

## 派閥モデルを用いたジレンマゲームにおける Green Beard Effect の考察

相良 博喜<sup>†</sup> 谷本 潤<sup>‡</sup>

<sup>†</sup>九州大学大学院総理工学府環境エネルギー工学専攻・修士課程

<sup>‡</sup>九州大学大学院総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 九州大学総理工学府環境工学研究室

E-mail: <sup>†</sup> uminchu@uminchu.jp, <sup>‡</sup> tanimoto@cm.kyushu-u.ac.jp

**概要** 多人数 2x2 ジレンマゲームモデルにおいて、選好度に基づく対戦相手の選別機構（派閥モデル）を導入することによりジレンマが解消されるが、相手を欺く Mimic エージェントの創発により、安定的な協調社会が構築されないことを確認した。また、“仲間”＝“協力者”の等式（Green Beard Effect）を仮定することにより安定的な協調社会が構築されることを確認した。

**キーワード** 囚人のジレンマ Green Beard Effect 派閥モデル

### Consideration of Green Beard Effect in the dilemma game using a cliquish model

Hiroki Sagara<sup>†</sup> Jun Tanimoto<sup>‡</sup>

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

<sup>‡</sup> Prof., Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Dr.Eng

E-mail: <sup>†</sup> uminchu@uminchu.jp, <sup>‡</sup> tanimoto@cm.kyushu-u.ac.jp

**Abstract** This paper reports that the Prisoner's Dilemma of a 2-players game can be diluted by the cliquish model, where a focal player selects his opponent by using his personal preference list that indicates preferable opponents to be matched in a game. Observed cooperation supported by the cliquish model is, however, not stable, since a mimic mutant can invade into a certain cliquish group. The mimic initially pretends a "Saint" agent (offering C even though his opponent offering D) when he is involved into the group. But after registered as a friend, he turns to offer D in order to exploit. Interestingly, assuming the Green Beard Effect where a friend must always cooperate to others in a cliquish group, a stable cooperative society can be built.

**Keywords** Prisoner's Dilemma Green Beard Effect Cliquish Model

#### 1. 緒言

ジレンマゲームにおける協調の創発は、生物学や社会科学の主要な問題として取り上げられている。ジレンマゲームは進化ゲーム理論を基礎とし、ここ 10 年非常に多くの研究が積み重ねられてきた。主には One-shot ゲームでは裏切り (Defect, 以下 D) 合いの Nash 均衡に陥ってしまう Prisoner's Dilemma (以下 PD) が取り上げられ、どのような付加機構を入れると、互いに協調 (Cooperation, 以下 C) 出来るかが、ジレンマゲーム研究の中心的な課題であった。

Axelrod<sup>1)</sup>は 2 人繰り返し囚人のジレンマゲーム (2-IPD) において、Tit-for-Tat 戦略 (しっぺ返し戦略, 以下 TFT) と呼ばれる相互的な戦略が、安定的な協調関係を構築できることを示した。2-IPD において TFT が有効なのは、C には C で応じ、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、TFT は戦略実行における過誤等の侵入に脆弱であるため、これ

らの戦略が考慮された場合、安定的な協調関係を保つことができないとされている<sup>2)</sup>。

また、TFT 戦略は不特定多数との対戦を強いられる環境では、安定的な協調関係を保つことができない (間接互恵を構築することは困難)。なぜなら、次の対戦相手が協調的である保証がないからである。

この点に関して、Hruschka<sup>3)</sup>らは人間社会における間接互恵をモデル化するため、対戦相手を認識し、限定的な相手とだけ対戦するモデルを構築している。プレイヤーはこれまで対戦した相手に関するリスト (友人の記憶) を有し、リスト内でより協調的な相手を指名して対戦する。リストから対戦相手を除外する協調率閾値を進化的に更新することにより、リスト内の対戦相手に対しては寛容で新規に対戦する相手に対しては裏切りを持って応じる「派閥戦略」が創発し、過誤に対してロバストなシステムが構築可能であることを示した。これは、人間社会システムにおける「友情」のような、中長期的な社会関係の比喩になっている。

Hruschka らのモデルでは、プレイヤーは派閥を組むことにより、非協調的なプレイヤーを排除しており、派閥内のプレイヤーの過誤に対して寛容な姿勢をとる

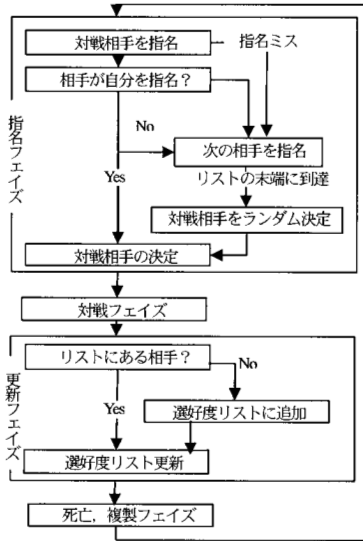


図1 フローチャート  
Fig.1 Flow Chart

形質を獲得している。これは派閥内のプレイヤーに対し、過誤がない限り協調的な振る舞いをする、つまり、仲間であることと協調者であることが同一である、Green beard effect<sup>9)</sup>が仮定されている。しかし、現実の生物社会に協調行為に関する Green beard effect は確認されていない。

ジレンマゲームにおいて、Green beard effect が協調関係の構築に大きくかかわっていることが示されている。谷本<sup>5)</sup>は、Green beard effect と他者との形態上の識別システムであるタグが協調の創発に有効であるかを 2x2 ゲームにおいて検討した。タグ・システムとは、対戦相手の表象（タグにより表される）を識別し、それに応じて行動を切り替える（戦略により表される）機構である。複雑な認知機構やメモリ、学習等の知性を前提にしないメカニズムであるため、一般の生物システムにおける互惠創発の一論拠と考えられてきた。しかし、谷本は Green Beard Effect の仮定なしにはタグ・システムによる協調の進化は困難であることを示している。

では、Hruschka らのモデルにおいて Green Beard Effect を仮定せずに間接互恵的な協調関係を構築できないのであろうか？本研究では、Hruschka らのモデルを Green Beard Effect を仮定しないモデルに拡張し、協調的な戦略が創発する条件について考察する。

また、Hruschka らは言及していないが、派閥モデルにおける、限定的な相手と繰り返し対戦が出来る機構は、社会ネットワークを意味していると考えられ、このネットワークの特性を考察することも興味深い。

## 2. モデルの説明

本モデルは  $N$  人のプレイヤーによって構成される多人数 2x2 繰り返しゲームである。モデルは大きく 4 つのフェイズに分かれている。流れを図 1 に示す。

まず、最初に対戦相手を次章で述べるアルゴリズムに基づいて指名し、相手との合意を得られればペアを

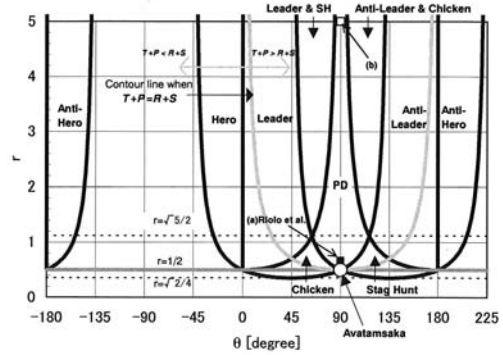


図 2  $\theta$  と  $r$  による 2x2 ゲームの表記。  
Fig.2 The scene of 2x2 game world.

組む指名フェイズ。次に対戦を行なって利得を得る、対戦フェイズ。そして、利得に応じて指名相手の優先順位を変更していく、適合度リスト更新フェイズ。最後に、利得に応じて、死亡、または自身の複製を作って、プレイヤーがもつパラメータを進化させる、死亡、複製フェイズがある。

この 4 つのシーケンス（以後、ラウンド）を GN 回繰り返す。これを一代とし、これを繰り返す。

### 2.1. 対戦相手の決定

プレイヤー  $i$  ( $i \in 1, \dots, N$ ) は過去に対戦した相手とその相手に対する選好度  $pref$  を記憶するリスト（以下、選好度リスト）を持つ。このリストは、可変長で、対戦後に後に述べる条件に基づいて追加、削除される。このリストの長さは可変長となっており、常に  $pref$  が大きい順にソートされている。

まず、各プレイヤーは選好度リスト上位、つまり選好度が大きい相手を指名する。また、過誤率  $em$  で指名ミスが発生し、指名できないとする。すべてのプレイヤーの指名が終了した段階で、指名した相手が自分を指名しているかチェックし、互いに指名していた場合、またはこれまでに指名されたことがある場合、ペアを作成する。ペアができなかった、または指名ミスをしたプレイヤーは、次に選好度が大きい相手を指名する。これを、すべてのプレイヤーがペアを見つけるか、選好度リストの最下位まで指名するまで繰り返す。ペアが作成できなかったプレイヤーは、ランダムに対戦相手が決められる。

### 2.2. 2x2 ゲーム

プレイヤーが行うゲームのゲーム構造は Tanimoto & Sagara<sup>6)</sup>の記法に基づき、以下の式で表される利得行列を用いる。

$$Payoff = \{P, R, S, T\} \quad \dots(1)$$

$$P = x_0 - 0.5 \cdot r_1 \cdot \cos(45) \quad \dots(2)$$

$$R = x_0 + 0.5 \cdot r_1 \cdot \cos(45) \quad \dots(3)$$

$$S = x_0 + r_2 \cdot \cos(45 + \theta) \quad \dots(4)$$

$$T = x_0 + r_2 \cdot \sin(45 + \theta) \quad \dots(5)$$

これにより、多様なゲームを  $\theta$  と  $r$  ( $r = r_2/r_1$ ) の 2 パラメータで表すことが出来る（図 2）。太線で囲われ

表 1 代表的な戦略の定義 表内記号は数値のパラメータの大きさを表している

Table 1 Definition of typical strategy. The sign in a table expresses the size of these parameter.

L :  $0 \leq \Delta, \chi, \delta < 1/3$  ...

M :  $1/3 \leq \Delta, \chi, \delta < 2/3$   $P \leq e, \tau < R$

H :  $2/3 \leq \Delta, \chi, \delta \leq 1$   $R \leq e, \tau$

Strategy	$\Delta$	$\chi$	$e$	$\tau$	$\delta$
Defector	H	H	*	*	*
Cooperator	L	L	*	*	*
Cliquer	L	H	L~M	L	M
Out-for-tat	L	H	M	H	L
Mimic	M	L	L	M	L~M

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。

互いに C を出すことによって R をとる R 互恵に注目するならば、検討対象のゲームとしては D 支配の PD が重要となるので、本論では  $\theta=90\text{deg}$  とする。

各プレイヤーは選好度リストに乗っている相手と対戦する場合対戦相手と PD ゲームを行う。プレイヤーが出す手 (C or D) は、プレイヤーが個別に持つ 2 つのパラメータ  $\Delta, \chi$  で決まる。 $\Delta, \chi$  はともに  $[0, 1]$  の実数である。各プレイヤーは、対戦相手が自分の選好度リストに載っている相手であれば  $\Delta$ 、そうでなければ  $\chi$  を D を出す確率として用い、手を決める。

$\chi$  はこれまで対戦したことのない相手に出す手を決めるパラメータであるため、進化計算により  $\Delta$  に比べ大きくなる (D を出しやすくなる) 傾向がある。

### 2.3. 選好度リストの更新

次に、対戦結果に基づいて選好度リストを更新する。プレイヤーは過去選好度の重み  $\delta$  と選好度許容閾値  $\tau$  なる 2 パラメータにより、次式に基づいてプレイヤー  $i$  の選好度リストを更新する ( $k$  は  $k$  ラウンド目であることを示す)。対戦が行われなかった相手の  $\text{Payoff}_{i,k}$  は 0 となる。

if  $i$  plays  $j$  then

$$\text{pref}_{i,j,k} = \delta_i \cdot \text{pref}_{i,j,k-1} + (1 - \delta_i) \cdot \text{Payoff}_{i,k}$$

else

$$\text{pref}_{i,j,k} = (1 - \beta) \cdot \text{pref}_{i,j,k-1} \quad \dots(6)$$

$\beta$  は 1 以下の定数で、対戦しなかった相手の選好度を割り引くもので、記憶の忘却を表す。

自分の選好度リストに載っていない場合、今回の利得がパラメータ  $e$  を越えていれば相手を新しく選好度リストに載せる。初期の  $\text{pref}_{i,j,k}$  は今回得た利得となる。

そして、自身の選好度リストの中で選好度が  $\tau$  以下のものは選好度リストから切り捨てる。

### 2.4. 死亡、自己複製

プレイヤーは対戦後、次式に示す適合度  $F_{i,k}$  を算する。そして、各プレイヤーは(5)式で定義される確率で死亡する。減少した分は、利得によるルーレット選択で選ばれたプレイヤーが自己複製することにより補充される。複製する際、選好度リストは引き継がれない。

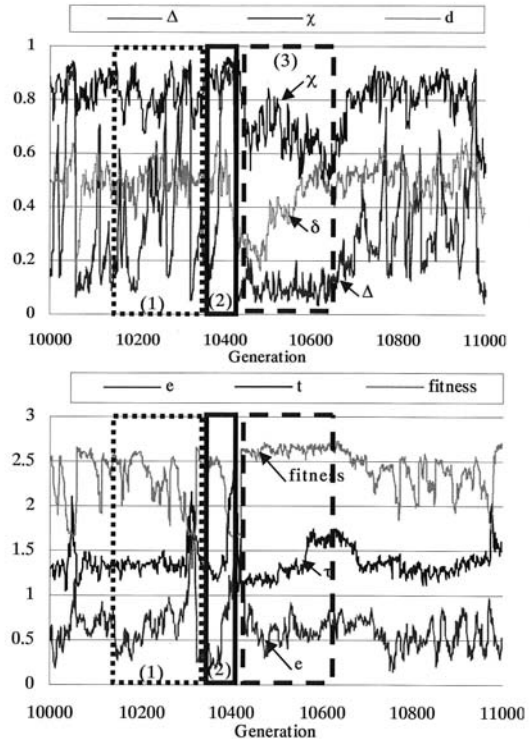


図 3 プレイヤーパラメータと利得の世代推移  
Fig.3 Evolution of strategy parameter and payoff

$$F_{i,k} = \gamma \cdot F_{i,k-1} + (1 - \gamma) \cdot \text{Payoff}_{i,k} \quad \dots(7)$$

$$p_i = C_d \left( 1 - \frac{F_{i,k} - F_{\min,k}}{F_{\max,k} - F_{\min,k}} \right) \quad \dots(8)$$

$C_d$  は死亡確率係数、 $F_{\max,k}$ 、 $F_{\min,k}$  は全プレイヤーにおける適合度最大、最小値、 $\gamma$  は過去の利得の重みである。また、 $p_m$  の確率でパラメータをランダムな値に変える突然変異を考慮する。

## 3. 実験

実験における共通のパラメータ設定は、 $N=1000$ 、 $P_m=0.001$ 、 $x_0=2.0$ 、 $r_1=2.828$ 、 $\theta=90\text{deg}$ 、 $\beta=0.05$ 、 $GN=N$  とした。また、1 エピソードは 20000 世代まで計算し、10000-20000 世代の平均値の 5 試行アンサンブル平均を擬似的な均衡の値とした (結果の章で示すように世代推移は安定的な均衡を示すものではなく、頻りに振動するため)。

### 3.1. 戦略定義

プレイヤーの戦略は以上で述べた  $\Delta$ 、 $\chi$ 、 $e$ 、 $\tau$ 、 $\delta$  の 5 つのパラメータによって定義される。

これらによって特徴づけられる代表的な戦略を表 1 に示す。Cliquer はこれは、Hruschka<sup>3)</sup>らの報告に登場する戦略で、 $\Delta$  が大きく、 $\tau$ 、 $\delta$  が小さいため、選好度リストにある相手 (以後、仲間と呼ぶ) に対して協調しつつ、仲間の裏切りに対して寛容な戦略である。

それに対し、Out-for-Tat は集団で行う TFT 戦略で、

表 2 (1)~(3)における対戦発生頻度とその利得  
Table.2 Frequency of play for each strategy pair.

Fig.3-(1)

Focal Strategy	Opp Strategy	Frequency	Focal Reward
Cooperator	Cooperator	0.1783	2.6168
Cliquer	Cliquer	0.4229	2.5271
Cooperator	Defector	0.0830	0.7876
Defector	Cooperator	0.0830	3.2869
Cliquer	Cooperator	0.1750	2.5568
Cooperator	Cliquer	0.1750	2.2163
Others		0.1408	2.1210

Fig.3-(2)-1

Focal Strategy	Opp Strategy	Frequency	Focal Reward
Cliquer	Cliquer	0.1827	2.7159
Mimic	Defector	0.1274	1.1181
Defector	Mimic	0.1274	2.5912
Mimic	Out-for-tat	0.1091	2.0451
Out-for-tat	Mimic	0.1091	1.9263
Mimic	Cliquer	0.1614	2.9126
Cliquer	Mimic	0.1614	0.9126
Others		0.4194	1.9980

Fig.3-(2)-2

Focal Strategy	Opp Strategy	Frequency	Focal Reward
Defector	Defector	0.3808	0.9800
Cliquer	Cliquer	0.1334	2.7366
Mimic	Defector	0.1305	0.7876
Defector	Mimic	0.1305	2.5869
Mimic	Out-for-tat	0.1125	2.0440
Out-for-tat	Mimic	0.1125	1.8845
Mimic	Cliquer	0.0925	2.9513
Cliquer	Mimic	0.0925	0.9010
Others		0.1503	1.6298

Fig.3-(3)

Focal Strategy	Opp Strategy	Frequency	Focal Reward
Cooperator	Cooperator	0.2783	2.8168
Cliquer	Cliquer	0.5229	2.5271
Cliquer	Cooperator	0.1750	2.5568
Cooperator	Cliquer	0.1750	2.2163
Others		0.0238	2.2410

仲間であっても裏切られたらすぐリストから排除する戦略である。

また、Mimic 戦略は最初は C を出すことにより、高い選好度で相手のリストに記載され、その後裏切ることにより相手から貪る戦略である。τ が小さい理由は、相手が仲間にも D を出す戦略であれば、リストから排除するためである。

#### 4. 結果および考察

##### 4.1. 協調崩壊の要因について

図 3 は  $r=1.0$ ,  $\gamma=0.3$  のあるエピソードにおける 10000-11000 世代での世代推移である。(A)に  $\Delta$ ,  $\chi$ ,  $\delta$ , (B)に  $e$ ,  $\tau$ ,  $F_{i,k}$  をプロットした。全体的に  $\Delta$  が小さく、 $\chi$  が大きい。これは、仲間に協調し、仲間以外に裏切る戦略が創発していることを示している。これは、Hruschka<sup>3)</sup>らの報告と同様の傾向である。しかし、彼らとの相違点として、利得が安定していない。また、このことが効いて、協調関係が保持できていないことが指摘される。図 3 では、大まかに分類して 3 つの相が見て取れる。第一に、利得が振動し、安定しない相 (1)。第二に、一挙に協調関係が崩壊する相 (2)。第三に、 $\chi$  が低下し、安定的に協調関係を保っている相 (3) である。これらの相は (1)→(2)→(3)→(1)、または (1)→(2)→(1) の周期で、試行中に頻繁に発生する。

それぞれの相における主だった戦略同士の対戦の

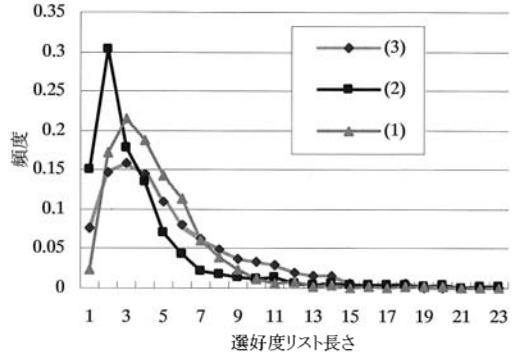


図 4 選好度リスト長分布

図 3 の (1), (2), (3) のそれぞれの区間の平均を取った Fig.4 Distribution of preference list length.

It is the average of each section of Fig. 3-(1),(2),(3).

発生頻度とそのとき得られた利得を表 2 にあらわす。(2)の相は、劇的に戦略分布が変化しているため、前半を (2)-1、後半を (2)-2 に分けてあらわしている。

まず (1) の相は、過半が Cooperator と Cliquer で占められているが、幾許かの Defector の存在が引き金となって短期間で細かく協調率が振動していると考えられる。

次に、時折、(1)の相から(2)の相へ劇的に変化する。このとき、表 2-(2)-1 のように、Mimic が発生している (1) ではほとんど存在しない。この戦略は劇的な変化の中、短時間だけ発生する戦略で、多数派にはなっていない。それは、Defector や Out-for-tat に対し、高い利得を上げられないからである。ただし、この戦略は Cliquer との対戦時に Cliquer の利得を著しく下げる。これは、Cliquer は仲間に対して寛容であるため、仲間になった Mimic の裏切りを許容してしまうからである。よって、Mimic は Defector が社会に蔓延する呼び水となっているといえる。その後、Mimic や Cooperator が絶滅すると Defector は貪る対象がいなくなるため、Cliquer や Out-for-tat に滅ぼされている。Defector がほとんど根絶されると、(3) のように Cooperator と Cliquer が高い利得を得ている。このときの選好度リスト長の頻度分布を図 4 に示す。(3) のとき、長いリストをもったプレイヤーが多くなっており、巨大なコミュニティを作成して Defector との対戦のリスクを避けることで、安定的になっていると考えられる。

この Mimic の戦略はタグモデルにおける Mimic と同様の働きをしている<sup>3)</sup>。同じタグであれば協調するというのは、Cliquer と同様の戦略である。このような戦略は仲間であることを詐称する戦略に対して脆弱である。

##### 4.2. Green Beard Effect の確認

Hruschka<sup>3)</sup> のモデルでは、Green Beard Effect を仮定することで安定的な協調関係を構築していた。それでは本モデルでも、 $\Delta=1$  (仲間は裏切らない) という仮定を導入すると、協調に至るのだろうか？

Hruschka<sup>3)</sup> のモデルでは  $\Delta=1$  が 0 の離散的な値のみに制限し、 $e_s$  の確率で C を D と出してしまう過誤を導入していた (本モデルは  $\Delta$  が実数であるため不要)。そこ

で、本モデルでは、 $\Delta=0.995$  か 0 しか取れないとした。また、 $\chi$  が 0.5 以下のプレイヤーのみ  $\Delta=0$  になると制約を課した。これは、Mimic が発生しないようにするためである。結果を図 5 に示す。

協調が安定的になることが確認された。これより、本モデルにおける協調関係構築には、Green Beard Effect の前提が非常に強い条件であると考えられる。タグによる間接互惠をサポートする上でも同様の事が云える<sup>3)</sup>。

#### 4.3. 各種パラメータを変えての検討

次に、ゲーム構造やパラメータを変化させて、検討した。 $\gamma$  と  $r$  を変化させたときの協調率を 10000-20000 世代の平均値の 5 試行アンサンブル平均したものを図 6 に示す。

$\gamma=0.2$  でピークを持つ分布となっている。 $\gamma=0.2$  のときの戦略分布を調べたところ、Mimic がほとんど現れていないことがわかった。これは、 $\gamma$  が小さくなることで、現在の利得がより重視されるようになり、仲間でないプレイヤーに C を出して仲間に入りこむ時点で淘汰されてしまう確率が高くなってしまふからだと思われる。

本モデルの Mimic は最初に相手に C を出して食らせる（先行投資）を行うことで、仲間になり、その後食ることにより、トータルで高い利得を得る戦略である。故に、先行投資した時点で淘汰されないために、プレイヤーは過去の利得をある程度保持する必要がある。よって、Mimic 戦略は過去の適合度（資産）を引き継がないようなモデルでは成立しないと言える。

これは、一般に身分を詐称するのにコストがかかるようなモデル（前述の先行投資のような仕組み）ならば、Mimic は過去に得られた利得（または未来に得られると予測される利得）でそのコストをまかなう必要があるため、過去の利得を参照できないモデルでは、戦略として成り立たないことを示唆しているといえる。

例えば、遺伝的アルゴリズムのような進化計算ならば、適合度はある世代の利得の積算になるため、Mimic は成立すると云える。

#### 5. まとめ

Hruschka らのモデルを Green Beard Effect を仮定しない戦略が定義できるモデル ( $\Delta=1$  に固定しない) に拡張し、Green Beard Effect と協調的な戦略が創発する条件との関係について考察した。これにより、以下の結果を得た。

(1) Green Beard Effect が安定的な協調戦略が創発するための強い条件であることを示した。Green Beard Effect を仮定しないモデルでは、仲間と見せかけつつ裏切りを行う Mimic 戦略が侵入する余地があるため、協調関係が崩壊してしまうことを示した。

(2) Mimic は先行投資を行うことで、仲間になり、その後食ることにより、トータルで高い利得を得る戦略であるため、過去（未来）の利得を参照する必要があることを示した。

#### 参考文献

- 1) Axelrod, R.; "The Evolution of Cooperation", New York: Basic book (1984).
- 2) Axelrod, R.; Douglas, D.; "The Further Evolution of

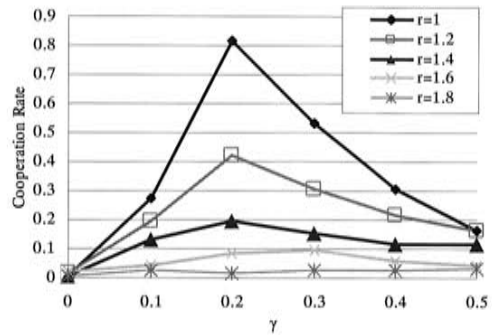
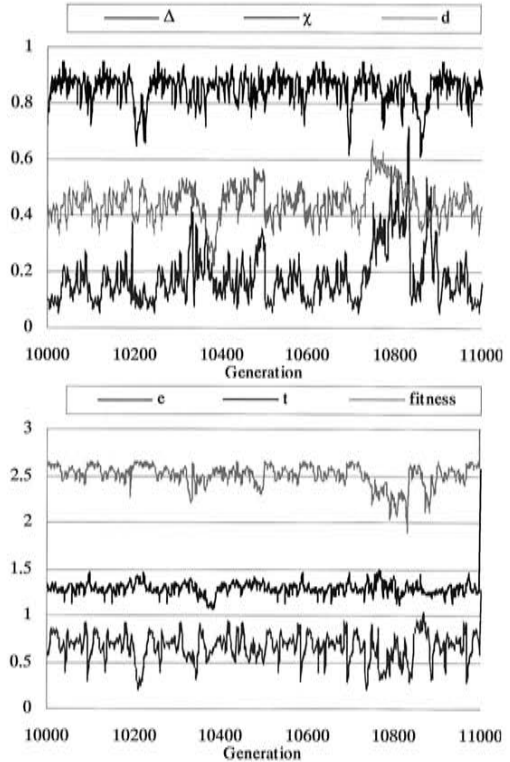


図 6  $\gamma$  と  $r$  を変化させたときの協調率 10000-20000 世代の平均値の 5 回試行アンサンブル平均。

Fig6. Cooperation rate by  $\gamma$  and  $r$ . 5 times trial ensemble average of 10000-20000 generation

- Cooperation", *Science*, 242, 1385-1390, (1988).
- 3) Hruschka,D.; Henrich,J; "Friendship, cliquishness, and the emergence of cooperation", *Journal of Theoretical Biology*, 239 1-15, (2006).
  - 4) Dawkins,R., "The Selfish gene", Oxford University Press, Oxford, (1989).
  - 5) 谷本 潤, "タグは協調創発に有効か?", 147 回知能と複雑系研究発表会, (2006)
  - 6) 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).