

ランキング学習による流れを考慮した自然な指し手の選択方法

杵渕哲彦[†] 伊藤毅志[†]

自然な指し手を選択する将棋 AI 実現のために、人間が用いる思考方法である流れに着目する。遷移確率を流れを表現するものとし、指し手選択の際に、既存の AI が考慮している評価値に加え、遷移確率も考慮することで、自然な指し手の選択を行う、ランキング学習を用いた手法を提案する。評価実験として、最善でない選択が一定数含まれる中級者の棋譜と提案手法の指し手の一致率を測定し、既存手法に比べて一致率の向上が確認された。

A Selection Method of Natural Moves in Shogi by Considering the Flow with Learning to Rank

TETSUHIKO KINEBUCHI[†] TAKESHI ITO[†]

In order to realize a computer Shogi which selects a natural move, we focus on the flow of moves. We assume that transition probability represent the flow of moves. We propose a method to select natural moves by considering transition probability and evaluation value with learning to rank when a computer Shogi selects a move. Result of experiments suggest that the proposed method works effectively with measuring a concordance rate of selection by the method and intermediates.

1. はじめに

ゲーム AI の研究ではこれまで強さの向上を大きな目標として様々な技術が進歩してきた。囲碁では 2016 年に Google が開発した AlphaGo がトップ棋士のイ・セドル氏に 4 勝 1 敗で勝ち越し、本研究で題材とする将棋でも、2015 年に将棋 AI がトップ棋士に統計的に勝ち越し可能性が高い事が示された。このように、現在では様々なゲームでゲーム AI が人間を超える強さを獲得している。

ゲーム AI の強さの向上に伴い、その過程で得られた知見を用いて、ゲームの内容を観戦者に向けて解説する事を目的とした解説文の生成や、対戦相手と拮抗した好勝負を繰り広げることを目的とした接待プレイなど、強さ以外に着目した研究がなされてきている。そこでの課題の一つとして、人間にとって自然な選択をすることが挙げられる。

人間の選択が全て自然であるとし、人間の選択結果を模倣することで自然な選択を実現しようとした場合、選択の多くが最善手であるような上級者の場合は、強さを求めた既存の AI である程度模倣が可能であると考えられる。一方で、その選択の多くが最善でない初中級者の場合、既存の AI で模倣するのは困難である。従って、そのような選択結果を模倣するためには、これまで用いられてきた強さと言う基準だけではなく、新たな基準を導入し、それらに基づいて選択を行う必要がある。

そこで人間が思考に用いる手法である「流れ」に着目する。人間は選択をする際、これまでの過程と現在の状態を把握し、今後の状態を予想した上で選択を行う。ここでよ

り良い選択を行いたい場合、人間は考慮すべき全ての情報を考慮して選択を行おうとする。しかし、情報が多すぎる、不確定な要素がある、などの理由からそれが完全には出来ない場合がある。その際に用いるのが、過去の経験的知識に基づいた前後関係を考慮した思考である流れである。例えば、「これまで攻撃が上手く繋がっている」という状況と、「そのような状況では次の攻撃も上手くいく」という経験的知識から、「攻撃する」という選択が行われる。また、「これまでお互いに攻め合っている」という状況と「攻め合いは続けるのが普通」という経験的知識から、現在採ることのできる選択肢の中からまず攻め合いを続けた場合どうかを考慮し、自分に都合の良い展開になるなら攻め合いを選択し、不都合が生じるなら次にそれ以外の選択肢を考慮するという思考が行われる。

本研究で題材とする将棋でも、この流れを用いた思考が行われている。伊藤はプロ棋士とコンピュータ将棋の思考過程の違いについて発話データと思考ログを比較し、プロ棋士は過去の経験的知識に基づいて前後関係を考慮し、直感的に指し手を生成した上で候補手を絞った狭く直線的な読みを行っているのに対し、コンピュータ将棋は膨大な先読みを寄ってその場で良い手を探索していると指摘している[1]。このように人間と現在のコンピュータ将棋の次の指し手を選択する思考過程は大きく異なっているため、人間の思考過程を経て選択された手をコンピュータ将棋に実現させるためには、流れを考慮する必要がある。

そこで本研究では将棋を題材とし、流れを考慮することで人間が行う候補手を絞った狭くて深い読みを模倣するこ

[†] 電気通信大学大学院 情報・通信工学専攻

とで自然な指し手を選択する事を目的とする。その基礎研究として、人間の指し手が全て自然であると仮定し、流れを考慮することで既存手法よりも自然な選択を行う事が出来る事を示す。具体的には、既存の AI の選択基準である $\alpha\beta$ 探索と評価関数を用いて求めた評価値に加え、流れを表現するものとして遷移確率も選択の際に考慮する。この遷移確率と評価値の 2 値をどの程度の重みで考慮するかは人間の棋譜を教師としたランキング学習を用いて決定する。評価実験として、最善でない手が多く含まれると考えられる中級者の棋譜との一致率を測定することで、提案手法の有効性の検証を行う。

2. 関連研究

将棋における人間の選択傾向を抽出した研究として、鶴岡らの実現確率探索が挙げられる[2]。実現確率探索は、ある局面が実現するであろう確率である実現確率を
(実現確率) = (親の局面の実現確率) × (指し手の遷移確率)
で再帰的に定義し、実現確率が閾値以上の局面を探索することで、実戦で生じやすい局面を中心とした探索を行う事を目的とした手法である。指し手の遷移確率は「直前の手で当たりになったかどうか」「成る手かどうか」と言った指し手の複数の特徴からロジスティック回帰によって算出している。本研究では、遷移確率自体がある局面でどのような手を選択されやすいかという人間の経験的知識を表したものであり、加えて「直前の手で当たりになった」と言った直前の相手の手を考慮していることから部分的な前後関係を考慮していると見なせると考え、指し手の遷移確率を流れを表す値として用いる。

また、将棋の評価関数の学習にランキング学習を用いた研究として末廣らの研究が挙げられる[3]。末廣らは、カーネル関数の一つである多項式カーネルを用いて、低次の特徴の単項式で既存の将棋 AI の学習に用いられている特徴や使われていない有用な特徴を非明示的に扱う事によって、優秀な局面評価関数を作成することを目的とした。ある局面でプロが指した手を正例、指していない手を負例とし、各局面での正例の値が負例の値よりも大きくなるような関数を求めるランキング学習を用いて評価関数を作成する手法を提案し、SVM と RankingSVM による学習結果を比較することでランキング学習の有効性を示した。本研究でも、各局面の全合法手を正例と負例に分け、同様の学習を行う事から、学習手法として有効とされたランキング学習を用いる。

3. 提案手法

3.1 概要

遷移確率と評価値に基づく値の 2 値を入力とし、各局面の全合法手の中で実際に指された手の値が最も高くなるようなランキング関数を、人間同士の棋譜を教師としたランキング学習によって求める。

3.2 遷移確率と評価値の算出

遷移確率の算出では、各局面に含まれる複数の特徴の特徴量を算出し、それらを用いたロジスティック回帰によって遷移確率を求める。

評価値の算出では、コンピュータ将棋で一般に用いられている $\alpha\beta$ 探索に基づくゲーム木探索を行い評価値を算出する。この際、多くの将棋 AI の評価値はおよそ -32600 ~ +32600 の間の値を取るが、ゲーム中のほとんどでは -1500 ~ +1500 程度の値を取るため、定義域に比べて実際の評価値の変動する領域は小さい。そのため、評価値をそのまま入力として用いた場合、各合法手間の評価値の差が認識しにくくなる可能性がある。そこでその差を認識しやすくする方法として、シグモイド関数によって評価値から勝率を算出して用いる手法が考えられる。この評価値に基づく 2 つの値のどちらがより提案手法に用いるのに適しているかを確認するため、予備実験として、評価値そのものを入力に用いた場合と、評価値から算出した勝率を入力に用いた場合とで、テストデータとの一致率の比較を行う。

3.3 ランキング学習

ランキング学習は、与えられた集合に対して正しい順位づけを行うための関数を求める手法であり、検索エンジンの検索結果の提示順序などに用いられている。ここでは、各局面の全合法手の集合の中で最も高い順位の手が、実際に選択された手となるよう学習を行うためにこの手法を用いる。合法手の中で実際に選択された手を正例、選択されなかった手を負例として、高い確率で正例の値が負例の値より大きくなるような関数を求める。ランキング学習には様々な手法があるが、本研究では、pairwise 手法であり、SVM を用いてランキング学習を行う手法である RankingSVM を用いる。

4. 予備実験

4.1 目的

提案手法において遷移確率と評価値の両者を考慮する際に、評価値をそのまま用いる場合と、シグモイド関数を用いて評価値から勝率を算出して用いる場合のどちらが有効であるかの検証を行う。

4.2 実験方法

将棋倶楽部 24 万局集のレーティング 1450~1550 の棋譜から無作為に 300 局抽出し、全ての局面の全合法手の遷移確率と評価値を計算した。200 局を教師データ、100 局をテストデータとする 3 分割交差検証を行い、一致率の比較を行った。遷移確率の計算は OpenShogiLib と GPS 将棋を用いて行い、評価値の計算は Bonanza6.0 を深さ 5 で探索させた結果を用いた。評価値 v に対する勝率 P は式(1)を用いて算出した。

$$P = \frac{1}{1 + \exp(-\frac{v}{600})} \quad (1)$$

ランキング学習には SVM を用いてランキング学習を行う RankingSVM のプログラムである SVMrank を用い、カーネル関数には線形カーネルを用いた。

4.3 結果

結果を表 1 に示す。表 1 より、遷移確率と式(1)を用いて評価値から算出した勝率を入力に用いた方が一致率が高い結果となった。

表 1 3 分割交差検証の結果

考慮した要素	一致率
遷移確率, 勝率	46.5%
遷移確率, 評価値	43.1%

4.4 考察

遷移確率と勝率を考慮した方が一致率が高かったことから、提案手法においては式(1)を用いて評価値から勝率を算出して用いる手法が有効であると考えられる。

5. 評価実験

5.1 目的

実験 1 では、提案手法におけるランキング学習の有効性を検証するために、SVM で学習した関数と RankingSVM で学習した関数をそれぞれ評価値として扱い、テストデータとの一致率の比較を行う。

実験 2 では、遷移確率と評価値の 2 値を考慮する提案手法の有効性を検証するために、評価値の最も高い手を選択する手法と遷移確率の最も高い手を選択する手法の 2 手法とのテストデータとの一致率の比較を行う。

5.2 実験方法

将棋倶楽部 24 万局集のレーティング 1450~1550 の棋譜から無作為に 1100 局抽出し、1000 局を教師データ、100 局をテストデータとした。RankingSVM の実装には SVMrank を、SVM の実装には LIBLINEAR を用い、用いるカーネル関数は共に線形カーネルとした。RankingSVM,

SVM 共に 5 分割交差検証を行い、コストパラメータの調整を行った。一致の判定については、上位 r 位以内の手の中にテストデータで選択された手が含まれていた場合、上位 r 位で一致したと見なす。また、同じ値の手があった場合には、その中でランダムに順位を決めることとする。

5.3 結果

5.3.1 実験 1

結果を表 2 に示す。いずれの r の場合も RankingSVM が SVM に比べて高い一致率を示しており、 $r=1$ でも RankingSVM が 4%程高い一致率を示している。

表 2 テストデータとの一致率

r	SVM	RankingSVM
1	0.4355	0.4758
3	0.6979	0.7541
5	0.8129	0.8529

5.3.2 実験 2

結果を表 3 に示す。評価値と遷移確率については $r=1,3$ では遷移確率が、 $r=5$ では評価値の方が高い一致率を示しているが、いずれの r の場合も、RankingSVM が両者よりも高い一致率を示している。

表 3 テストデータとの一致率

r	評価値	遷移確率	RankingSVM
1	0.4198	0.4358	0.4758
3	0.6854	0.6875	0.7541
5	0.8324	0.7862	0.8529

5.4 考察

実験 1 の結果より、提案手法においてランキング学習を用いることが有効であると考えられる。SVM の学習で得られたモデルによると、評価値の重みがおおよそ 0.031 であったのに対し遷移確率の重みがおおよそ 1.927 と大きかったため、遷移確率の最も高い手を選択する手法と近くなり、結果として一致率も RankingSVM には及ばず、遷移確率に近いものになってしまったと考えられる。一方で RankingSVM は学習の結果、評価値の重みがおおよそ 16.232、遷移確率の重みがおおよそ 9.842 と両方の要素をバランスよく考慮したことで、いずれか一方のみの場合よりも高い一致率が実現できたと考えられる。

6. おわりに

本研究では、流れを考慮することで人間が行う候補手を絞った狭くて深い読みを模倣することで自然な指し手を選択する事を目的とし、その基礎研究として、既存の AI で用いられている評価値に加えて、遷移確率の流れを表すものと見なして考慮することで、最善でない選択が一定数含まれる中級者の指し手を模倣する手法を提案した。

中級者の棋譜から、実際に指された手を正例、指されなかった手を負例として教師データを作成し、ランキング学習の 1 手法である RankingSVM を用いて評価値と遷移確率それぞれを考慮する重みを学習させ、同じく中級者の棋譜から作成したテストデータとの一致率を測定したところ、SVM を上回る一致率であった事から提案手法におけるランキング学習の有効性が、そして最も評価値の高い手を選択する手法と最も遷移確率の高い手を選択する手法の 2 手法を上回る一致率であったことから提案手法の有効性が、それぞれ示された。

本研究の問題点として、考慮する要素が評価値と遷移確率の 2 つのみと少なかったことが挙げられる。特に流れを表現するものとして遷移確率のみを用いたが、流れはこれまでの過程を考慮するのに対し、遷移確率は直前の相手の手までしか考慮することが出来ないなど、流れを表現するものとしては不十分な点がある。従って、隠れマルコフモデルのような、より長い時系列データを表現可能な要素を考慮に加えると言った工夫が考えられる。また、扱う局面情報を評価値一要素のみとしているが、局面の進行度のような評価値では表現できない要素も流れを考慮して自然な指し手を実現する上で必要であると考えられる。一致率のさらなる向上のために、考慮すべき要素の精査と導入、そして他の効果的な学習手法についても検討していく必要がある。

今後は、人間の手との指し手の一致率の精度を高めた上で、それを用いて人間が行う狭く直線的な読みの模倣を行う事で、人間が選択する自然な指し手の実現を目指していく。

参考文献

- [1] 伊藤毅志：コンピュータの思考とプロの思考—コンピュータ将棋の現状と展望，情報処理学会論文誌，Vol.48, No.12, pp.4033-4040, (2007)
- [2] 鶴岡慶雅：「激指」の最近の改良について—コンピュータ将棋と機械学習—，コンピュータ将棋の進歩 6，共立出版，pp.71-83 (2012)
- [3] 末廣大貴，畑埜晃平，坂内英夫，瀧本英二，竹田正幸：SVM による 2 部ランキング学習を用いたコンピュータ将棋における評価関数の学習，電子情報通信学会論文誌．D，情報・システム，J97-D(3), pp.593-600 (2014)