

# 人工知能を利用したシミュレーション

研究者：西川・和久

## 1 はじめに

自動車や医療など様々なところで人工知能が使われるようになり、研究が世界中で進められている。その中で少しでもその技術に触れられたらいいと考え、人工知能を利用したシミュレーションを行った。

## 2 基礎知識

### (1) 遺伝的アルゴリズム

遺伝的アルゴリズム(Genetic Algorithms:GA)は、生物界の進化の仕組みを模倣する解探索手法である。生物界の進化の仕組みとは「生物は、交叉、突然変異、淘汰を繰り返しながら、環境に適合するように進化していく」ということを指す。環境に適合する度合い(適応度)を数値で表せば、進化して生き残った個体の数値は徐々に高くなっていく。

一般的に、以下の流れで実装される。

- 1 集団の作成
- 2 評価
- 3 交叉・突然変異
- 4 次世代の個体

1 から 4 を繰り返し、もっとも適応度の高い解を出力する。

交叉・・・交叉は候補になっている解を組み合わせ、新たな解を生み出す操作である。

突然変異・・・突然変異は、ある確率により一部を変える操作である。交叉だけでは、限られた範囲の解しか生まれない。突然変異を加えることにより、交叉だけでは生まれない解を導き出せる。

淘汰・・・淘汰は適応度の高い解を選択する操作である。

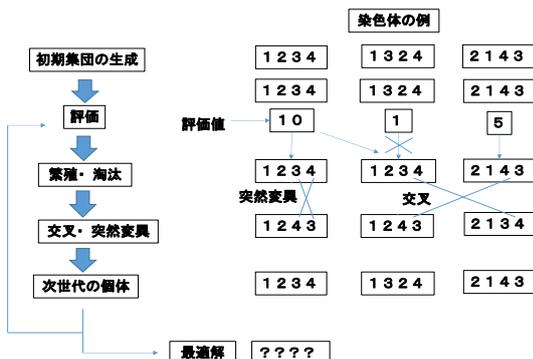


図1 遺伝的アルゴリズム流れ

### (2) Q ラーニング

一般の強化学習は一連の行動の最後に評価を与えこの評価にしたがって学習を進めていく。最後の評価値を用いて一連の動作の中のひとつの個別の行動について知識を学習して最適解を求めていく。

これと比べ、本研究で使用した Q ラーニングは、1 ステップごとに学習を行うことにより、学習効率を高速化することができる。

Q ラーニングは、報酬をより多く手に入れることができる方向に学習が進んでいく。

- 1 今ある情報を観測する。
- 2 行動を選択して行動する。
- 3 報酬がいいか悪いか学習する。

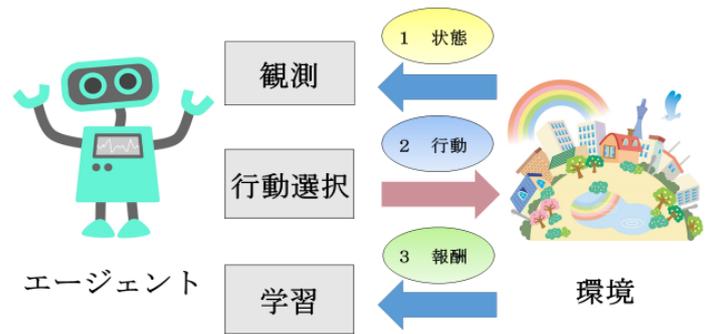


図2 Q ラーニングの学習方法

## 3 研究の成果

人工知能の基本的な知識を使い、2つのシミュレーションを行った。

### (1) 遺伝的アルゴリズムによる文字列探索手順

- 1 ランダムに生成された文字列を作成する。
- 2 正解文字列との正解率を求め、確率論方式を利用し文字列を選択する。
- 3 選択された2つの文字列を、ランダムで区切りを決め、前後で組み合わせる。
- 4 1 から 3 を、文字列が一致するまで繰り返す。図3が実行結果画面である。Target が求めたい文字列である。Best Phrase が、その時点でのもっとも Target に近いものである。そして、右にある All Phrase がその時点で作られたすべての文字列である。

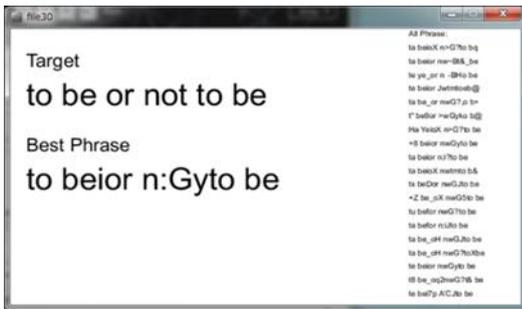


図3 制作したプログラム

(2) Q ラーニングによる迷路解析  
手順

- 1 評価を初期化する
- 2 行動を行って報酬を受け取る。\*
- 3 更新式で更新
- 4 終了条件 (ゴール) を満たしたらループを終了する。満たしていないなら、2に戻る。
- 5 終了条件 (学習回数) を満たしたら、ループを終了する。満たしていないなら、1に戻る。

\*行動を行うとき、e-greedy 法を用いる。今の状況 (今いるマス目) で評価が最大になるような行動を選択するか、ランダムに行動する。

迷路を一マスごとの座標にしてしまうと、エージェントは一マスごとにしか行動できない。それと比べ、今回作成したプログラムでは位置、速度、加速度をベクトルで表現したため、エージェントは二次元の空間上で行動することができる。

また、エージェントは通常、プログラム上には一つにしか存在しないが、今回のプログラムでは複数のエージェントを並行して実行させた。

エージェントは、マス目ごとに e-greedy 法によって、力 (加速度) が加えられる。マス目の壁に接触した場合、接触した側に行こうとする Q 値の値が下がる。

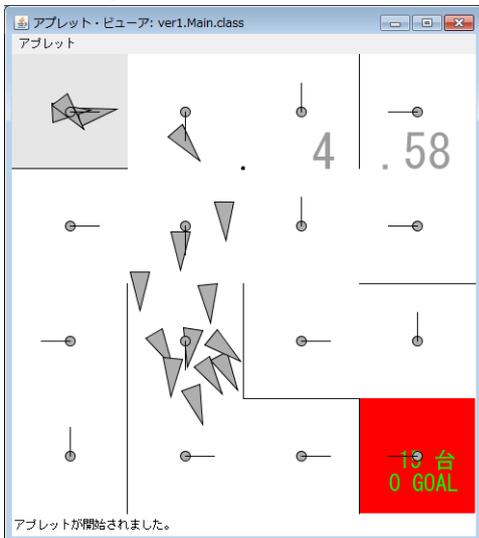


図4 初期画像

今回作成したプログラムは、マス目ごとではなく空間で行動するので収束 (ゴール) しない可能性が高い。そこで、一定時間ごとに Q 値を更新した。この Q 値を更新する際、今存在するマス目ごとの最大 Q 値を持つ行動にはペナルティを与え、その行動が低くなる可能性を高くすることで収束しやすくした。

そして、より精度の高い Q 値を出すために、遺伝子アルゴリズムを利用した。収集した多くの Q 値のデータの傾向を反映した Q 値を出力することができる。

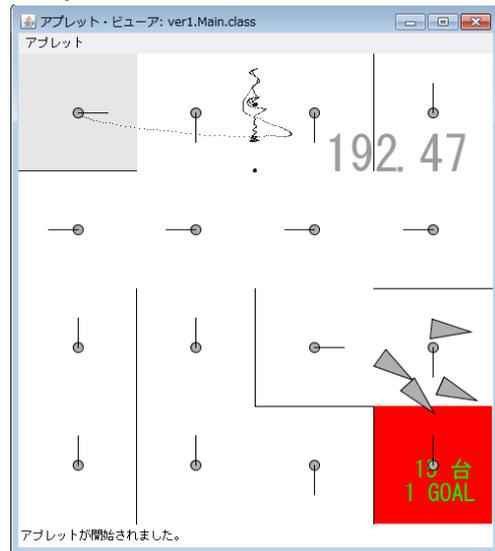


図5 学習後

#### 4 感想

【 西川 】

初めてやることばかりで、はじめはわからなくて、苦勞した。エラー文が消えないことがあり何が間違っていて、どこを直せばいいのかなど考えなくていけなかったのが大変だった。しかし、自分の思ったとおりに、実行結果が出たときはとても達成感があった。課題研究をやっている中で、少しは、わかるようになった。できたものでいろいろ試したが、できないこともあった。この課題研究でやってきたことがこれからは役立っていくと思うので学んだことを生かしていきたい。

【 和久 】

参考にした Q ラーニングのプログラムは見たことがないものが多く、どう実装するかわからなかったが、課題研究を進めて行くにつれてわかるようになってきた。Processing をはじめに使い、どうすればウィンドウが表示されるかもわからないほどだった。Processing を学んでいく中で今回の作成したプログラムの元になる、ベクトルで管理されているビークルを作成したが、ビークルには様々な特殊なメソッドが利用されており、理解することが難しかった。この課題研究で学んだことをこれからは生かしていきたい。