

## AIマップ

# 「GAの研究に取り組んで十余年」へのコメントと回答

## Comments on “Historical Roadmap of Research on Genetic Algorithms” and Their Answers

### コメント

◇ 元田 浩\*1

#### 1. はじめに

著者は機械学習やデータマイニングの研究をしているが、GAの研究が専門ではない。したがって、ここ十数年の進展も把握していないし、詳しい技術的なことは理解できない。正直なところ、多くの略号が出てきて、何度も読み返さないと、どれがどの手法のことだかなかなかわからず読むのに苦労した。まず、小林氏が最初にGAに触れたとき、その魅力を直感で感じ取り、それ以来十数年研究室で全力をあげて取り組んで来られ、その間、すばらしい成果を出してきたことに敬意を表したい。一仕事10年というが、真面目に取り組めば、最初は遅れをとっていても、世界のトップに躍り出る一つのいい実例である。適切なコメントはできないが、以下、著者が感じることを述べさせていただく。

#### 2. GAと知識ベース

GAの研究の進展の経緯を眺めていて、知識ベースの研究の進展の経緯とよく似ていると感じた。知識ベースシステム（エキスパートシステム）に最初に期待されたのは、悪構造でアルゴリズムが存在しない問題に対して有効に働く万能薬ということであった。しかし、知識獲得のボトルネックが叫ばれ、問題解決の過程を詳細に分析した結果、問題が構造化され、タスクとそれを実現する手法が区別され、それが概念モデルやビルディングブロックに集約され、結局のところ、細部まで透明なものをつくらなければだめであるという結論に達し、それをいかにシステムティックに実行するかの方法を世界中で模索してきた。

GAも複雑すぎて最適解が単純なアルゴリズムでは求められない問題を、解の基本要素の交叉、突然変異による世代交代を繰り返し、生物界の自然淘汰に似た過程を経て、膨大な探索空間の全部を網羅することなく、解に到達する方法を提供するものである。一般には最適解に到達する保証はないが、小林氏らは、高次元、非線形性、多峰性、悪スケール性など個別の問題の性質に対する深い洞察により、新しいコーディング、交叉、突然変異、世代交代の仕方を次々と考案してきた。多くの性質の違いの問題に対して、それぞれ適切な方法が見つかった段階で全体を眺めてみると、それぞれの方法の果たす機能が理解でき、解くべくして解けるしくみが出来上がる。そうすれば、GAの枠組み（どうすれば何ができて、何が難しく、どのような工夫が有効かなど）が理解できた気分になる。小林氏の自負もこのような境地に到達したことに起因するものであろう。

#### 3. メタGAの可能性

しかし、このような洞察から新たな提案に至る思考過程も最適化の一つである。GAが複雑な問題の最適化を標榜しているなら、自分自身の動きを分析して、小林氏らが提案するに至った新方式を多数の試行の結果、自力で見つけ出してくれることは不可能なのであろうか？ 最初に述べてある、初期の科学者が夢見たビルディングブロックのプリミティブは設計者が与えるものではなく、遺伝的操作によって発見し、その組合せ方も遺伝的操作に委ねるべき、という理想は、やはり夢なのであろうか？ 自然界は最適な進化過程を自力で見つけているように思える。メタGAなるものを考案する余地はないのであろうか？ 結局、それを設計するのでも人間なのであろうか？

#### 4. GAと欲張り探索

小林氏らが過去に解いてきた問題の多くは、巡回セールスマン問題やジョブショップスケジューリング問題などの離散組合せ最適化問題と複雑な関数の最大値探索やレンズ設計などの実数関数最適化問題である。これらは、欲張り探索などほかの方法では解けなかった（あるいは

\*1 大阪大学産業科学研究所

解けても時間がかかりすぎた) 問題を実用的な時間で解いたという意味でフレクスルーを実現したものといえる。小林氏らが挑戦した問題の一つに決定木の生成がある。生成し得る決定木の数は属性数の指数関数で増加するので、決定木の構築には欲張り探索が用いられる。決定木は欲張り探索が成功した代表的な例である。小林氏らは、決定木のサイズ最小化と予測誤差最小化の二つの目標をもつ多目的最適化問題として GA を適用したが、著者の感覚では、目的は、訓練例に過剰適合することなく、未知のデータに対して予測誤差を最小化することであり、小さな決定木がいいという性質は、その結果としてついてくるものである。真の誤差の評価が難しいので、見掛けの誤差(訓練例に対する予測誤差)とモデルの複雑さのトレードオフの問題を置き換えているに過ぎず、MDL 基準は代替手段である。実際、小さな決定木がいつもいいとは限らない。与えられた属性を使う決定木を GA で構築するメリットはあまりないように思える。むしろ、非常に多くの属性の中から重要なものを選択したり、もとの属性ではうまくクラス分離できない場合に、属性を組み合わせて新しい属性を構築するのに GA は威力を発する [Vafaie 98]。実際、属性構築は組合せ探索問題で、専門家主導、欲張り探索、GA が使われているが、今のところ、GA が良い結果を出している。

## 5. ランダム性

GA の特徴は制御されたランダム性であろう。ランダム性の効用は最近の機械学習でも注目されている。決定木に関しても、ランダム性を随所に入れた研究が増加している。それはバギングやブースティングに代表されるコミッティー学習の結果を見てもわかるように、最も簡潔なモデルが最適であるとは限らないからである [Domingos 98]。例えば、極端な例ではランダムに各ノードで属性を選択して構築されたランダム決定木を多数集めた、ランダムフォレストというものもある。驚くべきことに、集団では非常に良い精度を出す [Breiman 01]。統計の分野では粒子フィルタと呼ばれているものが、非線形システムの状態系列推定に使われている [Godsill 01]。GA の遺伝子が非線形微分方程式に相当しているようなものである。このような研究の流れを見ると、複雑がゆえにランダム性に頼らざるを得ないのか、それともランダム性の追求がもの本質なのかと考えさせられる。

## 6. 理解容易性と説明可能性

ランダム性にも関連するが、前章の、個々の要素の性能は悪くとも、それらを多数組み合わせた集団の性能は向上するという事実は、理解容易性や説明可能性を犠牲にすることを意味する。著者はかつて「明示的理解に魅

せられて」と題して AI マップの記事を書いた。そのときは人工生命を明示的理解に対する挑戦と取り上げた。最近では多方面から、この挑戦状をたたきつけられている。我々人間の認知的理解能力に整合する粒度の概念と推論ステップを使って結果を再解釈する必要性は今でも感じているが、システムの理解ができればよしとする考えも受け入れつつある。小林氏らの GA の応用研究では工学的問題として明確に定式化された最適化問題を扱っている点で、人工生命とは違うが、最適解がランダム過程で自然に生まれてくるのを待つという点では類似している。このような過程で得られた解の理解容易性と説明可能性に興味がある。

## 7. GA と知識獲得

前章で述べたことと矛盾する面もあるが、GA や類似の技術は、我々がまだよく理解できていない複雑な対象における有力な問題解決器であるがゆえに、逆に、対象に内在する構造を知識として明示化してくれる技術として使えることを期待したい。新しいコーティング技術を用いた遺伝的操作により複雑な構造データも扱えるようになりつつあるので、新しい知識獲得技術としての活路が GA に見いだせそうな予感がする。例えば、DNA 系列の中に潜む規則性を取り出し、生体制御に関する領域知識を獲得するのに使えれば非常に有効である。レンズ設計が該当するのかもしれないが、スキルの獲得など明示化されにくい知識の獲得にも有効であろう。

## 8. おわりに

以上、コメントというには心苦しいが、小林氏の AI マップの読後感を述べさせていただいた。小林氏の研究グループの今後のさらなる活躍を期待している。

## ◇ 参考文献 ◇

- [Breiman 01] Breiman, L Random forests, *Machine Learning*, Vol 45, No 1, pp 5-32 (2001)
- [Domingos 98] Domingos, P Occam's two razors The sharp and the blunt, *Proc of the Fourth International Conference on Knowledge Discovery and Data Mining*, p 37-43 (1998)
- [Godsill 01] Godsill, S J, Doucet, A and West, M Maximum a posteriori sequence estimation using monte carlo particle filters, *Annals of the Institute of Statistical Mathematics*, Vol 52, pp 82-96 (2001)
- [Vafaie 98] Vafaie, H and De Jong, K Evolutionary feature space transformation, H Liu and H Motoda (eds), *Feature Extraction, Construction and Selection A Data Mining Perspective*, pp 307-323, Boston Kluwer Academic Publishers (1998)

## ◇ 三木 光範\*2

## 1. はじめに

GA (遺伝的アルゴリズム) の研究は何回かの興隆を経て現在に至っている。著者は現在も一つの興隆期であると思っているが、それが何回目の興隆期なのかは研究者によって見解が異なるだろう。そもそも現在は興隆期ではないという研究者もいるかもしれない。著者が現在、一つの興隆期と思っているのは、GAの完成度がかなり高まり、いよいよ実社会での本格的な应用到耐え得るようになってきたからであり、また、種々の応用問題のクラスに応じて、かなり工夫されたメカニズムを組み込むことに注目が集まっているからであろうと思われる。

1985年に初めてGAの国際会議(ICGA)が米国ピッツバーグで開催されてから、1997年の第7回ICGAまで隔年開催、第8回が1999年に開催されたときからGP(遺伝的プログラミング)の国際会議と合同で開催され、名称もGECCOとなり、毎年の開催となっている。発表される論文数も多く、GAの研究はますます盛んとなってきている。

GAの研究がこれほどまでに多くなってきている理由は、研究費があまりかからず、しかも数学的、遺伝学的、生物学的、実験的、あるいは純アルゴリズム的など、多くの局面からのアプローチが可能であり、それぞれの立場から好きなことがいえ、かつ有用な方法だからである。

このようなGAの研究に対して、小林氏が1990年に“周回遅れでGAのコミュニティに参加した”わけであるが、それから十数年、氏のグループの研究者があげてきた成果は非常に大きく、GAで最も重要な交叉と世代交代のメカニズムを執拗に解き明かし、対象問題のクラスに応じて最も適切な方法を考案してきた。それによって、GAはまさに実問題の解法として耐えるものとなってきた。この小文では、小林氏の研究の歴史を読んで感じたことを述べたい。

## 2. GAに対する思い入れ

著者が初めてGAと出会ったのは1990年に米国ロングビーチで開催された米国航空宇宙学会の第31回SDM会議でDeb氏(現在インド工科大学、多目的GAの研究者として著名)が学生優秀論文賞の受賞講演を行ったときである。当時、著者の専門分野は複合材料構造の最

適設計であり、このような分野に、従来の非線形数理論計画法とは異なる新しい、そして革新的な方法が登場したことに驚いた。だが、GAの研究はこのときすでに10年以上の歴史をもっていた。

さっそく、GAの勉強を始めようかと思ったが、調べてみるとすでに研究はかなり進んでおり、しかも進化に対する理論的な考察など、数学的な解析が多く難解であったので、GAコミュニティに入ることは先延ばしにした。複合材料の積層設計ではGAを使わずとも、解析的に最適化が行える画期的方法を発見することができ、それにかかりきりで、著者がGAの研究に着手するのは複合材料最適設計の研究が一段落してからであり、2周遅れか、3周遅れになってしまった。

一方、小林氏はGAの研究に取り組むまでは強化学習の研究を続けていたが、対象問題を広げ、複雑化したときに強化学習の枠組みの中では限界があることを感じ、もっとメタな強化学習、すなわち進化的手法に魅せられたと思われる。GAがもつ、怪しげで、かつ、広大な未知の領域には強化学習や、それを発展させたマルチエージェントの相互作用に基づく学習にはない、群れの集団的発展過程のおもしろさ、いわゆる創発があるからである。

このことは、小林氏のグループが考案してきた多くの交叉および世代交代モデルに、ほとんど生物色や遺伝学的なおいが感じられないことに関係していると思われる。GAの研究者の中には、GAが生物の進化を模倣したアルゴリズムであることから、できるだけ遺伝学や生物学とのアナロジーを大切に考える人たちがいる。本物の染色体のメカニズムに似せたり、有性生殖を考えたり、あるいは免疫システムを組み込むことなどが行われている。小林氏はこうしたアプローチをとらず、あくまでGAを集団探索アルゴリズムと考え、複数の解に含まれる良質の部分解をどれだけうまく発生させることができるか、そして組み合わせることができるか、ということだけを考え、具体的な問題空間でそのメカニズムを検討してきた。

こうしたアプローチは、もしかすると、GAに別の思い入れをもっている研究者からは評価されないかもしれない。別の思い入れとは、まさに生物進化の思い入れであり、対象問題などを考えることなく、極めてプリミティブなコード化と交叉、および選択を考え、その進化過程で対象問題に適合したメカニズムが自然に発生してくるのを待つというアプローチである。

こうした考え方は創発現象の研究においては常に論争の種となっている。人為的なメカニズムをできるだけ少なくして、それでも“意図しなかった”現象が生じてくる、これこそ創発研究の醍醐味だというわけである。これはまさに手品の仕掛けのようなものであり、“意図しなかった”ように見える“種”をどれだけうまく仕込むか、というテクニクの競争になってくる。小林氏の

\*2 同志社大学工学部知識工学科。

GA に対する思い入れは、そのような方向にはなく、それが高性能な GA をつくり上げる原動力となったと思われる。

### 3. 交叉って何だろう

小林氏は、特に、問題に適合した交叉を研究してきた。では、そもそも交叉とは何だろう。一番明確な答えは「親世代の形質を組み合わせ、子世代をつくること」であろう。すなわち、親世代の個体群が有している部分解をうまく取り出し、それを組み合わせ、新しい解の集合をつくることである。

ここでの最も大きな問題は、1) どれが部分解か、2) どのように組み合わせるのか、である。小林氏のグループが提案した SXX, EAX, および JOX などは巡回セールスマン問題やジョブショップスケジューリング問題の最適解の生成過程を詳しく観察することで得られた優れた交叉法であり、優れた親の部分解をうまく組み合わせることに成功した。

巡回セールスマン問題やジョブショップスケジューリング問題などの組合せ最適化問題では、解が構造を有している。その構造を無視した交叉は無意味であり、このことは GA を未知の問題に適用する際に、まず対象問題の解の構造を徹底的に調べて、それに適合した交叉を考案する必要があることを教えてくれる。これは、ある意味では GA の欠点にもなり得る。

さらにいえば、組合せ最適化問題では適合度が高い解どうしの交叉のみならず、適合度が高い解と低い解の交叉も最適解を与える場合が多い。例えば巡回セールスマン問題では非常に良好な解どうしは本質的に異なった解であることも多く、それらの交叉には意味がない場合が多い。非常に良好な解に必要な部分解は、それだけでは適合度が上がらない解の中に存在する可能性も高い。これは人間の人生を考えるとよくわかる。ノーベル賞を受賞した科学者とトップクラスの野球選手の人生を交叉させても意味はないだろう。

連続最適化問題における交叉は、果たしてこれが交叉と呼んでよいのかというものだ。連続変数の場合でも、変数間で交叉させれば組合せ最適化問題と同じとなる。この場合は多次元の探索において小数の次元での良好な解を見だして、その組合せの中に最適解があるというアプローチであり、変数間に依存関係が少ない場合には有効である。しかしながら、実数ベクトル空間において提案されてきた多くの交叉は、複数の親個体の位置関係から、親の近傍に子個体を配置するというものであり、少なくとも操作的には交叉と呼べるものはない。もはや交叉は概念的なものとなり、複数のデータから別の複数のデータをなんらかの操作で生成すれば、それはすべて交叉といえることになる。

小林氏のグループが考案した UNDX ファミリーでは、

個体分布の統計量が保存されるという非常に重要な役割を果たしている。連続関数の最適値を集団で探索するとき最も重要な情報は適合度関数の景観（ランドスケープ）情報をできるだけ包括的に獲得し、かつその情報をもとに有望領域を絞り込み、次世代の集団に伝達することである。ここで統計的情報の保存と有望領域の絞り込みを明確に分離し、前者を交叉に、後者を世代交代に割り当てる機能分担の考え方は非常に興味深い。機能分担することで、それぞれの操作を最大限にチューニングすることが可能となるからである。

その意味で、現在、Goldberg らの研究室で精力的に行われている確率モデル GA [Pelican 99]こそが、個体分布の保存を行う機能に徹するという機能を最大限に引き出しているといえるだろう。ただ、このような統計量保存という考え方は個体数が統計的に意味ある程度に大きいということが前提条件であろう。2次元や3次元の関数なら個体数が100程度でも問題ないだろうと思われるが、10次元や20次元というような高次元の関数で個体数が100では、各次元に対して単純計算では個体数が2以下となり、意味のある統計量が得られるとは思えない。

したがって、統計的情報を活用した交叉を実現するには、かなりの個体数が必要となり、そのようなアプローチは、もしかしたら進化的探索という枠組みからは少し外れるのかもしれないという感じをもっている。すなわち、進化的探索という枠組みは、実は厳密な最適解を探索することをあきらめることで、非常に効率的に探索が行えるというコンセプトなのではないか、ということである。そして、その効率性を武器に、動的に変化する環境への対応を考えるのが進化的最適化の重要な意義ではないかとも思われる。もちろん、そのような“生物的思想”を捨て、あくまで工学的最適化器として、工夫することが重要であることはいままでもない。昨年、同志社大学を訪れた Goldberg 氏と議論したおり、彼は「現在の GA は、対象問題に関する知見を大いに取り込み、洗練された多くの複雑なメカニズムをもつようになってきているが、それでこそ性能が上がり、実用的にも耐え得るものとなる」と主張した。確率モデル GA の発達は、まさに「GA の生物的特徴」を捨てることで成し遂げられた。これからも大いに新しい交叉を、制約なく考えることが重要かもしれない。

### 4. 世代交代モデル

小林らの考案した MGG (世代間最小ギャップモデル) の性能は非常に高い。MGG の性能が高い理由は、1) 選択が局所的であり、全体的な選択を行わない、2) 子個体を非常に多く作成する、ことにある。前者は早熟を回避するうえで有効であり、後者は良好な部分解の組合せの可能性を著しく高めることにつながる。

筆者らに取り組んでいる分散GAでは、母集団の分割が早熟収束を回避することにつながるが、この母集団の分割を極限まで行ったものが2個体分散GA [廣安 02]である。このモデルは、今から考えるとMGGのコンセプトと同じであることがわかる。

一方、子個体を非常に多く生成することは、評価に費やす時間も増えることから、次の問題を解くことになる。すなわち、同じ計算量で、世代交代を多くすることと、世代交代は少なくとも生成子個体数を増やすこととのトレードオフの問題である。多少、世代数を少なくしても、子個体数を多くするほうがよいということがMGGの結果から考えられるが、これは、見方を変えれば、GAの途中で局所探索を入れたほうが性能が高くなるということを示していると考えられる。なぜなら、生成する子個体を非常に多くすることは、親個体から得られる可能性のある子個体の中で最良のものを見つけだそうとする努力だからである。もちろん、局所探索が失敗したときのための保険として生成子個体群と親個体の中から最良個体とルーレット選択個体を、新しい子個体としている。

GAの中に局所探索を組み込むアイデアは多く考えられ、いくつかは実行されてきた。そして、実際に性能の向上が図られている。これは、実は最適化問題の中に「局所解の組合せが最適解である」という問題が多いからである。これはビルディングブロック仮説とは多少異なる。ビルディングブロック仮説では、ビルディングブロックをもつ設計解の平均適合度は、そうでないものと比較して高いということを意味し、一方、局所解の組合せが最適解であるという仮説は、局所解にはビルディングブロックが含まれる可能性が高いということである。このため、MGGを機能分担仮説で解析すれば、局所選択と保険付きの局所探索であり、これを実現できる方法はMGG以外にもいくつも考えられるだろう。その意味では、対象問題に応じてより適切な世代交代モデルが考えられるだろう。

## 5. おわりに

GAに対する小林氏の一連の研究を眺めてみると「海図なき航海」という言葉とは裏腹に、非常に整理され、着実に問題の本質に迫っているように見える。これは、まさに小林氏の研究のすすめ方そのものがGAであるかのようである。無謀に取り組んだ初期解としてのスタート、そのうち、徐々にランドスケープが見えてきて、有望な領域に研究を絞り込む、しかし、そこだけに集中するのではなく、大域的最適化のために多様性を維持させるなど、こうした研究の進め方そのものが、性能の高いGAと似ていると思うのは筆者だけであろうか。

いずれにせよ、この小文を書かせていただいたお陰で得るところは非常に大きかった。GAに関して駆け出しの著者がこんなコメントを書かせていただいたことに感

謝申し上げるとともに、昨年秋、同志社大学で開催した情報処理学会の進化計算シンポジウムで小林先生に特別講演をいただいたことに厚くお礼を申し上げたい。

なお、小林氏の論述の中で「サブツアー交換交叉」の節に「2重円TSP」について「有力な局所解」という記述がある。しかし、著者は組合せ最適化問題の局所解というものは最適化における解の振動操作（GAの場合は交叉や突然変異）を規定したときに初めて現れるものであり、本質的に局所解というものが存在するのではないと考えている。このため、ここでの記述も、あくまで小林氏らが行った交叉や突然変異では局所解だったということであることを注意しておきたい。異なった交叉や突然変異を用いれば2重円に局所解はなくなるだろう。

## ◇ 参考文献 ◇

- [廣安 02] 廣安 2個体分散遺伝的アルゴリズム, 計測自動制御学会論文集, Vol 38, No 11, pp 990-995 (2002)  
 [Pelican 99] Pelican, M and Goldberg, D E A Survey of Optimization by Building and Using Probabilistic Models, Illgal Reort No 99018 (1999)

## ◇ 浅田 稔\*3

### 1. はじめに

最初に小林氏のグループによる十余年にわたる偉大な研究成果に関して、敬意を表する。これは疑いの余地がないもので、すでに論文として採択されている以上、その技術内容に関して、コメントをさしはさむものではないし、妥当ではなからう。それに評者はGAの専門家でもない。AIマップは研究者がAIに対するビジョンや哲学を語り、それがコメントされることで、議論を深め、AIのコミュニティの活性化を図るものと評者は理解している。語られる哲学の中から、何かしら原理を引き出し、新たに入ってくるであろう若い研究者に対して、メッセージを送る場と理解している。

小林氏は、最初に「『海図なき航海こそ、研究者のあるべき姿』と考える著者には、大仰なビジョン（将来像）や見え透いたロードマップ（行程表）を語って若者を誘惑する気はない」と述べ、最後にも「お仕着せのビジョンやロードマップをハナから信じない著者には、住みにくい世の中になった、何とかしなければならぬと思うこの頃である」とくくっている。ビジョンを語ろうとし

\*3 大阪大学大学院工学研究科

ない、もしくは語ることに意味がないと主張しているように見える。

この観点から、今回の AI マップに関するコメントは、通常の形態から離れるかもしれないが、お許しいただきたい。最初に「海図なき航海」の真意を問い、次に、生物進化とのアナロジーの可能性を問う。そして最後に、ビジョンを説くことの意味・意義について議論する。言葉尻を捉えた議論になるかもしれないが、誤解を恐れずに、評者のコメントを示す。

## 2. どこまで「海図なき航海」か？

小林氏の言を文字どおりに捉えれば、語られた哲学(信条)は「海図なき航海」であり、原理として引き出せるのは、「表現型空間における形質を人為的に抽出し、適切なコード化と交叉を設計することにより、遺伝子型空間において形質に対応するビルディングブロックを成長させるという考え」である。また、「10. おわりに」の最初で、「『海図なき航海』を地で行く、行き当たりばったりの研究を続けてきた割には、振り返ってみると、すべての研究が互いに絡み合い、補い合って、それなりに1本筋の通った研究を積み重ねてきたように思えるから不思議である」。評者は、GAの問題を最適化に特化し、コード化手法を設計者に委ねることで可能になったと考えている。前者で、問題領域の設定がなされ、それがゆえに後者で研究者の知性が試されている。成果を見ればわかるように、うまい設定とアプローチである。個々の問題に対する深い洞察と解法の一般化には感嘆するしかない。

この二つの観点が航海のガイドラインだとすれば、結果としてこれが海図ではなかろうかと思うのは評者だけであろうか？ つまり評者には「海図なき航海」には見えないのである。対象とする課題の表層的構造は変化しているが、海(最適化問題)を航行するためのコンパス(人為的コード化)をもってある方向に常に向かっているように見えるのである。断っておくが、「小林氏がうそをついている」というつもりはない。解釈の違いといえばそれまでである。ただし、AIマップとして明確にしたいのは、「海図なき航海」といっても、何らかの海図はあるのだ。それを明らかにしたい。「GAの主役は交叉であり、突然変異は脇役に過ぎない。監督としての世代交代モデルも重要である。しかし、主役がいてこそ監督も力を発揮できるとの考えから、形質遺伝性に優れたコード化・交叉の開発を主目標に、これをサポートする世代交代モデルを副目標において、GAのコミュニティにおける流行はあまり気にしないで、我々独自の姿勢と観点で研究に取り組んできた」と「10. おわりに」で述べられている。これが一本の筋であるならば、それが、GAの研究の将来に対して、どういう意味をもつか？さらにはAIの研究全体からの位置づけはどうか？と

聞いてみたいのである。

## 3. 最適化のツールがGAの生き残る道？

人工生命の研究は、一時期もてはやされたが、現状では衰退したかのように見える。生物進化の人工的再現には、夢があるように思えたし、その工学的応用も興味あるものを感じた。しかし、生物進化のアナロジーが表層的に留まり、中途半端なものになったと感じたのは評者だけであるまい。小林氏は、そんな中で、最適化手法としてのGAの潜在能力を見だし、そこに特化することで、種々の手法を編み出した。ただし、いずれの手法も静的な環境を想定してはいないだろうか？

もともとGAが人々を惹きつけた生物進化アナロジーの観点から考えると、動的な環境、特に他者の存在、社会性が知能の根源と目されている昨今[岡ノ谷 03]、GAの能力が知能の創発にどう関連するかは興味の尽きぬところである。例えば、多くの動物の類似性は、DNAなどに符号化されたユニバーサル言語を読む高度に保守的な発達システムであるにもかかわらず、コミュニケーションにはユニバーサル言語が欠けているように見えるのはなぜか[Hauser 02]？環境因子やある種の淘汰圧がこれらにコミットしているとすれば、固定化された環境のなかで最適化問題から動的な環境における進化プロセスへの応用としてのGAの生き残る道はないのか？

極端な議論になってしまったが、まずは開発されてきたモデルの中に、生物進化アナロジーの観点から、新たな進化モデルの提言が見いだせないか？もしくは、これまでの生物進化の新たな説明ができないか？といった興味がわく。例えば、UNDXファミリーは、生物進化でいえば、どのような戦略に対応しそうか？また、世代交代モデルであるMGGやDDAを具現化している(してきた)生物は見当たらないか？さらには、適応度景観としてのUV構造とその対処モデルISMなどは、生物進化の世界で対応する現象は何か？k-tablet構造は興味あるが、そのような構造を自然環境で見いだすことは可能か？

これらへの回答は、開発されてきたモデル群の能力と限界を異なる視点から眺めることになるであろうし、GAの能力と今後の発展を考えるうえで大切であろう。GAの第一人者としてぜひともコメントいただきたい。また、方法論の観点からは、「ビルディングブロックのプリミティブは設計者が与えるものではなく、遺伝的操作によって発見し、その組合せ方も遺伝的操作に委ねるべき」といった理想論に近づくための新たな方法論や道筋はないかと尋ねたい。

#### 4. 問題を考えることと解くこと

十余年にわたる一連の研究は、個々の課題に対して、「本質が何かを発見し、それをコード化する」といった戦略が随所に見受けられる。前者が「問題が何か」を考えることに対応し、後者がそれに基づいて解を獲得する過程とみなせる。当然だが、人工知能の問題に限らず、どういう問題を解くべきかを考えるのが先で、どう解くかは後である[國吉 03]。とするならば、個々の問題に立ち入った局面では、どういう問題を解くかが問われているとみなせるが、研究史全体から見たとき、対処療法的に見えるのは偏見であろうか？ 小林氏の十余年にわたるGAの研究成果が素晴らしいことは間違いがない。評者の偏見を払拭するうえでも、再度、一本の道筋の意味と意義を明確にしてもらいたい。

評者は人工知能とロボティクスの大いなる挑戦としてロボカップを提唱し、推進している[浅田 02]。まさしく「大仰なビジョン（将来像）や見え透いたロードマップ（行程表）を語って若者を誘惑している」のである。しかし、このことが、「国立大学の法人化への移行を控えて、学長の権限が大幅に強化される中、教授会の自治が封殺され、お仕着せのビジョンに合うロードマップを提示できなければ研究室の存続が問われる時代が到来しようとしている」に対応するビジョンになるとは思っていない。むしろ、こんな世の中であるからこそ、自らのビジョンを明示することが大切ではなからうか？ 先端研究のなかでビジョンを語ることで、大学経営の観点からビジョンを語ることは意味が違うのではなからうか？

「我々は『理解するとはどういうことか？』について、哲学のおよび方法的に問い直す時代に生きているように感じるこのごろである」[小林 99]と小林氏は語っている。方法論は垣間見たかもしれないが、哲学を聞いていない気がするのである。

#### ◇ 参考文献 ◇

- [浅田 02] 浅田 稔 編著・RoboCupSoccer・ロボットの行動学習 発達・進化、共立出版（2002）  
 [Hauser 02] Hauser, M D, Chomsky, N and Fitch, W T . The Faculty of Language What Is It, Who Has It, and How Did It Evolve ? , *Science*, pp 1569-1579 (2002)  
 [國吉 03] 國吉康夫：ロボットの知能—創発実体主義の挑戦—、計測と制御, Vol 42, No 6, pp 497-503 (2003)  
 [小林 99] 小林重信：編集後記, 人工知能学会誌, Vol 14, No 5, p 818 (1999)  
 [岡ノ谷 03] 岡ノ谷一夫．身体的「知」の進化と言語的「知」の創発, 人工知能学会誌, Vol 18, No 4, pp 392-398 (2003)

## 回 答

### ◇ 小林 重信\*4

#### 1. はじめに

「AIとは複雑さへの挑戦である」と著者は認識している。「複雑な対象や現象の明示的な理解を求めて」機械学習やデータマイニングに取り組んでいる元田氏。「複雑な問題を実用的な時間で解く」ことを目標に大規模並列計算に取り組んでいる三木氏。「AIとロボティクスの大いなる挑戦としてのロボカップ」を提唱、推進している浅田氏。いずれも明確なビジョンと哲学をもって複雑さに挑戦している方々から、著者の拙稿に対し非常に有益で示唆に富んだコメントをお寄せいただいたことを、とてもうれしくありがたく思っている。

さて、浅田氏から「ビジョンを語ろうとしない」、「哲学を聞いていない気がする」と質されたこともあり、コメントターの発言とは逆の順番で回答することにした。なお、3氏に対する回答を通じて多少重複があることをお許し願いたい。

#### 2. 浅田氏のコメントに対する回答

##### AI マップの意義について

「AI マップは研究者がAIに対するビジョンや哲学を語り、それがコメントされることで、議論を深め、AIのコミュニティの活性化を図るものと評者は理解している。語られる哲学の中から、何かしら原理を引き出し、新たに入ってくるであろう若い研究者に対して、メッセージを送る場と理解している」との浅田氏のご意見はごもっともで、そのとおりだと思う。「AI マップ」という記事は、十数年前、会誌編集委員であった著者が最初に提案したのがきっかけで実現したものであるが、当時の著者もAI マップの意義について浅田氏と同じような考えをもっていた。

AI マップに対する基本的な考え方は昔も今も変わらないつもりである。変わったところがあるとすれば、歴史を少し意識するようになったことかもしれない。歳をとると過去を整理して自分と折合いをつけるために、歴史を語るようになるといわれる。著者もそのような心境から歴史を語ってみようという気になったのであろうと自己分析している。語られる歴史の背後にはそれなりのビジョンや哲学が投影されているはずであり、行間に埋

\*4 東京工業大学大学院総合理工学研究科。

もれて明示的ではないかもしれないが、若い研究者に対して何らかのメッセージを送ったことになると思われている。これまで本誌に掲載された20編近いAIマップの記事を読み返してみると、半数が未来を語り、半数が歴史を語っている。全体としてほどよくバランスが取れているのではなからうか。

#### どこまで「海図なき航海」か？ について

海図というのは、海の深さ・浅さや潮流の方向など、航海をするのに必要なことを記したマップをいう。本来「海図なき航海」はハイオインフォマティクスのように新しい学際的な領域を開拓しようとしている研究者が使うべき言葉であろう。そこでは、研究テーマや方法論を自ら決めていくことが求められるからである。一方、AIのように領域が確立されてくると、コミュニティの中に共通の価値観や行動規範が形成されてくる。価値観や行動規範が共有されているからこそコミュニティは維持される。共通の価値観や行動規範はそのコミュニティ独自の海図（海の深さ・浅さ・潮流）を生み出す。ここで、海とはコミュニティがカバーする領域、深さ・浅さとはさまざまな研究テーマ・方法論のコミュニティにおける価値観、潮流とはそれらの移り変わりをいう。

海図は領域とコミュニティの健全な発展にとって必要であり、新規参入した若手研究者の手がかりとなるので有用である。しかし、一方では、海図は若手研究者を流行の研究に走らせ、創造性の芽を摘む危険性を内包していることに注意しなければならない。「海図なき航海こそ、研究者のあるべき姿」の真意は、「他人がつくった海図はあまり当てにしないで、自分が新しい海図をつくるのだという気概で研究に取り組むべき」にある。したがって、「(自分なりの)海図をつくるための航海こそ、研究者のあるべき姿」と言い換えてもよい。

さて、主題のGAに話を戻す。著者らは「GAを最適化のツールとして使えるようにする」という目的を掲げて、そのために「形質遺伝を重視したコード化・交叉の開発」および「多様性維持に優れた世代交代モデルの開発」の二つの目標を設定した。第1の目標については問題クラスごとに高性能なコード化・交叉の開発を目指すこと、第2の目標については汎用性の高いモデルの開発を目指すこと、を基本方針とした。このような目的と目標、方針はGAの研究に着手した当初から共同研究者の山村雅幸君と著者の間で共有していた。

「GAの主役は交叉であり、突然変異は脇役に過ぎない。監督としての世代交代モデルも重要である。しかし、主役がいてこそ監督も力を発揮できるとの考えから、形質遺伝性に優れたコード化・交叉の開発を主目標に、これをサポートする世代交代モデルを副目標において」という記述は、上記の目標や方針をわかりやすく表現したものである。

こうした目標や方針は航海のカイドラインあるいはコ

ンパスに相当するといってもよい。「行き当たりばったりの研究を続けてきた」とは「ロードマップをつくって計画的に研究を進めたのではなく、カイドラインを緩い拘束条件にしてあとは学生の興味と発想に任せて自由自在に研究させてきた」という意味である。そのような研究を進めていく過程において、節目ごとにうまく具合にそれまでの研究をまとめる仮説（機能分担仮説、統計量の遺伝、UV構造仮説など）がつけられ、それらが新たなカイドラインとなって、ある時期一気に研究が進展して、今日に至っている。そういう経緯を踏まえて「行き当たりばったりの研究を続けてきた割には、振り返ってみると1本筋の通った研究を積み重ねてきたように思える」と述べたのである。

「GAのコミュニティにおける流行はあまり気にしないで、我々独自の姿勢と観点で研究に取り組んできた」とは、「設定された目的・目標・方針をカイドラインとして、流行など周りの雑音はあまり気にしないで自由自在に研究を進めてきた」という意味である。「独自の姿勢と観点で研究に取り組む」ことを海図と解釈すれば、「海図なき航海には見えない」という浅田氏の指摘は、そのとおりである。海図を最初に述べた意味で使うとすれば「独自の姿勢と観点」があればこそ「海図なき（海図をつくるための）航海」に乗り出せるのであって、そうでなければ「他人のつくった海図を頼りに、後追い研究を続ける」以外なからうと著者は思っている。

「独自の姿勢と観点で研究に取り組む」ことが「GAの研究の将来に対して、どういう意味をもつか？」についてであるが、当初設定した「GAを最適化のツールとして使えるようにする」という目的はほぼ達成したと著者は判断している。引き続き別の観点からGAの研究を続ける積もりではいるが、GAのコミュニティからは少し距離をおいて、新たな展開を図りたいと考えている。

「AIの研究全体からの（GAの）位置づけはどうか？」という問いに対して簡潔に答えたい。「(知能の)複雑さへの挑戦」を標榜するAI研究の中で構成的アプローチの必要性と重要性は認識されているものの、全体としてはまだ十分ではない。最適化ツールとしてのGAの完成度が高まり、また並列計算環境が急速に整備されつつある時代状況を考えれば、GAを方法論とする構成的アプローチが活発化してくるものと予想し、かつそれを期待している。

#### 最適化のツールがGAの生き残る道？ について

著者には生物進化という観点からのGAに対する思い入れはない。ある時期「生物進化アナロジーの観点からGAが人々を惹きつけた」のは確かである。自然界の進化過程とのアナロジーでGAを理解することは魅力的であるが、危険でもあると著者は思っている。

「生物進化にヒントを得たとされるGA」という言い



方（実際、前稿の冒頭で著者はそう書いた）は、GAの創始者である Holland 氏の発想の源泉がそこにあったことに敬意を表しての枕言葉的な意味しかもたない、生物進化はつねに変化し続ける生態学的環境の中で繰り広げられるニッチ（生態学的地位）獲得競争の結果創出されるものであり、GAの枠組みでモデル化できるほど単純なものではないと著者は思っている。生態系は複雑系そのものであり、生物進化の解明のためにはさまざまな学問分野からの学際的なアプローチが必要と思われる。

したがって、「動的な環境における進化プロセスへの応用としてのGAの生き残る道はあるのか？」という問いに対しては、興味はあるが、やや否定的である。「開発されてきたモデルの中に、生物進化のアナロジーの観点から、新たな進化モデルの提言が見いだせないか？もしくは、これまでの生物進化の新たな説明ができないか？」という問いに対しては、世代交代モデルと適応度景観モデルの中には生物の適応戦略に該当するものが存在する可能性はあると考えている。例えば、鳥類などにおける近親交配を回避するための適応放散を世代交代と適応度景観の観点から分析してみると、何か新しい知見が得られるかもしれない。

GAは生物進化よりも人工交配（artificial selection）とのアナロジーで理解するのが適切と著者は考えている。人工交配は異種間交雑を繰り返して市場価値の高い種または個体をつくり出す合目的のプロセスであり、「最適化」という合目的性と「交叉」という方法論に共通性がある。ちなみに、ダーウィンの自然選択（natural selection）は人工交配とのアナロジーから生まれたといわれる。

GAの生き残る道の一つに社会システムのモデル化が考えられる。例えば、科学者や技術者はそれぞれの専門領域ごとにコミュニティを形成し、知識を共有・伝承しつつ、新しい知識を創出しているが、一部の天才を除いて、大部分の科学者・技術者が日常的に行っていることは知識の組合せとその改良であり、交叉と突然変異に対応づけられる。コミュニティへの参入・退出は世代交代モデルに対応づけられる。こう考えると、コミュニティの形成・成長・成熟・衰退の過程をGAでモデル化することができるのではなからうか。近年、政治・社会・経済システムの構成的研究として、Agent Based Approachが注目されるようになってきたが、その中でGAは重要な役割を担っている。

「ビルディングブロックのプリミティブは設計者が与えるものではなく、遺伝的操作によって発見し、その組合せ方も遺伝的操作に委ねるべき」といった理想論に近づくための新たな方法論や道筋はないかとのコメントに対しては、元田氏や三木氏からも同じようなコメントがあったので、両氏に対する回答の中で述べる。

### 問題を考えることと解くことについて

「研究史全体から見たとき、対処療法的に見えるのは偏見であろうか？」と浅田氏が感じたのは、著者の書き方が散文的であったことに起因するところが大きい。問題クラスごとにコード化・交叉を開発するという基本姿勢を対処療法的であると感じたのかもしれない。組合せ最適化というのは問題クラスごとにまったく異なる解構造をもっていることを考慮して、まずは個別のアプローチを取るのが著者らの基本戦略であり、その際の指針は“形質遺伝に優れたコード化・交叉”であり、実際そういう指針により既存手法を凌駕する解法を設計することに成功したわけである。性能の高い解法の設計に成功したのは各問題の解構造の本質を的確にモデル化することができたからであり、これは対処療法的なものではないと思っている。

また、さまざまな問題クラスの解法を設計する過程を通じて普遍的な原理を見だし、それらを仮説として取りまとめ、これらを指針として新しい交叉や世代交代モデルを設計してきた。例えば、「機能分担仮説」では交叉は親集団の分布を変えないサンプラーであるという交叉のあるべき姿を提案している[山村 98]。「統計量遺伝」は実数値GAのUNDXが機能分担仮説を満足することを証明するとともに、実数値GAの設計指針になっている[喜多 99]。「UV構造仮説」は通常のGAが最適解の発見に失敗する理由を明快に説明するとともに、新しい世代交代モデルを設計する指針になっている[池田 02]。

このように、著者らは特殊な問題を対象にその本質を追求し、その知見が集積された段階で共通原理として一般化し、一般化原理を指針としてふたたび特殊な問題の本質を追求するというサイクルを繰り返してきた。そういう観点から研究全史体を俯瞰すると「GAに関する知の体系」が構築されつつあると著者は思っている。

著者らの研究戦略を要約すると、最初に明確な目的・目標・方針を設定したうえで、研究者の自由な発想で研究を展開し、節目ではそれまでに得られた共通の原理を仮説や指針としてとりまとめ、それらをガイドとして次のフェーズでの研究の展開を図るのである。結果として、一本の道筋（Historical Roadmap of Research on Genetic Algorithms）が形成されたと思っている。

ビジョンや哲学を明示的に語るほどの見識や経験をもち合わせていないので、以上の説明をもって理解いただければと思う。

### 3. 三木氏のコメントに対する回答

#### GAに対する思い入れについて

三木氏が見抜いているように、著者には生物進化という観点からのGAに対する思い入れはまったくない。GAの取組みにおいてそういう視点は欠落している。前稿の冒頭で、著者がGAに魅力を感じたのは、1) 解表

現の自由度の高さ, 2) 微分情報を必要としない直接探索, 3) 集団で解を探索する多点探索, 4) 交叉オペレータ, の4点に集約されるからであると述べた。解表現の自由度が高いために, GA の設計者は問題の解構造を反映したコード化を自由に設計できる。多点探索は大域的探索において有利である。交叉は部分解の組合せによる探索の飛躍をもたらす。よって, GA は最適化手法として非常にユニークかつ魅力的な枠組みと思えたからである。

生物進化はとても興味深い。生物進化への思い入れから GA に関心をもつ研究者は, 三木氏が述べているように, 「極めてプリミティブなコード化と交叉, および選択を考え, その進化過程で対象問題に適合したメカニズムが自然に発生してくるのを待つというアプローチ」を採用しており, 元田氏が言及したメタ GA がこれに相当する。

著者は基礎研究としてのメタ GA の価値は十分認めるが, 生物進化の理解を指向しているのか, 汎用の最適化ツールを指向しているのか, によって研究のスタンスは大きく異なると考えている。前者の立場であれば, 人工生命の研究がなぜ停滞しているのかに対する反省を出発点とすべきであろう。後者の立場であれば, 達成すべき目標を明確化したうえで取り組むべきであろう。

三木氏は「こうした考え方は創発現象の研究においては常に論争の種になっている。人為的なメカニズムをできるだけ少なくして, それでも“意図しなかった”現象が生じてくる, これこそ創発研究の醍醐味だというわけである。これはまさに手品の仕掛けのようなものであり, “意図しなかった”ように見える“種”をどれだけうまく仕込むか, というテクニックの競争になってくる」と述べている。同感である。

平成7年度から3年間実施された重点領域研究「創発システム」においても同じような論争が繰り返された。そのような研究の中で後に残る成果はあまりなかったのではないかと著者は思っている。しかし, 基礎研究の芽を摘むようなことは避けるべきであろう。新たなブレークスルーと今後の成果に期待したい。

### 「交叉って何だろう」について

交叉に対する三木氏の見解には概ね同意できる。組合せ最適化問題は問題クラスごとに解構造が異なるので, 解構造を考慮に入れたコード化・交叉設計が必要かつ有効であり, そうしなければ既存の手法に太刀打ちすることはできないとの考えで取り組み, TSP (巡回セールスマン問題) における EAX や JSP (ジョブショッパスケジューリング問題) における JOX など優れた交叉を開発することができた。しかし, 一方では2重円 TSP や JSP の10難問 (10 tough problems) のように有力局所解が最適解の発見を妨げる騙し構造 (UV 構造) をもつ場合には, 交叉の工夫だけでは対処できず, UV 構造に

対応できる世代交代モデルが必要なことも明らかになった。

「適応度が高い解と低い解の交叉が最適解を与えることもあれば, 適応度が非常に高い解であっても本質的に異なる解どうしでは交叉が無意味な場合が多い」という三木氏の指摘はそのとおりである。2重円 TSP では C 型と O 型の二つの有力局所解があり, C 型解どうしまたは O 型解どうしの交叉はそれぞれ適応度の高い個体を生成する可能性は高いが, C 型解と O 型解の交叉は適応度の低い個体を生成する可能性が高い。一般に, 交叉近傍は突然変異近傍より大きく設計されるが, 解空間全体から見れば局所的なオペレータに過ぎない。交叉の設計に当たっては局所的な解構造を十分調べる必要があるが, 大域的な最適解を求めるためには適応度景観の分析に基づく世代交代モデルの選択が必要である。交叉が局所的なオペレータであるのに対し, 世代交代モデルは大域的なオペレータであり, この意味で両者は相補的な関係にあるといえる。

「連続最適化問題における交叉は, はたしてこれが交叉と呼んでよいのか」, 「交叉は概念的なものとなっている」との指摘, 生物とのアナロジーから見るとそのとおりである。機能分担仮説では, “交叉とは親集団の分布を変えないサンプラー”と位置づけている。機能分担仮説から見ると, UNDX ファミリーも Goldberg 氏のグループにおける確率モデル GA (分布推定アルゴリズム) も同類とみなされる。三木氏が指摘しているように, 確率モデル GA のほうが個体分布の保存を行う機能に徹底しているといえる。

しかし, 確率モデル GA では, 三木氏も指摘しているように, 高次元の関数に対して意味のある統計量を得るために集団サイズを十分大きく取らなければならないというジレンマに直面する。一方, UNDX ファミリーの交叉は確率モデルを明示的に学習することはしないで, 交叉を複数回繰り返すことによって確率モデルを非明示的に推定する一種の近似モデルといえる [佐久間 03]。

UNDX ファミリーの頂点に位置する LUNDX- $m$  + EDX は, 前稿でも述べたように, 数百次元の  $k$ -tablet 構造をもつ最適化問題の解を次元の半分程度の集団サイズで見いだすことができる。確率モデル GA がどの程度の次元にまで対応できるのか不明であるが, 確率モデルを明示的に推定する分布推定アルゴリズムと非明示的に推定する UNDX ファミリーのどちらに軍配が上がるのか, それぞれの今後の発展にたいへん興味がある。

「進化的探索という枠組みは, 実は厳密な最適解を探索することをあきらめることで, 非常に効率的な探索が行えるというコンセプトなのではないか」という三木氏の考えには理解を示すものであるが, 著者は「確率モデルを明示的に推定するという複雑な方法を導入しなくても, 機能分担仮説を設計指針として交叉を多親に自然に拡張するだけで, 確率モデルを非明示的に推定すること

ができ、しかも効率的かつ性能の高い探索が行えるまでにGAの完成度は高まっている」と主張したい。

「効率性を武器に、動的に変化する環境への対応を考えるのが進化的最適化の重要な意義ではないか」という三木氏の考えは、GAを人工生命研究のツールと考える立場の代表的な意見と思われる。著者はあくまでGAを広い意味での最適化のツールと考えたい。そのためには、静的な最適化問題に留まらず、強化学習が対象としている不確実な環境のもとで報酬に遅れを伴う最適制御問題の政策学習へ対応できるように守備範囲を広げることが必要と思われる。

著者らは、最近、重点サンプリングを用いたGAによって政策の学習を効率良く行う研究に取り組んで、良い感触を得ている。強化学習でよく使われるActor-Criticは探索手法としては局所探索法であり、多峰性に対応できないという欠点があるが、GAは対応できる。しかし、多点探索のために評価回数が増えてしまうというジレンマに直面する。重点サンプリングのGAへの導入はこの問題を大幅に緩和できることを確認している。これにより、強化学習問題を強化学習アルゴリズムによらずにGAで解く道が拓けてくる。多峰性への対処だけでなく、大規模性への対処という点でも、GAのほうが有利と見られる。この分野でブレークスルーが生まれるかもしれないと期待している。

#### 世代交代モデルについて

三木氏が取り組んでいる分散GAの極限モデルである2個体分散GAがMGGと等価であるという事実は興味深い。MGGの利点は並列化実装との親和性が非常に高いことにある。MGGにおいて子個体を多数生成するのは世代交代の選択圧がかかる前により良い個体を探索するという意図に基づく。したがって、これは親個体群の分布を探索領域とする局所探索にほかならない。「MGGを機能分担仮説で解析すれば、局所選択と保険つき局所探索であり、これを実現できる方法以外にもいくつも考えられるであろう」という三木氏の指摘はそのとおりだと思う。適応度関数の評価に大きなコストがかかる場合にはより効率的な局所探索を選択することも考えられよう。

#### 組合せ最適化における局所解について

「組合せ最適化問題の局所解というものは最適化における解の摂動操作を規定したときに初めて現れるものであり、本質的に局所解というものが存在するのではない」との指摘は、そのとおりである。そういう意味では、組合せ最適化問題に限らず、連続変数の関数最適化問題においても本質的に局所解が存在するものではない。しかし、便宜上ではあるが、局所解は探索オペレータとは独立に規定しているのが一般的と思われる。例えば、関数最適化問題では $\epsilon$ 近傍を用いて、巡回セールスマン問題

では2-opt近傍を用いて、局所解を定義するのが自然である。

なお、著者らが“有力な局所解”というとき、最適解とは異なる大谷構造の中央付近に存在する局所解の意味で使っている。2重円TSPではC型とO型という二つの有力な局所解が存在する。この場合、親個体群の分布を保存する交叉では同じ谷の中にある個体どうしを交叉させても子個体は同じ谷の中に生成されやすい。谷が大きいので2-optのような突然変異では谷から脱却できない。2重円TSPはUV構造の典型例である。

#### 4. 元田氏のコメントに対する回答

##### GAと知識ベースについて

「GAの研究の進展の経緯を眺めていて、知識ベースの進展の経緯とよく似ていると感じた」という元田氏のコメントに、同感である。著者も80年代を通じて知識ベースシステムの研究に携わった。元田氏のコメントと重複する部分もあるが、著者の経験から知識ベースの進展の経緯を総括しておきたい。

第一世代知識ベースシステムは専門家から獲得したif-then型の経験的知識をプロダクションシステムに実装すれば、認識-行動サイクルによって高度な問題解決が図れるという期待に基づくものであった。しかし、知識獲得がボトルネックとなり、また高次推論と呼ばれる諸技術も不完全な知識を補完できる代物ではないことが判明し、エキスパートシステム実現の夢はあっけなく崩壊した。

著者が80年代に数多くの事業所や生産現場を回った経験から感じたことは、文字どおり“専門家”と呼べる人はそれほど多くはなく、“担当者”が問題を抱え込んでいるケースが少なくなかったことである。これでは、知識獲得がうまくいったとしても、担当者システムしかできないであろうと思われた。知識獲得以前の問題として、対象問題のシステム分析が非常に重要であることを痛感した。

第二世代知識ベースに関する研究が経験的知識への依存を極力抑えて、対象問題の構造や挙動に対する深い理解に基づくモデル化が必要との考えから、問題解決過程のタスクへの分割とモデル化を指向したのは自然な成り行きであったと思われる。しかし、第二世代の研究によって知識ベースシステムの実用化が促進されたかという点、必ずしもそうではなかったように思う。

一方では、80年代後半にファジブームが起り、制御分野を中心に実用化例が数多く報告されるようになった。皮肉なことに、ファジィ制御に組み込まれた知識は第一世代知識ベースシステムの挫折で批判された経験的知識が主役であった。制御では状態変数や制御変数は連続値を取る場合が多く、if-then型の経験的知識と近似アルゴリズムとしてのファジィ推論をうまく組み合わせ

せて、理解可能性を犠牲にしながらも、操作性を重視したことが現場で受け入れられ、実用化につながったと思われる。

さて、著者は知識ベースに関する研究は基礎研究としての意義は認められるが、実用化の可能性は非常に低いと判断、90年を境に見切りをつけ、新しい問題解決パラダイムを模索する中でGAと遭遇し、GAの研究に着手した。simple GAに代表される第一世代GAでは、対象問題の性質に関する十分な考察なしに、ビットストリングによるコード化、一点交叉や一様交叉などの単純な交叉、選択圧の強い世代交代モデルが使われていた。第一世代知識ベースと同じような状況にあったといえる。

そこで、著者らは、simple GAを反面教師として、何が形質かという観点から問題の解構造を徹底的に調べて、問題クラスごとに適切なコード化・交叉を設計する戦略を取った。これらの開発経験を踏まえて、機能分担仮説が提唱された。機能分担仮説では“交叉は親集団の分布に従うサンプラー”とみなす。前稿で述べたように、“統計量の遺伝”は機能分担仮説を実数値GAに対して操作化したものである。これらを設計指針として開発された実数値GA (UNDXファミリー)を著者は第二世代GAと位置づけている。Goldberg氏のグループが中心となって精力的に研究が進められている分布推定アルゴリズム (Estimation of Distribution Algorithms)に基づくGAは、機能分担仮説とは独立に提案されたものであるが、機能分担仮説の良い具体例になっており、第二世代GAに属するといえる。

より複雑な適応度景観への取組みとして始まった一連の研究 (UV構造仮説と生得分離モデルISM, k-tablet構造と隠れ変数上の交叉LUNDX-mなど)を著者は第三世代GAの始まりと位置づけている。

以上をまとめると、形質遺伝重視を手がかりに巡回セールスマン問題から始まった著者らの個別的GAの研究成果は、機能分担仮説にまとめられ、この仮説を設計指針として一連の実数値GA (UNDXファミリー)が開発され、さらにはこれらを武器としてより複雑な適応度景観への挑戦というように、スパイラル的な発展を遂げて今日に至っている。

「GAの枠組み (どうすれば何ができて、何が難しく、どのような工夫が有効かなど) が理解できた気分になる。小林氏の自負もこのような境地に達したことに起因するものであろう」との元田氏のコメントが、これまでにGAを使って個別の問題をたくさん解いてきた結果、十分なノウハウが集積され、どんな問題にでも対処できるようになったのであろうという意味のものであれば、それは上述した経緯から誤解である。“理解できた気分になる”とか“このような境地に達した”といった情緒的なものではなく、理論と応用の両面からの裏づけのある「GAに関する知の体系」としてまとめつつあると著者らは認識していることを強調しておきたい。

### メタGAの可能性について

「ビルディングブロックのプリミティブは設計者が与えるものではなく、遺伝的操作によって発見し、その組合せ方も遺伝的操作に委ねるべき」という理想に近づくために、遺伝子間の相互作用を考慮したリンケージ学習が活発に研究されている。リンケージ学習の最近の研究動向は[倉橋03]の解説を参照されたい。著者はメタGAの実現に向けた基礎研究の重要性を支持する立場を取るが、問題クラスごとに開発されてきたGAと同等またはそれ以上の性能を達成するためにはまだかなりのフレキシブルが必要と感じている。

組合せ最適化問題は多種多様で、問題クラスごとにそれぞれを特徴づける特殊な解構造を有しており、リンケージ学習というシンプルな枠組みでそれらを見出すことは非常に困難と予想されるからである。この困難を突破したならば、科学革命が生じるであろう。メタGAは挑戦のしがいのある研究課題であることは確かである。

### GAと欲張り探索について

著者らのGAによる決定木の多目的最適化の研究に対する元田氏のコメントは、1) 決定木生成の目的は未知のデータに対して予測誤差を最小化すること、2) 木の複雑さは結果として定まるもの、3) 決定木では属性選択および属性構築が重要、の3点に要約される。

未知のデータに対して予測誤差を最小化することは原理的に不可能であるために、与えられた例題を訓練例とテスト例に分割し、訓練例に対して誤差を最小にする決定木を初期値として、テスト例に対する汎化誤差を最小にするように枝刈りを施すのが伝統的な方法であった。しかし、汎化誤差を最小にする決定木は唯一に定まるものではなく、またある許容誤差範囲内に存在する決定木は多数存在するのが一般的である。

著者は多目的最適化によって求めたパレート最適木集合の中から領域知識や専門家の知見に基づいて、理解可能・解釈可能という意味で合理的な決定木を選択することが意思決定者にとって好ましいとの立場を取る。受理可能な木の複雑さや汎化誤差は意思決定者の選好の結果定まるものと考え、訓練例に対して得られたパレート最適集合の各決定木に対してテスト例による汎化誤差を求めた後、木の複雑さと汎化誤差の2目的でパレート最適のふるいにかけて、小さな木や大きな木は脱落し、ある範囲の大きさの木だけが残る。そのような決定木を意思決定者に提示することが重要と考えている。

決定木では属性選択および属性構築が重要であるとの指摘については、そのとおりであると思う。先行研究からも属性選択及び属性構築においてGAの有効性は確認されているが、決定木生成の前処理としての属性選択および属性構築、後処理としての枝刈りによる汎化誤差の最小化という三つのフェーズに分けて処理するのではなく、多目的GAにより三つのフェーズを一括処理するこ

とが理想と著者は考えている。

### ランダム性について

「GAの特徴は制御されたランダム性」という元田氏のコメントは著者らの考えを的確に表している。ランダム性を“適切に制御する”ことがGA設計の重要なポイントである。機能分担仮説に基づく交叉、例えば実数値GAにおけるUNDXファミリーは親集団の分布を保存するように子集団の生成が適切に制御される。世代交代モデルMGGではランダム性を考慮した世代交代により多様性を維持しながら集団の分布を望ましい方向にシフトするよう適切に制御している。

「複雑がゆえにランダム性に頼らざるを得ないのか、それともランダム性の追及がものの本質なのかと考えさせられる」との元田氏のコメント、同感である。著者は“複雑”であるがゆえに、ランダム性に頼らざるを得ないが、ランダム性を適切に制御できなければ複雑さには立ち向かえないとの立場からGAの研究を進めてきた。“GAは適応的モンテカルロ法”という見方[山村 98]は近似アルゴリズムとしてのGAの本質を最も的確に表しているといえよう。

一方、ランダム性の追及がものの本質であるという考え方は統計力学的接近法に共通するものといえよう。統計力学的手法はマイクロなレベルでのランダムな挙動をマクロなレベルで理解することを目的として、平均場近似、スピングラスモデル、モンテカルロ法などの理論や手法を発展させてきたが、近年、大規模統計モデルに対して、統計力学的手法を用いて近似アルゴリズムを構成し、適用する試みが大きな成果を上げつつある[田中 03]。

ランダム性に頼った近似アルゴリズム(GA)とランダム性の本質を追及するために開発された近似アルゴリズム(統計力学的手法)が大規模情報処理において有力なツールとして注目されつつあるのは興味深い。はじめに述べたように、経験的知識に基づく第一世代知識ベースは失敗に終わったが、経験的知識を近似アルゴリズム(ファジィ推論)に結びつけたファジィ制御は実用化に成功している。これらの生まれも育ちもまったく異なる近似アルゴリズムの間には何らかの共通原理があるように思われる。

### 理解容易性と説明可能性について

GAは解の最適性を保証する解法ではないので、複雑な最適化問題に対してGAを用いてある解が得られたとき、これをどのように理解して受理したらよいのかという問題に直面する。近傍を適当に定義すれば、その範囲内での局所最適性のチェックは可能である。大域的最適性を証明することは原理的に困難であるが、何らかの方法で得られた解が“有力な局所解”であることを示せば、受理しやすいであろう。

素朴な方法は他手法との比較である。他手法ではGA

に比べて同等以上の解を得ることができない、あるいは時間がかかりすぎるということを示すことができれば、相対的に優位であるという理由でGAの解が受理される場合もあろう。しかし、解法としての相対的な優位性を示すためには比較手法として何を選べばよいのかという別の問題に直面するし、計算コストの問題も生じる。また、相対的な優位性を示すだけでは不十分な場合も考えられる。

適切に設計されたGAが大域的最適解の発見に失敗するのは、解空間がUV構造を有する場合である。GAによる最適化の前に、対象問題の適応度景観の分析を行い、最適解または有力な局所解が存在すると思われる有望な谷を同定し、各谷がUV構造のどのクラスに属するかを明らかにすることができれば、GAによって得られた解に対する理解容易性は高まるものと思われる。

UV構造の存在を確認するためには、景観分析が必要かつ有効との考えから、著者らは最近この問題に取り組みはじめたところである。基本的なアイデアは、解空間を適当に分割して、各部分空間に対して順方向GA(通常のGAによる局所解の探索)と逆方向GA(局所解から改悪方向への探索)を適用して、それぞれの収束特性データを収集する。収束特性の比較データは各谷の相対的な広さ、深さ、険しさなどを反映しているため、これらを分析することにより、適応度景観の複雑さが明らかになると見込まれる。このような景観分析の結果を踏まえて、最も有力な局所解が存在すると推定される領域を絞り込み、重点的に探索することが大域的最適解の発見につながるものと予想している。ベンチマーク問題とレンズ設計問題への予備的な適用からこのようなアプローチが有効との感触を得ている。

### GAと知識獲得について

「新しい知識獲得技術としての活路がGAに見いだせそうな予感がする」とのコメントにはまったく同感である。著者らは方法論としてのGAを鍛えるために、これまで明確に定義された最適化問題を中心に研究を展開してきた。かなり複雑な適応度景観をもつ大規模問題へ適用できるめどもついてきたので、これからは実問題への応用にも力を入れていきたいと考えている。見掛け上複雑な問題も適切に切り出せば、最適化問題として定式化される場合が多数あると思われる。

元田氏が例として上げられたバイオインフォマティクスの分野では膨大なデータの中から規則性または異常性を発見することが重要な意味をもつ問題が山積している。現実には、ポストドククラスの分子生物学の若手研究者がデータの洪水の中で日夜データマイニングのために格闘しているのが実情である。この分野でのGAによる知識獲得支援は彼らの生産性を飛躍的に向上させる可能性を秘めている。同僚の山村雅幸君は5年前よりバイオインフォマティクスの分野に参入して、GAで培った方法

論を駆使して活躍している。

方法論的観点からデータマイニング技術に貢献することも考えられる。例えば、データマイニングにおける部分問題としてのクラスタリングでは、複数の正規分布の重ね合せによってモデルを記述することがよく行われる。EM法は代表的な推定法であるが、高次元データに対して計算の不安定性や局所解に陥る問題が指摘されている。実数値GAは交叉によって正規分布を非明示的に推定する方法であり、EM法と融合させることにより、ロハストな手法を構築することができる。詳細は[佐久間03]を参照されたい。

## 5. おわりに

三人のコメンテータのコメントから多くのことを学ぶことができた。浅田氏からはビジョンと哲学を語ることの意味と意義を再認識させられた。本稿では著者の研究に対する基本姿勢や研究戦略は述べた。研究の大まかな方向性だけ提示して、あとは若い研究者の創造性を引き出すために、できるだけ制約を課さないで自由にやらせるのがよいと著者は信じている。

三木氏とは今年1月に同志社大学で開催された情報処理学会主催の「進化的計算シンポジウム」で初めてお会いしただけの間柄であるにもかかわらず、著者の考え方をすべて見通しておられることに驚かされた。コメントから察するに、三木氏はGAだけでなく周辺の動向も的確に把握されているせいであろうが、GAに対する思い入れは冷静かつ客観的であると感じた。互いに良き理解者として、良き好敵手として今後ともつき合っていきたいと思っている。

元田氏からは多様な観点からコメントをいただいた。機械学習とデータマイニングの分野で幅広く活躍されているだけあって、本質をついたコメントから考えさせら

れることが多くあった。著者は問題解決の方法論としてのおもしろさからGAの研究に取り組みはじめた。同じ理由から強化学習の研究にも十年近く取り組んでいる。データマイニングはGAと強化学習にとって魅力ある応用分野であり、方法論を鍛える意味でも興味があるので、これを機会にデータマイニングにも取り組んでみたいと思っている。

生命システム、生物システム、生態システム、社会システム、経済システム、経営システム、政治システムなどの複雑系の解明に向けて、構成的アプローチ(Analysis by Synthesis)の必要性和重要性は今後ますます高まってくると思われる。これらのシステムに共通するキーワードは“進化”であり、“適応”である。GAは構成的アプローチの最も強力な方法論である。実数値GA(UNDXファミリー)は実用レベルで使える日本発の世界に誇れる研究成果である。様々な分野での利用が広まることを期待して結びとする。

## ◇ 参考文献 ◇

- [池田02] 池田 心, 小林重信 GAの探索におけるUV現象とUV構造仮説, 人工知能学会論文誌, Vol 17, No 3, pp 239-246 (2002)
- [喜多99] 喜多 一, 小野 功, 小林重信 実数値GAのための正規分布交叉に関する理論的考察, 計測自動制御学会論文集, Vol 35, No 11, pp 1333-1339 (1999)
- [倉橋03] 倉橋節也, 勝又勇治, 寺野隆雄 ハイシアン最適化手法と分布推定アルゴリズムの動向, 人工知能学会誌, Vol 18, No 5, pp 487-494 (2003)
- [佐久間03] 佐久間 淳, 小林重信 確率分布推定に基づく実数値GAの新展開, 人工知能学会誌, Vol 18, No 5, pp 479-486 (2003)
- [田中03] 田中和之, 樺島祥介 大規模統計モデルに基づく情報処理—モノの理とコトの技の接点—, 計測と制御, Vol 42, No 8, pp 622-625 (2003)
- [山村98] 山村雅幸 モンテカルロ法による遺伝的オペレータの解析, SICEシステム/情報合同シンポジウム, pp 157-162 (1998)