

対話型遺伝的アルゴリズムにおける 表現型空間の自動生成手法の提案

The phenotype space automatic extraction method
for interactive Genetic Algorithms

要約

本稿では、対話型遺伝的アルゴリズム (interactive Genetic Algorithm: iGA) の解候補が既に存在する対象問題において、解候補の表現型が分布する空間 (表現型空間) を自動生成する手法を提案した。提案手法では、既存の解候補間の関連度を用いることで、iGA システムの開発者側が各解候補へのラベル付けなどの作業を行うことなく、自動的に iGA 探索の探索に必要な空間を構築する。解候補の関連度の算出には、Web 上に蓄積された嗜好情報を利用する。現在、ショッピングサイトやソーシャルブックマークなど Web 上のサービスには、多数のユーザの行動履歴が集積されており、その中には感性や嗜好に基づく情報が含まれている。本研究では、これらの情報を集合嗜好と呼ぶ。ユーザはこの集合嗜好から生成された空間内で、iGA によって最適解の探索を行う。提案手法の有効性を検証するため、ユーザの購買行動を反映しているとされる商品間の関連度をショッピングサイトから集合嗜好として取得し、表現型空間を生成した。生成した空間の特徴を検証することで、空間の自動生成が可能であることを確認した。また、ショッピングサイトを模擬した iGA システムを構築して被験者実験を行い、表現型空間から生成した設計変数空間上において、被験者が自身の嗜好に基づいた探索を行うことを確認した。

キーワード : 対話型進化計算手法, 対話型遺伝的アルゴリズム, モデリング, 主成分分析

Abstract

The paper proposes the automatic generation method of the phenotype-mapping space for the applications of iGA (interactive Genetic Algorithm). The proposed method constructs the search space for iGA from the degree of association between the solutions. The iGA system developers do not need to bear a burden of labeling each solution with the proposed method. The degree of association between the solutions is calculated from the preference information accumulated on the Web. Recently, online services such as shopping-sites and social-bookmark store a lot of users' action logs. These informations include the users' Kansei and preference. We call these types of informations collective preference. The users find their preference by iGA search in the space generated from the collective preference. To verify the effectiveness of proposed method, we obtained the degree of association between the products from the online shopping sites and made the phenotype-mapping space. This relationship reflects the users' buying behavior. It was confirmed that the proposed method is able to generate the space automatically by the analysis of the generated space. Furthermore, we performed the subjective experiments with the iGA system which simulated a shopping site. From the experimental result, it was verified that the subjects' search log in the generated space was personalized by their own preference.

Key words : interactive evolutionary computation, interactive genetic algorithm, modeling, principal component analysis

1 はじめに

対話型遺伝的アルゴリズム (interactive Genetic Algorithm: iGA) ^{1, 2, 3, 4)} は遺伝的アルゴリズム ⁵⁾ をベースとし、人間の嗜好、感性に基づいた最適化を行う手法である。iGA は最適化の対象となる解候補をユーザに複数提示し、それらを評価させ、評価値が高い解候補から新たな解候補を生成するという手続きを繰り返す。これにより、ユーザの感性に応じた解候補を探索することができる。これまでにも画像の感性検索 ^{6, 7)} や服飾デザイン ^{6, 8, 9, 10, 11)}、補聴器フィッティング ¹²⁾ など、人間の嗜好や直感的な判断が重視される問題に応用されてきた。

GA の探索の特徴に、生物のもつ遺伝子と形質との関係を探索に利用していることがあげられる。すなわち、GA の探索においては、解候補の遺伝子が各探索点として扱われ、これは遺伝子型と呼ばれる。一方で、遺伝子型が形質として表現されたものは表現型と呼ばれ、各解候補の評価はこの表現型に対して行われる。すなわち、問題が定式化された表現型の空間ではなく、別の遺伝子型の空間に解候補をマッピングして探索は行われ、評価の際は元の表現型の空間に戻して評価が行われる。よって、GA では遺伝子型と表現型の設計をうまく行うことで、効率良く解探索を行うことができる。

対象問題を iGA による最適化問題として定式化する場合、通常、表現型は明確である。一方で、探索を行うための設計変数の定義は、問題に依存するため困難である。例えば、服飾デザインを対象問題とした iGA システムを構築する場合、表現型は被服のデザインであり、生地の色や襟の形、袖の長さなどが設計変数として考えられる。しかし、長さや色といった設計変数は変数としての定量化が容易であるが、形や質感などは定量化に統計的なアンケートを必要とするなど、コストがかかる。また、靴やカバンなどに対象問題が変更されると、異なる設計変数を持つ最適化問題として定式化しなければならない。iGA では、設計変数が変更されるとシステムの大部分を変更する必要が生じる。iGA を適用させた多くの応用研究において、汎用的なシステムの構築が難しいのはこの理由による。

本研究では、Web 上のショッピングサイトにおいて、ユーザの好みを商品提示に反映させるシステムを iGA を用いて構築することを考える。しかし、先に述べたように対象問題ごとに設計変数が異なるため、対象となる商品毎に最適化問題を定式化し直し、その設計変数に応じてシステムを個別に構築する必要がある。これは、非常に大きなコストがかかり、現実的に

このようなシステムにおいて iGA を導入することを困難にしている。

そのため、対象問題に応じて設計変数を自動的に定義することが可能であれば、ユーザの感性や直感などを取り扱う多くの最適化問題で iGA を適用することが容易となる。本研究では、解候補の設計変数を定義するために、解候補の表現型が分布する空間を自動的に構築する方法を提案する。提案手法が適用される対象問題には、解候補が既にデータとして存在し、これらの解候補間の関連度も取得可能であれば、特定の対象問題に限らない。解候補間の関連度は、Web 上に蓄積された多数のユーザの嗜好情報から取得し、本研究ではこれらを集合嗜好と呼ぶ。例えば、ショッピングサイトを例とした場合、商品の 1 つ 1 つが表現型である。また、多くのショッピングサイトでは商品の説明やレビュー、タグ情報、カテゴリ分類、商品の推薦関係などの情報を公開している。これらの中にはユーザの嗜好や感性を反映しているものも含まれると考えられ、商品の関連度を算出する材料となる。集合嗜好によって生成された表現型の分布する空間は、人間の感性を反映しており、この空間から算出された各解候補の遺伝子型を用いることで、iGA による探索が有効に行われると考えられる。

本論文では、提案手法を説明した後、提案手法の有効性を確認するために、ショッピングサイトから取得した商品間の関連度を用いて iGA の探索空間を生成し、検討した。また、生成した表現型空間における設計変数の検討と iGA による被験者実験を行い、被験者の探索履歴について検証した。

2 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の概要

対話型遺伝的アルゴリズムは遺伝的アルゴリズムをベースとし、その評価系に人間の感性、嗜好を取り入れた対話型進化計算手法の一つである。GA は探索母集団中の複数の個体によって多点探索を行う最適化手法であり、iGA においても複数の個体がインタフェースを通じてユーザに提示され、ユーザは各個体に対して評価値を与える。その評価値を基に選択や交叉といった遺伝的操作を繰り返すことで、ユーザに提示する個体をよりユーザの感性、嗜好にフィットした組み合わせへと進化させる。

2.2 アルゴリズムと個体の表現

図 1 に iGA のアルゴリズムの流れを示す。遺伝的操作を行う際、各個体は遺伝子型と呼ばれる $\{0, 1\}$ のビットストリング ⁵⁾ や実数値 ^{13, 14, 15)} の配列によって表現される。一方で、評価部において各個体がとる

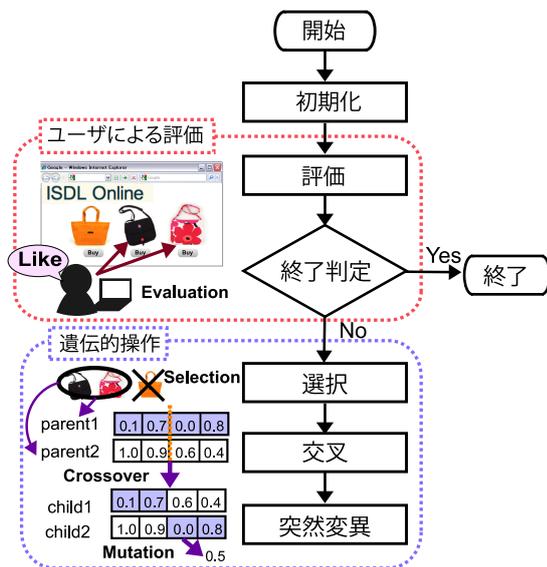


図1 iGAの流れ

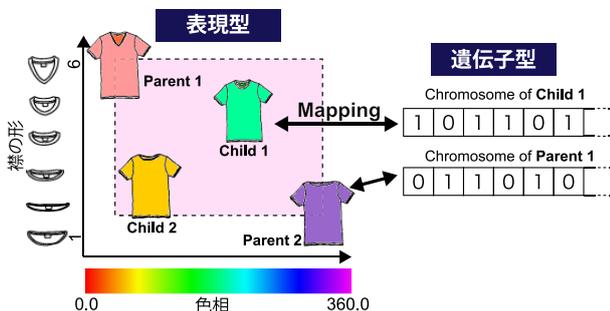


図2 Tシャツデザインシステムにおける遺伝子型と表現型の写像関係

表現を表現型と言い、GAではパラメータとして目的関数に入力され、iGAではユーザーが評価を行うデザインや音としてインタフェースを通して提示される。

図2にTシャツデザインシステムを模擬したiGAシステム¹⁶⁾における、遺伝子型と表現型の写像関係を示す。このシステムでは、襟の形と生地の色といった2つの特徴によって、1枚のTシャツが定義されるものとする。この特徴を数値化した設計変数から、ユーザーに提示される表現型のTシャツ画像が生成され、そして、遺伝的操作を行うための遺伝子型がコード化される。

2.3 対話型遺伝的アルゴリズムの課題

GAでは、交叉によって親個体の遺伝子型から子個体の遺伝子型を生成する。このとき、親と極端に異なる子個体が生まれるとランダムサーチと変わらないため、子個体は親個体の形質を受け継ぐか、もしくは平均的な形質をとるべきである。特に対話型遺伝的アル

ゴリズムでは、ユーザーが個体を評価するため、人間の直感に従う適切な子個体が生成されないと、ユーザーが違和感を覚え、満足度を低下させることになる。

よって、iGAでは遺伝子型を定義するにあたって、まず人間の感性から見て適切な近傍を持つように表現型が分布する設計変数空間を構築する必要がある。そして、この設計変数の軸に従って子個体を生成するように、遺伝子型および遺伝的操作を決定することで、親の形質を反映した、直感的に妥当な子個体が生成されるようになる。

しかし、システム実装者が実際に設計変数空間を構築する場合、以下のような課題が存在する。

1. 有力な設計変数の決定

個体を表現する各種のパラメータからGAによって最適化する設計変数を特定する。例えば、服飾デザインを対象問題とした場合、生地の色、模様や質感、袖の長さ、襟の形、ボタンの色や形など、衣類を表現する各種パラメータの中から目的に沿ったパラメータを選定する必要がある。

2. 各設計変数の定量化

色や長さといった要素は、数値的に表現しやすく、また人間の感性をある程度反映した指標が存在する。例えば、色であれば、RGB表色系などによって定量的な表現が可能である。しかし、模様や形などの要素は、汎用的な尺度がない場合が多く、その尺度構成には専門家の判断や、大量のユーザーアンケート等が必要となる。

3. 対象問題毎のモデリング

1, 2の課題は対象問題毎に解決される必要がある。例えば、衣類のデザインiGAシステムを構築する際に定義した設計変数を用いて、椅子のデザインiGAシステムを構築することはできない。それぞれの問題領域によって、有意な設計変数は異なる。

特に3の問題は、iGAにおける汎用的なシステムの構築を難しくしている。

また、商品推薦や感性検索などの場合、探索すべき候補が商品などの形で既に存在している。そのため、設計変数の決定、定量化を行った後、各候補の設計変数値を測定する必要がある。これは、模様や手触りといった予め電子化された設計変数の値でなかった場合、対象となる候補の数が多い程、値を測定するシステム実装者にとって大きな負担となる。

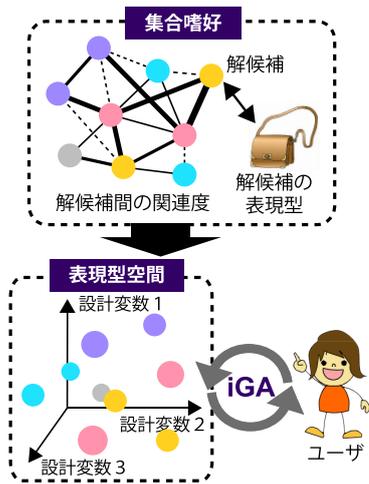


図 3 提案手法のイメージ

3 対話型遺伝的アルゴリズムのための表現型空間の自動生成手法の提案

3.1 提案手法の概要

2.3 節で述べた問題を解決するため、本研究では、商品推薦や感性検索など既に解候補が存在するアプリケーションを対象に、解候補の関連度から設計変数空間を自動的に生成する手法を提案する。従来の設計変数空間の生成は、まず用いる設計変数を決定し、各設計変数を軸とした空間と近傍を定義した上で、各解候補をマッピングする作業に等しかった。これに対し、提案手法では既存の解候補、および人間の感性から見た解候補間の関連度を既存のリソースから算出することで、解候補の表現型が分布する表現型空間を自動的に構築する。(図 3)

解候補の関連度を算出するため、多数のユーザーの感性や嗜好を反映した情報を Web から取得する。現在、Web 上には、多数のユーザーの行動履歴、またユーザーが発信した情報などが蓄積されている。オンラインショッピングサイトにおけるレビューやタグ情報、口コミサイトの批評やソーシャルブックマークなど、これらは少数の専門家による批評ではなく、多数のユーザーやベンダの発信する情報である。本研究ではそれらの多数のユーザーの嗜好を反映した情報を、集合嗜好として定義する。

提案手法によって生成された表現型空間上における解候補の座標から遺伝子型を算出することで、人間の感性を反映した iGA の探索が行えるものと考えられる。

3.2 提案手法の手順

提案手法では集合嗜好から解候補同士の関連度行列を求め、その各行を、各解候補の特徴ベクトルとして

とらえている。しかし、この特徴ベクトルの長さは、扱う解候補の数に依存している。よって解候補の数が莫大なものとなる場合、また、解候補の数が流動的である場合には、各行をそのまま設計変数として用いることは適切ではない。よって、提案手法では主成分分析¹⁷⁾を用いて関連度行列の次元数を適切なサイズに削減し、表現型空間における各解候補の座標とした。以下に、表現型空間の生成手順を示した。

Step.1 解候補間の関連度を定義し、その算出に必要な集合嗜好を Web 上から取得する。

Step.2 解候補の関連度を算出し、関連度行列を生成する。なお、行列の主対角線上は全て 1 とする。

Step.3 主成分分析により、関連度行列の次元数を削減する。削減後の次元数は、累積寄与率を元に決定する。

1. 関連度行列から共分散行列を算出し、その固有値、固有ベクトルを求める。
2. 固有値の降順に、予め定められた次元数分だけ固有ベクトルを抽出し、回転行列を生成する。
3. 元の関連度行列に回転行列を乗じ、表現型空間行列を得る。この行列の各行を表現型空間における各解候補の位置情報とする。

Step.1 の関連度の定義と集合嗜好の取得は、各対象問題毎に異なるため、システム開発者がそれぞれ対応する必要がある。次節では、検討のための対象問題として商品の推薦を扱い、解候補の関連度をショッピングサイトから取得した商品間の推薦関係と定義し、実際の推薦関係を Web 上から取得した。

また、Step.3 において最終的に得られた解候補の位置情報は、用いる交叉や突然変異のオペレータに合わせて遺伝子型に変換して、iGA システムに組み込む。本稿の 5 章における実験では、この位置情報をそのまま遺伝子型として用い、実数値 GA を行った。

4 生成された表現型空間の検証

本章では、提案手法によって生成した表現型空間上で iGA の探索が可能であるかを確認するため、まず生成した空間の特徴について検証した。主に空間上における解候補の分布の特徴、および表現型空間の次元数の 2 点から考察を行った。

4.1 検証データ

4.1.1 Amazon からの集合嗜好の取得

本研究では、解候補とその関連度の取得が可能な集合嗜好の一例として、オンラインショッピングサイト

の Amazon¹が公開する商品間の推薦関係に着目した。Amazon における商品間の推薦関係は、主に協調フィルタリング^{18, 19)}に基づく。協調フィルタリングとは、ある商品の組み合わせが購入される頻度が高い場合、一方の商品を購入したユーザにもう一方の商品を推薦するといった、ユーザの行動履歴から情報のフィルタリングを行う手法である。従って、商品自体の属性については着目しないため、従来のコンテンツフィルタリング^{20, 21)}のように商品に対して何らかの特徴値を設定する必要がない。この商品の推薦関係を解候補間の関連度として定義することで、Amazon におけるユーザの購買行動を反映した空間が生成されることが期待される。

4.1.2 表現型空間の生成

本稿の対象問題は書籍とし、Amazon 上におけるカテゴリ分類が異なる 3 種類の書籍の表現型空間を提案手法により生成した。表 1 に各表現型空間のパラメータを示した。

図 4 に、取得した商品の推薦関係から表現型空間における解候補の座標を求めるまでの流れを示した。各 Step について、以下、具体的な説明を述べる。なお、各 Step は 3.2 節にて述べた表現型空間の生成手順と同期している。

Step.1 解候補間の関連度を商品の推薦関係と定義し、Amazon の商品情報へのアクセスが可能な API²を用いて関連度を取得した。この API により、1 つの商品につき 10 商品までその推薦する商品が取得できる。実際に推薦関係を取得するにあたり、まず各空間の代表的な商品として、表 1 に示した始点の商品を定めた。これらは、それぞれのカテゴリにおいて、データ取得期間での売り上げ順位がトップであった書籍である。始点の商品から、表 1 に示すステップ数で推薦関係によってつながる商品の全てを取得し、互いの推薦関係を求めた。

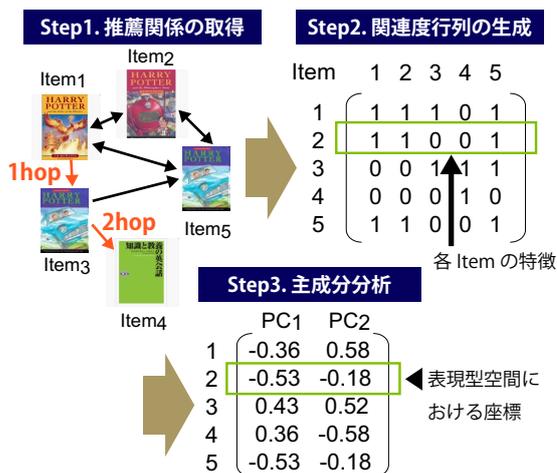


図 4 商品の推薦関係に基づく表現型空間の生成³

Step.2 取得した推薦関係から関連度を算出する手続きについて、本実験では商品 A が商品 B を推薦している場合、関連度を 1、推薦していない場合は関連度を 0 と定義した。

Step.3 生成した表現型空間の次元数は 5 である。これは 3 つの表現型空間において、累積寄与率が 0.1 を超える次元数である。一般には累積寄与率は 0.6 を超えることが望ましいと言われるが、本稿では後述する被験者実験での検討を容易にするため、次元数を低く設定している。

今回の実験では、関連度を商品の推薦関係として定義したが、この他にも、オンラインショッピングサイト上での商品のレビューやタグ情報など、テキスト間の類似度を関連度として定義することもできる。また、対象問題がブログ記事の推薦であれば、解候補の関連度を同じユーザにブックマークされている比率と定義し、ソーシャルブックマークにおける公開ブックマークの情報を取得するなどの対応が考えられる。

4.2 表現型空間における解候補の分布の特徴

今回、Amazon から取得した解候補とその関連度から生成された表現型空間では、著者と出版社が特定の

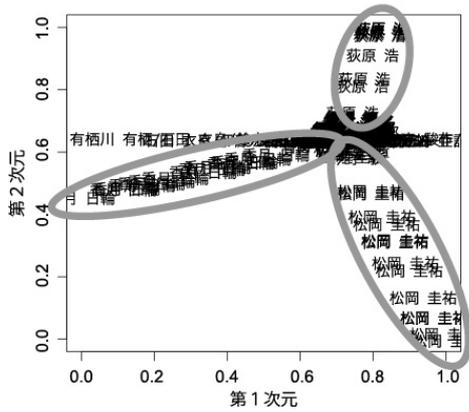
³ 商品画像は AmazonAPI より取得

¹ Amazon.com, Inc. <http://www.amazon.com/>

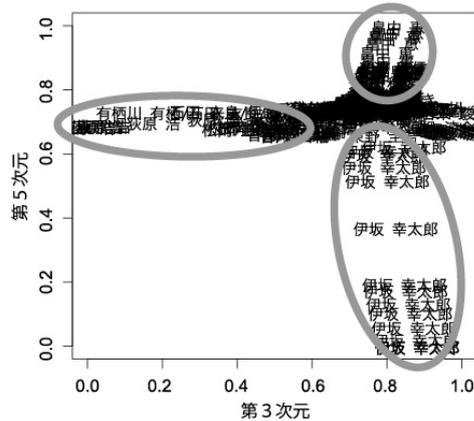
² Amazon Product Advertising API, 以降 AmazonAPI と表記

表 1 表現型空間のパラメータ

空間名	カテゴリ	始点の商品	取得期間	ステップ数	商品数
mystery	ミステリー・サスペンス・ハードボイルド	ブラックベアン 1988(下)	2009/12/31-2009/1/1	3	649
comic	少年コミック	鋼の錬金術師 24	2010/1/1-2010/1/2	4	309
science	コンピュータサイエンス	プログラマのための論理パズル	2010/1/3-2010/1/4	4	703



(a) 横軸：第1次元，縦軸：第2次元



(b) 横軸：第3次元，縦軸：第5次元

図5 mysteryの解候補の分布(著者)

次元において偏りを見せるなど、特徴的な分布が確認された。

4.2.1 解候補の分布における著者の偏り

表現型空間 mystery, comic では、解候補の分布に著者の影響が確認された。図 5(a) は表現型空間 mystery の第1次元を横軸に、第2次元を縦軸にとり、各解候補の座標にその著者名を印字した散布図である。図 5(b) は、第3次元と第5次元を軸としている。各軸に対し、著者を同じくする解候補が狭く分布している様子が確認できる。例えば、著者を「松岡圭祐」とする解候補は、第1次元に対して0.7から1.0の範囲内で分布している。

この分布の特徴は science では見られなかったが、comic では同様の傾向が確認された。図 6 にその例を示す。

この分布は、Amazon のユーザが同じ著者の書籍を買い揃えるという購買行動をとることが原因と考えられる。表 2 では表現型空間毎に、解候補である書籍の著者が空間の中で唯一ではない、即ち他の解候補と著者を同じくする書籍の数と、それらが全体に占める

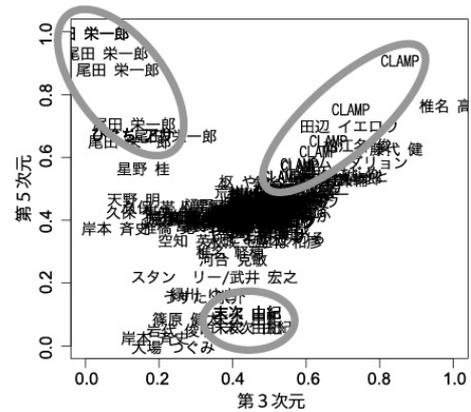


図6 comicの解候補の分布(著者)

割合を示した。mystery では、他の解候補と同じ著者を持つ解候補が575であり、解候補全体に占める割合0.886は、science の3倍近い。同じ著者の解候補は、互いに推薦し合う密な推薦関係を持つ傾向がある。そのため、主成分分析の結果、図 5に見られるようなはっきりとした分布を示す。science において著者の偏りが見られなかったのは、表 2 に示したように、著者を同じくする解候補が少なかったためと考えられる。

4.2.2 解候補の分布における出版社の偏り

表現型空間 science における解候補の分布では、出版社の偏りが見られた。science の第1次元、第4次元を軸とし、各解候補の出版社を印字したグラフを図 7 に示す。横軸に対して、複数の出版社が狭い範囲に分布している。これは、science に属する書籍の専門性が高いため、出版社の傾向を重視することが原因と考えられる。

このように、mystery, comic の表現型空間においては著者が、science では出版社が軸に対して狭く分布する現象が見られた。これは、著者や出版社を重視して iGA を行うユーザにとって、効率的な探索を行うことができ、嗜好性の高い解候補が提示される可能性を示している。

表 2 著者を同じくする解候補の比率

カテゴリ	同著者の解候補数	全体に占める比率
mystery	575	0.886
comic	216	0.699
science	213	0.303

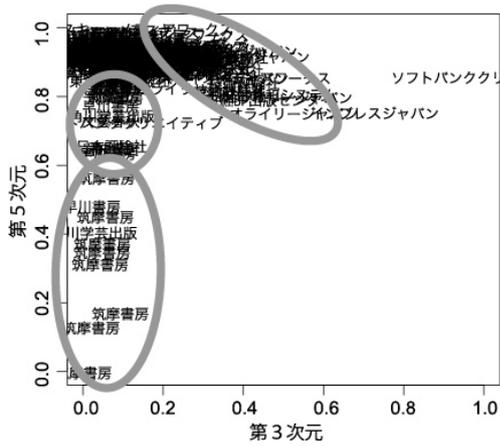


図7 scienceの解候補の分布(出版社)

4.3 表現型空間の次元数の検討

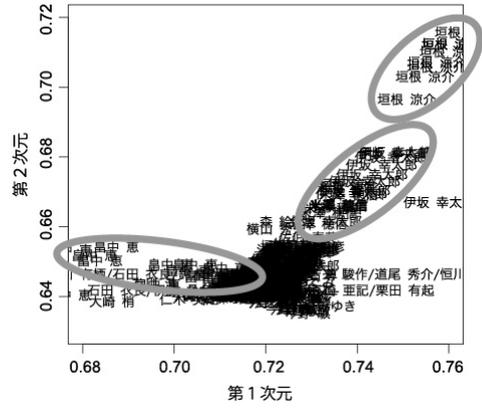
mysteryにおける解候補の分布から、表現型空間の適切な次元数とに関する考察を行った。

図8は図5(a)において解候補が密集した領域をそれぞれ12.5倍、100倍に拡大したグラフである。図8(a)では図5と同様に、分布における著者の偏りが確認されたが、図8(b)ではその傾向が見られなかった。これらの図8で確認されなかった著者の一部は、図5(b)など他の次元の分布において、偏りを示した。このように1つの主成分、すなわち1つの軸では有意な著者数が限られている。

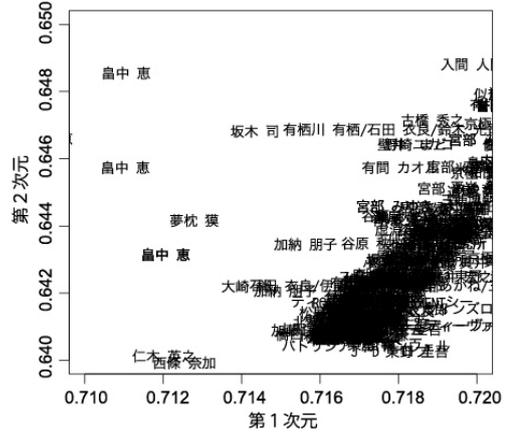
解候補の分布に対する著者の影響を調べるために、元の関連度行列に対して、主因子法による分析をおこなった。主因子法によって各解候補は、ある特定の軸、即ち主成分に対して持つ影響の度合いを因子負荷量として算出する。この因子負荷量が、危険率0.10において有意であった解候補を主成分毎に抽出し、その解候補の著者を列挙したものを表3に示した。表3より、各主成分に対して有意な因子負荷量を持つ著者の数は1から3である。また、全ての主成分に対して有意な

表3 有意な解候補の著者

次元	著者
1	松岡圭祐, 香月日輪, 畠中恵
2	松岡圭祐, 荻原浩
3	荻原浩, 垣根涼介
4	垣根涼介
5	伊坂幸太郎(共著者:石田衣良, 市川拓司, 中田永一, 中村航, 本多孝好), 畠中恵(共著者:柴田ゆう)



(a) 横軸:第1次元 [0.68-0.76], 縦軸:第2次元 [0.64-0.72]



(b) 横軸:第1次元 [0.710-0.720], 縦軸:第2次元 [0.640-0.650]

図8 mysteryの解候補の分布

著者はいなかった。このことから、より多くの著者について分布の偏りを示すには、より多くの主成分が必要となることが分かる。しかし、表現型空間の次元数を増やすことは、遺伝子型の次元数を増加させ、最適化アルゴリズムであるGAの解探索性能を低下させる。このトレードオフを解決する適切な次元数の決定手法、またはisomap²²⁾など、より低次元で元の関連度行列を表現する次元削減手法の検討が必要である。

5 表現型空間におけるiGAの探索の検討

5.1 実験概要

提案手法により生成された表現型空間にて、iGAによる探索が可能であるか検証するために、被験者実験を行った。実験では、4.1節で検討したmystery, comic, scienceの各表現型空間に対応する3つのiGAシステム(以降、mysteryシステム, comicシステム, scienceシステムと表記)を構築し、被験者に操作させた。各解候補の遺伝子型は、それぞれの表現型空間上における位置座標とした。

実験結果から、以下の2つの項目について検討した。

1. 個人の嗜好に応じた探索

提案手法によって生成した空間において、iGAによって個人の嗜好のランドスケープが構築されるか、また取得したランドスケープが被験者毎にパーソナライズされているかを検証した。

2. Amazon との探索経路の違い

本実験では、探索空間の生成に、Amazonの推薦関係を利用している。そのため、ユーザの探索履歴を調査し、元データと異なるiGA独自の探索が行われているかを検証した。

5.2 実験システム

5.2.1 評価インタフェース

本実験で構築したiGAシステムは、解候補を書籍とした表現型空間を実装するため、書籍のiGAシステムとなる。実験システムの評価インタフェースを図9(a)に示す。評価インタフェースでは、スクロールをせず一覧できる解候補数として16冊の書籍が提示される。この書籍の中から、被験者は自身の嗜好に合った書籍をクリックして選ぶ。このとき、一画面において選択できる商品の最大数は、用いたGAのアルゴリズムにより提示個体の半数と定めた²³⁾。被験者は1個以上の個体を選択して、画面上部右の“次へ”ボタンを押す。それにより、システムがユーザの評価に基づいてGAによる遺伝的操作を行い、インタフェース上の商品を更新する。

被験者に対して提供する商品の情報を、表4に示した。被験者が書籍の表紙にカーソルを合わせると、商品の詳細な情報が記載された詳細画面(図9(b))がポップアップされる。詳細画面は、カーソルが離れると消去される。

5.2.2 遺伝的操作

提案手法は、既に存在する解候補を扱うため、交叉や突然変異によって生成された子個体の遺伝子が、既存の解候補の遺伝子型と一致しない場合もある。その場合、生成した子個体の遺伝子型を、表現型空間上の座標が最も近い解候補の遺伝子型によって置換する。



(a) 評価インタフェース



(b) 詳細画面

図9 実験システム⁴

このとき、置換する解候補は、その世代の全ての親個体以外の解候補であるとする。これは直近に提示された個体について、被験者が記憶するためである。

単純な図形や色の最適化を行うといった、解候補がより単純化されたiGA²³⁾では、表示された解候補に対する被験者の記憶は曖昧である。しかし書籍など、文字情報を扱う問題や被験者の知識を連想させる対象問題では、被験者は直感だけでなく知識を活用して評価を行うため、提示された解候補の特徴をよく記憶する。そのため、同じものが世代間で連続して提示された場合、被験者は同じ解候補を評価することに違和感

⁴ 商品画像は AmazonAPI より取得

表4 被験者に提示する商品情報

ウィンドウ	提示する商品の情報
評価インタフェース (図9(a))	表紙
詳細画面 (図9(b))	表紙, 表題, 著者, 価格, 出版年, 出版社, カスタマーレビュー

を覚える、または、飽きてしまうといった意見が予備実験で得られた。

5.3 実験手順

被験者は各システムに対し、4世代ずつ iGA による評価を行い、それを1試行とする。3つのシステムの操作順序は、被験者間でカウンタバランスをとった。

被験者数は6名(男性5名、女性1名)である。実験は以下の手順で行った。

1. 被験者への教示

被験者に書籍を選択するモチベーションを与えるため、「明日は休日であり、その日に読書する本をオンラインブックストアで探す」という内容の教示を与えた。

2. iGA システムの操作

被験者が iGA システムを1試行操作する。

3. 満足度に関するアンケート

最終提示画面に対する満足度について、被験者が以下の5段階評価のアンケートに答える。

- (a) とても満足している
- (b) 満足している
- (c) 不満
- (d) とても不満
- (e) 特に思うことはない

4. 終了判定

全てのシステムについて試行した場合、実験を終了する。終了しなければ、2に戻る。

5.4 実験結果

5.4.1 個人の嗜好に応じた探索の検討

表現型空間において、個人の嗜好によってパーソナライズされた探索が行われるか検証するため、各

表5 GAパラメータ

パラメータ	値
個体数	16
次元数	5
世代数	4
交叉手法	$BLX - \alpha^{14)}$
交叉率	1
α	0.2
突然変異手法	一様突然変異
突然変異率	0.2

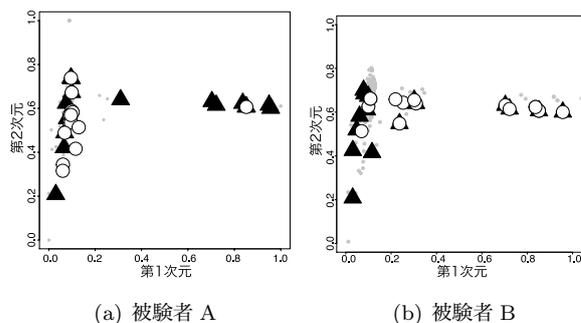


図10 最終世代の分布の比較 (mystery 第1次元-第2次元)

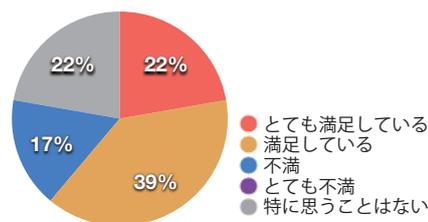


図11 満足度

ユーザの探索履歴から嗜好のランドスケープを取得し、被験者間で比較した。また、そのランドスケープが正しくユーザの嗜好を反映しているか、アンケートによって調査した。

図10は、mysteryシステムにおける被験者Aと被験者Bの探索履歴である。空間の第1次元、および第2次元において、第1世代に提示された解候補を黒三角、第4世代で提示されたものを白丸でプロットした。被験者A、Bの初期個体は同一である。しかし、最終世代においては被験者Aは第1次元に、被験者Bは第2次元に対して提示個体が狭い範囲に分布していた。この探索結果について、被験者アンケートの結果、被験者A、Bともに「満足している」と回答した。このことから、生成した表現型空間上において、被験者固有の嗜好のランドスケープを抽出できたものと考えられる。

軸に強く依存したランドスケープは少なかったが、他の被験者間においてもそれぞれ異なる探索結果が得られた。図11は、各被験者の空間毎のアンケート結果を総計したものである。「満足している」「とても満足している」との回答が全体の半数を超えており、探索の結果が被験者の嗜好に沿うものであったと確認できた。なお、不満と答えた被験者の意見には、「カテゴリに対して馴染みがなかったので判断し難い」といった対象問題そのものに対する不満などがあつた。



図 12 被験者 B の第 4 世代における提示画面 (comic)
5

5.4.2 Amazon との探索経路の違いの検討

iGA による探索が、表現型空間の元データである Amazon の推薦とは異なる探索経路をとることを確認するため、ユーザが好ましいと評価した候補の中に、Amazon における関連度が高い候補だけでなく、iGA によってのみ探索された候補が含まれているか検討した。

⁵ 商品画像は AmazonAPI より取得

図 12 は、被験者 B が comic システムを操作した際の第 4 世代の提示画面である。この世代において被験者 B が選択した個体を丸で示した。商品上部に A を示した個体は、第 3 世代で選択された個体と Amazon における関連度が高い商品（以降、関連個体と呼ぶ）である。即ち、被験者が選択した個体との推薦関係が、Amazon の API によって取得された個体を意味する。

被験者が第 4 世代で選択した個体にも、関連個体が含まれる。しかし、iGA によってのみ探索された個体（以降、非関連個体と呼ぶ）も被験者による評価を受けており、この被験者においては、設計変数空間の元となった Amazon の推薦関係とは異なる探索が iGA によって行われ、それが被験者に評価されたと言える。

表 6 は 2 世代目から 4 世代目の提示個体に占める、非関連個体の比率である。iGA によってのみ探索された個体が、各被験者、各システムにおいても、半数以上提示されている。また、非関連個体が、被験者の評価した個体中に占める比率を示した表が、表 7 である。評価された非関連個体の比率について、全体の平均をとると 0.390 であった。関連個体が、前の世代において被験者によって評価された個体と類似度の高い個体であることを考えると、非関連個体の比率として 0.390 は十分な値であると言える。

これにより、提案手法によって生成された空間における iGA では、元データにおける候補間の関連度とは異なる探索を行っていたと言える。また、その探索結果が被験者の評価に適うものであることが示された。

表 6 提示個体に占める非推薦候補の比率 (第 2 世代-第 4 世代)

システム	被験者 A	被験者 B	被験者 C	被験者 D	被験者 E	被験者 F
comic	0.583	0.479	0.521	0.813	0.875	0.667
mystery	0.646	0.750	0.479	0.521	0.542	0.938
science	0.375	0.604	0.750	0.729	0.708	0.708
average	0.535	0.611	0.583	0.688	0.708	0.771

表 7 被験者毎の評価個体に占める非推薦候補の比率 (第 2 世代-第 4 世代)

システム	被験者 A	被験者 B	被験者 C	被験者 D	被験者 E	被験者 F
comic	0.250	0.304	0.389	0.556	0.625	0.714
mystery	0.400	0.385	0.118	0.087	0.308	1.000
science	0.412	0.375	0.417	0.538	0.417	0.600
average	0.354	0.355	0.308	0.394	0.450	0.771

6 結論

本稿では、表現型がすでに複数のユーザによって評価され、それらの解の関係が得られている問題を対象とし、iGAにおける汎用的な設計変数空間の構築手法について提案した。今までのiGAシステムの設計変数空間は、対象問題が異なるごとに最適化問題を定式化し直し、実装側が手作業でシステムを再構築してきた。本研究では、既存のユーザの嗜好、感性に関する集合的なりソースを集合嗜好として定義し、これらから解候補の関連度を取得し、解候補が分布する表現型空間を生成した。この表現型空間を用いた探索空間では、ユーザの感性を反映したiGAの探索が可能であると考えられる。

提案手法の有効性を示すために、探索対象である個体の設計変数をWeb上の情報から生成するシステムを構築し、実験を行った。実験では、対象問題として商品推薦を用い、Amazonから取得した書籍間の推薦関係から関連度行列を生成し、主成分分析により各書籍の表現型を算出した。

提案手法により生成された表現型空間の特徴を検証するため、Amazonの書籍の推薦データを用いて、mystery, comic, scienceの3つの表現型空間を生成した。mystery, comicの設計変数空間においては著者の値が軸に対して狭い範囲に分布する現象が見られた。scienceでは出版社の値に同じ傾向が見られた。これは著者や出版社を重視したiGAを行うユーザが効率的な探索を行える可能性を示している。しかし、一方で著者などの解候補の情報をより適切に表現型空間に反映するには、設計変数の高次元化が必要となる。これはiGAにより有効な探索を行うための設計変数の低次元化には反する特徴だった。

また、生成された表現型空間においてユーザがどのような探索を行うのか検証するため、生成した表現型空間を用いた書籍のiGAシステムを構築し、被験者実験を行った。探索履歴から、被験者間で探索傾向が異なっており、被験者固有の嗜好のランドスケープが取得されていることを確認した。アンケート調査の結果において、探索結果に対して満足と答えたケースは半数を超え、被験者固有の嗜好に基づく探索が行われたと考えられる。また、表現型空間を作るための関連度を取得した元データとの違いを比較し、iGAによってのみ探索された結果が被験者の評価を受けることを確認した。

今後の課題としては、表現型空間を生成するにあたり、効率的に軸を抽出する手法と、また適切な次元数について調査を行う。そして、生成した空間での探索の精度を上げるiGAの遺伝的操作のオペレータにつ

いて検討する。

参考文献

- 1) Hideyuki Takagi. Interactive evolutionary computation: fusion of the capabilities of ec optimization and human evaluation. *Proceedings of the IEEE*, Vol. 89, No. 9, pp. 1275–1296, 2001. Invited Paper.
- 2) Hideyuki Takagi. Interactive ga for system optimization: Problems and solution. In *4th European Congress on Intelligent Techniques and Soft Computing*, pp. 1440–1444, Aachen, Germany, 1996.
- 3) 高木英行, 畷見達夫, 寺野隆雄. 対話型進化計算法の研究動向. *人工知能学会誌*, Vol. 13, No. 5, pp. 692–703, Sep. 1998.
- 4) 三木光範, 廣安知之, 小川泰正, 長谷佳明, 吉田昌太. 対話型遺伝的アルゴリズムにおける並列分散モデルの有効性. 第16回人工知能学会全国大会論文集, pp. 3A1–02, 2002.
- 5) David E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Professional, 1989.
- 6) Sung-Bae Cho. Towards creative evolutionary systems with interactive genetic algorithm. *Applied Intelligence*, Vol. 16, No. 2, pp. 129–138, 2002.
- 7) Hideyuki Takagi, Sung-Bae Cho, and Toshihiko Noda. Evaluation of an iga-based image retrieval system using wavelet coefficients. In *IEEE International Conference on Fuzzy Systems*, pp. 1775–1780, Seoul, Korea, 1999.
- 8) Maiko Sugahara, Mitsunori Miki, and Tomoyuki Hiroyasu. Design of japanese kimono using interactive genetic algorithm. In *2008 IEEE International Conference on Systems, Man and Cybernetics*, pp. 185–190. IEEE, October 2008.
- 9) 菅原麻衣子, 三木光範, 廣安知之. 好みのカラーイメージに基づく初期個体を生成させる対話型遺伝的アルゴリズム. *人工知能学会全国大会論文集*, 第22巻, pp. 2B1–02. 同志社大学, 2008.
- 10) Asuka Amamiya, Mitsunori Miki, and Tomoyuki Hiroyasu. Interactive Genetic Algorithm

- using Initial Individuals Produced by Support Vector Machine. *The Science and Engineering Review of Doshisha University*, Vol. 50, No. 1, pp. 34–45, 2009.
- 11) Hee-Su Kim and Sung-Bae Cho. Application of interactive genetic algorithm to fashion design. *Engineering Applications of Artificial Intelligence*, Vol. 13, No. 6, pp. 635 – 644, 2000.
 - 12) 渡辺芳信, 吉川大弘, 古橋武. 一対比較評価に基づく対話型遺伝的アルゴリズム (session 3). 情報処理学会研究報告. MPS, 数理モデル化と問題解決研究報告, Vol. 2007, No. 19, pp. 69–72, 2007.
 - 13) Lawrence David Davis. *Handbook of Genetic Algorithms*. Van Nostrand Reinhold Company, 1991.
 - 14) Larry J. Eshelman and J. David Schaffer. Real-coded genetic algorithms and interval-schemata. *Foundations of Genetic Algorithms*, Vol. 2, pp. 187–202, 1993.
 - 15) Alden H. Wright. Genetic algorithms for real parameter optimization. In *Foundations of Genetic Algorithms*, pp. 205–218. Morgan Kaufmann, 1991.
 - 16) 伊藤冬子, 廣安知之, 三木光範, 横内久猛. 対話型遺伝的アルゴリズムにおける嗜好の多峰性に対応可能な個体生成方法. 人工知能学会論文誌, Vol. 24, No. 1, pp. 127–135, 2009.
 - 17) I. T. Jolliffe. *Principal component analysis*. Springer, 1986.
 - 18) Badrul Sarwar, George Karypis, Joseph A. Konstan, and John Reidl. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web*, pp. 285–295. ACM New York, USA, 2001.
 - 19) G. Linden, B. Smith, and J. York. Amazon.com recommendations: item-to-item collaborative filtering. *Internet Computing, IEEE*, Vol. 7, No. 1, pp. 76–80, 2003.
 - 20) Ken Lang. Newsweeder: Learning to filter net-news. In *Proceedings of the 12th International Machine Learning Conference*, 1995.
 - 21) Michael Pazzani, Jack Muramatsu, and Daniel Billsus. Syskill & webert: Identifying interesting web sites. In *Proceedings of the 13th National Conference on Artificial Intelligence*, pp. 54–61, 1998.
 - 22) J. B. Tenenbaum, V. Silva, and J. C. Langford. A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science*, Vol. 290, No. 5500, pp. 2319–2323, 2000.
 - 23) Tomoyuki Hiroyasu, Nozomi Yamakawa, Fuyuko Ito, Mitsunori Miki, and Yasunari Sasaki. Discussion of the evaluation method and the generation method of the individuals for interactive genetic algorithm. Research note, IPSJ SIG Technical Report 2008-MPS-68, 2008.
1. 田中 美里
2008年同志社大学工学部知識工学科卒業。2010年同大学大学院工学研究科博士前期課程修了。現在、同大学生命医科学部研究補助員。最適化技術を応用した感性情報学の研究に従事。IEEE, 人工知能学会, 情報処理学会, 日本知能情報ファジィ学会各会員。
 2. 廣安 知之
1997年早稲田大学大学院理工学研究科後期博士課程修了。早稲田大学理工学部助手, 同志社大学工学部助手, 知識工学科専任講師, インテリジェント情報工学科准教授を経て2008年から生命医科学部教授。進化的計算, 最適設計, 並列処理, 設計工学, 医療画像工学などの研究に従事。IEEE, 情報処理学会, 電子情報通信学会, 計測自動制御学会, 日本機械学会, 超並列計算研究会, 日本計算工学会各会員。
 3. 三木 光範
1950年生。1978年大阪市立大学大学院工学研究科博士課程修了, 工学博士。大阪市立工業研究所研究員, 金沢工業大学助教授を経て1987年大阪府立大学工学部航空宇宙工学科助教授, 1994年同志社大学理工学部教授。進化的計算手法とその並列化, および知的なシステムの設計に関する研究に従事。著書は「工学問題を解決する適応化・知能化・最適化法」(技法堂出版)等多数。IEEE, 米国航空宇宙学会, 情報処理学会, 人工知能学会, システム制御情報学会, 日本機械学会, 計算

工学会，日本航空宇宙学会等会員，超並列計算研究会代表，経済産業省産業技術審議会委員などを歴任，知的オフィス環境コンソーシアム会長。

4. 佐々木 康成

2001年筑波大学大学院博士課程心理学研究科退学，2003年博士（学術），同志社大学工学部ポストドクター，同特別研究員（PD）を経て，2005年同知識工学科（現インテリジェント情報工学科）任期付講師，現在同大学理工学部研究員。実験心理学を背景としながら，空間の認知と感性評価，テクスチャの解析とユーザインタフェースへの応用，学習と潜在記憶，言語と音声知覚，注意機能等の認知科学的研究に従事，さらに情報分野との共同研究も幅広く展開。日本生理心理学会，日本音響学会，日本認知科学会，日本認知心理学会，日本光脳機能イメージング研究会等，各会員。

5. 吉見 真聡

平成16年慶應義塾大学理工学部情報工学科卒業。平成21年同大学院理工学研究科開放環境科学専攻後期博士課程修了。博士（工学）。平成18年度より日本学術振興会特別研究員（DC1）。現在，同志社大学理工学部助教。リコンフィギャラブルシステム，並列処理，知的システムの研究に従事。電子情報通信学会，人工知能学会各会員。

6. 横内 久猛

1971年徳島大学工学部電気工学科卒。同年（株）日立製作所入社。中央研究所にて眼底写真の自動認識の研究，がん細胞・血液像の自動認識装置の開発，カラーブラウン管検査自動化の研究，電子線描画露光装置の開発，デジタルX線装置の開発などに従事。1991年（株）日立メディコに転属。実時間デジタルX線装置開発，医療画像診断装置のマーケティング担当。2006年（株）日立メディコ退職。2008年同志社大学生命医科学部教授。医用画像工学・信号処理，ヒト脳機能計測・処理などの研究に従事。日本生体医工学会，日本医用画像工学会，映像情報メディア学会，日本ヒト脳機能マッピング学会各会員。