

並列処理による進化計算の高速化について

筒井 茂 義

I. はじめに

生物進化にヒントを得た探索手法である進化計算は、1970年代から基礎研究が始まった。1990年代になり多くの研究者の注目を浴びるようになり活発な研究が行われるようになった。進化計算の大きな特徴は、解空間が非常に大きく、従来手法では高速計算機を用いても数十年かかるような問題や、問題の構造が分からず数学的に定式化できないため従来手法では解けない問題などに対して、近似解を比較的高速に得られるという点にある。この手法では、正しい解とともに、人間が思いつかないような解に行き着くこともあり、「創発的探索」とも呼ばれる。

進化計算は、2000年代になって工学分野を中心に、実問題への適用が急速に進んでいる。身近な応用の一例をあげると新幹線のぞみ700系統の先端形状の最適化設計をあげることができる¹⁾。この例では高速走行で発生する騒音と空気抵抗を小さくする形状設計を進化計算の手法によってなされ、大きな成功を収めた。

進化計算の実問題への応用の推進力の一つには、近年のコンピュータ性能の飛躍的な向上があげられる。高価で高速なコンピュータを用いなくても、安価な高性能PCを利用することにより、従来困難であった実問題への応用が比較的容易に行えるようになった。従来数時間を要していた計算が数分で行えると、例えば配送計画の最適化問題に実時間で対応できる解を得ることに利用できる。

以下本稿では、一般的なユーザにおけるコン

ピュータ環境で進化計算の並列処理による高速化について、筆者が取り組んできた並列進化計算の手法^{2, 3, 4, 5)}について述べ、今後の研究課題について考察する。まずII章では、進化計算の最新の動向について概観する。つぎに、III章では、並列進化計算の各種モデルを概観する。IV章では、Java アプレットによるネットワーク環境型の並列進化計算モデルについて、また、V章ではマルチコア計算機による並列進化計算モデルについて述べる。最後にVI章ではグラフィックスプロセッサ (Graphics Processing Unit; GPU) を用いた超並列進化計算の試みについて述べ、その検討課題について考察する。

II. 進化計算の概要

1. 探索手法としての進化計算の特徴

今日、コンピュータは家電製品から自動車、事務処理、社会インフラのシステム制御など、あらゆる分野で用いられ、もはや使われていることが意識されないほどコンピュータはあらゆる分野に遍在し、我々の生活になくてはならないものとなった。

しかし、コンピュータは原理的な限界を持っている。それは、コンピュータはプログラムによって動作し、それ以外のことは基本的にできないということである。プログラムは人間によって作成される。コンピュータはプログラマが考えた以上のことはできない。コンピュータに思考、知性、適応といった、知的な能力を持たすことができないのか？コンピュータが発明された当初から、このような「人工知能」の研

究は、コンピュータ技術者の夢であった。しかし現実には、本当の意味での人工知能の実現は大変難しいということは明らかである。

さて、ここでコンピュータから離れて、生物進化を考えてみよう。約40億年前に原始生物が誕生したと考えられている。この過程で地球の大きな変動を何度も乗り越え、変動する環境に厳しい生存競争を通して適応することで生物は進化を遂げ、高等な生物へと進化した。当初は生物の変異は突然変異のみであったが、有性生物の出現により「交叉」という新たな変異の手段を獲得し、生物の多様化と進化が促進された。

生物進化における生物のこの優れた「適応機構」をコンピュータで模擬し、プログラムにより問題を直接解くのではなく、この適応過程を得た結果として解を見つけよう、このような発想で新たな「知的プログラミング」を構成しようというのが進化計算の源流の一つである「遺伝的アルゴリズム」(Genetic Algorithm, GA)の考え方である。ここで、生物に似せるものが問題の「解候補」である。すなわち、複数の解候補を生物の個体(Individual)にみ立てて、その集団(Population)を「進化」させ、より良い解候補の集団に進化させる。進化計算の基本的な手順を単純GA(Simple GA)と呼ばれるものを例に図1に示す。

単純GAでは、まず最初に複数の個体からなる初期集団 $P(t)$ (ただし、 $t=0$)を生成する。初期個体は一般にランダムに生成されるので、初期集団 $P(0)$ は低品質の解候補の集合である。つぎにそれらを評価(Evaluation)し、その中からよい個体を選択(Selection)する。選択された個体をペアリングし交叉(Crossover)と突然変異(Mutation)からなる遺伝的操作(Genetic Operator)を適用し次世代集団を生成する。これを設定した品質の解(個体)が得られるまで繰り返す。

進化計算はGAの他、本章2節で述べるように多くの手法があり、現在では多くの実際の問題に応用され、従来の手法では解くことが困難であった問題の解法に適用されている。ここで、進化計算の特徴をまとめると以下のようになる。

- (1) 解の探索は集団で行う。すなわち、探索手法から分類すると、多点探索である。これは山登り法やシミュレートアニーリングと大きく異なる点である。これにより局所解にトラップされることが少なくなる。
- (2) ブラックボックス探索ができる。生物は生存できるかどうかは環境への適応度である。これと同様進化計算では解候補の優劣の情報を用い、数学的な定式化ができない問題にも対処できる。

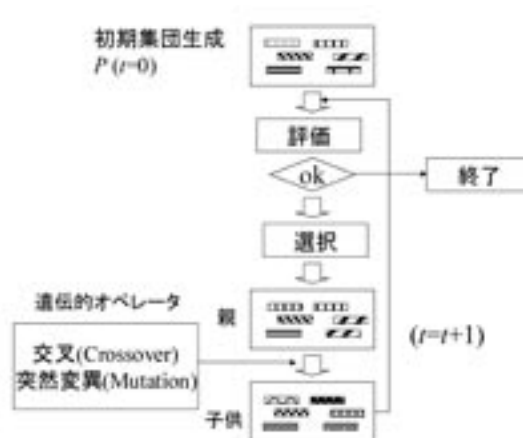


図1 進化計算の手順(単純GAの場合)

- (3) 上記(2)の特性は非常に大切な特徴であり、微分不可能な不連続な問題や非線形問題の解法にも有利となる。数学的に定式化できる問題でも、組合せ爆発により高速計算機を用いても実用的な時間で解けない解空間の大きな問題に対しても対応できる。
- (4) 「進化」の結果として解を得る。これにより人間が思いつかないような解を得るといって「創発的探索」ができる。

進化計算の応用事例を表1に示す。このように、進化計算は主に工学分野での応用が多い。しかし、最近では金融分野や、社会経済分野でも応用が始まっている。

2. 進化計算の研究の発展過程

進化計算は、本章2節で述べたように生物進化にヒントを得た探索から始まっているが統計的探索手法との融合など、その枠組みが広がっている。進化計算の最大の特徴は、集団で探索するという点にある。現在では、集団で探

表1 進化計算の応用事例

設計・計画問題	プリント基板の最適位置配置
	LSIのレイアウト設計
	ニューラルネットワークの合成
	通信網の動的予測ルーティング
	ジェットエンジンの設計
	エアコン制御の設計
	ハードウェア進化
	購入計画問題
制御問題	レンズ系の設計
	エレベータ群管理
	ガスパイプライン制御
	プロセス制御
スケジューリング問題 問題組合せ最適化問題	配電システムの損失最小化制御
	ジョブショップスケジューリング
	各種工程スケジューリング
	製品投入スケジューリング
	運送業乗務員スケジューリング
学習・その他	巡回セールスマン問題
	ロボットの行動学習
	概念形成問題
	戦略獲得問題
	DNA配列のシグナルパターン抽出
	DNAの配座解析
	画像解析・復元
タンパク質の構造推定	

索する手法を総称して進化計算 (Evolutionary Computation) と呼ばれている。以下では、進化計算の枠組みで捉えられる代表的な手法を簡単に述べる。

進化計算の源流は一般に、1960年代にRechenbergに提案され、その後Schwefelらにより発展させられた進化戦略 (Evolutionary Strategy, ES)、同じく1960年代のL. J. Fogelに始まる進化プログラミング (Evolutionary Programming, EP) および1950年代から研究が始まり1975年のHollandの研究⁶⁾で基本的な枠組みができた遺伝的アルゴリズム (Genetic Algorithm, GA) の三つが挙げられる。これらは、いずれも生物進化をヒントとする探索、適応化手法である。しかし、これらは1980年代までお互いに遭遇することなく研究されてきた。

GAの研究グループにより1985年に米国において、国際会議第1回 International Conference on Genetic Algorithms (ICGA)⁷⁾が開催され、以降2年おきに開催されてきた (1999年からより Genetic Programming Conference と統合して現在の毎年開催される GECCO となっている)。当初 ICGA は国際会議とはいえ、ローカルな研究仲間の会議であったが、1989年ごろからESの研究者も加わりはじめ、1990年代にはGA, ESを含む進化計算の中心的国際会議となった。

一方、ヨーロッパにおいては、もう少し広い枠組みを狙いとした国際会議第1回 Parallel Problem Solving from Nature (PPSN) が Schwefel らの提唱により ES の発祥国のドイツで1990年に開催され、GAの研究者も多く参加している。以降、2年おきに開催され、現在では進化計算に関するハイレベルな国際会議の一つとしての地位を確立している。1994年になって第1回 IEEE International Conference on Evolutionary Computation (ICEC) が開催された。1999年から新たに、IEEE と Society of EP との共催の形で Congress on Evolutionary Computation (CEC) が毎年開催されている。

このように1990年代になって従来の異なっ

たコミュニティが互いに本格的に遭遇し、研究交流が国際会議を通して始まった。この交流は、例えばGAのコミュニティが重視している交叉がESに取り入れられるというように方法論的イミグレーションを伴う。

1990年代に入り、従来の生物進化をベースとする手法に加え、群れの集団行動をモデルとするアントコロニー最適化 (Ant Colony Optimization, ACO)、粒子群最適化 (Particle Swarm Optimization, PSO) や進化計算に統計的手法を融合する分布推定型アルゴリズム (Estimation of Distribution Algorithm, EDA)^{8), 9)} などの手法が進化計算の仲間として加わってきた。

各コミュニティはときには強い自己主張を行い、またときには排他的になりつつも、これらにより進化計算コミュニティの多様性が維持され、相互に影響しながら発展してきた。

以上、進化計算の発展過程を簡単に述べたが、この過程を大局的に見ると研究の大きなブレイクスルーはほぼ10年を単位として進んでいる。多くの技術の発展、例えばインターネット技術の進展を見ても、一見急速な技術変化をしているように見えるが、大局的に見ると研究の節目はほぼ10年程度要している。興味ある一致である。

Ⅲ. 進化計算の並列化方式の分類

一般に、コンピュータによる計算の並列化に当たっては、どのような単位でプログラムを並列化するかが重要となる。どのようなプログラムでも、逐次法でしか実行できない部分と並列に実行できる部分からなる。逐次法で処理されるプログラムの作業量の割合を r とし、残りの $(1-r)$ が p 個のプロセッサで並列実行されるとすると、並列化による高速化は、高々、

$$speedup = \frac{1}{r + \frac{1-r}{p}} \quad (1)$$

である。これはアムダールの法則 (Amdahl's

law) として知られている式である。式 (1) に従うと、例えば $r=0.5$ の場合、プロセッサ数を無限にしても、2倍の高速化しか得られないことになる。このことから分かるように、並列化に当たっては、逐次法でしか実行できない部分を極力0に近づけなければ、プロセッサ数を増やしても高速化は得られない。並列化による高速処理を実現するには、このことを念頭にプログラム構造の設計が重要となる。

進化計算はII章でも見たように、もともと複数の個体に関する処理手順であり並列化に適しているといわれている。一般に進化計算の並列化方式には、以下のような4つのモデルが代表的である¹⁰⁾。

- (1) 単一集団によるマスタ/スレーブモデル：このモデルでは、マスタプロセッサにより集団が管理され、スレーブプロセッサにより各個体の評価が行われる。一般に、進化計算では個体の評価に要する計算量が大きく、この部分の並列化の効果が大きい。これはいわば効率的な負荷分散方式である。
- (2) 分散進化計算モデル：このモデルでは、進化の単位である集団を複数個配置する。それぞれの集団の処理が各プロセッサに割り当てられる。各集団は各プロセッサで独立に進化するが、一定の世代間隔で各集団の個体を集団間で交換する。このモデルは、島モデルとも呼ばれ、海洋の島々での進化過程にヒントを得たものである。分散進化計算モデルでは、各集団は異なった進化をすることと定期的な集団間での個体交換により、全体として多様性を持った集団の進化が期待でき、この結果として良好な解が高速に得られることが期待できる。
- (3) セル型進化計算モデル：小さな集団を2次元あるいは3次元上に配置し、隣り合う集団間にまたがって遺伝的オペレータを適用する。これは大陸における生物集団間の交流による進化過程にヒントを得たモデルである。大規模並列計算機に向けた方式である。

(4) ハイブリッドモデル：(1)～(3)の基本モデルを組み合わせたモデルであり、多くの変形方式が考えられる。

以上のモデルに加えて、本研究では、以下の並列化モデルを検討した。

(5) 個体レベルの並列化モデル：この方式では、各個体の一連の進化オペレーションを独立したプロセッサで行わせる。

以上の5つの並列化モデルには、使用する計算機環境や用いる進化計算の方式により各種の実現方式があり、与えられた計算機環境で最適な並列化方式を見つけることが並列進化計算の研究課題である。IV章で述べるネットワーク型並列計算は、ネットワーク環境上のプロセッサを用いたマスタ/スレーブモデルに関するものである。V章で述べるマルチコアプロセッサによる並列進化計算は、マルチコアプロセッサを用いて個体レベルの並列化モデルを実現したものである。また、VI章のGPUを用いた超並列進化計算は、個体レベルの並列化と分散進化計算モデルを組み合わせたものである。

最後に並列進化計算の研究の目的を再確認して本章を締めくくろう。並列進化計算の目的は大きく次の二つにまとめることができる。第一は、決められた品質の解を高速に得ること、第二は、決められた時間内で、よりよい品質の解を得ることである。本論文では前者、すなわち、決められた品質の解を高速に得ることを目的とし、その並列化方式について述べている。

IV. Java アプレットによる ネットワーク型並列進化計算

本章では、筆者が開発した高性能ACOアルゴリズムであるカニングアントシステム(cunning Ant System, cAS)をネットワーク環境で並列化した並列cAS(parallel cAS, p-cAS)を2次割当て問題(Quadratic Assignment Problem, QAP)の解法に適用した方式ならびに結果について述べる。

1. カニングアントシステム(cAS)と2次割当て問題(QAP)

ACOは、実際のアリの採餌行動における経路生成過程にヒントを得た探索手法であり、巡回セールスマン問題(Traveling Salesman Problem, TSP)など多くの組合せ最適化問題に適用され、有効な結果が得られている。アリはフェロモンを介したコミュニケーションを行いながら群れで行動し、ある種の秩序を形成する。ACOでは、この秩序形成過程を探索に用いる。

ACOの基本モデルは、DorigoらによるAnt System(AS)¹¹⁾と呼ばれるアルゴリズムに代表される。その後、多くの改良型ACOアルゴリズムが提案されている¹²⁾。

通常のACOアルゴリズムでは、各エージェント(アリ)により生成される解はフェロモン濃度 $\tau_{ij}(t)$ に基づいて確率的に生成される。ただし、 $\tau_{ij}(t)$ は、解の順序表現におけるノード i, j 間のフェロモン濃度である。著者が提案したcAS^{13), 14)}では、カニングアント(cunning ant, c-ant)と呼ぶエージェントを導入している。c-antは、従来のエージェントと異なり、新しい解候補を生成の際に、過去の探索において存在する解の一部(部分解)を借用し、解の残りの部分の生成にのみ従来と同様フェロモン濃度を利用する(図2参照)。このようなエージェントを抜け目のないアリという意味でカニングアントと呼んだ。狡をするエージェントという意味も言外に込められている。

cASは経路を一部借用するという方法により、探索過程においてポジティブフィードバックによる集中化を緩和し、初期収束を抑制する。この効果によりcASは優れた探索性能を示している。

さて、ここで解く問題は2次割当て問題(Quadratic Assignment Problem, QAP)である。QAPは、 L 個からなる部門を L 箇所のロケーションに式(2)で定義される値が最小になるような順列 ϕ を求める問題である。これは、企業における事業所や工場の立地の最適配置問

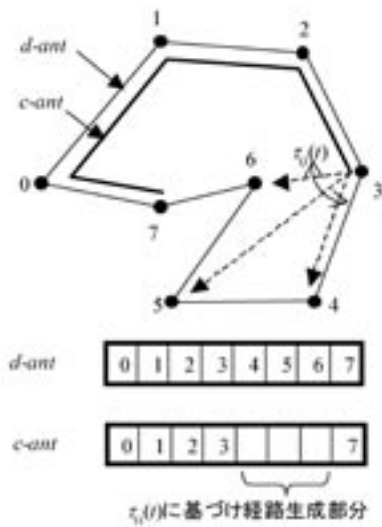


図2 カニングアントの解生成法

題やビルの部門の最適配置問題などを一般化したもので、組合せ最適化問題の中では、最も困難な問題に属する。

$$f(\phi) = \sum_{r=1}^L \sum_{s=1}^L b_{rs} a_{\phi(r)\phi(s)} \quad (2)$$

式(2)において、 $A = (a_{pq})$ および $B = (b_{rs})$ はそれぞれ $L \times L$ の距離マトリックスとフロー

マトリックスであり、 ϕ は $\{1, 2, \dots, L\}$ の順列である。マトリックス A と B は、それぞれ、場所 p, q 間の距離、部門 r, s 間の流量を表している。QAP は評価式が距離と流量の積となっており、両情報が相互に絡み合うため、TSP に比較して格段に困難な問題である。

2. ネットワーク環境における cAS による QAP の解法

1) cAS の並列化のためのネットワーク基本モデル

本研究における cAS の実装には Java を用いている。Java には RMI や Jini などネットワーク環境における分散処理のためのパッケージが準備されているが、本研究では cAS の並列化の基本アーキテクチャとして、クライアント/サーバモデルに基づくアプレットベースの基本モデルを構築した。この構造を図3に示す。

クライアントで実行されるプログラムはアプレットのクラスとしてサーバから転送される。すなわち、クライアントのプログラムは、ブラウザから Java アプレットとして実行される。クライアントとサーバ間の通信は、TCP ソケットプログラミングを用い、両者の情報交換は Serializable インタフェースを実装し

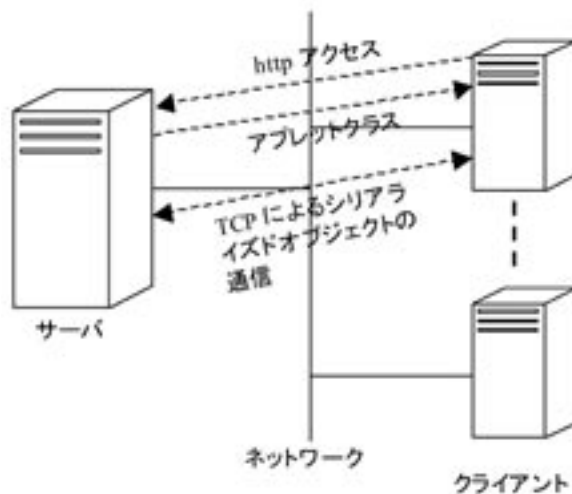


図3 cAS の並列化のためのネットワーク基本モデル

たオブジェクトを交換して行う。具体的には、ObjectInputStream, ObjectOutputStream オブジェクトの交換として通信する（このオブジェクトの詳細については Java 2 API¹⁵⁾ 参照）。

以上の方式により、ネットワークに接続され、Java アプレットの実行環境を有する任意のクライアントがサーバに http プロトコルによりアクセスすることで、簡単に並列計算環境を構築することができる。

なお、クライアント側で QAP の距離マトリックスやフローマトリックスなどのデータファイルの参照が必要となる。しかし、アプレットではファイル参照ができないので、これらのファイルは Java の一つのクラスに変換しておくことを前提としている。これにより各クライアントで必要となるファイルのデータは http プロトコルによりクライアントからアクセスできる。なお、サーバには Web サーバとしての機能が必要とする。

2) *p-cAS* の実装

表 2 は、*cAS* に局所探索としてタブーサーチを結合した場合の処理時間の分布を各サイズのテスト問題に対して示したものである。実行環境は 2 個の Opteron 280 (dual core, 2.4GHz, Socket940) プロセッサからなる WindowsXP マシン (主メモリ: 2 GB) である。この表より、探索の総時間のほぼ 99% 以上が局所探索に費

やされていることが分かる。そこで、今回の *p-cAS* では、局所探索を並列化して高速化を図るという方式を、図 3 で示したネットワーク基本モデルを用いて実装した。

p-cAS の実行環境は上述の Opteron マシンを 2 台用い (以下、それぞれ、マシン A, マシン B と呼ぶ)、両マシンを 1000BASE-T スwitchングハブで接続する。それぞれのマシンは 4 個のプロセッサを有する。マシン A をサーバとして用い、クライアント機能はマシン B に割り当てる。マシン A の Web サーバ用ソフトには、Apache¹⁶⁾ を用いた。

図 4 は、上記の環境での局所探索の並列化の方式である。用いるアント数 m (進化計算では一般に集団サイズ) は $m=5$ である。そこで、4 個の *c-ant* の局所探索をクライアントで行うこととし、このためマシン B には Java アプレットが実行できるブラウザを、4 つの独立したプロセスとして起動する。厳密には各ブラウザプロセスは、OS により 4 つのプロセッサ間で動的に切り替わって実行されるが、これらのプロセスは 4 つのプロセッサに論理的に割り当てられると考えてよい。また、残り 1 つの *c-ant* の局所探索はサーバプロセッサに割り当てる。なお、図 4 では、クライアントプロセッサ数を N_c として一般化して記述している。上記の説明は、同図において、 $N_c=4$, $m=5$ とした場合に相当し、局所探索を適用するために各プロ

表 2 局所探索を用いた *cAS* における QAP での計算時間の分布
(単位はミリ秒で、25 回の実験の平均値を示している)

QAP	サンプリング	ローカルサーチ	τ_{ij} の更新	その他	合計
tia40b (%)	3.8 0.1%	6833.8 99.7%	1.3 0.0%	12.3 0.2%	6851.2 100.0%
tia50b (%)	4.9 0.0%	13465.8 99.8%	2.3 0.0%	13.3 0.1%	13486.2 100.0%
tia60b (%)	7 0.0%	23448.7 99.9%	3 0.0%	15.7 0.1%	23474.4 100.0%
tia80b (%)	10.6 0.0%	56688.7 99.9%	5.9 0.0%	19.9 0.0%	56725 100.0%
tia100b (%)	14.8 0.0%	114411.7 100.0%	9 0.0%	21.4 0.0%	114456.9 100.0%
tail150b (%)	29.9 0.0%	422079.8 100.0%	19.9 0.0%	46.1 0.0%	422175.6 100.0%

セッサに割り当てられるエージェントの数 e は 1 である。

3. 結果と考察

図 5 に p -cAS の結果と考察結果をまとめる。データは 25 回の実験の平均である。ここで、 $speedup$ は、(並列化しない場合の cAS の実行時間) / p -cAS の実行時間を示す。クライア

ントとサーバ間の通信オーバーヘッド (T_{comm}) が無ければ $speedup$ の値は 5 に近い値となる。しかし同図に示しているように、通信オーバーヘッド (T_{comm}) が存在する。この通信オーバーヘッドのために tai40b および tai50b では $speedup$ 値は 1 より小さく、また tai60b でほぼ 1 となっている。一方、より大きな問題である tai80b, tai100b および tai150b では、 $speedup$

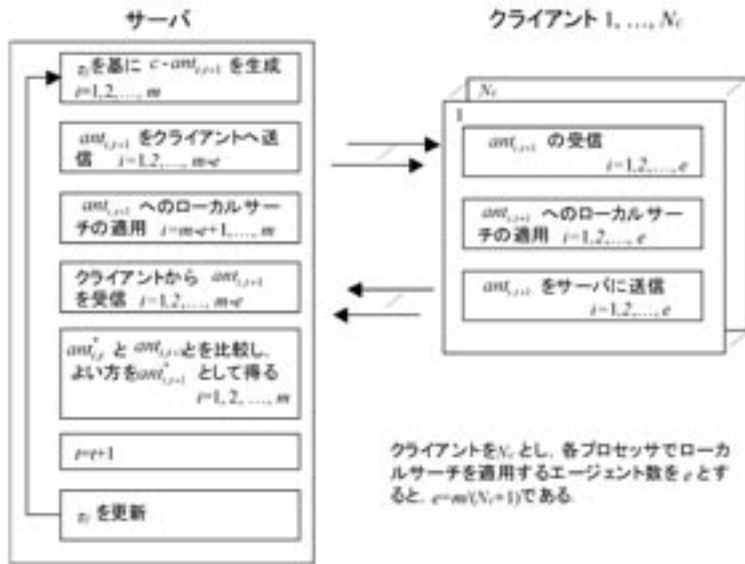


図 4 並列化cAS (p -cAS) の概要

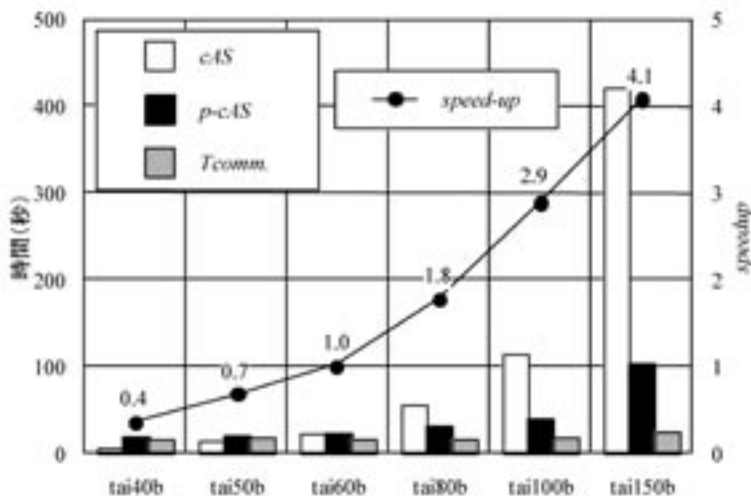


図 5 p -cASによる高速化の結果

新個体のメモリ領域にコピーする (Partial Solution Segment, PSS)。コピーの開始位置 p_{top} はランダムに決定する。

- (iii) 残りの l_s 個のノードは, EHM をもとに式 (4) に基づいてサンプリングする (Sampling Based Segment, SBS)。

先の EHBSA の研究¹⁷⁾では, l_s の決定に n カットポイント法を用いていたが, この方法では, l_s の平均値 $E(l_s)$ が L/n ($n = 2, 3, \dots$) に制限されるという問題があり, 改良型 EHBSA では, $E(l_s)$ が $L \times \gamma$ ($\gamma \in [0, 1]$) となるパラメータ γ を導入した。

$$f(l_s) = \begin{cases} \frac{1-\gamma}{L\gamma} \left(1 - \frac{l_s}{L}\right)^{\frac{1-2\gamma}{\gamma}} & \text{for } \gamma \in (0, 0.5] \\ \frac{\gamma}{L(1-\gamma)} \left(\frac{l_s}{L}\right)^{\frac{2\gamma-1}{1-\gamma}} & \text{for } \gamma \in (0.5, 1] \end{cases} \quad (3)$$

(iii) のサンプリングでは, 現在位置のノードが i であるとき, i の次の位置のノード j は, 次式の確率 p_{ij}

$$p_{ij} = \begin{cases} \frac{e_{ij}}{\sum_{s \in F(i)} e_{is}} & \text{if } j \in F(i) \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

で決定される。ただし, $F(i)$ は, 現在のノード i からつぎに選べるノードの集合である。この式に基づくサンプリングの計算量は一般に

$O(L^2)$ になり, 問題サイズが大きくなるとサンプリングに要する時間が問題となる。このような場合, 候補リスト¹⁸⁾を使うと計算量を $O(L)$ に近づけることができる。

EHBSA では, 1 世代に一個体を生成するモデルを用いたが, 本研究では, 図 8 に示すように, 1 世代で N 個体を生成する方式とし, 効率的な世代交代モデルに改良している。同図において, 新しい個体 I'_i は, 集団 $P(t)$ の個体 I_i をテンプレートとして用いる ($i = 1, 2, \dots, N$)。各 i のペア (I_i, I'_i) ($i = 1, 2, \dots, N$) を比較し, 良い個体を次の世代 $P(t+1)$ のメンバーとする。 I_i と I'_i とを比較するこの方式は, Mahfoud の deterministic crowding 法¹⁹⁾ のように多様性維持に効果がある。また, この方式では, 図 7 で示した γ も重要な設計パラメータとなる。

3) eEHBSA のシングルスレッド性能

eEHBSA の並列化されていない場合の基本性能を見るために, TSP を用いて評価した結果を述べる。ここでは, (1) 小規模問題: berlin52, pr76, (2) 中規模問題: att532, rat783, (3) 大規模問題: fl3795, rl5934 の 3 つの規模の問題を取り上げる。実験条件は以下の通りである。(1) の小規模問題に対しては, 集団サイズを $2L$, 最大解生成数を $2L \times 10000$ とし, 局所探索は適用しない。(2) の中規模問

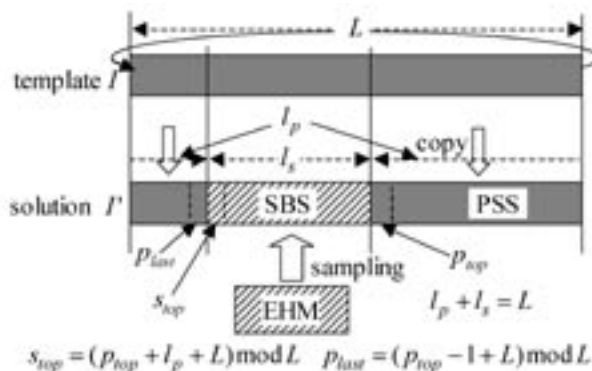


図7 個体の生成法

題に対しては、集団サイズを $L=15$ 、最大解生成数を $L \times 1000$ とし、局所探索には 3OPT を適用する。(3) の大規模問題に対しては、集団サイズを 4 とし、最大解生成数を $L \times 10$ とし、局所探索には TSP における最強の局所探索とされている Lin-Kernighan (LK) ヒューリスティックを適用する。

プログラミング言語には Java を用い、マシンには Core i7 965 プロセッサを用いる。LK コードには Concorde²⁰⁾ を用い、JNI を用いて eEHBSA コードと結合した。それぞれの問題に対して γ を 0.1 から 1.0 まで 0.1 刻みで実験を行った。各実験における試行回数は 20 回とした。各規模の問題に対して、#OPT (20 回の実験中で最適解を発見した回数) および T_{avg} (最適解を発見した実験において、最適解を発見するのに要した時間の平均) による結果をそれぞれ、図 9 に示す。

これらの結果から eEHBSA では、 γ の適切な設定が重要であることが分かる。また、いずれの規模の問題においても、 $0.2 \leq \gamma \leq 0.4$ において #OPT=20 であり、また、 T_{avg} も小さい値となっていることが分かる。

2. マルチスレッドプログラミングの実現法

eEHBSA の並列化方式として、同期マルチスレッドモード (synchronous multi-thread mode, SMTM) と非同期マルチスレッドモード

(asynchronous multi-thread mode, AMTM) の 2 つの並列化スレッドプログラミングを考える。使用言語は Java である。使用プロセッサは、5.1.3 項で述べた Core i7 965 である。このプロセッサは 4 つのコアを内蔵している。BIOS においてハイパースレッディング機能を有効に設定すると OS から見かけ上 8 個のコアが存在するが、ここではその機能は用いないとした。したがって、本研究では、4 つのスレッドを生成して eEHBSA の並列実行を評価する。eEHBSA の繰り返し過程は以下となる。

- (1) $t \leftarrow 0$. 個体 I_i ($i = 1, 2, \dots, N$) からなる初期集団 $P(t)$ を生成。
- (2) $P(t)$ の個体を評価 (局所探索を適用する場合は適用後評価し、 I_i も変更)。
- (3) $P(t)$ から EHM を作成。
- (4) 新個体 I'_i ($i = 1, 2, \dots, N$) を、EHM および I_i の部分分解から生成。
- (5) 各 i ($i = 1, 2, \dots, N$) に対して I_i と I'_i とを比較し、 I'_i の方がよければそれを I_i と置き換え、集団 $P(t)$ を更新する。
- (6) $t \leftarrow t + 1$ 。
- (7) 終了条件を満たせば終了、満たさなければ (3) へ。

各個体の一つの処理が一つのスレッドの処理となる。スレッドは 4 つとしているので、集団サイズ分の処理をこれらの 4 つのスレッドが処理を繰り返す。SMTM では、各ステップの処

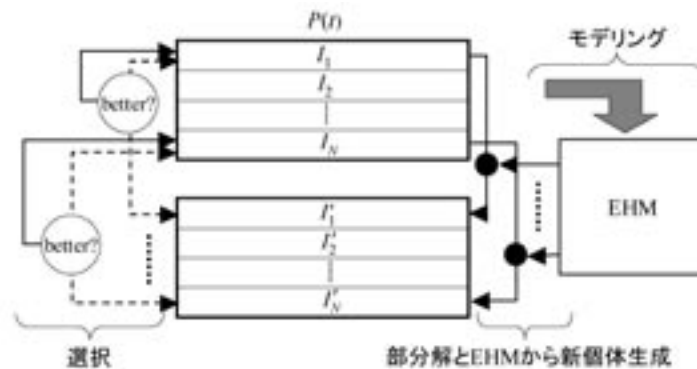


図 8 世代交代モデル

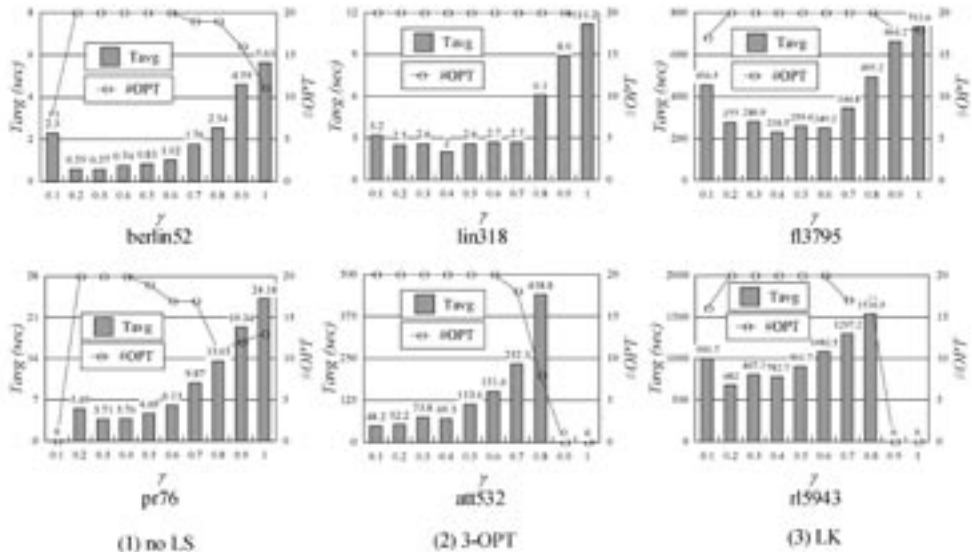


図9 eEHBSAのシングルスレッド性能

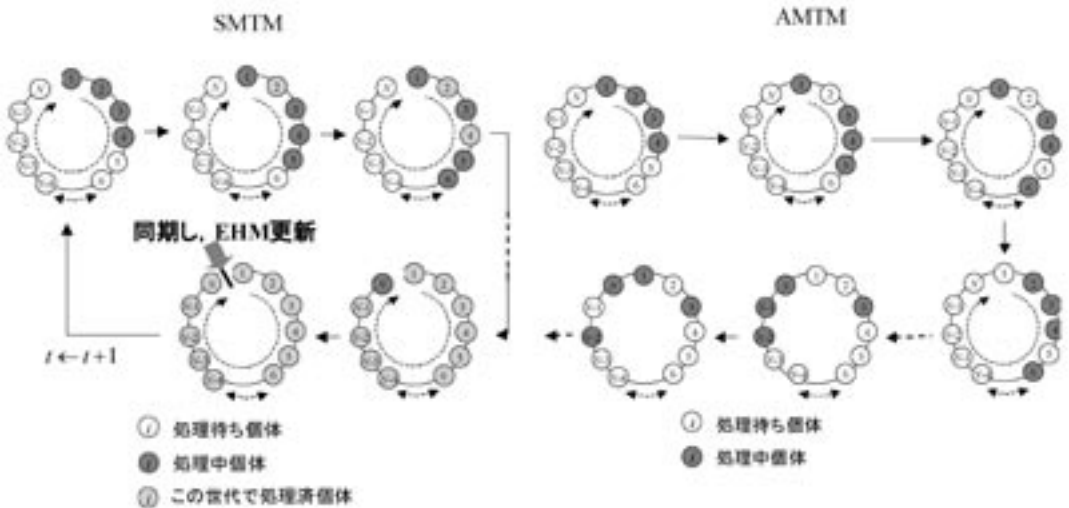


図10 AMTMおよびAMTMにおける4つのスレッドによる個体の処理過程の例

理を同期して行う。それに対して、AMTMでは上記(3)まで同期モードで動作し、それ以降は独立に未処理の個体を並列実行する。このとき、EHMの更新では、全ての個体が更新されてからではなく、 I_i と I'_i の入れ替えとなったときには、現在のEHMから I_i の属するエッジを取り出し(それらのエッジの数を-1する)、 I'_i に属するエッジを加える(それらのエッジ

の数を+1する)。この仕組みにより、AMTMでは処理の待ち合わせに伴う遅延は発生しない。ただし、処理が非同期に行われるので、各個体に対して各ステップの処理は完全な世代単位には行われなくなる。図10に4つのスレッドによる各個体の処理過程をそれぞれSMTMおよびAMTMに対して示す。

3. 結果と考察

各種サイズの TSP を用いて、2つの並列化方式の結果を表3に示す。実験回数は20とした。この実験では20回の実験で全て最適解を発見している (#OPT=20)。したがって、 T_{avg} は、最適解を見つけるまでに要した時間(秒)の20回の実験の平均である。 $speedup$ は、並列処理を行わない場合の T_{avg} をマルチスレッド方式による場合の T_{avg} で割ったものである。同表には、各 T_{avg} の95%信頼区間ならびに SMTM と AMTM の $speedup$ の比較の p -値も示した。

SMTMの結果を見ると、局所探索を用いない小規模問題と局所探索を用いる中、大規模問題で大きな違いがあることが分かる。中、大規模問題では、速度比が3.1~3.7であるのに対して、局所探索を用いない小規模問題では速度比が1.5~2.5と高速化の度合いが小さくなっている。一方、AMTMの結果を見ると、中、大規模問題ではSMTMの結果と大きな違いが

なく、同程度の高速化の結果が得られている。大きな違いは、小規模問題の場合である。この場合にも、SMTMの場合と異なりAMTMでは3.4~4.4と高速化の結果が得られている。局所探索を適用しない場合にはサンプリングに要する時間の割合が大きく(図11参照)、また、eEHBSAではこの時間は確率的に変動する。したがって、SMTMではスレッド間の待ち時間が長くなる。AMTMではこのような待ち時間はない。

VI. GPUを用いた超並列進化計算の試みと課題

ここ2,3年前より、いわゆるグラフィックボードとして一般のPCに用いられている画像処理用演算装置(Graphics Processing Unit, GPU)を用いて流体力学、医療用画像分析、統計データ処理などの科学技術計算の並列計算を実現する並列計算(GPU Computation, 以下

表3 マルチコア計算機による並列進化計算の結果

テスト問題	局所探索	シングル eEHBSA		並列 eEHBSA						p-値
		T_{avg} (sec)	SE	SMTM		$speedup$	AMTM		$speedup$	
				T_{avg} (sec)	SE		T_{avg} (sec)	SE		
		信頼区間		信頼区間		信頼区間				
oliver30	no	0.082	0.004	0.05	0.08	1.5	0.02	0.04	3.4	5.93E-07
		[0.074, 0.091]		[0.045, 0.063]			[0.020, 0.004]			
gr48		0.6	0.04	0.29	0.01	2.1	0.16	0.01	3.8	1.31E-09
		[0.53, 0.67]		[0.26, 0.31]			[0.15, 0.17]			
berlin52		0.57	0.03	0.29	0.02	2	0.13	0	4.3	1.58E-07
		[0.51, 0.64]		[0.25, 0.33]			[0.12, 0.14]			
pr76		3.71	0.14	1.46	0.07	2.5	0.86	0.11	4.3	8.01E-04
		[3.40, 4.01]		[1.31, 1.61]			[0.74, 1.20]			
lin318		2.57	0.64	0.84	0.06	3.1	0.76	0.04	3.4	0.31
		[1.23, 3.91]		[0.70, 0.98]			[0.68, 0.84]			
pcb442	12.49	2.21	3.12	0.28	4	2.92	0.26	4.3	0.606	
	[7.85, 17.12]		[2.54, 3.70]			[2.38, 3.47]				
att532	73.84	12.29	21.84	3	3.4	19.41	2.78	3.8	0.549	
	[48.11, 99.56]		[15.57, 28.11]			[13.81, 25.02]				
rat783	139.48	7.57	38.77	2.72	3.6	38.37	2.75	3.6	0.918	
	[123.63, 155.32]		[33.07, 44.47]			[32.62, 44.12]				
fl3795	280.88	34.7	105.55	19.08	2.7	85.51	12.28	3.3	0.384	
	[208.24]		[65.60, 145.50]			[59.81, 111.21]				
rl5934	807.27	76.04	252.44	26.51	3.2	209.62	16.91	3.9	0.183	
	[648.12, 966.43]		[196.96, 307.93]			[174.22, 245.01]				

SE: 標準誤差

GPU 計算)の研究が注目され、CPU に対して数 10 ~ 100 倍の高速化が図られたという有望な結果が報告されている^{21), 22)}。GPU 計算による進化計算の並列化の研究も今後大きく発展するものと思われる。

2009 年 7 月の ACM 主催の進化計算の国際会議 GECCO-2009 では、進化計算への GPU 計算の適用に関するワークショップ 2009 Computational Intelligence on Consumer Games and Graphics Hardware (CIGPU-2009) が開催され、筆者らの研究も含めて 8 件の研究が発表された。2010 年には、IEEE 主催の進化計算の国際会議 CEC において、CIGPU-2010 が開催される。

本章では、2 次割当て問題 (quadratic assignment problem, QAP) の解法に GPU 計算を用いて並列進化計算を実行した筆者らが試み

た研究結果⁵⁾について述べ、今後の研究課題について考察を行う。

1. GPU 計算による超多スレッドプログラミングの一実現法

本研究で用いた GPU は、NVIDIA GeForce GTX285 である。図 12 に本 GPU のアーキテクチャを示す。Thread Processor (TP) 演算装置の単位であり、8 個の TP を単位として 1 つの Multi-processor (MP) を構成している。MP 内の TP は、レジスタ並みの高速アクセスが可能な Shared Memory (SM) を介してデータを共有することができる。SM のサイズは 16KB である。MP の数は 30 である。異なる MP に属する TP 間のデータ共有は、VRAM 経由となり速度が遅くなる。プログラミング環境としては、CUDA (compute unified device

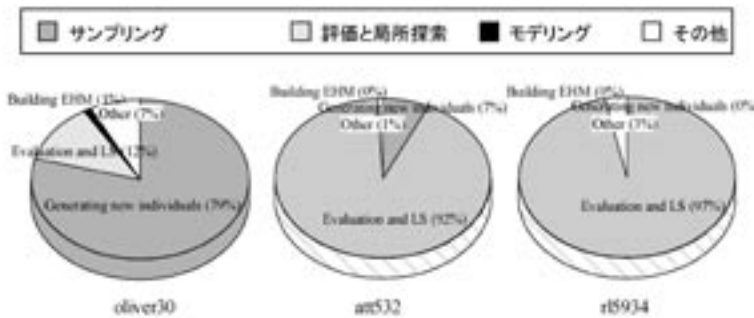


図11 eEHBSAにおける処理時間の分布

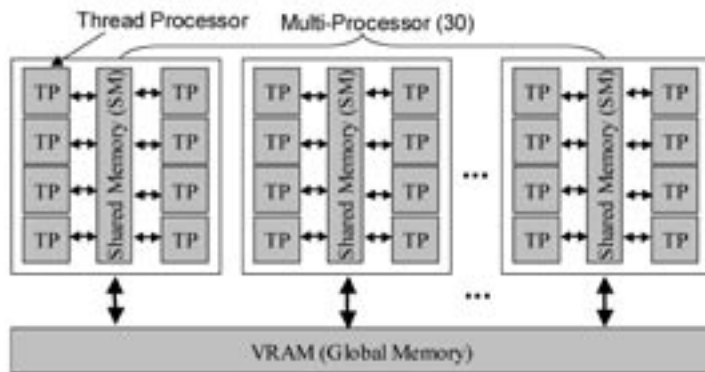


図12 NVIDIA GeForce GTX285のアーキテクチャ

architecture) と呼ぶ C 言語を拡張した統合開発環境があり、コンパイラやライブラリなどから構成されている²¹⁾。Windows 環境では、Microsoft Visual Studio に組み込んで使うことができ、本研究では、この方法を用いた。

CUDA では、合計 $8 \times 30 = 240$ 個の TP を動作待ち時間なく有効に動作させる（メモリアクセス遅延を隠す）ために数千という超多スレッドプログラミングを前提としている。このため、*grid*, *block*, *thread* という概念を導入し、物理的な構成を直接的に意識せずに超多スレッドプログラミングができる環境を提供している。スレッドの切り替えはハードウェアにより行われ、切り替え時間のオーバーヘッドは発生しない。この考え方は Pentium IV で導入されたハイパースレッドの機構と類似している。

GPU 計算を進化計算の並列化に応用する際の大きなネックは、SM のサイズにある。個体の集団を VRAM におくと、各 TP からのメモリアクセスの遅延が大きいため、高速化に不利となる。このため、SM (16KB) の活用が重要となる。本研究では、一つのサブ集団 $P(t)$ を SM に置くという分散型 GA モデルを考案し、高速化を図ることを考えた。QAP は大きな問題でも問題サイズが高々 150 程度である。したがって、順列表現である染色体の遺伝子には *unsigned char* 型を使用することができ、個体表現に要する SM の消費量を節約できる。

各サブ集団の世代交代モデルでは、図 8 で述べたのと同様、一つの個体の処理が CUDA 環境で 1 つのスレッドとなるよう工夫している。サブ集団サイズを N とすると、SM の消費量は、 $2L \times N$ バイトとなる。今回は $N = 128$ に固定し、最大問題サイズを $L = 56$ とした。交叉には PMX オペレータ、突然変異には 2 つの任意の遺伝子座の値を入れ替える SWAP オペレータを使用した。局所探索は適用していない。各サブ集団の世代交代モデルを図 13 に示す。

各 MP におけるサブ集団の個体は、図 14 に示すように 500 世代毎に VRAM を介してシャッフルする。この意味で、GPU 内の進化モデルは、分散 GA における島モデルの一形態といえる。個体数の総数は、 $128 \times 30 \times k$ ($k = 1, 2, \dots$) であり、これが CUDA プログラミング環境における総スレッド数になる。

2. 結果と考察

QAP の各テスト問題に対して、GPU 計算によるスレッド並列化の結果と、Core i7 965 プロセッサにおいてシングルスレッドで実行した結果とを、最適解を発見するのに要した時間 (T_{avg}) で比較したものを表 4 に示す。ここで、CPU における GA のモデルとして、論理的に GPU 計算の場合と同じとした場合 (GA-1) と島間の個体の交換をチューニングしたモデル

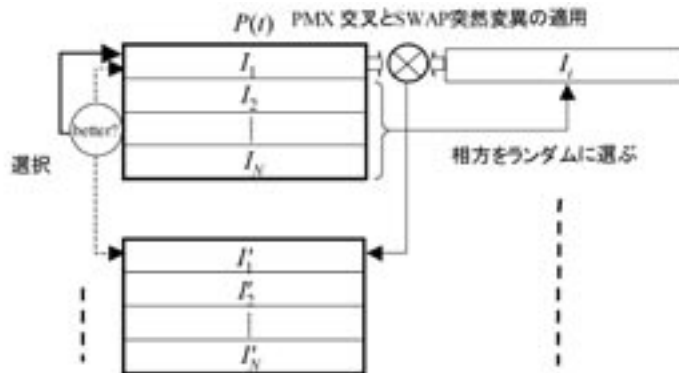


図13 GPU計算における各サブ集団の世代交代モデル

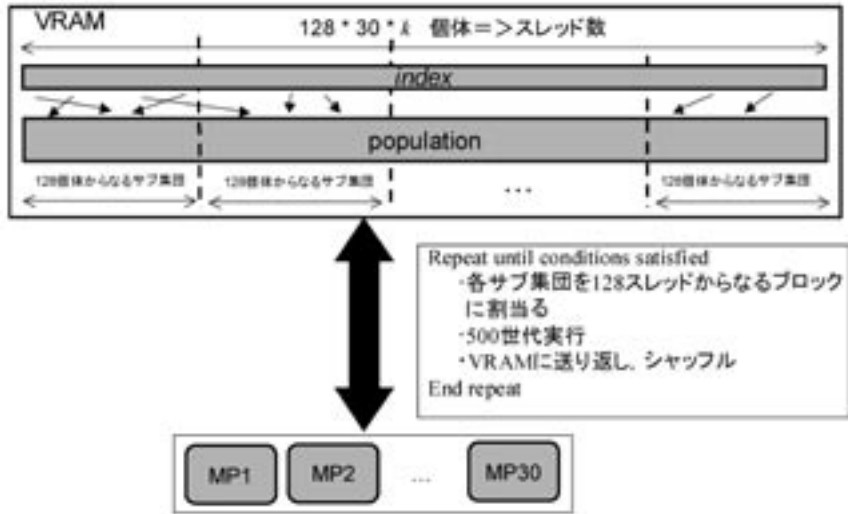


図14 GPU計算における進化計算の構成

(GA-2) の結果を示した。当然 GA-2 の結果の方が、GA-1 の結果よりも優れている。GPU 計算の GA-2 に対する *speedup* の値として 3 ~ 12 という結果が得られた。

画像処理や流体力学などへの応用では、CPU の数 10 倍 ~ 100 倍以上の高速化が期待できる¹⁰⁾。今回の結果はこれらの分野と比較すると十分とはいえないが、今後計算モデルの改良により、さらなる速度の向上は十分期待できる。以下では、進化計算において GPU 計算を用いる際に課題となる点について考察し、今後への研究課題を述べたい。

進化計算に GPU を用いる際に課題となる事項としては以下をあげることができる。

- 進化計算では、処理が確率的に行われ、これにより多くの分岐命令が現れる。GPU ではスレッドはブロックごとに MP に割り当てて実行されるが、各ブロック内のスレッドはワープに分割され (GTX285 の場合、32 スレッド単位) て SIMD 風に行われるので、スレッド間で True と False の処理がシリアルに行われ、このとき MP 内の TP にアイドル状態が発生する可能性がある。また、分岐がなく同じ処理が行われる場合でも、各スレッド間での処理時間が異なると、全体の処理時間は長い方に合合わせられ、TP にアイドル状態が発生する。したがって、このような問題を少なくする計算モデルを工夫する必要がある。

表 4 GPU計算による並列進化計算の結果

QAP テスト 問題	GPU 計算				シングルスレッドによる CPU 計算								<i>speedup</i>	
	総個体数	#OPT	T_{avg} (sec)	SE	GA-1				GA-2					
					総個体数	#OPT	T_{avg} (sec)	SE	総個体数	#OPT	T_{avg} (sec)	SE	GA-1	GA-2
tai20b	128×30×1	10	0.064	0.001	128×30×1	10	0.428	0.012	128×30×1	10	0.422	0.013	6.6	6.7
tai25b	128×30×1	10	0.169	0.005	128×30×1	10	1.386	0.043	128×30×1	10	1.286	0.046	7.6	8.2
kra30a	128×30×5	10	2.002	0.551	128×30×2	9	9.651	1.514	128×30×4	9	11.87	1.038	5.9	4.8
kra30b	128×30×5	9	1.332	0.244	128×30×5	8	23.399	3.831	128×30×4	9	16.745	3.721	12.6	17.6
tai30b	128×30×3	10	0.947	0.182	128×30×3	10	22.649	2.16	128×30×1	10	7.203	1.984	7.6	23.9
tai35b	128×30×4	10	2.51	0.234	128×30×3	10	22.649	2.16	128×30×1	10	7.203	1.984	2.9	9
ste36b	128×30×4	10	3.337	0.334	128×30×4	10	33.274	4.13	128×30×2	10	14.675	1.213	4.4	10
tai40b	128×30×1	10	1.088	0.028	128×30×1	9	6.016	0.162	128×30×1	10	5.811	0.153	5.3	5.5

SE: 標準エラー

ある。

- 各MPで共有される高速なメモリであるSMが一般に小さく、大きな問題に適用する場合に問題分割などの十分な工夫が必要となる。各MP間のデータ共有はVRAM経由となる。基本演算命令の実行時間は4クロックであるのに対して、VRAMアクセスの遅延は400から600クロックである(GTX285の場合)。このためMP間でのデータ共有や交換の方法に工夫が必要となる。
- 各スレッドで用いる乱数系列の管理に工夫を要する。
- スレッドで用いるローカル変数にはレジスタが割当てられ、高速な処理が行える半面、進化計算で多用される配列がローカル変数として使えない。したがって、ローカル変数の効率的な使用には十分な工夫が必要となる。

以上の点に関しては今後のGPUの進歩にも期待できる。事実、NVIDIAは公表している次期GPU(Fermi)では、SMの容量増加やVRAMアクセスに対するキャッシュの導入などが図られている²³⁾。しかし、GPUを並列進化計算に用いる研究テーマとしては、数百もの演算ユニットを如何に有効活用するかということが本質であり、進化計算のモデルまでさかのぼった研究を行うことが重要である。

VII. むすび

以上本論文では、一般的なユーザにおけるコンピュータ環境で、筆者が取り組んでいる並列化による進化計算の高速化に関する三つのアプローチを取り上げ、並列化の効果を述べた。また、今後の研究課題について考察した。

いずれも有効な結果が得られたが、特にGPUによる並列進化計算は、高速化の度合いも大きく、今後の研究に期待が持てると思われる。現時点で利用可能なGPUはハードウェアの制限もあり、並列進化計算手法として必ずしも十分であるとはいえないが、この分野でのハードウェアの技術進歩はCPU技術の進歩に

劣らず著しいものがある。したがって今後は、この進歩に十分期待でき、そのような進歩を利用する並列化手法の研究が一層進むものと思われる。

CPUとGPUとでは、それぞれ本来の利用目的も異なることから当然得て不得てな点を持っている。したがって、全ての計算をGPUで行わせるのは必ずしも効率的とはいえない。この観点からは、複数のGPUによる負荷分散やGPUとマルチコアCPUによる並列計算との機能分担方式、さらにはネットワークで結合した並列計算との融合など、新しい並列進化計算のモデルの検討など重要な研究課題がある。今後これらの研究課題に鋭意取り組んでいきたい。

謝 辞

本研究を進めるにあたり、前学長の大槻眞一先生にはご支援とともにいつも励ましのお言葉をいただきました。ここに紙面を借りて深く感謝の意を表します。

GPU計算の研究に関しては、大阪府立大学藤本典幸教授との共同研究の成果の一部です。なお、本研究の一部は、文部科学省基盤研究(C)、課題番号19500199の補助金に基づいて行われました。

参考文献

- 1) 伊藤：N700系新幹線車両の概要～遺伝的アルゴリズムを用いた先頭車両形状設計～、計測自動制御学会北陸支部講演会(2009)。
- 2) S. Tsutsui：Cunning Ant System for Quadratic Assignment Problem with Local Search and Parallelization, Proc. of the 2nd Intl. Conf. on Pattern Recognition and Machine Intelligence, pp. 269-278, Springer LNCS 4815 (2007)
- 3) 筒井 茂義, 劉 力綺, 小島 基伸：カニングアントを用いたACOの2次割当て問題への適用とその並列化, 情報処理学会論文誌：数理モデル化と応用, Vol. 49, No. SIG 4 (TOM 20), pp. 45-56, 情報処理学会 (2008)。
- 4) S. Tsutsui：Parallelization of an Evolutionary

- Algorithm on a Platform with Multi-core Processors, Proc. of the 9th Intl. Conf. on Artificial Evolution (EA'09), Springer (2009).
- 5) S. Tsutsui and N. Fujimoto : Solving Quadratic Assignment Problems by Genetic Algorithms with GPU Computation : A Case Study, Proc. of the GECCO 2009 Workshop on Computational Intelligence on Consumer Games and Graphics Hardware CIGPU-2009, pp. 2523-2530, ACM (2009).
- 6) J. H. Holland : Adaptation in Natural and Artificial Systems, The University of Michigan Press (1975).
- 7) J. J. Grefenstette (Ed). Proc. of the First Intl. Conf. on Genetic Algorithms. Lawrence Erlbaum Associates (1985).
- 8) H.Mühlenbein and G. Paaß : From recombination of genes to the estimation of distribution I. binary parameters, Proc. of the Parallel Problem Solving from Nature (PPSN IV), pp. 178-187 (1996).
- 9) P. Larranaga and J.A. Lozano (eds) : Estimation of distribution algorithms, Kluwer Academic Publishers (2002).
- 10) E. Cantu-Paz : Efficient and Accurate Parallel Genetic Algorithms, Kuwer Academic Publishers (2000).
- 11) M. Dorigo, V. Maniezzo, and A. Colorni : The Ant System : Optimization by a Colony of Cooperating Agents, IEEE Trans. on SMC-PartB, Vol. 26, No. 1, pp. 29-41 (1996).
- 12) M. Dorigo and Stützle, T. : Ant Colony Optimization, MIT Press, Massachusetts (2004).
- 13D3) S. Tsutsui : *cAS* : Ant Colony Optimization with Cunning Ants, Proc. of the 9-th Intl. Conf. on Parallel Problem Solving from Nature (PPSN IX), pp. 162-171, Springer LNCS (2006).
- 14) 筒井 : *cAS* : カニングアントを用いた ACO の提案, 人工知能学会論文誌, Vol. 22, No. 1, pp. 29-36, 人工知能学会 (2007).
- 15) Sun Microsystems, Inc. : Java 2 Platform, Standard Edition, v1.4.2 at API Specification. java.sun.com/j2se/1.4.2/docs/api.
- 16) Apache Software Foundation : Apache HTTP Server Project. httpd.apache.org.
- 17) 筒井 : エッジヒストグラムを用いる順序表現向き確率モデル GA の提案, 人工知能学会論文誌, Vol. 18, No. 4, pp. 173-182, 人工知能学会 (2003).
- 18) J. L. Bentley : Fast algorithms for geometric traveling salesman problems, ORSA Journal on Computing, Vol. 4, pp. 387-411 (1992).
- 19) S. Mahfoud : A Comparison of Parallel and Sequential Niching Methods, Proc. of the 6-th Intl. Conf. on Genetic Algorithms, Morgan Kaufmann (1995) .
- 20) Concorde TSP Solver, <http://www.tsp.gatech.edu/concorde.html> (2006) .
- 21) NVIDIA : CUDA Programming Guide 2.3, www.nvidia.com/object/cuda_develop.html (2009).
- 22) 青木 : フル GPU による CFD アプリケーション, 情報処理, Vol. 50, No. 2, pp. 107-115 (2009).
- 23) NVIDIA : Next Generation CUDA Architecture, www.nvidia.com/object/fermi_architecture.html (2009).