ショートノート

サブツアー交換交叉に対する 2 つのコメント

Two Comments on the Subtour Exchange Crossover Operator

柳浦 睦憲[*1]

Mutsunori Yagiura

永持 仁[*1]

Hiroshi Nagamochi

茨木 俊秀[*1]

Toshihide Ibaraki

*1 京都大学 工学部

Department of Applied Mathematics and Physics,

Faculty of Engineering, Kyoto University,

Kyoto 606-01, Japan.

Keywords: genetic algorithm, subtour exchange crossover.

Summary Given two tours of n cities, a pair of subtours of these tours consisting of the same set of cities is called a *common subtour*. The operation of subtour exchange crossover used in the genetic algorithm of Yamamura et al. [山村 92] is based on such common subtours. In this note, we present an $O(n^2)$ time algorithm to enumerate all common subtours, and show that the expected number of common subtours for two random tours is at most $4 + o(1)$.

1 まえがき

[山村 92] は, 行商人問題に対して遺伝アルゴリズムを適用し, 成果を収めた一例であるが, その中で提案されているサブツアー交換交叉を実現するためには, 共有サブツアーを列挙する必要が生じる. この列挙は, 次節で述べる共有区間列挙問題および共有相補区間列挙問題の両方を解くことで可能である. 本稿では, これらの問題について 2 つのコメントを与える.

その第 1 は, [山村 92] では共有サブツアーを全て列挙するのに $O(n^5)$ 時間を費やしている (但し, [山村 92] では, 巡回路の共通枝を縮約するなどの工夫をさらに加

え, 高速化をはかっている) のに対し, $O(n^2)$ 時間のアルゴリズムが可能であることを示す. 共有サブツアーの数は高々$O(n^2)$ であることから, この時間量はオーダー的に最適である. なお, [山村 92] と類似の交叉法が [Brady 85] ([Mühlenbein 88] にも紹介がある) で提案されており, その中でも共有サブツアーの列挙が必要になる. [Brady 85] には, 共有サブツアーの列挙が $O(n^2)$ で可能であることを意味する簡単なコメントはあるが, 具体的なアルゴリズムについては述べられていない. また, 小林ら [小林 93][小野 94] は, サブツアー交換交叉と同様の交叉法を, ジョブショップスケジューリング問題に適用しているが, その中では, 共有区間列挙問題に対する $O(n^3)$ 時間アルゴリズムを利用しているとの報告を受けている.

第 2 の話題として, 2 つの順列σ_A, σ_Bを, 一様にランダムに発生したとき, 共有区間数の期待値が高々$O(1)$ であること, および, 共有相補区間数の期待値が $o(1)$ であることを示す. これらの結果より, 共有サブツアーの個数の期待値も高々$O(1)$ (具体的には$4+o(1)$) であることが導かれる. 従って, サブツアー交換交叉を用いた遺伝

アルゴリズムでは, 一様にランダムに発生した解については, 新しい解が生成されにくい. これは, [柳浦 94] において, サブツアー交換交叉を 1 機械スケジューリング問題向きに手直しして用いた際, 実際に観測された現象である. 小林らの計算実験 [小林 93][小野 94][山村 92] においても類似の現象が観測されたとの報告を受けている. 但し, 遺伝アルゴリズムにおいては, 淘汰法の工夫や突然変異の導入などで, 候補解の多様性をはかることができ, また, 候補解としてランダムに生成されたものだけではなく, すでに良い近似解になっているものを用いると, より多くの共有サブツアーが含まれる傾向にあるので, これが直ちにサブツアー交換交叉の非効率性を示すわけではないことを注意しておく. 実際, [小林 93][小野 94][山村 92] における良好な結果は, これら遺伝アルゴリズムの総合的な効果によるものと考えられる. 本論文の結果は, サブツアー交換交叉の一つの理論的側面をとらえるものと理解される.

2 共有区間と共有サブツアー

本節では, 共有区間, 共有相補区間, および, 共有サブツアーの定義を与え, これらの関係について説明す

る. $N = \{1, \ldots, n\}$ の順列をσ: $N \to N$で表し, i 番目の要素が jであることを$\sigma(i) = j$ (または, $\sigma^{-1}(j) = i$) と記す. 2 つの順列をσ_A, σ_Bとする. σ_Aの区間 $[l_A, u_A]$ とσ_Bの区間 $[l_B, u_B]$ の対 $([l_A, u_A], [l_B, u_B])$ (但し, $1 \leq u_A - l_A \leq n-3$, $1 \leq u_B - l_B \leq n-3$) で, 条件

$$\{\sigma_A(i) \mid i \in [l_A, u_A]\} = \{\sigma_B(i) \mid i \in [l_B, u_B]\} \qquad (1)$$

を満たすものを共有区間と呼び, それらを全て列挙する問題を共有区間列挙問題という. また, 区間対 $([l_A, u_A], [l_B, u_B])$ (但し, $1 \leq u_A - l_A \leq n-3$, $1 < l_B < u_B < n$) で, 条件

$$\{\sigma_A(i) \mid i \in [l_A, u_A]\} = \{\sigma_B(i) \mid i \notin [l_B, u_B]\} \qquad (2)$$

を満たすものを共有相補区間と呼び, それらを全て列挙する問題を共有相補区間列挙問題という. 以下, $u_A - l_A + 1$ を, 共有区間, および, 共有相補区間の長さと呼ぶ. σ_Aの区間 $[l_A, u_A]$ として, その長さが 2 以上 $n-2$ 以下のものだけを考える理由は, これら以外は共有サブツアーを列挙する目的には不要となるからである. また, 共有サブツアーを列挙する際の重複を避けるため, 共有相補区間列挙問題において, $l_B = 1$ あるいは $u_B = n$ と

なる場合も考察の対象外とする.

次に, n の次が 1 と円状に連続している場合を考え, 円区間を

$$\langle s,t\rangle := \begin{cases} [s,\,t], & s \le t \text{ の場合} \\ [s,\,n]\cup[1,\,t], & s > t \text{ の場合}, \end{cases} \tag{3}$$

と定義する. また, $\langle s,t\rangle$ の補円区間を

$$\langle s,t\rangle^c := \langle t \pmod n + 1, s-2 \pmod n + 1\rangle \tag{4}$$

と定義する. 順列σ_Aとσ_Bが与えられたとき, 円区間対 $(\langle s_A,t_A\rangle,\langle s_B,t_B\rangle)$ (但し, $1 \le t_A - s_A \pmod n \le n-3$, $1 \le t_B - s_B \pmod n \le n-3$) で, 条件

$$\{\sigma_A(i) \mid i \in \langle s_A,t_A\rangle\} = \{\sigma_B(i) \mid i \in \langle s_B,t_B\rangle\} \tag{5}$$

を満たすものを, 共有サブツアーと呼ぶ. 共有サブツアーの長さを $t_A - s_A \pmod n + 1$ と定義する. サブツアー交換交叉は, 共有サブツアーを全て列挙し, その中の 1 つを等確率で選んで, σ_Aとσ_Bの対応する円区間を互いに交換することで新しい候補解を生成するというものである. このとき, 長さ 1, $n-1$, あるいは n の円区間 $\langle s_A,t_A\rangle$ については, 条件 (5) が成立するようなσ_Bの円区間が必ず存在するうえ, そのような円区間を

交換しても新たな解は生成されないため, 考察の対象から外す. 一方, 長さ k $(2 \leq k \leq n-2)$ の共有サブツアー $(\langle s_A, t_A\rangle, \langle s_B, t_B\rangle)$ が与えられると, その補円区間の対 $(\langle s_A, t_A\rangle^c, \langle s_B, t_B\rangle^c)$ も長さ $n-k$ $(2 \leq n-k \leq n-2)$ の共有サブツアーとなる. これらのことから, 共有サブツアーを全て列挙するためには,

(i) 共有区間 $\{\sigma_A(i) \mid i \in [l_A, u_A]\} = \{\sigma_B(i) \mid i \in [l_B, u_B]\}$

(但し, $1 \leq u_A - l_A \leq n-3, 1 \leq u_B - l_B \leq n-3$),

(ii) 共有相補区間 $\{\sigma_A(i) \mid i \in [l_A, u_A]\} = \{\sigma_B(i) \mid i \notin [l_B, u_B]\}$

(但し, $1 \leq u_A - l_A \leq n-3, 1 < l_B < u_B < n$),

を全て列挙し, (i) の場合は円区間対 $(\langle l_A, u_A\rangle, \langle l_B, u_B\rangle)$ および補円区間対 $(\langle l_A, u_A\rangle^c, \langle l_B, u_B\rangle^c)$, (ii) の場合は円区間対 $(\langle l_A, u_A\rangle, \langle u_B+1, l_B-1\rangle)$ および補円区間対 $(\langle l_A, u_A\rangle^c, \langle u_B+1, l_B-1\rangle^c)$ を出力すればよい. 但し, (i) と (ii) において, $\langle l_A, u_A\rangle$ が $l_A = 1$ または $u_A = n$ を満たすならば, その補円区間対は他の l_Aと u_Aの組合せで生成されるので, 重複を避けるため, 補円区間対を出力しない.

3 $O(n^2)$ 時間アルゴリズム

本節では, 共有区間列挙問題および共有相補区間列挙問題それぞれに対する $O(n^2)$ 時間のアルゴリズムを提案する.

共有区間列挙問題に対するアルゴリズムは,

$$l_A = \min_{j \in [l_B, u_B]} \sigma_A^{-1}(\sigma_B(j)), \tag{6}$$

$$u_A = \max_{j \in [l_B, u_B]} \sigma_A^{-1}(\sigma_B(j)) \tag{7}$$

とおくとき,

$$\text{条件 } (1) \Leftrightarrow u_A - l_A = u_B - l_B, \tag{8}$$

となる事実を利用したものである. 式 (6), (7), (8) は $O(n)$ 時間で計算できるので, この手続きをσ_Bの全ての区間に対して行えば, $O(n^3)$ 時間で共有区間を列挙できる. しかし, σ_Bの区間を順序よく列挙し, 式 (6), (7) の計算の反復を省くことにより, $O(n^2)$ 時間のアルゴリズムを構成できる.

アルゴリズム 1

for $l_B = 1, \ldots, n-1$ **do**

l_A:=u_A:=$\sigma_A^{-1}(\sigma_B(l_B))$;

for $u_B = l_B + 1, \ldots, \min\{n, l_B + n - 3\}$ **do**

l_A:=$\min\{l_A, \sigma_A^{-1}(\sigma_B(u_B))\}$;

u_A:=$\max\{u_A, \sigma_A^{-1}(\sigma_B(u_B))\}$;

if $u_A - l_A = u_B - l_B$ **then**

output $([l_A, u_A], [l_B, u_B])$

end for

end for.

σ_A^{-1}は, アルゴリズムを実行する前に用意しておけばよい ($O(n)$ で可能) ので, 全体の計算時間は $O(n^2)$ である.

共有相補区間列挙問題に対するアルゴリズムも同様に構成できる. 但し,

$$l_A = \min_{j \notin [l_B, u_B]} \sigma_A^{-1}(\sigma_B(j)), \tag{9}$$

$$u_A = \max_{j \notin [l_B, u_B]} \sigma_A^{-1}(\sigma_B(j)) \tag{10}$$

とおくとき,

$$\text{条件 } (2) \Leftrightarrow u_A - l_A + 1 = n - (u_B - l_B + 1), \tag{11}$$

となる事実を利用する. アルゴリズムを以下に示す.

アルゴリズム 2

for $u_B = 3, 4, \ldots, n-1$ **do**

l_A:=n, u_A:=0;

for $i = u_B + 1, \ldots, n$ **do**

l_A:=$\min\{l_A, \sigma_A^{-1}(\sigma_B(i))\}$;

u_A:=$\max\{u_A, \sigma_A^{-1}(\sigma_B(i))\}$

end for;

for $l_B = 2, 3, \ldots, u_B - 1$ **do**

l_A:=$\min\{l_A, \sigma_A^{-1}(\sigma_B(l_B - 1))\}$;

u_A:=$\max\{u_A, \sigma_A^{-1}(\sigma_B(l_B - 1))\}$;

if $u_A - l_A + 1$=$n - (u_B - l_B + 1)$ **then**

output $([l_A, u_A], [l_B, u_B])$

end for

end for.

アルゴリズム 2 の全体の計算時間は $O(n^2)$ である.

4 共有サブツアー数の期待値

2 つの順列σ_A, σ_B (要素数 n) を, 一様にランダムに発生したとき (即ち, 任意の順列が確率 $1/n!$で生成される), 長さ 2 以上かつ $n-2$ 以下の共有区間の数の期待

値 $E(n)$ が $O(1)$ であること, および, 長さ 2 以上かつ $n-2$ 以下の共有相補区間の数の期待値 $E^c(n)$ が $o(1)$ であることを示す. これらの結果より, 共有サブツアー数の期待値 $E^s(n)$ も $O(1)$ であることが分かる.

まず, 共有区間数について考える. l_Aと l_Bを定めたとき, 長さ kの区間 $[l_A, l_A+k-1]$ と $[l_B, l_B+k-1]$ が

$$\{\sigma_A(i) \mid i \in [l_A, l_A+k-1]\} = \{\sigma_B(i) \mid i \in [l_B, l_B+k-1]\} \tag{12}$$

を満たす確率は, $\{\sigma_A(i) \mid i \in [l_A, l_A+k-1]\}$ の k個の要素がσ_Bの区間 $[l_B, l_B+k-1]$ 内に並び, $\{\sigma_A(i) \mid i \notin [l_A, l_A+k-1]\}$ の $n-k$個の要素がσ_Bの区間$[l_B, l_B+k-1]$ の外に並ぶ確率に等しいから,

$$\frac{(n-k)!k!}{n!} \tag{13}$$

である. さらに, l_Bの取り得る値は, $l_B = 1, \ldots, n-k+1$ であることから, $\{\sigma_A(i) \mid i \in [l_A, l_A+k-1]\}$ の k要素が, σ_Bにおいて長さ kの区間内に現れる確率は,

$$\frac{(n-k)!k!}{n!} \times (n-k+1) \tag{14}$$

である. よって, l_Aの取り得る値が $l_A = 1, \ldots, n-k+1$ であることを考慮すると, 長さ kの共有区間数の期待

値は,

$$E_k(n) := \frac{(n-k+1)^2(n-k)!k!}{n!} \tag{15}$$

となり,

$$E(n) = \sum_{k=2}^{n-2} E_k(n) \tag{16}$$

となる. ここで, $E_k(n)$ 増減の様子を調べるため,

$E_k(n)/E_{k-1}(n) < 1$ となる条件を調べる.

$$\frac{E_k(n)}{E_{k-1}(n)} = \frac{k(n-k+1)}{(n-k+2)^2} < 1 \tag{17}$$

より,

$$2k^2 - (3n+5)k + (n^2+4n+4) > 0. \tag{18}$$

よって, 求める kの範囲は, $k < \alpha_-(n), \alpha_+(n) < k$である. ここに,

$$\alpha_-(n) = \frac{3n+5-\sqrt{n^2-2n-7}}{4}, \tag{19}$$

$$\alpha_+(n) = \frac{3n+5+\sqrt{n^2-2n-7}}{4} \tag{20}$$

である (但し, $n \geq 4$). $0 < \alpha_-(n) \leq n$ は, 容易に確かめられる. また, $n \geq 4$ では

$$(n-3)^2 \leq n^2 - 2n - 7 \tag{21}$$

が成り立つことから, $\alpha_+(n) > n$ がいえる. 以上より,

$E_k(n)$ は, $2 \leq k \leq \alpha_-(n)$ では kに関して単調減少,

$\alpha_{-}(n) \leq k \leq n$ では単調増加となる. 従って, $E_4(n) < \frac{24}{n^2}$ と $E_{n-2}(n) \leq \frac{24}{n^2}$ $(n \geq 4)$ に注意して,

$$E_k(n) \leq \frac{24}{n^2}, \quad (k = 4,5,\ldots,n-2, n \geq 4) \tag{22}$$

がいえる. また, (15) より,

$$E_2(n) = 2 - \frac{2}{n}, \tag{23}$$

$$E_3(n) = \frac{6(n-2)}{n(n-1)}. \tag{24}$$

結局, $n \geq 5$ のとき,

$$\begin{aligned} E(n) &\leq E_2(n) + E_3(n) + (n-5) \cdot \frac{24}{n^2} \\ &= 2 + o(1) \end{aligned} \tag{25}$$

となる.

次に, 共有相補区間数について考える. l_A と l_B を定めたとき, 長さ k の区間 $[l_A, l_A + k - 1]$ と長さ $n-k$ の区間 $[l_B, l_B + n - k - 1]$ が

$$\{\sigma_A(i) \mid i \in [l_A, l_A{+}k{-}1]\} = \{\sigma_B(i) \mid i \notin [l_B, l_B{+}n{-}k{-}1]\} \tag{26}$$

を満たす確率は, $\{\sigma_A(i) \mid i \in [l_A, l_A + k - 1]\}$ の k 要素が σ_B の区間 $[l_B, l_B + n - k - 1]$ の外に並ぶ確率に等しいから, やはり式 (13) となる. l_A の取り得る値が

$l_A = 1, \ldots, n-k+1$ であること, および, l_Bの取り得る値が $l_B = 2, \ldots, k$であることを考慮すると, 長さ kの共有相補区間数の期待値は,

$$E_k^c(n) := \frac{(k-1)(n-k+1)(n-k)!k!}{n!} \tag{27}$$

となり,

$$E^c(n) = \sum_{k=2}^{n-2} E_k^c(n) \tag{28}$$

となる. ここで, $E_k^c(n)$ 増減の様子を共有区間のときと同様の方法で調べると, $n \leq 6$ のときは kに関して単調増加, $n \geq 7$ のときは, $0 < \alpha^c(n) \leq n$ なる$\alpha^c(n)$ が存在し, $3 \leq k \leq \alpha^c(n)$ では kに関して単調減少, $\alpha^c(n) \leq k \leq n$ では単調増加となる. 従って, $E_3^c(n) \leq \frac{16}{n^2}$ $(n \geq 4)$ と $E_{n-3}^c(n) \leq \frac{24}{n^2}$ $(n \geq 4)$ に注意して,

$$E_k^c(n) \leq \frac{24}{n^2}, \quad (k = 3, 4, \ldots, n-3, n \geq 4) \tag{29}$$

がいえる. また, (27) より,

$$E_2^c(n) = \frac{2}{n}, \tag{30}$$

$$E_{n-2}^c(n) = \frac{6(n-3)}{n(n-1)}. \tag{31}$$

結局, $n \geq 5$ のとき,

$$\begin{aligned} E^c(n) &\leq E_2^c(n) + E_{n-2}^c(n) + (n-5) \cdot \frac{24}{n^2} \\ &= o(1) \end{aligned} \tag{32}$$

となる.

最後に, 2節の考察より, 長さ kの共有サブツアー数の期待値 $E_k^s(n)$ は,

$$E_k^s(n) \leq E_k(n) + E_{n-k}(n) + E_k^c(n) + E_{n-k}^c(n) \quad (33)$$

となる. よって, 以上の結果をふまえて,

$$\begin{aligned} E_k^s(n) &\leq o(1),\ (k = 3, 4, \ldots, n-3, n \geq 4), & (34) \\ E_2^s(n) &\leq 2 + o(1), & (35) \\ E_{n-2}^s(n) &\leq 2 + o(1), & (36) \\ E^s(n) &\leq 4 + o(1),\ (n \geq 5) & (37) \end{aligned}$$

である. 即ち, 一様にランダムに発生された 2 つの順列には以下の性質が成り立つ.

- 長さ 3 以上 $n-3$ 以下の共有サブツアーが存在することは希であり, その数の期待値は $o(1)$ である,
- 存在する共有サブツアーの長さは, 多くの場合 2 または $n-2$ であり, しかもその数の期待値は $4 + o(1)$ で抑えられる.

5　むすび

全ての共有サブツアーの列挙を $O(n^2)$ で行うアルゴリズムを提案し, また, 一様にランダムに生成した 2 つの順列の共有サブツアーの数の期待値が $4 + o(1)$ 以下であることを示した. なお, 以上の 2 つの議論は, 共有区間列挙問題と共有相補区間列挙問題の 2 つに分けず, 直接共有サブツアーを扱うことで, やや簡略化できる部分もあるが, サブツアー交換交叉のアイディアは, 円順列に限らず, 通常の順列に対しても適用できるため, 本報告の 2 つの結果も両者に適用できるように説明を行った.

謝辞　適切なコメントをいただいた, 東京工業大学の小林重信先生に厚く御礼申し上げます. なお, 本研究は一部文部省科学研究費によるものである.

参考文献

[Brady 85] R. M. Brady, "Optimization Strategies Gleaned from Biological Evolution," *Nature* **317**, pp. 804–806, (1985).

[小林 93] 小林, 小野, 山村, “遺伝的アルゴリズムによるジョブショップスケジューリングの性能評価,” 日本機械学会 [No. 930-80] シンポジウム講演論文集, pp. 27–32, (1993).

[Mühlenbein 88] H. Mühlenbein, M. Gorges-Schleuter and O. Krämer, “Evolution Algorithms in Combinatorial Optimization,” *Parallel Computing* **7**, pp. 65–85, (1988).

[小野 94] 小野, 山村, 小林, “遺伝的アルゴリズムによるジョブショップスケジューリング,” 第 19 回知能システムシンポジウム予稿集, pp. 23–28, (1994).

[柳浦 94] 柳浦, 茨木, “順序問題における遺伝的交叉法に対する一考察,” 電気学会論文誌 C, 6 月号 (1994).

[山村 92] 山村, 小野, 小林, “形質の遺伝を重視した遺伝的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 **7**, No. 6, pp. 117–127, (1992).

柳浦　睦憲 (非会員)

平 3 京大・工・数理卒. 平 5 同大学院修士課程了. 平 6 同大学院博士課程中退. 現在同助手. 組合せ最適化問題に対する近似解法の研究に興味を持つ.

永持　仁 (非会員)

昭 58 京大・農卒. 昭 60 同大学院・数理・修士課程了. 昭 63 同大学院博士課程了. 昭 63 豊橋技術科学大学助手. 平 2 京大助手, 平 5 同助教授. Simon Fraser 大学客員. グラフ理論, 組合せ最適化等の研究に従事. 工博.

茨木　俊秀 (正会員)

昭 38 京大・工・電気卒. 昭 40 同大学院修士課程了. 昭 44 京大助手, 昭 48 同助教授. 昭 58 豊橋技術科学大学教授, 昭 60 京大・工・数理工学科教授. Illinois 大学, Waterloo 大学, Simon Fraser 大学, Rutgers 大学等客員. 組合せ最適化, 計算の複雑さ等の研究に従事. 著書「アルゴリズムとデータ構造」,「最適化の手法」,「Resource Allocation Problems」など. 工博.