

# 逐次手番ゲームにおける心の理論の再帰レベルの影響

## Effects of Recursive Estimation with a Theory of Mind in a Sequential-Move Game

名古屋大学 大学院情報科学研究科 ○金井 裕史 有田 隆也

○Hiroshi KANAI Takaya ARITA

Graduate School of Information Science, Nagoya University

**Abstract:** An individual having a Theory of Mind (ToM) can read the minds of others. If we assume that individuals with a ToM also consider others and themselves to have a ToM, then there should be a recursive structure here. The purpose of our study is to investigate the property of recursive structure, based on a constructive approach. To achieve it, we constructed a computational model in which two agents take turns deciding next move based on recursive estimation in a two person sequential move game. The results of the computer experiments showed that deep recursive estimation is required to obtain a high payoff when the characteristic of the game is intermediate between competitive and cooperative.

### 1. はじめに

人間は他者の心を読み、行動を推測して自分の行動を決めている。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になくはない能力と考えられている。

一般に他者の心の状態を推測することのできる個体は「心の理論」を持っていると言われている。心の理論は1978年にPremack[3]がチンパンジーに関する研究の中で提案した概念であり、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、ふり・好みなどを推測する心の機能のことである。心の理論により他者も心の理論を持つと推測するとき、他者も他者に対して心の推測を行っていると言われている。この時、他者の心の推測の推測といった心の推測の入れ子構造が発生し、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。ヒトはこの入れ子構造の深さ(再帰レベル)に関して、他の

生物よりも深く考えることができる。

Premackによる提唱以来、心の理論に関する研究は盛んに行われてきた。Wimmerらは誤信念課題を用いてヒトが4歳以降に心の理論を獲得していることを明らかにしている[4]。またヒト以外の動物についても様々な研究が行われてきており、サル類や類人猿でもヒトほどではないが心の理論を持つという主張[6]や、心の理論の獲得に必要なと考えられている他の個体の視線を追う能力をカラス[7]などの動物も持っていることが示されている。心の理論に関する計算論的アプローチとしては高野らの研究が挙げられる[10]。高野らは生物の再帰的な予測による衝突回避行動を扱った計算論的モデルを構築し、再帰レベルの適応進化に焦点を絞った計算機実験を行った。実験結果から、心の理論の再帰レベルにおいて他者の心の状態の推測が起点となっている奇数レベルと、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レベルでは適応度に違いをもたらす可能性や、社会における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る可能性を示した。

複雑化した社会環境に適応するためには高度な知的能力が必要であり、その適応進化によって高度な知的能力を獲得したとする社会的知能仮説という考えがある[5]。心の理論もこの文脈で説明可能であろう。ヒトが形成する複雑化した社会環境では自己の行動で他の個体の行動を操作して利用する、逆に自分の都合が悪くなるようには操作されないといった駆け引きが存在し、その中で他の個体よりも優位に立つために他の個体の心を読んで行動を予測する能力は重要である。そして、複雑化した社会環境での心の理論の創発というシナリオにおいて、社会的相互作用の性質とそこにおける心の理論の再帰レベルの適応性には深い関係性が存在しているのではないかと推測される。

以上を踏まえ、単純化した計算論的モデルに基づく計算機実験によって、社会相互作用の性質とそれに適応的な再帰レベルの関係を解明することを本研究の目的とする。2 者間の相互作用のあり方は両者の対立度に強く依存するはずである。本研究では、対立度を協調から競合まで連続的に変えることが可能である逐次手番ゲーム(stackelberg ゲーム)[10]を用いて、社会的相互作用を表現する。

## 2. モデル

### 2.1. 逐次手番ゲームの流れ

本研究で扱う逐次手番ゲームの流れをFig. 1に示す。まず、ノード(I)で先手のエージェントAがstay又はswitchを選択する。もし、stayを選択すると、ノード(I)の利得が両者に入り(同図では3と1)終了し、switchを選択すると、次のノード(II)に移動し、後手のエージェントBがstay又はswitchを選択する。もし、stayを選択するとノード(II)の利得が両者に入りゲームは終了し(同図では1と3)、switchを選択すると、次のノード(III)に移り、今度はエージェントAの選択の番となる。以後、同様にどちらかのエージェントがstayを選択するまで各エージェントが交互に選択を繰り返す。あるいは、switchが続いて終端ノードに到達した

場合は終端ノードの利得を両者が得てゲームは終了する。

高い利得を得るためには相手の行動選択を予測する必要がある点がゲームの特徴である。例えば、Fig. 1ではエージェントAはノード(II)におけるBの行動決定を予測する必要がある。Bがノード(II)でstayを選択するならば、Aはノード(II)よりノード(I)の方が高い利得なのでノード(I)でstayを選び、Bがノード(II)でswitchを選択するならばノード(I)よりノード(IV)の方が高い利得なのでノード(I)でswitchを選ぶ。このように、逐次手番ゲームは相手の行動を予測することが重要であり、ヒトの再帰的な推論能力の検証に用いられてきた[1][2]。

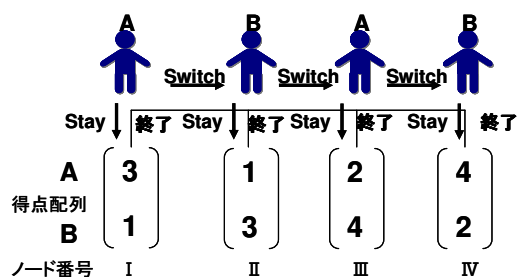


Fig. 1: 逐次手番ゲームの流れ

### 2.2. 利得の決定方法

利得の基本的な設定として、ゲームの意思決定ノード数を  $N$  としたとき、1 から  $N$  までの整数値を各エージェントが各ノードで得る利得に重複無く割り当てることとする。この割り当てが2 者間の対立度を定める。具体的には、2 者間の対立度は各ノードのAとBの利得の相関係数として理解できる。例えば、相関係数が1ならば全てのノードで両者の利得が等しくなって完全に協調的なゲームとなり、逆に相関係数が  $-1$  ならば全てのノードで両者の利得の和が同じになるので完全に競合的なゲームとなる。

ただし、同じ相関係数の値をとるノードの組であってもノードが並ぶ順序によってはゲームの結果は異なる。そこで、本モデルでは、並べ方の影響を排除するために、相関係数によって決定した各ノードのAとBの利得の組を1つのセットとし、ノード数分の全組み合わせ(意思決定ノード数を  $N$  として  $N!$  回行う)の得点配列でゲームを行って平均を取ることにした。以

下、全組み合わせのゲームを行うことを 1 step 回ゲームを行うと呼ぶ。

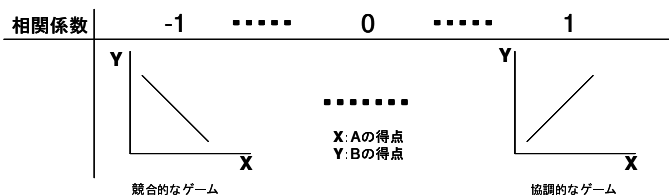


Fig. 2: 相関係数とゲームの性質の関係

### 2.3. エージェントの設定

エージェントは全てのノードの利得を知っているものとする。エージェントの行動決定までの処理は、1) 再帰的推測による相手の行動選択確率の決定、2) その選択確率に基づく自分の行動の選択確率の算出と stay 又は switch の決定、の 2 段階である。

#### 1) 再帰的な推測とその行動決定

再帰的な推測は Dennett の志向姿勢の考えに基づいて階層的に定義した。志向姿勢とは他の個体の行動を解釈する方法論で、他の個体を信念や欲求などを考慮して主体的に活動する合理的な活動体として扱い、他の個体の内的状態を階層的に表現している[9]。再帰的な推測の深さとそれに対応する行動の決定方法を以下のように定義した。

- ・ 0-level: 相手の行動を予測できない。単純に 50% の確率で switch 又は stay を選択する。
- ・ 1-level: 相手が 0-level で switch 又は stay の選択確率を決定すると予測する。その相手の行動選択の確率を用いて、switch を選択した時の利得の期待値と stay を選択した時の利得を使って switch と stay の選択確率を決定し、その確率に従って行動を決定する。
- ・ 2-level: 相手が 1-level で switch 又は stay の選択確率を決定すると予測し、switch を選択した時の利得の期待値と stay を選択した時の利得を使って switch と stay の選択確率を決定し、その確率に従って行動を決定する。

3 以上のレベルも同様に定義する。

#### 2) 行動の選択確率の決定方法

stay を選択した時の利得と switch を選択した時の利得の期待値からそれぞれの選択確率を以下の式から決定する。

$$P_{st} = \frac{\exp(f_{st}/T)}{\exp(f_{st}/T) + \exp(f_{sw}/T)} \quad (1)$$

$$P_{sw} = 1 - P_{st} \quad (2)$$

ここで、 $P_{st}$  は stay を選択する確率、 $P_{sw}$  は switch を選択する確率、 $f_{st}$  は stay を選択した時の利得、 $f_{sw}$  は switch を選択した時の利得の期待値、 $T$  は任意の定数である。 $T$  は選択のランダムさを調整する定数で、大きい値ほどランダムに選択するようになる。

行動の選択確率の算出方法は終端ノードに一番近い自分が行動決定するノードの行動選択確率から順々に決定していき、最終的に実際に行動決定するノードの選択確率を決定する。意思決定ノード数を  $N$  とした場合の具体的な手順を以下に示す。

- ① 最初に終端ノードに一番近い自分が行動決定するノードの行動選択確率を求める。そのノード番号が  $N-1$  の場合 (1) の式における  $f_{st}$  にノード番号  $N-1$  の利得  $f_{sw}$  にノード番号  $N$  の利得を代入して、選択確率を求める。ノード番号が  $N-2$  ならば  $f_{st}$  はノード番号  $N-2$  の利得で、 $f_{sw}$  をノード番号  $N-1$  の相手の行動確率を用いて求め、行動選択確率を決める。
- ② ①で確率を求めたノードより 2 個前のノードの行動選択確率を求める。このノードより先のノードの行動確率は①の結果と相手の行動確率を用いることで switch を選択した時の得られる利得の期待値を求めることができ、(1) と (2) の式を用いて行動の選択確率を求める。
- ③ ②で確率を求めたノードのさらに 2 個前のノードの選択確率を②と同じ方法で求める。

以上を繰り返して、行動を決定するノードまで戻り行動選択確率を算出する。

### 3. 実験

#### 3.1. 設定

以上のモデルを用いて実験を行った。基本設定として、意思決定ノード数は 7、 $T=0.1$  とした。相関係数は  $-1$ 、 $-0.5$ 、 $0$ 、 $0.5$ 、 $1$  を用いた。また 1 試行で 100 step ゲームを行うものとする。相関係数の各設定で 10 試行実験を行い、平均利得を比較した。

#### 3.2. 結果

##### 1) 高い利得を得たレベル

再帰レベルが 0 のエージェントAを以下ではA0と呼ぶ。A0 とB0、A0 とB1、...と全ての組み合わせのゲームを行い、どの再帰レベルの組み合わせが高い利得を得たかを調べた。相関係数 1 (Fig. 3 (i))、相関係数 0 (Fig. 3 (ii))、相関係数  $-1$  (Fig. 3 (iii))の結果を示す。各グラフのX軸はAの再帰レベル、各棒グラフはBの再帰レベル、Y軸は1 stepのAの平均利得を表す。

3 つのグラフに共通して見られる傾向として、A のレベルがBのレベルに対して1だけ高いレベルの場合、他のレベルの組み合わせより高い利得を得ていることがわかる。これはAがBの行動を正確に予測できているためである。

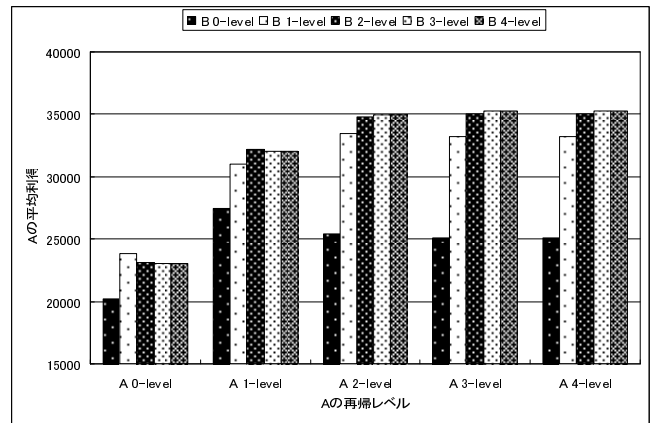
相関係数 1 と 0 の結果である Fig. 3 の (i)と(ii)を見ると、A と B のレベルが共に 0 の時の利得より共にレベルが 1 の時の利得が高く、共にレベルが 1 の時の利得よりも共にレベルが 2 の時の方が利得は高い。A 又は B のレベルが高いほど、高い利得を取っていることが分かる。

一方で、相関係数  $-1$  の結果である Fig. 3 (iii)を見ると、B0やB1の時のAの利得とB3やB4時のAの利得を比較するとB0やB1の場合の方が高い。相関係数  $-1$  ではBのレベルが低いほどAの利得が高くなる傾向があることがわかる。

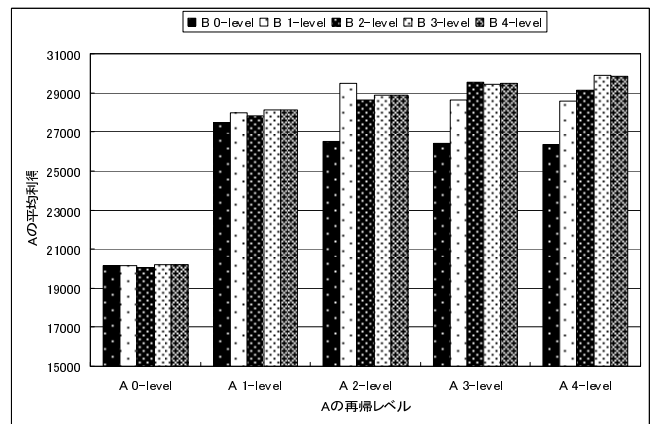
##### 2) ナッシュ均衡解との比較

逐次手番ゲームは、理論的に後ろ向き帰納法でナッ

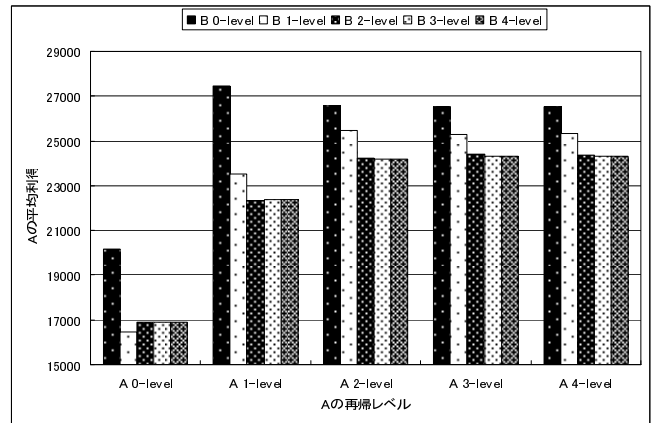
シュ均衡解が求まることが知られている。ナッシュ均衡解とはゲームとは理論における解の一種で、戦略を



(i) 相関係数 1



(ii) 相関係数 0



(iii) 相関係数  $-1$

Fig. 3: 1 step の A の平均利得

変更する誘因を持たない互いに最適な戦略のことである。戦略がナッシュ均衡解の時のゲームの利得を指標として用いた。Table 1に、1 step中にAとBが同一レベル場合のAの利得とナッシュ均衡解のAの利得が0.5 以上の差があるゲーム回数を示す。同表より、相関係数

がどの場合でも、レベルが高くなるほどナッシュ均衡解と異なる利得を取るゲーム回数が減少すること、つまりレベルが高くなるほどナッシュ均衡解と同じ行動を選択するようになることがわかる。また、相関係数0では、相関係数1 や-1 に比べてナッシュ均衡解と同じ行動を選択するようになるには高いレベルが必要なことがわかる。詳しく調べたところ、相関係数1 と-1 ではA4 B3の場合、相関係数0 ではA6B5の場合にナッシュ均衡解と同じ行動を選択するようになることがわかった。相関係数0 のような、協調と競合の中間の環境では、レベル2 や3 ではナッシュ均衡解のような合理的な行動選択をすることは容易ではなく、ナッシュ均衡解と完全に同じ行動を選択するには高いレベルが必要なことがわかる。

相関係数-1 で B のレベルが低い時に A が高い利得を取れた理由は、ナッシュ均衡解の利得を使って以下のように説明可能である。Table 2 は、A6B1 の場合における、ナッシュ均衡解よりも 0.5 以上 A の利得が高かったゲーム回数(+の段に示す)と 0.5 以上低かったゲーム回数(-の段に示す)を示す。これまでの解析より A6 はナッシュ均衡解と同じ行動選択をしていることが分かっているため、同表は B がナッシュ均衡解と異なる行動を選択した時の A の利得への影響を表していると考えられる。同表より、相関係数-1 の時ではナッシュ均衡解より高い利得を取ったゲーム回数は 725 回、低い利得を取ったゲーム回数は 0 回である。つまり、B がナッシュ均衡解と異なる行動を選択することで A はナッシュ均衡解より高い利得を得ていることがわかる。相関係数-1 では完全に利害が異なる利得の並びをしているので、B はナッシュ均衡解と異なる低利得の行動を選択すると、A がナッシュ均衡解より高い利得を取ることが期待できるためである。逆に、相関係数 1 ではすべてナッシュ均衡解より低い利得になることが分かる。相関係数 1 では完全に利害が一致する利得の並びをしているので B がナッシュ均衡解と異なる行動を選択すると A もナッシュ均衡解より低い利得を得ることになるからである。

次に、各ゲームでのナッシュ均衡解の A の利得に対する各レベルの A の利得の割合を求め、相関係数の変

化と共にその推移を示す(Fig. 4)。相関係数-1 又は-0.5 の時の A1B0や A2B1 の場合は100 を超えており、Table 1: 1step でナッシュ均衡解と利得が異なったゲーム回数

	レベル1	レベル2	レベル3	レベル4	レベル5	レベル6
相関係数 1	2954	458	18	0	0	0
相関係数 0	1745	734	260	60	16	0
相関係数 -1	1954	195	9	0	0	0

Table 2: B がナッシュ均衡解と異なる行動を選択した時の A の利得への影響(A6、B1 の時に 1step でナッシュ均衡解より高い利得又は低い利得だったゲーム回数)

	相関1	相関0	相関-1
+	0	198	725
-	1408	577	0

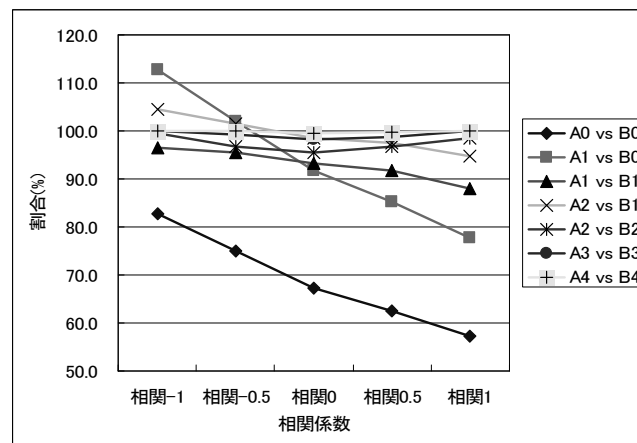


Fig. 4: 各相関係数でのナッシュ均衡解の A の利得に対する各レベルの A の利得の割合

そこから相関 1 に近づくとしたがって値が下がっていくことが分かる。つまり競合的なゲームでは相手が低いレベルの場合にナッシュ均衡解よりも高い利得をとることができるが、相関係数 0、0.5、1 と協調的なゲームになっていくにつれて得られる利得はナッシュ均衡の利得よりも低くなる。相関係数 0、0.5、1 ではナッシュ均衡解と同じ利得を得るレベルの時が最も高い利得を得るレベルということになる。相関係数 1 では A4 B3、相関係数 0 では A6 B5 の場合にナッシュ均衡解と同じ利得をとることから、相関係数 0 では最も高

い利得を得るためには相関係数 1 の場合よりも高い再帰レベルが必要だということがわかる。

#### 4. おわりに

本研究では逐次手番ゲームを用いて社会的相互作用の性質と心の理論における再帰レベルの関係を単純化した計算論的モデルを構築し、実験を行った。競合的環境では高い利得を取るには相手の再帰レベルが低いこと、協調的環境では両者共に再帰レベルが高いことが重要だということが示された。競合的と協調的の中間的な環境では、高い利得を得るために必要な再帰レベルが他の環境より高いことも示された。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高い再帰レベルを獲得した状況は、競合的環境と協調的環境の中間にあたる環境である可能性を実験結果は示唆している。

#### 参考文献

- [1] Hedden, T. and Zhang, J.: What do you think I think you think? Strategic reasoning in matrix games, *Cognition*, Vol. 85, pp. 1–36 (2002).
- [2] Colman, A.: Depth of strategic reasoning in games, *Trends in Cognitive Sciences*, Vol. 7, pp. 2-4 (2004).
- [3] Primacy, D. and Woodruff, G.: Does the chimpanzee have a theory of mind?, *The behavioral and Brain Sciences*, Vol. 1, pp. 515-526 (1978).
- [4] Wimmer, H. and Perner, J.: Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception, *Cognition*, Vol. 13, pp. 103–128 (1983).
- [5] Byrne, R. and Whiten, R. (Eds.): *Machiavellian Intelligence*, Oxford Univ. Press, Oxford (1988).
- [6] Premack, D.: 'Does the Chimpanzee Have a Theory of Mind?' Revisited, *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans*, Clarendon Press, Oxford, pp. 160-179 (1988).
- [7] Bugnyar, T., Stowe, M. and Heinrich, B.: Ravens, *Corvus corax*, follow gaze direction of humans around obstacles, *The Royal Society*, Vol. 271, pp. 1331-1336 (2004).
- [8] Dennett, D.: *The Intentional Stance*, MIT Press, Cambridge (1987).
- [9] Osborne, M. J. and Rubinstein, A.: *A course in game theory*, MIT Press, Cambridge (1994).
- [10] 高野雅典, 加藤正浩, 有田隆也: "心の理論における再帰のレベルの進化に関する構成論的手法に基づく検討", 認知科学, Vol. 12, No. 3, pp. 221-233 (2005).