

共生進化ウイルスを用いた GA の提案とそのスキーマ構築

A Proposal of GA Using Symbiotic Evolutionary Viruses and Its Adaptive Infection Technique

坂倉 義明 (立命館大学大学院) 亀井 且有 (立命館大学理工学部)

Yoshiaki SAKAKURA, Graduate School of Science and Engineering, Ritsumeikan University
Katsuari KAMEI, Computer Science, Ritsumeikan University

In this paper, we propose a unique genetic algorithm (GA) using symbiotic evolutionary viruses and discuss about its adaptive infection technique. Our idea is based on the scheme theorem. Generally, the search mechanism of GA is that genetic operations build schemes. However, it is difficult for genetic operation to deal with schemes directly. Our idea have both of genetic operations and virus operations which directly deal with schemes. Also, we apply the proposed GA to an optimal search problem, and show that solution search capability of the proposal method is better than that of SGA (Simple GA).

Key Words: Genetic Algorithm, The scheme theorem, Symbiotic evolutionary viruses

1. はじめに

筆者らはこれまで、GA におけるスキーマ定理に着想を得た共生進化ウイルスを用いた GA を提案してきた⁽³⁾⁽⁴⁾。本手法はスキーマ定理に着想を得ている。スキーマ定理とは、交叉や突然変異を行いながらルーレット方式の選択により、適応度の高いスキーマが増加するというものである⁽¹⁾⁽²⁾。しかしながら、選択、交叉はスキーマを直接的には扱っていない。また、突然変異はスキーマを直接的に扱っているが、そのランダム性から、スキーマの保持、有効利用といった動作は行わない。そこで、このスキーマ定理に着想を得、スキーマの保持および有効利用を目的とした手法として、久保田らの手法、半田らの手法が提案されている⁽⁵⁾⁽⁶⁾。久保田らの手法は、ウイルス感染によりスキーマを個体に直接的に伝播するモデルである。このとき、ウイルスは個体のスキーマを取り込み、それを他の個体に伝播する機能を持つ。これにより個体間の遺伝子列の水平伝播を実現している。また半田らの手法は、解探索を行う GA と共にスキーマを探索する GA をもつ。

一方本手法は、解候補として表現される個体群と個体の一遺伝子となり得るウイルス群からなる。本手法の進化は、ウイルスが個体に感染することによる個体の変化、その変化に基づくウイルスの学習という個体とウイルスの共生関係に基づく。このとき、ウイルスは自らの遺伝情報を個体に伝播する機能のみを持ち、ウイルスの学習とは、どのウイルスが解探索に有用なウイルスであるかの評価を得ることである。また本手法は、ウイルスの学習進度に合わせ、複数のウイルスを個体に感染させる適応的感染手法をもつ。これにより高オダのスキーマの構築・伝播をめざす。よって本手法は、ウイルスは自らを個体に伝播する機能だけを持つ点、ウイルスはそれにより学習を行う点、またその学習により有用なスキーマを探索する点で先に示した二つの手法と異なる。

以下、本手法を単峰性関数の最小値探索問題に適用する。そして、SGA (simpleGA) との比較実験を通して本手法の挙動・有用性について検証する。また、適応的感染手法について、その効果を検証する。

2. 遺伝的アルゴリズム

2.1 基本動作 GA の遺伝演算手順は、その使われ方により様々な形態がある。ここでは、基本的な GA の遺伝演算手順の概略を述べる。

まず、GA では、求めたいパラメータなどを遺伝情報として遺伝子で表す。この遺伝子の組み合わせを染色体とし、解候補となる一遺伝子とする。GA は、基本的に選択 (selection)、交叉 (crossover)、突然変異 (mutation) の 3 つの遺伝演算を行う。その手順は以下ようになる。

1. 初期集団の生成

2. 終了条件が満たされるまでループ

- (a) 適応度の評価
- (b) 選択
- (c) 交叉
- (d) 突然変異

まず、初期集団の生成を行う。一般には、決められた個体数の染色体をランダムに生成する。この際、個体数の決定や染色体の長さ、コーディングの方法などは設計者にゆだねられている。初期集団が生成されると、それぞれの個体について適応度の評価を行う。この適応度評価の方法は、取り扱う問題によって異なり、それぞれの問題の目的に対して適切な適応度関数を設定する必要があるが、確立された方法は存在しない。そして、各々の個体の適応度 f をもとに、交叉を行う個体をその適応度にあわせて選択する。選択には様々な手法が提案されているが、代表的な手法は適応度比例戦略である。これは、ルーレット戦略とも呼ばれ、各個体は適応度に比例して選択確率が与えられる。ある個体 i の適応度を f_i とすると、選択確率 $pselect_i$ は式 (1) で表される。

$$pselect_i = \frac{f_i}{\sum_{j=1}^n f_j} \quad (1)$$

その他よく用いられる手法として、エリート保存戦略がある。これは、集団中の最も適応度の高い個体をそのまま次世代に残す手法である。これにより、その時点で最も良い解が交叉や突然変異で破壊されないという利点がある。しかし、エリートの遺伝子が急速に集団内に広がる可能性が高いため、局所解に陥りやすい。そのため、一般的にこの戦略は、他の選択戦略と組み合わせられて使われる。交叉は、選択された 2 つの個体の染色体を組み替え、新たな個体を作る操作である。交叉方法は 1 点交叉 (単純交叉)、複数点交叉などがある。次に、ある確率で突然変異を加える。突然変異は、染色体のある遺伝子をほかの値に置き換えるものである。突然変異は、ある程度は必要であるが、高すぎるとランダムサーチになってしまう。また全くないと、初期集団が持つ遺伝子の組み合わせしか探索空間がないので、探索に限界がある。

2.2 スキーマ定理 GA において、個体にみられる特徴的な遺伝子構造をスキーマ (scheme) と呼ぶ。GA では、このスキーマを交叉により組み合わせることで、解の探索が進行すると考えられる。遺伝子をバイナリ表現した場合、スキーマは $\{0, 1, *\}$ を要素とした部分的なビット・パターンとして定義される。なお、* は、任意の値を表す don't care 記号である。例えば、スキーマ $H_1(1*0*01*)$ は、 $I_1(1100010)$ 、 $I_2(1001011)$ などの 8 つの遺伝子型 (3 つの '*' に対するビットの全ての組み合わせ) に含まれる。

さらに、スキーマの次元 (order, o) および定義長 (defining length, δ) は、次のように定義される。

- 次元: スキーマ H に含まれる don't care 記号 * 以外の記号の数 $o(H)$; 上の例では, $o(H_1) = 4$.
- 定義長: スキーマ H に含まれる don't care 記号を除く記号間の距離の最大値 $\delta(H)$; 上の例では, $\delta(H_1) = 5$.

$m(H, t)$ を, 世代 t において集団中に存在するスキーマ H の個数とする. さらに, $f(H)$ を, スキーマ H を含む個体の平均適応度とし, \bar{f} を集団中の個体全体の平均適応度とする. この時, 世代 $t+1$ に存在するスキーマ H の個数の期待値を式 (2) に示す.

$$m(H, t+1) \geq m(H, t) \frac{f(H)}{\bar{f}} \quad (2)$$

しかし, スキーマは交叉や突然変異で破壊される可能性がある. それらを考慮する必要がある. あるスキーマが交叉により破壊される確率は, 交叉確率を p_c とすると, $p_c \frac{\delta(H)}{l-1}$ で表せ, 突然変異による破壊の確率は, 突然変異確率を p_m とすると, $o(H)p_m$ で表せる. これらの遺伝演算によるスキーマ破壊を定量的に表したのがスキーマ定理である. このスキーマ定理を式 (3) に示す.

$$m(H, t+1) \geq m(H, t) \frac{f(H)}{\bar{f}(t)} \left[1 - p_c \frac{\delta(H)}{l-1} - o(H)p_m \right] \quad (3)$$

式 (3) では, 遺伝操作によって破壊されるスキーマのみが考慮されており, これらの演算によって新たに生成されるスキーマに関する項が含まれていない.

3. 共生進化ウイルスを用いた GA

3.1 感染操作 ここでは, 感染について説明する. 本手法では, ウイルスは遺伝情報および感染位置情報を持つ. 感染位置情報とは, そのウイルスが感染することのできる個体の遺伝子座をさす. ここで, 遺伝情報は 1bit のバイナリで構成され, その種類数は個体のコーディングに依存する. 本手法では, 候補をバイナリで表現できる問題に対しアプローチを行っているため, ウイルスの持つ遺伝情報は 0 もしくは 1 となる. よって, 同じ感染位置情報 p_k をもつウイルスは 2 つ存在する. 以下, これらのウイルスを対立ウイルスと呼ぶ. このように構成されるウイルス群を Fig.1 に示す. ここで, n は個体の遺伝子長となり, $k = 1, 2, \dots, n$ である. また感染とは, 遺伝情報 g_j , 感染位置情報 p_k を持つウイルス v_k の場合, 感染される個体 m の遺伝子座 p_k の遺伝情報をウイルスの遺伝情報 i に書き換えることとなる. これを Fig.2 に示す. このとき, ウイルスは評価値 $virusvalue$ を得る. これは感染による個体の適応度変化に依存する. 感染するウイルスを k , 感染された個体を m , その集合を S , 個体 m の感染前後の適応度をそれぞれ $fitness_m$, $fitness'_m$ とする. このときウイルス k の評価値 $virusvalue_k$ は Eq.(4) となる.

$$virusvalue_k = \frac{\sum_{m \in S} (fitness'_m - fitness_m)}{\# \{S\}} \quad (4)$$

3.2 適応的感染手法 ここでは, ウイルスの学習進度に合わせ, 複数のウイルスを 1 つの個体に感染させる手法について説明する. これを適応的感染手法と呼ぶ. Fig.3 に適応的感染手法を用いた感染処理手順を示す. Fig.3 において,

- 1 感染される個体をランダムに選択する.
- 2 すべてのウイルスから, 対立ウイルスの $virusvalue$ の差が一定以上のものを探す. ここで一定の差を持つ対立ウイルスとは, 対立ウイルスをそれぞれ v_k, v'_k , そ

Table 1 Simulation parameters

Gene coding	Gray coding
Selection	Roulette selection
Crossover	Two point crossover
Mutation	Bit revers mutation
Population size	100
Gene length	100
Crossover rate	0.9
Mutation rate	0.01
Termination condition	1000 generations
Number of infection times	5
Minimum $virusvalue$	$0.1 \cdot 10^{-10}$
Parameter C	100

それぞれの $virusvalue$ を $virusvalue_k, virusvalue'_k$ とするとき, $virusvalue_k \leq virusvalue'_k$ であったならば, $virusvalue_k - virusvalue'_k \leq virusvalue'_k \cdot C$ となる対立ウイルスのことを指す. なお, C はパラメータであり, あらかじめ決定しておく必要がある. そして, これらの対立ウイルス間で, $virusvalue$ が大きい方のウイルスを選択する.

- 3 手順 2 で選択されたものを除く対立ウイルスに対し, 対立ウイルス間の $virusvalue$ の差が小さければ小さいほど選択されやすい選択確率を用いて対立ウイルスを 1 つ選択する. 次にその対立ウイルス間で $virusvalue$ が大きければ小さいほど選択されやすい選択確率を用いてウイルスを選択する.
- 4 手順 2, 3 で選択されたウイルスの感染・評価を行う.

以上の手順で 1 つの個体に感染を行う. これを規定感染回数繰り返して, 1 世代における感染の終了とする.

3.3 GA の処理手順 感染を用いた GA の手順を Fig.4 に示す. Fig.4 において, まず GA と同様に初期集団の生成, 選択, 交叉, 突然変異を行い次世代の個体群候補を生成する. そしてその個体群に対し感染を行い, 次世代の個体群を生成する. これを終了条件が満たすまで行う.

4. 数値実験

4.1 実験環境 本論稿では, 関数の最小値探索問題を通して提案手法と SGA の比較実験を行い, 本手法の有用性, および適応的感染手法によるウイルスの挙動を検証する. 使用する関数は Sphere 関数を改良したものをを用いる. これを Eq.5 に示す. なお, n は次元数であり, 本実験では $n = 10$ とした. また定義域は $-5.11 \leq x_i \leq 5.12$ とした.

$$f(x) = - \sum_{i=1}^n (x_i + 1)^2 \quad (5)$$

また GA への適用について, 使用したパラメータ等を Tabel 1 に示す. ここで number of infection times は 1 世代に感染される個体数となる. minimum $virusvalue$ は, 個体を改悪する感染によりウイルスの選択確率が負になることを防ぐために設定した.

4.2 実験結果 Fig.5 に本手法と SGA の最良値の推移を示す. ここで, 探索初期段階では, 本手法, SGA 共に同じような最良値をとり, 探索が進むに連れて 90 世代目あたりから, 本手法が SGA より良い最良値をとることがわかる. また Fig.6 より, 手順 fixed viruses で選択されたウイルス数が探索が進むに連れて増加している. これによりウイルスの学習が行われてい

る事がわかる。また、個体に高オーダのスキーマが伝播されることがわかる。ここで Fig.7 より、同感染位置情報をもつ両方のウイルスの *virusvalue* が *minimum virusvalue* をとる感染位置数は探索初期段階で増加を示し、探索が進むに連れ 400 世代目あたりから減少しているのがわかる。一方 Fig8 より、同感染位置情報をもつウイルスの片方が *virusvalue* が *minimum virusvalue* をとる感染位置数は、探索初期段階から増加を示し、探索が進んでも減少する事が無い。これにより、同感染位置情報をもつウイルスの片方が有用なウイルスであれば、学習が進むにしたがってそのウイルスは良い評価を受けることがわかる。しかしながら、探索中期では有用なウイルスとそうでないウイルスが、おなじ *minimum virusvalue* をとることがある。これは、有用なウイルスが一度の個体を改悪する感染により、その *virusvalue* が *minimum virusvalue* となる可能性がある為である。

5. む す び

本稿では、共生進化ウイルスを用いた GA を提案し、その有用性・挙動について検証した。そこで、適応的感染手法にもとづく個体への高オーダのスキーマの伝播、有用なウイルスの学習の進行について確認した。しかしながら、探索中盤で有用なウイルスとそうでないウイルスの評価が同じになることがある。今後の課題として、これら有用なウイルスとそうでないウイルスに異なる評価を与えることの出来る評価・学習手法の検討が挙げられる。

参考文献

- (1) D.E.Goldberg. Genetic Algorithms in search, optimization and machine learning, Addison Wesley, 1989
- (2) 北野 宏明 編 "遺伝的アルゴリズム", 産業図書, 1993
- (3) 坂倉義明 亀井且有 "共生進化ウイルスを用いた進化的計算の提案", 第 17 回ファジィシステムシンポジウム講演論文集, pp.94-97, 2001
- (4) 坂倉義明 亀井且有 "共生進化ウイルスを用いた GA の提案とそのスキーマ構築", 第 46 回システム制御情報学会研究発表講演会講演論文集, pp.467-468, 2002
- (5) 久保田直行 福田敏男 "ウイルス進化論に基づく遺伝的アルゴリズムの巡回セールスマン問題への応用", 計測自動制御学会論文集, 第 34 巻第 5 号, pp.408-414, 1998
- (6) 半田久志 片井修 馬場則夫 榎木哲夫 小西忠孝 馬場充 "共生進化の機構を用いた遺伝的アルゴリズムの提案", 計測自動制御学会論文集 第 35 巻第 11 号, pp.1438-1446, 1999

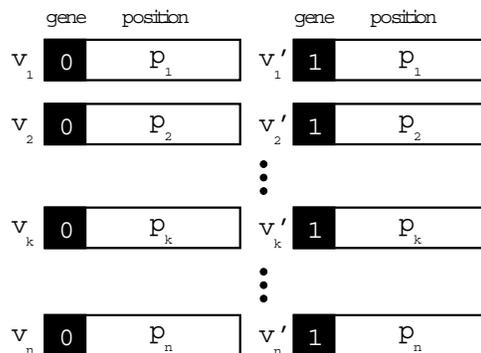


Fig. 1 Viruses

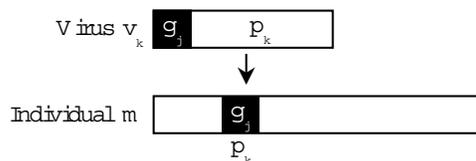


Fig. 2 The Infection

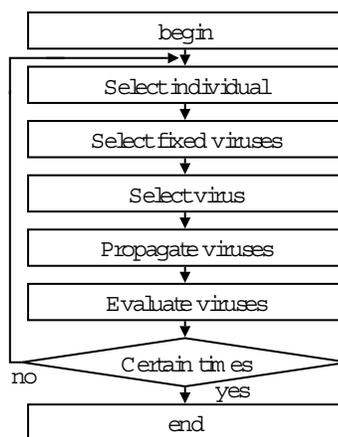


Fig. 3 Adaptive infection procedure

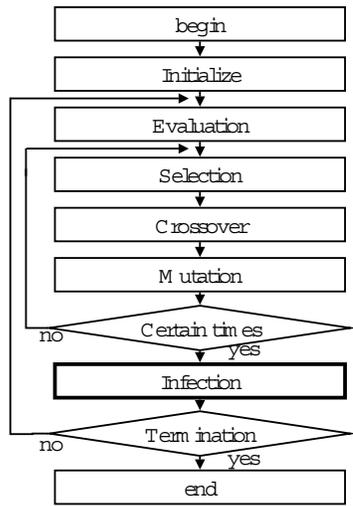


Fig. 4 GA procedure including infection

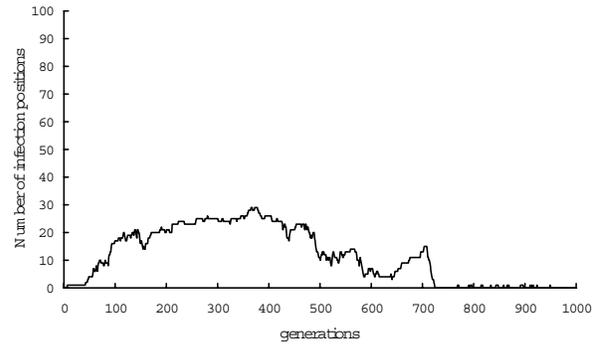


Fig. 7 The number of gene positions of each virusvalue is minimum

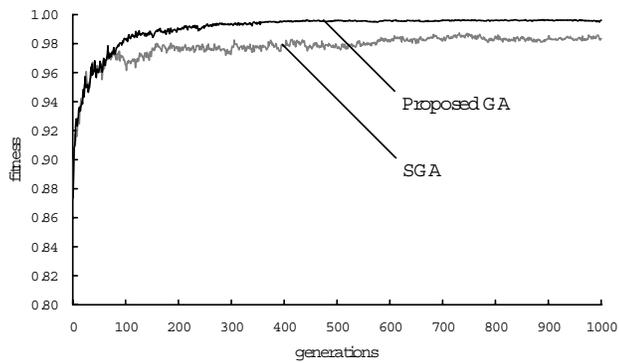


Fig. 5 Best fitness transition

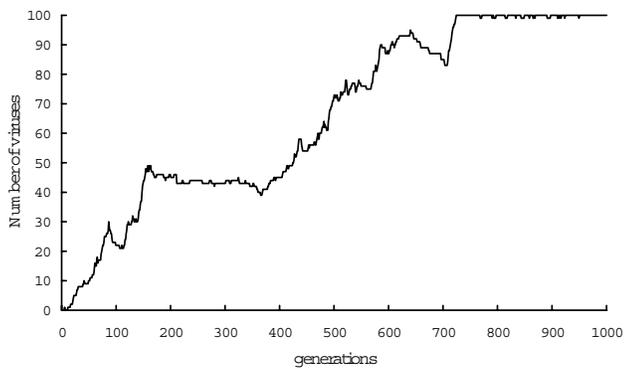


Fig. 6 The number of fixed viruses

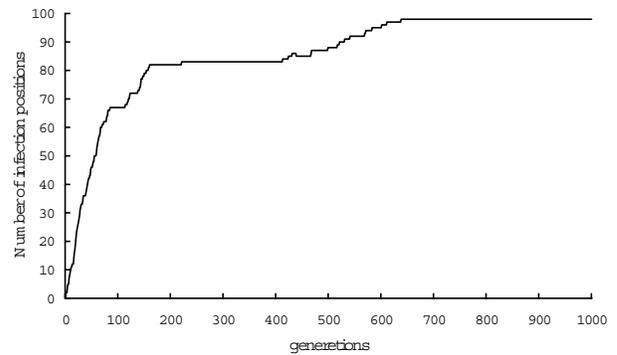


Fig. 8 The number of gene positions of either virusvalue is minimum