

実世界における遺伝的・ 進化的アルゴリズム

David E. Goldberg
University of Illinois at Urbana-Campaign

翻訳：安達 統衛
(株) 富士通研究所

★ なぜ遺伝的・進化的 アルゴリズムなのか

1992年以來、遺伝的アルゴリズム (Genetic Algorithms, GAs) — 自然淘汰と遺伝のメカニズムに基づいた探索手法 — の話をするために私は何度も日本を訪れてきた。はじめて来日した頃は、遺伝的・進化的アルゴリズム (Genetic and Evolutionary Algorithms, GEAs) を使っているのは一部の専門家に限られていたが、今日、日本の研究者や実務家はGEAの最先端研究や幅広い分野での応用を上手に進めている。従来の工学やOR分野の最適化問題のもとより、薬品設計、金融予測、データマイニング、作詩、作曲などの新しい分野に至るまで、遺伝的・進化的アルゴリズムは世界中で人々の注目を集めている。もちろん、ファッションやおもちゃなどと同様に科学技術にははやりすたりがあるので、GEAがツールとして永遠なものとなるのか、あるいは、現れては消えていった多くの手法と同様に一時的な

ものなのか、多くの実務家は関心を寄せている。それゆえ、遺伝的・進化的計算 (Genetic and Evolutionary Computation) の技術と理論の現在と将来について論じることはとても大切なことであろう。このエッセーで、以下の3つの問いをあげ、答えることとしたい。

- (1) なぜGEAは広く普及し知られるようになったか？
- (2) この普及は本物なのか、あるいは一時的なはやりなのか？
- (3) GEAの技法や応用に将来何が留意されているのか？

本文の主張の要点は、GEAがこれほど普及しているのは偶然でも流行でもなく、人々の生活を良くしていく上でますます重要な役割を果たすにちがいない、ということである。特に、GEAについて良い経験も悪い経験もしたことのある実務家にとっては、これは強い仮説と思われるかもしれないが、現在広く使われている技法は氷山の一角に過ぎず、たとえば小規模問題から実世界の実用規模問題にスケールアップした時に起こ

る課題が、新時代のGEAではたくみに回避できることが最新の研究から示唆されている。さらに、応用の舞台に転じれば、多くの物事を決める時に最初に考慮される要因はしばしば「経済性」であり、技術的なものではない。この面でもGEAは提供すべきものがたくさんある。

以下では、まず、GEAが人類の「斬新な何かを見出す行為 (以下、発明と呼ぶ)」と直接的な対応関係を持ち、この行為にはいくつかの「形態」があることを議論する。これがGEAが広く気に入られている理由の1つと思っている。次いで、ユーザがGEAを用いるようになる5つの「動機」について、上記の3つの問いに答えていく。実世界問題を抱えた現実のユーザは結局GEAをなぜ選ぶのか？ 5つの動機を詳しく調べることで、なぜ人々が今日GEAを利用しているのか、なぜ将来も使い続けるのだろうか、といったことをもっとよく理解できるであろう。

★ GEA と発明の原理

GEAの基礎的なメカニズムについては皆さんはすでによく知っていると思うので、誌面の都合もあって詳しくは述べない。簡単には、問題の解（最適ではないかもしれない解）(solution) をある規則で染色体(chromosome) と対応づけ、染色体の良さ悪さを適応度(fitness) と呼び、その染色体の表す解の良さ悪さ(普通は目的関数の値) で計る。さまざまな染色体からなる染色体集団と、遺伝的操作(genetic operator) と呼ばれる染色体集団を変化させる操作を使って、繰り返し染色体集団をより最適な方向に導いていく。遺伝的操作として、淘汰(selection)、組換え(recombination)、変異(mutation) が多く使われている。

では、これらの遺伝的操作によってなぜ探索がうまくできるのだろうか？ 個々の操作だけでは大した機構を持たずあまり機能しないのに、3つの操作を組み合わせることでどうしてうまく機能するのか？ 何かミステリーに思える。1983年³⁾以来、このミステリーを説明するために、私は「遺伝的アルゴリズムの正しさを直観的に基礎づける原理」について考察を続け、そして、GEAが人類の発明のやり方と直観的に同じ原理であると考えているに至っている。特に、淘汰と変異からなる過程を「改善」型の発明、淘汰と組換えからなる過程を「革新」型の発明とそれぞれ位置付けて、これらが人類の発明形態の一面をよく表すものであると考えている。

●淘汰+変異=改善型発明

淘汰と変異の組合せは山登り法となる。変異は現在の解の近傍に少し違った解を作り出し、淘汰はこれらの変化を高い確率で受け入れる。その結果、より良い解に向かって登っ

ていく。人類はこれをまったく自然に行っており、総合品質管理の用語では、この継続的な改善作業は「カイゼン(改善)」と呼ばれる。私がこの直観的原理を考察した時は、自分自身による観察からであったが、同様の考えを持っている人々がいた。たとえば、イギリスの作家・政治家のBulwer-Lytton¹⁾ は以下のように述べている：

発明は、優れたモデルから派生した優れたものか、優れたモデルを詳細にしたもの以上の何ものでもない。...もし、高尚で一般的ならば、模倣からオリジナリティが生まれる可能性が最も高い。

この定性的な表現はアルゴリズム的な記述からかけ離れているが、この文から淘汰や変異の響きを聞き取れるであろう。たしかに、局所近傍を調べていくのは改善のための強力な手段である。一方で、局所解を見つけた時に他にジャンプするうまい手段が見つからなければ、かなり限られた範囲に滞る傾向がある。

●淘汰+組換え=革新型発明

淘汰と組換えとを組み合わせれば、この種のジャンプが実現できるであろう。この組合せ効果を人類の革新型発明過程での効果に例えることで、このことが理解できる。革新型発明はどのようにしてなされるのだろうか？ 通常、一方からある概念——つまり良い解の特徴——を取り出し、他方からもある概念を取り出して両者を並置し、それにより並置した組合せが、個々の概念よりも良くなっているかもしれないと、思惑を立てるのが自然である。この考えも自分自身による観察からであったが、同じ思いの人々がいた。たとえば、フランスの数学者Hadamard⁷⁾ は以下のように述べている：

発見を純粋に偶然に帰着させる可能性はすでに排除されていること

はすぐに分かるであろう...実際、発明や発見は、数学や他の分野で、考えを組み合わせることによって生じる。

同様に、フランスの詩人・哲學家Valéryも似た観察をしている：

何かを発明するには2つものがある。1つは組合せを作り出すこと、もう1つは、作り出されたものの中から望みのものあるいは重要と思うものを選び再編成することである。

言語での記述は現代的ではないが、直観的原理に似たものが明瞭に語られていることが分かる。

次に、実世界においてGEAを利用する動機を調べていくことにしよう。

★ GEAを使う5つの動機

前章で、GEAが我々の発明の仕方・感性に強く訴えかけていること、またこれがGEAを受け入れる動機になっていることを提唱した。しかし、もっと実務的には、どんな動機でGEAを使うようになるのだろうか？ たしかにユーザの数だけ答えがあるが、一般化ができると思われる。ここでは5つのタイプの動機を挙げて調べていく：

1. 流行からの動機
(motive from buzz)
2. 自然界の現象からの動機
(motive from nature)
3. 人工システムからの動機
(motive from artificial system)
4. 対象問題への適切性からの動機
(motive from competence)
5. 経済性からの動機
(motive from economics)

●流行からの動機

新しいユーザを魅きつける要因の1つは、私が流行と呼ぶものである。GEAが印刷メディアや電子メディア

の注目を浴び、GEAのさまざまな発見や発明の評価が、人工生命や複雑系の普及により広まっている。これらの評価によって新しいユーザはGEAへと導かれるが、その評価だけですぐに問題が解けるわけではない。問題を提起し、いろいろな手法を取り入れ、やっと結果を得ることができる。だから、流行からの動機は長続きしない。

●自然界の現象からの動機

では、長続きするものは何か？我々の注意を引き留める要因の1つは、「科学的妥当性」である。Darwin以後、この惑星上のさまざまなしかもよく順応した形態を持った生命は自然淘汰と遺伝によって創造された、と広く認識されている。これから暗示されることは、選択肢の中から選びだすという自然界の探索アルゴリズムを「用いて」、人類の抱える問題の解法に応用することができるかもしれない、ということである。考えを持つことと実行に移すことは別なことであるが、この暗示は重要である。なぜなら、たとえ究極的な遺伝的アルゴリズムがまだ得られていないとしても、正しい道筋にいることを知る、何か存在証明になっているからである。たしかに、人々は鳥を観て人間が飛べることを夢見てきた。そして長い間試みはすべて失敗に終わった。飛ぶことができるものが存在するということが、双対的役割を持っている：(1)飛行機の詳しいデザインに対して科学的な示唆を提供すること、(2)失敗が積み重なるほど、探究や実験が持続すること。同じようにして、研究者や実務家は自然界の事例から示唆を受け、たとえ努力の結果が思うようにならなかった時でも続けようとするのである。

●人工システムからの動機

アイデアや存在証明の源として自然界は、主にGEAのデザイナーや研究者にとって、示唆に富むものである。しかし、実務家の動機は、伝統的な最適化やOR手法の限界に根ざしている。たくさん手法がすでにある。たとえば、線型制約を持つ線型問題には線型計画法を使うことができ、段階に分解できる問題には動的計画法が、非線型制約を持つ非線型問題には(ときどき)非線型計画法が使える。しかし、特定の問題のクラスに使える手法のリストを持っている事実それ自体が問題である。伝統的手法は特定の問題クラスによく機能するが、その手法の前提に合わない問題に対しては結果はみじめなものになってしまうにちがいない。幅広い問題クラスにもうまく機能してくれる人工的な探索・最適化の手順はないのだろうか？人工的な遺伝的・進化的手法はこの問いかけに対する潜在的な答えになっている。なぜなら、自然システムの進化は、生物種の間で不変なメカニズムによって生じ、進化の過程では、環境にはほとんどかわりなく、同一か類似の手順が使われているからである。多くの人々が、納得できる効率と品質を持った、まさに幅広く使える解法を求めてGEAに目を向けているのである。

●対象問題への適切性からの動機

ロバスト性 — 品質と効率が約束されていることが、なるほど、多くの実務家をGEAに引き込んでいる。が、応用の途上で奇妙な状況を経験することもある。その1つに小規模な問題ではGEAはまったくうまく動くのに、問題が大きくなるのは難しくなると、探索時間が増え、品質が落ちてくることである。このスケールアップ問題に対するユーザの反応はさまざまである。あるものは、う

まくいくまで、確率を変えて、遺伝的操作や染色体コーディング法をいじくり回す。あるものは、あきらめてGEAを捨てる。また、あるものは、なぜロバストなアルゴリズムがこんなにお粗末なスケールアップ結果を示すのか、疑問を抱く。

これらの難問は何年もの間都合よく隠されていたが、固定された交叉や変異を持った単純なGEAが使える範囲はかなり限定されていることを私は知った¹³⁾。この主張を支持する数理的解析を行い、順応的・自己順応的な遺伝的操作がこれらの困難を効果的に克服できることが分かった。このように、難しい問題を速く信頼性高く正確に解くことのできるGEAを「問題に適切なGEA (competent GEA)」と呼ぶことにしよう。

イリノイ大学ではこの流れに沿い、messy GA⁵⁾、fast messy GA⁴⁾、gene expression messy GA⁹⁾、linkage learning GA⁸⁾、linkage detection GA¹¹⁾、Bayesian optimization algorithm¹²⁾、などの一連の研究を展開している。これらはまだ現場で試されていないが、スケールアップ問題が軽減され、問題規模を表す変数の数の二乗以下程度の計算時間で難解な問題が解けるであろう。

●経済性からの動機

これまでの議論は、かなり崇高な理由であった。一方、多くの実務家にとっての要点はしばしば決算書の最後の数字にある。つまり、遺伝的最適化から得られる経済的利益に関心を抱いている。ある場合では、GEAを使うことで直接的な経済的利益を得ることができるかもしれないし、また違う状況下では、直接的な経済的利益はそれほどでもないが、GEAを選ぶことが重大となるかもしれない。3つの状況を考えてみよう。

a. 手法に対する投資の経済性

- b. モデルに対する投資の経済性
- c. GEA 高速化の経済性

a. 1つ目の理由は最適化手法に関してである。(1) GEAのような幅広く一定水準の機能を備えた手法、あるいは、(2) ORないしは伝統的最適化手法から持ってきた雑多な技法一揃い、を利用して解法を良くしようとする時、1つの手法を学習して使うことの方がいろいろな技法を使うよりも必要な投資は少なく済むはずである。いろいろな技法を使う場合には、技法それぞれを学ぶ必要がある他、どの技法を選ぶかも調べなければならない。

b. 手法に対する投資コストはたしかに大きい、多くのユーザにとっては投資の大半は問題の「モデリング」あるいは「シミュレーション」に費やされる。ほとんどの場合、複雑な問題では必然的に目的関数もかなり複雑になる。それらの目的関数は、有限要素モデルや非線型解法近似、離散事象シミュレーションなどに頼るかもしれない。最適化や設計にこのようなモデルを利用する前に、データ入力やテストケース実行、調整、そして解析など、問題のモデルを準備するためかなりの時間と労力を費やさなければならない。モデリングに投資した後に、最適化手法が適用できる形式にモデルを変形するためにさらに投資をしなければならない。多くの最適化手法ではこのモデル変換が必須である。一方、GEAでは、モデリングで得られた目的関数を実質的に修正あるいは変換することなく、そのまま評価に使うのである。

c. しかし、GEAに良いことばかりがあるわけではない。GEAは探索空間についてあまり前提を置かず、GEAと評価のインタフェースは目的関数

の値だけである(導関数や高次情報がいらぬ)。そのためGEAは何百何千回の目的関数の計算が必要になる。問題に適切なGEAを使えば計算回数を減らせるかもしれないが、それでも、大規模問題ではかなりの回数が必要になる。GEAを高速化する方法には、(1)空間の効率化(並列化)、(2)時間の効率化、(3)評価の緩和、(4)ハイブリッド化、が挙げられよう。これらについてはたとえば文献2)、6)、10)がある。

★ 計算論的発明と創造の黄金時代に向けて

先に述べたように、発明の計算論的な一表現形式としてGEAがたしかに役立ち、難解な問題を日課のようにすばやく信頼性高く正確に解くことができる competent GEAが完成すれば、計算論的発明の視点から人類の知性が議論されるようになるであろう。この展望には興奮するが、一方では危うさもある。この計算論的な発明と創造の時代について、私は次のことを予言する：

- a. competent GEAは問題解決の当たり前の部品となるだろう。
- b. competent GEAをデザインするのに使われる技法、分析、方法論を拡張することで、創造的なアルゴリズムが展開できるだろう。
- c. ビジネスでの勝者敗者はこれらの技法をすばやく習得しているかどうかで(少なくとも部分的に)決まるだろう。
- d. 計算論的発明と創造は、問題解決における効用以上の科学的重要性を持つであろう。それぞれ少し詳しく述べていこう。

a. GEAがビジネスの道具に含まれるようになるのはますます明らかになってきている。私は、世界の多く

の人々が、有限要素法のアナロジーでGEAの重要性を述べているのを見てきた。電磁場から流体力学、固体力学などの分野で、有限要素法が日常的に使われている。これと同様に実務家は、新しいものを産み出すときの計算論的なツールとしてGEAを考えるようになりつつある。

b. GEAの理論、手法、実装でなされてきた歩みはまだ最初の段階であると思う。今手元にある技法は、今後数カ月数年のうちに明らかになる何かの影だと信じている。このことを事例を挙げずに証明するのは難しいかもしれない。現存のGEAが計算論的発明の1形態であるというのは正しいが、現存のGEAが創造的であるというのは間違っていると思う。我々人類は「創造的」という言葉を革新的以上のものに対して使うようだが、私は知識を1つの領域から他の領域に「移す」ことのできる人々やものに対して創造的という言葉を使うことを提案したい。創造的書物や芸術は、しばしば他の業績に触れ、考えや概念を持ち込んでいる。創造的な科学的発見は自明ではなく、しばしば無関係と思われていたアイデアを移行することでなされている。同様に、1つの領域から他の領域にアイデアを移行できる「創造的アルゴリズム」を我々は手に入れることができると私は信じる。これらの考えはGEAの方法を構築し、たぶんニューラルネットやファジー論理などと統合されて、これまで想像できた以上の強力なシステムを我々に与えてくれるだろう。

c. 先の2つの提案が正しければ、ビジネスマン達がこれらの方法をすぐに採用しようとするのは明らかである。それにより企業や労働者は競争利益が得られるからである。企業での計算論的発明や創造について当

初は専門家が担当するだろうが、ツールがよくなるにつれて、広範囲に適用されるようになるであろう。現時点で、GEAがどこに採用されているのか、どのように広まっているのかを、企業自ら調べるべきである。特にこれらの概念は大学実習ではまだ広く教えられていないので、GEAの教育が入社数年間は必須である。

d. GEA やそれを拡張したものは問題を解く効用以上に、科学への影響力を持つだろう。GEAの助けによって発明的で創造的アルゴリズムが得られるのならば、人類の発明と創造性の「モデル」としてこれらのアルゴリズムを使うことにより、多くの科学的研究分野が恩恵を受けると思われる。たとえば、社会科学では、利害関係や、取引、社会的相互作用、心理的生活などの問題について、人々は互いにどんな関係にあるかをモデル化している。しかし、現在、これらの研究は人類の発明と創造性の明示的なモデリングなしには進展しない。そして、GEAのモデルはこのギャップを埋めてくれるにちがいない。実際、経済学、政策科学などでの人工的エージェントはこのやり方を見せ始めている。この傾向は、(1) これらの科学的目的において我々のモデルが有効であること、(2) 発明的創造的プロセスをメカニズムの視点で理解することによって社会科学分野における考え方を根本から変えてしまうこと、がもっとよく理解されると私は信じる。さらに、功利的アプローチにもかかわらず — いや、功利的アプローチであるゆえに — 直接的な科学的探究だけでは適切に理解できない重大な生物学的過程を明らかにするのをGEA研究が助けていることを、生物学者が、やや遅れて、理解するようになるだろう。

★ **GECCO :**
統合された国際会議

さて、本年7月13日から遺伝的・進化的計算の国際会議 GECCO-99 がアメリカ・オーランドで開催される。この分野の研究者や実務家がさまざまな最先端の成果を持ち寄る世界最強の国際会議である。またこの会議につき同地で AAAI-99 が開催される。進化戦略 (Evolutionary strategie) の創始者 Ingo Rechenberg や遺伝的アルゴリズムの創始者 John Holland の招待講演、25 の無料チュートリアル、15 のワークショップ、ポスターセッション、一般セッションなどが予定されている。参加案内など詳しくは <http://www-illigal.ge.uiuc.edu/gecco/> をご覧いただきたい。

Acknowledgements

My research was sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant F49620-97-0050. Research funding for my work was also provided by a grant from the U. S. Army Research Laboratory under the Federal Laboratory Program, Cooperative Agreement DAAL01-96-2-0003. The U. S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notification thereon.

The views and conclusions contained herein are my own and should not be interpreted as necessarily reprinting the official policies or endorsements, either expressed or implied of the Air Force Office of Scientific Research, U. S. Army Research Laboratory, or the U. S. Government.

参考文献

- 1) Asimov, I. and Shulman, J. (eds.) : Isaac Asimov's Book of Science and Nature Quotations, Weidenfeld & Nicolson, New York (1988).
- 2) Cantú-Paz, E. and Goldberg, D. E.: Modeling Idealized Bounding Cases of Parallel Genetic Algorithms, Genetic Programming 1997: Proceedings of the Second Annual Conference, pp.353-361 (1997).
- 3) Goldberg, D. E.: Computer-Aided Pipeline Operation Using Genetic Algorithms and Rule Learning, Doctorial Dissertation, University of Michigan, Ann Arbor (1983).
- 4) Goldberg, D. E., Deb, K., Kargupta, K. and Harik, G.: Rapid, Accurate Optimization of Difficult Problems Using Fast Messy Genetic Algorithms, Proceedings of the Fifth International Conference on Genetic Algorithms, pp.56-64 (1993).
- 5) Goldberg, D. E., Korb, B. and Deb, K.: Messy Genetic Algorithms: Motivation, Analysis, and First Results, Complex Systems, 3 (5) , pp.493-530 (1989).
- 6) Goldberg, D. E. and Voessner, S.: Optimizing Global-local Search Hybrids, GECCO-99: Proceeding of the 1999 Genetic and Evolutionary Computation Conference, in press.
- 7) Hadamard, J.: The Psychology of Invention in the Mathematical Field, Princeton University Press, Princeton NJ (1945).
- 8) Harik, G.: Learning Gene Linkage to Efficiently Solve Problems of Bounded Difficulty Using Genetic Algorithms, Doctorial Dissertation, University of Michigan, Ann Arbor (1997).
- 9) Kargupta, H.: Search, Evolution, and the Gene Expression, Messy Genetic Algorithm, Los Alamos Ntional Laboratory, Los Alamos NM (1996).
- 10) Miller, B. L. and Goldberg, D. E.: Optimal Sampling for Genetic Algorithms, ANNIE96: Proceedings of the Artificial Neural Networks in Engineering Conference, pp.291-297 (1996).
- 11) Munemoto, M. and Goldberg, D. E.: Identifying Linkage Groups by Nonlinearity/Nonmonotonicity Detection, GECCO-99: Proceedings of the 1999 Genetic and Evolutionary Computation Conference, in press.
- 12) Pelikan, M., Goldberg, D. E. and Cantu-Paz: BOA: The Bayesian Optimization Algorithm, GECCO-99: Proceedings of the 1999 Genetic and Evolutionary Computation Conference, in press.
- 13) Thierens, D. and Goldberg, D. E.: Mixing in Genetic Algorithms, Proceedings of the Fifth International Conference on Genetic Algorithms, pp.38-45 (1993).

本文は原著 IlliGAL Technical Report No.99013 を意識したもので、原著は <http://www-illigal.ge.uiuc.edu/> から参照できる。

(平成 11 年 4 月 30 日受付)