

要 旨

大貧民のモンテカルロ法プレイヤーにおける シミュレーション精度向上に関する研究

地曳 隆将

近年，大貧民の研究においてモンテカルロ法を適用したプレイヤーの研究が行われている．モンテカルロ法プレイヤーは，現在の盤面からゲーム終了までのシミュレーション（プレイアウト）を複数回行うことで各手役の評価値を算出する．プレイヤーの強さはプレイアウトの精度を向上させることで強化できる．

本研究では，大貧民モンテカルロ法プレイヤーのプレイアウトの精度を向上させるために 2 つの手法を取った．

1 つめの手法は，プレイアウト開始時に生成する相手手札を実際の盤面に近づけることである．大貧民は不完全情報ゲームであるため，相手プレイヤーの手札を知ることができない．そのため，プレイアウト開始時に生成した相手プレイヤーの手札と，実際の盤面の相手プレイヤーの手札の差がプレイアウト精度を下げる要因となる．プレイアウト開始時に生成した相手プレイヤーの手札を実際の盤面に近づけることでプレイアウト精度が向上する．

2 つめの手法は，プレイアウト中の各プレイヤーのプレイ方策を，実際のプレイヤーのプレイ方策に近づけることである．最も単純なモンテカルロ法である原始モンテカルロでは，プレイアウト中の各プレイヤーのプレイは乱数によってランダムに行われる．そのため，各プレイヤーのプレイとは提出手役に差ができる．各プレイヤーが提出する手役を真似できるようになればプレイアウト精度が向上する．

本研究では，この 2 つの手法が大貧民モンテカルロ法プレイヤーに与える影響を調査した．

1 つめの手法では，モンテカルロ木探索プレイヤーによって「最善手」を算出し，手札情報

の異なるモンテカルロ法プレイヤーが算出した「最善手と判断した手」と比較した。これによって、手札情報の精度がモンテカルロ法プレイヤーに与える影響を調査した。調査の結果、手札推定を行うことでモンテカルロ木探索プレイヤーの「最善手」の評価値にモンテカルロ法プレイヤーの評価値（モンテカルロ木探索プレイヤーの「最善手」と同じ手の評価値）は近づくことがわかった。一方、モンテカルロ法プレイヤーが「最善手と判断した手」にはほとんど影響は無く、「最善手」の推定には寄与していないことがわかった。

2つめの手法では2つの実験を行った。1つめの実験では、3層ニューラルネットワークを用いた評価関数の性能を調査した。棋譜でプレイヤーが提出した手役と、評価関数を用いて得られる手役の一致度を調査した。調査の結果、学習に使用する盤面データを増やすことで提出手役の一致率が上昇し、盤面データ数 15,000 程度で一致率が頭打ちになることが確認できた。盤面データ数 15,000 の評価関数では、未知の盤面に対する提出手役一致率がおよそ 69%となった。2つめ実験では、評価関数を適用したモンテカルロ法プレイヤーの性能を調査した。モンテカルロ法プレイヤーのプレイアウト中の提出手役を、3層ニューラルネットワークを用いた評価関数によって選択するプレイヤーを提案し、原始モンテカルロプレイヤーと比較実験を行った。調査の結果、原始モンテカルロプレイヤーよりも評価関数を適用したモンテカルロ法プレイヤーの方が強いことが確認できた。

キーワード 大貧民, モンテカルロ法, モンテカルロ木探索, 3層ニューラルネットワーク

Abstract

Improvement of Simulation Accuracy in Monte Carlo Daihinmin Players

Takamasa ZIBIKI

In recent studies of Daihinmin, the Monte-Carlo method has been applied to Daihinmin players. Monte-Carlo player calculates an evaluation value of each meld by performing a lot of playouts, which are random simulations to the end of the game. The strength of Monte-Carlo players depends on the accuracy of the playouts.

In this study, we took two approaches in order to improve the accuracy of the playouts of Daihinmin Monte-Carlo players.

In the first approach, we estimate the opponents' hands, from which playouts are performed, as close to the real ones as possible. Since Daihinmin is a game with imperfect information, it is impossible to know the opponents' hands. The difference between the estimated hands and the real ones would become a factor that reduces the accuracy of playouts. If we can estimate the opponents' hands more precisely, we can improve the accuracy of playouts.

The second method is to approximate the play strategies in the playouts. The simplest Monte-Carlo method, called primitive Monte-Carlo, performs playouts just with random numbers. The difference between the random plays and the real ones would become a factor that reduces the accuracy of playouts. If we can mimic the play strategies of the players, we can improve the accuracy of playouts.

In this study, we investigated the effects of these two approaches on Daihinmin

Monte-Carlo players.

For the first approach, we compared the melds computed by Monte-Carlo players with different hand informations with the best hand computed by Monte-Carlo Tree Search player with perfect information. We found after this experiments that partial information of opponents' hands may improve evaluation values but the effects on selecting the best hand are small.

The two experiments were performed for the second approach. The first experiment investigated the performance of the evaluation function developed with a three-layer neural network. We investigated how often we can select the real melds played in the game records using the evaluation function. It was confirmed that the concordance rate of submission hand combination increases by increasing the training data to be used for learning, and concordance rate saturates over 15,000 training data. In the evaluation function learned from 15,000 training data, submitted hand role concordance rate for the unknown boards has become approximately 69%. The second experiment investigated the performance of the Monte-Carlo player applying the evaluation function. The proposed player selected the meld computed with the three-layer neural network in the playouts. We compared the proposed Monte-Carlo player with the primitive Monte-Carlo player. After the experiments, we confirmed that the Monte-Carlo player with the evaluation function is stronger than the primitive Monte-Carlo player.

key words Daihinmin, Monte-Carlo Method, Monte-Carlo Tree Search,
Three-Layer Neural Networks