認知的インタラクションフレームワークに基づいた他者モデルの提案

Other's Mind Model Based on Cognitive Interaction Framework

大澤 正彦 [1, 2*] 奥岡 耕平 [1,3] 坂本孝丈 [4] 市川淳 [5] 今井倫太 [1]
Masahiko Osawa[1,2] Kohei Okuoka[1,3] Takafumi Sakamoto[4]
Jun Ichikawa[5] Michita Imai[1]

[1] 慶應義塾大学
[1] Keio University
[2] 日本学術振興会特別研究員（DC1）
[2] Japan Society for the Promotion of Science, Research Fellow (DC1)
[3] 株式会社 BLUEM
[3] BLUEM Inc.
[4] 静岡大学
[4] Shizuoka University
[5] 神奈川大学
[5] Kanagawa University

Abstract: 他者モデルとは，コミュニケーション相手である他者の心的状態のモデルであり，円滑なコミュニケーションにおいて重要な役割を担っている．本研究では，3 つの段階に分けた認知的インタラクションフレームワーク (CIF) に基づいた他者モデルを提案する．提案する他者モデルは，CIF の段階に対応する生物の行動原理に関して説明性がある．さらに著者らの一部による子どもを対象とした研究成果と照らし合わせ，定性的な分析を試みる．

1 はじめに

Human-Agent Interaction (HAI) 研究において，他者モデルと呼ばれる他者の心的状態や行動を予測するモデルは中心的な研究題材の 1 つである．他者モデルに関する研究は多岐にわたる．工学的観点からは，エージェントに他者モデルを実装することが試みられている [1, 2, 3]．心理学的観点からは，人がエージェントに対して想定している他者モデルについて頻繁に調べられており，どのようなエージェントに他者モデルを想定しやすいか [4, 5]，他者モデルを想定されたエージェントとのインタラクションがどのようなものになるか [6, 7]，といった論点となっている．

他者モデル研究の目指す将来的な 1 つのゴールは，ヒトの持つ他者モデルについてその情報処理メカニズムを明らかにし，工学的にエージェントに実装した上で，実際に人とのインタラクションにおいて応用することといえる．工学研究においては，他者モデルのアーキテクチャが提案され，そのダイナミクスについても頻繁に扱われる．しかしながら工学研究のほとんどはシミュレーションのタスクに閉じている．心理学的アプローチによって工学的に提案されたモデルを検証することも考えうるが，心理学的研究では定量性や客観性が担保された研究方法論の確立が求められているため，困難である．具体的には，インタラクション研究の心理学的な分析は，多くの場合図 1 に示すように，インタラクションが統制され，一部が切り取られた形での評価になる．しかしながら，他者モデルにおいて重要な時間的随伴性や相互適応のような時間的な要素は，このような研究方法論において扱いが難しいためしばしば軽視される [8]．

本研究では，従来の帰納的な科学的検証方法論に必ずしもよらない方法で，妥当な他者モデルを追求する方法として，演繹的な方法で他者モデルを構築する方法論を模索する．もし，演繹的な方法で妥当なモデル

*連絡先：慶應義塾大学大学院理工学研究科
〒 223-8522 神奈川県横浜市港北区日吉 3-14-1
E-mail: mosawa@ailab.ics.keio.ac.jp

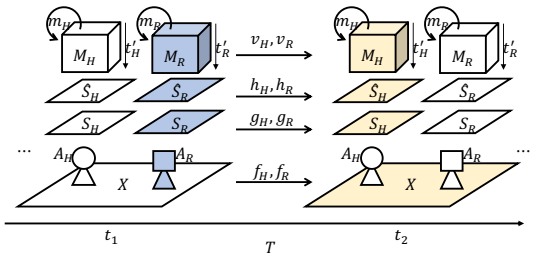

図 1: これまでの HAI 研究

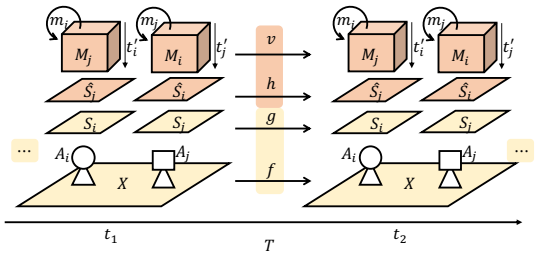

図 2: フレームワークが確立した後の HAI 研究

が提案できれば，すでに確立した科学的検証方法が適用できない場合でも，その妥当性を認められる可能性がある．具体的には，進化や発達の観点から，生物が実際に獲得しうる情報処理の仕組みについて考えていく．本来であれば進化と発達は分けて議論されるべきであるが，単純な情報処理機構が成熟し，その機構がより高度な情報処理に転用されていくという原則として抽象的に本論文では捉える．

提案する他者モデルは，著者らの一部が提案したインタラクションのフレームワーク [9] に基づく．同フレームワークを本論文では認知的インタラクションフレームワーク (CIF) と呼ぶ．CIF は進化的な背景から 3 段階に分かれており，提案する他者モデルは各 CIF の段階に対応する 3 つの段階がある．

また，構築した他者モデルを元に幼稚園児のリトミックを題材とした既存研究での知見と照らし合わせ，定性的な分析を試みる．

本研究が発展していく先に見据える目標は，CIF や他者モデルを整備することで，図 2 に示すように，インタラクション全体とその時間変化を捉えることのできる HAI 研究方法論を確立することにある．

2 背景

2.1 認知的インタラクションデザイン学（CID)

平成 26〜30 年度文部科学省科学研究補助金新学術領域研究（研究領域提案型）採択課題「認知的インタラクションデザイン学〜意思疎通のモデル論的理解と人工物設計への応用」[8] では, 他者の心的状態のモデルである他者モデルの理解やアルゴリズムの実装が目指された．同課題において特筆すべき点は成人, 幼児, 複数種の動物, 人工物といった多様なインタラクション研究に携わる研究者が集まり，多角的に他者モデルの原理について議論が行われていたことである．

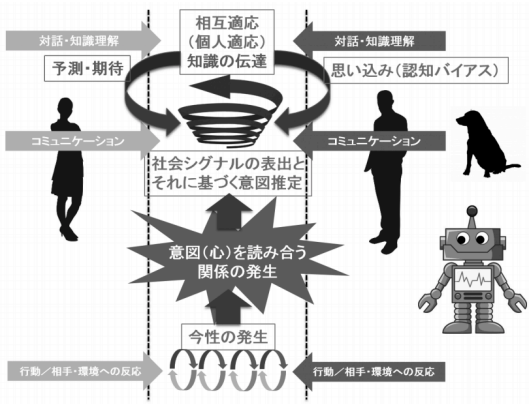

図 3: 認知的インタラクションデザイン学にて提案されたインタラクションモデル [8]

当該領域では，図 3 に示すインタラクションモデルを提案している．本研究は，CID の成果ともいえる図 3 に示したインタラクションモデルを，根本的なインタラクションのフレームワークから見直すことで，直接的な工学的有用性や心理実験における仮説となる情報処理モデルとしての役割を担えるものへと発展させようとするものである．

2.2 志向姿勢

哲学者 Dennett は，人間が対象の振る舞いを理解し，予測するために 3 つのスタンスがあると述べている [10, 11]．1 つめが物理スタンスであり，他者の行動を物理法則に帰属させて，他者を予測するスタンスである．2 つめが設計スタンスであり，他者の行動を設計に帰属させて，他者を予測するスタンスである．3 つめが意図スタンスであり，他者の行動を意図に帰属させて，他者を予測するスタンスである．いずれかのスタンスで他者とのインタラクションを捉えられれば，その他者を予測可能なインタラクティブな存在として認めやすい．

エージェントやロボットにおいては，設計スタンスか意図スタンスで捉えられることがほとんどである．例えば，ユーザが「おはよう」と発話したことに対するエージェントの返答を予測する際に,「おはようと言われたら，おはようと返すプログラムが入っているはずだ」と考えて「おはよう」という返答を予測すれば設計スタンスである．一方「おはようと言われたら，挨

拶を返そうとするだろう」と考えて「おはよう」という返答を予測すれば意図スタンスとなる．

本論文では，3 つのスタンスによる対象の予測モデルを下記のように表す．

物理スタンス $M_{physics}$

設計スタンス $M_{mechanism}$

意図スタンス $M_{intention}$

人は意図スタンスに感じる他者に対して，心を想定し擬人化する傾向があると言われている．そこで本論文では，対象を意図スタンスで捉えているときの予測モデル $M_{intention}$ を他者モデルと呼ぶ．

2.3 他者モデルのレベル

他者モデルに関する既存研究 [1, 2] では，しばしば推定する他者モデルの深度に応じてレベルが設定されている．本論文でも，他者モデルに関する議論のためにレベルの概念を下記のように導入する．

- レベル 0　行動主体が対象の行動を推定せず自己の意図のみに従って行動を決定する
- レベル $n(n \geq 1)$　対象をレベル $n-1$ の存在と想定してその心的状態や行動を予測し, 自身の行動を決定する

ここで，他者モデルを用いた推論は再帰的構造を持つ [1] ため，理論的には無限に計算が続きうる．

具体的には，レベル 1 では，他者モデルを用いて「他者の心的状態」を推定した上で自己の行動を決定する．レベル 2 では,「他者が想定している「自己の心的状態」」をも踏まえて自己の行動を決定する．さらにレベル 3 では,「他者が想定している「自己が想定している「他者の心的状態」」」に基づく．と無限に続いていく．

しかし実際には, 生物はいずれかのレベルで計算を打ち切って情報処理していると考えられるため, 同一の他者モデルを持つエージェント間でもそのエージェントの知的情報処理に使えるエネルギーや作業記憶の容量によって振る舞いが異なる可能性がある．同様に, 理論的に計算能力が同等であるモデルであっても情報処理コストが低いモデルの方が高度な情報処理を実現できる可能性も考えられる．

3 認知的インタラクションフレームワーク (CIF)

[9] では，認知的なインタラクションを記述するうえで，フレームワークとして 3 つの系を提案している．各系は仮定する変数と変数間の関係を示す関数で表され，$\mathcal{S}_1$，$\mathcal{S}_2$，$\mathcal{S}_3$ の順に仮定する変数や関数が多くなる．すなわち，インタラクションを記述するうえで考慮する行動決定のプロセスや認知プロセスが $\mathcal{S}_1$ から $\mathcal{S}_3$ にかけて複雑化する．

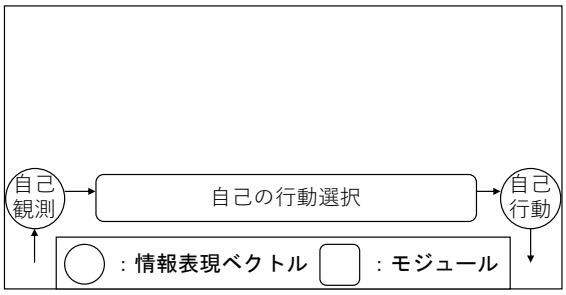

図 4: $\mathcal{S}_1$

各系を進化的な視点からみたコミュニケーションの階層性と照らし合わせると以下となる（詳しくは [9]）．

- $\mathcal{S}_1$: 物理的・予示的インタラクションを記述する系
- $\mathcal{S}_2$: 前言語的インタラクションを記述する系
- $\mathcal{S}_3$: 言語的・メタ認知的インタラクションを記述する系

本研究ではより具体的に，各系においてインタラクションを行うエージェントの行動決定に伴い用いられる情報表現について検討を行う．

- $\mathcal{S}_1$: 環境からの観測 o，o を反映させた自身の心的状態 s
- $\mathcal{S}_2$: o, s に加え，予測した相手の心的状態 $\hat{s}$
- $\mathcal{S}_3$: o, s, $\hat{s}$ に加え，$\hat{s}$ から予測した相手の行動 $\hat{a}$

3.1 $\mathcal{S}_1$

$\mathcal{S}_1$ の場合の情報処理に関して図 4 に示す．

$\mathcal{S}_1$ のインタラクションでは，エージェントは主に生物進化を通した適応（系統発生的適応）により獲得された行動パターンに従う．インタラクション最中に行動パターンは変化しない．つまり環境と心的状態が同一であれば，同一の行動を行う．例えば，捕食者-被捕食者の間のインタラクションや，なわばり行動などは $\mathcal{S}_1$ で記述される．

図 4 に示すように $\mathcal{S}_1$ におけるエージェントは，外界に対する観測と自身の心的状態のみによって行動決定を行う．すなわち，エージェント自身は相手の行動に対する情報処理を行う機構を持っていない．そのた

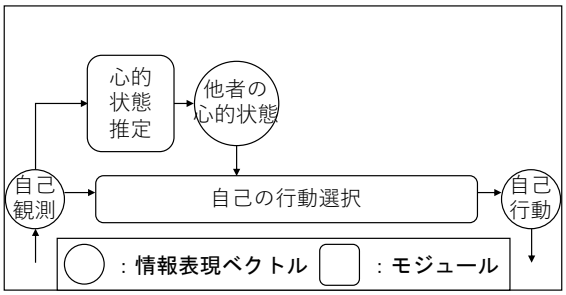

図 5: $\mathcal{S}_2$

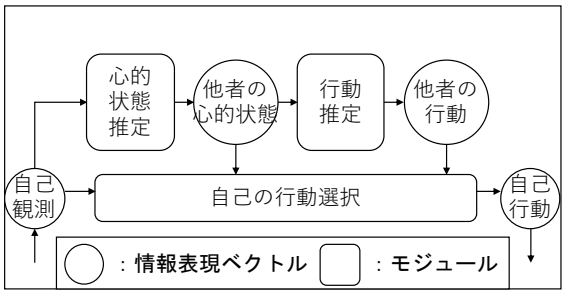

図 6: $\mathcal{S}_3$

め，外部から相手の行動に応じているように見える行動（例えば，接近行動に対する回避行動）の獲得過程はそのエージェント内のプロセスではなく，外部の事象（例えば，系統発生的適応）に依存することになる．

3.2 $\mathcal{S}_2$

$\mathcal{S}_2$ の場合の情報処理に関して図 5 に示す．

$\mathcal{S}_2$ のインタラクションでは，エージェントは前言語的なレベルでの社会性を持つ．すなわち，$\mathcal{S}_1$ と並行して，自身と対象の間の関係（二者間の関係）に応じて行動決定が行われる．$\mathcal{S}_2$ のインタラクションにおける相手の心的状態は行動により直接表出される．例えば，正直シグナルに代表される人同士の非言語インタラクションを表すためには，$\mathcal{S}_2$ による記述が必要と考えられる．

図 5 に示すように $\mathcal{S}_2$ におけるエージェントは，$\mathcal{S}_1$ に加えて，他者の心的状態を推定するプロセスを持つ．これにより，エージェント自身の行動に対する相手の反応を通して，エージェント自身の行動を学習することが可能となる．すなわち，そのエージェント自身の過去の経験を通して，相手の行動パターンに対する特定の行動パターンを返すことが可能となる．ただし，相手の行動生成過程について推定を行うプロセスを持たないことから，ここで対処できる行動パターンは社会的なシグナルといったコミュニケーションを行う相手一般に共通したものに限られる．

3.3 $\mathcal{S}_3$

$\mathcal{S}_3$ の場合の情報処理に関して図 6 に示す．

$\mathcal{S}_3$ のインタラクションでは，エージェントは高次の社会性を持つ．すなわち，$\mathcal{S}_1$，$\mathcal{S}_2$ と並行して，対象の行動を予測するためのモデル（他者モデル）に基づき，行動決定が行われる．他者モデルにより相手の心的状態を推定し行動を予測するプロセスがあり，相手の行動の予測結果に基づき自身の行動を計画することが可能となる．例えば，相手に配慮した行動や相手を欺く行動を表すためには，$\mathcal{S}_3$ の記述が必要と考えられる．

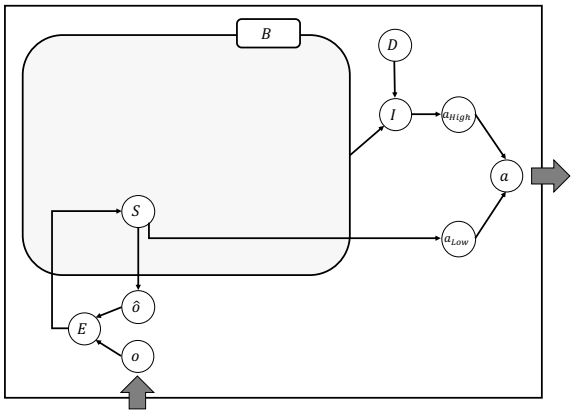

図 7: $\mathcal{S}_1$ における行動決定アーキテクチャ

4 CIF に基づく行動決定アーキテクチャと他者モデル

本章では前章で説明した CIF に基づき, 行動決定アーキテクチャを提案する. $\mathcal{S}_1$ の CIF から順に構築していき，その自然な拡張として $\mathcal{S}_2$, $\mathcal{S}_3$ を構築することで，進化や発達の過程で実際に実現しうる行動決定アーキテクチャの構築を目指す．

4.1 $\mathcal{S}_1$ における行動決定アーキテクチャと他者モデル

$\mathcal{S}_1$ における行動決定アーキテクチャを図 7 に示す．

$\mathcal{S}_1$ の場合，行動決定アーキテクチャ内の B は，環境からの観測 o や自身の心的状態 s によって構成される．

提案アーキテクチャは, BDI モデルと同様に信念 B と願望 D から意図 I を決定し, 意図 I を達成するための行動 a を決定し出力する. I の計算において，B に含まれる全ての変数の情報を利用できることを仮定して

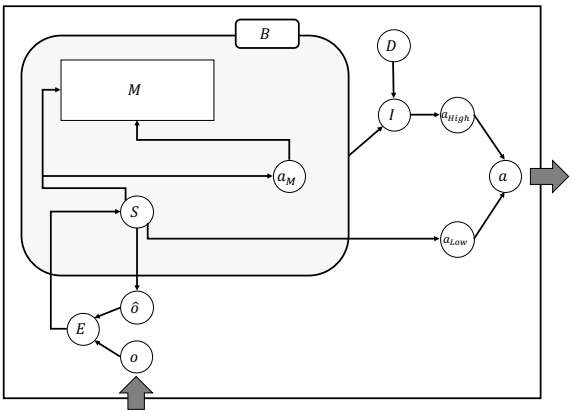

図 8: $\mathcal{S}_2$ における行動決定アーキテクチャ

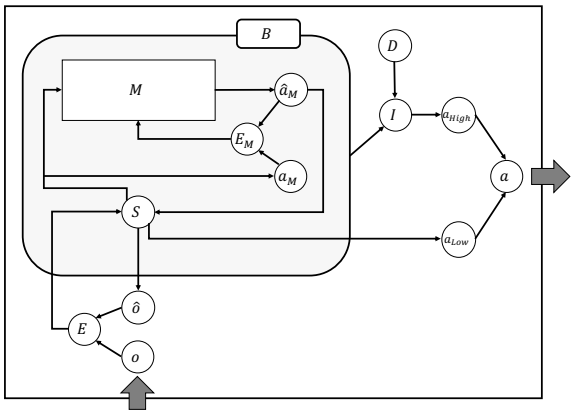

図 9: $\mathcal{S}_3$ における行動決定アーキテクチャ

いるが，図中では B 内の各変数から I の結線は省略しており，$\mathcal{S}_2$，$\mathcal{S}_3$ においても同様である．ただし $\mathcal{S}_1$ のフレームワークにおいては，B 内に含まれる変数は s のみであり，B と s は実質的に同じであるということに注意されたい．また, s の情報を用いて次時刻の観測 $\hat{o}$ を予測して実際の観測 o との予測誤差 E をフィードバックすることで, 環境の観測情報をエージェントの s に反映する．

この場合, 他者モデルは持っていない状態でありレベル 0 といえる．

4.2 $\mathcal{S}_2$ における行動決定アーキテクチャと他者モデル

$\mathcal{S}_2$ における行動決定アーキテクチャを図 8 に示す．

$\mathcal{S}_2$ の場合，行動決定アーキテクチャ内の B は，s に加え，他者に対する予測モデル M によって構成される．ここで他者モデルに対する入力は実質的に s のみであることに注意されたい．図には a_M も記載しているが，a_M は s のみから計算されている．モデルの中で他者の心的状態を推定しているが，他者の真の心的状態を知ることができないため，エラー訂正する仕組みを持たない．つまり $\mathcal{S}_2$ の設定においてこのモデルでは，他者の表出などから，他者の心的状態をクラスタリングしている状態といえる．この場合，学習のコストは大きく，高い性能が発現する可能性も低い．さらに，前言語的なインタラクションを想定しているため，心的状態にラベルづけもされていないだろう．

$\mathcal{S}_2$ の提案アーキテクチャにおいて, 他者モデルのレベルは理論上無限に計算可能であり, 「他者モデルの中に，他者が推定している自己の心的状態を予測する」というレベル 2 の情報処理を行うことも理論的には可能である．しかしながら，レベル 1 の場合に比べてありうる情報表現空間が膨大になるため，学習のコストが大きい $\mathcal{S}_2$ のアーキテクチャではレベル 2 以降の計算が現実的に行われるとは考えにくい．実際に，他者の行動を推定することや，推定結果を踏まえた推論といった知的処理ができないエージェントが，他者が想定する自己の状態といった情報を扱って行動しているとは考えにくいため，現実的にはレベル 1 までの他者モデルを扱っていると考えられる．

4.3 $\mathcal{S}_3$ における行動決定アーキテクチャと他者モデル

$\mathcal{S}_3$ における行動決定アーキテクチャを図 9 に示す．

$\mathcal{S}_2$ の設定と同様に，他者モデルに対する入力は s のみである．しかしながら他者の行動 $\hat{a}_M$ を予測しており，実際の他者の行動 a_M も観測できるため，モデルの予測エラー (E_M) を計算することができる．心的状態や行動を推定する関数は，ここで算出されたエラーをフィードバックすることで訓練できるため，$\mathcal{S}_2$ の設定よりも高速かつ高精度に獲得できると考えられる．他者モデルを高速に構築できる結果，レベル 2 以降の他者モデル構築が可能になったと考えられる．

これはつまり，他者の心的状態の予測から，他者の行動の予測という，いたって自然な能力拡張が，他者モデルという情報処理アーキテクチャに当てはめた場合に全体として大きな知的能力の拡張を起こす要因である, という仮説が提唱したといえる．

5 幼児研究からみるBDIモデルベース他者モデルの発達

本章では，子ども間のインタラクションを検討した研究から BDI モデルベース他者モデルの発達を議論する．以下では，Ichikawa らの 研究 [13] を取り上げ，CIF で

提唱された $\mathcal{S}_1$，$\mathcal{S}_2$，$\mathcal{S}_3$ と対応づけて各年代における他者モデルの特徴を述べる．

Ichikawa らの研究では，保育園を定期的に訪問し，音楽に合わせて身体を動かす表現活動であるリトミックの撮影を 10 数人で構成されるクラスごとに実施した．リトミックのなかでもピアノの演奏中に自由に走るという最も単純な遊びを対象に，子どもの社会性に基づく他の子どもに関わろうとする接触行動（例えば，手を握る，抱きつく）に着目した．

具体的な取り組みとして，アノテーションソフトを用いて動画から各子どもの接触行動を見つけてタグ付けして接触行動の頻度などを求める分析を行った．さらに，子どもの位置データを取得して，生物学やスポーツ科学の集団運動に関する先行研究 [14, 15, 16] で用いられている指標を参考に，子ども間の距離や走る方向などを分析した．

5.1 年少児の他者モデル

年少児クラスの動画に対してアノテーション分析を行ったところ，接触行動はほとんどみられなかった．一部の子どもは，さまざまな方向へ自由に走っていた．

このような特徴が観察される背景には，他者モデルが十分に発達していないことが関連する可能性がある．子どもは $\mathcal{S}_1$ の状態で他者モデルのレベル 0 に基づいて行動し，自らの願望や意図のみに従って走っていた，あるいは無意識的に走っていたと考えられる．

5.2 年中児の他者モデル

年中児クラスの結果を対象にアノテーション分析だけでなく，位置データを用いた分析も行ったところ，接触行動はほとんどみられなかったが，全体的に円状になって走る特徴が観察された．

円状に走る特徴がみられる背景には，他者モデルの発達が関連すると考えられる．$\mathcal{S}_2$ の状態で，他者モデルのレベル 1 に基づいて行動していることが予想される．他の子どもの心的状態を推定して自身の行動を決定している可能性がある．ただし，$\mathcal{S}_3$ の状態とは異なり，他者の行動自体は予測していないため，現れる集団遊び自体は複雑ではない．なお，発達心理学において他者の心的状態の推定に関する心の理論は 5 才ごろに発達すると言われている [12]．この知見は，年中児クラスで確認された他者モデルの発達と整合性があるといえる．

5.3 年長児の他者モデル

年長児クラスにおいても同様にアノテーション分析や位置データを用いた分析を行ったところ，接触行動が多くみられ，短時間で他の子どもに向かって近づく頻度が高い特徴が確認されており, 全体的に鬼ごっこのような集団遊びが形成されていた．

この時点では, 子どもの個人差や時間帯によって $\mathcal{S}_2$ や $\mathcal{S}_3$ が混在する，あるいは切り替わるような状態で他者モデルのレベル 1 やレベル 2 に基づいて走っていたと考えられる[1]．鬼ごっこのような遊びを成立させるためには，他の子どもの心的状態を推定し，行動を予測して反応することが求められる可能性がある．

Ichikawa らの実験の結果と対応して考察を行なった結果, 他者モデルの発達に伴い，同じような状況においても遊びがより戦略的になることが示唆された．ただし，どの程度正確に他の子どもの行動を予測しているか，あるいは行動から $\mathcal{S}_2$ や $\mathcal{S}_3$ の状態を識別するといった詳細まで踏み込んで議論することが難しい点に留意する必要がある．この理由として保育園の現場で撮影していることで，十分な統制がとれていないことが関係しており，外部観測から検討することの限界を示唆している可能性がある．

5.4 検討事項

Ichikawa らの研究 [13] において，外部観測された各年代の子どもの集団遊びに関する特徴を，CIF の状態 S および BDI モデルベース他者モデルでうまく説明できることが示された．なお，この研究では，発達について集団運動と認知を関連づけることで，質問紙調査や現場観察による記述では得ることが困難な知見が得られた．しかし，第 3 節で挙げた問題点がある．これをクリアするためには，他者モデルを数理モデルとして記述することで，集団運動のマルチエージェントシミュレーションが鍵になると考えられる．

例えば，年齢に伴って他者モデルが発達し，$M_{intention}$ の持つモデル M に $M_{intention}$ が導入されるようにパラメータを変化させ，Void モデル（例えば，[14]）のようにエージェント間でインタラクションさせることで，集団運動にどのような変化がみられるかを検討するアプローチが挙げられる．集団運動と他者モデルの関連について発展的な議論が期待される．

[1] 動画を見る限り，3 以上のレベルで子ども同士がインタラクションしているとは考えづらい．バスケットボールのプレーで起こるような駆け引きは観察されなかった．

6 おわりに

本論文では，3 段階の認知的インタラクションフレームワーク (CIF) に基づいた他者モデルを提案した．提案したモデルによって，各 CIF の段階で実現されている他者モデルに関する機能が発現する仕組みについて，情報処理の観点から説明可能であった．さらに，著者らの一部による子どもを対象とした研究成果を，本論文で提案したモデルに基づいて再解釈し，モデルの妥当性についても検討した．

参考文献

[1] 牧野貴樹, 滝久雄, 合原一幸: 利他的行動と再帰的他者推定. 生産研究 Vol. 62 No. 3, pp. 259–265 (2010)

[2] 横山絢美, 大森隆司: 協調課題における意図推定に基づく行動決定過程のモデル的解析., 電子情報通信学会論文誌 *A*, Vol. 92, No. 11, pp. 734–742, (2009)

[3] 鴫原 宏明, アッタミミ ムハンマド, 阿部 香澄, 長井 隆行, 大森 隆司, 岡 夏樹: 確率モデルに基づく他者モデル相互適応のモデル化., HAI シンポジウム, (2014)

[4] Premack, D.: The infant's theory of self-propelled objects., Cognition Vol. 36 No. 1, pp. 1–16, (1990).

[5] Opfer, J.E.: Identifying living and sentient kinds from dynamic information: The case of goal-directed versus aimless autonomous movement in conceptual change., Cognition Vol. 86, No. 2, pp. 97–122, (2002).

[6] 小川 浩平, 小野 哲雄: ITACO: メディア間を移動可能なエージェントによる遍在知の実現., ヒューマンインタフェース学会論文誌, Vol. 8, No. 3, pp. 373–380, (2006)

[7] 寺田和憲, 社本高史, 伊藤昭: 心の理論の枠組を利用した人工物から人間への意図伝達., ヒューマンインタフェース学会論文誌, Vol. 9, No. 2, pp. 23–33, (2007).

[8] 植田 一博: 『認知的インタラクションデザイン学』の展望: 時間的な要素を組み込んだインタラクション・モデルの構築を目指して., 認知科学 Vol. 24 No. 2, pp. 220–233 (2016)

[9] 坂本 孝丈, 竹内 勇剛: HAI 研究の体系化に向けたフレームワークの提案, HAI シンポジウム, (2020)

[10] Dennett, D.C.: The intentional stance., MIT press, (1989).

[11] ダニエル.C. デネット: 「志向姿勢」 の哲学: 人は人の行動を読めるのか., 白揚社 (1996).

[12] Gopnik, A., and Astington, J. W.: Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, Vol. 59, No. 1, pp. 26–37 (1988)

[13] Ichikawa, J., Fujii, K., Nagai, T., Omori, T., and Oka, N.: Quantitative analysis and visualization of children's group behavior from the perspective of development of spontaneity and sociality. *Proceedings of the 24th International Conference on Collaboration and Technology (CRIWG 2018)*, pp. 169–176 (2018)

[14] Couzin, I. D., Krause, J., James, R., Ruxton, G. D., and Franks, N. R.: Collective memory and spatial sorting in animal groups. *Journal of Theoretical Biology*, Vol. 218, No. 1, pp. 1–11 (2002)

[15] Kijima, A., Kadota, K., Yokoyama, K., Okumura, M., Suzuki, H., Schmidt, R. C., and Yamamoto, Y.: Switching dynamics in an interpersonal competition brings about "deadlock" synchronization of players. *Plos One*, Vol. 7, No. 11, doi:10.1371/journal.pone.0047911 (2012)

[16] Tunstrøm, K., Katz, Y., Ioannou, C. C., Huepe, C., Lutz, M. J., and Couzin, I. D.: Collective states, multistability and transitional behavior in schooling fish. *PLoS Computational Biology*, Vol. 9, No. 2, doi:10.1371/journal.pcbi.1002915 (2013)

[17] Baker, C. L., Jara-Ettinger, J., Saxe, R., and Tenenbaum, J. B.: Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour* Vol. 1 No. 4, fpp. 1–10 (2017)

[18] M.E. Bratman: Intention, Plans, and Practical Reason, *Harvard University Press*, (1987)

[19] Rao, Anand S., and Michael P. Georgeff: Modeling rational agents within a BDI-architecture. *KR 91*, pp.473-484, (1991)