

期待値法によるゲーム木探索

手光 隆[†] · 小出 敏弘^{††} · 小林 康幸^{††}

A Game Tree Search by Expected-Value Method

Takashi TEBIKA[†], Toshihiro KOIDE^{††}, Yasuyuki KOBAYASHI^{††}

Abstract

Minimax method, which is the most basic game tree search for zero-sum two-person game such as Chess, Othello, Shogi, and Igo, selects the best move within the limits of the search depth. But it is impossible for minimax method to deal with the rapid changes beyond the search depth. In this paper we propose expected-value method, which is an algorithm using expected values inspecting evaluation values beyond the search depth. We also show that expected-value method is more effective than minimax method by some simulations.

概 要

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は、先読みした範囲で最も良い手を選択するが、先読みした先での急激な変化には対応できない。我々は、この問題を解決するために、探索したさらに先を予測する値である期待値を用いて探索するアルゴリズムとして、期待値法を提案する。また、様々なシミュレーションによって、期待値法がミニマックス法より有効な探索法であることを示す。

1. はじめに

1997年5月11日IBM・ワトソン研究所のチームが開発したチェス専用マシン Deep Blueがチェスの世界チャンピオンである Gary Kasparov 氏に2勝1敗3分の成績で勝利した。この Deep Blue をはじめとする多くのゲームプログラムがミニマックス法というゲームの先読み戦術を用いている。

しかし、近年のゲーム・プログラミングに関する研究では、ミニマックス法とは異なる、より効果的な先読み戦術(思索的プレイ)についての研究がいくつか行われている。OM-search, 勝負手探索, またオセロに限定した方法では, Prob-Cut がそれである。ミニマックス法では先手・後手ともミニマックス法を用いることを仮定している(シングルエージェ

[†] (株)新日本海新聞社, 鳥取県
Sinnihonkaishinbunsha Co., Ltd., 137 2-tyoume Tomiyasu Tottori-shi Tottori, 680-8688 Japan

^{††} 島根大学総合理工学部, 鳥根県
Faculty of Science and Engineering, Shimane Univ., 1060 Nishikawatsu-cho, Matsue-shi, Shimane,
690-8504 Japan

ント)のに対し, 思索的プレイは先手・後手は異なる戦術を用いることを仮定している(マルチエージェント).

人間同士のゲームの対戦では, 後手が先手と異なる戦略を取ることを互いに前提にしてプレイしている. つまり, マルチエージェントの世界である. このため, 思索的プレイの研究は, 強いプログラムを作るという人工知能的な目標であるとともに, 人間の強いプレイヤーの思索方法を探求するといった認知心理学的な目標でもあるといえる.

われわれもその他の思索プレイと同様に, ミニマックス法とは異なるより効果的なゲーム木探索法である期待値法を提案する. われわれの提案する期待値法とミニマックス法が異なる最大のポイントはミニマックス法が先読みした範囲で最も良い局面を得られる手を選択するのに対して, 期待値法は探索した範囲のさらに先の手で良い局面が得られるであろう手を選択する点である.

まず, 第2章でミニマックス法といくつかの思索的プレイについて説明する. 第3章で期待値法の考え方について説明する. 第4章では仮想ゲーム木によって期待値法の有効性を示すことによって, 様々なゲームにおいて期待値法がミニマックス法より有効であることを示す. 最後に, 第5章で今後の課題について述べる.

2. ミニマックス法と思索的プレイ

本章では, ミニマックス法といくつかの思索的プレイについて説明する.

2.1 ミニマックス法

次の図によって, 現在, 最も一般的なゲーム木探索法であるミニマックス法 [1] について説明する. 図1を局面Rから3手先の局面までのすべての可能な指し手と局面をあらわしたもの(ゲーム木)とする. この図において, □は先手の手番の局面(先手ノード), ○は後手の手番の局面(後手ノード)をあらわしている. また, 線は親ノードから子ノードへの指し手をあらわしている.

つまり, 先手は局面Rから局面Aまたは局面Bへの指し手を進めることができる(図1参照).

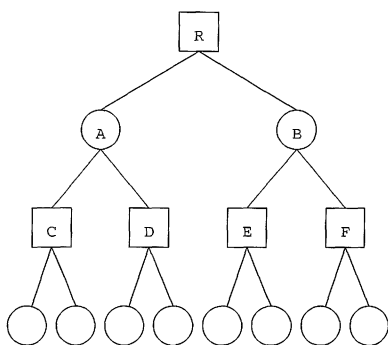


図1 深さ3のゲーム木

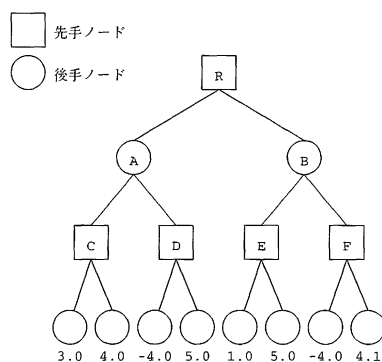


図2 評価値の与えられたゲーム木

先手ノードが局面 R において指し手を決定するために、3 手先まで深読みしたとき、3 手先の各局面の優劣を判断する基準として、あらかじめ決めておいた静的評価関数により、評価値を与える (評価値が大きいほど先手にとって有利であることを表す. 図 2 参照).

ノード C は先手の手番である. よって、この局面での指し手は先手にとって最も有利な局面になる指し手を選ぶはずである. つまり、最も大きな評価値を持つ局面を選ぶため、この局面の評価値は全ての子ノードの評価値のうち最も大きい評価値となる.

同様に、ノード D, E, F の評価値はそれぞれの子ノードの評価値のうち最も大きい評価値となる (図 3 参照).

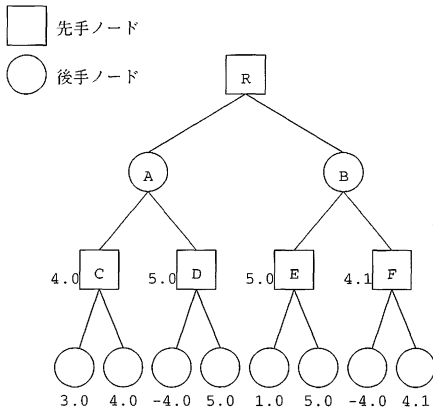


図 3 2 手先を探索後のゲーム木

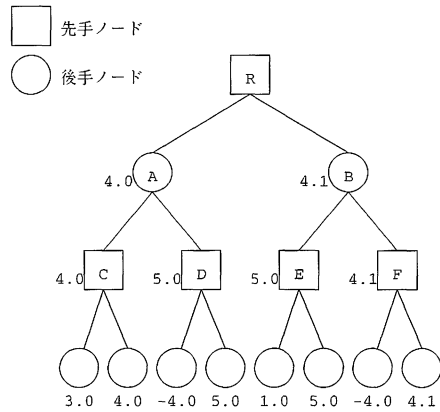


図 4 1 手先を探索後のゲーム木

ノード A は後手の手番である. よって、この局面での指し手は後手にとって最も有利、つまり、先手にとって最も不利な局面になる指し手を選ぶはずである. つまり、最も小さな評価値を持つ局面を選ぶため、この局面の評価値は全ての子ノードの評価値のうち最も小さい評価値となる.

同様に、ノード B の評価値はそれぞれの子ノードの評価値のうち最も小さい評価値となる (図 4 参照).

ノード R は先手の手番である. よって、この局面での指し手は先手にとって最も有利な局面になる指し手を選ぶはずである. よって、最も大きな評価値を持つ子ノードへの指し手がミニマックス法による選択手となる (図 5 参照).

P をノード, P_i を P の任意の子ノードとする. この時、ミニマックス法による $f(P)$ は、次の式で表される.

$$f(P) = \begin{cases} \max f(P_i) & P: \text{先手ノード} \\ \min f(P_i) & P: \text{後手ノード} \\ EV_f(P) & P: \text{葉ノード} \end{cases}$$

ここで $EV_f(P)$ とは静的評価関数によるノード P の値を表す.

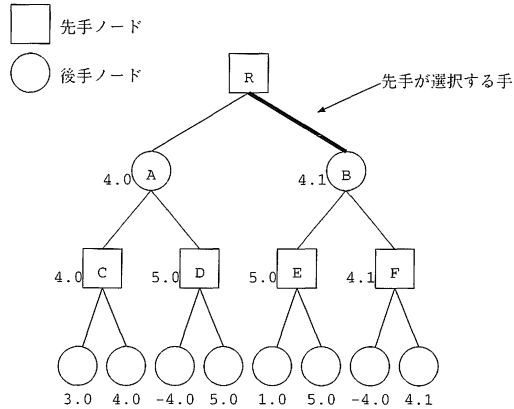


図5 選択手の決定

2.2 OM-Search

相手モデル探索 (Opponent-Model Search, OM-Search) は, 相手プレイヤーのモデルに関する完全な知識を仮定とした, ミニマックス法を拡張したゲーム木探索法 [2] である. また, OM-Search によるゲームプレイは思索的プレイ¹ である.

OM-Search では, 後手が用いる静的評価関数を予測して, その予測した静的評価関数と先手の静的評価関数を用いてゲーム木の探索を行う. 後手は予測した後手の評価関数のみを用いることを仮定する. 以下に, OM-Search の例を示す.

最初に, ミニマックス法と同様に, 局面 R から 3 手先局面までのすべての可能な指し手と局面をあらわしたゲーム木を作る (図 1 参照).

先手が局面 R において, 指し手を決定するために, 3 手先まで深読みしたとき, 葉ノードで表された先読みが最も深い局面に, 先手が 3 手先の各局面の優劣を判定する基準として, あらかじめ決めておいた先手の静的評価関数による評価値をあたえる. また, 後手が 3 手先の各局面の優劣を判定する基準として, 予測した後手の静的評価関数による評価値をあたえる (どちらの評価値も値が大きいほど先手にとって有利であることを表わす. 図 6 参照.).

ノード C は先手の手番である. よって, この局面での先手の指し手は先手にとって最も有利な局面になる指し手を選ぶはずである. つまり, 最も大きな先手の評価値を持つ局面を選ぶため, この局面の先手の評価値は全ての子ノードの先手の評価値のうち最も大きい評価値となる.

また, 後手は後手の静的評価関数のみを用いることを仮定している. よって, 後手は, 後手の静的評価関数で評価した局面のなかで, 最も先手が有利な局面を先手が選ぶと考えている. だから, この局面の後手の評価値は全ての子ノードの後手の評価値のうち最も大きい評価値となる.

同様にして, ノード D, E, F の先手の評価値はそれぞれの子ノードの先手の評価値のうち

¹ 思索的プレイとは, 必ずしも相手が最善手を打ってこないことを仮定したゲーム木探索法である.

最も大きい評価値となり、後手の評価値はそれぞれの子ノードの後手の評価値のうち最も大きい評価値となる (図7参照).

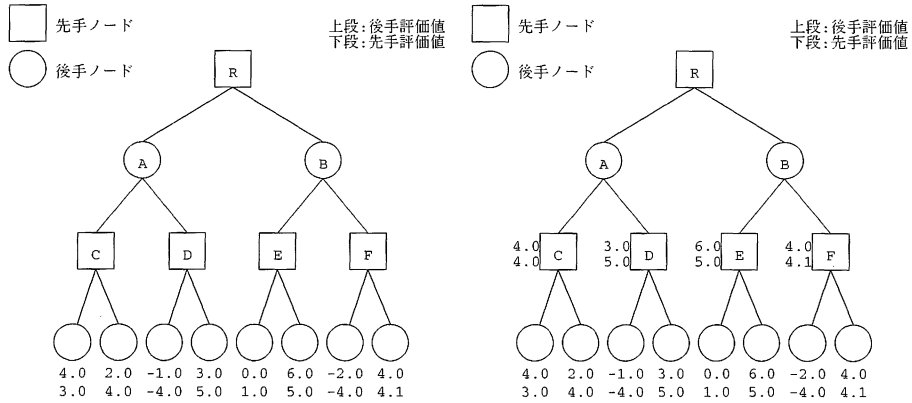


図6 両方の評価値が与えられたゲーム木

図7 2手先を探索後のゲーム木

ノード A は後手の手番である. よって, この局面での後手の指し手は後手にとって最も有利, つまり, 先手にとって最も不利な局面になる指し手を選ぶはずである. つまり, 最も小さな後手の評価値を持つ局面を選ぶため, この局面の後手の評価値は全ての子ノードの後手の評価値のうち最も小さい評価値となる²

また, この局面において先手は選択権がないので, この局面の先手の評価値は後手を選んだ局面の先手の評価値となる.

同様に, ノード B の後手の評価値はそれぞれの子ノードの後手の評価値のうち最も小さい評価値となり, 先手の評価値は後手が選択した手の局面の先手の評価値となる (図8参照).

ノード R は先手の手番である. よって, この局面での指し手は先手にとって最も有利な局面になる指し手を選ぶはずである. よって, 最も大きな先手の評価値を持つ子ノードへの指し手が OM-Search による選択手となる (図9参照).

P をノード, P_i を P の任意の子ノードとする. この時, OM-Search による先手の評価値 $F(p)$, 後手の評価値 $g(p)$ は, 次の式で表される.

$$F(P) = \begin{cases} \max F(P_i) & P: \text{先手ノード} \\ F(P_j) \text{ with } j \text{ where } g(P_j) = \min g(P_i) & P: \text{後手ノード} \\ EV_f(P) & P: \text{葉ノード} \end{cases}$$

$$g(P) = \begin{cases} \max g(P_i) & P: \text{先手ノード} \\ \min g(P_i) & P: \text{後手ノード} \\ EV_g(P) & P: \text{葉ノード} \end{cases}$$

² ノード C とノード D の評価値を比較し, その値の小さいノード D の後手プレイヤー評価値をノード A の後手プレイヤー評価値とし, 後手に選択されたノード D の先手評価値をノード A の先手評価値とする.

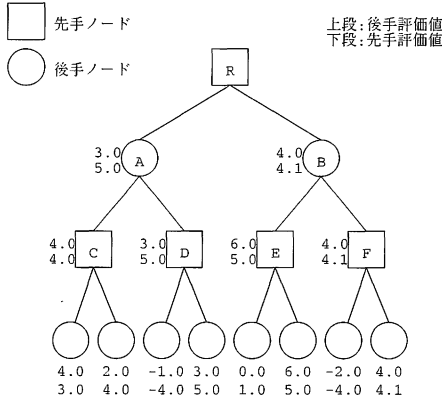


図 8 1 手先を探索後のゲーム木

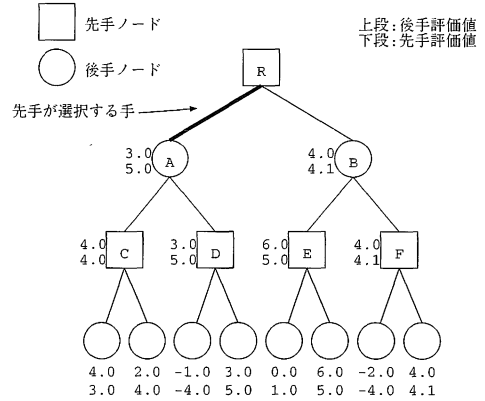


図 9 選択手の決定

ここで $EV_f(P)$ とは先手の静的評価関数によるノード P の評価値を表し、 $EV_g(P)$ とは先手が予測した後手の静的評価関数によるノード P の評価値を表す。

OM-Search によるゲーム木探索はミニマックス法による評価値より小さくなることはない [2]。これは図 5, 図 9 の比較でも明らかである。それは、先手の静的評価関数と先手が予測した後手の静的評価関数が異なるために起こる。

これをミス³と呼ぶ。実際、先ほどの例でも図 8 のノード A において、ミスが発生している。

2.3 勝負手探索

人間のプレーヤーは今の局面が不利と分かっているとき、「勝負手」と呼ばれる手を打つことがある。勝負手とは(最善手ではない)逆転をねらう手、または挽回をねらう手である。松原・飯田 [3] は、勝負手とは

- 次善手
- 戦線を拡大する手
- 後手が迷うような局面に導く手

とした。しかし、これらには、実際に勝負手を指すためのアルゴリズムがなく、現実的な探索法として用いるには問題があった。

堀田の勝負手探索 [4] では「読む確率」を用いて、実際に勝負手を指すためのアルゴリズムを示している。

³ 先手の静的評価関数は後手の評価関数より優れていることを仮定とする。

2.4 Prob-Cut

人間のプレーヤーは全ての手を探索しなくても良い手を見つけることができる。その方法をコンピュータで実現し、より短時間でより深いゲーム木の探索を行うゲーム木探索法が Prob-Cut[5] である。

Prob-Cut は今までに説明した思索プレイとは異なり、ミニマックス法と同じ手を打つことを目標としている。

Prob-Cut では、数手先の評価値分散を考慮に入れることによって、無駄な探索を省いている。これにより、ゲーム木の探索スピードはミニマックス法の約 6 倍になった (ミニマックス法と同じ手を選ぶ確率は約 95%)。

この Prob-Cut を用いたオセロゲームプログラム LOGISTELLO は現在、オセロゲームプログラムの中でも最も強いゲームプログラムである。

3. 期待値法

3.1 期待値法とミニマックス法の相違点

ゲーム木による評価関数を用いた先読み戦略を用いる場合、必ず水平線効果による影響を受ける。つまり、先読みした深さの次の手で局面ががらりと変わり、不利と考えていた局面が有利に、有利と考えていた局面が不利になる場合がある。

ミニマックス法では、このような水平線効果の影響を考慮したゲーム木探索法ではない。従って、評価関数の精度が悪くなればなるほど、水平線効果の影響を受けることになる。なぜなら、その次の手の局面ががらりと変わる可能性は評価関数の精度が悪いほど高くなるからである。

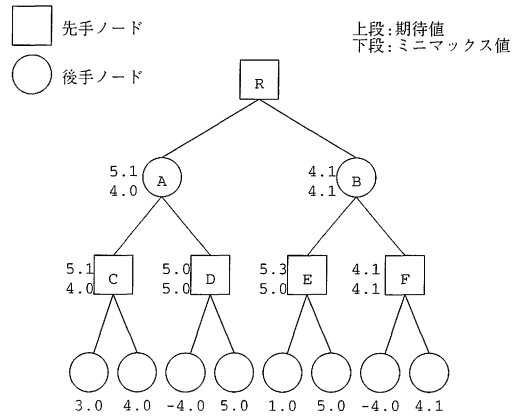
また、チェス、将棋、囲碁、オセロなど最終局面まで先読みすると、先読みする局面の数がとても大きくなるゲームでは、完璧な評価関数が存在しない。よって、必ず水平線効果の影響を受ける。

われわれの提案する期待値法は、この評価関数の不正確さを考慮したゲーム木探索法である。期待値法では葉ノードの評価値が a の時、葉ノードから n 手先まで探索したときの評価値のぶれは $N(a, \sigma^2)$ の正規分布となることを仮定している。期待値法はこの仮定より、葉ノードの n 手先で最もよい評価値がよくなるであろう手を選択する。我々はこの評価を期待値と呼ばれる値で行う。期待値法は水平線効果による影響が少なくなり、後の局面でより有利な局面を迎える手を打てるゲーム木探索法である。

また、期待値法では先手は評価値と期待値を、後手は評価値のみしか知らないことを仮定する。そのため、期待値を求めた後、各ノードでの評価値、期待値の決定は OM-Search と同じである。

図 10 において、ミニマックス法と期待値法の違いを示す。ノード R, A~F における下段の値は、各ノードから葉ノードまでのミニマックス値、上段の値は、各ノードから葉ノードのさらに 2 手先まで探索したときに得られるであろう値、つまり、期待値を表す。

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値が大きいノード B である。一方、期待値法で手を選択した場合、選択されるのは期待値が大きいノード A である (詳しいアルゴリズムは 3.4 節で示す)。

図 10 期待値法とミニマックス法の違い ($\sigma^2 = 8$)

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、期待値法では評価値と探索したゲーム木の深さのさらに 2 手先で得られるであろう値 (我々はこの値を期待値と呼ぶ。ノード上段の値.) を基準に手を選択している点である。

次節以降において、期待値の求め方、期待値法のアルゴリズムを示す。

3.2 期待値

期待値とは、探索するゲーム木の深さから更に n 手先まで探索したときに得られるであろうミニマックス値であることを仮定する。

この数手先の評価値は、葉ノードの評価値を平均とした正規分布であることを仮定する。

まず、葉ノードが c と d の 2 つの場合を考える (図 11 参照)。

c の評価値を f_c 、 d の評価値を f_d とする。ノード c をルートノードとするゲーム木 A のミニマックス値 (ノード C からさらに n 手先まで探索したときのミニマックス値) f_a は分散 σ^2 および平均 f_c 、ノード d をルートノードとするゲーム木 B のミニマックス値 (ノード D からさらに n 手先まで探索したときのミニマックス値) f_b は分散 σ^2 および平均 f_d で正規分布すると仮定する。つまり、 f_a は

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-f_c)^2}{2\sigma^2}}$$

という確率密度関数を持つ正規乱数であり、 f_b は

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-f_d)^2}{2\sigma^2}}$$

という確率密度関数を持つ正規乱数であるとする。

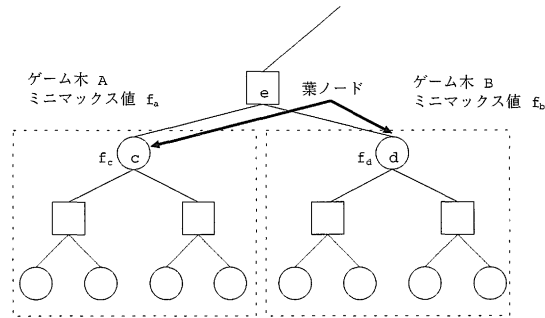


図 11 葉ノードが 2 つの場合の期待値

このとき、ノード e の期待値 $E(\sigma^2, f_c, f_d)$ は次のようにして求まる。

$$E(\sigma^2, f_c, f_d) = f_a * (f_a \geq f_b \text{ の確率}) + f_b * (f_b > f_a \text{ の確率})$$

つまり、

$$E(\sigma^2, f_c, f_d) = \int \int_{x \geq y} x g(x|f_c) g(y|f_d) dx dy \quad (f_a > f_b \text{ のとき}) \\ + \int \int_{x < y} y g(x|f_c) g(y|f_d) dx dy \quad (f_b > f_a \text{ のとき})$$

ただし、 $g(z|f) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(z-f)^2}{2\sigma^2}}$ とする。すなわち、葉ノードでの評価値が f の時、さらに深さ n まで探索した時のミニマックス値 z は正規分布 $N(f, \sigma^2)$ に従うことを仮定する。

これを一般化する。葉ノードの数が k 個、それぞれの評価値を f_1, f_2, \dots, f_k とすると、

$$E(\sigma^2, f_1, \dots, f_k) = \int \int \dots \int_{x_1 \geq x_2, x_1 \geq x_3, \dots, x_1 \geq x_k} x_1 g(x_1|f_1) g(x_2|f_2) \dots g(x_k|f_k) dx_1 dx_2 \dots dx_k \\ + \int \int \dots \int_{x_2 > x_1, x_2 \geq x_3, \dots, x_2 \geq x_k} x_2 g(x_1|f_1) g(x_2|f_2) \dots g(x_k|f_k) dx_1 dx_2 \dots dx_k \\ + \dots + \int \int \dots \int_{x_n > x_1, x_n > x_2, \dots, x_n > x_{k-1}} x_n g(x_1|f_1) g(x_2|f_2) \dots g(x_k|f_k) dx_1 dx_2 \dots dx_k$$

となる。

3.3 期待値の計算

前節で導いた計算式は、コンピュータで計算させるのに、適した計算式ではない。このため、我々はこの式に変換を施し、コンピュータで計算可能な式に変換した(葉ノードが 2 個の場合のみ)。

$$E(\sigma^2, l, m) = \int \int_{x \geq y} x \frac{1}{(\sqrt{2\pi\sigma})^2} e^{-\frac{(x-l)^2}{2\sigma^2}} e^{-\frac{(y-m)^2}{2\sigma^2}} dx dy + \int \int_{x < y} y \frac{1}{(\sqrt{2\pi\sigma})^2} e^{-\frac{(x-l)^2}{2\sigma^2}} e^{-\frac{(y-m)^2}{2\sigma^2}} dx dy$$

この式をより簡単に数値積分するために変換を行う. そのために, x, y という変数を u, v という変数に変換する. $\begin{pmatrix} u \\ v \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix}$ とする. ただし, $A = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$ という行列を用いて変換を行う. 逆行列は $A^{-1} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$ となる.

$\begin{pmatrix} x \\ y \end{pmatrix} = A^{-1} \begin{pmatrix} u \\ v \end{pmatrix}$ より, $x \geq y \Leftrightarrow u \geq 0$ となる. $\frac{D(x,y)}{D(u,v)} = \begin{vmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{vmatrix} = 1$ より,

$$\begin{aligned} (\sqrt{2\pi}\sigma)^2 E(\sigma^2, l, m) &= \int_0^\infty \int_{-\infty}^\infty \frac{1}{\sqrt{2}}(u+v)e^{-\frac{(u+v-\sqrt{2}l)^2}{4\sigma^2}} e^{-\frac{(-u+v-\sqrt{2}m)^2}{4\sigma^2}} dv du \\ &+ \int_{-\infty}^0 \int_{-\infty}^\infty \frac{1}{\sqrt{2}}(-u+v)e^{-\frac{(u+v-\sqrt{2}l)^2}{4\sigma^2}} e^{-\frac{(-u+v-\sqrt{2}m)^2}{4\sigma^2}} dv du \\ &= \int_0^\infty \int_{-\infty}^\infty \frac{1}{\sqrt{2}}(u+v)e^{-\frac{(v-\frac{l+m}{\sqrt{2}})^2}{2\sigma^2}} e^{-\frac{(u-\frac{l-m}{\sqrt{2}})^2}{2\sigma^2}} dv du \\ &+ \int_{-\infty}^0 \int_{-\infty}^\infty \frac{1}{\sqrt{2}}(-u+v)e^{-\frac{(v-\frac{l+m}{\sqrt{2}})^2}{2\sigma^2}} e^{-\frac{(u-\frac{l-m}{\sqrt{2}})^2}{2\sigma^2}} dv du \end{aligned}$$

ここで, 付録1の定理5より,

$$\begin{aligned} \int_{-\infty}^\infty (u+v)e^{-\frac{(v-\frac{l+m}{\sqrt{2}})^2}{2\sigma^2}} dv &= \sqrt{2\pi}\sigma(u + \frac{l+m}{\sqrt{2}}) \\ \int_{-\infty}^\infty (-u+v)e^{-\frac{(v-\frac{l+m}{\sqrt{2}})^2}{2\sigma^2}} dv &= \sqrt{2\pi}\sigma(-u + \frac{l+m}{\sqrt{2}}) \end{aligned}$$

よって,

$$\sqrt{2\pi}\sigma E(\sigma^2, l, m) = \int_0^\infty \frac{1}{\sqrt{2}}(u + \frac{l+m}{\sqrt{2}})e^{-\frac{(u-\frac{l-m}{\sqrt{2}})^2}{2\sigma^2}} du + \int_{-\infty}^0 \frac{1}{\sqrt{2}}(-u + \frac{l+m}{\sqrt{2}})e^{-\frac{(u-\frac{l-m}{\sqrt{2}})^2}{2\sigma^2}} du$$

次に, 付録1の定理6より,

$$\int_0^\infty \frac{1}{\sqrt{2}}(u + \frac{l+m}{\sqrt{2}})e^{-\frac{(u-\frac{l-m}{\sqrt{2}})^2}{2\sigma^2}} du = \frac{1}{\sqrt{2}}\{\sigma^2 e^{-\frac{(l-m)^2}{4\sigma^2}} + \sigma \frac{2l}{\sqrt{2}\sigma} \int_{\frac{m-l}{\sqrt{2}\sigma}}^\infty e^{-\frac{t^2}{2}} dt\}$$

付録1の定理7より,

$$\int_{-\infty}^0 \frac{1}{\sqrt{2}}(-u + \frac{l+m}{\sqrt{2}})e^{-\frac{(u-\frac{l-m}{\sqrt{2}})^2}{2\sigma^2}} du = \frac{1}{\sqrt{2}}\{\sigma^2 e^{-\frac{(l-m)^2}{4\sigma^2}} + \sigma \frac{2m}{\sqrt{2}} \int_{-\infty}^{\frac{m-l}{\sqrt{2}\sigma}} e^{-\frac{t^2}{2}} dt\}$$

よって、

$$\begin{aligned} \sqrt{2\pi}\sigma E(\sigma^2, l, m) &= \frac{1}{\sqrt{2}} \left\{ \sigma^2 e^{-\frac{(l-m)^2}{4\sigma^2}} + \sigma \frac{2l}{\sqrt{2}} \int_{\frac{m-l}{\sqrt{2\sigma}}}^{\infty} e^{-\frac{t^2}{2}} dt \right\} + \frac{1}{\sqrt{2}} \left\{ \sigma^2 e^{-\frac{(l-m)^2}{4\sigma^2}} + \sigma \frac{2m}{\sqrt{2}} \int_{-\infty}^{\frac{m-l}{\sqrt{2\sigma}}} e^{-\frac{t^2}{2}} dt \right\} \\ &= \frac{1}{\sqrt{2}} \left\{ 2\sigma^2 e^{-\frac{(l-m)^2}{4\sigma^2}} + \sqrt{2}l\sigma \int_{\frac{m-l}{\sqrt{2\sigma}}}^{\infty} e^{-\frac{t^2}{2}} dt + \sqrt{2}m\sigma \int_{-\infty}^{\frac{m-l}{\sqrt{2\sigma}}} e^{-\frac{t^2}{2}} dt \right\} \end{aligned}$$

すなわち、

$$E(\sigma^2, l, m) = \frac{1}{\sqrt{2\pi}} \left\{ \sqrt{2}\sigma e^{-\frac{(l-m)^2}{4\sigma^2}} + l \int_{\frac{m-l}{\sqrt{2\sigma}}}^{\infty} e^{-\frac{t^2}{2}} dt + m \int_{-\infty}^{\frac{m-l}{\sqrt{2\sigma}}} e^{-\frac{t^2}{2}} dt \right\}$$

ここで、

$$\int_{\frac{m-l}{\sqrt{2\sigma}}}^{\infty} e^{-\frac{t^2}{2}} dt \text{ および } \int_{-\infty}^{\frac{m-l}{\sqrt{2\sigma}}} e^{-\frac{t^2}{2}} dt$$

は、正規分布 $N(0, \frac{1}{\sqrt{2\pi}})$ における、 $\frac{m-l}{\sqrt{2\sigma}}$ の分布関数値であるから、 σ^2, l, m が与えられれば、容易に $E(\sigma^2, l, m)$ が求められる。

しかし、引数が4つ以上になった場合、この変換は複雑になり、計算機で計算可能な式に変換することは困難である。よって、次のようにして求める。

葉ノードの評価値がそれぞれ、 f_1, f_2, \dots, f_w とする (図12参照)。

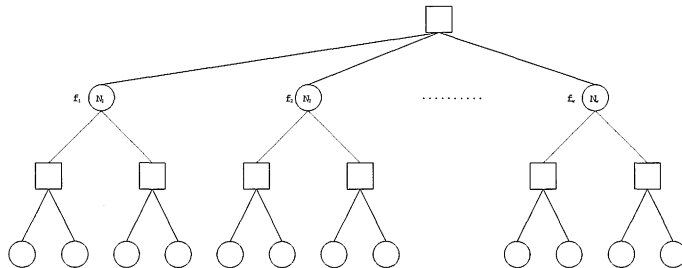


図12 3分木以上の期待値

葉ノードの評価値が f_i のとき、葉ノードからさらに n 手先まで探索した時のミニマックス値を F_i とすると、

仮定より F_i は、

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-f_i)^2}{2\sigma^2}}$$

という確率密度関数を持つ正規分布に従う ($N(f_i, \sigma^2)$)。葉ノードからさらに n 手先まで探索した場合、葉ノードの親ノードがとる値は、 $P = \max(F_1, F_2, \dots, F_w)$ である。

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

$$P_j = \max(F_{1j}, F_{2j}, \dots, F_{wj}) \quad (1 \leq j \leq m)$$

とし, $P_j (1 \leq j \leq m)$ も互いに独立であると仮定する.

$$S_m = \sum_{j=1}^m P_j, \quad \text{Var } P_j = \sigma^2$$

(σ は P_1 の標準偏差. $\sigma^2 < \infty$ は $w = 3$ の場合について付録 2 で示した.)

とすると,

$$\text{Var } (S_m) = \sum_{j=1}^m \text{Var } (P_j) = m \text{Var } (P_1) = m\sigma^2$$

このとき, 中心極限定理より, 任意の $x (-\infty < x < \infty)$ に対し,

$$|\Pr\left(\frac{S_m - E(S_m)}{\sqrt{m}\sigma} < x\right) - \Phi(x)| < \frac{A}{\sqrt{m}} \quad \dots (1)$$

が成り立つ. ただし, $\Pr()$ は () 内の条件が成立する確率, $E(S_m)$ は S_m の平均値, $\Phi(x)$ は $N(0, 1)$ の分布関数, A は定数である.

$$E(S_m) = \sum_{j=1}^m E(P_j) = mE(P_1)$$

だから, $V_m = \frac{S_m}{m}$ とおくと, (1) は

$$|\Pr\left(\frac{V_m - E(P_1)}{\sigma} < \frac{x}{\sqrt{m}}\right) - \Phi(x)| < \frac{A}{\sqrt{m}} \quad (-\infty < x < \infty) \dots (2)$$

となる. ここで, $\Phi(2.58) = 0.9951$, $\Phi(-2.58) = 0.0049$ を (2) に代入すると,

$$|\Pr\left(\frac{V_m - E(P_1)}{\sigma} < \frac{2.58}{\sqrt{m}}\right) - 0.9951| < \frac{A}{\sqrt{m}} \dots (3)$$

$$|\Pr\left(\frac{V_m - E(P_1)}{\sigma} \leq \frac{-2.58}{\sqrt{m}}\right) - 0.0049| < \frac{A}{\sqrt{m}} \dots (4)$$

すなわち,

$$-\frac{A}{\sqrt{m}} + 0.9951 < \Pr\left(\frac{V_m - E(P_1)}{\sigma} < \frac{2.58}{\sqrt{m}}\right) < \frac{A}{\sqrt{m}} + 0.9951 \dots (5)$$

$$-\frac{A}{\sqrt{m}} - 0.0049 < -\Pr\left(\frac{V_m - E(P_1)}{\sigma} \leq \frac{-2.58}{\sqrt{m}}\right) < \frac{A}{\sqrt{m}} - 0.0049 \dots (6)$$

となる。(5)+(6)を計算すると、

$$-\frac{2A}{\sqrt{m}} + 0.9902 < Pr(|\frac{V_m - E(P_1)}{\sigma}| < \frac{2.58}{\sqrt{m}}) < \frac{2A}{\sqrt{m}} + 0.9902$$

$$|Pr(|V_m - E(P_1)| < \frac{2.58\sigma}{\sqrt{m}}) - 0.9902| < \frac{2A}{\sqrt{m}} \dots (7)$$

以上より、 n が十分大きいときには $E(P_1)$ の近似値として V_m を用いることができる。また、この計算で発生する誤差は $O(\frac{1}{\sqrt{m}})$ となる。よって、期待値 $E(\sigma, f_1, f_2, \dots, f_w)$ は、 m を大きくすれば、徐々に収束する。以上の方法により、3分木以上の場合の期待値が算出できる。

3.4 期待値法アルゴリズム

我々は前節の期待値の計算により、葉ノードからさらに n 手先で得られるであろうミニマックス値を予測することが出来る。我々はこの期待値と静的評価関数で求めた評価値の両方を用いてゲーム木探索を行う。一方、後手は、期待値という考えを知らないため、評価値のみを用いてミニマックス法を使ってゲーム木探索を行うと考える。

よって、期待値法によるゲーム木探索では、先手は評価値と期待値、後手は評価値のみをゲーム木探索に用いることが出来ると仮定する。以下に、期待値法によるゲーム木探索の例を示す。

まず、ミニマックス法と同様に、局面 R から 3 手先局面までのすべての可能な指し手と局面をあらわしたゲーム木を作る (図 1 参照)。

先手が局面 R において、指し手を決定するために、3 手先まで深読みしたとき、葉ノードで表された先読みが最も深い局面に、3 手先の各局面の優劣を判定する基準として、あらかじめ決めておいた静的評価関数による評価値をあたえる (図 2 参照)。

ノード C は先手の手番である。後手は静的評価関数のみを用いることを仮定している。よって、後手は、静的評価関数で評価した局面のなかで、最も先手が有利な局面を先手が選ぶと考えている。だから、この局面の評価値は全ての子ノードの評価値のうち最も大きい評価値となる。

また、このノードは葉ノードの親ノードである。先手は葉ノードからさらに n 手先での局面の優劣を判定する基準として、全ての子ノードの評価値から葉ノードのさらに n 手先で得られるであろうミニマックス値 (期待値) を算出する。

同様にして、ノード D, E, F の評価値はそれぞれの子ノードの評価値のうち最も大きい評価値となり、期待値は全ての子ノードの評価値から算出される (図 13 参照)。

ノード A は後手の手番である。よって、この局面での後手の指し手は後手にとって最も有利、つまり、先手にとって最も不利な局面になる指し手を選ぶはずである。つまり、最も小さな評価値を持つ局面を選ぶため、この局面の評価値は全ての子ノードの評価値のうち最も小さい評価値となる。

また、この局面において先手は選択権がないので、この局面の期待値は後手が選んだ局面の期待値となる。

同様にして、ノード B の評価値はそれぞれの子ノードの評価値のうち最も小さい評価値となり、期待値は後手が選択した手の局面の期待値となる (図 14 参照)。

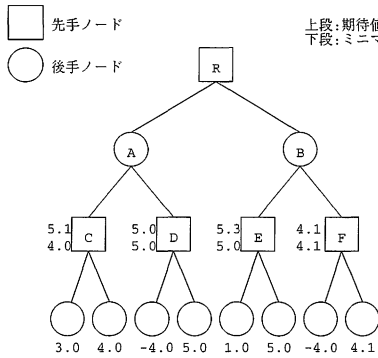


図 13 2手先を探索後のゲーム木

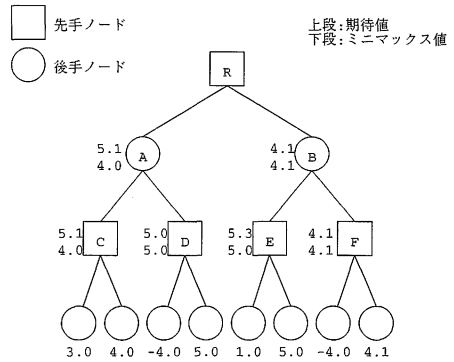


図 14 1手先を探索後のゲーム木

ノード R は先手の手番である. よって, この局面での指し手は葉ノードのさらに n 手先で先手にとって最も有利な局面になる指し手を選ぶはずである. よって, 最も大きな期待値を持つ子ノードへの指し手が期待値法による選択手となる (図 15 参照).

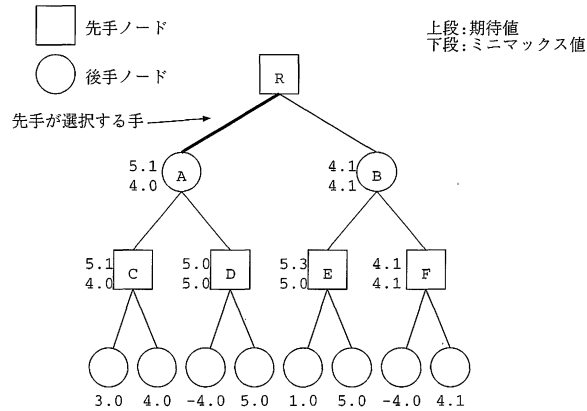


図 15 選択手の決定

P をノード, P_i を P の任意の子ノードとする. この時, 期待値法による各ノード P の期待値 $K(P)$, ミニマックス値 $g(P)$ は, 次の式で表される.

$$K(P) = \begin{cases} \max K(P_i) & P \text{ 先手ノード} \\ K(P_j) \text{ with } j \text{ where } g(P_j) = \min g(P_i) & P \text{ 後手ノード} \\ KV_f(P) & P \text{ 葉ノードの親ノード} \end{cases}$$

$$g(P) = \begin{cases} \max g(P_i) & P \text{ 先手ノード} \\ \min g(P_i) & P \text{ 後手ノード} \\ EV_f(P) & P \text{ 葉ノード} \end{cases}$$

ここで $KV_f(P)$ とは葉ノードにおける評価値をもとに算出されたノード P の期待値を表し, $EV_f(P)$ とは静的評価関数によるノード P の値を表す.

3.5 期待値法の枝刈

次に、期待値法の枝刈について考える。枝刈とはゲーム木全てを探索せず、指し手を決めるために必要ないノードを探索しないことである。実際に時間が制限されているゲームでより深くゲーム木を探索するために枝刈は必要である (図 16 参照)。

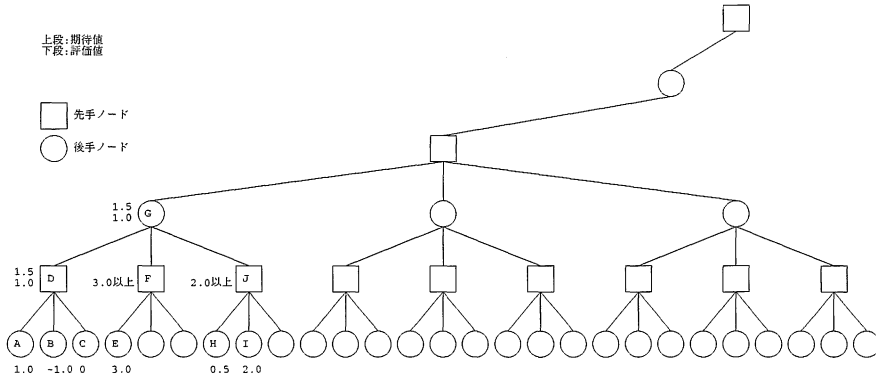


図 16 期待値法の枝刈

F ノードにおいて、子ノード E の評価値が 3.0 と分かった場合、G ノードにおいて F ノードの評価値・期待値が選ばれることはない。F ノードの評価値は子ノード E の評価値が 3.0 と分かった時点で、3.0 以上になることが確定し (F ノードは先手ノードより、全ての子ノードの評価値のうち最大の評価値を自らの評価値とするため)、D ノードの評価値 1.0 より必ず大きくなるのが分かる。G ノードは後手ノードなので、全ての子ノードの評価値のうち最小の評価値を自らの評価値とし、その選択された評価値を持つ子ノードの期待値を自らの期待値とする。

このことから、F ノードは最小の評価値を持つ子ノードではないので、評価値も期待値も G ノードの評価値・期待値とならないことがわかる。

よって、これ以上の探索は無意味であるため、E ノード以降のノードの探索は行わず、また、F ノードの期待値も求めない (枝刈)。

J ノードにおいても、I ノードの評価値が 2.0 と分かった時点で J ノードの評価値が 2.0 以上になることが分かるため、枝刈をすることができる。

2 分木における最も枝刈出来る例が図 17 である (点線部を枝刈)。

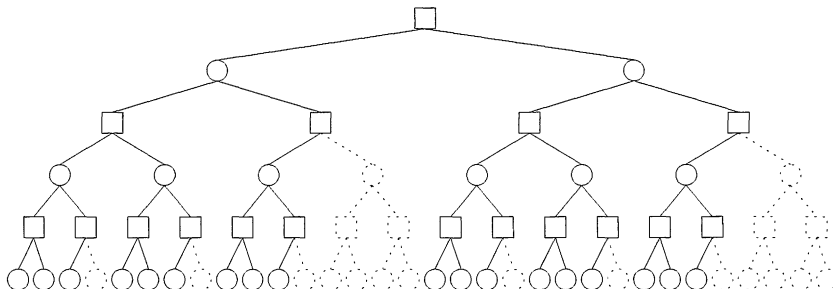


図 17 2 分木の枝刈 (最善の場合)

4. 仮想ゲーム木による実験

期待値法がミニマックス法より有効であるかどうかを見るために仮想ゲーム木を用いて実験を行った。

4.1 実験方法

4.1.1 比較方法

ここでは、後手がミニマックス法で指す時に先手が期待値法で指す場合とミニマックス法で指す場合にどちらの指し方がどの程度先手にとって有利になるかを実験で比較する。すなわち、先手が期待値法、後手がミニマックス法で指す場合と、先手も後手もミニマックス法で指す場合とで比較を行う。先手の指し手数を $T(=1,5,10,20)$ 、ただし、 $T=20$ 実行時の途中経過として $T=1,5,10$ の結果を出す、各指し手の選択における探索の深さを $D(=3)$ とする。またここで用いる仮想ゲーム木の枝の数は $B(=2,8)$ とする。

まず先手はルートノードから深さ $D(=3)$ までのゲーム木を探索して、期待値法及びミニマックス法で指し手を決める。以後選択した手についてのみ探索を行う。次に後手は、先手の指した深さ 1 のノードからさらに深さ D 、すなわちルートノードから深さ 4 までのゲーム木を探索してミニマックス法で指し手を決める。これを繰り返し先手が T 回目に打つ時に探索するときのルートノードからの深さ S は $D+2(T-1)$ となる。なぜなら、 $T-1$ 回先手後手が交互に指し手を進めルートノードから $2(T-1)$ の深さまで局面が進む。先手が T 回目に打つ時、その深さ $2(T-1)$ の局面からさらに深さ D まで探索することになるからである。最終的に先手が $2(T-1)$ で選択した手を S の 2 手先である深さ $D+2T$ で予測される評価値の大小によって比較をする。(図 18)

期待値法では水平線効果の影響を考慮に入れ、葉ノードの数手先で最も良い評価値が来るであろう手を選択することから実際にミニマックス法より数手先(ここでは 2 手先)で良い手を打っているかを見るために予測される評価値の大小によって両者を比較をする。

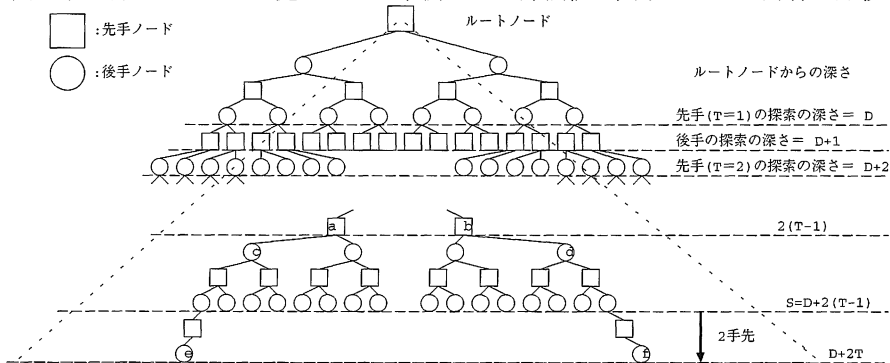


図 18 ゲーム木の例

よって、予測される評価値のことを以下では 2 手先の評価値と言うことにする。そして、それぞれ比較したものを

- 期待値法が選択した手の2手先の評価値が,ミニマックス法が選択した手の2手先の評価値より大きい場合… $K \circ M \times$
- 期待値法が選択した手の2手先の評価値が,ミニマックス法が選択した手の2手先の評価値より小さい場合… $K \times M \circ$

として集計し,両者の数によってどちらが有効であるか比較する.例えば 図 4-1 の場合先手の期待値法が a で選択した手が c,先手のミニマックス法が b で選択した手が d とすると,それぞれの2手先の評価値を e, f とする.その大小関係が $e > f$ ならば $K \circ M \times$, $e < f$ ならば $K \times M \circ$ ということになる.

4.1.2 アルゴリズム

仮想ゲーム木による実験のアルゴリズムは次のとおりである.

1. 深さ D , 完全 B 分木のゲーム木をつくる
2. 葉ノードに評価値 H を平均 0 , 分散 $\sigma^2 (= 5.0, 2.0, 1.0, 0.5, 0.1)$ の正規乱数によって与える
3. 葉ノードの評価値からその親ノードに期待値 E を与える
4. $T=1$ なので, 葉ノードの2手先の評価値 N を平均 葉ノードの評価値, 分散 1.0 の正規乱数によって与える
5. 期待値と評価値を深さ 1 まで以下のような方法であげる
 - 先手ノードでは, 子ノードの最大の評価値と期待値を親ノードの評価値と期待値にする.
 - 後手ノードでは, 子ノードの最小の評価値を親ノードの評価値とし, 最小の評価値を持つ子ノードの期待値を親ノードの期待値とする.
6. 2手先の評価値を深さ 1 まで以下のような方法であげる
 - 先手ノードでは, 子ノードのうちで最大の評価値を親ノードの評価値とする.
 - 後手ノードでは, 子ノードのうちで最小の評価値を親ノードの評価値とする.
7. それぞれ選択した手を打つ
8. 深さ $D+2t-1$ ($1 \leq t \leq T-1$) のノードに評価値 H を平均 深さ $D+2t-2$ の評価値, 分散 $(\frac{1}{D} \sigma)^2$ の正規乱数によって与える
9. 評価値を深さ $2t$ まで 6 と同じ方法であげる
10. それぞれ選択した手を打つ
11. 深さ $D+2t$ のノードに評価値 H を平均 深さ $D+2t-1$ の評価値, 分散 $(\frac{1}{D} \sigma)^2$ の正規乱数によって与える
12. 深さ $D+2t$ の評価値からその親ノードに期待値 E を与える
13. $T=5, 10, 20$ ($t=T-1$) のとき, 深さ $D+2t$ の2手先の評価値 N を平均 深さ $D+2t$ の評価値, 分散 1.0 の正規乱数によって与える
14. 期待値と評価値を深さ $2t+1$ まで 5 と同じ方法であげる

15. $T=5,10,20(t=T-1)$ のとき, 2手先の評価値を深さ $2T-1$ まで6と同じ方法であげる
 16. $T=5,10,20(t=T-1)$ のとき, それぞれ選択した手の2手先の評価値の大小を比較する

- 期待値法が選択した手の2手先の評価値: N_k
- ミニマックス法が選択した手の2手先の評価値: N_m

とすると,

- $N_k > N_m \Rightarrow K \circ M \times$
- $N_k < N_m \Rightarrow K \times M \circ$

として集計し, 終了する

17. $T \neq 20$ のとき, それぞれ選択した手を打ち, 8へ戻る

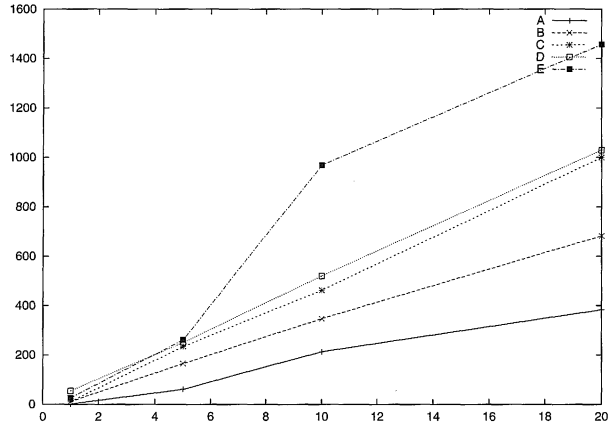
以上のプロセスを 10000 回行う.

4.2 結果と考察

前節のアルゴリズムによって得られた結果を表にしたものが表1である. また, 期待値法とミニマックス法で比較をするため $K \circ M \times$ と $K \times M \circ$ の差を出したものがそれぞれ図19, 図20である. この図は常に $K \circ M \times$ の数- $K \times M \circ$ の数にしている.

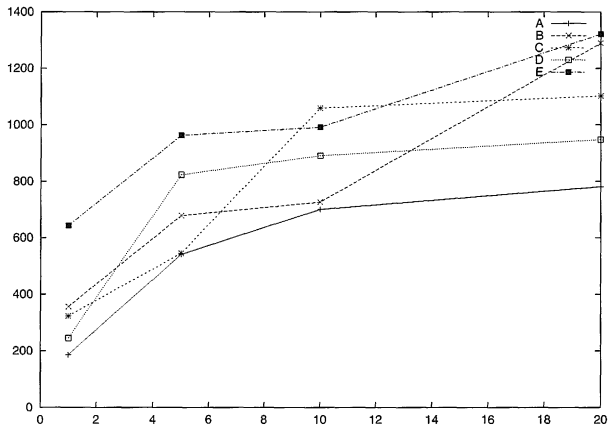
表1 $K \circ M \times$, $K \times M \circ$ の数

B	σ^2	T	1	5	10	20	
2	5.0	$K \circ M \times$	47	373	871	1698	
		$K \times M \circ$	46	312	658	1315	
	2.0	$K \circ M \times$	81	630	1417	2648	
		$K \times M \circ$	70	465	1070	1966	
	1.0	$K \circ M \times$	124	869	1861	3341	
		$K \times M \circ$	113	635	1399	2342	
	0.5	$K \circ M \times$	188	1116	2257	3865	
		$K \times M \circ$	133	866	1737	2836	
	0.1	$K \circ M \times$	346	1701	3397	4953	
		$K \times M \circ$	320	1440	2429	3496	
	8	5.0	$K \circ M \times$	756	2517	3959	5052
			$K \times M \circ$	570	1976	3259	4271
2.0		$K \circ M \times$	1213	3303	4613	5562	
		$K \times M \circ$	857	2625	3887	4273	
1.0		$K \circ M \times$	1545	3728	5117	5522	
		$K \times M \circ$	1222	3183	4058	4420	
0.5		$K \circ M \times$	1946	4296	5169	5461	
		$K \times M \circ$	1701	3446	4279	4514	
0.1		$K \circ M \times$	2856	4898	5418	5659	
		$K \times M \circ$	2213	3936	4427	4338	



line A: $\sigma^2 = 5.0$,line B: $\sigma^2 = 2.0$
 line C: $\sigma^2 = 1.0$,line D: $\sigma^2 = 0.5$,line E: $\sigma^2 = 0.1$
 (縦軸：差の数(件), 横軸：先手の手数T (回))

図 19 B=2 の時のKOM×の数-K×M○の数



line A: $\sigma^2 = 5.0$,line B: $\sigma^2 = 2.0$
 line C: $\sigma^2 = 1.0$,line D: $\sigma^2 = 0.5$,line E: $\sigma^2 = 0.1$
 (縦軸：差の数(件), 横軸：先手の手数T (回))

図 20 B=8 の時のKOM×の数-K×M○の数

いずれの場合でもKOM×の数がK×M○の数を上回った。また、Tの増加にもなって、KOM×とK×M○の差は開いていった。KOM×とK×M○の差は最大で、全体の14.57%の差が出た。T=1のときは、分散が小さくなっていてもほとんど両者の差は変わら

ないが、 T が大きくなるにつれてその差も大きくなっている。また、 B が大きい方が $T=1$ の段階から両者の差が広がっている。分散が大きいと期待値法もミニマックス法も同様の手を打ち、分散が小さくなるにつれて両者が異なる手を打つ数が増えていく。分散が小さくなると、評価値の精度が悪くなり期待値がばらつく。それにより間違っただ手を打っていると考えられる。葉ノードでは良い手だと思われていたものが2手先では悪い手に、またその逆になる可能性が高くなる、つまり水平線効果の影響を考慮に入れている分期期待値法の方が良い結果になっていると考えられる。

さらに、期待値法では期待値を求めている正規乱数の精度が悪く、この精度をあげればもっと期待値法にとって良い結果になると思われる。しかし、精度を10倍良くすると、実行時間は現在の100倍になるので、計算時間の関係でかなり困難になる。

4.3 オセロゲームを想定した実験

ここまで我々は様々なゲームにおける期待値法の有効性を示してきたが、実際のオセロゲームで有効であることを示すための準備としてオセロゲームに近い設定でこの仮想ゲーム木による実験を行った。

4.3.1 オセロゲームの様々な場合の分散

オセロゲームを想定した場合、必要となるのは次の各分散である。

- ゲーム木のルートノードの評価値に対する葉ノードの評価値の分散
- ゲーム木の葉ノードの評価値に対する2手先の評価値の分散

それぞれ、373件の棋譜データを用いて分散を調べた。(ただし、根ノードとして展開する局面は47手目とし、探索の深さは3とした) 結果は次の通り。

- 葉ノードの評価値の分散 … 189.08
- 2手先の評価値の分散 … 180.02

4.3.2 分散の変換

前小節での結果をもとに、オセロゲームに近い設定でシミュレーションを行うために分散の変換を行う。

まず、評価基準として2手先の評価値を用いる場合、その分散を1.0とした場合、下の表2がオセロゲームの設定に近い。

表2 分散の変換

	2手先の評価値の分散	葉ノードの評価値の分散
実際のオセロゲーム	180.02	189.08
シミュレーション	1.0	$189.08/180.02 = 1.05$

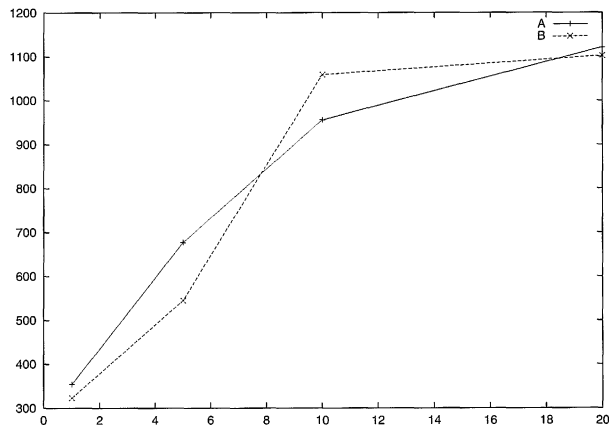
また、実際のおセロゲームの46手目終了時の局面における平均可能指し手数約8なので枝の数 $B=8$ とする。以上の設定で再び仮想ゲーム木による実験を行う。

4.4 オセロゲームを想定した実験の結果と考察

前節の設定によって得られた結果を評価基準によって表にしたものが表3である。また、仮想ゲーム木での実験と同様に期待値法とミニマックス法で比較をするため $K \circ M \times$ と $K \times M \circ$ の差を出したものが図21である。

表3 $K \circ M \times$, $K \times M \circ$ の数

T	1	5	10	20
$K \circ M \times$	1563	3793	5029	5523
$K \times M \circ$	1209	3116	4073	4402



line A:オセロの設定の場合, line B:4.2節の枝の数8,分散1.0の場合
(縦軸:差の数(件),横軸:先手の手数T(回))

図21 $K \circ M \times$ の数- $K \times M \circ$ の数

ここでも、 $K \circ M \times$ の数が $K \times M \circ$ の数を上回った。また、Tの増加にともなって、 $K \circ M \times$ と $K \times M \circ$ の差は開いていった。 $K \circ M \times$ と $K \times M \circ$ の差は最大で、全体の11.21%の差が出た。

この実験は、4.2節の枝の数8,葉ノードの分散を1.0とした仮想ゲーム木の実験の結果に近くなるはずである。図21で示した通り、ほとんど同じ結果となっている。よって、実際のおセロゲームにおいても期待値法はミニマックス法よりも有効であると言える。

5. まとめ

今回の研究により、われわれの提案する期待値法によるゲーム木探索が、ミニマックス法によるゲーム木探索よりも優れていることが、仮想ゲーム木による実験で示すことができた。

今後の課題として以下のようなことを挙げるができる。

- 実際のオセロゲームによる実験を行う
- 2手先よりさらに先の評価値を予測した場合の期待値法の有効性を調べる
- 葉ノードの親ノードが後手ノードとなる場合の期待値計算式の導出
- 3.5節で述べた枝刈りのアルゴリズムの開発と実用化

今後、これらの課題を解決し、さらに期待値法の有効性を増していく必要がある。

参考文献

1. Donald E.Knuth, Ronald W.Moore. An Analysis of Alpha-Beta Pruning. Artificial Intelligence 6,pp.293-326,1975.
2. Iida, H., Uiterwijk, J. W. H. M., Herik, H. J. v. d. and Herschberg, I. S. Potential Applications of Opponent-Model Search Part 1. The Domain of Applicability, ICCA Journal, Vol.16, No.4, pp.201-208,1993.
3. E.Knuth, Ronald W.Moore. An Analysis of Alpha-Beta Pruning. Artificial Intelligence 6,pp.293-326,1975.
4. 堀田 浩司. ゲーム木探索における勝負手探索. 平成10年度 修士論文,1998.
5. Buro, M. ProbCut:An Effective Selective Extension of Alpha-Beta Algorithm, ICCA Journal, Vol18, No.2 pp.71-76,1995.

付録 1

$$\text{定理 1} \quad \int x e^{-\frac{x^2}{2}} dx = -e^{-\frac{x^2}{2}}$$

$$\because f(x) = -e^x, \varphi(t) = -\frac{t^2}{2}, F(t) = -e^{-\frac{t^2}{2}} \text{ とおくと } F(t) = f(\varphi(t))$$

$$\text{定理} : F'(t) = f'(x) \varphi'(t) \text{ を用いて } (-e^{-\frac{t^2}{2}})' = -e^x \times (-t) = t e^{-\frac{t^2}{2}}$$

$$\text{定理 2} \quad \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}$$

$$\text{定理 3} \quad \int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}$$

$$\because (2) \text{ において } t = \frac{x}{\sqrt{2}} \text{ とおくと } x = \sqrt{2}t (= \varphi(t)), \varphi'(t) = \sqrt{2}$$

合成関数の積分定理より

$$\sqrt{2\pi} = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \int_{-\infty}^{\infty} \sqrt{2} e^{-t^2} dt \quad \therefore \int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}$$

$$\text{定理 4} \quad \int_0^{\infty} e^{-\frac{x^2}{2}} dx = \int_{-\infty}^0 e^{-\frac{x^2}{2}} dx = \sqrt{\frac{\pi}{2}}$$

$$\text{定理 5} \quad \int_{-\infty}^{\infty} (ax + b) e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \sqrt{2\pi}\sigma(am + b)$$

$$\therefore t = \frac{x-m}{\sigma} \text{ とおくと定理より } \varphi(t) = \sigma t + m, \quad \varphi'(t) = \sigma$$

$$\begin{aligned} (\text{左辺}) &= \int_{-\infty}^{\infty} \sigma(a(\sigma t + m) + b) e^{-\frac{t^2}{2}} dt \\ &= a\sigma^2 \int_{-\infty}^{\infty} t e^{-\frac{t^2}{2}} dt + \sigma(am + b) \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt \\ &= a\sigma^2 [-e^{-\frac{t^2}{2}}]_{-\infty}^{\infty} + \sigma(am + b)\sqrt{2\pi} \\ &= \sqrt{2\pi}\sigma(am + b) \\ &= (\text{右辺}) \end{aligned}$$

$$\text{定理 6} \quad \int_0^{\infty} (ax + b) e^{-\frac{(x-m)^2}{2\sigma^2}} dx = a\sigma^2 e^{-\frac{m^2}{2\sigma^2}} + \sigma(am + b) \int_{-\frac{m}{\sigma}}^{\infty} e^{-\frac{t^2}{2}} dt$$

$$\therefore t = \frac{x-m}{\sigma} \text{ とおくと定理より } \varphi(t) = \sigma t + m, \quad \varphi'(t) = \sigma$$

$$\begin{aligned} (\text{左辺}) &= \int_{-\frac{m}{\sigma}}^{\infty} \sigma(a(\sigma t + m) + b) e^{-\frac{t^2}{2}} dt \\ &= a\sigma^2 \int_{-\frac{m}{\sigma}}^{\infty} t e^{-\frac{t^2}{2}} dt + \sigma(am + b) \int_{-\frac{m}{\sigma}}^{\infty} e^{-\frac{t^2}{2}} dt \\ &= a\sigma^2 [-e^{-\frac{t^2}{2}}]_{-\frac{m}{\sigma}}^{\infty} + \sigma(am + b) \int_{-\frac{m}{\sigma}}^{\infty} e^{-\frac{t^2}{2}} dt \\ &= a\sigma^2 e^{-\frac{m^2}{2\sigma^2}} + \sigma(am + b) \int_{-\frac{m}{\sigma}}^{\infty} e^{-\frac{t^2}{2}} dt \\ &= (\text{右辺}) \end{aligned}$$

$$\text{定理 7} \quad \int_{-\infty}^0 (ax + b) e^{-\frac{(x-m)^2}{2\sigma^2}} dx = -a\sigma^2 e^{-\frac{m^2}{2\sigma^2}} + \sigma(am + b) \int_{-\infty}^{-\frac{m}{\sigma}} e^{-\frac{t^2}{2}} dt$$

$$\therefore t = \frac{x-m}{\sigma} \text{ とおくと定理より } \varphi(t) = \sigma t + m, \quad \varphi'(t) = \sigma$$

$$\begin{aligned} (\text{左辺}) &= \int_{-\infty}^{-\frac{m}{\sigma}} \sigma(a(\sigma t + m) + b) e^{-\frac{t^2}{2}} dt \\ &= a\sigma^2 \int_{-\infty}^{-\frac{m}{\sigma}} t e^{-\frac{t^2}{2}} dt + \sigma(am + b) \int_{-\infty}^{-\frac{m}{\sigma}} e^{-\frac{t^2}{2}} dt \end{aligned}$$

$$\begin{aligned}
&= a\sigma^2[-e^{-\frac{t^2}{2}}]_{-\frac{m}{\sigma}}^{-\frac{m}{\sigma}} + \sigma(am+b) \int_{-\infty}^{-\frac{m}{\sigma}} e^{-\frac{t^2}{2}} dt \\
&= -a\sigma^2 e^{-\frac{m^2}{2\sigma^2}} + \sigma(am+b) \int_{-\infty}^{-\frac{m}{\sigma}} e^{-\frac{t^2}{2}} dt \\
&= (\text{右辺})
\end{aligned}$$

付録 2

$$\begin{aligned}
E(P_j)^2 &= \int \int \int_{x>y, x>z} x^2 \phi(x-l) \phi(y-m) \phi(z-n) dx dy dz \\
&+ \int \int \int_{y>x, y>z} y^2 \phi(x-l) \phi(y-m) \phi(z-n) dx dy dz \\
&+ \int \int \int_{z>x, z>y} z^2 \phi(x-l) \phi(y-m) \phi(z-n) dx dy dz \quad (\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}})
\end{aligned}$$

被積分関数は常に非負だから、

$$\begin{aligned}
E(P_1) &\leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x^2 + y^2 + z^2) \phi(x-l) \phi(y-m) \phi(z-n) dx dy dz \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (l^2 + 1 + y^2 + z^2) \phi(y-m) \phi(z-n) dy dz \\
&= l^2 + m^2 + n^2 + 3 \left(\int_{-\infty}^{\infty} x^2 \phi(t-a) dt = a^2 + 1 \right) \quad \dots (8)
\end{aligned}$$

一方、

$$\begin{aligned}
E_1(l, m, n) &= \int \int \int_{x>y, x>z} x \phi(x-l) \phi(y-m) \phi(z-n) dx dy dz \\
&\geq \int \int \int_{x>y, x>z, x<0} x \phi(x-l) \phi(y-m) \phi(z-n) dx dy dz \\
&\geq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^0 x \phi(x-l) \phi(y-m) \phi(z-n) dx dy dz \\
&= \int_{-\infty}^0 x \phi(x-l) dx
\end{aligned}$$

$a(l) = \int_{-\infty}^0 x \phi(x-l) dx$ とおくと、定理 6 より、

$$a(l) = - \int_0^{\infty} t \phi(t+l) dt = -e^{-\frac{l^2}{2}} + l \int_{-l}^{\infty} e^{-\frac{t^2}{2}} dt$$

よって、

$$l \int_{-l}^{\infty} e^{-\frac{t^2}{2}} dt \geq \begin{cases} 0 & (l \geq 0) \\ l \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = \sqrt{2\pi} l & (l < 0) \end{cases}$$

以上より, $a(l) > -\infty \cdots (9)$ $E_1(l, m, n) > -\infty$. 同様にして,

$$E_2(l, m, n) = \int \int \int_{y>x, y>z} y\phi(x-l)\phi(y-m)\phi(z-n)dx dy dz > -\infty$$

$$E_3(l, m, n) = \int \int \int_{z>x, z>y} z\phi(x-l)\phi(y-m)\phi(z-n)dx dy dz > -\infty$$

ゆえに,

$$E(P_1) = \sum_{i=1}^3 E_i(l, m, n) > -\infty \cdots (10)$$

一方,

$$E_1(l, m, n) = \int \int \int_{x>y, x>z} x\phi(x-l)\phi(y-m)\phi(z-n)dx dy dz$$

$$\leq \int \int \int_{x>y, x>z, x>0} x\phi(x-l)\phi(y-m)\phi(z-n)dx dy dz$$

$$\leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_0^{\infty} x\phi(x-l)\phi(y-m)\phi(z-n)dx dy dz$$

$$= \int_0^{\infty} x\phi(x-l)dx$$

$b(l) = \int_0^{\infty} x\phi(x-l)dx$ とおくと,

$$a(l) + b(l) = \int_{-\infty}^{\infty} x\phi(x-l)dx = l$$

(9) より, $b(l) = l - a(l) < \infty$ よって, $E_1(l, m, n) < \infty$

同様にして, $E_2(l, m, n) < \infty, E_3(l, m, n) < \infty$

ゆえに

$$E(P_1) = \sum_{i=1}^3 E_i(l, m, n) < \infty \cdots (11)$$

(10),(11) より, $(E(P_1))^2 < \infty$

従って (8) を用いて, $Var P_1 = E(P_1^2) - (E(P_1))^2 < \infty$