

Evolution of Game Strategies through Genetic Fusion

Takashi Ikegami *

*The Graduate School of Science and Technology
Kobe University, Rokkodai, Nada-ku, Kobe 657, JAPAN*

Abstract

We present a model of evolving strategies in the noisy iterated Prisoner's Dilemma. Each strategy computes one's next move from past experience but miscomputes within the given probability.

The strategy is encoded on a binary tree structure, which is modified and complicated through our mutation and fusion algorithm. Two strategies represented as trees interact by forming a fusion tree: a new strategy with an extended memory length. Such a synthesized strategy combines and improves the parent strategies and is thus expected to become a better strategy. Because of this tree coding and interactive forms of strategies, a memory length can be extended potentially infinitely.

By using replicator equations, we simulate the ensemble of such evolving strategies. We are concerning with how the strategies with long memory emerge and examining the complex landscape of this game. Various kind of metastable strategies appear depending on a noise level. For example, a famous Tit For Tat only survives in the noiseless or high noisy world. It is shown that a relevant noise level evolves long memories and those strategies with long memories promote altruistic behavior.

□ はじめに

生命現象は、そのダイナミクスがもつ偶発性と発展性ゆえに興味深く、またきわめて扱いにくいものである。偶発性とは、例えば生物が個体レベルで示す盲目的・突発的行動、あるいは遺伝子レベルで示す突然変異や重複・組替えのことである。また発展性とは環境に対する適応能力、学習能力、技術・機能の発明・革新のことである。この偶発性と発展性はどう関わりあっているか。いかにしてシミュレートするか。これらが生物の理論的研究のひとつの方向である、と考えている。

生物進化はこの偶発性が自然選択を通じて発展性を与えているというのが、ダーウィン流の考えである。現代におけるひとつの問題点は、この偶発性の起源が単に熱的揺動によるものでなく、それ自体構造を持ち制御され方向づけられた「偶発性」なのではないか、という点である [1]。以前考えられていたよりも生

*E-mail address : ikegami@gradis.scitec.kobe-u.ac.jp

物個体は異種間で複雑に競合・協調し、遺伝子は他の遺伝子と複雑に相互作用しあう。こうした過程の中から生まれた偶発性が生物に進化を促す。シミュレーションにおいては『あらかじめ計算されているが非常に不安定なダイナミクス』というものによって進化のダイナミクスが特徴づけると期待される。

遺伝子の発展を担うダイナミクスは偶発的ではない。遺伝子重複による遺伝子の進化機構が進化を担っている、という学説が大野乾によって提唱されている[2]。最初に優れた遺伝子がある偶発性によって作り出し、以後それを重複コピーすることで進化してきた。新しく進化してきたように見える遺伝子も実は最初の遺伝子のコピーに過ぎない。この説は、遺伝子はよく動き回り種の壁も乗り越えてコピーされるといふことや、遺伝子に存在する多くの反復パターンの存在によってより確証されつつある。問題はこの過程がはたして本当に機能の革新・進化を担えるか、ということである。あるいは重複進化によって表現型の機能が向上するように遺伝子型と表現型の対応づけ（コーディング）がなされてきたのだろうか。

この問題に対しわれわれは計算機シミュレーションを行った[3]。遺伝子は計算機の中で1次元のビット配列として表わされ、ふたつの遺伝子は合体して新しい遺伝子を生成する。与える表現型は各遺伝子座対の相対的な距離の関数として定義され、多くの準安定状態を持った複雑な関数である。遺伝子は並列的に動き回り合体しながら、よりよい表現型を探しもとめる。表現型の善し悪しは、簡単のため遺伝子型の関数として一意に与えられているとする。合体を考慮することで遺伝子の動きまわる空間は無限に広がるので、最適な遺伝子型というものを限定はできない。

一般に探し方のアルゴリズムがあらわに書き下せない問題に対し、進化のダイナミクスは威力を発揮する。上の問題では、遺伝子は『モジュール進化』によって進化を加速することが見出された。モジュール進化とは計算機のプログラム中で汎用関数をコールするように、それと合体することで多くの遺伝子の表現型が向上する汎用型遺伝子の進化である。モジュール遺伝子の出現とともに全体が加速進化する。結果的に遺伝子は繰り返しパターンを多くもち、個々の遺伝子が突然変異で進化するよりも数桁速くいい表現型にたどり着くことができる。

ところでより面白いのは表現型間の相互作用によって個々の表現型の善し悪しが決まっている場合である。この状況はゲームによって表現することができる。

□ 2 ゲーム戦略の進化[4]

§ 2.1 ジレンマ・ゲーム

ここで扱うのは、利己的社会における協調性の進化の研究としてよく取り上げられる『繰り返し囚人のジレンマ・ゲーム (IPD)』である[5]。ここで囚人のジレンマ・ゲームは、2つの手CとDをもち、手DがCをdominateする以下の利得表によって定義される。対戦はすべてペアで行なうとする。

C	D	C	D	ここで列は相手の手、行は自分の手。また得点は R) (S + T)/2 及び T) R) P) Sを満た している。計算では、R = 3, S = 0, T = 5, P = 1。
C	R	C	R	
D	T	D	S	
プレイヤー-1		プレイヤー-2		
の利得表		の利得表		

相手がCをだすかDをだすか分からない状況で、1回だけこのゲームを行なうとすれば、より高い得点を得るにはDを出すのが合理的である。相互に協力的な解

(C, C)は成立することはない。このゲームを繰り返して行い総和を得点とすれば、状況は異なる。長期的展望から協力解が進化するるのである。もちろんただCを出し続けるだけでは相手に裏切りつづけられる(ずっとDをだす)ので、なんらかの作戦・戦略が必要となる。その中でもっともよく知られるのが『しつぺ返し(T f T)』である。T f T戦略は、〈Cでスタートし前回の相手の手を模倣する〉というものである。T f Tは簡単でありながら、対戦による被害を抑えつつ自分や他の協力的な戦略とは協力解を進化させるすぐれた戦略である。

さてペア対戦を行なう戦略の集団を考えよう。生物集団のアナロジーから、おのおのの戦略は単為生殖により自己増殖を行ない、その増殖率は獲得得点に比例するとする。この時どんな戦略が集団内に広まっていくだろうか。ゲームとしてIPDを選んだ場合協力解を進化させるT f Tの様な戦略は集団内にとどまられるだろうか。新しい戦略が次々と集団内に生成される場合はどうだろうか。

これらの興味深い問題は、いろいろなアプローチで特にここ10年の間に研究されてきた。戦略集団の時間発展を数学的に記述するものとしてある力学系モデル(TYZ方程式)が提唱され[6]、研究を促進した。ここでの研究もTYZ方程式を用いて個々の戦略の数の時間発展をシミュレートする。すなわち、

$$y(i) - y(i) = \Delta y(i) (\sum g(i, j) y(j) - \langle g \rangle) \quad (1)$$

$y(i)$, $y(i)$ はそれぞれ時間 t , $t+1$ における戦略 i の集団中の頻度である。従って集団中の全戦略に関する和は1に規格化される($\sum y(i) = 1$)。各戦略の数はその戦略の獲得得点($\sum g(i, j) y(j)$)と集団の平均得点($\langle g \rangle = \sum g(i, j) y(i) y(j)$)との差に比例して増加すると仮定する。

TYZ方程式は有限個の戦略のセットの上で定義されたものであったが、戦略の進化を課すことは方程式(1)を構成する変数 y の数自体を時間的に変化させていくことである。以下では遺伝合体と同じスキームを用いて戦略の進化をシミュレートする。

§ 2.2 記憶戦略の進化

対戦相手と自分の過去に打った手のパターンから次の一手を決定する戦略群をここでは扱うことにする。このような記憶戦略は(非均一な)2進木にコードすることができる。現時点(ルート)から木を遡って過去の対戦パターンを探索し、マッチするパターンが見つければCを、なければDを次の一手とする。ここでルートから数えて(2M-1)番目のノードはMステップ前の相手の手、(2M)番目はMステップ前の自分の手を表わす。例えば次の例では過去の手が{D C D ···}ならばCを次の一手とし、{D C C ···}ならばDを次の一手とする。また対戦パターンのマッチングは長さの短い方に合わせるので、初めの2回が{C D D C}ならばCを3回目における手とする。対戦開始の一手は2進木とは別に与えておく。

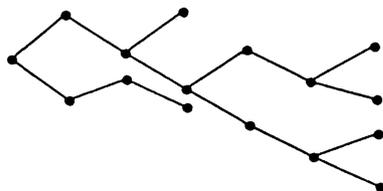


図1 2進木でコードされた戦略。ルートから数えて奇数番目は相手の手、偶数番目は自分の手をコードする。上側への分岐がC、下側への分岐はDを表わす。

この2進木は前の章で触れた点突然変異と合体によって進化する。
 ○点突然変異・・・ルート（木の始点）からリーフ（終点）に至るすべての経路について確率 μ でひとつのノードを取り替える（ $C \rightarrow D$ ； $D \rightarrow C$ ）。
 ○合体・・・ふたつの2進木を確率 ν で、一方の木の任意のリーフにもう一方の木のルートをつないで新しい2進木を生成する。

すなわち戦略同志は『合体』を介して相互作用し新しい戦略を生成する。生成された戦略は過去に遡って探索する記憶の長さを延長し、かつ親の戦略を受継ぐのである。

□3 シミュレーション

シミュレーションにおいてもひとつ重要な設定は『間違っただ手をうってしまう』確率 ε を考慮していることである。戦略の進化における間違いの確率 ε の重要性はリングレンらの仕事によって明確にされている[7]。無限回の繰り返しゲームを考えた場合、間違い確率 ε を考慮することで初回の手による依存性が消えると同時に全てが混合戦略となるため単一の戦略でもESSとなりうる。ESSとはある戦略（あるいはいくつかの戦略のセット）によって占められた集団が、任意の少数の戦略の侵入を許さないような戦略であるための条件を規定している。

われわれは有限回の繰り返しのもとにシミュレーションを行なうが、どの戦略もその繰り返しの長さを知らされていないため、ALL-D戦略が唯一のESSではない。また ε の大きさに対する依存性がでてくるものと期待される。実際のシミュレーションでは初期に比較的短い記憶(2ないし3)を持ったいくつかの戦略を用意し、時間とともに進化してくる戦略を調べる。

初期状態にかかわらず、通常ある1種類まれに2種類の戦略によって集団の90%以上が占有される状態が出現する。しかしこの状態は過渡的なものであり、いずれ新しく進化してきた戦略にとって代わられる。新しく進化してきた戦略はその時の集団で優勢だった戦略の突然変異種か、優勢戦略の自身との合体ないし優勢戦略と稀な変異戦略との合体によって生じたものである。この断続平衡状態がいくつか続いた後、優勢戦略が自分より「うまくやる」戦略を生み出せない最終状態に行き着いてしまう。この繰り返しゲームにはこのような進化の終点である準安定状態が無数にあると思われる。次ののは、記憶長3から始めた戦略群の進化プロセスの一例である。

戦略1と2はお互いのエラー（相手が誤ってCを出すところをDを出してしまう）を修復する機構をもっている。戦略1同志あるいは2同志にはその機構がないので、相手の存在によってのみ優勢な種でいられる。これは相互共生にある戦略であり、ジレンマゲームにおける相互共生ペアが他にも知られている。戦略3は初期にいた戦略であるが、この戦略は戦略1・2のどちらとも共生し（e.g.修復しあう）、かつ自己修復機構をもっている。この結果戦略3は優勢となるが、すぐに戦略1の変異種4によって駆逐される。戦略4は戦略3から周期的に搾取（Cに対しDを出す）する。戦略4の変異種5は戦略4に対しても5に対しても同様に高い得点を挙げるので、自分自身との対戦で高い得点を挙げられない戦略4は駆逐される。戦略6は戦略5と稀に存在していた戦略5の亜種と合体して戦略6をつくる。この時点で戦略の記憶長は6となる。しかしこの戦略はたちまち6の変異種7によって駆逐されてしまう。変異種7は記憶長6を持つが振るまいは殆どALL-Dである。この進化プロセスではこれが終点である（図2）。

この例では記憶は成長するが最終的に協調的振るまいは生成されなかった。次の例（図3）でも戦略としての記憶長は成長するが、最終的には非協力的な振るまいが残る。

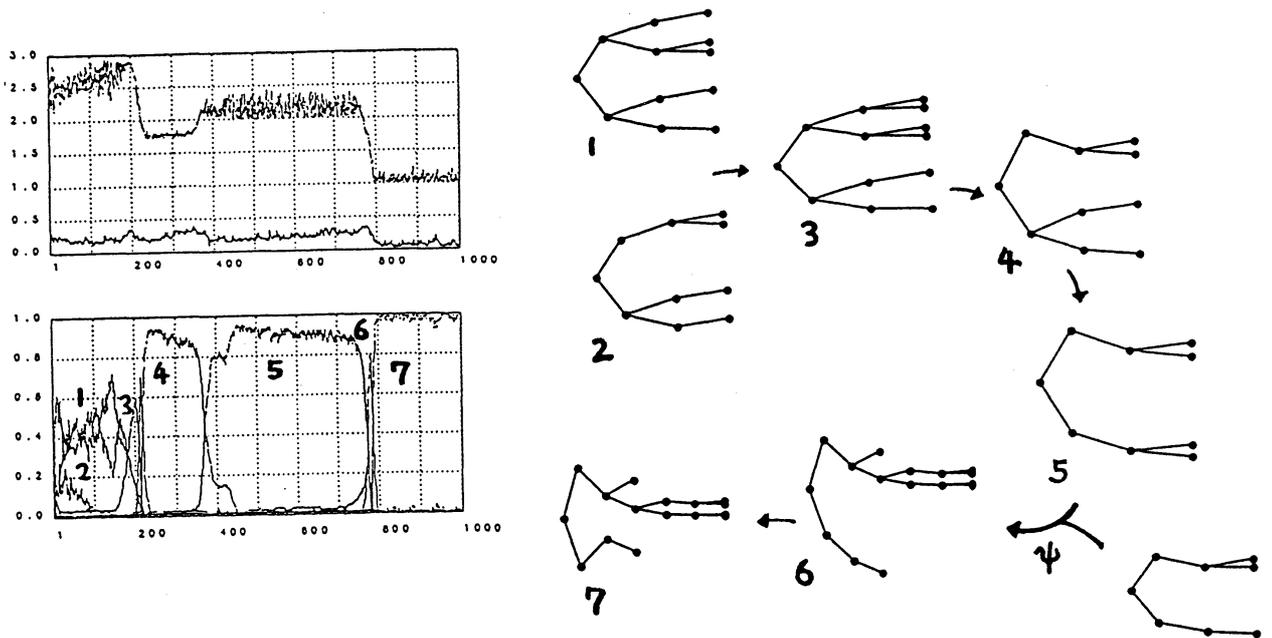


図2 記憶3からの進化の例。左上は平均の得点の時間発展。左下は戦略頻度の時間発展。戦略1,2共存から始まった進化過程は、最終的には自分自身との対戦でDを出しあう非協力的な戦略が進化してきた。

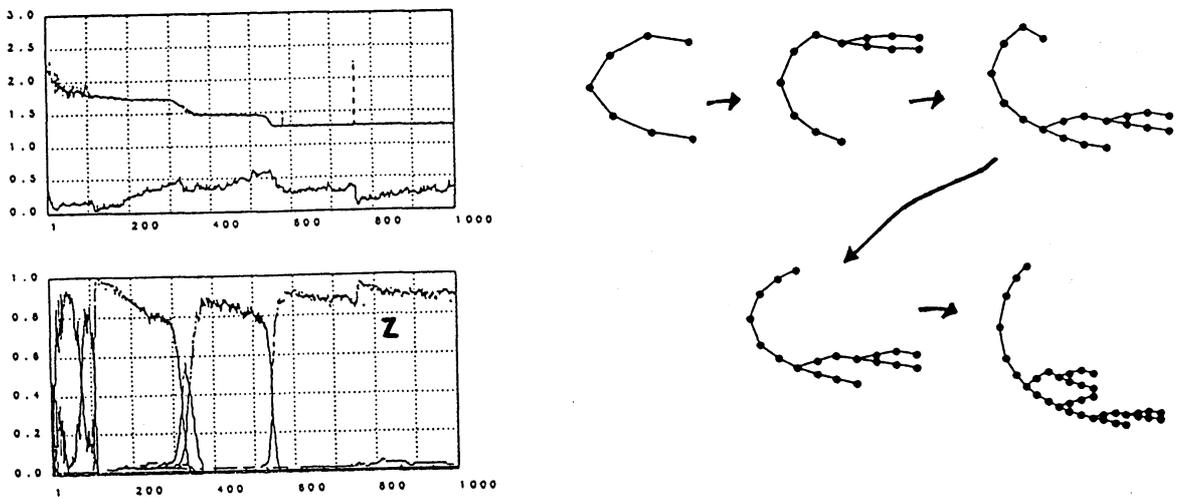


図3 記憶3からの進化の例。初期に置いた戦略が3回合体した後、1回変異したものがさらに重複合体して、記憶長12の戦略Zに進化する。

このように記憶の延長は相手を出し抜くために用いられ、自身との対戦においても非協力解を進化させてしまう。次の例ではある決った初期状態（しかし戦略として有力なものを含んだ）を用意し、この初期状態から生成される戦略の進化を観察する。

間違い確率 ϵ がゼロの時には初期状態に存在する T f T 戦略とその亜種が優勢となる。T f T 戦略から生成される亜種戦略の数は時間とともに増加する。この際記憶長の長い戦略も数多く出現するが、エラーがないため実効的に T f T 戦略

である。間違い確率を少しでもいれると多彩なT f T以外の戦略が進化してくる。

間違い確率が比較的小さいうち ($\varepsilon < 0.1$ (すなわち5回の対戦で1回どちらかがミスをする)) は、最終的に勝ち残る準安定な戦略として大体3つのパターン (A, B, C) がある。

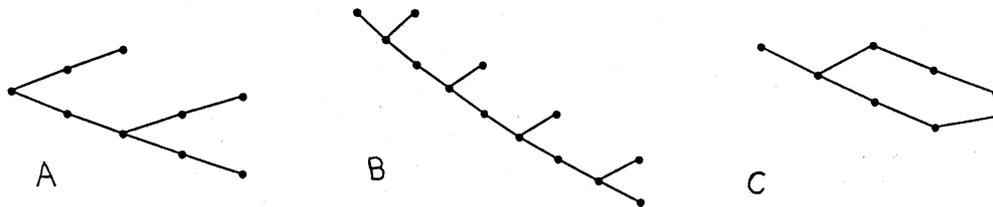


図4 ある初期状態 (記憶長2) の戦略群から進化してくる戦略は間違い確率 ε が小さいうちは、これらの3つのパターンに分かれた。

初期の集団から『前回に相手がCを出し自分はDを出した時はDを出し、それ以外はCを出す』という戦略xが優勢となる。この戦略が『前回に両方ともCを出すか、Dを出していればCを出す』戦略にとって代わられて、その戦略が重複合体して戦略Aが出現する場合と、戦略xからあまのじゃく戦略 (相手の前回の手と逆の手を出す) が進化してこれが重複合体して戦略Bをつくる場合がある。戦略Cはあまのじゃくの1回合体した戦略B (記憶長4) からの変異として観測される。このうち戦略パターンAはリングレンによって報告されている最良記憶戦略のひとつである。この戦略は高い修復機構を持ち自身との対戦で互いにCをだすのが定常状態で、どちらかがミスすると2回 (D, D) を繰り返したのち定常状態 (C, C) に復帰する。この戦略の記憶長は4である。一方戦略BとCは相互協調型の戦略ではなく、エラーを利用して階級分化 (搾取する側とされる側) をつくる戦略である。これらはあまのじゃく戦略の多重重複合体あるいはその変異種として出現したものである。この戦略が階級分化を引き起こす理由もここにある。自身との対戦においてどちらかがミスするまでは (D, D) が定常状態で、一方がいったん間違っってCを出すとそのプレイヤーは以後ずっとCを出し続け、もう一方はDを出し続けるからである。

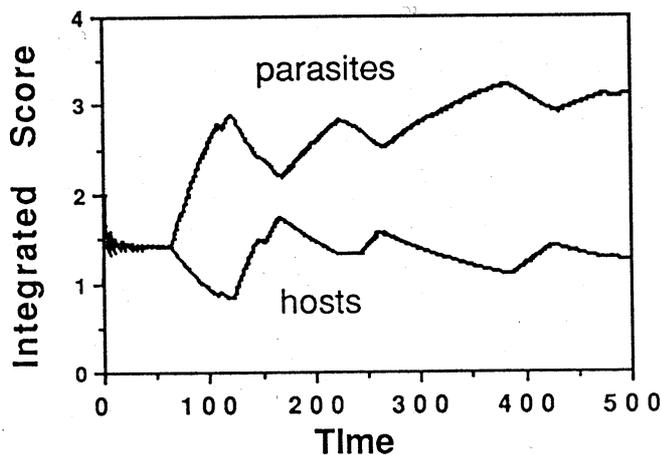


図5 戦略Bに関し、横軸をゲームの繰り返し回数とした場合の積算得点。初期の間違いが、同一戦略をとる二人のプレイヤーの得点を分岐させた。

このあまのじゃくタイプは他の戦略に対しても誤りでCを出し続ける危険性がある。しかし戦略B(記憶長 k)は k 回Dが続いたらCを出すというパターンを持っているために、自己重複合体戦略B(記憶長 $2k$)に確実に $k/2$ ステップ後にCを出すこととなり、この時戦略B($2k$)はDを出すので以後B(k)はCを出し続ける。他の戦略に間違っCを出してしまう前に、合体戦略B($2k$)が大勝して集団を占めてしまう。以後この繰り返しで、戦略Bが変異により自己合体できなくなるまで記憶長は大きくなる。

間違い確率 ε が大きくなると ε が小さい時に生じた進化は促進されなくなり、 ε の値の増加とともにあまのじゃく戦略、さらに戦略xで進化は止るようになる。T f T戦略も再び勝ち残ってくる。これは長い記憶をもつことの有利さが ε の増加とともに失われたからで、これもアイゲンのいうInformation Crisisの一例である。

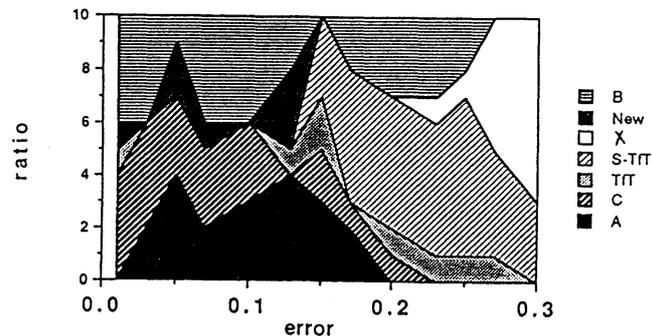


図6 ε と戦略の出現頻度。各 ε 値ごとに10回乱数列を変えて出現頻度を調べる。

□4 考察(進化と絶滅の可能性)

繰り返しジレンマ・ゲームでは、任意の戦略に対して大きく負けない万能純粋戦略、かつ自身と協調的にふるまうものとして有名なのはT f Tである。T f T戦略の記憶長はここでの表記で2である。リングレンによって見出された戦略(ここでは戦略A)は間違い確率のあるジレンマ・ゲームにおけるT f Tである。戦略Aの記憶長は4である。これらの戦略の特徴はアクセルロッドの言葉を借りれば、①礼儀正しさ②短気③寛容性を持っていることである。間違い確率 ε の存在下では自分自身と協調的に対戦している際に誤ってDを出してしまうことがある。この場合にもとの協調状態に復活させるような記憶パターンを備えている戦略は進化的に安定である。このような自己修復機構あるいはFault Tolerancyを戦略Aは備えている。ALL-C戦略はもちろん100%のFault Tolerancyを誇るが、これは非協力的な戦略に対しては無力である。Fault Tolerancyは寛容性に識別機構(敵か味方か)を加えたものだといえる。戦略Aが記憶長4でFault Tolerancyを獲得しているとしたら、記憶はどういう場合に進化するのだろうか。それがここでの問題であった。

結果からいえることは、記憶というものは戦略としての万能性に相反する性質であり、対戦相手の特殊性によって進化させられるものだという事である。ここでみられた長い記憶(>4)をもつに至った戦略は、自分自身との対戦でDを出し合う戦略Zのようなものか、戦略Bのように一方的な搾取関係を形成する場合であった。つねに一定頻度でT f TやALL-Dが存在するような集団ならばこれらの戦略は記憶長を長くして進化することはなかったと考えられる。戦略Aが非

特異的な万能戦略であるのに対し、戦略BやZは集団を占有している1種類の戦略に対してのみ高得点をあげることのできる特異戦略であった。かつ重複合体によりその戦略自身から進化してくるという状況が、特異な記憶パターンを進化させることになる。これは生物進化に見られるランナウェイ進化とよばれるもので、他の全く違う型の戦略や万能型の戦略の出現には弱く絶滅しやすい。一方で特異的な記憶パターンを進化させられる戦略が代々その集団を占有していくこととなる。

戦略の進化における万能性と特異性は、コンピューターにおける汎用性と特殊性と同じくトレード・オフの関係にある。しかし現実の生態系や社会にあてはめて見た場合、個々では特異的な進化を行ない、それが全体でみるといろいろな事態に対応できる万能なネットワークを構成しているはずである。これが共生ネットワークと呼ばれるものの一形態であろう。生物のもつ偶発性・発展性とその不安定なダイナミクスが構成する安定なネットワークとをつなげる研究は始まったばかりである[8]。

References

- [1] See e.g., R.E.Lenski, "Are Some Mutations Directed?", TREE 4 (1989)148; J.Marx, "A Mitey Theory for Gene Jumping", Science 253(1991)1092.
- [2] S.Ohno, *Evolution by Gene Duplication* (Springer-Verlag, 1970).
- [3] T.Ikegami and K.Kaneko, "Genetic Fusion", Phys. Rev. Lett. 65 (1990) 3352.
- [4] T.Ikegami, "The evolution of strategies and altruistic behavior in the repeated Prisoner's Dilemma game" (preprint).
- [5] See, e.g., R.Axelrod, *The Evolution of Cooperation*(Basic Books, NewYork 1984).
- [6] See e.g, J.Hofbauer aand K.Sigmund, *The Theory of Evolution and Dynamical Systems* (Cambrigid University Press, 1988).
- [7] U.Mueller, "Optimal Retaliation for Optimal Cooperation", J. Conflict Resolution 31 (1988) 692; K.Lindgren, "Evolution in a Population of Mutating Strategies" in *Artificial Life II* pp295. (eds. C.Langton et al., Addison Wesley, 1991).
- [8] K.Kaneko and T.Ikegami, "Homeochaos: Dynamical stability of Symbiotic Network with Population Dynamics and Evolving Mutation Rates", (to appear in Physica D).