

### 6.3 「ミニチュア経済 in コンピュータ」への小さなアプローチ—動的環境下での学習進化システム

山崎和子

yamasaki@rsch.tuis.ac.jp

東京情報大学経営情報学部

#### 6.3.1 プロローグ

日本の景気の回復基調が確かになるにつれ、ますます気になるのが、アメリカのダウである。前代未聞のインフレなき成長、ソフトランディングが可能なのか？私など数年前、ダウが2000\$の時にもうこれは売り時かと思ったのを覚えているが、なんと、それから5倍以上も上がった。

私の住む場所の近辺の地価は、バブル時の1/10である。1/10というのは、チューリップでもシテ株の値段でもなく、どこかの閉店セールで値段でもなく、りっぱな土地の値段がである。バブル時、日本全土の土地の値段が、アメリカ全土の土地の値段を上まわったと聞いて、当時多くの人がおかしいと声高に言うことができなかった。日本型経営、日本人の勤勉さが称賛され、「これがバブルだ。」と確信させないための理由はごろごろしていた。これが経済の常であろうか。

マイクロソフト、インテル、ネットスケープ、ヤフー、アマゾン.com、....思いつくまま株価数百倍というアメリカの会社をあげていってきりが無い。「ダウ10000\$は当然」というのはどうなのであろう。ネット関連の技術、ノウハウが急激に進歩する時代は、どこかいいかげんで、冒険的な人種に勝ち目があって、緻密で勤勉な人種は、技術が成熟し、応用の時代になるまで、雌伏せよというのであろうか。マイクロソフトに対する、独占の認定は何かの節目になるのだろうか？。この、地球上の資源が枯渇しつつある中で、クレジットで一手に消費している国が栄える（消費＝繁栄）というの、おかしいと思うのだが？それにしても、アメリカの判断が冷静に見えるのは、私も何かにだまされているのであろうか。

前置きが長くなったが、「日本全土の土地の値段が、アメリカ全土の土地の値段を上まわった」時代、みんな、たくさんの貯蓄を持っていた。さいふの中の千円札や100玉と同じお金もっていた。それが、今、どこにいつってしまったのだろうか。経済関連の貨幣の本は5万とあり、三段論法、マルクスが理系の人間をなやませるが、それとはちがって、我々のことばでお金を理解したい。千円札や100玉はお金の保存則を潜在意識に植え付けるが、お金の物の側面(価値の貯蔵手段)、媒介の側面(価値の交換手段)はどういう関係になっているのか？信用の創造とは？コンピュータシミュレーションをして、お金の正体を見たいと思った。小さな経済、中央銀行が1つ、その他銀行が数個、工場が数十個、それに商社、消費者兼労働者。このように、実際の人間、会社、生物などをまねたものをコンピュータのなかに作りシミュレーションをするやり方は、マルチエージェントと呼ばれているが、ただ、実際の社会をまねたものをコンピュータのなかというの、昔からあるようである。たとえば、人口の予測するのに数万人の人を作ってというのは、マイクロシミュレーションモデルとよばれ、10年前ごろ研究されていた。しかし、そこでは戸籍簿や住民台帳にあたるようなファイルをコンピュータの中につくり、それを参照して、プログラムが動いていた。それに比べて、マルチエージェントという時には、オブジェクト指向の広がりもあって、人は皮膚で周りの環境と接している塊であって(1つのプログラム単位として独立して)、環境を知覚し、考え判断をし、環境へ行動を返す、自律的存在、環境と対峙する存在であることが強調される。しかし、このことが、非常な困難さをもたらす。

私のところのゼミ生が、サッカーのゲームを作って、どこか有名なゲーム会社のコンクールで1等をとったが、監督の目で、中田を右に3歩走らせて、名波にパスしてとプログラムして、それらしい動きをつくることは、比較的簡単である。しかし、マルチエージェント流にいくと、選手同

士があうんのタイミングで動くように選手の思考判断メカニズムをプログラムしなくてはならない。それでこそ、この複雑な状況のなかで、人間がはじめから予期せぬ事態にも頑強に柔軟に対応できるプログラムができるというのである。社会のシミュレーションでは、人間がはじめから予期したように動くのでは、シミュレーションにならない。エージェントを人間らしくプログラムすることで、AIおよびロボティクスがぶつかった困難にここでも出合うことになる。たとえば、組織をオーガナイズするのが上手な人がいても、かれはそれをノウハウとしてすべて書き下すことはできない。彼が感覚的に動いている部分がむしろ重要なのである。また、人間は周りの環境の大事な部分を瞬時にとらえてそれだけを記憶しているが、ロボットには簡単にはできない。ロボットが周囲の環境のすべて頭のなかで追いかけるとしたら、それだけで、莫大なメモリーと時間を要してしまう。

たとえば、景気が自律反転をするような、自然な思考判断メカニズムを工場、商社、消費者兼労働者にもたせるにはどのようにしたらよいか。不良工場、商社がつぶれ、在庫調整が終り、経費を切り詰めるだけ切り詰めてそれに疲れ、、、。単純な思考判断だけでは、下降のスパイラルがいつまでも続くだけである。

また、こういうシミュレーションはパラメータを多く含むから、望むような結果をパラメータの調整で作ることが可能になる。

したがって、現在は多くのマルチエージェントのシミュレーションが1、2種のみエージェントの集合を扱っており、トイモデルに留まっている。(現在は株式市場などの個別問題に手がついている時期だという人もある。)

このような現状の中で我々はつぎのように考えた。

- 個々の問題を考えるのではなく、大局的な集団(社会)の性質を問題にしよう。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よう。
- 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しよう。
- 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しよう。

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)をする社会の動的な環境にレバントな性質を求めることにした。

社会のシミュレーションで遺伝子が何にあたるかと言えば、たとえば株式市場の場合では、エージェントの持っている意志決定方法(株売買タイミングの決定方法)であり、それぞれの方法について、パラメータの値を学習により適応させながら売買をしている。もうけの少ないエージェントは消え去りもうけの多いエージェントが生き残り仲間を増やす。(淘汰)個人の意志決定方法は、集団内で交換されることがある。(人間でいえば本を読んだり、セミナーで勉強したり)(交差)

動的環境下での進化学習システムの研究は未だ非常に少ない。佐々木[1]が発見した、「ダーウィン型とラマルク型集団の適応の差」について、我々は最近、彼らが示したその理由は正しくなく、別の非常に明解な理由があることを示した。次にその仕事の概略を説明する。

### 6.3.2 動的環境下での学習進化システム

進化学習の方法としては、ダーウィン型とラマルク型がよく知られている。ダーウィン型遺伝では個体が生涯の内に学習により獲得した獲得形質は遺伝しない。ラマルク型遺伝では獲得形質が遺伝する。実際の生物では、ダーウィン型遺伝が行なわれているが、工学的には適応の速いラマルク型遺

伝が使われることが多い。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。

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。それに対して、我々は、次のような問題意識を持った。

- 静的な環境から、動的な環境へ順次変えて実験を行った時、ラマルク型からダーウィン型への優位性の交代はどのようにおこるのか。
- 「うまく学習できる」特性とは、適応空間上で、何を意味するのか。

我々は、そのために、環境を連続的に、無限個の状態に変化させることができ、しかも適応空間上の考察を可能にするため、なるだけ、適応空間の次元の低いモデルを考え、実験を行なうことにした。また、モデルが一般性を持つように、もっともよく使われる、学習法、進化の方法を用いることにした。それで階層型ニューラルネットワークで、連続値の関数を学習する個体集団を作り実験を行った。そして、次のような結果を得た。

1. 動的環境下ではダーウィン型集団が適応性に優れている。しかし、それは、環境の不連続変化に起因している。
2. 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。
3. 環境の連続変化ではラマルク型集団が適応性に優れている、

またこれらをひきおこすメカニズムを次のように説明した。適応空間(重み空間)内で、学習により、ダーウィン型集団では、楔型の運動、ラマルク型集団では、往復運動がおこる。したがって、異なる環境いずれにも山登りの学習で適応できるための条件は、

ダーウィン型集団      体積  $< (2\lambda)^n$

ラマルク型集団      体積  $< \lambda^n$

という大きさの体積のなかに、異なる環境の適応状態が分布していることである。この2つの限界体積の比は一般的なnの値で、膨大な数になる。

また、[1]と同様の「毒と餌モデル」で実験をした。このモデルでは、環境を連続変化させることができず、また、適応空間(重み空間)の次元が大きいため、直接適応空間(重み空間)内の軌跡を視覚的に描くことができない。しかし、「毒と餌モデル」についても、(ここでは省略したが)「連続値関数モデル」と同様の結果を得、特に、結果2.の環境変化の周期や進化の速度に対する依存性が同じであることから、「毒と餌モデル」の実験結果は、「連続値関数モデル」と同様のメカニズムに起因していることが推測される。

### 6.3.3 連続値関数モデル

#### 6.3.3.1 モデルの概要

100個のエージェントで集団を作る、それぞれのエージェントはニューラルネットワーク(1入力層、2中間層、1出力層の階層型)を持ち、この7個の実数値の重みを遺伝子として持つ。環境は各エ

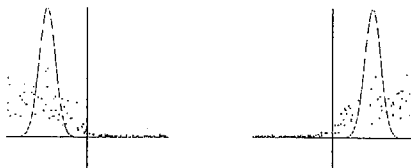
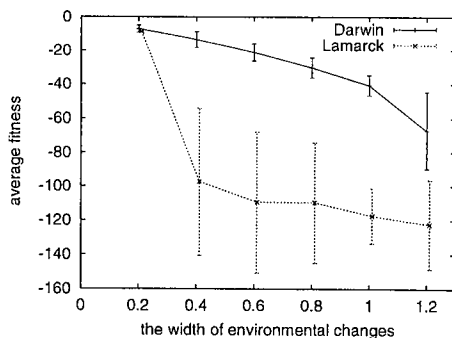
図 1: 環境(連続値関数)  $c=1.0$  B環境 A環境

図 2: 環境変化の幅と適応度

エージェントに  $-1.0 \leq x \leq 1.0$  の実数を提示する。各エージェント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。環境は各エージェントに正しい関数値を返す。この関数形を周期的に変えていく。

### 6.3.3.2 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。環境はガウス型の関数  $\exp(-50(x-x_0)^2)$  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。適応度の集団平均値と環境変化の幅  $c$  との関係を図2)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。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、結果1と同じ関数形で環境変化の幅を  $c = 1.0$  に固定して、今度は環境変化の周期を変える実験を行った。(図3)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3下)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。

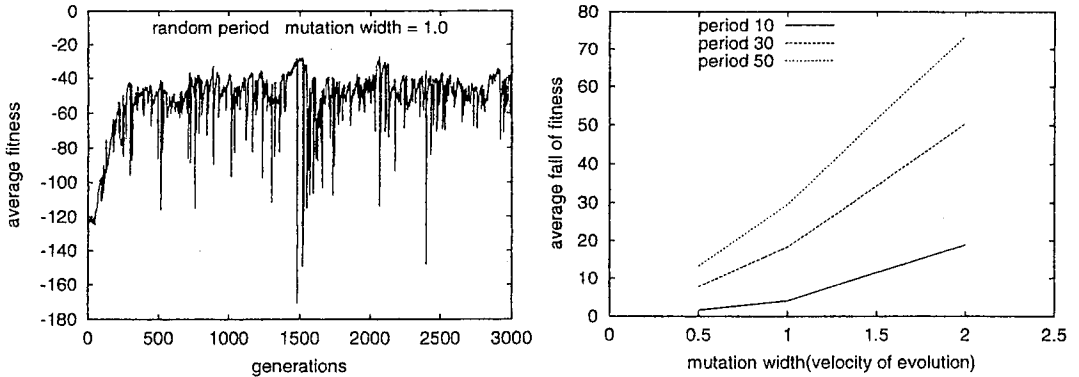


図3: 適応度の突如急降下と「環境変化の周期」および「進化の速度」

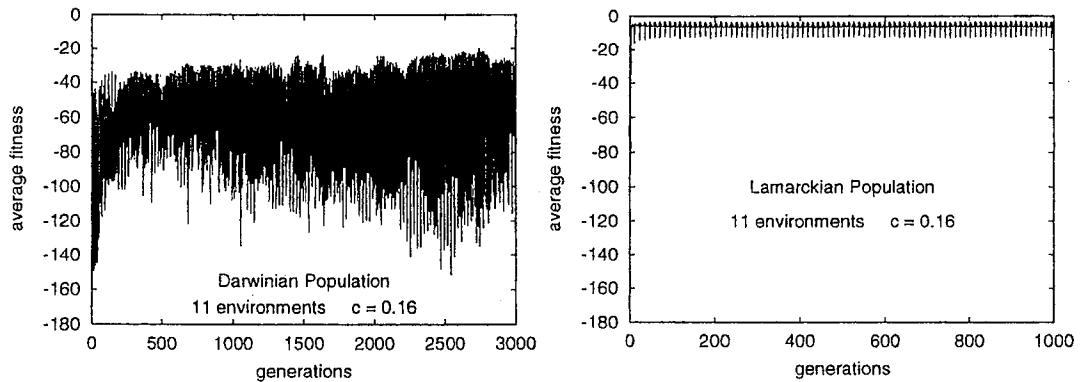


図4: 環境の連続的变化と適応度

- [結果3]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  
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 $n$ 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4)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。

### 6.3.3.3 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

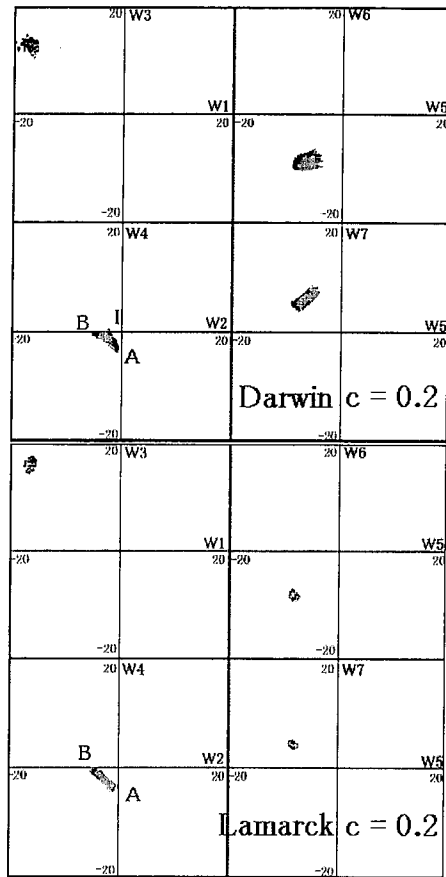


図 5: 重み空間における軌跡

ダーウィン型集団

ラマルク型集団

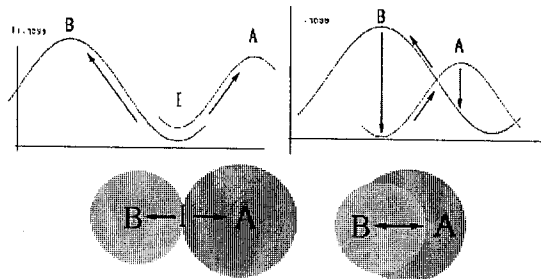


図 6: 山登り(学習)と環境の不連続変化

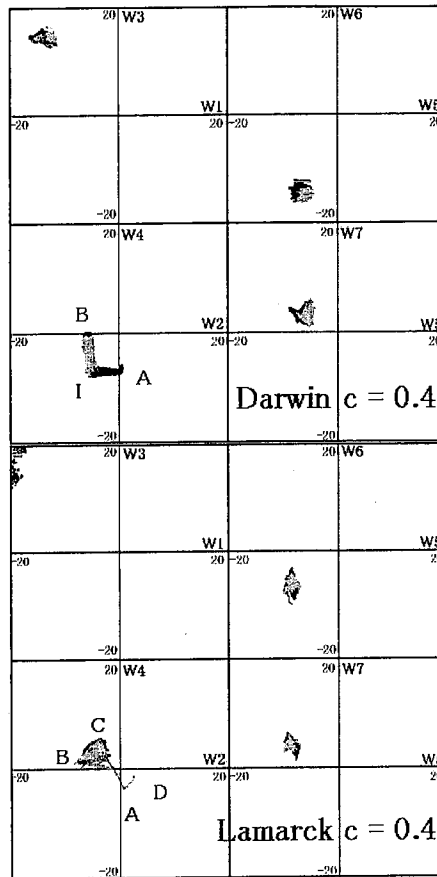


図 7: 重み空間における軌跡

(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5)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7)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。

環境変化の幅 $c=1.0$

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8)示す。楔型の運動がより鮮明に見ら

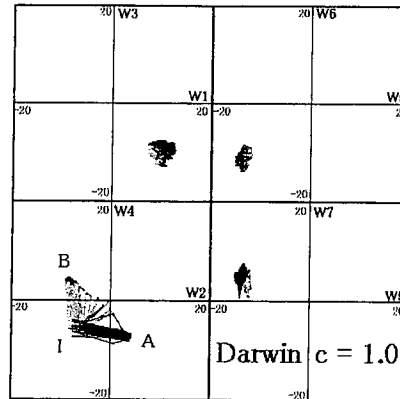


図 8: 重み空間における軌跡

れる。

これらを説明するために次のような考察をした。A環境に適した重みの地点Aとその吸引域circle A、B環境に適した重みの地点Bとその吸引域circle Bを、(図6)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6)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。地点Aを固定して、学習可能な地点Bの存在領域の体積は、吸引域circleの直径を $\lambda$ 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      存在領域の体積  $< (2\lambda)^n$

ラマルク型集団      存在領域の体積  $< \lambda^n$

のオーダーである。 $n$ はここでは7(存在領域の体積比は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、存在領域の体積比は約 $10^{30}$ 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、重み空間内での距離 $(2\lambda)^n$ 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。

- [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3)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9)に示す。P(1414世代目)時点では正常にA環境に適応している。Q(1415世代目)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世代目)時点まで、親から遺伝する重みの値が移動する。S(1483世代目)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がであるが、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10)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間隔で11環境に区切り、これを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



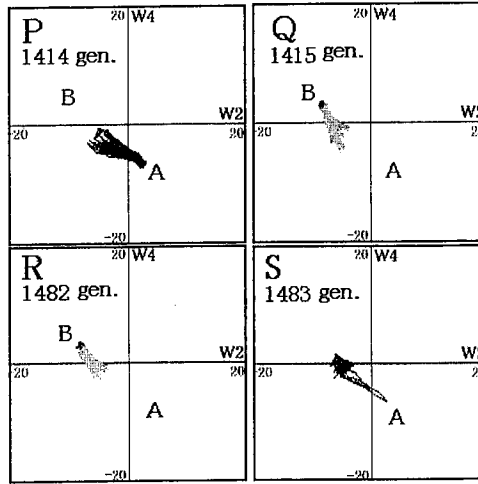


図 9: 適応度が突然急降する前後の重み空間における軌跡

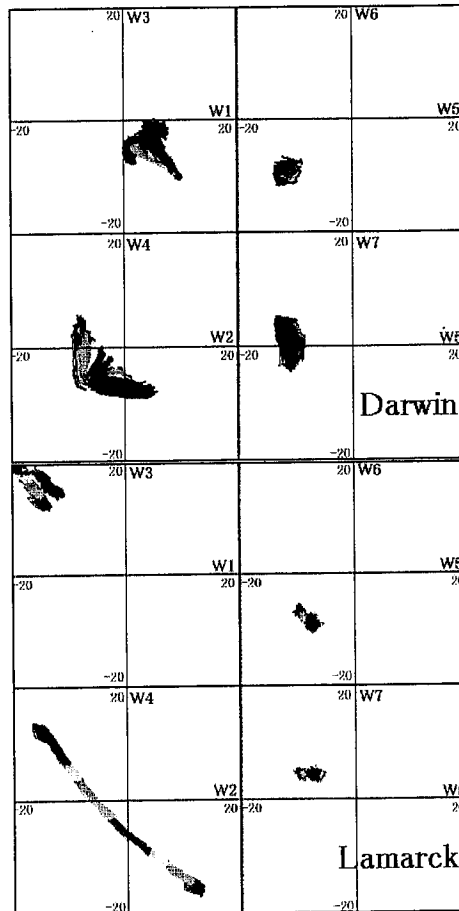


図 10: 連続的環境変化の時の重み空間における軌跡

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。

### 6.3.4 おわりに

この実験で示された、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。種の絶滅の規模がフラクタル性を持つという研究があるが、このモデルの過適応による適応の急激な下降から、フラクタル性を抽出できないか、現在研究中である。マルチエージェント系の研究で生物や原始の人間社会に対しては豊富な研究や成果があがっているが、現代の人間社会に関しては、個別問題あるいはトイモデルは散見されるものの、一般性のある研究は非常に少ない。また、社会経済モデルにとって、環境が変化することは、必要であるだけでなく、それが本質だとさえ言える。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。先にのべた「バブルの時のお金はどこに消えたか」という素朴な動機にはほど遠いが、我々はこのモデルを第一歩として、モデルの詳細に関わらずにどこまでのことが言えるか、研究を進めていくつもりである。

### 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:
- [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 al: 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] 山崎和子, 関口益照: 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: 情報処理学会、知能と複雑系研究会報告(2000):