

他者の内部状態を反映したインタラクティブなエージェント行動 選択手法の提案

Proposal of Interactive agent behavior selection considering internal states of others

田嶋 沙和子^{1*} 小山 宗三² 栗原 聡¹
Sawako TAJIMA¹, Shuzo KOYAMA² and Satoshi KURIHARA¹

¹ 慶應義塾大学理工学部

¹ Faculty of Science and Technology, Keio University

² 慶應義塾大学大学院理工学研究科

² Graduate School of Science and Technology, Keio University

Abstract: ロボットが人と円滑に共同作業を行うためには、人の行動や意図を推測し、人に合わせて行動を取る必要がある。本論文では、他者が介入した場合にも柔軟に自分の行動を決定する行動選択ネットワークを提案する。提案手法では、古典的プランニング手法を拡張し、他者の内部状態の推測と自分の行動のプランニングの両方を可能にした。積み木タスクを用いた実験を行い、他者の介入の仕方での自分の行動系列がどのように変化するか検証した。今後の課題としては、行動の実行とプランニングの並列化、大規模かつ複雑な環境での実験、複数のゴール設定への対応の3点が挙げられた。

1 はじめに

従来の製造ラインで用いられてきた産業用ロボットは、単純作業を単独で行うものがほとんどであった。しかし、近年の法改正による規制緩和により、人とロボットが同じ空間で作業を行うことが可能となり、協働ロボットの開発が進んでいる。ロボットが人と共同で作業を行うためには、人が想定外の行動をした場合にも臨機応変に対応する必要がある。一方で、現状の協働ロボットは人とインタラクションして作業を行うことが難しいという課題がある。将来的には、人の意図と行動を読み取り、ロボット同士あるいは人と協力しながら作業を行うことが出来るロボットが求められる社会になると考えられる。

私たちが他者と協調して行動するためには、他者の意図と行動を予測して動く必要がある。しかし他者の意図は直接目に見えるものではないため、私たちは他者の動作や表情から予測している。人と同じ空間で作業を行う協働ロボットは人の意図を読み取ることが求められ、2.2節にある様々なアプローチで他者の意図や行動を推測する研究が行われている。

本研究の目的は、他者が介入した場合にも柔軟に行動を選択可能なプランニング手法の提案である。従来

のプランニング手法では、他者と共同作業を行うことは想定されていなかったため、他者は環境の一部として扱われていた。一方、他者の行動予測アーキテクチャでは他者の行動予測を自分のプランニングに結び付けていなかった。そこで、本研究ではプランニング手法である Agent Network Architecture (以下 ANA) [1] と横山らの他者の行動予測を行うアーキテクチャ[2]を組み合わせてることによって、他者と円滑に協調作業を行うことを目指している。本論文では、ANAの自分の行動決定に関するモジュールを複製し他者の行動決定に関するモジュールとすることで、自分の行動決定プロセスを他者の行動予測に用いることを提案した。最終的に、提案システムを他者と共同作業をするストーリーにおいて検証し、その有効性を確かめた。

2 関連研究

2.1 プランニングの関連研究

2.1.1 ANA

ANA[1]はネットワーク状に結合された単位行動群が互いに活性伝播を行うことにより、プランニングを行うアーキテクチャである。ANAはSTRIPS[3]のように単位行動(モジュール)を持ち、それらを結ぶネットワー

*連絡先: 慶應義塾大学理工学部
〒 223-8522 神奈川県横浜市港北区日吉 3-14-1
E-mail: sawako_tajima.0727@keio.jp

クで構成されている。モジュールは action モジュールと belief モジュールの 2 種類に分けられる。action モジュールは「カップを持ち上げる」といった実際の行動を表し、belief モジュールは「物体を認識する」のような実際の行動ではないものを表す。各モジュールは、condition list, add list, delete list を持つ。condition list は実行するために必要な状態を要素とするリスト、add list は実行後に現環境に追加される状態を要素とするリスト、delete list は現環境から削除される状態を要素とするリストを表す。

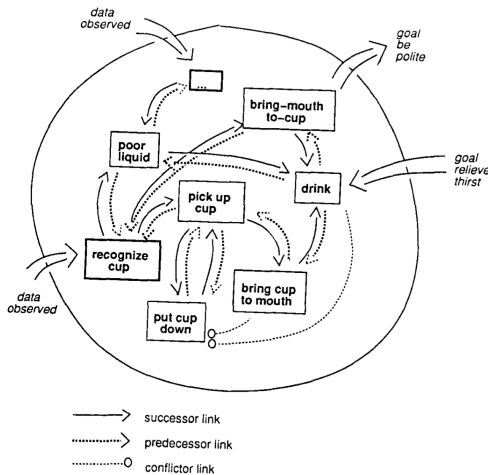


図 1: ANA のネットワーク [1]

「喉の渇きを癒す」ときの ANA のネットワークを図 1 に示す。ANA のネットワークは、predecessor link, successor link, conflictor link の 3 種類で構成されている。このネットワークを通じて、モジュール間で刺激を伝播して蓄積された刺激の量を活性レベルと言う。また各モジュールは、活性化に必要な最小限の活性レベルの値である閾値を持つ。活性レベルが閾値を超えると、そのモジュールは活性化され、モジュールが実行される。この一連の流れをセレクションと呼ぶ。セレクションを繰り返すことで、現環境から実行できるモジュールを選択し実行している。

2.2 他者の行動予測の関連研究

他者の行動予測に関する研究は「他者の振る舞いから、時系列パターンを学習する」アプローチと「自身の行動決定プロセスから他者の意図を推定する」アプローチの 2 つがある。前者は、自分の理解と他者の理解は独立するという考え方である。他者の行動パターンを学習する研究として、Schydlo ら [4] は動作パターン、坂本ら [5] は状態の変化を用いている。その他にも様々な手法 [6][7][8][9] が挙げられる。後者は、人は他

者の行動を模倣することで、他者の意図を理解することが出来るという考え方であり、横山ら [2] の研究が挙げられる。

2.2.1 インタラクシオンシーンにおける他者の意図推定

横山ら [2] は、協調行動における他者の行動予測の有効性を示した。横山らの研究では、他者とインタラクシオンする際に、人は大きく分けて受動的戦略と能動的戦略の 2 つの戦略を持つとしている。この前提を基に、自分の行動を決定するアーキテクチャを提案している。受動的戦略とは、他者の行動から意図を予測し、自分の行動を決定する行動戦略である。他者の行動から計算モデルを用いて、他者の意図を推測している。この計算モデルは、自分の行動を決定する計算モデルをそのまま使用しており、自分が他者の立場に立ったときにどのような意図を持つか推測することが出来る。一方で、能動的戦略では自分の行動から意図を他者に推測させ、その意図通りに行動してもらおうとする行動戦略を取る。受動的戦略とは異なり、他者が自分を観察しやすい行動を選択し、他者が自分の意図に合わせることが期待している。行動戦略に基づいて他者の行動予測は以下の 4 つのレベルに分類される。

- レベル 0: 他者の意図は考慮せず、自分の意図のみで自分の行動を決定する。
- レベル 0*: 他者に自分の意図通りに行動してもらえるように、自分の行動を決定する。
- レベル 1: 他者の行動から意図を推定した結果を基に、自分の行動を決定する。
- レベル 2: 「自分が他者にどう予測されるか」を自分が考えて、自分の行動を決定する。

レベル 0 は、基本的な行動戦略であり、他者の意図は予測せず、自分の目的を達成することを意図して行動する。レベル 1 は、他者の行動から意図を予測し、自分の行動を決定する。レベル 2 ではより深く予測しており、自分の意図を他者にどう予測されるかを考え、自分の行動を決定する。受動的戦略はレベル 1 と 2 に該当する。また、能動的戦略はレベル 0* に該当する。これは自分の意図通りに他者を行動させることを目的としている。

横山らのアーキテクチャでは、自分がどの行動を選択するかは条件付き確率から算出している。この条件付き確率は、Q-learning によって事前に学習された同時確率から求められており、未知の環境においても適した行動を取ることが出来る。

3 提案手法

提案するシステムは、2.1.1 項で挙げたプランニング手法の ANA に 2.2.1 項の横山らの他者の行動予測アーキテクチャを取り入れたものである。他者の行動予測と自分の行動決定の両方を ANA で行うことで、他者の行動を先読みして自分が取るべき行動を選択することを可能にした。プランニング手法に ANA を用いることには、以下の 4 つの利点がある。

1 つ目は、モジュール群に拡張性がある点である。実環境では、目的を達成する前に新しい行動を獲得することがある。自分の能力が向上したり、新たな道具が出現したりすることによって、実行出来る行動が増えた場合にも対応することが出来る。

2 つ目は、活性レベルと閾値の最適化に強化学習や遺伝的アルゴリズムを用いることができる点である。モジュールが活性化され、現環境で実行可能な場合、そのモジュールはすぐに実行されるため最適な行動を取ることが出来ない場合もある。伝播される活性レベルや閾値は、強化学習や遺伝的アルゴリズムを用いることで最適な行動に近付いていくと考えられる。

3 つ目は、自分の行動決定に関するモジュールの複製を他者の行動決定のモジュールとして扱うことが出来る点である。自分と他者のモジュールは動作自体は変わらず、動作を行う手が自分であるか他者であるかの違いである。そのため、複製したモジュールの condition list, add list, delete list の自分の手に関する状態を他者の手に関する状態に置き換えると、他者のモジュールとして使うことが出来る。他者のモジュールの複製は容易に出来るため、他者の人数が増えたときにも対応がしやすいという利点がある。

4 つ目は、他者の行動予測と自分のプランニングを ANA のみで行うことが可能な点である。上記で述べたように、他者のモジュールは自分のモジュールを複製したものであり、モジュールが持つ機能は殆ど同じである。そのため、他者のモジュール同士もネットワークで結ばれ、ネットワークを通じて刺激を伝播することが出来る。他者のモジュール間でセレクションを繰り返すと、他者のモジュールのいずれかが活性化されることがある。ここで活性化したモジュールは、自分が他者の立場に立ったときに選択する行動であり、他者の行動を予測することと等しい。

また、本研究は小山ら [10] の手法を用いて現環境を取得し続ける機能を従来の ANA に追加した。他者の行動を検知するには、環境の変化を常に取得する必要がある。例えば、机の上に置かれていたコップが消えて他者の手の中にコップがある状況では、「他者がコップを持ち上げた」という行動の実行を検知することが出来る。

他者モデルを適用した ANA は、横山らの他者の行動

予測のレベル 1 を想定しており、最初に他者が行動したかどうかを確認し、自分の行動を決定している。他者モデルを適用した ANA は、自分に関するモジュールとその複製である他者に関するモジュールで構成されている。以降では、自分に関するモジュール群とそのネットワークを自分 ANA、他者に関するモジュール群とそのネットワークを他者 ANA とする。

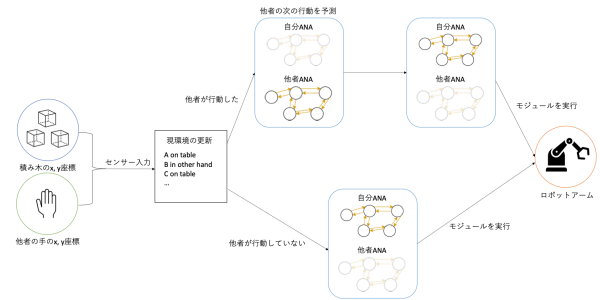


図 2: 他者モデルを適用した ANA

図 2 に他者モデルを適用した ANA を示す。センサー入力から積み木と他者の手の座標を取得し、その位置関係から現環境を更新する。現環境の取得後、自分 ANA が保存している環境と異なるときとそうでないときで処理が分かれる。自分 ANA が保存している環境と現環境が異なるとき、他者が行動したと判定する。これは、自分の行動による環境の変化は自分 ANA に保存されるが、他者の行動による環境の変化は保存されないためである。次に、他者が実行したモジュールを探索し、次に他者が実行する行動を予測した上で、自分が実行するべき行動を選択し実行する。一方、自分 ANA が保存している環境と現環境が変わらないとき、他者は何もしていないと判定される。この場合は従来の ANA と同様、自分 ANA 内で実行するべき行動を選択する。他者モデルを適用した ANA のアルゴリズムは以下の通りである。

他者モデルを適用した ANA のアルゴリズム

1. 現環境を取得する。
2. 環境の変化で他者がモジュールを実行したか判定する。
3. モジュールを実行したと判定された場合、以下の処理を実行する。
 - (1) 自分 ANA が保存している環境の要素を condition-list に持ち、現環境の要素を add-list に持つモジュールを他者 ANA 内で探索し、他者が実行したモジュールを決定する。
 - (2) 他者が実行するであろうモジュールを他者 ANA 内でセレクションを繰り返す。
 - (3) 他者 ANA で活性レベルが閾値を超え、現環境から実行可能なモジュールを仮に実行したとする。
 - (4) 実行したとされるモジュールの add-list を現環境に追加し、delete-list を現環境から削除したものを内部環境とする。
 - (5) 仮の環境下で自分が実行すべきモジュールを自分 ANA でセレクションを繰り返す。
 - (6) 自分 ANA で活性レベルが閾値を超えた場合、現環境から実行可能なモジュールを実行する。
 - (7) 実行したモジュールの add-list を現環境に追加し、delete-list を現環境から削除する。
 - (8) 実行したモジュールの活性レベルを 0 にする。
4. モジュールを実行したと判定されなかった場合、以下の処理を実行する。
 - (1) 自分 ANA でセレクションを繰り返す。
 - (2) 活性レベルが閾値を超えた場合、現環境から実行可能なモジュールを実行する。
 - (3) 実行したモジュールの add-list を現環境に追加し、delete-list を現環境から削除する。
 - (4) 実行したモジュールの活性レベルを 0 にする。
5. 1~4 を繰り返す。

これにより、従来のプランニングでは行えなかった、他者が行動した場合と他者が行動しなかった場合のどちらにも対応し、自分の行動を決定することが可能となった。そのため、他者と協調行動を取れるようになることが期待される。さらに、他者が行動したときに次の他者の行動を予測した上で自分の行動を決定するため、効率良く協調行動を行えると考えられる。

4 実験方法

4.1 実験ストーリー

本実験は、他者との共同作業を想定した積み木タスクで行った。またこの実験では、ロボットアームを自分、ロボットアームと協働作業を行う人間を他者とする。表 1 の 8 つの実験で他者とインタラクションを行い、ゴールを達成しようとするときの自分 ANA と他者 ANA の行動系列を調べた。他者と共同作業をするストーリーは大きく 3 つに分かれる。まず、他者がゴールの達成に協力する場合である。今回の実験では、実験 1~3 に該当する。次に、他者がゴールの達成を邪魔する場合である。これは実験 4~5 に該当する。最後に、他者が何も介入しない場合である。これは実験 6 に該当する。また本実験では、他者が協力する行動と邪魔する行動の両方を取った場合についても調べた。実験 7 は最初に他者が協力し、次に邪魔する行動を取る場合であり、実験 8 はその逆の場合である。

表 1: 実験ストーリー

	ストーリーの内容
実験 1	最初に他者が積み木 B を持ち上げる。
実験 2	最初に他者が積み木 A を持ち上げる。
実験 3	途中で他者が積み木 C を持ち上げる。
実験 4	途中で他者がフィールド内に置かれた積み木 B を取り除く。
実験 5	途中で他者が積み木 A を持ち上げ、フィールド内に置く。
実験 6	他者は何もしない。
実験 7	最初に他者が積み木 B を持ち上げた後、積み木 C を取り除く。
実験 8	途中で他者が積み木 B を取り除いた後、積み木 A を持ち上げる。

実験 1~3 では、他者の行動から意図を予測し、それに合わせた行動を選択できるかを確かめた。他者が次に実行しそうな行動とは異なる行動を選択する場合、他者の意図に合わせた行動を取ったことが確かめられる。また、他者が行動してから自分の実行すべきモジュールが活性化されるまでにかかるステップ数について環境取得型 ANA[10] との比較を行い、他者に合わせた行動を迅速に取ることが出来るかを確かめた。実験 4,5 では、他者に邪魔された場合の順応性を確かめた。他者に邪魔されたことでゴールから遠ざかった状態を自分で挽回する行動を取る場合、順応性を確かめられる。実験 6

では、他者が全く介入しない場合を調べた。この実験では他者に関するモジュールが追加された場合も、自分だけでゴールを達成できるかを検証した。また実験7,8では、他者の行動から意図に合わせた行動選択と他者に邪魔された場合の順応性の両方を確かめた。これは、上で述べた実験1~3と実験4,5の両方を満たす場合に確かめることができる。

4.2 実験環境

本実験では、テーブルとフィールドの2つに分かれているシミュレーション環境で行った。自分と他者は環境の全てを取得することができ、どの位置に積み木があっても手が届くものとした。図3に本実験の初期状態とゴールの状態を示す。初期状態はテーブル上に積み木A,B,Cが全て置いてある状態とし、ゴールの状態はBの上にAが積み、隣にCが置かれている状態とした。フィールド内に積み木を与えられた形通りに積めたときをゴールとした。



図 3: 初期状態とゴールの状態

5 実験結果

実験結果を表2に示す。また、表3は実験1~3で環境の変化を検知してから自分のモジュールが活性化されるまでにかかったステップ数についての結果である。

表 2: 各実験ストーリーの結果

	実験結果
実験 1	ゴール達成のために必要な行動であり、次に他者が実行すると予測された行動とは異なる行動を選択した。
実験 2	実験 1 と同様の結果が得られた。
実験 3	途中から他者が協力する行動を選択する場合にも、実験 1 と同様の結果が得られた。
実験 4	人が邪魔する行動をすぐに回避することが出来た。
実験 5	実験 4 と同様の結果が得られた。
実験 6	他者が全く介入しない場合であっても、自分の行動のみでゴールを達成することが出来た。
実験 7	他者が協力したときには実験 1, 邪魔したときには実験 4 と同様の結果が得られた。
実験 8	実験 7 と協力と邪魔の順番を入れ替えたときも、実験 7 と同様の結果が得られた。

表 3: 提案手法と環境取得型 ANA のステップ数の比較

	提案手法	環境取得型 ANA
	[steps]	[steps]
実験 1	1	3
実験 2	1	3
実験 3	1	1

6 考察

実験1~3で他者の協力の仕方による比較を行った。実験1,2では他者が持ち上げた積み木、実験3では他者が協力するタイミングを変化させたが、どれも他者の意図に合わせた行動を取ることが出来ており、大きな違いは見られなかった。また、実験4,5では邪魔のされ方による比較を行った。どちらも他者が行った邪魔する行動をすぐに挽回することが出来ていたが、実験4では不必要な行動を行う場面が見られた。実験4では、自分が実行したモジュールを再度実行する必要があり、そのモジュールが活性化されるまでにステップを要したと考えられる。本実験では、他者も自分と同じゴールを持つとしているため、他者に邪魔をされたときに他者の次の行動を正しく予測することが出来ないことも理由として挙げられる。

従来の ANA での結果と比較すると、実験1~5において従来の ANA では見られなかった行動系列が見られた。自分のみで行った場合、自分の手が空くまでは他の積み木に関するモジュールは実行することが出来ない。しかし他者と共同で行う場合、他者が代わりに実行することが出来る。そのため、自分モジュールと他者モジュール間に新たな link が張られるようになった。また従来の ANA の場合、ロボットは現環境から自分が実行すべき行動のみを選択していたが、提案手法では人間の次の行動を見越した上で、自分の行動を決めることが可能となった。

また環境取得型 ANA との結果を比較すると、表3より、実験1,2において提案手法は環境取得型 ANA よりモジュールの活性化までにかかるステップ数が少ないことが分かる。これにより、提案手法の方が迅速に人の意図に合わせた行動を選択することが出来ると考えられる。

さらに、本実験の結果と他者の行動予測手法での結果との比較を行う。[4]の動作パターンから他者の行動を予測する研究では、積み木タスクで同様の実験を行った場合、同じ積み木に関する行動は予測することが出来る。しかし、次の行動は異なる積み木に関する場合が多く、予測することが難しくなる。提案手法では他者が共通のゴールを達成するために行動することを想定しているため、他者が協力する実験1~3において次

の行動を正確に予測したと考えられる。ただし、この研究で用いられている手法は、行動が完了する前に動作パターンからどの行動かを判別するため、提案手法と組み合わせることでより速く行動を予測できると考えられる。

最後に、[5]の状態の変化によって他者の行動予測を行う手法との比較を行った。この研究の手法を本実験に適用し、他者と協力しながらゴールを達成することをタスクとして学習を行うことで、作成されたデータベースで検証することを考える。実験1~3は学習の前提と他者の介入の仕方が一致しているため、適した行動を選択することが出来ると考えられるが、実験4,5のように他者が邪魔する場合には対応できないと考えられる。提案手法では他者が邪魔をした場合にも、セレクションによって実行すべきモジュールの活性レベルが上がり、挽回する行動を即座に取る事が出来るようになった点で優れている。

7 結論

本研究では、他者モデルを適用した ANA を用いて、他者とインタラクションしたときの自分の行動系列がどのように変化するかを検証した。実験1~3より、他者が協力的な場合には、他者の行動から意図を予測し、それに合わせた行動を選択していることが確認された。実験4,5より、他者が邪魔をする場合にも、邪魔されたことに対してすぐに挽回していることが確認された。実験6より、他者が何もしない場合は、自分だけであっても目的を果たそうとすることが確認された。実験7,8より、他者が協力する行動と邪魔する行動を取る場合には、その両方に対応することが確認された。このように、提案手法が、同じゴールであっても自分の行動系列を他者の行動に合わせて動的に変化できる適応力を持ち、導入した他者 ANA による柔軟なユーザとのインタラクションを可能とすることを確認することができた。

7.1 今後の課題

本研究の他者モデルを適用した ANA では、自分がモジュールを実行している間に環境が変化した場合、対応することが難しい。理由として、自分のモジュールを実行しようと動いた際に他者によって実行されてしまった場合、モジュールの実行が完了するまでプランニングを行わないため、そのモジュールを実行してしまうからである。他者の行動を検知し、次に実行すべき行動に切り替えれば、円滑に協働作業を行うことが出来ると考えられる。従って、今後モジュールの実行とプランニングを並列で行うことが必要になる。

また本実験は1人の他者のみが介入する積み木タスクで行ったが、下記の NEDO プロジェクトに組み込むためには、多くの他者が介入する大規模かつ複雑な環境に対応する必要がある。そのため、他者 ANA を複数持つアーキテクチャを取り入れる必要がある。

最後に、複数のゴール設定への対応が挙げられる。現実世界では、ゴールが複数存在する場合がある。協調しながら、ゴールの候補のどれかを達成する場合には、人の行動からどのゴールを目指しているのかを判断し、ロボットの行動を決める必要がある。以上の3点を今後の課題としたい。

謝辞

本件は、NEDO・人と共に進化する次世代人工知能に関する技術開発事業「インタラクティブなストーリー型コンテンツ創作支援基盤の開発」の助成を受けたものである。

参考文献

- [1] Maes, P.: The Agent Network Architecture (ANA), *ACM SIGART Bulletin*, Vol. 2, No. 4, pp. 115-120 (1991).
- [2] Yokoyama, A. and Omori, T.: Modeling of Human Intention Estimation Process in Social Interaction Scene, *International Conference on Fuzzy Systems*, Barcelona, Spain, IEEE, pp. 1-6 (2010).
- [3] Fikes, R. E. and Nilsson, N. J.: Strips: A New Approach to the Application of Theorem Proving to Problem Solving, *Artificial Intelligence*, Vol. 2, No. 3-4, pp. 189-208 (1971).
- [4] Schydlow, P., Rakovic, M., Jamone, L. and Santos-Victor, J.: Anticipation in Human-Robot Cooperation: A Recurrent Neural Network Approach for Multiple Action Sequences Prediction, *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1-6 (2018).
- [5] 坂本覚仁, 松原崇充, 木戸出正繼: 状態予測-行動則モジュールデータベースを用いたヒューマン-ロボットインタラクションにおける行動則の適応, *電子情報通信学会論文誌. A, 基礎・境界 = The transactions of the Institute of Electronics, Information and Communication Engineers. A*, Vol. 95, No. 1, pp. 107-116 (2012).

- [6] Koo, S. and Kwon, D.-S.: Recognizing Human Intentional Actions from the Relative Movements between Human and Robot, *RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication*, Toyama, Japan, IEEE, pp. 939–944 (2009).
- [7] Tahboub, K. A.: Intelligent Human-Machine Interaction Based on Dynamic Bayesian Networks Probabilistic Intention Recognition, *Journal of Intelligent and Robotic Systems*, Vol. 45, No. 1, pp. 31–52 (2006).
- [8] Schrempf, O. C., Albrecht, D. and Hanebeck, U. D.: Tractable Probabilistic Models for Intention Recognition Based on Expert Knowledge, *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, San Diego, CA, USA, IEEE, pp. 1429–1434 (2007).
- [9] Shuji, T. and Takashi, O.: Identification and Learning of Other’s Action Strategies in Cooperative Task, *8-Th International Conference on Artificial Life and Robotics (AROB8th’03)*, pp. 40–43 (2003).
- [10] 小山宗三, 高屋英知, 加藤慶彦, 覚井悠生, 栗原 聡: インタラクティブ性のある適応型行動選択ネットワークの提案, 研究報告知能システム (ICS), Vol. 2020, No. 12, pp. 1–7 (2020).