

最適化ニューラルネットワーク間の 協調によるロボットの行動制御

- 倉庫番を例にして -

日大生産工(院) ○宮崎 海理
日大生産 松田 聖

1 はじめに

人間の行動制御は、目的に対する目標を設定し、目標を達成するための計画を行い、必要に応じてその計画自体を修正していくことによって行われると考えることができる。この人間的な『計画を適切に修正できる制御』が技術的に可能となれば、より柔軟な対応をできるシステムが構築できるであろう。その論理的計画を行うための簡単な例として倉庫番パズルを扱い、行動の制御を如何にして行っていくかを実験する。宮崎はBP法を用いた倉庫番パズルの解の計画を行った[2]が、柔軟な対応を産出することはできなかった。そこで、人間の知的活動はすべて何らかの最適化過程であるとの仮定に基づき、ホップフィールドネットワークの最適化を段階的に行うことによって、ロボットの行動を論理的に制御していく方法を考える。

2 倉庫番パズル

本研究でのモデルとなる倉庫番パズルの例が以下の図1である。

×	×	×	×	×	×	×	×
×			×				×
×			×				×
×							×
×		■	×				×
×		○	×				×
×			×	□			×
×	×	×	×	×	×	×	×

図1 . 倉庫番の例

倉庫番とは、特定の広さを持ったフィールドに、1ステップ時間で縦か横に1マスずつ移動ができるロボット（図中の○）、押すことができる物（■）、動かせない壁（×）と、物を最終的に移動させる目的地（□）が存在するパズルである。ロボットは物を進行方向へ押すことしかできず、引っ張ったり横にどけたりはできない。この中で、ロボットが物を如何にして目的地まで運ぶかを考えるパズルである。このパズルでは、ただ単純に目的地の方向へ物を押すだけではなく、壁を迂回する、また方向転換のために一度物から離れるなどの先を見越した計画と、状況が遷移する度に計画を修正する柔軟性が必要となる。

3 ホップフィールドネットワーク

ホップフィールドネットワークとは、ニューラルネットワークの一種である。1つのネットワークは、ニューロンと呼ばれるユニットが複数集まり、相互に結合することによって構成されている。ここに、エネルギー関数と呼ばれる、ネットワーク全体の目的に対する適合度を示す値を設定する。そして、各ニューロンが互いに影響し合い、エネルギーの減少する方向へと振る舞いを修正することによって、結果的にネットワーク全体でエネルギー最小、つまり最適化を目指すものである。

今回、このホップフィールドネットワークを、倉庫番を解くときに考えられる論理的ステップ1つ1つに対して構築する。それらのネットワークが論理的順序で最適化していくことによって、次に行うべき行動を決定する。以下の図2が、今回倉庫番を解くために使った論理とそのアルゴリズムである。

Robot Planning by Cooperation of Optimizing Neural Networks
- the case of Sokoban -

Kairi MIYAZAKI and Satoshi MATSUDA

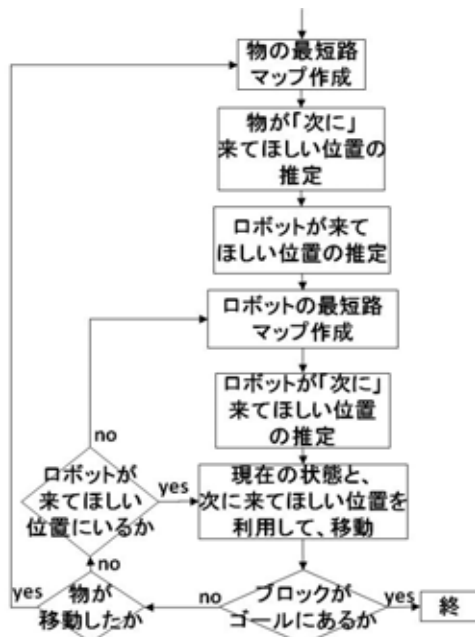


図2．倉庫番を解くアルゴリズム

まず物が目的地に着くまでの最短路を決定する。その最短路を通るときに、物が次に来るべき場所はどこなのかを決定し、そこからロボットが来るべき場所が決定する。ロボットの来るべき場所がロボットにとっての目的地となるので、ロボットに関する最短路を決定し、ロボットが次に来るべき場所を決定することができる。そして最後にロボットの移動によって、遷移した次の状態を決定する。一回の移動を終えて、物が目的地に着いていれば終了、着いていなければ次のロボットの移動を考えなければならない。その際、ロボットや物が動いている場合、それぞれ最短路を考え直す必要が生じる。そうすることによって、状態に応じた柔軟な計画の修正が行われるのである。

このアルゴリズムのそれぞれのステップに対してエネルギー関数を与える。以下に物の最短ルートマップに関するエネルギー関数を例として挙げる。

$$\begin{aligned}
 E_1 = & A \sum_i^M \sum_j^N M_b(i, j) \\
 & + B \sum_i^M \sum_j^N M_b(i, j) \\
 & \quad \times (M_b(i-1, j) + M_b(i, j-1) + M_b(i+1, j) + M_b(i, j+1) - 2)^2 \\
 & + C_1 (M_b(s_i^b - 1, s_j^b) + M_b(s_i^b, s_j^b - 1) + M_b(s_i^b + 1, s_j^b) + M_b(s_i^b, s_j^b + 1) - 1)^2 \\
 & + D_1 (M_b(g_i^b - 1, g_j^b) + M_b(g_i^b, g_j^b - 1) + M_b(g_i^b + 1, g_j^b) + M_b(g_i^b, g_j^b + 1) - 1)^2 \\
 & \quad \dots (1)
 \end{aligned}$$

M 、 N は盤面の大きさ、 M_b は物の最短ルートマップ、 s_i^b 、 s_j^b は物の現在の位置（スタートの位置） g_i^b 、 g_j^b は物の目的地（ゴールの位置）を表す。 A 、 B 、 C 、 D は定数である。式(1)は最短路の条件となる、「ニューロンの発火数が最小であること」、「スタート、ゴールからの隣はひとつだけ発火していること」、「そ

れ以外の発火しているニューロンの4方隣のうち2つは発火していること」を満たしていない場合、エネルギーが高くなるように作られている。このエネルギーが減少するように、ニューロンが全体の中での振舞いを変化させることによって、最適化が行われる。アルゴリズム内の他のステップにおいても同様に、条件を満たさない場合にエネルギーが高くなるように関数が設定され、修正されていくことで最適化が行われる。

4 シミュレーション結果

以上のホップフィールドネットワークを用いて、実際に倉庫番の問題を解くシミュレーションを行った。問題は、ロボットが1体、物と目的地が1つずつという設定とした。結果、ほとんどの問題では物を目的地に移動させることに成功した。図1のような、途中で一度、物の最短路から外れなければ解けない問題においても、最短路の計画修正によって解くことが可能であった。しかし、現在の最短路の条件ではうまく動かすことができなくなる問題については、そこからの改善を行う機構を持っておらず解くことができなかった。

5 まとめ

本研究では、論理的段階をそれぞれホップフィールドネットワークとして構築し、順序通りに最適化を行った。その結果、各ステップのネットワークがドミノ式に最適化することによってトップダウン的に協調し、最終的な結果を産出することができた。しかし解けない問題も存在した。今後は、うまく動かすことができなくなる最短路を、移動可能な場所からボトムアップ的に最短路を修正できるような機構を構築し、より協調的に移動を決定する方法を考えていく必要があるであろう。また、物や目的地が複数の場合に拡張した問題に対して、同じアプローチで協調し合い最適化が可能であるか探っていくことにより、論理的段階のネットワーク化が有効で拡張可能であることを示すことができると思われる。

「参考文献」

- 1) 松田 聖、巡回セールスマン問題の高次の最適化定式化、電気情報通信学会論文誌、vol.J83-D-、No.4 (2000)、pp.1162～1171
- 2) 宮崎 海理、タイルワールドにおけるニューラルネットワークを用いた計画行動の獲得、日本大学生産工学部第41回学術講演会数理情報部会講演概要(2008)、pp.123-124