

進化的計算理論のネットワーク的ダイナミクス解析に関する一考察

An Analysis on Evolutionary Computing Dynamics from Perspective of Complex Networks.

米田 貴雄[†]
Takao YONETA

吉井 伸一郎[†]
Shinichiro YOSHII

1. はじめに

遺伝的アルゴリズム (Genetic Algorithms, 以下 GA) は, 進化的計算理論の一手法として, 最適化, 適応, 学習などへ幅広く応用されてきた. しかし, 遺伝的アルゴリズムの問題解決プロセスについては, いわゆる確率的探索手法であるゆえ, 詳細メカニズムは明らかにされてはいない.

そこで, GA が多数の要素による相互作用を行っている系である点に着目し, 複雑ネットワークの視点からの解析を試みる. そして, ネットワークの特徴的な性質を利用し, 探索効率, 多様性維持性という点から GA の進化プロセスを解析する.

本稿では, 線形計画問題や多峰性のある既知の関数に対して極値を見つける問題を設定し, 異なる探索形態を持つ GA を用いてその解を探索する. 次に, 解の探索プロセスによる遺伝子型ネットワークを形成し, 複雑ネットワークの観点からネットワーク特徴量を観測する. そして, 提案する評価方法で評価し考察を行う.

2. 遺伝的アルゴリズムについて

本稿では以下の 3 つの世代交代モデルによる GA を扱う. また, 基礎的進化プロセスの性質を観測するため, 複製選択と生存選択のみに着目し, 交叉手法, 突然変異手法は固定とする.

2.1 世代交代モデル

世代交代モデルには, 基準となる SGA モデル, エリート保存選択とランキング選択の性質を持つ CHC モデル [1], より小さい個体群中においても多様性を維持できる MGG モデル [2] を用いる. これらのモデルの複製, 生存選択法を以下に示す.

Simple GA (SGA)

複製選択: ルーレット復元抽出.

生存選択: 無条件で親個体群と子個体群を入れ替える.

CHC

複製選択: ランダム非復元抽出.

生存選択: 親・子個体群の中から優良個体順に選択する.

Minimal Generation Gap (MGG)

複製選択: ランダム非復元抽出.

生存選択: 各家族のうち, 最良 1 個体およびルーレット選択により選ばれた 1 個体を選択.

3. 遺伝子型ネットワーク形成方法

ここでは, ノードを遺伝子型, リンクを遺伝子型の親子関係からなるネットワークとし, リンクは 2 つの親遺伝子型による交叉・突然変異により生成されるものとする. 遺伝子型の親子関係によるリンクの概念図を Fig.1 に示す.

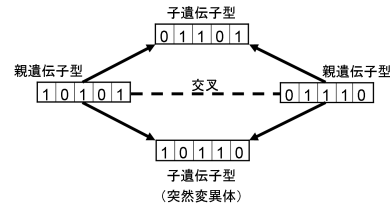


Fig. 1 遺伝子型の親子関係によるリンク

4. ネットワーク特徴量

ネットワークに関する最も単純な統計的特徴量として, 全体のノード数 N , リンク数 E がある. さらに, ネットワークの広がりや結合度などの構造的特徴を示す指標として, 平均パス長 L_i と平均クラustering係数 C_i がある [3].

4.1 平均パス長 (Average Path Length)

任意のノード i の平均パス長 L_i は, ノード i から全てのノードへの最短パス長 d_{ij} の平均である. ネットワーク全体の平均パス長 $\bar{L}$ は, その平均であり, 以下の式 (1) で表すことができる.

$$L_i = \frac{1}{N-1} \sum_{j \neq i}^N d_{ij} \quad \bar{L} = \frac{1}{N} \sum_i^N L_i \quad (1)$$

4.2 平均クラustering係数 (Clustering Coefficient)

任意のノード i のクラustering係数 C_i , および, ネットワーク全体の平均クラustering係数 $\bar{C}$ は以下の式 (2) で表すことができる. ただし, E_i はノード i の隣接頂点群に存在するリンク数である.

$$C_i = \frac{2E_i}{k_i(k_i - 1)} \quad \bar{C} = \frac{1}{N} \sum_i^N C_i \quad (2)$$

5. 評価方法

5.1 従来の評価方法と問題点

従来の評価方法として, 適応度を用いた最適解への収束性, 実行中のパフォーマンスを評価する方法や, 適応度分布の平均値や分散の時間変化から評価する方法などがある [2],[4]. このように, 従来の方法では適応度と適応度の変化に基づく評価方法が主である.

しかし, 本来, GA による進化は, 数世代にわたる多数の個体が相互作用を積み重ねることにより成されていると考えられる. つまり, 進化の点から見ると, 従来手法では適応度というミクロなパラメータのみでは GA の進化特性を明確に表すことはできない. このような見地から, 個体間の相互作用に着目した GA の進化特性に対する評価方法を提案する.

[†] 北海道大学大学院情報科学研究科

Table. 1 テスト関数

関数	上下制限約
$f_1 = \sum_{i=1}^{dim} x_i$	$-5.12 \leq x \leq 5.12$
$f_2 = \sum_{i=1}^{dim} x_i^2$	$-5.12 \leq x_i \leq 5.12$
$f_3 = 10n + \sum_{i=1}^{dim} [x_i^2 - 10\cos(2\pi x_i)]$	$-5.12 \leq x_i \leq 5.12$

Table. 2 シミュレーション結果

関数	モデル	N	E	$\bar{C}$	$\bar{L}$	最大適応度
f_1	SGA	604	1304	0.221	9.654	27.152
	CHC	49	69	0.173	3.182	30.72
	MGG	416	593	0.123	7.637	30.04
f_2	SGA	611	1340	0.206	8.865	401.7652
	CHC	54	78	0.185	3.664	402.6507
	MGG	409	575	0.133	4.109	402.4139
f_3	SGA	651	1383	0.199	9.796	133.3313
	CHC	97	203	0.414	3.420	132.8231
	MGG	414	543	0.118	5.740	135.4395

5.2 提案する評価方法

本稿では、個体間の相互作用による遺伝子型ネットワークの特徴量を用い、GAの探索効率、種の多様性という点からGAの進化プロセスを評価する。評価指標は $\bar{C}$, $\bar{L}$ であり、それぞれ探索の効率、種の多様性を示す。本稿のネットワーク形成手法では、 $\bar{C}$ は発見した遺伝子型が重複している割合を示すため、 $\bar{C}$ が低いほど、効率の良い探索が行われていることを意味する。また、 $\bar{L}$ は種の多様性が維持されている割合を示すため、 $\bar{L}$ が高いほど種の多様性が維持されていることを意味する。

6. 実験

対象問題は上下制限約のみの関数最適化問題とし、線形/非線形、単峰/多峰性の特徴を有するテスト関数 (Table.1) を用いる。

ここで、シミュレーションに用いるパラメータは、個体群数 10、染色体長 24、テスト関数次元数 $dim = 3$ 、交叉確率 $p_c = 0.75$ 、突然変異確率 $p_m = 0.01$ 、終了世代数 10、試行回数 10 とする。なお、交叉手法は一点交叉とする。

シミュレーション結果を Table.2 に示す。また、関数を f_3 とした場合の世代交代による平均パス長とクラスタリング係数の変化を Fig.3 に示す。Fig.2 はシミュレーションで得られた遺伝子型ネットワークの一部である。ノードの大きさは遺伝子型の適応度を表し、色は遺伝子型の発見世代数を表している。(高い適応度を持つ遺伝子型は大きく、新しい遺伝子型は黒く表示)。

7. 考察

7.1 遺伝子型ネットワークの対象非依存性

実験では特徴の異なる 3 つのテスト関数を用いたが、Table.2 に示すように f_3 の CHC モデル以外についてはネットワーク特徴量に大きな変化は見られなかった。つまり、対象問題によって遺伝子型ネットワークの構造は大きく異ならないということであり、GA は対象問題に依存せず、同様のプロセスによって探索を行っているということがわかる。

7.2 遺伝子型ネットワークの特徴

世代交代モデル別に形成された遺伝子型ネットワークの特徴量を見ると単純にノード数、リンク数の多さから SGA, MGG, CHC の順に幅広い探索が行われた事がわかる。しかし、SGA は発見された最大適応度が低いという点から、効率の良い探索が行われていないと言える。一方、SGA のネットワー

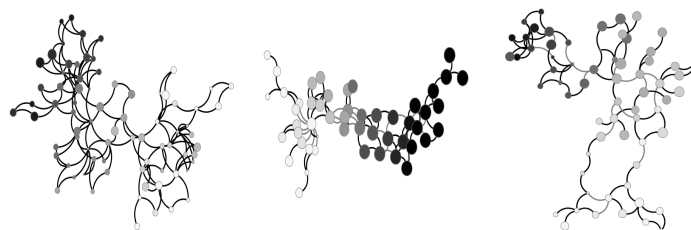


Fig. 2 遺伝子型ネットワーク (左から SGA, CHC, MGG)

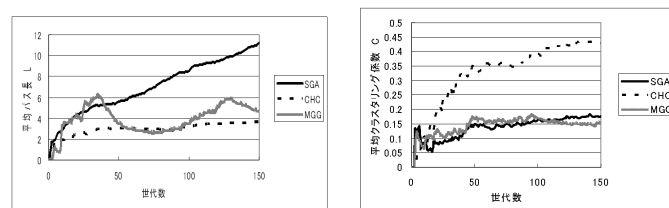


Fig. 3 平均パス長と平均クラスタリング係数

ク特徴量は $\bar{L}$ だけでなく、 $\bar{C}$ も高い値を示している。つまり、多様性を維持した探索であるが効率は良くないことを示しており、本稿の提案する評価方法の評価と一致している。

7.3 探索プロセスの変化に伴うネットワーク特徴量の変化

ネットワーク特徴量の世代別の変化を見ると、平均クラスタリング係数 $\bar{C}$ は収束する傾向にあるのに対し、平均パス長 $\bar{L}$ は収束するという以外の特徴が現れた。特に MGG では 2 つのピークが出現しており、このときの $\bar{C}$ を見ると、一つ目のピーク以降、値が上昇しており探索効率が悪くなっていることがわかる。しかし、その後、 $\bar{C}$ が減少、 $\bar{L}$ が上昇している点から、多様性を維持した効率の良い探索が行われ始め、結果としては最適解は 2 つ目のピーク付近で発見されていた。

以上の議論から、平均パス長 $\bar{L}$ 、平均クラスタリング係数 $\bar{C}$ を観測することにより、進化プロセスを特徴づけることができた。

8. おわりに

本稿では、GA の進化プロセスの解明には、個体間の相互作用の考慮が不可欠であるとの立場から、異なる形態を持つ GA による遺伝子型ネットワークの形成および、その特徴量解析を行った。そして、従来の適応度依存の評価方法ではなく、適応度非依存の評価方法を提案し、探索効率、多様性の維持率の点から、進化プロセスを評価した。

今後の発展として、適応度分布などを組み合わせた、より実用的な評価法を考える予定である。

参考文献

- [1] L.J. Eshelman : *Foundations of Genetic Algorithms: The CHC adaptive search algorithm*, Morgan Kaufmann Publishers, pp. 265-283, 1991.
- [2] 佐藤浩, 小野功, 小林重信 : "遺伝的アルゴリズムにおける世代交代モデルの提案と評価", 人工知能学会誌, Vol.12, No.5, pp.734-743, 1996.
- [3] D. Watts, S. Strogatz : Collective dynamics of small-world networks, *Nature*, Vol. 393, pp.440-442, 1998.
- [4] K.A. De Jong : *An Analysis of the Behavior of a Class of Genetic Adaptive Systems*, Doctoral dissertation, University of Michigan, 1975