

2 個体分散遺伝的アルゴリズムの並列モデルの検討

廣安 知之, 三木 光範, 谷村 勇輔, 佐野 正樹

同志社大学 工学部

E-mail: tomo@is.doshisha.ac.jp

本研究では2個体分散遺伝的アルゴリズム (DuDGA) の分散メモリ型並列計算機システムへの実装モデルを提案する。DuDGAは分散遺伝的アルゴリズムにおける島内の個体を2とし、各種の遺伝的操作に対して解の多様性が維持できるように改良を施したモデルである。本研究で提案するモデルではさらに島をいくつかのサブグループに分割しそれらを各プロセッサに割り当てる。提案モデルをテスト関数に適用することにより、その特性や有効性、パラメータの影響、細粒度モデルとの性能比較などの検討を行う。

Discussion on Parallel Implementation Model of Dual Individual Distributed Genetic Algorithms

Tomoyuki Hiroyasu, Mitsunori Miki, Yusuke Tanimura, Masaki Sano

Knowledge Engineering Department, Doshisha University

In this study, we propose a implementation model of Dual Distributed Genetic Algorithms (DuDGAs). In DuDGAs, there are only two individuals in each island and there are several islands are existed. In the proposed model, the islands are divided ed into sub groups. Each group is assigned to a processor and in each processor, DuDGA is performed for several iterations. After some iterations, one of island in each group is selected and is moved to the other group. This model was applied to some test functions. Through the numerical examples, the parallel efficiency of this model, the effects of the parameters and the comparison of the fine grained model were examined.

1 はじめに

我々は分散モデルの拡張として”2個体分散遺伝的アルゴリズム (Dual Individual Distributed Genetic Algorithm (DuDGA))”を提案した¹⁾。DuDGAはDGAにおける島内の個体数を2とし、多様性を維持するような遺伝的操作を改良したモデルである。いくつかのテスト関数に適用したところ、DuDGAはDGAやSGAと比較して少ない計算量で良好な解を求めることが可能であることが明らかとなった。

しかしながら、分散メモリ型並列計算機にてDuDGAを実装することを想定した場合、通常、島数よりもCPU数が少なくなるために、提案モデルをそのまま実装することができない。すなわち、分散メモリ型並列計算機に対してはDuDGAの実装モデルが必要となる。そこで、本研究では、DuDGAの分散メモリ型並列計算機の実装を想定したモデルを提案する。そのモデルにおいては、いくつかのパラメータが存在するために、そのパラ

メータが解の精度にどのような影響を与えるかについて数値計算例を通じて検討する。

2 2 個体分散遺伝的アルゴリズム (DuDGA)

遺伝的アルゴリズム (Genetic Algorithm: GA) における母集団をいくつかのサブ母集団に分割する。これらのサブ母集団は島と呼ばれる。島内では通常のGA操作を行う。数世代後、島内の数個体を選択し、他の島へ移動する。この操作が移住であり、移住により全体での解の多様性が保持される。これが島モデル (分散母集団モデル: DGA) の流れである。このモデルでは、通常のGAのパラメータに対して、移住を行うタイミングを決定する移住間隔、移住する個体数を決定する移住率、島数などの新たなパラメータの設定を必要とする。

2個体分散遺伝的アルゴリズム (Dual Individual Distributed Genetic Algorithms: DuDGAs) はDGAにおける各島の個体数を2として、多数の島でGAを行うアルゴリズムである。そして遺伝的

操作には多様性を維持するような改良を行っている¹⁾。

これらの設定により, 交叉率 $C_r = 1.0$, 島数 $N_i = Popsizel/2$, 移住率 $C_m = 0.5$ が自動的に決定され, ユーザの負担が軽減される。

3 DuDGA の分散メモリ型並列計算機への実装モデル

本研究では, 2 個体分散遺伝的アルゴリズムを分散メモリ型並列計算機に実装するために, 以下のようなモデルを提案する。

本モデルは以下のような流れで処理される。すなわち,

Step1 島群は複数のサブグループに分割される。基本的にそれぞれのグループが各プロセッサにわりあてられる。

Step2 それぞれのサブグループ内で通常の DuDGA が行われる。サブグループ内で個体の移住が行われる。

Step3 数世代後, 任意の島が選択され他のサブグループに移動される。

Step4 終了判定後, 条件を満たさない場合, ステップ 2 に戻る。

本モデルでは, サブグループ間で行われる島移動の際のみにプロセッサ間で通信が行われる。よって, 非常にネットワーク負荷の非常に低いモデルであると言える。図 1 に DuDGA のアルゴリズムの概略を示す。図では 2 プロセッサの利用の場合を示している。

4 数値計算による提案モデルの特性の検討とパラメータの解への影響

本章では, 前章で提案したモデルの特性や存在するパラメータの解への影響を数値計算を通じて検討する。

4.1 使用計算機環境とテスト関数, パラメータ

本章で行う数値計算を行った計算機環境は 16 ノードからなる PC クラスタシステムである。Pentium II (400 MHz), 128 Mbytes Memory, FastEthernet を搭載し, OS は Linux 2.2.10, 通信ライブラリは MPICH 1.1.2 を利用した。

また, 使用したテスト関数は以下に示した 4 種

GA Operator

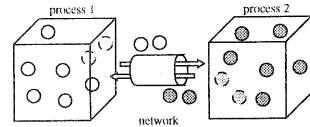
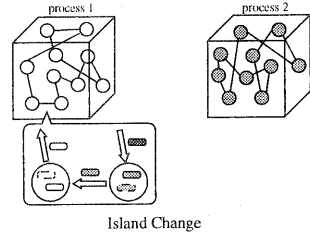
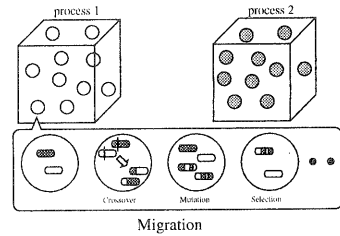


図 1: Parallel Implementation of DuDGA

類の関数である。

$$F_1 = 10N + \sum_{i=1}^N \{x_i^2 - 10\cos(2\pi x_i)\} \quad (-5.12 \leq x_i < 5.12) \quad (1)$$

$$F_2 = \sum_{i=1}^N \{100(x_{2i-1} - x_{2i}^2)^2 + (x_{2i} - 1)^2\} \quad (-2.048 \leq x_i < 2.048) \quad (2)$$

$$F_3 = 1 + \sum_{n=i}^{10} \frac{x_i^2}{4000} - \prod_{n=1}^N \cos\left(\frac{x_i}{\sqrt{i}}\right) \quad (-512 \leq x_i < 512) \quad (3)$$

$$F_4 = \sum_{i=1}^N \left(\sum_{j=1}^i x_j\right)^2 \quad (-64 \leq x_i < 64) \quad (4)$$

これらの関数はそれぞれ, Rastrigin 関数 (F1), Rosenbrock 関数 (F2), Griewank 関数 (F3), Ridge 関数 (F4) であり, 大域的最適解は 0 である。

4.2 分散と並列化の効果

DuDGA の並列モデルは, 逐次モデルと異なり移住の部分がプロセス間移住とプロセス内移住の

2段階に分かれている。このように複数のプロセスに分散化することによりモデルが変化し、また得られる解も影響を受けるものと考えられる。そこでモデルの違いにより、DuDGAの解の探索能力がどのような影響を受けるかを調べた。4つの対象問題について数値実験を行った。紙面の都合上、そのうちF1の結果を図4に示す。これらの結果は10回試行中の最良値の平均である。ここでは、総個体数を192,10設計変数、遺伝子長100bit、プロセス内移住間隔5、プロセス間移住間隔25である。グラフの横軸は各プロセスに散らばる全個体の評価回数である。GAの計算では主に評価計算の部分がホットスポットとなるため、評価回数が計算負荷の指標の1つとなる。グラフ中の並列モデルでは、192島を均等に各プロセスに割り当ててている。

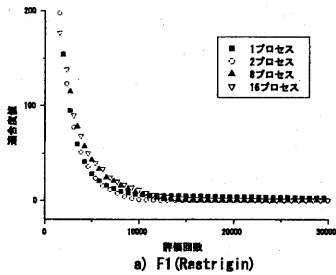


図 2: Distribution effects

いずれの関数においても、図4と同様にDuDGAの並列モデルが逐次モデルよりも良好な結果を示した。そしてF2を除くその他の3つの関数では、2プロセスの並列モデルが最も早く最適解を得た。これは2プロセスの移住トポロジが、適切な多様性の維持を作り出すことができたからであるといえる。一方F2の関数は、最適解自体が求まらなかった。つまり、いずれの結果も途中で局所解に収束してしまっている。結果として、この場合に最も多様性を維持できるモデルである16プロセスのDual GAが良い解を得ることができたと考えられる。

次に、実際のDuDGAの並列実行時の解の探索時間の調査を行った。ただしテスト関数は実問題に比べると、計算負荷が非常に小さい。そこで計算負荷を高めるために、個体の評価計算の部分を30回反復計算させて実験を行っている。この場合、通信のオーバーヘッドは無視できるほど非常に小さく

なる。ここでも紙面の都合上、F1の結果を図??に示す。解の探索性能だけに着目した場合には、F2以外では2プロセスの場合が最も効率が良いことが、並列実行した結果、どの問題においても16プロセスが高速に解を探索することができる事が分かった。その結果、本研究で提案するモデルは分散メモリ型並列計算機に適したモデルであると言える。

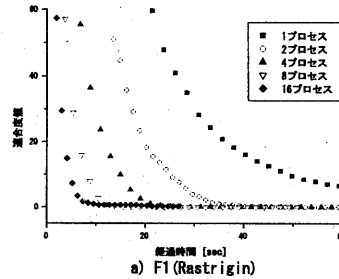


図 3: Elapsed time

4.3 移住パラメータによる影響

DuDGAでは、移住のパラメータが多様性の維持に大きく関わる。そこで移住のトポロジがリング状であった場合に、移住間隔を与える解の探索能力について調査を行った。実験ではプロセス数を16と固定した。プロセス内の移住間隔は5, 10世代を設定し、プロセス間の移住間隔は50, 100世代をそれぞれ設定した。これらの組み合わせとして計4パターンを計測し、10回試行中の最良値の結果をまとめたF1の結果が図??である。

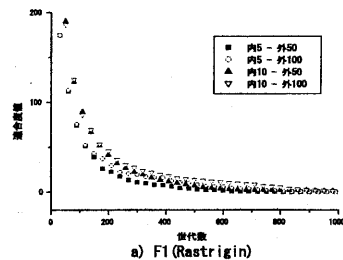


図 4: The effect of migration interval

F1やF3の問題に対しては、プロセス間の移住が小さい時に良い解を早く見つけた。一方、F2やF4の問題では、プロセス間の移住間隔を多くとった方が良い解を見つける結果となった。プロセス内の移住間隔は、いずれの問題においても短くした方

が良い結果を得た。これらのパラメータ設定は対象問題に依存すると考えられるが、今回着目した関数の性質にあまり相関のない結果となった。

4.4 Fine Grained Model との比較

DuDGA では母集団が非常に細かく分割されている。この点は Fine Grained Model と類似している。そこで、DuDGA と Fine Grained Model との比較を行った。

Maruyama らのモデル²⁾は解探索能力に優れた Fine Grained Model の一つである。以後、FG と省略する。このモデルの特徴は、通信量が少ないこと、近傍を定義しないことである。また、このモデルでは他のノードの個体の情報をバッファに格納する。本実験では、バッファサイズを 3 とし 8process、ルーレット選択、交叉率 (0.8)、突然変異率 (1/L) を採用している。本実験では個体数は DuDGA が 400、設計変数は 30、遺伝子長は 300 である。

図 5 に F1 の最良値の 30 試行の平均を示す。関数 F1 から F4 のどの結果においても探索の初期においては FG の方が DuDGA よりも解の収束が速いという結果となった。また、解の探索能力の優劣は問題によって異なった。一般に F1 および F3 は GA で探索が容易な問題、F2 および F4 は GA で探索が困難な問題である。GA での探索が容易な問題においては、DuDGA が FG を優越し、逆に、困難な問題においては、FG が DuDGA を優越した。

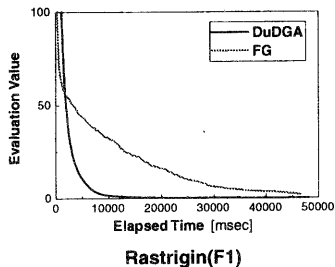


図 5: The comparison of DuDGA and FG

5 おわりに

本研究では、2 個体分散遺伝的アルゴリズム (DuDGA) の分散並列計算機への実装を想定したモデルを提案した。ここでは、先に提案した DuDGA モデルにおいて複数存在する島をいくつかのグループに分割する。このグループを各 CPU に割り当てる。各グループ内ではそれぞれ DuDGA の各操作、

島内での交叉、突然変移、選択を、島間での移住操作を行う。一定世代後、グループ間で、いくつかの島交換を行う。この島交換によって、グループ間の情報交換を行う。

本モデルをいくつかのテスト関数に適用しその有効性を検討した。その結果次のことが明らかとなった。

- 個体数を一定にした場合、適合度関数の計算量の観点から最も効率が良いのは 2 分割した場合である。
- それに対して、実際に実行時間が最も短いものはプロセッサ数だけ分割し、それらをプロセッサに割り当てた場合であった。
- 本モデルではプロセス内の個体の移住とプロセス間の島の移動が存在する。通信負荷は島の移動の際のみに発生する。個体の移住、島の移動のタイミングを決定するパラメータは解に影響を与えるが、その最適値は問題によって依存する。
- Fine Grained Model と比較した場合、GA で探索が容易な問題においては、本モデルを使用した方が良好な解が得られた。
- これらの結果を総合して、本研究で提案するモデルは PC クラスシステムなどの分散メモリ型並列計算機の実装モデルとして非常に有効であると言える。

Acknowledgement

This work was supported by a grant to RCAST at Doshisha University from the Ministry of Education, Japan

参考文献

- 1) 廣安, 三木, 佐野, 谷村, "2 個体分散遺伝的アルゴリズム", 情報処理学会研究報告, 2000-MPS-21, pp.37-40, 2000.
- 2) T. Maruyama, T. Hirose, and A. Konagaya, "A Fine-Grained Parallel Genetic Algorithm for Distributed Parallel Systems", *Proc. 5th International Conference on Genetic Algorithms*, pp.184-190, 1993.