

遺伝的アルゴリズムを用いた適応的温度調節を行う並列 SA のタンパク質への適用
吉田 武史

1 研究活動報告

今月は遺伝的アルゴリズムを用いた適応的温度調節を行う並列 SA(PSA/AT(GA)) をタンパク質へ適用し、有効性を検証した。本報告では今月行ったいくつかの実験と現在までに考察出来ている点を示す。

2 PSA/AT(GA) について

PSA/AT(GA) は並列 SA の温度スケジュールを GA を用いて決定する手法である。各 SA プロセスは独自の温度を個体として持ち、解探索過程で全個体に対して選択 突然変異 交叉を行う事で、重要温度領域に収束することを狙う手法である。

特徴としては選択操作において、解の動きを示す適合度を用いること、実数値である温度を表現するために $Temperature = expX$ に示す X をビット列で表現することがあげられる。

3 今月行った実験

PSA/AT(GA) では並列 SA に GA 操作を加えた手法であるため、一般的な SA パラメータに加え、GA オペレータが用いるパラメータを設定しなくてはならない。今月は以下に示すパラメータに関して検証を行った。

- GA オペレータの順序を変更して性能比較
- 交叉率
- 突然変異率
- 個体数 (SA プロセス数)

これらの実験結果より、PSA/AT(GA) では温度推移が重要温度領域より低い温度に収束することがわかった。また、総 MCsweep 数を同一にした独立型 PSA と比較すると PSA/AT(GA) の性能が劣ることもわかった。本発表ではこれらの実験のうち、個体数 (SA プロセス数) に着目した実験結果を示す。

4 個体数 (SA プロセス数) の検証

温度の個体数と得られるタンパク質の品質を検証するために数値実験を行った。実験では PSA/AT(GA) の個体数 (SA プロセス数) を 8, 16, 32, 64, 各 SA プロセスの MCsweep 数を 12000, 6000, 3000, 1500 に設定し、総 MCsweep 数を同一にして比較実験を行った。また比較に用いた独立型 PSA は個体数 16, 各プロセスの

MCsweep 数を 6000 としている。また、対象としたタンパク質は Met-enkephalin と (Ala)10 である。実験結果を Table 1 に示す。

Table 1 個体数と解探索能力の関係

Met-enkephalin					
	PSA	PSA/AT(GA)			
個体数	16	8	16	32	64
Mcsweep	6000	12000	6000	3000	1500
最良値	-12.01	-11.95	-11.78	-11.86	-11.70
平均値	-11.39	-10.89	-10.36	-10.83	-10.87
中央値	-11.52	-10.99	-10.33	-11.00	-10.89
最悪値	-10.37	-8.95	-8.57	-9.00	-10.02
成功率	76 %	50 %	27 %	50 %	47 %
(Ala)10					
	PSA	PSA/AT(GA)			
個体数	16	8	16	32	64
Mcsweep	6000	12000	6000	3000	1500
最良値	-9.76	-9.39	-9.46	-9.17	-9.00
平均値	-9.29	-7.08	-5.71	-5.22	-4.50
中央値	-9.49	-8.64	-6.44	-6.22	-4.51
最悪値	-4.81	2.77	2.01	0.39	0.61
成功率	2 %	0 %	0 %	0 %	0 %

この実験結果より、PSA/AT(GA) の解探索能力が独立型 PSA に比べ劣ることがわかる。また、個体数が少ない PSA/AT(GA) のエネルギー平均値が良好なことより、PSA/AT(GA) では十分な MCsweep を必要とすることがわかる。

一方、PSA/AT(GA) の温度履歴を検証した結果、どの個体数においても重要温度領域より低い温度に収束していることがわかった。これは現在用いている適合度の計算方法では、重要温度より低い温度の適合度が高くなり、選択操作によってそれらの温度が選択される確率が高くなるためと考えられる。

5 今後の課題

- MPS シンポジウムの準備
- PSA/AT(GA) のタンパク質への適用