

競争原理を用いた自律的学習及び自己組織化の研究

4K-4

菊地 英一 川守田 和男 平賀 譲
図書館情報大学

1 はじめに

遺伝的アルゴリズムを用いる際、まず問題となるのが個体の評価を行なう評価関数の作成である。評価関数の性質がシステムの性能を左右するにもかかわらず、試行錯誤的に評価関数の作成を行なっているのが現状である。本研究では対戦ゲームのように、個体間の優劣が対戦結果として直接得られる状態設定において、その結果を用いることで評価関数の明示的定義を回避することを試みる。また、遺伝的アルゴリズムを適用する場合、個体の評価は生成された個体毎に行なえる必要があったが、対戦原理を用いることでこの問題も解消できる。

個体は一つのニューラルネットワークとして表現し、それらに対戦させ優劣を決定する。その結果をもとに遺伝的アルゴリズムの手法を適用し、より優れた個体の生成を目指す。この考えをもとに、簡単な対戦ゲームを扱ったシステムを作成し、その適用可能性の検討を行なった。

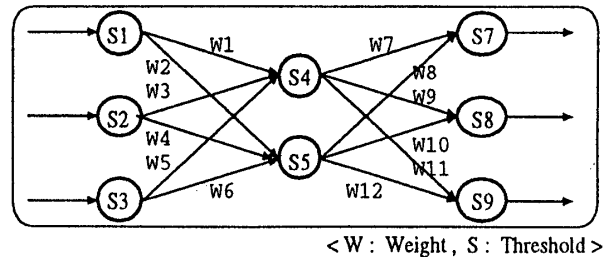
2 システムの概要

2.1 遺伝子のコーディング

遺伝子はニューラルネットワークの結合加重、及びしきい値をまとめたものとして表現し、対戦の各手ごとに対応したニューラルネットワークを生成する。よって、対戦が最大 n 手で終了する対戦においては、 $G1 \sim Gn$ の n 個のニューラルネットワーク（遺伝子）が対応し、これをまとめたものが一つの個体となる。 $G1_weight$ は結合荷重を、 $G1_threshold$ はしきい値をまとめたものである（図1）。

2.2 評価・対戦

個体の評価は個体同士を対戦させその勝敗をもとに行なう。実際には総当たりの対戦を行い、対戦の成績の良かった順にランクづけしそれぞれの個体に評価値を与える。したがって評価は完全に相対的で、また外的基準等は一切用いない。対戦は先攻の一手目の出力をもとに後攻の一手目の入力を決定し、同じように後攻の一手目の出力をもとに先攻の二手目の入力を決定する。以下同じように決着がつくまで繰り返す（図2）。



$G1 \begin{cases} G1_weight : W1 W2 W3 W4 W5 W6 W7 W8 W9 W10 W11 W12 \\ G1_threshold : S1 S2 S3 S4 S5 S6 S7 S8 S9 \end{cases}$

個体 : $G1 G2 G3 \dots Gn-1 Gn$

図1 ネットワークの構造と遺伝子-個体の表現方法

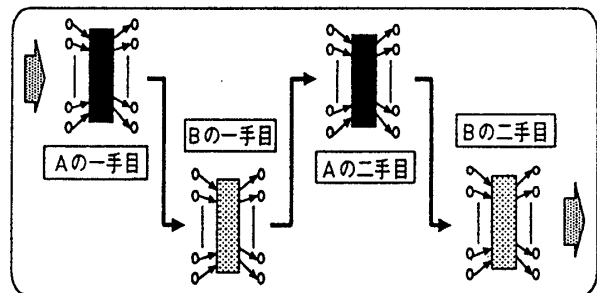


図2 対戦におけるデータの流れ

2.3 遺伝的操作

集団に対して各個体の評価をもとにした選択を行なった後、交差、突然変異を行なう。交差には二通りの方法がある（図1）。

[1] 遺伝子間交差

二つの個体をランダムに選択した後、遺伝子単位で交差点をランダムに決定し、一点交差を行なう。

[2] 遺伝子内交差

二つの個体をランダムに選択し、その中の遺伝子の一つ、ランダムに選択する。 $G1$ が選択された場合、 $G1_weight$ 、 $G1_threshold$ のそれぞれにおいて交差点をランダムに決定し、一点交差を行なう。始めは遺伝子間交差を行ない、世代が進むにつれ遺伝子内交差に換えていく。こうすることにより、最初はおおまかな学習を、後にはより細かい部分での学習を目指す。

突然変異にも二通りの方法がある。

- [1] 個体を1つランダムに選択し、遺伝子を1つランダムに選択する。G1が選択された場合、もう一度G1をランダムに生成し直す。
- [2] 個体を一つランダムに選択し、遺伝子を1つランダムに選択する。G1が選択された場合、さらにG1_weight、G1_thresholdそれぞれにおいて、一つをランダムに選択し、もう一度ランダムに値を生成する。

狙いは交差と同じである。

3 具体的な問題への適用

具体的な対象として、簡単な対戦ゲームを扱う。その一つとしてここでは○×ゲーム (tic-tac-toe) を挙げる。○×ゲームは最大9手で終了する対戦ゲームである。対戦は総当たり方式で行い、一回の対戦は互いに先攻後攻となるように2回勝負とする。

勝負の判定は

2勝0敗0分け, 1勝0敗1分け => 勝ち
 0勝2敗0分け, 0勝1敗1分け => 負け
 0勝0敗2分け, 1勝1敗0分け => 引き分けとする。

3.1 評価方法

実行は、二つの評価方法を用いて行なった。

<評価の例>

ある個体の総当たりの結果 => 20勝2敗18分け

* 評価1 (勝ち優先評価)

評価値 [132] = $5 \times (\text{勝ち数}) + (-2) \times (\text{負け数}) + 2 \times (\text{分け数})$

* 評価2 (分け優先評価)

評価値 [146] = $2 \times (\text{勝ち数}) + (-2) \times (\text{負け数}) + 5 \times (\text{分け数})$

3.2 実行結果

<結果1>

個体数50とし一回の実行は、世代数200まで行なった。表1の結果は、評価1、評価2の評価方法でそれぞれ50回実行し、各回最も評価値の高かった個体の平均を求めたものである。

	勝ち数	負け数	分け数	評価値
評価1	4.8	0.0	45.2	114.4
評価2	2.2	0.0	47.8	106.6

表1 実行結果1

<結果2>

これは、結果1の50回の実行のうち、評価1、2のそれぞれにおいてもっとも評価のよかった個体と人間が対戦した。個体との対戦は50回行ない、人間側は

先攻、後攻どちらの場合も敗けないように対戦した。表2はシステム側の成績である。

	勝ち数	負け数	分け数
評価1	0	44	6
評価2	0	34	16

表2 実行結果2

4 考察

○×ゲームは人間同士が最善に対戦すれば引き分けになるゲームである。実行結果からも優れている個体は勝ち数よりも分け数の方が圧倒的に多いことが分かる。勝ち優先評価の場合でも引き分けの多い個体の方が優れていた。この点ではゲームの性質が結果に現われており、個体に対し定石等の明示的な指示を与えることなく、システムは意味のある反応を示した。

しかし、結果2から、優れた個体の生成という点では不十分な結果に終わっている。結果1で負け数0という点では優れていると判断できるが、対戦を行なった集団のレベルが低かったために結果2のようにほとんどが負けとなった。これは、一種の局所解に陥ったためであり、これを回避するにはより多様性を持たせた集団を生成し世代交代を繰り返していく必要がある。また、結果2で分け数に違いが出た原因は、勝ち優先の評価を行なった場合に生成される個体は勝ちを急ぐ傾向がみられ防御が不十分であり、一方の分け優先の方は勝ちよりも防御を優先した結果であると考えられる。

以上のことから、現時点ではまだ良い結果は得られていないと判断できる。また、問題点として勝負の判定における引き分けの扱い方、実行時の個体数の少なさ等が挙げられる。今後この問題点を改善し、さらに交差方法、ネットワークの構造等に改良を加え、より適した形を考える必要がある。また、明示的な指示を個体に与えて学習を行ない、優れた個体を生成し、その個体の持つ値と本システムの生成する個体の値とを比較し、その結果をもとに本システムの適用可能性を検討する予定である。

参考文献

- [1] 桐谷滋. 入門と実習 ニューロコンピュータ. 技術評論社, 9 1989.
- [2] 甘利俊一. ニューラルネットの新展開. サイエンス社, 11 1993.
- [3] 八名和夫, 吉田久, 土橋健太郎, 寺崎健. ニューラルコンピューティング入門. 海文堂出版, 2 1993.