

ニューラルネットワークと決定木によるマルチエージェント学習

ーRoboCupへの適用の試みー

日大生産工(学部) ○高橋 昂大 日大生産工 松田 聖

1 まえがき

マルチエージェントシステムとは、複数のエージェントから構成されるシステムであり、単体のエージェントやシステムでは困難な課題を複数のシステム全体として達成する。このマルチエージェントシステムをより効率的に運用する為には各エージェントに協調行動を取らせる事が不可欠である。そこで本研究ではニューラルネットワークおよび決定木学習を用いて Robocup サッカーにおけるサッカーエージェントに協調行動をとらせることを目的とする。

2 RoboCup サッカーの概要

RoboCup サッカーは、人間のサッカーの試合と同じく、自分で考えて動く自律移動型ロボットを使った競技であり、それを研究課題とした関連技術を進歩促進させることを目的として行われている。そのなかで一部門にシミュレーションリーグがある。これはコンピュータ上でシミュレートされた 22 の独立したサッカーエージェントを用いて競技を行う。高さの概念が無い以外はほぼ実際のサッカーと同じ空間をシミュレートされている。本研究ではこの RoboCup サッカーシミュレーション 2D リーグのエージェントに二種類の学習を適用することにより、より高度なチームプレーを実現させることを目指す。

3 学習の概要

本研究において問題となるのが現在の状態を簡単に表現する指標がないことである。さらに、例えば状態を表現できたとしてもサッカーというモデルの性質上、状態の再現性が低いということも挙げられる。つまり現在のフィールド状況をどのように判断させるかが重要になってくる。さらに、エージェントには素早い意思決定が求められるため、多くの計算をする事ができない。ニューラルネットワークと決定木学習の二つの方法で、行動を決定する。

4 ニューラルネットワークの適用

ニューラルネットワークとは脳機能の特性を計算機上のシミュレーションによって表現した数理モデルである。結合荷重を修正する事により、より正確な出力値を出すモデルを形成する。サッカーのある点において、未来の挙動は現在の状態だけで決定する事ができる。このことから現在の状態を入力とし、出力を取るべき行動としたニューラルネットワークの構築を行う。選択された行動に対し評価をつけ、そこからバックプロパゲーションを用いてニューロンの結合荷重の修正を繰り返し行うことにより、最終的には現在の状況を得ること出来ればそこから最適な行動を選択できる。

Fig.1 にパスコースを決定するためのニューラルネットワークを示す。

Multi-agent Learning by neural network and decision tree

- Application to RoboCup -

Takahiro TAKAHASHI and Satoshi MATSUDA

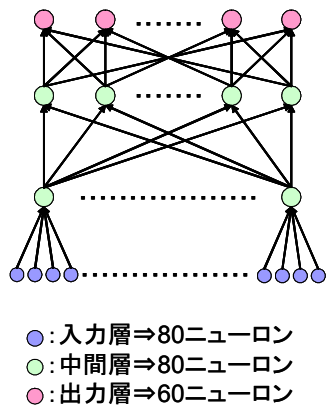


Fig.1 パスコース決定モデル

Fig.1 は本研究のニューラルネットワークモデルである。入力情報として味方エージェントの座標値とパスコースの成功確率を与え、各パスコースに対する評価を出力値としている。学習段階においてはパスの成否・ゴールとの距離・味方の位置から算出する評価値を教師信号として各出力値に対するニューラルネットワークの荷重結合を修正していく。実シミュレーション時には結合荷重が修正されたニューラルネットワークを用いて出力された評価が一番高いパスコースを選択し行動する。

5 決定木の適用

決定木とは、ある問題に対する未知のデータから、その問題の目標値に関する結論を導くための木構造をした予測モデルである。また、ある問題に対する既知のデータから決定木を生成するための機会学習法を決定木学習と呼ぶ。サッカーというモデルの観点から意思決定には正確さよりも素早さが求められることも少なくない。そのためより汎用的且つ高速な意思決定システムが求められる。これを解決するために行動の結果をデータリストとし、それを利用した決定木を生成する。この方法により行動とそれによる結果を総合的に分析したルールが作成されることになり、より汎用的かつ高速に行動を決定する事ができる。

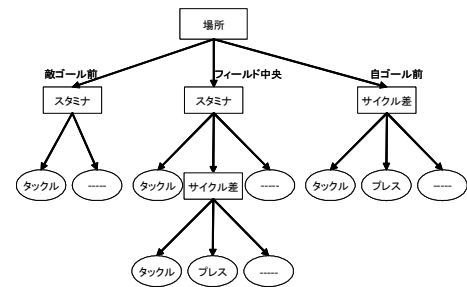


Fig.2 作成される決定木モデル例

Fig.2 は作成される決定木の一例である。「場所・スタミナ量・サイクル数・取った行動・結果」を要素とした守備行動の学習データを与え、決定木を作成する。実シミュレーション時には入力される「場所・スタミナ量・サイクル数」から最良の守備行動を選択する。

6 まとめ

現時点でパスコースに関してニューラルネットワークを用いた結果全体的に勝率・得点率の上昇がみられた。また、1000 パターンの状況から守備行動の決定木を生成しそれを基に試合をした結果ランダムな行動よりは改善されたがまだまだ学習する状況を増やす必要性があると思われる。今回採用した決定木学習もニューラルネットワークも入力情報としてにどのような情報を選択するかによって大きく結果が異なってくる学習法である。そのため入力情報を様々試してはいるが結果はほぼ同じで未だ最適とおもわれる結果を導く事が出来ていない。なのでより多くの試行を繰り返し結果を分析したい。

「参考文献」

秋山英久, ロボカップサッカーシミュレーション 2D リーグ必勝ガイド 秀和システム, (2006)

大島真樹, Java でつくる RoboCup サッカー選手プログラム 森北出版株式会社, (2005)