

# 囚人のジレンマゲームにおける Baldwin 効果

## The Baldwin Effect in the Iterated Prisoner's Dilemma

鈴木 麗璽 有田 隆也

Reiji SUZUKI Takaya ARITA

名古屋大学 大学院人間情報学研究科

Graduate School of Human Informatics, Nagoya University

The Baldwin effect is known as an interaction between learning and evolution, which suggests that individual lifetime learning can influence the course of evolution without Lamarckian mechanism. Since Hinton and Nowlan showed this by computer simulation, many studies have been made in the static environment. Our concern is to consider the Baldwin effect in dynamic environment, especially when there is no explicit optimal solution through generations and it only depends on interactions among agents. We adopt the iterated Prisoner's Dilemma as a dynamic environment and introduce phenotypic plasticity to strategies by using a meta-learning rule termed "Meta-Pavlov". In this simulation, the Baldwin effect was observed as follows: First, strategies with enough plasticity spread, which caused a shift from defective population to cooperative. Second, these strategies were replaced by the strategy [x00x] which has a modest amount of plasticity.

### 1. はじめに

進化と学習の相互作用に関する重要なトピックに、Baldwin 効果がある。我々は、Baldwin 効果が、最適解が決まらない動的な環境においてどのように働くかについて、表現型の可塑性を導入したくり返し囚人のジレンマゲームにおける戦略の進化実験を用いて解析している。進化実験の結果、Baldwin 効果が観察され、可塑性を利用した安定な戦略が得られたことを示す。

### 2. Baldwin 効果

Baldwin 効果は進化と学習が相互に与える影響を、学習のメリットとコストのバランスから説明するものである。一般には以下の2つの段階を経て、学習により獲得されていた形質が次第に生得的な形質へと進化していくものとされている。

第1段階：学習により生存上有利な形質を獲得した個体が次世代に多く子孫を残す。

第2段階：すべての個体が生存上有利な形質を学習に

より獲得した集団では、学習にかかるコストのためその形質を生得的に獲得している個体が次世代に多く子孫を残す。

これまで、Baldwin 効果に関する進化実験による研究は Hinton と Nowlan による先駆的実験[1]をはじめとして、最適解が固定されたものがほとんどであり、動的な環境において、Baldwin 効果がどのように働くかは未解明であった。

### 3. 進化実験

そこで、最適解が決まらず、特に個体間の相互作用のみを考慮した動的環境として、くり返し囚人のジレンマゲームの進化実験を取り上げ、戦略に学習（表現型の可塑性）を導入することで動的環境における進化と学習の相互作用について解析する。

#### 3.1 くり返し囚人のジレンマゲーム

くり返し囚人のジレンマゲームは、社会的集団に生じるジレンマ状態をシンプルに抽象化したモデルであり、様々な分野で多くの研究がなされているものである。本稿では、以下の利得行列を用いる。

表 1：囚人のジレンマゲームの利得行列

相手の手 ( )	協調	裏切り
自分の手 ( )	(C)	(D)
協調 (C)	(3, 3)	(0, 5)
裏切り (D)	(5, 0)	(1, 1)

(自分の得点, 相手の得点)

### 3.2 戦略の遺伝子表現

各個体の持つ戦略を戦略遺伝子列 GS と学習遺伝子列 GL の 2 つの遺伝子列の組で表現する。

戦略遺伝子列は Lindgren[2]のモデルと同様な、履歴に依存して次回の手を決定する戦略を定義する。

記憶長  $m$  の戦略は裏切りを 0、協調を 1 として以下のような 2 進数で表された履歴  $h_m$  を持つ。

$$h_m = (a_{m-1}, \dots, a_1, a_0)_2 \quad (1)$$

ここで  $a_0$  は前回の相手の手、 $a_1$  は前回の自分の手、 $a_2$  は前々回の相手の手...とする。

ある履歴  $k$  に対応して次回出すべき手を  $A_k$  (0 または 1) とすると、記憶長  $m$  の戦略は、

$$GS = [A_0, A_1, \dots, A_{n-1}] \quad (n=2^m) \quad (2)$$

と表すことができる。これを戦略遺伝子列とする。

さらに、各  $A_x$  に対してその表現型 (協調または裏切り) が可塑性を持つかどうかを  $L_x$  (0: 可塑性を持たない、1: 可塑性を持つ) として、学習遺伝子列を

$$GL = [L_0, L_1, \dots, L_{n-1}] \quad (3)$$

と定義する。

可塑性を持つ表現型は、対戦中にその表現型を用いた結果に応じて、メタパブロフ学習行列 (表 2) に基づいて表現型を変更 (学習) する。

表 2：メタパブロフ学習行列

相手の手 ( )	C	D
自分の手 ( )	C	D
C	C	D
D	D	C

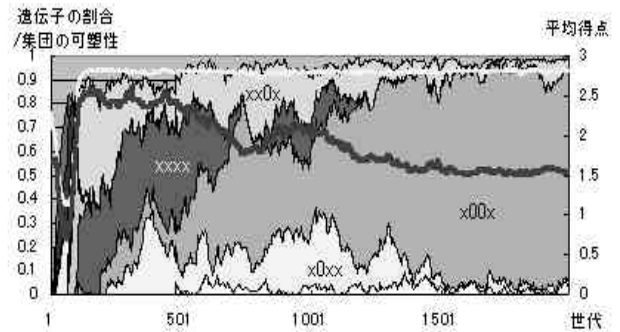
以下、各戦略を、可塑性を持つ学習遺伝子に対応する戦略遺伝子を  $x$  と置き換えた戦略遺伝子列でまとめて表現する (e.g.  $GS = [1000]$ ,  $GL = [1001]$  [x00x])。

### 3.3 遺伝的オペレーション

以上のような戦略を持つ集団について、ノイズありのくり返し囚人のジレンマゲームの総当たり戦を行い、その合計得点を各個体の適応度としたルーレット選択により次世代の集団を作り出す操作をくり返し、集団を進化させる。その際、遺伝子の一点突然変異を導入する。

## 4. 実験

記憶長 2 (初期集団はランダム) の集団において、パラメータとして突然変異率 0.001、個体数 1000、ノイズ率 2% 未来係数 0.99、世代数 2000 を用いて進化実験を行った。試行の約 75% で図 1 のような傾向を持つ進化が確認された。



(平均得点：白実線、集団の可塑性：黒実線)

図 1：記憶長 2 での実験結果

はじめに裏切りの戦略により平均得点が低下した後、集団の可塑性 (学習遺伝子列中に占める "1" のビットの割合) の増加とともに、高い平均得点を持つ集団へと進化した。これは、可塑性が裏切りの集団から協調的な集団へのシフトに有利な方向へと働いたことを示し、Baldwin 効果の第 1 段階と考えられる。

その後、高い平均得点を維持したまま、集団の可塑性は次第に低下し約 50% のところで安定し、集団の大部分を [x00x] 型の個体が占める結果となった。これは、十分協調関係が築かれた集団においては、ノイズによって可塑性がコストとして働くため、集団を維持するために最低限必要な可塑性のみが選択されたことを示しており、Baldwin 効果の第 2 段階と考えられる。

このような動的な環境においては、集団が安定するために最低限の可塑性を保つ必要があり、Baldwin 効果はそのバランスを探索する過程であると捉えられる。

## 5. おわりに

本稿では囚人のジレンマゲームの戦略に可塑性を導入することで、戦略集団の進化に生じる Baldwin 効果を示した。メタパブロフ型 [x00x] 戦略の解析[3]、記憶長や学習方式を固定しないオープンエンドな進化実験などを現在行っている。

## 参考文献

- [1] G.E. Hinton and S.J. Nowlan: How Learning Can Guide Evolution, Complex Systems, Vol.1(1987), pp.495-502.
- [2] K. Lindgren: Evolutionary Phenomena in Simple Dynamics, Artificial Life II(1991), pp.295-311.
- [3] 鈴木 麗壘, 有田 隆也: メタ・パブロフ: 進化と学習による適応を自動調節する囚人のジレンマ戦略, 第 1 回ゲーム情報学研究会(発表予定), 情報処理学会 (1999).