

F-023

囲碁学習支援のためのポリシーネットワークを用いた棋譜からの棋力推定

Estimating Player's Go Skill using Policy Network from Game Records

濱田 浩輝* 中村 貞吾*

Hamada Koki Nakamura Teigo

*九州工業大学 Kyushu Institute of Technology

1 はじめに

囲碁は、チェスや将棋に比べ盤面の評価が難しいとされるゲームである。それは囲碁が他のゲームと比べて、探索空間が大きすぎることや、駒を直接取り合うわけではないため局面の評価の明確な基準が少ないからである。そのため囲碁 AI の研究は、従来の探索法ではなく、モンテカルロ法や機械学習を用いた手法が主流となっている。その中でモンテカルロ法と機械学習を組み合わせた囲碁 AI アルファ碁 [1] がトッププロに勝利し、この時点で実力の上では十分と言える域に到達した。そこで、囲碁研究の新たな目標として学習支援が注目されている。囲碁の学習支援には、指導碁や感想戦などが挙げられるが、そのどちらもまず、相手の棋力というものがある程度わかっていないと、適切な解説や指導を行うことができない。本研究では、囲碁 AI で用いられているポリシーネットワークを用いて、学習者の棋譜中の着手を評価することで棋力を推定し、実力に合わせた適切な囲碁学習支援を行えるようにすることを目的とする。

2 棋譜からの棋力推定

人間は 1 局の棋譜からある程度の棋力を推定することができると言われていた。これは棋譜、つまりある局面に対する着手の集合から棋力を測ることが可能であるという事を示している。この棋力の推定にはプレイヤーが形勢判断をどのように行っているかという情報は含まれていないため、あくまで局面に対する次の一手の予測から、棋力の推定が行えているということである。AI を用いて次の一手の予測から棋力を推定するために、アルファ碁で用いられている局面の評価などを考慮しない次の手の予測にあたる部分である SL ポリシーネットワークによって実装する。

3 ポリシーネットワーク

アルファ碁は探索の幅を減らすための SL ポリシーネットワークと、局面の評価を行い探索の深さを減らすバリューネットワークをモンテカルロ法に組み込み、

効率的に探索を行っている。SL ポリシーネットワークは任意の局面が入力されると、その局面の合法手に対して、次の一手の打たれやすさを表す確率を出力する。その中から、最大確率の着手と実際の対局で選択された着手との比較より、対局者の棋力を推定する。本研究では、アルファ碁で使われている SL ポリシーネットワークと同じ、48 チャンネル、192 フィルタ、13 層の畳み込みニューラルネットワークを用いた囲碁 AI オープンソースである DeltaGo[2][3] を使用して、棋譜の解析を行う。

4 提案手法

同級 (段) 同士の対局を用いて、一手目から終局までの着手に対して、実際に選択された着手確率と最大確率との比をその着手の評価値として使用して解析を行う。7 段から 20 級までの各レベル毎に、棋譜中のすべての着手の評価値の平均値と分散を計算したものを図 1 に示す。

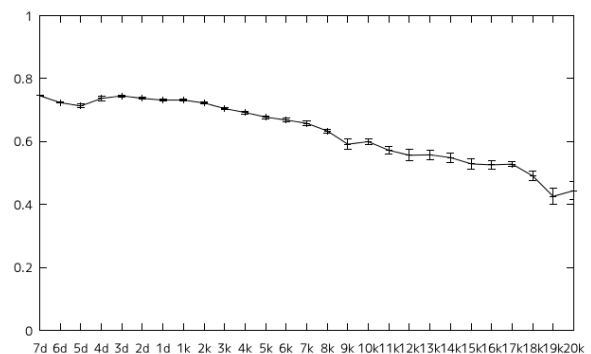


図 1: 各段 (級) 位の (着手確率/最大確率) の平均と分散

図 1 より、評価値の平均は高段者から低級者にかけて減少傾向にある一方で、その標準偏差は高段者から低級者にかけて増加傾向にあることがわかった。そこで、この特徴を棋譜からの棋力推定に利用する。

5 実験

以下の4段～17級までのデータを用いて3クラスに分類する。分類には平均からの分散を基準とした差を用いる。同級(段)同士の対局をポリシーネットワークを使用して解析を行う。

表 1: 4段～17級までの各クラスのデータ数

	データ数
1段～4段	7184
1級～9級	47770
10級～17級	4795
合計	59749

4段～17級までの各100個ずつのデータを用いて、分類を行う。各段(級)のそれぞれの棋譜について各クラス分類されたデータの個数を図2に示す。

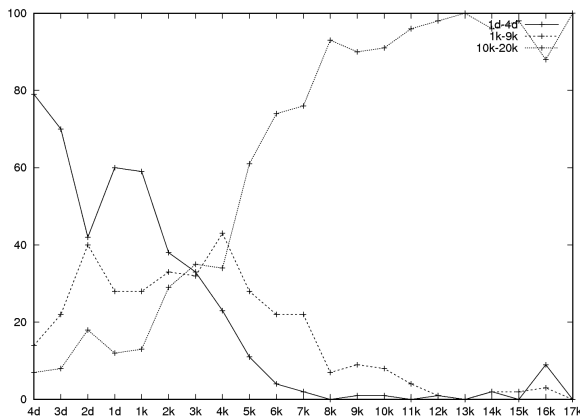


図 2: 各段(級)位の3クラスに分類されたデータの個数

図2に出力された分類されたクラス毎のデータ数と適合率を表2にまとめた。

表 2: 分類されたクラス毎のデータ数と適合率

	上級者	中級者	初級者
1段～4段	251(62.7%)	104(26.0%)	45(11.2%)
1級～9級	171(19.0%)	224(24.8%)	505(56.1%)
10級～17級	13(1.6%)	20(2.5%)	767(95.8%)
合計	435	348	1317

上級者、初級者の分類では、1段～4段の400局のデータに関しては、適合率は62.7%、10級～17級の800局のデータでは、適合率95.8%とかなりの確率で正しく分類することができていた。しかし、中級者の

データ900局に関しては、適合率24.8%と低くほとんどが初級者と間違っ分類されていた。これは、中級者に分類されているデータに偏りがあり、平均の値が高く見積もられていたためである。

6 まとめと今後の課題

SLポリシーネットワークを用いて上級者、中級者、初級者の3クラスに分類することができた。

今後の課題としては、今回の実験では、データの偏りや適切なクラスの区分を考慮せずに単純に平均と分散を用いてクラス分けを行っていたので、データの偏りが少なくなるように設定することや、何度も実験を繰り返して適切なクラスの区分を見つけてやる必要がある。また、1つの棋譜だけではなく複数の棋譜を用いてその平均をとってやることにより、棋力の推定の精度がどれだけ変わってくるのかという実験も行う必要がある。

謝辞

本研究は、JSPS 科研費 15K00506 の助成を受けた。

参考文献

- [1] David Silver, et al., “Mastering the game of Go with deep neural networks and tree search”, nature, vol.529, pp.484-489, 2016.
- [2] 大槻知史, 最強囲碁 AI アルファ碁解体新書 -深層学習, モンテカルロ木探索, 強化学習から見たその仕組み-, 翔泳社, 2017
- [3] DeltaGo: <http://home.q00.itscom.net/otsuki/delta.html/>