

将棋における cost function を用いた評価関数の調整

1N-4

薄井 克俊, 鈴木 豪, 乾 伸雄, 野瀬 隆, 小谷 善行
(東京農工大学 工学部 電子情報工学科)

1 はじめに

コンピュータ将棋を作る上で重要な構成要素の一つに、局面を評価する評価関数がある。

一般的に、コンピュータ将棋の評価関数は特徴量と重みを掛けたものを足しあわせる線形多項式の形が使われている。それぞれの評価要素の重みは人間が適当な値を与えているが、最適な値を与えるのは困難である。

本研究ではプロの指し手を手本として、コンピュータの指し手をプロの手に近づけることによって自動的に評価関数の重みを得ようと試みた。その手段として cost function[1]を用いた重みの調整システムの作成と実験を行った。

2 cost function を用いた調整システム

cost function とは、ある局面について、プロの指し手とコンピュータが評価関数によって得た指し手の差がどれくらいあるかを表すものである。

cost function は様々なものが考えられる。例えば、ある局面の可能手の中で、コンピュータの手とプロの手を指した後の、それぞれの局面の評価値を比べたものなどである。これらの評価値はその差が小さいほどよい。もし、評価関数の調整によって差を 0 にできれば、その局面においてコンピュータはプロの手と同じ手を指せることになる。

本研究は、cost function の値が最大（または最小）になるように評価関数の重みを調整し、最適な値を得ようとするものである。

2.1 cost function の設計

実験では cost function $C_p(w)$ に次のものを用いた。

$$c_p(w) = |\{j \mid E(p_d) \geq E(p_j)\}|$$

E: 評価関数（静的なもの、または先読みをした結果の値）

P: 学習する局面

P_d : プロの指し手 d を指したあとの局面

P_j : 指し手 j を指した後の局面

W: 評価関数の重み

$|\cdot|$: 集合の要素数

これはある局面においてプロの指し手の評価値が可能手の中で何位にあるかをあらわす。

一つの局面だけで調整したのでは正しい重みは得られない。このため、実験では 500~1000 局面を使って学習を行った。実際に cost function の値として使うのは各局面について計算された値を全て足したものになる。

2.2 山登り法による重みの調整法

cost function の値を学習局面に対して最小にするために、山登り法による重みの調整を行った。

係数ベクトル $w = (w_1, w_2, \dots, w_n)$

cost function $c(w)$

定数 k (1 以下)

勾配ベクトル $\text{grad } c(w)$

$$\text{grad } c(w) = \left(\frac{\partial c(w)}{\partial w_1}, \frac{\partial c(w)}{\partial w_2}, \dots, \frac{\partial c(w)}{\partial w_n} \right)$$

とし、

$$w_{n+1} = w_n - k \text{grad } c(w)$$

を反復計算することで係数ベクトルを求める。

3 評価関数の調整実験

3.1 実験方法

実験 1

盤上の駒の価値（基本点）は経験から得られた適当な値をまえて与えておく。その上で、

- 基本点に対する、持ち駒の価値の比
- 王との距離に応じて、駒に与えられるボーナス点の基本点に対する比

を評価要素とし、評価値を計算するようにした。

基本点は次のようにした。

王	飛	角	金	銀
10000	1000	900	700	600
桂	香	歩		と金
400	300	100		660
竜	馬	成銀	成桂	成香
1200	1100	690	680	670

探索の深さは0~4、学習する局面数は500~1000で実験を行った。

実験 2

評価要素と評価要素数を次のようにした。

盤上の駒の価値	14
持ち駒の価値	8
駒と王の距離	28
駒の盤上の位置	14
要素数計	64

探索の深さは0~2、学習する局面数は500~1000で実験を行った。

3.2 実験結果

実験 1 における「持ち駒の基本点に対する比」の調整の結果を図 1 に示す。初期値は 1(%)である。

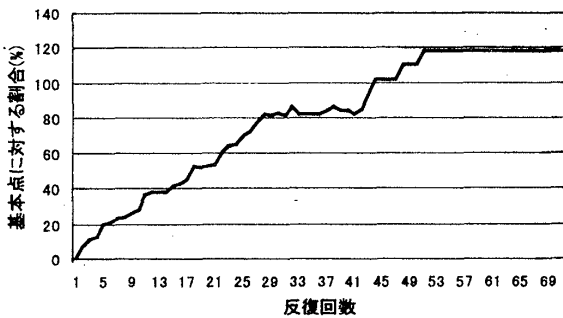


図 1 持ち駒の価値

実験 2 における駒価値調整の結果の一部を図 2 に示す。重みの初期値はすべて 1 である。

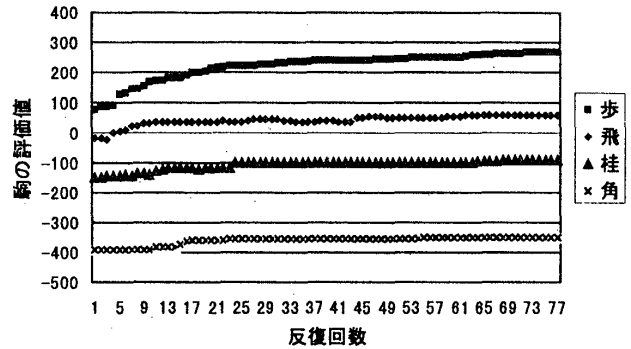


図 2 駒価値調整の様子

4 考察

- 実験 1 の駒の基本点に対する持ち駒の価値の割合については、経験的にわかっている値とほぼかわらない妥当な値といえる。
- 実験 2 では、大駒の価値が低くなり、一般の感覚とは一致しない。角同士が取れるのに取らない局面が多いと角を取る価値が下がるので、角の駒価値が低くなると考えられる。
- 駒の盤上における位置についての価値や駒と王との距離についての価値は、調整を進めても重みが正負の間で振動するものが多く、これといった値は得られなかった。これらの値は局面によって異なり、固定値にするのは困難である。

5 おわりに

本稿では、cost function を使った将棋の評価関数の調整実験を行った。その結果、

- 一部の重みはもっともらしい値を得られたが、多くの場合は現実離れした値だったり、値が得られなかったりした。
- cost function を使った評価関数の調整は、将棋でもある程度有効であることがわかった。

6 参考文献

[1] T.A.Marsland: EVALUATION-FUNCTION FACTORS, ICCA Journal vol.8 No2, pp47-57, ICCA, 1985