

GAにおける命題論理式の遺伝子表現

6Q-1

山口 哲司 小高 知宏 小倉 久和

福井大学

1 はじめに

遺伝的アルゴリズム(GA)は、生物の遺伝的な進化機構をモデル化した確率的なアルゴリズムであり、学習・最適化などの一手法として様々な方面の分野において研究、応用されている。近年では、行動戦略や判別ルールなどの知識獲得の研究も盛んに行われている[1]。このようにGAは、非常に汎用性に富んだ応用範囲の広い手法であるが、その反面、GAを設計する際には各々の問題に固有な特徴、性質といったものを考える必要がある。特に、GAを知識獲得に応用する際には、知識をどのように表現し、それをどう遺伝子としてコーディングするかが重要な検討課題と言える。

本研究では、GAによる知識獲得の応用として論理関数の式表現を獲得する試みを行った。論理関数とは、基本的なブール演算である和、積、補元の組合せで表される関数であり、ブール関数と呼ぶこともある[2]。論理関数は関数値として0あるいは1をとり、真理値表によって表すことができる。また逆に真理値表から論理関数の式表現を導くことができる。これをGAによって行なうというのが本研究の目的である。

論理関数の式表現は、別の見方をすれば、入力(変数値の組)に対して0であるか1であるかの判別を行なう規則であり、論理関数の式獲得は判別ルールの学習と考えることもできる。このような判別ルールの学習の対象として、我々は血圧波形から動脈硬化の有無を判別するという問題を検討している。これは、動脈硬化を起こしている人の血圧波形には何らかの特徴が見られ、そこから判別ルールを見つけられるのではないかというものである。本研究では、このような問題に対するGAの基本的な考え方を検討する。

2 GAの適用方法

2.1 遺伝子コーディング

ここでは論理関数の式表現として、例えば、

$$f = \overline{(x_1 \cdot x_2)} \cdot ((\overline{x_1} \cdot x_2 + x_3) \cdot \overline{x_1})$$

のように2項演算のAND、ORとNOTの組合せにより構成される式を扱う。これを図1のように木構造で表

したものを遺伝子として扱うことにする。

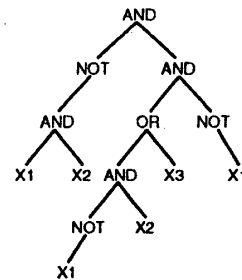


図1: 木構造を用いた遺伝子の例

遺伝子プールのサイズは30とし、初期集団の遺伝子においては、rootのノードにリテラルがきたり、木のサイズ(ノードの数)が20以上になるものは生成しない。

2.2 遺伝子の適応度

遺伝子の適応度は、遺伝子の表す式を計算した結果、問題の真理値表の関数値と一致した数とする。変数の数をnとすると2^n通りの変数値の組があるので、2^nが適応度の最高値となる。

2.3 選択

適応度による上位6つの遺伝子を無条件で次世代の集団に残す。さらに、12組のペアを生成し、交叉を行う。

2.4 交叉

交叉は2つの親遺伝子を選び、各親遺伝子中のノードを1つランダムに選択する。そして、選択されたノード以下の部分木を2つの遺伝子の間で交換する。

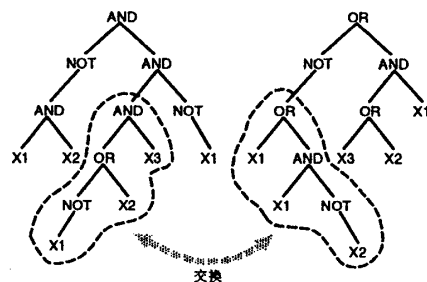


図2: 交叉の例

2.5 突然変異

突然変異は、交叉によって新しく生成された各遺伝子に対して 0.2 の確率で起こす。方法は、遺伝子中のノードをランダムに1つ選び、ノードのラベルがリテラルならば他のリテラルに、ANDならORに、ORならANDに置き換える。NOTに対しては突然変異は行っていない。

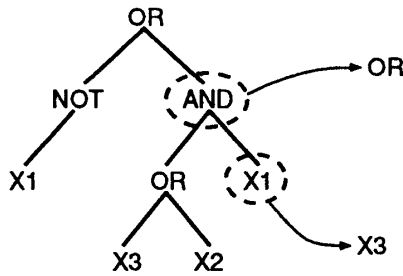


図 3: 突然変異の例

3 結果と考察

図4のグラフは、5変数の問題3種類に対して実行した結果である。横軸は世代、縦軸はその世代において最も優秀な遺伝子の適応度を示している。上のグラフは $f = x_1 \cdot x_2 + x_4 \cdot x_5$ という関数、中のグラフは多数決関数(5変数のうち3つ以上の変数の値が1ならば関数値が1という関数)、下のグラフは真理値表をランダムに作成した関数の場合の実行結果である。

最初の問題のように、式で表した場合に比較的簡単な形で表すことができる問題に対しては、図4の例のように真理値表を完全に満足する式を獲得できることもある。しかし、これはうまくいった場合の例であり、必ずうまくいくとは限らない。また、最終的に得られる式も様々である。これに対して多数決関数やランダムに作成した関数のように式表現が複雑になるような問題に対してはあまり良い結果が得られない。

今回検討してきたGAの特徴として次のことが考えられる。まず、この遺伝子コーディング法では論理関数の性質からも明らかなように、1つの真理値表(表現型)に対して複数の遺伝子(型)が存在する。このため、GAにおける探索の目的点が無数に存在することになる。

次に、GAでは遺伝子中の有益な部分構造という概念は非常に重要であると考えられる。これは交叉とも密接に関連することで、有益な部分構造が遺伝子の適応度を向上させ、交叉によって子孫に継承されていくような構造にコーディングされることが望ましいと考えられる。しかし、今回のように木構造を遺伝子として扱う場合、部分

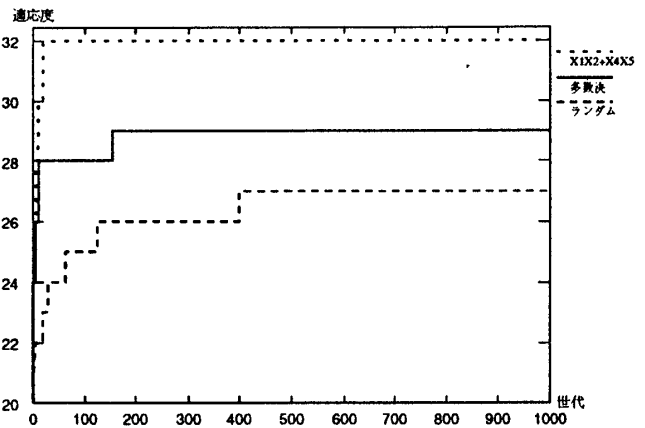


図 4: 実行結果

木の意味が上位ノードによって変化するため必ずしもも有益な部分構造が遺伝子全体の適応度に反映されない。このため、交叉が有効に作用していないと考えられる。

今回検討してきた突然変異は、ノードのラベルを交換するというものであるが、これは場合によっては木の表す意味、つまり遺伝子の表す式から得られる真理値表が大きく変化してしまうことがある。つまり遺伝子型と表現型で、突然変異による変化が違っている。突然変異の役割としては、集団中の多様性を維持することと、局所的な探索が挙げられるが、後者に関してはその役割を果たしていないと考えられる。

4 今後の課題

GAの設計において遺伝子のコーディング方法の検討は非常に重要な問題である。本来、有益な部分構造が遺伝子の適応度を向上させ、交叉によって子孫に継承されるような構造にコーディングされることが望ましいが、問題によってはそのようなコーディングが困難な場合もある。本研究で検討してきた問題もこのような類の問題であると考えられる。このような場合は、交叉、突然変異などのGAオペレータを工夫するというのも1つの手段であると思われる。この問題に対して遺伝子のコーディング方法としてより適切な方法を検討している。

参考文献

- [1] 北野 宏明:「遺伝的アルゴリズム」, 産業図書 (1993)
- [2] 小倉 久和, 高濱 徹行:「情報の論理数学入門」, 近代科学社 (1991)