

ニューラルネットワークを用いたパスレシーバ予測

Pass Receiver Prediction using Neural Networks

成木洋介

中島智晴

Yosuke NARIMOTO Tomoharu NAKASHIMA

大阪府立大学

Osaka Prefecture University

narimoto@ci.cs.osakafu-u.ac.jp, nakashi@cs.osakafu-u.ac.jp

Abstract

This paper describes a method that predicts a pass-receiver by neural networks for RoboCup Soccer simulation. In the proposed method, forward agents use a neural network to predict the pass receiver based on the field information such as the number of cycles for the forward players to reach the ball. We consider four methods for generating training data. In our computational experiments, we construct the neural networks that are learned from the generated training data and compare the performance of the trained neural network. We show the success rate of the prediction by each neural networks.

1 はじめに

ロボット工学や人工知能などの領域横断型研究プロジェクトとして RoboCup が知られている。RoboCup には様々なリーグが存在し、それぞれ活発に研究、開発が行われている。例えば、レスキューシミュレーションリーグでは、進化計算とファジィ識別により警官隊エージェントの意思決定を学習させる方法[1]や、サッカーシミュレーション 3D リーグでは、動的な状況において即時的な時空間解析を行う手法[2]、サッカーシミュレーション 2D リーグでは、ドロネー三角分割を用いた、人間の直感的な操作によるポジショニング[3]や、並列分散遺伝的アルゴリズムによるフォーメーション最適化[4]、強化学習によってプレイヤーエージェントの基礎的な行動を学習させる方法[5]、パレート最適解による協調的なディフェンスのポジショニング[6]などが提案されている。2D リーグでは、試合中にプレイヤーエージェントが取得できる情報に制限があり、また、各プレイヤーエージェントは独自の判断に基づいて行

動するので、状況の推移を予測することは困難である。しかし、例えば、ボールを持ったプレイヤーの次の動きを予測することができれば、試合を有利に展開することができると考えられる。そこで、本論文では、2D リーグにおけるニューラルネットワークによるパスレシーバの予測を行う手法を提案する。提案手法では、味方プレイヤーがパスをしたときの敵プレイヤーと味方プレイヤーの位置関係を学習用データとして収集する。次に、その学習用データを使ってニューラルネットワークを構築する。最後に、構築したニューラルネットワークにより、パスレシーバ予測を行い、自分がレシーバに選ばれるかどうかを判定する。数値実験では、異なる方法で学習用データを作成し、どの学習用データがパスレシーバ予測に最適であるかを調査する。

2 RoboCup

2.1 サッカーシミュレーション

サッカーシミュレーションリーグでは、実機を使用せずに、コンピュータ上の仮想フィールドでサッカー競技を行う。サッカーシミュレーションリーグには主に 2D リーグと 3D リーグがある。Figure 1, 2 に 2D リーグと 3D リーグの外観を示す。3D リーグでは 3 次元空間上に設置されたサッカーフィールド上で、ヒューマノイドロボットによるサッカー競技が行われる。3D リーグでは、ヒューマノイドに備えられている 22 の関節を制御することで、移動したりキックする。ヒューマノイドロボット同士が接触したり、バランスを崩すなどして転倒する状況が頻繁に発生する。また、歩行やキックなどの基本的な動作の優劣により勝負が決定する場合もあり、優秀な基本動作の開発に焦点が当てられていることが多いのが現状である。本論文で扱う 2D リーグでは、中央を原点とし、サイドラインに平行な x 軸、ゴールラインに平行な y 軸を持つ 2 次元フィールド上で競技が行われる。プレイヤーやボールは移動



Figure 1: サッカーシミュレーション 2D リーグ



Figure 2: サッカーシミュレーション 3D リーグ

可能な円で表される。3D リーグと対照的に、2D リーグでは基本動作は単純な関数で制御されており、基本的な動作よりも個々のプレイヤーの戦術や、優れた戦略を構築することが重要となっている。各プレイヤーはそれぞれ独立したプログラムで動き、試合は1サイクルを0.1秒とした離散時間で6000サイクルを1ゲームとする。各プレイヤーがサーバから取得可能な他プレイヤーの情報は、視覚情報と、メッセージによるコミュニケーションに限られている。また、視覚情報にはノイズがあり、メッセージの長さには制限がある。このような状況では、プレイヤーが他プレイヤーの情報を完全に把握するのは実質的に不可能である。また、個々のプレイヤーの運動能力にほとんど差がないため、ドリブルだけで敵陣へ攻め込むことは難しく、正確なパスを行うことが試合に勝利するために重要である。本論文では、レシーバがパスに気づいていないために起こるパスの失敗に注目し、レシーバ自身が周囲の状況から自分がレシーバに選ばれと事前に予測することができれば、パスの失敗を減少させることができると考え、パスレシーバ予測を行う。現在、サーバの最新バージョンは14.0.2であるが(2010年3月時点)、本論文では14.0.1を用いる。

3 パスレシーバ予測

3.1 予測手順

ある味方プレイヤーがボールをキックできる位置に移動したときに、その味方プレイヤーがパスを出すを仮定して、パスレシーバ予測を行う。パスレシーバの予測は、三層階層型ニューラルネットワークに学習させることで行う。まず、ボールがキックされてから各味方プレイヤーに到達するのに必要なサイクル数と、敵プレイヤーが各味方プレイヤーへのパスを妨害するのに必要な最短のサイクル数を計算する。このサイクル数を入力信号として、ニューラルネットワークによりレシーバに選ばれるプレイヤーを予測する。プレイヤーは自分がレシーバに選ばれと予測した場合、ボールに向かって移動する。

3.2 攻撃フォーメーション

前節で説明したように、ニューラルネットワークの入力信号は、ボールが各味方プレイヤーに到達するのに必要なサイクル数と、敵プレイヤーが各味方プレイヤーへのパスを妨害する時に必要な最短のサイクル数である。しかし、1体のプレイヤーが認識できる範囲は限られており、ニューラルネットワークの入力に必要な全てのプレイヤーの位置を把握することは困難である。そこで、パスレシーバ予測の対象となる味方プレイヤーを選別し、限定された状況だけで予測を行うために、本論文では専用の攻撃フォーメーションを作成する。Figure 3にフォーメーションの概形を示す。このフォーメーションは、敵ゴール側を前方とし、



Figure 3: 攻撃フォーメーション

攻撃を担当する5体のプレイヤーのうち、ボールに最も近い1体がボールを追い、他の4体のうち3体が前方に移動し、1体が背後に移動する。このとき、ボールが前方の3体のプレイヤーにパスされた時の周囲の状況から、その状況におけるパスの対象をニューラルネットワークに学習させるものとする。攻撃を担当する5体のプレイヤーは同じニューラルネットワークを使用するものとし、フォーメー

ションに参加していないプレイヤーはニューラルネットワークを使用せず、守備行動をするものとする。

3.3 ニューラルネットワークの設定

ニューラルネットワークの構成は、入力層ユニット数 6、中間層ユニット数 15、出力層ユニット数 3 とし、学習係数 0.01、慣性項係数 0.01 とした。入力信号は、前方 3 体の味方プレイヤーにパスするのに必要なサイクル数が 3 つと、各味方プレイヤーへのパスを敵プレイヤーが妨害するのに必要なサイクル数 3 つとした。Figure 4 は前方の味方プレイヤーが 2 体の場合の入力信号を表している。Figure 4

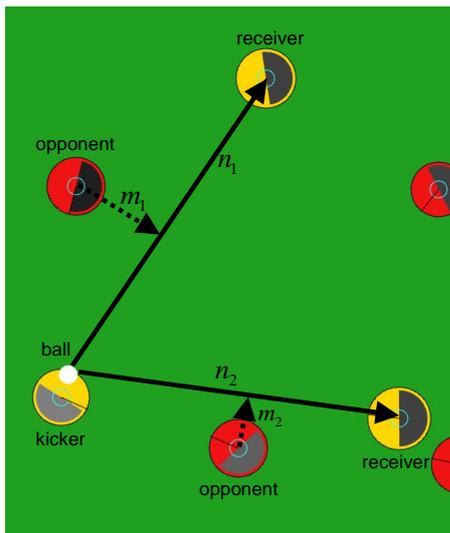


Figure 4: 前方の味方プレイヤーが 2 体の場合の入力信号

において、 n_1, n_2 は味方プレイヤーにパスするのに必要なサイクル数である。味方プレイヤーに到達するときのボールの速度が 1.2 となるように、現在位置から味方プレイヤーまでボールが移動するときのサイクル数を計算する。 m_1, m_2 は敵が妨害するのに必要なサイクル数である。敵プレイヤーは最初に 1 サイクルでターンをして、初速 0 の状態からパス経路までの最短距離を移動してパスに割り込むと仮定して、サイクル数を計算する。妨害サイクル数は敵プレイヤーの中で最も小さいものを入力信号とする。また、敵プレイヤーがボールと味方プレイヤーを結ぶ線分上にいる場合は、入力信号は 0 となる。入力信号として入力するサイクル数の上限を 10 サイクルとし、10 サイクルを越えている場合は入力信号を 10 とする。プレイヤーの位置情報が欠落しておりサイクル数を計算できない場合も、入力信号を 10 とする。

また、前方 3 体の味方プレイヤーに 3 つの出力層ユニットを割り当て、レシーバに対応する出力層ユニットの教師信号を 1、それ以外の出力層ユニットの教師信号を 0 とする。このとき、いずれのプレイヤーもレシーバとならなかった場合、全ての出力層ユニットの教師信号は 0 となる。出力層ユニットの出力が 0.5 以上だった場合、その出力層ユ

ニットに対応するプレイヤーがレシーバであると予測され、0.5 未満だった場合、出力層ユニットに対応するプレイヤーはレシーバではないと予測されるものとする。

3.4 学習用データ

データは試合中に以下の方法で収集する。攻撃を担当するプレイヤーのいずれかがボールを持ち、フォーメーションが形成されたとき、ニューラルネットワークによるパスレシーバ予測を行い、自分がレシーバであると予測したプレイヤーは、10 サイクルの間ボールに向かって移動する。10 サイクルの間に以下の状況が発生した場合、予測が間違っていたとして、入力信号とそれに対する教師信号を学習用データに追加する。

- 自分以外の味方プレイヤーがパスを受け取る
- ボールを持つプレイヤーがドリブルを行う
- 敵プレイヤーがボールを持つ
- 10 サイクルの移動でボールが持てなかった

また、自分がレシーバではないと予測し、自分がパスを受け取った場合も、学習用データに追加する。試合終了後に、ニューラルネットワークをもつ 5 体のプレイヤーが収集したデータを 1 つにまとめて 1 試合分のデータとする。

3.5 数値実験

本論文では、異なる方法で学習用データを収集することにより、N1 から N4 までの 4 種類のニューラルネットワークを構築し、どの学習用データが学習に最適であるかを調査した。ニューラルネットワークの結合強度と閾値の初期化は -0.5 から 0.5 の一様乱数によって行うものとし、ニューラルネットワークの学習はそれまでに収集した全ての学習用データを使用して行うものとした。

N1 を作成した手順を以下に示す。

Step 1: 初期化

Step 2: 試合によるデータ収集

Step 3: 初期化

Step 4: 学習

Step 5: Step 2 から Step 4 までを試合回数が 10 回になるまで繰り返す

N1 は試合による学習用データ収集を行うごとに初期化して学習をやり直した。学習後のニューラルネットワークを使用して試合を行うことで、学習用データ収集時のパスレシーバ予測の精度を高めながら、データを収集できるようにした。

N2 を作成した手順を以下に示す。

Step 1: 初期化

Step 2: 試合によるデータ収集を 10 回行う

Step 3: 学習

N2 は、全ての試合で初期化されたニューラルネットワークを使用してデータ収集を行い、データ収集が全て終わった後で学習を行った。

N3 を作成した手順を以下に示す。

Step 1: 初期化

Step 2: 試合によるデータ収集

Step 3: 学習

Step 4: Step 2 から Step 3 までを試合回数が 10 回になるまで繰り返す

N3 は試合が終わるごとに追加学習を行い、途中で初期化はしなかった。これにより、以前の試合で学習できなかったデータを次の試合で収集するようにした。

N4 を作成した手順を以下に示す。

Step 1: 初期化

Step 2: N3 の学習用データを使用して学習

N4 は N3 作成の際に収集した学習用データを全て使用して 1 回だけ学習を行った。

3.6 実験結果

まず、前節で説明した 4 つのニューラルネットワークによるパスレーバ予測を行うチームで、ニューラルネットワークの差によってパスレーバ予測がどのように変化するかを調査した。具体的には、trilearn_base-3.4[7]とそれぞれ 5 試合ずつ試合を行い、パスレーバ予測の成功回数を計測した。Table 1 に結果を示す。Table 1 より、成功

Table 1: パスレーバ予測の成功回数

	成功	失敗	成功率 [%]
N1	843	429	66.3
N2	772	426	64.4
N3	851	308	71.0
N4	888	488	64.5

率が最も高かったのが N3 で、他のニューラルネットワークより 5%程度高い。また、N3 と同じ学習用データを使用した N4 の成功率が、N3 よりかなり低くなっているのが分かる。

次に、パスレーバ予測を行うチームと、パスレーバ予測を行わないチームの比較を行った。具体的には、trilearn_base-3.4 とそれぞれ 5 試合ずつ試合を行い、平

Table 2: 平均得失点と勝利回数

	得点	失点	勝利回数
N1	4.6	7.2	0
N2	3.4	5.2	1
N3	5.4	4.6	3
N4	5.4	6.6	2
使用しない	2.2	5.4	0

均得失点と勝利回数を調査した。Table 2 に結果を示す。Table 2 より、N3 と N4 の平均得点が最も高く、N3 の平均失点が最も低いことが分かる。N3 は勝利回数も最も多く、N3 が最もよい結果を出している。また、N1 から N4 いずれの場合も、平均得点がニューラルネットワークを使用しない場合を越えていることがわかる。

また、同試合における各チームのボール保持サイクル数を調査した。Figure 5 に結果を示す。本論文では、ボー

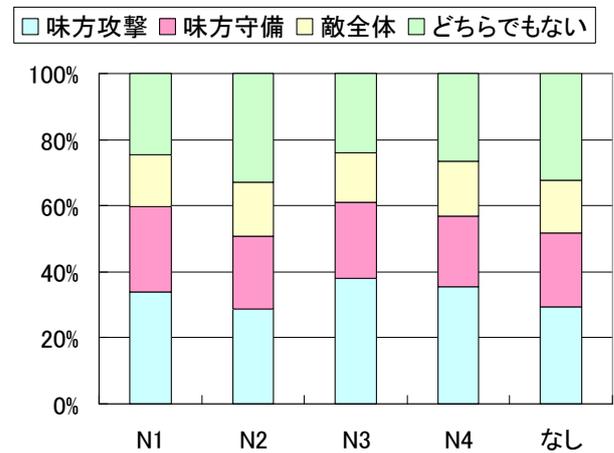


Figure 5: ボール保持サイクル数

ール保持サイクル数を、どちらかのチームがボールを相手チームに奪われずに保持していた時間とした。具体的には、あるプレイヤーがボールをキックできる状態で保持していた時間と、同一チームのプレイヤーの間でパスが行われていた時間とした。Figure 5 において、味方攻撃は攻撃を担当する味方プレイヤー、味方守備は守備を担当する味方プレイヤー、敵全体は敵チーム全体のボール保持サイクル数であり、どちらでもないはどちらのチームからもボール保持サイクル数として計測されなかった時間である。ただし、守備を担当する味方プレイヤーから攻撃を担当する味方プレイヤーへのパスの時間は、味方守備のサイクル数とした。Figure 5 より、攻撃を担当する味方のボール保持サイクル数は N3 が最も長く、ニューラルネットワークを使用しない場合の 1.21 倍となっている。これは、パスレーバ予測によって攻撃を担当する味方プレイヤーのパスの成功率が上昇しているためであると考えられる。また、N2 では攻撃を担当する味方プレイヤーのボール保持サイクル数

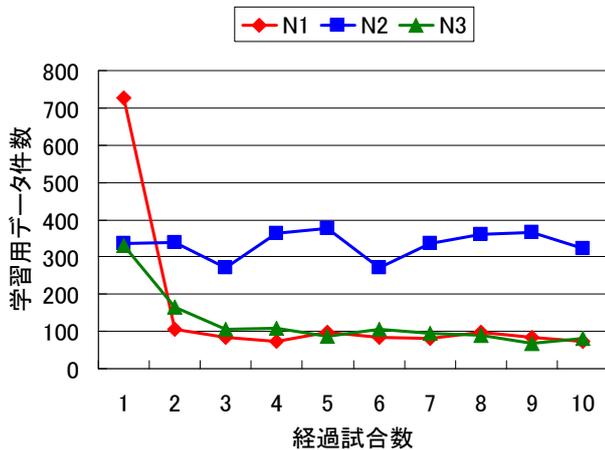


Figure 6: 1 試合ごとの学習用データ件数

は、ニューラルネットワークを使用しない場合とほぼ変わらない結果となっている。

また、Figure 6 に、N1, N2, N3 を作成する際の試合において得られた、1 試合ごとの学習用データの件数を示す。N1, N3 において、1 試合目のデータ数が非常に多く、N2 では全ての試合でデータ数が多くっており、ニューラルネットワークが初期化された状態で試合を行った場合、レシーバ予測の精度が低いためにデータ件数が多くなることがわかる。

そこで、学習の経過と予測失敗数の変化の関係を調査した。Figure 7, 8, 9 に、N1, N2, N3 を作成する際の試合中に発生した、1 試合ごとの予測の失敗件数の推移を示す。ここで、失敗パターン A を、自分がレシーバだと予測し自分がレシーバではなかった場合の予測失敗数、失敗パターン B を、自分はレシーバではないと予測し自分がレシーバだった場合の予測失敗数とする。Figure 7, 9 から、N1 と N3 において、失敗パターン A が 2 試合目に著しく減少し、2 試合目以降は緩やかに減少していることが分かる。一方で、失敗パターン B は初期段階からの発生が少なく、実験を通して変化が見られなかった。N2 では試合と試合の間で学習が行われていないため、失敗数は失敗パターン A, B とともに減少していない。

Table 1, Table 2, Figure 5 より、N3 が特に良い結果を残していることが分かる。これは、N3 がニューラルネットワークの学習を継続的に行うことで、学習用データを作成するための試合を行うときに学習できていないデータを効率よく抽出し、学習していたためと考えられる。

4 おわりに

本論文では RoboCup サッカーシミュレーション 2D リーグにおいて、ニューラルネットワークを用いて、パスレシーバの予測を行った。数値実験により、攻撃を担当するプレイヤーがパスレシーバを予測することで、ボール保持サイクル数が増加し、平均得点数が上昇することを示した。また、学習を継続的に行うニューラルネットワークが最も

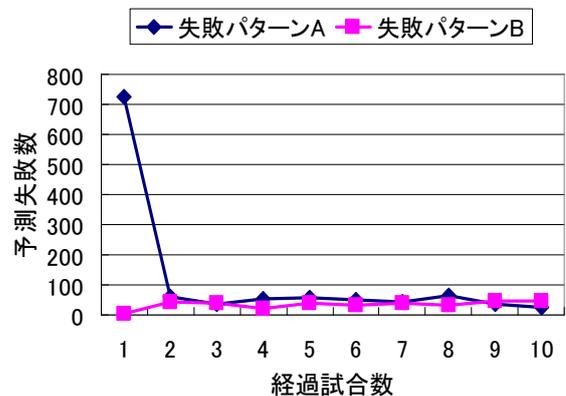


Figure 7: 1 試合ごとの予測の失敗数 (N1)

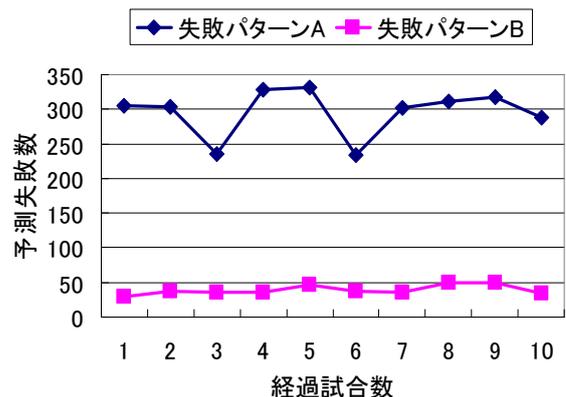


Figure 8: 1 試合ごとの予測の失敗数 (N2)

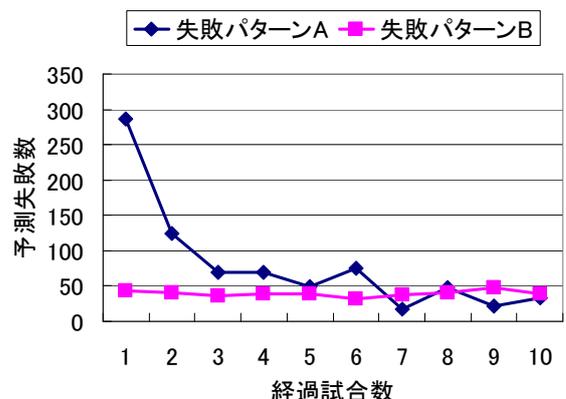


Figure 9: 1 試合ごとの予測の失敗数 (N3)

よい結果を出すことを示した。今後の課題としては、学習用データをさらに増やすことでニューラルネットワークがどのように変化するかを調査することなどが挙げられる。

参考文献

- [1] Ashkan Radmand, Eslam Nazemi, Mohammad Goodarzi, "Integrated Genetic Algorithmic and Fuzzy Logic Approach for Decision Making of Police Force Agents in Rescue Simulation Environment," RoboCup Symposium 2009, (2009)
- [2] Tobias Warden, Andreas D. Lattener, Ubbo Visser,

“Real-Time Spatio-Temporal Analysis of Dynamic Scenes in 3D Soccer Simulation” RoboCup Symposium 2008, (2008)

- [3] Hidehisa Akiyama , Hiroki Shimora , Itsuki Noda , “HELIOS2008 Team Description,” RoboCup2008 , (2008)
- [4] Yasushi Nakabayashi , Daisuke Hanyu , Masato Masuda , “NCL08,” RoboCup2008 , (2008)
- [5] Thomas Gabel , Martin Riedmiller , “Brainstormers2D Team Description 2009,” RoboCup2009 , (2009)
- [6] Vadim Kyrylov, Edde Hou, “Pareto-Optimal Collaborative Defensive Player Positioning in Simulated-Soccer,” RoboCup Symposium 2009, (2009)
- [7] UvA Trilearn, “<http://staff.science.uva.nl/~jellekok/robocup/>”