

DGA によるタンパク質の構造解析および PDGA の作成
岩橋 崇史

1 目標課題

作成した DGA により, タンパク質の構造解析を行った. また, MPI を用いた PDGA (Parallel Distributed Genetic Algorithm) を作成した.

2 研究の進捗状況

2.1 自作 DGA によるタンパク質の構造解析

構造解析を行うタンパク質は Met-enkephalin を用いた. 存在する 19 個の二面角が設計変数にあたる. 二面角の値から求まるエネルギー値が -11kcal/mol 以下になれば最適解とみなす.

Table 1 の設定で, Met-enkephalin の構造解析を 300 回試行した.

Table 1 設定

島	4	設計変数	19
個体数/島	100	ビット数/設計変数	9
移住個体数/島	50	交叉率	1.0
移住間隔	5	突然変異率	0.007
エリート保存/島	1	終了世代	1000

Table 2 に 300 回試行における実行結果を示す. Table 2 に示す平均値, 中央値および最良値は終了世代 (1000 世代) の値である. 300 回試行のうち最適解を 7 回得ることができた. 今回は Table 1 の設定でのみ Met-enkephalin の構造解析を行った. 今後, パラメータを変えて実行し, パラメータの変更が最適解発見率に影響があるか調べる必要がある.

Table 2 実行結果

平均値	-6.78533
中央値	-6.7208
最良値	-12.2109
最適解発見率	7/300

2.2 PDGA プログラミング

2.2.1 PDGA の作成

作成した PDGA は一つのプロセスが一つの島を担当する. そして, 一定世代ごとに各島内で選ばれた一定個体を別の島と交換する. この移住操作は, MPI を用いたプロセス間通信を用いる.

2.2.2 PDGA と ga2k の性能比較

作成した PDGA と ga2k の性能比較を Table 3 の設定で行った. 試行回数は 300 回である.

Table 3 設定

島	4	設計変数	10
個体数/島	100	ビット数/設計変数	10
移住個体数/島	50	交叉率	1.0
移住間隔	5	突然変異率	0.01
エリート保存/島	1	終了世代	1000

複数のテスト関数を対象問題としたが, その中で Rastrigin 関数と Griewank 関数を解いたときの 300 回試行中央値の解探索履歴を Fig. 1, Fig. 2 に示す.

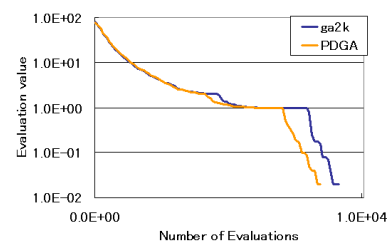


Fig. 1 Rastrigin 関数

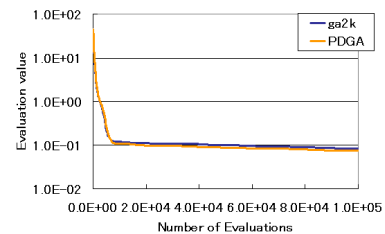


Fig. 2 Griewank 関数

Fig. 1, Fig. 2 において, PDGA と ga2k の解探索履歴に大きな差はないことがわかった. 他の関数においても同様の結果を得たため, PDGA が正常に動作していることが確認できた.

3 今後の課題

PDGA を用いて, タンパク質の構造解析を行う. また, それと平行して DGA によるタンパク質の構造解析において, パラメータを変更したときの最適解発見率の影響を調べる.