

## ゲーム性向「金持ち喧嘩せず」は進化するか？

—貯蓄とゲーム参加コストを付加した空間型囚人のジレンマゲームに観る複雑性の創発—

## Will the Game Propensity "Wealthier does not fight" evolve? : The Emergence of Complexity in a Spatial Prisoner's Dilemma with Resource-storing Mechanism and Game Participation Costs

内海 忍<sup>1</sup>, 晨 申<sup>1</sup>, 立川 雄一<sup>1,2</sup>, Md. Rajib Arefin<sup>3</sup>, 谷本 潤<sup>1,4</sup>  
 Shinobu Utsumi, Chen Shen, Yuichi Tatsukawa, Md. Rajib Arefin, Jun Tanimoto

## 概要

本研究では空間型囚人のジレンマゲームにおける標準的モデルに次の3点の拡張—(i)ゲーム参加コスト, (ii)利得の貯蓄機構, (iii)ゲーム性向；周囲よりも自身の総利得が大きい場合に次回ゲームへ参加 (or 不参加) する機構—を付加し, 通常の戦略更新に加えゲーム性向の更新も行う共進化モデルを構築した. これは, 協調か裏切りかという対戦相手の表面的な行動の模倣に留まらず, その裏に潜む相手の思考をも模倣する状況の再現を意図している. 解析結果からは, モデル構造の単純さに比して, ネットワーク互恵の再現に特化した通常空間型囚人ジレンマゲームでは観られない, 大変豊かな複雑性の創発が観察された. 本稿ではその中でも最も興味深い結果の一つ (最終的には同じ協調率に行き着く場合であっても, 貯蓄可能比率と参加コストの程度によって, その背後にある各個体が持つゲーム性向—金持ち喧嘩する or 金持ち喧嘩しない—は全く逆の方向に進化する) に焦点を絞り, 報告する.

## 1. 緒言

利己的な個体が構成された集団における利他的行動の発現は, 人間社会のみならず自然界の凡ゆる生物種に観察される普遍的現象である. この協調的群行動の自己組織化は, 生物学をはじめ, 経済学, 情報学, 人工知能など幅広い分野からの関心を集めてきたが, そのメカニズム解明の数学的枠組みとして進化ゲーム理論 (Evolutionary Game Theory, EGT) が適用されてきた[1–2]. EGT では, 現実の集団内の interaction を数理モデル等を用いて抽象的に表現するが, 各プレイヤーがゲームに関する利得関数を事前に把握し適切な戦略を“先回り”して分析し, 対戦するという一般的な非協力ゲーム理論の仮定に対し, EGT では繰り返されるゲームの結果 (利得) に応じて“後天的に”戦略を更新・保持していく個体を仮定する. そして, その個体によって構成された集団内の戦略分布に関する進化ダイナミクスを分析する. EGT では 1990 年代以降の急速な計算機の発達に伴い, マルチエージェント系によるシミュレーション (Multi Agent Simulation, MAS) —コンピュータ内に多数の個体 (エージェント) を発生させ, ローカルに生じる相互作用をボトムアップで設計していき, グローバルに起き

る現象を再現する—技法と組み合わせることで, これまでも数多くの研究成果が報告されている[3–4].

EGT を含め, ゲーム理論の代表例である囚人のジレンマ (Prisoner's Dilemma Game, PD) は, 現実社会の相互関係を抽象化したモデルとして理論的, 実験的観点から注目されてきた. PD では, 社会全体の利益最大化の為には全エージェントの協力 (Cooperation, C) が最適戦略であるにもかかわらず, 個人の利益最大化には常に裏切り (Defection, D) が最適戦略となるため, 個人視点と集団視点で最適な選択が乖離するというジレンマが生じる. その結果, 社会粘性の無い well-mixed な無限集団では全員が裏切り合う状況に行き着いてしまう. 当研究領域では, このジレンマを緩和・解消し, 社会全体の利益を最大化する為にはどのような機構を付加する必要があるかに関して, これまでに数多くの提案がなされてきた. Nowak によれば, これらの協調創発機構はすべて 5 つに分類できるとしているが[5], 特にそのうちの一つである“ネットワーク互恵”に関しては, 1992 年の空間型の PD (Spatial Prisoner's Dilemma Game, SPD) をはじめ, 多くの研究業績が蓄積されている[3–6].

近年の SPD 研究において新たに検討されている機構の一つとして, 利得の貯蓄メカニズムが挙げられる[7–9]. これは戦略更新において, これまで参照していた各エージェントの瞬時利得ではなく, 時間方向に瞬時利得を蓄積させた総利得を参照するとのアイデアである. この場合, END 期間 (e.g., [10–11]) に D 戦略による C 戦略クラスター (C-cluster) への搾取に対して耐性が生まれる為, 最終的にはより高い協調率となることが報告されている[7–9]. また, 利得貯蓄機構の他にも, 近年ではゲーム参加コストに焦点を当てた研究も注目されている[12–14]. 参加コストでは, 社会活動への参加に伴う利益損失の潜在的なリスクや, 所得税といった利得に対する一定の制限を抽象化して表現している場合が多い. よって, この二点 (貯蓄と参加コスト) は現実社会への比喩を考えたときに, 不自然な仮定ではなく, 寧ろ加味されて然るべき拡張であると考えられる.

Wang らが報告した[14]では空間型の公共財ゲーム (Public Goods Game, PGG) に関して, 利得貯蓄と参加コストを同時に加味したモデルを報告している. [14]では PGG への参加は集団内で比較的高利得を得ているエージェントのみに許され, また, 参加する場合には一定のコストが伴うという仮定が置かれている. つまり, 高利得を獲得している個体が, コストを払って更にゲームに参加するという, いわば“金持ち喧嘩する”の仮定が前提にある. これは, 富める者がその資金やノウハウを基に更に多くに財

1. 九州大学 大学院総合理工学府
2. エム・アール・アイ リサーチアソシエイツ株式会社
3. ダッカ大学 理学部 数学科
4. 九州大学 総合理工学研究院 環境工学部門

を築こうとするマネーゲームのような状況では顕著に観察されるゲームに対する意向性かもしれない。ところが、日本の諺にもあるように、これとは全く逆の発想“金持ち喧嘩せず”というゲーム性向も現実社会では考えら得る。これは、富める者は既に自分の所有する豊かな財産に満足しているために、更なる資金獲得のリスクを取るよりも、リスクを避けて守りに入ることを是とする思考である。実際に、このゲーム性向を仮定した SPD 研究も[15]のように報告されているが、一体どちらのゲーム性向の仮定が現実社会のモデル化において適切なのだろうか。また、このゲーム性向の差が、これまで報告されてきた SPD での進化ダイナミクスにどのような影響を与えるのだろうか。

著者らは、上記問いに動機付けを得て[6]の研究成果を報告している。[6]では、上記二種のゲーム性向の違いが、社会全体の協調率の進化にどう影響するかを解明すべく、SPD の標準的モデルをテンプレートに、参加コストと貯蓄機構を同時に加味した上で、戦略分布の進化ダイナミクスを MAS を用いて解析している。具体的には、社会全体で単一のゲーム性向しか持たない 3 つの社会；①全構成員が金持ち喧嘩“する”に基づくゲーム性向を持つ場合、②全構成員が金持ち喧嘩“せず”のゲーム性向を持つ場合、③ゲーム性向を持たない場合（全エージェントが総利得の大小に依らず、常に確率 1/2 で次回ゲームに参加）、に分類して同様の解析を行った結果を報告している。

しかし、結局のところ、現実社会のゲーム性向にはどちらを前提とすべきか。もしくは、両ゲーム性向を集団内に仮定した場合は、どちらかのゲーム性向が淘汰され、もう一方が生き残るのか。それとも両ゲーム性向が共存する社会に辿り着くのだろうか？

本研究ではこの本質的な問いに対する理論的な解の提示を目的とする。具体的には、上記[6]で拡張したモデルに対し、更に 1 点だけ拡張を加えたモデル、すなわち、通常の戦略更新に加えゲーム性向も更新してゆく、「共進化モデル」を構築し、MAS による解析を行なった結果を報告する。これは、協調か裏切りかという相手の表面的な行動のみを模倣するのではなく、その裏に潜む相手の思想も模倣する状況の再現を意図する。貯蓄と参加コスト、及びゲーム性向の共進化という 3 点を SPD の標準モデルに付加した本モデルの解析結果からは、モデルの単純さに比して、ネットワーク互惠の再現に特化した通常的空間型囚人ジレンマゲームではこれまで報告されてこなかった複雑で豊かな進化のダイナミクスが数多く観察された。本稿ではその中でも興味深い結果に焦点を絞り、報告する。

## 2. モデル

### 2-1. ゲーム

本研究の SPD モデルで想定するゲームには 2 人 2 戦略モデルを適用する。すなわち、各エージェントは 2 人組で対戦し、各々協調( Cooperation, C)か裏切り( Defection, D)の 2 戦略のどちらかを選択する。このとき、両プレイヤーは自身と対戦相手の戦略の組み合わせに応じた利得を獲得する。両者共に C 戦略の場合は、どちらも利得  $R$  (Reward) を、両者共に裏切り D 戦略の場合はどちらも利得  $P$  (Punishment) を得る。一方が C 戦略、他方が D 戦略の場合は、前者は利得  $S$  (Sucker) を、後者は利得  $T$  (Temptation) を獲得する。以上を利得行列で表現すると、次式；

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix} = \begin{pmatrix} 1 & -D_r \\ 1 + D_g & 0 \end{pmatrix}, \quad (1)$$

と表現できる。ここで、 $D_g$  はチキン型ジレンマ強度（相手を食らうとするジレンマ, Gamble intending dilemma),  $D_r$  は鹿狩り型ジレンマ強度(相手に食られまいとするジレンマ, Risk averting dilemma)であり、先行研究[3, 16–17]に倣い夫々  $D_g = T - R$ ,  $D_r = P - S$  で定義した。  $D_g$  のみが正となる利得構造では Chicken 型 (Snowdrift 型) のゲームクラスとなり、  $D_r$  のみが正となる時は Stag Hunt 型、共に正となる時は Prisoner's Dilemma 型 (PD), 共に負となる時はジレンマの無い Trivial のゲームクラスに分類される。式(1)右辺では、左辺の利得  $R$  を 1 に、  $P$  を 0 に固定し、  $D_g, D_r$  を用いて書き換えた利得行列であり、 4 つのゲームクラスを 4 変数  $(R, S, T, P)$  から 2 変数  $(D_g, D_r)$  で、一般性を失わずに取り扱いが可能となる。一般に PD 型のゲームクラスとは、  $T > R > P > S$ , 且つ、  $2R > T + S$  の領域、すなわち、  $0 \leq D_g(D_r) \leq 1$  の領域であるため、本稿ではジレンマ強度が共に正となる領域を考察対象とする。

### 2-2. ネットワーク

基盤グラフを全人口  $N = 10^4$  ( $100 \times 100$ ), 平均次数  $(k) = 8$  の 2 次元正方格子 (Moore 近傍) に採る。ここで、正方格子の端辺は向かい合う端辺同士で接続し、全てのエージェントが正確に 8 人の隣接エージェントを持つように設定する (周期的境界条件)。ノードをエージェントに、リンクを対戦ペアに見立て、各エージェントはゲームに参加する場合近接する 8 人のエージェントと 2-1. に示した対戦を行う。

### 2-3. ゲーム参加コストと利得貯蓄

緒言に記したように、本研究では参加コストと利得貯蓄を加味した上での、戦略とゲーム性向の共進化ダイナミクスを観察する。よって、参加コストを  $\sigma$  とすれば、本モデルでエージェント  $i$  が時刻  $t$  に獲得する瞬時利得  $\pi_i^t$  は次式；

$$\pi_i^t = \begin{cases} \sum_{j \in \{N_i\}} p_{i-j}^t - \sigma & (\text{if } A_i \text{ participates the game}) \\ 0 & (\text{Otherwise}) \end{cases}, \quad (2)$$

で表現できる。ここで、  $\{N_i\}$  はエージェント  $i$  が持つ隣人の集合であり、  $p_{i-j}^t$  はエージェント  $i$  と  $j$  が時刻  $t$  に対戦した場合に  $i$  が獲得する利得を意味する。また、エージェント (プレイヤー) がゲームに参加しない場合は、利得獲得の機会を一切の持たない代わりに、参加コストを支払う必要も無いとする (下段、  $\pi_i^t = 0$ ) 。

次に、利得貯蓄のメカニズムを導入する。次の時間ステップへの貯蓄可能比率を  $s$  ( $0 \leq s \leq 1$ ) とすれば、エージェント  $i$  のこの時間ステップ  $t$  における最終的な総利得  $\Pi_i^t$  は次式；

$$\Pi_i^t = s \cdot \Pi_i^{t-1} + \pi_i^t, \quad (3)$$

のように記述できる。ここで、  $\Pi_i^{t-1}$  は、エージェント  $i$  が前時間ステップ  $(t-1)$  に所持していた総利得である。つまり、  $s = 0$  の場合は各時間ステップで瞬時利得のみが勘定される従来型の SPD を再現し、  $s = 1$  の場合は各時間ステップで獲得した利得を全て次の時間ステップにキャリーオーバー可能な状況を再現する。

## 2-4. ゲーム参加確率とゲーム性向

各エージェントが自身の持つゲーム性向—金持ち喧嘩する or 金持ち喧嘩せず—に応じて、次回ゲームへ参加 or 不参加を判断する機構を導入する。エージェント  $i$  が時刻  $t$  で自身の総利得  $\Pi_i^t$  と隣接エージェントの平均総利得  $(\Pi_j^t)_{j \in \{N_i\}}$  とを非線形的に比較考量し、次ゲーム (時刻  $t+1$ ) への参加を判断すると考えれば、 $i$  がゲームに参加する確率  $P_{game,i}^{t+1}$  は、人間の意思決定プロセスを模擬する関数の一つとして SPD で戦略更新の際に適用される Fermi 関数を応用して、

$$P_{game,i}^{t+1} = \frac{1}{1 + \exp[\omega_{game,i}^t \cdot ((\Pi_j^t)_{j \in \{N_i\}} - \Pi_i^t)]} \quad (4)$$

と表現できる。ここで、式(4)中の感度パラメータ  $\omega_{game,i}^t$  こそが、本稿の要となるゲーム性向である。 $\omega_{game,i}^t$  はエージェント  $i$  が持つ時刻  $t$  までに進化したゲーム性向を表し、 $\omega_{game,i}^t > 0$  の場合は、自身が周囲と比較して高利得の場合  $([(\Pi_j^t)_{j \in \{N_i\}} - \Pi_i^t] < 0)$  に、 $P_{game,i}^{t+1} > \frac{1}{2}$  で、ゲームに参加しやすくなる (i.e., 金持ち喧嘩する)。一方、 $\omega_{game,i}^t < 0$  の場合は、自身が周囲と比較して高利得の場合  $([(\Pi_j^t)_{j \in \{N_i\}} - \Pi_i^t] < 0)$  に、 $P_{game,i}^{t+1} < \frac{1}{2}$  で、ゲームに参加しにくくなる (i.e., 金持ち喧嘩せず)。 $\omega_{game,i}^t = 0$  の場合は周囲との利得差に関わらず、ゲーム参加確率は常に  $P_{game,i}^{t+1} = \frac{1}{2}$  (常に次回ゲームに 50% の確率で参加) となり、ゲームに対する性向性を持たない。

## 2-5. 戦略更新・ゲーム性向更新

戦略、及びゲーム性向の更新に関しても Pairwise-Fermi を適用する。すなわち、エージェント  $i$  は時刻  $t$  における自身の総利得  $\Pi_i^t$  と隣接エージェントからランダムに選んだエージェント  $j$  の総利得  $\Pi_j^t$  を非線形的に比較考量し、エージェント  $j$  の方が高利得の場合は、エージェント  $j$  の持つ戦略、もしくはゲーム性向を高確率でコピーする。これを Fermi 関数の表式に当てはめれば、

$$P_{i \leftarrow j} = \frac{1}{1 + \exp\left[\frac{\Pi_i^t - \Pi_j^t}{\kappa}\right]} \quad (5)$$

となる。ここで、 $\kappa$  は Fermi 関数における熱力学温度であり、先行研究 (e.g., [3, 16–17]) に倣って、広く受け入れられている値の 0.1 に設定した。更新プロセスの実行に関しては、全エージェントが一斉に戦略・ゲーム性向を更新するシンクロ更新を用い、戦略更新とゲーム性向更新では、夫々独立に、参照する隣接エージェント  $j$  を選択する。これは、戦略更新がその時間ステップでゲームに参加したエージェントのみに許される行為であるのに対し、ゲーム性向の更新は毎時間ステップで全てのエージェントに許される為である。

## 3. 計算機実験

マルチエージェントによる計算機実験の手順は以下の通りである。はじめに、ネットワーク空間中に同数の戦略エージェントをランダムで配置し、進化計算を開始する。初期条件として、各エージェントのゲーム開始時の総利得は  $\Pi_i^0 = 0$  に、ゲーム参加確率は  $P_{game,i}^1 = \frac{1}{2}$  に、ゲーム性向値  $\omega_{game,i}^0$  は、一様乱数からランダムに振られた  $-10$  から  $+10$  までのいずれかの整数を付与する ( $\omega_{game,i}^0 \in \{x \in \mathbb{Z} \mid |x| \leq$

$10\}$  の、計 21 通りのゲーム性向値)。従って、初期条件での全エージェントの平均ゲーム性向  $\langle \omega_{game,i}^0 \rangle_{i \in \{N\}}$  は凡そ 0 に近い値をとる。

その後、協調率  $f_c^t$  と全エージェントの平均ゲーム性向値  $\langle \omega_{game,i}^t \rangle_{i \in \{N\}}$  が擬似均衡に達するまで進化計算を継続する。擬似均衡到達の判定は、協調率・平均ゲーム性向値の 100 時間ステップ平均値が、前回 100 時間ステップ平均値と比して協調率誤差 0.25% 以下、平均ゲーム性向値誤差 5% 以下の両方を同時に満たす場合とした。擬似均衡到達までを 1 エピソードとした場合の MAS 数値実験のアルゴリズムを図 1 に示す。また、結果が初期戦略配置や初期ゲーム性向値分布の影響を受けぬよう、独立した 100 エピソードのアンサンブル平均を以って 1 パラメータセットのシミュレーション結果とし、統計的頑健性を担保する。

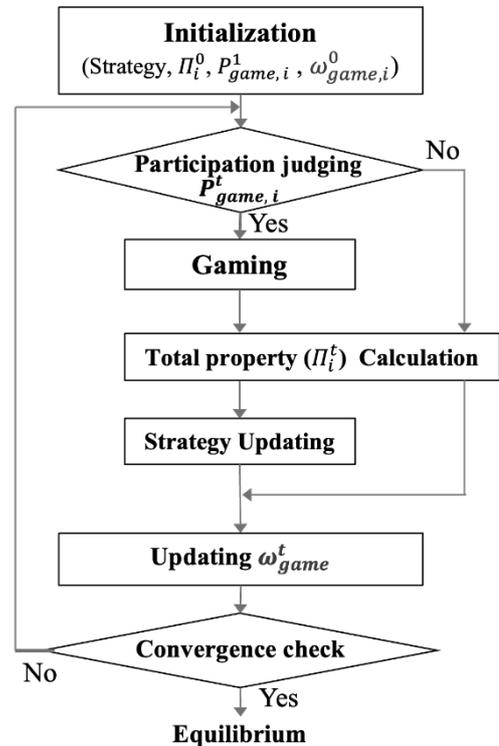


図 1 MAS 数値解析 (1 エピソード) の模式図

## 結果と考察

2 種のジレンマ強度 ( $D_g, D_r$ ) を系統的に変化させながら、9 通り (3 通りのゲーム参加コスト  $\sigma \times 3$  通りの貯蓄可能比率  $s$ ) の PD 型ゲームクラスに関して MAS 解析を行った結果を図 2 に示す。中央から左側の 9 パネルは各  $(\sigma, s)$  の組み合わせにおける協調率の結果 (各パネル左上のラベルに \*1 と記載) を、右側 9 パネルは集団の平均ゲーム性向値の結果 (各ラベルに \*2 と記載) を、横軸  $D_r$ 、縦軸  $D_g$  の強度マップ形式で表示している。左右 9 パネルはどちらも上段  $\sigma = 0$ 、中段  $\sigma = 1$ 、下段  $\sigma = 4$ 、左列  $s = 0.0$ 、中央列  $s = 0.2$ 、右列  $s = 0.8$  の結果を配置している。よって、上段左列  $(\sigma, s) = (0, 0)$  の設定は従来の SPD 研究における標準的モデルに (ゲーム性向の共進化部分を除いて) 最も近い状況を再現する。実際に  $(\sigma, s) = (0, 0)$  の協調率の結果を確認すると (a-1 参照)、ジレンマ強度の小さい領域でのみ協調が促進されているが (パネル内左下の青~緑領域)、これは先

行研究で報告されている内容とも一致し, “ネットワーク互惠”の影響であることが解明されている. (補足: 著者らの[6]では, 本モデルから共進化部分も除いた通常の戦略更新のみ行った結果を示しているが, 各時間ステップでのゲーム参加者数が標準モデルの約半数の場合であっても, 最終的な協調率は従来のSPDを正確にリカバーする事を報告している). このネットワーク互惠に拠る協調促進は, (a-1)に限らず, その他全ての結果について当てはまり(\*-1参照), その領域の集団の平均ゲーム性向値の進化に注目すると(\*-2参照), 常に高 $\omega_{game}$  (図中緑色) —金持ち喧嘩する—側へ進化していることも確認できる. このゲーム性向進化の事由は次節で解説するが, その前に図2の結果全体を俯瞰することで観察できる興味深い傾向について指摘しておく.

今, 協調率の進化は“参加コストと貯蓄の程度に依らず, 弱ジレンマ強度でのみ協調が促進され (図中左下領域), その他の殆どの領域ではD戦略が支配的となる裏切り合いの社会に陥る (図中赤色領域)”と大雑把に解釈できる. が, ゲーム性向の進化に注目すると, “参加コストと貯蓄のうち一方しか考慮されない場合は基本的に高 $\omega_{game}$  (図中緑色) —金持ち喧嘩する—側へ集団全体として進化するが, 両方が同時に加味される場合は, 高ジレンマ強度で真逆のゲーム性向: 低 $\omega_{game}$  (図中紫色) —金持ち喧嘩せず—側へ全体として進化してゆく”との傾向が観察される. 次節からは, 各解析結果に関するより詳細な情報を参照する事で, このゲーム性向の進化ダイナミクスを紐解く.

#### 4-1. 利得貯蓄のみ(a-c)

##### 4-1-1. 協調率の進化

まず, 参加コストを無視した場合の協調率の結果から見ていく. 図3には $(\sigma, s) = (0, 0)$ の結果うち2点 (図2内の点 a-I;  $(D_g, D_r) = (0.05, 0.05)$  と, 点 a-II;  $(D_g, D_r) =$

$(0.50, 0.90)$ ) のある1エピソードにおける4つの統計量の時系列データを表示している. 4つの統計量とは, 集団内の平均ゲーム性向値 $\langle \omega_{game} \rangle$  (灰色実線), 協調率 $f_c$  (緑実線), 空間内のCクラスター境界に位置するCエージェント (以下, Boundary\_C) の平均総利得 $\Pi_{ave}(Boundary\_C)$  (青線), Cクラスター境界に接するDエージェント (以下, Boundary\_D) の平均総利得 $\Pi_{ave}(Boundary\_D)$  (赤線) である. 各エージェントが境界に位置するか否かは, エージェントがC-Dリンクを一本でも持つか否かで判定した. 時系列データを見ると, 上段の(a-I)では $\Pi_{ave}(Boundary\_C)$ が $\Pi_{ave}(Boundary\_D)$ より常に高い値を示し, 全体の協調率は初期に減少した後増加していく. これはCクラスター境界の対戦(C-Dリンク)において, 平均総利得の高いC戦略がより周囲にコピーされる可能性が高いため, 当然の帰結であると云えよう. 故に, (a-II)の結果を見るとその逆の傾向が明瞭に観察できる;  $\Pi_{ave}(Boundary\_D)$ が $\Pi_{ave}(Boundary\_C)$ より常に高いため, C-Dリンクの対戦ではDがコピーされ易く, 結果としてCクラスターがDに駆逐され, 協調率は増加しない.

##### 4-1-2. ゲーム性向の進化

次に, ゲーム性向の進化に注目する. 再度図2を確認すると, 利得貯蓄のみ場合は集団の平均ゲーム性向値はPD内全域において高 $\omega_{game}$  —金持ち喧嘩する—側に収束していることが確認できる(a-2-c-2参照). この理由は, 以下の図4に示す状況がマイクロで起きている為と推察する. 図4には, D戦略によるC-clusterの侵食によって, C-clusterと接していたD戦略(Boundary\_D)が周囲を他のD戦略に囲まれ, 段階的にD-cluster内部のD戦略(以下, Inside\_D)へ移り変わる様子を示している. まず, 時刻 $t$ で, エージェント $i, j$ のようなBoundary\_Dを考える. これら

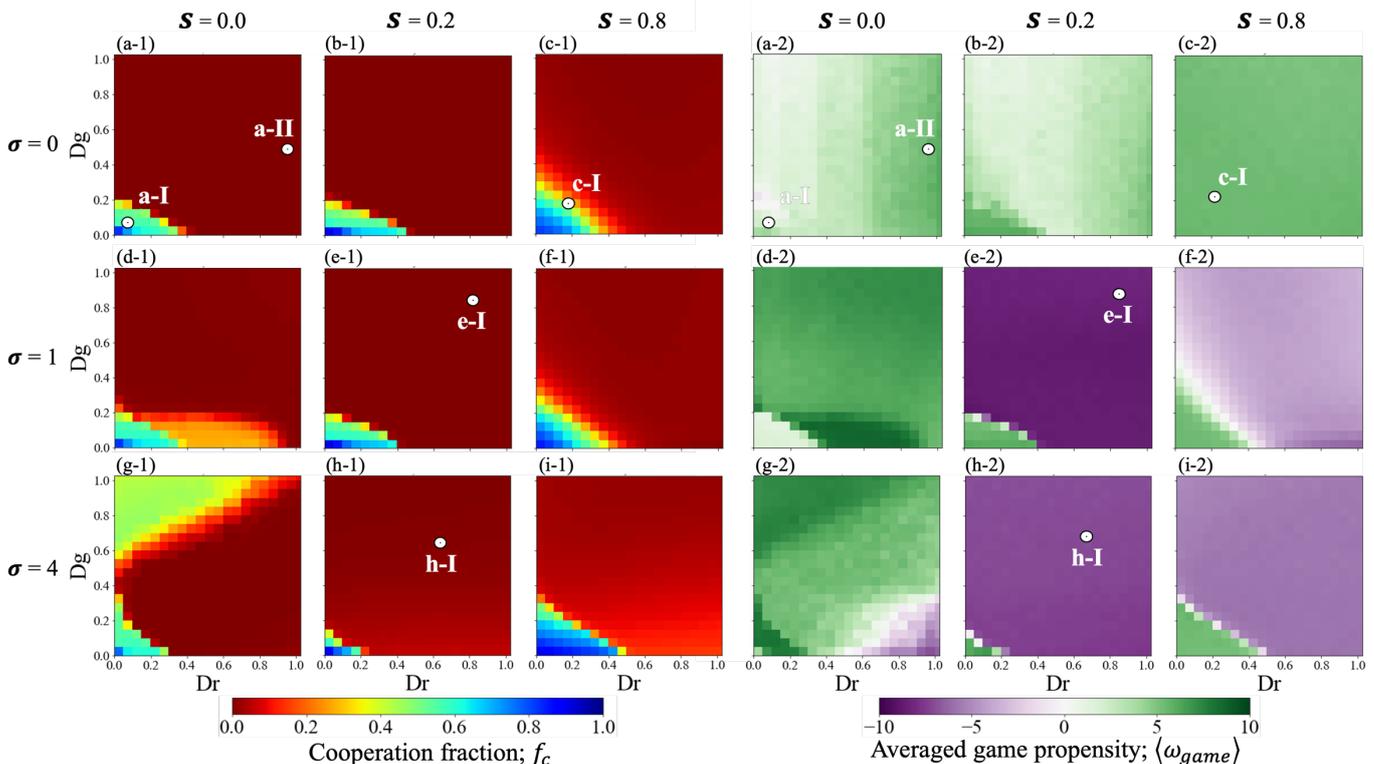


図2 Multi Agent Simulationによる協調率と平均ゲーム性向値の解析結果

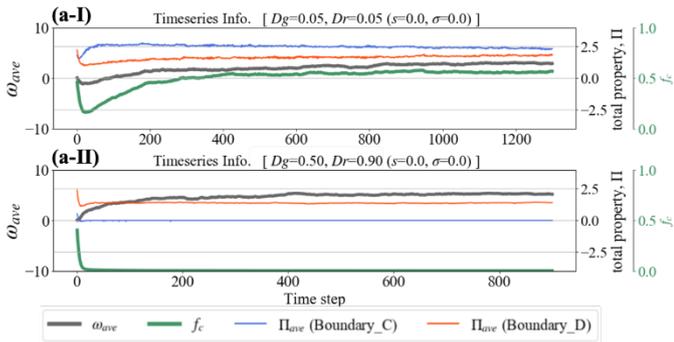


図 3 a-I, a-IIにおける統計量の時系列データ

Boundary\_D はゲームに参加すればどちらも  $3D_g$  の瞬時利得を得ることができ、周囲 (C とのリンクを持たない Inside\_D や、Boundary\_D に食らわれている Boundary\_C) よりも高利得となる。各エージェントは初期に一樣乱数で振られた  $-10$  から  $+10$  までのランダムな  $\omega_{game,i}^0 \in \{x \in \mathbb{Z} \mid |x| \leq 10\}$  を与えられている為、図 4 のように対照な周囲状況にあるエージェント  $i, j$  でも高  $\omega_{game}$  を保持して Boundary\_D になる場合 ( $i$ ) と、低  $\omega_{game}$  を保持して Boundary\_D になる場合 ( $j$ ) の 2 通りが考えられる。このとき、 $i$  は金持ち喧嘩“する”に従い、高確率で次の時刻  $t+1$  でもゲームに参加し、縮小した C-cluster から更に  $2D_g$  を稼ぐが、 $j$  は金持ち喧嘩“せず”に従い、時刻  $t+1$  ではゲームに参加せず、前回総利得を  $s$  だけ割り引いた総利得をキャリアオーバーする。すると、時刻  $t$  では総利得が同点であった  $i, j$  にも時刻  $t+1$  で利得差が生じる。今、戦略更新はゲームに参加したエージェントのみ許されるが、ゲーム性向の更新は全てのエージェントに許されるため、ゲームに参加しなかったエージェント  $j$  も周囲から自分より高利得のエージェントを参照した場合には、式(5)に従ってそのエージェントの持つゲーム性向値を高確率で模倣する。故に、 $j \leftarrow i$  のゲーム性向のコピー (より高  $\omega_{game}$  のコピー) が起き得る。換言すれば、参加コストが無い状況では、縮小する C-cluster から最後まで C を貪り尽くそうとする個体ほど高利得を維持できることに起因して、高  $\omega_{game}$  の拡大が起きるとのロジックである。これが集団内でゲーム性向値の平均が増加していく (金持ち喧嘩“する”側に社会が進化していく) メカニズムであると考察する。

#### 4-1-3. 総利得の分布

均衡での総利得分布等を参照し、参加コスト無しの場合には最終的にどのような社会構造に至るかを確認する。代表例として  $s = 0.8$  の結果を参照する。これは、高い貯蓄可能比率  $s$  によってエージェント間の総利得の差が拡大し、分布をより顕著に観察できる為である。図 5 には  $(\sigma, s) =$

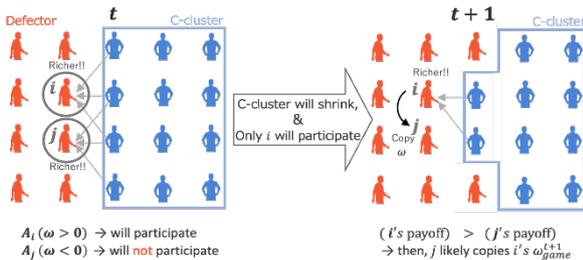


図 4 集団の平均ゲーム性向値増加のメカニズム

(0, 0.8) の結果のうち、図 2 に示す点  $c-I (D_g = 0.15, D_r = -0.15)$  のある 1 エピソードの均衡における総利得分布 (左の 3 ボックス) と戦略・ゲーム性向のスナップショット (右の 2 つ) を表示している。図 5 左側の 3 ボックスは、上段; 戦略別の総利得分布、中段; ゲーム性向別の総利得分布、下段; 戦略別のゲーム性向分布である。まず、中段・下段ボックスより、ゲーム性向は全体として正に進化している事が確認できるが、階級毎の内訳に特徴を持たないことから、ゲーム性向は戦略と総利得に対して相関を持たないことが確認できる。興味深いのは上段の戦略別総利得のヒストグラムである。今、集団内の総利得分布に注目すると、最も高利得を獲得しているエージェントの属性は、裏切りによって高利得を得ようとする D 戦略ではなく、C 戦略である。また、このヒストグラムの上部に表示している特徴量、CCLF\_rich (総利得上位 5% の C 戦略エージェントに関して、彼らの持つリンクのうち何割が C-C リンクかを示す指標、 $0 \leq \text{CCLF\_rich} \leq 1$ ) に注目すると、 $\text{CCLF\_rich} = 1.000$ 、つまり、この C 戦略エージェント達は、周囲を完全に C 戦略で囲まれていると判る (i.e., Inside\_C)。今、個々の持つゲーム性向値は集団全体で平均すると高  $\omega_{game}$ —金持ち喧嘩する—側のため、均衡ではこの Inside\_C は「高利得→更にゲームに参加→更に周囲の C から利得を獲得→高利得を維持→再度ゲームに参加→…」の繰り返しにあると考えられるが、この C-cluster の境界に位置する C (Boundary\_C) は Boundary\_D に食られ、また、Inside\_C よりも高利得を得ることは難しい為、常に比較的 low 利得となっている可能性が高い。高  $\omega_{game}$  “金持ち喧嘩する”、は裏を返せば、“金持ちでないなら喧嘩しない”、つまり、“貧乏喧嘩せず”である為、Boundary\_C は「周囲より比較的 low 利得→ゲーム不参加→戦略更新できず C 戦略を保持したまま→フリーズ」の状況にある可能性が高い。その証拠として、図 5 上段のヒストグラムでは C エージェントのおよそ半数は総利得 0 を示している。

小括すると、Boundary\_C は Boundary\_D に搾取され続けるが Inside\_C には協力し続ける、いわば岩の様な働きをしている。しかし、局所的に見れば裏切りによる搾取が起きているものの、社会全体で見れば最も高利得となるのは裏切り戦略ではなく、協調戦略であるという点で、この社会構造はそれなりに健全な構造とも云えよう。これは先行研究[6]に報告している内容とも重なる。

#### 4-2. ゲーム参加コストのみ(d,g)

紙面の関係上、本紙での説明は割愛するが、この条件設定から観察された興味深い結果としては、

- (d) :  $\Pi_{ave}(\text{Boundary\_C}) > \Pi_{ave}(\text{Boundary\_D})$  の場合でも協調率が低下する (4-1-2. に記述した協調率増加ロジックの唯一の例外)。また、C 戦略が紐状に並んで生き残る。
- (g) : 協調率の収束後に平均ゲーム性向値が増加していく。また、パネル左上領域でも協調が進化。

#### 4-3. 利得貯蓄+ゲーム参加コスト(e-f, h-i)

本節以降では、参加コストと利得貯蓄が同時に導入された場合を考える。図 2 を振り返ると、 $(\sigma \neq 0, s \neq 0)$  では協調率の結果だけを見ればこれまでと凡そ同じ傾向が観察されるが(図 2, \*-1 を参照)、その裏に潜む人々の思考 (ゲーム性向) の進化に注目すると、結果はこれまでと全く逆のゲーム性向—金持ち喧嘩せず—へ進化することが判明した

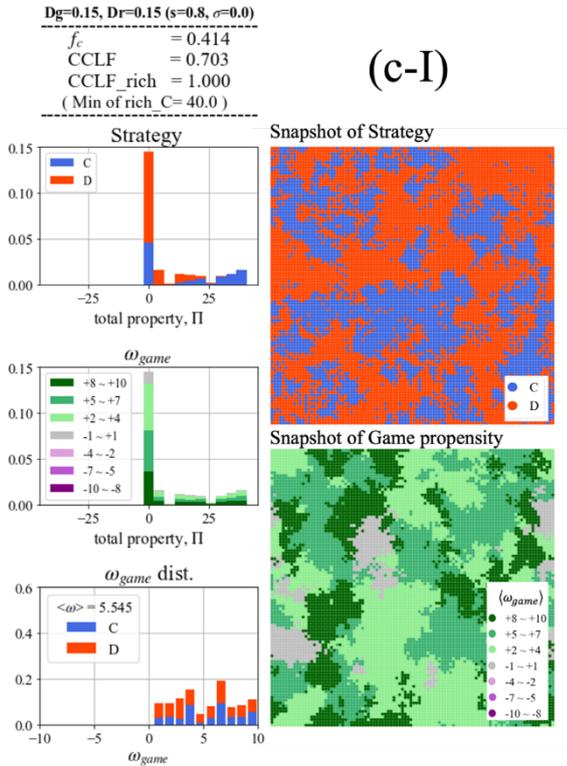


図 5 c-Iでの総利得分布とスナップショット

(図2, \*2を参照). 以下ではその結果の代表例として2カ所を抽出(図2の点e-I; ( $\sigma = 1, s = 0.2, D_g = 0.8, D_r = 0.8$ )とh-I; ( $\sigma = 4, s = 0.2, D_g = 0.6, D_r = 0.6$ ))し, 各詳細結果を基に, このゲーム性向の進化ダイナミクスを考察する.

4-3-1. ゲーム参加コスト (小) の場合(e-f)

図6には, 図2内の点e-I; ( $\sigma = 1, s = 0.2, D_g = 0.8, D_r = 0.8$ )における, ある1エピソード中の9つの統計量の時系列データと, 均衡での戦略・ゲーム性向のスナップショットを表示している. 9つの統計量とは, 図3で表示した4つの統計量に加え, 境界エージェント(異なる戦略の隣接エージェントを少なくとも一人持つ個体)の平均ゲーム性向値 $\omega_{ave}(Boundary)$ (ピンク破線)と内側エージェント(周囲を同じ戦略のエージェントのみに囲まれている個体)の平均ゲーム性向値 $\omega_{ave}(Inside)$ (水色破線), Cクラスター内部のCエージェント(Inside\_C)の平均総利得 $\Pi_{ave}(Inside_C)$ (薄青実線), Dクラスター内部のDエージェント(Inside\_D)の平均総利得 $\Pi_{ave}(Inside_D)$ (薄赤実線), 新規にD戦略からC戦略に切り替えたエージェント数 $New\_C/200$ (黄色実線, スケーリングの為に200で除している), の5つである.

まず, 協調率の結果では均衡で凡そ0に近い値を示しているが, これは4-1-1.に記した協調率増加のロジックで説明がつく( $\because \Pi_{ave}(Boundary\_C) > \Pi_{ave}(Boundary\_D)$ ではないから). 続いて, ゲーム性向の進化について考える. 図3では全エージェントのゲーム性向値を合算した $\omega_{ave}$ として, その時間推移を表示していたが, 図6では $\omega_{ave}$ に加えて, 境界・内側のエージェントに分けて $\omega_{ave}$ を表示している

(図6中ピンク破線の $\omega_{ave}(Boundary)$ と, 水色破線の $\omega_{ave}(Inside)$ ). すると, ゲーム開始直後から境界 or 内側で金持ち喧嘩する or しないのトレンドが完全に分岐するという興味深い傾向が観察された. 以下ではこの分岐に着目し, ミクロで起きている現象について掘り下げる.

表1には, エージェントを4つの属性(Boundary\_C, Boundary\_D, Inside\_C, Inside\_D)に分類した場合に, 夫々がミクロでどのように振る舞うかを, 時系列データ内の総利得順序と, 境界の内外で異なるゲーム性向値を基に分析し, まとめている. まず, Boundary\_Cから考えると, 彼らが高確率で隣接するのはBoundary\_Dである. そして, 総利得順序では常に $\Pi_{ave}(Boundary\_D) > \Pi_{ave}(Boundary\_C)$ である事から, Boundary\_Cは自身を比較的低利得と判断し, 境界のゲーム性向トレンド( $\omega_{ave}(Boundary) > 0$ ; 金持ち喧嘩する, i.e., 貧乏喧嘩せず)に従って, 次回ゲームに不参加の可能性が高い. すると, Boundary\_Cは戦略更新の機会を持たず, C戦略を保持したままフリーズしてしまう(以降もBoundary\_Cはこれを繰り返す). 一方で, Boundary\_Dは常に総利得順位が一位である為, 境界のゲーム性向トレンド( $\omega_{ave}(Boundary) > 0$ ; 金持ち喧嘩する)に従って, 次回ゲームに参加し, 隣接するBoundary\_Cから更に利得を獲得し続け, 高利得を維持すると考えられる. Inside\_Cは, ゲーム開始直後から薄青実線( $\Pi_{ave}(Inside\_C)$ )が途絶えている事から, 均衡付近では既に集団内に存在し得ないと判る. Inside\_Dに関しては後述するが, “勝ち逃げ”の状況にあると推察した. Inside\_D以外に関しては, 均衡における戦略分布のスナップショットを参照する事により, Inside\_Cが実際に存在しない事や, 少数のCが集団内に分散している事(Boundary\_C), そして, その周囲はBoundary\_Dである事を実際に確認できる. また, 図6のスナップショット2種を比較すると, 左の戦略スナップショットでフリーズしていると考えられるC戦略エージェントにちょうど対応する箇所でのみ, 右のゲーム性向スナップショットで高 $\omega_{game}$ のクラスター(緑色領域)が生き残っていることも確認できる. (それ以外では全て低 $\omega_{game}$ が拡大). これは, Boundary\_CがBoundary\_Dの持つゲーム性向—金持ち喧嘩する—をコピーしようとも, 自身がD戦略に転じることは極めて稀だからである. なぜなら, 自分から搾取を続けて高利得を維持するBoundary\_Dの“金持ち喧嘩する”という思考は, 裏を返せば“貧乏喧嘩せず”であり, 低利得のBoundary\_Cにとってはゲーム不参加を選択するゲーム性向として働き, 結果として戦略更新の機会を持ち得ないからである.

では, 何故低 $\omega_{game}$ がその他領域に拡大したのか. このメカニズムについて説明しているのが図7である. 図7では4-1-2.と同様に, Boundary\_DがC-clusterを侵食し, 次第にD中に埋もれてInside\_Dへ遷移していく状況を図示している. 4-1-2.では参加コストが発生しなかった為, 次第に縮小するC-clusterからも単純に最後まで貪り続ける(金持ちとなっても喧嘩し続ける—高 $\omega_{game}$ )エージェントの方がより高利得の状態を維持できる事に起因して, その遺伝子

表1 e-Iでの各エージェントの振る舞い

Boundary ( $\omega_{ave} > 0$ )	C	poor → will not participate → freed
	D	rich → will participate → richer
Inside ( $\omega_{ave} < 0$ )	C	∅
	D	“Win and Run”

(高 $\omega_{game}$ ) が集団内に拡大すると考えた。しかし、今回の設定では参加コストが発生する為、全員が裏切り合うような環境では、“参加しない方が高利得”という状況が起こり得る(参加コスト分の利得を回収出来ない場合は赤字となる為)。そうなった場合、縮小が進む前の比較的大きな C-cluster から高利得を獲得できているタイミングで上手くゲームから手を引き、次の時刻で今所有している総利得をキャリーオーバーした方が(つまり、“勝ち逃げ”をした方が)、高利得状態を維持できる。

小括すると、参加コストが発生する状況下では“勝ち逃げ”の状況—ただ闇雲にゲームに参加し続けるよりも、勝っている時にうまく手を引いた方が高利得を維持でき、この意味で“金持ち喧嘩せず”が有益な結果を齎らす—に起因して、低 $\omega_{game}$ が集団内に拡大したと推察する。

4-3-2. ゲーム参加コスト(大)の場合(h-i)

最後に、参加コストが増加した場合(h-i)の結果を考察する。図 8 には図 6 と同様のフォーマットで、h-I( $\sigma = 4, s = 0.2, D_g = 0.6, D_r = 0.6$ )における、ある 1 エピソードの詳細結果(9 つの特微量の時系列データと均衡でのスナップショット)を表示した。まず、図 2 の結果を振り返ると、(h-i)の場合も(e-f)と同様に、弱ジレンマ領域を除く全域で低 $\omega_{game}$ —金持ち喧嘩せず—側への進化が観察される(図 2, (e-2), (f-2)を参照)。が、注目すべきは $\omega_{game}$ 進化の中身である。全エージェントのゲーム性向値の平均 $\omega_{ave}$ を属性に分けて表示した $\omega_{ave}(Boundary)$ と

$\omega_{ave}(Inside)$ の時系列データを参照すると、境界・内側どちらのエージェントも低 $\omega_{game}$ 側に進化しており、先述の図 6 の場合と異なり、ゲーム性向の分岐は観察されない(実際に、均衡のゲーム性向スナップショットでは高 $\omega_{game}$ (緑色領域)が存在していない)。つまり、“金持ち喧嘩せず”に進化する理由として 4-3-1.で推察した“勝ち逃げ”のメカニズムは、もはや参加コスト(大)では成り立たないことになる。そこで、再度ミクロに起きている現象を、均衡スナップショット・時系列データから分析する。

図 8 の時系列データからは、総利得順位が(Boundary\_C) > (Boundary\_D) > (Inside\_D)であること、均衡での戦略スナップショットからは Boundary\_C が集団内に分散しているものの孤立していない(2 人ペア、もしくは数人程度の小クラスターを形成している)ことが判る。今、Boundary\_C がゲームに参加した場合の状況を考えると、「 $n$ 人のペアを形成し $\pi_i^t = nR(=n)$ の瞬時利得を獲得」→「境界(及び内側)のゲーム性向トレンド“金持ち喧嘩せず”に従い、次ゲームに不参加」→「次ゲーム終了後の総利得は、(瞬時利得:0)+(割引後の前回総利得)」となる。対して、集団内の大多数を占める総利得最下位の Inside\_D は、ゲーム不参加で利得ゼロか、ゲーム参加で参加コスト $\sigma$ 分の赤字のどちらかであると予測される。

今、注目すべき点は、次の 2 点である；(i) ゲーム不参加の可能性が高いと考察した Boundary\_C は“金持ち”ではないものの、ゲームに参加して赤字を出している Inside\_D よりは比較的高利得である点と、(ii) C 戦略は

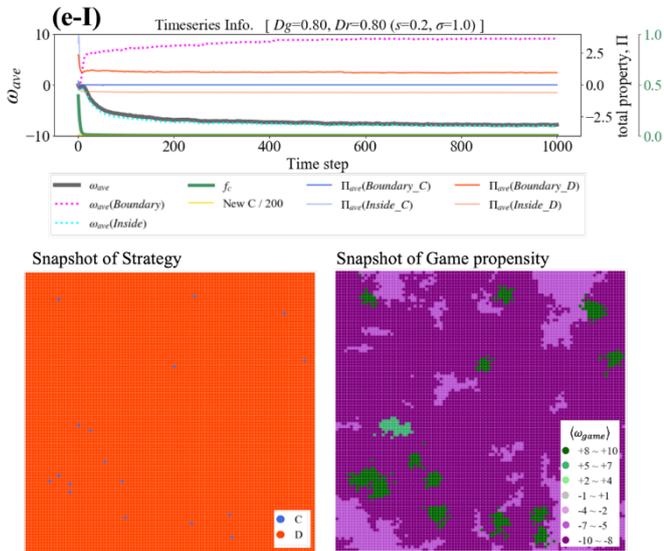


図 6 e-I における統計量の時間推移とスナップショット (均衡)

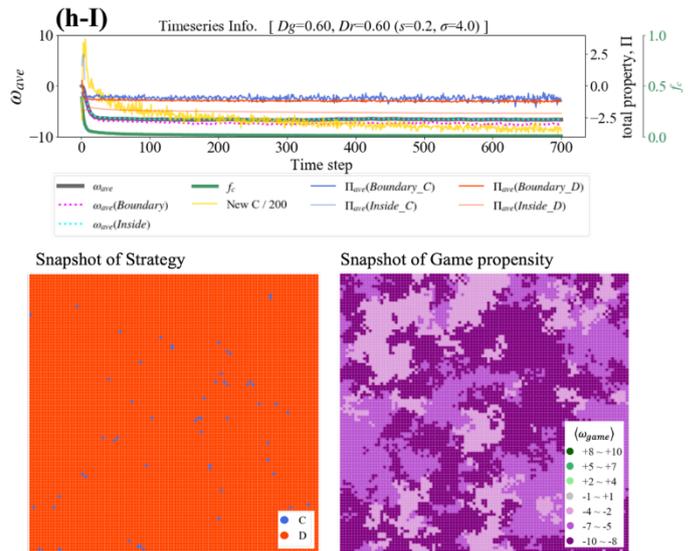


図 8 h-I における統計量の時間推移とスナップショット (均衡)

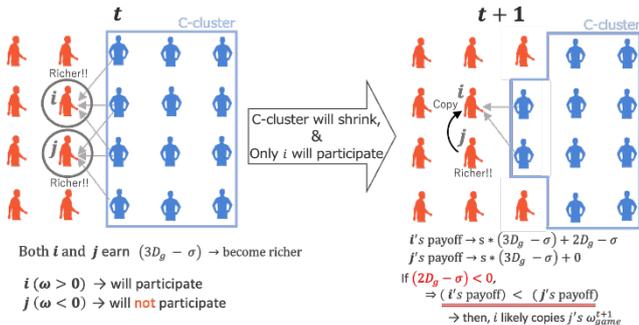


図 7 低 $\omega_{game}$ 拡大のメカニズム (その①)

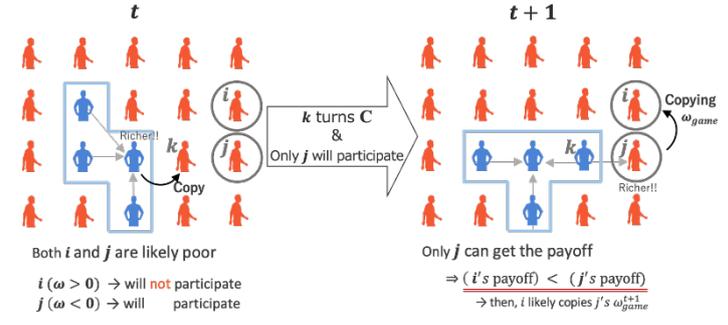


図 9 低 $\omega_{game}$ 拡大のメカニズム (その②)

フリーズしていない点である。これは上述 4-3-1.のマイクロで起きている現象と大きく異なる。(i)に関しては総利得の時系列データから、(ii)に関しても、時系列データ中の New C/200 (黄色実線)から、実際に確認できる。

((e-I)の黄色実線はゲーム開始直後から常に横軸にオーバーラップしているが、(h-I)では均衡付近でも常に全体の1~2%の新規Cエージェントの出現)。図9には、この新規Cエージェントはどのようにして出現するのか、また、なぜ集団のゲーム性向が低 $\omega_{game}$ へ進化しているかを図示した。まず時刻 $t$ (左側)ではC戦略が少人数のクラスターを形成した場合に、局所的に Boundary\_D よりも高利得を獲得できる状況を解説している( $D_g$ の程度によって、中心にCは右隣のDエージェント $k$ よりも高利得となる)。 $k$ は戦略更新の際に、この自分よりも高利得なCを参照した場合、 $t+1$ ではC戦略に切り替わる可能性をもつ(実際に、Boundary\_Cの平均総利得はBoundary\_Dの平均総利得よりも、常に若干高い)。では、この場合の $k$ に隣接するエージェント $i, j$ にも注目する。各エージェントの $\omega_{game}$ はゲーム初期に-10から+10まで幅を持たせて確率的に付与している為、時間進行の過程ではネットワーク中に金持ち喧嘩“する”でInside\_Dになるエージェント( $i$ )と、金持ち喧嘩“せず”でInside\_Dになるエージェント( $j$ )が存在するはずである。この時、Inside\_Dは常に総利得順位で最下位である為、次ゲームには、 $i$ ;不参加、 $j$ ;参加、となる可能性が高い。すると、 $t+1$ では、彼らの共通の隣人である $k$ がC戦略に切り替わっている為、ゲーム参加者の $j$ だけが $k$ からの利得( $1+D_g$ )を獲得でき、( $i$ の総利得) $<$ ( $j$ の総利得)となって、この時刻では $i$ が $j$ の持つゲーム性向である低 $\omega_{game}$ をコピーする可能性が生じる。

小括すると、参加コスト(大)の設定下では、C戦略が局所的に集まると、例えば数人のクラスターでもDエージェントよりも高利得を獲得できる為、集団中に新規のC戦略が出現し得る。すると、常に低利得であるInside\_Dの隣人が突如C戦略に切り替わる可能性があり、もしこのタイミングでInside\_Dが低利得でもゲームに参加していれば(貧乏でも喧嘩していれば)、DはCからの利得を獲得し、高利得に転じる事ができる。従って、貧乏でもゲームに参加し続け(i.e., 低 $\omega_{game}$ )隣人に新規のCが巡ってきた場合の備えをしている—つまり、“チャンス待ち”をしている—一方が良い結果を齎らす事に起因して、低 $\omega_{game}$ の遺伝子が集団内に拡大すると推察した。

#### 4. 結言

著者らの先行研究[6]では、人間の意思決定の背後には2種のゲーム性向—金持ち喧嘩する or 金持ち喧嘩しない—が潜んでいるという点に着目し、それぞれ単一のゲーム性向のみで構成される社会では協調率がどのように進化するかを、進化ゲーム理論における空間型囚人のジレンマゲームをテンプレートに解析し、報告した。しかし、現実の人間社会ではどちらか一方のゲーム性向だけでなく、どちらのゲーム性向も存在していると考えられる。では、集団内に両ゲーム性向が分布する場合、結局、人間のゲーム性向はどちらに進化していくのだろうか?本研究ではこの問いに動機付けられ、戦略・ゲーム性向が同時に進化する共進化モデルのSPDを構築した。解析結果からは、モデル構造の単純さに比して、ネットワーク互惠の再現に特化した通常の

SPDでは観られない、大変豊かな複雑性の創発が観察された。本稿ではその中でも特に印象的な結果;従来のSPDに近い状況では、集団全体で平均したゲーム性向は“金持ち喧嘩する”側に進化してゆく。が、利得貯蓄とゲーム参加コストを同時に考慮した場合は、協調率という表面的な指標ではこれまでと同じ結果に行き着くとしても、裏に潜む人々の思考は全く逆の世界—皆が“金持ち喧嘩せず”を前提とする社会—に進化する、を報告した。更に、この“金持ち喧嘩せず”の進化には二つの背景—「勝ち逃げ」が上手いくことによって、この思考が生き残る場合と、「チャンス待ち」が吉となって“金持ち喧嘩せず”翻って、“貧乏人喧嘩する”の意味で、この思考が遺伝子が集団内に拡大する場合—が存在することを明らかにした。

#### 謝辞

本研究の一部は科研費(19KK0262, 20H02314, 20K21062, 23KJ1724)により遂行された。また、本研究は九州大学情報基盤研究開発センター研究用計算機システムを利用した。記して謝意を表す。

#### 参考文献

- [1] J. M. Smith and G. R. Price, “The Logic of Animal Conflict,” *Nature*, vol. 246, no. 5427, pp. 15–18, (1973).
- [2] M. A. Nowak, “Evolutionary Dynamics.” *Harvard University Press*, (2006).
- [3] J. Tanimoto, “Fundamentals of Evolutionary Game Theory and its Applications,” *Springer*, (2015).
- [4] J. Tanimoto, “Evolutionary Games with Sociophysics,” *Springer*, (2018).
- [5] M. A. Nowak, “Five Rules for the Evolution of Cooperation,” *Science*, vol. 314, no. 5805, pp. 1560–1563, (2006).
- [6] S. Utsumi, Y. Tatsukawa, and J. Tanimoto, “Does a resource-storing mechanism favor ‘the wealthy do not fight’?—An approach from evolutionary game theory,” *Chaos, Solitons & Fractals*, vol. 160, p. 112207, (2022).
- [7] L. Yong-Kui, L. Zhi, C. Xiao-Jie, and W. Long, “Payoff-based accumulative effect promotes cooperation in spatial prisoner’s dilemma,” *Chinese Phys. B*, vol. 19, no. 9, p. 090203, (2010).
- [8] Y. Geng, C. Shen, H. Guo, C. Chu, D. Yu, and L. Shi, “Historical payoff promotes cooperation in voluntary prisoner’s dilemma game,” *Chaos, Solitons & Fractals*, vol. 105, pp. 145–149, (2017).
- [9] C. Sun and C. Luo, “Co-evolution of limited resources in the memory-based spatial evolutionary game,” *Chaos, Solitons & Fractals*, vol. 131, p. 109504, (2020).
- [10] Z. Wang, S. Kokubo, J. Tanimoto, E. Fukuda, and K. Shigaki, “Insight into the so-called spatial reciprocity,” *Phys. Rev. E*, vol. 88, no. 4, p. 042145, (2013).
- [11] K. M. Ariful Kabir, J. Tanimoto, and Z. Wang, “Influence of bolstering network reciprocity in the evolutionary spatial Prisoner’s Dilemma game: a perspective,” *Eur. Phys. J. B*, vol. 91, no. 12, p. 312, (2018).
- [12] C. Shen, C. Du, C. Mu, L. Shi, and Z. Wang, “Exit Option Induced by Win-Stay-Lose-Leave Rule Provides Another Route to Solve the Social Dilemma in Structured Populations,” *Front. Phys.*, vol. 8, (2020).
- [13] C. Shen, M. Jusup, L. Shi, Z. Wang, M. Perc, and P. Holme, “Exit rights open complex pathways to cooperation,” *J. R. Soc. Interface*, vol. 18, no. 174, p. 20200777, (2021).
- [14] J. Wang, W. Chen, F. Yu, J. He, and W. Xu, “Wealth-based rule favors cooperation in costly public goods games when individual selection is inevitable,” *Appl. Math. Comput.*, vol. 414, p. 126668, (2022).
- [15] M. R. Arefin and J. Tanimoto, “Impact of the baseline payoff on evolutionary outcomes,” *Phys. Rev. E*, vol. 104, no. 4, p. 044314, (2021).
- [16] J. Tanimoto, “Evolutionary Games with Sociophysics: Analysis of Traffic Flow and Epidemics,” *Springer*, (2019).
- [17] J. Tanimoto, *Sociophysics Approach to Epidemics*, *Springer*, (2021).