

確率モデル遺伝的アルゴリズム EHBSA における戦略、 パラメータの調査

Research on strategy and parameter about probabilistic model-building genetic algorithm EHBSA

1G01P046-7 酒井 大輔
Daisuke SAKAI

指導：上田 和紀 教授
UEDA Laboratory

概要

本研究では確率モデル遺伝的アルゴリズムの一種である EHBSA(edge histogram based sampling algorithm) における TSP 解法の戦略、パラメータを主に解の精度の観点から調べた。言語としては OCaml を使い、EHBSAWO、EHBSAWT とその並列版を実装した。実験の結果、集団の選択や並列化の際の戦略はエリート戦略が有効であることなどが分かった。また、並列化の結果逐次版のものよりも解の精度が 1.3 倍から 1.6 倍程度向上した。

キーワード: 確率モデル遺伝的アルゴリズム、巡回セールスマン問題、OCaml、並列化

Keywords: Probabilistic model-building genetic algorithm, TSP, OCaml, Parallelization

1 確率モデル遺伝的アルゴリズム

確率モデル遺伝的アルゴリズムとは、従来の GA(遺伝的アルゴリズム) のように交叉オペレータにより子個体を生成せず、集団の個体分布を確率モデルで近似し、そのモデルに基づいて子個体を生成する手法である。従来の GA に比べて高い探索能力を持つと期待されている。

2 EHBSA

EHBSA[1] は確率モデル遺伝的アルゴリズムの一種であり、順序問題に対する解を探索するのに用いる。確率モデル遺伝的アルゴリズムは集団の個体分布を確率モデルで近似するが、EHBSA ではこれを EHM(edge histogram matrix) という行列で行なう。つまり、世代 t の集団があったとき、これから EHM^t をつくりこれを用いて sampling することで、世代 $t+1$ の集団をつくる。この繰り返しにより、より良い解を探索する。

EHBSA のアルゴリズムの簡単な流れは次のようになる。

1. 個体をランダムに生成
2. 選択オペレータを用い、有望集団をつくる
3. 有望集団から EHM を作成
4. EHM を基に子個体をサンプリングする

5. 終了条件が満たされるまで、ステップ 2 から 4 をくり返す。

3 EHBSAWO, EHBSAWT

EHBSA には EHBSAWO(EHBSA without template) と EHBSAWT(EHBSA with template) という 2 種類の手法がある。EHBSAWO は第 2 節で述べたほぼそのままのアルゴリズムだが、EHBSAWT は集団の中から template としてひとつの要素を取りだしそれらをいくつかの cut によって分け、その中の 1 つの segment だけを sampling し他の segment はそのままコピーされる。これによって優秀なブロックが次世代にも遺伝すること、世代交代にかかる計算量が短縮されるなどの効果が望める。

4 実装した戦略

本研究では初期集団の選択、template の選択、sampling node の選択、EHM の戦略、局所的改善手法の有無、送受信の戦略に分けて戦略を実装した。また設定ファイルでこれらの組み合わせを指定し、実行できるようにした。

初期集団の選択においては、まったく選択しない no_population_select、優れた上位半分を使う (つまり同じ内容のものが 2 つずつ作られる) population_select_half、適応度の割合に応じて子孫を残す population_select_roulette、ランダムに 2 つ選び、そ

表 1: 並列版 EHBSAWO の実験結果

都市数	プロセス数	送受信数	並列版の解	逐次版の解	逐次版/並列版
200	7	5	4864.1	7657.1	1.57
400	8	10	8273.5	13525.5	1.63
600	6	10	15428.0	19317.1	1.25
800	8	10	17800.1	24701.0	1.38
1000	6	20	23639.5	30812.7	1.30

のうちの優れた方が受け継がれ、それを集団の数だけ繰り返す `population_select_tournament` の4種類の戦略を実装した。

`template` を選択するステップでは、ランダムに `template` を選ぶ `choose_template_randomly`、適応度に応じて選択する `choose_template_roulette`、最も良いものを `template` として選択する `choose_template_elite` の3種類を実装した。

`sampling node` を選択するステップでは、ランダムに `sampling node` を選ぶ `choose_node_randomly`、適応度に応じて選択する `choose_node_roulette` の2種類を実装した。

EHM を作成するステップでは EHBSA の論文に基づいたデルタ関数を使う `make_ehm_delta` と、独自の重みづけを行なう EHM を作成する `make_ehm_weight` の2種類を実装した。

また、2opt 法という従来の GA に使われているヒューリスティックスを使う `use_twoopt` と使わない `non_use_twoopt` の2種類を実装した。

並列版では送信用の戦略、受信用の戦略のそれぞれに対してランダムに送受信する `send_randomly`、`receive_randomly`、適応度に応じて送受信するものを選ぶ `send_roulette`、`receive_roulette`、上位の何個かを送る `send_elite`、自分の中で最も適応度の低いものと受信したものを取り換える `receive_elite`、ランダムに2つ選びそのうち適応度の高いものを送信する `send_tournament`、ランダムに1つ選び、それと受信したものを比較し、優れているときのみ交換する `receive_tournament` の計8種類を実装した。

5 実装したシステム

計測と分析を効率良く行なうためパラメータファイルを書き、その直積を計算しそれらすべてを実行するユーティリティと、同様のパラメータファイルを受け取り結果を自動的にグラフにするユーティリティを実装した。

6 実験結果

実験の結果、実効速度以外は EHBSAWO のほうが EHBSAWT よりも優れていた。また初期集団の選択法としてはエリート戦略、トーナメント戦略が適当であること、EHM の重みづけは有効であること、集団数は都市数の4分の1程度が最も解の精度が良くなること、エリート戦略を用いた並列化をすることにより、解の精度を向上させることが出来ること、などが分かった。

7 今後の課題

今後の課題としてはアルゴリズムの改良がある。EHBSA は世代交代にかかる時間は通常の GA よりも短い、解の精度自体は通常の GA に劣る。これは戦略にエリート戦略を用い EHM を重みづけするというかなり強引に探索領域をせばめる戦略のため、解が局所解に陥る可能性が高いせいである。また並列化したときに解の精度が大幅に向上することがあるのは、全く異なるかたちの集団の要素を取り入れることで探索空間が拡大することが理由として考えられる。このようなことを考えるとアルゴリズムの改善のためには、探索空間を大きく取れるような戦略、パラメータを増やす必要があるだろう。

参考文献

- [1] Shigeyoshi Tsutsui : Probabilistic Model-Building Genetic Algorithms in Permutation Representation Domain Using Edge Histogram , Proc. of the 7th Parallel Problem Solving from Nature - PPSN VII, pp.224-233 , 2002.