

5 J-5

# 経験に基づく類推木を用いたオセロ のパターン認識学習の理論及び実証

石川大介

神奈川大学理学部情報科学科

## 1 始めに

現在、最先端を行くオセロ（リバーシ）プログラム [1] [2] では、学習機能の搭載は欠かせない。その中でも、盤面の評価をするのに統計的なパターン認識の手法を用いているが、本稿ではこの学習をより人間に近い形で学習するために考案した理論を提示すると共に、その効果を実証した。

## 2 背景

1997年5月にチェスの人間チャンピオンである Kasparov と IBM のチェスマシン Deep Blue が対戦し、Deep Blue の勝利に終わった。しかし、Deep Blue には学習機能が一切無く、これは現在の人工知能及び学習の研究が、コンピュータチェスに対して何も貢献出来なかったのである [3]。つまり、実用（専門）的な領域に耐えうるような学習機能のシステム作りがいかに難しいかを示唆していると考えられる。逆に、学習機能が有効に働いたものは、オセロを筆頭にチェッカーやバックギャモンなどがある。

現在のオセロプログラムの学習機能の発端は、1990年のリーらの BILL からであり、盤面にパターン認識の手法を用いて静的評価関数を自動学習させた。それ以降、この手法が

広く使われ、現在では Michael Buro 作の logistello が最強で、1997年8月に世界チャンピオンの村上氏と6番勝負をしたが、logistello の全勝に終わった。

## 3 理論

この理論は「経験に基づく類推木により知識を得る」と定義され、盤面からあるパターンを抽出し、そのパターンが最終的にはどうなるかを、そこから予期される木を生成し、今までの経験を重み付けした木（類推木）によって評価するものである。

この理論の根拠として、オセロのルールはベクトル指向であり、そのベクトルに沿って抽出したあるパターンもまたローカルなオセロゲームとみなすことが出来る。よって、最終的に石数が多い方が勝ちであるため、それがあるパターンから類推木を用いて評価する。

### 3.1 評価関数

あるゲームの盤面を  $8 \times 8$  行列の  $B$  とした場合、その盤面からあるパターンを抽出した行列を  $P$  とし、この  $P$  を10進数に変換した値を  $p$  と置く。

類推木を

$$W = \begin{pmatrix} w_{11} & w_{12} & \dots & w_{1m} \\ w_{21} & w_{22} & \dots & w_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m1} & w_{m2} & \dots & w_{mm} \end{pmatrix}$$

という  $m \times m$  行列  $W$  とする。

Application of analogical reasoning tree based on experience to learning pattern in Othello

Daisuke ISHIKAWA

e-mail: dais@hbc.info.kanagawa-u.ac.jp

Department of Information Science, Kanagawa University

ある盤面を学習する際、その盤面から目的のパターンを抽出する。その時、あるパターンが  $a$  から  $b$  に変化していた時、

$$w'_{ab} = w_{ab} + 1$$

とする。この方法で学習して類推木に重みを付けるのである。

こうして出来た  $W$  を用いたパターンの評価関数は、

$$f(p) = \begin{cases} \sum_{i=0}^m \frac{w_{pi}}{s(p)} f(i) & [s(p) \neq 0] \\ v(p) & [s(p) = 0] \end{cases}$$

と定義する。但し、

$$s(p) = \sum_{i=0}^m w_{pi}$$

$$v(p) = \text{disc\_counter}(P) \quad (\text{石差を出す関数})$$

である。

## 4 実験

初手から8手目までランダムな局面(約40万)を生成し、その盤面から初手を始めるゲームを想定した。その各ゲームに対して、

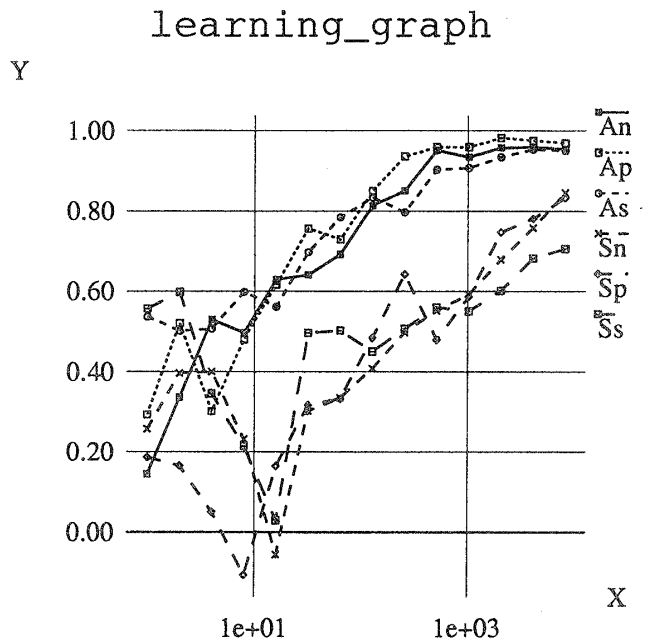
- ランダムにゲームを進行していく  $n$  タイプ(観戦学習)
- ランダムに着手するプレイヤーと対戦する  $p$  タイプ(実演学習)
- 自己対戦  $s$  タイプ(自己学習)

の3通りの学習スケジュールを設けた。今回学習させるパターンは辺のみである。

上記の学習スケジュールに沿って、類推木を用いて学習するプログラム A と、以前の一般的な方法である統計的な手法で各パターンの勝率を得るプログラム S とを比較した。

学習回数を  $x$ 、その時点に置ける強さを  $y$  とする。 $y$  は、学習回数が2の乗数毎に、ランダムゲームと1000回試合した結果の勝率差とする。こうして、8192回の学習結果を出した。

## 5 結果



グラフから分かるように、どの学習スケジュールタイプにおいても A は S に比べて約100倍の学習速度の加速が見られた。

## 6 終わりに

今回は辺のパターンのみの学習であったが、この方法で他のパターンを学習する場合、様々な問題が生じる。例えば対角線等は厳密なローカルゲームと見なせないからである。今後、それらの問題を解決すると共に、他の分野において有用かどうか検討する必要があると思われる。

## 参考文献

- [1] M.Buro: *An Evaluation Function for Othello Based on Statistics*, NEC Research Institute Technical Report.
- [2] M.G.Brockington: *Keyano Unplugged - The Construction of an Othello Program*, Department of Computing Science University of Alberta
- [3] 松原 仁: Deep Blue の勝利が人工知能にもたらすもの、人工知能学会誌、vol.12, No.5, pp.698-703(1997)