

構造的表現の適応学習とその応用

伊庭斉志

Hitoshi IBA

電子技術総合研究所知能情報部推論研究室

Electrotechnical Laboratory

Machine Inference Section

我々は、遺伝アルゴリズム (GA) に基づいて高次の知識表現を的確に扱い、適応的な学習を実現することを試みている。そのために、本稿では、GA のオペレータを拡張し、構造的な表現の学習を試み、多くの知識表現で用いられるグラフ (特に木) 構造の表現の獲得学習を行い有効性を確かめる。更に、グラフ理論の問題解決などへの適用を示す。その上で本手法の問題点と、今後の課題、プラン生成やプログラム合成などへの応用について考察する。

Learning of structural representations based on adaptive method

Hitoshi IBA

Electrotechnical Laboratory, Machine Inference Section
1-1-4 Umezono, Tsukuba-city, Ibaraki, 305, Japan

ABSTRACT

We present a new scheme for learning by discovery. Our approach is based on the evolutionary learning mechanism through the use of genetic algorithms. The purpose of this scheme is to establish adaptive learning for symbolic representations. Although most of genetic algorithms usually apply to bit or low-level representations, we try to extend them to handle high-level knowledge representations, such as graphs or structured trees. The validity of this extension is shown by experiments in acquisitions of desired graphs for graph-theoretical problem solving.

1. はじめに

本稿では、適応的手法に基づいて、構造的表現の学習を実現する。さらにその手法を知識の獲得・生成の分野に拡張する試みについて述べる。

近年、情報の進化や知識表現の変換(representational change)という考え方に基づいて学習の問題を扱う試みがなされている。このアプローチは、環境との相互作用や問題解決の過程を通して、自らの内部に持つ情報構造を変革し、適応的な学習を行うものである。本稿では、これらのアプローチ、特にその代表的な手法である遺伝アルゴリズムをもとにして、発見による学習を実現する試みについて述べる。本稿で扱う学習としては、

1. 学習のデータが大量にある環境
2. 教師的な価値基準を与える神託の存在
3. 試行錯誤を伴う誤りからの復帰の可能性

を想定している。これは例えばロボットが実世界下で環境のモデルを学習しながら的確に行動することに相当する[伊庭88]。

本稿では、構造的学習の対象としてグラフ・木表現を扱う。ここで用いられるグラフ理論の用語の詳細は[Lovasz79][Lu78, 84]を参照されたい。

2. GAと適応的学習

多量のデータからの適合的学習の手法としてGenetic Algorithm (遺伝アルゴリズム、以下GAと略す)が考え出され、様々の分野で応用されはじめている。この手法の特徴は、情報を「進化的アプローチ」で操作し最適化を実現することである。

我々は、GAのシステムをLISP上に開発し、実験を通してその有効性を確認した。さらに構造的表現の適応学習を行うための拡張を試みた。以下ではその概要について説明する。

ここでは、構造的な表現として、多くの知識表現で用いられるグラフ構造(特に木構造)を扱う。

木に対するGAのオペレータとして、以下を導入する。

- | | |
|---------------|------------|
| 1. gmutation | ノードのラベルの変更 |
| 2. ginversion | 兄弟の並べ換え |
| 3. gcrossover | 部分木の取り換え |

これらは従来のビット列を対象とするGAオペレータの構造表現への自然な拡張である (Fig. 1)。

Hollandらは、一次元的文字列表現を中心に扱うGAにおける最適問題の効率的解法能力を、次のスキーマ定理で主張した[Holland78]。

$$m(H, t+1) = m(H, t) \cdot [1 - P_c \cdot P_d] \cdot f(H) / f^* + [\text{Crossoverからの寄与}]$$

ここで、 $m(H, t)$ は探索空間内の部分集合H(スキーマと呼ぶ)を含む遺伝データ数の世代t

における期待値を示す。PcはCrossoverの起こる確率である。PdはCrossoverによりスキーマの破壊される確率であり、スキーマの定義長に比例する。f(H)はスキーマの有する平均の適合度であり、f*は遺伝データ全体の平均適合度である。

GAのオペレータによって、適合度の高い特定の部分構造（スキーマ）が保持され、オペレータの適用を通して次の世代の高い適合度に貢献する。特に探索に有効なのは、小さく一般的な高適合度の部分構造である（short, low-order, highly fit scheme）。これをBuilding block hypothesisと言う。さらに、N個の遺伝子データがあるとき、上で述べた適合選択を行う場合には、一般に少なくともN³の部分構造が割り当てられている（Implicit parallelism）。このようなことから、GAに基づく最適化の探索手法の効率性が保証される。

前述のグラフ操作においても、適切にグラフ上の距離と適合度を定めることで効率的にグラフ空間内での探索が実現される。次節では、こうしたことを示す実験結果について説明する。

3. 構造的表現の獲得学習

(1) 木の類似度に基づく学習[伊庭90]

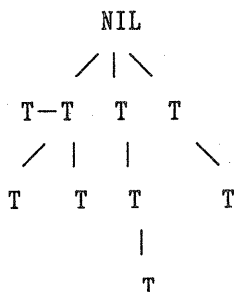
最も簡単な適応的学習の例として、ランダムにまたは意図的に生成した木（ノードはTまたはNil）に対しての獲得実験について説明する。

今、木の獲得のための適合度を次のように計算する。

$$\text{fitness}(T) = 1 / (1 + d(T, T^*))$$

ここでT*は獲得すべきモデル木である。但し、d(T, T*)はLuの距離と呼ばれ、木の類似性を示す指標として用いられ、木のノード数nに関してO(n³)で計算できる[Lu78, 84]。この適合度は1以下の正値をとり、1となったときに木が正しく獲得されたことを示す。

例えば、



のような木構造に対しての獲得の仮定をFig.2に示す。ここで各オペレータの適用確率は次のように設定した。さらに簡単のため各オペレータの適用部位はランダムに選択し、ginv

ersionは省略した。

Population size 20 (遺伝子データの数)
 Gcrossover probability 0.6 (gcrossoverの適用確率)
 Gmutation probability 0.033 (gmutationの適用確率)

図から分かるように、98世代で獲得は成功した。この実験での初期世代と獲得成功世代(98世代)での適合度の情報は次の通りである。

世代	最大適合度	最小適合度	平均適合度
初期世代	0.1667	0.0185	0.1146
98世代	1.0000	0.0909	0.3401

他のモデル(木T*)に対する実験データやその考察については[伊庭90]を参照されたい。

(2) Turanの定理の系

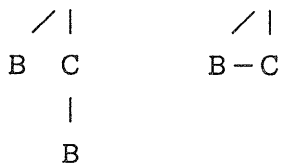
位数nの任意の単純グラフGとその補グラフG*には、合わせて $n(n-1)(n-5)/24 (= \text{num}(n))$ とおく)以上の三角形が含まれている。これはTuranの定理の証明に付随的に登場する命題である[Lovasz79, vol.3, p.99, ex.3.32(b)]。例えばn=9のとき、三角形の数は12以上となる。ここではこの数の下限を与えるようなグラフを適応的に構成してみる。そのために適合度としては、

$$\text{fitness} = \frac{1}{(\text{三角形の数}) - \text{num}(n) + 1}$$

を採用する。fitnessが1となる時が、下限を与えるグラフとなる。

さらに木の表現を用いてグラフを表わすために、ノードの重複使用を許す。

つまり、Aという木は、A というグラフの表現として考える。



これはGTYPEとPTYPEの二層の表現形式の導入に相当する[伊庭90]。

n = 9 に対しての獲得実験の様子を Fig. 3 に示す。このとき、Fig. 4 のようなグラフが 520 世代目で獲得された。用いられたパラメータを以下に記す。

Population size 20 (遺伝子データの数)
 Gcrossover probability 0.6 (gcrossoverの適用確率)
 Gmutation probability 0.033 (gmutationの適用確率)

この実験での初期世代と獲得成功世代 (520 世代) での適合度の情報は次の通りである。

世代	最大適合度	最小適合度	平均適合度
初期世代	0.0588	0.0152	0.0217
520世代	1.0000	0.0256	0.2093

この実験における探索空間は、36 (= ${}_9C_2$) のエッジ・ペアに対して、G にはいるか G^* に入るかであり、 2^{36} 程度となる。但し対照性や解の重複を除いた議論のため厳密な評価ではない。

(3) Ramsey 数に関連した問題

$r(m, n)$ 個の点を持つグラフが全て K_m (完全グラフ) または K_n^* (全非連結グラフ) を含むような整数 $r(m, n)$ の最小値は Ramsey 数と呼ばれている [Lovasz79, vol. 4, ch. 2]。 $r(3, 3) = 6$ であることはよく知られている (6 人の人間が集まれば互いに顔を知っている 3 人のグループか、お互い見知らぬ 3 人のグループを必ず含む)。 Ramsey 数を決定することは未解決の問題であるが、これに関連した命題として、 $n = n_1 + n_2$ のノードを持つグラフには、 K_{n_1+1} かつ $K_{n_2+1}^*$ を含むものがあることが導かれる。このような性質を持つグラフを適応的な探索で求めてみる。パラメータは以下の通りである。

Population size 20 (遺伝子データの数)
 Gcrossover probability 0.6 (gcrossoverの適用確率)
 Gmutation probability 0.033 (gmutationの適用確率)

適合度としては、

$$\text{fitness}(G) = \frac{1}{S(n_1 - \alpha(G)) + S(n_2 - \omega(G)) + 1}$$

を採用する。ここで $\alpha(G)$ はグラフの独立点の最大数、 $\omega(G)$ はグラフのクリーク数 (最大完全部分グラフの位数) を示す。関数 $S(x)$ は、 x が負のとき 0、非負のとき x を返す。fitness が 1 となるのが、 K_{n1+1} かつ K_{n2+1} を含むグラフとなる。

前二例と同様のパラメータを用いて、 K_5 かつ K_5^* を含む位数 9 ($= 4 + 4 + 1 = n$) のグラフを適応的に構成してみると、世代 20 から 60 のうちに成功する。他のデータについても同様に獲得は成功することを確認した。

4. 考察

前節で述べた実験では、GA のオペレータをグラフに拡張することを行い、その結果グラフにおける構造的な特徴をもとにして探索が行えることを示した。ここでの構造的な特徴と言うのは、厳密にはグラフの不変数と呼ばれるものである。それらには、染色数、独立数、被覆数などを含む。これらの不変数を適合度に組み込むことにより、グラフ全体のなす空間に適切に位相を導入し探索する手法を与える。例えば極小 m 染色グラフを構成する場合には、

1

$$\text{fitness} = \frac{1 + a \cdot (\text{辺数についての条件}) + b \cdot (\text{k-colorableか}) + c \cdot (\text{極小か})}{\dots}$$

という適合度を採用すればよいことになる。ここで a , b , c は 3 つの要素の重み付けを表わす正数である。

このような探索手法についての収束性や効率についてはまだよく分かっていない点が多い。3 節の (1) のように単純な場合には、線形性や相似性などの構造的特徴が顕著なものが容易に獲得される、という実験結果が得られている。今後の課題は、Holland のスキーマ定理の主張を構造的な表現の中で構成し直すことである。そのためには、building-block hypothesis や implicit parallelism をグラフ表現でどう捉えて行くか、が重要であると考えられる。

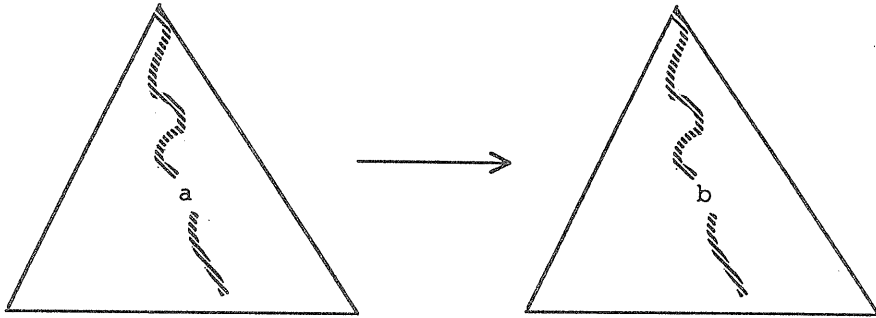
5. おわりに

本稿では、構造的表現の獲得のために GA を拡張した枠組を実現し、その有効性を示すための実験について説明した。本手法をより一般的な問題解決に応用することを現在試みている。具体的には、本手法に基づくプログラム合成、プラン生成のシステムを構築中である [伊庭91]。本手法の数学的な解析、収束性や効率の吟味は今後の興味ある課題である。

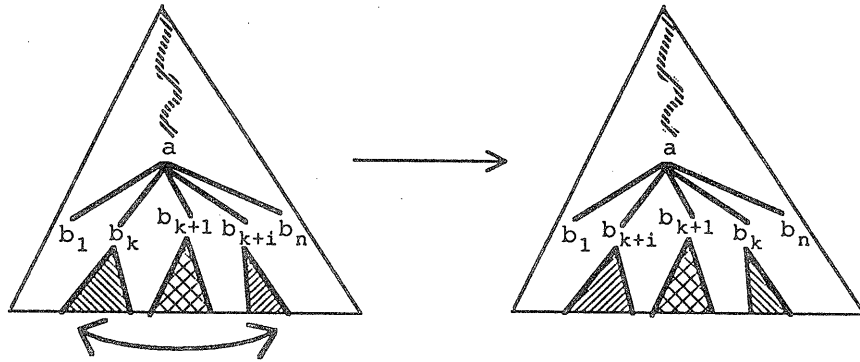
参考文献

- [伊庭88]伊庭斉志、松原仁、井上博允：“環境モデルにおける物体の見え方と見方”、人工知能学会誌、vol.3, no.4, 1988
- [伊庭90]伊庭斉志：“情報の進化と創造的学習について”、人工知能学会基礎論研究会、SIG-F/H/K-9001-10, 1990
- [伊庭91]伊庭斉志：“適応的学習に基づく知識の獲得と生成”、人工知能学会第5回全国大会、1991（発表予定）
- [Holland75]Holland, J.H.：“Adaptation in natural and artificial systems” University of Michigan Press, 1975
- [Lu78]Lu, S. and Fu, K.：“Error-correcting tree automata for syntactic recognition”, IEEE Tr. on computers, vol.c-27, no.11, 1978
- [Lu84]Lu, S.：“A tree-matching algorithm based on node splitting and merging”, IEEE Tr. on pattern analysis and machine intelligence, vol.PAMI-6, no.2, 1984
- [Lovasz79]Lovasz, L.：“Combinational Problem and Exercise”, Akademiai Kiado, Budapest, 1979
- （邦訳：組合せ論演習、全4巻、秋山仁・榎本彦衛他訳、東海大学出版会、1988）

gmutation



ginversion



gcrossover

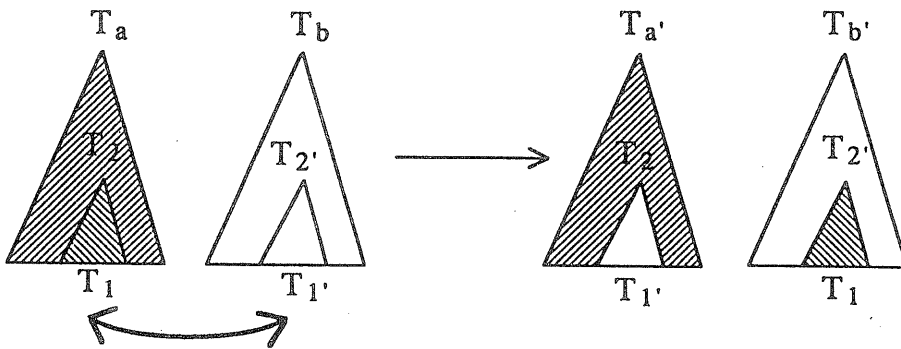


Fig. 1 GA operators on trees

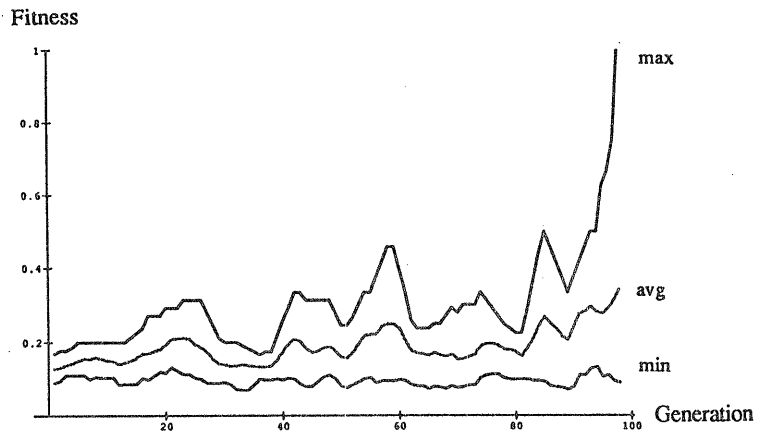


Fig. 2 Experimental result (1)

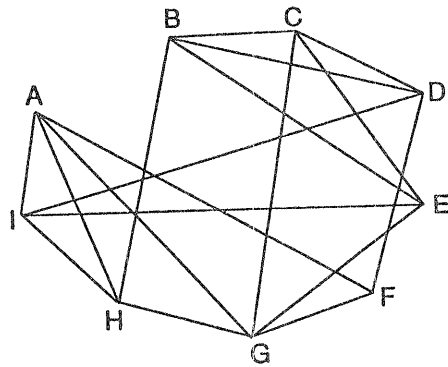


Fig. 3 Acquired graph

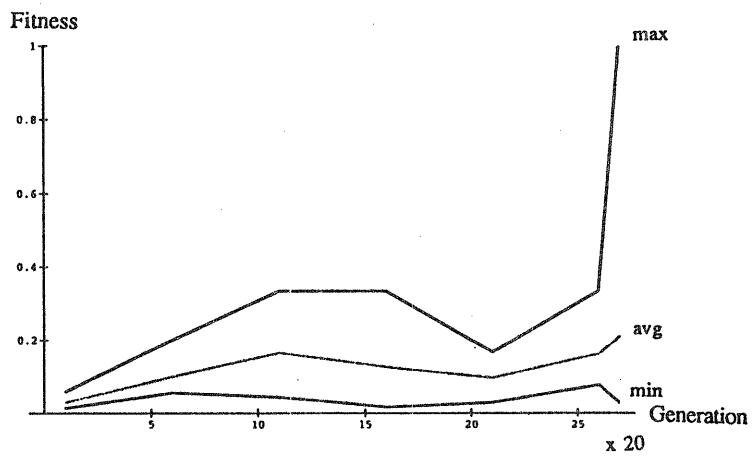


Fig. 4 Experimental result (2)