

二倍体遺伝と地理的隔離モデルを用いた 遺伝的アルゴリズムに関する研究

常念 一知 山田 新一 藤川 英司 志田 晃一郎
武蔵工業大学

1. はじめに

遺伝的アルゴリズム (Genetic Algorithm: 以下 GA) [1] は、「組み合わせ最適化問題」に特に有効な探索アルゴリズムである。

本研究は、河内らによって提案された二倍体遺伝と地理的隔離モデル [2] を取り上げる。彼らのニューラルネットワークの結合重みを決定する例題において、評価関数と個体、集団の構成を変えて調査を行ない、そのときもこのモデルが有効であることを確認した。

2節では、このモデルを検証するために用いたニューラルネットワークの構成を示し、3節では本研究で用いた二倍体のモデルについて述べる。さらに、4節では、これらの条件下で実験を行ない、その有効性について検証する。

2. 対象としたネットワークモデル

本研究で対象とするネットワークモデルは図1に示すような、入力論理和、論理積、排他的論理和の各関数を出力とするパーセプトロン型のネットワークである。各ユニットはしきい関数を持ち、結合係数-3, -2, -1, 0, 1, 2, 3の整数値をもつ。

各ユニットからの出力にはしきい値0.5のステップ関数を用い、ほかに中間層、出力層のユニットには、1の固定入力からの結合をもたせている。作りたいネットワークは、3つの出力がすべての入力に対して正解し、その上、結合数と中間層のユニット数が少ないものである。なお、重み0の結合と、入力または出力のないユニットはカウントしない。中間層の最大ユニット数を6とし、すべての結合に対応する39の遺伝子座を用意した。

評価関数として、ネットワークの3つの属性値(結合数、中間層ユニットの数、ネットワークに全入力パターンを試した正解率)のそれぞれに係数を掛けたものの和を用いた(式1)。

$$[\text{評価値}] = \frac{a(39 - [\text{結合数}])}{39} + \frac{b(6 - [\text{ユニット数}])}{6} + c([\text{正解率}]) \quad (1)$$

ただし、係数 $a = 50$, $b = 50$, $c = 900$ とする。

本研究に用いた GA の特徴を以下に示す。

- (1) 各遺伝子のとり得る値(結合重み)を整数にしている。
- (2) 二倍体と複数集団 GA による処理の2つの拡張を行なっている。
- (3) 親に選ばれる確率をすべての個体で同率とし、評価関数が最小値である個体を死滅させる。

3. 二倍体のモデル

3.1. 優性と劣性

二倍体は遺伝子を2本もった個体である。表現型として発生するとき、2本の染色体の情報を1本分に圧縮する必要がある。本研究では、2本の染色体に優劣の概念を与え、1本の染色体に相当するものを作成し、そこから表現型にあたるニューラルネットを得ている。このとき、0を優性遺伝子、0以外の遺伝子を劣性遺伝子とした。すなわち、対応する2つの遺伝子を調べ、0の遺伝子があればその値を0とする。両方とも0でなければ、表現型に関する影響は同等と考え、両者の算術的平均値をその表現型の値としている。0を優性とした理由は、結合重みが0、すなわち結合数が減少すると評価関数の値が大きくなると期待できるからである。

3.2. 二倍体における交叉と新たな個体の生成

一倍体では、1組の両親からの染色体をランダムな位置で交叉させることによって新たな子を生成する。それに対し、二倍体では図2に示すように、同じ量の情報をもつ2つの染色体がある。まず、父親となる1

個体の染色体を交叉して、1組の配偶子ができる。母方についても同じ操作が行なわれ、父方と母方からそれぞれ1つの配偶子が受精によって接合され、2つの染色体をもった新たな子の個体が得られる。

3.3. 突然変異

本研究における突然変異は点突然変異を用い、一定の突然変異率が確率的に起こす。突然変異は交叉の直後に行ない、変異の起こる遺伝子座をランダムに選び、遺伝子の値を乱数で-3, -2, -1, 0, 1, 2, 3から同確率で選ぶ。

3.4. 複数集団 GA における集団間の交流

本研究で用いた複数集団 GA は、通常の GA 処理を各部分集団内部で実行する。一定世代ごとに各部分集団から乱数で選んだ1個体と、隣の部分集団から乱数で選んだ1個体を交換する。図3にその様子を示す。

4. 実験

以上の二倍体遺伝と複数集団 GA の性能を評価するために次の4種の GA を対象とし、実験を行なった。

- (A) 一倍体, 単一集団 GA, 400 個体/集団
- (B) 一倍体, 4 集団 GA, 100 個体/集団
- (C) 二倍体, 単一集団 GA, 200 個体/集団
- (D) 二倍体, 4 集団 GA, 50 個体/集団

4種類の GA で集団当りの個体数を変えたのは、各種の GA で個体数の総数を合わせるためである。以上の4種の GA をランダムな初期配列でそれぞれ25回ずつ試行し、30000世代の GA 操作を上限とし、最適解、準最適解の探索成功率を調査した。その結果を表1に示す。

表 1: 各種 GA による探索成功率

	最適解	準最適解
(A) 一倍体・単一集団 GA	12 %	20 %
(B) 一倍体・4 集団 GA	20 %	32 %
(C) 二倍体・単一集団 GA	0 %	60 %
(D) 二倍体・4 集団 GA	88 %	92 %

最適解とは、本例題において考えられる限り最も単純なネットワーク構造であり、結合数は9つ、中間層ユニット数は2つであった。準最適解は、最適解に次いで単純な構造であり、結合数は11つ、ユニット数は3つであった。

どの GA においても準最適解に達することができたが、(A)～(C)では最適解探索成功率は低い値にとどまっている結果となった。一方、(D)二倍体・4 集団 GA においては準最適解、最適解探索ともに高い成功率を残し、初期配列状態に依存しにくいことを示している。

以上より、本例題における問題において、二倍体遺伝と地理的隔離モデルを用いた複数集団 GA は有効であるといえる。

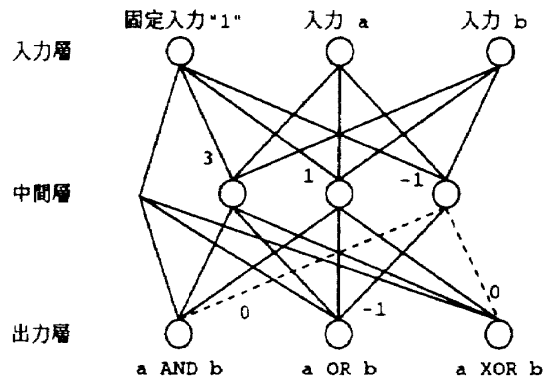


図 1: ネットワークモデルの遺伝子型と表現型

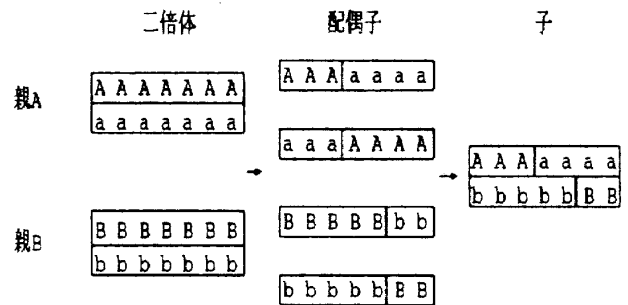


図 2: 二倍体モデルにおける交叉オペレータ

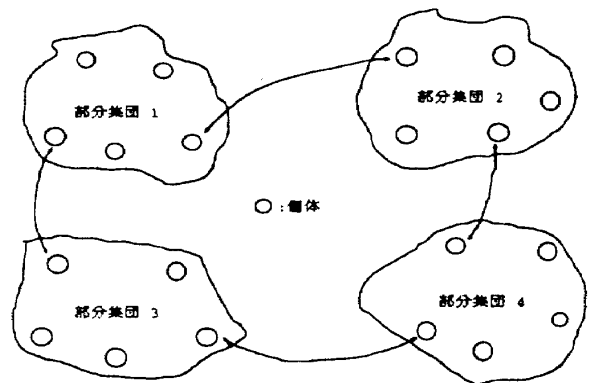


図 3: 部分集団中の個体交換

参考文献

[1] 北野 宏明: 遺伝的アルゴリズム, 産業図書 (1993).
 [2] 河内, 稲吉, 星野: 「二倍体遺伝と地理的隔離モデルを用いた遺伝的アルゴリズムによるニューラルネット構造の最適化」, 人工知能学会誌, Vol.7, No.3, page.509-517(1992).