

モンテカルロ木探索プレイヤーにおける評価値へのノイズ付加による影響

1200319 佐原 和就 【高度プログラミング研究室】

1 はじめに

モンテカルロ木探索は、コンピュータ囲碁や将棋をはじめとしたゲーム研究における探索手法の一つである。モンテカルロ木探索をより効果的にする手法として、局面評価やプレイアウトにおける評価関数の併用が広く用いられており、評価値の正確さがモンテカルロ木探索の動作に何らかの影響を及ぼすと考えられる。

本稿では、評価関数により算出された評価値に意図的にノイズを付加することにより、評価値の正確さによるモンテカルロ木探索の正確さを検証する。

2 本研究の調査内容と先行研究

2.1 先行研究

先行研究 [2] では、オセロプログラム Zebra における評価パターンをいくつか除去することで精度の低い評価関数を作成し、評価関数の違いによるモンテカルロ木探索への影響を研究している。結果は、評価パターンを除去すると勝率が従来の結果より低い勝率となっている。

プレイアウトへ評価関数を用いることで、プレイアウト中の良いとされる手が選ばれやすくなり、ある程度強いレベル同士でのプレイアウト対戦となる。

UCB1 値の算出過程に評価関数を用いることにより、局面評価において良いとされる手が有望な手と判断されやすくなる。

2.2 本研究の調査内容

本稿では、プレイアウトと UCB1 値で評価関数を用いることで算出された評価値に対し、意図的にノイズを付加することにより、精度の低い評価関数を作る。また、ノイズの大きさによっても何かしらの影響が現れる可能性があるため、分散を変えたノイズで検証を行った。

2.3 対戦相手と試合数

対戦相手は十分な強さを持つアルファベータ法プレイヤーを用いるものとする。また、試合数は 100 試合行う。

2.4 ノイズの種類

ここで用いるノイズは正規分布に従うノイズを使うものとする。Zebra では評価値が 512 で 1 コマに相当するため、ノイズの大きさは標準偏差 $\sigma = 1000$ (分散 $=\sigma^2$) から $\sigma = 10000$ まで 1000 単位で変更させながら対戦させる。また、ノイズの σ 値を局面の手数が増えるにつれ収束させた場合でも対戦させる。

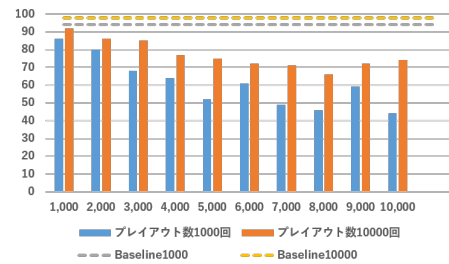


図1 UCB1 値とプレイアウトの評価値にノイズを付加した場合

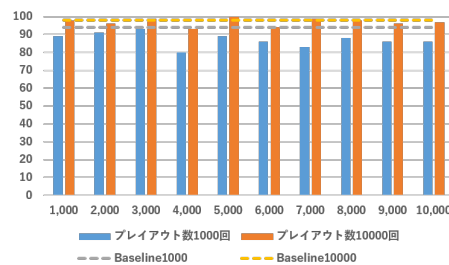


図2 手数が進む毎にノイズの分散が収束した場合

3 結果

実験結果、ノイズを付加することによる勝敗への影響は確認でき、ノイズの分散が増加するにつれ勝率が右肩下がりという結果が確認できた。しかし、 $\sigma = 9000$ 以上になると勝敗の大きな変化はみられなかった。

ノイズの σ 値を収束させた場合のモンテカルロ木探索では、上記のような大きな影響は終始確認できなかった。

4 まとめ

実験結果より、ノイズを付加することで勝敗への悪影響が確認された。また、 σ 値を大きくすることにより、より勝敗への影響が大きくなることもわかった。ある一定の σ 値を境に勝敗への影響がみられない原因として、ノイズが及ぼす悪影響の限界ではないかと考えられる。収束させた場合は、ゲーム前半に影響が及ぼされても終局近くで正しい評価値に収束していることで大きな影響は確認できなかったものと考えられる。

参考文献

- [1] Gunnar Andersson: Zebra, <http://radagast.se/othello/index.html>.
- [2] 北村直輝, 松崎 公紀: 評価関数の違いがモンテカルロ木探索プレイヤーの強さに与える影響. 第 59 回プログラミング・シンポジウム予稿集, pp 173-183, 2018.