ゲーム情報学 2－3
(2000. 3. 17)

ゲーム木探索における兄弟ノードを母集団とする偏差値による探索深度制御

本堂　敦、鈴木　豪、小谷善行
東京農工大　情報コミュニケーション学科
{kagyuu, go, kotani}@fairy.ei.tuat.ac.jp

概要

木探索において、兄弟ノードを母集団とする偏差値を用いてそのノードに連なる枝の探索の深さの制御を行う手法を提案する。偏差値を用いることによって、この手法はゲームの序盤・中盤・終盤のどこにでも適用することが出来る。また、兄弟ノードとの相対的な位置関係を探索深度の制御に繁栄させることが出来るようになる。

本稿では、この手法を将棋の木探索において用いることを考える。偏差値と探索の深さの関係は、遺伝的アルゴリズムによって求めた。そして、学習により求められた関係と既存の探索延長アルゴリズムの比較を行った。

A CONTROL METHOD OF THE DEPTH OF TREE-SEARCH USING DEVIATION VALUE

HONDOH Atushi, SUZUKI Tsuyoshi, KOTANI Yoshiyuki
Dept. of Comp.&Comm. Sci., Tokyo Univ. of A&T
{kagyuu, go, kotani}@fairy.ei.tuat.ac.jp

Abstract

I suggest the method to control depth of tree-search using deviation value. To use deviation value, we can apply same prescription to the early stage, the middle stage and the final stage of game. And this makes relative position in brother nodes reflects depth control.

In this paper, I take this method to Shogi. To regulate relationship between deviation value and search depth, I use Genetic Algorithm. And then I compared this new tree-search method with existing.

1. はじめに

将棋における木探索について、どの枝をどのくらいの深さまで探索すればよいかということについての研究は、過去に数多くなされてきた。既存の木探索の探索深度の制御は、あらかじめ決めておいた探索打ち切り深度になったら探索を終了して、そのノードを末端と見なして静的評価値を行い、その結果得られた評価値をその枝の評価値とする。しかし、このように決められた探索打ち切り深度で

探索を終えるような方法を用いた場合、各枝の特徴を細かく反映するような柔軟な探索延長を行うのは難しい。

本稿では、木探索を行う場合に、そのノードの兄弟ノードを母集団とする偏差値を用いてそのノードに連なる枝の探索延長を行う手法を提案する。そして、将棋のゲーム木において、偏差値と探索延長の深さの関係を得る学習を行った。

2.「通行税」と「所持金」による探索延長

柔軟な木探索延長を行うために、木探索に「通行税」と「所持金」という概念を導入する。これらは、既存の木探索手法において、最初に「深さ」を0としておいて、ノードを通過するたびに「深さ」を足していき「深さ」がある一定の値を超えたら探索を終了するのと逆になる。すなわち、最初にある一定の「所持金」を持って木探索を始めて、次のノードに移るときに、行き先のノードの特徴のよって決まる値である「通行税」を払っていき「所持金」が0より小さくなったら探索を終了する。

図1は、「所持金」と「通行税」による探索制御の例である。まず、所持金=100 で現在のノードであるk1から探索を始める、k1から派生するノードk2、k3を調べるときに、所持金からk2、k3の特徴によって決まる「通行税」を支払う。k2に行くときの「通行税」を50、k3に行くときの「通行税」を75とすると、k2を調べる段階での「所持金」はk1の時点での「所持金」100から「通行税」50をひいた100-50=50となり、k3を調べる段階での「所持金」はk1の時点での「所持金」100から「通行税」75をひいた100-75=25となる。以下、同様にして、「通行税」をひきながら「所持金」を減らしていくと、k5・k6・k7・k8・k9を調べる段階で、「所持金」が0未満になるので、これらのノードを末端と見なす。そして、このようにして得られた末端の静的評価値をMinMax探索（$\alpha\beta$木探索）を使って各々の枝の評価値に集約していって、現在ノードの次にどのような手を指したらよいかどうかを決定する。

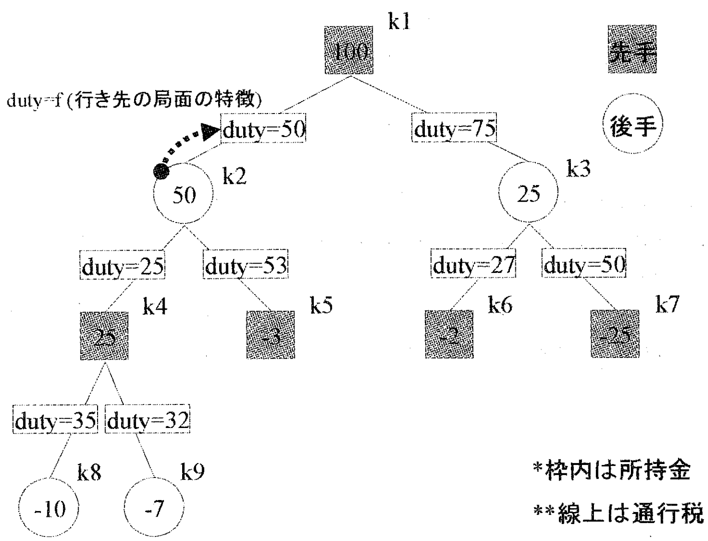

図1:通行税による探索深度制御の例

3. 偏差値による探索の延長

偏差値は、標準化スコアの一種である。標準化スコアは、n個の値からなる集合Vの要素v_i(i=0,1,2,・・・,n-1)の要素について、どんなVについても平均と標準偏差が一定になるように、v_iを操作したものである。標準化スコアを用いることにより、異なる条件下で得られたの各々要素v_iの全体に対する相対的な位置を知ることができる。偏差値は、標準化スコアの中でも平均が50、標準偏差が10に成るようにv_iを操作するものであり、以下のようにして求めることが出来る。

$$\text{平均}\quad \mu = \frac{\sum_{i=0}^{n-1} v_i}{n}$$

$$\text{分散(平均との差の2乗の平均)}\quad \sigma^2 = \frac{\sum_{i=0}^{n-1}(v_i-\mu)^2}{n}$$

$$\text{標準偏差(平均との差の絶対値の平均)}\quad \sigma = \sqrt{\sigma^2}$$

$$z\text{スコア(平均0.0　標準偏差1.0)}\quad z_i = \frac{v_i-\mu}{\sigma}$$

$$\text{偏差値(平均50　標準偏差10)}\quad d_i = z_i \times 10 + 50$$

このような偏差値の性質を利用をして、「通行税」を与える「ノードの特徴」にそのノードの兄弟ノードを母集団とする偏差値を使う。以下この延長手法を、「偏差値延長」と呼ぶ。

「偏差値延長」の利点の１つ目は、ゲームの進行によらずに使うことが出来ることである。例えば、将棋の場合、前半・中盤・後半と局面が進むに従って全体的に局面の評価値が異なってくるが、偏差値を利用することによって、全ての局面について、兄弟ノードのなかで、平均より悪い選択肢であるか・平均的な選択肢であるか・平均より良い選択肢であるかを統合的に扱えるようになる。

また、「偏差値延長」を用いることによって、ノードの駒取りに関する性質を探索延長に組み入れることが出来る。子ノードの偏差値のばらつきについて考えてみると、ノードの評価値は敵の駒を取ることによって大きく増えるので、偏差値のばらつきは次のような３つのパターンに分かれると考えられる。

(a)全ての子ノードが敵の駒を取らないときには、ノードの評価値はほぼ同じで、従って子ノードの偏差値は平均(50)近辺に固まることになる。
(b)ほとんどの子ノードが敵の駒を取らずに、少しの子ノードが敵の駒を取る場合には、敵の駒を取らない子ノードの偏差値(50)は平均より少し小さくなり、一方で敵の駒を取る子ノードの偏差値は非常に大きくなる。
(c)ほとんどの子ノードが敵の駒を取る場合には、取る敵の駒の価値によって、ノードの評価値は増えるため、子ノードの偏差値は、平均的に散らばる。

偏差値を「通行税」を与える「ノードの特徴」に使うことによって、このようなノードのばらつきの性質すなわち駒取りに関する性質をMinMax探索の探索深度の制御に反映させて、それぞれのノードの

状況の即した適当な探索深度の決定が出来るようになる。

4. 偏差値と通行税の対応の学習

本稿では、将棋における偏差値と通行税の関係を学習により得ることを考える。使用した将棋のベンチマークプログラムは、局面評価を持ち駒の価値の合計のみで行い、全幅探索を行い、取り合い延長を行う基本的なものを用いた。

学習には、遺伝的アルゴリズムを用いて、決められた1手当たりの総探索ノード数の中でどのような偏差値と通行税の関係が適当であるかを得ることとした。実験に利用したベンチマークプログラムの木探索のおける探索ノード数を調べる予備実験を行ったところ、1手当たりの総探索ノード数はおおむね10万ノードでの最大は49万ノードだった。そこで、本研究における1手当たりの総探索ノード数の上限は50万ノードとすることとした。

具体的な偏差値と通行税の関係は、偏差値{35、40、45、50、55、60、65}の代表7点に対応する通行税とした。木探索に用いる場合は、これら代表7点の間を線形一次補間したものを用いる。

個体数は、24 個体の群を3組とした。パラメータの初期値は、15（7手読みに相当）から、65(2手読みに相当)の範囲の乱数とした。遺伝的アルゴリズムにおける適応度は、群の中の毎回異なるランダムな相手と、12試合行った勝ち数とした。この適応度を元に、下位1/2を淘汰した。また、一方が適応度について上位1/2、もう一方が上位1/4に入るように両親を選んで交叉により交配させることにした。交差点は1点でその位置の決定には乱数を用いた。交配の際に、1%の頻度で、突然変異を起こしパラメータに-10から10の乱数を加えることとした。

また、4世代に1回、適応度が上位1/6の個体をそれぞれの隣の群に移すことで、局所解に陥る可能性を減らす様にした。

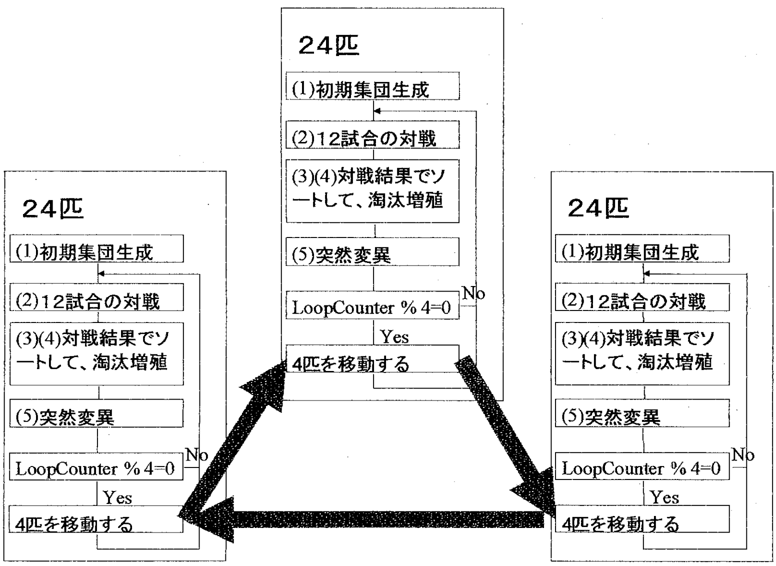

図2:学習部の実装

5. 学習結果

偏差値と、通行税の関係の学習を遺伝的アルゴリズムにより最適化してみた結果、50世代後には、ほぼグラフ5のように収束していた。50世代の各々の個体の各々のパラメータに付いて平均を取ってみると表5の様になった。さらに50世代の偏差値と通行税の対応の平均値に関して、1手目で偏差値がn（通行税m）の場合の平均探索手数を考えてみる。すなわち2手目以降が全て平均的な偏差値50（通行税28)のノードであったと仮定してみた場合、偏差値nのノードに対しておおむね

[（100 - m） ÷ 28] +1 （[a]はaを超えない最大の整数)

手読むと言うことが分かる。

その結果、評価値の悪いノード（偏差値40以下）には4手読みを行い、評価値が平均より少し劣るノード（偏差値45)には3手読みを行い、評価値が平均的なノードと平均より少し良いノード（偏差値50から55）には4手読みを行い、評価値が良いノード（偏差値60）には5手読みを行うが、評価値の偏差値65以上の非常にノードについては4手読みで十分というように学習が行われたことがわかる。

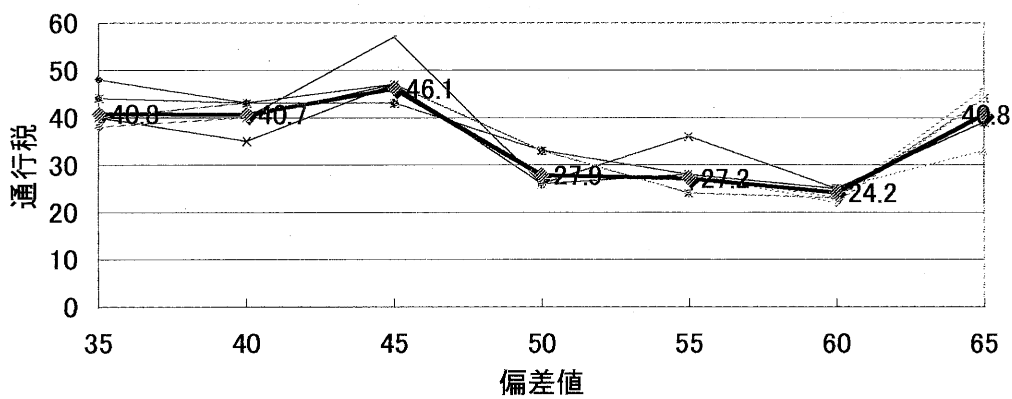

図3:50世代の偏差値と通行税の対応

偏差値	35 以下	40	45	50	55	60	65 以上
通行税	41	41	46	28	28	24	40
標準的な探索手数	4	4	3	4	4	5	4

表1:50世代の偏差値と通行税の対応の平均値

得られた偏差値と通行税の関係を使ったアルゴリズムと、既存の探索延長アルゴリズムを対戦させてみたところ、表2のような結果を得た。

対戦相手 ＼ 偏差値延長	先手	後手	合計
ベンチマークプログラム	58勝／100戦	55勝／100戦	113勝／200戦
ベンチマークプログラム+0.5手延長	55勝／100戦	57勝／100戦	112勝／200戦
ベンチマークプログラム+シンギュラー延長	49勝／100戦	56勝／100戦	105勝／200戦
合計			340勝／600戦

表2：既存の探索延長アルゴリズムとの対戦結果

互角なアルゴリズムが 100 試合対戦した場合の勝ち数の標準偏差 σ は 5 だから、100 試合行って 50 ± 2 σ (40 勝以上 60 勝以下)ならば、対戦したアルゴリズムは互角である(200 試合の場合は 86 勝以上 114 勝以下・600 試合の場合は 276 勝以上 324 勝以下)。すなわち表 2 からは、既存の探索延長アルゴリズムに比べて、本実験で得た探索延長アルゴリズムが強いとはいえない。

また、この対戦に関して末端ノード数(表 3)と探索の深さの頻度(図 4)を調べた。その結果、偏差値を使った探索延長手法は、ベンチマークプログラムの約4倍のノードの探索を行い、おおむね設定した50万ノード内で探索が行われるよう学習されたことが分かった。探索深度も偏差値を利用した延長では、3手読みが少なくなって、他の探索延長プログラムと比べた場合に、より深く読んでいることが分かる。

	ベンチマークプログラム	ベンチマークプログラム +0.5手延長	ベンチマークプログラム +シンギュラー延長	ベンチマークプログラム +偏差値延長
末端ノード数	49032716	79679693	123958469	197013619
ベンチマークプログラムとの比率	1.00	1.63	2.53	4.02

表3：既存の探索延長アルゴリズムの総末端局面数の比較(100対局)

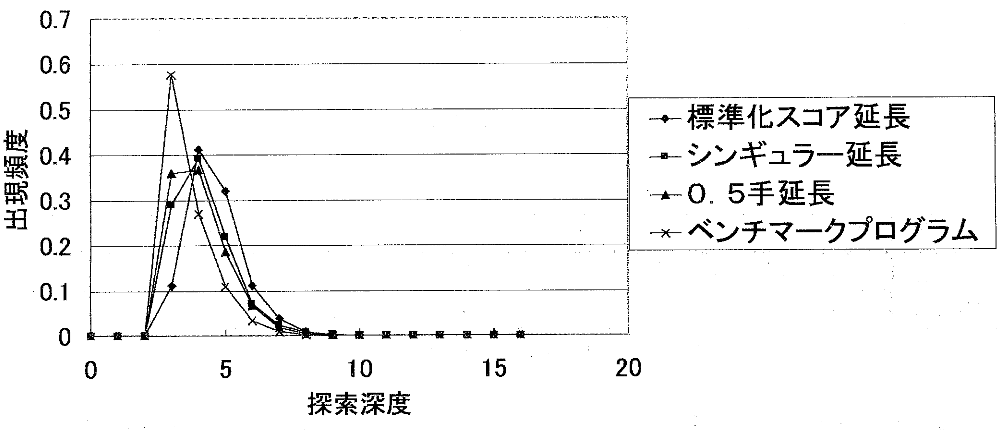

図4：探索延長アルゴリズムの探索深度

6. 考察

遺伝的アルゴリズムによる学習実験で得られた、偏差値と通行税の対応)を見て特徴的なのは、

(1)評価値の悪いノード（偏差値40以下）には4手読み

(2)評価値が平均より少し劣るノード（偏差値45）には3手読み

(3)価値が平均的なノードと平均より少し良いノード（偏差値50から55）には4手読み

(4)評価値が良いノード（偏差値60）には5手読み

(5)評価値の偏差値65以上の非常にノードについては4手読みで十分

ということである。

(1)は、非常に高い価値の駒が取れる場合、ある程度駒の取れる可能手が多い場合に、全く駒を取れないようなノードに相当する。このような場合、価値の高い兄弟ノードに連なる枝の評価値は一般的に(1)のようなノードに連なる枝の価値よりも高いので枝刈りが成されて(1)を比較的深く読むとしても探索ノード数はさほど増えないであろう。

(2)は、ある親ノードにおいて、駒が取れる可能手がたくさんある時に価値の小さな駒がとれる子ノードに相当するものである。このようなノードの場合、（少なくとも平均より）価値の大きな駒を取った方が良いであろうから、この枝を深く読まなくても良いと考えられる。

(3)は、最も頻繁に出現する偏差値であり、探索の基本的な深度になるので、限られ総探索ノード数（本研究の場合50万ノード）をなるべく使い切るような数になっているものと考えられる。

(4)は、兄弟ノードのほとんどで駒が取れる場合に比較的価値の低い駒が取れる場合に相当する。このようなノードは、評価値が安定せず、またこのノードで取った比較的価値の低い駒を利用できるので深く呼んだ方がよいと考えられる。

(5)は、評価値の大きい駒を取るか、その兄弟ノードの中でこのノードだけが駒を取るような場合に相当する。このようなノードを通る道は最終的にMinMax探索の結果を与える場合が多い。また、仮に偏差値延長による探索終了後に状況が大きく変わる場合でも取り合い延長によって評価値が是正されるので、さほど深く探索しなくても良いと考えられる。

(4)、(5)に関する学習結果は、なるべくそのプレーヤー（先手・後手）にとって良いノードを深く読むのがよいとされてきた今までの従来の考え方と異なる結果であった。学習結果は、兄弟ノードの中で評価値のずば抜けて高いノードはあまり深く読まなくてよく、むしろ評価値が平均より少し良いようなノードを深く読んだ方がよいことを示唆している。

7. まとめ

本研究は、通行税と所持金の概念を提案して、通行税の決定にそのノードの兄弟ノードを母集団とする偏差値を利用する手法を提案した。また、将棋において、その対応関係を学習によって得る実験を行った。その結果、シンギュラー延長や0.5手延長と同等の強さを持つ探索延長手法を得ることが出来た。

今後の課題としては、偏差値に用いるノードの評価値として、そのノードについて浅い深さの探索を行った結果を利用する事によってよりそのノードの特徴を正確に木探索の延長に反映させることが考えられる。また、駒の移動の仕方や、取る駒の種類、そのノードの評価値の順位なども偏差値と同

様に重要なノードの特徴であるから、それらと通行税の対応についても調べることが考えられる。

また、この手法は、将棋に特化したものではなく、ゲーム木探索に広く利用できるものである。従って、この手法をオセロやチェス・以後などの他のゲームにも適用することが可能である。

参考文献

[1]小谷善行、吉川竹四郎、柿木義一、森田和郎、 コンピュータ将棋、 サイエンス社、 1990.

[2]D．リービ、Mニューボーン共著・小谷善行監訳・飯田弘之、 乾伸雄、吉村信弘共訳、 コンピュータチェス、 サイエンス社、 1994.

[3]松原 仁、 コンピュータ将棋の進歩、 共立出版、 1996.

[4]松原 仁、 コンピュータ将棋の進歩２、 共立出版、 1998.

[5]山下 宏、 ０．５手延長アルゴリズム、 Game Programming Workshop in Japan '97、 1997、 46-54.

[6]Michael Buro、 PROB CUT : AN EFFECTIVE SELECTIVE EXTENSION OF THE $\alpha-\beta$ ALGLRITHM、 ICCA Journal 18(2)、 1995、 71-76.

[7]中川義久、 Singular Extensionにおける着手の性質、 Game Programming Workshop in Japan '97、 1997、 38-45.

[8]北野宏明編、 遺伝的アルゴリズム、 産業図書、 1992.