Title	有望さに基づくUCTの選択的探索手法とMonteCarlo囲碁への応用
Author(s)	矢島，享幸
Citation	
Issue Date	2011-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/9649
Rights	
Description	Supervisor:飯田 弘之，情報科学研究科，修士

Plausibility based Selective Search for UCT and its Application to Monte-Carlo Computer Go

Takayuki Yajima (0810063)

School of Information Science,
Japan Advanced Institute of Science and Technology

February 8, 2011

Keywords: computer Go, Monte-Carlo method, UCT, expansion condition.

Computer programs have already won to the top human player in Othello and chess. Although the classical method using the static evaluation functions and the minmax search has succeed a lot in these games, it does not perform sufficiently in the game of Go. The reasons why is that the search space of Go is quite larger than other classical games, additionally, the difficulty of the knowledge expression disturbs programmers in designing an accurate evaluation function.

In 2006, a Go program CrazyStone, which was based on MCTS (Monte-Carlo Tree Search), appeared and changed the situation. MCTS made it easier to develop a strong Go program than before. Instead of the static evaluation function, CrazyStone used random simulations (playout) to evaluate positions. It also combined a tree search algorithm with the simulations in order to find the best move in a given position.

MCTS has been widely used for Go, and been intensively studied in recent years. In MCTS algorithms, UCT (Upper Confidence bounds applied to Trees) is mainly used today, because if sufficient time is given, it guarantees to reach the same result as the minmax search would reach. UCT algorithms follow nodes which have the best expectations to a leaf node and do one or more playouts to the end. If the leaf node is visited more than a certain threshold, the node is expanded and its child nodes are appended to the search tree.

The search part and the MC part are widely studied in concerning the improvements of both quality and efficiency. For the MC part, there are well-known improvements such as RAVE (Rapid Action Value Estimation) or use of patterns, so that this thesis does not handle them. For the search part, pruning techniques have been mainly studied. However, these pruning techniques have a problem that a certain number of playouts is required before they bring the benefit, hence their effect is slight in the early stage of the search progress. In this thesis, we focus on the condition of **node expansion**, and tackle to the problem aiming for a more efficient search.

The most popular condition for node expansion is whether the node is visited more than a constant threshold. First, we examined the relation between the threshold and the performance of computer Go player. Experimental results show that a large threshold inhibits growth of the search tree, and a small threshold promotes the growth too much as reaching a limit of memory capacity. Hence, the best expansion threshold is reconfirmed to exist in the middle of them.

Next, we proposed three approaches to grow the search tree only around plausible moves for the search efficiency. These were based on *transition probability*, *prominent winning rate* and *visiting number estimation*. We also evaluated the effectiveness of these approaches.

Experimental results show that the effect of proposed approaches begin to appear from the relatively earlier stage of the search progress, compared to the pruning techniques. In the comparison of playing skill, the growth methods based on *transition probability* and *prominent winning rate* respectively show significantly better performance than the pruning techniques.

In this thesis, we achieved a certain level of search efficiency by modifying the node-expansion condition dynamically, but we are still in the way to establish the universal expansion method especially for a small threshold. On this research as the foundation, we would like to examine techniques to realize more effective UCT by controlling the search depth more flexibly.