

ニューラルネットワークによるサッカーエージェントの守備的ポジショニングの学習

入沢達矢 小谷善行

{irisawa,kotani}@fairy.ei.tuat.ac.jp
東京農工大学

概要

本稿では、RoboCupサッカーにおける守備的ポジショニング価値関数を階層型ニューラルネットワークによって設計する。過去に、サッカーの価値関数にニューラルネットワークを利用した研究はいつくかなされているがそれらのネットワークはいずれも全結合により設計されている。全結合の場合、ユニット間の依存関係の小さい部分もコネクションしている可能性がある。このようなコネクションはネットワークの出力に大きな影響を与えない無駄なコネクションである。サッカーにおける行動決定は実時間処理が必要なためこのような無駄を省いたネットワークを設計することが有効だと考えられる。そこで、本稿では部分結合による階層型ニューラルネットワークによりポジショニング価値関数を設計することを提案し、その有効性を検証する実験を行った。

Learning defensive positioning by Neural Network for Soccer Agents

Tatsuya IRISAWA, Yoshiyuki KOTANI

e-mail: {irisawa,kotani}@fairy.ei.tuat.ac.jp
Tokyo Univ. of Agr and Tech, 2-24-16 Nakamachi, Koganei, Tokyo, JAPAN

Abstract.

This paper, we design a defensive positioning value function using 3 layer neural network in RoboCup soccer. In times past, some studies concerning value function using neural network in soccer have done, and their are full connection network. In this case, it is possible that low dependence units connect. Such connection little influence of output units and it is says that there are ineffective connections. In soccer, action decision needs real time process and it is consider the network cut idle connection is effective. This paper, we propose to design partially connected layer neural network for positioning value function and experiment in test their efficiency.

1.はじめに

本研究の目的は、サッカーにおける守備的ポジショニング価値関数を階層型ニューラルネットワークによって設計することである。サッカーでは、状態空間が複雑であるため、単純にボールまでの距離や角度などの情報から、線形に価値関数を設計することは困難である。そのため、階層型ニューラルネットワークのような非線形の価値関数が有効だと考えられる。

過去に、サッカーの価値関数にニューラルネットワークを利用した研究はいつくかなされている [1, 2, 3, 4]。それらのネットワーク

はいずれも全結合により設計されている。全結合によりネットワークを設計すると、ユニット間の依存関係が小さくてもコネクションしている可能性がある。このようなコネクションは、ネットワークの出力に大きな影響を与えないので、無駄なコネクションだといえる。サッカーにおいて、行動決定は実時間で行う必要があるため、このような無駄を省いたネットワークを設計することが有効だと言える。そこで、本研究ではサッカーにおけるポジショニングの価値関数として、部分結合による階層型ニューラルネットワークの設計を提案する。

2. ネットワークの設計方法

本研究で取り扱うネットワークは部分結合による3層の階層型ニューラルネットワークである。ネットワークは、最初に全結合のネットワークをバックプロパゲーションにより重みを学習し、そこから不要な接続を削除することによって設計することができる。

2.1. コネクション削除の条件

図1に学習されたネットワークの一部を示す。今、ユニット*i*の値が1だとすると、ユニット*j*は0.1加算されるのに対し、ユニット*k*は0.0001加算されるだけである。このようにユニット間の重みの絶対値が十分に小さい場合は接続を削除しても出力への影響も小さいことがわかる。そこで、本稿では接続を削除できる閾値 ϵ_{cut} を設定し、絶対値がそれ以下の重みをもつ接続を削除することにした。

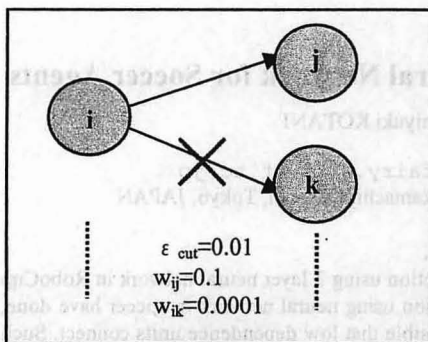


図1. コネクションの削除

2.2. 学習の繰り返し

2.1.の条件より接続を削除することにより出力の誤差が大きくなってしまいます。そこで、接続削除後に、もう一度重みを学習しなおすことによって誤差を減らすことができる。ここで学習されたネットワークには重みの絶対値が ϵ_{cut} 以下のものが再びでてくる可能性がある。そこで、重みの学習と接続の削除を学習の1ステップとし、複数ステップ繰り返すことによりネットワークを設計する。

3. 守備的ポジショニングの学習

本章では、学習の対象としている守備的ポジショニング問題の学習方法について述べる。

3.1. 少人数による部分ゲームからの学習

文献[5]によると、実世界のサッカーは、守備や攻撃の練習をする場合、少人数により練習を行う。これは、サッカーは局面全体を見ると11対11の戦いであるが、一定時間で区切るとボールの近くの区間にいるエージェントのグループ対グループでの戦いだからである。サッカーの試合はこのような部分ゲームの繰り返しにより構成されている。チームは部分ゲームに勝つことにより、最終的なゴールを決めることができるようになる。守備的ポジショニングはサッカーにおける部分ゲームの1つである。

ここで、それぞれの部分ゲームには目標が定められており、コーチは練習したい技術や戦術に応じて、部分ゲームを選択する必要がある。RoboCupサッカーの場合も同様に、学習させたい戦術に適した部分ゲームを選択することにより効率よく学習できると考えることができる。

3.2. 問題設定

本稿では、図2のような3対2の部分ゲームを設計した。

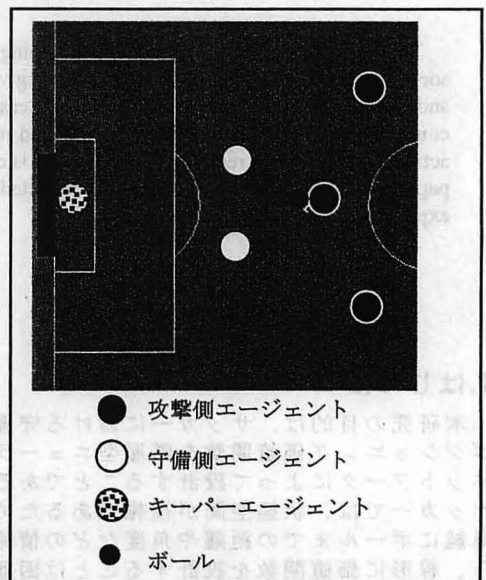


図2.3 対2の部分ゲーム

ここで、この部分ゲームにおける攻撃側と守備側の勝利条件を示す。

攻撃側勝利：

守備側のペナルティエリア内でボールをキープ

守備側勝利：

攻撃側からボールを奪う(ボールを蹴ることができる状態になる)

また、200ステップ経過したとき、フィールドからボールがでたとき、キーパーがボールをキャッチしたとき、オフサイドになってしまったときは、ノーカウントとした。

4. 学習実験

2,3章で述べた方法を用いて、守備的なポジションニングの学習実験を行った。

4.1 実験方法

まず、3.2.で述べた部分ゲームを行い、守備側エージェントの10ステップごとの行動と局面の特徴量の対を訓練データとして収集した。訓練データの報酬として、守備側勝利の場合1、攻撃側勝利の場合は0とした。つまり、1つの訓練データは(局面の特徴量、行動、報酬)となる。

次に、2章の方法を用いて訓練データからネットワークの設計した。ここで1ステップでの重みの更新回数 ϵ_{cut} は1000回、 ϵ_{cut} は0.8として学習を10ステップ繰り返した。

ネットワークの入力は4.3で述べる25個の局面の特徴量、出力は4.2.1で述べるそれぞれの行動に対する価値とした。実験は隠れ層のユニット数は16,32,64,128,256の5種類で行い、それぞれ重みの初期値を変えて10回ずつ行った。訓練データは20試合分、203のデータとした。

4.2 エージェントの設計

実験に用いたエージェントは、攻撃側エージェント、守備側エージェント、キーパーエージェントの3種類である(図2)。それぞれのエージェントは、RoboCupJapanOpen2000年度チャンピオンYowAIが提供しているライブラリを用いて作成した。

4.2.1 守備側エージェントの行動

サッカーには、ゾーンディフェンス、マンツーマン、フラットラインディフェンスなどさまざまな守備の戦略があるが、今回はゾーンディフェンスを作成した。今回作成したゾーンディフェンスは、ボールに一番近ければボールを追いかけ、それ以外は、ゾーンを作るという単純な設計にした。具体的には次の2つの行動を用意した。

- ・マーク対象の敵とゴールの中間に移動
- ・ボールを追いかける

4.3 局面の特徴量

ここで、ネットワークの入力に用いる局面の特徴量を示す。

- ・自分からボール(相対座標)
- ・自分から味方ゴール(相対座標)
- ・自分から一番近い敵(相対座標)
- ・自分から一番近い味方(相対座標)
- ・一番近い敵からボール(相対座標)
- ・一番近い敵から味方ゴール(相対座標)
- ・一番近い敵から一番近い味方(相対座標)
- ・一番近い味方からボール(相対座標)
- ・一番近い味方から味方ゴール(相対座標)
- ・一番近い味方から一番近い敵(相対座標)
- ・ボールから一番近い敵までの距離(相対座標)
- ・ボールから味方ゴールまでの距離(相対座標)
- ・味方の中でボールまで一番近いか(1,0)

4.4. 実験結果

表1にそれぞれの隠れ層において、正解率が一番高かった実験結果を示す。ここで、削減率とは、同じ隠れ層の数の全結合によるネットワークと比べてコネクション数が削減された割合であり、正解率とは訓練データの、成功事例の一致率と、失敗事例の不一致率を足したものである。

比較対象として表2に、全結合で10000回の学習を行った場合の実験結果を示す。これも、表1同様、10回行ったうち、最も正解率の高かった結果である。

表3に、テストデータに対する正解率を示す。これは表1,2で示したネットワークに対して行った結果である。

表1.作成されたネットワーク($\epsilon_{cut}=0.8$)

隠れ層のユニット数	16	32	64	128	256
接続数	108	166	320	600	1457
削減率	24.88%	18.53%	17.86%	16.74%	20.33%
最終誤差	45.80	27.49	46.43	79.75	96.01
正解率	75.37%	75.86%	75.86%	75.86%	52.71%

表2.全結合によるネットワーク

隠れ層のユニット数	16	32	64	128	256
最終誤差	48.50	49.50	120.49	120.50	120.50
正解率	74.88%	74.88%	73.89%	65.52%	58.13%

表3.テストデータの正解率

隠れ層のユニット数	16	32	64	128	256
$\epsilon_{cut}=0.8$	79.27%	80.49%	80.49%	79.27%	53.66%
全結合	80.49%	80.49%	79.27%	67.07%	59.76%

4.5 考察

表1,2より、接続数を削減したネットワークは隠れ層のユニット数がいくつでも全結合より低い最終誤差であり、隠れ層の数16から128でそれ以上の高い正解率を得ることができた。これは、全結合のネットワークは、学習する重みが多いため、10000回程度の学習では重みが収束していきなく、まだ学習が不十分なためだと考えられる。このことから、学習の途中の段階で不要な接続を削除することは、学習時間を速めることができるといえる。また削減率は、25%~15%程度であり、十分に行動決定の時間を短縮できたといえる。

表3より、テストデータの場合も、訓練データと同じ程度の正解率を上げることができている。このことから、作成されたネットワークはテストデータに対しても正しい出力をしていることがわかる。

5.まとめ

RoboCupサッカーにおける守備的ポジショニング関数を部分結合によるニューラルネットワークにより設計した。その結果、隠れ層の数と同じ全結合のネットワークと比べて、15%~25%の部分結合ネットワークで全結合

によるものより正解率の高いネットワークを設計することができた。また作成されたネットワークはテストデータに対しても、高い正解率を収めることができた。

参考文献

- [1]NODA Ituki, MATSUBARA Hitoshi, HIRAKI kazuo: learning Cooperative Behavior in Multi-agent Environment -a case study of choice of play-plans in soccer-, PRICAL' 96 pp570-579, 1996.
- [2] Peter Stone, and Manuela Veloso: A Layered Approach to Learning Client Behaviors in the RoboCup Soccer server, Applied Artificial Intelligence, 12, 1998.
- [3]安藤友人: サッカーエージェントにおける強化学習を用いたポジショニング, MACC' 97, 1997.
- [4]太田正幸: 協調行動の学習における合意の安定化手法, MACC' 97, 1997.
- [5]八重樫茂生: サッカー, 講談社, 1975.