

遺伝的アルゴリズムにおける遺伝子-評価値関係の可視化と遺伝的演算へのフィードバック†

石黒 英敬*・吉川 大弘*・古橋 武*

遺伝的アルゴリズム (Genetic Algorithm: GA) では一般的に、個体の評価値情報に基づいて淘汰・増殖などの遺伝的演算が施される。しかしこのとき、実際に探索が行われるのは遺伝子空間であるため、遺伝子空間と評価値空間における近傍関係の類似性が、効率的な解探索の上で重要となる。通常、これらの対応関係を決定づけるのがコーディングであるが、遺伝子空間と評価値空間との類似性を示す適切な指標が存在しないことで、用いたコーディングに対する評価や、適切な遺伝的演算方法の決定について、試行錯誤により行っていることも多い。本研究では、効率的な解探索を行う上で、遺伝子空間と評価値空間の類似関係を把握することを目的に、GAにおける探索中の個体を対象とした、遺伝子と評価値の関係の可視化手法の開発を目指す。本論文では、遺伝子空間、評価値空間それぞれにおける個体間の距離関係、そして遺伝子と評価値との非類似度を、同一空間内で表現する可視化手法を提案する。提案手法により、GAにおける遺伝子空間と評価値空間との類似性の把握支援が可能となることが期待される。また、得られた可視化結果を遺伝的演算にフィードバックすることで、求める評価値領域を効率的に探索可能となることを示す。

キーワード：可視化、遺伝的アルゴリズム、遺伝子空間、評価値空間

1. はじめに

組み合わせ最適化問題において準最適解を実用的な時間内に求める手法として、遺伝的アルゴリズム (Genetic Algorithm: GA) [1]が広く用いられている。GA では一般的に、個体の評価値情報に基づいて淘汰・増殖などの遺伝的演算が施される。しかしこのとき、実際に探索が行われるのは遺伝子空間であるため、遺伝子空間と評価値空間における近傍関係の類似性が、効率的な解探索の上で重要となる。なぜなら、これらの近傍関係が類似していない解空間では、評価値の高い個体の遺伝子空間における周辺探索が、評価値空間における周辺探索につながらず、評価値の高い個体に積極的に遺伝的演算を行うことの効果が薄れるためである。この対応関係を決定づけるのがコーディングであるが、遺伝子空間と評価値空間との類似性を示す適切な指標が存在しないことで、経験やカン、そして試行錯誤に基づいてコーディング方法を決定していることも多い。しかし用いたコーディング方法に対する評価や、そのコーディングにより決まる解空間の

ランドスケープに対して適切な遺伝的演算方法の決定についても、最終的に得られた評価値の比較をその決定の指標とすることが多い[2][3]。これは、特に問題が複雑化するほど、遺伝子空間、評価値空間がともに超多次元空間となり、それらの類似関係の把握が容易ではなくなるためである。そこで本研究では、効率的な解探索を行う上で、遺伝子空間と評価値空間の類似関係の把握を支援することを目的に、GAにおける探索中の個体を対象とした、遺伝子と評価値の関係の可視化手法の開発を目指す。

高木ら[4]は、インタラクティブ進化的計算において、評価値空間のランドスケープを自己組織化マップ (Self Organizing Map: SOM) [5]によって可視化してユーザに提示することで、評価値の高い個体の存在する領域を予測し、解探索の効率化を行う手法を提案した。この手法では、評価値の高い個体が集中しているような領域の1点をユーザ自身が選択し、その領域中の個体に対して遺伝的演算を施すことで、解の収束の高速化を試みている。この手法は、GAにおける遺伝子と評価値の関係を可視化するという点では本研究と類似してはいるが、遺伝子空間における近傍関係と評価値空間における近傍関係が類似していることを前提とした解の収束の高速化手法であり、これらの類似関係の把握そのものを目的とした本研究との位置付けは異なっている。また、個体の位置情報を2次元に射影し、評価値の軸と合わせて3次元で可視化を行ってい

† Visualization of Relationships between Genes and Evaluation Values in GA and Feedback into Genetic Operations
Hidetaka ISHIGURO, Tomohiro YOSHIKAWA and Takeshi FURUHASHI

* 名古屋大学大学院工学研究科計算理工学専攻
Dept. of Computational Science and Engineering, Graduate School of Engineering, Nagoya University

るため、評価値の軸が複数必要となる多目的最適化問題での適用は困難である。

本論文では、GAによる探索において、遺伝子空間内における他の個体との距離関係と、評価値空間内における距離関係との相関係数を用いて、遺伝子と評価値の非類似度を定義する。そして多次元尺度構成法 (Multi Dimensional Scaling: MDS) [6]を用いることで、遺伝子空間の距離関係、評価値空間での距離関係、遺伝子と評価値との非類似度を、同一空間内で表現する可視化手法を提案する。提案手法により、GAにおける遺伝子空間と評価値空間との類似性の把握支援が可能となることが期待される。本論文では、多目的最適化手法として代表的な Non-Dominated Sorting Genetic Algorithm-II (NSGA-II) [7]をベンチマーク関数に適用した際の個体群に提案手法を用い、得られた可視化結果を NSGA-II の遺伝的演算にフィードバックすることで、求める評価値領域を積極的に探索可能となること、そして効率的な遺伝的演算が実現できることを示す。

2. 提案手法

提案手法では、GAにおける個体間の遺伝子空間上での距離関係、評価値空間上での距離関係、そして遺伝子と評価値の非類似度に対して、MDSを適用することで同一空間内にマッピングする。提案手法では、各個体間の遺伝子空間上での距離 G_{ij} と、評価値空間上での距離 E_{ij} をそれぞれ計算し (G_{ij} , E_{ij} はそれぞれ個体 i と個体 j との距離)、(1)式により、それらの類似度 S_{ij} を求める。遺伝子距離や評価値距離は、例えばビットコーディングではハミング距離、実数コーディングではユークリッド距離を用いるなど、任意の距離指標で置き換えることが可能である。

$$S_{ij} = \frac{\sum_{k=1}^{N_p} (G_{ik} - \bar{G}_i) (E_{jk} - \bar{E}_j)}{\sqrt{\sum_{k=1}^{N_p} (G_{ik} - \bar{G}_i)^2} \sqrt{\sum_{k=1}^{N_p} (E_{jk} - \bar{E}_j)^2}}$$

$$\bar{G}_i = \frac{1}{N_p} \sum_{k=1}^{N_p} G_{ik}$$

$$\bar{E}_j = \frac{1}{N_p} \sum_{k=1}^{N_p} E_{jk}$$
(1)

(1)式において、 S_{ij} は個体 i の遺伝子と個体 j の評価値との類似度、 N_p は可視化に用いる個体数を表している。類似度 S_{ij} は、個体 i の遺伝子空間上での近傍関係(他個体との距離関係)が個体 j の評価値空間上で

の近傍関係と対応するほど(相関が高いほど)大きな値となる。なお、通常 S_{ii} は自己相関係数であるが、ここでは遺伝子 i と評価値 i との相関係数となるため、1になるとは限らない。類似度 S_{ij} は相関係数として -1 から 1 の値をとるが、マッピングの際には、他の距離 (G_{ij} , E_{ij}) と等価に扱うために、1 から S_{ij} を引いた非類似度 D_{ij} を用いる ((2)式)。

$$D_{ij} = 1 - S_{ij} \tag{2}$$

提案手法では、得られた3つの距離関係 G_{ij} , E_{ij} , D_{ij} を各世代の最大値で割ることで規格化した G_{ij}^* , E_{ij}^* , D_{ij}^* を [8], (3)式で示す目的関数 \mathcal{F} に用いることで、MDSを適用する。また、MDSで用いる距離行列を図1に示す。図のように、提案手法では、個体 i について、遺伝子 i (遺伝子空間における個体 i) と評価値 i (評価値空間における個体 i) という二つの異なるデータとみなしてMDSを適用する。これにより、マップ上には遺伝子 i と評価値 i が別々に付置される。(3)式において、 g_{ij}^* , e_{ij}^* , d_{ij}^* は、写像した可視空間上におけるそれぞれの距離を示している。提案手法では、(3)式の目的関数 \mathcal{F} を最小化することで、それぞれの距離関係を保持したまま、同一空間内へのマッピングを行う。可視空間上での遺伝子と評価値の距離の遠近が、それぞれ遺伝子空間または評価値空間における距離関係およびそれらの非類似度を表すため、遺伝子空間と評価値空間の類似関係が視覚的に把握可能となる。

$$\mathcal{F} = \frac{\sum_{i < j}^{N_p} (G_{ij}^* - g_{ij}^*)^2 + \sum_{i < j}^{N_p} (E_{ij}^* - e_{ij}^*)^2 + \sum_{i, j}^{N_p} (D_{ij}^* - d_{ij}^*)^2}{\sum_{i < j}^{N_p} (G_{ij}^*)^2 + \sum_{i < j}^{N_p} (E_{ij}^*)^2 + \sum_{i, j}^{N_p} (D_{ij}^*)^2} \tag{3}$$

図2に、提案手法による可視化結果のイメージ図を示す。図では、遺伝子空間において個体Aと個体Eが

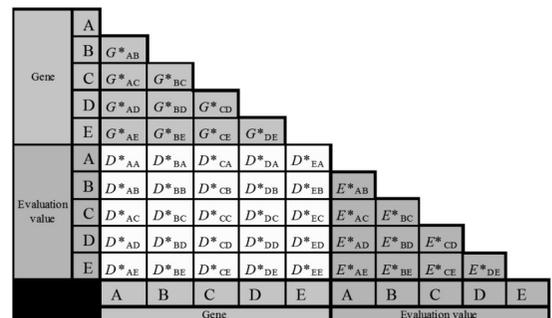


図1 MDSに用いる距離行列

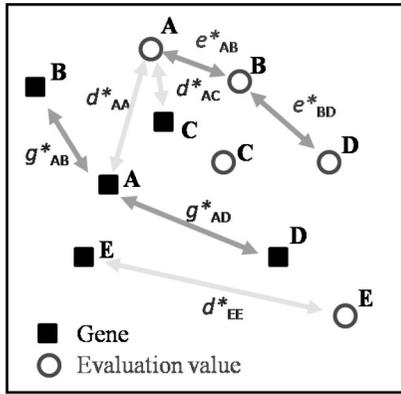


図2 MDSによる可視化のイメージ

近いこと、評価値空間において個体 A と個体 B が近く、個体 E とは離れていること、また、個体 A の評価値とは、遺伝子 A よりむしろ遺伝子 C の方が、その近傍関係が類似していることなどが把握できる。なお MDS では、各個体間の相対的な位置関係のみが意味を持ち、各座標や軸についての絶対的な意味はない。そのため、各個体間の相対的な位置関係を保持したまま、座標上で平行移動、回転または反転を行うことが可能である。

3. 実験

本論文では、多目的遺伝的アルゴリズムの代表的な手法の 1 つである NSGA-II を、多目的最適化のベンチマーク関数に適用した際の探索過程における個体群に対し、提案手法を用いて、それらの遺伝子と評価値との関係を可視化する。また、得られた可視化結果に基づいた、NSGA-II の遺伝的演算へのフィードバックを行う。

3.1 問題設定

本実験では、多目的最適化のベンチマーク関数として、(4)式に示す FON ベンチマーク関数(2 目的最小化問題)[7]を用いた。(4)式において、入力変数の数 N は 3 とした。図 3 に、FON ベンチマーク関数のパレート最適フロントを示す。

$$\begin{aligned} \min \{f_1\} &= 1 - \exp\left(-\sum_{i=1}^N \left(x_i - \frac{1}{\sqrt{3}}\right)^2\right) \\ \min \{f_2\} &= 1 - \exp\left(-\sum_{i=1}^N \left(x_i + \frac{1}{\sqrt{3}}\right)^2\right) \\ x_i &\in [-4, 4] \end{aligned} \quad (4)$$

染色体は、バイナリコードを用いて 3 変数を各 10 ビット

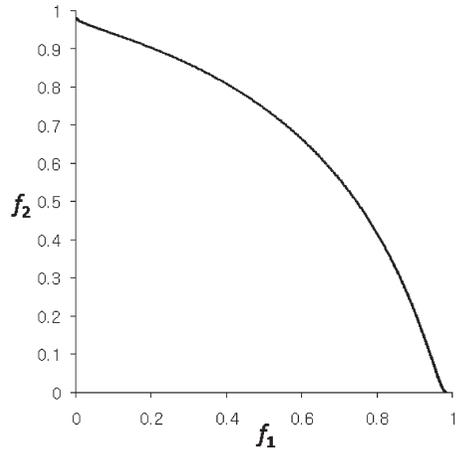


図3 FON ベンチマーク関数のパレート最適フロント

トで表し、計 30 ビットでコーディングした。また、NSGA-II における遺伝的演算パラメータとして、個体数を 50 個体、交叉率 0.9 の一点交叉、突然変異率 0.03 として各遺伝子座を選択し、1 ビット反転させる方法を用いた。なお、終了世代数は 30 世代とした。

3.2 NSGA-II の適用

NSGA-II では、評価値空間において疎な領域に存在する個体を親個体として選択し、積極的に遺伝的演算を施すことで、評価値空間における周辺領域を集中的に探索し、多様なパレート解を得る方法がとられる。ただしこのとき、優先して遺伝的演算を施されるのは“遺伝子”であるため、この方法は遺伝子空間上でのその個体の探索が評価値空間上でも疎な領域の探索と対応していることを前提とした方法であるといえる。そのため、その対応関係が成り立たない場合には、目的とする解探索が効率的には行えないと考えられる。例えば、評価値空間上で疎な領域をまたいで親個体を選択して交叉を行う場合に、遺伝子空間と評価値空間とが類似していない場合には、それらの親個体間の評価値空間上での疎な領域に子個体が生成されるとは限らない。

3.3 パレート解の遺伝子-評価値関係の可視化

初めに、ある世代で得られたパレート解(非劣解個体群)の遺伝子空間と評価値空間の対応関係を視覚的に把握し、得られた情報を用いて、その後の世代に対する遺伝的演算へフィードバックを行う。ここで(1)式における距離の定義として、遺伝子距離にはハミング距離、評価値距離にはユークリッド距離をそれぞれ用いた。

3.3.1 可視化結果

NSGA-IIによって得られた18世代目のパレート解 (図4) に対して提案手法を適用し、遺伝子と評価値の類似関係を可視化した (図5)。図4は、18世代目の個体群の評価値空間における分布を示す。ただし、図中のAとBは疎な領域に存在することで優先的に遺伝的演算が施された個体、Cは17世代目において劣解とみなされ淘汰された個体である。また、図5のひし形の点が遺伝子空間における各個体を、十字の点が評価値空間における各個体をそれぞれ表している。図5より、個体Aおよび個体Cは、遺伝子と評価値の距離が近く、類似度が高い、すなわち、遺伝子空間における近傍関係と、評価値空間における近傍関係が類似している個体であると考えられる。したがって、評価値空間において比較的疎な領域に属する個体Aの遺伝子の近傍を探索することは有効な演算であると考えられる。一方、個体Bも同様に評価値空間において疎

な領域に属しているが、遺伝子と評価値の距離が遠く、個体Bの遺伝子の近傍を探索しても、評価値空間において求める近傍探索と対応していないと考えられる。

3.3.2 選択へのフィードバック

図6 (a)に、通常通りNSGA-IIの親個体選択方法を用いて世代を進めた場合の、22世代目の個体群の分布の様子を示す。図6 (a)から、個体Aの周辺には新たな個体が生成されているものの、個体B周辺にはそれほど多くの個体が生成されていないことがわかる。個体Bは、疎な領域に属することで、遺伝的演算の親個体として選択されているものの、その周辺に新個体が生成されないことで親個体に選択され続け、結果的に演算効率を下げていると推測される。

一方、図5において、個体Bの“評価値”に近い“遺伝子”は、個体Cであることがわかる。すなわち、評価値空間における個体Bの近傍を探索するには、劣解として17世代目に淘汰された個体Cの遺伝子を選

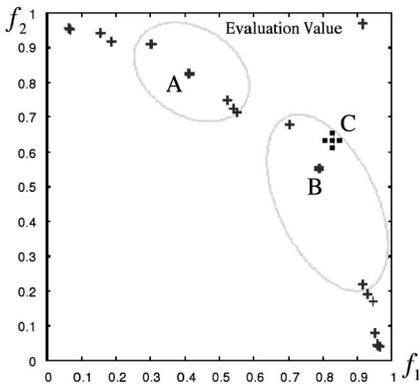


図4 パレート解の評価値分布 (18世代目)

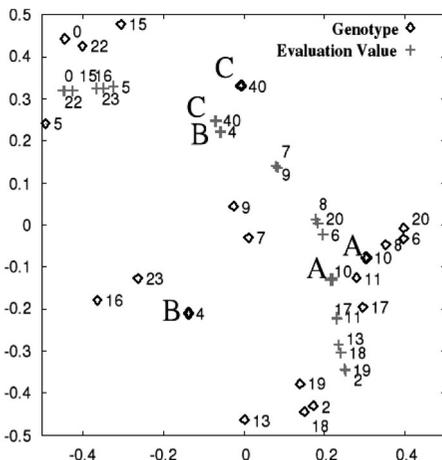
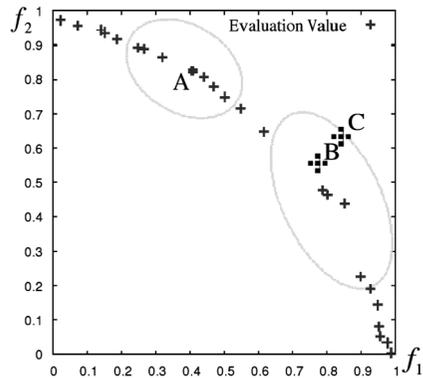
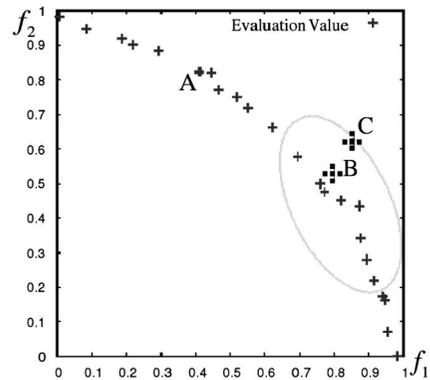


図5 提案手法による可視化結果 (18世代目)



(a) 通常



(b) フィードバック後

図6 評価値分布の比較 (22世代目)

択し、遺伝的演算を行うことが目的とする領域の探索につながると思われる。そこで、18世代目において、個体 B の代わりに個体 C を親個体として選択するという方法を用いた。図 6 (b) に、22 世代目の個体群の分布の様子を示す。図 6 (b) から、用いたフィードバックにより、評価値空間における個体 B の近傍にも新たな個体が多く生成されていることがわかる。

3.4 可視化による探索過程の特徴把握支援

次に、NSGA-II による探索過程を可視化し、遺伝的演算により生成された個体の特徴をとらえることで、より効率的な探索が可能となるよう、遺伝的演算へのフィードバックを行う。ここでは、ある世代の個体群と次世代に生成された子個体群とをあわせて可視化することで、探索過程における遺伝的演算の効果を把握する。

3.4.1 可視化結果

図 7 に、10~11 世代間における個体群の可視化結果を示す。図 7 において、四角の点は各個体の遺伝子を、三角は評価値を示しており、黒く塗りつぶされたものは次世代に残された個体、白抜きのものは淘汰された個体を表している。また、それぞれ色の濃い点は親個体を、色の薄い点は子個体を表している(その結果、マップ上において、色が薄い(白い)点が軸上にプロットされ、一部軸や数値が見えにくくなっている領域がある)。2 章で述べたように、MDS はその性質上、各点間の距離のみが意味を持つため、可視化結果は座標上で平行移動、回転または反転を行うことが可

能である。そこで図 7 では、マップ上にプロットされた評価値について、 f_1 に関して縦軸正方向、 f_2 に関して縦軸負方向にそれぞれ値が小さくなるよう、図を回転した結果を示している(ただし MDS は非線形写像であるため、厳密に上述の方向性を設定することは難しく、おおよその傾向をそろえた)。この評価値の方向性は、図 8 により確認することができる。図 8 において、(a) は f_1 の値が、(b) は f_2 の値がそれぞれ小さいほど、評価値およびそれらの遺伝子の色が濃くなるようにカラーリングを行っている。図 7 において点線で囲った領域に生成された子個体について、交叉における親個体を線で結んだものを図 9 に示す。図 9 より、図 7 の点線の領域には、図の上部の領域の個体と下部の個体、点線で囲われた個体と上/下部の領域の個体など、比較的ハミング距離の遠い個体どうしの交叉によって生成された子個体の遺伝子が多く含まれていることがわかる。これらの点のほとんどは薄い白抜きの点であることから、次世代に残されることなく淘汰された子個体であることがわかる。また、遺伝子と評価値との対応関係を見たところ、これらの個体の評

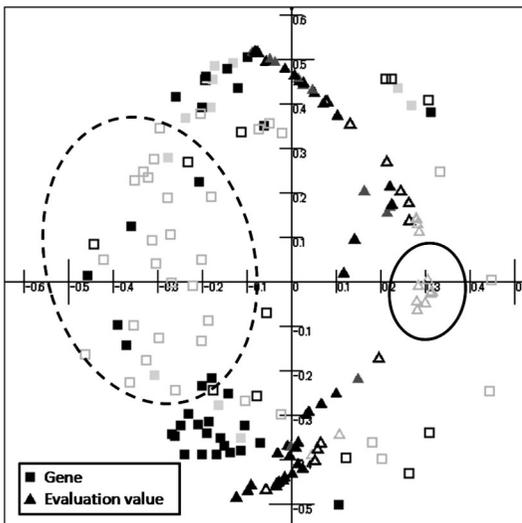
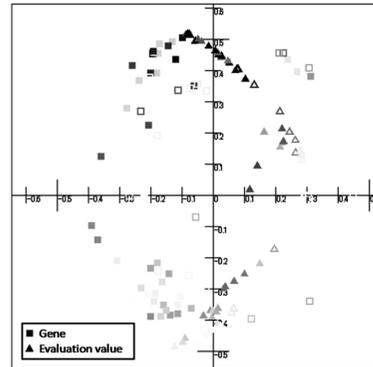
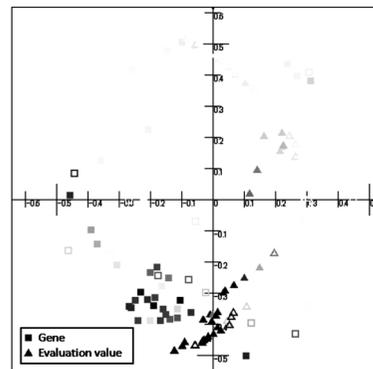


図 7 探索過程の可視化 (10-11 世代)



(a) f_1



(b) f_2

図 8 評価値の大小に基づくカラーリング

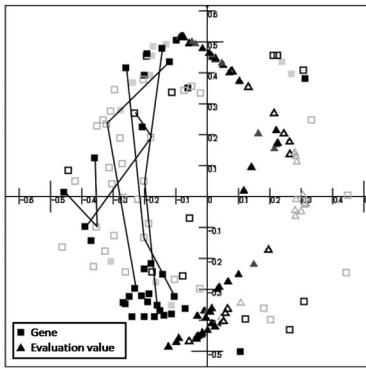


図9 交叉の親個体

価値の多くは、図7の実線で囲った領域に存在していた(実線で囲まれた個体の評価値は多数の点が重なってプロットされている)。図8より、これらの個体は f_1 , f_2 ともに大きい値であることがわかり、劣解の領域にできた子個体であると考えられる。さらにこの領域では、同じ個体の遺伝子と評価値とが互いに離れており、これらの個体の近傍関係は遺伝子空間と評価値空間とで類似していないことが確認できる。

図10に、29~30世代間における個体群の可視化結果を示す。図10において、ハミング距離の遠い個体どうしの交叉によって生成された子個体は、図の中央部に存在していた。図10から、これらの個体は、図7の結果と同様に淘汰されやすいことがわかる。また、これらの個体の遺伝子と評価値についても、図7ほどではないが類似度は低いことがわかった。しかし、マップ上で点線で囲われた領域内では色の薄い黒塗り

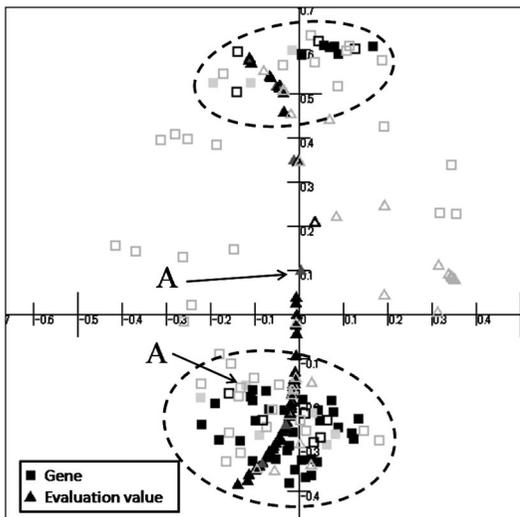


図10 探索過程の可視化 (29-30世代)

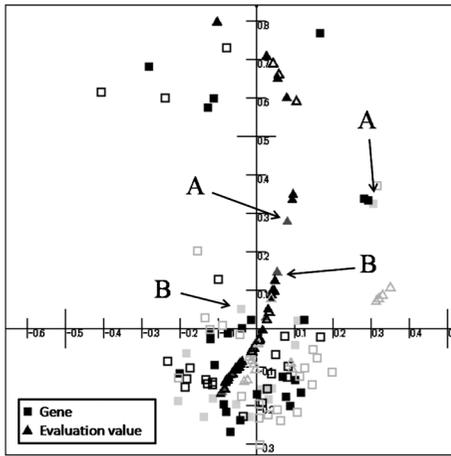
の点が多く残っており、生成された子個体が次世代に残されていることが確認できる。この領域内においては遺伝子と評価値との距離が互いに近く、個体の近傍関係が両空間で類似している領域だとわかった。また図中の個体 A は、評価値空間において疎な領域に生成された子個体だが、遺伝子空間上では別の領域にある親個体の近傍に生成された個体であった。

これらの結果より、FON ベンチマーク関数においては、評価値空間上で疎な領域に個体を生成するためには、ハミング距離の離れた個体どうしを交叉させることで、空間的に補間されることを期待するよりも、ハミング距離の近い個体どうしを積極的に交叉させる近傍探索を繰り返すことで、徐々に疎な領域を埋めていく方法が有効であると推測できる。なお本実験では、2目的最小化問題であることで、直接評価値を可視化可能であるが、4目的以上の多目的問題に提案手法を適用することで、図7、10のように評価値空間で疎/密となる領域を確認できることも、本手法の有効性の一つであると考えられる。

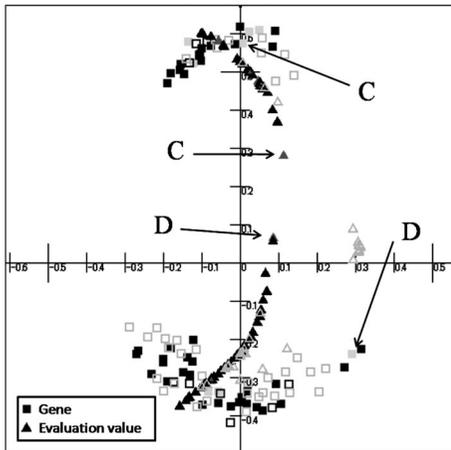
3.4.2 遺伝的演算へのフィードバック

前述のように、FON ベンチマーク関数において多様なパレート解を効率的に求めるためには、近傍探索が有効であると考えられる。そこでここでは、遺伝的演算へのフィードバックとして、交叉する親個体のペアを選ぶ際に、親個体群からランダムにペアを選ぶ方法(以降、通常交叉と呼ぶ)から、遺伝子空間上で探索個体群中のハミング距離の平均より距離が近くなるようにペアを選択する方法(以降、遺伝子近傍交叉と呼ぶ)へと変更した。なお近傍交叉については、先行研究において、様々な多目的関数や実問題の最適化において有効であることが報告されている[9][10][11]。これらの多くは、多目的最適化問題において評価値空間上での近傍の個体を交叉の際のペアとして選択するものであるが、3.4.1で得られた結果に基づき、本実験で用いる遺伝子近傍交叉は、遺伝子空間上での近傍個体を交叉させる方法とする。

図11に、遺伝子近傍交叉によって得られた探索過程のうち、(a)10~11世代間、(b)25~26世代間における個体群の可視化結果を示す。図11より、子個体が親個体の近傍に多く生成されており、またそれら子個体の多くが次世代に残されていることが確認できる。また、図7、10中央部に見られた、遺伝子空間上で親個体からハミング距離が遠く、結果的に淘汰されてしまう個体がほぼなくなっていることがわかる。そして個体 A から D のように、遺伝子空間上で親個体の近傍を探索することによって、評価値空間における疎な領



(a) 10-11 世代



(b) 25-26 世代

図 11 探索過程の可視化 (フィードバック後)

域に広がって生成された子個体も存在した。図11の結果から、遺伝子近傍交叉による探索が、期待する疎な領域のパレート解の獲得に対して有効に働いていることが確認できた。

3.4.3 フィードバック効果の数値的検証

ここでは、通常交叉と遺伝子近傍交叉それぞれの方法で得られたパレート解の比較を行う。パレート解の数値的比較方法として、優越個体割合(The Ratio of Non-dominated Individuals: RNI) [12]と被覆率 [13]のそれぞれ500試行の平均値を用いる。RNIとは、2つの手法により得られた解の優越度合いの比較であり、この割合が大きいほど、得られたパレート解がもう一方のパレート解を優越していることを示している。また被覆率とは、いかに真のパレート最適フロントを隙

表 1 RNI と被覆率の比較

RNI		被覆率	
通常交叉	遺伝子近傍交叉	通常交叉	遺伝子近傍交叉
0.454	0.546	0.750	0.822

間無く得られているかを評価する基準であり、1に近いほど多様で評価値の高いパレート解が得られていることを示している。ただし、被覆率を求めるためには各評価関数の最大値、最小値が必要となるが、FONベンチマーク関数においては各評価値の最大値1.0、最小値0.0は既知であるため、それらの値を使用した。領域分割数は25とし、25領域すべてに解が存在したときの被覆率を1とした。

表1に、30世代目のパレート解の、RNIと被覆率それぞれ500試行の平均値を示す。RNI、被覆率ともに、遺伝子近傍交叉の値が大きく、良好なパレート解が得られていることがわかる。また、RNI、被覆率ともに、平均値に対するt検定により、5%の危険率で有意な差があり、数値的にも、用いた遺伝子近傍交叉の効果があったことが確認できた。

3.5 コーディングの違いによる可視化結果の変化

ここでは、コーディング方法の違いに対する提案手法の適用例として、3.1で示したバイナリコードを実数コードとした場合について、提案手法による可視化結果の比較を行う。図12に、実数コードを用いてFONベンチマーク関数にNSGA-IIを適用した際の、30世代目に得られたパレート解の評価値分布を示す。三角の点は得られたパレート解を、実線はパレート最適フロントを示している。図12より、パレート解がパレート最適フロント上にほぼ隙間なく得られていることがわかる。

図13に、(a)バイナリコード、(b)実数コードそれぞれの場合におけるパレート解の可視化結果を示す。実数コードについては、(1)式における遺伝子空間上

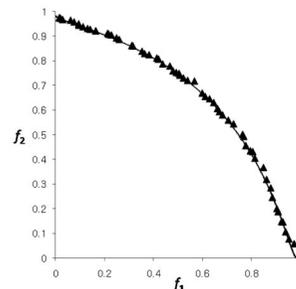


図 12 パレート解の評価値分布 (30 世代目)

での距離 G_{ij} にユークリッド距離を用いた。また、図13における軸や各点の表現は3.4で示したものと同様である(ただしパレート解のみの可視化であるため、親・子個体や淘汰などの情報は付与していない)。さらに、各個体の遺伝子と評価値とが線でつながれており、線の長さがその個体の遺伝子と評価値との非類似度を示している。

図13(a)より、バイナリコードの場合、図の中央部分において遺伝子と評価値とが離れて存在しており、この領域では遺伝子空間と評価値空間とが類似していないことがわかる。また一部の領域では、評価値空間上で近傍でも、遺伝子空間上では大きく離れた個体が存在していることがわかる。一方図13(b)より、実数コードの場合は、遺伝子と評価値とが互いに重なるほど近くに存在していることがわかる。用いたベンチマーク関数の性質上、当然のことではあるが、実数コードを用いることで、遺伝子近傍と評価値近傍とがほぼ対応し、評価値の高い個体の遺伝子空間での周辺探索が、評価値空間における周辺探索につながり、効

率的な遺伝的演算が実現できたと考えられる。

3.6 問題の性質の違いによる可視化結果の変化

前節まで、3.1で示した FON ベンチマーク関数に提案手法を適用してきたが、ここではさらに、単峰性、および多峰性の形状を有する他のベンチマーク問題におけるパレート解の可視化を行う。コーディング方法は、前節において評価値と遺伝子との類似関係が確認された実数コードを用いた。遺伝的演算の方法として、交叉は多親交叉であるシンプレックス交叉 (SPX) [14] を用い、交叉率は0.9とした。また、各遺伝子座 (1変数) の突然変異率は (1/設計変数の数) とし、各変数の範囲内での一様な乱数で値を生成させ、変更する方法を用いた。なお、NSGA-II の集団サイズは50とし、終了世代数は暫定的に100世代とした。遺伝子距離はユークリッド距離を用いた。

3.6.1 単峰性関数 (ZDT1)

解空間が単峰性関数となっているベンチマーク問題として、(5)式に示す Deb らの ZDT1 [15] に提案手法を適用した。この問題では、設計変数 x_i に対して、目的関数 f_1 、 f_2 がともに単峰性になっている。

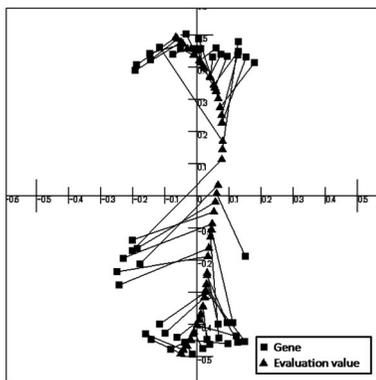
$$\begin{aligned} \min \quad & f_1 = x_1 \\ \min \quad & f_2 = g(x) \left[1 - \sqrt{x_1/g(x)} \right] \end{aligned} \quad (5)$$

$$\begin{aligned} g(x) &= 1 + 9 \left(\sum_{i=2}^N x_i \right) / (N - 1) \\ x_i &\in [0, 1], \quad N = 10 \end{aligned}$$

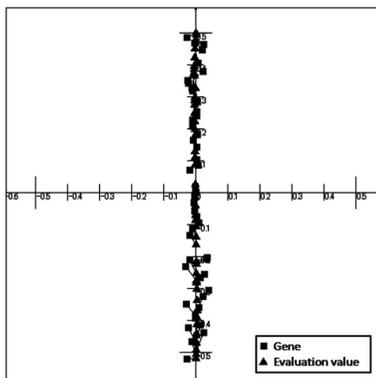
図14に、100世代目におけるパレート解の可視化結果を示す。図14(a)より、100世代目において得られたパレート解が、パレート最適フロントにほぼ近づいていることがわかる。また図14(b)の結果は、図13(b)と同様、個体が一直線上に並び、遺伝子空間と評価値空間とが類似していることを示している。この結果から、単峰性関数を実数コードで探索した場合については、パレート最適フロント付近まで探索が進むことで、各個体の遺伝子近傍と評価値近傍関係とがほぼ類似することが確認できた。

3.6.2 多峰性関数 (ZDT6)

次に、解空間が多峰性関数となっているベンチマーク問題として、同じく Deb らの ZDT6 [15] ((6)式) を用いた。この問題では、設計変数 x_i に対して、目的関数 f_1 、 f_2 がともに多峰性になっている。

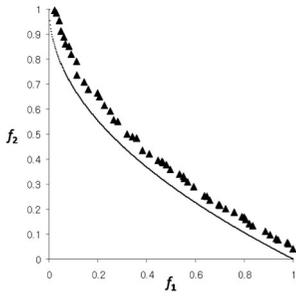


(a) バイナリコード

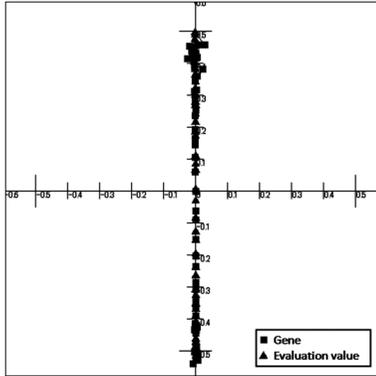


(b) 実数コード

図13 パレート解の可視化 (30世代目)

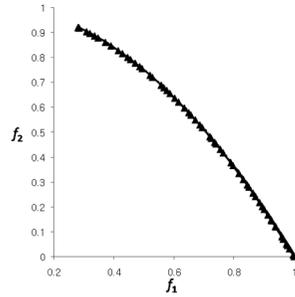


(a) 評価値

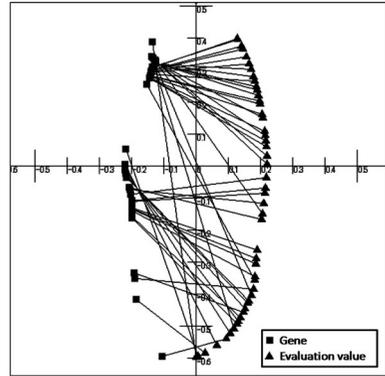


(b) 可視化結果

図14 パレート解の可視化 (ZDT1, 100 世代目)



(a) 評価値



(b) 可視化結果

図15 パレート解の可視化 (ZDT6, 100 世代目)

$$\begin{aligned} \min \quad & f_1 = 1 - \exp(-4x_1) \sin^6(6\pi x_1) \\ \min \quad & f_2 = g(x) \left[1 - (f_1/g(x))^2 \right] \\ & g(x) = 1 + 9 \left[\left(\sum_{i=2}^N x_i \right) / (N - 1) \right]^{0.25} \\ & x_i \in [0, 1], N = 10 \end{aligned} \tag{6}$$

図15に、100世代目の可視化結果を示す。図15(b)より、遺伝子と評価値とが離れ、それらの近傍関係があまり類似していないものの、図15(a)より、得られたパレート解がパレート最適フロント上で広く分布していることがわかる。また図15(b)から、特にパレート最適フロント端の部分において、遺伝子空間で近い位置にある個体が、大きく異なる評価値を持っていることがわかる。結果としてこの関数においては、遺伝子空間における周辺探索が評価値空間における多様な解の探索につながり、図15(a)のような評価値分布となっていたことが推測される。逆に狭い評価値空間領域が広い遺伝子領域に対応する解空間形状の場合には、評価値のみを基準とした探索では探索効率が低下し、3.3.2や3.4.2で示したようなフィードバックが有効となると考えられる。

4. おわりに

本論文では、GAにおける遺伝子空間と評価値空間の類似関係を把握することを目的に、遺伝子と評価値の対応関係の可視化手法を提案した。提案手法では、GAの探索過程における各個体について、遺伝子と評価値の非類似度を定義し、遺伝子空間での距離関係、評価値空間での距離関係および非類似度を同一空間内にマッピングすることで、それらの視覚的な把握が可能となる。本論文では、NSGA-IIを多目的ベンチマーク関数に適用した際の個体群に提案手法を用い、得られた可視化結果を示し、考察を行った。探索過程の特徴をとらえ、遺伝的演算へのフィードバックを行うことで、効率的に多様なパレート解を得ることが可能となることを示した。さらに、コーディングの違い、および問題の性質の違いによる可視化結果の変化も調べ、可視化結果においてそれぞれの特徴を把握可能であることを示した。

今後の課題として、コーディングの妥当性を評価し、適切なコーディング法の指針を得るための手法として提案手法を位置づけていくこと、および問題の性質とそのフィードバック方法に対する体系化について

検討するとともに、多数の目的関数を有する看護師スケジューリング問題(Nurse Scheduling Problems: NSP)[16]などの実問題に提案手法を適用していく予定である。

謝辞

本研究の一部は、平成20年度文部科学省科学研究費(若手研究(B)No.20700082)の補助を受けて遂行された。

参考文献

- [1] 坂和正敏, 田中雅博, “遺伝的アルゴリズム”, 日本ファジィ学会編, ソフトコンピューティングシリーズ, 朝倉書店, 1995
- [2] 佐藤浩, 小野功, 小重信, “遺伝的アルゴリズムにおける世代交代モデルの提案と評価”, 人工知能学会誌, Vol.12, No.5, pp.734-744, 1996
- [3] 廣安知之, 三木光範, 佐野正樹, 谷村勇輔, “2 個体分散遺伝的アルゴリズム”, 情報処理学会研究報告・MPS, Vol.2000, No.85, pp.37-40, 2000
- [4] Norimasa Hayashida, Hideyuki Takagi, “Visualized IEC: Interactive Evolutionary Computation with Multidimensional Data Visualization,” IEEE International Conference on Industrial Electronics, Control and Instrumentation (IECON2000), Vol.4, pp.2738-2743, 2000
- [5] T.Kohonen, “Self-Organizing Maps,” Springer, 2000
- [6] John W.Sammon, Jr, “A Nonlinear Mapping for Data Structure Analysis,” IEEE Transactions on Computers, Vol.C-18, No.5, pp.401-409, 1969
- [7] Kalyanmoy Deb, Amrit Paratap, Sameer Agarwal, and T. Meyarivan, “A Fast and Elitist Multi-Objective Genetic Algorithm: NSGA-II,” IEEE Transactions on Evolutionary Computation, Indian Institute of Technology, Vol.6, No.2, pp.182-197, 2002
- [8] 石黒英敬, 吉川大弘, 古橋武, “GA における遺伝子-評価値関係の可視化と探索へのフィードバックに関する検討”, 日本知能情報ファジィ学会第17回インテリジェント・システム・シンポジウム講演論文集, pp.303-306, 2007
- [9] 渡邊真也, “近傍個体の交叉に基づく多目的遺伝的アルゴリズムとその応用に関する研究”, 人工知能学会誌, Vol.19, No.1, p.103, 2001
- [10] 渡邊真也, 慶安知之, 三木光範, “近傍培養型遺伝的アルゴリズムによる多目的最適化”, 情報処理学会論文誌, 数理モデル化と応用, Vol. 43, No. SIG10 (TOM7), pp.183-198, 2002
- [11] H. Ishibuchi, Y. Shibata, “Mating Scheme for Controlling the Diversity-Convergence Balance for Multiobjective Optimization”, Proc. of 2004 Genetic and Evolutionary Computation Conference, part. I, pp.259-1271, 2004
- [12] 廣安知之, 三木光範, 渡邊真也, 迫田岳志, 上浦二郎, “多目的遺伝的アルゴリズムにおける各手法の比較”, 同志社大学理工学研究報告, Vol.43, No.1, pp.41-52, 2001
- [13] 廣安知之, 三木光範, 渡邊真也, 畠中一幸, “多目的分散遺伝的アルゴリズムにおけるシェアリング, 収束判定, 及び解の評価手法の検討”, 同志社大学理工学研究報告, Vol.40, No.4, pp.189-200, 1999
- [14] 樋口隆英, 筒井茂義, 山村雅幸, “実数値 GA におけるシンプレックス交叉の提案”, 人工知能学会誌, Vol.16, No.1, pp.147-155, 2001
- [15] E. Zitzler, K. Deb, L. Thiele, “Comparison of Multiobjective Evolutionary Algorithms: Empirical Results,” Evolutionary Computation, Vol.8, No.2, pp.173-195, 2000
- [16] 上野尚樹, 古橋武, “看護師長の作成手法を取り入れた看護師勤務表作成支援システム”, 日本知能情報ファジィ学会第14回インテリジェント・システム・シンポジウム講演論文集, pp.174-177, 2004
(2008年10月16日 受付)
(2009年3月4日 採録)

[問い合わせ先]

〒464-8603 愛知県名古屋市中千種区不老町
名古屋大学大学院工学研究科計算理工学専攻
吉川 大弘
TEL: 052-789-3167 3167(内線)
FAX: 052-789-3166
E-mail: yoshikawa@cse.nagoya-u.ac.jp

著者紹介



いしぐろ ひでたか
石黒 英敬 [非会員]

2007年名古屋大学工学部電気電子・情報工学科卒業。2009年名古屋大学大学院工学研究科博士前期課程計算理工学専攻修了。同年三菱電機株式会社入社。在学中は主として多目的遺伝的アルゴリズム(MOGA)における遺伝子-評価値関係の可視化に関する研究に従事。



よしかわ ともひろ
吉川 大弘 [正会員]

1997年名古屋大学大学院博士課程修了。同年カリフォルニア大学パークレー校ソフトコンピューティング研究所客員研究員。1998年三重大学工学部助手。2005年名古屋大学大学院工学研究科COE特任准教授。2006年10月同研究科准教授。現在に至る。主としてソフトコンピューティングとその応用に関する研究に従事。博士(工学)



ふるはし たけし
古橋 武 [正会員]

昭和60年名古屋大学大学院工学研究科博士後期課程電気系専攻修了。工学博士。平成16年名古屋大学大学院工学研究科計算理工学専攻教授。現在に至る。ソフトコンピューティング、感性工学に関する研究に従事。平成8年日本ファジィ学会論文賞受賞。IEEE、日本知能情報ファジィ学会、電気学会等の会員

Visualization of Relationships between Genes and Evaluation Values in GA and Feedback into Genetic Operations

by

Hidetaka ISHIGURO, Tomohiro YOSHIKAWA and Takeshi FURUHASHI

Abstract :

GA searches neighborhood of the genes that have high evaluation values as it is assumed that neighboring relationship of individuals in the genotype space is similar to that in the evaluation space. However, it is difficult to decide which coding, deciding the similarity between them, and which genetic operation to use because there is no criterion for the similarity and the landscape of search space. As a result, they are decided by trial and error in a lot of cases. This paper proposes the visualization method to grasp the relationship between genes and their evaluation values for efficient search. This paper applies the proposed method to a benchmark function of multi-objective optimization problem and shows that it enables us to grasp the similarity of genes between in the genotype space and evaluation space. It also shows that we could feed back the visualization result to genetic operations for more efficient search.

Keywords : Visualization, Genetic Algorithm, Genotype Space, Evaluation Space

Contact Address : **Tomohiro YOSHIKAWA**

*Dept. of Computational Science and Engineering, Graduate School of Engineering, Nagoya University
Furo-cho, Chikusa-ku, Nagoya, 464-8603, JAPAN*

TEL : 052-789-3167 (3167)

FAX : 052-789-3166

E-mail : yoshikawa@cse.nagoya-u.ac.jp