

駒の歩数にもとづく将棋の駒組み学習

佐藤 直人[†] 柴原 一友^{††} 但馬 康宏^{††} 小谷 善行^{††}

[†] 東京農工大学 工学部 情報コミュニケーション工学科 ^{††} 東京農工大学大学院 工学府

将棋においてコンピュータが駒組みを適切に行うために、人間の棋譜に出現する2つの駒の位置関係と駒の歩数（動かされた手数）を用いて駒組みを学習する手法を提案する。近年では評価関数のパラメータを機械学習により自動調整する手法がさかんに研究されているが、序盤に限っては人手によるパラメータの調整や定跡に頼っているといえる。評価関数で駒組みを行うのが困難である原因として、序盤は駒の損得が発生しにくく局面の明確な評価基準を設けることが難しい点が挙げられる。そこで本稿では、自陣の駒組みの良し悪しを、2つの駒の関係と“駒の歩数”という新しい概念を用いて人間の棋譜から学習する方法を提案する。その結果、囲いの良し悪しを判断するという点において実験では良い結果が得られた。

Learning Piece Formation Based on Piece Steps in Shogi

Naoto SATO[†] Kazutomo SHIBAHARA^{††} Yasuhiro TAJIMA^{††} Yoshiyuki KOTANI^{††}

[†] Department of Computer, Information and Communication Sciences,
Tokyo University of Agriculture and Technology

^{††} Department of Computer and Information Sciences, Tokyo University of Agriculture and Technology

We present the method to learn piece formation with the relations of pairs of pieces and piece steps in game-scores. In late years it is studied that method to adjustment parameter of evaluation function automatically by machine learning. However, piece formation depends on the opening theory and the adjustment of the parameter by hands. It is difficult to form the piece formation by using evaluation function because piece up or down is hard to occur in the opening, and to make a clear evaluation standard of the position. we suggest a method to learn the quality of its own piece formation from game-scores with relation of two pieces and a new concept “piece steps”. adjusted a parameter by learning piece steps and a relation of played piece and other piece in game-scores. As a result, a good result was provided by our experiments in a point to judge the piece formation.

1 はじめに

コンピュータ将棋において、与えられた局面から最善手を導くためにはその局面を評価する評価関数と、先読みを行う探索アルゴリズムの実装が必要である。現在、評価関数には玉の危険度や駒の自由度等多くの評価項目が設けられている場合が多いが、基本となるのは駒の損得である。末端局面において駒の取り合いが起こる場合には駒の取り合いが起こった後の損得を計算することが重要であり、1つのマスだけでなく複数のマスにおける駒の取り合いを評価する手法も研究されている¹⁾。しかし、駒の取り合いが発生しない序盤においては評価項目の設定および重み付けが困難なため、あらかじめ手の正しさが保証された定跡通りの手を指すように設計されていることがほとんどである。また、定

跡から外れた後の序盤戦略として、人間の持つ将棋の知識を評価項目として設定することが考えられているが、各評価項目の重み付けを適切に行うことは非常に大きな労力と試行錯誤が必要となる。

そこで、本稿ではコンピュータが駒組みを決定する方法として、盤上にある味方2つの駒の位置関係による価値と動いた駒の歩数を評価し、その評価値を最大化する手を選択する手法を提案する。駒の歩数とは、同一棋譜の中で一つの駒が動いた回数を指す。例えば、先手玉が5九から6八、7八へと移動した場合、7八にいる先手玉の歩数は2となる。駒の歩数を学習に用いた理由は、序盤において、動く前よりも動いた後の方がその駒自身の価値が高いと考えたことにある。なぜならば、将棋や囲碁のようなゲームでは手番を握っていること自体に価値

があり、その価値を行使して着手を行うのであるから、指された駒の価値は着手前よりも増していると考えられるからである。提案する手法が十分に効果的であったことを4章の実験で示す。

2 関連研究

棋譜学習を用いた現在までの研究として、序盤の駒組み抽出がなされている²⁾。この研究では玉の位置による勝率の違いや将棋の主な囲いの有用性を判定することに成功している。また、金子により駒の関係を利用して局面の評価を行った研究もなされており³⁾、定跡や人間の指し手を模倣することに成功している。この研究では、指し手の前後の局面の比較をもとに学習する親子モデルと指し手と他の候補手を比較して学習する兄弟モデルの2つの手法の比較分析も行われており、後者がより優れているという結果が得られている。実用的な評価関数の調整については、最適制御理論に基づいて兄弟モデルによりパラメータの重みを調整する手法⁴⁾が注目を集めている。兄弟モデルを用いた研究は十分な結果を得られているといえるが、親子モデルを用いた学習では金子の手法を超える研究報告は筆者の知り得る限りでは皆無である。金子の手法では一局を通して駒の損得のない静かな局面全てを対象として学習を行っているが、駒組みを行うことを目的とする序盤と玉を詰めることを目的とする終盤を、駒の関係という同じ評価項目で評価を行うことは必ずしも適切であるとはいえないと考える。なぜならば、序盤では、いくつかある候補手のうち、棋譜に現れた手が他の手に比べて圧倒的に良いとは断言できないからである。極端な例を出せば、初手7六歩が2六歩に比べて良いとはいえないだろう。したがって、序盤は兄弟モデルを用いて学習を行うべきではない。しかし、手を指した後の方が指す前よりも好形になっていることは1章で述べた通りである。そこで本稿では、駒の歩数という新しい概念を用いて序盤の2つの駒の関係を親子モデルで学習する手法を提案する。そして、本手法により正しい駒組みを手を学習することを目的とする。なお、駒を取る手が発生した後の局面においては、駒の関係の良し悪しにかかわらず手(相手陣に駒を打ち込む手やそれを防ぐ手等)を指す場合があるため、駒を取る手が発生する前の局面を学習対象とする。

3 学習アルゴリズムの設計

本章では、棋譜から駒組みを学習するためのアルゴリズムについて述べる。棋譜を1手ずつ読み込み、重みを更新することにより最終的な重みを定める。1局の棋譜中で、マス P_k に移動した駒 k の歩数が n_k ($n_k \geq 1$)回の局面において、マス P_i にある手番側の駒 i との位置による重み $w(P_k, k, P_i, i)$ の更新式を下式のように定める($w(P_k, k, P_i, i)$ を w と省略して表記する)。なお、手番側の指し手の合計を n 、学習率を α ($0 < \alpha \leq 1$)、手数による重みの逓減率を γ ($0 \leq \gamma < 1$)とする。

$$w \leftarrow w + \alpha(n_k \times (1 - \gamma)^n - w) \quad (1)$$

式(1)を、手番側の駒のある全てのマス P_i について行うことにより、その局面で動かされた駒と他の駒との、位置による重みを更新する。式(1)により、指された数の多い場所により高い価値が与えられるのと同時に、そのマスに駒が動かされた頻度と重みを学習率により制御することができる。また、歩数に逓減率 γ をかけることにより、棋譜の中で早く指された手の評価を高くしている。そうすることによって早く移動させるべき駒をより早く移動させることができる。これは、将棋プログラムYSSが採用している落とし穴方式⁵⁾と同じ原理であるが、YSSは人手で重み付けを行うのに対して提案手法は学習により重みを調整する点が異なる。式(1)により学習を行った結果をもとに実際に着手を決定するには、合法手集合に含まれる手の中でマス P_k に駒 k が移動したときの評価値 $E(P_k, k)$ が最大になるものを選択する。 $E(P_k, k)$ は以下の式で計算される。

$$E(P_k, k) = \sum_i w(P_k, k, P_i, i) \quad (2)$$

式(2)についても式(1)と同様に手番側の駒のある全てのマス P_i に対して計算を行う。これにより、各駒に対する位置による重みの総和が最も大きくなるような手を選択する。

4 実験結果

提案する手法の効果を検証するための実験結果を次節以降で示す。

4.1 実験条件およびパラメータの取得

学習に用いる棋譜は、将棋倶楽部24における棋譜48万局とした^{6) 7)}。棋力や戦型を考慮せず、全

での棋譜を学習に用いた。また、序盤の終了条件として駒を取る手が指された時点で終了とすることとした。この理由は2章の終わりで述べた通りである。加えて、学習回数を増やすため、後手の駒組みについても座標を点対称に変換することによって先手の駒組みとみなして学習を行った。学習の反復数は10回、式(1)における学習率 α は0.001から反復するたびに0.0001ずつ減少させた。また、減減率 γ は0.1に設定した。学習にはIntel(R)Core(TM)2 6700CPU(2.66GHz)で約8分を要した。

以上の実験条件により学習した重みの収束の様子を図1に示す。横軸は反復回数、縦軸は全ての重みにおける教師値($n_k \times (1-\gamma)^n$)と重み $E(P_k, k, P_i, i)$ の差の二乗和であり、この値を誤差とみなした。2回目の学習で誤差は大幅に減り、6回目の学習以降で誤差の大きな変化はないため、学習は収束しているといえる。

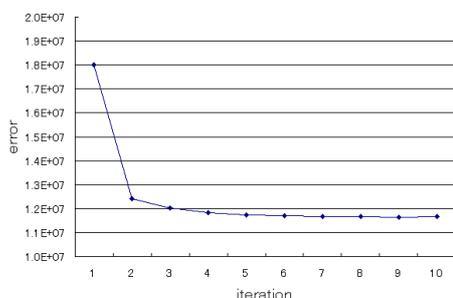


図1 反復数と誤差の変化

4.2 学習によって得られた重み

学習によって実際に得られた重みを、3種類の例をとりあげて紹介する。表1に各マスに対する玉の位置による重みを、表2に8八玉に対する各マスの金の位置による重みを、表3に9九玉に対する各マスの金の位置による重みを示す。

9	8	7	6	5	
0.769	0.929	1.025	0.585	0.382	八
1.052	0.640	0.524	0.407	0.086	九

表1より、玉の位置による重みについては、5九から遠ざかるたびに値が大きくなっており、最も高い点は穴熊の位置である9九となっていることから学習が成功しているといえるだろう。ここで、7

表2 8八玉に対する各マスの金の位置による重み

9	8	7	
0	0.011	0.025	七
0	玉	0.326	八
0	0	0.064	九

表3 9九玉に対する各マスの金の位置による重み

9	8	7	
0	0.006	0.098	七
0	0.183	0.464	八
玉	0.001	0.213	九

八の価値が8八よりも大きくなっているのは、7八にいる玉が8八へ行くためには8八にいる角をあらかじめ動かさなければいけないことと、7八から8八へ行くまでに他の駒を多く動かすため手数による減減が大きくなっているためと考えられる。

次に、表2より、8八玉に対する金の位置による重みも、7八が最も大きく、次いで7九、7七の順となっているため妥当と思われる。同じく表3についても穴熊の金の定位置である7九と7八の重みが高くなっているため妥当と思われる。

4.3 良形・悪形の判別

学習した重みが良形・悪形を判別できるかどうかを試す実験を行った。例題として用いたのは将棋プログラムKFEnd⁸⁾のページで掲載されている悪形チェックの問題であり、全て先手番である。学習した重みによる評価結果を述べる。

はじめに、例題1(図2)での悪手は3六飛、4六飛であるとされている。一方、評価結果は最善手が1六歩、次善手が4六歩となった。これら2つの手は好形であるだろう。したがって、学習した重みはこの例題において悪形を回避できている。

次に、例題2(図3)での悪手は7七金であり、7七銀が好形とされている。評価結果は、評価値の大きい順に7七銀、7七桂、7七金となった。よって、この例題についても学習した重みは悪形を回避できている。

最後に、例題3(図4)での悪手は3九玉と紹介されている。また、金子の研究では2六歩も悪手とされている。評価結果は、最善手が2六歩、次善手が5六歩となった。悪形となる手を最善手と判断して

			馬	馬
		王		
歩	歩	馬	歩	歩
		歩		
銀			飛	
歩	歩	歩		歩
			桂	香

図 2 悪形チェック例題 1

		歩		
歩	歩	馬	歩	歩
	銀	金		
香	桂		玉	

図 3 悪形チェック例題 2

しまっているが、この原因は全ての駒に対して同じ重みで関係性を評価している点にあると考えられる。例えば、2八に玉ではなく飛車がいる場合は2六歩は悪形とはならないのだが、提案手法では2八にいる駒が玉の場合と飛車の場合でそれほど評価値が変わらない。このため、出現数の多い2六歩が最善手と判断されたのだと考えられる。なお、次善手の5六歩は良形の手であろうと考えられる。

				歩
歩	歩	歩	歩	
金		銀	玉	
	金		桂	香

図 4 悪形チェック例題 3

5 おわりに

本稿では、2つの駒の関係性の重みを学習により調整する従来の手法に加え、駒の歩数を教師値として用いる手法を提案し、実験結果を通して効果を示した。学習した重みについては、1つの駒、2つの駒の関係ともに筆者の目で見て明確に判断できるような不正な値は発見できなかった。そして、学習した

重みによって好形と悪形の判断ができるかどうかを試す悪形チェックの実験でも、例題 1、例題 2 については完璧に正解を導くことができた。これは金子の手法における親子モデルの結果よりも優れた結果である。例題 3 において悪形を選択してしまう原因を解消するには、駒によって重みに差を設けることが考えられる。例えば、王が動いた場合の重みを歩が動いた場合の重みよりも高くする方が良い可能性がある。

また、手数による重みの逓減率 γ を本稿では 0.1 と固定しているが、これは予備実験の結果 0.01 で学習した結果よりも 0.1 で学習した結果のほうが優秀と筆者が判断したためである。 γ の値は学習結果に直結するためより多くの実験によって定めるべきである。

最後に、本稿では手番側の駒のみを評価対象としているが、より駒組みの精度を高めるためには、少なくとも相手の飛車や相手の王の位置を考慮すべきであると考えられる。これらは今後の課題である。

参考文献

- 1) Jeff Rollason, SUPER-SOMA -Solving Tactical Exchanges in Shogi without Tree Searching, LNCS, Vol.2063, pp.277-296, 2000
- 2) 金子知適, 田中哲朗, 将棋プログラムにおける棋譜を利用した囲いの評価, 夏のプログラミングシンポジウム予稿集, pp.63-70, 2003
- 3) 金子知適, 田中哲朗, 山口和紀, 川合慧, 駒の関係を利用した将棋の評価関数, 第 8 回ゲームプログラミングワークショップ, pp.14-21, 2003
- 4) 保木邦仁, 局面評価の学習を旨とした探索結果の最適制御, 第 11 回ゲームプログラミングワークショップ, pp.78-83, 2006
- 5) 山下宏, YSS-そのデータ構造, およびアルゴリズムについて. 松原(編), コンピュータ将棋の進歩 2, 第 6 章, pp.112-142. 共立出版, 1998.
- 6) 久米宏. 将棋倶楽部 24 万局集. ナイタイ出版, 2002
- 7) 久米宏, 最強の棋譜データベース, 成甲書房, 2004
- 8) 有岡雅章, 将棋プログラム KFEnd, Web, 2000 (<http://www31.ocn.ne.jp/~k fend/>)