

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

日大生産工（院） ○宮崎海理
日大生産工 松田聖

1. はじめに

人間がある目的を持った行動をするとき、そこには計画的な思考があると思われる。目的に則した目標を設定し、それを達成するための途中経過を計画し、実際に行動する。この論理的な思考を何がしかの技術で表現することができれば、柔軟な対応を強いられる実世界で適切な判断を下すことのできる機械を生み出すことができるだろう。そこで本研究では、ニューラルネットワークを用いて、その『途中経過を計画する』という部分の実現可能性を考えていく。

2. タイルワールド

今回、この計画的な行動というものを生み出す環境としてタイルワールドを利用する。次の図はタイルワールドの例である。

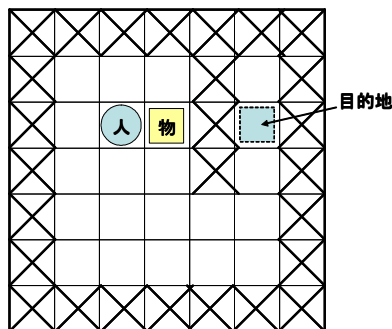


図 1. タイルワールドの例

タイルワールドとは特定の広さを持ったフィールドである。そこには、1ステップ時間で縦か横に1マスずつ移動ができる人、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。人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。この中で人が物を目的地まで運ぶ

ことを考える。この条件下では、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、先を見越した計画的な行動が必要となる。

3. 計画の実現方法

上記のタイルワールドにおいて行動を計画するための手段としてニューラルネットワークを利用する。ニューラルネットワークとは脳の神経細胞を模した学習方法である。ニューラルネットワークは特定の入力に対する特定の出力を学習し、また、それに似た入力に対して学習を踏まえた適当な出力をする汎化機能を持った機構である。そこで今回はこの汎化機能を利用し、現在のタイルワールドの状態を入力として、1ステップ時間後のタイルワールドの状態を出力するニューラルネットワークを構成する。ニューラルネットワークの位置づけのイメージを次の図に示す。

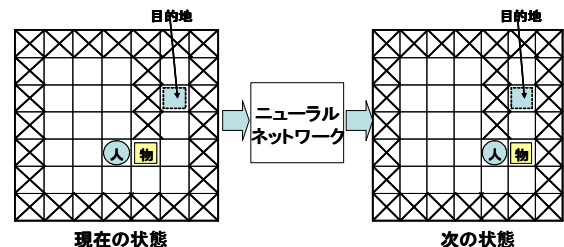


図 2. ニューラルネットの位置づけ

いくつかのケースを学習し終えたニューラルネットワークならば、学習したことがないケースを含む広範な行動計画が可能になることを期待している。

Planning robot's actions in the tile world by the neural network

Kairi MIYAZAKI, Satoshi MATSUDA

ニューラルネットワークに現在の状態を入力するにあたって、フィールド全体1マス1マスを数値に変換し入力することはできる。次の図は説明をするための簡単なイメージである。

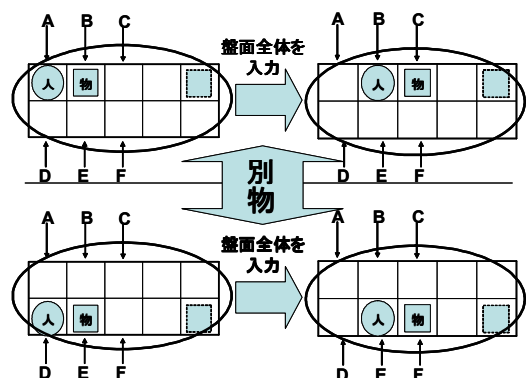


図3．盤面全体の入力

それによって、図の上のAとBの関係性は学習できるだろう。しかし、そのAからBへの移動は、下のDからEへ動くというのと同じ論理で導き出されていないだろうか。つまり、この入力方法では特定の場所と特定の場所との関係性しか学習することができないのである。このタイルワールドにおいて、問題となるのは人と、その他の物と目的地と壁の位置関係なのである。そこで、入力を人の周囲だけに注目させる。

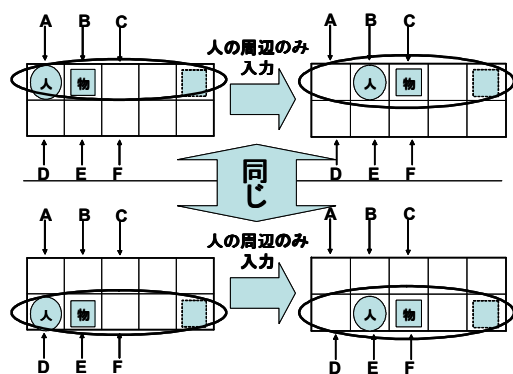


図4．注目した入力

こうして入力をずらすことによって、上と下はまったく同じ入力となる。すると盤面全体の入力では違った学習になっていたものが、同じ学習になっている。この図はイメージのために横方向だけを入力としたが、もちろん、本当は縦横方向両方でずらす。これによって人からの物や目的地、壁との関係のみを考えた入力を作ることができ、より汎化しやすいのではないかと考えている。

4．双方向探索

しかし、このニューラルネットワークだけでは、次の状態を予測しているだけで計画とは呼べない。計画とは最初から最後まで行動すべてを予測して初めて計画と呼べるのである。そこで、次の方法を取る。

まず、最初の状態からニューラルネットワークで1ステップ後の状態を予測する。次に、その1ステップ後の予測結果を再びニューラルネットワークに入力として入れる。すると最初から2ステップ先の状態を予測することができる。これを繰り返すことによって、最初の状態から終了状態に向けての計画を行うことができる。

ここで更に、このタイルワールドの問題では、終了状態というのが容易に予測できる（物が目的地の上であり、人がその隣にいればよい）ことから、終了状態から、その1ステップ前の状態を予測するニューラルネットワークも用意する。これも1つ目のニューラルネットワークと同じように再帰的に計画を行う。つまり、終了状態を入力とし、1ステップ前の状態を予測する。そして、その1ステップ前の状態を再び入力とし、2ステップ前の状態を予測するのである。その繰り返して終了状態から最初の状態に向けての予測ができる。

これら二つのニューラルネットワークがそれぞれに出力した結果で合致する部分があれば、最初から終了までの行動を計画することができたといえるのである。

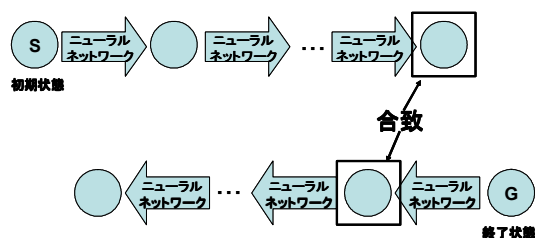


図5．ニューラルネット双方向探索

このタイルワールドの問題はステップ数が増えれば増えるほど、取りうる状態が比例的に増えていく。それを双方向探索によって、この問題が持っている取りうる可能性の発散を防ぐのが狙いである。

「参考文献」

北野宏明, 遺伝的アルゴリズム3,
p.302~327, 1997