

進化学習システムの動的環境における様々な適応メカニズムの発現

山崎 和子 関口 益照

東京情報大学 経営情報学部
〒265-8501 千葉市若葉区谷当町 1200-2
yamasaki@rsch.tuis.ac.jp

あらまし 動的環境下で進化学習する集団の特徴と様々な適応メカニズムの発現を研究した。佐々木は始めてダーウィン型集団とラマルク型集団の適応の差を示した。我々は単純な連続値関数モデルで、そのメカニズムを調べ次のような結果を得た。動的環境下ではダーウィン型集団が安定して適応する。しかし、それは、環境の不連続変化に起因している。ダーウィン型集団は、同じ環境が長く続いた時、過適応をし、適応度が突如急降下する時がある。また、それは、環境変化の周期と進化の速度の増大により亢進される。環境の連続変化ではラマルク型集団がすばやく適応する。さらに、我々は、これらをひきおこすメカニズムを簡単に説明した。また、2倍体の集団で実験をし、様々適応メカニズムが生成消滅することを示した。

キーワード 進化 学習 動的環境 マルチエージェント

Evolving Learning System Manifests Various Adaptation Mechanisms In the Dynamic Environment

Kazuko Yamasaki Masuteru Sekiguchi

Department of management information
Tokyo university of information sciences
Yatou town 1200-2, Wakaba ward, Chiba city, Chiba, Japan, 265-850
yamasaki@rsch.tuis.ac.jp

Abstract Characteristics of population under the dynamic environment for the evolving learning system and the expression of various adaptation mechanisms are studied. The Sasaki showed the differences between the Darwinian and Lamarckian type population for the first time. We examine it in simple "The Continuous Function Model" in order to investigate the mechanism of the phenomenon. Then we obtain the following results and mechanisms are briefly explained. The Darwinian type population is excellent in the adaptability, when the environment discontinuously changes, and the Lamarckian type population is excellent, when the environment continuously changes. If the same environment is continuing long, the adaptation of the Darwinian population under the dynamic environment breaks. That is the fitness may dive suddenly because of overadaptation. We also examine the population with 2 chromosomes. We obtain under various type of environments various adaptation mechanisms appear and vanish.

key words evolution learning Dynamic Environment multi-agent

進化学習システムの動的環境における 様々な適応メカニズムの発現

山崎和子 関口益照
東京情報大学 経営情報学部
yamasaki@rsch.tuis.ac.jp

概要

Characteristics of population under the dynamic environment for the evolving learning system and the expression of various adaptation mechanisms are studied. The Sasaki[1] showed the differences between the Darwinian and Lamarckian type population for the first time. We examine it in simple "The Continuous Function Model" in order to investigate the mechanism of the phenomenon. Then we obtain the following results and mechanisms are briefly explained. The Darwinian type population is excellent in the adaptability, when the environment discontinuously changes, and the Lamarckian type population is excellent, when the environment continuously changes. If the same environment is continuing long, the adaptation of the Darwinian population under the dynamic environment breaks. That is the fitness may dive suddenly because of overadaptation. We also examine the population with 2 chromosomes. We obtain under various type of environments various adaptation mechanisms appear and vanish.

1 はじめに

マルチエージェント系の研究の目的として、工学的利用目的の他に、生物あるいは人間の集団そのものの研究の手段があげられる。意思決定プロセスの簡単な生物や、生存の為の基本的要件（食糧の確保と生殖）に行動のほとんどが支配されていた原始の人間社会に対しては豊富な研究や成果がある。しかし、豊かな現代の人間社会は、嗜好、哲学、主義主張、蓄財、投機、戦略など、ありとあらゆることを包含し、簡単な解析を許さない。現在のところ、現代社会経済のマルチエージェント系の研究は、

個別問題あるいはトイモデルは見られるものの、一般性のある研究は非常に少ない。

そこでの困難をまとめてみると、エージェントを人間らしく行動させるための AI やロボティクスがぶつかったと同様の困難、人間は自分の持っているスキルをすべて書き下すことができないことやフレーム問題に加え、網の目のような社会のネットワークと多種多様な意思決定プロセス、それらの生成消滅があげられる。このような状況に対して我々は次のように考えた。

* 個々の問題を考えるのではなく、大局的な集団（社会）の性質を問題にする。そのために、個々のエージェントの判断はエージェント自ら学習により獲得させる。

* 集団（社会）の発展も集団（社会）自らが適応（進化）することによりなされるようにする。

* 経済においては、環境が動くということが、本質であるから、動的な環境を問題にする。

それで、個体レベルの適応（個体の生涯のうちに行われる学習）と集団レベルの適応（生物の歴史にわたり行われる進化）をもつ集団の動的環境下での性質を調べることにした。

学習系、進化系それぞれ1つずつとってみても、動的な環境下での研究は少ない。学習進化システムの動的環境下での研究としては、強化学習と進化を扱った Unemi[4]、や、ニューラルネットワークと進化を扱った Nolfi[3]Todd[2] 佐々木 [1] がある。Unemi[4] では、ボルツマン分布の温度に当たるパラメータを遺伝子に組み込むことにより、動的環境下では、学習の初期に思考錯誤をよくする（温度の高い）agent が生き残ることを示している。彼らはこの radical exploration をするようになることを、学習能力を獲得したと言っている。Nolfi[3]Todd[2] では、実ロボットが環境より、世代ごとに変化する決定的な情報と、変化しない確定的情報を得て進化学習するシステムを扱っている。さらに進化の中での学習の役割と

して、次の3つをあげている。1. ボードウィン効果により学習は進化の道しるべになる。2. 短いタイムスケールでの適応を担う。3. 個体が遺伝子長の制限を越えて複雑な表現型を作ることを可能にする。佐々木 [1] では、「ダーウィン型の集団の方が、静的環境下では効率的なラマルク型の集団よりも環境の変動に対して安定した挙動を示す。動的特性の強い環境では、「うまく学習できる」特性が遺伝していくボードウィン効果による間接的な機構の方が重要な役割を持つ」などが述べられている。また進化のみを扱った [8] では、動的環境下で、2 倍体をもつ集団が劣性遺伝子の中に過去の環境を記憶しておくことでよく適応することが実験で示されている。そこで、我々はまず、[1] に述べられている「ダーウィン型集団が安定した挙動を示す」がなぜ起こるのかより着手した。そして、我々はダーウィン型集団とラマルク型集団は、それぞれ異なる種類の動的環境に適応することができることおよび、その適応のメカニズムを明らかにした。次にそのような動的環境下で適応能力のある両集団の染色体を2倍体にし、あらたな適応メカニズムによる適応を可能にして実験を行った。そして、我々は、環境によって、集団が、1つメカニズムによる適応から別のメカニズムによる適応へ、適応メカニズムの転換をすることを実験した。これを、人間社会と照らし合わせて考えると、ダーウィン型、ラマルク型の集団の違いは、例えば、企業や団体を1つのエージェントとして、そのトップが交代する時、前任者のやり方を踏襲するか、試行錯誤でやり方を獲得するかの違いなどと、又、2倍体を利用した適応は、いわば、2大勢力が存在して、時流にあった方が政権をとることなどと、解釈が可能である。また、集団の中でこのような様々なメカニズムがスイッチオン、スイッチオフされているという切口中で社会を捕らえることもできる。以上のような方法で、我々はこの進化学習するシステムを、人間社会のプロトタイプとして、モデルの詳細に関わらずに、人間社会のダイナミクス論じられるのではないかと考えている。このような試みの基礎研究として、ダーウィン型集団とラマルク型集団の適応の差の

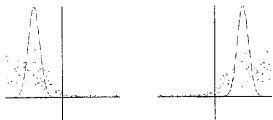


図 1: 環境 $c=1.0$ B環境 A環境

メカニズム、および、2 倍体をもつ集団の適応メカニズムについて述べる。

2 連続値関数モデル

2.1 モデルの概要

- 環境は各 agent に $-1.0 \leq x \leq 1.0$ の実数を提示する。
- 各 agent はニューラルネットワークで入力 x に対する関数値を予測する。(1 入力層、2 中間層、1 出力層の階層型)
- 環境は agent に正しい関数値を返す。
- agent は予測値と環境から返された値の差を用いて誤差伝搬法で学習する。
- fitness の値として、各 agent は誕生した時に 0 を持ち、誤差をこの値から引いてゆく。
- 100 個の agent で集団を作り、1 世代は 2000 ステップとし、世代ごとに fitness の値の低い agent は淘汰される。
- 各 agent は遺伝子として、ニューラルネットワークの 7 個の重みを実数値として持つ。
ダーウィン型の遺伝をする集団の agent は、自分で学習した重みの値は捨て去り、親から遺伝した重みの値を子に伝える。
ラマルク型の遺伝をする集団の agent は、自分で学習した重みの値を子に伝える。
- 世代交代の時、交差および突然変異を行う。
- 学習係数 0.75 慣性係数 0.8
交差確率 0.5 突然変異確率 0.05 突然変異の幅 ± 1.0

2.2 実験の結果

- [結果 1] 環境の不連続変化に対し、ダーウィン型の集団が広いパラメータの範囲で安定的に適応
環境は gauss 型の関数 $\exp(-50(x-x_0)^2)$ とし (図 1)、A 環境 ($x_0 = c/2$) と B 環境 ($x_0 = -c/2$) を 1 世代ごと交互にとる。環境変化の幅 c が 0.2、0.4、0.6、0.8、0.10、1.2、それぞれの時、ダーウィン型集団と

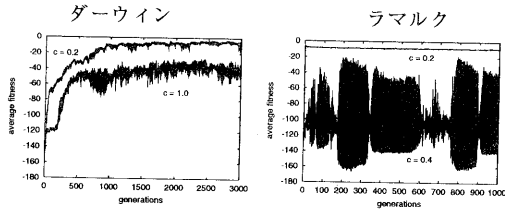


図 2: 環境変化の幅 c と適応度

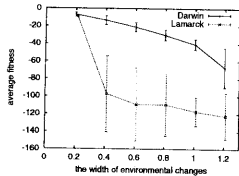


図 3: 環境変化の幅と適応度

ラマルク型集団で実験を行った。各世代の適応度の集団平均値を (図 2) に、適応度の集団平均値と環境変化の幅 c との関係 (図 3) に示す。ダーウィン型集団では環境変化の幅 c が 1.0 以下で安定的に適応する。それに対して、ラマルク型の集団では、0.2 の時には安定的に適応するが、0.4 で安定性は失われる。

● [結果 2] 適応度が突如急降下

結果 1 で安定的に適応したダーウィン型集団に、今度は環境変化の周期を変える実験を行った。(図 4) が示すように、周期が長くなった時、あるいは周期をランダムにした時 (図 5 左) に、しばしば、適応度が突如急降下するようになる。次にこの突如急降下が、進化の速度と、環境変化の周期に、どのように依存するか調べた。進化の速度をコントロールするパラメータとして、突然変異の幅をとり、突如急降下を量的に表す量として、環境が変化したステップの集団平均適応度の減少幅を各環境が変化したステップごとに計算し、その平均値を用いた。(図 5 右) は、周期 10、20、30、突然変異の幅 0.5、1、2 の時の減少幅の平均を表す。突然変異の幅 (進化の速度) が大きくなるにつれ、周期が大きくなるにつれ、減少幅は大きくなる。

● [結果 3] 環境の連続変化に対して、ラマルク型集団

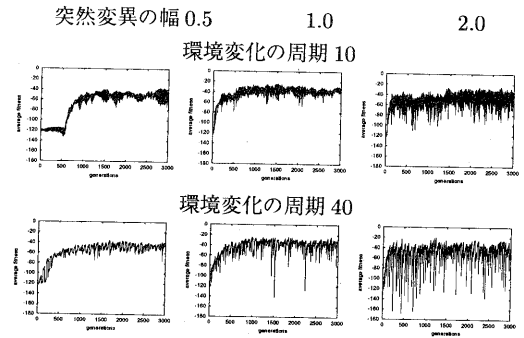


図 4: 適応度の突如急降下と「環境変化の周期」および「進化の速度」

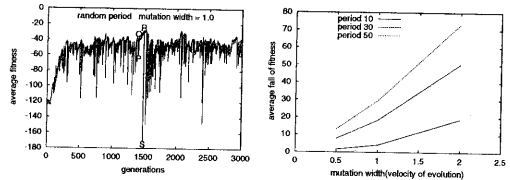


図 5: 適応度の突如急降下

が安定して適応

環境変化を連続変化に近づけるために、gauss 型の関数 $\exp(-50(x-x_0)^2)$ の x_0 を -0.8 から 0.8 まで、等間隔に n 個に区切りこの $n+1$ 個の環境を 1 世代ごとに巡回するモデルで実験を行なった。環境変化の間隔が細くなるほど (環境の数が増えるほど)、ダーウィン型集団は適応が悪くなり、ラマルク型集団は適応がよくなった。(図 6) に 0.16 間隔で 11 環境を 1 世代ごと巡回した実験のダーウィン型集団、ラマルク型集団の集団平均適応度を示す。

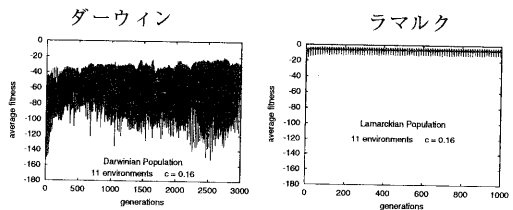


図 6: 環境の連続的变化と適応度

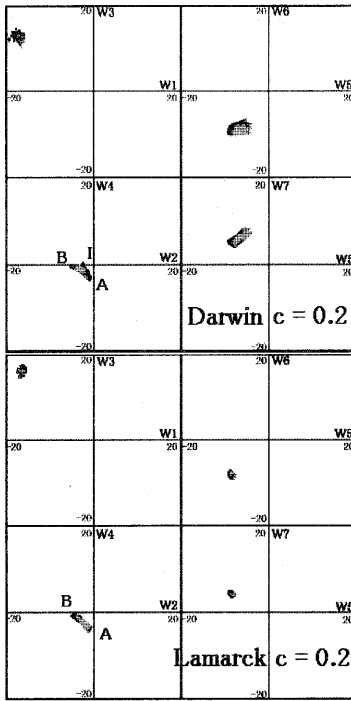


図 7: 重み空間における軌跡

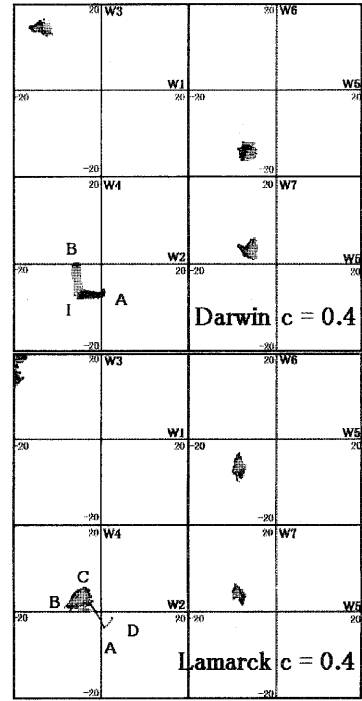


図 8: 重み空間における軌跡

2.3 適応のメカニズム

実験の結果にみられる現象がなぜ観察されるのか、ニューラルネットの7個の重みの空間(適応空間)で説明をする。ある世代100個のagentの誕生から死亡まで、7次元の空間内での軌跡を次のような4つの2次元グラフに描いた。

第2象限は W1-W3 空間	第1象限は W5-W6 空間
第3象限は W2-W4 空間	第4象限は W5-W7 空間

(W1, W3は入力層-中間層の重み、W2, W4は中間層の位相、W5, W6は中間層-出力層の重み、W7は出力層の位相。また、軌跡の両端のうち、黒印のある方が死亡時を表す)

- [理由1] なぜ、環境の不連続変化に対し、ダーウィン型の集団が広いパラメータの範囲で安定的に適応するのか?

環境変化の幅 $c=0.2$

両集団とも適応している、環境変化の幅 $c=0.2$ の時の、重み空間内の100agentそれぞれの軌跡を、(図

7)に2世代(A環境-濃色とB環境-淡色)にわたり重ねて描いた。(A環境-濃色の軌跡の上にB環境-淡色の軌跡が上書きされている。)ダーウィン型集団では、楔型の運動、 $I \rightarrow A$ 、 $I \rightarrow B$ をし、ラマルク型集団では、往復運動、 $A \leftrightarrow B$ をするのが、主に第3象限(W2-W4空間、中間層の位相の空間)で顕著にみられる。(ここで、I:初期値 A(B):A(B)境下で学習をし世代の終わりに到達した値)。

環境変化の幅 $c=0.4$

ラマルク型集団が適応できなくなった、環境変化の幅 $c=0.4$ の時の重み空間内の軌跡を(図8)示す。ダーウィン型集団では、第3象限(中間層の位相の空間)で、楔型の運動、 $I \rightarrow A$ 、 $I \rightarrow B$ がより顕著になるが、ラマルク型集団では、往復運動、 $A \leftrightarrow B$ ができなくなる。B点と別の値C点の間で往復運動をしている。つまり、B環境の時には学習できているが、A環境の時には学習できていない。

環境変化の幅 $c=1.0$

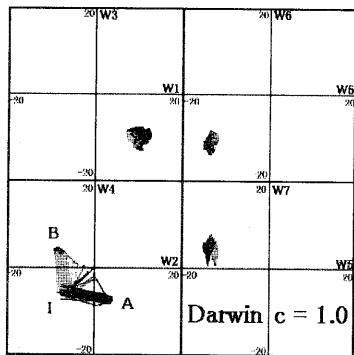


図 9: 重み空間における軌跡

環境変化の幅 $c=1.0$ の時の、ダーウィン型集団の軌跡を (図 9) 示す。楔型の運動がより鮮明に見られる。

これらを説明するために次のような考察をした。A (B) 環境に適した重みの地点 A (B) とその吸引域 circle A (circle B) を、(図 10) に示す。ラマルク型集団では、最小限、図のように、circle A と circle B は半分以上重なることが A → B、B → A と山登りの学習をするためには必要である。これに、対してダーウィン型集団では、最小限、(図 10) のように circle A と circle B が接していれば、その接点に、親から遺伝する重みの I 地点が、進化してくることによって I → A、I → B と山登りの学習をすることができる。このような山登り学習可能領域の体積は吸引域 circle の直径を λ とし、適応空間の次元 (重みの数) を n とすると、

ダーウィン型集団 学習可能領域 $< (2\lambda)^n$

ラマルク型集団 学習可能領域 $< \lambda^n$

のオーダーである。n はここでは 7 (学習可能領域の体積比は 128 倍) であるが、通常の実験用のニューラルネットワークで数 10 (100 の時、学習可能領域の体積比は約 10^{30} 倍)、一般の生物レベルでは大きい数であると思われるので、このラマルク型集団とダーウィン型集団の差は決定的になる。すなわち、ダーウィン型集団で、不連続変化する「複数の環境の組」に対する、(最適値でなくとも) 或る程度適した重みの値が、重み空間内の距離 $(2\lambda)^n$ より近い位置にある時のみ学習ができ、その時、その「複数の環境の組」の広い意味での一般性を学習したということになる。

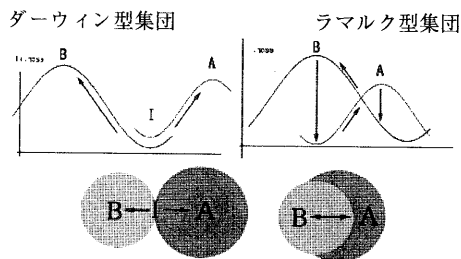


図 10: 山登り (学習) と環境の不連続変化

- [理由 2] なぜ、適応度が突如急降下するのか?それが、環境変化の周期および進化の速度の増大によって、亢進されるのか?

ランダムな周期の時、適応度の突如急降下がみられる (図 5 左) の P、Q、R、S、時点それぞれの、重み空間第 3 象限の軌跡を (図 11) に示す。P (1414 世代) 時点では正常に A 環境に適応している。Q (1415 世代) 時点で B 環境に交代する。それから、約 60 世代にわたり、B 環境が続いた為、ボードウイン効果により、R (1482 世代) 時点まで、親から遺伝する重みの値が移動する。S (1483 世代) 時点で A 環境に交代した時、もはや、適応することができない。(Q 時点から R 時点への親から遺伝する重みの値の移動はわずかであるが、それが決定的な違いをもたらす。) この、ボードウイン効果は、環境変化の周期および進化の速度に対して、正の効果を持つので、この突如急降下現象は、周期および進化の速度の増大によって、亢進される。

- [理由 3] なぜ、環境の連続変化に対して、ラマルク型集団が安定して適応するか?

(図 12) に gauss 型の関数 $\exp(-50(x - x_0)^2)$ の x_0 を -0.8 から 0.8 まで、0.16 間隔に区切り、この 11 環境を 1 世代ごと巡回する実験をした時の、両集団の重み空間内の軌跡を示す。(軌跡の濃淡の違いが環境の違いを表す。) ラマルク型集団は連続的に 11 世代に渡って長い距離をスムーズに動くのに対して、ダーウィン型集団は、初期の重みの値がアンカーになって、長い距離を動くことができない。また、世代内ステップごとに環境を変化させた時、ダーウィン型集団では、重み空間の軌跡が世代の途中で、折り返し戻ってくるなど、複雑な適応の仕方を示す。

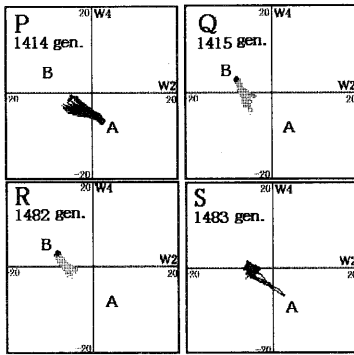


図 11: 適応度が突然急降下する前後の重み空間における軌跡

3 2倍体連続値関数モデル

3.1 モデルの概要

1倍体連続値関数モデルに加え、もう1本のニューラルネットワークの重みからなる染色体および、それらのどちらが発現するかを決定する遺伝子をビット値で持つ。そのビット値が突然変異する確率 0.01、染色体の交差確率 0として実験を行なった。

3.2 実験の結果

環境変化の幅 $c=1.0$ のダーウィン型集団で環境変化の周期を変えて実験を行ない次のような結果を得た。この集団は動的環境下で

1. 前 section でみたような親から遺伝した値をアンカーとして持つことによる適応能力
2. 1 方の染色体の中に過去の環境を記憶しておくことによる適応能力

を両方備えている。

- 環境変化の周期が小さい時には、1. による適応をする。1. では過適応をする環境変化の周期が大きい時には、2. による適応をする。(図 13) に環境変化の周期を 100 とし、2. による適応をした時の重み空間の軌跡を示す。淡色の点が劣性遺伝子を表す。A 環境で劣性であったものが B 環境で優性になり、そこから学習をして適応していることがわかる。

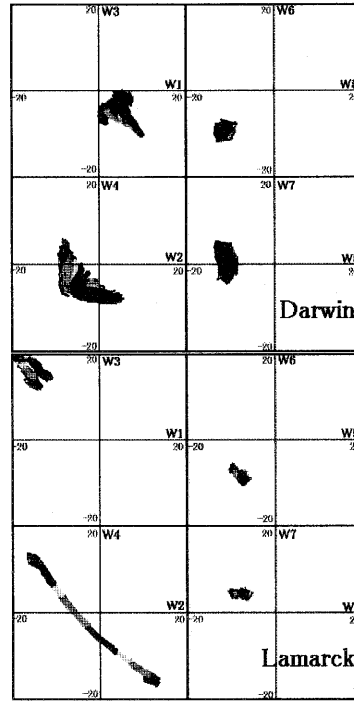


図 12: 連続的環境変化の時の重み空間における軌跡

- どちらの適応も可能な環境変化の周期では、1. による適応をする。
- 1. による適応をした集団を 1. では過適応をする環境変化の周期が大きい中にいれると 2. による適応を始める。
- 1. 又は 2. による適応をした集団をどちらの適応も可能な環境変化の周期の中にいれるとすでに獲得した適応を続ける。

4 おわりに

佐々木 [1] と同様の餌と毒のモデルでも同様の実験を行った。それらは、山崎 [5][6] に詳しく述べられている。餌と毒のモデルでは環境変化の幅を変えたり、連続的变化をさせることはできないが、それ以外の実験で、連続値関数モデルと同様の結果を得た。また、重みの空間の次元が大きいため直接重みの空間内の軌跡を描いて考察

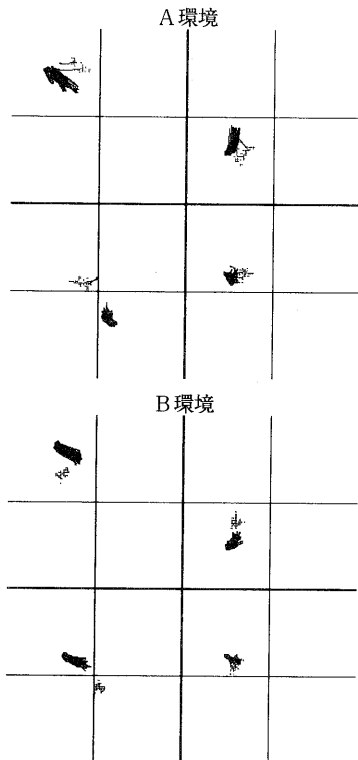


図 13: 2倍体による過去の環境の記憶による適応

することはできないが、同様の現象、同様のパラメータ依存性を示すということは、そのような現象を生じさせるメカニズムも、連続値関数モデルと同じであると推測できる。

また、変化する環境の中で、過適応が身を滅ぼすというのは、人類の歴史、我々の社会でもよく見られることで、興味深い。

ここでは、環境の動的変化はどのエージェントにも等しく、外生的に与えられているが、人間の社会では、エージェント間の相互作用が強く、外生的な環境の変化はこの相互作用によって、あるいは増幅され、あるいは遅延効果が生じ、と様々な形で個々のエージェントに伝わると思われる。このような、ダイナミクスの分析も大いに興味を持たれるところである。

東京情報大学情報学科のみなさんからたくさんのおアドバイスをいただき感謝いたします。

参考文献

- [1] 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: コンピュータソフトウェア, Vol.14, No.4 (1997), pp.33-46
- [2] Todd,P.M. & Miller,G.F.: Exploring adaptive agency: 2. Simulating the evolution of associative learning: From animals to animats: Proceeding of the First International Conference on Simulation of Adaptive Behavior. Cambridge,MA: MIT Press.
- [3] Nolfi,S. & Parisi,D.: Learning to Adapt to Changing Environments in Evolving neural Networks: Adaptive Behavior,Vol.5,No.1(1997),75-98:
- [4] Unemi,T. et all: Evolutionary Differentiation of Learning Abilities - a case study on optimizing parameter values in Q-learning by a genetic algorithm
- [5] 山崎和子, 関口益照: 動的環境下での進化学習システムの特徴と適応のメカニズム: ソフトウェア科学会誌、投稿中:
- [6] 山崎和子, 関口益照: 動的環境下での進化学習システムの特徴と適応のメカニズム: ソフトウェア科学会、マルチエージェントと協調計算ワークショップ (1999):
- [7] Yamasaki,K & Sekiguchi,M : The Clear Explanation of Different Adaptive Behaviors between Darwinian Population and Lamarckian Population in the Changing Environment:Proceeding of the 5th inInternational Symposium on Artificial Life And Robotics(2000)
- [8] Goldberg, D.E. and Smith, R.E., Nonstationary Function Optimization using Genetic Algorithms with Dominance and Diploidy, in *Genetic Algorithms and Their Applications: Proc. the Second Int. Conf. on Genetic Algorithms (ICGA87)*