囚人のジレンマにおける連続戦略から離散戦略への切り換えによるネットワーク互恵

岸本憲幸[1]，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　総合理工学研究院　エネルギー環境共生工学部門

概要

Zhong & Kokubo ら（*BioSystems*）が報告した離散戦略に対する連続戦略の特性の違いに着目し，ゲームエピソード初期の定期には連続連略を，その後，離散戦略に切り換えることで，ネットワーク上の囚人ジレンマゲームにおける協調が enhance されることを発見した．この機構は，連続戦略ゲームを一定期間行うことで，初期にランダム配置される協調と裏切り戦略をいわば“再配置”して協調エージェントをクラスター化するものである．

Shifting from continuous to discrete strategy enhances the network reciprocity for Prisoner’s Dilemma Games.

Noriyuki Kishimoto[1], Satoshi Kokubo[1], Jun Tanimoto[1], Aya Hagishima[1]

[1]Interdisciplinary Graduate School of Engineering Sciences, Kyushu University

Abstract

Following to the work by Zhong & Kokubo et al., we find non trivial enhancement of network reciprocity for Prisoner’s Dilemma Games by shifting from continuous strategy, imposed to agents only for a certain beginning period, to usual discrete strategy system for the following period. This particular protocol can be said one of the measures to make a neutral situation by initial randomized assignment of defectors and cooperators to more cooperation favored situation by forming clusters by cooperators despite same cooperation fraction.

1　緒言

利己的個体の相互作用の中で協調的群行動の自己組織化をもたらす機構を解明することが進化ゲーム理論研究の大きな目的である．最近の研究トレンドで特に関心が寄せられているのが，囚人ジレンマ (Prisoner’s Dilemma, PD) におけるネットワーク互恵の力学機構の解明と更なるenhance効果をもたらすプロトコルである．ネットワーク互恵は1992 年に空間型PD (SPD)の報告がされて以来[1]，高等な情報処理機能やメモリを前提としない協調創発機構として理論生物学や統計物理学を中心に多くの研究が蓄積されてきた．ネットワーク互恵とは，1bitの戦略 (協調；Cooperation (C) か裏切り；Defection (D) か) しか持たないエージェントたちのゲーム対戦相手と戦略適応相手を限定することで協調を創発させる機構である．これは，空間構造がエージェント間の匿名性を減少させ，well-mixedな状況から社会粘性を増大させることにより協調創発を可能にしている (well-mixedな集団内の匿名性を圧縮することで，協調戦略の期待利得が裏切り戦略のそれを凌駕するように変容させる機構である) と考えられる．従来，PDを含む 2 人 2 戦略 (2×2)ゲームではエージェントの戦略はCかDかのbinaryな離散戦略で定義されてきたが，ごく最近，Zhong & Kokuboら[2]はネットワーク上の進化ゲームでは，完全裏切り(s=0) と完全協調 (s=1) の中間的戦略を許容する連続戦略による均衡は離散戦略のそれとは異なると報じている．本研究では，この連続戦略ゲームと離散戦略ゲームの進化上の特性の違いに注目し，さらなるenhanceをネットワーク互恵にもたらす枠組みを提示するものである．

その大略を要括すれば以下のようなものである．

連続戦略ゲームは離散戦略ゲームよりも初期段階での裏切り戦略の侵襲が弱く，協調クラスターが形成されやすい[3]．一方で，連続戦略ゲームでは中間

的な戦略値を取り得る為，裏切り戦略からの侵襲を耐え凌いで形成された協調クラスターは「ほどほどの協調クラスター」であり，増殖しても協調率はさほど高くならない．この点で，初期 D の侵襲以降協調増加に転じたモードにあっては離散戦略の方が優位である．そこで，離散戦略と連続戦略ゲーム両者の長所を生かすため，ゲームのエピソード初期には連続戦略でゲームを行い，適当な次期に離散戦略にシフトするゲームの枠組みを考える．これは，現実社会との対比で云えば，ゲーム初期の過渡的状況下ではゼロ or ナッシングの極端な意志決定は控え，ゲームの帰趨が見えてきた段階でやおら 2 値的な戦略に切り換えて意志決定を行うことに対応するだろう．蓋し，現実的にはあり得そうな考え方だと思われる．

2　モデル

2.1　離散戦略，連続戦略と切り換え

固定ネットワーク上の 2×2 ゲームを考える．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[4]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, D_rで表す．ここで, D_g, D_r について若干の説明を加えておく．D_g=T-R, D_r=P-Sと置くことで，4 つのゲームクラス(Chicken, Stag Hunt(SH), Prisoner’s Dilemma(PD), Trivial)は，D_g, D_rの正負に一致する．すなわち，D_g>0 は相手を貪ろうとするギャンブル型ジレンマ(チキン型ジレンマ)の，D_r>0 は相手に貪られまいとするリスク回避型ジレンマ(鹿狩り型ジレンマ)の有無を表し，それぞれの条件が単独に満たされるときChicken, Stag Hunt(SH)のクラスとなり，同時に満たされるとき Prisoner’s Dilemma(PD), いずれも満たさないときにはジレンマのないTrivialのゲームクラスとなる．また，Replicator Dynamicsそのものと内部均衡点は，R, T, S, Pに代わってD_gとD_rだけで表式可能である．故にD_gとD_rは well-mixedな無限サイズ集団を前提にする際のジレンマの強さを表していることになる．

R=1, *P*=0 で固定すると，ゲーム構造は

$$G = \begin{array}{c} \\ \mathrm{C} \\ \mathrm{D} \end{array} \begin{array}{c} \begin{array}{cc} \mathrm{C} & \mathrm{D} \end{array} \\ \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{array} = \begin{array}{c} \\ \mathrm{C} \\ \mathrm{D} \end{array} \begin{array}{c} \begin{array}{cc} \mathrm{C} & \mathrm{D} \end{array} \\ \begin{pmatrix} 1 & -D_r \\ 1+D_g & 0 \end{pmatrix} \end{array} \qquad (1).$$

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[2]．

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

この定義は，*P*, *R*, *S*, *T* の 4 つの端点の間を補間したもので，離散戦略ゲームを最もシンプルに拡張した構造になっている．

全てのエージェントは，初期，一様乱数で振られたランダムな連続戦略でゲームを開始し，以下述べる切り換え条件を満たした時点で，離散戦略に切り替える．従って，連続から離散戦略への切り換えはシンクロに行われる．切り換えは，タイミング 2 パターン，方法 3 パターンの計 6 通りについて検討する．

<u>タイミング</u>

①下降→上昇：協調率がゲームを開始してから最初に前回時間ステップより上昇した時点で切り換え．

②下降→上昇かつ協調率>0.5：協調率がゲームを開始してから最初に前回時間ステップより上昇し，かつ協調率>0.5 の時に切り換える．

<u>切り換え方法</u>

(1)四捨五入：エージェント毎に切り換え時の連続戦略値を四捨五入して，離散戦略に変換する．

(2)個別確率：エージェント毎に連続戦略値に応じて確率的に 0, 1 の離散戦略に変換する(戦略値を s とすると, s の確率で 1(C),　1-*s* の確率で 0(D)).

(3)一括確率：同一の連続戦略値を有するエージェントをグループ化し，グループ毎に戦略値に応じて確率的に 0, 1 に変換する．これは，同一連続戦略値を有するエージェントは互いにコピーし合って，近接してクラスターを形成しているだろうとの考えによる．

2.2　ネットワーク

全エージェント数 *N* を 4900, 平均次数<*k*>は 4, 8, 24 の 3 パターンとした．ネットワーク構造は 2 次元格子グラフ (Lattice), B-A アルゴリズム[5]で生成した Scale-Free グラフを用いた. 合計 6 通りである．

2.3　戦略更新方法

エージェントの利得は全隣人とのゲームで得た利得合計値とする．戦略更新方法は，(a)全エージェントが隣人とのゲームを終えた後，Imitation Max (IM, 隣人中の最大利得エージェントの戦略を確定的にコピーする) で戦略適用を行なう Synchronous (同期)と，(b)各エージェントがランダム順にゲームと戦略適用を続けて行う Asynchronous (非同期) の 2 通りで行なった．

2.4　実験方法

数値実験では戦略切り換えの組み合わせ 6, ネットワークトポロジー6, 戦略更新法 2 通りの合計 72 通りを悉皆的にシミュレーションする．1 つの組み合わ

せに対して，PDクラス（D_g∈[0,1], D_r∈[0,1]においてそれぞれ 0.1 刻み）の 11×11=121 ポイントのゲーム構造についてシミュレーションを行い，各ポイント 100 試行のアンサンブル平均を解析データとする．協調enhance効果を評価する指標として，PDの全 121 ポイントの各 100 試行の均衡協調率を算術平均した値(AllPD)を用いる．

なお，比較のため，エピソード全ステップを通じて離散戦略，連続戦略を用いるゲームについてもシミュレーションする．

エピソード初期における各エージェント戦略は s_i∈[0,1]の一様分布で付与（離散戦略固定ゲームではCorDを 0.5 の確率でランダムに付与）．各エピソードは，協調率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 1000 時間ステップの最終値を採用する．

3 まとめ

3.1 結果及び考察

図1に，各条件の協調率を連続戦略の比として，切り換え方法毎にまとめて示した．1以上が連続→離散への切り替えが，全期にわたって連続戦略ゲームで行うケースに比べて協調のenhance効果があったことを意味する．ネットワークトポロジーとシンクロ・アシンクロによる相違はあるが，総じて，”下降→上昇かつ協調率>0.5&四捨五入”で切り換えると良好な結果である．また，特異的に，Lattice&<*k*>=24のトポロジー，シンクロ更新のもとで，”下降→上昇&一括確率”の戦略切り替えを行うと非常に大きなenhance効果が観察された．本稿以下ではこの2点について考察する．なお前者に関しては，Scale-Free &<*k*>=8のトポロジー，シンクロ更新を例に取り上げる．

まず，前者に関して図2にD_g - D_r相図に示した100アンサンブル平均均衡協調率（同条件連続戦略，離散戦略ゲームも併示），図3にD_r=0, D_g =0.1の100試行のうち、高い均衡協調率に至った上位10試行の協調率時系列を示した．図2より，比較的ジレンマの弱い領域で戦略切り換えは協調enhance効果があることが分かる．これは，ジレンマが弱いため初期のD侵襲には耐えて協調増加のモードに推移するようになるが，全期を連続戦略によると中間戦略を取り得るため，そこそこの協調で均衡してしまうのに対し，離散戦略に切り換えることで，ほぼ完全な協調に至り得るようになったからだと推量される．また，ジレンマの強いチキンとの境界に近いゲームでは，戦略切り換えの結果は全期・連続戦略ゲームと同様に中間的な協調の均衡に至っており，離散戦略ゲームより良好な結果である．(図には示していないが）この領域では，多くの場合，戦略切り替えは生起せず，提示モデルではほぼ全期・連続戦略ゲームとして推移する．以上から，提案のプロトコルは，ジレンマ強（特にチキン型ジレンマ）領域では連続戦略ゲームのメリットを，ジレンマ弱領域では連続戦略ゲームの欠点を切り替えによって克服していることが分かる．特に後者に関しては，初期間の連続戦略ゲームにより協調的（完全に協調ではない）なクラスターを方々に形成させ，四捨五入で離散戦略に切り換える際にその一部のCクラスターを温存することで，初期に戦略がランダム配置される影響をより協調がfavorされる方向に転換している効果だと言えよう．0.5以上で切り替えると設定したから，このことは，離散戦略で通常のネットワーク互恵ゲームをする際に，プレゲーム期間として連続戦略ゲームを行って同じ初期協調率0.5でも，よりCクラスターが形成された状態に調整し，しかるのち再スタートさせるゲームだと解釈することが出来る．

次にLattice&<*k*>=24のトポロジー，シンクロ更新のもとで，”下降→上昇&一括確率”の戦略切り替えを行ったケースの結果を上記同様，図4と図5 (D_r=0.5, D_g =0.4) に示す．このケースでは，戦略切り換え時点の協調率0.5以上との条件は付されていないために，ごく低いグローバル協調率で連続から離散戦略に切り換えられる．その際，四捨五入でなく同じ実数値の連続戦略グループ毎に確率的にCorDを決めるので，エージェント数*N*が十分大きければ，周囲が全て裏切りでもCが増殖出来るクラスターサイズの条件を満たしたCクラスターが確率的に創成されることになる．そのようなクラスターが1つでも2つでも出来れば，切り換え後にCは大増殖を遂げ，切り換え時点の協調率が低かろうが，周囲のDをどんどんC化して遂には協調均衡に至ることになる．この”周囲が全て裏切りでもCが増殖出来るクラスターサイズの条件”は当然次数大ほど厳しく，小ほど緩く小さなクラスターでよい．よって，次数だけ小さい（4および8）同条件下のケースでは，全期・離散戦略ゲームの初期ランダム配置でもこの条件を満たす可能性が相対的に高くなる．これは，全期・離散戦略ゲームにおける初期協調率への依存性が相対的に小さいこと言っているから，前パラグラフのレトリックを使うと，プレゲームとしての連続戦略ゲームを行うメリットは*k*=24に比べて大きくないことになる．以上のことから，Lattice&<*k*>=24のトポロジーで特異的に大きなenhance効果が顕れたものと推量される．

3.2 結語

PDにおけるネットワーク互恵に関連して，戦略として完全な協調，裏切りの中間値を許容する連続戦略でエピソードの初期にゲームを行い，離散戦略に切り換える枠組みを提示し，数値実験の結果，切り替えの条件によって大きな協調enhance効果があることを報告した．

参考文献

[1] Nowak, M.A., May, R.M., Nature 359 (1992) 826–829.

[2] Zhong, W., Kokubo, S., Tanimoto, J., BioSystems, In Press (2011).

[3] Tanimoto,J., Sociobiology 58 (2) (2011) 315-325.

[4] Tanimoto, J., Sagara, H., BioSystems 90(1) (2000) 105-114.

[5] Barabasi, A.L., Albert, R., Science 286, (1999) 509–512.

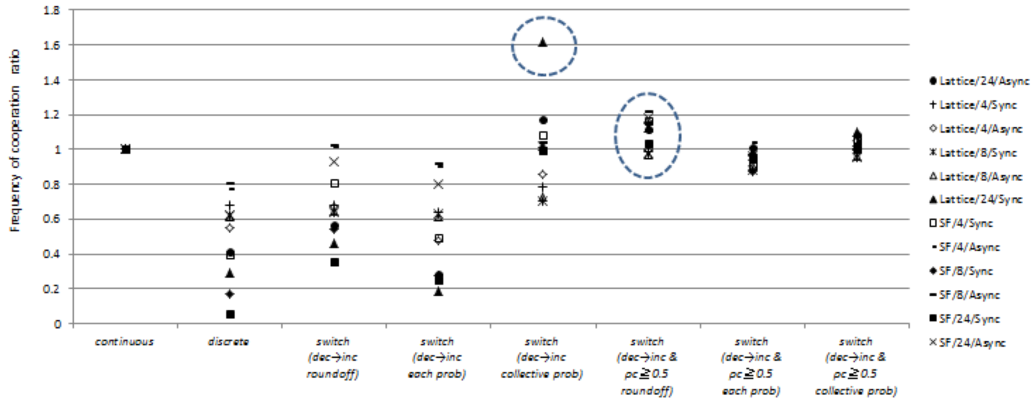

図1. 戦略切り換え法ごとに観た連続戦略に対する協調enhance効果(AllPD領域).

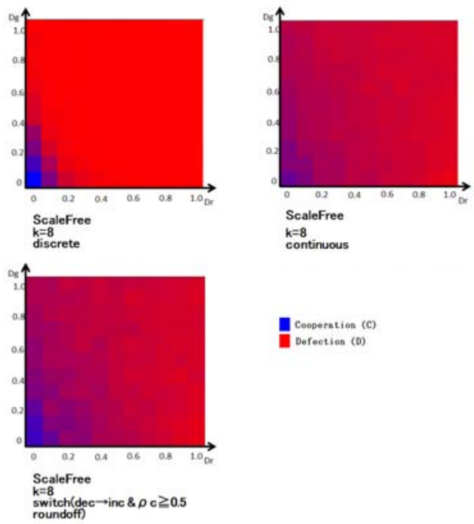

図2. 均衡協調率の100アンサンブル平均(下降→上昇かつ協調率>0.5 &四捨五入, Scale-Free, <*k*>=8, シンクロ更新（左上：離散戦略ゲーム, 右上：連続戦略ゲーム, 左下：戦略切り換えモデル).

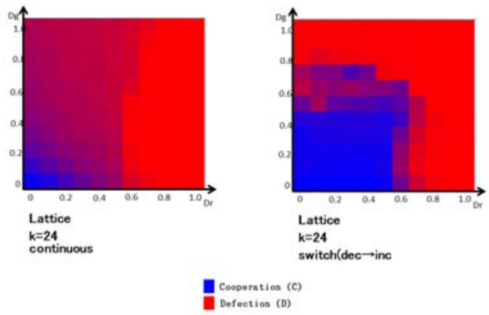

図4 平均衡協調率の100アンサンブル平均(下降→上昇&一括確率, Lattice, <*k*>=24, シンクロ更新（左：連続戦略ゲーム, 右：戦略切り換えモデル).

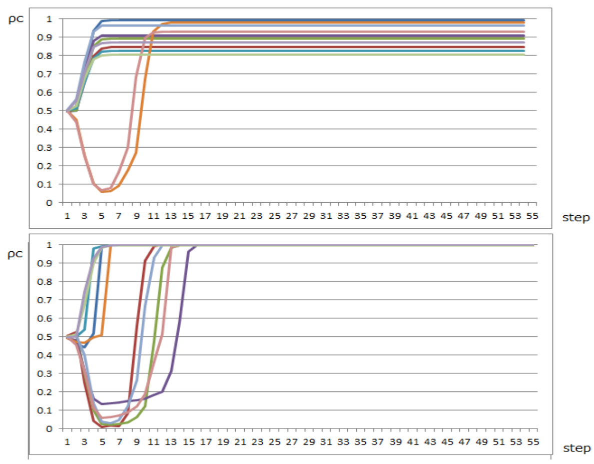

図3. 下降→上昇かつ協調率>0.5&四捨五入, Scale-Free, <*k*>=8, シンクロ更新のD_r=0, D_g=0.1のエピソードにおける協調率推移(上：連続戦略ゲーム, 下：戦略切り換えモデル）.

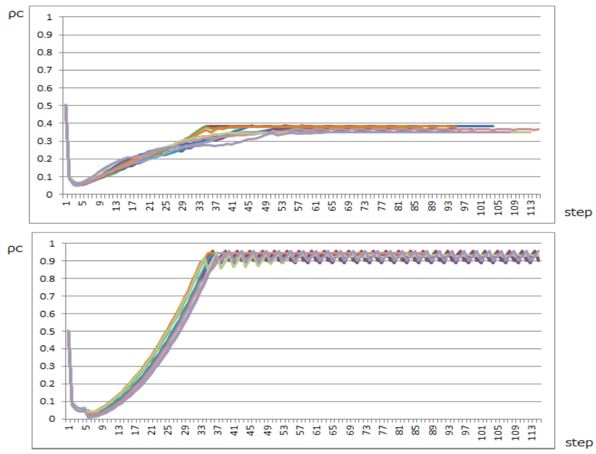

図5. 下降→上昇&一括確率, Lattice, <*k*>=24, シンクロ更新のD_r=0.5, D_g=0.4のエピソードにおける協調率推移(上:連続戦略ゲーム, 下：戦略切り換えモデル）.