

4E-4

マクロオペレータの生成による
状態空間探索における学習の研究成田 昌弘 中西 正和
慶應義塾大学理工学部

1 はじめに

機械学習には様々な分野があり、その対象も多岐にわたるが、経験したことを知識表現化することにより、システムを改良することと、一般的にいうことができると思われる。

本研究における学習とは、問題解決という行為を通して、以後出現する類似した問題をより効率良く解決するための知識を、問題解決の経験から獲得してシステムの効率を向上させることを目的としたものである。人間はいくつかの手順からなる複雑な操作をどの操作に次第に慣れていくことにより、ひとかたまりの作業として習得する。この研究のプログラムでは、このような行為に基づいた学習を作用素の生成として表現する。そのプログラムは、問題の最終状態の副目標を達成するごとにそれまでの手順を一般的な形式の作用素に変換し、その作用素を次の副目標の達成、および他の問題の解決のために使用する。このような操作によって、どのようにシステムの向上がなされるかを調べるのが、本研究のねらいである。

本研究では8パズルにおける状態空間探索の時間を、どれだけ短縮できるかという点に注目して、手続きの習得による学習による効果を考察した。

2 状態空間探索

2.1 状態空間探索

定理の証明のように、問題が明確に定義され、それを解くために使うことのできる手段がいくつか与えられている場合には、問題を解くということは状態空間における探索として表現される。問題は状態空間、初期状態、終了状態、およびある状態から別の状態へと遷移させる作用素の集合から構成される。状態空間探索は状態空間に関する知識の有無により、盲目的探索とヒューリスティック探索に分けることができる。

2.2 ヒューリスティック探索

状態空間に関する常に正しいとは限らないが多くの場合において有効な知識をヒューリスティック知識と呼ぶ。このような知識を利用して効率的に探索をすることを目的とした探索法をヒューリスティック探索という。ヒューリスティック探索は、盲目的探索と比較して少ない計算量で解を得ることができる。

本研究においては、最良優先探索というヒューリスティック探索の一種を使用している。8パズルのヒューリスティック知識としては、タイルの本来あるべき位置と現在ある位置との距離の総和を使用している。

Learning by macro operator in state space search

Masahiro NARITA

KEIO, University.

Masakazu NAKANISHI

KEIO, University.

3 マクロオペレーター

3.1 マクロオペレーター

マクロオペレーターとは単一のオペレーター(作用素)とみなすことのできるオペレーターの列である。普通のオペレーターが一回の状態遷移を行なうのに対して、マクロオペレーターは二回以上の状態遷移を行なう。

マクロオペレーターを使用した古典的なシステムとしてSTRIPS[1]がある。STRIPSはロボットの計画生成を行なうシステムであるが、マクロオペレーターを生成して別の問題の解決に使用することができる。

本研究のマクロオペレーターは通常的作用素と同一の構造をしている。副目標が達成されるごとに解から新たな作用素を生成し、もしこれが今までにないものであるならば、新たな知識として獲得する。

3.2 作用素の構成

本研究において作用素は次のような形のリストで表現されている。

$$\begin{aligned} &((x_1 x_2 * 0 x_3 * * * *)) \\ &(x_3 x_1 * 0 x_2 * * * *) \\ &(0 3 2 1)) \end{aligned}$$

このリストの第1要素は、作用素の条件部と実行部を表現し、第2要素はこの作用素で行われるタイルの移動を表現している。

4 実験および結果

4.1 実験

本研究のプログラムの学習効果を調べるために、UNIX上のKCLにより以下のような実験を行った。

- 1 評価用の問題を学習機構を使わずに解かせた。
- 2 練習用の問題を1問解かせて作

用素を習得させた

3. 評価用の問題を学習機構を使わずに解かせた。
- 2, 3を繰り返した。

4.2 結果

10問の問題について、その所要時間の変化を調べた結果、2問目あるいは3問目の問題を解き作用素を獲得した時点で最も短くなるが、それ以降はゆるやかに所要時間が長くなった。

5. 考察

学習の初期においてはマクロオペレーターは効果をあげ、短時間で問題を解くことを可能にしたが、次第に作用素の数が増大するにつれて、逆に所要時間は長くなっていった。

状態空間探索に必要な計算コストは規則を選択するコストと規則を適用するコストにわけられる。作用素の習得は規則を選択するコストを減少させた一方で規則の適用のコストを増大させる。このプログラムで次第に所要時間が長くなったことは、作用素が増えたことによる規則適用コストの増大によるものと考えられる。

参考文献

- [1] Fikes, R.E., Hart, P.E., and Nilsson N.J.: Learning and executing generalized robot plans, *Artificial Intelligence, Vol.3*, pp.251-288, 1972
- [2] Nilsson, N.J.: (邦訳) 白井良明、辻井潤一、佐藤泰介: 人工知能の原理, 日本コンピュータ協会, 1983