心の理論における再帰のレベルはいかに進化してきたか?

*高野 雅典　有田 隆也
名古屋大学 大学院情報科学研究科
〒 464-8601 愛知県名古屋市千種区不老町 情報科学研究科棟 有田研究室
takano@alife.human.nagoya-u.ac.jp
http://www2.create.human.nagoya-u.ac.jp/

Abstract: 心の理論とは他者の心の状態を推測する心の機能である．他者の心を推測するとき，他者も他個体に対して心の推測を行うと仮定すると，心の推測の入れ子構造が発生し，他者の心の状態や自分の他者への影響を推測できる．集団で生活する生物は，複雑な社会的環境で生ずる他者との駆け引きの中で他者よりも優位に立つことが重要であり，他者の心を推測する能力などが適応進化によって創発してきたと考えられる．この心の理論による他者の推測の再帰性に関して知見を得ることが本研究の目的である．本稿では，衝突回避行動を用いた計算論的モデルを用い，他者の心の推測の再帰性を表現し，動物から人間への心の進化における再帰の構造の進化に関するシミュレーション実験を行った．その結果，他者の心の推測をする・しないの差が再帰のレベルの適応性に影響しうること，他者との関係性の大きさが再帰のレベルの適応性の支配要因となっている可能性があることなどが示された．さらに，それらから心の理論における再帰のレベルの進化は，記憶や処理コストに関わる制約だけでなく，処理内容そのものに関わる制約を少なからず受けていることが示唆される．

1　はじめに

我々は他者と相互作用し合う環境の中で他者の行動を予測し自分の行動を決めている．さらに，自分の行動が他者に与える影響がどのようなものであるかということも考えることができる．例えば，対戦競技では，味方や敵の心を推測したり欺いたりすることが重要な要素であることは明らかである．また，人混みの中を移動するとき，意図する方向に一直線に向かうのではなく，向かってくる他者を避けることや他者が自分を避けることなどを考えて行動している場合もある．このような例にとどまらず，他者の心の状態を推測することは，社会に順応した生活を送るための条件であると考えられる．

一般に，他者の心の状態を推測できる個体は,「心の理論」を持っていると言われる．心の理論とは，ある個体が自己または他者の心の状態，目的，意図，知識，信念，思考，疑念，推測などを推測する心の機能のことである [8]．そして，心の理論により他者が心の理論を持つと推測するとき，他者も他個体に対して，心の推測を行っていると推測できる．このとき，心の推測の入れ子構造が発生し，他者がどのような状態なのか，自分が他者にどのような影響を与えるのか，を推測することができる．

集団で生活する生物は，同種の個体は同じ空間を利用し，同じ種類のものを食べ，同種の異性と交尾して子孫を残していく．そのため，自分と同種の個体は協力者であると同時に，競合相手でもある [7]．このような複雑な社会的環境で生ずる他者との駆け引きの中では，他者よりも優位に立つことが重要であった．このような，社会環境で発揮される知能のことを社会的知能と言う [2, 1]．とすると，他の個体の行動を予測することに始まり，社会的知能の一つとして，心の理論は進化してきたのではないかと考えられる [9]．そしてさらに他者の行動を操作する能力と，自分の都合の悪くなるよう操作されないように他者の心を推測する能力が互いに強化しあって進化してきたことが一つのシナリオとして考えられる [5]．つまり，社会的相互作用の中で他者の行動を予測する能力，さらに他者の予測を予測する能力などが適応進化によって創発してきたというものである．特に，高度な社会的知能を持つチンパンジーの可能な他者の再帰的な推測はレベル 2 程度と言われる [6] 一方で，ヒトはレベル 5 前後まで可能と言われており [3, 4]，他種に比べて際立って高い．

従来の心の理論に関する研究において，心の推測の入れ子構造について我々が経験的に 3 次以上の他者の心の推測も可能であると感じているのにもかかわらず，ほとんどが 2 次までを対象としていた [4]．本研究は，心の理論における再帰のレベルの進化に焦点を絞り，再帰的な推測という機能を仮定した上で，計算論的モデルを構築する．そして，シミュレーション実験により，ヒトの独自性を際立たせている心の理論の

再帰レベルの進化に関する知見を得るのが目的である．本稿では，個体が集団の中で他者の移動を再帰的な心の処理により推測し，その上で自分の移動を決定するというものを考え，冒頭で述べた人混みの中での移動という，物理的に着地した群集の「衝突回避行動モデル」を構築する．このモデルは社会的な相互作用を象徴的に表す最小モデルを狙ったものであり，群集の群れ行動をリアルに再現することが目的ではない．

2　モデル

2.1　基本設計

本研究では,「他者の心の推測とその推測に基づく行動」の再帰性を図 1 のような記述を用いて表現する．X, Y, Z の 3 者が存在し（X_n, Y_n, Z_n はレベル n の志向システム），得た情報 E に応じて行動 B を取る例を表す．ここで，E_{in} はレベル n の個体 i が得る環境情報，B_{in} はレベル n の個体 i が行う行動を表すとする．この他者の行動予測は，予測する他者の行動予測を内部で呼び出すことで再帰的に無限の深さまで記述可能である．

- レベル 0
  X_0 knows E_{x0} and takes action B_{x0} based on it.
- レベル 1
  X_1 knows
  {
  Y_0 knows E_{y0} and takes action B_{y0} based on it.
  Z_0 knows E_{z0} and takes action B_{z0} based on it.
  E_{x1}
  } and takes action B_{x1} based on it.
- レベル 2
  X_2 knows
  {
  Y_1 knows
  {
  X_0 knows E_{x0} and takes action B_{x0} based on it.
  Z_0 knows E_{z0} and takes action B_{z0} based on it.
  E_{y1}
  } and takes action B_{y1} based on it.
  Z_1 knows
  {
  X_0 knows E_{x0} and takes action B_{x0} based on it.
  Y_0 knows E_{y0} and takes action B_{y0} based on it.
  E_{z1}
  } and takes action B_{z1} based on it.
  E_{x2}
  } and takes action B_{x2} based on it.

図 1: 他者の心の推測とその推測に基づく行動の再帰表現

2.2　衝突回避行動モデル概要

衝突回避行動モデルとは，基本設計に基づき構築されたものであり，集団中における，他者の移動を再帰的な心の処理により推測し，その上で自分の移動を決定する．環境内の複数のエージェントは，スタート地点とゴール地点をそれぞれランダムに設定され，スタート・ゴール間を早く移動することを目的とする．エージェント同士が衝突すると，移動速度が低下するため，他者の行動予測に基づいた効率よい他者との衝突回避が問題となる．エージェントはそれぞれ「視界」を持ち，視界内に存在するエージェントを内部モデル化し行動予測を行う．さらにその内部モデルが他者の行動予測を行い，内部モデルの持つ視界内にエージェントが存在する場合，内部モデルが入れ子構造的に生成される．つまり他者の行動予測が再帰的に行われる．このように行った他者の行動予測に基づいて，エージェントは他者と衝突しないように，かつゴールへ近づくことができるように次の自分の行動（移動ベクトル）を決める．個体の適応度は要した時間をスタート・ゴール間の距離で割ったものとする．実験では視界は試行ごとに固定し，全エージェントで同一とした．

3　シミュレーション実験の結果

シミュレーション実験を行った結果，第一に，予測レベルの大小と回避行動の強さの間に図 2 に示すような興味深い関係があることがわかった．簡単のために，2 者間がすれ違う際の衝突回避における各レベルの予測と行動を考える．レベル 0 は他者を全く考慮しないので予測を行わずに直進する．レベル 1 は他者が全く避けない（レベル 0 行動）と予測するので，極端に他者を避ける．レベル 2 は他者が自分を避けてくれる（レベル 1 行動）と予測するので，レベル 1 よりは小さい回避行動を取って前進する．このように考えると，奇数レベルは偶数レベルにくらべて回避行動を取りやすいことがわかる．しかも，この傾向はレベルが低い時の方がより強く出るため，図 2 のような傾向が行った各実験で示された．また，レベルが高くなるにつれ行動の傾向が徐々に弱まるため，高レベル同士の衝突回避では予測と実際の行動の誤差が小さくなる．

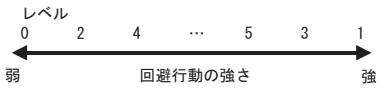

図 2: 予測のレベルと避ける力の強さの関係

第二に，図 2 の軸における各レベルの行動傾向の適応性は視界に依存することが示された．視界が狭いと，至近距離から衝突回避を始めることになるため，回避行動を強く取らないと衝突してしまう．そのため，回避行動の強い個体（レベル 1, 3, 5）のほうが適応度が高くなる．逆に，視界が広いとある程度離れた状態から衝突回避を始めるため，回避行動を強く取ると，他者を避けすぎることになり，効率が悪い衝突回避行動になってしまう．そのため，あまり回避行動を行わない直線的に移動する個体（レベル 0, 2, 4）のほうが適応度が高くなる．そして，両者の境界の領域では，適度な回避行動，つまり，奇数にせよ偶数にせよレベルが高いほうが適応度が高くなる．興味深いことに，視界の広さがこの中間領域にある場合に適応度が最も高くなった．

4　議論

ここでは，心の理論における再帰レベルの基本的性質や，ヒトの有する心の理論への進化過程に関して，本実験の結果が意味することや示唆することを議論する．

まず，奇数レベルと偶数レベルの差はレベル 0 とレベル 1 の質的なギャップに起因すると言うことができる．他者の心の状態の推測が起点となって，その推測から自分の行動を決める奇数レベルと，自分の心の状態を他者が推測することが起点である偶数のレベルではその結果に本質的な違いが生ずるのであると考えられる．特に，奇数レベルの予測は,「他者が『他者を意識しない自身の欲求だけを考えて行動をする』」という予測に基づいている．そのため，相手の欲求に沿った行動に自分自身の行動を合わせがちにである．それは，自分自身の欲求に沿った行動を取りづらくなることを意味する．

また,「視界の広さ」は，他者の行動が自分に影響を及ぼすような環境において，自分の行動を決定する際に，どの程度，相手の行動を想定するかという，予測における相互作用の影響の大きさを決めるパラメータである．視界が狭いとき，つまり，利害関係がぶつかりやすいとき，あるいは，他個体のことに対する配慮の小さい集団における社会的関係においては，奇数の志向性のレベルの推測が他者との衝突をうまく避けることができ，適応的であると考えられる．逆に，社会的相互作用が強い時，偶数の志向性のレベルの推測をすると自分自身の目的に近い行動を取りやすくなり，適応的であると言える．そして，相互作用の強さが両者の境界にあるとき，奇数偶数に関係なくより高いレベルが高い適応度が得ることができる．

このようなことから，再帰のレベルの進化について，次のようなことを考えることができる．

- 他者との関係性（相互作用）が弱い段階では，レベル 1 が適応的となり高いレベルへは進化しない．

- 他者との関係性（相互作用）が強い段階においては，適応的なレベルは 0, 2 であり，高いレベルへは進化していかない．
- 他者との関係性が強すぎず，弱すぎない中間の特定の領域に達したとき，レベルの奇数，偶数の差が最も小さくなり，さらに高いレベルへの選択圧が生じると考えられる．高いレベルへの進化は，他者との関係性，あるいは相互作用が大きすぎず，小さすぎない，特定の中間領域にあるときに，レベルの奇数，偶数の適応度差が小さくなり，同時に高いレベルへの選択圧が生じるときに可能となる．
- ヒトの歴史において，個体間の相互作用の強さは徐々に強くなり，それが知能の進化を駆動してきた [4] と考えると，有利なレベルは，図 2 において，右から左へと変わってきたと考えられる．ヒトと他種との差は，この特定の領域に達したかどうかの差である可能性がある．
- 高いレベルへ進化していくと考えられるこの特定の中間領域とは，集団の適応度が最も高くなる「理想的」な状態である可能性がある．

5 おわりに

構成論的手法に基づいて，動物から人間への心の進化における再帰の構造の進化に関する計算論的モデルを構築し，シミュレーション実験を行った．他者の心の推測をする・しないの差が再帰のレベル，特にその偶数・奇数レベルの適応度の差に影響しうること，他者との関係性の大きさが再帰レベルの適応性の支配要因となっている可能性があることなどが示された．人間の独自性を際立たせているもののひとつである心の理論における再帰の深さの進化は，記憶や処理コストに関わる制約だけでなく，処理内容そのものに関わる制約を少なからず受けていることが示唆される．

参考文献

[1] R Byrne. *Thinking Ape.* Oxford University Press, Oxford, 1995. (小山 高正, 伊藤 紀子 訳 (1998). 『考えるサル』. 大月書店.).

[2] R Byrne and A Whiten, editors. *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans.* Oxford University Press, Oxford, 1988. (藤田 和生, 山下 博志, 友永 雅巳 監訳 (2004). 『マキャベリ的知性と心の理論の進化論』. ナカニシヤ出版.).

[3] D Dennett. *The Intentional Stance.* MIT Press, Cambridge, 1987. (若島 正, 川田 学 訳 (1996). 『志向姿勢の哲学』. 白揚社.).

[4] R. Dunbar. On the origin of the human mind. In P. Carruthers and A. Chamberlain, editors, *The Evolution of Mind*, pp. 238–253. Cambridge University Press, 2000.

[5] J. R. Krebs and R. Dawkins. Animal signals: Mindreading and manipulation. In *Behavioural ecology: An evolutionary approach*, pp. 380–402. Blackwell, Oxford, 1984.

[6] 松沢哲郎. チンパンジー・マインド–心と認識の世界–. 岩波書店, 1991.

[7] 小川秀司. 霊長類における同種間での他者理解の進化. 心理学評論, Vol. 40, No. 1, pp. 22–38, 1997.

[8] D. Premack and G Woodruff. Does the chimpanzee have a theory of mind? *The Behavioral and Brain Sciences*, Vol. 1, pp. 515–523, 1978.

[9] C Zimmer. How the mind reads other minds. *Science*, Vol. 300, No. 16, pp. 1079–1080, 2003.