

平成 30 年 6 月 28 日現在

機関番号：55201

研究種目：基盤研究(C) (一般)

研究期間：2012～2017

課題番号：24500192

研究課題名(和文) 名人を超える将棋序盤アルゴリズムの開発

研究課題名(英文) Development of Shogi opening algorithm that exceeds the Meijin

研究代表者

橋本 剛 (HASHIMOTO, TSUYOSHI)

松江工業高等専門学校・情報工学科・准教授

研究者番号：40420335

交付決定額(研究期間全体)：(直接経費) 4,000,000円

研究成果の概要(和文)：コンピュータ将棋は一般に終盤が強いが、序盤にはまだ弱点が多い。プロは良い「形」の評価ができるが、コンピュータは多くの駒位置パターンからなる「形」を評価できていない。「形」を理解するプロ並みの序盤を実現するには多くの駒位置からなる評価が必要になる。だが評価要素が増えるとメモリと時間が指数的に増大するため、実現は難しい。本研究では多駒を使った高頻度のパターンを評価させるため、「乱数を用いる評価項目自動抽出法」を提案し、高頻度の評価要素を自動で抽出することに成功した。だが、実用化にはかなりの高速化が必要であることがわかり、新たなデータ構造を提案し一定レベルの高速化に成功した。

研究成果の概要(英文)：Computer shogi is generally strong in the end game, but there are still weak points in the opening game. Although professional players can evaluate a good "shape", the computer can not evaluate "shape" consisting of many pieces position pattern. In order to realize professional-looking play in the opening game that understands the "shape", evaluation consisting of many pieces positions is required. However, as the number of evaluation factors increases, memory and time increase exponentially, which is difficult to realize. In this research, in order to evaluate a high frequency pattern using multiple frames, we propose "automatic evaluation item evaluation using random numbers" and succeeded in automatically extracting high frequency evaluation factors. However, it was found that considerable speeding up was necessary for practical use, we proposed a new data structure and succeeded in increasing the speed to a certain level.

研究分野：ゲーム情報学

キーワード：コンピュータ将棋 評価項目自動抽出法 序盤 機械学習

1. 研究開始当初の背景

チェスでは 1997 年に DEEP BLUE が世界チャンピオンに勝利したが、将棋は持ち駒がありチェスより遥かに複雑で、順調に強くなっているものの名人に勝てるレベルには達していなかった。そんな中、2010 年に情報処理学会のコンピュータ将棋システム「あから 2010」が清水市代女流王将と対戦し、コンピュータ将棋が公の場で初めてプロ棋士から勝利を上げ、大きな話題となった。申請者は序盤の作戦担当となり、対戦相手の棋譜を解析し、相手が苦手と思われる意表を突く作戦(4 手目 3 三角戦法)を立案し、成功を収めた。コンピュータ将棋はプロに近いレベルにまで達したが、名人と戦い勝利するまでにはまだまだ克服しなければいけない問題が残っていた。中でもコンピュータが最も苦手なのが序盤であった。終盤の実力は当時すでにプロも脅かす実力で、特に詰め将棋が得意であることは知られており、数百手の問題も一瞬で解け、人間ではまったく敵わないレベルに達している。だが、序盤はまだアマチュアレベルであった。プロレベルの序盤は駒の損得などのわかりやすい基準ではなく、長年の経験から得られる駒組みの微妙な相性など評価しにくい要素が重要で、コンピュータに理解させることが難しい。表 1 に当時のコンピュータ将棋の実力をまとめた。そこで、本研究は序盤を名人レベルに到達させる事を目指とした。

表 1: 研究開始当時のコンピュータ将棋実力

序盤	アマチュア二段程度
中盤	アマトップクラス
終盤	プロレベル
詰み	名人以上

2. 研究の目的

コンピュータ将棋の序盤は当時まだアマチュアレベルで、名人に勝つためには強い序盤力が必要であった。プロは良い「形」の評価ができるが、コンピュータは多くの駒位置パターンからなる「形」を評価できていない。評価関数の機械学習により中盤以降は強くなったが、「形」を理解するプロ並みの序盤を実現するには多くの駒位置からなる評価が必要になる。だが評価要素が増えるとメモリと時間が指数的に増大するため、実現は難しい。本研究では多駒を使った高頻度のパターンを評価させ、かつ実用のために高速化を実現することを目標とし、プロ以上の序盤が実現できるよう工夫をし、名人に勝つコンピュータ将棋の完成を目指す。

3. 研究の方法

コンピュータ将棋の序盤指し手をプロレベルにするため、「乱数」「いいかげん」な学習

をキーワードとする方法を考える。3 駒以上の位置評価による局面評価関数は容易に計算爆発を起こすため、まともに扱くと計算時間かメモリを使いすぎ、うまく行かないことが多いと予想される。だが、まともに計算しても対象が膨大すぎる時に、乱数を使って一見いいかげんで乱暴そうな方法がうまく行くことが良くある。この発想を、局面評価関数の駒位置組み合わせに使う。乱数を使ったいいかげん学習により、3 駒以上のパターンから教師データ(プロ棋譜から派生する局面)で出現頻度の高い組み合わせだけを登録、学習する局面評価関数を作成する。ただし計算時間が十分早く実用的なシステムにするため、研究期間中「修正 学習 自動対戦」を根気良く繰り返す。実装は申請者が開発する TACOS と、オープンソースの Bonanza で行う。

4. 研究成果

(1)「形」を理解するプロ並みの序盤を実現するには多くの駒位置からなる評価が必要になるが、評価要素が増えるとメモリと時間が指数的に増大するため、現在コンピュータ将棋でよく行われているように駒組み合わせをすべて評価する方法(all 型と呼ぶ)では実現し難い。この問題を解決するには、無駄な組み合わせを排除し重要な組み合わせのみを評価する方法(freq 型と呼ぶ)が必要である。現在では、freq 型を手動で設定しているプログラムはあるが、有効な自動獲得法は提案されていない。そこで乱数を用いた駒組み合わせの出現頻度カウントによる評価項目の抽出法、「評価関数自動抽出法」を考案した。出現頻度を調べるために対局中に現れる駒組み合わせの出現回数をカウントする必要があるが、全ての組み合わせをカウントすることは不可能なため、カウントする組み合わせはランダムに選択した。選ばれた駒組み合わせを分析した所、ほとんど動かない香や桂馬などの組み合わせがカウント回数上位となり、囲いなどに現れる駒の組み合わせはカウント回数中位に多く現れた。次に実際に抽出された項目を用いて freq 型評価関数を Bonanza に実装し、all 型のみのプログラムと対戦実験を行った。その結果、探索時間は多く必要としたものの、王と他の 1 駒を評価するプログラムには大きく勝ち越しその有用性を示すことが出来た。

コンピュータ将棋は、Bonanza メソッドの考案により非常にレベルが高くなった。現在では、プロ棋士との対局も盛んに行われており、トッププロレベルに迫っている。しかし、中盤以降に比べて序盤が弱いという弱点も抱えている。コンピュータ将棋の性能は評価関数の質に大きく影響される。正確に盤面評価を行うほどコンピュータ将棋は強くなる。評価関数のパラメータは機械学習により自動調整されているが、40 駒からなる将棋の盤面をそのまま評価するためには膨大な評価

項目が必要となる。現状では部分局面に分割することにより、近似的に評価を行なっている。しかし、現在の多くのプログラムは駒組み合わせ全てを保持するオール（all）型の評価関数を用いているため、3 駒程度の少ない駒数にしか分割できない。そのため序盤で特に重要な盤面の細かな違いを認識することが出来ない。高次元組み合わせ評価を行うためには、重要な駒組み合わせのみを評価するフレック（frequency）型評価関数が必要である。フレック型評価関数実現のためには、重要な評価項目を抽出する必要がある。しかし、現在重要な評価項目を自動抽出するための有効な手法は存在しない。そこで本研究では、乱数を用いた盤上の駒組み合わせのランダムカウントにより、評価項目を自動抽出する手法を提案する。実際に提案手法により抽出した項目を用いてフレック型評価関数を設計し、Bonanza に組み込んだ。そして、オール型評価関数と対局実験を行い、性能向上を確認した。

(2)

プロ並みの良い「形」を評価させることを目指し、「乱数を用いる評価項目自動抽出法」を提案し、多くの駒位置からなる評価要素を乱数を使い自動で抽出することに成功し成果を収めた。だが、提案手法を実装すると予想以上に探索に時間がかかり、まだ実用には至っていない。これは、駒組み合わせのデータがリストを用いた木構造に保存されており、評価値参照に非常に時間がかかるためである。また、木構造は差分計算に向いていないため、本質的に高速化に向いていない。そこで本年は高速化を目指して木構造ではない新しいデータ構造評価値参照の差分計算方法を考案した。その実装のため駒組み合わせ参照の新しい仕組みを提案する。新しいデータ構造に必要な条件は、1 つの駒の種類と位置から、それが含まれる駒組み合わせが高速に参照できること、一度に駒組み合わせ全てを取得できることである。そこで、駒組み合わせとその評価値を格納する評価値配列を作成し、評価値配列のインデックスをすべての駒に対応してまとめたインデックスリストを作成した。すべての駒に対してインデックスリストをあらかじめ作成しておくため、駒情報から評価値配列に高速でアクセスできる。この手法を実装し実験を行ったところ、一定レベルの高速化には成功したことが確認できた。

(3)

本研究ではコンピュータ将棋にプロ級の良い「形」を評価させることを目指し、「乱数を用いる評価項目自動抽出法」を提案し、多くの駒位置からなる評価要素を、乱数を使い自動で抽出することに成功し成果を収めた。また、プロ棋譜に現れる駒組み合わせをランダムに抽出したことにより、どのような組み

合わせが多く現れるかのデータが明らかになった。組み合わせ上位は桂や香などを含む、初期配置と差があまりないものが多かった。組み合わせ中位には、矢倉や美濃囲いなど頻出する囲いに関する要素が多く見られる、などの知見が得られた。同様の手法で、他のゲームなどへの応用が期待できる。

評価項目を自動で得る事ができる手法として、ATARI のゲームや alphaGO で有名になった深層学習を用いる強化学習が近年注目されているので、最終年度にはテトリス、弾幕シューティングを題材に評価項目を与えない深層学習を用いた強化学習を行った。その結果、単純な強化学習では膨大な学習時間が必要で難しいが、学習対象を小さいものから大きいものへと変化させることでうまく学習が出来るというアイデアを提案し、実験によりその効果を確認できた。

#### 5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕(計 0 件)

〔学会発表〕(計 5 件)

野村直也、橋本剛、弾幕シューティングゲームにおける汎用的機械学習法、第 39 回情報処理学会ゲーム情報学研究会、2018

青木勢馬、橋本剛、テトリスを題材にしたスケールダウンを利用した学習手法の開発、ゲームプログラミングワークショップ、2017

野津佑太、後藤嵩幸、橋本剛、将棋 AI における評価項目自動抽出法の高速化手法、ゲームプログラミングワークショップ、2016

品川和哉、橋本剛、ナポレオンを題材とした機械学習に有効な特徴の研究、第 65 回電気・情報関連学会中国支部連合大会、2014

後藤 嵩幸、橋本 剛、コンピュータ将棋における高次元組み合わせ評価のための評価項目自動抽出に関する研究、第 31 回情報処理学会ゲーム情報学研究会、2014

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

名称：  
発明者：  
権利者：  
種類：  
番号：  
出願年：  
国内外の別：

取得状況（計 0 件）

名称：  
発明者：  
権利者：  
種類：  
番号：  
取得年：  
国内外の別：

〔その他〕  
ホームページ等

#### 6. 研究組織

##### (1) 研究代表者

橋本 剛 (HASHIMOTO Tsuyoshi)  
松江工業高等専門学校・情報工学科・准教授  
研究者番号：40420335

##### (2) 研究分担者

( )

研究者番号：

##### (3) 研究協力者

( )