

ニューラルネットワークによるモンテカルロ囲碁の枝刈りの効果と検証

† 佐藤 慎也, ‡ 山本 修身

† 名城大学 理工学研究科 情報工学専攻, ‡ 名城大学 理工学部 情報工学科

1 はじめに

古くから世界で親しまれてきた戦略性の高いゲームの一つに囲碁がある。囲碁では、 9×9 または 19×19 の盤面に黒石と白石を交互に配置していく。この盤面上に、地と呼ばれる陣地を作り相手より多くの地を作り上げることが囲碁の目的である。2015年現在最も強い囲碁プログラムにはモンテカルロ法を用いた木探索アルゴリズム、モンテカルロ木探索 (Monte-Carlo Tree Search) [1] が使用されている。モンテカルロ木探索を用いた囲碁プログラムはモンテカルロ囲碁と総称されて世の中に多く存在している。本稿では機械学習のひとつであるニューラルネットワーク (NN) を用いて囲碁における悪手の判別を試みる。これにより、モンテカルロ囲碁の探索に直接の枝刈りおよびモンテカルロ木探索においてサブルーチンとして用いられるプレイアウトにおける枝刈りを行う。本稿のねらいはデータの表現が囲碁に適している NN を用いて悪手の枝刈りを行い、探索効率を向上させることでより良い手を発見するアルゴリズムを実現することである。

2 悪手を判別する NN の構成

本稿では、悪手を判別するための入力層 81 ノード、出力層 81 ノードの NN の作成をする。入出力層はそれぞれの碁盤の 9×9 座標 81 個に対応している。NN の入力層 (IN) と出力層 (OUT) は以下のように表現した。

$$IN_i = \begin{cases} 0 & (\text{石なし}) \\ 1 & (\text{黒石}) \\ 2 & (\text{白石}) \end{cases}, OUT_i = \begin{cases} 0 & (\text{良手}) \\ 1 & (\text{悪手}) \end{cases} \quad (0 \leq i \leq 80) \quad (1)$$

隠れ層はノード数や層の数によって表現力が違うため、それぞれ変更して実験を行い決定することにした。

また、本稿では教師データとしてインターネット対局場 KGS [2] の棋譜を元に教師データを作成した。ここで、棋譜で着手された良手、それ以外の座標は悪手として座標は (1) のように教師データを作成する。本稿では、36,754 対局 (1,778,191 局面) を教師データとして用いた。教師データのうち 8 割を訓練用データ、2 割を評価用のデータとして使用した。学習した後、NN は出力ノードで 0 以上 1 以下の実数を出力する。このとき、出力が 1 に近いほど悪手、0 に近いほど良手であると解釈する。この出力から悪手がどこにあるかを判定する。

3 モンテカルロ囲碁へ適用と結果

前節で述べた悪手を判別する NN をモンテカルロ囲碁に応用した場合の効果について検証する。効果の検証は悪手を考慮したモンテカルロ囲碁と考慮しないモンテカルロ囲碁で実際にゲームを繰り返し行い、考慮したモンテカルロ囲碁の勝率からどの程度強いプログラムとなったかで検証した。モンテカルロ囲碁への応用では NN の出力が悪手の箇所を枝刈りする。NN

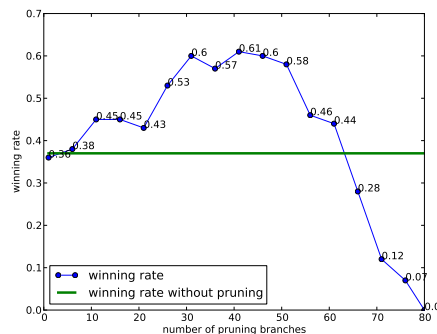


図 1: 枝刈り前と枝刈り後の対局における枝刈り後の勝率の変化。0.37 で引かれた線は枝刈りなしの黒の勝率 (基準線)。

の出力が 1 (悪手) に x 手を枝刈りしたときの効果で比較して評価する。盤面上のそれぞれの点を NN の出力値順にソートした列を p_0, p_1, \dots, p_{80} として、それぞれの出力値を $v(p_i)$ と書けば、

$$v(p_0) \leq v(p_1) \leq \dots \leq v(p_x) \leq \dots \leq v(p_{80})$$

となっている。このとき、 p_0, \dots, p_{x-1} が良手、 p_x, \dots, p_{80} が悪手と定義する。実験では悪手の枝刈りによってプログラムがどの程度強くなったかを評価する。そのため、枝刈り前と枝刈り後のプログラムで繰り返し対局を行った。本稿では黒番を枝刈り適用後のプログラム、白番を枝刈り適用前のプログラムとした。モンテカルロ木探索での枝刈りは悪手のノード (局面) を探索から排除することで実現する。図 1 にて枝刈りの個数を変化させたときの勝率の変化を示す。このグラフを見ると、注意として白と黒のもとの勝率が同じでないという点がある。グラフから得られる結果として、枝刈り前の黒の勝率から見て枝刈り後の勝率は 0.37 から 0.61 までで最大 0.24 程度向上させることができた。

4 まとめと今後の課題

本稿では悪手を識別するための NN を作成し、これをモンテカルロ囲碁へ木の枝刈りという形で適用した。作成した NN の評価を行ったところ、理想的な出力には及ばないものの近い形の NN を作成することができた。NN で判別した悪手を枝刈りしたところ、枝刈り適用前のプログラムとの対戦成績が勝率 0.24 向上したことを確認できた。今後の課題としては学習方法についての工夫、局面に応じた NN の適用が考えられる。

参考文献

- [1] Bruggmann, B.: Monte Carlo Go. Technical report. Physics Department, Syracuse University, 1993.
- [2] The KGS Go Server. <https://www.gokgs.com/>