

# 囚人のジレンマにおける戦略適用範囲を拡大した切替戦略 によるネットワーク互惠

岸本憲幸<sup>1</sup>, 小窪聡<sup>1</sup>, 谷本潤<sup>2</sup>, 萩島理<sup>2</sup>

<sup>1</sup>九州大学大学院 総合理工学府 環境エネルギー工学専攻

<sup>2</sup>九州大学大学院 総合理工学府 環境エネルギー工学専攻 エネルギー環境共生工学部門

## 概要

Kishimoto らが報告した[3]ゲームエピソード初期の定期は連続戦略, その後, 離散戦略に切り換える“切替戦略”を, 戦略適用範囲を拡大したネットワーク上の囚人ジレンマゲームに適用し, 協調が enhance されることを発見した. そのメカニズムは, 切替戦略における“初期連続戦略”の機構により, 高ジレンマ下でも C クラスタが消滅せず生き残り, 生き残った C クラスタが戦略適用範囲の拡大により容易に成長可能となり, 大きな enhance 効果をもたらすことによる.

## Network reciprocity for Prisoner's Dilemma Games with expanding adaptation by the strategy-shifting from continuous to discrete

Noriyuki Kishimoto<sup>1</sup>, Satoshi Kokubo<sup>1</sup>, Jun Tanimoto<sup>1</sup>, Aya Hagishima<sup>1</sup>

<sup>1</sup>Interdisciplinary Graduate School of Engineering Sciences, Kyushu University

## Abstract

Following to the work by Kishimoto et al. [3], we find non trivial enhancement of network reciprocity for Prisoner's Dilemma Games with expanding adaptation neighborhood by the strategy-shifting from continuous to discrete. On the proposed model, agents form clusters with mid-cooperative strategy at the early stage in an episode even if the dilemma strength is strong. And small number of them change to entire cooperative clusters (C-clusters) after the strategy shifting. It follows that those small number of C-clusters survived can easily expand by imposing lots of neighboring defectors to have C. Why this mechanism works is discussed from detailed dynamics.

## 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惠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の提示である. ネットワーク互惠は 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. ネットワーク互惠とは, 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の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. これに着目して Kishimoto らはゲームエピソード初期の定期は連続戦略, その後, 離散戦略に切り換える“切替戦略”を提案し, enhance 効果があることを報じた[3]. 本研究では, 戦略適用範囲を拡大したネットワーク上の囚人ジレンマゲームに先述の切替戦略を適用して, さらなる enhance をネットワーク互惠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.

切替戦略とは, ①離散戦略ゲームよりも初期段階 (END 期間, 後述) での裏切り戦略から侵襲を低抑化出来る[4]連続戦略ゲー

ムの長所, ②裏切り戦略からの侵襲を耐え凌ぎ協調率が増加していく段階 (EXP 期間, 後述) では完全協調戦略を持つため協調がより大きく enhance されるとの離散戦略の長所を両者採り入れたアイデアである。一方, 戦略適用範囲を拡大したネットワークにおいては, クラスタ境界の D エージェントが, クラスタ内部の R を取り合っている高利得 C エージェントを適用範囲に多く含むようになり, 隣人集合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 IM (後述) に依るならば, 境界 D エージェントが C 化しやすくなり, 結果としてクラスタの拡大につながる。本研究は, この両者を組み合わせてさらに大きな enhance 効果を期待するものである。

## 2 モデル

### 2.1 離散戦略, 連続戦略, 切替戦略

固定ネットワーク上の  $2 \times 2$  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。  $2 \times 2$  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  $D_g, D_r$  で表す。  $R=1, P=0$  で固定すると, ゲーム構造は

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

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 [2]。

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

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。

切替戦略は, 連続戦略でゲームを開始後, 切替条件を満たした時点で離散戦略に切り替える。切替タイミングは, 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とし (後述の END  $\rightarrow$  EXP への遷移ステップに相当する), 切替方法としては, 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, 近接してクラスタを形成しているから, これらに一括して D (0) か C (1) を付与することで C か D かのクラスターに変換することを意味する。

本稿では, 離散戦略と切替戦略について実験を行なった。

### 2.2 ネットワーク

全エージェント数  $N$  を 10000, ネットワーク構造は 2 次元格子

グラフ (Lattice) で, 対戦範囲, 戦略適用範囲ともに次数  $k=8$  の従来モデルと, 対戦範囲  $k=8$  (Moore 第 1 近傍) で戦略適用範囲  $k=24$  (Moore 第 2 近傍) とした拡張モデルの 2 パターンとした。拡張モデルでは, 戦略適用範囲だけを Moore 第 2 近傍に拡大している。

### 2.3 戦略更新方法

エージェントの利得は対戦範囲の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, 全エージェントが隣人とのゲームを終えた後, Imitation Max (IM, 戦略適用範囲の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) かつ Synchronous (同期) で行なった。切替戦略において, 連続戦略から離散戦略への切り換えも Synchronous (同期) に行なった。

### 2.4 実験条件

あるジレンマ強さのゲーム構造において, アンサンブル平均 100 回と採って解析対象とする。各エピソードは, 離散戦略においては, 初期に協調率 0.5 でランダムに C, D エージェントを配置, 切替戦略においては, 初期に一樣乱数で振られたランダムな連続戦略を持つエージェントを配置する。各試行は, 戦略値の平均の振動が十分小さくなり疑似均衡とみなせるまで続ける。振動が大きく均衡に達さなかった場合は, 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。

### 2.5 Enduring (END) and Expanding (EXP) periods

図 1 に示すように END と EXP 期間を定義する。初期協調率 0.5 から始める典型的な進化プロセスの場合, 一般に相前後する 2 段階が存在する。すなわち, 協調率が急激に下がる段階, 協調率が徐々に上昇する段階である。前者を C エージェントが D エージェントの侵襲に耐え忍ぶ (enduring; END) 期間とし, 後者を END を生き残った C クラスタが拡大していく (expanding; EXP) 期間とする [6] [7]

## 3 結果及び考察

図 2 に PD 全領域の平均協調率を示す。離散戦略, 切替戦略とも拡張モデルは従来モデルに比べて高い協調率を示している。また同じ拡張モデルでも, 切替戦略の方が離散戦略に比べて高い協調率となる事が分かる。以降, 上記の結果について考察する。まず拡張モデルの従来モデルに対する優位性について, 離散戦略から考察する。図 3 (a), (b) に 100 アンサンブル平均均衡協調率の  $D_g - D_r$  相関を示す。相関の右下すなわちチキン型ジレンマ弱, 鹿狩り型ジレンマ強の領域では, 従来モデルでは協調率ゼロに吸引されているのに対し, 拡張モデルでは高い enhance 効果を示している。この領域の代表例として  $D_g=0.8, D_r=0.0$  における従来モデル, 拡張モデルの代表エピソードのスナップショットを図 4 に示す。図 4 を見ると, 従来モデルでは生き残ったクラスタが安定状態に入っているのに対し, 拡張モデルでは各クラスタが拡大, 結合し大きなクラスタを形成している事が分かる。これにより, 従来モデルでは低い協調率, 拡張モデルでは高い協調率が達成されたと考えられる。これは緒言にも述べたとおり, ネットワークの戦略適用範囲を上げたことによりクラスタ境界の D エージェントが, クラスタ

タ内部のCを取り合っている高利得Cエージェントを適応範囲に多く含むようになり、境界DエージェントがC化し、クラスタが拡大したからだと考えられる。

次に切替戦略における拡張モデルの従来モデルに対する優位性、ならびに離散戦略に対する優位性について考察する。

再び図3 (c), (d) の $D_g$ - $D_r$ 相図に戻ると、チキン型ジレンマ弱、鹿狩り型ジレンマ強の領域において、拡張モデルが従来モデルに比べて高い協調率を示している。これは離散戦略の時と同様であるが、特徴的なのは、相図上部すなわちチキン型ジレンマ強の部分において切替戦略かつ拡張モデルでは、ある程度の協調率を保っていることである。この領域の代表例として、図5に切替戦略 拡張モデル $D_r=0.8$ ,  $D_g=0.7$ における代表エピソードのスナップショットを示す。図5を見ると、END期間にC-D中間的戦略のクラスタが多数形成され、これらのいくつかは切替時に完全協調クラスタに変換される。これは緒言でも述べたとおり、切替戦略のEND期間に連続戦略ゲームである為、離散戦略ゲームよりも裏切り戦略の侵襲が弱く(完全裏切り者は稀で中途半端な裏切り者しかいないため)、協調的クラスタ(無論、あくまで”中途半端な”協調者のクラスタ)が形成されやすいとされる。次にEXP期間に着目すると、複数のクラスタが拡大、分裂を繰り返していることが分かる。これは強ジレンマであるが故、複数のCクラスタが拡大し境界が近づくと、境界Dエージェントの利得が高くなり、境界CエージェントがD化しクラスタが縮小分裂し、その後縮小分裂したCクラスタは再び他のCクラスタとの境界が近づくまで拡大する、ということを繰り返すからである。以上により、切替戦略 拡張モデルでは、チキン型ジレンマが強いゲーム構造において、ある程度の協調率を保ち得るのだと考えられる。

上記のように、離散戦略 拡張モデルに対して大きく優位性を持つ切替戦略 拡張モデルであるが、( $D_g$ - $D_r$ 相図では読み取れないが)  $D_r \geq 0.6$ ,  $D_g = 0.3$  の領域において、わずかに(平均協調率でおおよそ0.1程度) 離散戦略 拡張モデルの方が高い協調率を示している。これについて考察する。この領域は $D_r$ 強であるが故にEND期間において生き残る拡大可能なクラスタの数が離散戦略、切替戦略とも少なくなっている。離散戦略の場合は、これらのクラスタがEXP期間で拡大し高協調率を得るが、切替戦略の場合は、生き残ったクラスタの数が少ない為、連続から離散戦略への切替によって、クラスタが消滅してしまい協調率ゼロに吸引するエピソードがいくつか現れる。これにより、全エピソードの平均均衡協調率を比べると離散戦略の方がわずかに高協調率を示すと考えられる。これが $D_g = 0.4$ になると、離散戦略において拡大可能なクラスタが生き残れないエピソードが多く現れてくるため、再び切替戦略が優位になると考えられる。

最後に、切替戦略 拡張モデルである程度の協調率で平衡状態に達する $0.6 \leq D_g \leq 1.0$ の相での興味深い現象に言及しよう。先に述べたように、この相ではほとんどのエピソードがクラスタ間の相互干渉により拡大縮小を繰り返し中程度の協調率に落ち着くが、稀に高い協調率に至るエピソードも確率的に生じる。このエピソードのスナップショットを図6に示す。スナップショットを見ると、切替によって拡大可能なクラスタが多数形成され、それらが拡大し前記のように相互干渉して縮小するが、その結果、1つのクラスタ以外は全て再拡大することなく消滅し、拡大可能なクラスタが1つのみ生き残るという状態に至る。ひとたびこの状況になれば、相互干渉するCクラスタが存在しない為、この生き残ったCクラスタが拡大しシステム全体をC化して最終的にほぼ全員協調する相が達成される。 $0.6 \leq D_g \leq 1.0$ の領域で全エピソードの統計を取ると、約1.6%のエピソードがこの状況になり高い協調率を達成している。

## 4 結語

PDにおけるネットワーク互恵に関連して、先に報告した切替戦略を、戦略適用範囲を拡大したネットワーク上の囚人ジレンマゲームに適用し、協調がenhanceされることを発見した。また本提案モデルにおいて、強ジレンマ下で稀に高い協調率を達成するエピソードがあるという興味深い事象を示した。

## 謝辞

本研究の一部は科研費萌芽研究( #25560165)、ファイザーヘルス研究財団による。記して謝意を表する。

## 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] 岸本憲幸, 小窪聡, 谷本潤, 萩島理, 第17回交通流のシミュレーションシンポジウム (2011)
- [4] Tanimoto, J., *Sociobiology* 58 (2) (2011) 315-325.
- [5] Tanimoto, J., Sagara, H., *BioSystems* 90(1) (2000) 105-114.
- [6] Shigaki, K., Tanimoto, J., Wang, Z., *Physical Review E* 86 (2012) 031141
- [7] Wang, Z., Kokubo, S., Tanimoto, J., *Physical Review E* 88 (2013) 042145

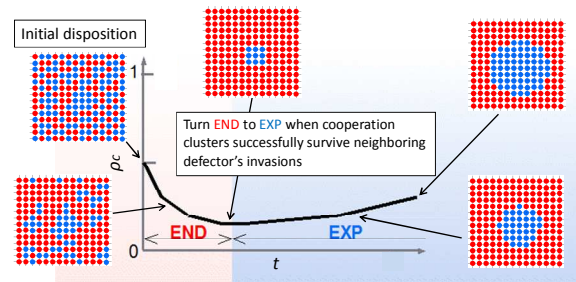


図1. ENDおよびEXPの概念図。Enduring (END) period: 初期のCエージェントが急速にDエージェントになり少数の小さなCクラスタが形成される。Expanding (EXP) period: Cクラスタが拡大し始め、境界のDエージェントをC化していく。

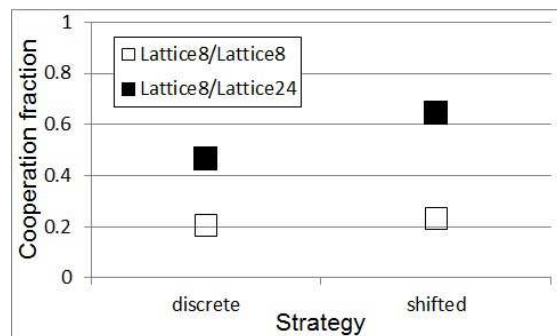


図2. PD全領域の平均協調率

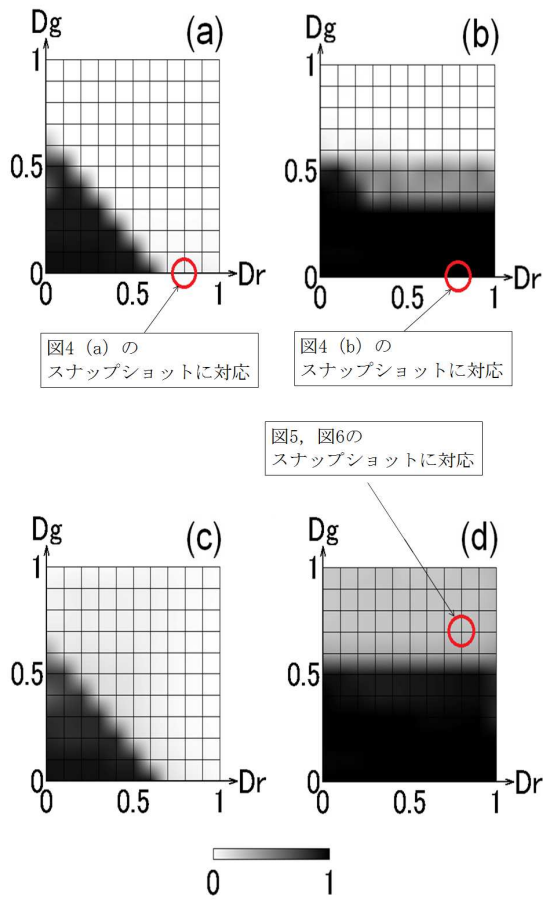


図3. 100 アンサンブル平均均衡協調率の  $D_g$ - $D_r$  相図

- (a) 離散戦略 従来モデル
- (b) 離散戦略 拡張モデル
- (c) 切替戦略 従来モデル
- (d) 切替戦略 拡張モデル

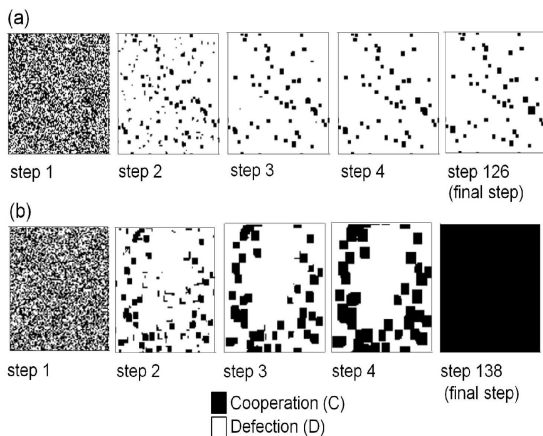


図4. 代表エピソードのスナップショット ( $D_r=0.8$ ,  $D_g=0.0$ )

- (a) 離散戦略 従来モデル
- (b) 離散戦略 拡張モデル

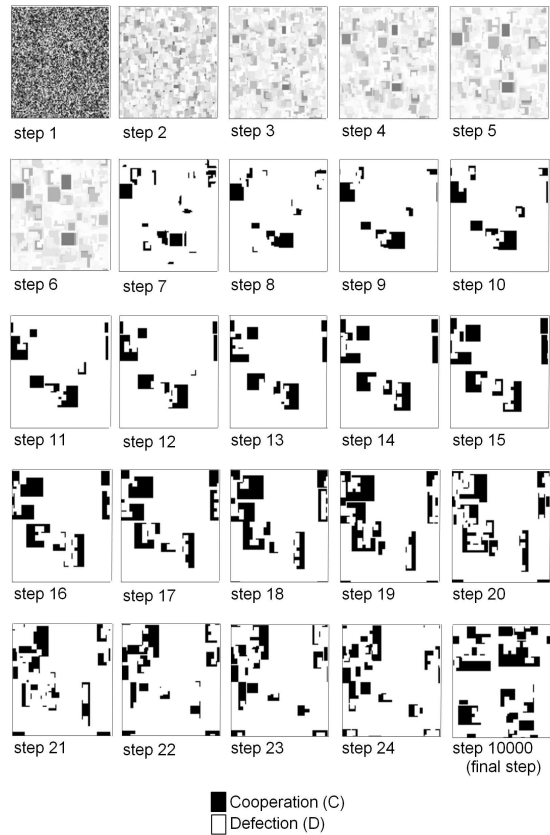


図5. 代表エピソードのスナップショット ( $D_r=0.8$ ,  $D_g=0.7$ )

切替戦略 拡張モデル  
step6 まで連続戦略, step7 以降離散戦略

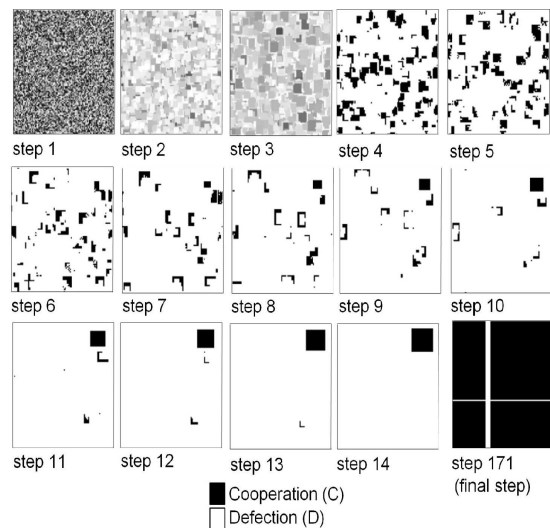


図6. 代表エピソードのスナップショット ( $D_r=0.8$ ,  $D_g=0.7$ )

切替戦略 拡張モデル (高協調率に達するエピソード)  
step3 まで連続戦略, step4 以降離散戦略