

記号的コミュニケーションシステムの形成過程において 解釈学的循環を構成するモデルベース強化学習の役割

The role of model based reinforcement learning with hermeneutic circle in formation of symbolic communication systems

金野 武司^{†1}, 橋本 敬^{†2}, 奥田 次郎^{†3}, 鮫島 和行^{†4}, 李 冠宏^{†2}

Takeshi Konno, Takashi Hashimoto, Jiro Okuda, Kazuyuki Samejima, Guanhong Li

^{†1} 金沢工業大学 工学部 電子情報通信工学科

^{†2} 北陸先端科学技術大学院大学 知識科学系

^{†3} 京都産業大学 コンピュータ理工学部 インテリジェントシステム学科

^{†4} 玉川大学 脳科学研究所 基礎脳科学研究センター

^{†1} Information and Communication Engineering, Dept. of Electronics, Kanazawa Institute of Technology

^{†2} School of Knowledge Science, Japan Advanced Institute of Science and Technology

^{†3} Faculty of Computer Science and Engineering, Kyoto Sangyo University

^{†4} Basic Brain Science Research Center, Brain Science Institute, Tamagawa University

konno-tks@neptune.kanazawa-it.ac.jp, hash@jaist.ac.jp, jokuda@cc.kyoto-su.ac.jp,

samejima@tamagawa.ac.jp, adam.li@jaist.ac.jp

Abstract

In most cases, humans do not directly link their own behavior to symbolic messages from the other person but rather to symbolic meaning based on the inferred intent of the other person. At such times, the inferred intent of another person plays a hypothetical role in identification of the meaning of a message. Therefore, it is pointed out that the method forms a hermeneutic circle with another person. This study uses model-based reinforcement learning to model the cognitive mechanism of subjects forming such circles. This model was applied to participant behavior data from a task in which pairs created artificial languages, and by using likelihood to evaluate the model's applicability, the predictive ability of the model to describe cognitive mechanisms was investigated. The results demonstrated a pattern of high likelihood in the initial stages of the task, and a gradual decrease moving toward the final stages. These results suggest the possibility that the mechanism of behavioral decisions, in response to symbols, changes along the way.

Keywords — experimental semiotics, symbolic communication, denotation and connotation, model based reinforcement learning

記号的なコミュニケーションにおいて人間は、「字義通りの意味」だけではなく「言外の意味（意図）」を伝える [1]. 例えばテーブルで食事をとっている状況で、「お塩とれる？」と聞いたとき、字義通りには「塩（の瓶）をとることが可能か」を意味していても、「塩をとって欲しい」という言外の意味が込められていることを受け手は理解する。コミュニケーションにおいて人間がこのような二重の意味をやりとりすることは自然で当たり前のことだが、自閉症におけるコミュニケーション障害は、それが当たり前のことではないことを示している [2].

表現が二重の意味を持つことを前提としたコミュニケーションは、「字義通りの意味」の特定なくして「言外の意味」を適切に推論することができず、また「言外の意味」が適切に推論できなければ、多義性を持つ「字義通りの意味」を相手との間で特定することができないという解釈学的循環が生じる問題がある [3, 4]. すなわち、一連の会話の中では、個々の語や文の「字義通りの意味」がわかり、そこから発話の「言外の意味」を推論することだけが問題なのではなく、「言外の意味」が分かった場合に、既存の発話やその後の発話における語や文の「字義通りの意味」を特定することもまた問題となる。

これまでに我々が設計・実施してきた人工言語の共創実験 [5] では、そうした「言外の意味」と「字義通りの意味」の相互循環的な取り決めプロセスが明示

的に生じる。この実験で参加者は、日常的なコミュニケーション手段（会話やジェスチャ、あるいは表情など）の使用が制限された状態で、実験者が用意する図形のやりとりを通じて、設定された協調課題を解くための人工的な言語の作成に取り組む。これにより参加者は、図形の「字義通りの意味」を相手と取り決めなければならない状況に置かれる。

この取り決め過程において、参加者が相手の使う図形の「字義通りの意味」を、推定される「言外の意味」から特定するのは直感的に自明なことである。しかし計算モデルを構築すると、それはかならずしも必要とされる仕組みではないことに気づかされる [6]。なぜなら、自分の状況に対する行動の結果から報酬を受け取って学習を進める強化学習を考えた場合、相手の図形と行動の直接の対応関係を学習する方法が考えられるからである。強化学習において、この方法はモデルフリー強化学習に対応し、前出の方法はモデルベース強化学習に対応する。

計算モデルの構築を通じたこれまでの分析により、モデルベース強化学習は、「言外の意味」が課題に対して適切に設定され、互いがそれを誤解することがなければ、モデルフリーよりも課題を解くための記号システムを素早く共創できることが分かった。また、モデルベース強化学習の1つの懸念は、相手の「言外の意味」を誤解した際に図形が持つ「字義通りの意味」の特定も間違え続けるのではないかというものであるが、「言外の意味」を様々に設定した場合においても、モデルベース強化学習はモデルフリーよりも同じ時間で実現される課題の成功率が悪くなることがなかった [6]。

そこで我々は、実験参加者の行動データにそれらのモデルを適用し、その当てはまりの良さを尤度を用いて評価することで、それぞれのモデルが持つ認知機構の説明力を調べた。実験では20ペア（40名）が課題に取り組み、その中で60回（ラウンド）の図形の交換と行動選択が行なわれた。結果、13ペアが協調課題を安定的に解く記号システムの共創に成功した（以降、これを成功群とする）。

この課題を解く主体の認知機構を説明する計算モデルとしては以下の3つのタイプを構築した。ただし、ここで構築したモデルの主体は、自分の「言外の意味」をあらかじめ設定し、それに基づいて図形を相手に送るものとした。

- 相手から送られてくる図形に対して、「言外の意味」に基づく「字義通りの意味」の特定を行わず、図形と行動の直接の対応関係を学習する（特

定なしモデル）。

- 相手から送られてくる図形に対して、「言外の意味」に基づき「字義通りの意味」を特定する（特定あり1表象モデル）。
- 1表象モデルと同様に相手から送られてくる図形の「字義通りの意味」を特定する。ただし、自分の「字義通りの意味」と相手の「字義通りの意味」を別の表象として独立して特定する（特定あり2表象モデル）。

初めのモデルがモデルフリー強化学習に、後の2つがモデルベース強化学習にそれぞれ対応する。ただし、モデルベース強化学習は、相手の「言外の意味」の解釈が複数あったとしても、そのうちの一つであると仮定して「字義通りの意味」を特定することとした。

この3つの計算モデルを、認知実験により得られた行動データに対する当てはまりの良さ（尤度）で調べた。尤度は、人間の行動データの系列があるとき、ある時点までのデータを学習した結果として、その次の時点の行動をどの程度の確率で予測できるかを計算するものである。この尤度を計算する際に、計算モデルでは「言外の意味」を決定する必要がある。これは行動データから推定することが望ましいが、相互循環的に変化する性質を考えると、それは非常に難しいと考えられる。そこで、計算モデルが想定する「言外の意味」（ここでは2つ）について、その全ての組み合わせ（二者がそれぞれに自分と相手に関する「言外の意味」を持つ）で尤度を計算し、最も高い尤度になる「言外の意味」を使用した。

それぞれの計算モデルで12ラウンドごとの平均尤度を計算し、特定なしモデルと特定あり1表象モデルの尤度に関して分散分析を行なった¹。結果、課題の成功群（13ペア）と失敗群（7ペア）の尤度に有意な差が認められ（ $F(1,18) = 53.05$, $MSE = 0.026$, $p < .001$, $\eta^2 = 0.303$ ）、それぞれの尤度の高さは課題の成功率と有意に相関していた（特定なしモデル： $r(18) = .95$, $p < .001$ 、特定あり1表象モデル： $r(18) = .83$, $p < .001$ ）。これは、計算モデルが想定する仕組みが、課題の成功を実現していることを示唆する。

続けて、成功群において学習の進行に伴う尤度の変化を調べた。成功群における計算モデルごとの平均尤度の推移を図1に示す。これを見ると、特定なしモデルと特定あり1表象モデルの尤度は課題の序盤と終盤で入れ替わっているように見える。分散分析をすると、モデルと期間の間に有意な交互作用が見られた（ F

¹ 3要因混合計画（成功群と失敗群、特定なしモデルと特定あり1表象モデル、5つの期間）

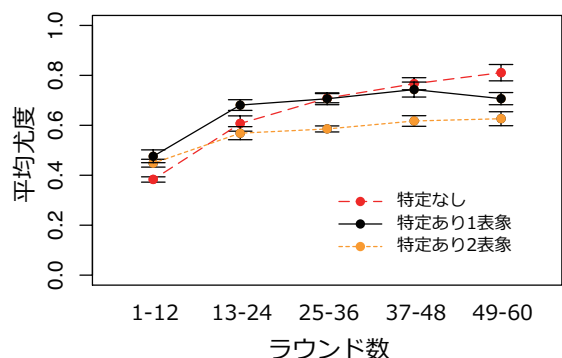


図1 成功群における計算モデルごとの平均尤度の推移

(4,72) = 15.80, $MSE = 0.002$, $p < .001$, $\eta^2 = 0.031$).
 そこでさらに最初の期間と最後の期間で単純主効果を計算すると、いずれも有意であることが確認された (初め 12 ラウンド: $F(1,18) = 41.87$, $MSE = 0.001$, $p < .001$, $\eta^2 = 0.310$. 終り 12 ラウンド: $F(1,18) = 19.98$, $MSE = 0.003$, $p < .001$, $\eta^2 = 0.365$). これは、課題の序盤においては他者の記号からその「字義通りの意味」の特定に基づいた学習 (モデルベース強化学習) が行なわれているが、終盤になって記号システムが形成されてしまえば、そういった特定は行なわれなくなる (モデルフリー強化学習へ移行する) ことを示唆する [7].

特定あり 2 表象モデルの尤度は、課題の後半において 1 表象モデルおよび特定なしモデルよりも低くなっている。これは、人間が記号の「字義通りの意味」の特定において、自分が使う記号システムと相手が使う記号システムを分けずに 1 つのシステムとして学習することを示唆している。認知負荷の観点および集団での記号システムの共有を考えればこれは自然な方略であろうと考えられる。ここで興味深いのは、特定あり 1 表象モデルを使う場合 (成功群は実際に課題の序盤で使っていると思われる)、相手と同じ記号の使い方をしなければならず、必然的に相手の記号の使い方を学習する仕組みが必要になる点である。しかも、相手の記号の使い方と自分の記号の使い方を同じにするという点で、その学習は相手の視点に立った学習になると考えられる。我々はこの仕組みが役割反転模倣 [1] によって実現できることを確認している [8].

さらにもうひとつ興味深い点は、メッセージの理解と生成が 1 つの記号システムによって行なわれると、それぞれの関数関係が逆転することである。つまり、記号システムによって自分がメッセージを送る際には、意味から記号が生成されるのに対して、相手のメッ

セージを受け取った際には、記号から意味が特定されることになるのである。このような逆転は、意味と記号の 1 対 1 の対応関係を要請すると共に、論理的な間違いを含む可能性がある。なぜなら、論理命題として p ならば q を学習したとき、そこから q ならば p を導くのは論理的には間違っているからである。この議論は、子どもの語彙学習において重要性が指摘される対称性バイアス [9] の議論に関連している。

以上の考察より、二重の意味をやりとりする仕組みは解釈学的循環を構成しても、自他を分けずに作成する記号システムの共創を早める効果があり、人間もその方略を使用しているようであると言える。ただし、この方略は記号システムが成立してくるにつれて使われなくなるようである。

謝辞

本研究は、JSPS 科研費基盤研究 (A) 「意図共有と意味創造を伴うコミュニケーション創発の進化的構成論による研究」/ 課題番号 26240037 の助成を受けた。ここに記し謝意を表します。

参考文献

- [1] Tomasello, M.: *Constructing a Language: A Usage-Based Theory of Language Acquisition*, Harvard University Press (2003)
- [2] Frith, U.: *AUTISM: Explaining the Enigma*, Blackwell (1989).
- [3] Carston, R.: *Thoughts and Utterances: The Pragmatics of Explicit Communication*, Blackwell (2002).
- [4] 津田 一郎: ダイナミックな脳—カオスの解釈, 岩波書店 (2002).
- [5] 金野 武司, 森田 純哉, 橋本 敬: 言語的コミュニケーションシステムの創発に関する実験的アプローチ, 計測と制御, Vol. 53, No. 9, pp. 801-807 (2014).
- [6] 金野 武司, 橋本 敬, 李 冠宏, 奥田次郎: 記号コミュニケーションにおける言外の意味の推論に基づく先行的行動価値修正メカニズムの役割, 日本認知科学会第 32 回大会予稿集, pp.477-486, (2015).
- [7] Daw, D N., Niv, Y., Dayan, P.: Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control, *Nature Neuroscience*, Vol. 8, pp. 1704-1711 (2005).
- [8] 金野 武司, 森田 純哉, 橋本 敬: 記号コミュニケーションシステムにおいて言外の意味を成立させるメカニズムの計算モデルによる検討, 計測自動制御学会 システム・情報部門 学術講演会 2014 (SSI2014), pp. 401-406, (2014).
- [9] Sidman, M.: Symmetry and equivalence relations in behavior, *認知科学*, Vol. 15, No. 3, pp. 322-332 (2008).