

Reputation System における加罰的 Norm の創発についての考察

相良 博喜[†] 谷本 潤[‡]

[†]九州大学大学院総合理工学府環境エネルギー工学専攻・修士課程

[‡]九州大学大学院総合理工学府環境工学研究院・教授・工博

〒816-8580 福岡県春日市春日公園 6-1 九州大学総合理工学府環境工学研究室 都市建築環境工学研究室

E-mail: [†] uminchu@uminchu.jp, [‡] tanimoto@cm.kyushu-u.ac.jp

概要 本モデルでは、one-shot の 2x2 ジレンマゲームにおいて、Norm-Reputation System という概念を用いることにより、加罰的な社会システムを創発させた。また、その社会システムはある程度のジレンマ構造が無いと創発しないことを確認した。

キーワード 囚人のジレンマ Norm Reputation System 加罰

Emergence of Punishment Norm in Reputation System

Hiroki Sagara[†] Jun Tanimoto[‡]

[†] Graduate Student, Interdisciplinary Graduate School of Engineering Sciences, Kyushu University

[‡] Prof., Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Dr.Eng

E-mail: [†] uminchu@uminchu.jp, [‡] tanimoto@cm.kyushu-u.ac.jp

Abstract The Norm Reputation System is established to solve a Prisoner's Dilemma in a one-shot 2-players game, which is applicable to support an indirect reciprocity. The Norm Reputation System consists of a 2 layer co-evolution mechanism in which an agent makes his own action strategy evolve and the group he is belonging makes group's norm evolve. The norm regulates an action of agents who belong to the group by "offering D against defectors is good action" etc. The Norm Reputation System encourages a cooperative society by inventing particular norms that approve social punishment rules. This indirect reciprocity backed by the norm allowing social punishment can be evolved more significantly when a game has moderate dilemma than weaker dilemma.

Keywords Prisoner's Dilemma Norm Reputation System Punishment

1. 緒言

ジレンマゲームで協力が創発するメカニズムはこれまで血縁淘汰 (kin election)¹⁾ や互恵的他利主義²⁾ により説明が試みられてきた。

Axelrod ら³⁾ は 2 人繰り返し囚人ジレンマ (Prisoner's Dilemma, 以下 PD) ゲーム (2-IPD) において Tit-for-Tat (TFT) が有効なのは、協力 (Cooperation, 以下 C) には C で応じ、裏切り (Defect, 以下 D) には D で応じるしっぺ返しが有効に機能し、将来にわたって繰り返し対戦する相手に対しては、自ら C を出すことで相手の C を期待するとの直接的互恵 (direct reciprocity) 関係が創発し得るからであると説明した。

しかし、これは繰り返し同一の相手と相互作用を行える直接互恵を前提としている。では、繰り返しのない 1 回きりの相互作用を不特定多数の相手と行う状況において利他的行動はどう進化するのであろうか。このような間接互恵の創発は生物学的にも、社会学的にも興味深い問題である。

Nowak ら⁴⁾ は間接互恵の創発の問題に対し、Image Scoring という機構を提案した。これは「協力した人を良い人、協力しなかった人を悪い人」とラベルづけし、この情報 (Image Score, 以下 IS) を社会全体で共有し、この情報に基づいて行動することにより、協調的な行動を創発させるというものがある。

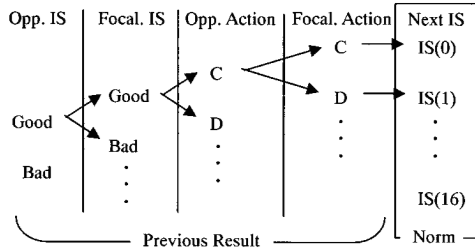


図1 Norm の表現 Norm は 16bit で表され、各 bit は次の IS を表す
 Fig.1 Expression of Norm .Norm is expressed with 16 bits and each bit expresses the next IS.

しかし、このルールでは、協調的な戦略が安定的に進化することはできない。なぜなら、Image Score に基づいて行動する戦略 (Discriminator) は All-C に侵入を受けてしまい、All-C は All-D の侵入を許容してしまうからである⁹⁾。

そこで、Ohtsuki⁹⁾らは Norm (規範) という、誰を良い人とみなし、誰を悪い人とみなすかの動的なルールを取り入れ、「良い人が悪い人に協力しなかった場合、その人は良い人のままでいられる」(なぜなら悪い人を罰したのだから)という要素が規範にないと、間接互惠性が進化できないことを ESS 解析により理論的に示した。そして、「leading8」と呼ばれる最高レベルの協力を安定的に引き出すことのできる 8 種類の規範を発見した。

また、Chalub ら⁹⁾によって、Norm を個人とは別のスケールの「島」と呼ばれる集合体ごとに進化させることにより、その規範に「leading 8」と同様の加罰のシステムが創発することが確かめられている。しかし、このモデルは群島に生息する生物種を想定したモデルで、島間における Norm の伝播をタカハトゲームによって定義される利得に応じて定義しているため、プレイヤーが行っているゲームの影響とタカハトゲームの影響を分離することができない。

また、これらの既往のモデルでは異なる Norm との対戦が考慮されていない。Ohtsuki らは同一 Norm のみの社会に別の Norm の個体が侵入してきたとしても、他個体の利得にほとんど関係しないため、ESS を論じる際には無視できるとしている。

しかし、実社会には多様な規範や価値観の人々が存在し、相互作用を行っている。このような状況を考察することは十分価値があると考えられる。

本研究では、Chalub らのモデルを参考に、Norm と個人が持つ戦略を共進化させるモデルを作成し、異なる Norm との対戦が許容された社会における協調創発について考察する。

2. モデル

本モデルは、 N 人のプレイヤーが存在する社会を想定する。各プレイヤーは 16bit で表現される Norm (2.1 参照) と呼ばれ

表 1 プレイヤーの戦略

		IS_{opp} / IS_{focal}		Name
G / G	G / B	B / G	B / B	
C	C	C	C	All-C
D	D	D	D	All-D
C	C	D	D	Discriminator
D	D	C	C	Bad World Discriminator

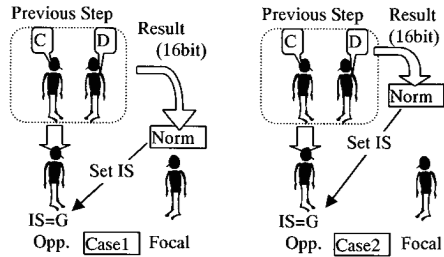


図 2 IS の決定方法 Case1 は相手の前回行動を参照する。Case2 では相手の前回相手が決めた IS を参照する。

Fig.2 Setting method for IS. (Case1) Refer to a partner's last action. (Case2) Refer to IS decided by partner's

social norms shared by one group to which they belong. This group is called an island. Also, each player has an IS (16bit) which indicates the information of the previous game result and Norm. Action (C or D) is 4bit expressed strategy (2.1 described later) based on the opponent's and own IS.

各プレイヤーは島とは別に定義されるグループ (以後、対戦グループ) に所属しており、同じ対戦グループに所属するプレイヤーとランダムに相手を変え回りながら One-shot 2x2 ジレンマゲームの対戦を行う。用いるゲーム構造は 2.4 で述べる。

対戦グループはプレイヤーごとに個別に定義されている。対戦グループには同じ島に所属するプレイヤーが P_{same} 残りは異なる島に所属するプレイヤーで構成されている。島と対戦グループはエピソード内で不変である。

プレイヤーの戦略の更新は島ごとに GA アルゴリズムを用いて行う。世代内で得られた平均利得を適合度に、島ごとに繰り返しルーレット選択による淘汰を行い、各プレイヤーの戦略を更新する。これらの個体の戦略には一定の確率 $p_{crossover}$ で交叉イベントが発生し、交叉イベントが発生した戦略同士で多点交叉を行い、次世代に交際交叉によって生成された個体を残す。また、一定の確率 $p_{mutation}$ で突然変異が発生し、戦略はランダムに変更される。

2.1. Norm の定義と進化メカニズムについて

島がそれぞれ持つ Norm は 16bit の情報で構成され、4bit の前回行動 (前回の相手と自分の IS と前回の相手と自分の行動) 別に相手に設定する IS を記憶している。これに基づいて相手

の IS を決定する(図 1).

Norm の進化はその Norm グループに所属するプレイヤーの利得の総和を適合度として、適合度が高い Norm を適合度が低い Norm に対し、それぞれ bit ごとに(1)式の確率で各 bit をコピーすることによって起こる。高い方の適合度を $fitness_{high}$ 、低い方の適合度を適合度を $fitness_{low}$ とすると

$$P_{high \rightarrow low} = \alpha \frac{fitness_{high} - fitness_{low}}{fitness_{high} + fitness_{low}} \quad (1)$$

コピーイベントはすべての Norm グループから 2 つの Norm グループを選ぶすべての組み合わせ ($N_{island} C_2$) において、それぞれ W_p の確率で発生する。また、コピーを行う際、 $P_{norm_mutation}$ の確率で Norm の bit を反転する。

2.2. プレイヤーの戦略の定義について

対戦相手は自分の IS_{ego} と相手の IS_{opp} に基づいて、行動を C と D から選択する。よって戦略は 4bit で定義できる。ここで気をつけなければならないのは、IS の G と B という呼び名はあくまでラベルであり、その言葉に意味はない。例えば、B が良い行動を表すラベルとなっている社会も十分存在しうる。表 1 にいくつかの戦略の定義例を挙げる。

2.3. IS の設定について

IS を決定する際に、いつの時点で IS を設定するか、誰の Norm を基準に設定するかで、Norm や IS の性質 (情報の伝播の仕組み) が変わってくる。これには 2 つケースが考えられる。その二つを図 2 に示す。

まず、対戦を行う前に自分の Norm を基準として相手の前回行動を元に相手の IS を決定し、それを今回の対戦に利用する方法がある。これは、前回行動を情報として社会から受け取ることができて、それを元に自分のルールや規範(Norm)に照らし併せて相手の善悪(IS)を判断していることになる。以後、これを Case1 とする。

次に、対戦を行った後に自分の Norm を基準として相手の

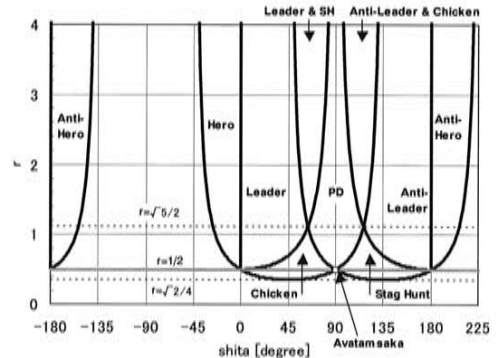


図 3 θ と r による 2x2 ゲームの表記

Fig.3 The scene of 2x2 game world.

今回行動を元に相手の IS を決定し、その IS を次回対戦相手を利用するという方法がある。この場合、IS を情報として社会から受け取っている。この場合、相手の IS を決定したのは第三者のプレイヤーであり、異なる Norm と対戦する場合は、自分の持つ Norm とは全く違った Norm によって判断された可能性がある。これは、情報が噂のような曖昧なものに丸められて送られてきていることになる。以後、これを Case 2 とする。

2.4. 2x2 ゲームのゲーム構造

すべてのゲーム構造は次式のようにパラメータ θ [deg] と $r \equiv r_2 / r_1$ で表すことが出来る⁷⁾。

$$P = x_0 - 0.5 \cdot r_1 \cdot \cos(45) \quad \dots(1)$$

$$R = x_0 + 0.5 \cdot r_1 \cdot \cos(45) \quad \dots(2)$$

$$S = x_0 + r_2 \cdot \cos(45 + \theta) \quad \dots(3)$$

$$T = x_0 + r_2 \cdot \sin(45 + \theta) \quad \dots(4)$$

これにより、多様なゲームを θ と r ($r = r_2 / r_1$) の 2 パラメ

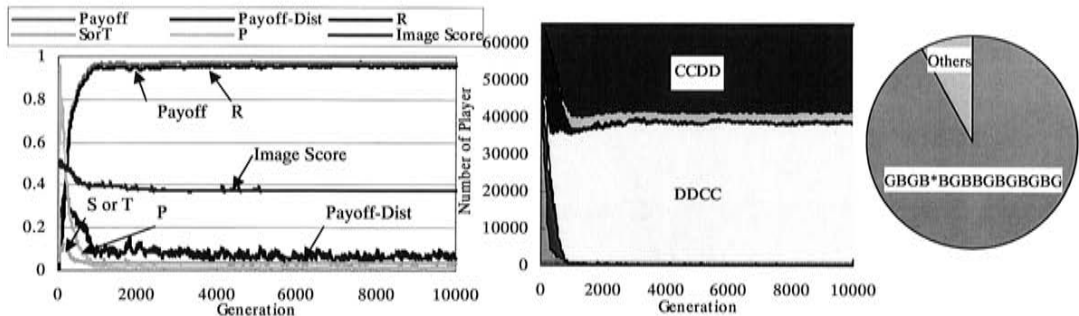


図 4 $P_{same}=1.0$ $r=1.6$ における結果

(左) 利得, P,R,S,T の生起確率 平均 IS (中) 戦略分布 (右) 10000 世代における Norm 分布

Norm 分布は比率が 20% を超える Norm の全体における割合を示す

Fig. 4 Result in $P_{same}=1.0$ $r=1.6$. (Left) Payoff. Occurrence probability of P,R,S,T. Average of IS. (Center) Distribution of player's strategy. (Right) Distribution of Norm in 10000 generation.

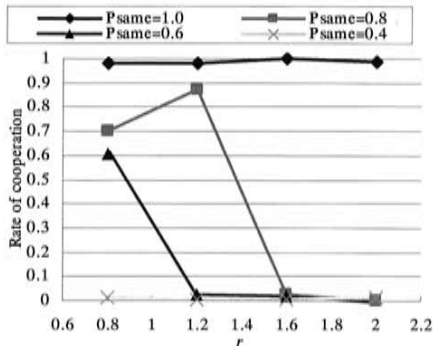


図5 10000 世代時の平均協調率 (Case1)
Fig.5 Average of Cooperative rate in 10000 generation. (Case1)

ータで表すことができる (図2)。太線で囲われジレンマゲーム名が冠されている領域以外のエリアが、ジレンマのない Trivial ゲームである。

(x_0 は解可能域がシフトするだけで P, R, S, T の相対関係上意味がないので、実質的には 2 パラメータとなる)。これら 2 パラメータで表されるゲーム空間を表すと図3のようになる。太線で囲われ、ジレンマゲーム名が冠されている領域以外が、ジレンマのない Trivial ゲームである。

3. 実験結果および考察

実験は島のプレイヤー数 $N_{agt_island}=128$ 、対戦グループのプレイヤー数 $N_{agt_battle}=128$ 、島数 $N_{island}=512$ 、対戦グループ数 $N_{battle}=512$ で行った。これより、全プレイヤー数は $N_{agt_island} \cdot N_{island}=65536$ に設定した。 $p_{crossover}=0.3$ 、 $p_{mutation}=0.01$ 、 $p_{norm_mutation}=0.01$ 、 $\alpha=1.0$ 。シミュレーションは 10000 世代行なった。

本モデルでは、IS は focal プレイヤー自身が直接操作できないため、基本的には R 互恵が焦点となる。従って、本稿では囚人のジレンマのゲーム構造 (θ は 90°) に固定し r を変化させた。

3.1. 異なる Norm との対戦がない場合

$P_{same}=1.0$ における結果を、図4に示す。なお、異なる Norm との対戦がない場合、2.3 で述べた 2 つの Case は違いが無くなる。

異なる Norm との対戦がない場合、協調率はほぼ 1 で安定し、協調社会が構築された。また、そのときの戦略は CCDD、または DDCC となっており、相手の IS に応じて手を出し変える Discriminator が支配的となっている。Norm は GBGB*GBGBGBGBGBGB が全体の 98% を占めており、非常に高頻度に存在していることが分かる。これは「good が bad に裏切れば IS=good (8~11bit)」かつ「good が good に協調すれば IS=good (0~3bit)」。Bad 関連では「bad が bad に裏切れば IS=good (14~15bit)」かつ「bad が bad に協調すれば IS=bad

(12~13bit)」となっており、悪人に協調することを良しとしない価値観が構築されている。この Norm は 3) で見られた Norm と同じである。つまり、この Norm は 2 次のフリーライダー (悪者を懲らしめない; punishment をただ乗りする) 対策になっているといえる。それだけでなく寛容さ、つまり IS が bad なプレイヤーも悪者を懲らしめると good に復活することが可能な safety ネットが考慮されている点がより重要で、この枠組みがあって高い協調率が達成される。さらに、この Norm は Good が悪人を表す島においても、Bad が悪人を表す島とまったく同様の効果をもたらす。これにより、Norm が社会全体に広がって行きやすい。この点も、この Norm が支配的になる原因の一つと云える。

以後、これに類似した Norm を加罰 Norm とする。

3.2. 異なる Norm との対戦がある場合

3.2.1. Case1

Case1 における、異なる Norm との対戦を考慮した結果を図5に示す。まず、 P_{same} が小さくなる、つまり、異なる Norm との対戦比率が高くなると協調率が低くなる傾向が見られる。これは、異なる Norm との対戦にはいくつかの困難があるからだと考えられる。例えば、Good の意味が異なった島同士での対戦の場合、お互いが協調に対して良い IS を返す Norm であったとしても、一方にとっては悪い IS をつけられてしまうことになる。

また、 $P_{same}=0.8$ において r が大きくなる、つまりジレンマゲームが難しくなるに従って協調率が高くなっている。図6に利得の世代推移を示す。 $r=0.8$ では、早い段階で協調社会が

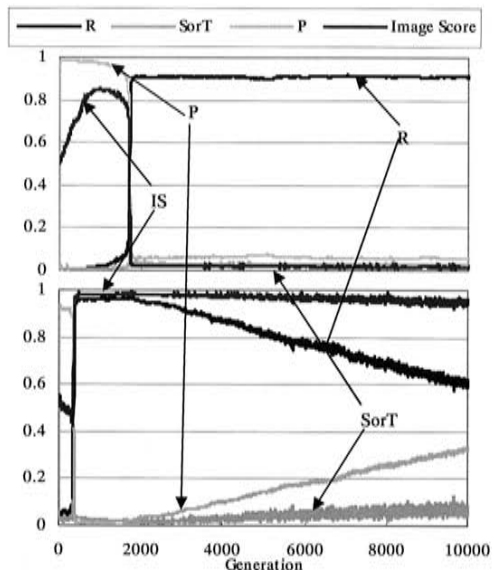


図6 Case1 $P_{same}=0.8$ (上) $r=1.2$ (下) $r=0.8$ における世代推移
Fig.6 Result in Case1 $P_{same}=0.8$ (Upper) $r=1.2$ (Under) $r=0.8$.

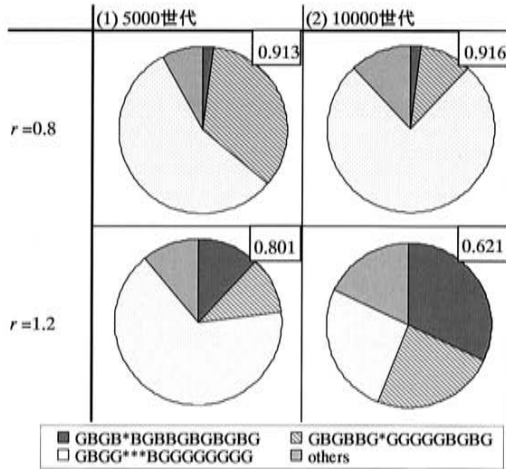


図7 Case1 $P_{same}=0.8$ における Norm の分布
 比率が10%以上の Norm と加罰 Norm を分類した。
 右上の数字はその時点の協調率

Fig.7 Distribution of Norm in Case1 $P_{same}=0.8$. The Ratio classified more 10% of Norm, and Punishment Norm. The number at the upper right of a pie chart is the cooperative rate at the time.

構築されているが、その後崩壊している。しかし、 $r=1.2$ では、協調社会が構築されるのは遅いが、長く協調社会が保持されているのがわかる。図7に5000世代と10000世代において、それぞれの社会で多数を占める Norm を示す。GBGBB*GGGGGBGBG は2.1で述べた加罰 Norm と似ているが、相手が Bad で自分が Good であれば相手に Good をつけるという Norm になっている。GBGB***BGGGGGGGG は、相手が Bad で D を出さない限り Good をつけるという Norm である。 $r=1.2$ において、5000世代で加罰 Norm がある程度存在しており、10000世代においても協調社会を保持できている。

$r=0.8$ では、5000世代の時点で加罰 Norm の比率は低く、無条件に Good を出す Norm が大半を占めていた。この時点では協調社会が構築されているが、10000世代では、協調社会が崩壊している。

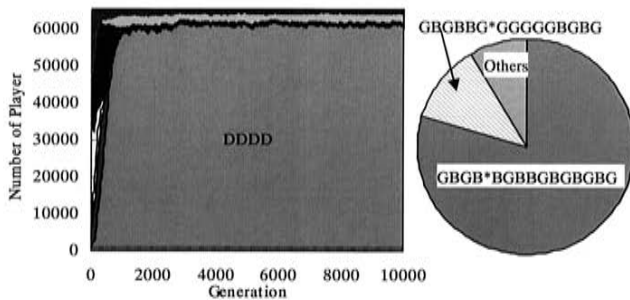


図9 (左)戦略分布 (右)10000世代における Norm 分布
 Fig.9 (Left) Evolution of strategy distribution. (Right) Distribution of Norm.

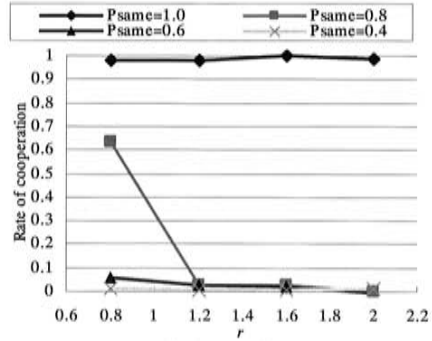


図8 10000世代時の平均協調率 Case2

Fig.8 Average cooperative rate in Case2.

先ほども述べた要に $r=0.8$ では、早い段階で協調社会が構築されているが、ジレンマゲームが難しくない故に、この協調社会は無条件に皆に Good をつける、Good から D を出された時のみ Bad をつけるという甘いルールでも協調できてしまい、安定できなかったと云える。それに対し、 $r=1.2$ では、しばらく D プレイヤーも混在する社会が続いたため、加罰 Norm が発生したと考えられる。

3.2.2. Case2

Case2の結果を図8に示す。Case1と比べ、協調率が低くなっている。これは、ISの評価が他者によって行われるため、決定する際に使われる Norm が自分と異なる可能性がより高くなるためと思われる。

$P_{same}=0.8$, $r=1.2$ のある試行の10000世代における戦略分布と Norm の分布を図9に示す。ほとんどが加罰的な Norm によって構成されている。しかし、プレイヤーの戦略はほとんどが All-D であるため、利得を得ることはできていない。なぜ、このような現象が起こったのであろうか？

各島の平均利得を調べたところ、ある一つの島のみ協調関係を保って高い利得を得ていることがわかった。これよりこのことは次のように説明できる。

初期にほとんどの島はジレンマを克服できず、All-D 戦略ばかりとなるが、戦略の多様性が残っているうちに、加罰 Norm を獲得した島があったとする。当然、Discriminator 戦略により、その島は高利得をあげる事ができる。そのため、Norm だけは他の島に伝播していくが、他の島は All-D 戦略が大半を占めており、加罰 Norm を有効利用できないためこのような現象が起こったと推測できる。

これは、Chalubら⁹⁾のモデルで考慮されていた、島間における戦略の交換が行われなために起こる現象と考えられるため、島間でプレイヤーの移住が行われれば、ジレンマを

		r			
		0.8	1.2	1.6	2
$P_{err}=0.00$	$P_{same}=1.0$	0.979	0.981	0.999	0.990
	$P_{same}=0.8$	0.700	0.873	0.025	0.000
	$P_{same}=0.6$	0.608	0.025	0.020	0.000
	$P_{same}=0.4$	0.012	0.009	0.006	0.012
$P_{err}=0.05$	$P_{same}=1.0$	0.979	0.981	0.999	0.990
	$P_{same}=0.8$	0.837	0.944	0.025	0.000
	$P_{same}=0.6$	0.648	0.025	0.020	0.000
	$P_{same}=0.4$	0.012	0.009	0.006	0.012

図 10 10000 世代時の平均協調率 Case1 Action Error 有り
Fig.10 Average cooperative rate in Case1,using action error.

克服できる可能性がある。

3.2.3. Action Error の効果

簡単なジレンマゲームにおいて、加罰 Norm が進化しない理由として、初期に R ばかりを取りに行く社会が構築されてしまうと、前回 Bad なプレイヤーに対して裏切りをするよりも、盲目的に協調をする方が高い利得を挙げられるようになってしまうからだと考えられる。そこで、ある一定の確率で協調を出すところで裏切る、Action Error (出し間違え) を導入してみる。

Case1 において、Action Error を導入した結果を図 10 に示す。 P_{err} は Action Error が発生する確率である。若干ではあるが、Action Error を入れた方が協調率は高くなっている。Action Error が 5% のケースでは、 $r=0.8$ において、Action Error がいない場合は創発しない加罰 Norm が創発していることを確認した。

これらの事より、Action Error は加罰システムの創発サポート効果があると考えられる。このような効果がある理由として、次の事が考えられる。

Action Error がいない場合、早い世代に協調に至ってしまうと、それ以上 Norm の進化は行われず、加罰システムのない Norm が社会に蔓延し、遺伝的浮動 (遺伝子のばらつき) の余地がなくなる。そのような社会では、裏切りプレイヤーの介入を許してしまう。Action Error がある場合は、協調社会が構築されても、一定数の裏切りを出すプレイヤーが存在するため、加罰を行う Norm が進化すると考えられる。

これは、協調的なプレイヤーばかりの社会では、協調を出すことを前提にしたルールしか構築されないため、社会が成熟するには、ある程度の裏切りを行うプレイヤーがいて、社会に脅威を与えている必要があることを示唆している。

4. 結論

本稿では Norm Reputation システムを用い、Norm と戦略の二段階の進化機構をもつゲームモデルを作成した。これにより、以下の結果を得た。

(1) Norm Reputation システムを用いることにより、加罰的なシステムが創発することを確認した。

(2) 異なる Norm との対戦が許容される場合、安定的な相互関係は創発しなかった。また、行うジレンマゲームが簡

単なほど、安定的な相互関係を築き得ないことがわかった。これは、簡単なジレンマゲームだと、加罰 Norm を進化させずに協調社会を構築してしまい、結局裏切りプレイヤーの侵入を許してしまうためだと思われる。

(3) Action Error を導入すると、協調率が高くなるケースがあることを示した。これは、協調的なプレイヤーばかりの社会では、協調を出すことを前提にしたルールしか構築されないため、社会が成熟するには、ある程度の裏切りを行うプレイヤーがいて、社会に脅威を与えている必要があることを示唆している。

参考文献

- 1) Hamilton, W.D.; The evolution of altruistic behavior, *Am. Nat.* 97, 35 4-356 (1963).
- 2) Trivers, R.; The evolution of reciprocal altruism, *Q. Rev. Biol.* 46, 35-57 (1971).
- 3) Axelrod, R.; *The Evolution of Cooperation*, New York: Basic book (1984).
- 4) Nowak, M.A., Sigmund, K., Evolution of indirect reciprocity by image scoring, *Nature* 393, 573-577, (1998).
- 5) Ohtsuki, Hisashi, & Iwasa, Yoh. 2004. How should we define goodness?—reputation dynamics in indirect reciprocity. *J. Theoret. Biol.*, 231(1), 107–120 (2004).
- 6) F.A.C.C. Chalub, F.C. Santos and J.M. Pacheco, The Evolution of Norms, *Journal of Theoretical Biology*, Volume 241, 233-240(2006)
- 7) Tanimoto, J., Sagara, H., "Relationship between dilemma occurrence and the existence of a weakly dominant strategy in a two-player symmetric game", *BioSystems*, *In Press*, (2006).