

# 論文紹介 進化型繰り返し囚人のジレンマにおける 最適戦略の探究

糸井良太, 田中美栄子  
情報処理学会研究報告, Vol. 88, No. 3, 2012.

川口 晏璃

April 28, 2023

## 背景

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ようという試みは、Lindgren によって導入された遺伝子列の 2 倍化と 2 分裂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発展した。

## 目的

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もつ特徴を調べることを目的とした。

## 基本概念

- ▶ 相手（協力）：自分（協力）得点  $R$ ,（裏切る）得点  $T \rightarrow$  裏切る選択をする方が合理的
- ▶ 相手（裏切る）：自分（協力）得点  $S$ ,（裏切る）得点  $P \rightarrow$  裏切る選択をする方が合理的

つまり、相手の行動に関わらず自分の最も合理的な選択は裏切り行動になる。相手も合理的な選択は裏切り行動となり、どちらも裏切った場合の利得  $P$  となり、両者協力した場合の利得  $R$  より小さくなってしまう。損を承知で裏切り合うほかない、というのがジレンマである。

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

※  $S+T < 2R, S < P < T < R$

図 1: 囚人のジレンマの利得

# 繰り返し囚人のジレンマ

4/11

終わりを告げずに囚人のジレンマを繰り返し行う繰り返し囚人のジレンマ（Iterated Prisoners Dilemma：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ことがアクセルロッドの実験によって証明された。また優秀な戦略には3つの性質をも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ば自分もすぐに協力する

しっぺ返し戦略はこの3つの性質を兼ね備えていた。

## Lindgren モデル

戦略を進化させることができ、複数の戦略を同時に使用してシミュレーションを行う場合、このモデルを使用することで無限にある戦略から検索領域を絞ってシミュレーションを行うことができる。

## 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。高さ  $m$  の 2 分木の葉として表現。  $m$  は参照する歴史（過去の自分と相手の行動）の数。  $m = 1$  は直前の相手の手のみを考慮する戦略。  $m = 2$  は直前の自分の手までを考慮する場合に相当する。

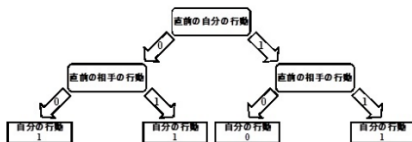


図 2: 過去の手と戦略の関係 ( $m = 2$  の例)

戦略は人口という属性をもち、自身を含めた全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。

本研究で用いるモデルは、対戦終了を確率的に決定しているため、対戦回数が毎回異なる。よって、戦略  $i$  が戦略  $j$  と対戦した回数を  $c_{i,j}$  と定義すると、平均値の算出は、以下の式を用いる。

$$S_i(t) = \sum_j^N \frac{g_{i,j}(t)x_i(t)}{c_{i,j}(t)} \quad (1)$$

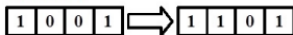
また、戦略  $i$  の人口比率  $X_i$ 、戦略  $i$  の獲得平均利得  $S_i$ 、全体の平均利得  $\bar{S}$  としたときの人口の変化は以下の式で表される。

$$X_i(t+1) - X_i(t) = \alpha(S_i(t) - \bar{S}(t))X_i \quad (2)$$

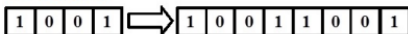
$\alpha$  は増加率を表す。人口が戦略の強さの指標となる。

- 1 点変異：戦略の1箇所が反転する
- 2 複写変異：戦略情報の長さが2倍になる（文字列そのままコピー）
- 3 分離変異：戦略情報の長さが半分になる（2分割のしたどちらか一方をランダムで選ぶ）

## 点変異



## 複写変異



## 分離変異

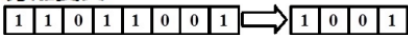


図 3: 3 種の突然変異

シミュレーションを行う上での実験条件，初期戦略，用いた利得表は以下の通りである．

総人口	1000	増減率	10
対戦終了確率	0.5%	ノイズ発生確率	1%
点変異	$2 \times 10^{-3}\%$	複写変異	$1 \times 10^{-4}\%$
分離変異	$1 \times 10^{-4}\%$		

図 4: 実験条件

00	全て裏切り行動
01	直前の相手の行動と同じ行動
10	直前の相手の行動とは逆の行動
11	全て協力行動

図 5: 初期戦略

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

図 6: 利得表



シミュレーションの結果，3つの戦略が規則正しいパターンで出現する結果が得られた。

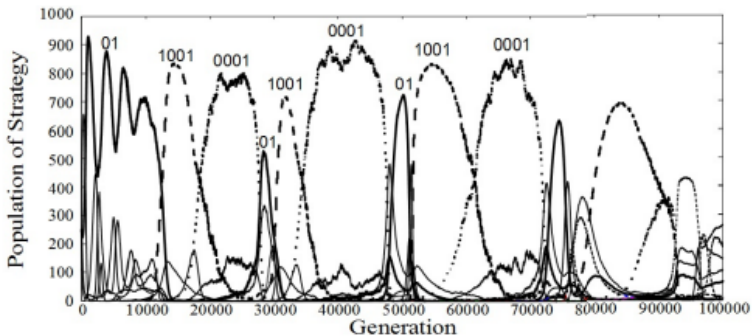


図 7: シミュレーション結果

1001 戦略, 0001 戦略の 2 つの戦略は長く生き残る戦略の遺伝子情報として頻繁に表れる.

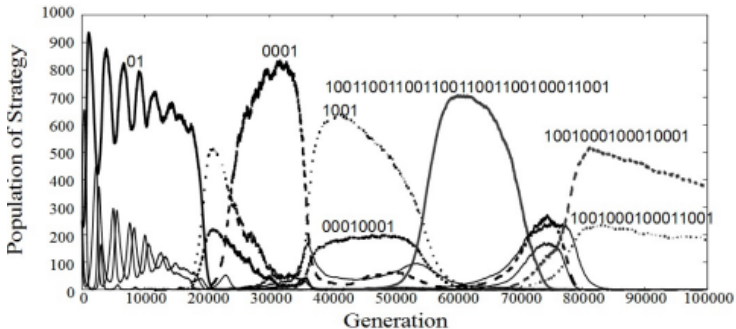


図 8: 長く生き残る戦略

- 1 01 戦略：3つの条件を兼ね備えた戦略（しっぺ返し戦略）
- 2 1001 戦略：01 戦略よりもノイズに強い性質．ただし，裏切り自体に弱いため，裏切りやすい戦略には勝てない
- 3 0001 戦略：裏切りやすい戦略．ただし 01 戦略に弱い．

このことから，01 戦略 > 1001 戦略 > 0001 戦略 > 01 戦略 > ... という3棘みの関係性が構築されていることが分かる．また，長期間生存している戦略は，01, 1001, 0001, 0000 の4つで構成されている確認できた．

よって，01 戦略，1001 戦略，0001 戦略の3つの戦略が動的環境における有用な戦略かと考えられる．

## 今後の課題

遺伝子情報の組み合わせによる戦略の強さの変化や，初期パラメータによる変化などの実験を行い，精密な最適戦略の分析を行っていく．