遺伝アルゴリズムを枠組としたメタ戦略の構築法[†]

—グラフ彩色問題を実例として—

田 川 聖 治*・金 重 憲 治**

井 上 克 巳*・羽根田 博 正*

A Design of Genetic Algorithm-Based Meta-Heuristic Method:
An Application to the Graph Coloring ProblemKiyoharu TAGAWA*, Kenji KANESIGE**,
Katsumi INOUE* and Hiromasa HANEDA*

In this paper, a meta-heuristic method that combines the global search power of Genetic Algorithms with the local search power of local optimization algorithms is described. First of all, a metric function between two solutions, or phenotypes, is defined by the shortest Hamming distance between sets of isomorphic genotypes. The phenotypic distance is useful to analyze and control the behavior of genotypes in the search space from the view point of the problem space. Then, by using the phenotypic distance, a new crossover technique named Harmonic Crossover is proposed in which children always come to the position between their parents in the problem space. Finally, in order to get maximum efficiency of the meta-heuristic method, the best timing in employing the local optimization algorithms is discussed. The experimental results indicate that the local optimization algorithms should be synchronized with the Harmonic Crossover to find high quality solutions in reasonable time.

Key Words: meta-heuristic, genetic algorithm, phenotypic distance, harmonic crossover

1. はじめに

工学的に有用な多くの問題は最適化問題に帰着され、それらのほとんどがNP困難となる。そこで、これらNP困難な問題に対して様々な解法が考案されてきた。しかし、一般的かつ強力な解法は得られておらず、「NP=P問題」が否定的に予想されていることから、今後、NP困難な問題に対する万能な解法が開発される可能性はきわめて低い¹⁾。

メタ戦略とは最適化問題を解くための幾つかの戦術(技法)を有機的に結合させたものであり、個々の問題に適した技法の編成が可能であることから、NP困難な問題の現実的な解法として注目されている。また、複数の異なる技法の相互作用によって、それぞれの性能を越えた探索能力が発揮されるという意味において、メタ戦略は最適化問題の創発的な解法となりうる。メタ戦略の枠組としては、焼きまし法(SA:

Simulated Annealing)、タブー探索(TS: Tabu Search)、遺伝アルゴリズム(GA: Genetic Algorithm)などがある²⁾。SAとTSは近傍の探索を基盤に構築され、従来の局所探索法の拡張として体系化されている³⁾。一方、GAは生物進化にアナロジーを持ち、個体集団に基づく多点探索や、交叉による跳躍の探索によって、近傍に束縛されない大域的な探索が可能である。反面、GAは良い解の近傍における系統的な探索が苦手であり、GAを枠組としたメタ戦略を構築する際には、適当な局所探索法を組込む必要がある。

メタ戦略における異なる技法の統合では、対象とする問題の解空間をどのように把握するかが肝要である。一般的なGAでは、問題の解候補(個体)を遺伝子型にコード化することで、解空間は探索空間に写像される⁴⁾。これまで、GAにおける探索空間の特徴付けは、スキーマ(schema)統計⁵⁾、エピスタシス(epistasis)分散⁶⁾、ウォルシュ(Walsh)関数⁷⁾、エントロピー(entropy)⁸⁾などにより行われてきた。また、GAの挙動の理論的な解析手法として、スキーマ解析、マルコフ(Markov)連鎖、遺伝演算子の相関係数などが提案されている⁴⁾。しかし、これらの解析手法は完成されたものではなく、GAの特性の一側面を述べているにすぎない。さらに、GAを枠組としたメタ戦略を構築する上で、探索空間を対象とするGA理論の最大の難点は、解空間での近傍構造

[†] 第8回FANシンポジウムで一部発表(1998・10)

* 神戸大学工学部 神戸市灘区六甲台町1-1

** 神戸大学大学院自然科学研究科 神戸市灘区六甲台町1-1

* Faculty of Engineering, Kobe University, Nada-ku, Kobe

** Graduate School of Science and Technology, Kobe University, Nada-ku, Kobe

(Received January 18, 1999)

(Revised July 27, 1999)

に基づく局所探索法との理論的な整合性がないことである。一方、隣接点で定義される近傍構造を、GAの特徴である交叉の議論に導入することにも無理がある。

本論文では、近傍構造を包括する距離空間の概念に基づき、GAと局所探索法を効果的に組合せたメタ戦略の構築法について考える。また、具体例としてグラフ彩色問題（GCP：Graph Coloring Problem）を対象とする。GCPは実用的な組合せ最適化問題であり、NP困難な問題の代表的なベンチマークの一つでもある。このため、GCPには幾つかの局所探索法が報告されている⁹⁾。はじめに、個体間に構造的な距離を導入することで、GCPの解空間とGAの探索空間をともに距離空間と見做し、両者の関係を明らかにする。これにより、遺伝子型の記号操作を本質とするGAの汎用性を失うことなく、個体集団の多様性や遺伝演算子の作用など、GAの挙動を解空間の観点から定量的に評価できる。次に、個体間の距離を利用した調和交叉と突然変異を提案し、調和交叉によれば両親の形質が子に継承されることを証明する。さらに、GAを枠組とするメタ戦略の構築法を示すとともに、その性能を計算機実験によって評価する。最後に、GAによる大域的探索に局所探索法を組込む適切なタイミングの検討を通して、創発的な探索のメカニズムについて考察する。

2. グラフ彩色問題の記述

グラフ彩色問題（GCP）は、次のように記述される。

〔問題1〕 グラフ $G=(V, E)$ が与えられる。頂点 $v_i \in V$ を任意の k 色で色分けすることで、頂点集合 V を k 個の部分集合 V_p ($p=1 \sim k$) に分割する。ただし、どの枝 $e \in E$ の両端も同一の部分集合に含まれてはならない。このとき、最少の彩色数 k とそれを与える V の分割方法を求めよ。

本論文では、上記のGCPをペナルティー関数法により解くものとする⁹⁾。すなわち、頂点集合 V の任意の分割方法 $s = \{V_p \mid 1 \leq p \leq k\}$ に対して、(1) 式の目的関数 $g(s)$ を最小化する。このとき、 s がGCPの実行不可能解であることも許して、 $E_p \subset E$ を両端が V_p に含まれる枝の集合とする。

$$g(s) = - \sum_{p=1}^k |V_p|^2 + \sum_{p=1}^k 2|V_p||E_p| \quad (1)$$

(1) 式の目的関数 $g(s)$ において、彩色数 k が少ない分割方法ほど第1項の値は小さくなる。また、(1) 式の第2項はGCPの実行不可能解に対するペナルティーであり、分割方法 s がGCPの実行可能解であるとき、 $|E_p| = 0$ ($p=1 \sim k$) から第2項の値は零となる。さらに、あとで述べるように、(1) 式の局所最適解はGCPの実行可能解となる。

3. GAを枠組としたメタ戦略

3.1 解空間と探索空間

3.1.1 個体の表現方法

最適化問題にGAを適用する場合、その候補は個体（表現型）に対応づけられる。本論文では、GCPの表現型を頂

Fig. 1 Phenotype of GCP

点集合 V の任意の分割方法 $s = \{V_p \mid 1 \leq p \leq k\}$ とする。また、表現型を適当な記号列にコード化したものが遺伝子型である。本論文では、 s の遺伝子型を頂点数 $n = |V|$ に等しい n 次元の配列 $A = (a_1, \dots, a_n)$ として、頂点 $v_i \in V_p$ であるとき要素 $a_i \in A$ の値を $a_i = p$ とする。このとき、GCPでは彩色数 k のみに着目するので、一つの表現型に対して $k!$ 個の遺伝子型が存在する。たとえば、Fig. 1 のように七つの頂点を $k=3$ で色分けした場合の表現型に対応する遺伝子型 A^x ($x=1 \sim k!$) を (2) 式に示す。

$$\begin{cases} A^1 = (1, 1, 2, 3, 3, 2, 2) \\ A^2 = (1, 1, 3, 2, 2, 3, 3) \\ A^3 = (2, 2, 1, 3, 3, 1, 1) \\ A^4 = (2, 2, 3, 1, 1, 3, 3) \\ A^5 = (3, 3, 1, 2, 2, 1, 1) \\ A^6 = (3, 3, 2, 1, 1, 2, 2) \end{cases} \quad (2)$$

GCPにおいて、一つの表現型に対応する $k!$ 個の遺伝子型 A^x ($x=1 \sim k!$) は、同値類 $[A]$ であると考えられる。表現型と $[A]$ は一対一に対応するので、以降、遺伝子型の同値類 $[A]$ を用いてGCPの表現型を表わす。

ところで、遺伝子型は計算機上に実装される個体のデータ構造を規定するとともに、GAの汎用性を主張するための重要な仕様である。上記の遺伝子型 A は、GCPに限らず、解の構造が集合の分割方法となる様々な最適化問題で採用されている¹⁰⁾。また、同じ同値類に属する遺伝子型の判定などは、リナンバリング (renumbering) 処理¹⁰⁾により遺伝子型を特定の標準形に変換することで容易に行える。

3.1.2 個体間の距離

遺伝子型と表現型に対して、それぞれ距離関数を定義する。これにより、表現型の全体集合である解空間と、遺伝子型の全体集合である探索空間をともに距離空間と見做す。

はじめに、二つの遺伝子型 $A = (a_1, \dots, a_n)$ と $B = (b_1, \dots, b_n)$ の距離を、(3) 式のように一般的なハミング (Hamming) 距離で定義する。

$$\delta_g(A, B) = \sum_{i=1}^n h(a_i, b_i) \quad (3)$$

ただし、 $h(a_i, b_i) = 0$ ($a_i = b_i$)、 $h(a_i, b_i) = 1$ ($a_i \neq b_i$)。

次に、表現型 $[A]$ と $[B]$ の距離を、対応する遺伝子型 $A \in [A]$ と $B \in [B]$ の最短距離で定義する。

$$\delta_p([A], [B]) = \min\{\delta_g(A, B) | A \in [A], B \in [B]\} \quad (4)$$

ここで、(3) 式と (4) 式から、表現型と遺伝子型の距離には、(5) 式のような関係が成り立つ。

$$\delta_p([A], [B]) \leq \delta_g(A, B) \leq n \quad (5)$$

ただし、 $A \in [A]$ 、 $B \in [B]$ である。

《定理1》 表現型の距離は「距離の公理」を満たす。

公理0: $[A] = [B] \Leftrightarrow \delta_p([A], [B]) = 0$

公理1: $\delta_p([A], [B]) \geq 0$

公理2: $\delta_p([A], [B]) = \delta_p([B], [A])$

公理3: $\delta_p([A], [B]) + \delta_p([B], [C]) \geq \delta_p([A], [C])$

(証明) 遺伝子型の距離 (ハミング距離) が距離の公理を満たすことから容易に導ける¹¹⁾。□

ところで、(4) 式で定義した表現型の距離は、(6) 式に示すように、より簡潔に定義することも可能である。

$$\delta_p([A], [B]) = \min\{\delta_g(A, B) | B \in [B]\} \quad (6)$$

ただし、 $A \in [A]$ は任意の遺伝子型である。さらに、表現型 $[A]$ の彩色数は $[B]$ のそれよりも多くはなく、同値な遺伝子型の総数は $|[A]| \leq |[B]|$ であるとする。

表現型の距離を (6) 式の定義どおりに求めると、遺伝子型の距離を $O(n)$ で $k!$ 回 ($k = |[B]|$) 評価する必要があるため、時間計算量は $O(nk!)$ となる。しかし、彩色数 k が多くなると、 $O(nk!)$ は実用的な計算量とは言えない。

《定理2》 表現型の距離は $O(n + k^3)$ で計算できる。

(証明) 表現型の距離を求める問題を、最大割当て問題¹²⁾に変換して解くものとする。付録Aに示すように、二つの問題は等価であり、問題の変換は $O(n)$ で行える。また、最大割当て問題に対しては、 $O(k^3)$ のアルゴリズムが知られている¹²⁾。したがって、上記の方法で表現型の距離を求めれば、全体の時間計算量は $O(n + k^3)$ となる。□

3.2 調和交叉

GCP の遺伝子型に適用可能な交叉法として、一点交叉 (1X)、二点交叉 (2X)、一様交叉 (UX) などが挙げられる¹⁰⁾。ここでは、GCP の表現型の距離に基づき、従来の交叉における形質遺伝性¹³⁾を改善する手法として、新たに調和交叉を提案する。ただし、従来の交叉法も含めて、二つの親 A 、 B から一つの子 C が生成されるものとする。

【調和交叉】

手順1: 両親 A 、 B の距離 $\delta_g(A, B)$ が最短となるように、 B を同値な遺伝子型 $B^z \in [B]$ に変換する。

このとき、 $\delta_g(A, B^z) = \delta_p([A], [B])$ である。

手順2: 従来の交叉 (1X, 2X, UX など) を両親 A 、 B^z に適用して、子 C を生成する。■

たとえば、ランダムに生成したビットマスク $M = (m_1, \dots, m_n)$ 、 $m_i \in \{0, 1\}$ に基づき、UX を (7) 式の両親 A 、 B に適用すると、(8) 式に示す子 C が得られる。この

とき、両親の遺伝子型の距離は $\delta_g(A, B) = 6$ である。

$$\begin{cases} M &= (0, 1, 0, 1, 0, 1, 0) \\ A &= (1, 1, 2, 2, 3, 3, 2) \\ B &= (1, 2, 3, 4, 2, 2, 4) \end{cases} \quad (7)$$

$$C = (1, 2, 2, 4, 3, 2, 2) \quad (8)$$

ただし、 $c_i = a_i$ ($m_i = 0$)、 $c_i = b_i$ ($m_i = 1$)。

一方、調和交叉 (H-UX) を (7) 式に示した A 、 B に適用する。はじめに、 B を (9) 式のように同値な遺伝子型 $B^z \in [B]$ に変換する。このとき、 B^z は両親の表現型の距離を与え、 $\delta_g(A, B^z) = \delta_p([A], [B]) = 2$ となる。

$$B^z = (1, 3, 4, 2, 3, 3, 2) \quad (9)$$

次に、ビットマスク M に基づき、両親 A 、 B^z に UX を適用すると、(10) 式に示すような子 C が得られる。

$$C = (1, 3, 2, 2, 3, 3, 2) \quad (10)$$

《定理3》 調和交叉を A 、 B に適用して C が得られるとき、対応する表現型は次の関係を満たす。

$$\delta_p([A], [C]) + \delta_p([B], [C]) = \delta_p([A], [B]) \quad (11)$$

(証明) 付録Bを参照のこと。□

定理3より、調和交叉で得られる子 C は、解空間において両親 $[A]$ 、 $[B]$ の中間に位置し、明らかに両親の構造的な形質を継承している。ここで、遺伝子型の距離を用いて形質遺伝性を議論することは適当でない。なぜならば、個体の形質は表現型に付随するものであり、遺伝子型ではコード化の影響によって形質が破壊されているためである。

3.3 突然変異

調和交叉によれば両親を結ぶ線分上に子が生成されるため、解空間の景観を考慮した効率的な探索が期待できる。反面、調和交叉は個体集団の多様性を喪失させ、初期収束を招く恐れもある。そこで、調和交叉によって得られた子に突然変異を施すことで、両親から継承した形質を損なわない程度の摂動を与える。すなわち、両親 A と B の調和交叉により子 C が得られたものとする。ここで、 $a_i \in A$ と $b_i \in B^z$ を比較して $a_i \neq b_i$ ならば、 $c_i \in C$ の値を $1 \sim k+1$ の範囲でランダムに発生させた整数によって書き換える。

《定理4》 調和交叉により A 、 B から得られた C に突然変異を施して D とするとき、表現型は次の関係を満たす。

$$\max\{\delta_p([A], [D]), \delta_p([B], [D])\} \leq \delta_p([A], [B]) \quad (12)$$

(証明) 付録Cを参照のこと。□

定理4より、両親から子が継承した形質は、突然変異を施しても失われえないと言える。また、提案した調和交叉と突然変異を概念的に図示すると Fig. 2 となる。

さらに、上記の突然変異を施す確率 p_m は、両親 $[A]$ 、 $[B]$ 間の距離のシェアリング関数⁴⁾によって調節する。

$$p_m = \max \left\{ 0, 1 - \frac{\delta_p([A], [B])}{\delta_{share}} \right\} \quad (13)$$

ただし、 δ_{share} は初期集団内の個体間の平均距離とする。

Fig. 2 Harmonic crossover and mutation

3.4 局所探索法

局所探索法は近傍構造を基盤として構築され、GCPにおいてはmove近傍がある³⁾。すなわち、頂点集合 V の分割方法 $s = \{V_p \mid 1 \leq p \leq k\}$ のmove近傍 $N(s)$ とは、ある頂点 $v_i \in V_p$ を別の部分集合 V_q ($1 \leq q \leq k+1$, $q \neq p$)に移すことによって得られる分割方法 s' の集合である。また、(1)式の目的関数について、ある解 s のmove近傍に $g(s) > g(s')$ となる解 s' が存在しないとき、その解 s は局所的最適解と呼ばれる。さらに、このような局所的最適解は、GCPの実行可能解であることが証明されている⁹⁾。

GCPに対する局所探索法では、現在の解 s のmove近傍を走査して、 $g(s) > g(s')$ となる解 s' が存在すれば、直ちにその解 s' へ移るという動作を繰り返す。そして、解 s が局所的最適解となれば、探索を終了するものとする。このように、局所探索法は解の近傍を系統的に探索することで、局所的最適解(GCPの実行可能解)を確保する。

GAにおいて、GCPの解(表現型)に対応する任意の遺伝子型を A とする。このとき、表現型 $[A]$ のmove近傍に属する表現型の遺伝子型 B は、遺伝子型 A のある要素 $a_i \in A$ の値 p を別の値 q に変更することで得られる。さらに、両者の遺伝子型の距離は $\delta_g(A, B) = 1$ であり、(5)式の関係から、表現型 $[A]$ のmove近傍 $N([A])$ に含まれる表現型 $[B]$ は、解空間において(14)式の関係を満たす。

$$\forall [B] \in N([A]) \quad \delta_p([A], [B]) = 1 \quad (14)$$

上記の局所探索法の実装では、目的関数の反復計算が最も負担となるため、その高速化を図っている。(1)式に示した $g(s)$ の計算において、すべての頂点に隣接する頂点を調べて $|E_p|$ ($p = 1 \sim k$)の値を求めると、時間計算量は $O(n^2)$ である。そこで、現在の目的関数値 $g([A])$ からの変化量のみを評価することで、近傍の表現型 $[B] \in N([A])$ に対する目的関数値 $g([B])$ の時間計算量を削減している。すなわち、 $g([A])$ の計算過程で $|V_p|$, $|E_p|$ ($p = 1 \sim k$)の値を記憶しておけば、移動した一つの頂点に隣接する頂点のみを調べれば済むので、 $g([B])$ の時間計算量は $O(n)$ となる。

3.5 GAのハイブリッド化手法

GAは良い解の近傍における系統的な探索が苦手であるため、実用的なGAの構築では何らかの局所探索法を組み込む必要がある。また、前節で述べたように、実行可能解を保証するために局所探索法が不可欠となる場合もある。ここで、GAと局所探索法の汎用的なハイブリッド化手法として、ラマルク主義に基づく遺伝的局所探索(GLS: Genetic Local Search)がある⁴⁾。すなわち、GAにおける集団内の個体を初期値として局所探索法を実行し、得られた局所的最適解を用いて、その初期値となった個体を書き換える。

本論文ではGLSにおいて、調和交叉の際に両親の表現型の距離を計算し、あらかじめ定められた交叉距離 L_c よりも短ければ、得られた子(突然変異後の個体)を初期値として局所探索法を実行する。ここで、 L_c ($0 \leq L_c \leq n$)の値を変えることで、GAに局所探索法を組み込むタイミングを調節できる。たとえば、 $L_c = n$ とした場合は、すべての子に対して局所探索法が適用される。適応度に基づき選ばれた両親が解空間で近接しているならば、その付近に良い解が存在する可能性が高い。また、定理3と定理4から、調和交叉によって得られた子は、解空間において常に両親の中間に位置する。したがって、交叉距離 L_c を小さくすれば、GAによる大域的探索で絞り込まれた有望な領域において、局所探索法を集中的に実行できる。反面、GAが局所探索法とは独立に行われるようになるため、両者の相互作用は弱くなる。

メタ戦略の枠組となるGLSの世代交代モデルでは、個体集団の多様性を保持するため、定常状態モデルを基礎とし、同世代の集団内に同値な個体が含まれることを禁止する。ただし、二つの個体が同値であるとは、表現型の距離が零となることである。さらに、調和交叉においては、異なる両親を適応度に基づくルーレット選択によって決める。ここで、個体 $[A]$ の適応度 $f([A])$ は、その目的関数値 $g([A])$ を用いて(15)式のように定義する。目的関数 $g([A])$ は負の値になることもあるが、適応度 $f([A])$ は常に非負である。

$$f([A]) = g_{max} - g([A]) + |g_{min}| \quad (15)$$

ただし、ある世代 t における個体集団 $P(t)$ 内での目的関数の最大値を g_{max} 、最小値を g_{min} とする。

提案するGLSの世代交代モデルを以下に示す。

【世代交代モデル(GLS)】

- 手順1: ランダムに $|P|$ 個の個体(遺伝子型)を生成して初期集団 $P(0)$ を作り、世代 $t = 0$ とする。ただし、初期集団 $P(0)$ 内に同値な個体は存在させない。
- 手順2: 各個体 $A \in P(t)$ の適応度 $f([A])$ を計算する。
- 手順3: 適応度に基づくルーレット選択によって、集団 $P(t)$ から二つの異なる個体 A, B ($A \neq B$)を選ぶ。
- 手順4: 両親 A, B の調和交叉により子 C を生成する。
- 手順5: 確率 p_m に基づき、 C に突然変異を施す。
- 手順6: $\delta_p([A], [B]) \leq L_c$ ならば、 C (あるいは突然変異を施された C)を局所探索法で改良する。

手順7: 集団 $P(t)$ に C と同値な個体が存在しなければ, C を $P(t)$ 内で適応度が最小の個体と置き換える.

手順8: 終了条件を満たせば, $P(t)$ 内で適応度が最大の個体を解として終了する. そうでなければ, 世代 $t = t + 1$ として手順2に戻る. ■

ただし, 手順8における終了条件としては, 世代数 t が最終世代 T に達する, 何世代か最良の個体が更新されない, 規定回数の局所探索法が実行されるなどが考えられる.

4. 計算例と考察

GCP のベンチマーク集 DIMACS¹⁵⁾ から, 大きさの異なる三種類の例題#1 (DSJC125.1: 頂点数 125; 枝数 1472), #2 (zero1.i.1: 211; 4100), #3 (school1.nsh: 352; 14612) において, 調和交叉と GLS での交叉距離 L_c を評価する.

4.1 調和交叉の評価

定理3で示したように, 提案した調和交叉では両親の構造的な形質が子に引き継がれる. はじめに, Manderick¹⁴⁾らが提唱している相関係数を用いて, 調和交叉が親子の関係に与える影響を調べる. ランダムに生成された両親に対する目的関数値の平均を $\bar{g}_p$ とし, その両親の交叉で生成された子の目的関数値を g_c として, 交叉における親子間の相関係数 ρ を (16) 式のように定義する. 相関係数 ρ ($-1 \leq \rho \leq 1$) の値が 1 (-1) に近いほど, $\bar{g}_p$ と g_c には正 (負) の相関が強く, 相関の存在が認めにくい場合には $\rho \simeq 0$ となる.

$$\rho(\bar{g}_p, g_c) = \frac{\text{Cov}(\bar{g}_p, g_c)}{\sigma(\bar{g}_p) \sigma(g_c)} \quad (16)$$

ただし, Cov は共分散, σ は標準偏差である.

ランダムに生成した 10^4 組の両親を交叉させて求めた相関係数を Table 1 に示す. ただし, 従来の交叉法 1X, 2X, UX を調和交叉に拡張したものを, それぞれ H-1X, H-2X, H-UX と表す. Table 1 では, すべての例題において, 調和交叉の相関係数が本来の交叉よりも向上している. 相関係数は目的関数値による形質遺伝性の評価指標であることから, 表現型の距離で定義された構造的な形質が, 各例題の解空間の景観を反映していることが確認できる.

次に, GLS の世代交代モデルで $L_c = 0$ として局所探索法を除き, 調和交叉と突然変異の効果を調べる. 例題#1 において, 突然変異を用いない UX, H-UX と, H-UX に突然変異を取り入れた Hm-UX を比較した結果を Fig. 3 に示す.

Fig. 3 Best value of objective function

Table 2 Performance of harmonic crossovers

	cross.	$\bar{g}$	$\sigma(g)$	$\bar{e}$	T
#1	UX	-1905.6	109.5	0.0	0.8×10^5
	H-UX	-1807.8	208.9	0.0	
	Hm-UX	-2319.4	119.1	0.0	
#2	UX	-6532.0	877.0	32.5	1.0×10^5
	H-UX	-4518.2	1084.0	22.5	
	Hm-UX	-11180.2	2433.8	36.9	
#3	UX	-494.2	36.8	121.3	1.3×10^5
	H-UX	-2156.2	112.4	14.3	
	Hm-UX	-2122.0	90.2	19.0	

従来の UX に比べて, H-UX では初期収束が起きているが, Hm-UX では明らかに探索性能の改善が見られる. 上記の各交叉法について, すべての例題における実験結果を Table 2 に示す. Table 2 では, 10 回の試行で得られた最良解に対する目的関数値の平均 $\bar{g}$ と標準偏差 $\sigma(g)$, 制約条件を満たさない枝数の平均 $\bar{e}$, 最終世代 T を示す. ただし, 集団サイズ $|P| = 200$, 最終世代 T は個体集団が十分に収束した世代数とした. Table 2 から, 例題#1 と例題#2 では Hm-UX によって最良の解が得られており, 調和交叉に突然変異を導入して, 新たな形質を積極的に獲得する必要がある. 一方, 規模の大きな例題#3 では, 形質遺伝性のみを重視した H-UX も効果的である. ところで, 例題#1 以外は $\bar{e} \neq 0$ であり, GCP の実行可能解は得られていない.

4.2 交叉距離の評価

GLS の世代交代モデルで交叉距離 L_c を変化させ, GA に局所探索法を組込む適切なタイミングを調べる. 例題#1 において, $L_c = n, 0.45n, 0.15n, 0.05n$ ($n = 125$) とした場合と, GA のみ ($L_c = 0$) の性能を比較した. ただし, 調和交叉 Hm-UX を使用し, 集団サイズは $|P| = 200$ である. 各世代 t で得られた最良の目的関数値を Fig. 4 に示す. 同様に, 集団 $P(t)$ における個体間の平均距離を Fig. 5 に示す. Fig. 5 からは, 個体集団の多様性を評価できる.

Fig. 4 と Fig. 5 によれば, 交叉距離 L_c を大きくするほど, 最良解への収束が急激となり, 個体集団の多様性も早期に失

Table 1 Correlation coefficient of crossovers

	cross.	ρ	cross.	ρ
#1	1X	0.4899	H-1X	0.5458
	2X	0.4836	H-2X	0.5407
	UX	0.3657	H-UX	0.4509
#2	1X	0.5344	H-1X	0.5373
	2X	0.5404	H-2X	0.5669
	UX	0.4315	H-UX	0.4594
#3	1X	0.5120	H-1X	0.5427
	2X	0.5110	H-2X	0.5409
	UX	0.3682	H-UX	0.4534

Fig. 4 Best value of objective function

Fig. 5 Average of phenotypic distance

われているにも関わらず、最終的に得られる解は良くなっている。これらの結果から、局所探索法の実行を制限して個体集団の多様性を長く維持するよりも、最初から局所探索法をGAに組込む方が望ましいと言える。その理由として、GCPに対するGLSでは、局所探索法が実行不可能解を実行可能解に修正する働きを持つためと考えられる。すなわち、実行不可能解を含む領域を探索空間とするGAは探索効率が悪く、局所探索法で得られた局所的最適解（実行可能解）に対して、調和交叉による大きな摂動を加える方が、大域的な探索を効率的に行えるものと推察される。それならば、提案したメタ戦略において、GAと局所探索法の相互作用による創発的な探索を実現するためには、局所探索法が実行される時点で、個体集団の多様性が高いことが重要である。

そこで、集団サイズ $|P|$ と同じ回数の局所探索法が実行された時点で、GLSを終了するものとし、すべての例題における実験結果をTable 3に示す。Table 3では、GLSの集団サイズを $|P| = 200$ とし、交叉距離 L_c を変えて、10回の試行で得られた最良解に対する目的関数値の平均 $\bar{g}$ と標準偏差 $\sigma(g)$ 、GLSが終了した世代数の平均 $\bar{t}$ 、最期の集団 $P(\bar{t})$ に

Table 3 Evaluation of crossover distance

	L_c	$\bar{g}$	$\sigma(g)$	$\bar{t}$	$\bar{\delta}_p$
#1	n	-2876.0	52.6	200.0	86.6
	$0.45n$	-2862.0	27.5	2801.0	60.3
	$0.15n$	-2725.6	74.6	9992.6	18.4
	$0.05n$	-2652.2	99.2	19559.2	7.6
	(MLS)	-2825.2	54.1	—	—
#2	n	-14663.4	3.3	200.0	53.5
	$0.45n$	-14384.2	319.8	1211.3	42.4
	$0.15n$	-13267.0	936.6	9559.4	28.4
	$0.05n$	-13951.6	764.6	36973.2	11.6
	(MLS)	-14423.0	117.0	—	—
#3	n	-6122.2	242.0	200.0	259.1
	$0.45n$	-5232.2	189.9	12113.2	79.6
	$0.15n$	-5108.2	113.7	26758.8	28.4
	$0.05n$	-5049.0	165.6	41559.5	10.8
	(MLS)	-5530.6	198.9	—	—

おける個体間の平均距離 $\bar{\delta}_p$ を示す。同様に、MLSはランダムに生成した $|P|$ 個の初期解に対して、局所探索法のみを適用した結果である。ここで、すべての解法において、局所探索法の実行回数は $|P| = 200$ 回となることに注意されたい。Table 3から、 $L_c = n$ とした場合に最良の解が得られており、最期の集団 $P(\bar{t})$ における多様性も高いことが確認できる。一方、 $L_c = 0.05n$ とした場合は、膨大な世代数を費した割には良い解が得られていない。また、最期の集団 $P(\bar{t})$ では多様性が著しく低下している。さらに、 $L_c = 0.05n$ としたGLSの性能はMLSにも劣り、不適切なハイブリッド化は逆効果となる場合もあることを示唆している。

5. おわりに

本論文では、GAの特性を考慮した個体間の距離を定義するとともに、近傍構造を包括する距離空間の観点から、GAを枠組とするメタ戦略の構築法について考えた。はじめに、GCPの解空間において表現型の距離を定義した。次に、調和交叉と突然変異を提案し、これらの遺伝演算子の解空間における作用を明らかにした。最後に、GAと局所探索法の効果的な相互作用を促すためには、個体集団の多様性が高い段階で、GAに局所探索法を組込む必要があることを示した。個体集団の多様性はGAの探索能力を映すものであるが、多様性の維持を目的としてはならない。個体集団の多様性が失われる過程として、創発的な探索は進むからである。

ところで、提案した表現型の距離や調和交叉は、遺伝子型の冗長性を利用したものである。組合せ最適化問題にGAを適用する場合、その遺伝子型の表現方法が冗長となる事例は少なくない^{11), 16)}。したがって、これらの技法はGCPのみならず、多くの組合せ最適化問題に適用できる。

GAを含む創発的システムでは、その挙動について理論的な予測が困難なものが多く、理論的な方法での信頼性の保証が難しいと言われている¹⁷⁾。本論文で導入した個体集合の距離空間という概念は、GAの挙動解析のみならず、GAの理論的な発展においても寄与できるものと期待する。

参考文献

- 1) 渡辺治: NP 型問題の近似解法の可能性について, オペレーションズ・リサーチ, **2**, 105/111 (1994)
- 2) C. R. Reeves (Ed.): Modern Heuristic Techniques for Combinatorial Problems, John Wiley & Sons (1993)
- 3) 室田一雄編/久保幹雄: 離散構造とアルゴリズム IV, 第5章 メタヒューリスティック, 近代科学社 (1995)
- 4) 三宮, 喜多, 玉置, 岩本: 遺伝アルゴリズムと最適化, 朝倉書店 (1998)
- 5) 北野宏明編/相澤彰子: 遺伝的アルゴリズム 2, 第1章 スキーマ貪欲な遺伝的探索アルゴリズム, 産業図書 (1995)
- 6) Y. Davidor/大橋, 濱岡共訳: 遺伝的アルゴリズム, 第9章 遺伝的アルゴリズムにおけるエピスタシス, 培風館 (1996)
- 7) D. E. Goldberg: Genetic algorithms and walsh functions: part I - II, Complex Systems, **3**, 129/171 (1989)
- 8) 森, 吉田, 喜多, 西川: 遺伝アルゴリズムにおける熱力学的選択ルールの提案, システム制御情報学会論文誌, **9-2**, 82/90 (1996)
- 9) D. S. Johnson, C. R. Aragon, L. A. Mcgeoch and C. Schevon: Optimization by simulated annealing: an experimental evaluation; part II, graph coloring and number partitioning, Operations Research, **39-3**, 378/406 (1991)
- 10) D. R. Jones and M. A. Beltramo: Solving partitioning problems with genetic algorithms, Proc. of the 4th ICGA, 442/449 (1991)
- 11) K. Tagawa, Y. Kanzaki, D. Okada, K. Inoue and H. Haneda: A new metric function and harmonic crossover for symmetric and asymmetric traveling salesman problems, Proc. of 1998 IEEE ICEC, 822/827 (1998)
- 12) 伊理, 白川, 梶谷, 篠田: 演習グラフ理論—基礎と応用—, 第4章 ネットワーク手法, コロナ社 (1983)
- 13) 北野宏明編/小林, 山村: 遺伝的アルゴリズム, 第2章 遺伝的アルゴリズムにおけるコード化と交叉, 産業図書 (1993)
- 14) B. Manderick, M. de Weger and P. Spiessens: The genetic algorithm and the structure of the fitness landscape, Proc. of the 4th ICGA, 143/150 (1991)
- 15) ftp://dimacs.rutgers.edu/pub/challenge/graph/
- 16) S. Ronald, J. Asenstorfer and M. Vincent: Representational redundancy in evolutionary algorithms, Proc. of 1995 IEEE ICEC, 631/636 (1995)
- 17) 喜多一: 工学の立場から見た創発システム, 計測と制御, **35-7**, 532/535 (1996)

付 録

A. 定理2の証明(補足)

表現型の距離を求める問題が, $O(n)$ で最大割当て問題に変換でき, 二つの問題が等価であることを示す.

遺伝子型 $A = (a_1, \dots, a_n)$, $B = (b_1, \dots, b_n)$ について, それぞれの彩色数を $k(A)$, $k(B)$ ($k(A) \leq k(B)$) とする. A , B から $k(A) \times k(B)$ の重み行列 $W = [w_{qp}]$ を (A.1) 式のように定義する. この時間計算量は $O(n)$ である.

$$w_{qp} = |\{i \in I \mid a_i = q \wedge b_i = p\}| \quad (\text{A.1})$$

ただし, 添字集合 $I = \{1, 2, \dots, n\}$ とする.

最大割当て問題では, W の各行と各列の要素を高々一つずつ選んでできる行列 $Z = [z_{qp}]$, $z_{qp} \in \{0, 1\}$ について, (A.2) 式の目的関数 $\varepsilon(Z)$ を最大化する.

$$\varepsilon(Z) = \sum_{z_{qp}=1} w_{qp} \quad (\text{A.2})$$

一方, 遺伝子型の距離は (A.3) 式のように定義できる.

$$\delta_g(A, B) = n - \sum_{i=1}^n \bar{h}(a_i, b_i) \quad (\text{A.3})$$

ただし, $\bar{h}(a_i, b_i) = 1$ ($a_i = b_i$), $\bar{h}(a_i, b_i) = 0$ ($a_i \neq b_i$).

まず, B の幾つかの色番号を $p \rightarrow q$, ($1 \leq q \leq k(A)$) と変更して $B^x \in [B]$ とするとき, $\delta_g(A, B^x) = \delta_p([A], [B])$ が成り立つとする. ここで, 色番号の変更 $p \rightarrow q$ に対応する行列 $W = [w_{qp}]$ の q 行 p 列の要素を選んで $Z = [z_{qp}]$ を作る. このとき, 最大割当て問題の最適解を Z^* とすると,

$$\delta_p([A], [B]) = n - \varepsilon(Z) \geq n - \varepsilon(Z^*) \quad (\text{A.4})$$

逆に, 最大割当て問題の最適解 $Z^* = [z_{qp}]$ において $z_{qp} = 1$ のとき, B の色番号を $p \rightarrow q$ と変更することによって得られる同値な遺伝子型の一つを $B^y \in [B]$ とすると,

$$\delta_g(A, B^y) = n - \varepsilon(Z^*) \geq \delta_p([A], [B]) \quad (\text{A.5})$$

(A.4) 式と (A.5) 式より (A.6) 式が成り立ち, 最適解 Z^* が表現型の距離 $\delta_p([A], [B])$ を与える.

$$\delta_p([A], [B]) = n - \varepsilon(Z^*) \quad (\text{A.6})$$

B. 定理3の証明

遺伝子型 $A = (a_1, \dots, a_n)$ の各要素 a_i ($i = 1 \sim n$) を, その値 (対立遺伝子) と位置 (遺伝子座) を表す遺伝子と考える. さらに, 遺伝子型 A は文字列ではなく, 遺伝子 a_i の集合 $A = \{a_i \mid 1 \leq i \leq n\}$ であると見做すと, ハミング距離として定義した遺伝子型 A と B の距離は, 異なる遺伝子の総数で決まり, (B.1) 式のように表すことができる.

$$\delta_g(A, B) = n - |A \cap B| \quad (\text{B.1})$$

A と B の調和交叉により C が生成されるとき, B は同値な $B^z \in [B]$ に変換され (B.2) 式と (B.3) 式が成り立つ.

$$\delta_p([A], [B]) = \delta_g(A, B^z) \quad (\text{B.2})$$

$$A \cup B^z \supseteq C \supseteq A \cap B^z \quad (\text{B.3})$$

(B.3) 式から, 分配則を用いて (B.4) 式が導ける.

$$C = (A \cup B^z) \cap C = (A \cap C) \cup (B^z \cap C) \quad (\text{B.4})$$

ここで, (B.3) 式を考慮して集合の要素数を考えると,

$$\begin{cases} |C| &= |(A \cap C) \cup (B^z \cap C)| \\ &= |A \cap C| + |B^z \cap C| - |A \cap B^z| \end{cases} \quad (\text{B.5})$$

さらに, (B.1) 式での定義と $|C| = n$ より,

$$\delta_g(A, C) + \delta_g(B^z, C) = \delta_g(A, B^z) \quad (\text{B.6})$$

(5) 式, (B.2) 式, (B.6) 式より,

$$\delta_p([A], [C]) + \delta_p([B], [C]) \leq \delta_p([A], [B]) \quad (\text{B.7})$$

一方, 表現型の距離は「距離の公理」を満たすので,

$$\delta_p([A], [C]) + \delta_p([B], [C]) \geq \delta_p([A], [B]) \quad (\text{B.8})$$

(B.7) 式と (B.8) 式より, (B.9) 式が成り立つ.

$$\delta_p([A], [C]) + \delta_p([B], [C]) = \delta_p([A], [B]) \quad (\text{B.9})$$

C. 定理 4 の証明

(B.3) 式の子 C に突然変異を施して D とすると,

$$D \supseteq A \cap B^z \quad (C.1)$$

さらに, 集合の演算によって,

$$\begin{cases} D \supseteq A \cap B^z \wedge A \supseteq A \cap B^z \\ \Rightarrow A \cap D \supseteq A \cap B^z \\ \Rightarrow |A \cap D| \geq |A \cap B^z| \end{cases} \quad (C.2)$$

(B.1) 式に示した遺伝子型の距離の定義より,

$$\begin{cases} \delta_g(A, D) = n - |A \cap D| \\ \leq n - |A \cap B^z| = \delta_g(A, B^z) \end{cases} \quad (C.3)$$

(5) 式, (B.2) 式, (C.3) 式より,

$$\begin{cases} \delta_p([A], [D]) \leq \delta_g(A, D) \\ \leq \delta_g(A, B^z) = \delta_p([A], [B]) \end{cases} \quad (C.4)$$

同様にして, (C.5) 式が導出できる.

$$\delta_p([B], [D]) \leq \delta_p([A], [B]) \quad (C.5)$$

(C.4) 式と (C.5) 式より, (C.6) 式が成り立つ.

$$\max\{\delta_p([A], [D]), \delta_p([B], [D])\} \leq \delta_p([A], [B]) \quad (C.6)$$

[著 者 紹 介]

田 川 聖 治 (正会員)

1991 年神戸大学大学院修士課程修了. 93 年神戸大学工学部助手. 最適化アルゴリズム, 進化的計算, ロボットの計算機制御などの研究に従事 (博士 (工学)). 日本ロボット学会, システム制御情報学会, 電気学会, IEEE 各会員.

金 重 憲 治

1998 年神戸大学工学部電気電子工学科卒業. 同年, 同大学院自然科学研究科博士前期課程入学. 遺伝アルゴリズムを中心とした最適化手法などの研究に従事.

井 上 克 巳

1984 年京都大学大学院工学研究科数理工学専攻修士課程修了. 松下電器産業 (株), (財) 新世代コンピュータ技術開発機構, 豊橋技術科学大学工学部情報工学系を経て, 97 年神戸大学工学部助教授. 人工知能, 論理プログラミング, 計算機科学に関する教育研究に従事 (京都大学博士 (工学)). 情報処理学会, 人工知能学会, AAAI 各会員.

羽根田 博 正 (正会員)

1972 年カリフォルニア大学 (バークレイ校) 大学院博士課程修了. 同年神戸大学工学部助教授, 82 年同教授. 97 年 IEEE 産業エレクトロニクス部門学会会長. 計算機援用システム設計・解析などの研究に従事 (Ph.D.). システム制御情報学会, 電気学会, 電気情報通信学会, 情報処理学会, ACM, IEEE (フェロー) などの会員.