

動学的意思決定状況における Rough Reasoning のモデル化 ムカデゲームへの適用

今野 直樹

Naoki Konno

東京工業大学大学院 価値システム専攻

Department of Value and Decision Science, Tokyo Institute of Technology

本発表では展開形ゲーム状況においてプレイヤーの推論の誤りを記述するモデルを提案する。人間の意思決定は様々な箇所エラーが起こる可能性がある。特にこれまで扱われてきた行動の際のエラーだけではなく、推論の際にもエラーが起こっていると考えられる。この推論に関するエラーの性質を明らかにするために動学的な意思決定状況に注目し、そこで粗い推論のためにエラーが発生する確率が、結果の利得の差とどれだけ先の選択なのかという木の長さの二点に依存すると仮定することにより、推論エラーの影響を分析するためのモデルを提案する。さらにこのモデルをムカデゲームに適用することにより、推論能力の不完全性もまたムカデゲームにおいて Nash 均衡からの逸脱が観察される一因であることを示す。最後に適度な合理性が高い頻度での非合理的行動や協力行動を示す可能性をこのモデルにより示す。

Keywords: ムカデゲーム, Quantal Response, 社会厚生, 限定的な推論

1. はじめに

研究プロジェクト課題「エージェントベースモデリングとその応用」において、より現実的なエージェントのモデル化という観点から、特にエージェントの推論能力の不完全性に注目することによりモデル化を行う。さらに本モデルを、従来の理論と実験結果との乖離が指摘されてきた典型的な例であるムカデゲームに適用することにより、推論能力の不完全性という要因もまたナッシュ均衡からの体系的な逸脱を引き起こす一因であることが明らかにする。さらにこの結果の現実的な意味として、適度な合理性を持った主体のほうがより非合理的な主体よりも、高い頻度で非合理的な行動をする確率が高くなる場合がある一方、合理的な主体からなる社会よりもかえって高い社会厚生を達成しうることを指摘する。

古くは Selten[2]、Myerson[3]らが均衡の精緻化という視点から、より起こりそうな均衡を選ぶために、人間のエラーに注目した均衡のふるいわけを行っている。このように伝統的なゲーム理論では、各プレイヤーは状況を完全に認識し、全ての状況に対し、エラーなくそれを比較検討し、最

適な行動を選択できるとして、モデル化するものが大半だった。しかしながら、合理的なエージェントに対する解概念が、しばしば直観的に奇異に映る結果を導くこともあることが指摘されている[2]。

このように均衡からの逸脱が観察される理由としては様々なものが考えられる。本発表では完全情報ゲームに焦点を当てるが、囲碁やチェスといったこのクラスに属するゲームは、理論的には、後ろ向き帰納法により均衡戦略を求めることができることが知られている。しかしながら現実にはそれを実行することは不可能である。また実際のプレイヤーもそのような思考はしていない上、ゲームソフトのプログラムでも後ろ向き帰納法をそのまま使っているのはごく一部だけであるのが現状である。

また今日の夕食を何にしようか？といった日常的な意思決定を思い起こしてみても、どこの店に入るのかといった段階で、すでに無限に近い選択肢が理論的には存在するはずである。しかしながら実際には有力な代替案やシナリオ数個を想定し、それを比較検討するといったプロセスで意思決定が行われることが多いと想像される。

このように一般の意思決定状況は、有力なシナリオをいくつか推論によって選び出し、最後にそれらの中で選択を行うという、推論と、決定の二つのプロセスを持っていると考えられる。決定プロセスにおける人間の限定互応理性の表現として McKelvey and Palfrey[4] は、戦略形ゲーム状況において、プレイヤーは「利得が高い戦略ほど高確率で選択したがるが、最適な戦略を確率 λ で選ぶことはできない」というモデルを提案し、そのエラー率の関数が各プレイヤー間の共有知識であるならば、均衡が存在することを示した。

これに対し推論プロセスはプロセス自体のモデル化が困難であるため、完全な厳密さが求められるわけではないエラーも起こしうる普通の意味決定に関する数理的なモデル化はこれまでほとんど扱われてこなかった。今回は推論プロセスが最も扱いやすい完全情報の展開形ゲーム状況に注目することにより、粗い推論プロセスによる意思決定のモデル化を試みる。本モデルでは各プレイヤーは後ろ向き帰納法により、目の前のノードを評価すると仮定する。その際のエラーの起こる確率については、利得に関しては戦略形の場合と同様に、高い利得を導く行動を最適応答と認識する可能性が高いという仮定は自然なように思える。それに加えて、遠い先の判断ほど誤りやすい、すなわち木の深さにも推論精度は依存すると考えられる。本稿ではこれら二つの要素に注目し、限定的な推論による意思決定のモデルを提案する。

次に、従来の理論と実験結果が異なることで知られている代表的なゲームである、ムカデゲーム[1]にこのモデルを適用する。ムカデゲームは二人が交互に協力または終了のどちらかを選択する有限二人完全情報ゲームである。協力を選ぶと自分の利得が低下するものの、相手の利得はそれ以上に増加する。従ってムカデゲームは協力が望まれるような社会状況の一種のモデルと考えることができる。この状況では両者が協力を繰り返せば、パレートの二人にとってより望ましい状態を導くことができうる。しかしながら、このゲームには最終期が存在するために、最終期では終了戦略を選択することが合理的な行動になる。よって相手が合理的なプレイヤーならば最終期では終了行動が予想される。このことから最終期より一期前の時点でそこで意思決定を行うプレイヤーにとって、協力行動を選択したほうが高い利得を得られると予想されるため、最終期よ

り一期前の時点で終了行動を選択することが合理的な行動となる。このような推論を繰り返すことにより、一期目でいきなり終了行動を選択するという結果がこのゲームの均衡解による結果として導かれる。

しかしながら実際に被験者を用いた実験では、かなりの割合で協力行動が採用されることが報告されている[6]。この結果は、ゲームの構造を両者がしっかりとした形では認識していないこと[7][8]や利他主義者がある程度の割合で存在するという要因[6]によって説明されてきた。しかし限定的な推論能力もまた、このような協力行動を発生させる一因であることをシミュレーションを用いて確認する。

さらに推論能力のレベルが変わったときに、達成される結果の社会厚生はどのように変化するのかということについて考察を行う。

2. 完全情報の展開形ゲーム

本稿では真の客観的ゲームが完全情報の有限展開形ゲームの場合のみ扱う。完全情報の有限展開形ゲーム状況は $G = (I, N, A, \alpha, P, r_i)$

で与えられる。ここで、

I : 意思決定主体の集合

N_T : 終点ノードの集合

N_D : 決定ノードの集合

$N = N_T \cup N_D$: ノード集合

A : 行動の集合

$\alpha: N - \{n_1\} \rightarrow N_D$: 最初のノード n_1 以外のノードに対し、その前のノードを対応させる関数

P : 決定ノードに対し、そのノードで意思決定を行うプレイヤーを対応させるプレイヤー関数

$r_i: N_T \rightarrow \mathbb{R}$: 各終点ノードに対しプレイヤーの利得を割り当てる利得関数

3. 限定的な推論モデル

完全情報の客観的ゲームを前にした限定的な推論能力を持つプレイヤーは本節で定義するようなヒューリスティクスに従い意思決定を行うと仮定する。

まず以下の用語を定義する。

N_2 : n_1 から到達できるノードの集合

N_m^1 : 最後の決定ノードの集合

$a(n)$: ノード n から n_1 までの深さ

: 推論パラメーター. 非負の実数値をとる. 値が大きいほど推論能力が高いことを意味する. また利得の単位に依存する.

$r(n_d) = (r_1(n_d), \dots, r_j(n_d), \dots, r_I(n_d))$:

n_1 で選択を行うプレイヤー i_1 にとって, ノード $n_d \in N_D - \{n_1\}$ のあと各プレイヤーが最適に行動したときに達成されると予想する利得ベクトル.

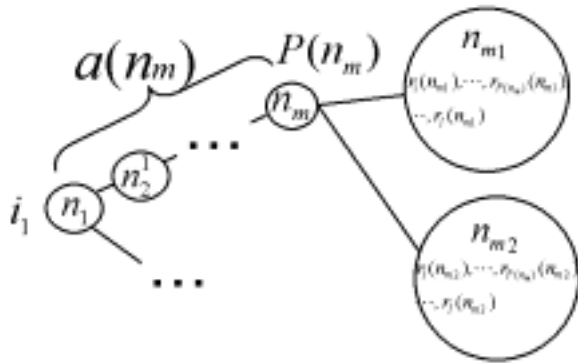


図1 i_1 の推論

このとき各プレイヤーは,

- (1) プレイヤー i_1 は各ノード $n_2 \in N_2$ に対する仮想利得を求めため, 後ろ向き帰納法を試みる. ここで $n_2 \in N_T$ の場合は, その利得を正しく認識していると仮定する. $n_2 \in N_D$ なら最後の決定ノードから彼の推論能力の範囲での後ろ向き帰納法を試みる.
- (2) $n_m \in N_m^1$ に対し, $r(n_m)$ として $r(n_{m1})$ の値が割り当てられる確率が以下のように与えられると仮定する.

$$e^{\frac{r_{P(n_m)}(n_{m1})}{a(n_m)}\sigma} / \sum e^{\frac{r_{P(n_m)}(n_{mj})}{a(n_{mj})}\sigma}$$

これだけではイメージが難しいと思うので, 図2のようなゲームを考える. ここで, プレイヤー1は自分が右の戦略を採用して, プレイヤー2の決定ノードになった場合どのような利得が達成されるのかということに関しては, 確率

$$\frac{e^\sigma}{1+e^\sigma}$$

で最適反応である下の結果である(0,3)という利得の組が達成される. 一方

$$\frac{1}{1+e^\sigma}$$

という確率で最適反応の結果ではない(3,2)という利得の組を生み出す結果が達成されると予想する.

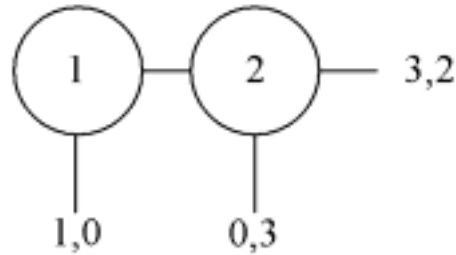


図2 ゲーム例1

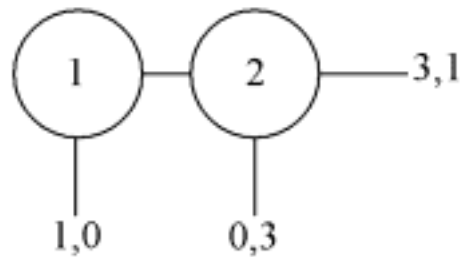


図3 ゲーム例2

ここでもしゲームが図3のように2の決定ノードにおいて, 最適なものとそうでないものとの差が大きい場合には, 最適な結果が起こると正しく認識する確率は

$$\frac{e^{2\sigma}}{1+e^{2\sigma}}$$

となり, これは図2の際の

$$\frac{e^\sigma}{1+e^\sigma}$$

より大きい. このようにこのモデルでは利得差が大きいほど推論で誤る可能性が低いと仮定している.

次に, もしゲームが図4のようにより深いところから推論する必要がある場合を考える. このときプレイヤー1が,

プレイヤー3の決定ノードに到達した場合にそこで3にとって最適な行動の結果である下の結果が到達されると確率

$$\frac{e^{\frac{\sigma}{2}}}{1 + e^{\frac{\sigma}{2}}}$$

で推論する.これは図2の

$$\frac{e^{\sigma}}{1 + e^{\sigma}}$$

と比較すると低い精度になっている.

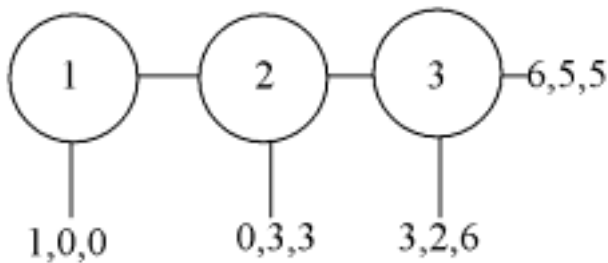


図4 ゲーム例3

なお σ が大きくなるほど推論精度が高くなっていることがこれらの例からも確認できる.また利得の表示単位が変われば, σ もそれに応じて変動するという,単位に関する情報も含んだパラメーターでもある.

(3) すべての $n_m \in N_m^1$ に対してこのような手順で仮想利得が与えられたならば,次に n_m を終点ノードとみなして新たにもう一段階前にある決定ノードに対し同様の operation で仮想利得を割り当てる.

(4) すべての $n_2 \in N_2$ に対して仮想利得が定まったならば,1 は,その中で自分にとって最もよいと考えるものを選択する.

(5) 1 の行動の結果到達したノードにおいて意思決定を行うプレイヤー i_2 は i_1 の推論とは独立に,同様のルールで自分の意思決定を行う.

(6) このようなルールに従い終点ノードに至るまでゲームがプレーされる.

このモデルにおける分析結果として,我々はどの結果がどのような確率で起こるのかということに関する終点ノ

ード上の確率分布が得られる.

4. ムカデゲームへの適用

次にこの意思決定が,従来の均衡戦略による行動結果とどのような違いがあるのかムカデゲームに適用し,確認を行う.ムカデゲームは,完全情報の二人有限展開形ゲームである.

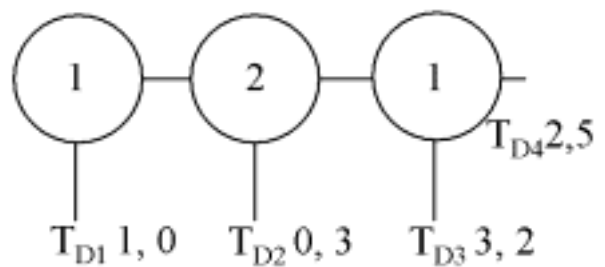


図5 3階ムカデゲーム

お互いが交互に Pass (P) または Take (T) の2通りの行動のどちらかを選択する.(P)を選択した場合は,自分の利得が減少する一方相手の利得はそれ以上に増加する.(T)を選択するとその時点で得ている利得のままゲームが終了する.

このゲームは様々な利得構造で研究されてきているが本稿では[1]に従い,(P)を選択した場合は,自分の利得が1減少する一方相手の利得は3増加すると仮定する.なお決定ノードが n 個あるようなムカデゲームを n 階ムカデゲームと呼ぶ.なお1が最初に(T)を選んでゲームが終了したときのノードを T_{D1} ,一回目で1が(P)を選択したあと,2が(T)を選択したときのノードを T_{D2} とする.

ムカデゲームは有限回で終了するため,後ろ向き帰納法により,全ての決定節において両者が,(T)を選択し,ゲームは一回目で終了するという行動戦略の組が,唯一の部分ゲーム完全均衡となる.一方被験者を用いた実験では少なからぬ割合で協力的行動が観察されていくものの,協力的行動の発生率はゲームが終わりに近づくに従い高くなることが報告されている[6].

そこでまず一回目でゲームが終わる確率がどのようになるのかということについて、ゲームの規模である n と推論能力である σ を変化させシミュレーションを行ったものが図6である。なお FCTF は一回目でゲームが終了する頻度を表している。

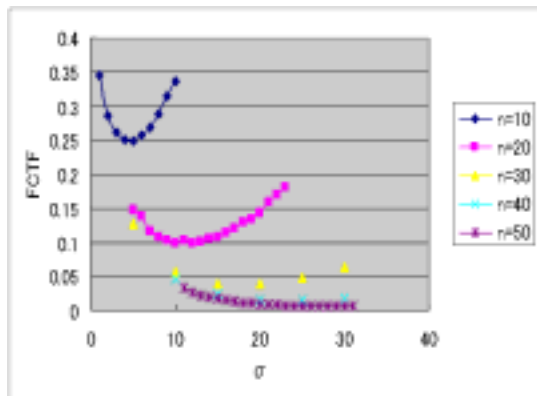


図6 一回目でゲームが終了する頻度

これを見るとわかるように、いずれの n に対しても、推論能力が増えていくに従って、ある点までは、FCTF が低下していく。しかしある転換点が存在し、転換点以降は逆に FCTF は増加していく。このことからわかるように FCTF が最小値をとるのは、 n の値がいずれであっても完全にランダムに相当する $\sigma = 0$ の場合でもなく完全に合理的な場合に相当する $\sigma = \infty$ の場合でもない、適度な合理性を持つ場合であると予想される。さらに転換点は n の増加に従って、 σ が大きい方向に移動し、また転換点のときの FCTF の値自体も n の増加とともに低下していくことがわかる。

次に、実際にどのような結果が導かれるのかという視点から、10 階のムカデゲームにおいてそれぞれの終点ノードにどのような頻度で到達したのかということに関する結果が図7である。

被験者を用いた実験結果と同様に、一期目で確実に終わるわけではなくある程度の割合で協力行動も観察されることがわかる。さらに協力行動が選択される頻度もゲームが進むごとに低下している。このことからプレイヤーの推論能力の不完全性もまた、実際に協力行動がとられるための一因であると考えられる。

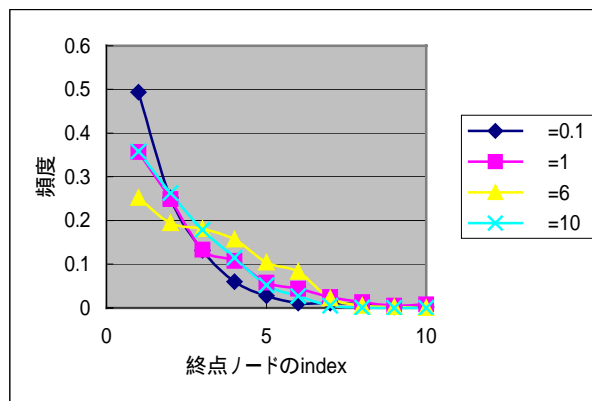


図7 10 階ムカデゲームにおける到達頻度

さらにその不完全性の程度に関しては、この 10 階のムカデゲームにおいては $\sigma = 6$ の、プレイヤー同士のほうが、より合理的な $\sigma = 10$ のプレイヤー同士の場合や、よりランダムな $\sigma = 0.1$ のプレイヤー同士の場合よりもゲームが続きやすいという結果が得られた。なお n を変えた場合もこのような山型のシミュレーション結果が得られることが確認されている。

さらに達成される平均利得をまとめたものが以下の表1、表2である。ここで APP1 はプレイヤー1 の平均利得を、APP2 はプレイヤー2 の平均利得を表している。

表1 10階ムカデゲームにおける平均利得

	0.1	1	6	10
APP1	1.359	1.824	2.163	1.534
APP2	1.685	2.488	2.919	2.158

表2 30階ムカデゲームにおける平均利得

	0.1	1	6	10
APP1	1.396	1.771	7.327	9.789
APP2	1.756	2.471	8.187	10.785

まず表1を見ると上の議論と同様に、適度の合理性を持ったプレイヤー同士のゲームのほうがランダムなプレイヤー同士や、合理的すぎるプレイヤー同士の場合よりも高い平均利得を達成していることがわかる。また回数が増加した場合は、より合理性の高いプレイヤー同士の場合のほうが、利得の上昇幅が大きいという傾向があることがわかった。

これらのことから、プレイヤーの推論が本稿で定義したような意味での粗さを伴う場合は以下の二つのような事

象が起こりうることを暗示していると考えられる。第一に協力的行動を選択することは非合理的な行動に見える。ここで図 6 から、完全にランダムなプレイヤーよりも、適度な合理性を持ったプレイヤーのほうが、かえって非合理的な戦略を高い確率で採用する場合がある。第二に、適度な合理的な主体からなる社会のほうが、その高い頻度での非合理的選択の結果として高い社会厚生を達成できる可能性がある。

それでは次に n の増加について考えてみる。 n が増加するということは、問題の潜在的な協力機会が増えることを意味している。表 1 と表 2 を比較すると、問題が拡大した場合には、合理性の低いプレイヤーは、それほど恩恵を受けることができず、合理性のより高い主体ほどその恩恵を受けやすい傾向にあることがわかった。正確に言うと、 n の増加と共に、社会厚生が最大となる も大きくなる。またそのとき達成される社会厚生のレベルも高くなる。

ムカデゲームは協力行動が望まれるが両者共に望んでいるもの逸脱の誘引が存在している意思決定状況のモデルとみなすことができる。このような状況では協力行動が絶対に実現しないのかということ、確かに Nash 均衡戦略をお互いが実行するならば、協力行動は実現しない。しかしながら、もしも各プレイヤーが本稿のモデルのような意思決定を行っているならば、一定程度の協力行動が生起することが予想できる。

確かに(T)を取ったときに適度な罰則体系を導入することにより、パレート最適な結果の達成が保証されるようなゲームに元のゲームを変換することができるかもしれない。しかしながら、罰則体系の導入には通常社会的なコストが必要である。従ってそのような変換が必ずしも社会厚生を増加させるとは限らない。もしも罰則体系を導入せずとも一定程度の協力行動が実現できるならば、そちらのほうが社会的に見てある種望ましい状態ということもできると考えている。

5. おわりに

本稿では第一に、推論の際のエラーを数理的に表現するモデルにおいて、従来の利得だけではなく、構造の深さという点にも着目し、限定的な推論による意思決定を表現するモデルを提案した。第二に、ムカデゲームにおける実験

結果の新たな解釈を提案した。第三に、推論能力の不完全性が社会における協力状態の実現にどのような影響を与えるのか議論した。

我々が意思決定状況に直面した場合、そもそもその状況が実際にはどのように認知されているのかということ自体が、重要であると考えられる。観察者の目から見て、均衡と思えるような戦略の組が、意思決定に直面している主体たちにとっては、そのように認知されていないことも少なくないと考えられる。本研究の枠組みは、そのような現実の意思決定状況の一つのモデルとなっていると考える。

またムカデゲームでは、協力がある程度続く可能性が高いと我々が認識する理由としては、これまで議論されてきた、利他性などの影響もある程度は存在すると考えている。しかしながら、従来の解釈に加え、推論の誤りという視点からも、このような直観が導かれうると示したことは、意義あることと考えている。さらに、 n 階ムカデゲームという、協力が望まれるある社会において、協力行動の起こりやすさについて、合理性の程度と問題の潜在的な規模という文脈から、示唆される新たな解釈を提示した。

本稿では、主体の合理性が等しい場合のみ議論している。この仮定の下での比較で主体の合理性の増加が必ずしも社会厚生の上昇に結びつくとは限らないことを例示した。今後は様々な合理性をもつ主体たちが混在し、random matching を行うような社会でもなお、今回の結果と同様の性質が保たれるのか検討したいと考えている。さらにこのモデルは一回限りのゲームの結果としてモデル化しているけれども、このような社会状況が繰り返されたとき進化的視点でもどのような戦略が安定するのかということも、今後の課題である。

参考文献

- [1] R. Rosenthal. *Games of Perfect Information, Predatory Pricing and the Chain-Store Paradox*. Journal of Economic Theory 25, 92-100, 1981.
- [2] R. Selten. *The-Chain-Store Paradox*. Theory and Decision 9, 127-159, 1978.
- [3] R.B. Myerson. *Refinements of the Nash Equilibrium Concept*. Int. Journal of Game Theory 7, 73-80, 1978.
- [4] R.D. McKelvey and T.R. Palfrey. *Quantal Response*

Equilibria for Normal Form Games. Games and Economic Behavior 10, 6-38, 1995.

[5] R.D. McKelvey and T.R. Palfrey. *Quantal Response Equilibria for Extensive Form Games*. Experimental Economics 1, 9-41, 1998.

[6] R.D. McKelvey and T.R. Palfrey. *An Experimental Study of the Centipede Game*. Econometrica 60, 803-836, 1992.

[7] R. Aumann. *Correlated Equilibrium as an Expression of Bayesian Rationality*. Econometrica 55, 1-18, 1992.

[8] R. Aumann. *On the Centipede game: Note*. Games and Economic Behavior 23, 97-105 1998.