが備わっていないことが問題となる。われわれは、その情動は設計者が埋め込むのではなく、ロボットが学習を通じて獲得する必要があると考え、第一レベルのロボットが注視体験を蓄積する過程の中に、情動的やりとりを組み込んだ実験を設計した。その実験では、参加者である人が、自らの見ようとする対象を見る動作の前に、今から見ようとする対象（数字）の好みを、好き・中立・嫌いの3段階でカードによって提示する手続きがとられた（図2）。



図2 数字の好みのロボットへの提示

ロボットは、視線方向とその後の注視対象の頻度分布の作成と同時に、提示された好みとその後の注視対象の頻度分布を作成した。そして、一般化された好みの状態を作り出すために、複数人とのインタラクションによって得られた頻度分布を平均化して、想起のための確率分布を用意した。これにより、実験は2段階に分けて実施された。はじめの段階（第一テスト）では実験に参加する全員が反射的に行動するロボットとインタラクションし、そこで得られた注視体験を集約・平均

化した上で、再度全員がそのロボットとインタラクションした（第二テスト）。また、第一レベルのロボットとの違いを確認するために、実験参加者を2群に分け、一方を第一レベルのロボットと（好みの提示なし群）、他方を好みの蓄積を追加したロボット（好みの提示あり群）とインタラクションさせた。

さらに、この実験には、第三レベルのロボット製作に関わる親の諾否の要素を組み入れた。第一テストでは、人はロボットがどの対象を見ているのかがわからない状態で視線によるインタラクションを行ない、第二テストでは、一度の交互凝視が行なわれるたびに、ロボットが注視した対象がどの数字であったのかがロボットの背後に設置したモニタに表示される形で分かるようにした。以上をまとめると、実験は好みの提示（あり、なし）×注視対象の確認（あり、なし）の2要因混合計画であった。

（2）成果報告とそのフィードバック

第一レベルのロボットが備える特徴を整理し、第三レベルに至るまでのロボットの製作指針に関する議論をまとめ、それを科学哲学誌で発表した[雑誌論文]。

この製作指針に基づいて前記の実験をデザインし、参加者10人で得られた結果について、主に実験設計の考え方に関する妥当性を議論することを目的として、ベルリン（ドイツ）で開催された国際会議（CogSci2013）にてポスター発表を行なった[学会発表]。この発表においては、ロボットの内的メカニズムの豊かさに対する人の印象評価の違いについての結果の提示と議論が行なわれた。特に、好みの提示を行なった群と行なわなかった群のいずれでも、第一テスト（注視対象の確認なし）に比べて第二テスト（注視対象の確認あり）でロボットの理解度に関する人の評価は低下したが、ロボットの意図性の有無に関する評価は、好みの提示を行なった群では低下しない傾向にあることが確認された。この結果を引き起こした原因を、われわれは内的メカニズムの豊かさ（好みに関する頻度分布の形成と、それに基づく注視対象の想起メカニズム）に求めようとしたが、残念ながらロボットに蓄積された経験の性質からは、行動特徴の差異につながる条件間の差異を見つけることはできなかった。

この結果に関する議論から得たわれわれの知見は2つある。1つは、ロボットに好みのような感情を、学習を通じて獲得させることを考えた場合、単に複数人の好みを集約・平均化するだけでは、注視対象を想起する際の特徴づけ（視線方向に複数の対象があるような場合に、相手がどの対象に注意を向けているのかを特定するバイアス）にはなりにくいということである。またもう1つは、好みに応じた想起分布の形成というだけでは、視線の情報に好みの情報が加えられただけになってしまい、そこに推論の仕組みが入り込

めないということである。想起した好みは、ロボットに内的に生起する情動状態に位置付けると共に、その生起状態に対する対象の選択価値の学習によって、想起すべき対象が判断される仕組みが必要であるという考えをわれわれは得た。

この考えを具体的させるために、われわれは得られた結果の詳細な分析と議論を重ね、それを岐阜大学サテライトキャンパスで開催された国内会議（HAI2013）で口頭発表した[学会発表]。この会議は全ての発表がシングルトラックで行なわれると共に、われわれの成果は40分の時間枠が設けられた討論セッションで、計6件採択されたうちの1件として発表された。また、分析結果に対する議論をより進めた内容について、ケベックシティ（カナダ）で開催された国際会議（CogSci2014）でポスター発表を行なった[学会発表]。

これらの発表内容においてわれわれが注力したのは、条件間でのロボットに対する人の評価の違いがロボットの内部メカニズムの違いによって生じたのではないとしたら、実験手続きの中にその違いを生じる要因があったのではないかということであった。これを確認するため、われわれはロボットの内部メカニズムによる行動特徴の差異を脇に置き（表出した行動特徴に差異がないことを確認した上で）、好みの提示あり/なしと注視対象の確認あり/なしにどのような印象評価の違いがあったのかを分析した。

この分析から、まず、第一テストと第二テストの間（注視対象の確認なし/あり）で、注視対象を確認する第二テストにおいて、互いの注視対象の一致度合いが高くなるにも関わらず、人が評価するロボットの理解度や親和性が低下する傾向が確認された。これは、ロボットが注視した対象を確認できるようにしたことで、人とロボットの間で注視対象の不一致が生じたことによる効果であると考えられた。つまり、視線によるインタラクションにおいて、注視対象を確認しない第一テストでは、人はロボットと同じものを見ていたかと思いついていたが、それが思い込みであったことに第二テストで気づいたのではないかと考えられたのである。ここで興味深いのは、この変化に伴って、ロボットへの印象評価のうち、機械的・人間的的印象評価と、簡単・複雑の印象評価が、共に第二テストの方で、より人間的かつ複雑な方向に変化したことである。この結果は、ロボットの意図的な行動が人の期待と異なることに対して、人は人間性や複雑さを感じるのではないかということを示唆する。

他方、第二テストの注視対象を確認する条件で引き起こされた親和性の低下は、本実験で扱った好みのような情報をやりとりすることで緩和できる可能性がある。なぜなら、好みの提示を行なわなかった群と行なった群の間では、同じように第一テストと第二テ

ストでの理解度の評価の低下は見られたが、意図性の有無に関する評価や親和性に関する評価の低下は起こらない傾向にあることが確認されたからである。これは、感情表出とその交換が、注視対象の不一致が顕在化することによって引き起こされる親和性の低下を抑止した可能性を示唆する結果であるとわれわれは考えている。上記の成果によってわれわれの研究は、先述の国内会議（HAI2013）の翌年に開催された同会議（HAI2014）において、Outstanding Research Awardを受賞することになった。

(3)まとめ

われわれはこのプロジェクトを通じて、ロボットにおいて意図的主体性（自らの見ようとするものを見るという意図を持ったロボット）を実現する方法を検討し、その意図的主体性に基づく推論の実現には情動要因を導入する必要があることを見出した。そして、その主体的な行動が人の期待と異なる行動を生成することで、ロボットに人間性を備えさせる可能性があることを確認した。また、人とロボットの間の感情的な情報の交換が、主体性を発揮するロボットの親和性を維持する可能性があるという知見を得た。

これらの知見によって、今後はロボット内部に情動状態を生起させ、その状態に基づいて他者意図の推論を行なうメカニズムを実現し、その影響および効果を確認することができると考えられる。また、多くの哲学的議論を通じて構築された実験デザインは、条件をより厳密に統制し、多くの参加者からデータを得る実験への接続を可能にするものと考えられる。

5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕(計 36 件)

大平 英樹、感情的意思決定を支える脳と身体の機能的関連、心理学評論、査読有、57巻、2014、pp. 94-119.

柏端 達也、欲求の反映と人称的バイアス、哲学（三田哲学会編）、査読有、第133集、2014、pp. 1-21.

Briedis M, Nagataki S., The problem of naturalizing phenomenology: a radiologist case, Proceedings of the 15th annual international conference of SPM, 査読有、vol.1, 2013, pp.13-16.

服部 裕幸、情緒とは何か？、アカデミア（人文・自然科学編）、査読無、第6号、2013、pp. 1~11.

Ohira, H., Matsunaga, M., Murakami, H.,

Osumi, T., Fukuyama, S., Shinoda, J., & Yamada, J., Neural mechanisms mediating association of sympathetic activity and exploration in decision-making, Neuroscience, 査読有、Vol.246, 2013, pp.362-374 .

金野 武司、柴田 正良、回帰的意図理解をめざす共同注意ロボット、科学哲学、査読有、44 巻、2012、pp . 29-45

柴田 正良、記憶喪失と世界喪失---水槽脳になったばかりの人が持つ記憶は元の世界を指示できるか？---、哲学・人間学論叢（金沢大学哲学人間学研究会）査読無、第3号、2012、pp . 17-26

Nakao T., Ohira, H., Distinction between Externally vs . Internally Guided Decision-Making: Operational Differences, Meta-Analytical Comparisons and Their Theoretical Implications, Front Neurosci, 査読有、Vol . 6, 2012, pp . 1-26.

S . Nagataki, S . Hirose, On What Mediates Our Knowledge of the External World, 査読有、Glimpse--Phenomenology and Media, Vol.13, 2011, pp . 99-106.

〔学会発表〕(計 46 件)

三浦 俊彦、多義性の誤謬としての点滅論法、日本科学哲学会第 47 回大会、2014 年 11 月 15 日、南山大学（愛知県・名古屋市）

T .Konno, S .Nagataki, M .Shibata, T .Hashimoto, H .Ohira, Effects of uncovering gaze target mismatch in human-robot joint visual attention on evaluation of understanding and impressions of robot, The 36th annual meeting of the cognitive science society (CogSci2014), 2014 年 7 月 23 日～26 日, Quebec City, Canada

金野 武司、長滝 祥司、柴田 正良、ロボットとの共同注視場面において注視対象の確認手続きが与える主観評価への影響、HAI シンポジウム 2013、2013 年 12 月 8 日、岐阜大学駅前サテライトキャンパス（岐阜県・岐阜市）

Shoji Nagataki, Masayoshi Shibata, Takeshi Konno, Takashi Hashimoto, Hideki Ohira, Reciprocal Ascription of Intentions Realized in Robot-human Interaction, The 36th annual meeting of the cognitive science society (CogSci2013), 2013 年 8 月 1 日, Berlin, Germany .

柏端 達也、知覚の共同性と公共性について ‘ joint attention ’ の話を中心に、応用哲学会第 5 回年次研究大会、2013 年 4 月 20 日、南山大学（愛知県・名古屋市）

大平 英樹、意思決定に伴う脳と身体の機能的関連、第 14 回日本ヒト脳機能マッピング学会、2012 年 07 月 05 日～2012 年 07 月 06 日、京王プラザホテル札幌（北海道・札幌市）

S . Nagataki, S . Hirose, Mediated Mind, SPM The 14th Annual International Conference, 2012 年 2 月 28 日, San Diego, USA

金野 武司、心の哲学とロボット工学の出会い～ロボットの持ち得る意図的主体性は人の心的表象の理解にどのように迫れるか、「複雑系科学と応用哲学」沖縄研究会第 1 回大会、2011 年 8 月 30 日、琉球大学（沖縄県・那覇市）

橋本 敬、言語とコミュニケーションの進化～実験アプローチの展開、「複雑系科学と応用哲学」沖縄研究会第 1 回大会、2011 年 8 月 30 日、琉球大学（沖縄県・那覇市）

〔図書〕(計 16 件)

服部 裕幸、信原 幸弘、柏端 達也、他、勁草書房、新・心の哲学 情動篇、2015、252 頁（服部 31-66、柏端 159-195）

大平 英樹、三浦 俊彦、中村 靖子、他、春風社、虚構の形而上学:「あること」と「ないこと」のあいだで、2015、457 頁（大平 317-360、三浦 361-442）

Shoji Nagataki, Marco Pina, Nathalie Gontier, et al., Springer Verlag, The Evolution of Social Communication in Primates: A Multidisciplinary Approach, 2014, 326 頁（Nagataki 147-161）

長滝 祥司、村田 純一、他、東京大学出版会、知の生態学的転回 2 : 技術 身体を取り囲む人工環境、2013、301 頁（長滝 209-232）

柏端 達也、河野 哲也、他、東京大学出版会、知の生態学的転回 3 倫理～人類のアフオーダンス、2013、334 頁（柏端 183-207）

服部 裕幸、柴田 正良、美濃正、他、大隅書店、これが応用哲学だ！、2012、312 頁（服部 78-86、柴田 135-143）

柏端 達也、柴田 正良、戸田山 和久、他、世界思想社、応用哲学を学ぶ人のために、2011、380 頁（柏端 71-83、柴田 123-134）

Shibata, M., J. L. Krichmar, H. Wagatsuma, et al., Cambridge University Press, *Neuromorphic and Brain-Based Robots*, 2011, 376 頁 (Shibata 345-361)

大平 英樹、子安 増生、他、新曜社、*ミラーニューロンと<心の理論>*、2011、231 頁 (大平 195-220)

〔その他〕

ホームページ等

<http://siva.w3.kanazawa-u.ac.jp/index.html>

<http://members.jcom.home.ne.jp/miurat1/>
<http://www.lit.nagoya-u.ac.jp/~ohiralab/top.htm>

6. 研究組織

(1) 研究代表者

柴田 正良 (SHIBATA, Masayoshi)

金沢大学・その他部局等・理事

研究者番号：20201543

(2) 研究分担者

大平 英樹 (OHIRA, Hideki)

名古屋大学・環境学研究科・教授

研究者番号：90221837

橋本 敬 (HASHIMOTO, Takashi)

北陸先端科学技術大学院大学・知識科学研究科・教授

研究者番号：90313709

金野 武司 (Konno, Takeshi)

北陸先端科学技術大学院大学・知識科学研究科・特任助教

研究者番号：50537058

長滝 祥司 (NAGATAKI, Shoji)

中京大学・国際教養学部・教授

研究者番号：40288436

三浦 俊彦 (MIURA, Toshihiko)

和洋女子大学・言語・文学系・教授

研究者番号：10219587

柏端 達也 (KASHIWABATA, Tatsuya)

慶應義塾大学・文学部・教授

研究者番号：80263193

服部 裕幸 (HATTORI, Hiroyuki)

南山大学・人文学部・教授

研究者番号：40110754