

進化的計算のためのビジュアルインタフェース EMACI を用いたシミュレーション

大山喜冴 1) 伊藤貴之 1) 山下富義 2) 小山田耕二 2)
1)お茶の水女子大学 2)京都大学

Simulations Using a Visual Interface EMACI for Evolutional Computation

Kisa Ohyama 1) Takayuki Itoh 1) Fumiyoshi Yamashita 2) Koji Koyamada 2)
1) Ochanomizu University 2) Kyoto University

This paper presents EMACI, a visual interface for evolutionary computation such as genetic algorithm. It represents distribution and change of multiple parameters, by applying two-tone pseudo coloring and parallel coordinates techniques. EMACI also provides a GUI for interactively controlling configurations of the evolutionary computation algorithms. We applied EMACI to heart cell simulation, which mimics biological dynamic model of the cells. This paper presents that we obtained better results by applying EMACI.

1. はじめに

進化的計算は、生物や社会の進化をモデル化した計算手法により、最適化問題をはじめとする情報科学の諸問題を解くものである。進化的計算の代表例として、遺伝的アルゴリズム(GA: Genetic Algorithm)があげられる⁽¹⁾。GAは、解空間構造が不明で、全探索が不可能なほど広大な解空間を持つ問題に有効な最適化手法である。GAはダーウィンの進化論をモチーフに、環境に適合する度合い(適合度)を数値で表し、その適合度を最適化問題の目的関数と考える。適合度の値が進化と共に大きくなることから、最適化問題における最適解に近づくことに相当する。

本報告では、GAをはじめとする進化的計算の反復処理を改善するためのビジュアルインタフェース

EMACI (Evolutional Multi-parameter Analysis and Control Interface) を提案する。EMACIはGAの解探索過程の可視化機能、およびユーザが可視化結果を観察しながら意図的に突然変異確率・交叉確率を調節するためのGUIから構成される。EMACIを用いることで最適化問題を扱うユーザは、より短時間で準最適解を発見できると考えられる。

本報告では、心筋細胞シミュレーション⁽²⁾のパラメータ最適化にEMACIを適用した事例を紹介し、EMACIの有効性について検証する。

2. 関連研究

2.1 遺伝的アルゴリズム

GAは、解の候補を遺伝子で表現した「個体」を複

数用意し、世代が進むに連れ適合度の高い個体を優先的に選択して、交叉・突然変異などの操作を繰り返しながら、解を探索するアルゴリズムである⁽¹⁾。

GA の解探索性能は遺伝的演算法だけでなく、そのパラメータの個体数・突然変異確率・交叉確率などにも大きく依存する。そのため、遺伝的演算法やパラメータが解探索性能に及ぼす効果を知るために、適合値の改善効果を指標とするのが一般的である。近年ではGAの解探索の特性を理解するために、GAの反復処理の過程を可視化する研究もいくつか発表されている⁽³⁾。本報告の提案手法もその一種であると考えられる。

2.2 擬似 2 色表示手法

擬似 2 色表示手法 (TPC: Two-Tone Pseudo Coloring) は、2 つの色を帯領域に割り当てて、その各々の色の占める幅を制御することで、任意のスカラ値を表現する可視化手法である⁽⁴⁾。特に時系列データにおいて、数値変化を正確に、また微小な変化が読み取りやすいように可視化できるのが特徴である。この手法による表示例を図 1 に示す。

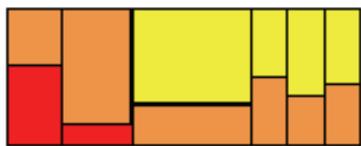


図 1：擬似 2 色表示手法による表示例

2.3 平行座標プロット

平行座標プロット (PCP: parallel coordinate plot) は多変数データを視覚化するのに有用な統計グラフの一つである⁽⁵⁾。PCP では、変数の座標軸を平行に配置し、全ての変数で最小値が下端に、最大値が上端になるように変数値をプロットする。そして隣接する座標軸上の変数値をそれぞれ線分で結ぶ。平行座標軸の間隔を調節することにより、すべての変数の座標軸を一画面に表示できることから、同時に全ての辺量を表示できる。

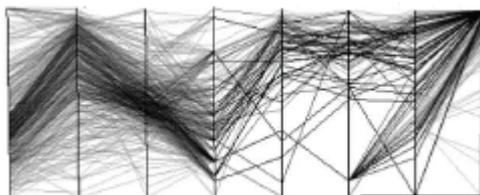


図 2：平行座標プロットによる可視化結果

2.4 心筋細胞シミュレーション

細胞や臓器、また血液循環・エネルギー代謝といった生体機能をコンピュータ上に再現することで、人体の複雑なメカニズムを解明する研究が、近年活発に進められている。本報告では、心筋細胞を再現するシミュレーション⁽²⁾を最適化問題の題材とする。心筋細胞の信号伝達は電位の変化に基づいており、その電位の変化は細胞膜を通して出入りする各種イオンの動きと深い関係にある。つまり、このイオンの動きを再現することで、心筋細胞の信号伝達の特性をコンピュータ上で再現できると考えられる。このシミュレーションによって、現実の心筋細胞の働きに近い特性を再現するためには、イオンの動きを決定するパラメータを正しく設定する必要がある。言い換えれば、心筋細胞シミュレーションの最適化は、イオンの動きを決定するパラメータの最適化問題に帰着される。この最適化問題を解くための計算手法として、進化的計算は有力であると考えられる。

3. 提案手法

本報告で提案する EMACI は、進化的計算における反復計算の過程、反復計算の任意のタイミングにおける詳細情報、反復計算を制御する GUI、の 3 つのブロックで構成される。現時点での我々の実装は、GA を題材として、反復計算における各世代の遺伝子の変遷と、ある 1 世代における全遺伝子の分布を可視化する。さらに、交叉・突然変異などの操作を制御する GUI を備える。以下に、EMACI を構成する各ブロックについて説明する。

3.1 時系列表示部

本報告では、GA が求める解を遺伝子で表現した「個体」の中で、最も適合度の高い個体をエリート解とみなす。また本論文では遺伝子の適合度をスコア値とし、遺伝子を構成する値をパラメータ値とする。時系列表示部では、GA の反復処理の過程において、世代ごとに更新されるエリート解のスコア値とパラメータ値の変化を時系列表示するものである。

図 3 に示すウィンドウの上半分が時系列表示部である。時系列表示部では TPC を用いて、各世代のスコア値とパラメータ値を表現する。

なお提案手法の名前「EMACI」は、GA の反復計

算の進行に沿って、時系列表示部が絵巻物のように左右に広がって表示される様子にも起因している。

3.2 一世代表示部

一世代表示部は、ある一世代について、非エリート解を含めた全遺伝子の全パラメータを、PCP を用いて表示するものである。ユーザは時系列表示部の任意の世代にカーソルをあてることで、任意の世代を選び、その世代における解候補の遺伝子の分布を観測できる。これにより、個体同士の相対的位置関係を理解できるので、解の収束や集散の様子などを把握しやすくなる。

図 3 に示すウィンドウの下半分が一世代表示部である。本手法で用いている PCP の座標軸は、各パラメータ値に対応している。

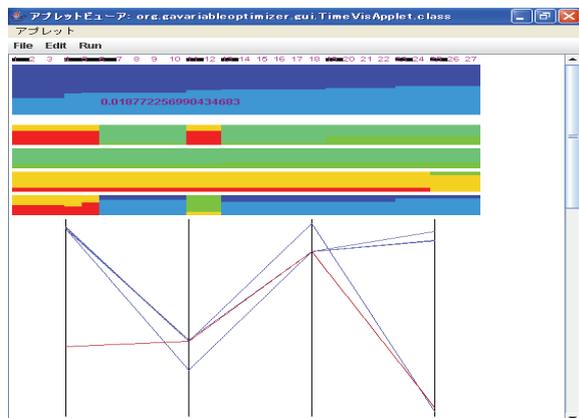


図 3：時系列表示部と一世代表示部

3.3 GUI 部

EMACI では、解の進化が停滞し、何世代にもわたってエリート解のスコア値が改善されないようであれば、ユーザに対してポップアップウィンドウを提示し、突然変異確率・交叉確率・遺伝子数の再設定を促す仕組みを搭載している。

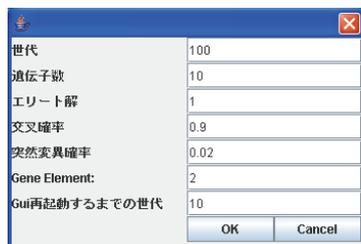


図 4：GUI 部のポップアップウィンドウ

図 4 は EMACI の GUI 部におけるポップアップウィンドウである。このウィンドウが表示されたら、ユーザは時系列表示部や一世代表示部を参照して

GA の反復過程を観察し、状況に応じて意図的に各変数を再設定できる。

3.4 シナリオ

GA をはじめとする進化的計算が対象とする最適化問題は、その解空間が非常に大規模であることや、数値分布が非常に複雑であることが多く、収束までに数分から数時間といった長時間を要することも多いと想定される。そのような長時間にわたる反復計算の一部始終を監視し続けるのは、非現実である場合が多い。そこで我々は、以下のような使い方を想定して EMACI を実装している。

EMACI は GA と連動しており、反復計算の一世代ごとに GA から計算結果を受け取り、時系列表示部の表示を更新する。この間、必ずしもユーザは時系列表示部を監視し続けなくてもよい。よってユーザは GA による最適化計算を実行させながら、並行して別の作業を進めることができる。そして GA のエリート解のスコア値が改善されないようであれば、EMACI は警告音とともにポップアップウィンドウを起動し、ユーザに各変数の再設定を促す。この時点で初めてユーザは時系列表示部に目を向け、マウス操作によって一世代表示部を確認し、各変数を再設定する。例えば、GA による最適化計算が局所解付近で停滞している可能性が高いと判断したら、一時的に突然変異確率を高くすることで、解空間の他の領域を積極的に探索させることができる。このような操作の繰り返しによって、EMACI はユーザに、より早い段階での最適化計算の収束を導くものである。

4. 実行結果

著者らは EMACI の有効性を実証するために、心筋細胞シミュレーションのパラメータ最適化問題を題材にして、GA と EMACI を用いてスコア値が最大となるパラメータ値を求めた。著者らの実験では、パラメータ値に ICaL, IK1, Iks, Ikr の 4 つのイオン属性値を用い、サンプル解との類似度をもってスコア値を算出した。

本実験では第一段階として、同一の心筋細胞シミュレーションのパラメータ最適化を、GUI 部を操作せずに 10 回反復した。その結果、10 回中 2 回だけ、際立ってスコア値の高い結果が出た。その 2 回にお

けるパラメータ値は非常に近く、残りの 8 回のパラメータ値はそれとは遠かった。このことから、第一段階にてパラメータ最適化が準最適解に到達したのは 10 回中 2 回と判断した。また、この 2 回において準最適解に到達したと判断される世代数は、それぞれ 103 世代、132 世代であった。

続いて実験の第二段階として、同一の心筋細胞シミュレーションのパラメータ最適化を、GUI 部を操作しながら 5 回反復した。この結果、第一段階で算出された準最適解に近いパラメータ値を得たのは 5 回中 3 回であり、第一段階よりも高い確率で準最適解を得ることができた。この 3 回のスコア値は、第一段階と比べて 3 回とも高い値であり、より最適解に近い準最適解を得られたことがわかる。また、この 3 回において準最適解に到達したと判断される世代数は、それぞれ 93 世代、40 世代、69 世代であり、第一段階よりも短い世代数で準最適解に到達できたことがわかる。

以上の実験結果を表 1 に示す。この結果から、EMACI を用いて GUI 部を操作することで、より高い確率で、かつ早い世代で準最適解の近傍に到達し、最適解に近い準最適解を得られていることがわかる。

表 1 準最適解に到達したときのスコア値と世代数

	スコア値	世代数
第一段階(1)	0.025632	103
第一段階(2)	0.025718	132
第二段階(1)	0.025766	93
第二段階(2)	0.025896	40
第二段階(3)	0.025943	69

5. まとめ・今後の課題

本報告では、GA をはじめとする進化的計算に対して、反復処理の過程で得られる解探索過程を可視化させ、可視化結果を観察しながら意図的に突然変異確率・交叉確率を調節することができるビジュアルインタフェース EMACI を提案した。EMACI による可視化結果から、解候補である遺伝子間の相対的位置関係を理解し、解の収束や集散の状況を把握できると考えられる。これを参考にして、解の収束や集散に応じて適応的に突然変異確率や交叉確率を操作

することで、より早い段階で準最適解を求めることができると考えられる。

本報告ではまた、細胞心筋シミュレーションへの適用例を示した。この実験結果から、GUI による各変数の再設定を用いることで、GA の反復処理より早い世代で準最適解を発見できることが実証された。

今後の課題として、パラメータや反復回数が非常に多くて一画面に表現しきれないような最適化問題を想定した表示機能の改良があげられる。

また、統計的に有意なだけの実験回数を重ねる、心筋細胞シミュレーション以外の多くの最適化問題に対して EMACI を適用する、多くの被験者に EMACI を試用してもらう、などの方向でさらなる実験と検証を進めたい。

謝辞

京都大学 細胞・生体機能シミュレーションプロジェクト関係者の方々に感謝の意を表します。

参考文献

- (1) D. E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wasley Publishing Company, 1989.
- (2) 皿井, 野間, simBio: 生物学的ダイナミックモデル開発基盤, 日本エム・イー学会雑誌(BME), 18, 3-11, 2004.
- (3) N. Hayashida and H. Takagi, Acceleration of EC Convergence with Landscape Visualization and Human Intervention, Applied Soft Computing, 1, 4F, 245-256, 2002.
- (4) T. Saito, H. N. Miyamura, M. Yamamoto, H. Saito, Y. Hoshiya, T. Kaseda, Two-Tone Pseudo Coloring: Compact Visualization for One-Dimensional Data, IEEE Information Visualization '05, 173-180, 2005.
- (5) A. Inselberg, B. Dimsdale, Parallel Coordinate: A Tool for Visualizing Multi-Dimensional Geometry, IEEE Visualization, 361-370, 1990.