

ランダムな試行の学習によるルービックキューブの評価関数の構成

草間 正喜[†] 山本 修身[†]
[†] 名城大学理工学部情報工学科

1. はじめに

スライディングパズルをコンピュータで効率的に解く場合、A*や IDA*などのヒューリスティックサーチアルゴリズムが利用される。その際、探索が効率的に行われるかどうかは評価関数の性質によって決まる。本研究ではルービックキューブを IDA*などのヒューリスティックサーチで解くための評価関数を目標の状態からのランダムな試行を学習することによって得られたニューラルネットワーク(NN)を用いて作成した。15 パズルに関する同様の試みは[1]において提案されており、本研究はその手法をルービックキューブの解法に応用したものである。

2. ルービックキューブについて

ルービックキューブは一般的に3×3×3の立方体で、各面は白(0), オレンジ(1), 黄(2), 赤(3), 緑(4), 青(5)の6色のいずれかが設定されている。1つの面を回転させることによって状態を変化させることができ、この回転を繰り返すことで目標の状態に変化させることがこのパズルの目的である。今回は図1のように立方体を展開し色に対応する番号で管理した。

3. 評価関数の構成

ルービックキューブは回転させた方向と逆方向に回転させることによって、元の状態に戻すことができる。ある状態に至る状態変化と回転数を目標状態から回転させることで求めることができる。ランダムに回転させたときその回転が最小回転数であるとは限らない。ある程度正しい値を求め一部のパターンを頻出させないために以下の4つの条件を適応して移動させる。1. 連続3回同じ面を同じ方向に移動しない、2. 一部連続2回同じ方向に移動しない、3. 元に戻す回転をしない、4. 平行な2つの面の回転には制限を設ける、以上の4つである。目標状態からランダムに30回転させる試行を新しい状態が500,000個になるまで行い、それを学習データとする。既出の状態が現れた場合最小回転数を記録した。このようにして得られた状態と回転数を用いて、NNの学習を行った。

次に、構築したNNの構成について示す。入力層、中間層5層、出力層で構成され、入力層は54か所設定されている色を6ビットで表しているため324ノード、中間層は10000、5000、2500、1000、1000ノード、出力層はsoftmax関数による0から30までの発生確率31ノードとした。

4. 実行結果

前章で述べた内容で評価関数を構成した。計算環境を以下に示す。OS: Windows 10, CPU: Intel Core i7 8565U, メモリ:8GByte, 学習データ作成, IDA*, NN 計算: Node. Js 12.18.3(JavaScript), 学習 python 3.8.5 上の tensorflow 2.3.0を用いた。学習には13時間必要であった。ランダムな状態を各回転数10個ずつ作成し、作成した評価関数と簡易な評価関数(ブロックの面ごとに目標の色にできる最小回転数を合計し12で割ったもの)をIDA*用いて探索した時のそれぞれの平均探索時間、平均ノード数を表1.に示す。その結果として探索するノード数は削減することができたが、探索時間が増加した。計算時間が増加した原因はNNの計算に時間がかかるためだと考えられる。

5. まとめ

本稿では、目標状態からルービックキューブをランダムに状態遷移させてできる状態と回転数の組をNNに学習させることによってIDA*アルゴリズムで解くための評価関数を構成し評価した。計算時間やノード数を削減する方法として、平面ではなく立体で扱うことで類似の状態を検出することができると考えている。

参考文献

[1] 山本修身, 伊藤康太:ランダムな試行の学習による15パズルの評価関数の構成, 電気学会論文誌C(電子・情報・システム部門) Vol. 139, No. 12, pp. 1420-28(2019).

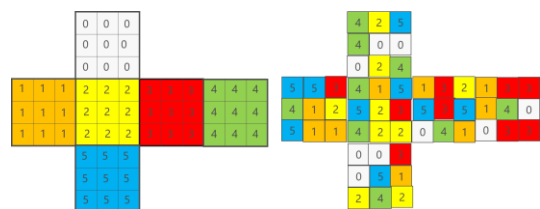


図1 左:目標状態 右:ランダムに回転させたある状態

表1 探索時間とノード数の性能評価

回転数	構成した評価関数		簡易な評価関数	
	探索時間(s)	ノード数	探索時間(s)	ノード数
5	1.845	81.8	0.001	282.4
6	1.045	51.6	0.002	1965.0
7	4.962	224.6	0.007	9500.2
8	11.348	495.0	0.041	73072.4
9	54.892	2240.6	0.741	832847.9