

オセロプログラムの評価関数の改善について
 Improvement of the evaluating function of Othello program

大筆 豊
 Yutaka Ohfude

鳥取環境大学 情報システム学科
 Department of Information Systems, Tottori University of Environmental Studies

Abstract: In order to develop game program such as Othello, Chess, Go and so on, it is needed to construct evaluating function which evaluates the state of the game. I select following five parameters for evaluating function to develop Othello program: Board Position(BP), Centering(CT), Straight(ST), Piece : Koma Number(KN) and Candidate Number(CN). In order to improve these parameters and evaluating function, I make to play computer Othello programs which have different parameters or evaluating function, and determine strong parameters which win the game. In this way, I can improve parameters and evaluating function, gradually.

1. はじめに

オセロやチェスといったゲームプログラムでは、最終局面まで読ませることは、ゲームツリーが膨大になることから、現実上不可能である。このため、何手か先読みをさせ、その時点での局面を評価することが必要になる。このため評価関数を定義し、それに基づき定量化する必要があるが、評価関数をどのように作ればいいのか、評価関数のパラメータをどのように決めればいいのかについては明確でない。

本研究では、それぞれ異なる評価関数を持つコンピュータプログラム同士を対戦させ、強さの評価を行なっている。具体的には、異なる評価関数(A, Bとする)を持つ2つのプログラムを先攻、後攻(A:B, B:A)でそれぞれ100回づつ戦わせ、両方で勝ち数が多いとき強いと仮定している。

2. 評価パラメータ

オセロを戦う上で経験的な戦略がある。これらの先験的な戦略を数値化し、それら

を組み合わせることで評価関数とすることにした。

評価パラメータとして盤位置(BP)、中心性(CT)、直線性(ST)、駒数(KN)、候補数(CN)と名づける5つを選んだ。この他、説明上、現在の手数(t)、先読み数(p)および風潰し先読み数(s)を用いる。sは、最終局面に近づいたとき最終局面まで読ませる先読み数である。

① 位置(BP) : 駒が盤のどの位置にいるかにより評価するもので、図1. に示す重みをつけることとしている。重みは暫定的にCP=6、NP=-2としている。

$$BP = \sum_{i=0}^7 \sum_{j=0}^7 BP_table[i,j] * K[i,j]$$

	0	1	2	3	4	5	6	7
0	CP	NP	0	0	0	0	NP	CP
1	NP	NP	0	0	0	0	NP	NP
2	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0
6	NP	NP	0	0	0	0	NP	NP
7	CP	NP	0	0	0	0	NP	CP

図1. 盤位置の評価テーブル
 Fig.1 Points of Boad Position

ここで BP_table は図 1 の座標 (i,j) における重み、K[i,j]は(i,j)に(白駒(○)、空(・)、黒駒(●))があるかにより(1、0、-1)を与える。(白の立場で評価している)

正確には、隅に隣接する位置(NP)の重みには補正を加えるがここでは省略する。

- ② 中心性 (CT) : 序盤は味方の駒が内側に入り、敵の駒が外回りにいくようにすると良いと言われているので取り入れる。中心性の評価値は、白(○)、黒(●)から駒の重心への距離の2乗の総和の差で計算させる。

$$CT = \sum_{i=0}^7 \sum_{j=0}^7 K[i,j] * ((i-Gx)^2 + (j-Gy)^2)$$

(ここで Gx=3.5, Gy=3.5)

- ③ 直線性 (ST) : 味方の駒がなるべく直線に並んでいる方が良いと言われているもので、長さ3以上の直線駒について次のように計算する。

$$ST = (K[i,j] * (\text{長さ3以上の直線駒の長さ})^2)$$

の総和 (0 ≤ i,j ≤ 7)

- ④ 駒数 (KN) : 序盤は味方の駒数を少なくし、最後になってから多くを反転させた方が良いといわれている。KNの効果は手数(t)に従属すると考えられる。

$$KN = \sum_{i=0}^7 \sum_{j=0}^7 K[i,j]$$

⑤ 候補数 (CN) : ある局面で次に打てる可能性のある手の数を表す。敵が選択できる候補数はなるべく少なく、逆に味方の候補数を多くすればいいと言われている。

3. 評価関数

局面を評価する評価関数 F は、2. で述べた5つのパラメータを変数とする関数

$$F = F(BP, CT, ST, KN, CN)$$

で表すことになるが、その関数形をいかにすべきかが課題であるが、ここでは F を各パラメータの1次式であると仮定する。

$$F = W_{BP} * BP + W_{CT} * CT + W_{ST} * ST + W_{KN} * KN + W_{CN} * CN$$

ここで W_{xx} は各パラメータの重みである。

異なる2つの評価関数のどちらが強いかを判定するのに、コンピュータ同士複数回(ここでは100回戦づつ)対戦させるため、評価の値を乱数で±5%の範囲で変動させている。すなわち元の評価値を F とするとき、変動後の値 H を $H = F * (1 + 0.1 * (r - 0.5))$ とした。r は 0 から 1 までの一様乱数である。

4. コンピュータによる対戦実験

本研究では先験的な5つのパラメータの重みを変動させながら段階的にパラメータおよび評価関数の値を改善させることを意図している。

4-1 対戦実験の前提条件

今、2つの評価関数を A と B とするとき、A、B をそれぞれ何回か先攻(A:B)、

表1 BP(盤位置)の強さの確認
Table1: BP: Strongest Parameter

対戦	勝	引分	負	駒総数	結果	
BP:CT	96	0	4	4485	1919	BP>>CT
CT:BP	5	1	44	1574	4879	
BP:ST	92	3	5	4376	2023	BP>>ST
ST:BP	3	0	79	1627	4773	
BP:KN	92	4	4	4396	2004	BP>>CN
KN:BP	4	1	95	1647	4752	
BP:CN	92	0	8	5389	2009	BP>>CN
CN:BP	15	0	85	2156	4241	

後攻(B:A)で対戦させ、先攻・後攻の両方でAの勝数が多いとき「A>B」または「B<A」と記述する。勝負が明確でない時「A?B」で表す。

対戦の条件は次の通りである。

- ①: 対戦回数を100回戦とする。
- ②: 先読み数(p)は3とする。
- ③: 風潰し先読み数(s)は9とする。

これらの条件は、一回の対戦が5分程度で終わるようにということで選択した。

4-2 パラメータ効果の確認

(1) BPの強さの確認

オセロの実戦体験からも、隅を取るかどうかは勝負に大きく影響する。すなわち、盤位置(BP)の効果が一番高と仮定し、BPと他のパラメータとの対戦を行なった(表1)。表1で勝・負・引分の数は先手側から見たものである。予想どおり、BPが絶対的な強さを示している。(表1では参考のため100回戦の駒の総数をつけて

表2. BP以外のパラメータの効果の測定
Table2: Test of effect of other parameters

対戦	勝	引分	負	駒総数	結果	
φ:φ	52	9	39	3362	3025	
CT:φ	41	0	59	3054	3330	CT?φ
φ:CT	51	0	49	3137	3220	
ST:φ	46	6	48	3062	3335	ST?φ
φ:ST	46	4	50	2885	3509	
KN:φ	54	3	43	3250	3149	KN?φ
φ:KN	82	4	14	3800	2598	
φ:CN	63	2	35	3588	2807	CN>>φ
φ:CN	16	4	80	1951	4442	

注:先攻、後攻ともBPは効かしている

いる。駒の総数が6400にならないのは60手以前に勝負のついた対戦があるからである。)

(2) 他のパラメータの効果のチェック

BPの効果は明確になったが、他のパラメータが本当に効果があるのかを確認する必要がある。BPが絶対的に効くので、これは必ず評価関数に入るはずであるので、BPに他のパラメータを1つずつ付け、対戦させた。BPの効果は他のパラメータの値に乱されないようにその重みを

$$W_{BP} = 100$$

としている。特定のパラメータの値を大きくすることを「効果の分離」と呼ぶことにする。それぞれのパラメータを加えた対戦結果を表2に示す。BPは常に効かせているので、たとえば表で「KN:φ」とあるのは「BP+KN:BP」のことである。

表2では、候補数(CN)の効果は認められるものの、他のパラメータは効果ははっきりしない。

5. パラメータの有効成分の抽出

4章で示したように、中心性(CT)、直線性(ST)、駒数(KN)の効果ははっきりしない。これらのパラメータは手数(t)に依存するもとであると仮定し、何手目で効かせれば効果があるかを調べてみた。

表3 KN: 区間(10)の負効果の測定
Table3: Interbal check of KN effect

区間	対戦	先攻		後攻		結果		
		勝	負	勝	負			
1-10	φ:φ	52	3	39	52	3	39	OX:?
		27	1	72	44	6	50	XO:?
11-20	KN:φ	52	3	45	32	2	62	OO:薬
21-30		86	2	12	65	2	33	OX:?
31-40		47	9	44	57	2	41	OX:?
41-50		47	8	45	61	6	33	?X:?
51-60		60	6	34	42	9	49	OO:薬

注:先攻、後攻ともBPは効かしている

表4 KNの区間(5)の負効果の測定

Table4: Interbal check of KN effect

区間	対戦	先攻			後攻			結果
		勝	分	負	勝	分	負	
	φ:φ	52	3	39	52	3	39	○X
0-5	KN[-]:φ	17	3	80	85	0	15	××:毒
6-10		70	4	26	71	2	27	◎X:?
11-15		45	3	52	26	2	72	?◎:?
16-20		61	6	33	68	5	27	◎X:?
21-25		79	2	19	67	5	28	◎X:?
26-30		61	3	36	69	4	27	◎X:?
31-35		59	5	36	53	2	45	◎X:?
36-40		59	6	35	59	7	34	◎X:?
41-45		50	4	46	53	1	46	◎X:?
46-50		43	6	54	73	3	25	XX:毒
51-55		57	4	39	55	6	39	◎X:?
56-60		60	4	36	46	9	52	◎X:?

注:先攻、後攻ともBPIは効かしている

5-1 KNの有効成分の抽出

KNの効果を10手毎の区間で W_{KN} を負にして効かせた結果を表3に示す。先攻・後攻の両方に聞く場合を「薬」として、両方とも逆効果のとき「毒」として効くとしている。表3では区間(11-20)、(51-50)に薬として効いているように見えるがはっきりしない。そこで5毎の区間で調べてみた(表4)。「薬」部分がなくなり、こゝろどろ「毒」があるだけである。同様に、 W_{KN} を正に効かせて見たものを調べ対比したものが表5である。「薬」として効く部分はないものの何らかの効果がありそうなので、効果がありそうな区間を(重みを正として)徹底的に調べたものが表6である。実際に薬として効く区間を図2に示す。図2では負の効果のある部分も示している。面白いのは区間(6-6)と(12-12)す

表5 KNの区間(5)効果の要約

Table5: Summary of KN effect

区間	負(-)	正(+)	判断
1-5	XX	X◎	詳細調査
6-10	◎X	X◎	詳細調査
11-15	?◎	××	詳細調査
16-20	◎X	◎X	無効
21-25	◎X	◎X	無効
26-30	○×	○×	無効
31-35	○×	○X	無効
36-40	○×	○×	無効
41-45	○×	○×	無効
46-50	××	××	判断不能
51-55	○×	○×	無効
56-60	○×	○×	無効

表6 KNの区間の正効果の測定

Table6: Interbal check of KN effect

区間	対戦	先攻			後攻			結果
		勝	分	負	勝	分	負	
	φ:φ	52	3	39	52	3	39	○X
1-5	KN[:]:φ	37	2	61	22	0	78	×◎:?
1-6		37	2	61	22	0	78	×◎:?
1-7		48	4	48	21	0	71	?◎:?
2-6		37	1	62	26	1	73	×◎:?
2-7		48	8	44	21	2	77	◎◎:薬
2-8		56	0	44	53	4	43	○×:?
2-9		51	2	47	50	5	45	○×:?
3-6		37	5	58	25	2	73	×◎:?
3-7		61	4	35	24	2	74	◎◎:薬
3-8		56	0	44	53	4	43	○×:?
3-9		57	5	38	54	5	41	◎X:?
4-5		47	3	50	20	1	79	×◎:?
4-6		61	4	35	24	2	74	◎◎:薬
4-7		58	8	26	21	1	78	◎◎:薬
4-8		54	5	41	49	2	49	◎?:?
5-5		52	9	39	54	5	41	◎X:?
5-6		48	6	46	25	0	75	◎◎:薬
5-7		50	4	46	14	2	84	◎◎:薬
5-8		48	2	50	51	7	42	××:毒
6-6		57	3	40	16	0	84	◎◎:薬
6-7	57	1	42	22	5	73	◎◎:薬	
6-8	63	5	32	46	3	51	◎◎:薬	
6-9	10	2	88	50	3	47	××:毒	
7-7	60	9	31	60	7	33	◎X:?	
7-8	49	7	44	51	3	46	◎X:?	

注:先攻、後攻ともBPIは効かしている

なわち6手目と12手目だけにKNを効かせても効果があることが分かった。結果としてKNで一番効果のありそうな部分は

$$KN : [+](4-7) + [-](12-13)$$

であることが分かった。

5-2 他のパラメータの有効成分の抽出

他のパラメータの有効範囲をKNと同様に調べその有効範囲は

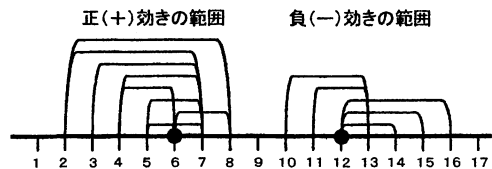


図. 2 KNの効きの範囲

Fig. 2: Extent of effective KN

- ① BP : [+](1-60)
- ② CT : [+](6-6)+ [+](25-25)
- ③ ST : [+](4-7)+ [-](10-10)+ [+](12-15)
- ④ KN : [+](4-7)+ [-](12-13)
- ⑤ CN : [+](1-60)

であることが分かった。

6. 評価関数の作成

(パラメータの組合せ調査)

5章の実験で、個別のパラメータが、区間を上手く選択すれば効果があることが分かったが、これらのパラメータを実際に合成させて評価関数を作ることを考える。

6-1 各パラメータの効果の再確認

表7に5章で抽出したパラメータの強さを再確認する。BPが絶対的に強いこと、BPに各パラメータを加えた時有効であることが明らかになった。BP以外のパラメータの強さを確認すると表8のようになり、BP、CNの強さははっきりしているがCT、ST、KNは三三すくみのような関係で強さがはっきりしない。

$$BP \gg CN \gg (CT, ST, KN)$$

6-2 CT, ST, KNの調査

CT, ST, KNを単純に加えても効果ははっきりしないので、5-2節で示した

表7 BPの強さの再確認

Table7: Strength of parameter BP

対戦	先攻			後攻			結果
	勝	分	負	勝	分	負	
BP : CT	96	0	4	5	1	94	◎◎:BP>>CT
BP : CT	92	3	5	3	0	97	◎◎:BP>>ST
BP : KN	92	4	4	4	1	95	◎◎:BP>>KN
BP : CN	92	0	8	15	0	85	◎◎:BP>>CN
BP:BP	50	5	45	50	5	45	○×
BP+CP:BP	100	0	0	23	1	78	○○:CT有効
BP+ST:BP	69	1	30	25	0	75	○○:ST有効
BP+KN:BP	62	0	38	33	1	66	○○:KN有効
BP+CN:CN	64	1	35	12	3	80	○○:CN有効

表8 BP以外のパラメータの強さの再確認
Table8: Strength of parameters except BP

対戦	先攻			後攻			結果
	勝	分	負	勝	分	負	
CT:ST	24	2	74	56	6	39	X×:CT<ST
CT:KN	50	5	45	27	1	72	◎◎:CT>KN
CT:CN	14	0	86	75	2	23	XX:CT<<CN
ST:KN	28	0	72	46	4	50	X×:CT<ST
ST:CN	11	0	89	71	2	27	XX:ST<<CN
KN:CN	22	4	74	65	4	31	XX:KN<<CN

注:先攻、後攻ともBPは効かしている

CT, ST, KN (CSKと略記)の有効な範囲の共通部分を中心にして実験し、結局6手目、25手目に効かせた場合が一番効果があることが分かった。本研究の課題としての評価関数は

$$(BP + CN) (1-60) + (CT + ST + KN) ((6-6)+(25-25))$$

を一応の結論とする。

7. 考察および今後の課題

コンピュータ同士の対戦で、はじめに仮定したパラメータが強さに関して何らかの効果を持つことが確認でき、また、パラメータを段階的に変更することにより徐々に強くなった。

一手だけに効かせても効果があるのは分かったことが収穫と言える。ただし、何故そうなるかのその因果関係は明確ではない。ゲームのプログラムのように複雑な問題では結果としての現象面だけをとらえられても、その理論的な裏付け(因果関係)までは分からないかも知れない。

本研究で行なった実験では、次のような多くの仮定を前提としていており、まだきちんと検証しているわけではない。以下のことは今後の課題とする。

- ① 本当に100回戦の勝敗で評価関数の強さの評価が出来るのか、統計的な検定で確認する必要がある

- ② パラメータは独立であるか（特にCT, ST, KNは相互に従属しているように見える）
- ③ 他に評価関数のパラメータとなり得るものがないか
- ④ BPの評価値として図1で示したものでいいのか、これも実験で最適解あるいは改善解が見つかるかもしれない
- ⑤ 評価関数を各パラメータの1次式で近似しているがこれでよいのか
- ⑥ コンピュータ同士での対戦で強くなったとしても、本当に（人間と対戦して）強くなったのか。
- ⑦ 4-2(2)で特定のパラメータの重みを大きくする「効果の分離」という方法を取ったが、これが妥当であるかは明らかではない。

8. おわりに

本研究では、実験報告という域をでてい

ない。しかし、複雑なシステムのパラメータを改善する（理想的には最適化する）ため、コンピュータにより繰り返し実験すると言う手法がゲームのプログラム作りにも有効ではないかとの示唆は得られたと思う。

本研究では、評価関数の改善に努め、プログラムを強くするための実戦結果からの定石を適用すること、先読みを多くするための高速化については注力していない。

これらをあわせ本当に強いプログラムを作成するための研究を継続したい。

参考文献・資料

- [1] 坂口実：“ゲームの理論”、森北出版、1969.
- [2] M. Mitchell Waldrop: “Complexity”、Penguin Science、1992.
- [3] Seal Software：リバーシのアルゴリズム、工学社、2003年.
- [4] 長谷川五郎：リバーシの勝ち方、河出書房新社、2003年.