

フリーセルの最短手順の解の探索のための ディープラーニング

Deep learning for search for shortest solutions of Freecell

和氣 卓史[†] 玉井 慎吾[†] 神保 秀司[†]
Takafumi Wake Shingo Tamai Shuji Jimbo
[†] 岡山大学大学院自然科学研究科

1. はじめに

近年、囲碁を始めとした探索空間が膨大なゲームにおいて、機械学習を取り入れた探索手法が有効であることは AlphaGo などのゲーム AI によって実証されている[1]. 本研究では、ニューラルネットに質の良い着手を学習させることで、一人遊びのカードゲームであるフリーセル[2]について、ディープラーニングによる最善手の推測と最短手順の解の効率的導出を試みた. また、本研究には、ResNet[3]による深層ニューラルネットや近年画像認識分野でも注目されはじめている Self-Attention 層を持つニューラルネットといった複数のモデルを構築し、比較する実験も含まれている. 実験に用いたニューラルネットの概要として、盤面情報を一次元の学習データとし、従来の総当たり探索によって求められた最短着手の情報を教師データとした教師あり学習を行っている.

2. 局面に応じたフリーセルの着手の学習

2.1 フリーセルとは

フリーセルとは一人用トランプゲームであるソリティアの一種であり、ジョーカーを抜いた52枚のカードをランダムな順番で8列(タブロー)に配置し、ホームセルと呼ばれるゴールにすべてのカードをスート(種類)毎に昇順で並べなおす完全情報ゲームである. タブローのカードは天(図では一番下)のカードのみ動かすことができ、カードの移動は色が交互に降順となるように動かさなければならないといった制約がある. また、ホームセルに一度置いたカードは動かすことはできず、フリーセルと呼ばれる、自由に一枚のカードを置くことのできるスペースが4つ存在する. 図1に、フリーセルの初期局面の一例を示す.

2.2 方針

本稿では、ソルバープログラムによって得られる局面毎の着手情報を教師とした教師あり学習を行うディープラーニングモデルを構築する. 入力データとした局面情報は列の切り替えを区切り文字で表した一次元配列で表現する. これは、盤面

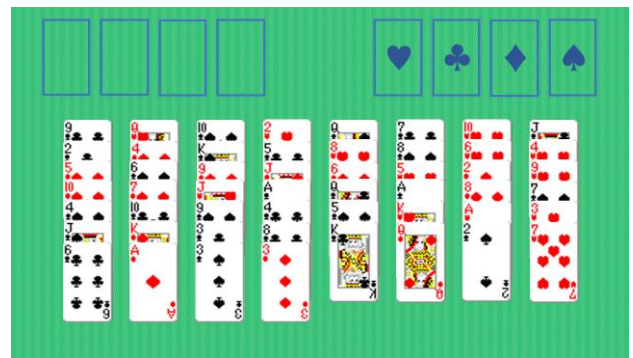


図1 フリーセルの初期局面

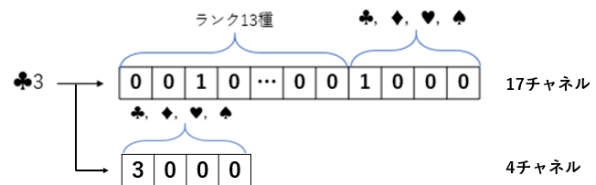


図2 カード情報の例

の情報を二次元情報として取り入れた場合ゲームに意味のないタブロー同士の横の並びの特徴を強くとらえてしまいかねないからである. また、カードの情報はスート4種とナンバー(ランク)13種を区別できるように、one-hot表現を二つつなぎ合わせた17チャンネルでの表現と、スート4種をチャンネルとし、ナンバーの情報はそのままにした4チャンネルでの表現をモデルに応じて使い分けている(図2). 変換したカード情報はフリーセル、タブロー(左のタブローから順番に)、ホームセルの順に一次元配列としてならべ、カード情報を全て0とした区切りカードを各セルやタブローの間に挿入する. データセットとして開発中のソルバープログラムによって生成された訓練用データ220万局面分、検証用データ22万局面分を使用した. モデル開発にはディープラーニングフレームワークであるKeras[4]を使用し、バックエンドとしてTensorflowを用いている.

深層畳み込みニューラルネットの構築にはResNet (Residual Network)を採用し、さらに、通常のResNetのような縦に深いネットワークではなく、

フィルタ数を増やすことで横に広いネットワークを構築し、層数の多い通常の ResNet 並みの性能を持ちつつ並列処理をより行いやすくした WideResNet[5]を構築する。また、Self-Attention 層[6]を取り入れたネットワークの構築も行い、比較実験を行った。

2.3 実験

評価指標としては正答率 (accuracy) を用いることが望ましいが、フリーセルには同時に複数の着手が最善となる状況も存在する。しかし、学習済みモデルを用いた探索の際、少ない数の着手候補を得ることができれば探索の効率化は望めるため、各ラベルの予測した確率分布内での上位ラベル 4 つの中に正解ラベルが含まれている割合を表す top_4_accuracy を今回の評価指標とした。

作成した WideResNet は畳み込み層 33 層を含み、構造は図 3 に示している。図の右側に示す通り、2 層の畳み込み層を含む残差ブロックを構成し、残差ブロック 4 つ毎に、フィルタ数を倍にしている。図中の k は WideResNet のネットワークの広さの係数とし、本実験では 3 とした。他にもいくつかのモデルを作成し、一部に Self-Attention 層を導入したものと比較を行った。

2.4 実験結果

WideResNet モデルから学習結果として得られたグラフが図 4 であり、top_4_accuracy は本実験での最高値となる 0.819 であった。Self-Attention 層を導入したモデルについては、畳み込み層が少数 (5 層程度) のモデルに導入した場合は畳み込み層のみのモデルよりも高い結果が得られたが、層数が増えるとむしろ精度が悪化する結果となった。またその際、訓練データの予測精度と検証データの予測精度の乖離がより顕著にみられた。

2.5 考察

WideResNet による学習は学習データの増減で精度の変化が大きく見られたため、更なるデータ追加による精度向上が見込める。また、正則化やハイパーパラメータの調整も最適であるとは言えないため、検証が必要である。

Self-Attention 層が浅めのネットワークで有効であり、深いネットワークで有効でなかった原因としては、Self-Attention ではすべての位置との関係性を考慮に入れてしまうため、本稿の盤面データを一次元配列として扱った理由と同じく、タブロー同士の入れ替わりを考慮せずに位置関係ごと特徴をとらえてしまったことが考えられる。最終的な精度がそれほど高くない浅いネットワークでは得られる特徴による精度向上の効果が大きかったが、高精度な結果を得ようとする際には学習データに対するか適合の面が大きくなっていくのではないかとと思われる。

3. まとめ

本稿では、最短手順探索を行うフリーセルソルバーへの適用を目的とした、ニューラルネットの構築を行った。予測手の評価は正解ラベルが予測された確率分布の上位 4 つに含まれている割合 (top_4_accuracy) で行い、検証データの最高精度は

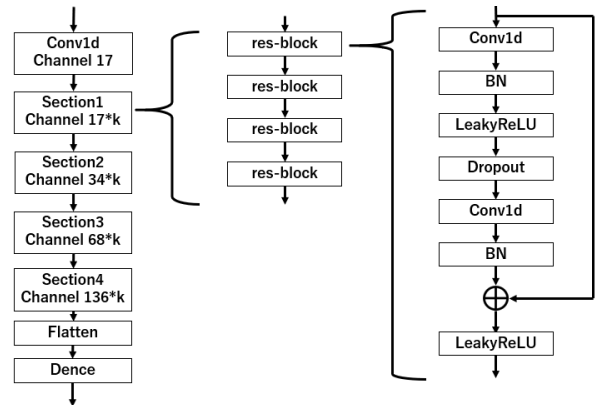


図 3 ネットワークの構造

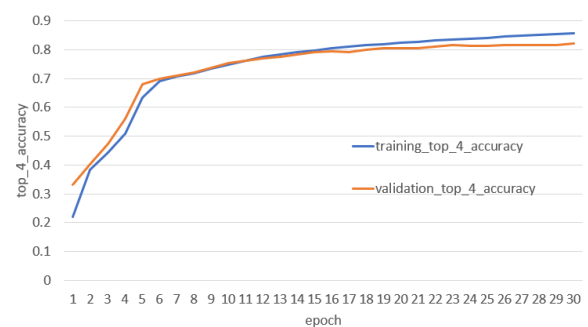


図 4 学習結果

WideResNet による 0.819 であった。本モデルをベースとした精度向上の余地は十分にあると思われる。Self-Attention 層を用いたネットワークから精度の高い結果を得ることはできず、その原因はフリーセルの盤面の特徴によるものであると考えられる。ただし、本技術の導入方法に問題がある可能性もあるため、更なる理解と追加の検証を今後も行っていく。

謝辞

本研究の一部は、株式会社アースライズカンパニーの支援により実施された。また、本研究は九州大学情報基盤研究開発センター研究用計算機システムの一般利用を利用した。

参考文献

- [1] 齊藤康己. アルファ碁はなぜ人間に勝つたのか. ベストセラーズ, 2016.
- [2] 新谷敏朗. カードゲーム「フリーセル」の性質に関する考察. 福山大学工学部紀要=The Memoirs of the Faculty of Engineering, Fukuyama University pp. 63-68, 2018.
- [3] He, Kaiming and Zhang, Xiangyu and Ren, Shaoqing and Sun, Jian. Deep residual learning for image recognition, Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778, 2016.
- [4] Chollet, François and others. Keras. <https://keras.io>, 2015.
- [5] Zagoruyko, Sergey and Komodakis, Nikos. Wide residual networks. arXiv preprint arXiv:1605.07146, 2016.
- [6] ZHANG, Han, et al. Self-attention generative adversarial networks. arXiv preprint arXiv:1805.08318, 2018.