

遺伝的アルゴリズムの分散並列化に関する研究* (踏み石モデルによる分散遺伝的アルゴリズムの検討)

三木 光範¹, 廣安 知之¹, 中村 康範²

A Study on Parallel Distributed Genetic Algorithms (Discussion on a Stepping Stone Population Model of Distributed Genetic Algorithms)

Mitsunori MIKI³, Tomoyuki HIROYASU³, and Yasunori NAKAMURA⁴

³ Doshisha University, Dept. of Knowledge Engineering, Kyo-Tanabe, Kyoto, 610-0321, Japan

This paper discusses about the characteristics of distributed genetic algorithms (DGAs). Among the several types of distributed models of GAs, this paper focuses on a stepping stone population model. The effects of the number of subpopulations, subpopulation immigration rate, and immigration interval on the performance of the DGA are examined. Solving standard mathematical problems and a typical structural optimization problem, it is found that there are several advantages in the stepping stone population model of DGAs: this model is very suitable for parallel computation, and the optimum solutions can be obtained efficiently due to the partitioning of the population.

Key Words : Genetic Algorithms, Distributed Processing, Parallel Computing, Stepping Stone Population Model, Optimum Design

1. はじめに

遺伝的アルゴリズム(Genetic Algorithms: GA)は生物の遺伝と進化を模擬したアルゴリズムであり,確率的探索,学習,最適化などに広く応用されている⁽¹⁾⁽³⁾.特に近年では,工学的な問題の解決のためにGAを利用する研究が活発に行われている⁽⁴⁾.例えば,GAによるトラス構造設計⁽⁵⁾や構造の位相決定問題におけるGAの利用⁽⁶⁾⁽⁷⁾,材料設計や選択にGAを利用する研究⁽⁸⁾⁽⁹⁾など様々な問題の有用な解法として広く用いられている.さらに,対象とする問題が離散探索空間であるような従来のものから,連続空間に対応できるようなGAの開発⁽¹⁰⁾や離散・連続空間,双方が混合した問題への適応⁽¹¹⁾といった研究も行われている.

しかしながら GA では良い解を探索するためには,数多くの個体を用意し,多くの世代にわたって繰り返し計算を必要とするため,計算負荷が高くなる.そのため,計算時間を短縮することは重要な課題となっている.

この問題に対する一つの解決法として並列計算機を利用することが挙げられる.GAを並列化することによって計算時間が短縮されることは明白であろう.

GAを並列化する手法はいくつか考えられるが,その1つに母集団を分割する分散 GA による手法がある.分散GAでは分割された母集団間での複数個の個体の移住時以外にはプロセッサ間通信が発生せず,したがって並列化効率が高くなり計算時間短縮の効果は大きい.こうした大きな利点とは別に分散GAには個体進化における早熟を避け,世代が進んでも個体の多様性を維持することができ,得られる最適解の高品質化が達成できるという利点が報告されている⁽¹²⁾.しかしながら,工学的な問題に対する分散GAに関する研究はほとんど見られず,さらに詳細な研究が必要である.こうした観点から本研究では分散GAの基本的な特性を明らかにする目的で,構造最適化の分野における問題を対象に,分割母集団の数,個体数,移住間隔などの影響を調べることにする.

2. 並列分散GA

GAの並列化についてはNangとMatsuo⁽¹³⁾が種々の並列化モデルについてレビューを行っている.それによるとGAを並列化するには次の3つのアプローチが考えられる.

- 1)母集団の並列化
- 2)評価の並列化
- 3)選択,交差,突然変異を行う再生とよばれる遺伝

*原稿受付 1998年11月6日

¹ 正員,同志社大学(610-0321 京田辺市多田羅都谷1-3)

² 正員,大阪産業大学(570-8530 大東市中垣内3-1-1)

子操作の並列化

このうち、3)は計算の粒度が小さすぎることや、世代ごとに全プロセッサで同期をとる必要があること、そのときにプロセッサ間通信が多く発生することなどの理由から、プロセッサを各個体に与えるように特別に設計された並列計算機以外では好ましい方法ではない。一方、2)は適合度の計算にかなりの計算コストがかかる場合に適した方法であり、これは解析の並列化であってGAの並列化とはいえない。このような理由により、GAの並列化は基本的には1)の母集団の並列化、すなわち母集団を分散化して並列化するというアプローチで行われることが多い。

母集団を分割するGAをここではDGA (Distributed GA: 分散GA) と呼び、通常の母集団が集中したGAをCGA (Canonical GA: 単一母集団GA) と呼ぶこととする。分散GAでは分割された母集団内で独立した世代交代が進み、適切な世代間隔で、あるいはランダムな世代間隔で、ある数の個体が別の母集団に移る、すなわち移住を行う。この移住をどのように行うかで次の3種類のモデルがある。

- 1) 島母集団モデル (Island Population Model)
- 2) 踏み石母集団モデル (Stepping Stone Population Model)
- 3) 近傍母集団モデル (Neighborhood Population Model)

ここで、1)では移住はランダムに行われ、2)では近傍母集団間で移住を行う。3)では母集団内の個体数は1であり、単一の個体が隣接する個体とのみ交差を行う。ここで、3)のモデルはGA専用設計されたCPUでない限り極めて多数のプロセッサが必要となり、また計算の粒度が小さすぎて効率が極めて悪くなる。

こうして、分散GAにおいては島母集団モデルおよび踏み石母集団モデルが基本的なモデルとなる。ただし、両者の違いは単に移住方式が規則的かランダムかというだけの相違である。このため分割された母集団のことをこれ以後は島と呼ぶことにする。

これまでにGAの並列化に関してなされてきた研究は主として先に述べた並列化の各種のアプローチや、上で述べた母集団分割の方法などに関する種々のモデル化に関するものが多い。GAの並列化の研究の歴史は浅く、現状では並列化に関する基本的な方式の提案が中心となる。したがって、各モデルでの詳細な研究はこれからの課題となる。特に島母集団モデルおよび踏み石母集団もでるに基づく分散GAにおいては次に示す3点が論点となるものと考えられる。すなわち、

- 1) 島の数

2) サブ母集団における個体数

3) 移住スキーム、移住率、移住間隔

である。本研究においては、第4章において一般的な数学的関数および構造物の最適設計問題を対象として数値実験を行い、これらの検討を行うこととする。

3. 本研究における分散GAスキーム

本研究においては次に示すような分散GAのスキームを考える。

全体の個体数 $N(N=630)$ を一定とし、これを島数 P ($P=3, 7, 15, 30$) で分割した個体数を各島における個体数とする。各島内でのGAは単純な一点交差のGAとし、予備的な検討により交差率0.6、突然変異率0.01とした。

分散遺伝的アルゴリズムの最も基本的な工学問題への応用の際の効果を検討するために、移住方式として踏み石モデルを採用した。先に述べた通り、踏み石モデルにおける移住先をランダムにしたものが島母集団モデルであると考えられるからである。踏み石モデルにおいては、各島がリング状に連結され、あらかじめ決められた隣の島へ1方向に順次移住する。移住においては島の個体数に移住率を乗じた数の個体をランダムに選択し、一定の世代交代後に同期をとって移住させる方式とした。すなわち、移住においてはエリート個体を考慮しない。この一定の世代交代の回数を移住間隔と呼ぶ。

一方、各島でのGAにおいてはエリート保存戦略を用いた。すなわち、各島における各個体の適合度を計算し、もっとも適合度が低い個体を前世代のエリート個体と入れ換える。その後、その世代でのエリートの保存を行う。次に移住を行う。ここでは該当する個体群を先に送り出し、そののち同じ数の個体群を別の島から受け取る。そのあと、交差、突然変異、およびブレット方式による淘汰を行う。

4. 数値実験と考察

4.1 計算の対象とした問題 本研究においては2種類の問題を対象とする。すなわち、標準的数学問題と構造最適設計問題である。まず、数学問題に前章で説明したスキームを用いて解を探索し、分散GAの特徴を把握する。続いて、工学的な設計問題においても同様に解を探索して分散GAの特徴を把握し考察を行う。標準的数学問題ではGAの研究で一般的に用いられるDeJongの関数のうち、もっとも最適解が得られにくいものを2種類採用している。その他のDeJongの関数では最適解がすぐに得られてしまい、手法の比

較ができない。

使用した計算機は MIMD型の並列計算機nCUBE2E (米国nCUBE社, 64プロセッサ) であり, 個体総数は 630 とし, これを 1, 3, 7, 15, 30 島に分割した。後に示す結果はいずれも 100 世代後のものであり, 計算結果はすべて 10 回の試行の平均である。

4.2 標準的数学関数問題(その1) 式(1)は2変数の問題で非常に多峰性の高い問題である。本研究ではDeJongの原関数の複雑度を2倍にしたものを用いている。変域は -65 から 65 までとし, 長さ 34 ビットの染色体を用意しバイナリ表現をすることにより分解能を 0.001 とした。

$$f = -\left(\frac{1}{500} + \sum_{j=1}^{49} \frac{1}{j + \sum_{i=1}^2 (x_i - a_{ij})^6}\right)^{-1} \dots\dots\dots(1)$$

Fig. 1 では横軸に島数, 縦軸に各島で得られる適合度のうちの最大適合度を示している。ただし, 適合度は式(2)で表される適合度関数によって求められる値であり最大値約 1.0 が最良値である。

$$f = -\frac{1}{S} \dots\dots\dots(2)$$

これより明らかなように, 単一GAでは最適解が得られないのに対して, 移住率によって多少ばらつきがあるが分散 GA の方が良い解を得ていることがわかる。

Fig. 2は式(2)の関数fに関するGAの最大適合度と世代の関係を示したもので, 10回の試行平均を表している。これより明らかなように, 単一母集団では個体の早熟が生じ, その局所最適解から突然変異で抜け出すことが困難であることがわかる。

これらの結果からこの問題において, 分散GAでは単一母集団GAと比較して初期収束することなく良解が得られることがわかる。特に, 島数が多いときには移住率は問題とならないことも明らかとなった。

4.3 標準的数学関数問題(その2) 式(3)は単純

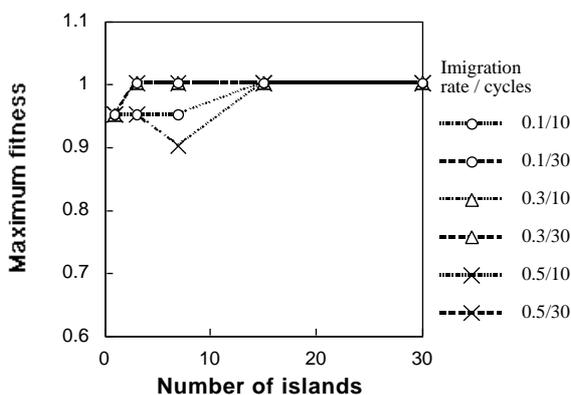


Fig. 1 Maximum fitness for function f

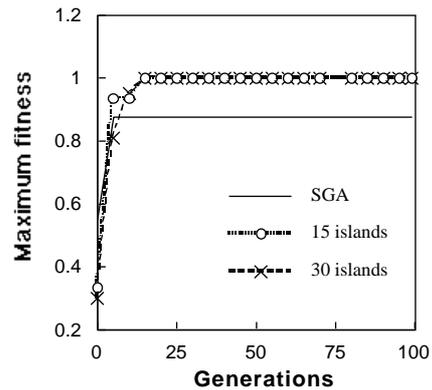


Fig. 2 History of the maximum fitness for function f

な関数に平均 0, 標準偏差 1 のガウスノイズが乗っている問題である。この関数は 30 変数で, 一つの変数の変域を -1.27 から 1.27 までとし, 分解能を 0.01 として変数に 8 ビットを与え, 染色体の長さは 240 ビットとした。この関数を適合度関数とし, この関数の値を適合度としている。この関数は同じ染色体および遺伝子的に近接した染色体でもノイズの影響で異なる適合度を持つために, ビルディングブロックはある程度形成されるがそれ以上の形成は困難なため GA による最適解の探索は極めて難しい問題である。

$$Q = -\left\{\sum_{i=1}^{30} ix_i^4 + \text{GAUSS}(0,1)\right\} \dots\dots\dots(3)$$

縦軸に各島で得られる適合度のうちの最大適合度を, 横軸に島数を表した計算結果を Fig. 3 に示す。

島数が 3 では単一母集団のものと比較してある程度良好な解が得られているが, 島数がそれ以上に増えると解が極端に悪くなる。これは各島における個体数の影響によるものと考えられる。すなわち, 式(3)は比較的単純な単峰の関数にノイズが乗っている問題であるが, GA でこの問題を探索する場合, 母集団内の染色体で表現することのできる式(3)の第 1 項の最大値と最小値との差が第 2 項のガウスノイズよりも小さくなる

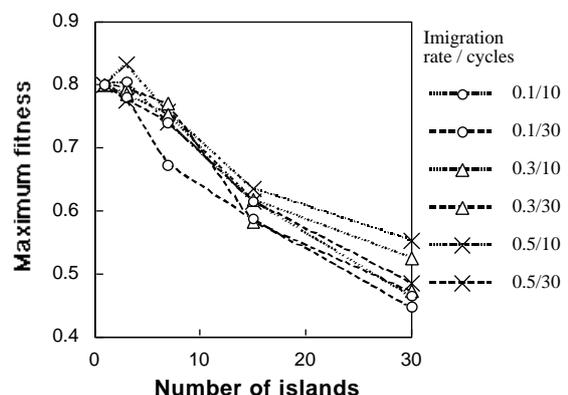


Fig. 3 Maximum fitness for function Q

場合が生じる。この場合には容易にはそこから抜け出すことができず、大局的最適解を得るためには十分な数の個体数が必要となる。

よって、今回のように全体の計算量を固定して分散GAを行う場合には、島数を増やした時に十分な個体数が存在するかどうかを吟味する必要があることがわかる。

4.4 構造最適設計問題 この問題はFig. 4に示す11部材からなるトラス構造物の最小体積問題である。Fig. 4における各数字はそれぞれ節点番号および部材番号を示す。最小体積問題は構造設計問題の基本的問題の1つであり、ここで得られる分散GAの特性は、同種の最適設計問題において、ラーメン構造や板構造にも拡張可能であると考えられる。目的関数は部材の体積の合計であり、制約条件は各部材の引張応力および座屈応力、ならびに節点6における変位制約とした。これと同種の問題は多いが、ほとんどの問題は部材の引張と圧縮応力制約のみで、座屈応力制約や変位制約を考慮しているものは少ない。これによって問題は各段に難しくなる⁽¹⁴⁾。

設計変数は円形部材の断面積とし、変域を1から4000mm²とし、これを12ビットで表現する。最小化すべき評価関数は次式に示す。

$$H = w_v V_T + P_t + P_b + P_d \dots \dots \dots (4)$$

ここで、 V_T は部材体積の合計、 w_v は重み係数、 P_t はそれぞれ変位(d)、引張(t)、および座屈(b)の制約条件に関するペナルティ関数である。変位に関するペナルティは変位の2乗に適切な重み係数をかけたものとし、引張応力および座屈に関するペナルティは制約条件が満足されないときに一定値を足す方法をとった。この方法を用いない場合には局所制約を完全に満足させる解を見出すことは難しくなる。

Fig. 4に示したトラスの最適設計の結果をFig. 5および6に示す。Fig. 5は全個体中の最大適合度の値と島数の関係を示し、Fig. 6は各島における最大適合度の値

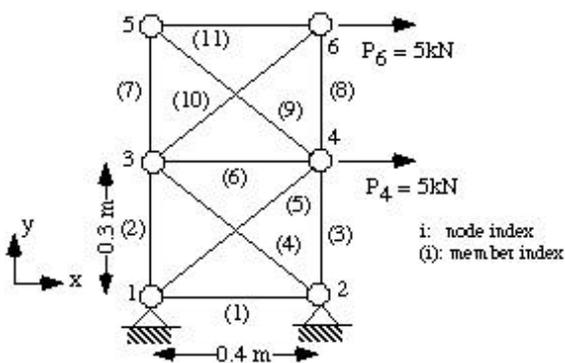


Fig. 4 A 11-member truss structure

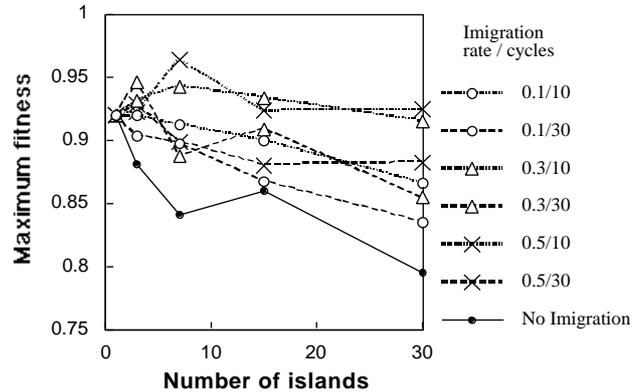


Fig. 5 Maximum fitness for truss optimization

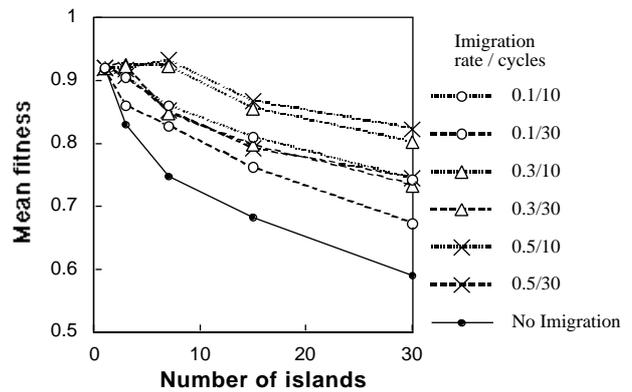


Fig. 6 Mean fitness for truss optimization

の島平均（図中では Mean fitness と表記）と島数の関係を示したものである。これより、島数、移住間隔、および移住率を適切に選べば適合度の最大値も島内最大値の島間平均値とともに母集団の分散化によって上昇することがわかる。また、Fig. 5より、適合度の最大値の上昇は7島で移住間隔10世代、移住率0.5で最大となっていることもわかる。この条件で1000世代の計算を行ったところ適合度は1.0636となった。Fig. 7に得られた最適解を示す。この図において、各部材の太さが断面積比を表し、最も太い部材（部材番号4）の断面積が1416mm²である。一方、計算時間はたとえば7島では約1/7になり、分散GAの並列化は極めて優れた方法であることがわかった。

これらの図から移住を行わない分散GAと移住を行った結果とを比較するとこの問題においても島数を増大することによって1島あたりの個体数が減少してしまい、良好な解が得られなくなっていることがわかる。しかしながら、そのような条件下でも、移住を行うことで解が高品質化していることがわかる。したがって、母集団を分割して個体を適切に移住させることにより、通常の単一母集団GAのときに必要な個体数よりも1島においては少ない個体数で解が得られると言える。

分散GAは解の高品質化とは別に先に述べた通り並

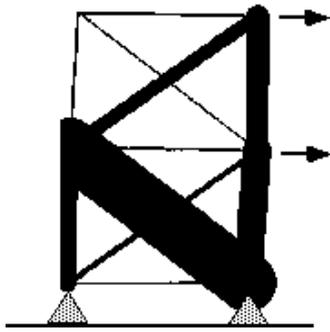


Fig. 7 Example of optimum solution

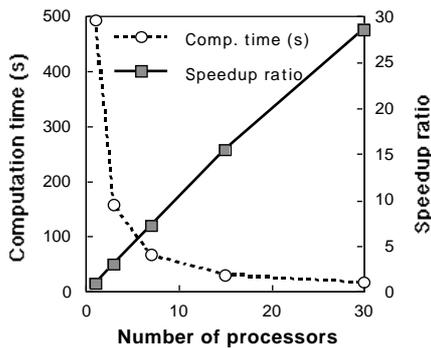


Fig. 8 Computation time and speed up ratio

列処理によって計算時間の短縮化に貢献する。Fig. 8 には横軸に並列計算機で使用したプロセッサの数を取り、縦軸は計算時間および速度増加の比率を示している。プロセッサの数は分散GAにおける分割した母集団の数、すなわち島数に相当する。速度増加の比率は1プロセッサを使用する際に必要な時間を各プロセッサを使用する際に必要な時間で除した値である。これより、用いるプロセッサの数に比例して高速化が計られ、しかも並列化効率はほぼ100%であることがわかる。これは、プロセッサ間通信が移住時にだけ発生し、それ以外の時間は各プロセッサは独立して処理を進めることができるためであり、分散GAは並列化に極めて適した方法といえる。

これまでの結果から分散GAは最適解の品質を向上させることが分かった。次にその理由を考察する。

単一GAでは解の早熟により局所最適解に収束していることが予想される。これに対して、分散GAでは分散した母集団内では個体は局所最適解に収束している可能性が高いが、それらの局所解は各母集団ではそれぞれ異なっていると考えられ、移住操作によってこれらの局所最適解同士が交叉し、よりすぐれた解が得られるものと考えられる。

このことを確かめるには、得られる各島の適合度の最大値や各島での適合度の最大値の島数による平均、あるいは適合度の全母集団による平均ではなく、解そのもの、すなわち個体の表現型を個別に調べることが

必要である。

そこで、単一母集団における上位30の適合度を持つ個体と、30島モデルにおける移住間隔10世代、移住率0.5の条件で得られる各島での適合度の最大値を持つ合計30個体の解の詳細を比較する。Fig. 9および10は単一GAで良好な解が得られた試行における49および99世代での上位30の適合度を持つ個体の設計変数の分布を示したものである。

これより、単一GAでは49世代目での630個体のうち適合度が高い個体はすでに局所最適解に収束していることがわかる。すなわち、単一GAでは、初期値に依存している局所最適解が見つかったと、その解の適合度が高いために高い確率で選択が行われ、母集団の中で個体の多様性が急速に失われることがわかる。一方、99世代目では突然変異の作用でその局所解と少し異なる解が出現していることがわかる。

これらは良好な解が得られた試行に関する結果であり、悪い解しか得られなかった試行では99世代においても49世代と同様、解の多様性が全く見られないものが多くあった。こうして、単一GAでは個体の多様性を保持するのが困難で、これによって大域的最適解の

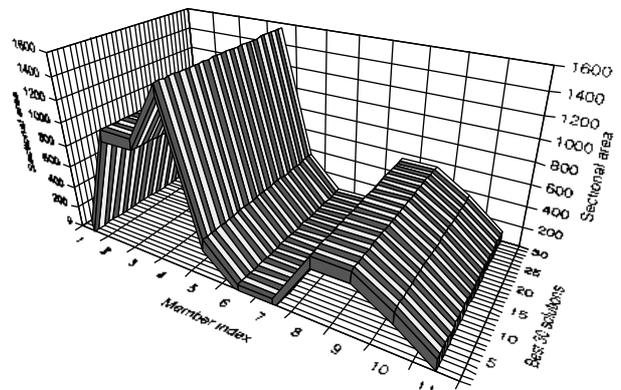


Fig. 9 Distribution of design variables in SGA (49 generation)

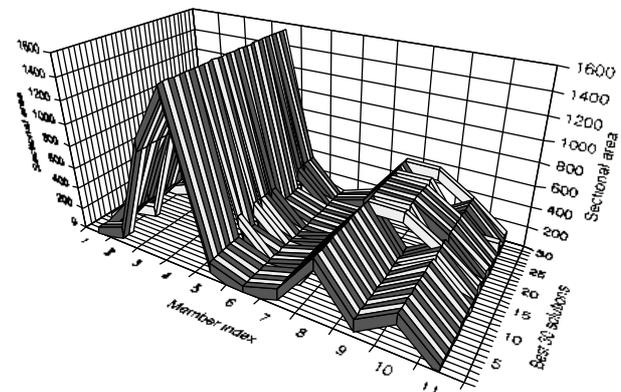


Fig. 10 Distribution of design variables in SGA (99 generation)

探索が困難となっていることがわかる。

一方、分散GAでの同様の結果をFig. 11および12に示す。これを見ると、30島の各島における適合度の最大値を持つ個体は49世代においては大きくばらついており、個体の多様性がいつまでも保持されていることがわかる。また、99世代では移住の影響で大域的最適解に収束している様子も分かる。また、99世代においても単一母集団のGAと比べて、かなりの個体の多様性を保持していることもわかる。

これらの結果より、分散GAではそれぞれの母集団において異なった局所解が見いだされ、それが移住によって交叉し、大域的な最適解に発展することが確認できた。

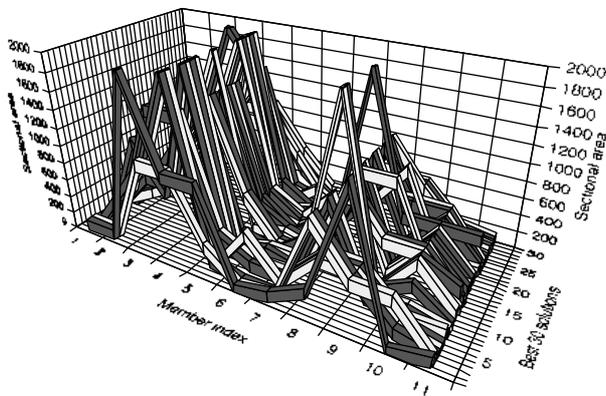


Fig. 11 Distribution of design variables in DGA (49 generation)

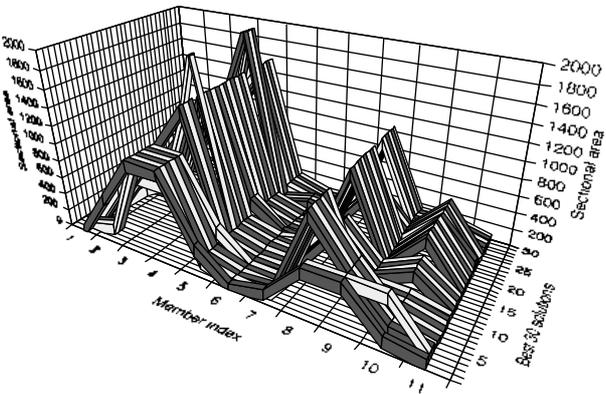


Fig. 12 Distribution of design variables in DGA (99 generation)

5. 結 言

本研究では、分散GAの基本的な特性を明らかにするために、以下の検討を行った。

- [1] GAを並列化する際にどのようなモデルが可能であるかを分類し、分散GAの論点を明確にした。
- [2] 数学関数問題および構造最適設計問題に対して、踏み石モデルの分散GAを適用し、その特性を把握し、考察を行った。

その結果、単一GAに対して分散GAは次のような利点を持つことが明らかとなった。

- [1] 並列処理に適しており、並列化効率はほぼ100%となり、計算の高速化が可能となる。
- [2] 適切な移住率および移住間隔を選択する事により、分散GAは単一GAより高い品質の最適解をもたらす。
- [3] 単一GAと同様、分散GAでも各島に十分な数の個体数が必要である。ただし、各島に必要な個体数は、移住の効果により単一GAの際に必要な個体数よりも少ない。
- [4] トラス構造の数値計算を通じてある種の構造最適設計問題は分散GAを利用することで、高い品質の最適解が高速に得られる可能性がある。
- [5] 解の高品質化は分割された母集団での進化によりそれぞれ異なった局所最適解が探索され、それが移住により組み合わせられることで大域的な最適解になる。一方、単一母集団では局所最適解に収束する早熟現象が見られる。

最後に本研究遂行にあたり、数値実験に協力いただいた大阪産業大学大学院、谷上克己君（現在：島津エス・ディー（株））に心から感謝する。

参 考 文 献

- (1) Goldberg, D. E., Genetic Algorithms in Search, Optimization and Machine Learning, (1989), Addison-Wesley Publishing Company.
- (2) 北野宏明編, 遺伝的アルゴリズム (1993), 産業図書.
- (3) 田中・坂和, 遺伝的アルゴリズム (1995), 朝倉書店.
- (4) 嘉数・皆川, 機械の研究, 45-5 (1993), 521.
- (5) 朝山・他2名, 機論, 62-597, A(1996), 1234.
- (6) 稲川・他2名, 機論, 61-587, C(1995), 2901.
- (7) 尾田・他2名, 機論, 61-582, A(1995), 460.
- (8) 鹿・他2名, 機論, 61-584, A(1995), 805.
- (9) 轟・他3名, 機論, 61-587, A(1995), 1453.
- (10) 古川・矢川, 機論, 61-586, A(1995), 1409.
- (11) 荒川・萩原, 機論, 64-621, C(1998), 1626.
- (12) Belding, T.C., The Distributed Genetic Algorithm Revisited, Proc. 6th International Conf. on Genetic Algorithms, (1995), 114.
- (13) Nang J., and Matsuo, K., A Survey on the Parallel Genetic Algorithms, 計測と制御, 33-6, (1994), 500.
- (14) Miki, M., et al., AIAAPaper-96-1584, (1996).