順序型GAの遺伝子表現と操作
―セールスマン巡回問題を例に―

小野俊彦

セールスマン巡回問題（TSP）に代表される順序問題に対する GA の適用方法について考えてみる．例えば 100 都市の TSP の場合に問題になるのは最小の巡回距離になる 100 都市の巡回順序であるので，この問題の常識的な表現方法は 1 から 100 までの整数の順列ということになる．従って，遺伝子の表現方法は 1 から 100 までの長さ 100 の整数列ということになる．もちろん，これとは異なった方法も考えられる．いずれにしても，これらの表現方法では交叉や突然変異などの遺伝子操作はバイナリビットストリングの場合とは異なってくる．実際問題を考えてみると，このような順列表現に適した問題が意外に多いことがわかる．これら遺伝子操作法は遺伝子の表現方法によっても変わってくる．特に大きく影響を受けるのは交叉方法であるので，これについて現在用いられている各種の方法につい述べる．

交叉においては子が親のそれぞれの形質，この場合では巡回経路，をどの程度受け継ぐことができるかが重要となる．TSP では巡回経路の順番が問題であり，どこから開始するかは関係がない．しかし，問題によっては順番のみならず，その位置が重要となる場合があるので，各方式の評価において TSP とその他では異なってくることもあり得る．なお，説明は主として文献 [1] を参考にしつつ行う．また TSP を対象に説明しているが，以下に述べることは他の順序・組合せ問題にも多くの場合にそのまま適用できる．順序問題に対する遺伝子の表現方法としてはパス表現，隣接表現，順序表現の 3 種と別に行列による表現があるが，ここでは前 3 者を対象とする．これら方式により交叉の方式が変わってくるので，各表現別に交叉方法及び突然変異の方法について説明する．表 1 は遺伝子表現法とそれに用いられる交叉方法とを一覧表にまとめたものである．

1　パス表現 (Path Representation)

パス表現 は巡回する都市の番号を巡回する順序に並べた表現方法であり，最も自然な表現法といえる．この表現法では，経路:

$$5-1-7-8-9-4-6-2-3$$

表 1: TSP の遺伝子表現と交叉方法

遺伝子表現方法	交叉方法
パス表現 (PR : Path Representation)	部分射像交叉 (PMX : Partially-mapped)
	順序交叉 (OX : Order)
	循環交叉 (CX : Cyclic)
	サブツア交換交叉 (Subtour exchanging)
	辺再組合せ交叉 (ER : Edge recombination)
隣接表現 (AR : Adjacent Representation)	交互辺交叉 (Alternating-edge)
	部分経路交叉 (Subtour-chunks)
	ヒューリスティック交叉 (Heuristic)
順序表現 (OR : Order Representation)	通常の交叉方法

を次のようなリストで表す.

(5 1 7 8 9 4 6 2 3)

パス表現法には以下に述べる３種の交叉法が考えられている.

1.1 部分写像交叉 (PMX: Partially-mapped crossover)

部分写像交叉 は Goldberg らにより提案された方法である．親の一方からはその部分経路をそのまま受け継ぎ，他の親からも残りの都市について，できるだけ多く親の順番を受け継ぐことを目的にしている．そのために，２つのランダムに選んだ切断点間の部分経路を選び，如何に示す交叉を行う．例えばつぎに示す２つの親の経路の交叉を考える.

$p_1 = (1\ 2\ 3 \mid 4\ 5\ 6\ 7 \mid 8\ 9)$

$p_2 = (4\ 5\ 2 \mid 1\ 8\ 7\ 6 \mid 9\ 3)$

まず縦棒 $\mid$ にて示した２点で挟まれた部分を交叉するとつぎに示す子が得られる．ここで $*$ は今のところ未決定であることを示す.

$c_1 = (*\ *\ * \mid 1\ 8\ 7\ 6 \mid\ *\ *)$

$c_2 = (*\ *\ * \mid 4\ 5\ 6\ 7 \mid\ *\ *)$

このことは次の入れ替えを定義していることにもなる.

$1 \leftrightarrow 4,\ 8 \leftrightarrow 5,\ 7 \leftrightarrow 6,\ 6 \leftrightarrow 7$

残りの未決定の部分について，既定のものと衝突を起こさないもの（即ち，既定のものに含まれないもの）はそのまま入れると，つぎの経路が得られる.

$c_1 = (*\ 2\ 3 \mid 1\ 8\ 7\ 6 \mid\ *\ 9)$

$c_2 = (*\ *\ 2 \mid 4\ 5\ 6\ 7 \mid 9\ 3)$

最後に残った部分には前記の入れ替えの定義を参照して衝突を起こさないように入れる．即ち，c_1 については１番目の $*$ には 1 の代わりに 4 を入れ，８番目の $*$ には 8 の代わりに 5 を入れる．同様の操作を c_2 についても行うと次の経路が得られる.

$o_1 = (4\ 2\ 3 \mid 1\ 8\ 7\ 6 \mid 5\ 9)$

$o_2 = (1\ 8\ 2 \mid 4\ 5\ 6\ 7 \mid 9\ 3)$

この方法は類似性と順序を保存している点に特徴がある.

1.2 順序交叉 (OX:Order crossover)

順序交叉 は Davis により提案されたものである．片方の親からは一部をそのままに受け継ぎ，残りの部分については相対的な順番は保存しつつ受け継ぐ 方法である．例として次の２つの親からの交叉を考える.

$p_1 = (1\ 2\ 3 \mid 4\ 5\ 6\ 7 \mid 8\ 9)$

$p_2 = (4\ 5\ 2 \mid 1\ 8\ 7\ 6 \mid 9\ 3)$

先ず２切断点間はそのままコピーする.

$c_1 = (*\ *\ * \mid 4\ 5\ 6\ 7 \mid\ *\ *)$

$$c_2 = (* \ * \ * \mid 1\ 8\ 7\ 6 \mid \ * \ *)$$

次に第２切断点から開始して右回りに順序を保ちながら他の親をコピーする．その際に既定のものと衝突するものがあれば除く．例えば，子 c_1 について見てみると，第２の親 p_2 の順序は

$$9-3-4-5-2-1-8-7-6$$

となるが子 c_1 と衝突するもの（4, 5, 6, 7）を除くと

$$9-3-2-1-8$$

となるので，これを c_1 に挿入するとつぎのとおりになる．

$$c_1 = (2\ 1\ 8 \mid 4\ 5\ 6\ 7 \mid 9\ 3)$$

同様にして， c_c がつぎのように得られる．

$$c_2 = (3\ 4\ 5 \mid 1\ 8\ 7\ 6 \mid 9\ 2)$$

この方法はパス表現法の特質である都市の順番の重要性，即ち

$$9-3-4-5-2-1-8-7-6$$

$$4-5-2-1-8-7-6-9-3$$

は同じ巡回経路であることを利用している．

1.3 順序交叉 (OX) の修正方式

Syswerda は順序交叉 (OX) につぎの２つの修正案を提案している．
修正案１：順序に基づく交叉 (Order-based crossover)：つぎの例をもとに説明する．

$$p_1 = (1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9)$$

$$p_2 = (4\ 1\ 2\ 8\ 7\ 6\ 9\ 3\ 5)$$

- まず親 p_2 より任意の位置にある都市を選ぶ．例えば 3,4,6,9 番目にある 2,8,6,5 を選ぶ．つぎに p_1 のこれらの数値に相当する位置（即ち 2,8,6,5 番目）以外にある要素をそのまま子 c_1 に受け継ぐ．従って

  $$c_1 = (1\ *\ 3\ 4\ *\ *\ 7\ *\ 9)$$

- 残りの場所には p_2 の $2, 8, 6, 5$ をこの順序に入れる．従って最終的に

  $$c_1 = (1\ 2\ 3\ 4\ 8\ 6\ 7\ 5\ 9)$$

  となる．同様に c_2 については次のようになる．

  $$c_2 = (3\ 1\ 2\ 8\ 7\ 4\ 6\ 9\ 5)$$

修正案２：位置に基づく交叉 (Position-based crossover)：原案と殆ど同じであり，違いは複写する都市を順番ではなくランダムに選ぶことである．

1.4　循環交叉 (CX: Cyclic crossover)

循環交叉 は Oliver により提案された方式である．次の親の場合を例に説明する．

$$p_1 = (1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9)$$

$$p_2 = (4\ 1\ 2\ 8\ 7\ 6\ 9\ 3\ 5)$$

まず，順番に残すものを決める．第 1 の子 c_1 について，親 p_1 の最初の都市の 1 より開始し，つぎは親 p_2 の 1 番目の都市は 4 であるので 親 p_1 の 4 を取り出す．つぎは p_1 の 4 の位置に相当する p_2 の都市は 8 であるので，p_1 の 8 を取り出す．この結果つぎの子ができる．

$$c_1 = (1\ *\ *\ 4\ *\ *\ *\ 8\ *)$$

この様にして c_1 を作って行くと次のリストが得られる．

$$c_1 = (1\ 2\ 3\ 4\ *\ *\ *\ 8\ *)$$

このリストで最後に入れられたのは 2 であるが，この位置の p_2 の要素は 1 であり最初に戻り，これ以上は進めない．従って残りの都市は反対側の親 p_2 のものをそのまま入れると，つぎのリストが得られる．

$$c_1 = (1\ 2\ 3\ 4\ 7\ 6\ 9\ 8\ 5)$$

他の親も同様の方法を用いると，つぎのリストが得られる．

$$c_2 = (4\ 1\ 2\ 8\ 5\ 6\ 7\ 3\ 9)$$

この方法では各要素は親の位置をそのまま受け継いでいる．

1.5　サブツア交換交叉 (Subtour exchanging crossover)

アブツア交換交叉 は山村 ら [2] により提案された交叉法である．2 つの親の間で交換されるサブツアに含まれる都市集合が一致するときのみ交叉する．例えば次の親の場合で 2 点交叉を考える．

$$p_1 = (1\ 2\ |\ 3\ 4\ 5\ |\ 6\ 7\ 8\ 9)$$

$$p_2 = (2\ 7\ 6\ 9\ |\ 4\ 3\ 5\ |\ 1\ 8)$$

両親の | で囲まれた範囲をみると都市集合がいずれも同じである，交叉を行うとつぎの子が得られる．

$$c_1 = (1\ 2\ 4\ 3\ 5\ 6\ 7\ 8\ 9)$$

$$c_2 = (2\ 7\ 6\ 9\ 3\ 4\ 5\ 1\ 8)$$

これに右回り，左回りを考えてさらに 2 つの子が作られる．

1.6　辺再組合せ交叉（ER : Edge recombination crossover）

辺再組合せ交叉 は Whitley により提案された方式であり，基本的には親の保有する辺を子に受け継ぐことにある．この方式では親の 95 % の辺を子に移している．この操作では辺に関する情報を扱う．例えば経路

$$(3\ 1\ 2\ 8\ 7\ 4\ 6\ 9\ 5)$$

では辺は

$$(3\ 1), (1\ 2), (2\ 8), (8\ 7), (7\ 4), (4\ 6), (6\ 9), (9\ 5), (5\ 3)$$

となり，各辺は距離の情報を持つことになる．目的関数は正しい経路を構成する辺の距離の合計を最小にするように働くので，都市の位置が重要でないと同様に辺の方向も重要ではない．辺 (3 1) と (1 3) とはただ都市 1 と都市 3 が直接接続していることを示しているに過ぎない．辺再組合せ交叉の基本的な考えは，子は親の両方に存在する辺から構成すべきである，にあり，これには両方の親の経路から作られた辺リストを利用する．各都市は少なくとも２つ，最大４つの都市と接続している．親の辺リストがつぎのとおりであったとする．

$$p_1 = (1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9)$$

$$p_2 = (4\ 1\ 2\ 8\ 7\ 6\ 9\ 3\ 5)$$

これより辺リストはつぎのとおりになる．
都市 1:辺で接続している都市:9 2 4
都市 2:辺で接続している都市:1 3 8
都市 3:辺で接続している都市:2 4 9 5
都市 4:辺で接続している都市:3 5 1
都市 5:辺で接続している都市:4 6 3
都市 6:辺で接続している都市:5 7 9
都市 7:辺で接続している都市:6 8
都市 8:辺で接続している都市:7 9 2
都市 9:辺で接続している都市:8 1 6 3
この親の辺リストから子の構成を次のように行う．

1. 最初の都市を親の１つより選ぶ．これを都市 1 とする．
2. 辺リストより都市 1 に接続する都市は 9,2,4 であるので，この中で最小の辺の数を持つ都市を選ぶ．都市 2,4 が３つの都市，9 が４つの都市と接続しているので 2,4 よりランダムに選んで都市 4 とする．
3. 今度は都市 4 について同じ操作を繰り返す．以下同じ．

このようにして次の子の辺リストが得られる．

(1 4 5 6 7 8 2 3 9)

実験によると辺の失敗は非常に少なかった（$1 - 1.5\%$）．

その後，辺再組合せはさらに改善された．その考えは共通部分経路である．辺リストの中に繰り返し現れる辺がある．前の例で述べると辺 (4 5) がこれに相当する．都市 4 に接続する都市は都市 3,5,1 であり，辺リストで見ると辺 (4 5) は２回現れている．これを表すために辺リストに接続する都市には記号” $-$ ”を付ける．例えば上記の例では
都市 1：辺で接続している都市：9 -2 4
都市 2：辺で接続している都市：-1 3 8
（以下略）
この様な都市には優先権を与えるようにした．これにより性能が改善したといわれている．辺再組合せ交叉法をみると，パス表現法が経路の重要な性質の表現力に乏しく，これを辺リストで補っていることを示している．

1.7　辺に着目したヒュリスティック操作

これまで論じて来た方式では都市（その位置と順序）を考えてきたが重要なのは経路の中の都市の位置ではなくむしろ都市間の接続関係である．従って巡回問題のビルディングブロックとして都市の位置表現ではなく辺（Edge）表現を用いることが考えられる．Grefenstette は辺に着目したヒューリスティック操作を開発した．この操作はつぎのように動作する．

1. 子の都市として任意の都市１つを選ぶ．

2. この都市に接続する辺を４つ（各親より２つずつ）選ぶ.
3. 上記各辺のコストに基づき確率分布を定義する．以前訪問した都市に関連した辺には零を割り振る.
4. １つの辺を選択する．少なくとも１つの辺が非零確率なら上記分布に基づき，そうでない場合は未訪問都市よりランダムに選ぶ.
5. 選定した辺の他端の都市がつぎの都市になる.
6. 経路が完成なら停止．そうでない場合はステップ２に戻る.

報告によるとこのような操作は親の辺の 60%を伝えている．即ち残り 40% は任意に選ばれていることになる.

1.8　突然変異

突然変異の方式は順序問題では一般の方式と異なった方式が必要となる．すなわち順序を突然変異で変更することになる．その方式としては以下のものが用いられている．いずれの方式がよいかは一概に言えないので，シミュレーションの結果などを参考に決めることになる.

- 交換　ランダムに選んだ２つの都市を互いに入れ換える方式である.
- 挿入　ランダムに選んだ１つの都市を取り出してその位置を詰め，同じくランダムに選んだ１つの都市の所に挿入し，それ以降をずらす方式である.

2　隣接表現 (Adjacent Representation)

隣接表現 は巡回経路をリストで表し，i 都市から j 都市へ進む場合にリストの i 番目に j として表す方法である．例えば

(2 4 8 3 9 7 1 5 6)

は次の巡回路を表している.

$1-2-4-3-8-5-9-6-7$

この方法の場合，正しくない経路を表す場合がある．例えば

(2 4 8 1 9 3 5 7 6)

は次の様にループを作ってしまう.

$1-2-4-1$

また通常の交叉演算子がそのままの形では使用できないので，次の３種の修正アルゴリズムが考えられている.

- 交互辺交叉 (Alternating-edge crossover)：先ず第１の親からランダムに１つの辺を選び，つぎに第２の親からこれに続く辺を選び，再び第１の親からこれに続く辺を選ぶ，というように両方の親から交互に辺を選んで交叉を行う．もし途中でループを作る場合は残りの辺から（ランダムに）１つを選んだ後，再び上記の方法を続ける．例えば

  $p_1 = (2\ 3\ 8\ 7\ 9\ 1\ 4\ 5\ 6)$

  $p_2 = (7\ 5\ 1\ 6\ 9\ 2\ 8\ 4\ 3)$

  に対して，つぎの様に選ぶ.

  $c_1 = (2\ 5\ 8\ 7\ 9\ 1\ 6\ 4\ 3)$

  この場合，辺 (7 8) がループを作ったので，代わりに (7 6) を選んだのが唯一のランダム選択である.

- 部分経路交叉 (Subtour-chunks crossover)：最初に第 1 の親からランダムな長さの部分経路を選び，つぎに第 2 の親からこれに続くランダムな長さの部分経路を選ぶという方法で交叉を行う．もしある辺がループを作る場合には残り部分より新しくランダムに辺を選ぶ．

- ヒューリスティック交叉 (Heuristic crossover)：先ず開始都市を任意に選ぶ．つぎに両方の親のこの同じ都市を出発する辺のうち短い方を選ぶ．この辺の到達都市をつぎの出発都市として上記を繰り返す．もし途中でループを生じる場合は残りの都市よりランダムに選ぶ．

  このヒューリスティック交叉の効果は親の経路の短い部分経路をつなぎ合わせることにある．しかし望ましくない辺の交わりを作ることもあるのでローカル経路の最終調整の段階で使用することは望ましくない．

隣接表示の特徴はセマータ解析が可能なことである．例えばスキーマ

$(* \ * \ * 3 \ * \ 7 \ * \ * *)$

は辺 (4 3) 及び (6 7) を持つ全ての経路を表している．しかし，この方法の欠点は全ての操作に対して比較的貧しい結果となっていることである．交互辺交叉ではよい経路を壊すことがしばしば起こる．部分辺交叉では交互辺交叉に比べて壊す割合は低いが性能はまだよくない．ヒューリスティック交叉が最もよい性能を示しているといわれている．

3 順序表現 (Ordinal Repesentation)

順序表現 は n 都市のリストという形で表す，即ちリストの i 番目の要素は i から $n-i+1$ の範囲の数字のひとつである．この表し方を具体的な例で示す．先ず基準となる順序リストを定める．これがつぎの通りであるとする．

$C = (1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9)$

これを用いて，巡回経路

$1-2-4-3-8-5-9-6-7$

をつぎのリスト ℓ で表す．

$\ell = (1\ 1\ 2\ 1\ 4\ 1\ 3\ 1\ 1)$

このリストは次の通り解釈する．

- リスト ℓ の最初の数値は 1 であるのでリスト C より 1 番目の 1（即ち都市 1）を採用し，リスト C より 1 を除く．

- リスト ℓ の次の数値もまた 1 であるので，リスト C の 1 番目の 2（即ち都市 2）を採用し経路はつぎのとおりになる．同時に 2 をリスト C より除く．

  $1-2$

- リスト ℓ の次の数値は 2 であるので，リスト C の 2 番目の 4 を採用して経路はつぎのとおりになる．同時に 4 をリスト C より除く．

  $1-2-4$

- リスト ℓ の次の数値は 1 であるので，リスト C の 1 番目の要素 3 を採用して経路は次のとおりになる． この 3 をリスト C より除く．

  $1-2-4-3$

以下これを繰り返す.

この方法の特長は通常の交叉が使えることである．例えば二つの親がつぎのとおりであったとする：

$$P_1 = (1\ 1\ 2\ 1 \mid 4\ 1\ 3\ 1\ 1)$$

$$P_2 = (5\ 1\ 5\ 5 \mid 5\ 3\ 3\ 2\ 1)$$

これはそれぞれ次の経路を表している.

$$1 - 2 - 4 - 3 - 8 - 5 - 9 - 6 - 7$$

$$5 - 1 - 7 - 8 - 9 - 4 - 6 - 3 - 2$$

縦棒 | で示した点で交叉すると次のリストが得られる.

$$c_1 = (1\ 1\ 2\ 1 \mid 5\ 3\ 3\ 2\ 1)$$

$$c_2 = (5\ 1\ 5\ 5 \mid 4\ 1\ 3\ 1\ 1)$$

これはつぎの経路を表している.

$$1 - 2 - 4 - 3 - 9 - 7 - 8 - 6 - 5$$

$$5 - 1 - 7 - 8 - 6 - 2 - 9 - 3 - 4$$

交叉点の左側の経路は親のものが保存されているが，右側の経路は壊されており，このように一部しか継承されてない．この方式も結果はよくないといわれている.

参考文献

[1] Michalewicz,Z : Genetic Algorithm + Data Structures = Evolution Programs, Springer-Verlag, 1992.

[2] 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,Vol.7,No.6,1049/1059,1992.