

2R-07 遺伝的交叉を用いた並列SA(分散メモリ型並列機への実装モデルの検討)

廣安 知之[†] 三木 光範[†] 角 美智子[†] 小掠 真貴^{††}

[†]同志社大学工学部 ^{††}同志社大学大学院

1 遺伝的交叉を用いた

並列シミュレーテッドアニーリング

遺伝的交叉を用いた並列シミュレーテッドアニーリング (Parallel Simulated Annealing using Genetic Crossover : PSA/GAc) は、図 1 に示すように、並列に実行している各 SA の解の伝達時に、遺伝的アルゴリズム (GA) のオペレータである遺伝的交叉を用いたものである [1].

このモデルでは、独立して複数の SA によって解探索を行う。一定ステップ後、解の伝達時に並列に実行している SA からランダムに親として 2 探索点を選択し、設計変数間交叉を行い、子と呼ぶ新しい探索点を生成する。設計変数間交叉とは各設計変数の間でのみ交叉を行うことをいう。もとの親と生成した子との 4 探索点のうち評価値の高い 2 探索点を選択して、この 2 点から次の探索を続けるという方法である。そのため、SA の局所探索能力と GA の持つ大域的探索能力を有する。GA のオペレータを用いた SA であるため、本研究では SA の探索点の総数 (SA の並列数) を個体数と呼んでいる。

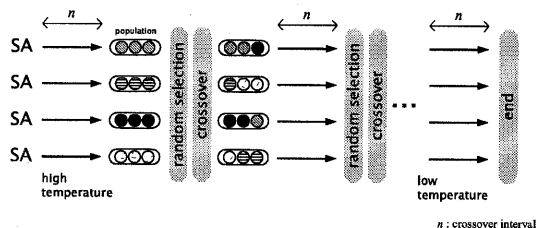


図 1: Simulated annealing using genetic crossover [1]

2 並列モデルの提案

PSA/GAc を PC クラスタのような分散メモリ型並列計算機に実装するには、新たな実装モデルを必要とする。本研究では、いくつかのモデルを考案した。そのうち 2 種について以下に説明する。

Parallel Simulated Annealing using Genetic Crossover

[†] Tomoyuki HIROYASU(tomo@is.doshisha.ac.jp)

^{††} Maki OGURA(ogu@mikilab.doshisha.ac.jp)

Department of Knowledge Engineering and Computer Science, Doshisha University (†)

Graduated School of Knowledge Engineering and Computer Science, Doshisha University (††)

1-3 Miyakodani, Tatara, Kyotanabe, Kyoto 610-0321, Japan

2.1 モデル 1

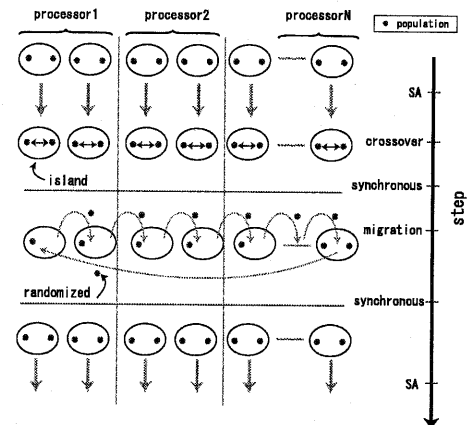


図 2: model1

図 2 において、小さな点は個体を示し、大きな円は個体の集合 (島) を示す。このモデルでは、島ごとに 2 つの個体が存在する。1 プロセッサ当たりには、全島数をプロセッサ数で分割した数だけの島が存在する。

ある世代までは、全ての個体が並列に SA を行う。そして解の伝達前に、並列に実行している島の中に存在する 2 個体を親として、設計変数間交叉を行う。交叉終了後、全ての島で同期を取り、各島の 2 個体のうち 1 個体をランダムに選択し、隣の島に移住させる。各島において、送信した個体数と受信する個体数は共に 1 個体であることから、常に島に存在する個体数は 2 個体で一定である。全ての島が個体情報の受信を完了した時点で、全ての個体が SA 処理を再開する。

2.2 モデル 2

交叉処理のみを行うプロセッサを 1 個設け、他の SA 処理を行うプロセッサと区別する。図 3 において、小さな点は個体を示し、大きな円は個体の集合 (島) を示す。このモデルでは、SA を行う 1 プロセッサ当たりには 1 個の島が存在する。そして各島には、全個体数を SA 処理を行うプロセッサ数で分割した数だけの個体が存在する。ある世代までは、すべての個体が並列に SA を行う。そして解交換周期に達したプロセッサから順に、その島に存在する個体の中から 1 個体選択して、交叉処理を行うプロセッサに送信する。受信するプロセッサ側では、受け取った順に 2 個体を親として設計変数間交叉を行い、得られた 2 個体を 1 個ずつ

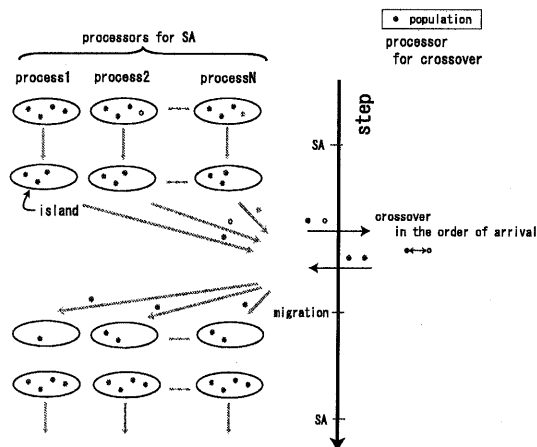


図 3: model2

送信されてきた SA 処理を行うプロセッサに送信する。交叉後の 1 個体を受け取ったプロセッサから SA の処理を再開する。

3. 各モデルによるタンパク質構造解析

タンパク質の構造解析は、最適化問題の 1 つである。タンパク質の立体構造はエネルギーの最小状態に対応しており、そのためエネルギーを最小とするような構造を最適化手法を用いて求めることが可能である。PSA/GAc をこのタンパク質構造解析問題に適用し、各モデルの解探索能力 (Success rate) 及び計測時間の比較検討を行った。用いた SA のパラメータを、表 1 に示す。各個体数 (Population size), MCsweep 数 (MCsweeps) について 20 試行し、その解取得確率を解探索能力と定義した。4 プロセッサと 6 プロセッサを用いて試行を行った。モデル 2 では、SA の処理を行うプロセッサを 4 プロセッサと 6 プロセッサとした。

表 1: Parameter of PSA/GAc

Parameter	Value
Population size	16, 24
MCsweeps	4000, 5000
Initial temperature	2.0
Cooling rate	0.999
Crossover interval	32
Migration interval	32

個体数が 24 個で、4000MCsweeps の場合の結果を、以下に示す。図 4, 図 5 はそれぞれ解探索能力、計測時間を示している。

解探索能力は、モデル 1, モデル 2 共にプロセッサ数が増えると高くなっている。また 4 プロセッサ, 6 プロセッサ共にモデル 1 に比べてモデル 2 が高い解探

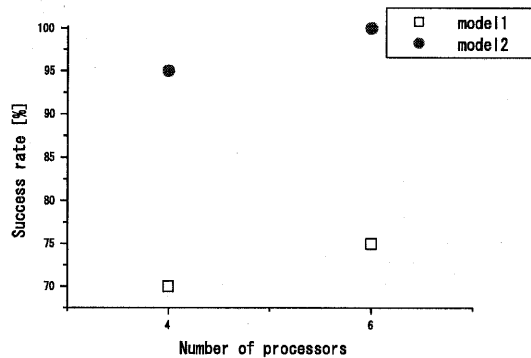


図 4: Success rate (Population size 24, 4000MC-sweeps)

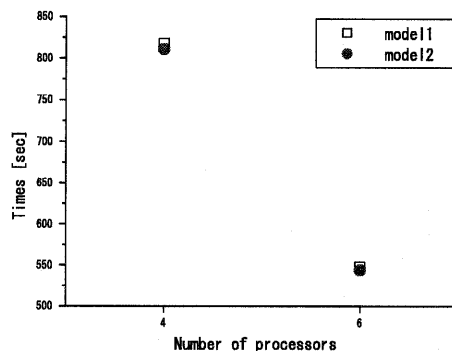


図 5: Times (Population size 24, 4000MCsweeps)

索能力を示した。計測時間においては、大きな違いは見られなかったが、両モデル共に高い並列化効率を示した。また、個体数と MCsweep 数を変化させた場合にも同様の結果が得られた。

4. 結 論

本論文では、分散メモリ型の並列計算機に実装するモデルをいくつか考案し、解探索能力や計測時間の面から比較を行った。タンパク質構造解析問題に適応したところ、モデル 2 は解探索能力が高く、必要とする計測時間も短く、最もよいふるまいをした。常に 100% 近い確率で最適構造を得ることができたため、モデル 2 は特に探索能力の面で優れていることが明らかとなった。

5. 謝 辞

本研究は文部省からの補助を受けた同志社大学の学術フロンティア研究プロジェクトにおける研究の一環として行った。

参考文献

- [1] Tomoyuki HIROYASU, Mitsunori MIKI, and Maki OGURA. Parallel simulated annealing using genetic crossover. *Proceedings of the IASTED International Conference PARALLEL AND DISTRIBUTED COMPUTING AND SYSTEMS*, Vol. 1, pp. 139-144, 11 2000.