

多人数グループ内での間接的互惠性の進化： Standing 評判基準の効果

鈴木 真 介[†] 秋 山 英 三[†]

社会的ジレンマにおける協力の形成・維持は人間社会や生態系にとって重要な問題であり、社会生物学、社会心理学、経済学など様々な分野で研究されている。間接的互惠性に関する一連の研究では、評判の効果を導入することにより、間接的互惠性に基づく協力関係が成立しうることが示されている。しかし、これらの研究で前提となっているのは主に2者間の相互作用である。現実の社会や生態系では多くの場合、3者以上の間の相互作用も重要になる。本研究では、様々な評判基準や行動基準の下での「評判が存在する状況における多人数グループ内（3者以上の相互作用）での協力的行動の進化」を分析した。その結果、以下のことが明らかになった。(1) すべての評判基準、行動基準の組合せの下で、グループの人数が増加すると協調関係を形成・維持するのが困難になる、(2) 他人の行動の理由を思い語り2者間の協力的行動の進化を安定化する評判基準（*standing*）が、他人の行動の理由を思い語らない評判基準（*image scoring*）に比べて、必ずしも多人数間の協力的行動を促進するとは限らない。

Evolution of Indirect Reciprocity in Sizable Groups: the Effect of Standing Reputation Criterion

SHINSUKE SUZUKI[†] and EIZO AKIYAMA[†]

Evolution of cooperation in social dilemmas has been of considerable concern in various fields such as socio-biology, socio-psychology, and economics. Recently works on the indirect reciprocity theory states that cooperation can evolve through reputation even if pairs of individuals interact only a few times. To our knowledge, most indirect reciprocity model presumed dyadic interaction. However, social dilemmas in real communities usually involve three or more individuals. We investigate the evolution of cooperation in large communities where reputation of individuals affects the decision making process, under the various reputation and action criteria. In this paper, we found the following: (1) under all the reputation and action criteria, formation and maintenance of cooperation becomes difficult as group size increases even if the effect of reputation is included; (2) *standing* reputation criterion, which reflects motivation of others' action and stabilizes the evolution of cooperation under dyadic interaction, does not always facilitate the evolution of cooperation under three or more person interaction, compared with *image scoring* reputation criterion which does not reflect motivation of others' action.

1. 研究の背景と目的

1.1 社会的ジレンマ

社会的ジレンマにおける協力的行動の進化は人間社会や生態系にとって重要な問題であり、社会生物学、経済学、社会心理学など様々な分野で研究されている。

社会的ジレンマとは、「共有地の悲劇⁶⁾」のように、個人合理性と社会合理性が対立する状況である。このような状況では、各個人にとっては合理的な行動が社会的には不合理な結果を招いてしまう。例として、魚

の乱獲問題を考える。それぞれの漁師は魚をたくさん捕れば捕るほど利益が得られる。しかし、全員が魚を捕り過ぎると魚は枯渇してしまい、翌年には、ほとんど魚が捕れない状況に陥ってしまう。つまり、各漁師の合理的な行動（できるだけ多くの魚を捕ること）により、全員にとって望ましくない（社会合理的でない）状況、つまり漁獲資源の枯渇に陥ってしまう。社会的ジレンマは、過剰人口問題、資源乱獲問題など現実社会の様々な場面で見られる。社会的ジレンマ状況における協力的行動の進化のメカニズムを研究することが、これらの問題の解決に役立つと考えられている。

1.2 互惠主義

社会的ジレンマにおける協力的行動の進化は、しばし

[†] 筑波大学大学院システム情報工学研究科
Graduate School of Systems and Information Engineering,
University of Tsukuba

ば互恵性の理論によって説明されてきた。互恵性の理論は Trivers¹⁸⁾ によって提唱されたものであり、「協力行動をするならば、その相手からも協力行動を返してもらえ」というメカニズムにより協力行動は進化する」と主張した。

Axelrod²⁾ はこの互恵性の概念をもとに、「繰返し 2 人囚人のジレンマゲーム」を用いて社会的ジレンマにおける協力行動の進化を分析した。Axelrod は、「2 者が繰返し対戦する」とき「しっぺ返し戦略」に基づく協力行動が社会に広がり安定化することを示した。なお、Axelrod の研究では、相互作用は 2 者間で行われ、かつ長期間続くと仮定されている。

1.3 n 人 (多人数) 囚人のジレンマゲーム

しかし、現実の社会で見られる相互作用、そしてそこでの協力行動は必ずしも 2 者間とは限らない (e.g. 会社の同僚、友人関係など)。そこで、Joshi⁷⁾、Boyd³⁾ は Axelrod の互恵性の理論に基づいて「繰返し n 人囚人のジレンマゲーム」を分析した。その結果、グループの人数 (n) が増加すると協力行動の進化に必要な条件が急激に厳しくなることが分かった。つまり、グループの人数が増加すると協力行動の進化は困難になるのである。

この結果は、Axelrod の互恵性の理論では多人数グループ内での協力行動を説明できないことを示唆する。多人数グループ内での協力行動の進化を説明するためには、何か別のメカニズムを考慮する必要がある。

1.4 評判の効果

そのメカニズムの有力な候補の 1 つとして考えられるのが「評判」による協力形成である。たとえば、インターネット・オークション・サイトの多くでは、参加者の評判を公開することで不正行為を抑制している。また、被験者実験においても、評判が人間の協力行動に大きな影響を与えることが確かめられている^{9),19)}。

1.4.1 間接的互恵主義

上記のような「評判」と「協力行動の進化」の関係を保った研究に「間接的互恵性」に関する研究がある。

間接的互恵性の理論は Alexander¹⁾ によって提唱され「相手に協力行動をすれば、相手以外の他の誰かから協力行動を返してもらえ」というメカニズムにより協力行動は進化する」と主張した。間接的互恵性の理論は「同じ相手との相互作用の繰返し」を前提としないため、前述の互恵性の理論よりも適用範囲が広いと考えられている。

Nowak^{10),11)} は「評判の効果を導入することにより、同じ相手と何度も相互作用しなくても間接的互恵性に基づく 2 者間の協力行動が進化する」ことを示した。彼らのモデルでは、各プレーヤの評判は「image scoring 評判基準」に従って変化するとされていた。image scoring 評判基準の下では、プレーヤが協力行動をとると評判が良くなり、裏切り行動をとると評判が悪化する。

Nowak らの研究の後、image scoring 評判基準の下では間接互恵的協力行動は進化的に安定にならないことが様々な研究により指摘された^{4),8),12),13)}。そして、それらの研究は「裏切り者への裏切りを正当だと評価する」評判基準 (*standing*)¹⁵⁾ を採用するとき、間接互恵的協力行動が進化的に安定な戦略となることを示した。

1.5 多人数の社会的ジレンマにおける評判の役割

評判と間接的互恵性に関する上記の研究では、2 者間の相互作用を前提としている (e.g. 個人 A から個人 B、個人 C から個人 A への寄付行動など)。しかし、1.3 節に述べたように、現実社会で見られる協力行動の進化を扱うためには、多人数間の協力行動を扱う必要がある。

Suzuki and Akiyama^{16),17)} は「評判が意志決定に影響を与える状況における、多人数グループ内での協力行動の進化」を、計算機シミュレーションを用いて分析した。具体的には、Nowak¹⁰⁾ の間接的互恵性のモデルを n 人ゲームに拡張し、協力行動の進化とグループの人数との関係を分析した。その結果「評判が存在する状況でも、グループの人数の増加にともなって協力行動の進化は難しくなる」ことが示された。なお、Suzuki and Akiyama のモデルでは、各プレーヤの評判は「image scoring 評判基準」に従って変化するとされている。

1.6 本研究の目的

Suzuki and Akiyama のモデルでは image scoring 評判基準が使われていたが、現実社会では image scoring 以外の評判基準もしばしば見られる。たとえば、ある人が裏切り者に対して裏切り行動をとった場合「裏切り者を罰したのだから正当な行為である」と判断されることもありうる。

そこで、本研究では、Suzuki and Akiyama の評判付き n 人囚人のジレンマゲームのモデルに「裏切り者に対する裏切り」を正当な行為と見なす評判基準 (*standing*) を導入した。*standing* 評判基準が 2 者間の協力行動の進化を安定化することはすでに示されて

対戦の初回は協力し、以降は相手に裏切られたら裏切り、相手が協力したら協力する戦略。

いるが^{(4),(8),(12),(13)}，多人数グループ内での協力行動の進化にどのような影響を与えるのかについてはまだ明らかにされていない．本研究では，standing 評判基準の導入が多人数グループ内での協力行動の進化にどのような影響を与えるのかについて，image scoring 評判基準と比較することで調べた．

2. モデル

まず， N 人のプレーヤからなる集団を考える．1 世代は m 回のラウンドから構成される．各ラウンドでは，集団からランダムに選ばれた n (≥ 2) 人がグループを形成し， n 人囚人のジレンマゲームを行う．

n 人囚人のジレンマゲームでは，それぞれのプレーヤが「協力」か「裏切り」かを選択する．協力者の利得， π^C ，裏切り者の利得， π^D ，はそれぞれ以下のとおりである：

$$\pi^C = (i/n)b - c \quad (1)$$

$$\pi^D = (i/n)b \quad (2)$$

ここで， i はグループ内の協力者数である．本研究では，社会的ジレンマ状況を扱っているため以下を仮定する：

- $c > b/n$ ：あるプレーヤが協力行動をとったときのグループのメンバ 1 人あたりの利得の増加分， b/n ，は協力のコスト， c ，よりも小さい．よって，各プレーヤにとっては「協力しない（裏切る）」方が得である（裏切りは個人合理的）．
- $b > c > 0$ ：あるプレーヤが協力したときの社会全体の利得の増加分， b ，は，協力のコスト， c ，よりも大きい．よって，各プレーヤが「協力」を選択した方が，社会全体の利得の和は増加する（協力は社会合理的）．

各プレーヤは評判， s ($s = -5 \sim 5$)，を持つ．各プレーヤの評判の初期値は 0 であり，最大値，最小値はそれぞれ 5，-5 である．なお，最大値 $s = 5$ を持つプレーヤの評判はそれ以上増加はしない（最小値に関しても同様）．また，各プレーヤの評判は集団のメンバ全員に正確に知られているとする．本研究では，評判 s の値を決める基準として image scoring 評判基準と standing 評判基準を考える．それぞれの評判基準の下で，各プレーヤの評判は以下のように変化する．

- image scoring 評判基準：協力行動をとれば，そのプレーヤの評判は 1 単位増加し，裏切り行動をとれば 1 単位減少する．
- standing 評判基準：協力行動をとれば，そのプレーヤの評判は 1 単位増加する．そのプレーヤ以外のグループ・メンバの評判の平均が 0 以上のとき，裏切り行動をとると評判が 1 単位減少する．そのプレーヤ以外のグループ・メンバの評判の平均が 0 より小さいとき，裏切り行動をとっても評判は変わらない（裏切り者への裏切りは正当な行為と評価されるから）．

また，各プレーヤは，他人および自分の評判に基づいて自分の行動を決定する．本研究ではその基準（行動基準）として以下の 4 つを考える．

- DIS 行動基準：各プレーヤは「他人の評判の平均が高い（良い）」とき協力する．プレーヤの戦略は k ($k = -5 \sim 6$) で表され，戦略 k を持つプレーヤはグループ・メンバの評判の平均が k 以上のとき協力行動をとる．
- OR 行動基準：各プレーヤは「他人の評判の平均が高い（良い）」，または「自分の評判が悪い」とき協力する．つまり，自分の評判を考慮する行動基準である．プレーヤの戦略は k ($k = -5 \sim 6$) と h ($h = -5 \sim 6$) で表され，戦略 (k, h) を持つプレーヤはグループ・メンバの評判の平均が k 以上，または，自分の評判が h より小さいとき協力行動をとる．
- MIN 行動基準：各プレーヤは「他人の評判の最小値が高い（良い）」とき協力する．つまり，最も評判の悪いメンバの評判のみを考慮する行動基準である．プレーヤの戦略は k ($k = -5 \sim 6$) で表され，戦略 k を持つプレーヤはグループ・メンバの評判の最小値が k 以上のとき協力行動をとる．
- MAX 行動基準：各プレーヤは「他人の評判の最大値が高い（良い）」とき協力する．つまり，最も評判の良いメンバーの評判のみを考慮する行動基準である．プレーヤの戦略は k ($k = -5 \sim 6$) で表され，戦略 k を持つプレーヤはグループ・メンバの評判の最大値が k 以上のとき協力行動をとる．

さらに，世代の終わりには各プレーヤは子孫を残す．各プレーヤが残す子孫の数は，その世代（ m ラウンド）で各プレーヤが獲得した利得の合計（「適応度」）によって決まる．適応度が高いプレーヤほど，子孫をたくさん残せる可能性が高くなる．本研究では，この過程に遺伝的アルゴリズム（トーナメント・セレクション）

「裏切り者に対する裏切り」を正当な行為と見なす評判基準を採用することにより，「協力者が自らの評判を下げず，裏切り者を罰する」ことが可能になるからである．

当然，実際の人間は様々な評判基準を使っていることが考えられるが，本研究では「安定な協力状態を実現する最も単純な評判基準^{(4),(12),(13)}」として，standing を採用した．

表 1 アルゴリズム
Table 1 Algorithm.

1)	N 人からなる初期集団を作る (戦略はランダムに決まる).
2)	集団から n 人がランダムに選ばれる.
3)	これら n 人が n 人囚人のジレンマゲームを行う.
4)	2), 3) を m 回繰り返す.
5)	各個人は適応度に応じて子孫を残す.
6)	次世代の集団の評判を 0 に戻す.
7)	2) から, 6) までを繰り返す.

表 2 基本的なパラメータ
Table 2 Basic parameter settings.

集団の人数	$N = 200$
1 世代のラウンド数	$m = 4 \times N$
協力のコスト	$c = 1$
協力の利益	$b = \alpha \times n$ ($\alpha = 0.85$)
評判	$-5 \leq s \leq 5$
戦略 k	$-5 \leq k \leq 6$
戦略 h	$-5 \leq h \leq 6$

ン)を使用した. なお, 次世代のプレーヤの評判は 0 に戻るが, 戦略は親の戦略を継承する.

もちろん, プレーヤが戦略を決定・修正する方法はほかにも考えられる (強化学習など). しかし, 本研究では協力行動の進化 (または, 優れた戦略の模倣による協力行動の普及) に注目しているため, 上記のような「適応度が高いプレーヤほど多くの子孫を残す」という方法 (遺伝的アルゴリズム) を採用した. また, 本モデルには「戦略の突然変異」は組み込まれていない. よって, 実現した結果の安定性については議論していない. これらは今後の課題である.

計算機シミュレーションのアルゴリズムの概要は表 1, 各パラメータの値は表 2 を参照. 今後は特に言及がなければ, これらのアルゴリズム, パラメータを使用する.

この計算機シミュレーションを上記の評判基準 2 種類, 行動基準 4 種類のすべての組合せ 8 通りについて行った. 以後, 評判基準として image scoring, 行動基準として DIS が使われているケースを image-DIS と表す (同様に standing-OR, image-MAX など). なお, image-DIS は Suzuki and Akiyama¹⁶⁾ と同一のケースである.

3. 結 果

計算機シミュレーションは, 評判基準 2 種類, 行動基準 4 種類のすべての組合せ 8 通りの状況それぞれにおいて「最終的に生き残る戦略」, 「グループの人数と協力率の関係」, 「standing 評判基準導入の影響」に注目して行った.

3.1 DIS 行動基準

3.1.1 生き残った戦略の分布

本節では, DIS 行動基準の下で「どのような戦略が生き残るのか」, また「その生き残った戦略は協力的社会を実現するのか」に着目する. 図 1, 図 2 はそれぞれグループの人数が少ないとき ($n = 20$) とグループの人数が多いとき ($n = 60$) の, 生き残った戦略の分布を表す.

図 1 より, image-DIS の下ではグループの人数が少ないとき, $k = -2$ 付近の戦略が生き残る可能性が高いことが分かる. なお, 世代の開始時点ではすべてのプレーヤの評判は 0 なので, すべてのプレーヤが戦略 $k \leq 0$ を採用しているとき, その世代でとられる行動はすべて協力行動になる. よって, 戦略 ($k \leq 0$) を協力的戦略と呼ぶ (同様に, 戦略 ($k > 0$) を裏切的戦略と呼ぶ). また, 協力的戦略の中で, 戦略 ($k = -2$) のように裏切り者 (評判の悪いプレーヤ) に対しては裏切ることで罰を与え, 協力者 (評判の良いプレーヤ) には協力する協力的戦略を互恵的協力戦略と呼ぶ.

一方, グループの人数が増加すると, $k = 6$ 付近の戦略が生き残る可能性が高くなっていることが図 2 から分かる. この戦略 $k = 6$ はつねに裏切り行動をとる (プレーヤの評判の最大値は 5 だから).

つまり, image-DIS の下では, グループ人数が少ないときは互恵的協力戦略により協力社会が実現し, グループ人数が多いときは裏切的戦略により裏切社会が実現する. なお, 協力 (裏切) 社会とは, すべてのプレーヤが協力 (裏切) 行動をとる社会である. また, この傾向は standing-DIS の下でも同様に見られた. なお, 図 1, 図 2 は 1,000 回のシミュレーションの平均であり, 1 回のシミュレーションでは 1 つ, または 2 つの戦略のみが生き残る. 以後の節 (OR, MIN, MAX 行動基準) でも同様である.

3.1.2 グループの人数と協力率の関係 (image scoring と standing 評判基準の比較)

前節では生き残る戦略に注目したが, 本節ではグループの人数と協力率の関心に注目する. 図 3 より, image-DIS と standing-DIS の両ケースで, グループの人数の増加にともなって協力率が低下していること

表 2 の設定ではラウンド数がグループの人数に依存しないため, グループの人数が変化すると各プレーヤが 1 世代で行うゲームの回数が増える. しかし, このことは, 本研究のパラメータの範囲 (表 2) では, 協力行動の進化に影響を及ぼさないことが分かっている¹⁷⁾.

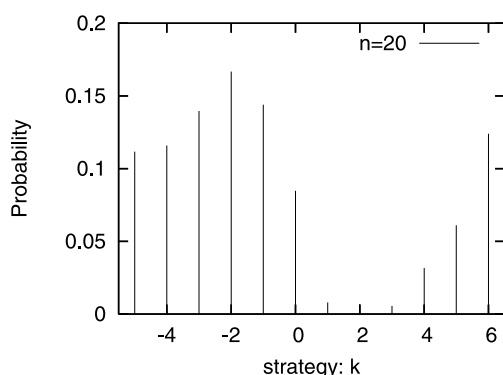


図 1 生き残った戦略の分布 (image-DIS), 横軸: 戦略の種類, 縦軸: それぞれの戦略が生き残る確率 (シミュレーション 1,000 回の平均), グループの人数 $n = 20$

Fig. 1 Survival probability of each strategy under image-DIS (the average values of 1,000 computer simulation runs), horizontal axis: strategies, vertical axis: survival probability of each strategy, group size, $n = 20$.

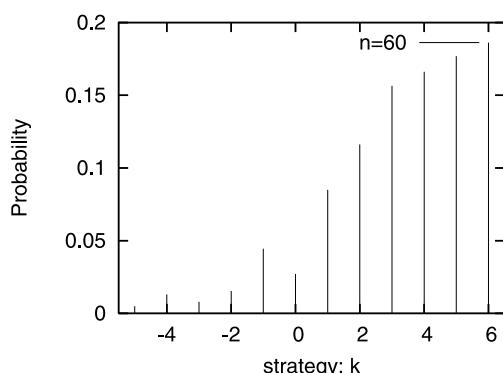


図 2 生き残った戦略の分布 (image-DIS), 横軸: 戦略の種類, 縦軸: それぞれの戦略が生き残る確率 (シミュレーション 1,000 回の平均), グループの人数 $n = 60$

Fig. 2 Survival probability of each strategy under image-DIS (the average values of 1,000 computer simulation runs), horizontal axis: strategies, vertical axis: survival probability of each strategy, group size, $n = 60$.

が分かる。

ただし, 2 種類の評判基準を比べると standing 評判基準の下での協力率が image scoring 評判基準の下での協力率を上回っている。つまり, standing 評判基準を導入すると協力行動の進化が促進されるのである。

3.2 OR 行動基準

3.2.1 生き残った戦略の分布

本節では, OR 行動基準の下で「どのような戦略が生き残るのか」に着目する。図 4, 図 5 はそれぞれグループの人数が少ないとき ($n = 4$) とグループの人

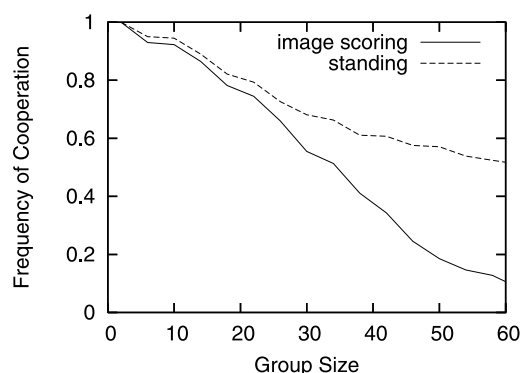


図 3 グループの人数と協力率の関係 (DIS 行動基準, 実線: image scoring 評判基準, 破線: standing 評判基準, シミュレーション 1,000 の回平均)

Fig. 3 Relationship between group size and the frequency of cooperation under DIS (the average values of 1,000 computer simulation runs, solid-line: under image scoring, dashed-line: under standing).

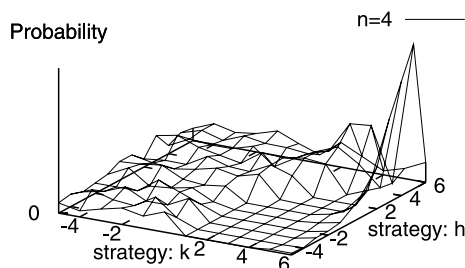


図 4 生き残った戦略の分布 (image-OR, シミュレーション 1,000 の回平均) グループの人数 $n = 4$

Fig. 4 Survival probability of each strategy under image-OR (the average values of 1,000 computer simulation runs), group size, $n = 4$.

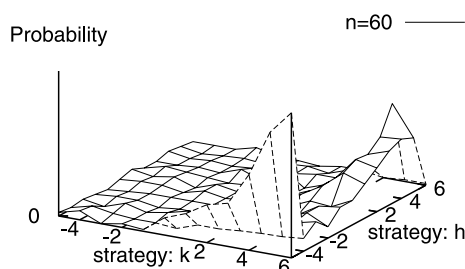


図 5 生き残った戦略の分布 (image-OR, シミュレーション 1,000 の回平均) グループの人数 $n = 60$

Fig. 5 Survival probability of each strategy under image-OR (the average values of 1,000 computer simulation runs), group size, $n = 60$.

数が多いとき ($n = 60$) の, 生き残った戦略の分布を表す。image-OR の下ではグループの人数が少ない

図 4 では, グループの人数が少ない時として「 $n = 4$ 」のケースを示した (DIS 行動基準, 図 1, では $n = 20$ のケース)。結果の特徴がより顕著に現れているからである。

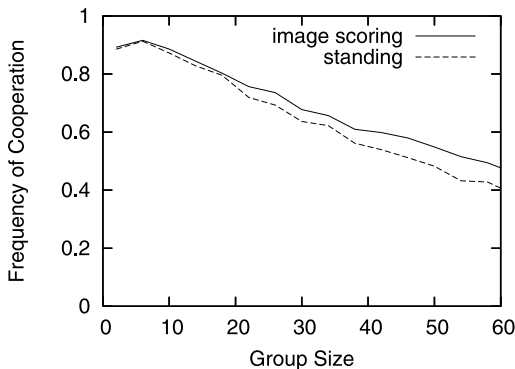


図 6 グループの人数と協力率の関係 (OR 行動基準, 実線: image scoring 評判基準, 破線: standing 評判基準, シミュレーション 1,000 の回平均)

Fig. 6 Relationship between group size and the frequency of cooperation under OR (the average values of 1,000 computer simulation runs, solid-line: under image scoring, dashed-line: under standing).

とき, 戦略 ($k = 6, h = 4$ or 5) が生き残りやすいということが図 4 から分かる. この戦略は他人の評判をいっさい考慮せず, 自分の評判をつねに 4 か 5 以上に保とうとする戦略である. なお, この戦略は高い協力率を実現するが, 自分の評判のみを考慮し, 評判の悪い相手を罰したりはしないため, 互恵的戦略ではない. また, 図 5 より, グループの人数が多いときは裏切的戦略 ($k > 0, h = -5$) が生き残る可能性が高いことが分かる.

つまり, image-OR の下では, グループの人数が少ないときは, 協力的だが互恵的ではない戦略が生き残る可能性が高く, グループの人数が多いときは, 裏切的戦略が生き残る可能性が高い. なお, standing-OR に関しても同様の傾向が見られた.

3.2.2 グループの人数と協力率の関係 (image scoring と standing 評判基準の比較)

次に, OR 行動基準の下での, グループの人数と協力率の関係に注目する. 図 6 より, image-OR, standing-OR の両ケースで, グループの人数の増加にともなって協力率が低下していることが分かる. ただし, 2 種類の評判基準を比べると, standing 評判基準の下での協力率が image scoring 評判基準の下での協力率を下回っている. つまり, OR 行動基準の下では, standing 評判基準を導入すると協力行動の進化が促進されず, 逆に阻害されるのである.

3.3 MIN 行動基準

3.3.1 生き残った戦略の分布

本節では, MIN 行動基準の下で「どのような戦略が生き残るのか」に着目する. image-MIN, standing-

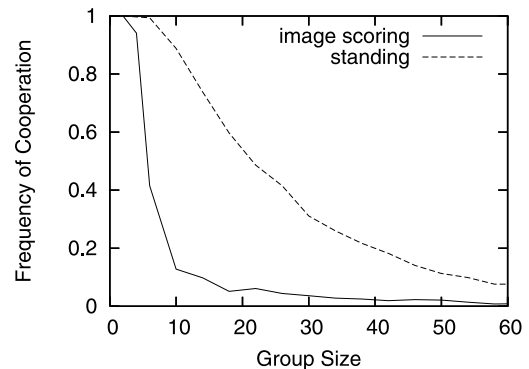


図 7 グループの人数と協力率の関係 (MIN 行動基準, 実線: image scoring 評判基準, 破線: standing 評判基準, シミュレーション 1,000 の回平均)

Fig. 7 Relationship between group size and the frequency of cooperation under MIN (the average values of 1,000 computer simulation runs, solid-line: under image scoring, dashed-line: under standing).

MIN において生き残る確率の高い戦略の傾向は DIS 行動基準におけるものと同じであった (図は省略). つまり, グループの人数が少ないときは, 互恵的協力戦略が生き残る確率が高く協力社会が実現され, グループの人数が多いときは, 裏切的戦略が生き残る確率が高く裏切社会が実現されるのである.

3.3.2 グループの人数と協力率の関係 (image scoring と standing 評判基準の比較)

次に, MIN 行動基準の下での, グループの人数と協力率の関係に注目する. 図 7 より, MIN 行動基準の下では, image scoring 評判基準, standing 評判基準の両ケースでグループの人数の増加にともなって協力率が低下していることが分かる. ただし, 両評判基準を比べると, standing 評判基準の下での協力率が image scoring 評判基準の下での協力率を上回っている. つまり, standing 評判基準を導入することで協力行動の進化が促進されるのである.

3.4 MAX 行動基準

3.4.1 生き残った戦略の分布

本節では, MAX 行動基準の下で「どのような戦略が生き残るのか」に着目する. image-MAX, standing-MAX においても, 生き残る確率の高い戦略の傾向は DIS 行動基準におけるものと同じであった (図は省略). つまり, グループの人数が少ないときは, 互恵的協力戦略が生き残る確率が高く協力社会が実現され, グループの人数が多いときは, 裏切的戦略が生き残る確率が高く裏切社会が実現される.

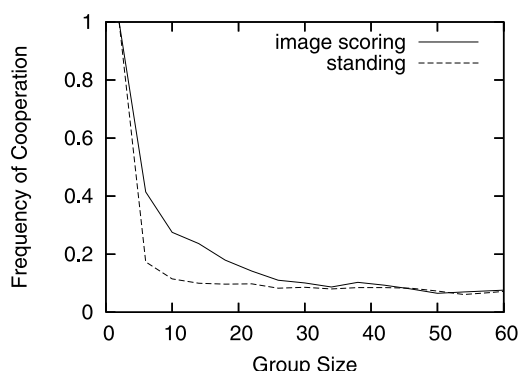


図 8 グループの人数と協力率の関係 (MAX 行動基準, 実線: image scoring 評判基準, 破線: standing 評判基準, シミュレーション 1,000 の回平均)

Fig. 8 Relationship between group size and the frequency of cooperation under MAX (the average values of 1,000 computer simulation runs, solid-line: under image scoring, dashed-line: under standing).

3.4.2 グループの人数と協力率の関係 (image scoring と standing 評判基準の比較)

次に, MAX 行動基準の下での, グループの人数と協力率の関係に注目する. 図 8 より, image-MAX と standing-MAX の両ケースで, グループの人数の増加にともなって協力率が低下していることが分かる. ただし, 2 種類の評判基準を比べると, standing 評判基準の下での協力率が image scoring 評判基準の下での協力率を下回っている. つまり, standing 評判基準を導入すると協力行動の進化が促進されず, 逆に阻害されるのである.

4. 議 論

我々は日常生活において, 他人に評判を付ける際に相手の行動の理由を思い詰ることがある. たとえば, ある人が裏切者に対して裏切り行動をとった場合「裏切者を罰したのだから正当な行為である」と判断することもありうる.

本研究では, Suzuki and Akiyama¹⁶⁾ の評判付き n 人囚人のジレンマゲームのモデルに「相手の行動の理由を思い詰る」という要素を組み込んだ. 具体的には, 裏切者に対しての裏切りを正当な行為と評価する standing 評判基準を組み込んだ.

standing 評判基準は 2 者間の協力行動の進化を安定化することがすでに示されている^{4), 8), 12), 13)} が, 多人数グループ内での協力行動の進化にどのような影響を与えるのかについては, まだ分かっていない. そこで, 本研究ではそれについて調べた. その結果, 以下のことが明らかになった.

4.1 グループの人数が増加すると協力行動の進化が困難になる

2 種類の評判基準と 4 種類の行動基準のすべての組合せで, グループの人数の増加にともなって協力関係の形成・維持が困難になることが分かった. 協力行動が進化するためには「協力的戦略を巻き添えにすることなく, 裏切的戦略には裏切ることで罰を与える」ことが必要である. 評判付き n 人ゲームでは特定の裏切的戦略に罰を与えることができず, グループの人数が増えれば増えるほど, 1 人の裏切的戦略に罰を与えるために多くの協力的戦略を巻き添えにしてしまう. よって, 協力行動の進化はグループの人数が増えると困難になる. つまり「他人の行動の理由を思い詰る評判基準 (standing 評判基準)」や「自分の評判を考慮する行動基準 (OR 行動基準)」を組み入れても, 多人数グループ内での協力行動の進化を完全に説明することはできないのである. 多人数グループ内での協力行動の進化を説明するためには, さらに別の行動基準や評判基準, もしくはまったく別のメカニズムを導入する必要がある.

また, グループの人数の増加にともなう協力率の低下の程度は, 各個人がどのような行動基準を採用しているのかによって異なっている. 各個人が「他人の評判の平均が良い, または, 自分の評判が悪いとき協力する」という行動基準 (OR) を採用しているとき, 協力率の低下は最も緩やかになる. そして, 各個人が「他人の評判の平均が良いとき協力する」という行動基準 (DIS) を採用しているときには, 協力率は比較的急激に低下する. さらに, 各個人が「最も評判の悪い他人の評判が良いとき協力する (MIN 行動基準)」, もしくは「最も評判の良い他人の評判が良いとき協力する (MAX 行動基準)」という基準を採用しているケースでは, 協力率は最も急激に低下する.

上記のような「協力率低下の程度の違い」を生じさせるメカニズムを明らかにするために, まず「他人の評判だけを考慮する 3 つの行動基準 (DIS, MAX, MIN)」の下における協力率低下の程度の違いについて考察する. 協力行動の進化には「協力的戦略を巻き添えにすることなく, 裏切的戦略には罰を与える」ことが重要だが, MAX 行動基準の下では各個人が行動を決める際「最も評判の良い相手」に注目するため, 十分に裏切的戦略を罰することができない. 反対に, MIN 行動基準の下では各個人が行動を決める際「最も評判の悪い相手」に注目するため, 必要以上に裏切的戦略に罰を与え多くの協力的戦略を巻き添えにしてしまう. 一方, DIS 行動基準の下では各個人が「他人の

評判の平均」に注目するため、裏切的戦略への罰（裏切り）が MAX 行動基準と MIN 行動基準の間くらい頻度で行われ、「裏切的戦略への罰」と「協力的戦略を巻き添えにしないこと」のバランスがとれる。その結果、協力率低下の程度は、MIN, MAX 行動基準の下で大きくなり、DIS 行動基準の下では小さくなる。最後に OR 行動基準の下では、各個人が「他人の評判の平均」に加え「自分の評判」も考慮する。この場合、他人の評判に関しては DIS 行動基準と同様だが、裏切的戦略を罰した協力的戦略は悪化した自分の評判をすぐに回復しようとする。そして「自分の評判」を考慮し回復させようとする分だけ、協力率の低下の程度が DIS 行動基準の下に比べ、緩やかになる。

4.2 standing 評判基準が image scoring 評判基準

と比べて協力行動を促進するとは限らない
2 者間の協力行動の進化を安定化することが知られている standing 評判基準が image scoring 評判基準と比べて、必ずしも多人数間の協力行動を促進するとは限らない、ことが分かった。

具体的には、standing 評判基準が多人数グループ内での協力行動の進化に与える影響について以下のことが明らかになった：(1) 各個人が「他人の評判の平均が良いとき協力する（DIS 行動基準）」、もしくは「最も評判の悪い他人の評判が良いとき協力する（MIN 行動基準）」という行動基準を採用しているケースでは、standing 評判基準を導入することで協力行動の進化を促進できる；(2) しかし、各個人が「他人の評判の平均が良い、または、自分の評判が悪いとき協力する（OR 行動基準）」、もしくは「最も評判の良い他人の評判が良いとき協力する（MAX 行動基準）」という行動基準を採用しているケースでは、standing 評判基準を導入することで協力行動の進化が逆に阻害される。

上記の結果が生み出されるメカニズムを説明するため、まず standing 評判基準導入の一般的な効果について考察する。image scoring 評判基準の代わりに、standing 評判基準を導入することには以下の 2 つの効果があると考えられる：

- 互恵的協力戦略の適応度を上げる：互恵的協力戦略は評判の悪い者（裏切的戦略である可能性が高い）には裏切ることで罰を与え、評判の良い者（協力的戦略である可能性が高い）に対しては協力する。image scoring 評判基準の下では、互恵的協力戦略の裏切的戦略に対する裏切り（罰）が不当な行為であると判断されてしまう。その結果、裏切的戦略を罰した互恵的協力戦略自身の評判が悪

化し、互恵的協力戦略間で「裏切りの連鎖」が発生してしまう。たとえば、裏切的戦略に罰を与えた互恵的協力戦略 A の評判が悪化することにより、別の互恵的協力戦略 B から裏切られ、その結果 B の評判も悪化して、別の互恵的協力戦略 C に裏切られ、今度は C が・・・というプロセスで裏切りの連鎖は生じる。その結果、互恵的協力戦略の適応度が低下する。一方、standing 評判基準の下では、裏切的戦略への罰（裏切り）は正当な行為であると判断される。そのため、互恵的協力戦略による裏切りの連鎖は生じない。よって、standing 評判基準は image scoring 評判基準と比べて、互恵的協力戦略の適応度を上げることができる。

- 裏切的戦略の適応度を上げる：裏切的戦略は評判の悪い者（裏切的戦略である可能性が高い）に対しても評判の良い者（協力的戦略である可能性が高い）に対しても裏切り行動をとる。image scoring 評判基準の下では、その裏切り行動が評判の悪化につながり、将来に他人から協力してもらう可能性の減少につながってしまう。一方、standing 評判基準の下では、裏切り行動のうち、評判の悪い者に対して行われたものが正当な行為と判断される。つまり、裏切的戦略の裏切り行動の一部が、将来に他人から協力してもらう可能性の減少につながらないのである。よって、standing 評判基準は image scoring 評判基準と比べて、裏切的戦略の適応度を上げてしまう。

前者の効果が強い場合は standing 評判基準を導入することで協力行動の進化が促進され、後者の効果が強い場合は standing 評判基準を導入することで協力行動の進化が阻害されると考えられる。

次に、standing 評判基準が DIS, MIN 行動基準の下では協力行動の進化を促進し、OR, MAX 行動基準の下では阻害するメカニズムについて考察する。

各個人が「他人の評判の平均が、ある閾値より良いとき協力する（DIS 行動基準）」、もしくは「最も評判の悪い他人の評判が、ある閾値より良いとき協力する（MIN 行動基準）」という行動基準を採用しているケースでは、協力社会は互恵的協力戦略によって実現される（3.1.1 項, 3.3.1 項参照）。よって、このケースでは、standing 評判基準を導入することで、互恵的協力戦略の適応度が増加し協力率が上昇した。

裏切的戦略の適応度も上がったが、互恵的協力戦略の適応度の増加分の方が上回った。

一方、各個人が「最も評判の良い他人の評判が、ある閾値より良いとき協力する (MAX 行動基準)」という行動基準を採用しているケースでは、協力社会は互惠的協力戦略によって実現されているにもかかわらず (3.4.1 項参照), standing 評判基準を導入することで協力行動の進化が阻害される。MAX 行動基準の下では image scoring 評判基準を採用していたとしても互惠的協力戦略による裏切りの連鎖は起こらない。MAX 行動基準の下で各プレーヤが意志決定を行う際、グループ・メンバの評判の最大 (最良) 値のみを考慮するため、ある 1 人の互惠的協力戦略の評判が悪化したとしても、別の互惠的協力戦略の行動にあまり影響を及ぼさないからである。互惠的協力戦略による裏切りの連鎖が起こらないので、MAX 行動基準の下では image scoring と standing のどちらの評判基準を採用していても互惠的協力戦略の適応度は変わらない。その結果、standing 評判基準導入の「裏切的戦略の適応度を上げる」という効果のみが効き、協力行動の進化が阻害される。

また、各個人が「他人の評判の平均がある閾値より良い、または、自分の評判がある閾値より悪いとき協力する (OR 行動基準)」という行動基準を採用しているケースでは、協力社会は互惠的協力戦略ではなく、協力的だが互惠的ではない (評判の悪い者を罰しない) 戦略により実現されていた (3.2.1 項参照)。このケースでは、standing 評判基準の持つ「互惠的協力戦略の適応度を上げる」という効果は効かず、「裏切的戦略の適応度を上げる」という効果のみが効いてくる。その結果、standing 評判基準を導入することで協力行動の進化が阻害される。

5. おわりに

近年、インターネットを介したオンライン市場が急速に発展している。しかし一方で、オンライン市場での不正行為 (e.g. 代金不払い、商品不渡しなど) が社会問題化している。現に「オンライン取引での不正行

為の数は実際の店舗で行われる取引に比べてきわめて多い」という調査結果も報告されている⁵⁾。

オンライン市場では、長期間の相互作用に基づく互惠的協調関係を築くことが難しい。取引の多くが「1 回限り」で、同じ相手と何度も取引が行われないからである。たとえば、インターネット・オークション・サイト eBay で行われる取引の 89% は「1 回限り」の取引である¹⁴⁾。このような状況では、長期間の相互作用を前提としない「評判を介した間接互惠的協調関係」が重要になってくる。実際、前述の eBay などでは、参加者の評判を公開することで間接互惠的協調関係の成立を促進し、不正行為を抑制している。それでは、インターネット社会のように多くの相互作用が 1 回限りであり、さらに相互作用が多人数間で行われるようなケースでは、どのような基準で評判を付けると参加者の自発的な協調行動を引き出すことができるのだろうか？

本研究では、image scoring と standing という 2 つの評判基準を比較して、上記の問題を考えた。計算機シミュレーションによる分析の結果、2 者間の協調関係を安定化することが知られている standing 評判基準が image scoring 評判基準と比べて、必ずしも多人数間の協調関係の成立を促進するとは限らないことが分かった。どちらの評判基準が協調関係の成立を促進するのかについては、各参加者が評判を用いてどのように意志決定をするのかによって変わってくるのである。

参 考 文 献

- 1) Alexander, R.D.: *Darwinism and Human Affairs*, University of Washington Press, Seattle (1979).
- 2) Axelrod, R.: *The Evolution of Cooperation*, Basic Books, New York (1984).
- 3) Boyd, R. and Richardson, P.J.: The Evolution of Reciprocity in Sizable Groups, *J. Theor. Biol.*, Vol.132, pp.337–356 (1988).
- 4) Brandt, H. and Sigmund, K.: The logic of reprobation: assessment and action rules for indirect reciprocity, *J. Theor. Biol.*, Vol.231, pp.475–486 (2004).
- 5) GartnerG2, Online transition fraud and prevention get more sophisticated.
www.gartnerg2.com/rpt/rpt-0102-0013.asp
- 6) Hardin, G.: The tragedy of commons, *Science*, Vol.162, pp.1243–1248 (1968).
- 7) Joshi, N.V.: Evolution of cooperation by reciprocity within structured demes, *J. Genet.*, Vol.66, pp.69–84 (1987).

なお、本研究は「インターネット社会のように多くの相互作用が 1 回限りであり、さらに相互作用が多人数間で行われるようなケース」における「協力行動の進化」、または「優れた戦略の模倣による協力行動の普及」に焦点をしばったものであり、現実社会の制度設計 (e.g. インターネット・オークション) などに直接応用することを意図したものではない。また、「世代がいつせいに代わり、評判が 0 に戻る」という設定も現実社会との対応を考えると不自然な点であり、今後、現実社会の制度設計に直結した研究を行ううえでは考慮すべき点である。ただ、本研究では既存の数学的モデル^{4),12),13)} との対応・比較をするために、「世代がいつせいに代わり、評判が 0 に戻る」という設定を採用した。

- 8) Leimar, O. and Hammerstein, P.: Evolution of cooperation through indirect reciprocity, *Proc. R. Soc. Lond.*, Vol.B 268, pp.745–753 (2001).
- 9) Milinski, M., Semmann, D. and Krambeck, H.J.: Reputation helps solve the Tragedy of common, *Nature*, Vol.415, pp.424–426 (2002b).
- 10) Nowak, M.A. and Sigmund, K.: Evolution of indirect reciprocity by image scoring, *Nature*, Vol.393, pp.573–577 (1998a).
- 11) Nowak, M.A. and Sigmund, K.: The dynamics of indirect reciprocity, *J. Theor. Biol.*, Vol.194, pp.561–574 (1998b).
- 12) Ohtsuki, H. and Iwasa, Y.: How should we define goodness? —reputation dynamics in indirect reciprocity, *J. Theor. Biol.*, Vol.231, pp.107–120 (2004).
- 13) Panchanathan, K. and Boyd, R.: A tale of two defectors: the importance of standing for evolution of indirect reciprocity, *J. Theor. Biol.*, Vol.224, pp.115–126 (2003).
- 14) Resnick, P.R. and Zeckhauser, R.: Trust among strangers in internet transactions: Empirical analysis of ebay's reputation system, *The Economics of the Internet and E-Commerce*, Baye, M.R. (Ed.), Vol.11. *Adv. Appl. Microeconomics*, Elsevier Science, Amsterdam, The Netherlands.
- 15) Sugden, R.: *The Evolution of Rights, Cooperation and Welfare*, Oxford: Blackwell (1986).
- 16) Suzuki, S. and Akiyama, E.: Reputation and the evolution of cooperation in sizable groups, *Proc. R. Soc. Lond.*, Vol.B 272, pp.1373–1377 (2005).
- 17) Suzuki, S. and Akiyama, E.: Reputation and the Evolution of Cooperation in Sizable Groups, *Proc. 9th Workshop on Economics and Heterogeneous Interacting Agents (WEHIA-04)* (2004).
- 18) Trivers, R.: The evolution of reciprocal altruism, *Q. Rev. Biol.*, vol.46, pp.35–57 (1971).
- 19) Wedekind, C. and Milinski, M.: Cooperation Through Image Scoring in Humans, *Science*, Vol.288, pp.850–852 (2000).

(平成 17 年 9 月 30 日受付)

(平成 18 年 3 月 2 日採録)



鈴木 真介

昭和 55 年生。平成 17 年筑波大学大学院システム情報工学研究科社会システム工学専攻博士前期課程修了。現在、同博士後期課程に所属。日本数理生物学会、日本認知科学会

各会員。



秋山 英三 (正会員)

昭和 44 年生。平成 10 年東京大学大学院総合文化研究科広域科学専攻博士課程修了。博士(学術)を取得。平成 13 年 8 月から筑波大学社会学系講師。現在、筑波大学大学院システム情報工学研究科助教授。日本物理学会、日本経済学会、各会員。