

コンピュータ囲碁におけるだまし構造探索の改良

D-8 Improvement of Deceptive Structure Search in Computer Go

内田 環太郎 浅香 緑

Kantaro UCHIDA Midori ASAKA

群馬大学大学院理工学府

Graduate School of Science & Technology, Gunma University

1. はじめに

コンピュータ囲碁の探索法にはモンテカルロ木探索が用いられている[1]. その中で最も強い手法は UCB (Upper Confidence Bound) アルゴリズムを応用した UCT (UCB applied to Tree) 探索である[2]. しかし UCT 探索は, だまし構造(少数の手が他の手とかけ離れて良い盤面)に弱点を持つ[1]. 本研究は, UCT 探索に各盤面の平均勝率を保持した代理ノードという仮想的な候補手を追加する. だまし構造のような盤面では代理ノードの選択割合が低くなる特性を持つ. この特性を利用しつつ, さらに UCB 判定式を改良することで, だまし構造を持つ候補手の探索優先度の増加をおこなった.

2. モンテカルロ木探索

モンテカルロ木探索は, ランダムに終局まで進め勝敗を得ること(プレイアウト)を数多くおこなうことで, 各候補手の勝率を得る手法である. より良い勝率の手で探索を進めるために, UCB アルゴリズムが応用されている. 以下に UCB 判定式(1)を記す.

$$\bar{x}_j + C \sqrt{\frac{2 \log n}{n_j}} \quad \text{式(1)}$$

式(1)において \bar{x}_j は j 番目の候補手の平均勝率, C は0~1の定数, n は盤面全体のプレイアウト数, n_j は j 番目の候補手のプレイアウト数である. UCB 判定式が最大となる候補手を次にプレイアウトする候補手として選ぶ.

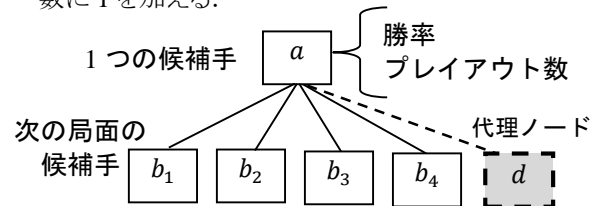
3. 提案手法

3.1 代理ノードの機能

だまし構造の有無を容易に判断するために代理ノードを導入する. 代理ノードは各候補手の暫定的な平均勝率を保持する. UCB 判定式による勝率重視の探索をおこなった場合, だまし構造のような盤面では代理ノードがあまり選択されない. 代理ノードの機能は以下のとおりである.

- ①探索する候補手の次の候補手において, 1 回ずつプレイアウトを終えたとき代理ノードを生成する(図 1).
- 図 1 において a を探索する候補手, b_i を次の局面の候補手, d を代理ノードとする.
- ②各ノードは勝率とプレイアウト数を保持する. 代理ノードは代理ノード生成時の a の勝率・プレイアウト数として閾値を保持する.
- ③UCB 判定式により代理ノードが選択された場合, プレ

イアウトはおこなわず, 代理ノードの保持するプレイアウト数に 1 を加える.

図 1 次の局面の候補手 b_i と代理ノード d

3.2 UCB 判定式の改良

だまし構造を持つ候補手により早く探索を進めるために, UCB 判定式に式(2)を追加する. 式(2)は代理ノード選択割合が小さいほど大きい値を取る. 以下に式(2)を記す.

$$C_1 \left(\frac{1}{i+1} - \frac{n_{dairi}}{n_j + n_{dairi}} \right) \quad \text{式(2)}$$

式(2)において C_1 は0~1の定数, i は次の盤面の候補手数(図 1 における b_i の数), n_{dairi} は代理ノードの保持するプレイアウト数である. この式は代理ノードが 1 度選択された次の UCB 判定時から適応する. すべての探索を終えたときは 1 番勝率の良い候補手を実際に打つ手として選ぶ.

4. 評価

だまし構造を持つ盤面において各プレイアウト数で, だまし構造の探索割合を取得した(図 2). UCT に比べ, だまし構造を持つ候補手により多く探索がおこなわれる.

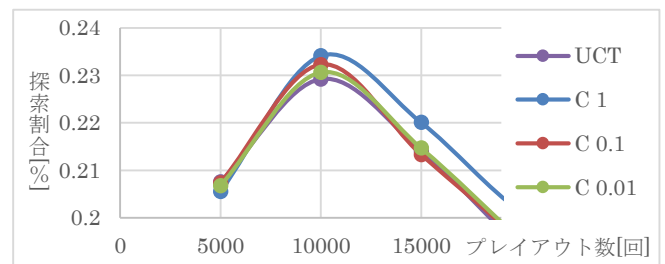


図 2 だまし構造を持つ候補手の探索割合

5. 今後の課題

だまし構造が複数続くような盤面では, 改良式が逆効果をもたらす場合がある. その対策をおこなう.

参考文献

- [1] 松原・美添・山下, “コンピュータ囲碁 モンテカルロ法の理論と実践,” 共立出版, 2012.
- [2] L. Kocsis and C. Szepesvari, “Bandit based Monte-Carlo planning,” In 17th European Conf. on Machine Learning, pp. 282-293, 2006.