

Title	局所的多様性の損失を考慮した評価関数を用いたGAのTSPへの適用
Author(s)	永田, 裕一
Citation	人工知能学会論文誌, 21(2): 195-204
Issue Date	2006
Type	Journal Article
Text version	publisher
URL	http://hdl.handle.net/10119/7827
Rights	Copyright (C) 2006 人工知能学会. 永田 裕一, 人工知能学会論文誌, 21(2), 2006, 195-204.
Description	

局所的多様性の損失を考慮した評価関数を用いた GA の TSP への適用

New Approach of a Genetic Algorithm for TSP Using the Evaluation Function Considering Local Diversity Loss

永田 裕一
Nagata Yuichi

北陸先端科学技術大学院大学 情報科学研究科
Graduate School of Information Science, Japan Advanced Institute of Science and Technology
nagatay@jaist.ac.jp

keywords: genetic algorithm, TSP, EAX, local diversity loss, evaluation function

Summary

The edge assembly crossover (EAX) is considered the best available crossover for traveling salesman problems (TSPs). In this paper, a modified EAX algorithm is proposed. The key idea is to maintain population diversity by eliminating any exchanges of edges by the crossover that does not contribute to an improved evaluation value. For this, a new evaluation function is designed considering local diversity loss of the population. The proposed method is applied to several benchmark instances with up to 4461 cities. Experimental results show that the proposed method works better than other genetic algorithms using other improvements of the EAX. The proposed method can reach optimal solutions for most benchmark instances with up to 2392 cities with probabilities higher than 90%. For an instance called fnl4461, this method can reach an optimal solution with probability 60% when the population size is set to 300 – an extremely small population compared to that needed in previous studies.

1. はじめに

巡回セールスマン問題 (traveling salesman problem, 以下 TSP) は最も良く知られた NP 困難な組合せ最適化問題の一つである [久保 94]。 $G = (V, E, w)$ を N 個の頂点からなる重み付完全グラフとする。ここで V, E, w はそれぞれ頂点集合, 枝集合, 枝の重み集合とする。TSP における解は, グラフ G 上で最も短い巡回路長をもつハミルトン閉路と定義される。ここで, 巡回路長はハミルトン閉路を構成する枝の重み和で定義される。また, TSP における近似解はグラフ G 上のハミルトン閉路で定義され, 近似解の評価値は同様にハミルトン閉路を構成する枝の重み和で定義される。近似解を実行可能解, 解を最適解と呼ぶこともある。

これまで, 困難な組合せ最適化問題に対する近似解法として, simulated annealing, tabu search, genetic algorithm, あるいはそれらを拡張したさまざまな手法が提案されているが, TSP はそれらの手法のベンチマークとして最も多くの適用がなされている問題である [Reinelt 94, Johnson 97]。また, 新しいアルゴリズムのアイデアを試すテスト問題として利用されることも多い。それゆえ, TSP に対してこれまでに提案されている近似手法の洗練度も高く, 新しい手法で従来手法の性能を越えること

が困難な問題でもある。近似解法の性能は近似解の精度と近似解を得るための計算コストにより評価される。遺伝的アルゴリズム (genetic algorithm, 以下 GA) は, 他の手法と比較して, 集団で探索を行う点と交叉オペレーターにより複数の解候補を組み合わせて新しい探索点を生成する点が大きな特徴である。TSP は GA のベンチマークとしても, これまでもっとも多くの適用がなされてきた問題といえる。

ある特定の最適化問題へ GA を適用する場合, GA を設計する上で最も重要となるのが, 交叉, 突然変異, 世代交代モデルの設計である。特に交叉の設計は GA の性能を決定する上で最も重要な要素であり, TSP を対象とした GA においてはこれまでに多くの交叉が提案されている [山村 92, Grefenstette 85, 前川 95, Whitley 89, 永田 99]。これまでに提案された TSP の交叉においては, [永田 99] で提案された EAX がよい探索性能を示すといくつかの論文 [Tsai 02, 池田 02] で述べられており, [Watson 00, Nagata 03] では EAX が良い性能を実現する要因を解析している。さらに著者ら [永田 00] や池田ら [Ikeda 02] は, EAX を改良してさらに良い探索効率を実現するための拡張法を提案した。また, [Tsai 02] では EAX に neighbor-join (NJ) mutation と呼ばれる突然変異を組み合わせて, EAX を用いた GA の近似性能を改善した。

世代交代モデルは主に複製選択 (*selection for mating*) と生存選択 (*selection for survival*) から構成される。前者は集団からどの個体を親として選択して子個体を生成するかを、後者は次世代を構成する集団を決定するために、現集団と生成された子個体からどの個体を選択、あるいは排除するかを決定する機構である。GA を使った探索では、集団の多様性を適切に保ちつつ探索を重点化することが要求される。そこで、世代交代モデルの設計では探索領域の絞り込み (*exploitation*) と開拓 (*exploration*) のトレードオフを考慮する必要がある。一般的に、*exploitation* として目的関数の高い個体を選択圧をかけ、*exploration* として目的関数が集団内で劣る個体を選択圧をかけたり、探索空間内で類似した個体が増え過ぎないようにするといった工夫がなされる。

これまで多くの世代モデルが提案されているが、特に TSP を対象とした交叉を用いた GA に対して提案された世代モデルとしては、[永田 00] では生存選択において、生成された子個体を両親の近い方と置換える方法、[前川 97] では集団の多様性指標としてエントロピーを定義し、これを維持するような生存選択をする方法、[Tsai 03] では複製選択において集団中で比較的類似した個体のみを両親として選択する方法、などが提案されている。

本論文では TSP の近似解法として EAX を交叉とした GA を用い、世代交代モデルの生存選択において、*exploitation* と *exploration* のトレードオフを明示的に考慮した新しい手法を提案する。同様の手法として [前川 97] があるが、多様性の指標として集団全体から計算されるエントロピーを定義している。一般には集団の多様性は集団全体で評価されるべきであるが、どのように多様性を定義すべきであるかは明らかとはいえず、多様性の計算に要する計算コストが大きいことも問題である。本論文では生存選択として生成子個体が両親と置換えられるモデル (家族内淘汰) を用いる。そして、ある子個体を選択した場合の多様性の変化を、生成された子個体と生成した親の間 (家族) の局所的な関係のみで評価する手法を提案する。この場合、局所的な関係のみで多様性の変化を評価できるため、この計算に要する計算コストは問題にならないという利点がある。本論文では、提案する手法を TSP のベンチマークに適用し、実験により提案手法の有効性を確認する。さらに、提案手法が TSP の近似解法としてこれまでに提案されている GA よりも優れた近似性能を実現することを実証する。

本論文の以下の構成は、二章で本論文に關係する EAX に関する従来研究について概説し、三章では提案手法について述べる。そして、四章において実験結果について述べ考察を行う。五章は結語である。

2. 関連研究

本章では、関連研究について述べる。まず交叉 EAX について概説する。さらに、EAX の性能を改善するための拡張について述べる。

2.1 EAX

EAX の概要を以下と図 1 に述べるが、詳細は文献 [永田 99] を参照されたい。オリジナルの EAX では、以下のステップ 3 で *AB-cycle* を確率 1/2 でランダムに選択して *E-set* を構成する。この方法を EAX-Rand と呼ぶことにする。

[EAX algorithm]

1. 両親を tour-A, tour-B と表す。 G_{AB} を tour-A と tour-B を重ね合わせて構成される多重グラフとする。 G_{AB} 上で、両親で共通する枝は多重枝として扱う。
2. G_{AB} 上の枝を *AB-cycle* へと分割する*1。ここで、*AB-cycle* とは G_{AB} で tour-A と tour-B の枝を交互にたどって得られる閉路のことをいう*2 (図では 12 個の *AB-cycle* が示されている)。
3. 生成された全ての *AB-cycle* から幾つかを何らかの基準に従い選択する。ただし、一対の多重枝のみからなる *AB-cycle* は候補から外す。*E-set* を選択された *AB-cycle* に含まれる枝の集合と定義する (図では 3 種の *E-set* が示されている)。
4. *E-set* を tour-A に適用し、中間個体を生成する。この操作は、*E-set* に含まれる tour-A の枝を tour-A から取り除き、これに *E-set* に含まれる tour-B の枝を付け加えることをいう。
5. 部分巡回路からなる中間個体をつなぎ合わせ、完全な巡回路へと修正する。どの部分巡回路同士をつなぐか、どの枝を切り離して巡回路同士を接続するかはヒューリスティックを用いて行う。

2.2 EAX の拡張法

EAX-Rand はステップ 2 で生成された *AB-cycle* をランダムに選択して *E-set* を構成したが、その他に EAX を改良する方法としていくつかの方法が提案されている。

§ 1 EAX-1AB

著者らは *E-set* を一つの *AB-cycle* のみから構成する方法を提案した [永田 00]。それゆえ、中間個体として tour-A に近い (tour-A の枝を多く含む) 個体が生成される。この EAX を EAX-1AB と呼ぶことにする。4.2 節で示す実験において、この方法を用いた GA は EAX-Rand を用いた場合よりも優れた探索性能をもつことが示される。

*1 G_{AB} に含まれる枝は余すことなく *AB-cycle* へと分割されるが、一般にこの分割は一意ではない。

*2 図 1 で示されている *AB-cycle* の例はすべて単純閉路であるが、そうでない場合も存在する。

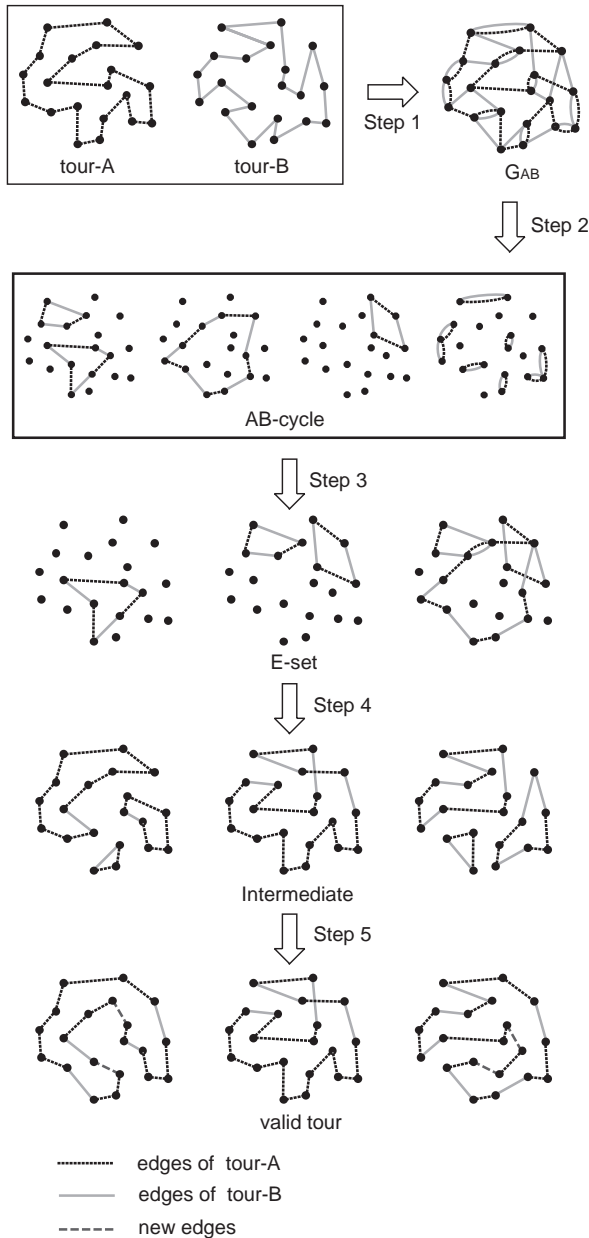


図 1 EAX アルゴリズムの概要

§ 2 dMSXF

池田 [Ikeda 02] らは EAX-1AB に、多段階交叉 [山田 97] という手法を変更した決定的多段階交叉という手法を組合せた。この方法は EAX-1AB を逐次的に tour-A に適用して解候補を改善していく方法で、以下のような手続きで行われる。

1. A, B を両親とする。 $A_1 = A, k = 1$ とする。
2. μ 個の子個体を A_k と B を両親として EAX-1AB を用いて生成する。
3. ステップ 2 で生成された子個体から最良個体 (巡回路の最も短い個体) を選択する。
4. A_{k+1} をステップ 3 で選択された最良個体とする。 $k = k + 1$ とし、 $k = k_{max}$ となるか、 A_{k+1} が B に等しくなれば終了。 そうでなければ、ステップ 2 へ

戻る。

この交叉は EAX-dMSXF と呼ばれる。一つの両親からは $k_{max} \times \mu$ 個の子個体が生成される。池田 [Ikeda 02] らの実験では、EAX-dMSXF を用いた GA は EAX-Rand や EAX-1AB を用いた場合に較べて良好な性能を示すと報告している。それらの実験では、 (k_{max}, μ) の値として (4, 6) または (5, 8) が用いられた。

この方法は、パス再結合 (path-relinking) 法 [Glover 94] で用いられる局所探索に交叉 (EAX-1AB) を用いたものと捉えることができる。パス再結合法では新たな解候補を生成する際に、集団から選択された 2 つの解候補の一方の解候補からスタートし、局所探索を用いて他方の解候補へ近づくように、順次新たな解候補の生成と更新をしていく方法である。

3. 提案手法

3.1 接近法

以下に述べる世代交代モデルが用いられるものとする。このモデルは [Ikeda 02] において TSP に適用された場合に良好な結果をもたらしており、本論文の実験でも TSP に適用した場合に良い性能を示すことを確認したので用いた。

[世代交代モデル I]

0. N_{pop} と N_{cross} を固定値とし、それぞれ集団サイズ、各両親から生成される子個体の数とする。
1. $t = 0$ とする。ランダムに N_{pop} 個の解候補を生成し、これを $x_1, x_2, \dots, x_{N_{pop}}$ とする。
2. 個体に付けられたインデックスをランダムにつけ直す。
3. $i = 1, \dots, N_{pop}$ のそれぞれに対し、 x_i と x_{i+1} を両親として選択する。ただし、 $x_{N_{pop}+1} = x_0$ 。
4. それぞれの両親 (x_i, x_{i+1}) に対し、交叉により N_{cross} 個の子個体を生成する。
5. それぞれの両親 (x_i, x_{i+1}) から生成された子個体と親 x_i の中から最良個体を選び、 x_i と置換える。
6. もし、集団内の最小巡回路長の改善が 20 世代 (ステップ 2 からステップ 5 までで 1 世代) にわたり改善しなかった場合終了、そうでない場合は $t = t + 1$ として、ステップ 2 へ。

以下では上記の世代モデルのステップ 4-5 において、両親のペア (x_i, x_{i+1}) をそれぞれ A, B と表示する。また距離 $d(A, B)$ を A, B 間で異なる枝の数と定義する。例えば、 B を A から 2-opt 近傍 [Croes 58, Reinelt 94] で一回遷移させた解候補とすると、 $d(A, B) = 2$ となる。ここで C を A, B から交叉で生成された子個体とすると、これらの個体間の距離の間には次のような関係が成立する。

$$\begin{aligned}
 d(A, B) &= d(A, C) + d(B, C) - n_{new} - n_{sr} \\
 &= n_a + n_b + n_{new} - n_{sr} \tag{1}
 \end{aligned}$$

ただし、 n_{new} は両親 A, B に含まれない枝で新たに子個

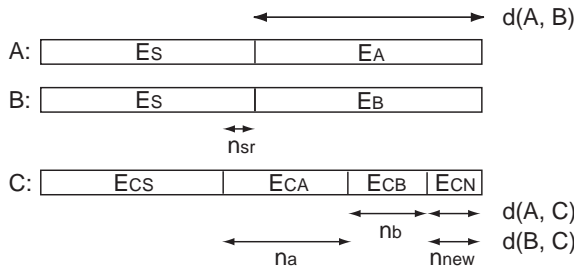


図 2 $d(A, B)$, $d(A, C)$, $d(B, C)$ の間の関係. E_s は両親 A, B に共通して含まれる枝の集合, E_A, E_B はそれぞれ, A または B のみに含まれる枝の集合. $E_{CS}, E_{CA}, E_{CB}, E_{CN}$ は子 C を構成する枝の集合で, それぞれ, E_s, E_A, E_B から継承されたか, あるいは新しく生成された枝であることを示す. n_a, n_b, n_{new} は図中対応する集合に含まれる枝の数, n_{sr} は両親 A, B に共通する枝で子個体 C に継承されなかった枝の数を表す.

体 C に導入された枝の数, n_{sr} は両親 A, B に共通する枝で子個体 C に継承されなかった枝の数である. n_{new} 本の枝は EAX アルゴリズムのステップ 5 において, 両親に含まれない新しい枝として子個体に導入される. n_{new} や n_{sr} は多くの場合個体を構成する全枝数の 2% 以下であることが実験的に分かっている. n_a, n_b はそれぞれ子個体 C に含まれる親 A または B のみに含まれる枝の数である. 図 2 に式 (1) の説明図を示す.

生存選択 (世代交代モデル I のステップ 5) において, 親 A を子 C で置換えることは, 集団内の 2 個体 x_i, x_{i+1} 間の距離 $d(x_i, x_{i+1})$ を変化させるが, EAX を用いた場合はこの値は減少する傾向にあり, 集団内の 2 個体間を類似させる要因となる. これを局所的多様性の損失 (LDL) と呼ぶこととし, その損失量を以下のように定義する.

$$LDL = d(A, B) - d(C, B) = n_b - n_{sr} \quad (2)$$

交叉に EAX-Rand を用いた場合, 中間個体は TSP の緩和条件 (各頂点に 2 本の枝が接続している) の下で, 親 A の枝と親 B の枝をランダムに組み合わせられて構成される. そのため, 子 C は親 A の枝と親 B の枝をほぼ均等に含むよう生成され, $n_a \approx n_b$ となる. 式 (1), (2) と n_{new}, n_{sr} が十分小さいことを考慮すると, EAX-Rand を用いた場合の LDL は以下の傾向を持つ.

$$LDL_{Rand} \approx \frac{1}{2}d(A, B) \quad (3)$$

一方, EAX-1AB を用いた場合, 中間個体を生成する段階で, 子 C は親 A の枝を親 B の少数の枝で置換えて生成されるため, $n_b < n_a$ となる傾向を持つ. n_{new}, n_{sr} が十分小さいことを考慮すると, EAX-1AB を用いた場合の LDL は以下の傾向を持つ.

$$0 < LDL_{1AB} < LDL_{Rand} \quad (4)$$

すなわち EAX を用いた場合, LDL はたいていの場合正の値となるが, EAX-1AB は EAX-Rand に比べ LDL の小さな子を生成する傾向があるといえる.

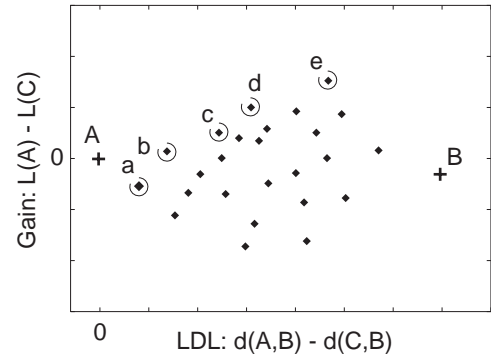


図 3 両親 A, B から生成された子個体のゲイン (Gain) と局所的多様性の損失 (LDL) の例

親 A が子 C で置換えられることにより, 集団内の 1 個体 x_i の巡回路長 $L(x_i)$ は改善 (減少) する. この改善量をゲイン (Gain) と呼ぶこととし, 以下のように定義する.

$$Gain = L(A) - L(C) \quad (5)$$

多くの世代交代モデルでは生存選択において子個体を選択する場合, 近似解最小化の観点からは最も近似解の評価値の高いものを選択する. 例えば, 上記の世代交代モデル I のステップ 5 において親 A と置換えられる子個体を C_s とすると, 最も大きなゲインを持つ個体を C_s として選択する. 一方, 集団の多様性維持の観点からは, 生存選択においてなるべく集団の多様性を維持するような子個体を C_s として選択することが望ましい. 例えば, ランダムに選択した子をおある割合で選択するという方法もある [佐藤 97]. しかし, 集団の多様性としてどのような指標を用いるべきかについては, これまでのところ明確な指針があるとはいえないのが現状である. 本論文では, 生存選択における多様性維持の指針として, 局所的多様性の損失 (LDL) を小さくすることを提案する. 局所的多様性の損失が集団全体の多様性と探索性能に与える影響については 3.2 節で示す.

本論文で提案する手法では生存選択において子個体を選択する際, 大きなゲインを得ることと, 局所的多様性の損失を小さくすることの二つの指標を考慮する方法を提案する. 例えば, ある両親のペア A, B から生成された子個体が, 図 3 に示されるような LDL と Gain の分布を持っていた場合, 5 個のパレート点が存在する (図中, a-e). これらの候補から適切に子個体を選択することができれば, GA の探索性能が改善されることが期待できる. どのような基準で子個体を選択するかについては 3.3 節で述べる.

これまでに, 生存選択における個体の評価に近似解の評価値と集団の多様性を陽に考慮した手法として, 前川ら [前川 97] の提案した TDGA がある. この手法においては, 選択された子個体が生存した場合の集団の多様性への影響を計算する際に, 選択子個体と集団全体との関

係を計算した．一方，本論文で提案する局所的多様性の損失を考慮する手法では，選択子個体が生存した場合の多様性への影響は両親と子個体間の局所的な評価を用いるため，その計算は容易であることが特徴である．

3.2 予備実験

本節では前節で述べた生存選択における二つの指標，(i) ゲイン (Gain) を大きくすること，(ii) 局所的多様性の損失 (LDL) を小さくすること，それぞれが EAX を用いた GA において集団の多様性を維持し，探索性能を改善する上で有効であることを実証する．

§1 各交叉の比較

図4に EAX-Rand, EAX-1AB, EAX-dMSXF によって生成された子個体に対し，ゲイン (縦軸) と局所的多様性の損失 (横軸) の例を示す．問題は rat575 (都市数:575, 最適巡回路長: 6773) を用い，図中 A, B は GA の探索途中からサンプルした両親である．それぞれ巡回路長は 6834, 6823 を持ち，距離 $d(A, B)$ は 91 を持つ．EAX アルゴリズムを適用することにより，この両親から 29 個の AB-cycle が生成された．EAX-Rand では AB-cycle をランダムに選択して E-set を構成し，100 個体を生成した．EAX-1AB では 29 個体を生成した．EAX-dMSXF では $(k_{max}, \mu) = (4, 6)$ を用い，24 個の個体を生成した．ただし，それぞれの交叉において同一個体が生成される場合もあるため，プロットした点の数は生成した子の数とは異なる．

図4に示すように，生存選択において EAX-Rand を用いて最も Gain の大きな子個体を選択することで他の交叉方法より大きなゲインを得られるが，もっとも局所的多様性の損失の大きな交叉といえる．一方 EAX-1AB では得られるゲインは少ないものの，局所的多様性の損失はもっとも小さくなっているのが分かる．EAX-dMSXF は，他の二つの交叉の中間的な性質をもつ交叉といえる．図4は探索の一局面的のみを示したものであるが，各交叉の間で概ねこのような関係がある．LDL に関しては式 (4) でも述べた．Gain に関しては，4章の表2で示すように，GA の一試行当りで最良解に到達するまでに要する世代数は EAX-Rand, EAX-dMSXF, EAX-1AB の順で少ないことから，この順で選択子個体の Gain が大きいといえる．

§2 Gain の影響

LDL を一定にして Gain を変化させたときの探索性能の変化について調べるために，以下の GA の比較を行う．[設定 1] 前節の世代交代モデル I を用い，交叉として

EAX-Rand を用いる． $N_{cross} = 100$ として各両親から 100 個の子個体を生成し，生存選択において Gain が k 番目に大きな子個体を選択する． $Gain > 0$ なら親 A と置換える． $Gain \leq 0$ の場合は， $Gain > 0$ となる子が存在すれば，その中から最も Gain の小

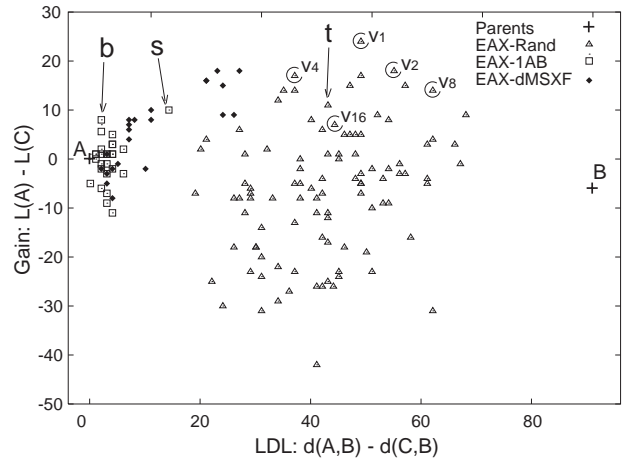


図4 両親 A, B から生成された子個体のゲイン (Gain) と局所的多様性の損失 (LDL) の関係

さな子個体を選択して親 A と置換える*3．この選択方法を Best- k 選択と呼び， k の値としては 1, 2, 4, 8, 16 を用いる．

図4には 3.2.1 節で述べたように両親のサンプル A, B から EAX-Rand で生成された 100 個体が示されているが，Gain と LDL の間に相関はほとんどないことが分かる*4．図4中 v_k は生成された 100 個の子個体のうち Best- k 選択で選択された子個体を示しているが，LDL は k に依存しない．よって，小さな k を用いるほど LDL は同程度で Gain の大きな子個体を選択して親 A と置換えることができる．

表1に実験結果を示す．ここで集団サイズ (N_{pop}) は 300, 適用問題として att532, rat575, pcb1173 を用い，初期解の生成は 2-opt [Croes 58, Reinelt 94] を用いた局所探索により生成した．各設定に対し 50 試行を行った．表1に示されるように， k が増加するにしたがって最適解を得られる回数と解の質の平均値が悪化している．すなわち，生存選択において LDL を一定とした場合，Gain が大きくなることで探索性能が改善していることが分かる．

Gain の大きな子個体を生成するためには N_{cross} を大きくして Gain 最大子個体を選択すればよいが， N_{cross} を大きくすると計算時間は増加してしまう．EAX-Rand を用いる場合，多くの問題では N_{cross} を 50 程度とするのが実験的には適切であった．

§3 LDL の影響

次に，Gain を一定にして LDL を変化させたときの探索性能の変化について調べるために，次の二つの GA の

*3 このようにならないと，探索の終盤では k が大きいとほとんどの場合 $Gain \leq 0$ となり，探索が進まなくなる．

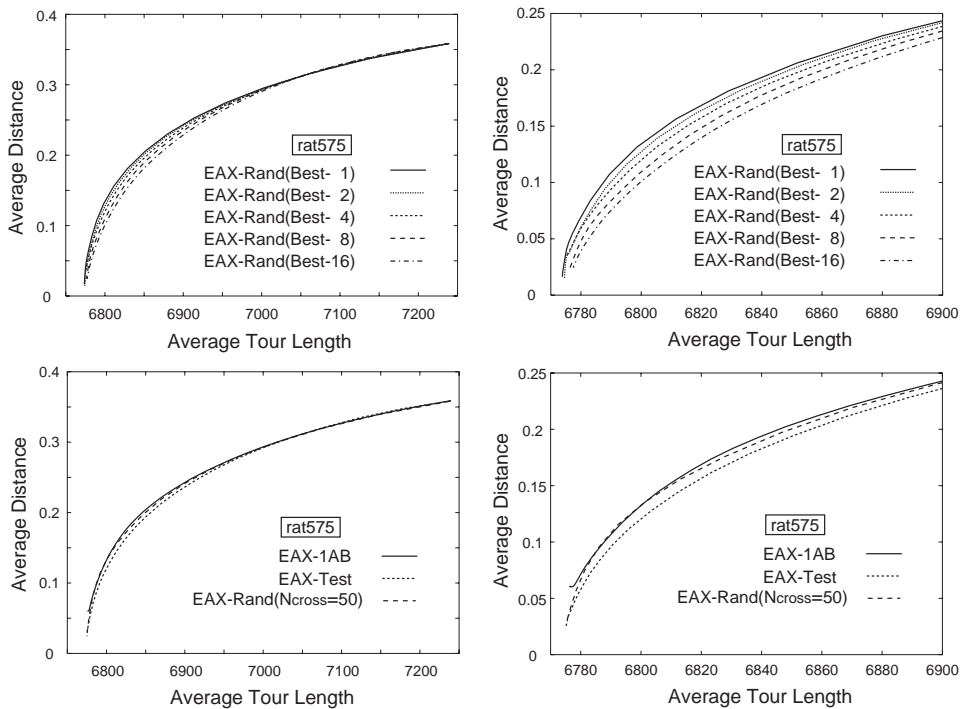
*4 正確には，個々の親のペア A, B から生成される子個体において Gain と LDL の間には相関がある．例えば $L(B) < L(A)$ の場合は，LDL と Gain には正の相関がある．ただし，平均的には $L(A) = L(B)$ であるので，LDL と Gain は無関係であるといえる．

表 1 EAX-Rand(設定 1) を用いた場合で、生存選択において Best-k 選択を行った場合の性能の比較 (左) . EAX-1AB (設定 2) と EAX-Test (設定 3) の性能比較 (右) . 集団サイズは 300 を用い、50 試行を行った . opt. は最適解へ到達した回数, err. は最適解からの誤差平均, gen. は各試行で最良解を得た時点での世代数の平均を表す .

		att532			rat575			pcb1173		
Best-k		opt.	err.(%)	gen.	opt.	err.(%)	gen.	opt.	err.(%)	gen.
EAX-Rand	1	34	0.008	13.0	20	0.011	15.0	37	0.002	19.4
	2	24	0.015	15.2	5	0.020	16.9	33	0.003	22.7
	4	21	0.016	18.1	3	0.024	19.7	22	0.007	27.3
	8	10	0.027	23.6	0	0.032	24.7	4	0.014	34.0
	16	1	0.045	32.5	0	0.046	34.1	0	0.027	49.0

		att532			rat575			pcb1173		
Ncross		opt.	err.(%)	gen.	opt.	err.(%)	gen.	opt.	err.(%)	gen.
EAX-1AB	30	26	0.015	35.1	35	0.004	42.3	37	0.002	63.3
EAX-Test	50	18	0.023	22.1	1	0.025	23.0	8	0.012	34.4

図 5 GA の探索の進行度 (集団の評価値の平均) に対する集団の多様性 (集団の平均距離) の推移 . 上図は、EAX-Rand(設定 1) を用いた場合で、Best-k 選択を行った場合の比較 . 下図は、EAX-1AB (設定 2) と EAX-Test (設定 3) の比較 . 右図は探索終盤における拡大図 . 距離は最大で 1 になるように規格化されている .



比較を行った .

[設定 2] 前節の世代モデル I を用い、交叉として EAX-1AB を用いる . N_{cross} を 30 とし、生存選択において Gain 最大の子個体を選択 . $Gain > 0$ なら親 A と置換える .

[設定 3] 前節の世代モデルを用い、交叉として EAX-Rand を用いる . N_{cross} を 50 とし、仮想的に EAX-1AB を適用した場合に得られる最大の Gain に最も近い Gain を持つ子個体を選択 . $Gain > 0$ なら親 A と置換える . これを EAX-Test と呼ぶ .

図 4 には両親のサンプル A, B から、生存選択において、設定 2 で選択される子個体を s, 設定 3 で選択される子個体を t で示している . 設定 2 と設定 3 では Gain はほぼ同じであるが、設定 2 の方が LDL の小さな生存選択の方法となることが分かる . 表 1 に実験結果を示す . 表に示されるように、EAX-1AB(設定 2) の方が EAX-Test(設定 3) より多くの最適解を発見しており解の質の平均値も

良い . すなわち、生存選択において Gain が同じであれば LDL を小さくすることで探索性能が改善されることが分かる .

§ 4 Gain と LDL の集団の多様性への影響

生存選択において Gain が大きな子個体を選択すること、LDL が小さな子個体を選択することが集団の多様性へ与える影響について考察する .

通常、GA の探索の進行に従い集団の多様性は失われるが、本論文では探索の進行度の指標として集団の巡回路長の平均値を、集団の多様性の指標として集団内のすべての 2 個体間の組み合わせにおける距離平均を考える . 多様性維持の観点からは、同一の探索の進行度であれば、集団の多様性は高い方が好ましいといえる .

図 5 に設定 1-3 を用いた場合の探索の進行度合い (横軸) と集団の多様性 (縦軸) の遷移を示す . グラフは各設定における 10 試行の平均で示している . また、紙面の都合上 rat575 における結果のみを示しているが、その他

の問題においても同様の傾向を示す。

設定 1 の比較から、生存選択で Gain が大きな個体を選択する方が集団の多様性を維持して探索をすすめることができることが分かる。また設定 2 と 3 の比較から LDL が小さな個体を選択することで集団の多様性を維持して探索をすすめる効果があることが確認できる。同じ探索の進行度合いを持つ 2 つの集団を比較した場合、多様性が高い集団ほど最適解への到達可能性は高いと考えられる。ただし、ここで用いた、探索の進行度と集団の多様性の指標は計算の簡便性から採用したものであり、適切な指標であるという保証はなく、参考程度に解釈するのが妥当であろう。

3.3 提案方法

3.1 節ではゲイン (Gain) の最大化と局所的多様性の損失 (LDL) の最小化という二つの指標を考慮することを提案し、3.2 節ではそれぞれの指標が探索性能を改善する上で重要であることを示した。提案手法では両者の関数として一次元化した評価関数を用いることとする。本論文では単位局所的多様性損失当りのゲインを生成子個体の評価関数として以下のように定義する。

$$eval(C) = \begin{cases} \frac{Gain}{LDL} & (Gain > 0, LDL > \epsilon) \\ \frac{Gain}{\epsilon} & (Gain > 0, LDL \leq \epsilon) \\ 0 & (Gain \leq 0) \end{cases}$$

ただし、 ϵ は十分小さな正数である。

この評価関数を用いて生存選択 (世代交代モデル I のステップ 5) において、評価値が最大となる子個体を選択することで、親 A のゲインに貢献する枝だけを選択的に親 B から交叉により受け継ぐこととなり、ゲインに貢献しない余計な枝の交換を制限できることが期待できる。これによって集団は必要最小限の局所的多様性の損失で巡回路の改善を行うことができると期待できる。ただし、 $Gain > 0$ かつ $LDL \leq \epsilon$ となる子個体が生成された場合には、局所的多様性の損失はないものと見なし、それらの個体の中でゲインが最大の子個体が最も高い評価値を得る*5。また、評価値が 0 である場合 ($Gain \leq 0$) は、巡回路長最小化が目的であるので、生存選択において親 A をそのまま残すこととした。

これまでに提案された EAX-Rand, EAX-1AB, EAX-dMSXF を用いて生成された子個体を比較した場合、提案した評価関数を用いて最も高い評価値を持つ子個体は EAX-1AB で生成される傾向があることが実験的に得られた。例えば図 4 において、最も高い評価値を持つ個体は b でマークされているが、これも EAX-1AB で生成されていることが分かる。そこで、提案手法では交叉として EAX-1AB を用いることにする。本論文では提案した

評価関数、交叉として EAX-1AB を用いる方法を総合して EAX-Dis と呼ぶこととする。

4. 実験と考察

本章の実験において、提案手法 EAX-Dis を用いた GA を TSP のベンチマークへ適用し、性能を EAX-Rand, EAX-1AB, EAX-dMSXF と比較することでその有効性を確認する。また、提案手法を従来提案されている EAX を用いた TSP に対する性能の良い GA [Tsai 03] や EAX を TDGA と組み合わせた場合との比較も行う。

4.1 実験設定

GA は C++ を用いて実装され、計算実験は 1.7GHz Xeon Processor (実行は 1CPU) 上で行った。世代交代モデルは 3.1 節で述べたモデルを用い、集団サイズ (N_{pop}) は 300 とした。各交叉における N_{cross} と、EAX-dMSXF におけるパラメーター (k_{max}, μ) は予備実験の結果から各手法に対して適当と思われるものを用いた*6。また、サイズの大きな問題に対しては N_{pop} を 1000 にした実験も行った。探索序盤における解の改善は容易であるため、初期解の生成は 2-opt [Croes 58, Reinelt 94] を用いた局所探索により生成した。

各試行において集団内の最短巡回路長の改善が 20 世代にわたり停滞した場合終了条件とした。しかし、交叉に EAX-1AB および EAX-Dis を用いた場合、終了条件においても集団は十分収束していない状況、すなわち改善の余地を残したまま探索が打ち切られていることが確認された。この理由は、この二つの交叉は EAX アルゴリズムのステップ 3 において $E-set$ を構築する際に、探索終盤では集団内の個体が類似するため N_{cross} にくらべ少数の $AB-cycle$ しか生成できず、EAX-1AB では EAX-Rand にくらべ少ない種類の子個体しか生成できないからである。そこで、EAX-1AB および EAX-Dis を用いた場合には、GA が終了条件に達した後は EAX-Rand を用い、再び終了条件を満たした場合に探索の終了とした。

4.2 実験結果と考察

表 2 に探索性能に関する実験結果を示す。提案手法では集団サイズ 300 を用いて、都市サイズ 2392 までの問題において u1432 を除いて*7 90% 以上の確率で最適解へ到達しているのが分かる。また、都市数が 3000 以上の問題については集団サイズ 1000 での実験も行った。その結果、それらの問題でもほぼ 100% 最適解が得られる

*6 EAX-Rand, EAX-1AB, EAX-Dis に関しては N_{cross} を 10, 30, 50, 100 を比較した結果ロバストに良い性能を示すものを、EAX-dMSXF に関しては [池田 02] に従った。

*7 u1432 は頂点が格子状に並んでいる個所が多く、同一の巡回路長に対し多くの異なる巡回路が存在する。よって、探索集団を有望な探索領域に絞り込むところが困難となり、GA での探索には難しい問題であると考えられる。

*5 探索の序盤を除き、そのような場合はほとんど起らない。

結果が得られた。表から分かるように同一集団サイズで比較を行った場合、提案手法は比較手法に比べると明らかに高頻度で最適解を得ることに成功していることが分かる。

図 6 は図 5 と同様に各手法を用いた場合の集団の探索の進行度に対する集団の多様性維持の度合いを示した。提案手法は比較手法に比べ最も多様性を維持した状態で探索が進行しており、結果としてすべての問題において最も多い最適解への到達回数を実現しているといえる。

一方、表 2 を見ると、提案手法の多様性維持能力により収束までに要する世代数は増加している。しかし、計算実行時間は世代数の増加ほどには増加していないことがわかる。これは EAX-Dis あるいは EAX-1AB を用いた場合、探索の後半になるにつれて両親が類似してくる（すなわち多様性が失われてくる）ため、生成できる *AB-cycle* の数が少なくなり、各両親からの生成子個体数は実際には N_{cross} よりも少なくなるからである。また、EAX-1AB や EAX-Dis は交叉が局所的に行われるため、交叉に要する計算コストが EAX-Rand や EAX-dMSXF にくらべ少ないことも要因の一つである。

表 3 には、比較手法の計算コストが提案手法と同一となるよう比較手法の集団サイズ調節した場合で、最適解への到達回数を比較した結果を示す。集団サイズの調節は表 2 で示した計算時間に反比例するように調節した。EAX の TSP への適用においては、集団サイズを増やすことで得られる解の質は向上し、計算コストの増加は集団サイズにほぼ比例するためこのような比較を行った*8。同一計算コストでの比較においても、提案手法 EAX-Dis は比較手法との比較において明らかな優位性を示していると言える。

さらに、提案手法は上記の比較手法以外で従来提案されている TSP に対する性能の良い GA [Tsai 03] などに比較しても高い性能を示す。[Tsai 03] では EAX を用いた GA に NJ mutation と呼ばれる突然変異と、HSGA と呼ばれる複製選択において集団中で比較的類似した個体のみを両親として選択する方法を導入することで、TSP に対して良い探索性能が実現されることを示した論文である。表 4 に HSGA の性能を示す。ただし、表中でのデータは [Tsai 03] からの抜粋である。表 2 との比較により、提案手法は TSP に対して非常に高い近似能力をもつと結論できる。例えば、[Tsai 03] では、集団サイズを問題の頂点数（あるいは頂点数の半分）に設定している。一方、提案手法では集団サイズは 300、あるいは pcb3038, fnl4461 においては 300 または 1000 と設定しているにもかかわらず、u1432 を除いて HSGA と同等か高い確率で最適解を得ることに成功している。計算時間においては

*8 集団サイズ以外にも N_{cross} を増加させることで解の質を向上させることができる。表 2 で示されている実験では、 N_{cross} は予備実験により適切なパラメータとなっており、 N_{cross} を増加させても集団サイズを増加させるほどの解の質の改善は見られないため、集団サイズを調節することとした。

明らかに提案手法が EAX-HSGA と比較して少ないことがわかる。

また、TDGA を [前川 97] に基づき EAX-rand を交叉として実装した結果も表 4 に示す。ただし、TDGA を規模の大きな問題へ適用した場合の計算コストは膨大であったため、小さなサイズの問題に対してのみ適用した。集団サイズとしては 100 と 300 を用いた。TDGA を用いた場合は集団サイズ 300 では、最適解到達回数は提案手法と比較して同等かやや低いが、計算コストは圧倒的に大きくなっている。集団サイズを 100 にしても計算コストは依然高く、最適解を得る確率は大きく減少する。

5. ま と め

本論文では交叉 EAX を用いた GA の TSP に対する近似性能を改善するために、生存選択において多様性維持を考慮した子個体の評価関数を導入した GA を構成し、提案手法を TSP のベンチマークへ適用することで有効性を示した。

提案手法では世代交代モデルとして、両親から生成された子個体の中で最も評価の高い子個体を一方の親と交換するという局所的な生存選択の方法を採用し、子個体の評価指標として評価値のゲイン (Gain) と局所的多様性の損失 (LDL) という概念を導入した。Gain と LDL はそれぞれ生成子個体を一方の親で置換えた場合における近似値の改善量、集団中の 2 個体 (両親) 間の距離の減少量で定義される。予備実験において、Gain が大きな子個体を、あるいは LDL が小さな子個体を選択して生存選択を行うことが、集団の多様性を維持して探索を進めることに貢献し、結果として GA が高い確率で最適解を発見できることを示した。

提案手法では、二つの指標を一次元化したものとして、単位 LDL 当りの Gain が最大になるとなるような子個体の選択方法を提案し、交叉 EAX を用いて GA を構成した。そして、提案手法を TSP のベンチマークに適用した結果、生存選択において二つの指標を独立に用いるよりも良い性能を示すことを確認した。また、LDL の計算が両親と子個体間の局所的な関係のみで計算されるため、計算コストが無視できるほど小さいことも確認した。さらに、提案手法が TSP の近似手法として従来提案されている性能の良い GA の性能を上回ることを示した。

提案手法は EAX を用いた GA の TSP への適用に限らず、一般の最適化問題への適用に対しても容易に応用することができる。その際に必要となるのは、個体間の距離を定義することのみである。さらに、LDL は局所的に定義されるため、クラスタリング GA のような集団がクラスター化された GA に対する多様性維持手法としても利用が期待できる。今後は本手法を別の最適化問題へと適用し、有効性を確認していきたいと考えている。

表 2 EAX-Rand, EAX-1AB, EAX-dSDXF, EAX-Dis の性能比較. 各設定で 50 試行を行った. opt. は最適解へ到達した回数. err. は最適解からの誤差平均. gen. は各試行で最良解を得た時点での世代数の平均. time(s) は一試行当りの実行時間の平均 (秒) を表す. CPU は 1.7GHz Xeon Processor を使用.

EAX-Rand						EAX-1AB							
instance	Npop	Ncross	opt.	err.(%)	gen. time(s)	instance	Npop	Ncross	opt.	err.(%)	gen. time(s)		
att532	300	50	33	0.0097	14.6	19	att532	300	30	26	0.0152	35.1	12
rat575	300	50	17	0.0141	16.6	26	rat575	300	30	35	0.0047	42.3	17
u724	300	50	45	0.0026	15.3	40	u724	300	30	45	0.0022	35.7	19
vm1084	300	50	11	0.0099	19.0	67	vm1084	300	30	24	0.0083	49.3	56
pcb1173	300	50	35	0.0033	21.3	72	pcb1173	300	30	37	0.0020	63.3	48
u1432	300	50	12	0.0259	20.6	212	u1432	300	30	12	0.0283	81.8	89
vm1748	300	50	23	0.0224	24.3	323	vm1748	300	30	18	0.0308	88.7	200
pr2392	300	50	29	0.0043	27.3	413	pr2392	300	30	40	0.0015	117.1	328
pcb3038	300	50	0	0.0165	38.6	622	pcb3038	300	30	8	0.0077	182.8	794
	1000	50	0	0.0095	35.7	1859		1000	30	12	0.0047	177.6	2321
fnl4461	300	50	0	0.0546	66.0	2637	fnl4461	300	30	2	0.0065	347.3	3095
	1000	50	0	0.0237	58.8	7064		1000	30	4	0.0021	313.8	8742

EAX-dMSXF						EAX-Dis							
instance	Npop	k _{max} , μ	opt.	err.(%)	gen. time(s)	instance	Npop	Ncross	opt.	err.(%)	gen. time(s)		
att532	300	4, 6	31	0.0103	20.1	21	att532	300	30	48	0.0010	49.5	17
rat575	300	4, 6	45	0.0014	25.2	25	rat575	300	30	50	0.0000	61.6	24
u724	300	4, 6	46	0.0016	20.0	28	u724	300	30	50	0.0000	54.0	29
vm1084	300	4, 6	23	0.0095	24.4	66	vm1084	300	30	47	0.0011	62.2	77
pcb1173	300	4, 6	28	0.0031	35.3	93	pcb1173	300	30	49	0.0000	85.1	70
u1432	300	5, 8	22	0.0200	45.3	140	u1432	300	30	26	0.0105	110.8	150
vm1748	300	5, 8	36	0.0072	32.3	211	vm1748	300	30	50	0.0000	96.7	253
pr2392	300	5, 8	42	0.0019	39.8	302	pr2392	300	30	50	0.0000	164.2	464
pcb3038	300	5, 8	4	0.0090	70.2	691	pcb3038	300	30	34	0.0013	262.1	1129
	1000	5, 8	20	0.0027	72.0	2146		1000	30	49	0.0000	252.9	3074
fnl4461	300	5, 8	2	0.0820	217.1	3387	fnl4461	300	30	29	0.0013	535.3	4730
	1000	5, 8	9	0.0019	194.6	10369		1000	30	50	0.0000	527.8	11843

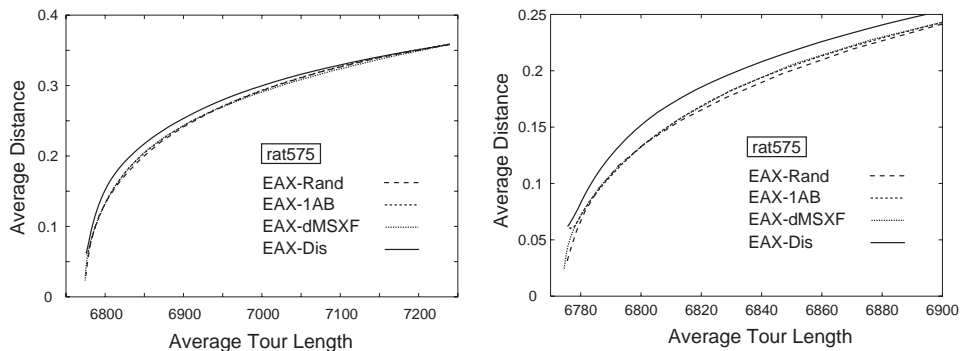


図 6 GA の探索段階 (集団の評価値の平均) に対する集団の多様性 (集団の平均距離) の推移. EAX-Dis とその他の手法との比較. 右図は探索終盤における拡大図. EAX-1AB と EAX-dMSXF のグラフはほぼ重なっている.

表 3 EAX-Rand, EAX-1AB, EAX-dMDXF, EAX-Dis の性能比較. 計算コストが各手法で同じになるよう, 集団サイズを調節している. 50 試行中での最適解へ到達した回数 (opt.) を示す.

instance	EAX-Rand			EAX-1AB			EAX-dMSXF			EAX-Dis		
	Npop	Ncross	opt.	Npop	Ncross	opt.	Npop	k _{max} , μ	opt.	Npop	Ncross	opt.
att532	288	50	24	428	30	24	248	4, 6	36	300	30	48
rat575	272	50	9	444	30	35	294	4, 6	37	300	30	50
u724	298	50	41	423	30	50	332	4, 6	49	300	30	50
vm1084	310	50	19	398	30	31	344	4, 6	27	300	30	47
pcb1173	308	50	37	424	30	40	358	4, 6	40	300	30	49
u1432	295	50	17	418	30	20	316	5, 8	26	300	30	26
vm1748	330	50	29	400	30	24	350	5, 8	38	300	30	50
pr2392	408	50	35	388	30	46	404	5, 8	45	300	30	50
pcb3038	435	50	1	444	30	2	414	5, 8	8	300	30	34
	1290	50	4	1356	30	16	1408	5, 8	25	1000	30	49
fnl4461	538	50	0	458	30	2	418	5, 8	1	300	30	29
	1676	50	0	1354	30	12	1142	5, 8	8	1000	30	50

表 4 EAX-HSGA, EAX-TDGA の性能. EAX-HSGA は [Tsai 03] からの引用であるが, 表 2 に含まれている問題と共通しているものについてのみ示す. opt. は 20 試行での最適解到達回数. time(s) は一試行当りの実行時間の平均 (秒) であるが, CPU は Pentium IV 1GHz を使用. EAX-TDGA は EAX-Rand を用いて [前川 97] に従い実装し, 50 試行を行った結果. CPU は 1.7GHz Xeon Processor を使用.

EAX-HSGA					EAX-TDGA				
instance	Npop	opt.	err.(%)	time(s)	instance	Npop	opt.	err.(%)	time(s)
att532	532	20	0.0000	97	att532	100	50	0.0000	688
vm1084	542	19	0.0033	955		300	50	0.0000	1980
pcb1173	587	18	0.0009	1177	rat575	100	8	0.0124	616
u1432	716	20	0.0000	1875		300	43	0.0020	2206
pr2392	1196	20	0.0000	4968	pcb1173	100	11	0.0100	3012
pcb3038	1519	20	0.0000	11617		300	48	0.0004	10641
fnl4461	2232	16	0.0024	50033					

◇ 参 考 文 献 ◇

- [Croes 58] G.A. Croes: A method for solving traveling-salesman problems, *Operations Research*, Vol. 6, pp. 791–812 (1958).
- [Grefenstette 85] J. Grefenstette, R. Gopal, B. Rosmaita and D. Gucht: Genetic Algorithms for the Traveling Salesman Problem, *Proceedings of the 1st International Conference on Genetic Algorithms*, pp. 160–165 (1985).
- [Glover 94] F. Glover: Genetic algorithms and scatter search: Unsuspected potentials, *Statistics and Computing*, Vol. 4, pp. 131–140, 1994.
- [Ikeda 02] K. Ideda and S. Kobayashi: Deterministic Multi-step Crossover Fusion: A Handy Crossover Composition for GAs, *Proceedings of the seventh International Conference on Parallel Problem Solving from Nature*, pp. 162–171 (2002).
- [池田 02] 池田, 小林: GA の探索における UV 現象と UV 構造仮説, *人工知能学会誌*, Vol. 17, No. 3, pp. 239–246 (2002).
- [Johnson 97] D.S. Johnson and L.A. McGeoch: The traveling salesman problem: a case study, *Local Search in Combinatorial Optimization*, E. Aarts and J.K. Lenstra, eds., pp. 215–310, London: John Wiley and Sons (1997).
- [久保 94] 久保: 巡回セールスマン問題への招待, *オペレーションズ・リサーチ*, vol. 39, No. 1 pp. 25–31 (1994).
- [前川 95] 前川, 玉置, 喜多, 西川: 遺伝的アルゴリズムによる巡回セールスマン問題の一解法, *計測自動制御学会誌論文集*, Vol. 31, No. 5, pp. 598–605 (1995).
- [前川 97] 前川, 森, 玉置, 喜多, 西川: 熱力学的選択ルールを用いた巡回セールスマン問題の遺伝的解法, *計測自動制御学会誌論文集*, Vol. 33, No. 9, pp. 939–946 (1997).
- [Nagata 03] Y. Nagata: Criteria for designing crossovers for TSP, *Proceedings of the 2004 Congress on Evolutionary Computation*, pp. 1465–1472 (2004).
- [永田 99] 永田, 小林: 巡回セールスマン問題に対する交叉: 枝組み立て交叉の提案と評価, *人工知能学会誌*, Vol.14, No.5, pp. 848–859 (1999).
- [永田 00] 永田, 小林: 遺伝的アルゴリズムを用いた巡回セールスマン問題の解法, *東京工業大学知能システム科学専攻博士論文*, (2000).
- [Reinelt 94] G. Reinelt, *The Traveling Salesman: Lecture Notes in Computer Science 840*, Springer-Verlag Berlin Heidelberg (1994).
- [佐藤 97] 佐藤, 小野, 小林: 遺伝的アルゴリズムにおける世代交替モデルの提案と評価, *人工知能学会誌*, vol. 12, No. 5, pp. 734–744 (1997).
- [Tsai 02] H.K. Tsai, J.M. Yang, and C.Y. Kao: Solving Traveling Salesman Problems by Combining Global and Local Search Mechanisms, *Proceedings of the the 2002 Congress on Evolutionary Computation*, pp. 1290–1295 (2002).
- [Tsai 03] H.K. Tsai, J.M. Yang, Y.F. Tsai and C.Y. Kao: A heterogeneous selection genetic algorithm for traveling salesman problems, *Engineering Optimization*, Vol. 35, No. 3, pp. 297–311 (2003).
- [Watson 00] J. Watson, C. Poss D. Whitley and so on: The Traveling Salesrep Problem, *Edge Assembly Crossover*, and 2-opt, *Proceedings of the fifth International Conf on Parallel Problem Solving from Nature*, pp. 823–833 (2000).
- [Whitley 89] D. Whitley, T. Starkweather and D. Fuquay: Scheduling Problems and Traveling Salesman: The Genetic Edge Recombination Operator, *Proceedings of the 3rd International Conference on Genetic Algorithms*, pp. 133–140 (1989)
- [山田 97] 山田, 中野: 遺伝的局所探索法によるジョブショップスケジューリング問題の解法, *情報処理学会論文誌*, vol. 38, No. 6, pp. 1126–1138 (1997).
- [山村 92] 山村, 小野, 小林: 形質遺伝性を重視した遺伝アルゴリズムに基づく巡回セールスマン問題の解法, *人工知能学会誌*, Vol. 10, No. 6, pp.1049–1059 (1992).

〔担当委員: 小野田 崇〕

2005 年 8 月 16 日 受理

—— 著 者 紹 介 ——



永田 裕一(正会員)

1994 年 9 月 東京工業大学応用物理学卒業, 2000 年 3 月 同大学院総合理工学研究科知能システム科学専攻博士後期課程修了. 工学博士. 同年 4 月 同大学院総合理工学研究科リサーチアソシエイト, 2001 年 4 月より 北陸先端科学技術大学院大学情報科学専攻助手, 現在にいたる. 遺伝的アルゴリズム, 機械学習の研究に従事