

# バクテリア型進化的計算における分割幅導出法

山本 康高<sup>†</sup> 吉川 大弘<sup>†</sup> 古橋 武<sup>†</sup>  
篠木 剛<sup>†</sup> 鶴岡 信治<sup>†</sup>

近年、生物の進化過程を模倣した進化アルゴリズム (Evolutionary Algorithm: EA) がさかんに研究されている。古橋らは、バクテリアの進化メカニズムにヒントを得たバクテリア EA (Bacterial Evolutionary Algorithm) を提案し、EA の持つ局所探索能力の低さという問題点を改善している。このアルゴリズムは、分割された染色体の一部分に対し、集中的に突然変異演算を施すことで、染色体の局所部分の改善を促進するものであり、通常の突然変異演算と比較して有効な探索が行えることが報告されている。しかしこれまで、バクテリア EA に対する理論的な検証は行われておらず、シミュレーションなどを行うことでしかその効果を表現することができなかった。また、バクテリア EA の課題として、染色体分割の際、適切な幅が試行錯誤を通してでしか決められないという問題があった。そこで本論文では、NK-Landscape 問題を例題として、バクテリア EA に対する探索性能を、遺伝子間の相互関係の強さ (エピスタシス) や突然変異率などのパラメータを変数として用いることで、理論的に定式化する。本理論式により、試行錯誤的なシミュレーションを行うことなく、分割幅の最適化およびバクテリア EA の有用性の検討が可能となることを示す。

## Derivation Method of Division Width for Bacterial Evolutionary Algorithm

KOSUKE YAMAMOTO,<sup>†</sup> TOMOHIRO YOSHIKAWA,<sup>†</sup>  
TAKESHI FURUHASHI,<sup>†</sup> TSUYOSHI SHINOGI<sup>†</sup> and SHINJI TSURUOKA<sup>†</sup>

Recently, Evolutionary Algorithms (EAs), which are based on the biological evolutionary process, are widely studied in a lot of fields. However, their probabilistic behavior makes the theoretical investigation difficult. As a result, the search abilities of EAs have been tested through simulations in most of the reports. The parameters of genetic operations have also been decided experientially or by trial and error. This paper mathematically analyzes the effects of bacterial mutation of Bacterial Evolutionary Algorithm (BEA). NK-Landscape Problem is employed for the investigation of this analysis in this paper. The search ability of bacterial mutation is formulated with the parameters for the operations and the interactions between genes called "Epistasis". The optimization of the division width, which is one of the most important parameters in BEA, and comparative study with conventional mutation operations are made possible by this formula. It is shown that the bacterial mutation surpasses conventional ones in search performance.

### 1. はじめに

近年、生物の進化過程を模倣した進化アルゴリズム (Evolutionary Algorithm: EA) がさかんに研究されている<sup>1)~3)</sup>。EA には EP (Evolutionary Programming), ES (Evolutionary Strategy), GA (Genetic Algorithm) などがあり、問題の解候補をコーディングした染色体に対し、遺伝的演算を施すことで、準最適解を実用的な時間内で獲得することを目的とした手法である。しかし EA は、大局的な探索能力は高いも

の、局所探索能力が低いという問題点がある。これは、長い染色体に対し、一様な確率で遺伝的演算を施すことにより、局所部分の改善が必ずしも効率良くはなされないことに起因している

これに対し、バクテリアの進化メカニズムにヒントを得たバクテリア EA (Bacterial Evolutionary Algorithm) が、古橋らによって提案されている<sup>4)~6)</sup>。このアルゴリズムは、染色体を複数の部分に分割し、それらの部分に対し集中的に突然変異演算を施すことで、染色体の局所部分の改善を促進するものである。バクテリア EA における突然変異 (バクテリア演算) は、EA における通常の突然変異と比較して、有効な探索が行えることが報告されている。しかしこれらの研究

<sup>†</sup> 三重大学工学部

Faculty of Engineering, Mie University

では、その確率的挙動のため、手法に対する理論的な検証ができず、シミュレーションなどを行うことでしかその効果を表現することができなかつた。また、バクテリア EA の課題として、染色体分割の際の適切な幅を、経験的に、もしくは試行錯誤を通してしか決められないという問題がある。

そこで本研究では、バクテリア演算で用いられるパラメータを変数とし、バクテリア EA に対する探索能力を定式化することにより、バクテリア EA の課題である有効な分割幅を、理論的に導出することを試みる。バクテリア EA においては、染色体の部分情報の伝達方法として、GTO (Gene Transfer Operation) を用いており、これは、EA における交叉と同様に、大局的な探索を行ううえで有効に働くことが知られている。しかし、GTO により、染色体情報が大きく変化するため、この演算に対する探索性能を計ることは困難であると考えられる。そこで本論文では、上述の目的に対する基礎的段階として、GTO を用いないバクテリア EA、すなわち、バクテリア突然変異の探索性能についての検討を行う。また、このバクテリア EA では、複数の染色体を持つ GA などと異なり、染色体が 1 つであることから、理論的検証が比較的容易になると考えられる。

一方、染色体中での遺伝子間の相互関係が EA の探索能力に影響を与えるという研究が報告されている<sup>7)</sup>。この遺伝子間の相互関係の強さはエピスタシスと呼ばれる。エピスタシスは、適切な分割幅決定を困難にしている要因の 1 つであると考えられ、探索能力の定式化を行ううえで考慮されなければならないが、通常の問題においては、そのエピスタシスを算出することは困難であるため、本研究では、遺伝子間の相関パターンを設計者により設定可能な NK-Landscape 問題<sup>8)</sup>を例題として用い、NK-Landscape 問題における探索能力の定式化を試みる。

本論文では、導出された理論式に対する妥当性についての検討を行い、本理論式により、試行錯誤的なシミュレーションを行うことなく、適切な分割幅の決定、および手法の有用性の検討が可能となることを示す。さらに、NK-Landscape 問題に対する山登り法とバクテリア EA の解探索手法としての性能比較についても、本理論式により評価が可能となることを示す。

## 2. バクテリア EA

バクテリア EA は、バクテリアの進化メカニズムを模倣したアルゴリズムであり、染色体を複数の部分に分割し、それら分割された部分に対して集中的に突然

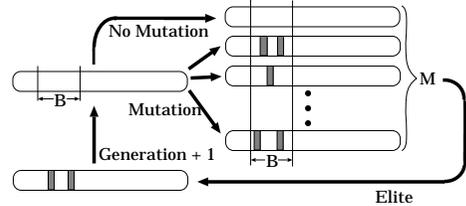


図 1 バクテリア EA の概略図  
Fig. 1 Bacterial mutation.

変異演算を施すことで、染色体の局所改善を可能としている。バクテリア EA の概略図を図 1 に示す。バクテリア EA では、初めに染色体を  $M$  個に増殖させる。その際、増殖した  $M - 1$  個の各染色体において、分割幅  $B$  内の各遺伝子に対して確率  $P_m$  で突然変異を施す。各染色体を評価し、 $M$  個の染色体から選択されたエリートを元の染色体と置き換える。この一連の動作を 1 世代とする。こうして新たに生成された染色体に対し、局所部分の突然変異とエリート選択を繰り返す。ただし、突然変異を加える分割幅  $B$  の位置は、各世代ごとにランダムに選ぶものとする。

## 3. NK-Landscape 問題

NK-Landscape 問題では、染色体の各遺伝子は、それぞれ他の遺伝子の持つ値との組合せにより得点が決まる。通常、この遺伝子どうしの関わりの強さをエピスタシスと呼ぶ。NK-Landscape 問題では、このエピスタシス ( $K$ ) をパラメータとして持っているため、設計者が  $K$  の値を設定することによりエピスタシスを任意に決定できるという特徴がある。ただしここで遺伝子どうしの関わりの強さは、各遺伝子の得点が決まる遺伝子の組合せの数、すなわち 1 つの遺伝子の得点が影響を受ける遺伝子の数 ( $K$ ) で定義される。本論文では、染色体長を  $N$  とし、各遺伝子の得点は、その遺伝子を含め右側に連なる  $K$  個の遺伝子との組合せで決まることとする。また、その得点は、各遺伝子につき、 $2^K$  の組合せすべてに対する得点テーブルをあらかじめランダムに設定する。染色体の評価値は、各遺伝子の得点の総和とし、染色体は、環状になっているものとする。すなわち、染色体の最後尾の遺伝子の右には最初の遺伝子があるとし、得点の計算を行う。図 2 に、具体例として  $K=5$  としたときの評価値の計算方法を示す。図の一番左の遺伝子は、それを含む  $K$  個 (図の例では 5 個) の遺伝子群 01011 によって得点が決まる。あらかじめ作成しておいた遺伝子群と得点の対応表を用いることで、一番左の遺伝子の得点 (513) が決定される。この操作を染色体中のす

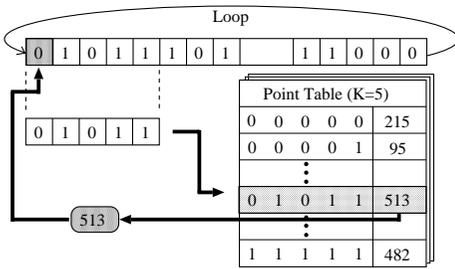


図 2 NK-Landscape 問題  
Fig. 2 NK-Landscape problem.

すべての遺伝子 ( $N$  個) に行い、それらの得点の合計を染色体の評価値として用いる。

4. NSS 式

NK-Landscape 問題にバクテリア EA を適用する際、各パラメータが探索に与える影響について考察する。バクテリア EA による分割幅  $B$  が大きい場合、探索空間が非常に大きくなることで、局所探索の能力が落ち、逆に  $B$  が小さい場合は探索空間は小さくなるが、局所解に陥りやすくなると考えられる。つまり、有効な探索を行うことのできる適切な  $B$  が存在することになる。また、これ以外にも、増殖数  $M$  が大きい場合、一度に分割幅  $B$  内の最適化が進むため、局所的な探索は進むものの、それらが他の遺伝子に与える影響により、染色体全体としての最適化が困難になると考えられる。また、分割幅  $B$  が小さいときなどは、 $M$  を大きくしても、増殖した染色体が同一の染色体になる確率が高くなることにより、無駄な探索を増加させてしまう。一般に、世代数は大きい方が、より多くの解を探索することが可能となるが、特に評価に時間のかかる問題設定などにおいては、時間的制約などにより、求める評価値が得られるまでの世代数や、評価値が飽和するのに必要な世代数を求めることができれば、世代数を設定するうえで有効であると考えられる。また、突然変異率  $P_m$  や探索空間の大きさに影響を及ぼす染色体長  $N$  についても、探索能力に影響を与える原因となると考えられる。さらに、これらの値を一定にしたとしても、適用する問題により、分割幅  $B$  などの適切なパラメータ値が異なることが経験上知られている。これは、問題によって染色体中の 1 つの遺伝子の変化が、他の遺伝子群に影響を与える強さ (エピスタシス:  $K$ ) が異なるためだと考えられる。したがって、エピスタシス  $K$  も、探索に影響を与える一要因となる。

4.1 探索可能解個数

本論文では、解の探索能力を示す値として、探索可

能解個数 (Number of Searchable Solutions:  $NSS$ ) を定義し、その定式化を試みる。解の探索能力を決定する指標としては、前述した分割幅  $B$ 、増殖数  $M$ 、世代数  $G$ 、染色体長  $N$ 、突然変異率  $P_m$ 、エピスタシス  $K$  といったパラメータが関係してくると思われる。そこで、これら 6 つの変数を用いて  $NSS$  を定式化することで、バクテリア EA の探索能力を定量化する。以下に  $NSS$  の概要について述べる。

初めに、1 回の試行における解探索の回数を考える。その回数は増殖数  $M$  と世代数  $G$  を用いることにより  $\{(M - 1) \times G + 1\}$  で求めることができる。しかしこの回数は、同じ解を重複して探索することや、すでに改善された遺伝子を変化させようとする探索の無駄、改善された遺伝子から影響を受ける遺伝子が改善しにくくなるといった、探索能力の減衰要素が考慮されていないため、実際の探索可能な解個数とは一致しない。それらの要素を考慮し、有効な探索解個数を  $NSS$  として定義する。本論文では、探索能力に影響を与えると考えられるすべてのパラメータ、分割幅  $B$ 、増殖数  $M$ 、世代数  $G$ 、染色体長  $N$ 、突然変異率  $P_m$ 、エピスタシス  $K$  を変数とすることにより、探索可能解個数  $NSS$  を理論的に定式化する。

探索可能解個数  $NSS$  は、1 世代あたりに探索可能な解の個数 ( $\nu$ ) と、第  $n$  世代における探索能力の減衰率 ( $\alpha_n$ ) との積を、各世代における探索個数として、世代について和をとることにより、式 (1) のように求める。本論文では、探索能力の減衰要因として、3 つの減衰率 ( $\gamma_i, \gamma_p, \gamma_g$ : それぞれ 4.1.2, 4.1.3, 4.1.4 項で説明) を用いる。これらの減衰率を用い、式 (2) により世代  $n$  における探索能力の減衰率 ( $\alpha_n$ : 導出方法は 4.1.5 項で説明) を求める。

本論文では、上述の  $NSS$  の値、すなわち、探索可能な解個数が大きいほど、有効な解を獲得できる確率が高いと仮定する。実際には、解空間の形状により、局所解への陥りやすさなどが異なり、探索した解の個数と獲得された解の有効性とは、単純な比例関係が成り立つとは限らないが、一般的に解空間の形状は未知であるため、ここではこれらの間には線形性が成り立つものとする。すなわち、分割幅  $B$  などのパラメータを用いて  $NSS$  を定量化することにより、最も高い  $NSS$  を与える分割幅を求めることで、実際にバクテリア EA を用いる際の適切な分割幅に対する指針を得ることが可能となる。

$$NSS = \sum_{j=0}^{G-1} \nu \alpha_j \tag{1}$$

$$\alpha_0 = 1$$

$$\alpha_{n+1} = \{1 - (\gamma_i + \gamma_p)\varepsilon_n\}(1 - \gamma_g)\alpha_n \quad (2)$$

以下に、式 (1) で用いられている各変数および導出法について説明する。また、各項で導出された式については、本論文末尾に式一覧を付録として掲載する。

4.1.1 有効増殖数  $\nu$

バクテリア EA では、染色体は  $M - 1$  個の新しい染色体を生成するが、確率上、増殖の際に、同じ染色体 (部分列) への増殖、すなわち染色体どうしの重複が起こるため、実際には  $M - 1$  個の解探索を行っていないことがある。そのことを考慮し、実質的な新しい染色体としての増殖数  $\nu$  を求める。まず、分割幅  $B$  内に突然変異が起こらない確率  $p$  を求める (式 (3-1))。

$$p = (1 - p_m)^B \quad (3-1)$$

次に増殖により生成される新しい染色体がそれまでの増殖により得られた染色体と同一の染色体となる確率  $q$  を求める (式 (3-2))。

$$q = \sum_{l=0}^B {}_B C_l \{(1 - p_m)^{B-l} p_m^l\}^2 \quad (3-2)$$

これらの  $p, q$  を用い  $\nu$  を導出する。1 個目の増殖は、元の染色体に対して、1 つの遺伝子でも突然変異が行われることが新しい染色体を生成する条件となり、その確率は  $1 - p$  である。さらに次の染色体は、元の染色体とは異なり、かつ 1 個目の増殖により生成された染色体とも異なる必要がある。その確率は  $(1 - p)(1 - q)$  により与えられる。次からの染色体の増殖に対しても同様の考え方を適用することにより、生成される染色体がそれまでに増殖された染色体と同一の染色体にならない確率を求める。これらの確率は、分割幅  $B$  内での突然変異により、新しい染色体として生成される確率であるため、その確率を実質的な増殖個数と見なし、 $M - 1$  個分の各確率の和を求めることにより、有効増殖数  $\nu$  を得る (式 (3))。

$$\nu = \sum_{l=0}^{M-2} (1 - p)(1 - q)^l \quad (3)$$

4.1.2 影響遺伝子による能力減衰率  $\gamma_i$

バクテリア EA では、分割幅  $B$  内で局所的な探索を行うため、染色体中の部分的な最適化が行われる。このとき、影響を与える (受ける) 遺伝子情報をもとに最適化が行われる。したがって、その後染色体上においてその近傍に分割幅  $B$  が選択された場合、それらの影響を受ける遺伝子を変えることは、先の最適化に用いられた情報を壊すこととなり、結果として評価値が下がることにつながる。これは、ある解の探索 (改

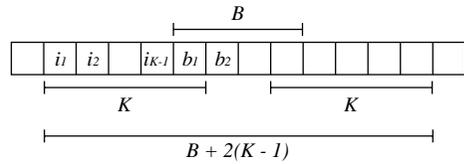


図 3 影響遺伝子数  $\lambda$  (分割幅の外側)

Fig. 3 Number of influenced genes  $\lambda$  (outside of division width).

善) により最適化された遺伝子が、その後の探索に影響を与えていると考えることができる。ここでは、この影響を近似的に 2 つの要素に分けて導出する。すなわち、(i) 分割幅  $B$  内の变化により、何個の遺伝子に影響を与えるか、(ii) 分割幅  $B$  をランダムに決定したとき、(i) の遺伝子が何個  $B$  内に存在し、その影響により、探索がどの程度困難になるかである。(i) は、NK-Landscape 問題におけるエピスタシスが  $K$  個であることを考慮すると、 $B + 2(K - 1)$  (図 3) で近似することが可能であるが、この式は特に分割幅  $B$  が大きいときに誤差が生じる。本論文では、(i)  $\lambda$  を求め、その後、(ii)  $\lambda$  を用いることで上述の影響による減衰率  $\gamma_i$  を求める。

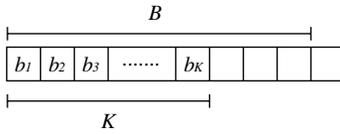
4.1.2.1 影響遺伝子数  $\lambda$

初めに、 $\lambda$  の導出法について述べる。まず、 $B$  の外側の影響遺伝子数について導出する。影響遺伝子数は分割幅  $B$  内の突然変異箇所に依存する。たとえば、図 3 中の  $i_1$  は、分割幅  $B$  内における  $b_1$  が変更される場合のみ影響を受ける。つまり、 $i_1$  が影響を受ける確率は、 $b_1$  が突然変異を受ける確率に等しい。このとき、 $i_1$  は本式では  $p_m$  個の影響遺伝子であると定義する。また、 $i_2$  は  $b_1, b_2$  のどちらか一方、または両方が変更された場合に影響を受ける。そのときの影響遺伝子数は先ほどの定義により、 $1 - (1 - p_m)^2$  と求められる。同様に分割幅の両側についてこれらの確率を求め、それらを足し合わせるにより、 $B$  の外側の影響遺伝子数を求める (式 (4-1))。ただし、式 (4-1) において、式中の加算範囲や指数部が複雑になっているのは、 $K$  や  $B$  が染色体  $N$  に対して大きい場合に、分割幅の外側の遺伝子数が  $K - 1$  に満たない場合や、環状に構成されている染色体を 1 周して、逆側の分割幅内の遺伝子からの影響を受けることを考慮しているためである。なお、 $C_1, C_2$  については本項で後述する。

$$2 \sum_{l=\max(1, K-C_2)}^{K-1} \{1 - (1 - p_m)^{E_1}\} \quad (4-1)$$

$$E_1: \min(l + \max(0, K - (C_1 + C_2 + (l - \max(1, K - C_2)))) - 1, B)$$

また、式 (4-2) は分割幅  $B$  以外の遺伝子数が奇数

図4 影響遺伝子数  $\lambda$  (分割幅内)Fig. 4 Number of influenced genes  $\lambda$  (inside of division width).

の場合を考慮した式であり，導出の考え方は式 (4-1) と同様である．

$$C_1 \{1 - (1 - p_m)^{\min(\max(0, 2(K - C_2 - 1)), B)}\} \quad (4-2)$$

次に，分割幅  $B$  内の影響遺伝子について考える．その遺伝子そのものが突然変異を受けた場合に与える影響については，4.1.3 項に述べる  $\gamma_p$  において考慮しているため，ここではその遺伝子は突然変異を受けない場合に，周りの遺伝子の変化により生じる影響遺伝子数を求める．図4の  $b_1$  は， $b_2 \sim b_K$  までのいずれかの遺伝子に対して突然変異が加えられた場合，次世代において影響遺伝子となる． $b_2 \sim b_K$  の遺伝子に対して，突然変異が加えられる確率は  $1 - (1 - p_m)^{K-1}$  によって与えられる．また，上述のように，自らの遺伝子（図4中の  $b_1$ ）の変更は影響遺伝子とは数えないため，上の式に自らの遺伝子は変更されない確率を乗算した  $(1 - p_m)\{1 - (1 - p_m)^{K-1}\}$  が  $b_1$  における影響遺伝子数となる．このようにして  $B$  内のすべての遺伝子に対し，影響を受ける確率を求め，それらの和をとる（式 (4-3)）ことにより， $B$  内の影響遺伝子数を求める．ここで，式 (4-3) において，加算範囲や指数部が複雑になっているのは， $B$  が大きい場合や， $B$  よりも  $K$  が大きい場合などに生じる分割幅の逆側の遺伝子からの影響や，影響遺伝子数の重複分を考慮しているためである．また，確率上の等価性から， $B$  内を左右対称に2分割して式の簡略化を行っている．式 (4-3) が2倍されているのはそのためである．

$$2 \sum_{l=1}^{\frac{B-C_0}{2}-1} (1 - p_m) \{1 - (1 - p_m)^{E_2}\} \quad (4-3)$$

$$E_2: \min\{K + \min(l - 1 + \max(0, K - (N - B + l - 1) - 1), K - 1), B\} - 1$$

さらに，分割幅  $B$  内の遺伝子数が奇数である場合の中心の遺伝子分を式 (4-4) により求める．

$$C_0 (1 - p_m) \{1 - (1 - p_m)^{\min((2K-1), B)-1}\} \quad (4-4)$$

ここで  $\lambda$  導出時に用いた  $C_0, C_1$  は，それぞれ以下に示す  $C_0$  ( $B$  内の遺伝子数奇数：1，偶数：0)， $C_1$  ( $B$  の外側の遺伝子数 奇数：1，偶数：0) となる値であり， $C_2$  は式の複雑性をさけるために用いている．

$$C_0 = \frac{(-1)^{B+1} + 1}{2}$$

$$C_1 = \frac{(-1)^{N-B+1} + 1}{2}$$

$$C_2 = \frac{N - B - C_1}{2}$$

以上により求められた式 (4-1) ~ (4-4) の和をとることにより影響遺伝子数  $\lambda$  を得る（式 (4)）．

$$\lambda = 2 \sum_{l=\max(1, K-C_2)}^{K-1} \{1 - (1 - p_m)^{E_1}\} + C_1 \{1 - (1 - p_m)^{\min(\max(0, 2(K - C_2 - 1)), B)}\} + 2 \sum_{l=1}^{\frac{B-C_0}{2}-1} (1 - p_m) \{1 - (1 - p_m)^{E_2}\} + C_0 (1 - p_m) \{1 - (1 - p_m)^{\min((2K-1), B)-1}\} \quad (4)$$

#### 4.1.2.2 影響遺伝子による能力減衰率 $\gamma_i$

次に  $\lambda$  を用いた  $\gamma_i$  の導出法について述べる．影響遺伝子による探索能力の減衰は，次世代においての分割幅  $B$  が  $\lambda$  内に含まれる，もしくは1部分が重なる場合に生じる．まず， $B$  が完全に  $\lambda$  に含まれる（または  $\lambda$  が  $B$  に含まれる）場合について求める． $B$  が  $\lambda$  に完全に含まれる場合の数は， $\lambda - B + 1$  通りとなり，そのような位置に  $B$  が選択される確率は  $(\lambda - B + 1)/N$  となる．このとき， $B$  内の遺伝子に突然変異が加えられる確率は  $1 - (1 - p_m)^B$  となる．本項で述べたように，前世代においては，影響遺伝子の情報を基にして解の改善（遺伝子の変化）がなされていると考えられるため，この場合において影響遺伝子を変化させることは，前の改善と同等以上の解の改善を行わない限り，解の劣化，すなわち有効な探索とはならないと考えられる．ここでは，前の改善，すなわち，元の染色体と増殖された染色体  $1 + \nu$  個の染色体中でのエリート染色体と，同等以上の改善となる確率を  $1/(1 + \nu)$  で表す．したがって，影響遺伝子による探索能力の減衰は  $\nu/(\nu + 1)$  となり， $B$  が  $\lambda$  に完全に含まれる場合の，探索能力減衰率は式 (5-1) により求められる．ただし，式 (5-1) では， $K$  が小さい場合や  $B$  が大きい場合などに生じる  $\lambda$  が  $B$  よりも小さくなる場合を考慮して定式化を行っている．

$$\frac{1}{N} (|\lambda - B| + 1) \{1 - (1 - p_m)^{\min(\lambda, B)}\} \frac{\nu}{\nu + 1} \quad (5-1)$$

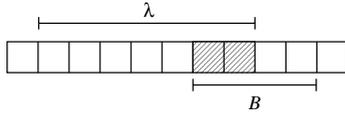


図5 影響遺伝子による能力減衰率  $\gamma_i$

Fig. 5 Decrement rate of search ability  $\gamma_i$  by influenced genes.

次に、選択された  $B$  の一部分が  $\lambda$  に重なっている場合について述べる。本論文では、 $\lambda$  の外側では探索能力の減衰は生じないとしているため、図5の斜線部に突然変異が加えられた場合に探索能力が減衰する。斜線部分が  $1 \sim B-1$  個まで、それぞれ  $1/N$  で  $B$  が選択された場合、斜線部に突然変異が起こる確率は  $1 - (1 - p_m)^l$  ( $l: B$  と  $\lambda$  の重なり遺伝子数) となる。これにより、式(5-2)のようにそれぞれの確率の和を計算し、上述の改善確率を掛け合わせることで、探索能力の減衰率を求める。また式(5-2)は、上述と同様に  $\lambda$  と  $B$  の大小関係を考慮し、 $\lambda$  の両側分について求めている。

$$\frac{2}{N} \sum_{l=1}^{\min(\lambda, B)-1} \{1 - (1 - p_m)^l\} \frac{\nu}{\nu+1} \quad (5-2)$$

さらに、式(5-2)においては、染色体長  $N$  に対し  $\lambda$  が大きい場合、 $B$  が  $\lambda$  の両側にまたがっていることで重複して求めている確率分が生じるため、式(5-3)に示す重複分を差し引くことで、式(5)の影響遺伝子による探索能力減衰率  $\gamma_i$  を得る。

$$\frac{1}{N} \sum_{l=1}^{\max(0, E_3-1)} \{1 - q_m^l\} \{1 - q_m^{E_3-l}\} \frac{\nu}{\nu+1} \quad (5-3)$$

$$q_m = 1 - p_m$$

$$E_3: \min(\lambda, B) - (N - \max(\lambda, B))$$

$$\begin{aligned} \gamma_i = & \frac{1}{N} \frac{\nu}{\nu+1} \left[ (|\lambda - B| + 1) \{1 - (1 - p_m)^{\min(\lambda, B)}\} \right. \\ & + 2 \sum_{l=1}^{\min(\lambda, B)-1} \{1 - (1 - p_m)^l\} \\ & \left. - \sum_{l=1}^{\max(0, E_3-1)} \{1 - (1 - p_m)^l\} \{1 - (1 - p_m)^{E_3-l}\} \right] \quad (5) \end{aligned}$$

#### 4.1.3 同一遺伝子変更による能力減衰率 $\gamma_p$

ある世代において、バクテリア演算によりある遺伝子が増殖されたとき、その遺伝子は、増殖された染色体の中でのエリートとして改善されたことになる。したがって、その遺伝子そのものを再度変更することは、

探索の無駄になると考えられる。そこで、それらの遺伝子を再度変更する確率を用い、探索能力が  $\gamma_p$  だけ減衰していくとする。ここで  $\gamma_p$  の導出方法について述べる。バクテリア EA において前世代と同一の遺伝子が増えられたのは、その遺伝子を含む分割幅  $B$  が選択され、かつ特定の遺伝子に突然変異が加えられた場合である。分割幅  $B$  が前世代と完全に一致する確率は、染色体長  $N$  を用いると  $1/N$  となる。その中で同一の遺伝子に突然変異が起こらない確率は  $(1 - p_m^2)^B$  で与えられ、これらにより、式(6-1)が導出される。

$$\frac{1}{N} \{(1 - p_m^2)^B\} \quad (6-1)$$

また、分割幅  $B$  が前世代と重なりを持たない確率は  $(N - 2B + 1)/N$  で与えられ、この場合は、同一の遺伝子が増えられることはない。しかし、分割幅  $B$  の長さが染色体長  $N$  の半分より長い場合、必ず重なる部分が生じ、その場合は、その重なり部分に対し、同一の遺伝子に突然変異が起こらないという確率を用い、式(6-2)を得る。

$$\frac{1}{N} \{(|N - 2B| + 1)(1 - p_m^2)^{\max(2B - N, 0)}\} \quad (6-2)$$

さらに、前世代の分割幅  $B$  と一部分が重なる場合についても同様の求め方により、式(6-3)を導出する。加算範囲が複雑な形をしているのは、 $B$  が染色体長の半分より長い場合においては、式(6-2)により計算が行われていることにより、その重複を避けるためである。

$$\frac{1}{N} \{2 \sum_{l=\max(2B - N + 1, 1)}^{B-1} (1 - p_m^2)^l\} \quad (6-3)$$

以上により求められた式(6-1)~(6-3)の和を1から引くことにより、同一遺伝子変更による探索能力減衰率  $\gamma_p$  を得る(式(6))。

$$\begin{aligned} \gamma_p = & 1 - \frac{1}{N} \{(1 - p_m^2)^B \\ & + (|N - 2B| + 1)(1 - p_m^2)^{\max(2B - N, 0)} \\ & + 2 \sum_{l=\max(2B - N + 1, 1)}^{B-1} (1 - p_m^2)^l\} \quad (6) \end{aligned}$$

#### 4.1.4 世代間染色体重複による能力減衰率 $\gamma_g$

進化的計算では、解探索が進むにつれ、それまでに探索した解を重複して探索する確率が高くなる。そこで探索空間と、1世代に有効増殖数  $\nu$  個だけ解探索を行ったときの未探索空間との比により、世代ごとに、探索が重複しない確率が  $\gamma_g$  だけ減衰していくとし、この能力減衰率  $\gamma_g$  を求める。

ここで、探索空間について考える。通常、染色体長  $N$  であれば、その探索空間は  $2^N$  で表されるが、染

染色体を分割幅  $B$  で分割し、集中探索を行う場合、すべての解を一様に探索していないため、実際の探索可能な解空間（探索の行いやすさを考慮した解空間、以降、探索可能解空間）は異なると考えられる。探索可能解空間は突然変異率に依存する。例として、分割幅  $B$  が 3 の場合について考える。ここで  $p_m = 0.5$  の場合、すべての  $B$  内の部分列は等価に探索が行えるため、探索可能な新しい解は、元の染色体を差し引き  $2^B - 1$  により求められる。しかし、 $p_m = 1.0$  の場合では、新しく得られる解の種類としては、元の染色体内で選択された  $B$  の遺伝子がすべて反転した解しか得られない。つまり、これらは探索可能解空間が突然変異率により異なっていることを意味している。そこで本論文では、このことを考慮し、突然変異率の違いによる解の生成されやすさを考慮した解空間を求める。具体例として、 $B$  内の遺伝子が {000} である場合について考える。この分割幅内に  $p_m = 0.2$  で突然変異を与えた場合、新しく生成される解のパターンは、遺伝子が (i) 1 つ変化する場合 {001, 010, 001}, (ii) 2 つ変化する場合 {011, 101, 110}, (iii) 3 つとも変化する場合 {111} とに分けられ、それぞれの場合についての生起確率は  $0.2^2 \times 0.8$ ,  $0.2^1 \times 0.8^2$ ,  $0.2^3$  で与えられる。ここで本論文では、(i) の解は、(i)~(iii) までの生起確率中  $0.2^2 \times 0.8$  の確率で 3 種類の解候補を持つと定義する。すなわち、生起確率が探索の行いやすさを、それぞれの場合の数が探索可能解空間を表すとする。(ii), (iii) についても同様に求め、これを一般式で表すと次式のように表される。

$$\frac{\sum_{l=1}^B {}_B C_l p_m^l (1-p_m)^{B-l}}{\sum_{l=1}^B p_m^l (1-p_m)^{B-l}} \quad (7-1)$$

式 (7-1) により、上述の例の、 $B = 3$ ,  $p_m = 0.2$  のときの探索可能な解個数は約 2.9 個となる。 $B$  の選択箇所の場合は  $N$  であるため、これを式 (7-1) に乗算することにより、探索可能解空間を得る。また、探索可能解空間を用いることにより、有効増殖数  $\nu$  との比が求められ、世代間染色体重複による能力減衰率  $\gamma_g$  を得る (式 (7))。ただし、各世代ごとの染色体の重複は、 $\nu$  により考慮されている。

$$\gamma_g = \frac{\nu}{N \times \frac{\sum_{l=1}^B {}_B C_l p_m^l (1-p_m)^{B-l}}{\sum_{l=1}^B p_m^l (1-p_m)^{B-l}}} \quad (7)$$

#### 4.1.5 探索能力減衰率 $\alpha_n$ ・解改善確率 $\varepsilon_n$ 探索可能解個数 $NSS$

以上の減衰項を組み合わせることにより、第  $n$  世代目における探索能力の減衰率  $\alpha_n$  となると考えられる。

ここで、バクテリア演算による各世代の解改善確率について考える。初期世代については、 $\nu$  個の増殖により、元の染色体と合わせて  $\nu + 1$  個の新しい染色体が生成される。初期世代では、探索能力の減衰はなく、また、評価値に確率的な優劣はないため、初期世代でのバクテリア演算による解改善確率  $\varepsilon_0$  は  $\nu/(\nu + 1)$  となる。それ以降の世代では、世代を追うごとに解の改善が困難になっていくと考えられるため、減衰率  $\alpha_n$  を用いて  $\varepsilon_n$  を求めている (式 (8))。

$$\varepsilon_n = \frac{\nu \alpha_n}{\nu \alpha_n + 1} \quad (8)$$

4.1.2 項で述べた  $\gamma_i$  と 4.1.3 項で述べた  $\gamma_p$  については、前の世代で解が改善されたときのみ生じる減衰項であるため、これらの和に解改善確率  $\varepsilon_n$  を掛け合わせた減衰項と、 $\gamma_g$  による減衰項の積により、探索能力減衰率  $\alpha_n$  (初期世代では 1.0) の漸化式が求まり (式 (2))、各世代ごとの有効増殖数  $\nu$  との積を世代数だけ足し合わせることで、探索可能解個数  $NSS$  を得る。

#### 4.2 山登り法への適用

本  $NSS$  式を山登り法の探索性能評価に用いることについて考察する。山登り法では、バクテリア EA における増殖数  $M$  を 1、突然変異率  $P_m$  を 1.0、分割幅  $B$  を 1 としたものと同等であると考えられる。よって式 (1)~(8) にこれらの値を代入することにより、各変数は次のように求めることができる。

$$\begin{aligned} \nu &= 1 & \gamma_p &= \frac{1}{N} \\ \lambda &= 2K - 2 & \gamma_g &= \frac{1}{N} \\ \gamma_i &= \frac{1}{2} \frac{\lambda}{N} & \varepsilon_n &= \frac{\alpha_n}{\alpha_n + 1} \end{aligned}$$

EA では、多くの解探索手法が提案されているが、それらの手法は類似した演算を行っている場合が多い。そのため、本  $NSS$  式は、今後、バクテリア EA に特化せず、上述のようにパラメータの値を代入したり、変数式を多少拡張したりすることで、式の基本的な構成や、用いている変数式を他の解探索手法にも適用していくことが可能となると考えられる。

## 5. シミュレーション

### 5.1 シミュレーション環境

本論文では、NK-Landscape 問題にバクテリア EA および山登り法を適用し、本  $NSS$  式の検討を行う。染色体長  $N$  を 50、各遺伝子のエピスタシスの強さ  $K$  を 5, 10, 15、バクテリア演算の際の染色体の増殖数  $M$  を 6, 101 (新しく生成される解がそれぞれ 5

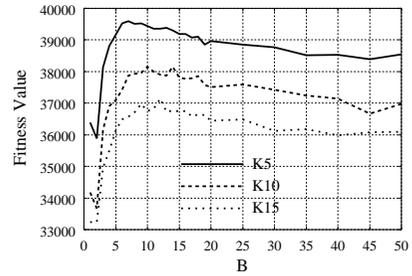
個, 100 個), 染色体の分割幅  $B$  を 1 から 50 としてシミュレーションを行った. また, 突然変異率  $P_m$  は分割幅  $B$  に関係なく確率上一定の個数変化するように設定した. 具体的には,  $P_m = 2/B$  として与えた. 山登り法については, 毎世代ランダムに選ばれた 1 つの遺伝子を変化させ, 評価値が改善されれば元の染色体と置き換えるという方法を用いた. 各遺伝子に与えられる得点は, その遺伝子を含め右側に連なる  $K$  個の遺伝子との組合せで決まるものとし, それらの組合せ  $N \times 2^K$  通りに対して, 最大 999 までの整数値でランダムに設定した.

5.2 結果と考察

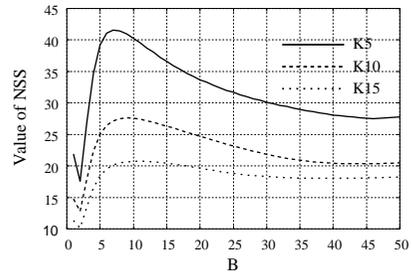
シミュレーション結果をそれぞれ, 図 6, 図 7, 図 8 に示す. 図 6(a) および図 7(a) は, バクテリア EA において, 増殖数  $M$  を 6, 101 と設定した場合における, エピスタシス  $K$  を 5, 10, 15 と変化させたときの, 10,000 世代後のそれぞれの分割幅  $B$  に対する最終評価値を示している. また図 8(a) は, 山登り法とバクテリア EA それぞれの評価値の推移を示している. ここでは, これら 2 つの手法を正当に比較するため, 横軸は世代数ではなく評価回数としてある. すなわち, 山登り法では 1,000 世代, バクテリア EA では増殖数 6 で 200 世代 ( $(M - 1) \times G = 1000$ ) までの評価値の推移である. 図 6~8(a) のシミュレーション結果はそれぞれ 50 回試行の平均値である.

図 6(b) は, 式 (1) にシミュレーションと同様のパラメータを代入したときに得られた  $B$  に対する  $NSS$  の値である. 図 6(a) と図 6(b) を比較すると, その推移が対応していることが分かる. 図 6(a) において, 評価値が最大となるときの  $B$  の値が, 図 6(b) において  $NSS$  の値が最大となる付近とほぼ一致している. このことから, 本  $NSS$  式を用いることで, シミュレーションによる試行錯誤なしに, 適切な分割幅  $B$  に対する指針が得られると考えられる.

また, 増殖数  $M$  を 101 とした場合においても, 図 7(a) と図 7(b) を比較すると, その推移が対応していることが分かる. このとき, 図 7(a) のシミュレーション結果において, 評価値の最大値を与える  $B$  が  $M = 6$  の場合と比較して, 若干右側にずれている. これは, 4 章で述べたように, 大きい  $M$  に対して  $B$  が狭い場合, 一度に局所的な最適化が行われ, 染色体全体としての最適化が困難になったためだと考えられる. これに対し, 本  $NSS$  式による図 7(b) の結果でも, 同様に  $NSS$  の値の最大値を与える  $B$  の値が右側にずれている. このことより, 本  $NSS$  式が分割幅  $B$  と増殖数  $M$  の関係についても適切に評価できると

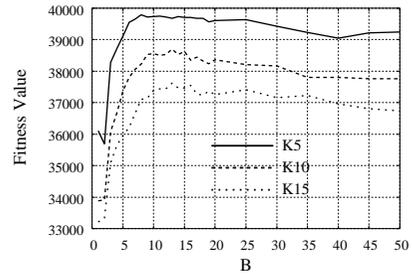


(a) シミュレーション結果

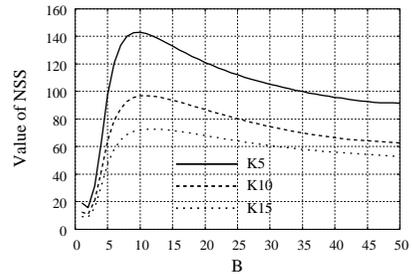


(b)  $NSS$  式による  $NSS$  の値

図 6  $K$  を変化させたときの評価値,  $NSS$  の値 ( $M = 6$ )  
Fig. 6  $NSS$  and fitness value ( $M = 6$ ).



(a) シミュレーション結果

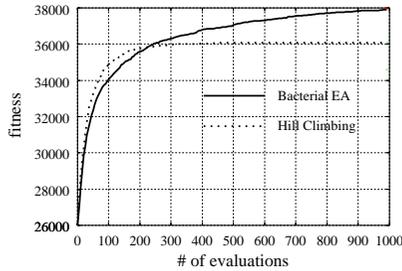


(b)  $NSS$  式による  $NSS$  の値

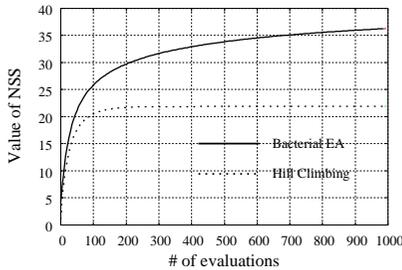
図 7  $K$  を変化させたときの評価値,  $NSS$  の値 ( $M = 101$ )  
Fig. 7  $NSS$  and fitness value ( $M = 101$ ).

考えられる.

さらに, 図 8(b) は, 世代に対する山登り法およびバクテリア EA の  $NSS$  の推移を示している. 図 8(a) と図 8(b) を比較すると,  $NSS$  の推移が, シミュレー



(a) シミュレーション結果



(b) NSS 式による NSS の値

図 8 山登り法とバクテリア EA の評価値, NSS の推移  
Fig. 8 Transition of NSS and fitness value.

シミュレーションにより実際に得られた評価値の推移にそれぞれ対応していることが分かる。したがって、この NK-Landscape 問題に対しては、実際のシミュレーションを行うことなく、バクテリア演算の方が有効であることが NSS 式により示すことが可能であると考えられる。また、この結果から、本 NSS 式またはそれを拡張することにより、パラメータの最適化だけでなく、手法そのものの性能評価や、他の手法との性能比較についても、本 NSS 式が適用できる可能性が示されたと考えられる。

## 6. まとめと今後の課題

本論文では、バクテリア EA における適切な分割幅の導出法および探索性能の評価について考察した。NK-Landscape 問題を例題として用い、バクテリア EA における解の探索能力を示す指標として探索可能解個数  $NSS$  を定義し、エピスタシス  $K$ 、分割幅  $B$ 、増殖数  $M$ 、世代数  $G$ 、染色体長  $N$ 、突然変異率  $P_m$  を用いて定式化を行った。シミュレーションを行い、本 NSS 式を用いることで、適切な分割幅を得る指針が得られることを示した。さらに、山登り法に対して本 NSS 式を拡張し、本問題に関して、バクテリア演算が、山登り法よりも有効であることを示すと同時に、本式が、手法に対する性能評価や、他の手法との性能比較についても適用可能である可能性を示した。本研

究は、限られた問題設定ではあるものの、理論的に探索性能の評価を行うことに対する試みである。

現段階では、本 NSS 式により得られた結果は、シミュレーション結果と若干のずれが生じている。これは、本式に近似的な要素が含まれることや、NSS 式における変数に、不十分な部分や不足している部分が存在すること、また問題設定によっては、本論文で仮定している「NSS の値と実際にバクテリア EA を適用した際の評価値との線形性」が成り立たない可能性があることに起因していると考えられる。そこで、今後も、本 NSS 式についての再考察を行っていき、本式の精度の向上を図りたい。また、本論文では、山登り法に対して本式の適用を行ったが、今後は、現在の理論式を拡張していき、SA・GA などのより複雑な手法への拡張を行うとともに、NK-Landscape 問題だけでなく、一般的な問題設定への適用を行っていく予定である。その際には、本論文では取り入れなかった染色体間での相互干渉による演算(バクテリア EA においては GTO)を本 NSS 式に取り入れることが必要となると考えられる。また、一般的な問題への拡張には、本論文において既知として用いたエピスタシスの導出が必要となってくるため、これらに対する検討を行っていきたい。それにより、たとえば NSS 式を用いることにより、提案する新しい手法に対して、その有効性を理論的に検証することが可能となることが期待される。

## 参考文献

- 1) 北野宏明(編): 遺伝的アルゴリズム 1-4, 産業図書 (1993-2001).
- 2) Reeves, C.R. (編): モダンヒューリスティックス, 日刊工業新聞社 (1997).
- 3) Holland, J.H.: *Adaptation in Natural and Artificial System*, The Univ. Michigan Press (1975), and MIT Press (1997).
- 4) 古橋 武, 中岡 謙, 前田 宏, 内川嘉樹: 局所改善型遺伝的アルゴリズムの提案とファジールールの発見, 日本ファジィ学会誌, Vol.7, No.5, pp.978-987 (1995).
- 5) 宮田裕次郎, 古橋 武, 内川嘉樹: 遺伝的アルゴリズムの新しい局所改善法の提案, T. IEE Japan, Vol.116-C, No.1 (1996).
- 6) Yoshikawa, T., Furuhashi, T. and Uchikawa, Y.: Knowledge Acquisition of Fuzzy Control Rules for Mobile Robots Using DNA Coding Method and Pseudo-Bacterial GA, Lecture Notes in Artificial Intelligence, Vol.1285, pp.126-135 (1997).

7) 新貝昌也, 田中 清, エルナン・アギレ, 杉村 立夫: NK-Landscape 問題を用いたエピスタシスに対する改良 GA の性能検査, 第 10 回インテリジェント・システム・シンポジウム講演論文集, pp.451-454 (2000).

8) De Jong, K.A., Potter, M.A. and Spears, W.M.: Using Problem Generators to Explore the Effects of Epistasis, *Proc. 7th ICGA*, pp.338-345, Morgan Kaufmann(1997).

## 付 録

$$NSS = \sum_{j=0}^{G-1} \nu \alpha_j \quad (1)$$

$$\alpha_0 = 1$$

$$\alpha_{n+1} = \{1 - (\gamma_i + \gamma_p)\varepsilon_n\}(1 - \gamma_g)\alpha_n \quad (2)$$

$$\nu = \sum_{l=0}^{M-2} (1-p)(1-q)^l \quad (3)$$

$$p = (1-p_m)^B$$

$$q = \sum_{l=0}^B {}_B C_l \{(1-p_m)^{B-l} p_m^l\}^2$$

$$\begin{aligned} \lambda = & 2 \sum_{l=\max(1, K-C_2)}^{K-1} \{1 - (1-p_m)^{\min(l+\max(0, K-(C_1+C_2+(l-\max(1, K-C_2)))-1), B)} \\ & + C_1 \{1 - (1-p_m)^{\min(\max(0, 2(K-C_2-1)), B)}\} \\ & + 2 \sum_{l=1}^{\frac{B-C_0}{2}-1} (1-p_m) \{1 - (1-p_m)^{\min(\{K+\min(l-1+\max(0, K-(N-B+l-1)-1), K-1)\}, B)-1}\} \\ & + C_0 (1-p_m) \{1 - (1-p_m)^{\min((2K-1), B)-1}\} \end{aligned} \quad (4)$$

$$C_0 = \frac{(-1)^{B+1} + 1}{2}$$

$$C_1 = \frac{(-1)^{N-B+1} + 1}{2}$$

$$C_2 = \frac{N-B-C_1}{2}$$

$$\begin{aligned} \gamma_i = & \frac{1}{N} \left[ (|\lambda-B|+1) \{1 - (1-p_m)^{\min(\lambda, B)}\} + 2 \sum_{l=1}^{\min(\lambda, B)-1} \{1 - (1-p_m)^l\} \right. \\ & \left. - \sum_{l=1}^{\max(0, \min(\lambda B) - (N - \max(\lambda B)) - 1)} \{1 - (1-p_m)^l\} \{1 - (1-p_m)^{\min(\lambda B) - (N - \max(\lambda B)) - l}\} \right] \frac{\nu}{\nu+1} \end{aligned} \quad (5)$$

$$\gamma_p = 1 - \frac{1}{N} \{ (1-p_m^2)^B + (|N-2B|+1)(1-p_m^2)^{\max(2B-N, 0)} + 2 \sum_{l=\max(2B-N+1, 1)}^{B-1} (1-p_m^2)^l \} \quad (6)$$

$$\gamma_g = \frac{\nu}{N \times \frac{\sum_{l=1}^B {}_B C_l p_m^l (1-p_m)^{B-l}}{\sum_{l=1}^B p_m^l (1-p_m)^{B-l}}} \quad (7)$$

$$\varepsilon_n = \frac{\nu \alpha_n}{\nu \alpha_n + 1} \quad (8)$$

$\alpha_n$  : 探索能力減衰率

$\nu$  : 有効増殖数

$\lambda$  : 影響遺伝子数

$\gamma_i$  : 影響遺伝子による能力減衰率

$\gamma_p$  : 同一遺伝子変更による能力減衰率

$\gamma_g$  : 世代間染色体重複による能力減衰率

$\varepsilon_n$  : 解改善確率

(平成 14 年 1 月 31 日受付)  
 (平成 14 年 4 月 4 日再受付)  
 (平成 14 年 5 月 28 日採録)



山本 康高(学生会員)

1978 年生。2001 年 3 月三重大学工学部電気電子工学科卒業。同年 4 月三重大学大学院工学研究科電気電子工学専攻博士前期課程に進学、現在に至る。進化的計算手法およびファジィモデリングに関する研究に従事。日本ファジィ学会会員。



吉川 大弘

1970 年生。1997 年名古屋大学大学院博士後期課程修了。同年カリフォルニア大学バークレー校ソフトコンピューティング研究所客員研究員。1998 年三重大学工学部電気電子工学科助手。主としてソフトコンピューティングとその応用に関する研究に従事。博士(工学)。電気学会, 日本ファジィ学会会員。



古橋 武

1985 年 3 月名古屋大学大学院博士課程修了, 1990 年 12 月より同大工学部助教授, 2001 年 1 月より三重大学情報工学科教授, 現在に至る。工学博士。94 年日本ファジィ学会論文賞受賞。主として, ソフトコンピューティングに関する研究に従事。



篠木 剛(正会員)

1954 年生。1977 年東京工業大学理学部情報科学卒業。1979 年東京工業大学大学院理工学研究科修士課程修了。同年株式会社富士通研究所入社。1985 年より 1 年間米国オレゴン大学客員研究員。1998 年三重大学大学院工学研究科博士後期課程修了。工学博士。1998 年三重大学工学部講師。1999 年助教授。LSI のテストや設計支援, 並列/分散計算機システム, 探索アルゴリズム, 画像処理等に興味を持つ。電子情報通信学会, IEEE 各会員。



鶴岡 信治(正会員)

1954 年生。1979 年名古屋大学大学院博士前期課程修了。同年三重大学工学部電子工学科助手。現在同大工学部電気電子工学科教授。1991 ~ 1992 年米国ミシガン大学デアボーン校客員助教授。文書構造理解, 医用画像処理, 時系列データの統計的解析の研究等に従事。工学博士。電子情報通信学会, 計測自動制御学会, 日本 ME 学会, 人工知能学会, 映像情報メディア学会, IEEE 各会員。