

## 学習と進化の相互作用を用いたエージェントの行動ルール生成

黒木麻理 乾伸雄 小谷善行  
東京農工大学

{mary.k, nobu, kotani}@fairy.ei.tuat.ac.jp

### 概要

社会的関係について研究するときエージェントベースモデルを利用して社会的関係を分析する。エージェントは自身の学習や集団の進化によってさまざまな関係を築いていく。学習と進化の相互作用に関する話題として Baldwin 効果がある。これは集団における学習が進化に影響を与えるというものである。本稿では、簡単な捕食モデルにおいて、エージェントがよりよい行動をするように行動ルールを学習、進化させ、どのような行動が生成されるか調べる。エージェントは周りの4つの環境情報と乱数を加えた5つの要素について行動ルールを適用し移動方向を決定する。学習によって多くのエサがとれるように、行動ルールにおける5つの要素の重みを更新する。また、エージェントに進化を取り入れ学習の要素を変更する。実験の結果、学習と進化の相互作用によって、両方とも取り入れた集団がより多くのエサをとることができた。また、学習・進化の補完の様子もみられた。

### Action Rule Generation of Agents Using Interaction of Learning and Evolution

KUROKI Mari INUI Nobuo KOTANI Yoshiyuki  
Tokyo University of Agriculture and Technology

### Abstract

When we study social relations, we analyze them using an agent base model. Agents build various social relations by learning or evolution. The Baldwin effect is subject about the interaction of learning and evolution. It explains influence between leaning and evolution. In this paper, we analyze what action rules agents learn and evolve to do better actions in the easy predation model. Agents make their moves from action rule and five elements (four environment information and random numbers). They change five elements of action rule to get more foods. And agents change action rule by evolution, too. Agents can get more food using interaction of learning and evolution than which using only learning or evolution. And we observe that learning and evolution make up each other.

### 1. はじめに

近年、マルチエージェントを用いた研究が多く取り上げられており、人工知能やゲーム理論の分野では積極的に研究が進められている。また、社会科学の分野でもエージェントベースモデルを用いてシミュレーションを行うというアプローチも取られている。代表的な例として Sugerscape モデル[1]が挙げられる。これはエージェントが消費する食料が空間に分布されており、さまざまな局所ルールを定めることで食料をめぐる社会的行動をボトムアップで構成する。このようにエージェントモデルでは、エージェントや空間の挙動についてルールを定めることで各エージェントが個々の判断によって行動し、その結果ルールによっては協調的な行動や社会的な行動を発生させることが可能である。

エージェントは外部からの環境を取り入れ、自身の行動ルールに基づいて行動を決定する。このルールは常に一定ではなく環境に適応するように変化する。ルールの変化の方法として、学習と進化がある。前者はエージェントの一生を通じた個体レベルでの適応であり、後者は世代を通じた集団レベルでの適応である。エージェントの行動ルールはこの二つが作用して変化すると考えられる。

本稿では、簡単なルールを定めたエージェントモデルを用いて、エージェントの学習と進化の相互作用について実験する。個体の学習と集団の進化は別々の次元における環境の適応方法であるが、学習と進化のバランスによって相互作用を起こす可能性もある。学習が進化にもたらす影響や進化による学習の補完について検証する。

## 2. Baldwin 効果

学習と進化の相互作用に関する理論の一つに Baldwin 効果[2]がある。これは学習により行動が変化することで新しい淘汰圧が生まれ、遺伝子が進化していくことをいう。現在の一般的な定義では次の2つの段階に分けて考えられる[4]。

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：十分多くの個体が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もともと持っている個体が次世代に多く子孫を残す。

第1段階は学習のメリットが選択圧となる場合、第2段階は学習のコストが選択圧となる場合である。学習のメリットとコストのバランスによって Baldwin 効果が起こると考えられており、集団における個体の学習が集団の進化に方向性を与える。

## 3. エージェントモデルの概要

実験は文献[3]で用いたモデルを改良して行った。エージェントの機能を単純化し、各エージェントは学習や進化によって効率のよい行動を生成する。このモデルは仮想空間上にエージェントとエサが複数配置された捕食モデルとなっている。仮想空間は  $N \times N$  マスのトーラス状であり、それぞれのマスにはエージェントかエサが1つしか入れない。

### 3.1 エージェントの定義

エージェントは仮想空間上を移動しエサをとる。エージェントは次のような機能を持つ。

- ・毎周期に上下左右どちらかの空マスに1マス移動する。
- ・体力を持ち毎周期に1ずつ減少する。体力が0になると移動することはできない。
- ・視野を持ち各方向（上下左右）の視野の範囲内にあるエージェントとエサを認識する。獲得する環境情報は、視野の範囲内で一番近いエサおよびエージェントの距離とエサおよびエージェントの数である。例えば、視野3を持つエージェントAの上方向の視野は図1の太線の範囲となる。この場合、対象○は認識できるが対象●は認識できない。

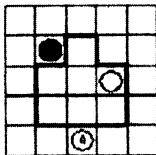


図1 エージェントの視野

エージェントの移動方向は、方向ごとに認識した環境情報から評価値を計算し最も大きい方向へ移動する。評価値は次の評価式を用いて計算する。

$v$  方向の評価値

$$\begin{aligned}
 &= W_{fd} \times \text{一番近いエサまでの距離} \\
 &+ W_{ad} \times \text{一番近い他エージェントまでの距離} \\
 &+ W_{fn} \times \text{エサの数} \\
 &+ W_{an} \times \text{他エージェントの数} \\
 &+ W_{rd} \times \text{乱数} \\
 &v \in \{ \text{上, 下, 左, 右} \}
 \end{aligned}$$

また、各エージェントは初期値遺伝子 GI を持つ。上記の評価式に用いる重みの初期値を遺伝子として定める。

$$GI = [ iw_{fd} \quad iw_{ad} \quad iw_{fn} \quad iw_{an} \quad iw_{rd} ]$$

例えば、 $w_{ad}$  の初期値が 1.0、その他の重みの初期値が 0.0 のエージェントの場合は  $GI = [ 1.0 \ 0.0 \ 0.0 \ 0.0 \ 0.0 ]$  となる。

### 3.2 エサの定義

エサは仮想空間上に一様に分布している。エサは仮想空間を移動することはできない。エージェントにとられるとその場からなくなる。エサは毎周期にエサが発生する確率（エサ発生確率と呼ぶ）にしたがって空のマスに発生する。

### 3.3 エサの取得

空間上にあるエサはそのエサの4近傍にいるエージェントがとることができる。図2ではエサBに対して太線の範囲にいるエージェント（○印）が対象になり、その他のエージェント（●印）は対象にはならない。エサをとるためには対象のエージェントの合計体力が、取得条件体力以上でなければならない。エサをとるとエージェントの体力は、取得による増加分を周囲のエージェントの数で割った値だけ増加する。

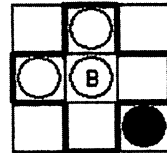


図2 エサの取得

## 4. エージェントの学習と進化

エージェントはエサをより多くとるように個体ごとに学習を行う。また、エージェント集団を更新するため一定周期に遺伝的な進化を行う。

#### 4. 1 行動ルールの学習

エージェントが一定周期でより多くのエサをとるように移動方向を決定するエージェントの評価式を更新する。更新の流れとしては次のようになる。

- (1) 実験モデルを 1 世代実行する。
- (2) もし今世代の実行でのエサ取得数が前世代の実行より少ない場合、重みを前世代の実行の値に戻す。
- (3) 重みからランダムに 1 つ選択し 0.1 増加または減少する。
- (4) (1)に戻る

エージェントの前世代のエサ取得数と比較して、今世代のエサ取得数の方が多い場合は評価式をそのまま更新する。少ない場合は前世代の重みに戻して異なる重みを更新する。

#### 4. 2 エージェントの進化

エージェント集団を更新するためにエージェント全体を対象に遺伝的な進化を行う。進化はエージェントの死亡と発生の 2 つで構成される。

エージェントの死亡条件は次の二種類がある。

- ・一定年齢以上の場合（寿命による条件）
  - ・エサ取得数が低い場合（適応度による条件）
- 条件に当てはまるエージェントからそれぞれランダムに選択し発生するエージェントと置き換える。

新しいエージェントはエージェント全体から両親をルーレット選択によって 2 体選択して発生する。初期値遺伝子は両親からランダムに選択して作成する。また、発生したエージェントは一定の確率で突然変異が起こる。突然変異が起こった場合、初期値遺伝子の値がランダムに 1 つ変化する。

#### 5. 実験

捕食モデルの環境は次のように設定する。

- ・仮想空間の大きさは 50×50 マス。
- ・エージェントの個体数は 50、初期体力は 150、視野は 15、初期値遺伝子はすべて 0.0。
- ・エサの初期個体数は 1000、エサ取得の体力条件は 200 以上、取得による体力の増加は 200。
- ・300 周期で 1 つの実行とし（世代と呼ぶ）、行動ルールを更新する。
- ・集団の進化は 20 世代ごとに行う。死亡の年齢は 50 世代以上のエージェントをランダムに 5 体、エサ取得数の下位 5 体の計 10 体を置き換える。
- ・突然変異は 5% で起こり、初期値遺伝子は 0.5 増加または減少する。
- ・実行は 1500 世代行う。
- ・エサ発生確率は次のように変化する（図 3）。

- 0~1000 世代：領域 a は 1/500、領域 b は 1/2500
- 1001~1500 世代：領域 a、b ともに 1/2500
- 1501~2000 世代：領域 a は 1/500、領域 b は 1/2500

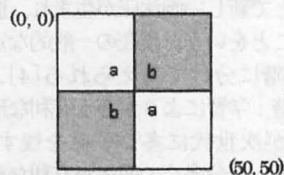


図 3 仮想空間でのエサの発生分布

実験の結果は次のようになった。はじめに全体のエサ取得数を図 4.1、エージェントの平均年齢（何世代生きたか）を図 4.2 に示す

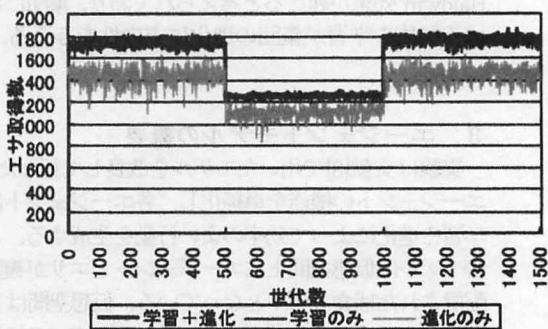


図 4.1 全体のエサ取得数

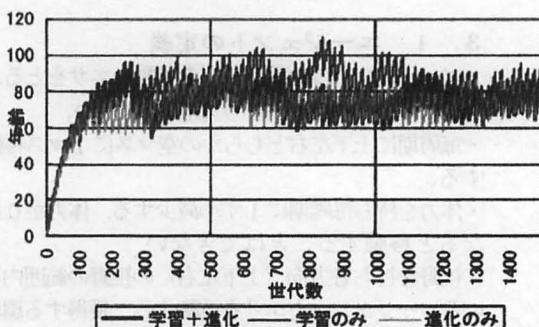


図 4.2 エージェントの平均年齢

全体のエサ取得数についてみると、進化のみの場合より、学習を行う場合のほうがエサ取得数が多い。よって、エージェントが学習を行うことでエサをより多くとることができる。また、エージェントの平均年齢をみると、学習だけの場合より進化も行っている場合のほうが平均年齢は低い。よって、エージェントが進化することで若い世代のエージェントでもエサがとれるようになる。

また、エージェントごとのエサ取得数を図5に示す。

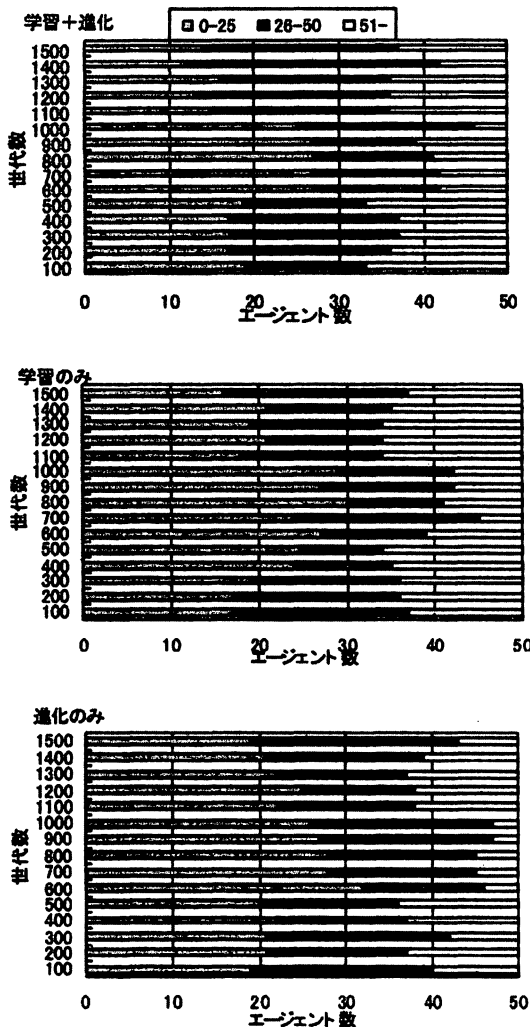


図5 各エージェントのエサ取得数の変化  
(上：学習+進化，中：学習のみ，下：進化のみ)

各エージェントのエサ取得数をみると、学習と進化を合わせた場合が片方ずつの場合と比べてエサの取得が少ないエージェントの数が少なくなっている。これは古い世代では学習が進み、若い世代では進化の影響でエサがとれるようになったためと考えられる。学習のみの場合においてエサの取得が多いエージェントの数が多のは、学習が進んだ古い世代のエージェントが多いためである。

次にエージェントが更新した重みの平均の変化を図6に示す。

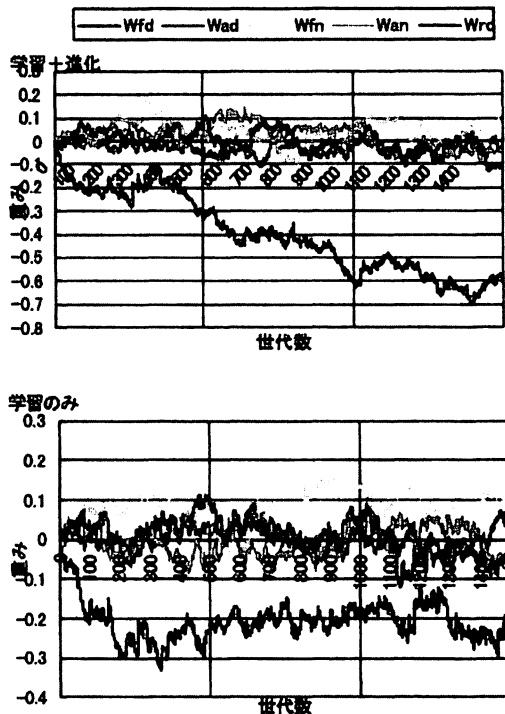


図6 平均の重みの変化  
(上：学習+進化，下：学習のみ)

重みの平均は学習と進化の場合のグラフでは、進化の影響で  $w_{fn}$  が大きく負に向かっていく。また、エサ発生確率が変化する 501~1000 世代の間において  $w_{an}$  の値が正になっている。これはエサの距離だけでなく、他エージェントの数も行動決定に取り入れることを示している。逆に、 $w_{fn}$  の値は負になっており、エサが発生しにくい環境ではエサの数が少ないので大きく関わらないことになる。学習の傾向については図6の2つのグラフから  $w_{fn}$  が負になる傾向があることが分かる。

エージェント集団の進化によって変化した GI について、500 世代ごとの GI の組み合わせとそれぞれのエージェント数は表 1 のようになる。

表 1 500 世代ごとの GI 別エージェント数  
(上：学習+進化，下：進化のみ)

	500 世代後	1000 世代後	1500 世代後
<b>学習+進化</b>			
[0,0,0,0,0]	38	3	0
[-0.5,0,0,0,0]	6	39	37
[0,0,0,0.5,0]	5	0	0
[-0.5,0,0,0.5,0]	1	4	0
[-0.5,0,0,-0.5,0]	0	0	8
その他	0	2	5

	500 世代後	1000 世代後	1500 世代後
<b>進化のみ</b>			
[0,0,0,0,0]	30	22	5
[0,-1,0,0,0]	6	0	0
[0,0,0,0,0.5]	6	13	4
[0,0,0,-0.5,0]	3	4	0
[-0.5,0,0,0,0]	0	0	15
[-0.5,0,0,0.5,0]	0	0	6
[-0.5,0,0,-0.5,0]	0	0	4
その他	5	11	16

500 世代後ではどちらの場合も  $GI=[0,0,0,0,0]$  となるものが多く、多くのエージェントにおいて進化が十分進んでいない状態であったが、1000 世代後、1500 世代後と世代が進むと数種類の遺伝子に分かれるようになった。特に学習と進化の場合、1500 世代後にすべての遺伝子において  $w_{gi} = -0.5$  になった。また、エサ発生確率が小さい場合には  $w_{an} = 0.5$  となるエージェントがみられたが、環境が変化して  $w_{an} = -0.5$  となり、環境の変化に対応して遺伝子が変わったことが分かる。学習と進化の場合と進化のみの場合の遺伝子の種類を比べると進化のみの場合のほうが遺伝子の種類が多い。また、学習と進化の場合で発生した遺伝子を持つエージェントのエサ取得数を表 2 に示す。500 世代後は進化があまり進んでいない状態なので  $GI=[0,0,0,0,0]$  から学習を開始したほうがエサ取得数は多い。しかし、環境が変化したり世代を経て進化が進んでいくと、はじめから学習する場合より進化によって学習のコストが減少する遺伝子を持つエージェントが多くエサをとる。

表 2 各遺伝子を持つエージェントのエサ取得数  
(80 世代の学習後)

500 世代後	エサ取得数
[0,0,0,0,0]	53.5
[-0.5,0,0,0,0]	57
[0,0,0,0.5,0]	48
<b>1000 世代後</b>	<b>エサ取得数</b>
[0,0,0,0,0]	17.8
[-0.5,0,0,0,0]	33.8
[-0.5,0,0,0.5,0]	33.2
<b>1500 世代後</b>	<b>エサ取得数</b>
[-0.5,0,0,0,0]	38.7
[-0.5,0,0,-0.5,0]	45.1

## 6. 考察

エージェント全体でみた場合、エサ取得数は学習を用いた集団のほうがエサ取得数は多くなる。これは各エージェントが効率のよい行動ルールを学習するのでエサ取得数が多くなるからである。学習だけを用いた集団と進化も取り入れた集団を比較すると、学習と進化を行う集団のほうがエージェントの平均年齢が低い。これは世代をかけて学習を行わなくても進化によって重みの初期値がエサのとれる方向に変化しているためであり、進化によって学習のコストを軽減している。逆に、進化が十分進んでいない状態だと学習のメリットがエージェントの生き残りに関わってくる。学習のメリットとコストが関係する Baldwin 効果がみられる。

エージェントをエサ取得数で分類し 3 つの場合（学習+進化、学習のみ、進化のみ）を比較した場合、やはり学習を行う集団のほうがエサ取得数の多いエージェントが多くなる。しかし、進化が進むと進化のみの集団も多くのエサをとれるようになるが、進化のコストがかかる。学習だけの集団と進化も行う集団を比較すると、この 2 つで異なる部分はエサ発生確率に関する環境変化が起こる部分である。学習のみの集団の場合はエサ発生確率が小さくなる期間にエサ取得数の少ないエージェントの数に大きな変化がみられるが、学習と進化を行う集団は大きな変化はみられない。また、環境変化の起こる 1000 世代終了後も環境変化前のように戻っている。進化のみの集団も学習のみの集団に比べて変化が小さい。このことから集団の進化は環境の変化に対応する能力を持っていることになる。

重みの学習は  $w_{gi}$  が負になる傾向がある。これはエサの距離が近い方向に進むことになる。 $w_{fn}$  や  $w_{an}$  についても学習の影響があり、これらは他の重みの

学習によって変化するが  $w_m$  は正になる傾向がある。しかし、エサが極端に発生しない環境になると空間にあるエサが少ないので図 6 上のグラフのようにエサの数は気にしないように学習される。このような重みの学習に伴い、遺伝子の進化の傾向も表れてくる。5つの遺伝子のうち、 $iw_{m1}=0.5$  となるパターンが一番多く、学習のコストを減少させる方向に進化する。また、環境によって  $iw_m$  が変化し学習を助ける方向に働き、学習が進化に影響を与えている。さらに、学習と進化を行う集団の遺伝子の種類と進化のみの集団の遺伝子の種類を比較すると、学習と進化を行う集団のほうが遺伝子の種類少ない。こゝでも学習は進化に方向性を与えていることが分かる。

## 7. おわりに

本稿では簡単なエージェントモデルを用いて、エージェント個体の学習とエージェント集団の進化における相互作用を検証した。進化が不十分な場合は学習によって補うことができ、環境の変化などで学習が困難な場合は進化によって学習を助けることができる。また、学習によって進化に方向性を与えたり、進化によって学習のコストを減少することができる。このように、学習と進化には相互作用があることが分かった。今後の課題として、環境が周期的に変化する場合やエージェントの学習に制限がある場合での学習と進化の関係を調べていきたい。

## 参考文献

- [1] Joshua M. Epstein, Robert Axtell 著 服部正太, 木村加代子 訳: 人工社会, 共立出版, 1999.
- [2] J.M. Baldwin : A New Factor in Evolution. *American Naturalist*, vol.30, pp.441-451, 1896.
- [3] 黒木麻理, 乾伸雄, 小谷善行: 山登り法を用いたエージェントの行動ルールの自動生成, *Game Programming Workshop 2002*, pp.147-150, 2002.
- [4] Peter Turney, Darrell Whitley, Russell W. Anderson : Evolution, Learning, and Instinct: 100 years of the Baldwin Effect, *Evolutionary Computation*, vol.4, no.3, pp.4-8, 1996.
- [5] 鈴木麗聖, 有田隆也: 進化と学習の相互作用 - 繰り返し囚人のジレンマゲームにおける Baldwin 効果 -, *人工知能学会誌*, vol.15, no.3, pp. 495-502, 2000.
- [6] David Ackley, Michael Littman : Interactions Between Learning and Evolution, *Artificial Life II*, pp.487-509, 1992.

[7] 永川成基, 有田隆也: 進化と学習の相互作用におけるミームの影響に関する人工社会モデル, 第 26 回知能システムシンポジウム資料, pp.193-198, 1999.

[8] 山本祐歌, 佐々木貴宏, 所真理雄: 学習と進化によるエージェント集団の動的環境への適応, 第 7 回マルチエージェントと協調計算ワークショップ, 1998.

[9] 松浦賢一, 嘉数侑昇: 非均質エージェント系における組織的行動の生成: 2次元追跡問題における考察, *情報処理学会論文誌*, vol.38, no.6, 1997.

[10] 岩永佐織, 生天目章: 選択的な相互作用と異質なエージェントの集合行為の評価, *電子情報通信学会論文誌*, vol.J86-D-1, no.8, pp.505-513, 2003.

[11] 伊庭斉志: 進化論的計算の方法, 東京大学出版会, 1999.