ゲーム木探索で得られる形勢判断の信頼性の推定とその応用

03110412　亀甲 博貴

東京大学工学部電子情報工学科
近山・鶴岡研究室

2013年2月18日

背景

コンピュータゲームプレイヤの研究

- コンピュータに人間同様の知性を持たせる試み
- ルールが明確で評価が容易

二人零和有限確定完全情報ゲーム

- オセロ・チェス・将棋・囲碁 etc..
- 先読み $\longrightarrow$ コンピュータと相性がよい
- リーフ局面に評価値を与え，ゲーム木探索によって評価値が最大になる手を探す

探索の性質

探索で得られる評価値が高い $\stackrel{?}{=}$ 自分に有利

- 評価値は常にある程度の不確かさがある
- 先読みしきれない数手先で突然不利になるかも

評価値の高い局面が常に盤石とは限らないのでは

$\longrightarrow$ 評価値の信頼性の推定を目指す

提案手法

以下の手法を提案する

- 評価値の信頼性の定義と推定
- 探索への応用
  - ProbCut への応用
  - 勝負手生成への応用

探索へ応用することで，棋力の向上が期待できる

将棋プログラム「激指」を用いて手法の有効性を示す

評価値の信頼性の定義

浅い探索で得られる評価値の信頼性を考える

深い探索での評価値に近い　：信頼できる

深い探索での評価値から遠い：信頼できない

評価値差がどのような分布を取るかを評価値の信頼性と定義する

浅い探索と深い探索の評価値差の分布

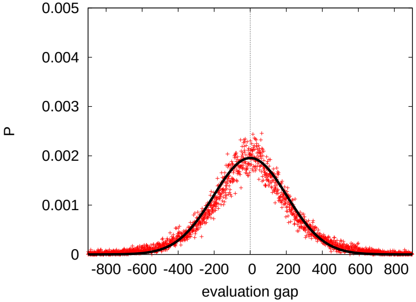


Figure: 中盤の局面

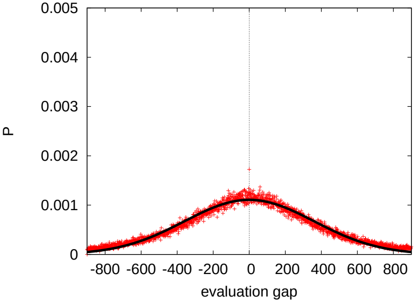


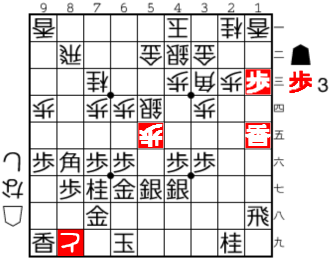
Figure: 終盤の局面

期待値 $\mu = 0$ の正規分布で近似する
$\rightarrow$ 標準偏差を求めれば分布がわかる

進行度

ゲームの進行具合を示す．激指においては

- 0-127 の整数値
- 初期局面は 0 で，大きいほど終局に近い
- 駒が自陣からどれだけ遠いか（5 段以上）
駒が成っているか
持ち駒であるか
で計算



左図の赤い駒が進行度に強く影響する

進行度 45

標準偏差と進行度の関係

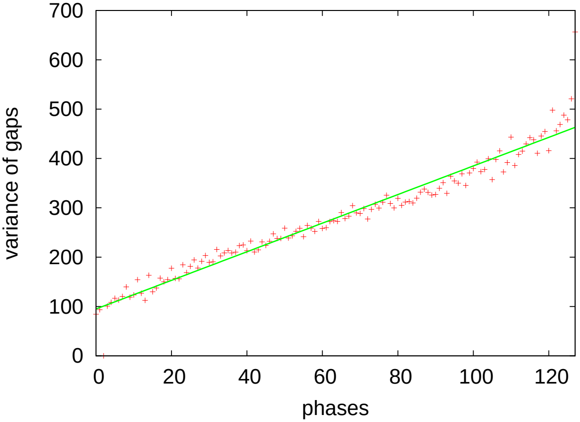


Figure: 進行度のみ

縦軸：標準偏差　横軸：進行度

標準偏差の推定

$$
\begin{aligned}
\boldsymbol{\sigma}_i &= \boldsymbol{w}^T \phi(\boldsymbol{X}_i) \\
\boldsymbol{w} &: \text{重みベクトル} \\
\boldsymbol{X}_i &: \text{注目する局面} \\
\phi(\boldsymbol{X}_i) &: \text{局面 } \boldsymbol{X}_i \text{ を表す特徴ベクトル}
\end{aligned}
$$

最尤推定により$\boldsymbol{w}$を得る

$$
\begin{aligned}
\boldsymbol{likelihood}(\boldsymbol{X}_i \mid \boldsymbol{w}) &= \frac{1}{\sqrt{2\pi\boldsymbol{\sigma}_i^2}} \exp\left(-\frac{\boldsymbol{g}_i^2}{2\boldsymbol{\sigma}_i^2}\right) \\
\boldsymbol{g}_i &: \text{評価値差}
\end{aligned}
$$

用いる特徴

- 盤面上の駒の行位置と種類
- 持ち駒中の駒の種類と枚数
- 盤面上の駒がピンしている・されている
- 互いの玉の周りの駒
- 盤面上の駒の列位置と種類
- 互いの玉の周囲のマスに効いている駒の数は敵と味方でどちらがどれだけ多いか

標準偏差の推定

推定された標準偏差で整列し，データを 128 分割
各データ群ごとに推定された標準偏差の平均と実際の
標準偏差をプロット

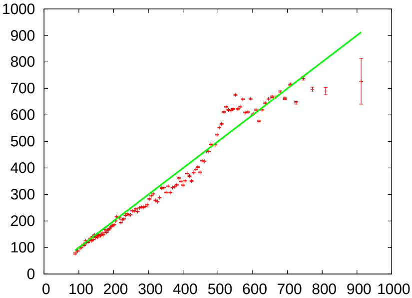


Figure: 進行度の各要素

縦軸：標準偏差　横軸：推定された標準偏差

標準偏差の推定

推定された標準偏差で整列し，データを 128 分割
各データ群ごとに推定された標準偏差の平均と実際の標準偏差をプロット

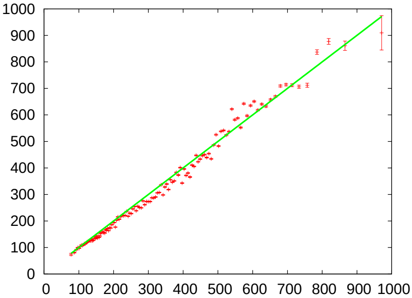


Figure: 示した特徴全て

縦軸：標準偏差　横軸：推定された標準偏差

用いた特徴と尤度の関係

用いた特徴	学習率	訓練データ	開発データ
進行度	0.005	-7.367	-7.363
進行度要素	0.005	-7.074	-7.071
提案手法	0.005	-7.054	-7.053

対数尤度の平均
特徴を増やすと尤度が上昇している
＝データをよく表現できている

$\alpha\beta$ 探索と ProbCut

$\alpha\beta$ 窓

評価値がこの範囲に含まれないなら明らかに他の指し手に劣っている範囲

劣っていることが証明されるから枝刈り

$\longrightarrow$ $\alpha\beta$ 探索

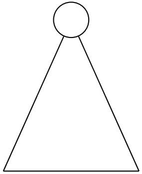
証明されてないけど高確率で劣っている手を枝刈り

$\longrightarrow$ ProbCut

ProbCut [Buro, 1995]

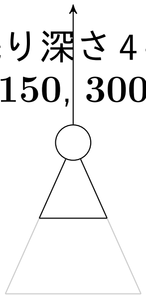
残り深さ 8 手

(150, 300)

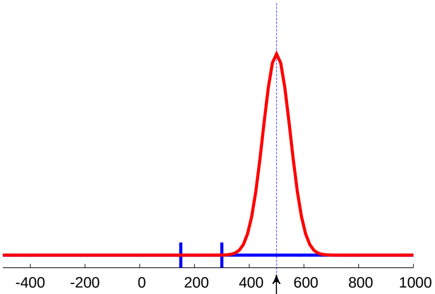


ProbCut [Buro, 1995]

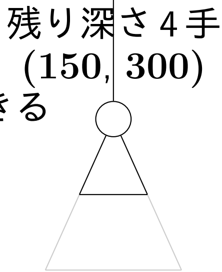
浅い探索を試す
→ 評価値 500



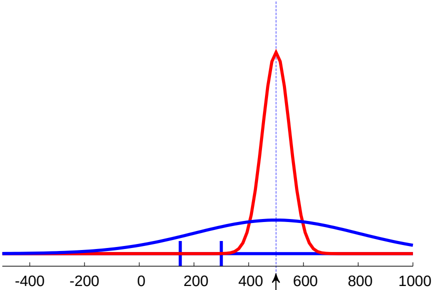
ProbCut [Buro, 1995]



浅い探索を試す
→ 評価値 500
分布が狭い：枝刈りできる

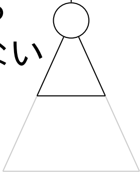


ProbCut [Buro, 1995]

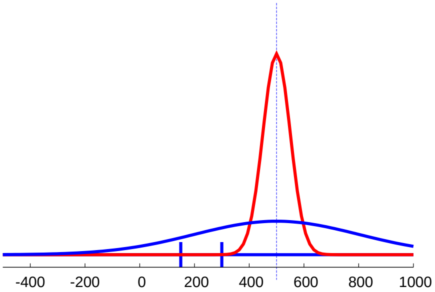


浅い探索を試す
→ 評価値 500
分布が狭い：枝刈りできる
分布が広い：枝刈りできない

残り深さ 4 手
$(\mathbf{150, 300})$



ProbCut [Buro, 1995]



深い探索で得られる評価値の分布を知る必要がある

ProbCut の実装

深さ 8 以上 11 未満のノードが対象
浅い探索で得られる評価値が閾値からマージン以上離れていたら枝刈り
マージンに推定された標準偏差 σ を用いる

ProbCut：探索ノード数比較

1,000 局面でプロの指し手と比較

手法	一致数	時間（s）	ノード数	ノード数/秒
ProbCut 無し	574	1,427	399M	279,978
進行度のみ	569	1,285	361M	280,957
提案手法（1σ）	570	1,244	346M	277,973
提案手法（2σ）	574	1,364	385M	281,966
提案手法（3σ）	572	1,411	399M	282,902

探索時間が短縮できている

ノード数を削減できている

棋力を落とさずに ProbCut に成功している

ProbCut：探索ノード数比較

1,000 局面でプロの指し手と比較

手法	一致数	時間（s）	ノード数	ノード数/秒
ProbCut 無し	574	1,427	399M	279,978
進行度のみ	569	1,285	361M	280,957
提案手法（1σ）	570	1,244	346M	277,973
提案手法（2σ）	574	1,364	385M	281,966
提案手法（3σ）	572	1,411	399M	282,902

探索時間が短縮できている
ノード数を削減できている
棋力を落とさずに ProbCut に成功している

ProbCut：探索ノード数比較

1,000 局面でプロの指し手と比較

手法	一致数	時間（s）	ノード数	ノード数/秒
ProbCut 無し	574	1,427	399M	279,978
進行度のみ	569	1,285	361M	280,957
提案手法（1σ）	570	1,244	346M	277,973
提案手法（2σ）	574	1,364	385M	281,966
提案手法（3σ）	572	1,411	399M	282,902

探索時間が短縮できている
ノード数を削減できている
棋力を落とさずに ProbCut に成功している

ProbCut：対戦による棋力比較

対戦プレイヤ	1 手	結果	勝率
提案手法（3σ） vs. ProbCut 無し	3s	478-426	0.529
提案手法（1σ） vs. 進行度のみ	3s	439-457	0.490
提案手法（2σ） vs. 進行度のみ	3s	449-443	0.503
提案手法（3σ） vs. 進行度のみ	3s	441-481	0.478

5%有意水準は約 51.6%

ProbCut を行わない相手に対しては有意に勝ち越し

進行度を用いた ProbCut とは有意な棋力差見られず

棋力差が現れるほど枝刈り量の差がない

ProbCut：対戦による棋力比較

対戦プレイヤ			1手	結果	勝率
提案手法（3σ）	vs.	ProbCut無し	3s	478-426	0.529
提案手法（1σ）	vs.	進行度のみ	3s	439-457	0.490
提案手法（2σ）	vs.	進行度のみ	3s	449-443	0.503
提案手法（3σ）	vs.	進行度のみ	3s	441-481	0.478

5%有意水準は約51.6%
ProbCutを行わない相手に対しては有意に勝ち越し
進行度を用いたProbCutとは有意な棋力差見られず
棋力差が現れるほど枝刈り量の差がない

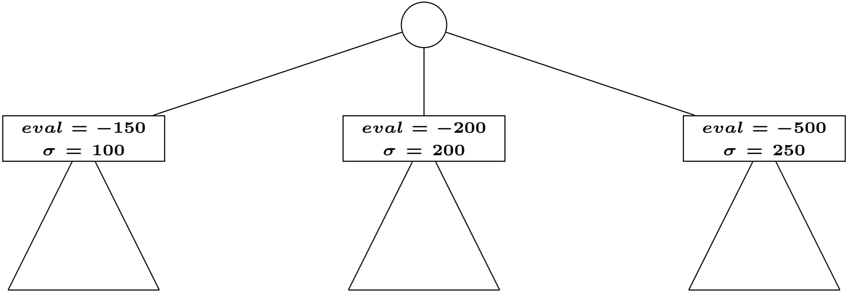
勝負手の定義

探索で得られた評価値は他の手に劣っていても，先読みしていない所で評価値が良くなる確率の高い指し手

数手先で評価値が正になる確率を指標とする

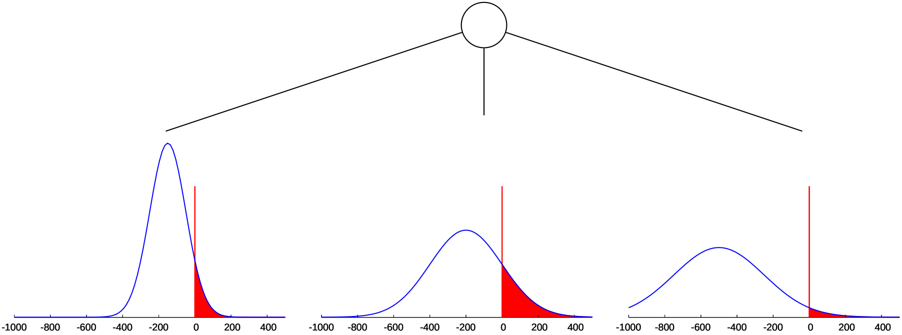
- 不利な局面 $\longrightarrow$ 逆転を目指す勝負手
- 有利な局面 $\longrightarrow$ 逆転を許さない堅実な手

勝負手の生成手法



1段目の各ノードをルートにして部分木を作成
各部分木で通常探索
それぞれの部分木は数手先の評価値が正になる確率を評価値として返す

勝負手の生成手法



1段目の各ノードをルートにして部分木を作成
各部分木で通常探索
それぞれの部分木は数手先の評価値が正になる確率を評価値として返す

勝負手生成：評価方法

- 棋譜を一定手数読む
- 1 手目の通常探索と勝負手生成探索をそれぞれ深さ 12 で 1 回行い，局面を 2 つ作る
- それぞれ残りの対戦を深さ 6 で行う
- 結果に違いが出たかを確認する

同じ 1,000 棋譜を用いるが，一定手数未満の棋譜は除外する

勝負手生成実験

80 手目の後手が勝負手を指す
先手から見た評価値

初期手数	初期局面の評価値	敗 → 勝	勝 → 敗	変化なし
79	〜−1000	5	3	3
79	−1000〜−500	16	14	61
79	−500〜0	33	40	138
79	0〜500	50	67	143
79	500〜1000	36	31	103
79	1000〜	15	33	79
79	合計	155 (17.6%)	188 (21.4%)	537

多くの局面で結果を負け → 勝ちに変えることに成功
同様に多くの局面で勝ち → 負けになってしまっている
初期局面の評価値ではその 2 つを区別できない

勝負手生成実験

120 手目の後手が勝負手を指す
先手から見た評価値

初期手数	初期局面の評価値	敗 → 勝	勝 → 敗	変化なし
119	〜−1000	0	1	34
119	−1000〜−500	7	6	26
119	−500〜0	10	8	36
119	0〜500	6	8	50
119	500〜1000	9	9	38
119	1000〜	12	6	65
119	合計	44 (13.3%)	38 (11.5%)	249

79 手に比べて結果が変わる局面数が少ない
適用するのが遅いと逆転が難しくなる

まとめ

評価値の信頼性の指標として探索深さ間の評価値差の標準偏差を用いることを提案し，これを推定した．また ProbCut，勝負手生成への応用手法を提案した．

- 標準偏差の推定をよい精度で行えた
- ProbCut に実装して対戦による有意な棋力差は見られなかったものの，ノード数の削減量における有効性を確認した
- 勝負手生成に応用して，一定数の局面で勝敗を変えることに成功した

今後の課題

- 今回使用した特徴以外の特徴の検討
- Multi ProbCut への拡張
- 勝負手生成を適用する局面の決定方法

尤度・目的関数・勾配

$$
\begin{aligned}
likelihood(X_i) &= \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{g_i^2}{2\sigma_i^2}\right) \left(\sigma_i = w^T\phi(X_i)\right) \\
\log likelihood(w) &= \log \prod_i \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{g_i^2}{2\sigma_i^2}\right) \\
&= \log \prod_i \frac{1}{\sqrt{2\pi(w^T\phi(X_i))^2}} \exp\left(-\frac{g_i^2}{2(w^T\phi(X_i))^2}\right) \\
&= \sum_i \log \frac{1}{\sqrt{2\pi(w^T\phi(X_i))^2}} \exp\left(-\frac{g_i^2}{2(w^T\phi(X_i))^2}\right) \\
&= \sum_i \left\{-\frac{g_i^2}{2(w^T\phi(X_i))^2} - \log(w^T\phi(X_i)) - \log\sqrt{2\pi}\right\} \\
\mathrm{grad} \log likelihood(w) &= \frac{\partial \log likelihood(w)}{\partial w} \\
&= \sum_i \frac{\partial}{\partial \sigma_i}\left\{-\frac{g_i^2}{2\sigma_i^2} - \log\sigma_i - \log\sqrt{2\pi}\right\}\frac{\partial \sigma_i}{\partial w} \\
&= \sum_i \left\{\frac{g_i^2}{\sigma_i^3} - \frac{1}{\sigma_i}\right\}\phi(X_i)
\end{aligned}
$$