

進化計算シンポジウム 2014 に参加して

鳥山直樹
Naoki TORIYAMA
電子情報学科 4年

1. はじめに

2014年12月20日から21日に広島県で開催された進化計算学会2014において「遺伝的プログラミングを用いた木構造間特徴量を考慮した並列モデルの検討」の発表を、ショートプレゼンとポスターセッションの形式で発表した。

2. 発表内容

遺伝的プログラミング (GP) とは、1992年 John Koza によって提案された進化計算手法であり、生物の進化をモデルとしたものである。GP において、遺伝子の多様性を広げることはより良い解を得るための重要な要素のひとつである。対象問題に適する遺伝的オペレータである交叉、突然変異および選択が与えられたとき、更に GP の遺伝的多様性を増加させる方法として、鳥モデルの適用は最も基本的な方法であるといえる。鳥モデル GP の性能を向上には効果的な個体の移住法が必要であり、本研究では移住個体の選択について考えた。GP は解を木構造で表すことができるため、移住個体の選択に関し、評価値以外の木構造特有の特徴量を考えることができる。しかしながら、この特徴量を求めるには多くの計算量が必要であるため、これまで鳥モデル GP に適用されてこなかった。しかし、近年 MPI や GPGPU などを用いることによって、高速に並列計算が行えるため、木構造の特徴量を用いた研究には将来性がある。本研究では木の編集距離 (リーベンシュタイン距離) に着目した鳥モデル GP の移住法を提案した。

一方、個体の多様性を促進と木構造のノード数の急激な増加 (プロット) には関係があり、ノード数

の増加抑制を考慮しない個体の多様性の促進はプロットを引き起こす。そこで、我々は個体間の類似度に着目し、各島の評価値が最も高い最良個体との編集距離に

基づき島の個体を選択を行うことで、探索と探査のバランスのとれた移住法を考えた。

鳥モデルにおいて、母集団は島と呼ばれる M 個のサブ母集団に分割され、島は並列に解の探索を行う。 $L^m (m = 1, \dots, M)$ は m 番目の島を表す。 L^m には N 個の個体 $\{p_n^m = 1, \dots, N\}$ があり、その個体集合を P^m とする。母集団全体の個体集合 P を

$$P = \bigcup_{m=1}^M P^m$$

とし、各島はラベル付き順序木で記述される個体で構成される。また GP は解が木構造であるため、評価値以外の特徴量を持つ。本研究では木の編集距離に注目し、それを用いた3つの手法を提案した。1つ目の Proposed 手法は、島間の移住個体を類似個体のみにするによって、遺伝子の多様性を制御しプロット抑制を試みる手法である。移住個体の選択の手順を以下に示す。

P0: L^m に L^m 以外の個体をアーカイブとして保存する。

つまり $P = P^1 \cup P^{m-1} \dots \cup P^{m+1} \dots \cup P^M$ である。

P1: L^m の中で、評価値が最も高い個体を基準個体 p_B^m とする。

P2: P の中で p_B^m と類似している個体を集め、個体集合 p_s^m を作成する。

P3: L^m の個体集合の P^m を基準個体と p_s^m に置き換える。

P4: L^{m+1} 以降も同様の手順を行う。

Proposed 2 手法は、1つ目の手法を拡張し、移住個体の重複を除くことによって多様性を向上をさせる手法である。Proposed 3 手法は、2つ目の手法に加え、基準個体の選択を改良することによって多様性の向上とノードサイズの抑制を図る手法である。

3. 結果

Fig. 1 は提案法の評価値について、Fig. 2 は提案法のノードの大きさについて比較したものである。提案法は2, 3と続くにつれて遺伝子の多様性を促進するようにしている。一般的な手法において、多様性を向上させることは、ノードの大きさの増加につながる。

しかしながら提案法では、個体間の類似度を用いることで、評価値を向上させることができ、さらに、ノードの大きさも抑えることができた。よって、レーベンシュタイン距離を特徴量として用いる

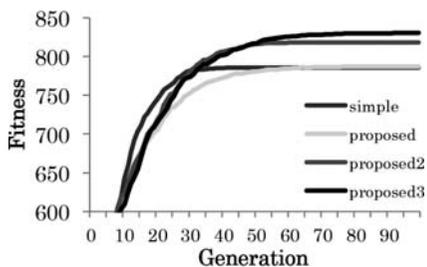


Fig. 1 提案法の評価値について

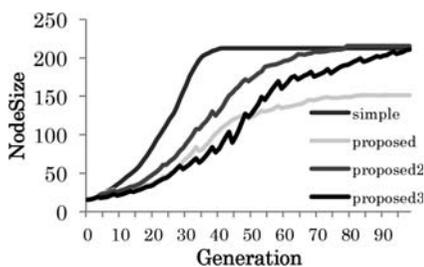


Fig. 2 提案法のノードの大きさについて

ことは、評価値の向上とブロート抑制のバランスを取ることに於いて有効であることを示した。

4. 学会を終えて

初めての学会ということでもとても緊張していたが、落ち着いて自分の研究を説明することを心がけた。また、様々な大学や企業から参加者が集い、同じくポスター形式で発表をするため、質問や議論を交えることで、とても貴重な経験や知識を得ることができた。今回得た経験をこれからの研究生活だけでなく、様々なことに活かしていきたいと思う。

最後に今回の発表を行うにあたって、ご指導いただいた小野景子講師、ならびに研究室の皆様に深く感謝致します。

参考文献

- 1) D. Andre and J. R. Koza. A parallel implementation of genetic programming that achieves super-linear performance. In Proc. of PDPTA'96, Volume III, pp.1163-1174, 1996.
- 2) V. F. Fernández, G. G. Galeano, and P. J. Antonio Gómez. Comparing synchronous and asynchronous parallel and distributed genetic programming models. In Proc. of Euro GP'02, pp.326-336, 2002.
- 3) W. B. Langdon and R. Poli. Foundations of genetic programming. Springer, 2002.
- 4) W. F. Punch. How effective are multiple populations in genetic programming. In Proc. of GP'98, pp.308-313, 22-25 1998.
- 5) K. Yanai and H. Iba. Estimation of distribution programming based on bayesian network. In Proc. of CEC'03, Vol.3, pp.1618-1625, dec.2003.