

認知アーキテクチャを利用したコミュニケーションシステムの発生に関するシミュレーション

Simulation of Generating Communication Systems based on a Cognitive Architecture

森田 純哉*¹ 金野 武司*¹ 橋本 敬*¹
Morita Junya Konno Takeshi Hashimoto Takashi

*¹北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology

We conducted a simulation study exploring a formation of communication systems. Imitation and instance models, which could perform a coordination game, were implemented in ACT-R architecture. As a result, we observed a good fitting with human behavioral data in the imitation model. We also found differences of BOLD predictions between the two models. The imitation model produced higher BOLD predictions in the declarative module that was assumed to correspond to the prefrontal cortex. This result is consistent with past psychological/brain studies concerning perspective taking, implying imitation requires high cognitive load.

1. はじめに

認知アーキテクチャの研究とは、当初より知覚や運動、記憶、思考など複数の認知機能を統合することを目指してきた [Newell 94]. 近年では、認知アーキテクチャによる脳活動のシミュレーションが盛んになされている。たとえば、Anderson らによって開発された ACT-R (Adaptive Control of Thought-Rational [Anderson 07]) は、アーキテクチャを構成するモジュールを脳部位に対応づけ、fMRI により計測される BOLD 反応を予測する。この機能を用いることで、課題の習熟にともなう前頭前皮質の活動低下などの現象が再現されている [Anderson 07, Qin 03].

認知アーキテクチャを利用した脳活動のシミュレーションは、認知と脳の統合的な理解を導く利点がある。ACT-R を含む多くの認知アーキテクチャは、人間の意識状態をバッファ (短期的な情報の貯蔵スペース) に保持されるコンテンツ (記号群) とみなす。バッファのコンテンツは、プロダクションルールを介して推移し、それによって複雑な課題を遂行する際の認知プロセス (意識の流れ) が再現される。認知アーキテクチャの出力と脳計測データを対応付けることは、記号的意識の背後にある神経活動の解明に繋がる。

しかし、これまで ACT-R を用いた脳活動シミュレーションは、個人内部で完結するプロセスのみを対象とし、二者間の相互作用によって形成されるプロセスを検討してこなかった。人間の活動を包括的に理解し、認知科学・人工知能の分野を発展させるためには、認知・脳・社会という 3 つの異なるレイヤーを結びつける必要がある。

このような問題意識から、本発表では、二者間でのコミュニケーションシステムの成立過程をシミュレーションによって検討する。コミュニケーションは、人間社会の基礎を構成する活動であり、その成立過程を検討することは、人間の社会性の原点を探ることに繋がる。本研究では特に、コミュニケーションシステムの成立過程における模倣に注目し、行動実験の結果を再現し、背後にある脳活動を推測することを目指す。

2. 行動実験

本研究におけるシミュレーションは、著者らの先行研究 [Konno 13] にて報告された行動実験を再現するものである。実験課題は、独立したラウンドの繰り返しから構成される調整ゲームであった。各ラウンドにおいて、2 名の参加者はペアとなり、それぞれ別の初期位置に配置されるキャラクターを、共通の位置に移動させることを目指した。

ゲームの環境に、キャラクターが配置可能な位置は 4 つ存在し (2 × 2 の配置)、参加者は、各自のキャラクターを、1 ラウンドにつき 1 回のみ動かすことができた。キャラクターは、上下左右方向のみ動かすことができ、斜め方向には動かすことができなかった。よって、それぞれの参加者には、初期位置からキャラクターを移動できない位置が存在した。ラウンド中、パートナーのキャラクターの位置を見ることはできず、ラウンドが終了した時点 (二者が共に移動を終えた時点) で、参加者はパートナーの初期位置と移動先を知らされた。

参加者は移動に先立ち、1 回のみメッセージを交換できた。メッセージは、意味や運用ルールが事前に定められていない図形を 2 つ組み合わせることで構成された。メッセージの交換は非同期になされ、一方の参加者がメッセージを送信すると、直ちにパートナーにメッセージが届けられた。メッセージを先に受け取った参加者は、受け取ったメッセージを考慮しつつ、自分がパートナーに送るメッセージを決定できた。つまり、本実験課題では、メッセージの送信順序を利用したターンテイクを行うことができた。これを利用することで、たとえば、参加者達は、現在位置の送信者、移動先の送信者などの役割を分担することができた。

メッセージに利用する記号の意味は事前に定まっていなかったものの、ラウンドの繰り返しにより、記号の意味に関する合意がペア内で形成され、継続的に共通の位置への移動できるようになる。逆に言えば、この課題の成功に記号の意味に関する合意の形成は不可欠であり、この課題におけるコミュニケーションシステムの成立は、継続的な共通の位置への移動によって判定できる。行動実験では、共通の場所に移動したラウンドにおいて 2 点、移動に失敗したラウンドにおいて -1 点の得点をペアに与えた。この得点が 50 点に達するか、60 分が経過した後課題を終了した。

連絡先: 森田純哉, 北陸先端科学技術大学院大学, 石川県能美市旭台 1-1, 0761-51-1707, j-morita@jaist.ac.jp

3. モデル

ACT-R を利用することで、前節の行動実験を再現するモデルを構築した (詳細は, [Morita 12] を参照). モデルは, 2体の ACT-R エージェントが, 実験と同型の課題環境を介して相互作用するものである. それぞれのエージェントは, 視覚モジュール, 運動モジュール, 宣言的モジュール, プロダクションモジュール, ゴールモジュールによって構成される.

エージェントは, 事例ベース学習, プロダクションコンパイル, 強化学習など複数の学習メカニズムを保持する. 事例ベース学習は, 移動に成功 (2つのキャラクターの位置が一致) したラウンドにおいて, ゴールモジュールに蓄えられた情報 (お互いの初期位置, 移動先, メッセージ) を, 事例として宣言的モジュールに格納する. 格納された事例は, 以後のラウンドにおいて, 移動先の決定やメッセージの作成に利用される. 事例が繰り返し利用されることで, その事例はプロダクションルールに変換される (プロダクションコンパイル). さらに, 強化学習によって, 成功率の高いプロダクションルールが優先的に用いられるようになる.

エージェントは, 事例の利用の仕方として, 2つの方略を保持する. 一方の方略は, 現在のラウンドにおいて認識される情報と合致する事例を検索し, その事例と同じように移動先やメッセージを決定する (事例方略). 他方の方略は, 自分とパートナーの役割を入れ替えた状態で, 現在のラウンドにおいて認識される情報と合致する事例を検索する (模倣方略),

パートナーからのメッセージを未だ受け取っていないとき, 事例方略は, 現在のラウンドにおける自分の初期位置と事例における自分の初期位置が一致する事例を検索し, 過去の自分と同様に送信するメッセージと行き先を決定する. それに対して, 模倣方略は, 現在のラウンドにおける自分の初期位置とパートナーの初期位置が一致する事例を検索し, 過去のパートナーと同様のメッセージと行き先を決定する. パートナーからのメッセージが既に届いているとき, 事例方略は, 現在のラウンドにおいてパートナーから受け取ったメッセージと同じメッセージを自分が受け取った事例を検索し, 事例と同様に自分のメッセージと行き先を決定する. 模倣方略は, パートナーから受け取ったメッセージを自分が送信した事例を検索し, 事例におけるパートナーと同様にメッセージと移動先を決定する. 事例方略が, 現在の自分を過去の自分にあてはめる推論であるのに対し, 模倣方略は, 現在の自分を過去の他者の立場にあてはめる推論といえる [Tomasello 99].

本研究におけるシミュレーションでは, このような模倣方略が, コミュニケーションシステムの発生において, どのように寄与するのかを探る. そのために, 事例方略のみをもつモデル (事例モデル) と事例方略に加えて模倣方略をもつモデル (模倣モデル) を比較する. 事例モデルは, 事例の検索が失敗した後に, ランダムに行き先やメッセージを決定する. 模倣モデルは, 事例方略が失敗した後に, 模倣方略を行い, それが失敗した後に, 行き先やメッセージをランダムに決定する.

4. シミュレーションの結果

4.1 行動の再現

図 1 は, 課題の開始から終了条件に至るまでの各ラウンドにおいて, 位置の移動に成功した試行の割合 (成功率) を示している. 模倣モデル, 事例モデルともに試行数は 100 である. 図中の点線 (Human data) は, [Konno 13] における行動実験の結果を示している. 成功率の集計では, 終了条件に到達した試行のみを対象とした. 終了条件に至ったラウンドは試行毎に

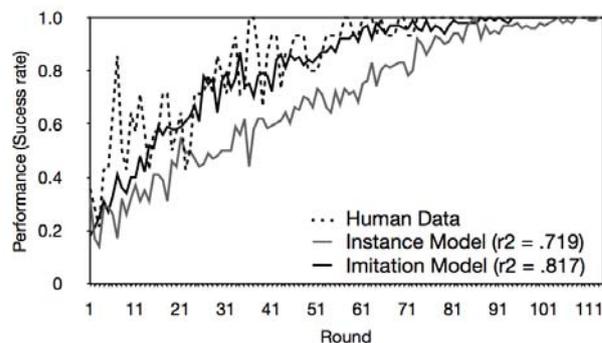


図 1: 各ラウンドにおける成功した試行の割合.

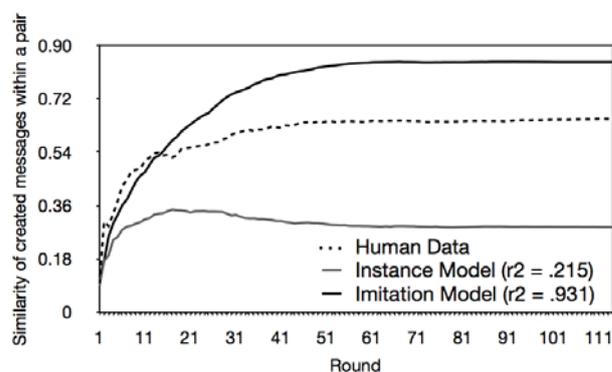


図 2: 各ラウンドまでに送信されたメッセージのペア内での類似度.

異なったが, 終了条件到達以後のラウンドは, 位置の移動に成功したものと扱った. なお, 模倣モデルと事例モデルはともに全ての試行において, 終了条件に至り, 行動実験では 21ペア中, 14ペアが終了条件に至った. 図より, 事例モデルに比べ, 模倣モデルが, 行動実験における成功率の推移をよく再現していることがわかる. ここから, 模倣モデルは行動実験における成功ペアと同程度の速さでコミュニケーションシステムを成立させたといえることができる.

図 2 は, ペア内でやり取りされるメッセージの類似度の推移を示している. そのラウンドまでに使用された図形の組み合わせの頻度から単位ベクトルを構成し, 二者間での内積を計算することでメッセージの類似度とした. 図 1 のデータと同様, 終了条件に至ったペアのみを分析の対象とした. 実験データと模倣モデルは, ラウンドの進行に伴い, メッセージの類似度を上昇させた. それに対し, 事例モデルは類似度を上昇させなかった. この結果からも, 模倣モデルが事例モデルに比べ, 人間のデータをよく説明することが示される. ただし, 模倣モデルに比べ, 人間のデータの最終的な類似度は低いものになった. この差については, 5 節において考察する.

4.2 fMRI データの予測

コミュニケーションシステムの発生における模倣は, どのような脳内活動によって成し遂げられるのであろうか. 2 節に示した実験では, 課題遂行中の脳活動を計測しなかったため, この問いに直接答えることはできない. 本研究では, 3 節において構築されたモデルによって課題遂行中の脳活動を推測する.

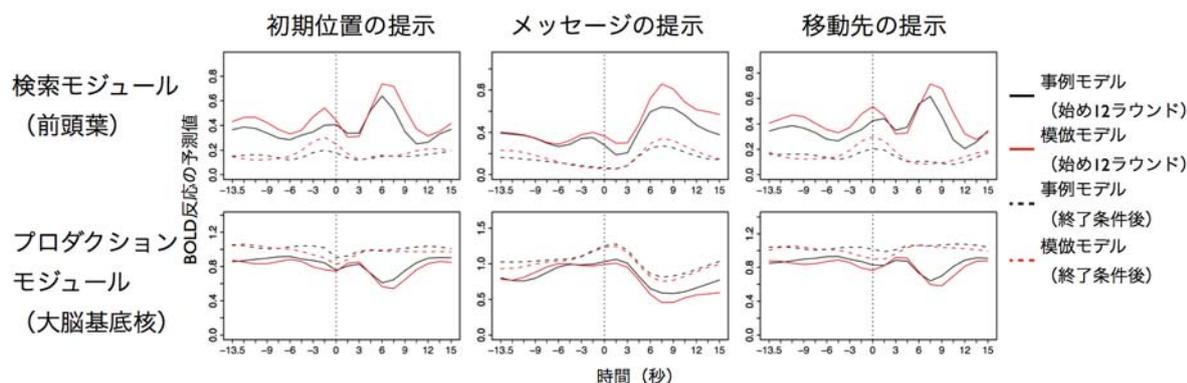


図 3: 模倣モデルによる BOLD 反応の予測.

ACT-R による BOLD 反応の予測は、各モジュールにおけるバッファの活動履歴（トレース）を変換することでなされる。この機能を用いることで、実装されたモデルと同様のプロセスを人間が行っていた場合の脳活動を検討することができる。

図 3 に、事例モデルと模倣モデルのトレースから予測される BOLD 反応を示した。モデルには、5つのモジュールが用いられたが、ここでは、宣言的モジュールとプロダクションモジュールの結果を示す。これら2つのモジュールは、ACT-R を用いた過去の研究で度々検討されてきた [Anderson 07]。それらの研究では、宣言的モジュールは前頭前皮質に、プロダクションモジュールは大脳基底核に対応する仮定されている。

図中の各グラフに示される縦の破線は、イベント（初期位置の提示、メッセージ受信、移動結果の提示）の生起時刻を示している。事例モデル（黒線）と模倣モデル（赤線）のそれぞれ 50 試行を対象に、試行中のイベント生起時刻を同定した。そして、イベント前後 15 秒にわたり、各時点での BOLD 反応予測値を平均した。それぞれのグラフにおいて、実線は課題の始め 12 ラウンドの平均値、破線は終了条件に到達した後に、同様の手続きで実施したテスト課題*1の平均値を示している。

図より、終了条件後（破線）に比べ、課題の始め（実線）では、宣言的モジュールが活発に活動したことがわかる。それに対し、プロダクションモジュールは、課題の始めに比べ、終了条件後において活動が活発であった。この結果は、1 節において述べた過去の ACT-R の研究と整合的である。過去の研究では、課題の習熟に伴う前頭前皮質の活動低下は、宣言的知識の自動化（プロダクションコンパイル）により引き起されると説明されている [Anderson 07, Qin 03]。

事例モデルと模倣モデルの差異は、宣言的モジュールにおける BOLD 反応の強さに見ることができる。特に始め 12 ラウンドの段階で、模倣モデルは、事例モデルに比べ、宣言的モジュールでの BOLD 反応を強くしている。3 節で述べたように、模倣モデルは、事例方略が失敗した後に、模倣方略による事例の再検索を行う。よって、模倣モデルは、事例モデルに比べ、宣言的モジュールへの負荷が高くなったと考えることができる。

5. まとめ

本研究におけるシミュレーションでは、コミュニケーションシステムの成立に関わる人間の行動を再現し、BOLD 反応を予測した。行動のシミュレーションによって、模倣モデルが行動実験のデータをよく再現することが示された。この結果より、人間の社会性は、現在の自分を過去の他者の立場にあてはめる推論によって成り立っていることが示唆された。

BOLD 反応の予測では、事例モデルと模倣モデルはともに、終了条件後で宣言的モジュールの活動を低下させた。この結果は、終了条件後のコミュニケーションは、前頭前皮質に負荷のかからない自動化されたものであることを示している。このことは、我々が、特別な認知的負荷をかけずに母語を話せることと整合的である。

また、模倣モデルは、事例モデルに比べ、課題初期において、宣言的モジュールを活発に活動させた。模倣モデルは、事例の検索を 2 つの方法で繰り返すことができた。これにより、模倣モデルは、少ないラウンド数で、コミュニケーションシステムを成立させることができたと考えられる。ただし、このメリットは、宣言的モジュールに負荷をかけるものであったといえる。

模倣モデルにおける宣言的モジュールへの負荷は、自己と他者に対応付ける推論の困難さを示している可能性がある。実際、コミュニケーションに関わる過去の研究は、パートナーの立場に自分を置く推論が、反応時間を増大させ [Lin 10]、前頭前皮質を含む領域を賦活させることを示している [Dumontheil 10]。模倣が人間にとって困難な推論であるとすれば、メッセージの類似度（図 2）における模倣モデルと人間のデータとの差も解釈可能である。行動実験において、実験参加者は、認知的負荷を嫌い、模倣が可能な場面であっても模倣を行わなかったのかもしれない。

上記の議論は、本研究において構築された ACT-R モデルによって導かれる推測である。今回示したモデルは、模倣の 1 つの実装であり、別のモデルを考えることも可能である。また、ACT-R による BOLD 反応の予測は、実際の脳計測データとの対応付けによって、検証されるべきである。今後、ACT-R による脳計測データとの対応付けを行うことで、認知・脳・社会のレイヤーの包括的な理解に近づくと考える。

*1 12 ラウンドによって構成。詳細は [Konno 13] および [Morita 12] を参照

謝辞

本研究は文部科学省科学研究費補助金 (課題番号 21120011) の助成を受けた。

参考文献

- [Anderson 07] Anderson, J. R.: How can the human mind occur in the physical universe? New York: Oxford University Press (2007)
- [Dumontheil 10] Dumontheil, I., Kuster, O., Apperly, I. A., and Blakemore, S: Taking perspective into account in a communicative task, *NeuroImage*, Vol. 52, pp. 1574-1583 (2010)
- [Konno 13] Konno, T., Morita, J., and Hashimoto, T.: Symbol communication systems integrate implicit information in coordination tasks, in Yamaguchi, Y. ed., *Advances in Cognitive Neurodynamics(III)*, Springer, pp. 456-460 (2013)
- [Lin 10] Lin, S., Keysar, B., and Epley, N.: Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention, *Journal of Experimental Social Psychology* 46, pp. 551-556 (2010)
- [Morita 12] Morita, J., Konno, T., and Hashimoto, T.: The role of imitation in generating a shared communication system, in *Proceedings of the 34th Annual Meeting of the Cognitive Science Society*, pp. 779-784 (2012)
- [Newell 94] Newell, A.: Unified Theories of Cognition. Harvard University Press (1994)
- [Qin 03] Qin, Y., Sohn, M.T., Anderson, J. R., Stenger, V. A., Fissell, K., and Goode, A.: Predicting the practice effects on the blood oxygenation level-dependent (BOLD) function of fMRI in a symbolic manipulation task *Proceedings of the National Academy of Sciences* 100 (8), pp. 4951-4956 (2002)
- [Tomasello 99] Tomasello, M: The cultural origins of human cognition. Harvard University Press (1999)