

異文化型島モデルを用いた遺伝的プログラミング

1 はじめに

遺伝的プログラミング (Genetic Programming: GP)[1] は遺伝的アルゴリズム (Genetic Algorithm: GA) の拡張で、遺伝子型を配列から木構造にかえることによって、構造的表現を可能にしたものである。この手法は現在、ハードウェアの設計、カオス系列の推定などで盛んに応用されている。GP による探索は効率のよいことで知られているが、この手法は GA 同様パラメータの設定が難しく、初期収束による探索の停滞などの問題点も報告されている。

一方 GP の基になっている GA の並列分散処理の手法の一つに島モデル (Island Model) がある。このモデルは並列計算を関する手法であるが、単一母集団で行なう GA と比較して精度の良い解が得られるという報告がなされており、逐次処理で行なう場合に対しても有効な手法である。

GP における島モデルの適用に関する研究もいくつか報告されている。しかしながらこれらの報告は島の数に着目してはいるが島全体の個体の差を考慮したものではない。

本論文では GP 独自の島モデルとして母集団を意図的に二つとし、次世代を残すパラメータに変化をつけることにより、人為的に島の環境を変えることで探索効率を改善する異文化型島モデルを提案し、いくつかの問題に対しこれを適用してその有効性を検証する。

2 島モデル

島モデルとは GP の基である GA における荒粒度の並列分散処理手法である。この手法は集団を複数の島と呼ばれるサブ集団に分割し、各島ごとに独立に探索を進めていく。さらにある一定の世代毎でいくつかの個体を他の島へコピーする移住 (Migration) という操作を行なう。

島モデルが単一集団の GA より有効である理由として移住の操作が大きいと考えられる。各島では単一集団より個体数が少ないので通常の方法より多様性が失われやすいが、移住の操作により各島での多様性が失われることになっても移住してきた個体により再び多様性を回復することが期待できる。このためより大域的な探索ができると考えられるからである。

2.1 島モデル GP

GP における島モデルの適用は GA と異なる特性を持つことが知られている。過去の報告では GA では島内の個体数は 10 程度あたりが望ましく、移住率も 10% 以上で解の改善が最も見られている [3]。これに対し GP での各島の個体数は対象の問題の性質や困難さに依存しており、探索空間を広げると単一集団の GP より劣ることが起きると報告があり、移住率に関してもある問題では移住率を 12% にしたとき単一母集団より探索効率が劣っていたという例も報告されている [4]。すなわち GP における島モデルの適用は GA に比べると有効性が低いといえる。

このような現象が起きる原因として木の爆発的な成長であるブロー現象が考えられる。この現象は個体の集団数が少ないほど起こりやすく移住による多様性維持の効率よりブロー現象による探索の停滞の影響の方が大きい島数を増やすことが GA より効果的ではないと考えられる。また移住率を増やすことにより起こる精度の低下はブロー現象により肥大化した遺伝木を持つ個体が多数移住してくるため各島でのブロー現象が伝播され全体的な多様性も失われやす

くなるためと考えられる。

3 異文化型島モデルの提案

島モデルは GA に対しては有効な拡張手段であるが GP においては必ずしも良い結果のみが返ってくるとは限らず、パラメータの調整次第では単一母集団に劣るという結果も現れている。

本論文では GA では無く GP での適用に適する島モデルの形成を提案する。前述したように GP での島モデルによる探索の性質の違いは

1. 各島の個体数が少なくなるためブロー現象が起きやすくなる。
2. 移住率を上げた場合、ブロー現象が伝播する確率が高くなる。

ということが主原因であると思われる、すなわち GP の探索においては、移住による多様性維持の効果がブロー現象によって阻まれるため不十分であると考えられる。

多様性を維持するための手法は突然変異を始め幾つか提案されているがいずれも探索の効率を下げるというデメリットがある。そこで多様性の維持を移住のみに頼らずに、多様性維持のためにパラメータを調整した島を用意して探索効率と多様性を同時に確保する異文化型島モデルを提案する。すなわち島を探索と多様性維持の役割分担させることで全体として成績と多様性を同時に上げることが目的とする。ここで探索効率に重点を置く島を「成績重視の島」、多様性維持に重点を置く島を「個性尊重の島」と呼ぶことにする。

本実験では 2 島のモデルを提案する。

3.1 成績重視の島

成績重視の島では多様性を考えずに解の精度を上げることだけを考える。まず突然変異の操作は行わず、交叉のみで次世代を生成する。

3.1.1 交叉方法

GP における交叉の一つに深さ依存型交叉という手法がある。この手法は GP はスキーマ理論に基づく遺伝木の根に近い部分の選択確率を高くすることにより有用なスキーマを破壊される確率が低くなるものとする。岩崎ら [2] はこの手法に着目し島モデルにこの交叉と移民世代によるパラメータ変化を適用した結果従来の島モデルより探索効率に改善が見られたことを報告している。

本実験では成績重視の島ではこの交叉を用いることにする。

この交叉は以下の手法で行なうものとする。

1. 交叉の対象となる個体の遺伝木の深さを d とする。
2. $0 \sim d$ の数値をランダム選ぶ
3. 選んだ数値の深さの要素をランダムで一つ選んで交叉点とする。

さらにエリート保存として現在の世代の個体と生成された個体を成績順に整列させその上位の個体を世代にする。

3.2 個性尊重の島

個性尊重の島では解の精度を考えずに多様性の維持に重点を置くことにする。具体的には交叉以外にも突

然変異による個体生成も適用し、エリート保存は行わないことにする。

また次世代の生成ごとにランダムで生成した個体を作成し現在の個体と一部交換することとする。

3.3 移住とその他のパラメータ

移住に関しては毎世代、個性尊重の島の上位個体と成績重視の島の下位個体を交換する形で行うことにする。これは個性尊重の島で成績重視の島より成績の良い個体が生成された時その個体を保存するためである。また成績重視の下位個体を個性尊重の島にコピーすることにより個性尊重の島と成績重視の島での解の差異が大きくなるようにする効果が期待できる。

交叉および突然変異の適用時における個体の選択方法はトーナメント選択を用いる。これはトーナメントサイズと呼ばれる定数だけ集団からランダムに個体を選択しその中で最も成績の良い個体を選択する方法である。この選択方法の利点として解の値の差に対する影響が少なく適合度にペナルティを課したときに有効に働くことが期待できる。

本報告ではトーナメントサイズは成績重視の島では高く設定し、個性尊重の島では低く設定することにする。

4 検証

単一母集団の GP との解と比較することでその有効性を検証する。

なお、本実験では同じ適合度なら要素数の少ない個体を優先するために適合度には全て 1.0×10^{-6} のペナルティを課すことにする。また個性尊重の島におけるランダムな個体を挿入する割合は全て島の個体数の 20% とする。

比較する単一母集団の GP は個体数を島モデルの総個体数と同じとしそれ以外のパラメータは成績重視の島と同じものと、同じく総個体数が同じ通常の GP を使用する。

本論文では比較に使用した問題として人工蟻の探索問題、11 マルチプレクサ問題、からみ螺旋問題、巡回セールスマン問題での適用を行なった。以下に巡回セールスマン問題の説明と比較結果を詳述する。

4.1 巡回セールスマン問題

NP 困難問題として有名な巡回セールスマン問題 (TSP) への適用 [5] を行なう。

本実験ではベンチマーク問題集から 76 都市問題である Pr76 問題と 100 都市問題である kroA100 問題を扱う。

非終端記号は { PROG(ノードの数は3~5) : ノードを逐次実行, TURN2(node1 node2) : 都市数が 1/2 以下なら node1 それ以外は node2 を実行, TURN3(node1 node2 node3) : 都市数が 1/3 以下なら node1 を 2/3 以下なら node2 をそれ以外は node3 を実行する, MOD2 (node1 node2) : 都市数が 1 偶数なら node1 それ以外は node2 を実行, MOD3(node1 node2 node3) : 都市数を 3 で割った余りが 1 なら node1 を 2 なら node2 をそれ以外は node3 を実行する, PRIME(node1 node2) : 都市数が素数なら node1 それ以外は node2 を実行する }。

終端記号は { NNM : 最近傍法による都市の追加, NAM : 最近追加法による都市の追加, NIM : 最近挿入法による都市の追加, CIM : 最安挿入法による都市の追加, FIM : 最遠挿入法による都市の追加 } とする。適合度は生成された巡回路の長さとする。

5 実験結果

移民率を 5% と 20% に設定し、各問題を 50 回試行した図 1 に Pr76 問題の結果を示す。

巡回セールスマン問題の結果を考察すると Pr76 問題に関してはまず成績重視のパラメータのみの GP は通常の GP より非常に速く収束するが収束後の解の

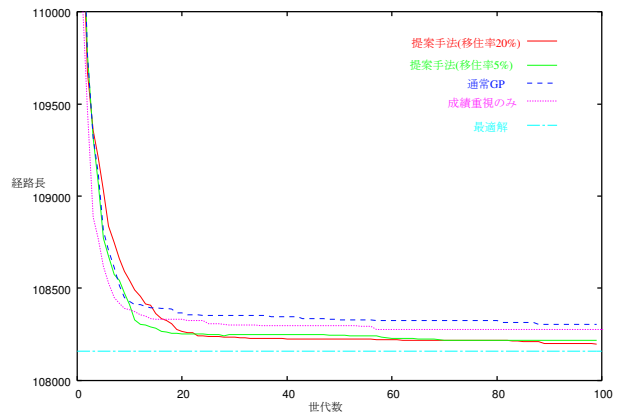


図 1: 巡回セールスマン問題の実行結果 (Pr76)

改善が小さく 100 世代目では通常の GP とほとんど変わりが無くなっている。

一方、提案手法は移住率 5% 時には通常の GP とほぼ同じ軌跡を描きながら 20% 時には通常の GP よりゆっくりと解の改善が見られ 100 世代目では共に通常の GP より良好な解の改善が見られた。これは提案手法が通常の単一母集団の GP に比べて多様性の確保ができていたことを表している。

6 結論

GP における島モデルとして異文化型島モデルを提案し幾つかの問題に適用して提案手法の有効性を確認した。

その結果、巡回セールスマン問題では結果移住率 20% 時には適用した問題に対して解の改善が見られており提案手法の有効性が確認できた。

また 11 マルチプレクサ問題やからみ螺旋問題、人工蟻の探索問題でも通常の GP や成績重視の島と同じパラメータの GP に比べてその良好な解の改善が見られた。

移住率 20% というのは従来の島モデル GP においては解の改善が見られなくなる数値であり提案手法は従来の GP に比べて移住率を高くすることによって解の改善が可能な手法であると考えられる。

今後の課題として 2 島のモデルだけではなく多島のモデルによる提案などの研究が考えられる。

参考文献

- [1] J.R.Koza, Genetic Programming, MIT Press, 1992.
- [2] 岩崎 誠, 伊庭 斉志, “移民世代数と深さ依存型交叉を用いた並列分散 GP”, 情処学論, vol.43 no.SIG10(TOM7), pp.146-156, 2002.
- [3] 廣安知之, 三木光範, 上浦二郎, “分散遺伝的アルゴリズムにおけるパラメータの検討”, 第 8 回 MPS シンポジウム講演論文集, pp.147-154, 2001.
- [4] D.Andre, J.R.Koza, “A Parallel Implementation of Genetic Programming That Achieves Super-Liner Performance”, Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications, vol.III, pp.1163-1174, 1996.
- [5] 原 裕一, 山内 仁, 金川明弘, 高橋浩光, “遺伝的プログラミングによる巡回セールスマン問題の解法について”, FIT2005 一般論文講演集 第 1 分冊, pp.47-50, 2005.