

令和 4 年 6 月 13 日現在

機関番号：32665

研究種目：研究活動スタート支援

研究期間：2020～2021

課題番号：20K23351

研究課題名（和文）エージェントが個人の性格へ高速かつ高精度に適応できる3段階他者モデルの構築

研究課題名（英文）Construction of a three-level other's mind model that enables fast and highly accurate personalization

研究代表者

大澤 正彦（OSAWA, Masahiko）

日本大学・文理学部・助教

研究者番号：40875803

交付決定額（研究期間全体）：（直接経費） 2,200,000円

研究成果の概要（和文）：提案アーキテクチャの根幹となる未学習データの検出技術の研究では、双曲空間を利用することの優位性を示し、17th IEEE TOWERS Undergraduate Student Awardを受賞した。また実際に他者モデル技術の実装を進め、複数の観点で応用・評価した。例えば、開発した典型他者モデルを活用して、繰り返し囚人のジレンマと呼ばれるタスクを行い、他者の性格に合わせた振る舞いができるかを評価した。これらの成果に基づき、現在は企業との共同研究へ発展もしている。

研究成果の学術的意義や社会的意義

本研究は、人とエージェント（ロボットやチャットボットなど擬人化されたテクノロジーの総称）が協調する未来を実現する上で重要な研究である。現在の機械学習技術の発展は目まぐるしいが、人と共に学習し成長するには、学習の速度や必要とする学習回数、データ数の違いなど、課題が多く存在する。本研究は、人同士が学び合う際に行なっているであろう心の読み合いの原理を計算機技術として実現し、人とエージェントの間でも実現することを目指す研究である。提案アーキテクチャは高速かつ高精度に人に適応できる可能性があり、実用化すれば人とエージェントの関係性を大きく発展させられる。

研究成果の概要（英文）：Novelty detection is a fundamental study of the proposed architecture. We showed the superiority of using a hyperbolic space and received the 17th IEEE TOWERS Undergraduate Student Award. We also proceeded with the implementation of the other's mind model and applied and evaluated it from several perspectives. For example, we performed a task called the "Repeated Prisoner's Dilemma" using the typical other's mind model that we proposed. We then suggested that it can be personalized. Based on these results, we are currently conducting joint research with a company.

研究分野：人工知能

キーワード：ヒューマンエージェントインタラクション 他者モデル 機械学習 適応

1. 研究開始当初の背景

- (1) 他者モデルは、他者の心的状態や行動の推定モデルであり、生物の他者モデルを調べる研究や、他者モデルを実装する工学研究が多く行われている。本研究は他者モデルの工学研究である。
- (2) 既存の他者モデルの工学研究の問題点は、個人の性格に応じた他者モデル形成を高速かつ高精度に実現できない点である。この問題は、大量のデータ数や学習時間を要する機械学習が、人とのインタラクションに適用困難なことに起因している。
- (3) 本研究の学術的「問い」は、「個人の性格への高速かつ高精度な適応を実現する工学的方法はどのようなものか。」である。

2. 研究の目的

- (1) 研究目的は、インタラクション中の他者モデル更新による個人適応技術の確立である。

3. 研究の方法

- (1) 本研究では個人に適応できる他者モデルを構築する。ここで他者モデルは、心的状態/行動推定器と、推定された他者の心的状態/行動ベクトルからなる(図 1)。また性格は、他者モデル内で扱う心的状態/行動の起こりやすさと定義する。
- (2) 中心的なアイディアは、人間が形成していると思われる 3 つのモデルを段階的に構築する(図 2)ことである。

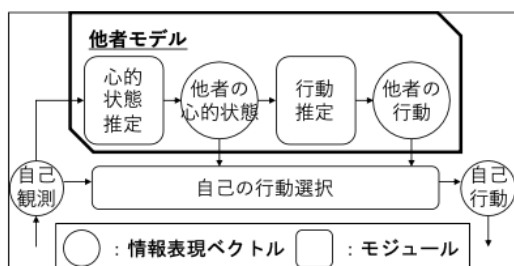


図 1: 他者モデルも持ったエージェント

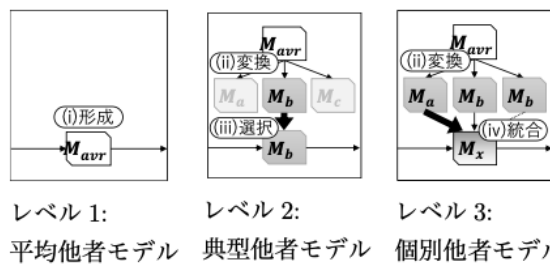


図 2: レベル 1~3 の他者モデルの概要

- (3) 最終目標である個別他者モデルは、典型的な性格ごとの典型他者モデルの線形的な統合(図 2(iv))で形成し、個々の典型他者モデルは平均他者モデルを変換(図 2(ii))することで形成する。他者モデル形成法として 2 つのアプローチを想定する(図 3)。1 つはモデルベースアプローチ(MB)であり、扱う心的状態をあらかじめ選別する。もう 1 つはモデルフリーアプローチ(MF)であり、心的状態の選別は行わず、心的状態の表現を機械学習により獲得する。MB はモジュールの作り込みが容易であり実現可能性が高く、MF は実現難易度は高いが MB よりも高精度となりうる。

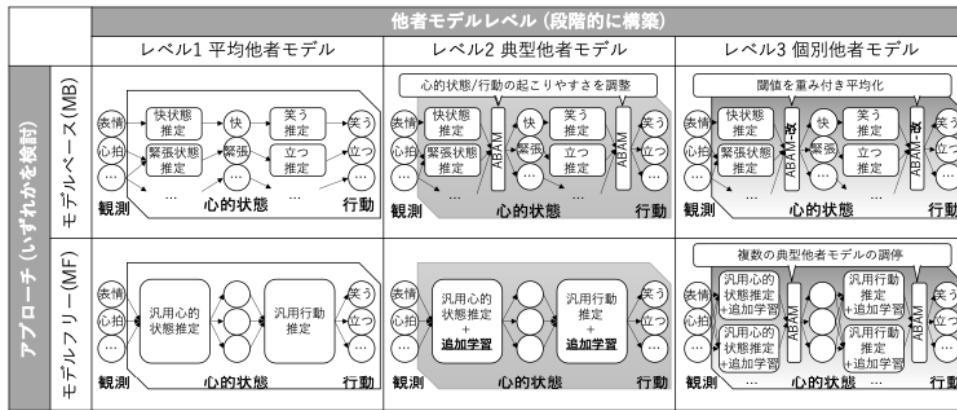


図 3: 2 種類のアプローチによる 3 段階の他者モデル

- (4) 他者モデルや心的状態推定に関する既存研究と比較して、以下の 2 点で学術的独自性と創造性がある。
- (ア) 従来は構築済みの単一モデルを用いていたが、個別の他者モデルを動的に形成する点。
 - (イ) 自ら開発した手法の応用による、報告者ならではの方法論である点。

4. 研究成果

- (1) 未学習データの検出は、課題で作成する複数の他者モデルのうち、どの他者モデルを利用すべきかを決定するために重要な技術要素である。我々は、双曲空間を用いることで効果的な未学習データの検出が行えるのではないかと考えた。なぜなら、双曲空間では空間の外に行けばいくほど空間が広くなるという性質を持っているため、学習データに含まれるクラスは空間の外に配置され、未学習のデータはどのクラスとも均等に距離をとる原点付近に集まるのではないかと考えた。実験の結果概ね仮説通りの傾向が見てとれたほか、一般的によく使われるユークリッド空間と比較すると、未学習データの検出に有効である傾向が示唆された。当該研究は 17th IEEE TOWERS Undergraduate Student Award を受賞した。現在は、どのような性格の他者かを提案した技術を活用して高速に判断し、適切な性格に合わせた典型他者モデルを当てはめていく方法論について実験している。
- (2) 開発した典型他者モデルを活用して、繰り返し四人のジレンマと呼ばれるタスクを行い、他者の性格に合わせた振る舞いができるかを検討した。簡単に作り込まれた 3 つの性格に適應できることまでを確認しており、今後より発展的な条件へと展開していく予定である。さらに典型他者モデルの切り替えは、4(1)で説明した未学習データの検出方法を活用していく予定である。
- (3) 提案アーキテクチャの応用例として、シャーデンフロイデと呼ばれる人の不幸を喜ぶ情動を扱った。シャーデンフロイデの認知モデルを、我々が開発した他者モデルを中心とした認知アーキテクチャによって構築することができると考え、シャーデンフロイデを有する種と有さない種でそれぞれ学習課題を行った。結果シャーデンフロイデを有する種は有さない種に比べて協調的な振る舞いをする可能性が見られた。例えば、いわゆる「独り占め」が起こった際に、シャーデンフロイデがある種では独り占めしたエージェントが他者から嫌がらせを受けるため、それを避けるために独り占めが起こりにくくなるといった作用が想定される。
- (4) 研究の過程でヒトの「共感」のメカニズムを参考にすることで、さらに発展的な学習が行える他者モデルを提案できるのではないかとという新たな着想を得た。具体的には、共感には「情動的共感」と「認知的共感」があり、それぞれの役割が、自分自身の行動モデルである事故モデルと、他者モデルの相補的な学習を下支えしているのではないかと考えている。現在はアーキテクチャの設計を進めている。
- (5) 日本認知科学会全国大会では、2020 年から 3 年連続で当該研究プロジェクトに関連するオーガナイズドセッションを主催し、多くの研究者と濃密な議論が実現された。

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計13件（うち招待講演 3件 / うち国際学会 1件）

1. 発表者名 小椋 有唯, 奥岡 耕平, 大澤 正彦
2. 発表標題 エージェントへのあたたかさのインストールに向けて
3. 学会等名 HAIシンポジウム
4. 発表年 2021年

1. 発表者名 阿部 将樹, 奥岡 耕平, 大塚 拓幹, 大澤 正彦
2. 発表標題 エージェントとの共同視聴によるユーザの孤独感解消への影響
3. 学会等名 HAIシンポジウム
4. 発表年 2021年

1. 発表者名 石川 幸太郎, 飯野 直樹, 磯部 光裕, 中島 亮一, 大澤 正彦
2. 発表標題 パーソナリティ特性に基づく球体の動きに感じるアニメーションの分析
3. 学会等名 HAIシンポジウム
4. 発表年 2021年

1. 発表者名 菊池 華世, 野寄 修平, 大澤 正彦
2. 発表標題 喋る空気清浄機の印象の評価
3. 学会等名 HAIシンポジウム
4. 発表年 2021年

1. 発表者名 坂本 孝丈, 大澤 正彦, 市川 淳
2. 発表標題 認知的インタラクションフレームワークの説明
3. 学会等名 日本認知科学会全国大会 (招待講演)
4. 発表年 2020年

1. 発表者名 奥岡 耕平, 大澤 正彦, 今井 倫太
2. 発表標題 BDIモデルベース階層型他者モデルの検討
3. 学会等名 日本認知科学会全国大会 (招待講演)
4. 発表年 2020年

1. 発表者名 田足井 昇太, 芦原 佑太, 青木 賢治, 大澤 正彦
2. 発表標題 双曲空間を利用した未学習クラスデータの検出
3. 学会等名 電子情報通信学会 ニューロコンピューティング研究会
4. 発表年 2020年

1. 発表者名 Shota Tatarai, Yuta Ashihara, Kenji Aoki and Masahiko Osawa
2. 発表標題 Hyperbolic Space Embedding for Openset Recognition
3. 学会等名 The 17th IEEE Transdisciplinary-Oriented Workshop for Emerging Researchers
4. 発表年 2020年

1. 発表者名 下田 音里, 田足井 昇太, 佐々木 康輔, 大森 隆司, 大澤 正彦
2. 発表標題 シャーデンフロイデが協調性に与える影響評価
3. 学会等名 HAIシンポジウム
4. 発表年 2022年

1. 発表者名 長原 令旺, 田足井 昇太, 佐々木 康輔, 大森 隆司, 大澤 正彦
2. 発表標題 繰り返し囚人のジレンマゲームを題材とした典型他者モデルの切り替えによる個人適応
3. 学会等名 HAIシンポジウム
4. 発表年 2022年

1. 発表者名 佐々木 康輔, 及川 颯斗, 田足井 昇太, 宮田 章裕, 大森 隆司, 大澤 正彦
2. 発表標題 共感プロセスを参考にした自己・他者モデルの相補的な学習方法の提案
3. 学会等名 HAIシンポジウム
4. 発表年 2022年

1. 発表者名 大澤 正彦
2. 発表標題 相互適応を中心に体系化するエージェント研究の現在と展望
3. 学会等名 日本認知科学会全国大会 (招待講演)
4. 発表年 2021年

1. 発表者名 Masaki Abe, Kohei Okuoka and Masahiko Osawa
2. 発表標題 Examining the Factors that Make Co-Watching with Agents Effective
3. 学会等名 9th International Conference on Human-Agent Interaction (国際学会)
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------