

期待値を用いたゲーム木探索

小林 康幸[†] 坂口 敏洋^{††} 傍島 教文^{†††}
難波 伸也[†] 黒岩 大史[†]

ミニマックス法は、チェス、オセロ、将棋、囲碁などの零和2人ゲームにおける最も基本的なゲーム木探索法である。ミニマックス法では、葉ノードに評価値を与え、子ノードの値の最大値または最小値として親ノードの返り値を求めている。葉ノードに与えられた評価値がゲームの結果を正確に反映している場合には、このミニマックス法による探索は最善の手を選択するが、葉ノードに与えられた評価値がゲームの結果を正確に反映していない場合には、ミニマックス法の返り値を求める方法は、最善の方法とは限らない。ミニマックス法の返り値を求める方法に対する様々な改善案が提案されている。本論文では、我々は、探索した葉ノードからさらに2手先を予測することができる期待値法を提案する。また、オセロゲームによる実験によって、期待値法がミニマックス法に対し改良されていることを示す。

A Game Tree Search Using Expectation

YASUYUKI KOBAYASHI,[†] TOSHIHIRO SAKAGUCHI,^{††}
NORIFUMI SOBASHIMA,^{†††} SHINYA NANBA[†] and DAISHI KUROIWA[†]

Minimax is the most basic game-tree search method for zero-sum two-person games such as Chess, Othello, Shogi, and Go. Minimax uses evaluation values at leaf nodes and calculates returned values by taking the maximum or minimum of the successors. If evaluation values given at leaf nodes exactly reflect the result of the game, minimax selects the best move. However, if evaluation values given at leaf nodes do not properly reflect the result of the game, minimax may not be the best approach. Various improved methods for minimax have been proposed. This paper proposes the so-called Expectation method which at the level of the leaf nodes takes some probability information of two additional ply into consideration for every position. Experiments within the domain of Othello show that the Expectation method is an improvement of minimax.

1. はじめに

2人零和完全情報確定ゲームでのゲーム木探索法として、ミニマックス法およびその効率的探索法としての α - β 法^{1),2)}が用いられてきた。

ミニマックス法では、探索したゲーム木の葉ノードに、あらかじめ決めておいた静的評価関数によって評価値を与え、先手(またはmax)ノードには子ノードの値の最大値を、後手(またはmin)ノードでは、子ノードの値の最小値を親ノードの返り値としている。

ミニマックス法で用いられる静的評価関数の項目に、

兄弟ノードの評価値やさらに先を探索した情報に加えようとする試みがある。他の兄弟ノードより極端に良い評価値を得るノードについてはさらに先読みをする、非凡拡張という方法が提案されている³⁾。また、コンピュータチェスにおいて、一方が王手にさらされているとき、相手のミスの可能性を期待するためのボーナス/マルス関数の中に兄弟ノードの情報を取り入れる試み⁴⁾や、探索する深さまでの各深さにおける評価値を葉ノードの静的評価関数に取り入れる試み⁵⁾もある。

一方、ミニマックス法とは異なる方法で返り値を求める提案もされている。ミニマックス法では、子ノードに与えられた評価値または返り値の最大値または最小値を親ノードの返り値とするのに対して、先手ノードには、子ノードの値の中で最も大きい方から M 個の値の、後手ノードには、子ノードの値の中で小さい方から N 個の値の関数によって返り値を得る方式が提案されている⁶⁾。また、すべての子ノードの値によ

[†] 島根大学

Shimane University

^{††} 三菱電気メカトロニクスソフトウェア株式会社

MITSUBISHI ELECTRIC MECHATRONICS SOFTWARE CORPORATION

^{†††} 株式会社 FFC システムズ

FFC SYSTEMS LIMITED

る一般化平均を用いて返り値を得る方式⁷⁾、葉ノードに先手の勝つ確率を与え、積によって返り値を求める積伝搬^{8),9)}、積と最大値/最小値の平均を得る平均伝搬^{10),11)}なども提案されている。

本論文では、葉ノードからその親ノードへの返り値として、葉ノードからさらに深く探索したときの返り値の期待値を求めることを提案する。期待値は、すべての子ノードの値による関数である。一般化平均⁷⁾は、子ノードの値の最大値から最小値までの値を得るが、期待値は先手ノードでは、子ノードの値の最大値より大きい値を、後手ノードでは、子ノードの値の最小値より小さい値を得る。

まず、2章で期待値を用いた探索である、期待値法を提案する。3章ではオセロゲームを用いて、期待値法がミニマックス法よりどの程度改良できるかを示す。最後に、4章で今後の課題について述べる。

2. 期待値法

2.1 期待値法のアルゴリズム

期待値法では、以下のことを仮定する。

探索したゲーム木の葉ノードの評価値を h とするとき、その葉ノードからさらに深さ 2 の探索をしたときの返り値 g は、平均 h 、分散 σ^2 の正規分布 $N(h, \sigma^2)$ に従う。

この仮定により、期待値法では、探索するゲーム木の深さからさらに 2 手先まで探索したときの返り値を予測することができる。以下に、例を用いて期待値法によるゲーム木探索のアルゴリズムを示す。

図 1 のように、局面 R から探索する深さ 3 までのすべての可能な指し手と局面を表したゲーム木を作り、各葉ノードに評価値を与える。

図 2 において、局面 R から先手が局面 A を、さらに後手が局面 C を選んだとする。先手は局面 C ではさらに深さ 3、すなわち葉ノードからさらに 2 手先まで探索したときの返り値を予測する。

ノード C_i ($i = 1, 2$) の評価値を h_i 、ノード C_i をルートノードとする深さ 2 のゲーム木 (ゲーム木 A またはゲーム木 B) のミニマックス法による返り値を g_i とする。仮定により、 g_i は正規分布 $N(h_i, \sigma^2)$ に従う。すなわち g_i は

$$f(x|h_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-h_i)^2}{2\sigma^2}\right) \quad (i=1,2)$$

という確率密度関数を持つ確率変数である。

このとき、ノード C から深さ 3 のミニマックス法で探索したときの返り値の期待値 $E(\sigma^2, h_1, h_2)$ は次のようにして求まる。

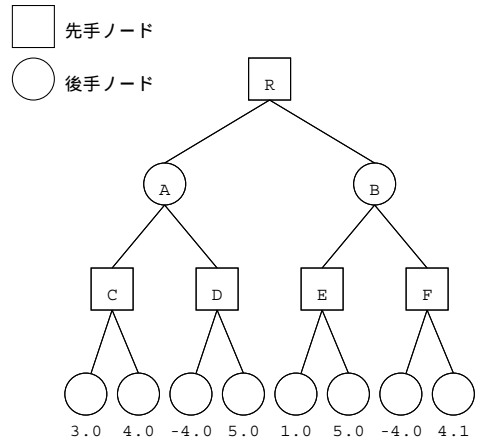


図 1 深さ 3 のゲーム木
Fig.1 A game tree of depth 3.

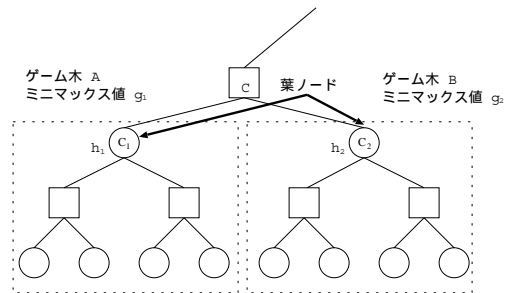


図 2 葉ノードが 2 つの場合の期待値
Fig.2 Expectations in the case of two leaf nodes.

$$E(\sigma^2, h_1, h_2) = g_1 \times (g_1 \geq g_2 \text{の確率}) + g_2 \times (g_2 > g_1 \text{の確率})$$

つまり、

$$E(\sigma^2, h_1, h_2) = \int \int_{x \geq y} x f(x|h_1) f(y|h_2) dx dy + \int \int_{x < y} y f(x|h_1) f(y|h_2) dx dy$$

図 3 において、ノード C の子ノードの評価値 3.0, 4.0 を用いて、ノード C でさらに深さ 3 のミニマックス探索をしたときの返り値の期待値 $E(\sigma^2, 3.0, 4.0)$ を求める (図 3 では $\sigma = 8.0$ で計算している)。同様に、ノード D, E, F の期待値も求める。

図 4 において、ノード C, D, E, F よりさらに上のノードに対しては、ミニマックス法と同様の方法で返り値を求め、局面 R における先手の選択手を決定する。

一般に P をノード、P の子ノードの数を k 、 P_1, P_2, \dots, P_k を P の子ノードとする。このとき、期待値法による各ノード P の返り値 $K(P)$ は、次の式で表される。

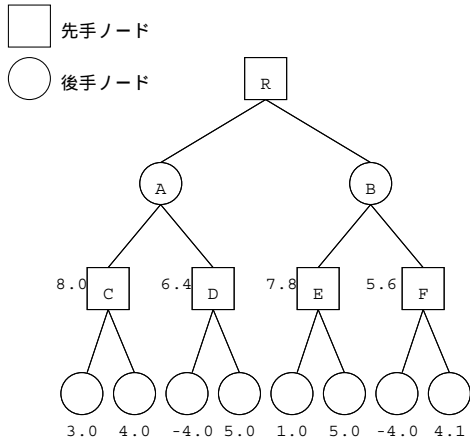


図3 期待値の計算
Fig. 3 Calculation of expectations.

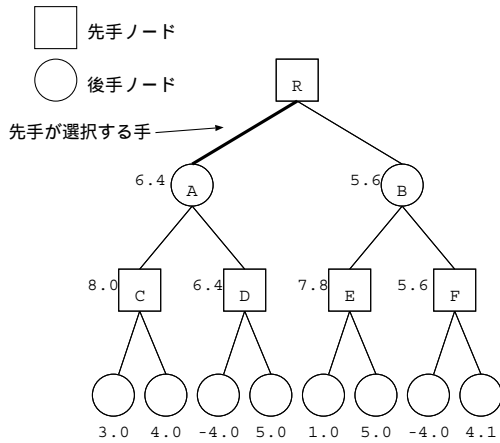


図4 期待値を用いたミニマックス法
Fig. 4 Minimax method using expectations.

$$K(P) = \begin{cases} \max_{1 \leq i \leq k} K(P_i) & (P: \text{先手ノード}) \\ \min_{1 \leq i \leq k} K(P_i) & (P: \text{後手ノード}) \\ E(\sigma^2, h_1, h_2, \dots, h_k) & (P: \text{葉ノードの親ノード}) \end{cases}$$

ただし、 h_i はノード P_i の評価値を表し、 $E(\sigma^2, h_1, h_2, \dots, h_k)$ は、葉ノードの親ノードが先手ノードのときは、

$$E(\sigma^2, h_1, h_2, \dots, h_k) = \iiint \dots \int_{x_1=\max(x_1, \dots, x_k)} x_1 f(x_1|h_1) f(x_2|h_2) \dots f(x_k|h_k) dx_1 dx_2 \dots dx_k + \iiint \dots \int_{x_2=\max(x_1, \dots, x_k)} x_2 f(x_1|h_1) f(x_2|h_2) \dots f(x_k|h_k) dx_1 dx_2 \dots dx_k + \dots$$

表1 期待値の近似値の平均と標準偏差

Table 1 Means and standard deviations of approximations for expectation.

k		m=100	m=10,000	m=1,000,000
10	平均	12.32098	12.32588	12.31491
	標準偏差	0.45793	0.07168	0.01735
100	平均	20.05479	20.07026	20.07223
	標準偏差	0.33935	0.11895	0.01154

$$+ \iiint \dots \int_{x_k=\max(x_1, \dots, x_k)} x_k f(x_1|h_1) f(x_2|h_2) \dots f(x_k|h_k) dx_1 dx_2 \dots dx_k$$

であり、葉ノードの親ノードが後手ノードのときは、 $E(\sigma^2, h_1, h_2, \dots, h_k)$

$$= \iiint \dots \int_{x_1=\min(x_1, \dots, x_k)} x_1 f(x_1|h_1) f(x_2|h_2) \dots f(x_k|h_k) dx_1 dx_2 \dots dx_k + \iiint \dots \int_{x_2=\min(x_1, \dots, x_k)} x_2 f(x_1|h_1) f(x_2|h_2) \dots f(x_k|h_k) dx_1 dx_2 \dots dx_k + \dots + \iiint \dots \int_{x_k=\min(x_1, \dots, x_k)} x_k f(x_1|h_1) f(x_2|h_2) \dots f(x_k|h_k) dx_1 dx_2 \dots dx_k$$

である。

期待値は、モンテカルロ法によって求めることができる。

葉ノードの評価値をそれぞれ、 h_1, h_2, \dots, h_k とする。葉ノードの評価値が h_i のとき、葉ノードからさらに深さ 2 のミニマックス法で探索したときの返り値を F_i とすると、仮定より F_i は、正規分布 $N(h_i, \sigma^2)$ に従う。したがって、葉ノードの親ノードの期待値 $E(\sigma^2, h_1, h_2, \dots, h_k)$ は、 $P = \max(F_1, F_2, \dots, F_k)$ または $P = \min(F_1, F_2, \dots, F_k)$ の平均 $E(P)$ である。

k 個の正規乱数の組 $(F_{1j}, F_{2j}, \dots, F_{kj})$ (F_{ij} は $N(h_i, \sigma^2)$ に従う) を m 組 ($1 \leq j \leq m$) 発生させ、

$$P_j = \begin{cases} \max(F_{1j}, F_{2j}, \dots, F_{kj}) & (1 \leq j \leq m) \\ \text{(親ノードが先手ノードの場合)} \\ \min(F_{1j}, F_{2j}, \dots, F_{kj}) & (1 \leq j \leq m) \\ \text{(親ノードが後手ノードの場合)} \end{cases}$$

$$V_m = \sum_{j=1}^m \frac{P_j}{m}$$

とおくと、 m が十分大きいときには $E(\sigma^2, h_1, h_2, \dots, h_k)$ の近似値として V_m を用いることができ、こ

表 2 $\mathbf{h} = (10, 21, 29, 32)$ に対する $M_p(\mathbf{h})$
 Table 2 $M_p(\mathbf{h})$ for $\mathbf{h} = (10, 21, 29, 32)$.

p	-32	-16	-8	-4	-2	-1	0	1	2	4	8	16	32
$M_p(\mathbf{h})$	10.4	10.9	11.9	13.9	16.6	18.7	21.0	23.0	24.5	26.5	28.3	29.7	30.7

の計算で発生する誤差は $O(\frac{1}{\sqrt{m}})$ となる. 表 1 に乱数を m ($m=100, 10,100, 1,000,000$) 組発生させ, $E(\sigma^2, h_1, h_2, \dots, h_k)(h_1 = h_2 = \dots = h_k = 0, k = 10, 100, \sigma = 8)$ の近似値を求めた場合, その近似値がどの程度ばらつくかを 100 回の試行によって各近似値の平均と標準偏差を示す.

2.2 期待値の性質

我々の提案する期待値法は, 異った深さのミニマックス探索による返り値の間に強い相関があるという事実を利用したゲーム木探索法である.

Buro は, オセロゲームのデータを用いた線形回帰分析により, 深さ 4 の探索による返り値 v' から深さ 8 の探索による返り値 v を推測している¹²⁾. その結果 v は, 正規分布 $N(w, \sigma^2)$ に従っており, 28 手目終了時および 44 手目終了時の局面で, それぞれ ($\sigma = 0.542, w = 1.036v' - 0.009$), ($\sigma = 0.884, w = 0.956v' - 0.067$) の値を得ている. すなわち, w の予測式における v' の係数が 1 の近く定数項が 0 に近いため, v はおおよそ正規分布 $N(v', \sigma^2)$ に従っている.

期待値法では葉ノードの評価値が h のとき, 葉ノードから 2 手先まで探索したときの評価値は $N(h, \sigma^2)$ の正規分布に従うことを仮定する. σ^2 は過去のゲームの統計によって得られた値とする. 期待値法はこの仮定より, 葉ノードのさらに 2 手先まで探索したときの返り値の期待値を求めることができる.

図 5 は, ミニマックス法と期待値法によって選択手が異なる例を示す.

ノード A~F における下段の値が, 葉ノードからのミニマックス法による返り値, 上段の値が期待値法による返り値を表す. 図 5 ではミニマックス法で手を選択した場合, 選択されるのはミニマックス法による返り値が大きいノード B である. これに対し, 期待値法で手を選択した場合, 選択されるのは期待値法による返り値が大きいノード A である.

ミニマックス法と期待値法の違いは, ミニマックス法が葉ノードにおける評価値(ノード下段の値)を基準に手を選択しているのに対し, 期待値法では探索したゲーム木の深さのさらに 2 手先まで探索したときに得られる返り値の期待値(ノード上段の値)を基準に手を選択している点である.

Rivest は, 子ノードの評価値を $\mathbf{h} = (h_1, \dots, h_k)$

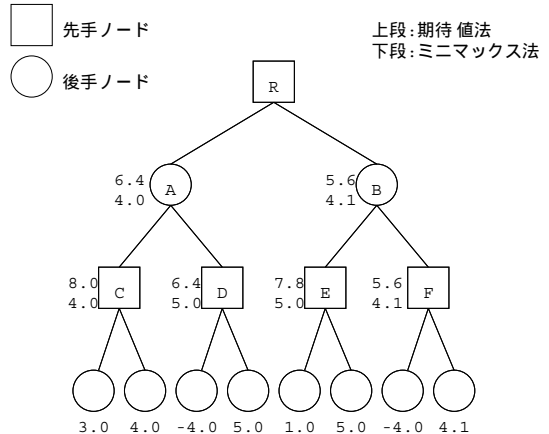


図 5 期待値法とミニマックス法 ($\sigma = 8$)

Fig. 5 Expectation method and minimax method ($\sigma = 8$).

表 3 $E(\sigma^2, 10, 21, 29, 32)$

Table 3 $E(\sigma^2, 10, 21, 29, 32)$.

σ	1	2	4	8
先手ノード	32.01	32.24	33.09	35.58
後手ノード	9.98	9.97	9.88	8.78

(ただし, $h_i > 0(1 \leq i \leq k)$) としたとき, それらの親ノードへの返り値を次の式によって一般化平均を求めることを提案している⁷⁾.

$$M_p(\mathbf{h}) = \left(\frac{1}{k} \sum_{i=1}^k h_i^p \right)^{\frac{1}{p}}$$

ただし, p は $-\infty < p < \infty$ を満たす定数である.

この一般化平均 $M_p(\mathbf{h})$ は, 以下の性質を持っている.

$$\lim_{p \rightarrow \infty} M_p(\mathbf{h}) = \max(h_1, \dots, h_k)$$

$$\lim_{p \rightarrow -\infty} M_p(\mathbf{h}) = \min(h_1, \dots, h_k)$$

$$p < q \Rightarrow M_p(\mathbf{h}) \leq M_q(\mathbf{h})$$

これらのことから分かるように, 一般化平均は, 子ノードの評価値の最大値よりも小さく, 最小値よりも大きい値を返り値としている.

これに対して期待値は, 先手ノードではノードの評価値の最大値よりも大きく, 後手ノードでは最小値よりも小さくなる.

表 2 に, $\mathbf{h} = (10, 21, 29, 32)$ に対する $M_p(\mathbf{h})$ を示す(文献 7)より). また, 表 3 に, 先手ノードと後手

表 4 先手ノードの $E(8^2, 20, h_2, h_3)$
Table 4 $E(8^2, 20, h_2, h_3)$ for the max node.

$h_2 \setminus h_3$	-20	-10	0	10	20
-20	20.02	-	-	-	-
-10	20.10	20.20	-	-	-
0	20.17	20.25	20.44	-	-
10	21.18	21.20	21.26	21.98	-
20	24.46	24.52	24.59	24.88	26.70

ノードにおける期待値 $E(\sigma^2, 10, 21, 29, 32)$ を示す．次に，表 4 に子ノードが 3 個で，それらの評価値が $(20, h_2, h_3) (20 \geq h_2 \geq h_3)$ の場合の，先手ノードの期待値 $E(8^2, 20, h_2, h_3)$ を示す．これらのいずれの場合もミニマックス法での返り値は 20 となるが，期待値は 20 より大きく，また， h_2 および h_3 が 20 に近いほど大きくなる事が分かる．

3. オセロゲームによる実験

期待値法がミニマックス法よりどの程度改良できるかをオセロゲームの実験で調べる．静的評価関数は，黒石の数-白石の数，とする．まず，先手が期待値法とミニマックス法でそれぞれ 1 手打つ場合に期待値法がミニマックス法よりどの程度改良されるかを比較する．次に，先手が期待値法とミニマックス法でそれぞれ T 手 ($T = 2, 3, 5, 7$) 打つ場合に期待値法がミニマックス法よりどの程度改良されるかを比較する．

3.1 分散の推定

期待値法では，探索したゲーム木の葉ノードの評価値を h とするとき，その葉ノードからさらに深さ 2 の探索をしたときの返り値 g は，平均 h ，分散 σ^2 の正規分布 $N(h, \sigma^2)$ に従うことを仮定した．ここでは，オセロゲームでこの仮定が成り立つことを実験で確かめるとともに， σ^2 を推定する．実験は次のように行った．

- (1) オセロゲームの過去の対戦棋譜 100 局面の S 手 ($50 \leq S \leq 54$) 打ち終わった各局面に対し，深さ 3 のゲーム木を作成し，そのゲーム木の各葉ノードに対して評価値 $h_i (1 \leq i \leq N, N$ は葉ノードの数) を求める．
- (2) 各葉ノードをルートノードとして，深さ 2 のミニマックス法で探索し，返り値 g_i を求める．
- (3) $h_i - g_i$ のヒストグラムを作成するとともに，分散の推定値 $\hat{\sigma}^2 = \frac{\sum_{i=1}^N (h_i - g_i)^2}{N}$ および標準偏差の推定値 $\hat{\sigma}$ を求める．

図 6 は， $S = 54$ での $h_i - g_i$ のヒストグラムである．おおよそ正規分布 $N(0, 8.33^2)$ に従っていることが分かる．分散と標準偏差の推定値を表 5 に示す．

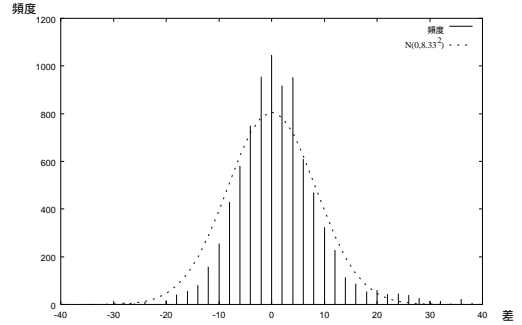


図 6 評価値と 2 手先からの返り値の差の分布
Fig. 6 Distribution of the difference between evaluations and back-up values from the depth 2.

表 5 2 手先の分散と標準偏差

Table 5 Variances and standard deviations of depth 2.

	分散	標準偏差
53 手から 55 手	51.98	7.21
54 手から 56 手	69.39	8.33
55 手から 57 手	90.30	9.50
56 手から 58 手	121.84	11.04
57 手から 59 手	183.95	13.56

なお，他のオセロゲームの対戦棋譜 100 局面に対しても，53 手から 55 手の標準偏差の推定値は，7.23 および 6.99 が得られた．したがって 2 手先の分散と標準偏差はオセロゲームの進行手数によって変わるが，進行手数が同じなら局面によっては大きく変動しないと考えられる．

3.2 1 手での実験

先手が期待値法とミニマックス法でそれぞれ 1 手打つ場合に期待値法がミニマックス法よりどの程度改良されるかを実験で比較する．

以下に実験の方法を示す．

- (1) オセロゲームでの過去の対戦棋譜 10,000 局面 (3.1 節で用いた棋譜とは異なる) に対して， $S (S = 50, 51, 52, 53, 54)$ 手終了の各局面に対して，次の手を先手番とし，ミニマックス法と期待値法のそれぞれの探索法で深さ $D (D = 3)$ まで探索し， $S + 1$ 手目の選択手を決める．期待値の計算では，3.1 節で求めたオセロの進行手数ごとの分散の推定値 (たとえば， $S = 50$ では 53 手から 55 手の分散の推定値) を用いた．また，乱数の発生回数を 100 とした．
- (2) 2 つの探索法を以下の 2 項目について比較する．
 - (a) ミニマックス法で深さ $D + 2$ まで探索し，(1) で決定されたミニマックス法と期待値法の選択手への返り値の大小で探索法

表6 2手先での評価(1手)
Table 6 Evaluation by depth 2 (one ply).

S	50	51	52	53	54
K M ×	1086	1095	1148	1292	1196
K × M	786	715	742	795	746
評価指数1	3.00	3.80	4.06	4.97	4.50
評価指数2	58.0	60.1	60.7	61.9	61.6
差の平均	0.266	0.432	0.500	0.735*	0.626

*は有意水準5%のt-検定で有意

表7 最終値での評価(1手)
Table 7 Evaluation by the result of the game (one ply).

S	50	51	52	53	54
K M ×	1139	1212	1203	1219	1151
K × M	988	829	819	735	815
評価指数1	1.51	3.83	3.84	4.84	3.36
評価指数2	53.5	59.4	59.5	62.4	58.5
差の平均	0.294	0.659	0.624	0.737*	0.443

*は有意水準5%のt-検定で有意

の優劣を比較する。期待値法の選択手への
の戻り値が大きいときは $K M \times$ とし、
その逆のときは $K \times M$ とする。また、
戻り値の差(期待値法-ミニマックス法)
の平均を求める。

- (b) ミニマックス法でゲーム終了まで探索し、
上と同様の方法で比較する。

すなわち、(a) では期待値法がどの程度2手先
まで予測して選択手を決めているか、また (b)
では期待値法の選択手がゲームを勝利に導くた
めにどの程度貢献しているかを調べる。

実験結果を表6および表7に示す。評価指数1お
よび評価指数2は、以下の式で与える。

$$\text{評価指数1} = \frac{K M \times - K \times M}{10,000} \times 100$$

$$\text{評価指数2} = \frac{K M \times}{K M \times + K \times M} \times 100$$

評価指数1は、実験局面数10,000に対して期待値法
の選択手がミニマックス法の選択手より良い手であっ
た局面数の割合を、評価指数2は、期待値法とミニ
マックス法の選択手が異なった局面に対する期待値法
の勝率を示す。

1手の実験結果から次の点が確認できる。

- (1) いずれの場合でも $K M \times$ の数が $K \times M$
の数を上回り、評価指数1では、期待値法がミニ
マックス法より3.00%から4.97%改善され、
評価指数2では、期待値法の勝率が58.0%から
61.9%になっている。
- (2) 平均値の差の有効性を示すためのt-検定では、
 $S=53$ の2手先および最終値で、2群の平均値

表8 乱数の発生組数による比較

Table 8 Comparison by generated instances for the
Monte Carlo method.

m	10	30	50	100	150	200
K M ×	1340	1239	1217	1196	1161	1153
K × M	1194	930	815	746	703	703
評価指数1	1.46	3.09	4.02	4.50	4.58	4.50
評価指数2	52.9	57.1	59.9	61.6	62.3	62.1
差の平均	0.262	0.548	0.609	0.626	0.635	0.639

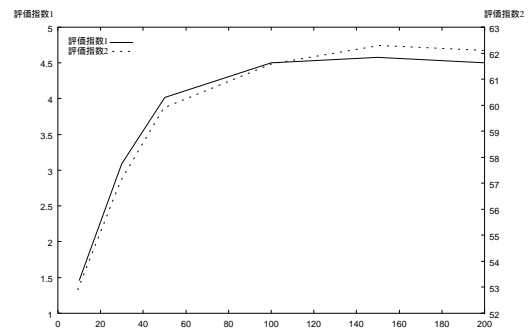


図7 乱数の発生組数による比較

Fig. 7 Comparison by generated instances for the Monte
Carlo method.

には有意水準5%で差がある」ことが示された。

なお、 $S = 54$ について、期待値を求めるための乱
数の発生組数 m を変化させた場合の実験結果を表8
および図7に示す。評価指数1および評価指数2は、
 $m \geq 100$ ではあまり増加しない。

3.3 複数手の実験

先手が期待値法とミニマックス法でそれぞれ T 手
($T = 2, 3, 5, 7$) 打つ場合に期待値法がミニマックス
法よりどの程度改良されるかを実験で比較する。以下
に実験の方法を示す。

- (1) 1手での実験と同じ対戦棋譜10,000局面に対
して、後手をミニマックス法での探索に固定
し、先手を期待値法で探索した場合とミニマッ
クス法で探索した場合を比較する。探索の深さ
は D ($D = 3$) とする。先手、後手、と交互
に選択手を決めてゲームを進め、先手が T 手
($T = 2, 3, 5, 7$) 探索し、 T 手目に選択した(期
待値法とミニマックス法による)手を決定する。
 T 手目の探索局面をオセロゲームの52手終了の
局面に揃えるために、 $T = 2, 3, 5, 7$ のそれぞ
れに対し、対戦の開始を、 S ($S = 50, 48, 44, 40$)
手終了の局面とした。ここでも、期待値を求め
るための乱数の発生回数は100とした。なお、
ここでは標準偏差の推定値を8.5に固定した。
- (2) 2つの探索法での T 手目の選択手を以下の2項

表 9 2手先での評価(複数手)
Table 9 Evaluation by depth 2 (multiple ply).

T	2	3	5	7
K M ×	2128	3150	4445	4937
K × M	1514	2153	3255	3934
評価指数 1	6.14	9.97	11.90	10.03
評価指数 2	58.4	59.4	57.7	55.7
差の平均	0.993*	1.729**	2.717**	3.309**

* は有意水準 5% の t -検定で有意

** は有意水準 1% の t -検定で有意

表 10 最終値での評価(複数手)

Table 10 Evaluation by the result of the game (multiple ply).

T	2	3	5	7
K M ×	2197	3137	4462	4897
K × M	1599	2226	3349	3976
評価指数 1	5.98	9.11	11.13	9.21
評価指数 2	57.9	58.5	57.1	55.2
差の平均	1.198**	1.862**	2.658**	2.719**

** は有意水準 1% の t -検定で有意

目について比較する。

- (a) 期待値法とミニマックス法のそれぞれで T 手目を探索する際に与えられた局面から深さ $D + 2$ の探索をし、(1) で決定されたミニマックス法と期待値法の選択手への返り値の大小で探索法の優劣を比較する。期待値法の選択手への返り値が大きいときは $K \quad M \times$ とし、その逆のときは $K \times M$ とする。また、返り値の差(期待値法-ミニマックス法)の平均を求める。
- (b) 期待値法とミニマックス法のそれぞれで T 手目を探索する際に与えられた局面からミニマックス法でゲーム終了まで探索し、上と同様の方法で比較する。

表 9 および表 10 に示す、 T 手打つ場合の実験結果から次の点が確認できる。

- (1) 2手先および最終値において、評価指数 1, 評価指数 2, 平均値の差のいずれの評価でも期待値法がミニマックス法より優れていることが分かる。
- (2) 2手先および最終値ともに、 T が大きくなるにつれ、平均値の差は大きく開くが、評価指数 1 では、 $T = 7$ より $T = 5$ の値の方が大きくなっている。また、 T が大きくなるにつれ期待値法とミニマックス法での選択手が異なるケースが増え、評価指数 2 は減少する傾向が見られる。
- (3) 1手の実験結果と比較すると、評価指数 1

および平均値の差では、2手先および最終値ともに、複数手の方が大きくなっている。しかし、評価指数 2 は複数手の方が減少しているケースが多い。

- (4) 平均値の差の有効性を示すための t -検定では、2手先および最終値のいずれでも、「2群の平均値には有意水準 1% または 5% で差がある」ことが示された。

4. まとめ

本論文では、葉ノードからさらに2手先まで探索したときの返り値を期待値として求めることを提案した。期待値法では葉ノードからその親ノードへの返り値がミニマックス法と異なる。オセロゲームでの実験により、期待値法が、ミニマックス法より改良された探索法であることを示すことができた。

ミニマックス法は α - β 枝刈りができるが、期待値法で葉ノードの親ノードに期待値を求める場合には α - β 枝刈りができない。しかし、期待値法でも探索の深さ D に対し、深さ $D - 1$ から深さ 1 まで返り値を求める方法がミニマックス法と同様なので α - β 枝刈りができる。今後 α - β 枝刈りによるミニマックス法と期待値法の探索時間の比較をし、探索時間を同じにして期待値法の優位性を示す必要がある。

今回は、探索した深さよりさらに2手先まで探索したときの期待値を用いた期待値法について考察および実験を行った。今後 n 手先 ($n \geq 3$) まで探索したときの期待値についても検討する必要がある。また、ミニマックス法に対する多くの他の改善方法と期待値法との比較も必要である。

参考文献

- 1) Slagle, J.M. and Dixon, J.K.: Experiments with Some Programs that Search Game Trees, *J. ACM*, Vol.16, No.2, pp.189-207 (1969).
- 2) Knuth, D.E. and Moore, R.W.: An Analysis of Alpha-Beta Pruning, *Artificial Intelligence*, Vol.6, No.4, pp.293-326 (1975).
- 3) Anantharaman, T.S., Campbell, M.S. and Hsu, F.-h.: Singular Extensions: Adding Selectivity to Brute-Force Searching, *Artificial Intelligence*, Vol.43, No.1, pp.99-109 (1990).
- 4) Uiterwijk, J.W.H.M. and Herik, H.J.v.d.: Speculative Play in Computer Chess, *Proc. Advances in Computer Chess Conference 7*, University of Limberg, Maastricht, pp.79-90 (1995).
- 5) Jansen, P.: Problematic Positions and Specu-

- lative Play, *Computers, Chess, and Cognition*, pp.169–182, Springer-Verlag, New York (1990).
- 6) Slagle, J.M. and Dixon, J.K.: Experiments with the M&N Tree-Searching Program, *Comm. ACM*, Vol.13, No.3, pp.147–154 (1970).
 - 7) Rivest, R.L.: Game Tree Searching by Min/Max Approximation, *Artificial Intelligence*, Vol.34, No.1, pp.77–96 (1988).
 - 8) Pearl, J.: Heuristic Search Theory: Survey of Recent Results, *Proc. International Joint Conference on Artificial Intelligence (IJCAI-81)*, pp.24–28, Morgan Kaufmann, Los Altos, CA. (1981).
 - 9) Baum, E.B. and Smith, W.D.: A Bayesian Approach to Relevance in Game Playing, *Artificial Intelligence*, Vol.97, Nos.1-2, pp.195–242 (1997).
 - 10) Nau, D.S., Purdom, P. and Tzeng, H.C.: An Evaluation of Two Alternatives to Minimax, *Uncertainty in Artificial Intelligence*, North Holland, Amsterdam, pp.505–509 (1986).
 - 11) Nau, D.S., Purdom, P. and Tzeng, H.C.: Experiments on Alternatives to Minimax, *International Journal of Parallel Programming*, Vol.15, No.2, pp.163–183 (1986).
 - 12) Buro, M.: ProbCut: An Effective Selective Extension of Alpha-Beta Algorithm, *ICCA Journal*, Vol.18, No.2, pp.71–76 (1995).
 - 13) 河田竜夫：確率と統計，朝倉書店 (1970).

(平成 14 年 2 月 19 日受付)

(平成 14 年 9 月 5 日採録)



小林 康幸 (正会員)

1971 年大阪市立大学数学科卒業。1973 年大阪市立大学大学院理学研究科修士課程 (数学専攻) 修了。理学博士。1974 年広島大学理学部勤務。現在島根大学総合理工学部数理・情報システム学科教授。データベース，ゲーム理論，計算機統計学の研究に従事。応用統計学会，日本統計学会，日本計算機学会各会員。



坂口 敏洋

2002 年島根大学総合理工学部卒業。同学部在学中にゲーム理論の研究を行う。2002 年三菱電気メカトロニクスソフトウェア株式会社に入社，現在に至る。



傍島 教文

2002 年島根大学総合理工学部卒業。同学部在学中にゲーム理論の研究を行う。2002 年株式会社 FFC システムズに入社，現在に至る。



難波 伸也

2002 年島根大学総合理工学部卒業。同学部在学中にゲーム理論の研究を行う。現在，島根大学大学院総合理工学研究科博士前期課程数理・情報システム学専攻に在学中。



黒岩 大史

1991 年弘前大学理学部情報科学科卒業。1996 年新潟大学大学院博士課程自然科学研究科物質科学専攻修了。同年より島根大学総合理工学部数理・情報システム学科勤務。最適化理論，ゲーム理論等の研究に従事。博士 (理学)。日本数学会，日本オペレーションズ・リサーチ学会各会員。