

環境情報の非対称性が Legible Motion の主観的評価に与える影響

The effect of information asymmetry on the subjective evaluation of legible motions

宝田 悠[†], 福地庸介[‡], 今井倫太[‡], 高橋 達二[†]

Yu Takarada, Yosuke Fukuchi, Michita Imai, Tatsuji Takahashi

[†] 東京電機大学 [‡] 慶應義塾大学

[†]Tokyo Denki University, [‡] Keio University

tatsujit@mail.dendai.ac.jp

概要

機械学習技術の発展に伴い、人工エージェントと人間との協働の実現が期待されている。協働では参加するメンバ間の相互理解が重要となる。Fukuchi et al. が提案した PublicSelf は、エージェントの目標を人に伝達する動きである legible motion を生成することができる。本研究では人とエージェントの間に情報の非対称性が発生する場面において legible motion を生成した際の観測者への影響を検証する実験を行った。結果、情報の非対称性を考慮することによって人がエージェントの目標を推測する際の精度を向上できることが確認できた。

キーワード: 強化学習, Legible Motion, PublicSelf, Bayesian Theory of Mind

1. はじめに

機械学習技術は目覚ましい成長を遂げており、その応用分野は多岐にわたっている。機械学習技術の現実世界への適用が期待される例として、機械学習によって行動を獲得するエージェントと人間との協働があげられる。

効果的な協働を実現するには参加するメンバ間の相互理解、特に互いの目標を知ることが必要になる。また、現実への人工エージェントの介入を検討する際には、ロボットアームなどのハードウェアの使用が考えられる。相手の目標を把握することは人工エージェントと人間で物理的に衝突してしまうなどの危険性を排除する為にも重要な要素である [1]。

Fukuchi et al.[5] が提案している PublicSelf モデルは、行動をする人工エージェント(行為者)とそれを観察する観測者の間の関係において、観測者から行為者の行動がどう見えているかを推測する行為者のモデルである。PublicSelf モデルを使うことによって行為者の目標を観測者に伝達する動き、legible motion を生成することができる。Fukuchi et al. は legible motion

の生成の際に、行為者と観測者の間で環境の状態に関する情報の非対称性を考慮することの必要性を主張し、また PublicSelf モデルによって生成される legible motion が情報の非対称性のある場面でも有効に行為者の目標を伝達できていることを検証するため、実験協力者に様々なシナリオ及びエージェントの行動方針を組み合わせ記録した動画を提示し評価をする実験を行った。実験結果として、情報の非対称性を考慮することで PublicSelf モデルによって生成される legible motion がより明確に目標を目指しており予測しやすく感じるといった主観評価への影響が予想されたが、情報の非対称性を考慮した動きと考慮しない動きでの人間の主観評価には差異が確認できなかった。

この結果に関して、実験デザインの問題から十分な検証がされていなかったことが問題として挙げられる。代表例としては、実験参加者へ提示される刺激にはランダム刺激が多く、行動方針ごとの主観評価がランダム刺激に左右されてしまうということがある。他にも、実験参加者へ求める主観評価の項目の再検討も同様に実験デザイン上の問題点として挙げられる。

本研究ではこれらの要素を改善し再度実験を行い、情報の非対称性が観測者の主観評価に与える影響を考察し検証を行なった。結果、観測者が行為者の目標を推測する際の精度の向上が確認され、主観評価への影響は確認できなかった。

2. 背景

2.1 Legible motion

人と人工エージェントが同じ環境で共存していくためには、人がエージェントの振る舞いを理解できるようになることが必要である。人が予期しないエージェントの振る舞いは致命的な事故に繋がる。また人とエージェントによる効果的な協働を実現するためには、エージェントの振る舞いの背後にある目標を人が理解できることが重要となる。人はエージェントの目

標が理解できてはじめて、タスクを分担する、エージェントを手助けするといった行動を選択できるようになるためである。

しかし、一般的な強化学習によって行動を獲得するエージェントの場合、周囲の人にはエージェントの行動決定がブラックボックス化してしまい人がエージェントの目標を理解できなくなるという問題が生じる。行動を決定する強化学習器が、エージェントの振る舞いの背後にある目標に関して明示的な表現を持たないことが多いためである。

これまで実世界へのプロジェクション [2] や自然言語 [7] など、様々な手段を通じてエージェントの目標を人に伝達する手法が提案されている。Dragan et al. [4] は、人が他者の行動からその背後にある目標を推測しようとする（心を読む）能力に着目し、動きによってエージェントの目標を伝達する legible motion の生成手法を提案している。例えば図3で、青いエージェントがりんごかなしのどちらかを目標に移動している状態を考える。図3左ではエージェントがりんごとなしの間に向かって直進し、りんごとなしの直前で向きを変えりんごに向かっていく。エージェントの動きを見ている我々には、エージェントが向きを変えるまでエージェントの目標を推定することが困難だといえる。一方、図3右ではエージェントがりんごの側へ回り込んでいるため、図3左に比べてエージェントの目標をより素早く推定できると考えられる。

Dragan et al. は legible motion を「エージェントの軌道 ξ を元に人が推測するエージェントの目標 g が、実際の目標 g^* と一致する確率を最大化する軌道」と定義している。

$$\text{LegibleMotion}(g^*) = \operatorname{argmax}_{\xi} \frac{\int P(g = g^* | \xi) f(t) dt}{\int f(t) dt}, \quad (1)$$

ここで、 f は軌道の初期に大きく重みをつけることでエージェントの目標がより素早く推定されるようにする関数である。ベイズ則によると $P(g|\xi) \propto P(\xi|g)P(g)$ であり、 $P(\xi|g)$ は、行動するエージェントが目標に対してコストが小さくなる動きほど選択されやすい、というエージェントの行動の合理性を仮定することで推定される。

$$P(\xi|g) \propto e^{-(C[\xi] + \min(C[\xi']))} \quad (2)$$

$C[\xi]$ は軌道 ξ に必要なコストであり、 ξ' は ξ の時点から目標 g を達成する軌道である。

Dragan et al. は人とロボットとの協働に legible motion を応用しその有効性を示している [4]。しかし、

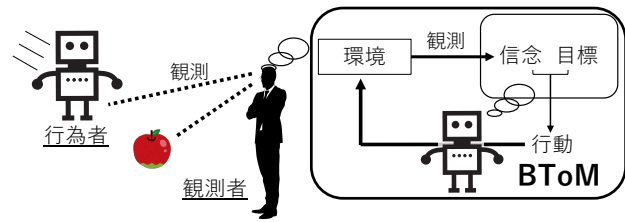


図1 Bayesian Theory of Mind

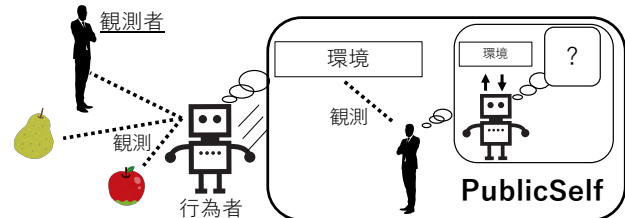


図2 PublicSelf モデル

これまでの legible motion の研究では協働を行う環境が限定的で、人とエージェントが環境の情報を完全に共有していることを前提としていた。実際の協働では人とエージェントにはそれぞれ部分観測性があるため、一方が知っている環境の情報を他方が知らないという情報の非対称性が生じ得るが考慮されてこなかった

2.2 Bayesian Theory of Mind

Bayesian Theory of Mind (BToM)[3] は、人が他者の振る舞いから信念や目標といった心的状態を帰属する過程をベイズ推論としてモデル化したものである(図1)。本稿では、行動を見せているエージェントを“行為者”，行為者の行動を観測しているエージェントを“観測者”と呼ぶ。BToM は行為者に心的状態を帰属している観測者をモデル化したものである。

BToM は、行為者の行動選択に合理性を仮定することで行為者の目標を推定する点では Dragan et al. の legible motion と共通している。一方 BToM では行為者の部分観測性を前提としている。BToM では、行為者は自身の観測をもとに環境に関する信念を形成し、信念のもとで目標を合理的に達成する行動を選択しているという過程がモデル化されている。

3. PublicSelf モデルによる Legible motion の生成

3.1 PublicSelf モデル

PublicSelf モデル [5] は「観測者が行為者に帰属する心的状態」を推定する行為者をモデル化したもので、BToM の入れ子構造が1つ深くなったものである(図2)。

PublicSelf モデルは、時刻 t までの行為者の観測 $o_{:t}$ と行動 $a_{:t}$ をもとに観測者が行為者に帰属する行為者の目標 g^2 を推定する。

$$\begin{aligned}
 & P(g^2 \mid o_{:t}, a_{:t}) \\
 \propto & \sum_{\substack{b_t^2, b_{t-1}^2, b_t^1, b_{t-1}^1 \\ o_t^2, o_{t-1}^2, s_t, s_{t-1}}} P(a_t \mid b_t^2, g) P(b_t^2 \mid b_{t-1}^1) P(o_t^2 \mid b_t^1) \\
 & P(b_t^1 \mid b_{t-1}^1, o_t^1) P(o_t^1 \mid s_t) P(o_t \mid s_t) \\
 & P(s_t \mid s_{t-1}, a_{t-1}) P(g^2 \mid o_{:t-1}, a_{:t-1})
 \end{aligned} \tag{3}$$

ここで、 b は信念を表す。上付きの添字¹はその変数が行為者が観測者に帰属する心的状態、添字²は、行為者が推定する「観測者が行為者に帰属する心的状態」であることを示している。式3は前向きアルゴリズムによって計算できる。

3.2 Legible motion の生成

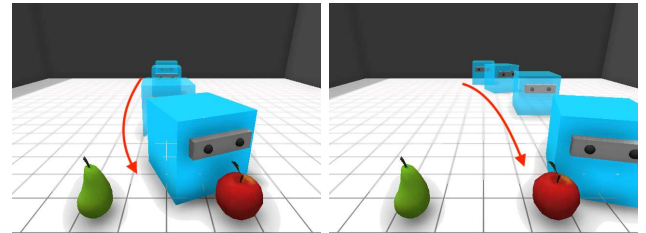
式3によって「観測者が推測する行為者の目標が、実際の行為者の目標と一致する確率」を最大化する行動が選択できる：

$$\operatorname{argmax}_{a_t} P(g^2 = g^* \mid o_{:t}, a_{:t}) \tag{4}$$

ただし、式4によって選択される行動がエージェントの実際の目標を達成するためのコストを大きく増加させる行動を選択しないよう、選択可能な行動は制限をかける。これは例えば、エージェントの目標が誤解される確率 $P(g^2 \neq g^* \mid o_{:t}, a_{:t})$ を減少させる行動を取ることで $P(g^2 = g^* \mid o_{:t}, a_{:t})$ が大きく上昇することがある一方、そういった行動が実際の目標の達成を阻害してしまうことがあるためである。

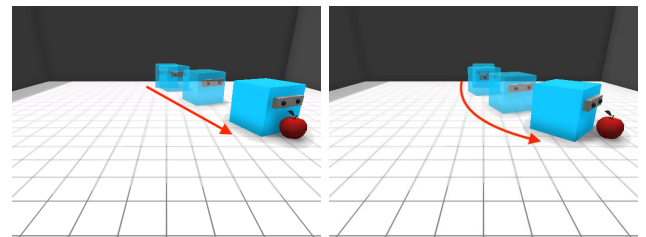
4. シミュレーション環境における実装

PublicSelf モデルによる legible motion の生成を評価するためシミュレーション環境を用意し、環境内のエージェントに対して legible motion の生成手法を実装した。環境には行為者であるエージェント、観測者である人とりんご、なしが存在する。エージェントの目標 g はりんごとなしの二択である。エージェントの行動は、深層強化学習によってりんごかなしに向かうよう学習した Original、従来手法と同様に観測者と行為者に生じる情報の非対称性を考慮せずに legible motion を生成する False-projective、そして情報の非対称性を考慮する PublicSelf の3種類のモデルから生成した。Original が学習するのは、目標に向かう際



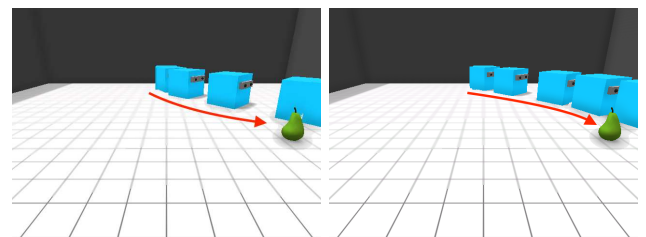
(a) Original motion (b) False-projective, PublicSelf

図3 center シナリオ



(a) Original motion, PublicSelf (b) False-projective

図4 side_visible シナリオ



(a) Original motion (b) False-projective, PublicSelf

図5 side_invisible シナリオ

のコストを最小化する行動である。評価は `center`, `side_visible`, `side_invisible` という 3 種類のシナリオで行なった。図 3,4,5 は、観測者の視点から見たそれぞれのシナリオを示している。本章では行為者がりんごを目標としている場合を考える。

`center` シナリオではりんごとなしが観測者の目の前にあり観測者と行為者の両者がりんごとなしを観測できるため、観測者と行為者の間には情報の非対称性が存在しない。Original の動き (図 3 左) では行為者が観測者の方向へまっすぐ進み直前でりんごに向きを変えるため、観測者は終盤になるまで行為者の目標を推定することができない。一方 False-projective と PublicSelf の動き (図 3 右) では最初から行為者がりんごの側に回り込んでいるため、観測者は早い段階から行為者の目標がりんごであることを推測しやすくなっていると言える。

`side_visible` シナリオは、エージェントの目標であるりんごのすぐ隣になしがあるが、観測者の視界には入っていない状況である。Original と PublicSelf では、行為者は図 4 左のように一直線に目標に向かっていく。一方 False-projective (図 4 右) は、りんごの側に回り込むことで行為者がなしを目標にしていると誤解される可能性を減少させようとしている。しかし観測者の視点からはなしが見えていないため、りんごの側に回り込む動きは有効にエージェントの目標を伝達できていないと考えられる。

`side_invisible` シナリオでは、エージェントの目標であるりんごがなしのすぐ隣にあるものの、観測者の視界からはなししか見えないという状況である。Original (図 5 左) はなしの側へ膨らんだカーブを示した。一方 False-projective と PublicSelf (図 5 右) はりんごの側に回り込むことで、行為者の目標がなしでないことを示している。

5. PublicSelf モデルによって生成された Legible motion の評価実験

5.1 先行研究における評価実験

Fukuchi et al.[5] は観測者及び行為者間の情報の非対称性が発生する場面において、PublicSelf モデルで生成した Legible motion が情報の非対称性を考慮することで観測者に与える影響を検証すべく実験を行なった。実験は大学生 12 名 (男性 6 名, 女性 6 名, 20-24 歳, $M=22.6$, $SD=1.83$) を対象に実施され、謝礼として実験参加者一名につき 750 円の謝礼金が支払われた。

5.1.1 実験概要

実験はディスプレイの映像観察とキーボードによる入力を用いて行われた。

実験は以下のような手順に沿って行われた。

1. 実験参加者に対して、実験の流れや操作方法を口頭で説明した。
2. 第一段階の練習試行を行った。実験参加者は画面の表示領域にひらがなで断続的に「りんご」もしくは「なし」を表示した画面を観測し、表示した文字に対応したキーを押下する。この試行は実験参加者が任意のタイミングで終了するまで行った。
3. 第二段階の練習試行を行った。実験参加者はテストシナリオとして行為者であるエージェントの挙動を観察し、エージェントの目指す目標の推定を行い、対応するキーを押下することで回答を行なった。テストシナリオは 2 つ提示し、2 つ目の刺激提示終了と同時にこの練習試行は終了した。
4. エージェントの行動を Original, False-projective, PublicSelf の 3 モデルから選び、本実験を行なった。手順 3 と同様の形式でエージェントの挙動を刺激として提示し、エージェントの目標の推定に対応するキーでの回答をするよう促し、キーの押下及び離上タイミングを記録した。刺激は、`center`, `side_visible`, `side_invisible` を含む 5 つのシナリオとダミー刺激 4 つの計 9 つを用意し、ランダムな順番で提示した。
5. 刺激提示終了後、フォーム形式のアンケートの回答を促した。以下にリッカート尺度の質問項目を示す。
 - Q1. エージェントの動作は簡単に予測できた。(Predictability)
 - Q2. エージェントの動作は一貫していた。(Consistency)
 - Q3. エージェントの意図は明確だった。(Clarity)
6. 提示する行動のモデルを残っているものに変更し、再度手順 4 から 5 を繰り返した。提示するモデルの順序はカウンターバランスをとった。

5.1.2 要改善点

この実験において各行動モデルの挙動が実験参加者へ与える影響を検証したが、後述の点より適切な検証

を行うことができなかった。問題として考えられる点を以下に示す。

1. 実験実施上の環境に関して、実施時間が長かった事
2. 提示刺激に関して、ダミー刺激が多かった事
3. 刺激観察実験後の主観アンケートに関して、質問内容の検討が十分に行われていなかった事

1に関して、この実験では各モデルの実験刺激数は9シーンであり、観察が終了し次第主観アンケートを実施し次のモデルの実験刺激を観察する、といった流れを繰り返す形式となっていた。多くの刺激観察を行ったことにより、実験参加者の疲労度や集中力が回答へと影響していたことが考えられる。2に関して、各モデルの挙動の法則性への気づきが回答に影響することがないようにダミー刺激をランダムに混ぜて提示したが、その数は全刺激数27シーンに対して12シーンと、およそ4割強を占める数であった。各モデルのダミー刺激を含んだ一連の刺激提示が終了する度に実施していた主観アンケートは、提示されたダミー刺激やその提示順が実験参加者の印象に大きく影響し、モデルによって生成された挙動そのものに対する主観の回答を得られなかったことが考えられる。3に関して、この実験では3つの項目のリッカート尺度による質問を行っていたが、他の収集データと用途が重複していたり項目数が少なかったりといったことが要改善点として挙げられる。

5.2 本実験における改善点

今回行った実験は、前節で述べた点に関して改善を行ったものである。以下に変更点を示す。

1. 各モデルのシミュレーション刺激を実験参加者ごとにカウンターバランスし、比較対象とするモデルのどれか1つに関しての実験刺激の観察のみを行う形式に変更
2. ダミー刺激の提示を廃止し、分析対象のシーン3つのみを提示する形式に変更
3. 主観アンケートの内容を変更

1に関して、全てのモデルの刺激を全実験参加者に提示していた形式から、モデルを1つ選択し提示する形式へと変更することで、実験参加者の疲労度をはじめとする負荷を軽減し回答への影響を軽減した。2に関して、上記1によりダミー刺激を提示する必要がなくなった為削除した。3に関して、Legible motionの先行研究である Dragan et al.[4] を参考に、本実験で該当する質問項目を選択し設定した。

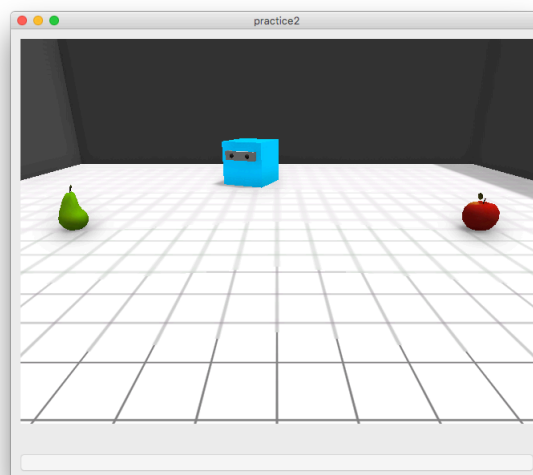


図6 実験プログラムの画面

5.3 実験

5.3.1 実験概要

先行研究での要改善点を改善した内容で再度実験を行なった。実験は大学生20名を対象として行い、実験上不具合が生じたデータ3件を除外した17件を収集し分析を行った。本実験の目的は、情報の非対称性を考慮した PublicSelf モデルによって生成された Legible motion が Original や False-projective と比較して観測者に与える影響を検証することである。

5.3.2 実験で使用したプログラムについて

実験において刺激提示に使用したプログラムを説明する。メインで使用する画面を図6に示す。

画面上部には、実験刺激であるエージェントの挙動を表示する区画がある。実験刺激は毎秒5フレームで表示され、実験参加者はその刺激に対する推定を F, J キーを用いて回答する。キーにはそれぞれりんごとなしのいずれかが割り当てられているが、キーとの対応は実験参加者ごとにランダムに選択される。

画面中央には、実験参加者がキーを押している間に該当のキーが対応する目標を表示する区画がある。例えば、実験参加者がりんごに対応しているキーを押下している間「あなたの入力：りんご」といった表示がされる。

画面下部には、実験もしくは練習試行開始時に押下するボタンが配置されている。各フェーズ終了時には終了した旨が表示される。

5.3.3 実験手順

実験手順を以下に示す。

1. 実験参加者に対して、実験の流れや操作方法を口頭で説明した。
2. 第一段階の練習試行を行った。実験参加者は画面の表示領域にひらがなで断続的に「りんご」もしくは「なし」を表示した画面を観測し、表示した文字に対応したキーを押下するよう教示した。この試行は実験参加者が任意のタイミングで終了するまで行った。
3. 第二段階の練習試行を行った。実験参加者はテストシナリオとして行為者であるエージェントの挙動を観察し、エージェントの目指す目標の推定を行い、対応するキーを押下することで回答を行った。テストシナリオは2つ提示し、2つ目の刺激提示終了と同時にこの練習試行は終了した。
4. 実験を行なった。手順3と同様の形式でエージェントの挙動を刺激として提示し、エージェントの目標の推定に対応するキーでの回答を促し、キーの押下及び離上を記録した。刺激は1モデルに関する挙動で、被験者間でのカウンターバランスをとった。目標の配置別3つのシナリオを用意し、ランダムな順番で提示した。
5. 刺激提示終了後、フォーム形式のアンケートの回答を促した。以下にリッカート尺度の質問項目を示す。
 - ロボットは信頼できると感じた。(Trust)
 - 私はロボットの動きに不快な印象を受けた。(Safety/Comfort,R)
 - ロボットの動きは合理的で頭がいいと感じた。(Ability)
 - ロボットの動きは意外性があった。(Predictability,R)
 - ロボットが何を狙っているのかを予測することは簡単だった。(Legibility)
 - ロボットは自身の意図(ゴール地点)を明確にするように動いていた。(Legibility)
 - ロボットは、観測者(私)が目標を推測する事を助けるような方法で動こうとしていた。(Legibility)

以上の項目への回答をもって実験終了とした。

5.3.4 結果

実験参加者が刺激の観察と同時に入力した推定の精度に関して、各実験参加者ごとに得点を計算した。得点は式5の通り与えた。

$$score = \begin{cases} +1 & (\text{正解時}) \\ -1 & (\text{不正解時}) \\ 0 & (\text{その他}) \end{cases} \quad (5)$$

center, side_visible, side_invisible 各シナリオごとの得点の平均をそれぞれ、図7, 8, 9に示す。また、刺激提示後に行った主観アンケート結果を図10に示す。

5.3.5 考察

center シナリオにおける実験参加者の得点に関して、Original はおよそ8000ms経過するまでの得点平均が0.00から-0.25の間にある。得点が上昇する付近でエージェントが果物に近づいて目標に向き直るシーンであることから、その時点以前では適切な推測が立てられなかった事が考えられる。その一方で、PublicSelf と False-projective は刺激提示1500ms付近から得点の上昇が見られ、このことから早期のタイミングでの推測可能性に影響を与えていると言える。しかし、PublicSelf は False-projective と比較して後半(3000-12000 ms)の得点が低い。刺激の終了まで得点が低い状態が続いていることは、操作ミスもしくは持続して誤った推測をしていたかを判断するのは、今回の実験ではモデルそれぞれの実験データが少数であることから、実験参加者一人の結果への影響力が非常に大きいため難しい。この問題に関しては、データ数を増やして引き続き検討を行う必要がある。

side_visible シナリオでは、Original は center シナリオと比較して早期に得点が上昇しており、False-projective と PublicSelf は center シナリオと同様のタイミングで得点が上昇している事が観察できる。違いとしては、False-projective は得点の上昇するさまが緩やかである事が挙げられる。この点より、side_visible シナリオにおいては PublicSelf モデルはより早く正確に意図の伝達していたといえる。

side_invisible シナリオではより複雑な得点変化がみられる。全体的に、刺激の前半部分はどのモデルも得点がマイナスである事がみとられ、このことより前半はどのモデルも共通して正しくない意図を伝えてしまっているといえる。False-projective と PublicSelf が Original と差をつけられなかった理由として、提示

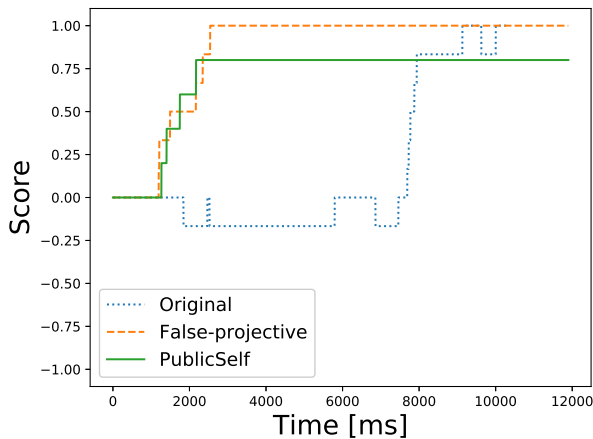


図7 center シナリオにおける実験参加者の平均得点

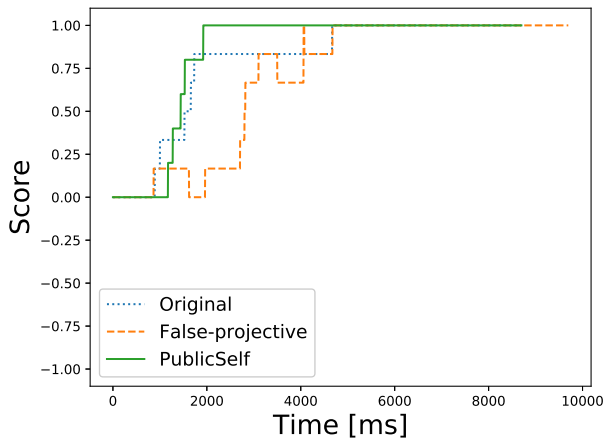


図8 side_visible における実験参加者の平均得点

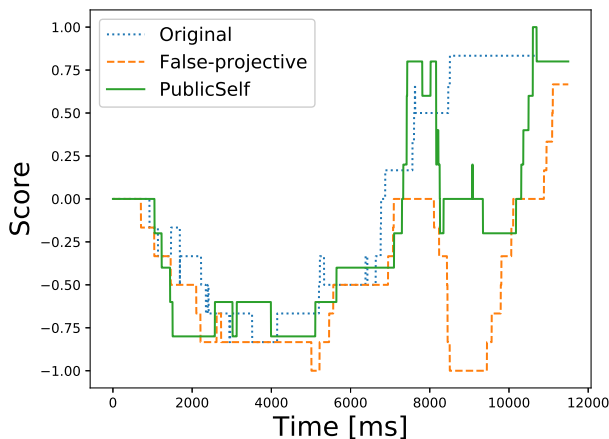


図9 side_invisible における実験参加者の平均得点

する映像の遠近感が十分でなく、False-projective と PublicSelf において行為者が回り込んでいることが伝わりにくかった可能性が考えられる。一方、刺激後半部分 (7000-12000 ms) で特筆すべきは、PublicSelf と False-projective の得点変動である。いずれも得点の変動のタイミングや正誤の方向が概ね一致している、違いとしては得点の値そのものが挙げられる。得点が一度下がった後に上昇しているという共通点が存在するが、PublicSelf は False-projective よりも先に正しい意図を伝達しているという事がいえる。このシナリオでは、Original の後半得点が安定して高い。このことから、side_invisible のような横方向に並んだ目標の、観測者から見えない方の目標を目指す状況では、直線的な動きは終着点により推測しやすく曲線的な動きは推測が困難であったのではないかと考えられる。PublicSelf モデルにおいては、後半に誤った意図の伝達をしてしまう点を解消する事ができれば side_invisible シナリオでも Original と同等の性能を発揮する事ができるため、得点を下げた原因についてさらなる調査を行う事が今後の課題として挙げられる。

主観評価アンケートに関しては、特筆すべき特徴に Original の ability に関する回答平均が他 2 つのモデルに対して大きく下回っている事がある。理由として center シナリオでの挙動が考えられる。図 7 で表されるように、Original と False-projective および PublicSelf の得点の変動タイミングは著しく異なっている。実際、center シナリオの提示刺激を観察するとエージェントは目標が設置してある付近に接近するまでは、観測者から見てちょうど中央を直線上に移動していることから、この得点変動は説明ができる。観測者にとって判断が難しい行動を刺激全体の時間の大半に渡ってとっていたことが、合理性に欠けているという印象を与えていたと考察する。

また、主観アンケートの結果において False-projective と PublicSelf の間には顕著な傾向や差異はほとんど見られなかった。これは、今回の実験形式と質問項目があまりマッチしていなかった事が原因として考えられる。観察のみを行って回答する本実験の形式では、今回の質問項目において主観に対する影響は小さいものであったのではないかと、いう事である。作成する際に参考にした質問は元々、現実においてロボットの動作を実際に協働することで確かめた上で回答する質問であったためである。質問項目を刺激提示形式に合わせて作成する事も解決策としては存在するが、PublicSelf は現実環境での人間との協働を前提

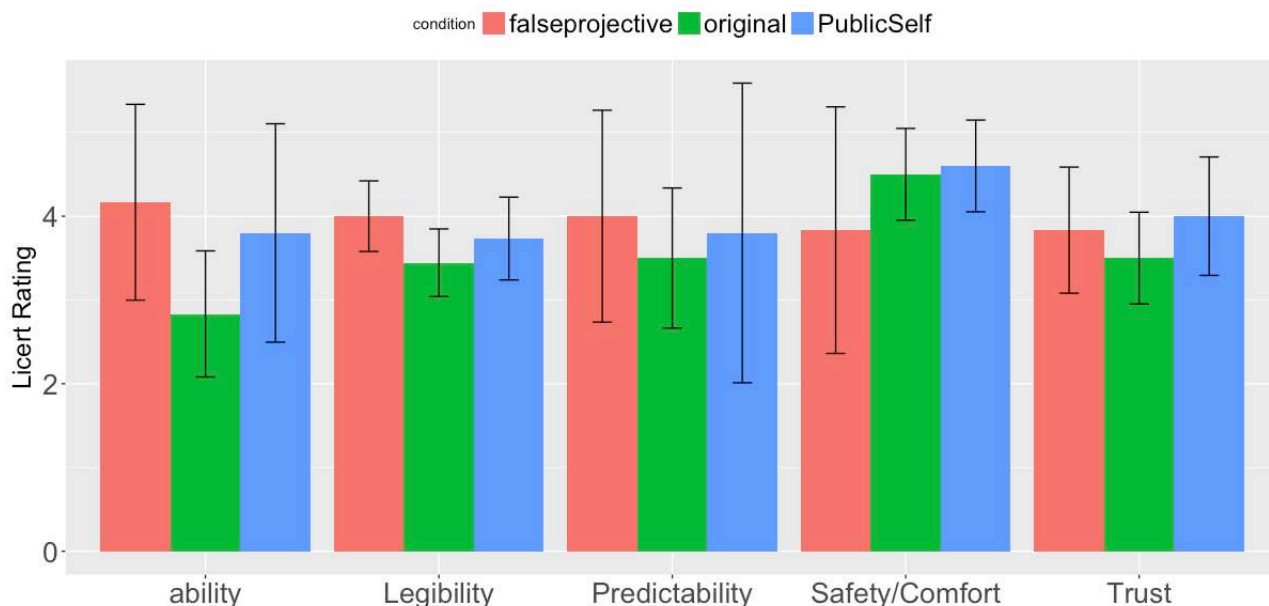


図 10 主観アンケートの回答

としたモデルであるため、想定する状況により近い実験設定である方が性能を評価する上で望ましいと考える。

この問題に関しては、より想定に近い状況への没入感を向上させるために、刺激の提示形式をインタラクティブな形式にすることが挙げられる。Dragan et al.の実験では回答者自身の意思は観測者（この場合回答者と同一個体）の行動に反映される、よりインタラクティブな形式の観察を行っていた [4]。対して本実験はモニタ上で実施する実験であるが、インタラクティブな形式での刺激提示をすることでより適切な評価を得られるのではないかと考える。

6. おわりに

本研究では、Fukuchi et al. が提案した PublicSelf モデル [5] に関して、観測者と行為者間で生じる非対称性を考慮した挙動が、観測者への与える効果を検証することを目的とした実験を行った。結果として、観測者の推測精度の向上は見られたが心理的影響についての確認はできなかった。実験をモデルが前提としている状況に近づけるなど詳細な評価実験の必要性が認識された。今後の展望としては、詳細な評価実験を行うことや、人間との協働において必要な要素を組み込むことなどが挙げられる。

7. 参考文献

- [1] Amodei, D., Olah, C., Steinhardt, J., Christiano, P.F., Schulman, J., Mané, D.(2016). “Concrete problems in ai safety.” *CoRR* abs/1606.06565
- [2] Andersen, R. S., Madsen, O., Moeslund, T. B., & Amor, H. B. (2016, August). Projecting robot intentions into human environments. In 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN) pp. 294-301. IEEE.
- [3] Baker, C.L., Jara-Ettinger, J., Saxe, R., Tenenbaum, J.B.: Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour* 1, 0064 EP (2017). URL <http://dx.doi.org/10.1038/s41562-017-0064>
- [4] Dragan, Anca D., et al. (2015) “Effects of Robot Motion on Human-Robot Collaboration.” *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI '15* pp. 5158. doi:10.1145/2696454.2696473.
- [5] Fukuchi, Y., Osawa, M., Yamakawa, H., Takahashi, T., Imai, M.(2018). “Bayesian inference of self-intention attributed by observer.” , *Proceedings of the 6th International Conference on Human-Agent Interaction, HAI '18*, pp. 310. ACM, New York, NY, USA. DOI 10.1145/3284432.3284438. URL <http://doi.acm.org/10.1145/3284432.3284438>
- [6] Hayes, B., Scassellati, B.(2013) “Challenges in shared-environment human-robot collaboration.” *Collaborative Manipulation Workshop at the ACM/IEEE International Conference on Human-Robot Interaction (HRI 2013)*, vol. 8, p. 9
- [7] Hayes, B., & Shah, J. A. (2017, March). Improving robot controller transparency through autonomous policy explanation. In 2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI

- (pp. 303-312). IEEE.
- [8] Kahn, G., Villafior, A., Ding, B., Abbeel, P., Levine, S. (2018) “Self-supervised deep reinforcement learning with generalized computation graphs for robot navigation.” *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 18 (2018). DOI 10.1109/ICRA.2018.8460655
- [9] Kalashnikov, D., Irpan, A., Pastor, P., Ibarz, J., Herzog, A., Jang, E., Quillen, D., Holly, E., Kalakrishnan, M., Vanhoucke, V., Levine, S. (2018) “Scalable deep reinforcement learning for vision-based robotic manipulation.” *A. Billard, A. Dragan, J. Peters, J. Morimoto (eds.) Proceedings of The 2nd Conference on Robot Learning, Proceedings of Machine Learning Research*, vol. 87, pp. 651673. PMLR (2018). URL <http://proceedings.mlr.press/v87/kalashnikov18a.html>
- [10] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M.A., Fidjeland, A., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., Hassabis, D. (2015) “Human-level control through deep reinforcement learning.” *Nature* 518(7540), 529533
- [11] Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., van den Driessche, G., Graepel, T., Hassabis, D. (2017) “Mastering the game of go without human knowledge.” *Nature* 550, 354 . URL <http://dx.doi.org/10.1038/nature24270>