

# エージェントの行動エラーが空間型囚人のジレンマゲームにおける協調進化に与える影響

小笠原崇<sup>1</sup>, 谷本潤<sup>2</sup>, 福田枝里子<sup>1</sup>, 萩島理<sup>2</sup>, 池谷直樹<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 総合理工学研究院 エネルギー環境共生工学部門

## 概要

本論では、ネットワーク上の空間型2×2囚人のジレンマゲームにおいて、エージェントの行動エラーがネットワーク互恵に与える影響について考える。初期協調率と行動エラー確率を変えた実験から、エージェントが行動エラーを起こしやすくなるにつれ、協調が促進され、ある以上の確率で行動エラーを起こすと、初期依存性が解消されることを示す。また、ネットワーク互恵の本質から、この協調促進および初期依存性解消の機構について論じる。

## Effect of agent's action error for evolution of cooperation in spatial prisoner's dilemma games

Takashi Ogasawara<sup>1</sup>, Jun Tanimoto<sup>1</sup>, Eriko Fukuda<sup>1</sup>, Aya Hagishima<sup>1</sup>, Naoki Ikegaya<sup>1</sup>

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

## Abstract

By a series of systematic simulations, we elucidate how agent's action error affects on network reciprocity in spatial prisoner's dilemma games. When an agent is prone to mistake her/his strategy, we can observe that cooperation is enhanced larger than usual network reciprocity and the dependency on initial cooperation level is disappeared. From the standpoint of the principle of network reciprocity, we report how and why this mechanism is brought about.

## 1 緒言

自然界に見られる協調行動の創発と維持を理解するため進化ゲーム理論が適用される。とりわけ、囚人のジレンマゲーム (Prisoner's Dilemma Game:以下 PD) は現実の相互的作用を比喩するモデルとして理論的、実験的観点から注目されてきた。1992年の空間型囚人のジレンマゲーム (Spatial Prisoner's Dilemma Game:以下 SPD) をはじめ[1]、これまで SPD に関係するネットワーク互恵[2]について、数多くの研究が行われている。2人2戦略(2×2)PDでは、各エージェントが協調 (Cooperation:以下 C) か裏切り (Defection:以下 D) のいずれかの戦略を持ち、ゲーム対戦において、自身の戦略にしたがった行動をとるが、不注意による誤りで、自身の戦略とは異なる行動をとること (行動エラー) は十分に起こり得る。Daiらは、エージェントが自身の戦略に反した行動を確率的にとることで、協調が促進されると報告している[3]。しかし、その協調促進機構の詳細については論及されておらず、彼らが用いたゲーム構造は、PDのサブクラスのひとつである Donor &

Recipient Game (DRG) [4]のみであるため、PD全域において、エージェントの行動エラーがネットワーク互恵に与える影響は明らかになっていない。

一方、著者らの研究グループでは、ネットワーク互恵の本質が、各エージェントが初期状態においてランダムに配置され、ゲーム対戦と戦略適応を繰り返しながら、Cエージェントによって形成されるクラスター (Cクラスター) がどのように拡大するかが重要であるということを示した[5, 6]。Cクラスターが初期に高い利得を得たDエージェントに侵襲され、協調率が低下する Enduring 期間 (END) と、その期間で生き残ったCクラスターとの境界に存在するDエージェントがC戦略をコピーすることでCクラスターが拡大し、協調率が上昇する Expanding 期間 (EXP) に分けて論じることで、文献[5, 6]では、ネットワーク互恵の本質が理解可能であることを指摘している。

また、Shigakiらは、協調が促進される試行では、初期協調率が小さい程、ENDの終わりで残るCクラスターの数は少なくなるが、これが、EXPでCクラスターの順

調な拡大を促進し、より高い均衡協調率を達成させる、すなわち、EXPにおける協調促進効果の大きさは、初期状態に依存することを明らかにした[7]。

このような背景のもと、本論では、PD全域において初期協調率と行動エラー確率を変えてシミュレーションすることにより、行動エラーが協調促進および初期依存性に及ぼす影響について解析し、ネットワーク互恵の本質から、ENDおよびEXPに分けて、協調促進機構を明らかにする。

## 2 モデル

### 2.1 2×2 ゲーム

エージェント数  $N=10000$ 、次数  $k=8$ 、周期境界条件を持つ2次元正方格子のネットワーク上に配置されたエージェントの2×2ゲームを考える。各エージェントはCもしくはD戦略のどちらかを持つ。2×2ゲームの利得構造は  $P$  (Punishment),  $R$  (Reward),  $S$  (Saint),  $T$  (Temptation) と定義され、これらの利得を獲得する自他の手組みは、それぞれ D-D, C-C, C-D, D-C で表される。Tanimoto および Sagara[8]に倣い、 $D_g = T - R$ ,  $D_r = P - S$  を再定義し、 $D_g$  を Chicken (CH) 型のジレンマ、 $D_r$  を Stag-Hunt (SH) 型のジレンマと呼ぶ。これらの値が大きい程、ジレンマが強いゲームである。ここで、 $R=1, P=0$  とすると、ゲーム構造は、

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

本論では、 $D_g, D_r \in [0, 1]$  を独立に変化させて、PD全域を考察対象とする。

### 2.2 戦略適応とその更新方法

各エージェントは、行動エラー確率  $\varepsilon$  ( $0 \leq \varepsilon \leq 1$ ) で自身の戦略に反した行動をとり、全隣人とゲームを行い、式(1)にしたがって利得を得る。戦略適応方法は、自身を含め隣人全員の中で最大利得を持つ戦略をコピーする。ただし、最大利得を持つエージェントが複数存在する場合、そのエージェントの中からランダムに戦略をコピーする。戦略の更新方法は、全エージェントが利得獲得後、一斉に戦略適応を行うシンクロ更新を用いる。

### 2.3 実験方法

各試行において、初期協調率  $P_{c\_ini} = 0.05, 0.5, 0.95$  でランダムに C, D エージェントを配置する。ゲーム対戦および戦略適応は、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。摂動し続ける場合は、10000 時間ステップを上限とし、最終の 100 時間ステップの平均をとる。以上に示す一連の試行を独立に 100 回行い、その平均を結果とする。

## 3 結果および考察

図1は、 $\varepsilon$  に対する PD 全域の平均の擬似均衡協調率  $P_{c\_ave}$  である。 $\varepsilon < 0.5$  では、 $\varepsilon$  が大きくなるにつれ、 $P_{c\_ave}$  が大きくなる傾向が見られる。 $0.5 < \varepsilon \leq 1$  では、各エージェントは、自身の戦略とは逆の行動をとりやすく、

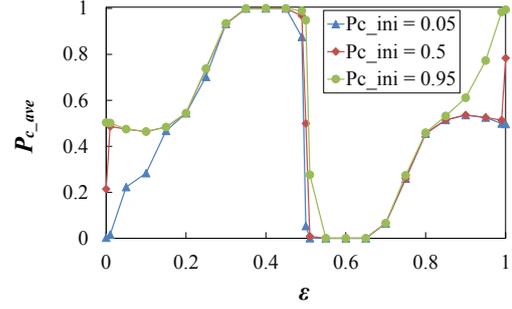


図1.  $\varepsilon$  に対する  $P_{c\_ini}$  別の PD 全域における  $P_{c\_ave}$ .

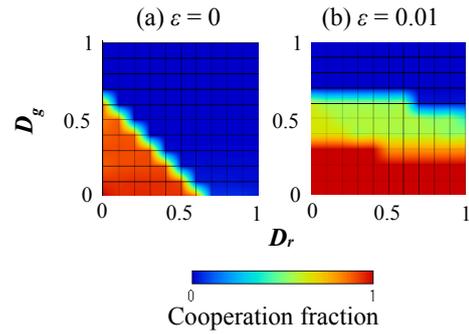


図2.  $P_{c\_ini} = 0.5$  における PD 全域の協調率. (a)  $\varepsilon = 0$ , (b)  $\varepsilon = 0.01$ .

特に、 $\varepsilon=1$  では、完全に反転戦略を持ち、C と D の特性が逆転する。そのため、 $1-\varepsilon$  に対する初期裏切率 ( $1-P_{c\_ini}$ ) ごとの裏切率 ( $1-P_{c\_ave}$ ) は、 $\varepsilon$  に対する  $P_{c\_ini}$  ごとの  $P_{c\_ave}$  と等しくなり、図1において、 $\varepsilon=0.5$  で対称となる。また、 $0.15 \leq \varepsilon \leq 0.85$  において、 $P_{c\_ini}$  に依存せず、 $P_{c\_ave}$  が等しくなることから、初期依存性が解消されることが分かる。

まず、 $\varepsilon < 0.5$  における協調促進の機構について考察する。図2は、 $P_{c\_ini} = 0.5$  における PD 全域の協調率を示す。 $\varepsilon=0.01$  で行動エラーが起こる場合は、行動エラーが起こらない場合 ( $\varepsilon=0$ ) と比べ、CH側のPD領域 CH-PD ( $0.4 \leq D_g \leq 0.6, D_r=0$ ) では、協調率が低下するが、SH側のPD領域 SH-PD ( $D_g=0, 0.7 \leq D_r \leq 1$ ) やジレンマ強さが中程度のPD領域 Middle-PD ( $0.4 \leq D_g \leq 0.6, D_r \geq 0.3$ ) では、協調が促進されることが確認できる。図3は、CH-PD ( $D_g, D_r$ ) = (0.6, 0) および SH-PD ( $D_g, D_r$ ) = (0, 1) における、典型的な試行の時系列スナップショットである。CH-PDでは、 $\varepsilon=0.01$  の場合、 $\varepsilon=0$  と比較すると、ENDの終わり(5ステップ目)で残ったCクラスタが、EXP(10ステップ目以降)で拡大するとき、Cクラスタの境界から多くの亀裂が生じている(30ステップ目)。これは、 $D_g$  が大きいため境界部のCエージェントがDエージェントに変わることにより起因している。均衡時では、Cクラスタに入った亀裂が原因で、Cクラスタが分裂し、数多くのDエージェントが分裂したCクラスタ同士の間に残りやすくなるため、協調があまり促進されない。一方、SH-PDでは、ENDの終わり(5ステップ目)で少数のCクラス

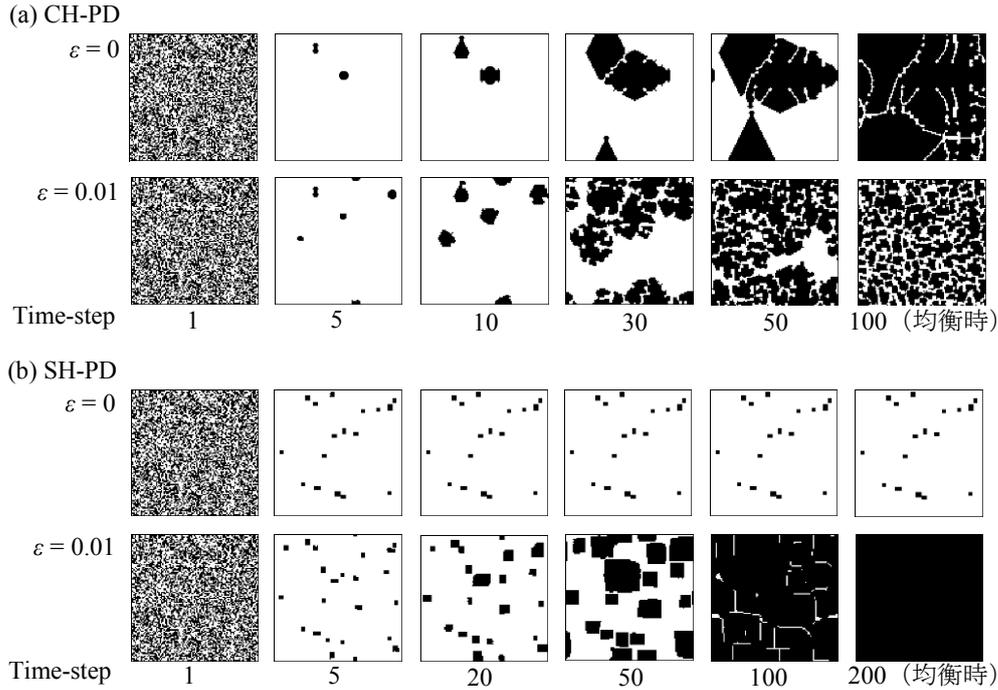


図 3. 行動エラーが起こらない場合 ( $\varepsilon=0$ ) と行動エラーが起こる場合 ( $\varepsilon=0.01$ ) の典型的試行の時系列スナップショット. 黒が C エージェント, 白が D エージェント. (a) CH-PD ( $D_g, D_r$ ) = (0.6, 0), (b) SH-PD ( $D_g, D_r$ ) = (0, 1).

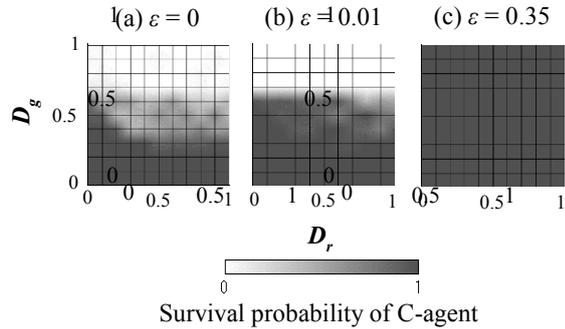


図 4.  $P_{c\_ini} = 0.5$  における PD 全域の C エージェントの生存率. (a)  $\varepsilon=0$ , (b)  $\varepsilon=0.01$ , (c)  $\varepsilon=0.35$ .

タが残るが, 行動エラーが起こらない場合, C クラスタは拡大できず, 極めて低い協調率に留まるのに対し, 行動エラーが起こる場合, 残った C クラスタが最終的に集団全体まで拡大するため, 均衡時の協調率は 1 となる. 図 4 は,  $P_{c\_ini}=0.5$  における PD 全域の C エージェントの生存率 (すなわち, 均衡協調率が 0 とならなかった試行の割合) を示す. 行動エラーが起こらない場合 (図 4(a)), Middle-PD では, C エージェントは生き残ることが難しく, たとえ生き残ったとしても, 図 3(b) 上図に示したように, C クラスタが拡大しないため, 極めて低い協調率に至る. 一方, 行動エラーが起こる場合, 図 4(b)を見ると, Middle-PD でも, C エージェントは高確率で生き残ることが確認できる. C エージェントが生き残った試行では, 図 3(a)下図に示したように, C クラスタが拡大する. したがって, 行動エラー

が起こるとき, 上述した機構によって, CH-PD では, 協調があまり促進されないが, SH-PD や Middle-PD では, 協調が促進されるため,  $P_{c\_ave}$  は上昇する.

$\varepsilon$  が大きくなり,  $\varepsilon = 0.25$  程度になると, C エージェントが D エージェントに対して行動エラーを起こし, D エージェントによる搾取を防げる可能性が高くなるため, C エージェントにとって有利になる. そのため, CH-PD や Middle-PD において, 分裂した C クラスタ同士の間に残る D エージェントを C エージェントに変える効果が強く働きやすくなる. また,  $0.35 \leq \varepsilon < 0.5$  では, ジレンマが強い PD 領域 Strong-PD ( $0.7 \leq D_g \leq 1, 0 \leq D_r \leq 1$ ) においても, C エージェントが生き残り (図 4(c)), D エージェントが C エージェントに変わる効果が働くので,  $P_{c\_ave}=1$  となる. 以上の結果を PD 領域ごとに区別すると, 表 1 にまとめられる.

次に,  $\varepsilon$  がある大きさで, 初期依存性が解消される結果について考察する.  $P_{c\_ini}$  が小さい程, END の終わりで残る C クラスタの数は少なくなるが, これが, EXP において, C クラスタの順調な拡大を促進し, より高い均衡協調率に至ることを保証する[7]. 行動エラーが起こる場合も同様に, END の終わりで残る C クラスタの数は  $P_{c\_ini}$  に依存するが, 均衡時の協調率は,  $P_{c\_ini}$  に依存せず, 等しくなる (図 5). これは, 上述した CH-PD で協調があまり促進されない原因と同様に, END の終わりで少数の C クラスタが残ったとしても, EXP において, C クラスタが拡大するときに生じる亀裂が原因で, C クラスタが分裂するためである. よって, 協調が促進される試行において, 行動エラーが起こる場合, EXP における協調促進効果の大きさは, END の終わり

表 1.  $\varepsilon$  に対する各 PD 領域における協調促進の大きさと C クラスタの状態.

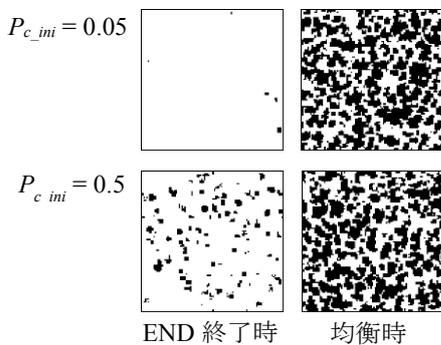
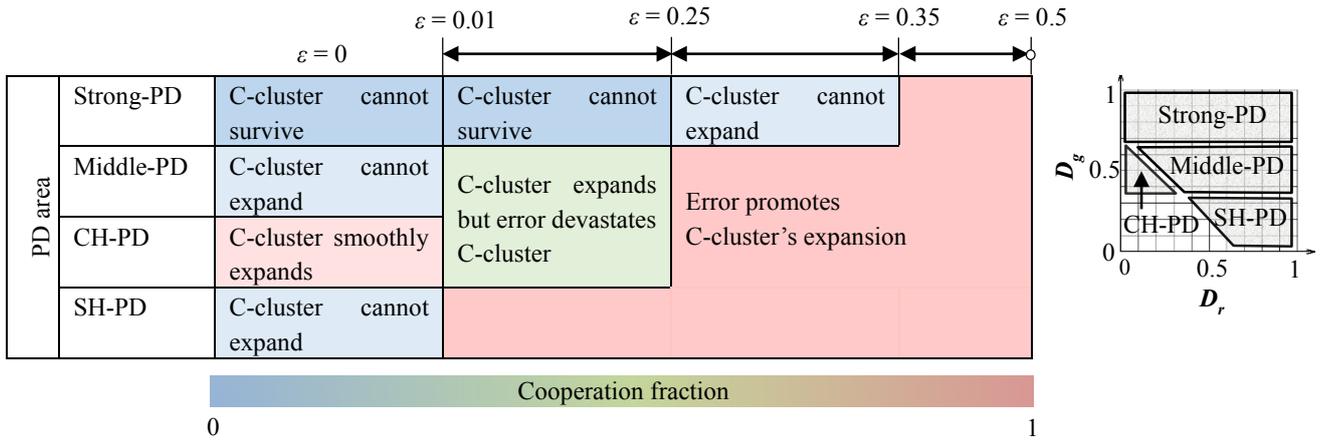


図 5.  $\varepsilon = 0.15$ , CH-PD における  $P_{c\_ini}$  別の END 終了時および均衡時のスナップショット. 黒が C エージェント, 白が D エージェント.

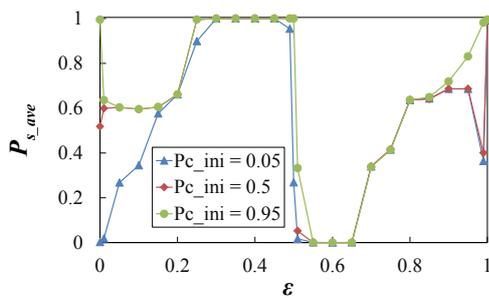


図 6.  $\varepsilon$  に対する  $P_{c\_ini}$  別の PD 全域における  $P_{s\_ave}$ .

で残る C クラスタの数に依存しない, すなわち,  $P_{c\_ini}$  に依存しない. 図 6 は,  $\varepsilon$  に対する PD 全域の平均 C エージェント生存率  $P_{s\_ave}$  である.  $P_{s\_ave}$  は  $P_{c\_ave}$  と同様に,  $0.15 \leq \varepsilon \leq 0.85$  で,  $P_{c\_ini}$  に依存せず, 等しくなる. よって, 行動エラーが起こる場合, EXP における協調促進効果の大きさに初期依存性はないので, END で D に吸引されず, 少なくとも 1 つ以上の C クラスタを生き残らせることが, 初期依存性の解消に大きな影響を与える. また,  $0.35 \leq \varepsilon < 0.5$  では,  $P_{c\_ini} = 0.05$  において, ほとんど生存できなかった C エージェントを必ず生き残らせることから, 行動エラーは C エージェントの生存

に対して, 効果的に働くと言える.

#### 4 結語

PD 全域において,  $P_{c\_ini}$  と  $\varepsilon$  を変えてシミュレーションすることにより, 行動エラーが協調促進および初期依存性に及ぼす影響について解析し, ネットワーク互恵の本質から, END および EXP に分けて, 協調促進機構を明らかにした.

行動エラーは, SH-PD および Middle-PD において, END の終わりで残った C クラスタを EXP で拡大させる効果を持つ. 一方, CH-PD では,  $\varepsilon$  が小さい場合, EXP で C クラスタが分裂し, 協調があまり促進されないが,  $\varepsilon$  が大きくなると, EXP において, 分裂した C クラスタ同士の間にいる D エージェントが C エージェントに変わるため, 協調が促進される. また, 行動エラーは, C エージェントの生存率を大きくする効果を持つが, これが, 初期依存性の解消に大きな影響を与える.

#### 謝辞

本研究の一部は科研費萌芽研究 (#25560165), 立石科学技術振興財団による. 記して謝意を表する.

#### 参考文献

- [1] Nowak, M.A., May, R.M., *Nature* 359 826–829 (1992).
- [2] Nowak, M.A., *Science* 314 1560–1563 (2006).
- [3] Dai, Q., Li, H., Cheng, H., Li, Y., Yang, J., *New Journal of Physics* 12 113015 (2010).
- [4] Hauert, C., Szabó, G., *American journal of Physics* 73 405 (2005).
- [5] Shigaki, K., Tanimoto, J., Wang, Z., Kokubo, S., Hagishima, A., Ikegaya, N., *Physical Review E* 86 031141 (2012).
- [6] Wang, Z., Kokubo, S., Tanimoto, J., Fukuda, E., Shigaki, K., *Physical Review E* 88 042145 (2013).
- [7] Shigaki, K., Wang, Z., Tanimoto, J., Fukuda, E., *PLoS ONE* 8 e76942 (2013).
- [8] Tanimoto, J., Sagara, H., *BioSystems* 90(1) 105-114 (2007).