

# 情報の非対称性を含む Legible Motion に対する印象的評価の検証 Evaluation of legible motions considering information asymmetry

宝田 悠<sup>†</sup>, 福地庸介<sup>‡</sup>, 今井倫太<sup>‡</sup>, 高橋 達二<sup>†</sup>

Yu Takarada, Yosuke Fukuchi, Michita Imai, Tatsuji Takahashi

<sup>†</sup> 東京電機大学 <sup>‡</sup> 慶應義塾大学

<sup>†</sup>Tokyo Denki University, <sup>‡</sup> Keio University

tatsujit@mail.dendai.ac.jp

## 概要

本研究では, 人工エージェントと人間の効果的な協働の実現に焦点を置いている. 協働では参加メンバー間の相互理解が重要となる. Fukuchi et al. は, 人間が人工エージェントを正しく理解しているかを推定するため, 他者の視点から見た自己の心的状態 (信念, 目標) を推測する PublicSelf モデルを提案している. また, PublicSelf モデルを応用することによって, 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と期待される. 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 PublicSelf モデルが効果的な legible motion を生成できるかを検証した. 特に Fukuchi et al. で行われた実験の問題点を改善することで, 生成された legible motion に対する人間の主観的評価を検証した. 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, 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との協働があげられる. 本研究はこの, 人間と人工エージェントとの協働を背景としている.

効果的な協働を実現するには参加するメンバー間の相互理解, 特に互いの目標を知ることが必要になる. また, ロボットアームなどの身体を持つ人工エージェントには, 人間と物理的に衝突してしまう危険性がある. 人間がエージェントの目標を把握できるようにすることは, 事故の危険性を排除する為にも重要な要素である [1].

Fukuchi et al. 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) とそれを観察する人間 (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効果的に伝達する動き, legible motion を生成することができる [5]. 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実験を行った. 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期待されたが, 実験結果からは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されなかった.

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ダミー刺激が多く, 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共存するには, 人がエージェントの振る舞いを理解でき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、役割分担やエージェントの補助など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することが困難であ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が、その中で Dragan et al. 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 [4]。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況を考える。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ている。エージェントの動きを観測している人間には、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難しいといえる。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早い段階から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  $\xi$  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測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。

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

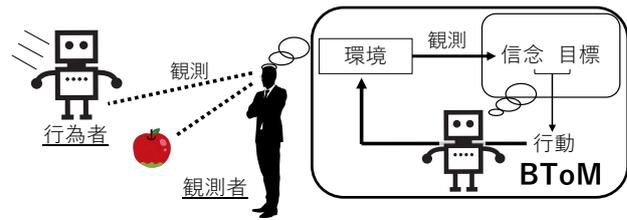


図1 Bayesian Theory of Mind

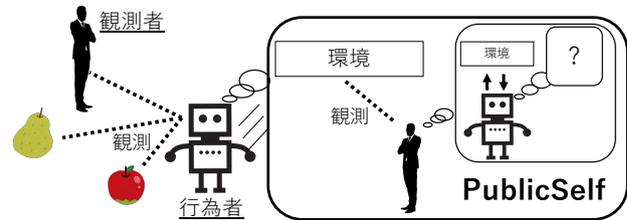


図2 PublicSelf モデル

$C[\xi]$  は軌道  $\xi$  に必要なコストであり、 $\xi'$  は  $\xi$  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]。が、これまでの legible motion の研究では協働を行う環境が限定的であり、人とエージェントが物の位置などといった環境の情報を完全に共有していることを前提であった。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把握してい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) 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)[3]。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し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にしている。また 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」を推定する行為者のモデルで、BToMの入れ子構造を1つ深くしたものである(図2)[5].

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.

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

ここで、 $b$  は信念、上付きの添字<sup>1</sup>はその変数が行為者が観測者に帰属する心的状態、添字<sup>2</sup>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」を示している. この式3は前向きアルゴリズムによって計算することができる.

#### 3.2 Legible motion の生成

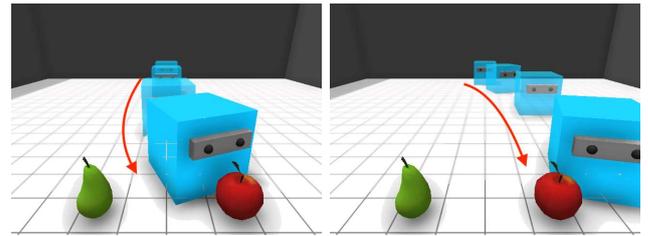
式3によって「観測者が推定する行為者の目標が、実際の行為者の目標と一致する確率」を最大化する行動を選択することができる:

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

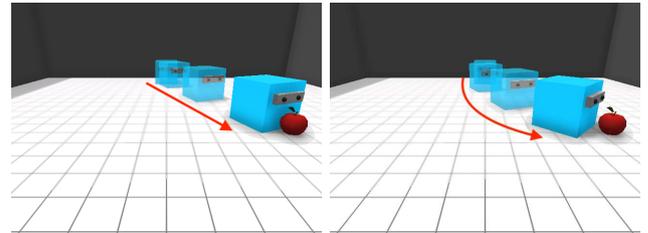
ただし、式4によって選択される行動がエージェントの実際の目標を達成するためのコストを大きく増加させる行動を選択しないよう、選択可能な行動には制限をかける. これは例えば、エージェントの目標が誤解される確率  $P(g^2 \neq g^* | o_{:t}, a_{:t})$  を減少させる行動を取ることで  $P(g^2 = g^* | o_{:t}, a_{:t})$  が大きく上昇することがある一方、そういった行動が実際の目標の達成を阻害してしまうことがあるためである.

### 4. シミュレーション環境における実装

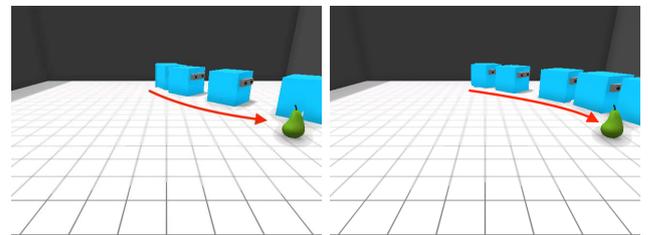
PublicSelf モデルによる legible motion の生成を評価するためシミュレーション環境を用意し、環境内のエージェントに対して legible motion の生成手法を実装した. 環境には行為者であるエージェント、観測者である人(カメラ)とエージェントの目標  $g$  であ



(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

る林檎、梨が存在する. エージェントの行動は、深層強化学習によって林檎か梨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 は、目標に向かう際のコストを最小化する行動を学習している. 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. 本章では行為者が林檎を目標としている場合を考える.

center シナリオでは林檎と梨が観測者の目の前にあり観測者と行為者の両者が林檎と梨を観測できる. そのため、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林檎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林檎の方向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林檎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林檎のすぐ隣に梨があるが、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、林檎の方向に回り込むことで行為者が梨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梨が見えていないため、林檎の方向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` シナリオにおいて、エージェントの目標である林檎は梨のすぐ隣にあるものの、観測者の視界からは見えないというシチュエーションである。Original(図5左)は梨方向へ膨らんだカーブを示し動いた。一方False-projectiveとPublicSelf(図5右)は林檎の方向に回り込むことで、行為者の目標が梨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、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た実験を行なった。しかし、この実験ではPublicSelf モデルとその他のモデルの間に主観的評価の差異が見られなかった。その理由として、以下のような実験上の問題点が挙げられた。

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

#### 5.1.1 実験概要

Takarada et al. では、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を目的として、これらの点に関する改善を行った上で実施した。以下にFukuchi et al. で行われた実験からの変更点を示す。

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

項目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。項目2に関して、上記項目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。項目3に関して、legible motion の先行研究である Dragan et al. を参考に、実験で該当する質問項目を選択し設定した[4]。

実験は大学生20名を対象として行った。結果として、各シナリオにおける意図推定過程、特に`side_invisible` シナリオでの複雑な推定過程を観察することができた。また、刺激観察後に行った主観評価アンケートでモデルごとに全てのシーンを観察後評価を行ったところ、情報の非対称性有無による評価の差異が見られなかった。

## 5.2 実験

### 5.2.1 実験概要

本研究では、観測者 - 行為者間に生起する情報の非対称性が発生する場面において、行為者が情報の非対称性を考慮することが観測者に与える影響を検証することを目的とする。前回の実験と比較した本実験の変更点を以下に示す。

1. 刺激提示の形式を、単一のシナリオでの各モデルの動作をランダムな順番で一つずつ提示する形式から、各モデルの動作を並べて同時に観察する形式に変更した。
2. `side_invisible` シナリオでの提示刺激について、各要素の位置を変更した。
3. クラウドソーシングで実施可能な形式に変更し、実験参加者を募集した。

項目1について、実験参加者ごとの評価のばらつきを抑える目的で変更を行った。項目2について、特に複雑な状況である`side_invisible` シナリオにおいて、エージェントの動き出しや最終的な目的地が見え辛いために判断が困難である可能性を考え、エージェントとカメラの距離を近づける、カメラからより見え

やすい位置に目標を移動するなどの変更を行った。項目 3 について、実験参加者のバイアスを極力排除し多様な回答を得るため変更を行った。

### 5.2.2 実験手順

実験は Web ページをブラウザで表示する形式で行った。実験手順を以下の通りである。

1. クラウドソーシングサイトにて、実験参加者に実験ページへの遷移を指示した。
2. 実験ページで実験参加者に対し説明文章を提示し、実験内容に関する各注意事項の確認有無をチェックボックスで入力するよう促した。
3. はじめに、本実験で実際に使用する刺激と同様のレイアウトで IMC(Instructional Manipulation Check) を実施した。説明文に記載されている指示通りの選択肢を選択しなかった場合にはエラー文章を表示し、IMC 用の問題を表示し続けた。
4. 今回比較する 3 つのシナリオのうち 1 つに関して、動画形式の提示刺激を横に並べて表示した。この時動画を再生する回数に上限は設けなかった。
5. 手順 4 で観測した 2 つの動き方に関して、どちらがどれくらいわかりやすいかを 5 段階評価で入力した。
6. 手順 4 から手順 5 に関して 3 つのシナリオ全てを評価するまで繰り返した。各シナリオの提示順、提示するモデルの提示位置(左右)はそれぞれカウンターバランスをとった。

### 5.2.3 結果

実験参加者はクラウドソーシングで募集した 50 人(男性 33 人, 女性 17 人)に対して、作業時間の上位及び下位各 5% のデータを除いた計 45 件を分析の対象とした。実験結果は以下の通りである。実験参加者が回答した各モデルの評価値を t 検定により検定した。 $center(t(43) = 7.76, p < .05)$ ,  $side\_visible(t(43) = -19.94, p < .05)$ ,  $side\_invisible(t(43) = 7.52, p < .05)$  シナリオにおいて、各モデルの評価平均値に有意差が認められた。本実験で得られた評価データの平均値を図 6 に示す。ここで、1 つの刺激に 2 つのモデルが当てはまる場合同一の点数が記述されている。また、一対比較法による各モデルの比較結果を図 7 に示す。

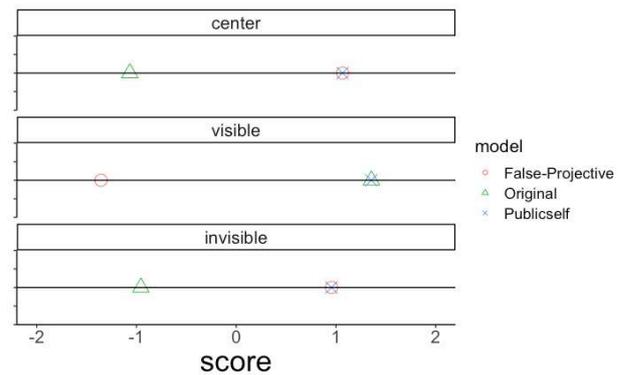


図 6 シナリオ毎の回答値の平均

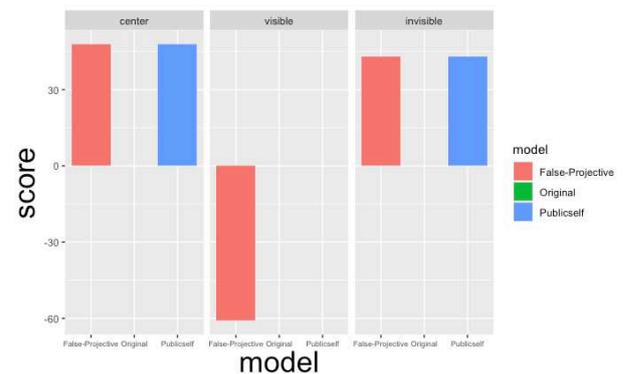


図 7 Sheffe の一対比較法を用いた評価の相対的スコア

### 5.2.4 考察

今回の実験では、実験参加者の認知的負荷軽減のため同じ動きを生成する場合は 1 つの刺激として提示していた。 $center$  シナリオ、 $side\_invisible$  シナリオでは PublicSelf モデルと False-Projective モデルを同一の刺激とし、 $side\_visible$  シナリオでは Original モデルと PublicSelf モデルを同一の刺激としている。そのため、 $center$  シナリオと  $side\_invisible$  シナリオでは legible Motion の有無による評価の差を、 $side\_visible$  シナリオでは情報の非対称性考慮による影響を見ることが出来る。また、図 6 より各モデルの平均値の差は  $center$  シナリオでは 2.12,  $side\_visible$  シナリオでは 2.71,  $side\_invisible$  シナリオでは 1.91 ほど生じている。このことより、これらのシナリオ下においては legible Motion と情報の非対称性を考慮することは人間にとってわかりやすい動きをしているということが言える。これは前回の実験であるリアルタイム推測実験から得られた結果とは一部異なるものではあるが、前回と似た傾向がみられた  $center$  シナリオ及び  $side\_visible$  シナリオにおいては、より顕著に差があらわれていると言えるだろう。その一方で、 $side\_invisible$  シナリオの結果は前回の実験の結果と異なり、各モデル間の差がみられてい

る。本実験と前回の実験の差異としてエージェントの位置を変更したことが挙げられる。本実験ではよりカメラに近く、観測しやすい位置へと移動していた。このことから、PublicSelf モデルを現実世界へ適用するための条件の一つとして、観測者からどれほど見えているかという連続的な捉え方をする必要はあるとは言えないだろうか。

また、現在は主観的評価のみを対象としてクラウドソーシングで検証をしており、リアルタイムで判断するような実際の推測自体を検証する実験も今後行っていく必要がある。

## 6. おわりに

本研究では、Fukuchi et al. が提案した PublicSelf モデルに関して、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、情報の非対称性を考慮することが主観的評価を向上させていることがわかった。今後の展望としては、より現実世界への適用に必要な要素をモデルに組み込むことや現実により近い状況での実験・検証などが挙げられる。

## 7. 参考文献

### 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. 51–58. 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. 3–10. 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. 1–8 (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. 651–673. 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), 529–533
- [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>