

# 異種エージェント間における協調のための 相互意図理解の研究

加藤 慶彦<sup>1,a)</sup> 栗原 聡<sup>1,b)</sup>

概要：科学技術の発展により、ロボットが産業目的ではなく、生活の補助として使用されるようになってきた。しかし、現在のロボットの行うことのできるインタラクションはあらかじめプログラムされたものであり、ロボットの身体性で行うことのできるインタラクションのうちの一部しか使用することができていないため、ロボットと共生をするほど浸透はしていない。人間はお互いの意志や意図を伝えあい、協調を行って社会を形成し、生活を行っている。その意志や意図を伝える手段は言語のみではなく、表情、仕草、様々な要素が複雑に絡み合って形成される。例えば、人混みを歩く際の動きが挙げられる。人間は人が密集して歩行している中でも相手の足の動き、目が見ている方向、身体の向き、顔の表情など、様々な要素を観察し、ぶつからないように進む事ができる。そうして形成された社会に、ロボットが受け入れられる為には、人間との意思疎通をし、協調行動を行えるようになる必要がある。しかし、人間の意思疎通モデルは極めて多量のセンサー、文化的に構築された自然言語も相まって極めて複雑になっており、それに対応するインタラクション系の設計をいかにするのか。に対する答えは見つかっておらず、意思疎通モデルの一部しか利用することができていない。また、ロボットの身体性や性能は多種多様であり、それらに合わせて1つ1つ設計をい行うには限界がある。よって、ロボットの側から人とのインタラクション（意思疎通の手段）を獲得する必要がある。本研究では意思疎通の手段の獲得がどのような過程進むのか、を明らかにしたいと考える。しかし、ロボットがインタラクションを獲得するまで人間がロボットと一緒に試行錯誤を行うのは不可能である。そこで、本研究では、お互いの意思疎通の手段が不明なロボット同士を用いて、それらが意思疎通を行い、協調することが可能となるのが実験を行う。

## Intention Understanding for Cooperation between Two Different Agents

YOSHIHIKO KATO<sup>1,a)</sup> SATOSHI KURIHARA<sup>1,b)</sup>

### 1. はじめに

近年、センサやアクチュエータ、計算機の発展により、ロボットによって行うことのできるが増えた。それにより、今まで工業の分野においてのみ使用されていたロボットが、サービス業のようなロボットの使用を専門としない人間も使用するような環境でも使用されるようになってきた [1]。しかし、ロボットの家庭への進出、つまりロボット

との共生の実現は様々な理由で難しい。

その理由の1つとして、人間とロボットの協調の難しさが挙げられる。人間はお互いを助け合い、協調をすることによって社会を構築し、生活を営んでいる。ロボットがその人間と人間の協調の中にとけ込み、社会を構築する一部となるためには、人間との協調を行えるようになることが必要である。この協調とは、相手の意図を理解し、その意図に合わせて最適と考えられる行動を選択した結果生まれる行動であり、そのためには、ロボットが人間の意図を理解する必要がある。しかし、現状のロボットでは、人間が一方向的に機械の操作を覚え、人間の側だけがロボットに合わせて行動を行っている。それに対し、HCI(Human Computer Interaction)の分野では、ロボットに対し、人

<sup>1</sup> 電気通信大学 大学院情報理工学研究科  
Graduate School of Informatics and Engineering - The University of Electro-Communications, Chofu, Tokyo 182-8585, Japan

a) y.kato@uec.ac.jp

b) kuri@acm.org

間が入力できる情報を増やし、ロボットが人間に対して合わせることでできる分野を増やす研究が行われてきた [2]。しかし、人間の意図は複雑であり、言葉、表情、仕草等、様々なチャネルを用いて表現されており、それを元に他の人間は意図の理解を行っている。この複雑な意図を身体性の異なるロボットが理解するためには、ロボットが自発的に人間の意図を理解する必要がある。そこで、ロボットが人間の意図を理解することができるのかどうかを調べる必要があるが、人間とロボットでのやり取りの観察を通して意図の理解を観測することは現実的でない。

そこで、本研究では、人間を想定した既にある意図を持って決まった動きを行うエージェントとロボットを想定した何も学んでいないエージェントにタスクを課し、前者のエージェントの意図を後者のエージェントが理解できるかの検証を行う。

## 2. 協調

協調という言葉は様々な現象を対象として使用されており、一口に協調と言っても様々な意味を持つため、各研究分野で用いられる協調という言葉の意味を確認した後、本研究で用いる定義の説明を行う。

### 2.1 国語辞典による定義

協調という言葉はデジタル大辞林 [3] では、『名 ](スル)互いに協力し合うこと。特に、利害や立場などの異なるものどうしが協力し合うこと。「労資が協調する」「協調性』と説明されている。国語辞典では、『協力し合う』という行動の中に、『利害や立場などのことなるものどうし』という制限を追加している。

### 2.2 ロボティクス分野における協調

油田 [4] は自律移動ロボット同士の間で協調という行動を定義する上で、ロボット同士の独立性の定義を行うことが必要であると説明し、独立性のレベルを3種類定めた(表1, 図2.2)。その上で、協調の種類を積極的協調と消極的協調の2種類に分けた。消極的協調はロボット同士がぶつからないように動くような単体で行動するときよりも負の作用を起こさないことを目的に行われ、積極的協調は他のエージェントが動きやすいように環境の整備を行う等の単体で行動するときよりも正の作用を起こすことを目的として行われる。また、協調において使用される通信をインプリシット通信とエキスプリシット通信の2つに分類した。インプリシット通信は人間の行動における「阿吽の呼吸」のような明確なシンボルのやり取りが行われない通信を指す。人間が人混みでぶつからないように歩くことができるのは、相手の動きを観察し、相手の次の動作をその動きから次の動作を予測し、行動を行っているからである。このとき、相手の動きが通信の役割を行っている。また、

エキスプリシット通信はインプリシットにおける相手の動きの観察を予測して自分の行動を示すシンボルを見せることによる通信である。

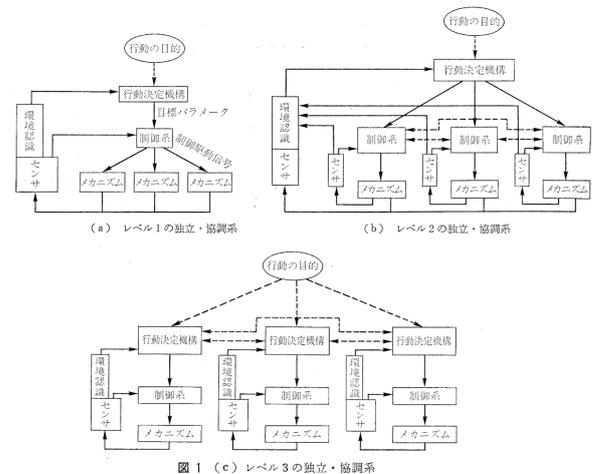


図1 油田による協調の分類

### 2.3 本研究における協調

本研究では基本的にロボティクスにおける定義に従う。また、国語辞典における意味は『利害や立場』という言葉の解釈次第になるが、マルコフ決定過程において、異なる行動価値関数を持っているエージェントという表現に置き換えられると考える。そこで、本研究における協調の定義として油田によるレベル3かつ積極的協調の定義を協調と呼ぶ。協調に至るまでには少なくとも3ステップがあり、本研究ではステップ1の再現を目指す。

- (1) 他のエージェントを含む環境の次の状態の予測ができる
- (2) 自らの行動による他のエージェントの行動の変化が予測できる
- (3) 相互的に行動を変化させ、最適化する

## 3. 関連研究

### 3.1 協調の研究

エージェントを用いた協調の研究はマルチエージェント最適化の分野で行われている。

Joelら [5] は性能の異なる2種類のエージェントを使用し、マルチエージェント環境においてエージェント同士がどのような条件で協調を行ったり、競争したりするのかについて実験を行った。実験は Gathering と呼ばれるリングを取り合うタスクと Wolfpack と呼ばれるターゲットを追い込んで捕まえるタスクにおいて行った。Wolfpack と呼ばれるタスクは追跡問題 [6] とも呼ばれており、マルチエージェントによる協調の創発を観察する際に用いられる [7]。その結果、エージェントは Gathering において競争を行い、Wolfpack において協調を行った。また、タスクが複雑で

表 1 油田による協調の分類

レベル	独立性	制御機構
レベル 1	独立した機械的自由度をもつ .	単一のコントローラおよびコントロールプログラムがすべての自由度を制御する . = 集中制御 =
レベル 2	各々の機械 (自由度) が別個のコントローラによって制御される .	集中した意思決定部にセンサ情報や環境認識の結果を集中し, 個々の機械の動作を決定した後, 分散的に制御する . = 集中・分散制御 =
レベル 3	個々の機械が独立したセンサや外界認識機構とコントローラをもち, 自律的に行動を決定して動作する .	分散 (自律的) 行動決定と分散制御 . = 自律制御 =

あるほど, エージェントが協調を行うことを示した. 本研究では, Joel らの特定した協調を行うタスクを元にタスクの設定を行う .

### 3.2 意図理解の研究

意図理解の研究は記号創発の分野において記号と意図のヒモ付を行う記号接地と呼ばれる分野において行われている. Steels ら [8] は意図する概念に合った記号を確率的に選択するテーブルを持つエージェントを複数用いて, 意図理解の実験を行った. その結果, 複数のエージェントにおいて, 概念-記号のテーブルが統一され, それぞれの意図理解が行われることが観察された. また, Steels らによるその後の研究 [9][10][11] では, 2 体のロボットを用いて画像, 音声を用いての概念-記号テーブルの完成が観察された. 本研究では, Steels らの意図理解の再現のために使用したテーブルを用いた実験を元に 1 対 1 の関係における意図理解に焦点をあてた実験を行う .

## 4. 提案実験

本研究では意図理解の行われた上での協調の実現を目指す. そのため, 実験 1 で意図理解のモデルの検証を行い, その上で実験 2 によって創発された協調における意図理解の観察を行う. ロボットと人間での協調の実現を目指した実験であるため, 2 体のエージェントをロボットに見立てた観察対象である「対象エージェント」と人間に見立てた元々プログラムされた動きを行う「環境エージェント」と名付けて実験を行う .

### 4.1 実験 1-意図理解単体の再現-

実験 1 では, 「対象エージェントが意図を理解できたら, それに合わせた行動を行うことが必ずできる」という仮定で行った. 環境エージェントに上下左右の 4 種類の行動に対して, 4 種類の色を 1 対 1 の対応で定めるテーブルを持たせ, 上下左右, 色はそれぞれ, 0,1,2,3 と番号で保持した. そして, その中からランダムに行動を選択し, それに対応する色を対象エージェントの入力とした. 対象エージェント

表 2 実験 1 における色と場所のペア

色テーブル	0	1	2	3
行動テーブル	1	3	0	2

表 3 実験 1 における対象エージェントの確率テーブル

色	0				1			
場所	0	1	2	3	0	1	2	3
割合	314	1000	10	15	10	13	17	970
色	2				3			
場所	0	1	2	3	0	1	2	3
割合	936	11	10	13	10	11	966	10

は色に対する行動選択のテーブルを持ち, 行動選択のテーブルより, 確率的に行動を行うよう設定した. 各テーブルの初期値を 100 にした対象エージェントと対象エージェントを用いて環境エージェントの意図を対象エージェントが当てられるか確認し, 当てられたらそのテーブルの確率を+1, 当てられなかったらそのままという試行をテーブルのどれかの値が 1000 を超えるまで繰り返した. 以上の操作を 6 回繰り返し, 結果として対象エージェントのテーブルを取り出して確認し, 環境エージェントの意図である移動先を対象エージェントが色から判断できているかを評価した .

### 4.2 結果 1

結果として取り出された対象エージェントのテーブルのうち 1 回目のものを表 2, 表 3 に示す. 環境エージェントとしてあらかじめ指定した位置-色のペアを指すテーブルの位置の数字が大きいことが観察された. 他の 5 回も同様の結果を得られた .

### 4.3 実験 2-意図理解から協調行動の観察-

実験 2 は, 対象エージェントが意図理解をし, 協調を行う過程の再現を目的として行った. タスクとして, 追跡問題をモデルとした環境エージェントと対象エージェントによってターゲットを挟み込むことによって報酬を得られる問題を用意した (図 4.3). このタスクは環境エージェントが対象エージェントの意図理解を行っていないとクリア

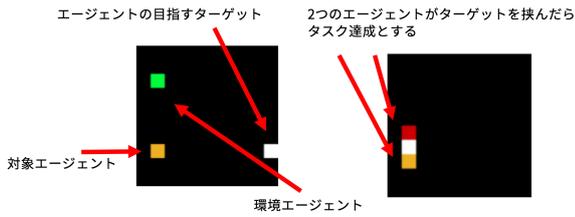


図 2 実験 2 におけるタスク

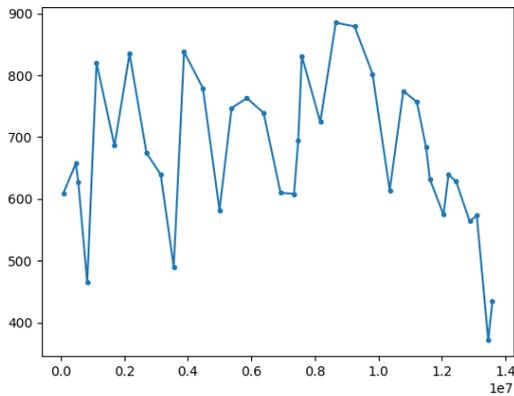


図 3 実験 2 におけるエージェントのかかったターン数の平均

できないタスクとして用意した．環境エージェントには実験 1 と同様に場所と色のテーブルを持たせ，ターゲットに対してランダムに選択された位置へと移動するようにプログラムを行った．また，対象エージェントには，Actor-Critic を使用した学習器を持たせ，学習器によりチューニングされた行動選択器を用いて行動選択を行わせた．1000 ターン以内で報酬を得られるまで繰り返した．1000 ターン以内に報酬を得られなかったら報酬を 0 とし，環境のリセットを行った．以上の環境をセットとし，20 セットの環境を用いて並列で学習を行った．また，学習同時にテストを行った．タスクの完了までにかかった期間の長さおよび対象エージェントの行動の観察を行い，評価を行った．

#### 4.4 結果 2

13000000 エピソード完了時におけるエージェントがタスク完了までにかかった時間を図 4.4 に示す．タスク完了までにかかった時間は同じモデルを使用して 100 回の試行を行い，その平均を取っている．また，タスクの完了を達成した際の対象エージェントのラスト 8 回分の行動を図 4.4 に示す．

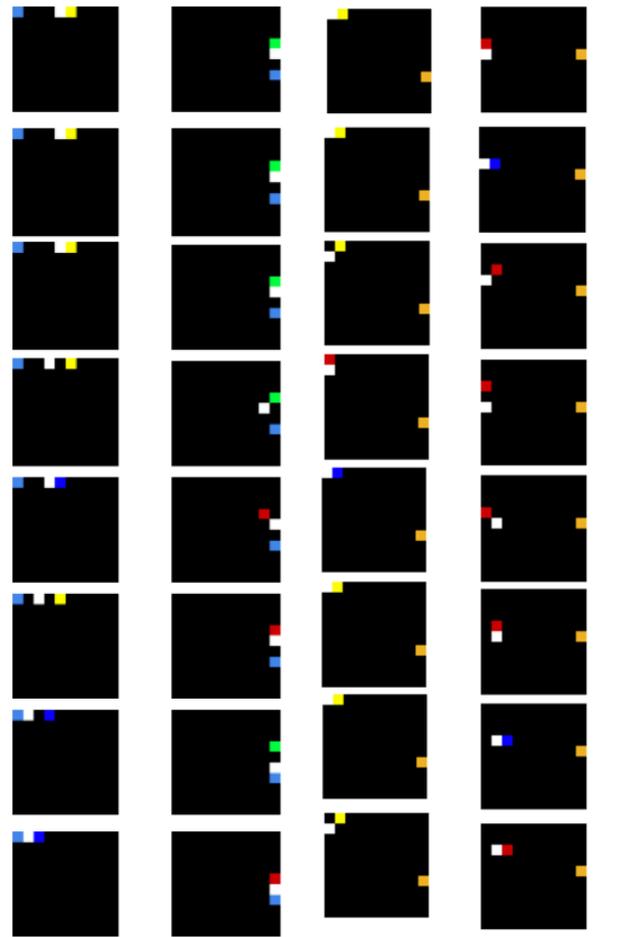


図 4 左 2 つ：タスクを達成した際のエージェントのラスト 8 ターンの動き，右 2 つ：タスクを達成できなかった際のエージェントのラスト 8 ターンの動き

## 5. 考察

実験 1 の結果より，報酬を意図の予測と実際の意図との合致に与え，試行を繰り返すことにより，対象エージェントによって環境エージェントの意図の理解を実現できるということが分かった．また，実験 2 は対象エージェントが環境エージェントの意図（移動先）を色から理解し，それに協調して動くまでを想定した実装であったが，図に示されるとおり，タスクの完了を達成した際も対象エージェントは動いておらず，対象エージェントが協調行動を獲得することは無かった．原因としては以下の 2 点が挙げられる．

- (1) ターゲットがランダムに動いている点
- (2) 環境エージェントの動き方の問題

1 つ目はターゲットのランダム運動によって対象エージェントへの入力がランダムに変動してしまっている点である．対象エージェントに学習をさせたいのは環境エージェントが色を変化させながらターゲットを追い込む行動であるため，ランダム運動を行うことは学習においてノイズを混ぜていることになってしまっている．改善策としては

ターゲットの行動をプログラムされた逃避行動にすることが挙げられる。

2つ目は環境エージェントの動き方の問題である。環境エージェントはターゲットに張り付くように行動を行っている。ターゲットは環境エージェントが居る方向には進まないため、必然的に環境エージェントと反対方向に進んでいき、最終的には壁まで追い詰められる。すると、対象エージェントが挟み込むために反対側に入れなくなってしまいうため、タスクの達成が妨害されてしまう。改善策としては、フィールドをトラス構造にすることが挙げられる。

## 6. まとめ

本研究では、ロボットが社会進出した際の人間との協調行動の学習を目的として、未知のエージェントの意図を理解できるか実験を行った。実験は2段階に分け、1段階目では意図理解を対象として色から意図である移動先の理解を獲得した。2段階目では意図理解した上での協調行動の獲得を目指して、WolfPack ゲームをベースとした観察対象となる対象エージェントとあらかじめプログラムされた環境エージェントがターゲットを両側から挟むと成功というタスクを用いて実験を行った。その結果、環境エージェントのプログラムされた行動、ターゲットが行うランダム行動によって対象エージェントの学習が阻害されていることが分かった。

## 参考文献

- [1] 総務省: 情報通信白書. 入手先 <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h28/html/na000000.html> (参照 2017-1-14).
- [2] 石黒浩: HRI と HRL, 認知科学, Vol. 19, No. 3, pp. 269-275 (2012).
- [3] 小学館: デジタル大辞泉. 入手先 <https://kotobank.jp/word/%E5%8D%94%E8%AA%BF-478664> (参照 2017-1-27).
- [4] 油田信一: 複数の自律移動ロボットの協調行動, 日本ロボット学会誌, Vol. 10, No. 4, pp. 433-438.
- [5] Leibo, J. Z., Zambaldi, V., Lanctot, M., Marecki, J. and Graepel, T.: Multi-agent Reinforcement Learning in Sequential Social Dilemmas, *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, International Foundation for Autonomous Agents and Multiagent Systems, pp. 464-473 (2017).
- [6] BENDA, M.: On Optimal Cooperation of Knowledge Sources, *Technical Report*, (online), available from (<https://ci.nii.ac.jp/naid/10006535376/>) (1985).
- [7] 荒井幸代, 宮崎和光, 小林重信ほか: マルチエージェント強化学習の方法論: Q-Learning と Profit Sharing による接近, 人工知能学会誌, Vol. 13, No. 4, pp. 609-618 (1998).
- [8] Steels, L.: A self-organizing spatial vocabulary, *Artificial life*, Vol. 2, No. 3, pp. 319-332 (1995).
- [9] Steels, L.: The synthetic modeling of language origins, *Evolution of communication*, Vol. 1, No. 1, pp. 1-34 (1997).
- [10] Steels, L.: Grounding symbols through evolutionary

language games, *Simulating the evolution of language*, Springer, pp. 211-226 (2002).

- [11] Steels, L., Kaplan, F., McIntyre, A. and Van Looveren, J.: Crucial factors in the origins of word-meaning, *The transition to language*, Vol. 2, No. 1, pp. 4-2 (2002).