

## Bayesian approach の探索効率改善手法\*

6Q-05

丸山 孝志, 鶴岡 慶雅, 横山 大作†

東京大学

工学系研究科‡

近山 隆§

東京大学

新領域創成科学研究科¶

## 1 はじめに

現在、ゲーム木の探索手法としては、探索深さと局面の静的評価の 2 つを利用する方法が主流である。しかし、この従来手法では、現在探索している局面が元の現局面にどのくらい影響を与えようであるかが全くわからない。そのため、どこまで深く探索すればよいのか、あるいは、どのノードから展開していくべきなのか、といったことは基本的に判断できない。

そこで、現在探索している末端局面が元の現局面に与える影響の度合いを定式化し評価することで、上記の事柄を考慮した探索手法がいくつか提案されている。このような手法の 1 つが Bayesian approach と呼ばれる手法である。この手法の特徴的な点は、局面の評価値を確率分布で表し、各局面の評価誤差は独立であるという仮定の下に、その確率分布を根ノードに向かって伝播させていくという点で、伝播によって更新された局面評価の確率分布に基づき、次に展開すべき末端を決定する手法である。

しかし、実際のゲームにおいては探索空間中の近傍局面間には相関性があり、評価誤差は独立ではないと考えられる。そこで本稿では、簡単な予備実験により、近傍局面間の評価誤差が独立ではないことを確認した後、エキスパートプレイヤーのゲームから得られる統計的データに基づき、相関性を考慮した Bayesian approach の探索効率改善手法を提案する。

## 2 Bayesian approach

Bayesian approach とは、E.B.Baum と W.D.Smith が提案した手法で、ノードの評価値は真の評価値の確率分布を持っているというモデルを採用していて、この葉ノードの確率分布をベイズの定理を利用して根ノードまで伝播させる [1][2]。

まず、確率分布を持つ評価値の伝播手法を簡単に説明し、その後その問題点を述べる。

## 2.1 確率分布を持つ評価値

あるノードの静的なスカラー評価値と、さらに深く探索した際に得られるスカラー評価値とを比較することを考える。深く探索すればするほど、評価値は正確になるはずだから、静的な評価値に対して、深く探索した評価値を近似的に真の評価値と考えることができる。よって、両者の差はその局面での評価誤差を表しているといえる。

ここで、似たような局面における評価誤差の性質が似ていることは容易に予想できるので、何らかの特徴により分類した局面ごとに、より深く探索した場合との評価誤差を繰り返し測定することで、各分類局面ごとの最適な誤差分布を求めることができる。

これを利用すると、実際の個別局面の確率分布を持つ評価値は、その局面が属する分類局面の誤差分布を、その局面の静的な評価値の分だけ補正したものとして得ることができる。

## 2.2 確率分布の伝播

さて、上記のように求められた各ノードの確率分布を持つ評価値は、根ノードに向かって伝播される。実際の確率分布は計算しやすいようにディラックのデルタ関数の和の形で近似表現するが、この場合の葉ノードの確率分布が根ノードまで伝播され、根ノードの確率分布が決定される考え方を図 1 に示す。

ここで重要な点は、各ノードの評価誤差は独立した確率事象であるという仮定をおいている点である。図 1 の (a) で示されたような葉ノードの確率分布が与えられた場合、この仮定により、各葉ノードで取り得る評価値の全組合せを考えて、その各組合せに対してスカラー評価値を通常の実数法により伝播させて (b) ~ (e)、得られたスカラー値にその組合せの確率を重み付けして統合すればよいことがわかる。(f)

なお、ノードの評価値の分布をその確率によって重み付け平均したもの、つまり、分布がデルタ関数で表されている場合には期待値がそのノードを評価する値と考えられるので、上記のように探索木全体の確率分布が求められれば、根ノードの子ノードのうちで最も期待値が高いノードを選べば、根ノードでの最善手の

\*Improving search efficiency of Bayesian approach

†Takashi Maruyama, Yoshimasa Tsuruoka, Daisaku Yokoyama

‡School of Engineering, The University of Tokyo

§Takashi Chikayama

¶School of Frontier Science, The University of Tokyo

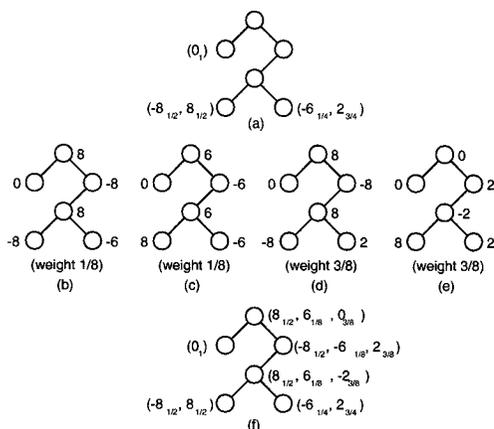


図 1: 確率分布の伝播

選択ができることになる。

## 2.3 問題点

Bayesian approach では、評価誤差の学習段階において分類局面ごとの相関性は当然織り込まれることになる。しかし、各ノードの評価誤差は独立であると仮定しているため、実際には存在するはずの、探索空間中の近傍に位置するノード間の相関性は全く考慮されていない。

## 3 近傍局面間の相関性の測定

### 3.1 実験手法

ここで簡単な予備実験により、探索空間中の近傍に位置する局面間に相関性を測定する。

選択したある局面の全ての子ノード、つまり、兄弟ノードにおける評価誤差の分布を測定し、相関性の有無を確認する。評価誤差としては、その局面の静的評価値（静止評価値）と、さらに深く探索した際に得られる評価値（近似的な真の評価値）の差を用いた。

なお、評価関数は、当研究室で開発しているコンピュータ将棋プレイヤー「激指」の評価関数を利用した。この評価関数は、駒得、駒の効率、玉の堅さ、手番などの要素から算出される一般的な評価関数であり、100点 が歩 1 枚分に相当する評価値となっている。

### 3.2 結果

当然、選択した局面によって様々な評価誤差の分布が得られたが、その一例として図 2 を示す。この結果は、駒がぶつかっている典型的な中盤局面から得られたもので、ゲーム中に何度も現れるような局面である。

もし、評価誤差に相関性がない場合には 0 を中心と

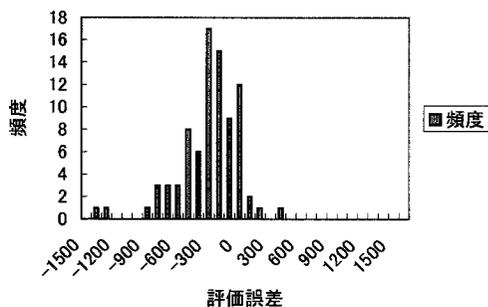


図 2: 評価誤差の分布例

した正規分布が得られるはずだが、この例では約 -300 を中心とした分布になっている。すなわち、兄弟ノード間の評価誤差に相関性があることを示している。

## 4 提案手法

実験の結果より、相関性がないという仮定は誤りであることがわかった。そこで、兄弟ノード間の評価誤差の相関性を組み込むことができれば、探索効率の改善が期待できるので、確率分布を伝播させる際に、確率に修正を加えることで相関性を組み込む手法を提案する。

この手法を簡単な例を使って説明する。今 2 つのデルタ関数の和で表される確率分布を持つ 2 つの兄弟ノードがあり、これを親ノードに伝播することを考える。確率分布の伝播計算に際して、片方の子ノードが取り得る 2 つの評価値のうち低い方を取る場合には、もう一方の子ノードにおいても同様に低い方の評価値を取る確率が高くなるように適切に確率を修正することで、相関性を組み込むことができるはずである。

今後は、確率分布の修正の適切なモデルを構築して、実験によりその有効性を確認する予定である。

## 参考文献

- [1] E. B. Baum, W. D. Smith, "A Bayesian approach to relevance in game playing", Artificial Intelligence, Vol. 97, No. 1-2, pp. 195-242, 1997.
- [2] W. D. Smith, C. Garrett, E. B. Baum, R. Tudor, "Experiments with a Bayesian game player", <http://www.neci.nj.nec.com/homepages/eric/>.