

最後通牒ゲームにおける社会的意思決定の認知モデリング —利己的な他者に対する適応過程の ACT-R シミュレーション— Modeling Social Decision-Making and Adaption: An ACT-R Simulation of the Ultimatum Game

大麻 紀真[†], 下條 志巖[†], 林 勇吾[†]

Kishin Oasa, Shigen Shimojo, Yugo Hayashi

[†] 立命館大学

Ritsumeikan University

cp0150sr@ed.ritsumei.ac.jp

概要

繰り返し最後通牒ゲームにおいて、記憶された行動履歴から提案者の利他性を推定し、承認・拒否を切り替える応答者モデルを ACT-R アーキテクチャを用いて構築した。シミュレーションの結果、記憶減衰率の高いモデルは提案者の利己的な振舞いや戦略の変更に迅速に適応し、経済合理的なモデルを上回る報酬を得た。一方、減衰率の低いモデルは過去の印象に固執し適応に失敗した。相互作用における不公平回避・互惠的意思決定の創発とその個人差を、宣言的記憶の検索と忘却のメカニズムに基づいて体系的に記述できる可能性が示唆された。

キーワード：ACT-R, 最後通牒ゲーム, 不公平回避行動, 互惠性, 社会的選好, 認知モデリング

1. 背景

人間同士の相互作用は短期の経済合理性を超えた社会的意思決定によって支えられており、互惠性、すなわち親切な相手には親切に、不親切な相手には不親切に対応する性質はその代表例である。利他行動は一見すると行為者にとって不利益だが、将来に他者からの返報が期待できる場合、長期的には有益になりうる (Trivers, 1971)。しかし、協力にただ乗りする非利他主義者に対しては返報が期待できないため、互惠的関係を成立させるために個人は協力的なパートナーと利己的なフリーライダーとを見分ける必要がある (Stephens, 1996)。

このような社会的意思決定を検討する枠組みとして経済ゲームが用いられてきた。最後通牒ゲームでは、ある量の財について提案者がその分配方法を提示し、応答者がそれを承認するか否かを選択する。拒否した場合は両者とも報酬を得られない。合理的には応答者は取り分が 0 でない限り承認すべきであるが、実際には不公平な提案は高い確率で拒否されることが知られている (Güth et al., 1982)。このことは、個人が一時

的なコストを支払っても理不尽な相手を罰し、公平さを維持しようとする不公平回避行動 (Fehr & Schmidt, 1999; Fehr & Gächter, 2002) をとることを示しており、これはゲームが繰り返し実施される状況において特に顕著である。

互惠性研究は人間が協力的なパートナーを選ぶ際に、結果の公平性だけでなく他者の行動から推察される意図も評価することを指摘 (Falk & Fischbacher, 2006) しており、個人が他者の行動からその意図を読み取り、相手が自分に対して好意的なら自分も相手に対して好意的な行動をとるというモデルを提唱している (Rabin, 1993)。すなわち、経済ゲームにおける個人の意思決定は相手の行動への単なる反応ではなく、観察された行動から意図や方針といった内的状態を推論し (Theory of Mind (Premack & Woodruff, 1978)), それに応じて自身の行動を変化させる、動的かつ相互作用的な過程といえる。

上記を踏まえると、繰り返し最後通牒ゲームにおいて応答者は、相手の提案傾向からその利他性のタイプを推定し、利他的な相手には寛容に、利己的な相手には厳格に応答することで互惠的関係を調整している可能性がある。なぜなら、利他的な相手からは将来の返報が期待できる一方、利己的な相手には譲歩しても継続的に搾取される恐れがあり、相手の理不尽な行動を罰しようとする動機が強まるからである。このような推定は、相手の過去の提案行動を想起することによって形成されと考えられる。例えば、以前に提案者が利己的な提案をしてきた記憶を想起したなら、応答者はかれが利己的であると推定することになるだろう。このとき、より古い記憶に対する重みづけの仕方、すなわち履歴の参照における時間窓は記憶の減衰特性に依存する。記憶の減衰 (忘却) は本来欠陥ではなく、環境においてある情報が再び必要になる確率を最適に予測するための適応 (Anderson & Schooler, 1991) である。この記憶の減衰特性の個人差によって、例えば

「過去の利己的な行動をいつまでも根に持つ」とか「過去の出来事は水に流し、現在の相手の振る舞いを重視する」といったような、応答者における個別の性格を説明しうる。

本研究では、こうした人間の意思決定を計算論的にモデル化するにあたり、人間が問題解決に用いる知識を宣言的記憶（事実や経験の記憶）と手続き的記憶（条件と行動のルール）の2つに区別して扱う。すなわち、過去の提案行動を記憶（あるいは忘却）する過程を宣言的記憶、そこから文脈に基づいて実際に承認・拒否の行動を選択する過程を手続き的記憶の働きとして捉え、その上で前者の特性が個人による社会的意思決定を予測すると考える。

2. 目的と仮説

このように、互惠性や不公平回避、その現れ方における個人の性格的特徴は、明示的なルールを仮定せずとも、記憶の検索と忘却（減衰）という基本的な認知過程によってその創発を説明できる可能性があるが、この点について検討した例は少ない。

本研究では認知アーキテクチャ ACT-R(Anderson, 2007)を用いて、繰り返し最後通牒ゲームにおける応答者の意思決定を宣言的記憶の検索による提案者のタイプ推定と、手続き的記憶による承認・拒否の判断という二段階のプロセスとしてモデル化することで、この説明の妥当性を検討する。作成したモデルについて記憶減衰率を操作し、複数の戦略を持つ提案者エージェントとの繰り返し最後通牒ゲームのシミュレーションにおいて、互惠関係の形成や搾取の回避がどのように変化するかを、提案を拒否しない経済合理的なベースライン・モデルと比較して検討する。

仮説として、不公平回避を行うモデルは経済合理的モデルより多くの報酬を獲得すると予想する。さらに、繰り返し四人のジレンマにおいてしつぱ返し戦略が有効であることが示されているように(Rapoport, 1989), 相手の行動が変化する環境ではその変化を迅速に検知し適応することが有利である。したがって、不公平回避モデルの中では記憶減衰率が大きいモデルが、小さいモデルより多くの報酬を獲得すると予想する。

3. 方法

3.1 手続き

ゲームは 1000 ポイントの報酬を提案者と応答者で分け合う方式をとり、その分配方法は表 1 に示す 7 通りであった。それぞれの選択肢は対応する 0~6 のインデックスで管理された。提案エージェントは後述す

表 1 提案できる分配方法とそれぞれの取り分

proposal options	share	
	responder	proposer
0	900	100
1	800	200
2	700	300
3	500	500
4	300	700
5	200	800
6	100	900

る 3 つの条件ごとに、定められた方針に従ってこれらのうちの 1 つを提案し、ACT-R による応答者モデルが「承認」または「拒否」の応答を決定した。提案と承認のサイクル 1 回を 1 ラウンドとし、人間を対象として同様のゲームを扱った先行研究 (Hayashi & Okada, 2017) に倣い 15 ラウンド繰り返した。

3.2 提案エージェント

エージェントは第 1 ラウンドに提案 3（両者が 500 ポイントを得る）を選択し、第 2 ラウンド以降は、前回より提案者の取り分を 1 段階増加させる raise、同一の提案を維持する hold、1 段階減少させる lower のいずれかの提案行動をとった。エージェントが提案行動を選ぶ方針の違いから、以下の 3 条件を定義した。

(1) **利己的条件:** 前ラウンドの相手の応答に依存し、承認されていれば raise、拒否されていれば lower を選ぶ。相手が譲歩する限り自身の利益を追求することで、応答者に対し最大利益を獲得しようとする。

(2) **裏切り条件:** 第 7 ラウンドまで hold（提案 3）を継続し、第 8 ラウンド以降は利己的条件と同じ戦略へ切り替える。

(3) **改心条件:** 第 7 ラウンドまでは利己的条件と同じ戦略をとる。第 8 ラウンド以降は提案 3 へ収束するよう lower または raise を選択し、到達後は hold を継続する。

3.3 応答者モデル

応答者モデルは、各ラウンドの提案と応答をエピソードチャンクとして宣言的記憶に格納した。チャンクにはそのラウンドにおける相手の提案、提案が持つ利己的・利他的・中立的の属性（提案 3 が中立、それよりモデルの取り分が多いものを利他的、少ないものを利己的とした）、および自分がした応答が対応するスロットの値として保持された。モデルはそれらのチャンクから一つを想起し、そのラウンドにおける相手の提案の属性を現在の相手のタイプとした。

3.3.1 宣言的記憶の想起

宣言的記憶に格納されたチャンクを想起する際に、モデルはそれぞれのチャンク i について以下の式によって与えられる活性値 A_i を計算し、その値が最も大きいチャンクを想起する。

$$A_i = B_i + S_i + \epsilon_i \quad (1)$$

ここで、 B_i はベースレベル活性値、 S_i は文脈からの拡散活性値、 ϵ_i はノイズを表す。ベースレベル活性値 B_i は、そのチャンクが過去にどの程度頻繁に、かつ最近利用されたかに対応し、以下のベースレベル学習 (Base-Level Learning) の式によって計算される。

$$B_i = \ln \left(\sum_{j=1}^n t_j^{-d} \right) \quad (2)$$

n はそのチャンクが過去に作成・想起された回数、 t_j は j 番目の参照からの経過時間、 d は記憶の減衰率を表す。すなわち、より最近の記憶や、頻繁に利用される記憶ほど思い出されやすい傾向にある。減衰率が大きいほど古い記憶の活性値は急速に低下するため、これに対応するモデルパラメータ「:bll」を変化させることで、モデルが意思決定にあたって直近のラウンドにおける相手の行動にどれだけ重きを置くか、すなわち「忘れっぽい」性格か、あるいは「根に持つ」性格かを操作できる。本研究では高減衰率 (:bll = 0.6) と低減衰率 (:bll = 0.3) の2条件を設定した。

拡散活性値 S_i はある知識の利用によって関連する別の知識が利用されやすくなる、活性化拡散の影響に対応する項である。本モデルではこの機能を利用しなかったため、パラメータは初期設定のままとした。

また、 ϵ_i は想起プロセスにおける確率的な揺らぎを表現するノイズであり、ロジスティック分布に従う。この分散を規定するパラメータ「:ans」について、本研究では標準的な値である 0.1 を使用する。

3.3.2 応答の選択

同一の提案に対する承認・拒否の異なる判断を、競合する別のプロダクション (手続き的記憶) として表現した。競合するプロダクションの中からどれが実際に適用される (発火する) かは、以下の式で計算される各プロダクションのユーティリティ (効用) 値によって決定される。 U_i^* はそのプロダクションにあらかじめ規定されたベースユーティリティ値であり、 ϵ_i はロジスティック分布に従うノイズである。

$$U_i = U_i^* + \epsilon_i \quad (3)$$

相手のタイプごとにそれぞれのプロダクションにお

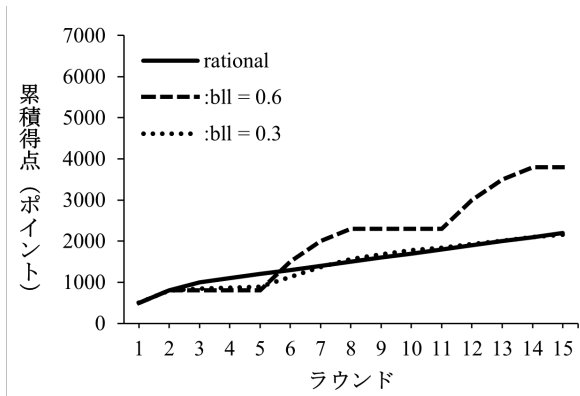


図1 「利己的」条件における各モデルの累積得点

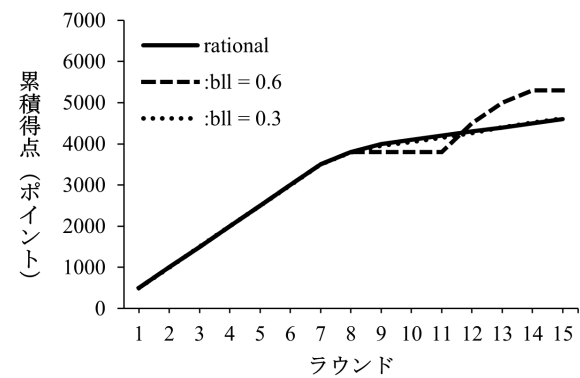


図2 「裏切り」条件における各モデルの累積得点

ける U_i^* のバランスを操作することで、利己的な相手に対して拒否を選択しやすくなる不公平回避型のモデルと、相手の利他性に関わらず承認を選ぶ、経済合理性に従うモデルを実装した。具体的には、前者においてはベースユーティリティ値に対応するパラメータ「:u」を承認のプロダクションで 5.0、拒否で 10.0 に設定し、後者においてはその逆とした。ノイズの分散を定める「:egs」パラメータは 0.5 に設定し、これはそれぞれのプロダクションにおける :u の差に対して十分小さいので、モデルは相手のタイプに応じて基本的には決まった応答を返していたといえる。

4. 結果と考察

提案者のタイプに関わらず承認行動の効用が高い経済合理的 (rational) モデルと、不公平回避行動を取る2つのモデル (高減衰率 :bll = 0.6, 低減衰率 :bll = 0.3) について、3種類の提案エージェントとのゲームをそれぞれ 100 回ずつ実行した。ラウンドごとの累積獲得ポイントの平均値を図1~3に示す。

利己的エージェント条件では、序盤は拒否コストを支払わない rational モデルが有利であったが、第6ラウンド以降は高減衰率群が rational モデルを上回り、ラウンドの進行とともに差が拡大した。この結果は、

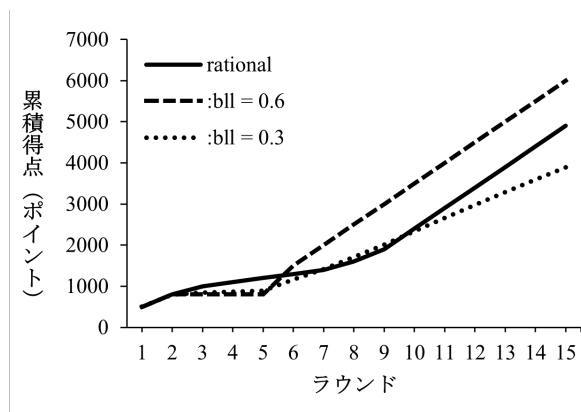


図3 「改心」条件における各モデルの累積得点

不公平回避行動が短期的には損失を伴う一方で、利己的提案者に拒否を返すことで譲歩を引き出し、長期的利益を向上させようことを示している。一方、低減衰率群は相手が譲歩した後も過去の利己的提案に基づく印象を保持し続けたため、承認への切り替えが遅れ、得点は rational モデルと同程度に収束した。最後の数ラウンドを含むいくつかの部分で高減衰率群による得点増加の鈍化が見られるが、これはモデルがしばらく承認を繰り返したことで提案エージェントが要求を引き上げ、それに対応したモデルが再び拒否を増加させたことに因る。これを受けてエージェントは再び譲歩を強いられるため、モデルの獲得する得点は増加する。本研究で用いたモデルやエージェントはあくまで単純なものであり、設計上は相互作用を 15 ラウンド以上に延長した場合も同様のサイクルが繰り返され、差が開き続けることが決定論的に予測される。

「裏切り」条件では、第 7 ラウンドまでは全モデルが同程度の報酬推移を示したが、提案者が利己的戦略へ切り替えた後、高減衰率群のみが拒否行動を増加させ、長期的報酬を向上させた。低減衰率群では、それまでの公平な行動履歴が高い活性値を伴って保持された結果、提案者の行動変化を十分に検知できず、裏切りへの適応が遅れたと考えられる。

「改心」条件では、高減衰率群が最も高い報酬を獲得した一方、低減衰率群は rational モデルを下回った。これは、序盤に形成された利己的提案者としての印象が維持され、提案者が公平な提案へ転じた後も拒否が継続した（第 9～11 ラウンド付近）ためと解釈できる。すなわち、過去の印象を保持し続けることは搾取回避に寄与する一方で、協力関係の修復を阻害しうる。

以上より、記憶減衰率の違いは、搾取の回避だけでなく協力関係の修復可能性にも影響し、社会的相互作用における適応的意思決定の重要な決定因となることが示唆された。

5. まとめ

本研究では ACT-R アーキテクチャを用い、繰り返した最後通牒ゲームにおける応答者の不公平回避行動を宣言的記憶の検索と忘却に基づく計算モデルとして実装した。シミュレーションの結果、高い記憶減衰率、すなわち相手の直近の行動を重視する性質が、搾取の回避および協力関係の修復の双方に有利な適応的特性であることが示唆された。本研究の結果は、互惠性や不公平回避が社会規範や感情といった高次の機能のみならず、宣言的記憶の検索と忘却という基本的な認知メカニズムから創発するという見解を支持する。同時に、相互作用の履歴を参照する時間窓として機能し、印象形成の更新速度を規定する記憶減衰率のパラメータを導入することで、「根に持つ」「水に流す」といった個人の特性を同一の枠組みにおいて記述し、その機能を検討できた点は重要な貢献といえるだろう。ただし、本研究のモデルは提案に利他的・利己的の属性を固定的に付与した点や、タイプ推定を単一のエピソード想起に依存させた点など、探索的な研究としてのいくつかの単純化を含んでおり、人間のより複雑な認知過程を十分に反映できていない可能性がある。また、人間の実験データとの直接比較を行っていないため、拒否率や反応時間などの詳細な行動指標についてどの程度説明できるかは今後の課題である。今後も引き続き設計を洗練させ、柔軟で人間らしい意思決定をより精緻に記述できるモデルを構築していく必要がある。

文 献

- Anderson, J. R. (2007). *How Can the Human Mind Occur in the Physical Universe?* Oxford University Press.
- Anderson, J. R., & Schooler, L. J. (1991). Reflections of the Environment in Memory. *Psychological Science*, 2(6), 396–408.
- Falk, A., & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, 54(2), 293–315.
- Fehr, E., & Gächter, S. (2002). Altruistic Punishment in Humans. *Nature*, 415, 137–40.
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3), 817–868.
- Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, 3(4), 367–388.
- Hayashi, Y., & Okada, R. (2017). Compound effects of expectations and actual behaviors in human-agent interaction: Experimental investigation using the Ultimatum Game. *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*, pp. 2168–2173.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. *Behavioral and Brain Sciences*, 1, 515–526.
- Rabin, M. (1993). Incorporating Fairness into Game Theory and Economics. *The American Economic Review*, 83(5), 1281–1302.
- Rapoport, A. (1989). *Prisoner's Dilemma*, pp. 199–204. London: Palgrave Macmillan UK.
- Stephens, C. (1996). Modelling Reciprocal Altruism. *The British Journal for the Philosophy of Science*, 47(4), 533–551.
- Trivers, R. (1971). The Evolution of Reciprocal Altruism. *Quarterly Review of Biology*, 46, 35–57.