

人間の嗜好における判断要素間のトレードオフ関係の検証

小林 祐介, 廣安 知之, 三木 光範

1 はじめに

人間の感性を扱う研究の一つに、対話型遺伝的アルゴリズム (iGA) がある。iGA は、提示される複数の個体に対しユーザの嗜好を基に評価を与え、その評価を用いて解探索を行う最適化手法である。この際、従来の iGA では、総合的な判断基準によって評価を与えているが、その総合的な判断基準の中には、複数の判断要素を有すると考えられる。例えば、人が T シャツのデザインを選択する際、「好きなもの」や、「気に入ったもの」を選択している。しかし、その「好き」や「気に入った」という大まかな嗜好には「かわいい」や「さわやか」などの判断要素を複数有すると考えられる。それらの判断要素間には、「ある判断要素を改善すると、他の判断要素が改悪する関係」であるトレードオフ関係にあることが多いと考えられる。そこで、その「かわいい」や「さわやか」等の複数の判断要素のそれぞれを 1 つのサブ目的とみなす事により、単一目的最適化問題を多目的最適化問題に拡張する。これにより、判断要素間のトレードオフ関係を検証する事が可能となる。

本研究では、iGA による嗜好から、各判断要素を抽出し、その判断要素間のトレードオフ関係を検証するための基礎研究として、iGA に多木低遺伝的アルゴリズム (MOGA) を適用した多目的対話型遺伝的アルゴリズム (MOiGA) を用いて判断要素間のトレードオフ関係の検証を行う。

2 多目的対話型遺伝的アルゴリズム

2.1 対話型遺伝的アルゴリズム

対話型進化計算法 (Interactive Evolutionary Computation: IEC) の一つである iGA は、遺伝的アルゴリズム (GA) の遺伝操作のうち「評価」の部分人間が行う事によって解の探索を行う。一般的に iGA では単一目的最適化問題を扱うが、その際ユーザがシステムによって提示された個体に与える評価は、ユーザの嗜好を反映したものである。そのため、従来の GA に比べ、人の感性という複雑な構造の解析により適しているといわれている。この単一目的最適化問題を対象とした iGA に多目的最適化問題を対象とした MOGA の手法を適用した手法が MOiGA である。

2.2 多目的最適化問題

多目的最適化問題とは、複数の評価基準のもとで最適解を求める問題である。しかし、一般的に多目的最適化問題では、トレードオフ関係にある事が多く、そのような場合には唯一つの最適解は存在しない。そのため、多目的最適化では、「パレート最適解」という概念を用いて

探索を行う。

パレート最適解とは、「ある目的関数値を改善するためには、少なくとも他の 1 つの目的関数値を改悪せざるを得ないような解」と定義されている。一般にパレート最適解は複数存在することが多く、目的関数間に存在するトレードオフ関係を知る上でも、パレート最適解を数多く求める事が重要となる²⁾。Fig. 1 は 2 目的最小化問題におけるパレート最適解の例である。

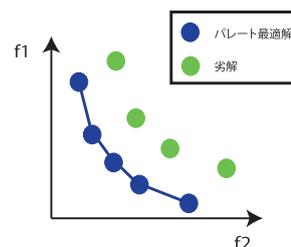


Fig.1 パレート最適解 (出典：自作)

近年、GA を多目的最適化問題に適用した、MOGA に関する研究が数多く行われている。その理由として、GA が多点探索であり、一度の探索でパレート解集合を求められることがあげられる²⁾。

2.3 多目的対話型遺伝的アルゴリズム

MOiGA は、iGA に MOGA 手法を適用したアルゴリズムである。iGA での評価は、ユーザの嗜好の総合的な評価であるが、MOiGA では、総合的な評価に含まれる複数の判断要素に対して評価を付ける。例えば、T シャツのデザイン問題について考えると、T シャツのデザインの評価とは「かわいさ」や「さわやかさ」などの判断要素についてユーザが総合的に判断したものである。これらの各判断要素を最適化の目的として捉えることで、単一目的最適化問題を多目的最適化問題へと拡張することが可能である。これより、iGA に MOGA 手法を適用することで、ユーザの嗜好を各判断要素毎に反映し、各判断要素間のトレードオフ関係について検証を行う事が可能である。

MOiGA と通常の iGA の大きな違いとして、目的数の増加による個体の適合度値の扱いがある。iGA の場合にはユーザが与える評価値を直接適合度値として扱うことが可能であるが、多目的 iGA においては複数の目的に対する評価値を有するため、直接適合度値として扱う事が困難である。そこで、本研究では非優越ソートによって各個体をランク付けし、そのランクを適合度値として扱う。

また、iGA におけるエリート戦略と同等な仕組みとして、MOiGA ではアーカイブを用いて探索途中における非劣解を保存している。保存される解はランクと混雑距

離を考慮して選択されるため、精度と多様性に優れた解を次世代に残し、探索する事が可能となる。

3 システムを用いたテスト実験

3.1 対象問題概要

本実験では、人間の感性間のトレードオフ関係を検証するため、簡単なテスト問題である、2色旗を用いて実験を行った。また、トレードオフ関係のはっきりしている目的として、「暑そうな国の国旗」と「寒そうな国の国旗」という2目的を用い、2目的最大化問題として扱う。初期個体12個体をランダムに生成し、Fig. 2に示したインタフェースによりユーザに提示する。各世代毎にこのインタフェースを更新し、新たな個体を提示する。

第1世代目

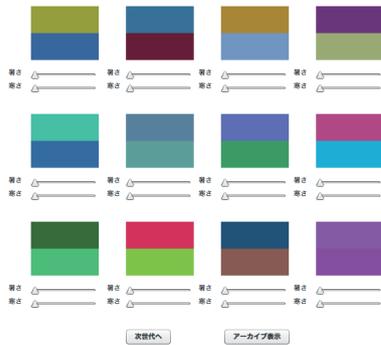


Fig.2 初期生成インタフェース (出典：自作)

また、Fig. 3は提示個体の例である。

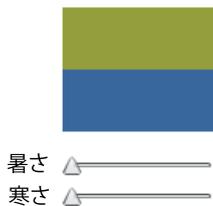


Fig.3 提示個体の例 (出典：自作)

本実験では、一つの個体に対し、目的別に評価を行うため、スライダーを用いている。この際、Fig. 3に示している各個体の下に配置されるスライダーを用いて、1~100までの評価値を付けることが可能である。これらの各評価を考慮したランク付けを行い、ランクを適合度値とする。

設計変数は旗の上下の色としており、色の表現には、色相、彩度、明度を扱うHSB表色系を用いた。また、アーカイブ保存では、適合度値の良い個体から順に12個体保存する。もし、適合度値が同じ個体の場合は混雑距離の大きいものを優先し、保存する。実験終了条件を10世代目の評価と設定した。また、実装において、iGAに一般的なMOGAの手法であるNSGA-II¹⁾を適用した。

3.2 問題点

3.1節に記したアルゴリズムにおいて、以下のような問題点が存在する。

● アーカイブ内の個体の更新

3.1節に記したアルゴリズムにおいて、アーカイブに保存される個体は、適合度が良く、混雑距離が大きいものである。しかし、この手法では、探索初期の段階で良い個体に対し最高の評価値が付けられた場合、後の世代でさらに良い個体を提示しても、探索初期の段階で付けられた最高の評価値より良い評価値を付ける事ができず、探索初期の個体より良いはずの後の世代の個体の評価値が、探索初期段階の個体の評価値より悪い評価値になってしまう (Fig. 4)。

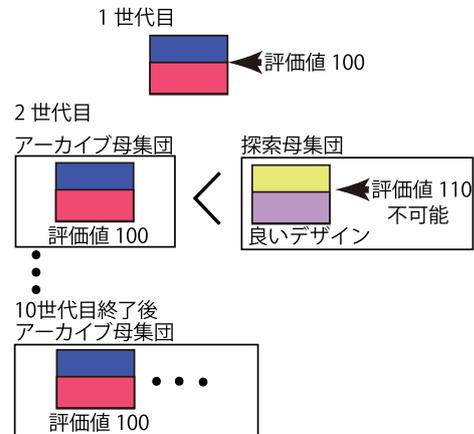


Fig.4 アーカイブ内の個体の更新における問題の概念図 (出典：自作)

Fig. 4に示すように、例えば1世代目で最高の評価値である100を付けられた個体は、2世代目にその個体よりも良い個体も出現しても、1世代目に付けられた最高の評価値である100よりも良い評価値を付ける事ができないため、1世代目の個体が最後までアーカイブに保存されたままになってしまう。つまり、ユーザの評価が、その世代の評価である相対的な評価ではなく、探索全体を考慮に入れた絶対的な評価として扱われている事が問題点である。

● 評価手法

3.1節に記したアルゴリズムでは、目的別にスライダーを用いて評価を行っている。この手法では、12個体に対しそれぞれ目的数である2回の評価を行うため、10世代終了までに総計240回の評価を行う必要がある。これでは、ユーザの疲労度が考慮されていないと考えられる。

3.3 改良点

3.2節で述べたようにNSGA-IIをそのままiGAに適用するだけでは、問題が存在する。そこで、問題点を解決するため、アルゴリズムの以下の改良を行った。

● アーカイブ内の個体の評価値の更新

アーカイブ内の個体の更新を行うため、絶対評価として扱われていた評価値を相対評価として扱う必要がある。その手法として、各目的における評価値のスケーリングを行う。以下にアルゴリズムの流れ

を示す。

1. アーカイブに保存された個体から、各目的間数値が最大の個体を抽出する。
2. 抽出された個体を次世代に提示する。
3. その個体の評価値が変更された場合、アーカイブ内の全個体の各評価値に対し、スケーリングを行う。

以上の1~3の処理を繰り返す事により、アーカイブ内の個体の評価値の更新が行われ、絶対評価として扱われていた評価値を相対評価として扱う事が可能となる。スケーリングを表す図を Fig. 5 に示す。

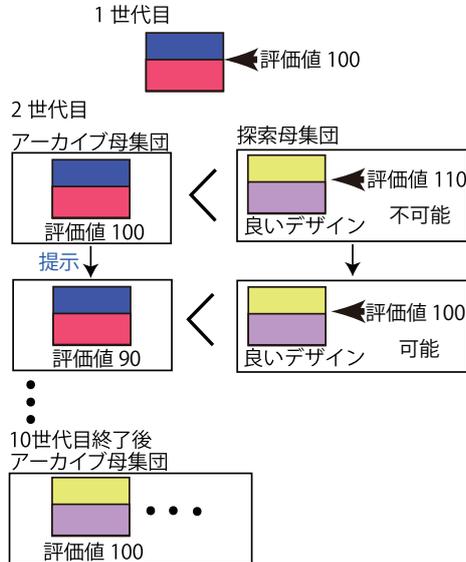


Fig.5 スケーリングの概念図 (出典：自作)

例えば、Fig. 5 に示すように、1 世代目に最高の評価値である 100 が個体に付けられても、次の世代にその個体を提示し、評価値を変更可能にすることで、他に良い個体が出現しても、良い個体に対して、1 世代目に評価値 100 を付けた個体よりも良い評価値を付ける事が可能となる。また、その個体の 1 世代目に付けられた評価値と、2 世代目に付けられた評価値を利用して、1 世代目に保存されたアーカイブ内の個体に対して、スケーリングを行う事が可能であると考えられる。つまり、次世代に提示された各目的間数値が最大の個体の評価値が変更されれば、アーカイブに保存された他の個体の評価値も、変更された目的に対する評価値が変更される。このスケーリングを行う事で、絶対評価として扱われていた評価値を相対評価として扱う事が可能となる。

● 評価手法

ユーザの疲労度軽減のため、評価回数を減らし、簡単に評価できる手法を提案する。それは、ドラッグ&ドロップによる評価手法である。改良表評価手法のインタフェースを Fig. 6 a), b) に示す。

- また評価の手順を以下に記す。
1. 提示されている個体のグラフィックをドラッグする。
 2. インターフェース上の目的関数空間を表す平面

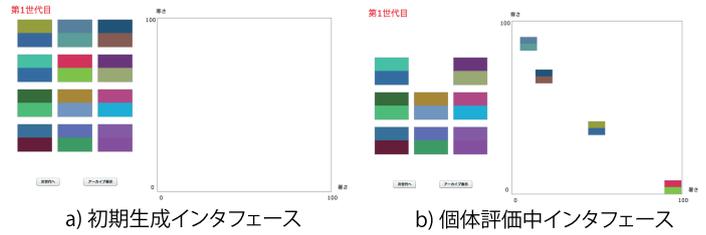


Fig.6 評価手法改良後のインタフェース (出典：自作)

空間にドロップを行う。

3. ドロップされた個体の座標により、各目的に対する評価値が決まる。

この操作により、従来の手法より評価回数を 120 回程度の約半分の回数で行う事ができる。また、スライダーを用いた評価では、評価を付けづらく時間がかかるという事があり、ドラッグ&ドロップの方が操作がやりやすいと考えた。これはユーザの疲労度の軽減に繋がると思われる。

3.4 被験者実験

3.3 節までに説明した MOiGA の被験者実験を行った。今回の実験では、改良前のシステムとスケーリング手法を実装したシステムとドラッグ&ドロップを実装したシステムの比較を行った。どのシステムも、初期個体を同一のデザインとし、終了条件を 10 世代目の評価終了後と設定した。

Table1 に、実験に用いたパラメータを記す。

Table1 実験に用いたパラメータ

Parameter	Value
Generations	10
Populations	12
Archive size	12
Way of Selection	Crowded Tournament
Tournament size	2
Way of Crossover	BLX- α
Crossover Rate	1.0
Mutation Rate	1.0/Chromosome Length

実験手順としては、年齢 21~24 歳の被験者 6 名 (男：4, 女：2) に各システムの実験を行っていただき、一つのシステムの実験が終わり次第、アンケートを実施した。また最後に、システムの比較に関するアンケートを実施した。

3.5 被験者実験における結果と考察

まず、改良前のシステムとスケーリング手法を実装したシステムの比較実験を行った。この比較では、以下の事を検証する。

1. 「暑そうな国の国旗」という目的に関して、目的に沿った旗が作成できたかどうか。

- 「寒そうな国の国旗」という目的に対して、目的に沿った旗が作成できたかどうか。
- 「暑そうな国の国旗」と「寒そうな国の国旗」の両目的に沿った旗が作成できたかどうか。

以上の3つの項目に対して、アンケートを行った。システムの比較アンケートの結果を Fig. 7, Fig. 8, Fig. 9 に示す。

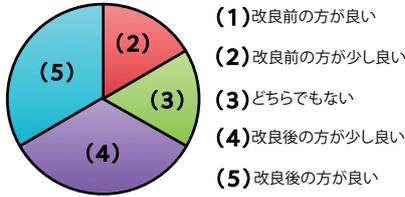


Fig.7 改良前システムと改良後システムの比較 (暑そうな国の国旗)(出典：自作)

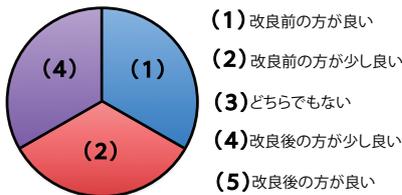


Fig.8 改良前システムと改良後システムの比較 (寒そうな国の国旗)(出典：自作)

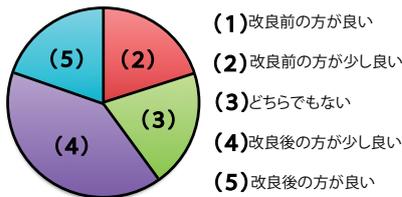


Fig.9 改良前システムと改良後システムの比較 (「暑そうな国の国旗」と「寒そうな国の国旗」の両目的)(出典：自作)

以上の結果より、改良前システムとスケーリング手法を適用したシステムは、ほとんど同等な結果を得る結果となった。これは、被験者の目的に対するイメージが薄く、どのような国旗に対してどのように評価するかが曖昧であったためだと考えられる。

次に、スケーリング手法を適用したシステムとスケーリング手法と評価手法にドラッグ&ドロップを適用したシステムを比較する。ドラッグ&ドロップを適用したシステムは疲労度を考慮したシステムであるため、この比較では、被験者の疲労度に関して比較を行った。アンケートの結果を Fig. 10 に示す。

この結果より、ドラッグ&ドロップを評価手法として適用したシステムの方が、被験者の疲労度が大きくなりにくく、疲労度を考慮した評価手法であると考えられる。

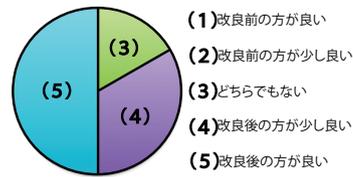


Fig.10 改良前システムと改良後システムの疲労度の比較 (出典：自作)

4 まとめ

今回、MOGA を iGA に適用したが、アーカイブに保存される個体の更新、ユーザに対する疲労度の考慮などの問題点が存在したため、その問題点を改良するアルゴリズムを実装した。また、問題点を改良できているのかどうかのテストを行うためシステムを用いたテスト実験を行った。完全に競合する目的である「暑そうな国の国旗」と「寒そうな国の国旗」という2目的最大化問題をテスト対象問題として、実験を行った。結果として、スケーリング手法を適用したシステムには、まだ問題点が存在し、検討する必要がある事がわかった。ドラッグ&ドロップを用いた評価手法は、ユーザに対して疲労度の軽減に繋がる事がわかった。

5 今後の展望

今回は、被験者数が十分で無かったため、次回は被験者数を増やし、実験を行う必要がある。また、完全に競合する目的間でのトレードオフ関係を検証したものであったが、競合しない目的での被験者実験を行い、トレードオフ関係の検証を行う。また、対象問題が2色旗では、目的に関してイメージをしにくいという意見があったため、対象問題をよりイメージし易い対象問題に変更する必要があると考えられる。また、提示個体、アーカイブに保存される個体における設計変数空間での多様性についても検討する必要があると考えられる。

参考文献

- K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," IEEE Trans. on Evolutionary Computation, 6 (2), pp. 182-197, 2002.
- 渡邊真也, "遺伝的アルゴリズムによる多目的最適化に関する研究", 同志社大学大学院工学研究科知識工学専攻博士論文, 2003