

ネットワーク上の進化ゲームにおける連続戦略および 混合戦略がもたらす協調創発機構の解明

小窪聡¹, 谷本潤², 萩島理²

¹九州大学大学院 総合理工学府 環境エネルギー工学専攻

²九州大学大学院 総合理工学府 環境共生工学部門

概要

ネットワーク上でジレンマゲームを行いその協調創発のダイナミクスを理解しようとの試みが数多報告されてきた。既往研究ではエージェントは完全な協調もしくは裏切りで定義されるバイナリーな2値戦略を有するとされてきたが、現実社会を考えるとこの前提は妥当とは云いがたい。本稿では、ネットワーク上のジレンマゲームでは、これら離散戦略による社会ダイナミクスの進化的帰結と連続値を許した戦略によるそれとがドラスティックに異なることを、包括的かつ系統的な数値実験を通じて報告する。

A study on how Continuous and Mixed strategies differently enhance network reciprocity from Discrete strategy.

Satoshi Kokubo¹, Jun Tanimoto¹, Aya Hagishima¹

¹Interdisciplinary Graduate School of Engineering Sciences, Kyushu University

Abstract

Cooperation in the Prisoner's Dilemma (PD) played on various networks has been explained by so-called network reciprocity. Most previous studies have presumed that agents can offer either the cooperative (C) or defective (D) strategy. Such a discrete strategy choice seems unrealistic in the real world because actual options might be continuous rather than discrete. This study finds that there is considerable inconsistency in equilibrium among discrete, continuous, and mixed strategy games in network PD games.

1 緒言

囚人のジレンマゲーム(PD)に基づく進化ゲーム理論は、自然界に見られる多くの他利的協調行動の創発を理解するアプローチとして注目されている。1992年に空間型PD (SPD) の報告がされて以来[1], 膨大な数のSPD研究, すなわちネットワーク互惠の効果に関する検討が数多行われてきた。ネットワーク互惠では、次の2つの効果が協調創発を可能にしていると考えられている：(1) ゲーム対戦相手の限定 (2) 局所的な戦略適応 (エージェントはリンク接続している隣人エージェントの戦略のみをコピーする)。これら2つの効果によって、エージェントがC

戦略もしくはD戦略のどちらかしかない単純な戦略しか持ち得ない場合 (1bit記憶) であっても、囚人のジレンマゲーム (PD) でCエージェントが生き残れるようになる。

ところで、既往研究では、エージェントは完全な協調もしくは裏切りで定義されるバイナリーな2値戦略を有する (離散戦略) とされてきたが、現実社会を考えるとこの前提は妥当とは云い難い。なぜなら、実際には完全な協調者や裏切り者はほとんど存在せず、むしろその中間の戦略を選択するエージェントが大多数だと考えられるからである。つまり、エージェントは、戦略値が $[0,1]$ の実数で定義される連続戦略もしくは混合戦略を有すると考えるべきであ

る。

ネットワーク互恵が付加されていない進化ゲームにおいて、離散戦略と、連続値を許した連続戦略と混合戦略の均衡点が一致することが証明されている[2][3]。しかし、ネットワーク互恵を付加したジレンマゲームにおいては、それら 3 戦略の均衡点に関する議論が行われてこなかった。この背景から、筆者らはネットワーク上の進化ゲームにおける 3 戦略の均衡に注目し、それがドラスティックに異なることを報告した[4]。しかし、3 戦略の均衡点の差異がどのように生じ、何故生じるか、ダイナミクスの本質を理解するには十分な検討が為されていない。

本研究の目的はこれらの諸点を、包括的かつ体系的な数値実験により明らかにすることである。

2 モデル

2.1 離散戦略・連続戦略・混合戦略

2 人 2 戦略ゲーム、つまり、 2×2 ゲームを原型とする。離散戦略を想定する場合、エージェントは離散手である協調 (Cooperation, C) もしくは裏切り (Defect, D) のどちらかを選択する。 2×2 ゲームの利得構造を P (Punishment, 自他の手組は D-D, 以下同様), R (Reward, C-C), S (Saint, C-D), T (Temptation, D-C) で表す。Tanimoto & Sagara[5]に倣って、チキン型ジレンマ、鹿狩り型ジレンマを夫々 $D_g (=T-R)$, $D_r (=P-S)$ で表す。上記 2 つのジレンマは夫々、相手を出し抜いて自分が高利得を上げようとの意図、出し抜いて相手を低利得に突き落とそうとの意図を表している。 $R=1$, $P=0$ で固定すると、ゲーム構造は

$$G = \begin{matrix} & \begin{matrix} C & D \end{matrix} \\ \begin{matrix} C \\ D \end{matrix} & \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{matrix} = \begin{matrix} & \begin{matrix} C & D \end{matrix} \\ \begin{matrix} C \\ D \end{matrix} & \begin{pmatrix} 1 & -D_r \\ 1+D_g & 0 \end{pmatrix} \end{matrix} \quad (1).$$

本稿では、 $D_g \in [0,1]$, $D_r \in [0,1]$ の PD ゲームのクラスを考察対象とする。

連続戦略を想定する場合、エージェント i は $s_i \in [0,1]$ で定義される実数値を戦略として有する ($s_i=0$, $s_i=1$ はそれぞれ完全な C 戦略と D 戦略を表している)。本稿では、エージェント i がエージェント j とゲームを行った際に得る利得 $\pi(s_i, s_j)$ を、次式で定義した。

$$\begin{aligned} \pi(s_i, s_j) &\equiv (S-P)s_i + (T-P)s_j \\ &\quad + (P-S-T+R)s_i s_j + P \\ &= -D_r s_i + (1+D_g)s_j + (-D_g + D_r)s_i s_j \end{aligned} \quad (2).$$

この定義は、 P, R, S, T の 4 つの端点の間を単純に補間したものである。

混合戦略を想定する場合、連続戦略と同様、エージェント i は $s_i \in [0,1]$ で定義される実数値を戦略として有する。しかし、各エージェントはその戦略値に基づき、確率的に離散手である C もしくは D を出す。すなわち、エージェント i は $\text{Rnd}[] < s_i$ の場合に C を出し、それ以外は D を出す。ここで、 $\text{Rnd}[]$ は一様分布に従う乱数を表している。

2.2 ネットワーク

全エージェント数 N を 4900 とし、平均次数 $\langle k \rangle$ が 8, 12 となるようにネットワークを生成する。ネットワーク構造は 2 次元格子グラフ (Lattice), サイクルグラフに $p=0.2$ の確率でショートカットを加えた次数分布のない Small-World (HoSW) と次数分布の存在する Watts-Strogatz スモールワールド (HeSW) [6], BA アルゴリズム[7]で生成した Scale-Free グラフ (最小次数は $\langle k \rangle / 2$ となるようにした) を用いた。エージェントはこれらのネットワーク上で進化ゲームを繰り返す。

2.3 実験方法

全ての単位ステップ時間において、各エージェントは隣人とゲームを行い、その合計利得と隣人の利得に応じて戦略を更新する。戦略適応方法は、隣人中最大利得のエージェントの戦略をコピーする Imitation Max を用いた。初期協調率 (協調者の戦略比) は 0.5 とする。連続戦略と混合戦略では $s_i \in [0,1]$ の一様分布を想定している。各試行は、戦略値と利得の平均の摂動が十分小さくなり擬似均衡と見なせるまで続ける。摂動が大きく均衡に達しなかった場合は、上限である 10000 時間ステップの最終 100 ステップの平均データを算出する。PD クラス ($D_g \in [0,1]$, $D_r \in [0,1]$ においてそれぞれ 0.1 刻み) の $11 \times 11 = 121$ ポイントのシミュレーションを行い、各ポイントにおいて 100 試行実行する。本稿では解析する際、以下の 4 つの PD 領域を定義した: (1) PD の全 121 ポイント (AllPD), (2) $D_g = D_r$ の所謂 Donor & Recipient ゲームと呼ばれる 11 ポイント (DRG), (3) $D_r = 0$ の PD とチキンゲームの境界ゲームにおける 11 ポイント (BCH), (4) $D_g = 0$ の PD と鹿狩りゲームの境界ゲームにおける 11 ポイント (BSH)。

3 結果及び考察

図 1 に、離散戦略、連続戦略、混合戦略の各ジレンマ領域の平均協調率を示す。図 1 から、well-mixed におけるゲームでは等しいと証されている、離散戦略、連続戦略、混合戦略の均衡が、ネットワークゲームにおいては大きく異なることがわかる。AllPD 領域と DRG 領域ではどのネットワークにおいても、離散戦略と比べ連続戦略は高協調率か同程度の協調率を示している、混合戦略は連続戦略よりも高い協調率を示している (図 1a, b)。しかしながら、チキン型のジレンマのみが存在する BCH 領域と鹿狩り型のジレンマのみが存在する BSH 領域の結果を比較すると、どちらのジレンマが存在するかにより、協調 enhance 効果は 3 種類の戦略で異なる性質を示していることがわかる (図 1c, d)。そこで、BCH 領域と BSH 領域に注目して考察を行う。

まず、BCH 領域における 3 戦略の比較を行う。BCH 領域ではどのネットワークにおいても、離散戦略に比べると、連続戦略と混合戦略の均衡は高協調率を示している (図 1c)。ここからは、 $\langle k \rangle = 8$ の Lattice を例に挙げて考察していく。図 2 に BCH における 3 戦略の均衡到達後の平均協調率を示す。離散戦略では全員裏切りの均衡に到達しているジレンマが強いゲームの場合でも、連続戦略と混合戦略では、ある程度の協調率が保たれていることがわかる。空間型 PD においてチキン型のジレンマが強くなると内部均衡に吸引されるダイナミクスが強くなる。そのために、完全な協調と裏切りの両極端の戦略しか存在し得ない離散戦略では、 D_g がある閾値を超えると、大多数が協調戦略を選択する状態から全員が裏切り戦略を選択する状態へと崩落してしまうのに対し、中間戦略が存在する連続戦略と混合戦略では、 D_g が大きくなってもある程度の協調率が保たれると考えられる。図 3 は、連続戦略と混合戦略の均衡到達後における戦略分布の 100 アンサンブル平均を示しており、ジレンマがごく微弱な場合を除きピークを示す戦略値は 0 と 1 の中間値であり、ジレンマが強くなるにつれて徐々に低くなっていくことがわかる。

次に、BSH 領域における 3 戦略の比較を行う。離散戦略と連続戦略における均衡の協調率に比して、混合戦略のそれは非常に高い値を示している。これは混合戦略が戦略値に応じて確率的に戦略を選択することと深く関係している。例えば、エージェント i

とエージェント j がゲームを行うとする。離散戦略と連続戦略では、 $s_i > s_j$ の場合、必ずエージェント j の利得の方が高くなる。しかし、戦略値に応じて確率的に協調か裏切りを選択する混合戦略では、 $s_i > s_j$ の場合でも、稀にエージェント i が裏切りをエージェント j が協調を選択することにより、エージェント i の利得の方が高くなる可能性もある。混合戦略では、この「逆転勝利」により、初期に裏切り戦略からの侵略を逃れることができた協調戦略によるクラスターが他の 2 戦略よりも比較的多く形成されるので、BSH で高い協調率を示すのだと考えられる。

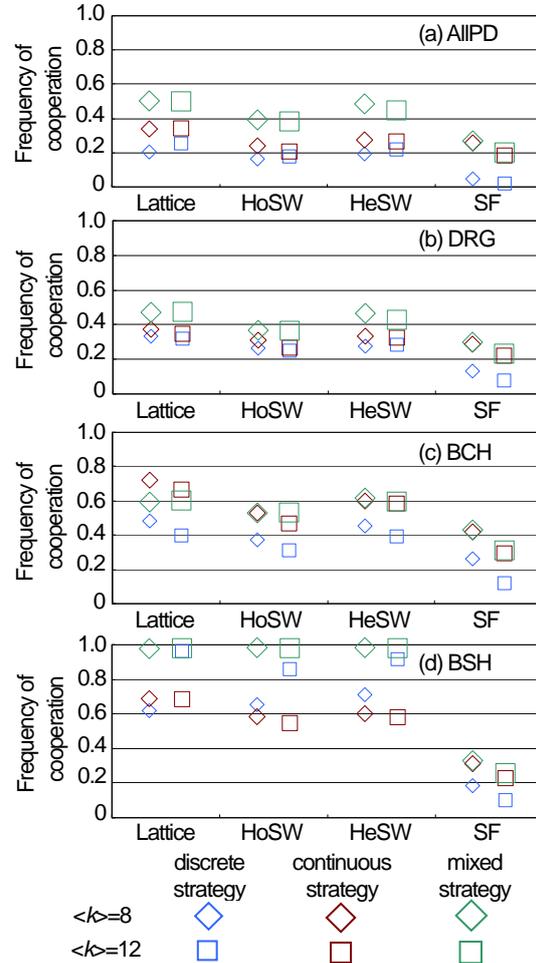


Figure 1: Averaged cooperation proportions of three strategy games for (a) AllPD, (b) DRG, (c) BCH, and (d) BSH within the limits of $0 \leq D_g \leq 1$ and $0 \leq D_r \leq 1$.

ただし、SF の BSH 領域を見ると、他のネットワークとは異なった性質を示している。平均次数 8 と 12 のどちらにおいても、連続戦略と混合戦略の協調率

はほぼ同じ値を、そして離散戦略よりもわずかに高い値を示している (図 1d)。そこで、 $\langle k \rangle = 8$ の SF に注目して、図 4 に BSH 領域における試行毎の協調率と 100 アンサンブル平均協調率を示す。均衡は bi-stable 的ではなく、様々な内部均衡に到達していることがわかる。これは、SF には高い次数を持ったハブエージェントが存在し、戦略の伝搬はハブエージェントを通じて広がっていく、ということが原因であると考えられる。連続戦略と混合戦略ではエージェントは様々な戦略値を持った者が存在する。つまり、ハブエージェントが初期にどの戦略値を有するかによる均衡への影響が大きいため、図 4 のような結果を示すと考えられる。

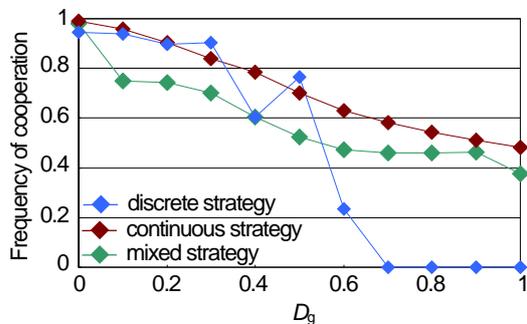


Figure 2: Averaged cooperation proportion for discrete strategy game, continuous strategy game, mixed strategy game. Games are played on lattice networks with the average degree $\langle k \rangle = 8$. The dilemma subclass is BCH.

4 結語

ネットワーク上の進化ゲームにおいて、戦略を離散戦略、連続戦略、混合戦略で定義し、比較・考察を行った。その結果、ネットワーク互恵が付加されていない進化ゲームでは生じないとされている均衡点の差異が、ネットワーク互恵が付加された進化ゲームでは生じることが解った。さらに、包括的かつ体系的な数値実験により、チキン型ジレンマと鹿狩り型ジレンマのどちらのジレンマが存在するかによって、均衡点の差異が異なる性質を示すことを明らかにした。また、SF 以外のネットワークと SF ネットワークで均衡点の差異が異なる性質を示すことから、ネットワークトポロジーによっても均衡点の差異の性質が異なることを明らかにした。

本稿の結果は、既往研究で積み上げられてきたジレンマゲーム研究に関する知見を根底から覆す成果に結びつく可能性がある。

参考文献

- [1] Nowak, M.A., May, R.M., *Nature* 359 (1992) 826–829.
- [2] Vincent, T.L., Cressman, R., *Theor. Pop. Biol.* 58 (2000) 173.
- [3] Day, T., Taylor, P.D., *Evol. Ecol. Res.* 5 (2003) 605.
- [4] Zhong, W., Kokubo, S., Tanimoto, J., *BioSystems*, In Press (2011).
- [5] Tanimoto, J., Sagara, H., *BioSystems* 90(1) (2000) 105–114.
- [6] Watts, D.J., and Strogatz, S.H., *Nature* 393 (1998) 440.
- [7] Barabasi, A.L., Albert, R., *Science* 286, (1999) 509–512.

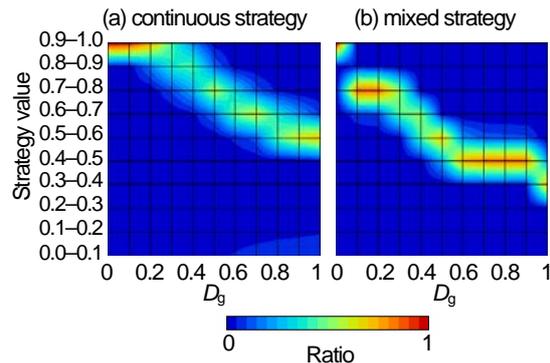


Figure 3: Strategy distributions at equilibria for BCH based on (a) continuous and (b) mixed strategy games. Games are played on 8-neighbor lattices ($k = 8$) with 4900 agents.

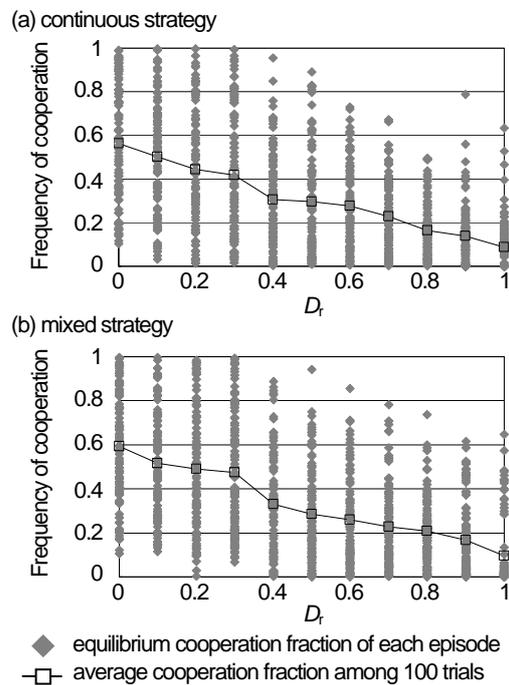


Figure 4: Equilibrium cooperation proportion of each episode and average cooperation fraction of 100 trials based on (a) continuous and (b) mixed strategy games. Games are played on SF networks with the average degree $\langle k \rangle = 8$.