

# 15パズルのための評価関数のランダムな学習による構成について

伊藤康太†

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

本稿では、15パズルをIDA\*アルゴリズムなどのヒューリスティックサーチで解くための効率的な評価関数をゴール状態からの逆回しの試行による学習により構成した。その学習にはニューラルネット (NN) を用いた。また、この評価関数を同様な試行を行うことで構成されたデータベース (DB) による評価関数 [1] と、許容的な評価関数であるマンハッタン距離に基づく評価関数とで、その性能を比較した。比較的簡単に解ける15パズルで試し、うまくできれば比較的小さな変更で、他の同様な状態変化を行うルービックキューブなどのスライディングブロックパズル (SBP) にも応用できると考えられる。

## 2 15パズルと比較に用いる評価関数について

15パズルはSBPの一種で、 $4 \times 4$ の枠の中に1から15までの数が書かれたコマを入れて、空いたマスを用いてスライドさせることでコマの配置を変化させることができる。ここでは図1左図をゴール状態とする。また、比較に用いるDBによる評価関数 [1] は、パターンを入力としたときゴールまでのスライドの回数 (ステップ数) を出力するものである。すべてのパターンは約  $4 \times 10^{12}$  個存在しそれをキーとしてDBを構成することは不可能なので、DBはパターンの一部を用いて構成されている。

## 3 ランダムな学習による評価関数の構成とNNの構成

15パズルのようなSBPでは、ゴール状態からランダムに状態変化させたコマの動作を逆回しすることで、ある状態からゴール状態に戻る動きを求められる。さらに多くの試行を行うことでより最適な動きを得られる。そこで、それぞれの動きに現れる状態を入力とし、ゴール状態からその状態に至るまでのステップ数を出力とした教師データをNNに学習させる。ゴール状態から始め  $N$  回ランダムに動かす試行を  $M$  回行い (ここでは  $N = 200, M = 10$  とする)、それを教師データとする。このとき、何回かの試行で同じ状態が現れた場合、ステップ数の最小値で更新する。さらに、この教師データを一旦入力としてNNに入れ、出力値が0以上かつ教師データより小さい場合、その教師データは削除する。残った教師データをまとめてNNに学習させる。このとき、ステップ数  $s$  ではなく  $|s - MD(p)|$  を記録することで、小さな値をNNに学習させることとする。

次に、本稿で構築した3層からなる単純なNNの構成

1	2	3		14	5	13	3
4	5	6	7	2	1	8	10
8	9	10	11	4	9	12	15
12	13	14	15	6	11	7	

図1 左: ゴール状態 右: ある状態

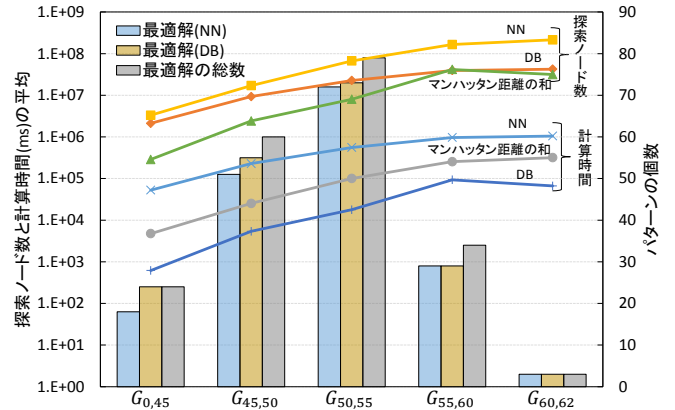


図2 評価関数別の探索ノード数と計算時間の平均 ( $G_{i,j}$  は最適解  $d$  が  $i < d \leq j$  の範囲にあるパターンの集まり)

ついて示す。これは入力層、隠れ層、出力層で構成され、入力層は空きスペース (0) とコマ (1~15) の計 16 ノード、隠れ層は 255 ノード、出力層はステップ数を出力する 1 ノードとした。また、重みの更新方法として誤差逆伝搬法を、活性化関数として ReLU 関数を用いた。

## 4 計算機実験

Python (Chainer) で学習モデルの構成を行い、C++ でIDA\*と学習モデルによる出力値  $NN(p)$  の計算を行った<sup>1</sup>。評価関数として、 $V(p) = MD(p) + NN(p)$  を用いる。ただし、出力値は小数点第一位を四捨五入し、出力値が0未満の場合は0とする。学習回数は  $10^5$  回とした。また、IDA\*にDBを使い構成した評価関数と、マンハッタン距離の和を用いた評価関数を使うことで、ランダムに作成した200個のパターンを解いたときの探索ノード数と計算時間の計算結果を図2に示す。その結果全体として、探索ノード数は57%削減でき、計算時間は7倍に増加した。また、最適解が得られた割合は87%であった。計算時間が増加した原因は  $NN(p)$  の計算に時間がかかるためだと考えられる。

## 5 今後の課題

このようなNNによる評価関数のランダムな学習による構成方法はまだ実験段階で、NNの層数などを変更することでよりよい評価関数が得られ、また、アルゴリズムの改良で  $NN(p)$  の計算が速くなると考えている。

## 参考文献

[1] 伊藤康太, 山本修身: 15パズルのための評価関数の「学習」による構成. 平成29年度電気・電子・情報関係学会 東海支部連合大会 (2017)

<sup>1</sup>CPU: Core i7-6700K, メモリ: 16GB, OS: Windows 10, 言語処理系: Python 3.6.3, Visual C++ 2017 14.10.