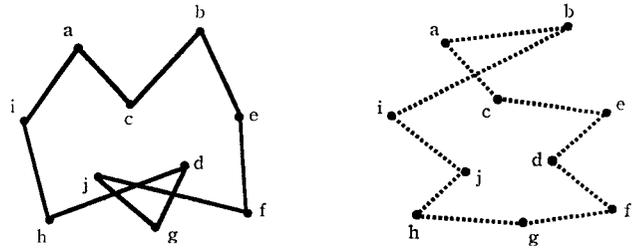


遺伝的アルゴリズムの基礎と応用〔Ⅱ〕

小林 重信

前回は、コード化／交叉の設計、適応度の設計を適切に行ない、さらに、実装パラメータの調整を適切に行なえれば、大域的探索と局所的探索のバランスがとれた探索 (well-balanced search) が実現されることを論じた。

今回は、GAの組合せ最適化問題の代表例として、巡回セールスマン問題および看護婦スケジューリング問題を取りあげ、バランスのとれた探索を実現するための設計上および実装上の創意工夫を紹介する。



親A (a, c, b, e, f, j, g, d, h, i) 親B (a, b, i, j, h, g, f, d, e, c)

図 1 10都市問題の解候補とパス表現

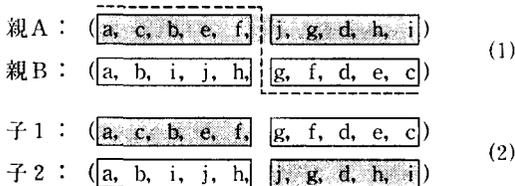
5. 巡回セールスマン問題への適用

都市間の巡回コストが平面上のユークリッド距離によって与えられる巡回セールスマン問題 (Traveling Salesman Problem; TSP) を考える。

解候補となる巡回路をツアーと呼び、ツアーに即した巡回コストの総和である目的関数を適応度関数として設定する。

5.1 従来方法

都市名を遺伝子として、起点とする都市から巡回する順番に都市名を列挙した文字列を染色体とする表現をパス表現 (path representation) という。図 1 に示す 10 都市問題の解候補である 2 つの親に対して、5 番目と 6 番目の遺伝子座の間を切断箇所として 1 点交叉を適用すると、次のような子が生成される。



子 1 と子 2 は “すべての都市を 1 度ずつ訪問する” という制約を満たしていないので TSP の実行可能解ではない。このような個体は致死遺伝子をもつ個体として最悪の適応度が割り当てられ、次の世代で淘汰されることになる。

この例に示すように、パス表現と 1 点交叉の組合せによって生成される個体の多くは致死遺伝子をもつ可能性が非常に高い。すなわち、この方法はコード化の評価規範の健全性を満たしていない。

Grefenstette は、致死遺伝子の生成を抑制するための工夫として、順序表現 (ordinal representation) [1] を提案している。順序表現では、アルファベット順にソートされた都市リスト (a, b, c, d, e, f, g, h, i, j) を予めもち、次に巡回する都市が残りの都市リスト (未訪問都市リストと呼ぶ) の中で何番目に相当するかを調べ、その番号を遺伝子とし、起点とする都市から順に並べた文字列を染色体とする。

たとえば、
 パス表現 : (a, c, b, e, f, j, g, d, h, i) (3)
 順序表現 : (1, 2, 1, 2, 2, 5, 2, 1, 1, 1)

図 2 は順序表現された親 A と親 B から、1 点交叉により、実行可能な解が生成される様子を示す。この方法によって生成される解候補は必ず実行可能である。図 2 において、太い実線と太い破線で示されるサブツアーは、それぞれ図 1 の親 A と親 B から継承したものになってい

こばやし しげのぶ

東京工業大学 総合理工 知能科学専攻

〒227 横浜市緑区長津田 4259

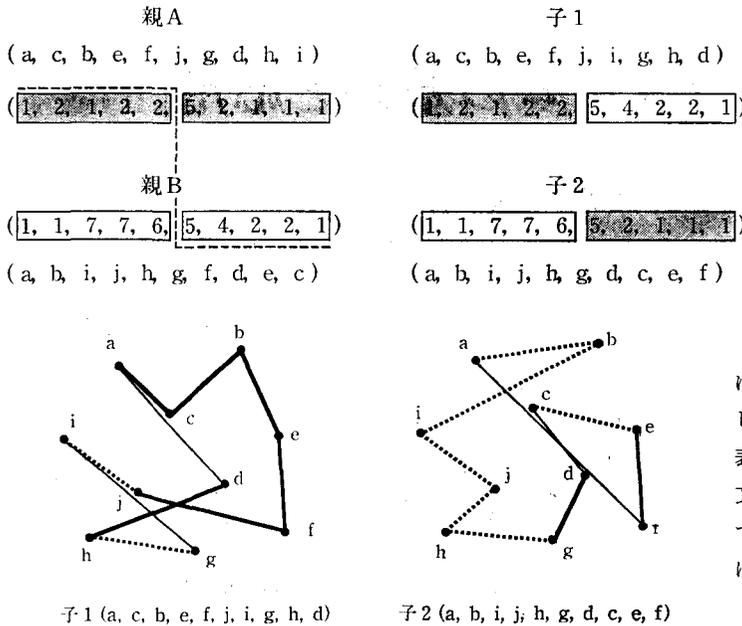


図 2 順序表現と1点交叉

る。すなわち、子1と子2は、それぞれ親Aと親Bから、交叉箇所の前半部分を継承しているものの、後半部分については子1は親Bの断片を、子2は親Aの断片を継承しているに過ぎない。

本方法はコード化の評価規範をすべて満たしているが、交叉の評価規範である形質遺伝性を部分的にしか満たしていない。本方法では、交叉箇所が染色体の前の方に設定された場合、両親の形質の大部分が失われてしまう危険性を内包している。Goldbergによる部分写像交叉 (partially mapped crossover)[2]も同じ短所をもつ。

実際、本方法にもとづくGAは、突然変異からのみなるランダムサーチと同程度の性能しか示さないことが知られている。さらに、ユークリッド距離を巡回コストとするTSPでは、交差路をもつ経路を本方法では解消できないことが知られている。

5.2 サブツアー交換交叉

筆者らはサブツアーをなるべく破壊せずに遺伝させる交叉方法としてサブツアー交換交叉[3]を提案している。

パス表現をもちいて、都市の訪問順序を自然に表現する。巡回コストとしてユークリッド距離を用いるとき、巡回コストは巡回の向きに対して対称であるので、同一のツアーに2通りの表現が可能である。図2の親Aについて、次のような右回りの表現と左回り表現が可能であ

る。
右回り表現：(a, c, b, e, f, j, g, d, h, i)

(4)

左回り表現：(a, i, h, d, g, j, f, e, b, c)

ここで、左右の区別は便宜上のものである。

このコード化は、完備性かつ健全性の評価規範を満たしている。さらに、コード化は非冗長であることが望ましいので、片方を選んで両方の表現を代表させるものとする。ただし、これらは文字列としては異なるので、交叉を適用する際には両方の組合せを考慮しなければならない。

たとえば、通常の1点交叉や本方法で用いる2点交叉では、2つの親、A、Bのあいだに次の2通りの文字列の組合せ

を同時に考える。

親A (右回り) と親B (右回り) での交叉 (5)

親A (右回り) と親B (左回り) での交叉

なお、このとき次の組合せ(6)によって生成される子は、(5)によって生成されるものと同じである。

親A (左回り) と親B (右回り) での交叉

(6)

親A (左回り) と親B (左回り) での交叉

TSPにおいて子孫に遺伝すべき形質はサブツアーである。染色体上の2点で挟まれた範囲 (サブツアー) を交換する2点交叉 (two points crossover) を考える。しかし、1点交叉の場合と同様に、普通の2点交叉では、致死遺伝子をもつ子を生成し得る。コード化は健全であることが望ましいので、交換されるサブツアーに含まれる都市集合が一致するときに限って、交叉させることとする。そのために、親Aと親Bでは、交叉のための切断箇所は異なることを許容する。この方法に従う交叉を、通常の2点交叉と区別するために、サブツアー交換交叉 (subtour exchange crossover) [3]と呼ぶ。

図3にサブツアー交換交叉による個体の生成の様子を示す。親Aのサブツアー (f, j, g, d, h) と親Bのサブツアー (j, h, g, f, d) は、都市の訪問順序は異なるが、都市集合が一致しており、交叉の条件を満たしている。パス表現のサブツアーの下に書かれた有向アークは右回りと左回りの区別をしており、交換の対象となるサブツア

のみ右回りと左回りを考えればよく、この交叉により、図4に示すように、4通りの子が生成される。これらの子はいずれも、TSPの制約充足解であり、親の形質すなわちサブツアーをできる限り継承していることに注意されたい。この例からも明らかなように、サブツアー交換交叉は形質遺伝性において優れている。

5.3 サブツアー交換交叉にもとづく解法

サブツアー交換交叉は、都市集合が一致するときのみ交叉が許されることから、条件を満たすサブツアーを見いだす探索に負荷がかかることが短所である。したがって、大規模問題への適用に際しては、探索コストを軽減するための工夫が必要である。

【染色体の縮約】

2つの解候補のあいだに、都市集合だけでなく訪問順序も一致するサブツアーが存在するとき、これらにサブツアー交換交叉を適用することは無意味である。そこで、次のような染色体の縮約を考える。

- 1) 2つの親に共通するサブツアーを見いだす。
- 2) 共通サブツアーを1つの仮想都市とみなし、染色体を縮約する。
- 3) 縮約された染色体に対してサブツアー交換交叉を適用して子を生成する。
- 4) 生成された子を逆変換してもとの染色体表現に戻す。

例として、次の2つの親を考える。

親A: (a, j, h, i, f, [d, b, c], [e, g], a)
 親B: (a, [e, g], f, i, h, [d, b, c], j, a) ⁽⁷⁾

ここで、[d, b, c]と[e, g]は親Aと親Bの間で共通するサブツアーであり、縮約の対象とされる。世代交替の進行とともに、個体間で共通するサブツアーが蓄積されてくるのが予想されるので、サブツアーの縮約は全体の計算量を節約する効果が期待できる。

【世代交代モデル】

突然変異は、それ自身強力な探索手段であるが、コー

パス表現

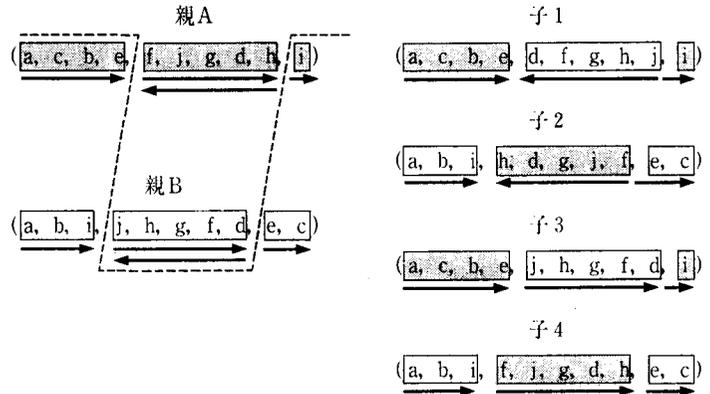


図3 パス表現とサブツアー交換交叉

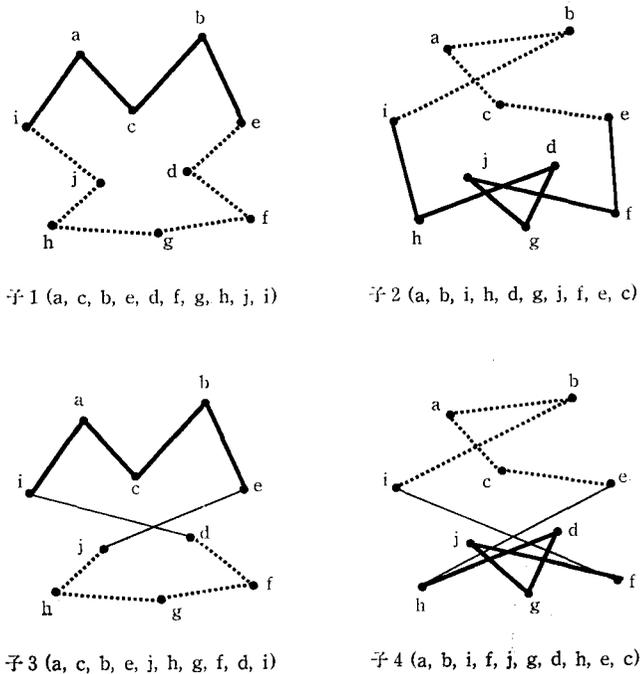


図4 サブツアー交換交叉の適用結果

ド化と交叉が適切に設計されていれば必ずしも必要なものではないとの考えから、本モデルでは使わないことにする。また、通常は、集団内に同一の染色体をもつ個体の存在を許容しているが、初期収束の原因にもなることから、個体の重複は認めないことにする。

以上の考えにもとづいて設計された解法は以下の手順からなる。

【アルゴリズム】

- (1) ランダムな染色体をもつM個の個体からなる集団

をつくり、第0世代とする。

(2) 第N世代の集団より、可能な親のペアをすべてつくり、各ペアに対して、交叉確率に従い交叉の可否を決定する。

交叉させる場合のみ以下の処理をほどこす。

①染色体を縮約する。

②交叉回数だけサブツアー交換交叉を適用する。

(3) 生成された子の個体集団について、重複する個体を削除したあと、適応度の分布にもとづいて淘汰を行ない、集団サイズMの第N+1世代とする。

(4) (2)~(3)を打切り世代数までくりかえす。

5.4 解法の評価

提案したアルゴリズムの有効性を確認するために、次のような数値実験を行なった。1×1の方形の範囲内に都市をランダムに配置するものとし、50, 100, 200, 300, 500個の都市について、乱数の初期値を変えて、それぞれ100, 100, 35, 21, 1回の試行を行ない、オンライン性能を測定した。集団サイズを10, 交叉確率を0.4, 交叉回数を10に設定した。これは、可能な親のペアが

${}_{10}C_2=45$ 通り、1世代で生成される子の数の平均値が $45 \times 0.4 \times 40 = 720$ 個であることを意味する。

なお、打切り世代数をそれぞれ 500, 1,000, 2,000, 3,000, 10,000世代とした。

図5に本実験のオンライン性能を示す。横軸は世代、縦軸はツアー距離を表わしている。ここで、ツアー距離は、各都市数ごとに、初期世代における最良ツアーを試行に関して平均した値を100ポイント、全世代を通じての最良ツアーの平均値を0ポイントとなるように標準化し

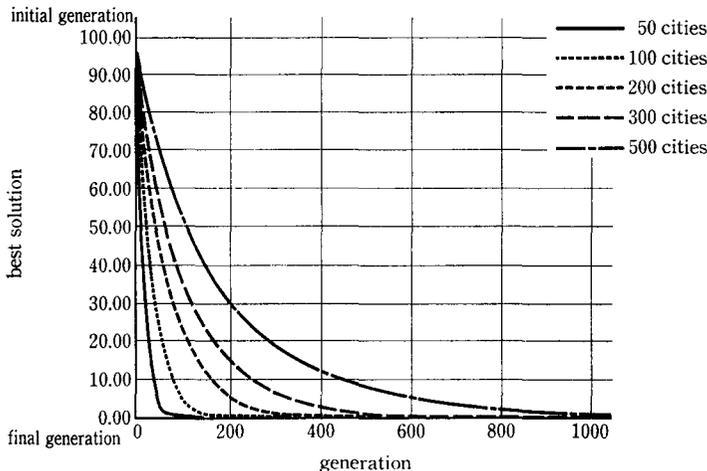


図5 本解法のオンライン性能

表1 最良値および交差路の消失世代

都市数	初期世代の最良値	全世代の最良値	交差路の消失世代
50	23.0	5.86	49
100	48.3	8.36	137
200	99.5	11.9	342
300	149.3	14.2	579
500	252.4	17.7	1210

である。各曲線は、各都市数に対して、各世代における最良値をプロットしたものであり、各プロットの分散は高々数ポイントである。表1は各都市数に対する初期世代の最良値、全世代を通じての最良値、交差路の消失世代の各平均値を示している。

たとえば、文献[1]では、順序表現よりも改良された方法を用いて、50, 100, 200都市の同様の問題について実験を行なっている。そこでは、集団サイズ50とした試行で、それぞれ約300, 400, 5000世代かけても交差する経路を排除できないことが報告されている。

表1からも明らかのように、典型的な問題設定においては、形質を保存するGAの効果は明らかに優れている。図6に、500都市の問題に対する初期世代における最良ツアーと全世代における最良ツアーを示す。最良ツアーは8,084世代以降に現われている。左肩の数値は巡回距離を表わしている。

5.5 考察

GAのTSPへの適用から得られた知見を以下に要約する。TSPはGAにおけるコード化/交叉問題を考える

上で格好の題材であるといえる。致死遺伝子の抑制を重視した従来のコード化/交叉[1], [2]は、形質遺伝性という評価規範への配慮が十分でないことから、最適化という本来の目的を達成することができず、ランダムサーチ程度の性能にとどまっている。

TSPにおいて優れた形質とは部分的に成功しているサブツアーである。それらをビルディングブロックとして蓄積し、交叉により組合せていく過程は、GAに期待される創発的振舞い(emergent behavior)によく適合している。サブツアー交換交叉はそのような過程を自然モデル化するものである。

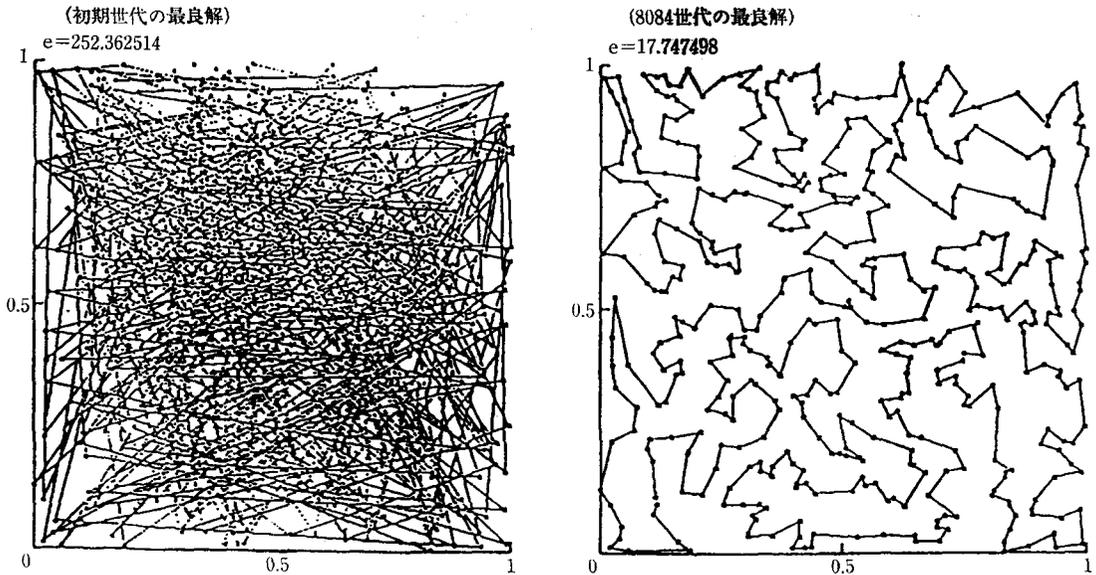


図 6 本解法による500都市問題の最良解

サブツアー交換交叉にもとづく解法では、突然変異を排除して、集団サイズを10と小さく設定したにもかかわらず、500都市問題でも初期収束の現象を回避することができた。これは、集団内に染色体の重複を認めなかったこと、および1世代当たり、平均して、18対の親から合計720個の子を生成し、確率的な淘汰を通じて10個体を次世代に残すことにより、集団の多様性が十分維持されたことによるものと考えられる。

サブツアー交換交叉は、2つの個体の間で交換可能なサブツアーを見いだすために $O(n^2)$ の計算を必要とする。しかし、2つの個体のあいだで共通するサブツアーを仮想都市として縮約することにより、計算コストの軽減が期待できる。実際の計算に要する時間の半分近くは最初の数世代に費やしており、ビルディングブロックの形成に伴い、染色体の縮約効果が劇的に働いてくることが観察された。

6. 看護婦スケジューリング問題への適用

スケジューリング問題の一例として看護婦スケジューリング問題(Nurse Scheduling Problem; NSP)がある。ORの分野でも、20年以上前から、ゴールプログラミングを用いた接近が試みられてきたが、制約条件が多数あり、評価規範が多目的であることから、適切な解法

の構築が難しいとされていた問題である。ここでは、NSPへのGA的接近法[4]を紹介する。

6.1 NSPとは

NSPとは、病院の1病棟を担当する看護婦(15~20人)の1月分の勤務スケジュールを決定する問題である。図7に、NSPの制約条件と評価規範の主要なものを示す。看護婦の一般的な勤務形態は日勤(8:00~16:00)、準夜勤(16:00~24:00)、深夜勤(24:00~8:00)の3交代制を基本としている。毎日、所定の数の日勤者、夜勤者(準夜勤および深夜勤)を確保することは絶対的な制約条件とされる。希望休暇を充足すること、夜勤の前後は指定の公休日とすること、長期間の連続夜勤勤務は回

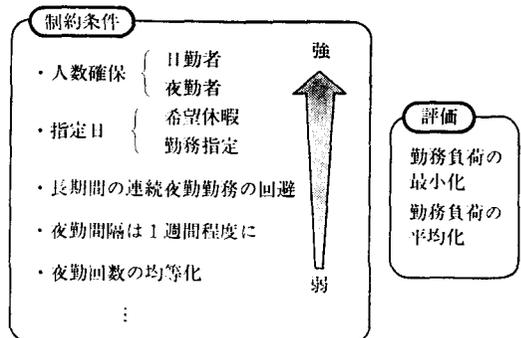


図 7 NSPの制約条件と評価規範

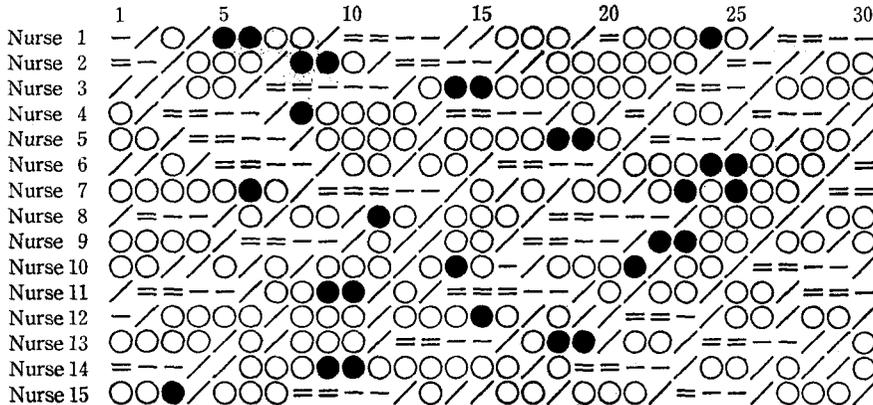


図 8 勤務スケジュールの実例

避すること、夜勤の間隔は一定日数以上空けること、夜勤回数は均等化することなどが主たる制約条件である。その他に、特別な勤務指定の遵守、新人同士や相性の悪い者同士での夜勤勤務は回避すること、などの制約がある。

目的関数は、各看護婦の勤務負荷を最小化すること、および勤務負荷を看護婦のあいだで均等化すること、の2つが設定されるのが普通である。

図 8 は実際の勤務スケジュールの一例を示している。婦長は、毎月、月末になると、3晩ほどを費やして、翌月の勤務スケジュールを作成しているが、制約や評価が多様なことから、最適な勤務スケジュールを作成することは難しく、場当たり的なスケジュールで妥協せざるを得ないのが現状といわれている。

6.2 GAによる解法の設計

以下の方針に従って、GAによるモデル化を行なうこととした。

- 1)各看護婦をそれぞれ1個体に対応づける。
- 2)看護婦の総数を集団サイズとする。
- 3)1つの染色体は各看護婦の1月分の勤務スケジュールを表わすものとする。
- 4)各看護婦の勤務スケジュールの評価値を各個体の適応度とする。
- 5)個体の適応度の分散を集団の適応度とする。
- 6)交叉を探索の主オペレータとし、突然変異は行なわない。

以上の方針にもとづいて、図 9 に示す枠組みに従って、次のようなGAの解法を構成した。

【アルゴリズム】

- (1) 初期世代

縦制約および希望休暇だけを充足する勤務スケジュールを作成し、これを初期世代とする。

- (2) 親のペアの選択

各個体の適応度に比例した確率にもとづき、ルーレット選択によって、交叉対象となるペアを1つ選択する。

- (3) 交叉

選択により選ばれたペアを親として、2点交叉により、交叉個所を変えることにより、可能な子のペアをすべて生成する。

- (4) 子のペアの選択

(3)によって複数生成された子のペア群に対して、パレート最適集合をつくり、その中からランダムに1つの子のペアを選択する。

- (5) 世代交替

(4)で選択された子のペアを親のペアに置き換えて

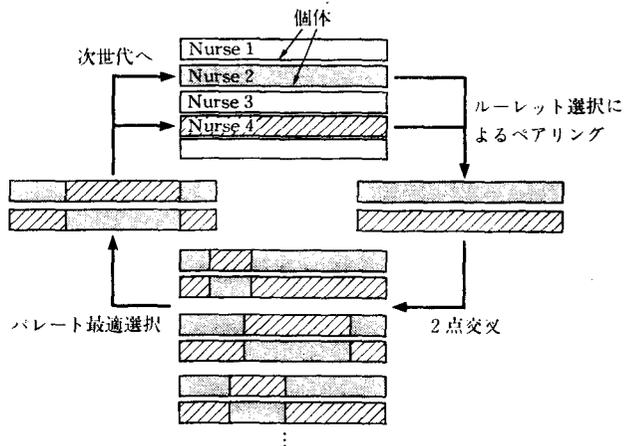


図 9 GAによるNSPの解法の図式

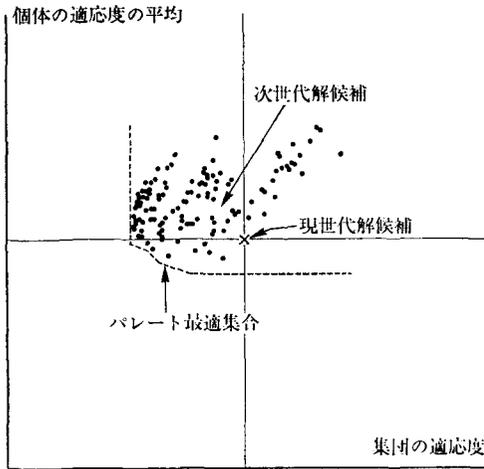


図10 世代交替モデル

次世代とし、(2)に戻る。

なお、(1)における縦制約とは、毎日、所定の数の日勤者および夜勤者を確保しなければならないことをいう。(3)において、希望休暇など各個体に固有の勤務指定日は交叉の対象外とする、(4)におけるパレート最適集合の意味を図10によって説明する。

図10の縦軸および横軸は、それぞれ個体の適応度の平均（各看護婦の勤務スケジュールの評価値の平均）および集団の適応度（個体の適応度の分散）を表わす。図の中央に位置する×印は交叉を行なう前の集団の評価値をプロットしたものであり、一方、小さな重点で示される印は交叉により生成された子のペアを親のペアに置き換えた場合の集団の評価値を表わしている。黒点群の中で、左下方に別の点が存在しない点の集合はパレート最適集

合を形成する。図10において破線上に位置する点がパレート最適な候補である。

本解法の世代交替は、2点交叉によって生成される子のペアのうちパレート最適になっているものだけを親のペアと置き換える候補としようとするものである。代替案として、個体の適応度の平均と集団の適応度のトレードオフ比を設定し、最適な候補を唯一決定する方法も考えられるが、トレードオフ比の決定という新たな問題を生じること、および予備的な実験では局所最適解に陥りやすいことが観察されている。

6.3 解法の評価

提案したモデルの有効性を確認するために、数値実験を行なった。図11に示す15人の看護婦の勤務スケジュールを初期世代として設定した。“==—”を標準夜勤パターン、“==—”、“—”、“—”を許容夜勤パターンとし、各パターンにペナルティを課した。夜勤間隔は11日を基準として、この間隔からのズレに対しペナルティを課した。

図12の実線は本モデルによる収束の様子を示したものである。図の横軸および縦軸は、それぞれ集団の適応度および個体の適応度の平均を表わしている。左下隅の点が最適点である。収束の途中、1,000~4,000世代にかけて、試行錯誤的に探索が行なわれている様子が観察される。

この図の左上にある破線は、パレート最適集合の中より、集団の適応度および個体の適応度の両方が改善されるものだけを次世代の候補とした場合を示しており、集団の適応度の改善は行なえたものの、局所最適解にすぐに陥ってしまうことが観察された。

図13は、本モデルによって9,644世代後に得られた解

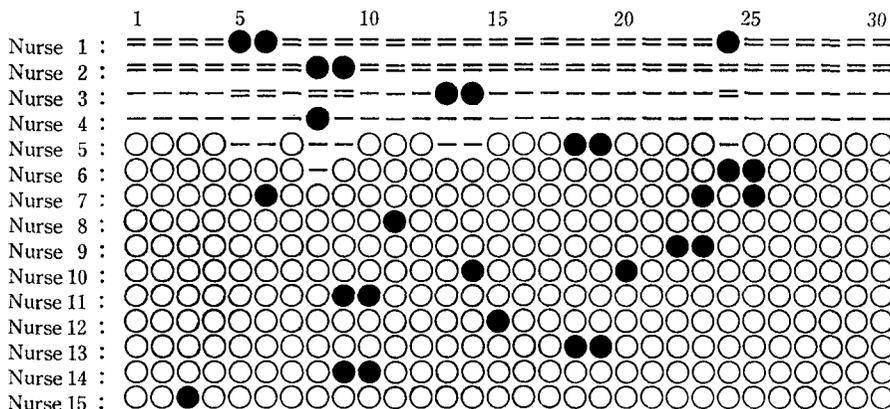


図11 初期世代のスケジュール表

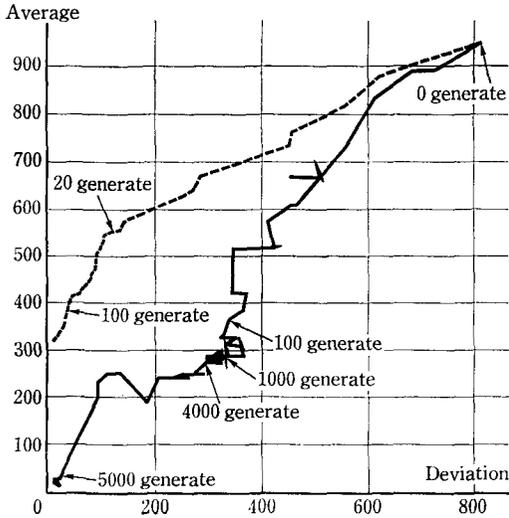


図12 NSPの収束過程

スケジュールである。図8と比較して、良好なスケジュールとなっていることが理解されよう。個体の評価値のバラツキも許容範囲内と判断される。

6.4 考察

NSPは複雑な制約をもつため、全看護婦を対象としたスケジュール表を1つの個体とすると、致死遺伝子を生じないコード化/交叉の設計はきわめて困難である。本解法で示したように、看護婦ひとり分のスケジュールを1個体とし、全スケジュールを集団とするやや変則的なGAを用いることによって、致死遺伝子を抑制することができる。

NSPは多目的問題であり、無理やりスカラー化せずに、各個体の適応度の最小化と集団の適応度の最小化を同時に考慮したパレート最適集合の中から次世代の解候補を選択する方法は局所解からの脱却に有効である。

実際のNSPでは、特別な勤務指定の遵守、新人同士や相性の悪い者同士での夜勤勤務の回避など多数の制約を考慮に入れる必要があるが、ここで紹介した解法の基本的な枠組みを変更する必要はなく、適応度関数の設計および交叉箇所の限定などによって対処することが可能である。

今回は、TSPおよびNSPの2つをとりあげ、GAにもとづくバランスの取れた探索を実現するための設計上および実装上の創意工夫を紹介した。

TSPについては、サブツアー交換交叉および染色体の縮約という考え方が有効であることを示した。サブツアー交換交叉のアイデアはジョブジョブスケジューリング問題へも適用可能[5]であり、次回に紹介する。

NSPについては、看護婦ひとり分のスケジュールを1個体として扱い、各個体の適応度の最小化と集団の適応度の最小化を同時に達成するために、パレート最適の考えにもとづき、バランスのとれた探索を行なうことが有効であることを示した。本解法は、一般性を有することから、他の勤務スケジューリング問題に対しても適用可能である。

GAの組合せ最適化への本格的な応用は始まったばかりであり、現在は、試行錯誤にもとづく経験を蓄積している段階にすぎないが、染色体によって特徴づけられた解候補を集団で維持し、それらの交叉によって新たな解候補を生成するアイデアは組合せ最適化への探索戦略としても斬新であり、今後の発展が期待される[6]。

次回、はジョブジョブスケジューリング問題への適用およびGAの挙動に関する理論を紹介する。

参考文献

[1] Grefenstette, J., Gopal, R., Rosmaita, B.

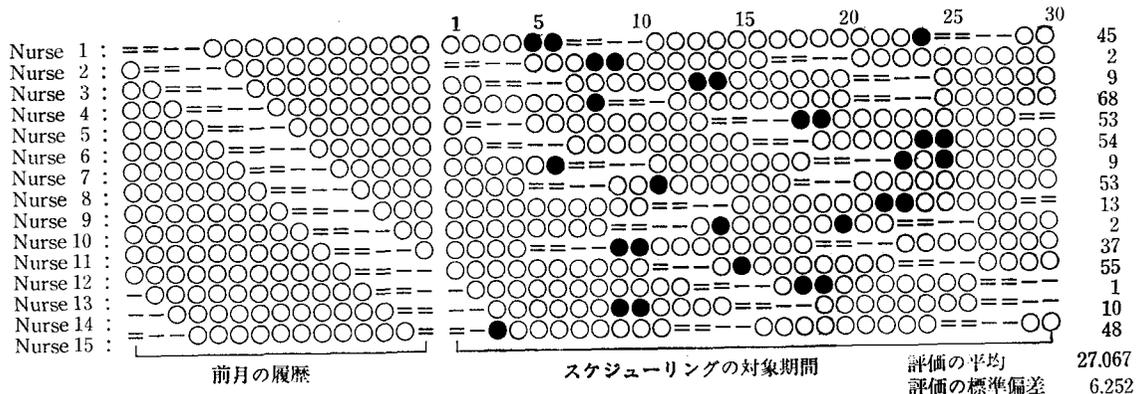


図13 最終世代のスケジュール表

- J. and Van Gucht, D.: Genetic Algorithms for the Traveling Salesman Problem, Proc. of ICGA '85, pp.16—168 (1985).
- [2] Goldberg, D. and Lingle, R.: Alleles, loci, and the travelling salesman problem, Proc. of ICGA—85 (1985).
- [3] 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もと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, pp.1049—1059(1992).
- [4] 太田, 山村, 小野, 小林: 遺伝的アルゴリズムを用いた Nurse Scheduling Problem の解法, 第16回知能システムシンポジウム講演資料集, pp.123—128 (1992).
- [5] 太田, 山村, 小野, 小林: 遺伝的アルゴリズムを用いた Job-Shop Scheduling Problemの解法, 第17回知能システムシンポジウム講演資料集, pp.27—32 (1993).
- [6] 山村, 小林: 遺伝的アルゴリズムによる組合せ最適化, シミュレーション, Vol.12, No.1, pp.4—10 (1993).

追悼

北川敏男先生を偲んで

東洋大学 北原 貞輔

本年3月13日、北川敏男先生には、83歳でご逝去されました。謹んで哀悼の意をささげます。

先生は、昭和9年、東京大学理学部数学科をご卒業後、大阪大学を経て九州大学理学部に助教授として着任され、昭和18年に教授に昇進、統計数学講座、計画数学講座をご担当になられました。この間、推測統計学を開拓、さらに推測過程論や管理過程論を展開され、多くの研究成果を挙げられましたことは、特に記すまでもないことと思います。

昭和25年から6期18年間にわたって、先生は、日本学術会議会員として日本の科学研究の発展に指導的立場で貢献されました。昭和28年には、品質管理に関してデミング賞を受賞され、31年には、国際統計協会正会員に推挙されました。その間、インドの経済開発計画に参画、32年に、歴史学のトインビー博士や物理学のオープンハイマー博士とともにカルカッタ大学の名誉博士号を受けておられます。また、日本OR学会フェロー、米国数理統計学会フェロー、情報処理学会会長などを勤められ、科学の発展に多方面から貢献してこられました。

九州大学では、評議員や図書館長、理学部長、基礎情報学研究施設長などを歴任され、昭和48年に停年退

官、同時に富士通㈱国際情報社会科学研究所長に就任され、それを国際的評価の高い研究所に育成されました。

北川先生の卓越した先駆的な識見は、他の追従を許さないものがあったように思われます。昭和20年代中葉に、早くもウィーナーのサイバネティクスを日本に紹介され、現代科学者たちのかなり多くが、なお構造的立場に立っているのに対し、すでに20年代から過程論思想を展開されていたことは特質すべきことではないでしょうか。その延長線上で、初めて情報学という用語を使用され、制御・営存・創造の3座標軸を提起されたことは、あまりにも画期的なものでした。そこには死んだ世界の科学から生きた世界の科学への思想転換の要が示唆されているからです。

現在の自然環境破壊にも、先生は、いたく心配しておられ、昨年10月にお会いした折にも、多くの有用なご意見をいただき、地球の未来に対する高邁なご識見に驚くと同時に、生きた世界観をお持ちであったことを痛感させられました。人類存続の危機に直面している今日、偉大な指導者を失くしたことは痛恨のきわみであります。心からご冥福をお祈り申し上げます。