

ノイズ付き進化型繰返し囚人のジレンマにおける 長寿戦略の探究

糸井 良太¹ 田中 美栄子^{1,a)}

受付日 2012年4月19日, 再受付日 2012年6月7日,
採録日 2012年7月13日

概要: 戦略の自動進化を取り入れた繰返し囚人のジレンマモデルにおいて, 特に2進表現による1次元遺伝子配列の複写変異や分離変異を点変異と組み合わせることで遺伝子の長大化の効果を考察することを目的としたLindgrenモデルがあり, その結果として生き残りやすい戦略に共通した特徴のあることが指摘された. 我々はこのモデルに依拠して長大な遺伝子配列が出現するまでシミュレーションを行った結果をもとに, 長期間生存する遺伝子の配列パターンに共通する性質, 遺伝子長32以下の範囲内で, 長寿戦略の約60%が[1001 0*0* 0*0* 0001]という形の遺伝子に対応することを突き止め, さらにこの形の遺伝子の性質が1. 自分から裏切らない(最右遺伝子が1=協力), 2. 裏切られたらすぐに報復する(4つ組の第3要素が0=裏切り), 3. 裏切りが続くと自分から協力行動を行う(最左遺伝子が1=協力)の3要素であることを見出した. このことは, Lindgrenの先行研究で報告されている[1**1 0*** 0*** *001]の遺伝子配列に比べて, より明確に長寿戦略の遺伝子構造を同定することができたといえる. これらの戦略は単純なしっぺ返し戦略より強く, 多様な戦略との対戦に勝利した結果, 長期間生存し続けることのできるロバストな戦略であるといえる. このような戦略の出現・生存に対する条件について考察する.

キーワード: 囚人のジレンマ, 戦略の自動進化, Lindgrenモデル, 長寿戦略, 遺伝子構造

On the Sustainable Strategies in the Evolvable Iterated Prisoner's Dilemma with Noise

RYOTA ITO¹ MIEKO TANAKA-YAMAWAKI^{1,a)}

Received: April 19, 2012, Revised: June 7, 2012,
Accepted: July 13, 2012

Abstract: In the realm of iterated prisoners' dilemma equipped with evolutionary generation of strategies, a model has been proposed by Lindgren that allows elongation of genes by means of doubling and fission of one-dimensional genetic arrays multiplied by point mutation of genes. We conducted numerical simulations of this type of models under various conditions, and found that approximately 60 percent of long-lived strategies share the same structure represented by [1001 0*0* 0*0* 0001]. Such strategies are characterized by the following three features: (1) to cooperate as long as the opponent cooperates (2) to revenge immediately if defected (3) to offer cooperation after continuous defections. In other words, we have discovered the type of gene structure of long-lived strategies in more strict manner compared to the original structure suggested by Lindgren. Such strategies are stronger than a simple TFT, and robust strategies that survive under many matches with various kinds of strategies. We consider when and how such strategies are generated in the history of simulation.

Keywords: prisoner's dilemma, evolutionary generation of strategies, Lindgren model, sustainable strategies, gene structure

¹ 鳥取大学大学院工学研究科情報エレクトロニクス専攻
Department of Information and Electronics, Graduate
School of Engineering, Tottori University, Tottori 680-8552,
Japan

^{a)} mieko@ike.tottori-u.ac.jp

1. はじめに

ゲーム理論は, 互いに影響を与え合う複数の主体の間で
生じる戦略的な相互関係を研究するためのツールであり,

人間行動や経済の動きを理解するための手法として重要である。2人ゲームの最もよく知られた例である「囚人のジレンマ」は、2人のプレイヤーの合理的選択が、全体では良い結果にならない様子をモデル化したものである。二酸化炭素排出問題や核保有問題、また価格競争など、身近な問題を考える際にもモデルとして役立つものであり、昔から研究されてきた [1], [2], [3]。歴史的に有名なアクセルロッドの実験においてはしっぺ返し戦略が最強であった [4]。しかし、環境を変えると最強戦略もまた変化することや、最初からしっぺ返しが存在したのか、何らかのプロセスを経て発生してきたのか、また何らかの要因により消滅するのかなどの問題を考えると、もっと視点の時空を広げ、人工生命的な観点で様々な戦略の生成消滅の様子を観察することは大変興味のある研究課題といえる。

戦略を $\{0,1\}$ 列からなる遺伝子と見立て、遺伝的アルゴリズム (GA) などを用いて自動進化させることによって、強い戦略を自動生成させる試みは、Lindgren [6] によって導入された。遺伝子列の複写変異と分離変異を点変異と組み合わせることで、より環境に適応した長い遺伝子列を創成しようとする、人工生命的試みへと発展していった。

一方、田中ら [7] は、しっぺ返し戦略を最初から存在したものではなく、適者生存のルールによって自動生成されるというモデルを考察し、シミュレーションによってしっぺ返しが生成され生存する様子を観察した。

我々は Lindgren モデルの初期条件を変えて長期間シミュレーションする途中、カンブリア紀になぞらえることのできる、多数の異なる遺伝子が発生する時期のあることを見出した。また、その生成消滅の様子と、エントロピーの増減とを結び付けて、このような事象の発生する条件について考察した [8]。

本稿では再び Lindgren モデルのシミュレーションの結果から、生成された長い遺伝子列の持つ特徴を調べた。その結果、生存する長い遺伝子には共通した特徴があり、実はそのほとんどが、0001 (報復力の強いしっぺ返し戦略)、1001 (通常 Pavlov 戦略とよばれるもの)、そして 01=0101 (しっぺ返し戦略) を要素として持つことを発見し、また、長期間生存する戦略がこれらの遺伝子要素の組合せであることを見出した結果を報告する。

2. 囚人のジレンマ

2.1 基本概念

囚人のジレンマは一般的に表 1 の利得表が用いられる。パラメータは R, P, S, T の 4 つで、 $S+T < 2R$, $S < P < R < T$ となるように設定されている。この意味するところは、

- 相手が協力を行うと仮定したとき、自分が協力した場合に得点 R を得る一方、裏切れば得点 T を得る。 T の方が R より大きいので裏切り行動を選択するのが合理的である。

表 1 囚人のジレンマの利得表

Table 1 Payoff table of prisoner's dilemma.

自分, 相手	協力	裏切り
協力	R, R	S, T
裏切り	T, S	P, P

※ $S+T < 2R$, $S < P < R < T$

的である。

- 相手が裏切りを行うと仮定したとき、自分が協力した場合に得点 S を得る一方、裏切れば得点 P を得る。 P の方が S より大きいので裏切り行動を選択するのが合理的である。

つまり、相手の行動にかかわらず自分の最も合理的な選択は裏切り行動になる。当然、相手も合理的な選択を行うならば裏切り行動になり、この場合の両者の利得はいずれも P となってしまい、両方で協力し合ったときの利得 R より小さくなってしまふ。損を承知で裏切り合うほかない、というのがジレンマなのである。

2.2 繰返し囚人のジレンマ

1 回だけの囚人のジレンマにおいてプレイヤーはお互いに裏切り合ってしまう。しかし、終わりを告げずに囚人のジレンマを繰り返し行う繰返し囚人のジレンマ (Iterated Prisoners Dilemma: 以下 IPD) を行うことによって、協力行動を行う戦略でも勝ち残る事例がアクセルロッドの実験によって見出された。また、優秀な戦略は以下に示す 3 つの性質を持っていることが確認された。

- 自分からは裏切らない。
- 相手の裏切りにはすぐに裏切りで反撃する。
- 相手が協力してくれればこちらもすぐに協力する。

最も成績の良い戦略であった「しっぺ返し戦略 (Tit for Tat: 以後 TFT)」は、直前に相手が行った行動を模倣する戦略であり、この 3 つの性質を兼ね備えていた [4]。

TFT を上回る戦略として、パブロフ戦略 (以後 Pavlov) が知られている。Pavlov は、得られた得点が R か T なら前回と同じ行動を行い、 S か P なら異なる行動を行う戦略である。この戦略はノイズ (エラー) を含むような環境 (協力を出す場面で、間違っで裏切りを出してしまうことがありうる環境) において、しっぺ返し戦略を上回ることができる [5]。その様子を表 2 に示す。

時刻 $t = 1$ において、(TFT, TFT) と (Pavlov, Pavlov) の組合せで始まったとする。また、(TFT, TFT) と (Pavlov, Pavlov) のどちらにも、協力関係がなりたっていたとする $\{(TFT, TFT)=(C, C)=(Pavlov, Pavlov)\}$ 。

$t = 2$ において、ノイズ (ϵ で表す) が発生すると、どちらか一方が間違えて D を出してしまう $\{(TFT, TFT)=(C, 'D')=(Pavlov, Pavlov)\}$ 。

$t = 3$ において、TFT は直前の相手の行動を模倣する

表 2 TFT 対と Pavlov 対の行動におけるノイズの影響

Table 2 Effect of noise on the actions of a TFT pair and a Pavlov pair.

TFT	VS	TFT	t	Pavlov	VS	Pavlov
C		C	1	C		C
C		'D'	2	C		'D'
D		C	3	D		D
C		D	4	C		C
D		C	5	C		C

※ ' ' はノイズを表す

ため、一方は D、一方は C を出す。{(TFT, TFT)=(D, C)}。対して、Pavlov は、D を出した側は、利得 T を得たため、D をもう 1 度出す。しかし、もう片方は、利得 S しか得られないため次の行動を C から D に変える {(Pavlov, Pavlov)=(D, D)}。

$t = 4$ において、TFT は、 $t=3$ と同じように相手の行動を模倣する {(TFT, TFT)=(C, D)}。Pavlov は、お互いに D だったので、利得は P になり、前回と異なる行動をする {(Pavlov, Pavlov)=(C, C)}。

$t = 5$ において、TFT は $t=4$ と同じように相手の行動を模倣する {(TFT, TFT)=(D, C)}。Pavlov はお互いに C なので利得 R を獲得し、前回と同じ行動をする。{(Pavlov, Pavlov)=(C, C)}

TFT が (C, D), (D, C) を繰り返すのに対して、Pavlov は (C, C) の関係に戻っている。よって、TFT 側は、ノイズが発生すると利得が $\frac{S+T}{2}$ になるのに対して、Pavlov は R になる。結果、Pavlov 側が高い利得を得る。つまり、ノイズを考慮した環境において、TFT は決して安定な戦略とはいえなくなる。

3. 進化型 IPD モデルにおける戦略の自動進化

3.1 進化型 IPD モデル

Lindgren は戦略を生物の遺伝子に見立てることにより 1 次元のバイナリ文字列で表し、遺伝的アルゴリズムに従って進化させるマルチエージェントモデルを考えた [6]。これは人工生命研究の中で注目され、多くの研究者の興味を引き付けた [9], [10], [11], [12]。生物が単純な生物から現在のよう非常に多様で複雑に進化してきたように、このモデルでは 1 次元配列で表現される戦略は、点変異、複写、分離などの変異によって戦略を進化させることができる。このようにして、他の進化的アルゴリズムと同様に、非常に大きな探索空間の中から環境に適合して高い利得を勝ち取ることのできる戦略を効率良く自動検出できる。

Lindgren が行った先行研究においては、最も単純な戦略からより長い戦略に進化する工程が報告されている。その結果、進化的に安定な戦略 (Evolutionarily Stable Strategy: ESS) として [1 * 1 0 * * * 0 * * * * 001] 構造の戦略が述べられている。

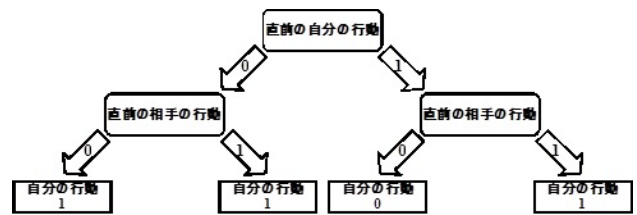


図 1 過去の手と戦略の関係 (m=2 の例)

Fig. 1 A strategy and the history of actions (m=2).

本研究では、このモデルを用いて実験を行った。ただし、Lindgren のモデルでは、繰り返し対戦を行う回数は定数であったが、本研究では、対戦の終了を確率的に決定している点が異なっている。

3.2 戦略の 2 値表現とノイズ

戦略は裏切りを表す '0' と協力を表す '1' の 2 値文字列で構成され、高さ m の 2 分木の葉として表現される。この 2 分木は、0 なら左の子に遷移し、1 なら右の子に遷移する。 m は参照する歴史 (自分と相手の過去の行動) の数であり、 $m=1$ は直前の相手の手のみを考慮する戦略となり、高さ 1 の 2 分木となる。この場合の可能な戦略は、00, 01, 10, 11 の 4 種類のみである。00 は相手の行動によらず 0 を出す。逆に 11 はいつも 1 を出す。01 は直前の相手と同じ手を出す戦略であり、TFT である。10 は直前の相手の手とは反対の手を出す逆しっぺ返し (Anti Tit for Tat: 以後 A-TFT) である。 $m=2$ は直前の自分の手までを考慮する場合に相当する。図 1 に 1101 戦略を例として示した。ただし、今回の実験では m の最大値は 5 としている。さらに、必ずしも戦略どおりに行動するわけではなく、一定確率で行動選択を失敗する。この失敗をノイズと定義する。

3.3 人口動態

戦略は人口という属性を持ち、自身を含めたすべての戦略と総当たりで対戦を行い、獲得した利得の平均値と戦略全体での利得の平均値の差を評価値として、人口を変動させていく。文献 [6] では、対戦回数が定数であったため、 t 世代目の戦略 i の獲得平均利得 S_i は、戦略 i の人口比率を x_i (戦略 i の人口/全体人口)、戦略 i が戦略 j より獲得した利得を $g_{i,j}$ 、 $N(t)$ を t 世代目に存在する戦略の種類の数と定義すると、平均値の算出には以下の式を用いている。

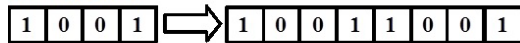
$$S_i(t) = \sum_{j=1}^{N(t)} g_{i,j}(t)x_j(t) \quad (1)$$

しかし、本研究で用いるモデルは、最低 1 回は対戦が行われるが、対戦終了は確率的に決定するため、対戦回数が毎回異なる。よって、最低 1 回は対戦が行われ、対戦終了は確率的に決定する。そのため、戦略 i が戦略 j と対戦した回数を $c_{i,j}$ と定義し、平均値の算出には以下の式を用いる。

点変異



複写変異



分離変異

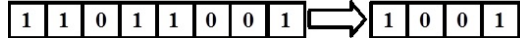


図 2 本稿で考察する 3 種の突然変異 (点変異, 複写変異, 分離変異)
Fig. 2 Three types of mutations used in our model (point mutation, doubling, fission).

$$S_i(t) = \sum_{j=1}^{N(t)} \frac{g_{i,j}(t)x_j(t)}{c_{i,j}(t)} \quad (2)$$

また, 全体での利得の平均 $\bar{S}$ は以下の式によって求められる.

$$\bar{S}(t) = \sum_{i=1}^{N(t)} S_i(t)x_i(t) \quad (3)$$

戦略 i の人口動態は, 式 (2) と式 (3) によって求められた戦略 i の獲得利得 $S_i(t)$ と全体利得平均 $\bar{S}(t)$ から, 以下の式で表される.

$$x_i(t+1) - x_i(t) = \alpha(S_i(t) - \bar{S}(t))x_i(t) \quad (4)$$

α は増加率を表す. 式 (4) より, $\bar{S}$ 以上の利得を得ている戦略は人口を増やし, $\bar{S}$ 未満の戦略は人口を減らすことが分かる. つまり, 人口が戦略の強さの指標となる.

3.4 突然変異

次の 3 つの変異を一定確率で用いて, 新しい戦略を出現させる.

- 点変異・・・戦略の 1 カ所が ‘0’→‘1’ もしくは, ‘1’→‘0’ に反転する.
- 複写変異・・・戦略情報の長さが 2 倍になる.
- 分離変異・・・戦略情報の長さが半分になる.

具体的な例を図 2 に示す. 複写変異時に追加される情報は元の戦略情報のコピーであり, 戦略自体は元の戦略と同一である. また, 分離変異は戦略を 2 分割し, どちらか一方をランダムに選んで新しい戦略としている.

4. 実験

4.1 実験目的と実験条件

3 章で述べた進化型 IPD モデルを用いて, 戦略, 人口分布が時間とともに変化する動的な環境における長寿戦略の分析を行う. また, シミュレーションを行ううえでの実験条件, 初期戦略, 用いた利得表を表 3, 表 4, 表 5 にまとめる. 利得表より, 最大平均値はお互いに協力したときの 3 であり, 最少平均値はお互いに裏切りあったときの 1 である. したがって, 平均利得が 3 に近いほど協力関係が構

表 3 実験条件

Table 3 Experimental condition.

総人口	1000	増加率 α	0.1
対戦終了確率	0.005	ノイズ発生確率	0.01
点変異	2×10^{-5}	複写変異	1×10^{-5}
分離変異	1×10^{-5}		

表 4 初期戦略

Table 4 Initial strategy.

戦略	性質	人口比率
00	すべて裏切り行動	0.25
01	直前の相手の行動と同じ行動	0.25
10	直前の相手の行動とは逆の行動	0.25
11	すべて協力行動	25%

表 5 利得表

Table 5 Payoff table.

自分, 相手	協力	裏切り
協力	3, 3	5, 0
裏切り	0, 5	1, 1

築されているといえる.

4.2 動的環境における囚人のジレンマ

突然変異やノイズの発生確率, 対戦終了を決定する乱数の SEED が異なる 50 回のシミュレーションを行い, そのうち 37 回において, 特定の戦略による 3 すくみの関係性が見られた. 動的な環境において, その他の戦略を差し置いて, 特定の 3 つの戦略が何度も人口を大きく伸ばしていることから, この 3 つの戦略が有用な戦略ではないかと考えられる. そのパターンの一例を図 3 に示す. 図は縦軸が人口数を表し, 横軸が経過世代数を表している.

- (1) 表 4 に示した初期戦略から 01 戦略 (TFT) が人口を大きく伸ばす.
- (2) 01 戦略の代わりに 1001 戦略 (Pavlov) が台頭する.
- (3) 1001 戦略の代わりに 0001 戦略が台頭する.
- (4) 0001 戦略の代わりに 01 戦略が台頭する.
- (5) (2) から (4) を繰り返す.

表 4 に示す初期戦略の中では, 01 戦略が安定して勝つことができるため, 01 戦略が人口を伸ばす. 01 戦略が人口割合の大半を占めているときに, 1001 戦略が出現すると, 表 2 に示すように, 自身どうしの対戦のとき, 01 戦略はノイズによって C (協力) と D (裏切り) を交互に出すようになり, 平均利得が 2.5 となるのに対して, 1001 戦略はノイズが発生してもすぐに協力を出し合う関係に戻ることができる. そのため, 平均利得は 3 のままである. また, 1001 戦略と 01 戦略が対戦を行いノイズが発生しても得られる利得はともに同じである. つまり, 1001 戦略は 01 戦略に勝つことができる. そのため, 01 戦略の次に

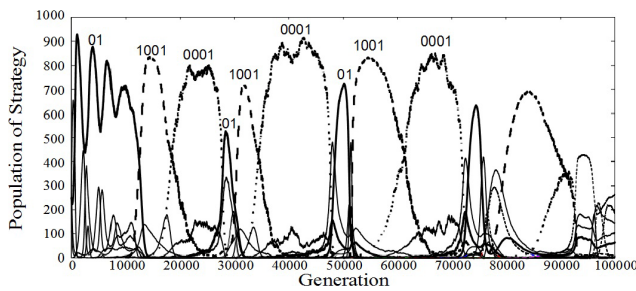


図 3 3すくみ (1001-0001-01) の繰返しパターン

Fig. 3 Triplet pattern (1001-0001-01) repeats in the long run.

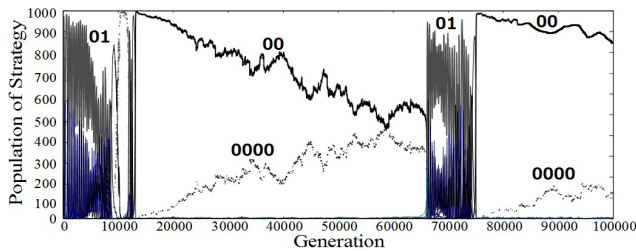


図 4 00 戦略 (ALL-D) の出現による 3すくみの崩壊

Fig. 4 Collapse of the triplet pattern by the emergence of ALL-D strategy.

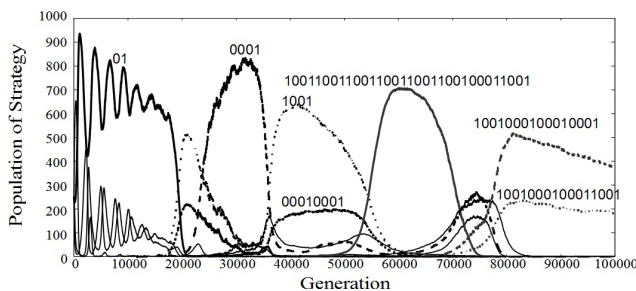


図 5 複雑に進化した戦略の出現による 3すくみの崩壊

Fig. 5 Collapse of the triplet pattern by intricately evolved strategy.

1001 戦略が人口を大きく伸ばしている。しかし、1001 戦略は、表 2 に示したように、相手が裏切ったとしても協力を行ってしまふ。そのため、0001 戦略のように裏切り行動を行いやすい戦略に非常に弱いという特徴を持つ。そのため、0001 戦略が出現すると同時に衰退し始める。0001 戦略は自分または相手が裏切りなら、必ず次手が裏切りになる報復志向の強い TFT 戦略 (Retaliation oriented Tit For Tat: 以後 R-TFT) といえる。そのため、ノイズのある環境では、協力行動を続けることが難しい。逆に、01 戦略は裏切りやすい戦略に対して、得点を取られにくい特徴を持つため、0001 戦略の次に 01 戦略が台頭したと考えられる。

図 4 と図 5 に 3すくみの関係が崩壊した事例を示す。図 4 は 3すくみが発生せず、00 戦略 (ALL-D) が支配的になった事例の 1 つである。1001 戦略は裏切り続ける戦略に弱いため、00 戦略が存在する場合、すぐに淘汰されてしまふ。また、00 戦略が支配的になると、01 戦略であったとしても 00 戦略相手に勝ち残るのは非常に難しい。そ

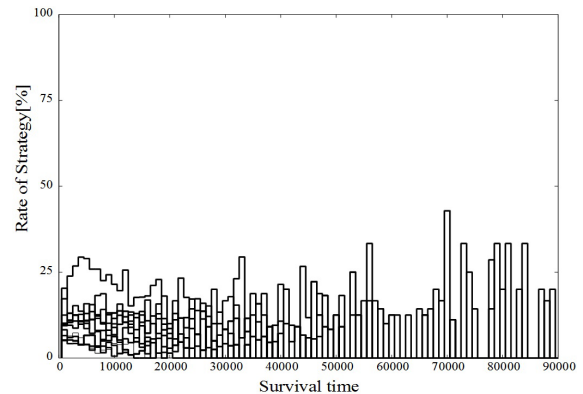


図 6 主要 4 要素 (TFT, Pavlov, R-TFT, ALL-D) を遺伝子に含まない戦略の比率とその戦略の生存期間

Fig. 6 The rate of strategies not containing the important four elements (TFT, Pavlov, R-TFT, ALL-D).

のため、図 4 では、00 戦略が支配的になる 10000 世代から 01 戦略が勝ち上がる 70000 世代まで 00 戦略の支配が続いている。このように、0001 戦略の代わりに 00 戦略が現れると、3すくみは発生しない。図 5 は、より長く複雑に進化した戦略が出現することにより、3すくみの関係が崩壊した事例の 1 つである。これは、長く複雑に進化したことにより、01 戦略や 1001 戦略、0001 戦略に対応できる戦略が出現したためだと考えられる。

4.3 長期間生存する戦略に現れる特徴

長大な遺伝子のうち、比較的寿命の長いものに注目すると、図 5 に見るように、00011001 や 1001000100011001 といった、1001 と 0001 を要素として持つものの人口が増大しており、1001 戦略・0001 戦略を遺伝子に持つ戦略は、動的環境において長期的に勝ち残ることができるのではないかと考えられる。そこで、長期的に勝ち残る戦略が、どのような遺伝情報を含んでいるかの検証を行った。

50 回のシミュレーションで出現した戦略 7,927 個 (重複あり) の中から、戦略長 4 以上の戦略を対象に検証を行う。各戦略を長さ 4 の遺伝子情報で区切り、その遺伝情報が 0000 から 1111 のどのいずれであるかを生存期間別に調査した。特に、3すくみの関係にあった TFT, Pavlov, R-TFT に着目する。その結果を図 6, 図 7, 図 8 に示す。各図の横軸は生存していた期間を表し、縦軸は割合 (各期間の特定遺伝子を含んでいる戦略数/各期間の生存戦略数) を表す。

着目する 4 種類 (TFT, Pavlov, R-TFT, ALL-D) を遺伝子に含まない戦略の生存期間別の割合を図 6 に示す。これによると、全期間を通して 25%前後にとどまることが確認できる。

遺伝子情報 Pavlov を要素として含む戦略の全戦略に占める割合を図 7 の左に示す。主要 4 要素を含まない戦略と比較して、明らかに大きな比率で長寿戦略に Pavlov が含

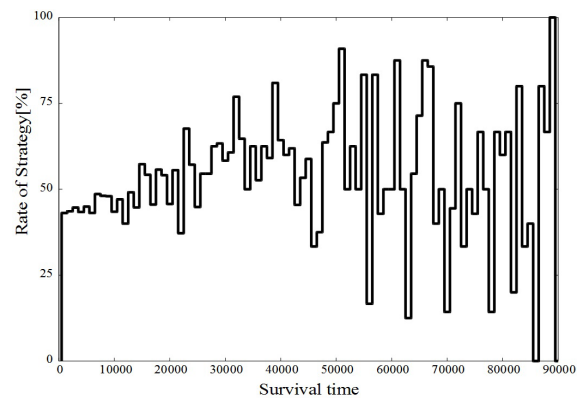
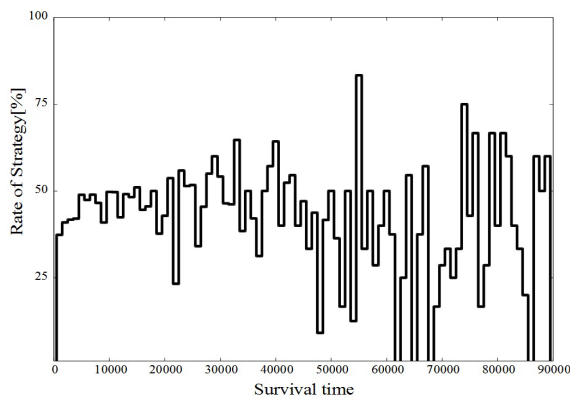


図 7 遺伝子情報 Pavlov (左) と R-TFT (右) を要素として含む戦略の全戦略に占める割合とその生存期間との関係

Fig. 7 The rate of strategies containing Pavlov (left) and R-TFT (right) as a function of their survival times.

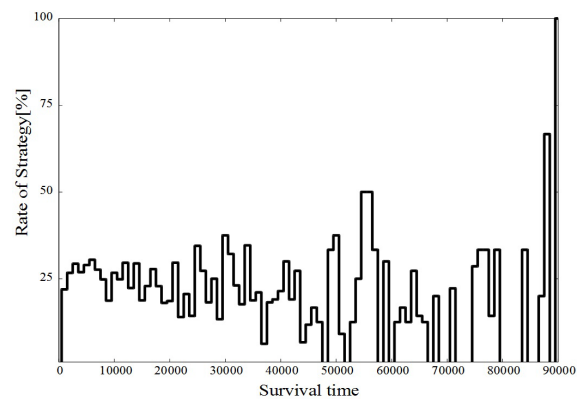
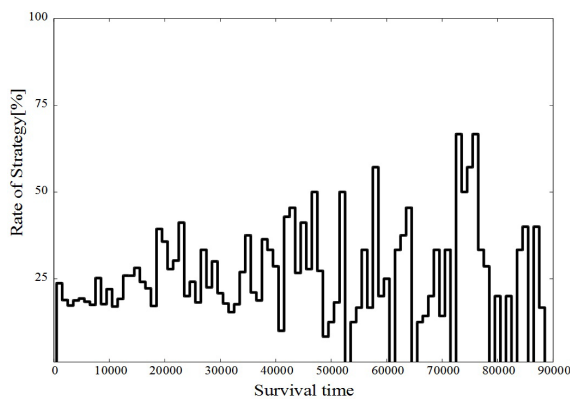


図 8 遺伝子情報 ALL-D (左) と TFT (右) を要素として含む戦略の全戦略に占める割合とその生存期間との関係

Fig. 8 The rate of strategies containing ALL-D (left) and TFT (right) as a function of their survival times.

まれることが見て取れる。また、図 7 の右の図の R-TFT を遺伝子情報に含む戦略数の割合、図 8 の ALL-D を遺伝子情報に含む戦略数の割合と TFT を遺伝子情報に含む戦略数の割合についても、図 6 に比べ非常に大きい値を示していることが確認できる。特に Pavlov, R-TFT の 2 つの戦略が、他の戦略に比べて高い値であった。このことから、長期間生存する戦略ほど、Pavlov, R-TFT を含んでいると考えられる。

50 回のシミュレーションで出現した戦略のうち、6 万世代以上生き残った戦略は 54 種類存在した。その中で、戦略を 4 つで区切ったとき、左端が Pavlov であった戦略 ([1001 ****] や [1001 **** **** **** ****] など) の数は 39 種類存在し、右端が R-TFT である戦略 ([**** 0001] や [**** **** **** **** 0001] など) は 38 種類存在した。また、この 2 つの条件を満たす戦略は 32 種類 ([1001 **** **** 0001] など) であった。表 6 にその一例を示す。

4.4 実験結果および考察

動的環境での囚人のジレンマの結果、図 3 に示すよう

表 6 長期生存戦略の例

Table 6 Examples surviving strategies.

戦略	生存期間
0101	89364
1001 0001 0001 0001	88546
0001 0001	88504
1101 0001 0001 0001	88096
1001 0001 0001 0001	88055
1001 0001 0001 0001 1001 0001 0001 0001	88016
1001 0001 0000 0001	87539

に、強い戦略が次々と移り変わる様子が確認できる。TFT は 2 章で述べたように、3 つの優秀な戦略である条件を兼ね備えた戦略である。Pavlov は TFT よりノイズに強い性質を持ち、ノイズが存在する環境での Pavlov どちらの対戦の場合、TFT よりも高い利得を得ることができる。ただし、Pavlov は裏切り行動主体の戦略にはきわめて弱いという特徴を持っているため、R-TFT のような、裏切りやすい戦略には勝つことができない。しかし、R-TFT は、TFT からは高い利得が得られない。このことから、

TFT>Pavlov>R-TFT>TFT>... という 3 すくみの関係性が構築されていることが分かる. 図 3 の結果は, この 3 すくみの関係がもたらした結果といえる.

長期的に生存する戦略の 59%が, [1001 **** **** 0001] の遺伝子構造を持つことが確認できた. また, 図 6 から図 7 より, 長期的に生存する戦略の遺伝子には, Pavlov や R-TFT が多いことも確認できる. このことから, 長期的に生存する戦略の遺伝子配列は, [1001 0*0* 0*0* 0001] であると考えられる. これは, Lindgren が発見した [1001 0*** 0*** *001] にも符合する結果である [6]. [1001 0*0* 0*0* 0001] を持つ戦略の特性として, 次の 3 つがあげられる.

1. 自分からは裏切らない.
2. 裏切られたらすぐに報復する.
3. 裏切りが続くと自分から協力行動を行う.

1, 2 の特徴はアクセルロッドの実験と同じ結果である [4]. また, 3 の特徴は Pavlov と同じ特徴である. しかし, [1001 0*0* 0*0* 0001] を遺伝子配列に持つ戦略は, こちらから協力を出す間隔が Pavlov に比べて長い. そのため, Pavlov よりも裏切り戦略に点を取られにくい. また, 2 の特徴は [1001 0*** 0*** *001] では確認できない特徴である. これらの特徴から, [1001 0*0* 0*0* 0001] を遺伝子配列に持つ戦略は, ノイズ付きの動的環境における繰返し囚人のジレンマにおいて長寿戦略になったと考えられる.

5. おわりに

戦略および人口分布が刻々と変化する動的な環境において, 戦略がどのように変化していくのかを, IPD を用いたシミュレーションを用いて分析を行った. その結果, アクセルロッドの実験において優秀な成績を残した TFT, ノイズに強い Pavlov, 報復傾向の強い TFT である R-TFT の 3 つの戦略が他の戦略にはない動きを見せた. また, より複雑に進化した長期生存戦略においても, これらの戦略は遺伝子情報として残っていることが確認でき, また, その組合せは, [1001 0*0* 0*0* 0001] であった. これらの結果から, 動的な環境において有用な戦略は, TFT, Pavlov, R-TFT で構成された, [1001 0*0* 0*0* 0001] の遺伝子構造を持つ戦略でないかと考えられる.

本稿での結果は定性的な比較によって得られたことが多いため, 定量的知見を得るためにはさらに多量のシミュレーションを重ねる必要があると考えられる

参考文献

- [1] Novak, M.A. and Sigmund, K.: Evolution of indirect reciprocity by image scoring, *Nature*, Vol.393, pp.573-576 (1998).
- [2] Roberts, G. and Sherratt, T.N.: Development of cooperative relationship through increasing investment, *Nature*, Vol.394, pp.175-178 (1998).
- [3] Yao, X. and Darwen, P.: How important is your requi-

tation in a multi-agent environment, *IEEE-SMC1999*, pp.575-580 (1999).

- [4] Axelrod, R.: *The Evolution of Cooperation* (1984). 松田耕治 (訳): つきあい方の科学, ミネルヴァ書房 (1998).
- [5] Nowak, M.A. and Sigmund, K.: A strategy of win-stay, lose-shift that outperforms tit-for-tat in Prisoner's Dilemmagame, *Nature*, Vol.364, pp.56-58 (1993).
- [6] Lindgren, K.: Evolutionary Phenomena in Simple Dynamics, *Artificial Life II*, pp.295-312, Addison-Wesley (1990).
- [7] Tanaka-Yamawaki, M. and Murakami, T.: Effect of reputation on theformation of cooperative network of prisoners, *New Advances in Intelligent Decision Technologies, SCI199* pp.615-623, Springer (2009).
- [8] 糸井良太, 田中美栄子: 進化型 IPD における共存社会の形成情報処理学会研究報告, Vol.2011-MPS-83, No.2 (2011).
- [9] 佐々木貴宏, 所真理雄: 進化的エージェント集団の動的環境への適応, 一般社団法人日本ソフトウェア科学会, コンピュータソフトウェア, Vol.14, No.4, pp.365-378 (1997).
- [10] 鈴木麗聖, 有田隆也: 囚人のジレンマゲームにおける Baldwin 効果, 人工知能学会第 13 回全国大会論文集, pp.277-278 (1999).
- [11] 田中美栄子, 中武耕治: 囚人のジレンマ型問題における戦略の進化, 宮崎大学工学部紀要, Vol.30, pp.319-326 (2001).
- [12] 根路銘もえ子, 遠藤聡志, 山田孝治, 宮城隼夫: 繰返し囚人のジレンマゲームにおける競合共進化戦略の解析に関する考察, 電子情報通信学会技術研究報告, CST, コンカレント工学, Vol.99, No.418, pp.65-70 (1999).



糸井 良太 (学生会員)

1989 年生. 2007 年鳥取大学工学部知能情報工学科入学. 2011 年鳥取大学大学院情報エレクトロニクス専攻修士前期課程入学.



田中 美栄子 (正会員)

1950 年生. 1974 年京都大学理学部卒業, 1979 年名古屋大学大学院満期退学, 1983 年 Rochester 大学博士課程修了 (Ph.D. in Physics). CCNY, SUNY, NASC, 相山女学園大学, 宮崎大学工学部を経て, 現在, 鳥取大学大学院工学研究科情報エレクトロニクス専攻知能情報工学講座教授. 主たる研究テーマは経済物理学, 複雑系科学. 日本物理学会, IEEE, 応用数理学会各会員.