

# 他者意図の推定に基づく協調行動の計算モデル化 ～ 他者の推定レベルに合わせた行動決定方式の選択～

Computational Modeling of Cooperative Behavior Based on Estimation of Other's Intention

長田悠吾\*1 石川 悟\*2 大森隆司\*3 森川幸治\*4  
Yugo Nagata Satoru Ishikawa Takashi Omori Koji Morikawa

\*1北海道大学大学院情報科学研究科

\*2北星学園

Graduate School of Information Science and Technology, Hokkaido University

Hokusei Gakuen University

\*3玉川学園学術研究所

\*4松下電器産業株式会社 先端技術研究所 知能情報技術研究所

Tamagawa University Research Institute

Advanced Technology Research Laboratory, Matsushita Electric Industrial Co.,Ltd.

We propose a computational model of cooperative behavior based on an estimation of other's goal. In the model, we define levels of estimation: Level 0 means no-estimation, Level 1 means estimation of the other's goal, and Level 2 means estimation of "the other's estimation about self". A computer simulation in which two hunters chased preys in cooperation showed that the proposed model can cooperate smoothly with the other that has various levels of estimation by controlling its own level of estimation.

## 1. はじめに

近年、ペットロボットや掃除ロボットなど、家庭での実用を目的としたロボットが開発されている。また、擬人化されたエージェントによってユーザを補助するユーザインタフェースも提案されている。これらのロボット及びエージェント（以下、両者をまとめてエージェントと呼ぶ）は、ユーザである人間とのインタラクションを通してサービスを提供する。

このインタラクションにおいて、エージェントには、ユーザの現在の意図、即ち、目標を推定し、この目標に合わせた協調的な振る舞いが求められる[小林 03]。例えば、掃除ロボットがユーザの邪魔にならない様に掃除を行うためには、ユーザの進行方向を予測して衝突を避けたり、ユーザが作業している机の下の掃除は後回しにするといった判断が必要である。

その一方で、ユーザ自身もエージェントの目標を推定し、その振る舞いを予測することも想定される。例えば、掃除ロボットが机の方向へ移動しているときには、ユーザは「掃除ロボットは机の下を掃除しようとしている」と推定することができる。

このように、対象に目標などの心的状態を帰属させる態度を Denett は志向姿勢 [Dennett 96] と呼んでおり、人間は非人間・非生物の対象に対しても、志向姿勢をとってしまうと言われている。外見が人間に近いヒューマノイドや擬人化エージェントに対しては、特に志向姿勢を取り易いことが予想される。

さらに人間は他者との関わりの中で、相手の推定の内容を推定するといった「深さ」を伴った推定を行う傾向も持っている [Ohtsubo 06]。例えば、掃除ロボットに対し「自分が机に座っているときには、掃除ロボットは自分が作業中であることを推定し、机の掃除には来ないだろう」といった推定である。本研究では、このような推定の「深さ」を推定のレベルと呼ぶ（図 1）。

この掃除ロボットのように、ユーザとのインタラクションの中で円滑に作業を進める為には、エージェントは様々なレベルで推定を行うユーザに対応しなければならぬが、その一般的な方法は現時点では知られていない。そこで本研究では、相手の

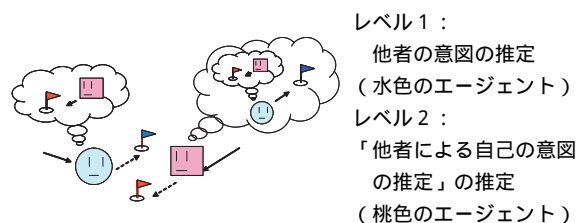


図 1: 他者意図の推定のレベル

目標を様々なレベルで推定し合う状況で、協調的タスクを効率的に解決する行動決定過程の計算モデル化を目的とする。

## 2. 他者の目標の推定手法

本研究では、他者の目標を推定するために、他者の行動を予測できる他者モデルを用いる。他者モデルは、ある目標  $G$  を達成するために他者が状態  $s$  で行動  $a$  をとる確率  $Pr(a|s, G)$  の形で表される。

過去  $T$  時間分の他者の行動履歴

$$m_o(t) = \{(s_o(t), a_o(t)), \dots, (s_o(t-T+1), a_o(t-T+1))\}$$

が与えられたとき、その行動の他者の目標に対する尤度

$$L(G) = \sum_{(s,a) \in m_o(t)} Pr(a|s, G)$$

を計算し、尤度の高い目標  $G$  を他者の目標の推定値とする。この推定値の決定は、Boltzman 選択法を用いて確率的に行われる。

本研究では、他者モデル  $Pr(a|s, G)$  は、自己自身の行動決定方策と同じものを用いた [長田 05]。自己の行動決定方策は強化学習 [Sutton 98] を用いて学習的に獲得させた。

## 3. 推定のレベル

### 3.1 定義

ここで、推定のレベルを定義する（図 1）。

連絡先: 長田悠吾, 北海道大学大学院情報科学研究科, 〒060-0814 札幌市北区北 14 条西 9 丁目, TEL 011-706-6815, FAX 011-706-7831, yugo@complex.eng.hokudaia.ac.jp

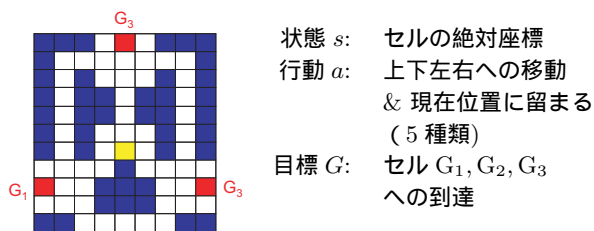


図 2: 予備的シミュレーションのタスク環境

まず、他者の過去の行動履歴  $m_o$  からの「他者目標の推定」をレベル 1 (Lv.1) の推定と定義する。そして、レベル 1 の推定に基づいて自己の目標を決定する行動決定戦略をレベル 1 の戦略と呼ぶ。

次に、自己自身の過去の行動履歴  $m_s$  からの、レベル 1 の推定を行う他者エージェントの推定内容の推定、すなわち、「自己の目標が他者からはどのように推定されるか」の推定をレベル 2 (Lv.2) の推定と定義する。そして、レベル 2 の推定に基づいて自己の目標を決定する行動決定戦略をレベル 2 の戦略と呼ぶ。

最後に、他者の目標を推定せず、他者の意図に応じた自己目標の変更をしない行動戦略を便宜上レベル 0 の戦略と呼ぶ事にする。

### 3.2 予備的シミュレーション

以上の 3 つの戦略の特性を調べる為に、予備的な計算シミュレーションを行った。シミュレーションのタスク環境は、図 2 に示す単純な迷路である。各エージェントが設定できる目標は、環境中のゴール ( $G_1, G_2, G_3$  のセル) の何れかへに到達の 3 種類である。この環境中で、2 体のエージェントが互いに異なるゴールに到達するとタスク解決となる。従って、このタスク解決には、他者が達成するであろう目標の予測しその目標とは異なるゴールを自己の目標に選ぶという協調的振る舞いが必要になる。

レベル 0, レベル 1, レベル 2 の全ての戦略の組合せについてシミュレーションを行った結果、このタスクの効率的な解決には 2 体のエージェントの内、少なくともどちらか一方がレベル 1 の戦略をとる必要があることが分かった [長田 05]。このことは、2 体のエージェントによる協調的タスクの解決にはレベル 1 の推定、すなわち他者の目標の推定が有効であることを意味している。ただし、タスクを解決できる組合せのうちで、2 体のエージェントが両方ともレベル 1 の戦略をとる場合には、他の組み合わせに比べて所要ステップ数が多い、即ちパフォーマンスのロスが大きい事がわかった (図 3)。

### 3.3 レベル調整の必要性

レベル 1 の戦略同士でのパフォーマンスは以下の理由で低下する。レベル 1 の戦略は、レベル 0 の戦略に従うエージェントの目標を推定することで自己の目標を決定する。また、レベル 2 の戦略は、レベル 1 の推定を行うエージェントの推定内容を推定することで自己の目標を決定する。従って、組み合わせられた戦略のレベルの差が 1 の場合には、レベルが高い方の戦略がレベルが低い方の戦略の行動決定を予測することができる。これに対し、レベル 1 の戦略同士では、両エージェントが互いに、相手が目標を変更しない (レベル 0 の戦略に従う) ことを想定して、自己の目標を変更してしまう。

他者目標の推定に基づく協調的タスク解決において、このようなパフォーマンスロスを避けるには、状況に応じて自己の戦略のレベルを調整することで、自己と他者の戦略のレベルの

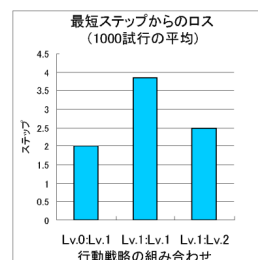


図 3: 予備的シミュレーションの結果：最短ステップ数と比較して余計にかかったステップ数

差を 1 にすることが有効であると本研究では考える。次節からは、レベル調整のモデルについて説明する。

## 4. レベル調整のモデル

予備的シミュレーションで、タスクの協調的解決には他者目標の推定 (レベル 1 の推定) が有効であった事をふまえ、レベル 1 の戦略を、状況に応じて戦略のレベルを調整できるモデルへと拡張する。

### 4.1 確信度に基づくレベル 0 の戦略の選択

レベル 1 の戦略同士でパフォーマンスロスが発生する原因の 1 つは、レベル 1 の戦略が自己の目標の変更を頻繁に行う事にある。各戦略の尤度差が十分大きくなる前に、他者の目標の推定値を決定すると、他者目標の推定の誤り、及び推定のし直しが頻繁に発生してしまう。これに伴い、自己の目標の変更も頻繁に起こる。自己目標の頻繁な変更は、今度は他者による自己目標推定の誤り、及び推定のし直しが頻繁に発生する原因になる。以上の悪循環がパフォーマンスロスを引き起こす。

この悪循環を回避するため、本モデルでは他者目標推定の「確信度」を導入する。確信度  $C$  は、他者目標の尤度差

$$C = L(G_{1st.}) - L(G_{2nd.})$$

で計算される。ここで、 $G_{1st.}$  は最も尤度が高い目標、 $G_{2nd.}$  は 2 番目に尤度が高い目標である。確信度  $C$  が閾値  $\theta$  より小さい場合には、他者目標は確信的ではないとして他者目標の決定を行わず、自己の目標を変更しない。すなわち、レベル 0 の戦略に従って行動する。

これにより、自己自身の誤った目標決定を回避できる。また、自己の目標の頻繁な変更を減らす事で他者からも目標を推定されやすくなる事が期待される。

### 4.2 尤度に基づくレベル 1 の戦略とレベル 2 の戦略の選択

レベル 1 の戦略同士でパフォーマンスロスが発生する原因のもう 1 つは、両エージェントが同時に自己の目標を変更してしまう事にある。確信度を導入しても、両エージェントの確信度がほぼ同時に高くなったときには、両エージェントは同時に目標を変更してしまい、お互いに間違った行動決定をしてしまい、パフォーマンスロスを引き起こす。

目標の同時変更を回避するには、「他者目標の推定に基づき自己が目標を決定する」(レベル 1 の戦略)と、「他者が目標を変更するという予測に基づいて目標を決定する」(レベル 2 の戦略)との、どちらかを適切に選択する必要がある。本モデルでは以下の基準でそれを選択する。

レベル 1 の戦略では、自己目標は他者目標に基づいて決定される。従って、レベル 1 の戦略は、現在の他者目標の尤度が

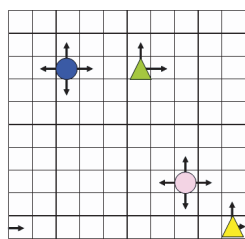


図 4: ハンター課題の環境  
(実際のサイズは 20 × 20)

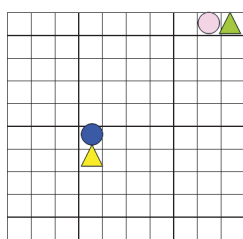


図 5: ハンター課題の解決場面の一例

高いときに有効である。また、レベル 2 の戦略では、他者が自己目標に基づき目標を決定することを想定している。従って、レベル 2 の戦略は自己目標の尤度が高いときに有効であると見込まれる。以上の理由から、本モデルでは戦略の選択の基準として尤度を採用し、以下のように自己のレベルを決定する。

1. 他者の目標についての確信度  $C$  を計算する。 $C < \theta$  なら自己はレベル 0 の戦略に従う。そうでないのなら 2 へ。
2. 他者目標を決定し、自己が目標を変更する必要があるかどうかを調べる。必要なければレベル 0 の戦略に従う。必要があれば 3 へ。
3. レベル  $i$  の戦略を確率的に選択する。本実装では以下の確率分布を使用した。

$$Pr(i) = \frac{\alpha_i \exp(\beta' L_i)}{\sum_{j=1,2} \alpha_j \exp(\beta' L_j)} \quad (i = 1, 2).$$

ここで、 $L_1$  は他者目標の尤度、 $L_2$  は自己目標の尤度である。また、 $\beta' \in [0, 1]$  は逆温度パラメータ、 $\alpha_i \in [0, 1]$  は重み ( $\alpha_1 + \alpha_2 = 1$ ) である。

## 5. シミュレーション実験

### 5.1 設定

提案モデルの効果を評価するために、シミュレーション実験を行った。タスクは、2 体のエージェントが協力して獲物を捕まえるハンター課題である。20 × 20 のグリッド状のトラス空間に、ハンター 2 体と区別可能な獲物 2 体が配置される (図 4)。2 体のハンター (図中 ) がそれぞれ互いに異なる獲物 (図中 ) に接触するとタスク解決となる (図 5)。ハンターは上下左右への移動と停止の 5 種類の行動を選択できる。獲物は上 1/5、右と停止それぞれ 2/5 の確率で行動する。このタスクにおける目標は、どちらの獲物に接触するか 2 種類である。このタスクの円滑な解決には、エージェントは、他者が接触するであろう獲物とは別の獲物への接触を目標に選ぶ必要がある。

### 5.2 結果

2 体のエージェントの戦略の組合せに対応した、タスク解決までの平均ステップ数の 100 試行の平均を図 6 に示す。円の大きさは平均ステップ数に対応している。×印は制限ステップ以内にタスクを解決できなかった事を示している。図 6 の中で、Model1 は確信度に基づく他者目標推定の抑制は行いが、プランの選択は行わない (確信度  $C \geq \theta$  の時は常にレベル 1 の戦略を選ぶ) 戦略であり、Model2 は提案モデルの内、確信度に基づく他者目標推定の抑制とプランの選択の両方を行う戦略である。なお、パラメータは  $T = 10$ ,  $\theta = 5$  (Model1),  $\theta = 1.5$  (Model2),  $\beta' = 1.0$ ,  $\alpha_1 = 0.6$ ,  $\alpha_2 = 0.4$  を用いた。

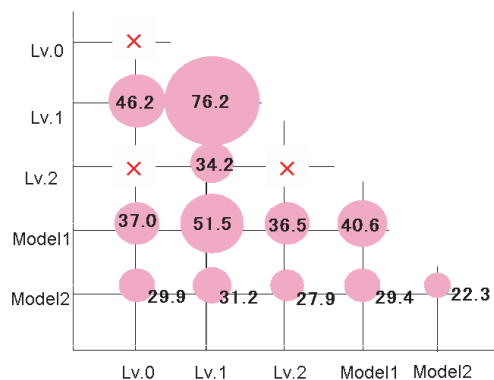


図 6: 計算機シミュレーションの結果：タスク解決までの平均ステップ数

図 6 から、このタスクにおいてもレベル 1 の戦略同士ではパフォーマンスロスが発生することがわかる。このパフォーマンスロスは Model1 では軽減されている。これは、確信度に基づいた他者目標推定の抑制が有効であったことを意味している。しかし、Model1 でも、レベル 1 の戦略及び Model1 が相手のときは長いステップ数を必要としている。これに対し、Model2 では全ての戦略に対し、少ないステップ数でタスクを解決できていることがわかる。

## 6. まとめ

本稿では、相手の目標を相互に推定し合い協調的にタスクを解決する場面において、自己自身の推定のレベルを制御することで、様々なレベルで推定を行う他者とも効率的にタスクを解決する行動決定方式をモデル化し、計算機シミュレーションによる評価を行った。

今後の課題の 1 つは、より複雑かつ現実的な協調タスク場面における提案モデルの検証である。また、今回は、エージェント - エージェントの協調問題においてモデルを検証したが、人間とエージェントとの効率的な協調作業場面へのモデルの拡張とその検証がもう 1 つの課題である。

## 参考文献

- [Dennett 96] Dennett, Daniel Clement. "Kinds of minds : toward an understanding of consciousness," New York, N.Y. : Basic Books, 1996. (邦訳：ダニエル・C・デネット．心はどこにあるのか．土屋俊訳．草思社．1997.)
- [小林 03] 小林一樹, 山田誠二: 人間と小型移動ロボットによる協調的な掃除, 第 21 回ロボット学会学術講演会 (2003).
- [長田 05] 長田悠吾, 石川悟, 大森隆司, 森川幸治: 他者推定に基づく協調行動過程のモデル化, 日本神経回路学会第 15 回全国大会講演論文集, 133-134, (2005)
- [Ohtsubo 06] Ohtsubo, Y., Rapoport, A : Depth of reasoning in strategic form games. Journal of Socio-Economics, 35(1), 31-47, (2006)
- [Sutton 98] Sutton and A. G. Barto : Introduction to Reinforcement Learning, MIT Press (1998), (邦訳：強化学習, 三上貞芳, 皆川雅章 訳, 森北出版, (2000)).