

Title	ボードゲームの着手評価関数の機械学習のためのパターン特徴量の選択と進化
Author(s)	Nguyen, Quoc Huy
Citation	
Issue Date	2014-09
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/12287
Rights	
Description	Supervisor:池田 心, 情報科学研究科, 博士

氏名	NGUYEN QUOC HUY		
学位の種類	博士(情報科学)		
学位記番号	博情第 306 号		
学位授与年月日	平成 26 年 9 月 24 日		
論文題目	Selection and evolution of pattern features for supervised learning of an action evaluation function in board games (ボードゲームの着手評価関数の機械学習のためのパターン特徴量の選択と進化)		
論文審査委員	主査	池田 心	北陸先端科学技術大学院大学 准教授
		飯田 弘之	同 教授
		白井 清昭	同 准教授
		Nguyen Minh Le	同 准教授
		Le Hoai Bac	同 客員准教授

論文の内容の要旨

The target of our research is the artificial intelligence of computer games. The the artificial intelligence makes computer games be more interesting. There are several areas of games that AI is contributed to, but we focus on two-player deterministic perfect-information games, because the rules of these games are simple, a lot of people know them, and these systems have been used as the simple testbed of AI research. Our goal is to make strong computer players of board games.

In general background, game tree search is frequently used for making computer game player. State evaluation functions are often necessary to evaluate a node, and action evaluation functions are also often used to enhance the search. Machine learning is quite effective to learn good evaluation functions automatically, compared to manual design, and it is already known that selection of features used for machine learning is very important for the performance.

In our specific case, we focus on the Monte Carlo Tree Search for Othello game as a game tree search, it needs a good action evaluation function using some patterns. We employ Bradley-Terry Minorization-Maximization as the learner, because it is effective in the game of Go and attracts attention. In our case, only the patterns are used as the features. But there are a lot of possible patterns, then working with pattern shapes is our motivation because we believe that there are many hidden good pattern shapes which have not been discovered.

From the reasons above, our purpose is to establish the way to obtain good pattern shapes automatically from given game records. Thus, the research questions are how to obtain better pattern shapes and how to reduce the optimization cost.

We want to optimize the pattern shapes, so two approaches are proposed to evaluate them. One is to use statistic method to measure the hopefulness of a PS before learning. Another is to use the very

early result of BTMM further with less learning data. Two approaches are our contribution, and they have some reasonable results.

The game of Othello is used as an environment of experiments in this thesis, but the methods can be applied to almost all board games such as Go, Connect-6, Lines of Action, Hex, Shogi, which patterns are used. From many experiments, we have some following conclusions.

Evolution of PSs is done not by using MCTS performance but by light-BTMM for reducing cost. If MCTS performance needs 100 games to evaluate a pattern shape and each game needs 3 minutes, then we need 300 minutes to evaluate a pattern shape. Instead of using MCTS performance, if the heavy-BTMM is used, then it needs 8 minutes to evaluate a pattern shape.

However, the light-BTMM needs nearly 1 minute to evaluate a pattern shape. It is clear that using light-BTMM reduces cost significantly. MLE of the optimized patterns is improved compared to the often-used patterns by evolution. Because the MLE of the often-used patterns is -1.5 before optimization, but the MLE of the optimized patterns is -1.4 after optimization. This means that we hope the MCTS performance will be improved if the optimized patterns are applied.

The MCTS performance itself is also improved a lot. The winning ratio between two MCTS programs using the often-used patterns is 50%. But the winning ratio is increased to 70% if the MCTS using the optimized patterns, compare with the MCTS using the optimized patterns.

There are many strange patterns after evolving, this means that it is difficult for programmers to optimize pattern features by manual.

Key words: Feature selection, pattern features, action evaluation function, supervised learning, board games

論文審査の結果の要旨

近年注目を集めるモンテカルロ木探索(MCTS)では、効率化のために行動(着手)の評価関数を援用することが多い。本論文は、その評価関数の機械学習に用いる中心的な特徴量である盤面のパターンについて、どの範囲のパターンを参照するかを自動で最適化する方法について研究したものである。パターンの重みの学習については多くの研究があるが、それらでは参照の範囲は人手で固定されており、範囲そのものを自動で最適化しようとする試みは新規で意欲的である。

本論文では、第二章でモンテカルロ木探索、第三章で評価関数の機械学習、第四章で確率的最適化についての定式化と既存研究の紹介を行ったのち、第五章ではパターン範囲群の確率的最適化(evolution)、第六章では良いパターン範囲の低コストでの選別(selection)を提案した。これらの効果は著名なボードゲームであるリバーシを対象とした

実験を通して確かめられた。

パターン範囲群最適化の目的は MCTS による強いプログラムを作ることである。しかし、最適化のために強さを対戦等から計測することは非常に高コストである。そこで本研究では、(1)機械学習の結果が良いほど MCTS も強くなる (2)少ない学習データ・少ない繰り返し数の機械学習（軽い学習）の結果が良いほど本来の機械学習の結果も良くなる、という仮定をおいた。そのうえで、パターン範囲群の良さを対戦を通してではなく、「軽い学習」の評価関数値を用いて計測し、Simulated Annealing による最適化を行った。その結果、仮説(1)(2)が妥当であり、人手で設計するのが困難な高度なパターン範囲が得られ、MCTS の強さも大きく向上することが確認できた。

Simulated Annealing で生成されるランダムに変異したパターン範囲の中には、明らかに有望でないものも含まれる。これらは「軽い学習」で評価する前に、より低コストな指標によって評価して除外したい。本研究では、パターン範囲の良さを予測するためのいくつかの統計量に着目し、それらを統合した指標を提案した。その上で、この予測により良いと評価されたパターン範囲群が実際にそれ以外のものよりも学習結果・MCTS の強さ両方において優れていることを確認した。

以上、本論文は、ボードゲームに頻繁に登場する盤面パターンという特徴量について、その参照範囲を自動かつ高速に最適化するための手法を提案、効果を実証した。本研究の枠組みは MCTS やリバーシのみならず多くの探索手法・対象ゲームに適用可能であると思われる。ゲーム情報学の世界に大きな貢献をしたものであると判断できる。よって博士（情報科学）の学位論文として十分価値のあるものと認めた。