

動的位置決定法に対する相互結合型ニューラルネットワークの適用*

玉城 斉 遠藤 聡志 山田 孝治†

琉球大学工学部‡

1 はじめに

マルチエージェントシステムにおける問題解決においては、動的に変化する環境の下で各エージェントを効果的に配置することが重要である。これは、各エージェントが自律的に移動位置を決定する動的な位置決定問題として扱うことが出来る。

本研究では、エージェントの位置関係と協調行動の関係について検討を行なうため、マルチエージェント協調問題として動的な位置決定問題を採用し、相互結合型ニューラルネットワーク及び競合共進化を用いたマルチエージェントシステムに対する動的な位置決定法を提案する。提案手法を SoccerServer[1] に適用し、エージェントの位置獲得について評価、検討する。

2 動的な位置決定問題

2.1 目的と制約

目的 エージェントが環境の変化に応じて互いの位置関係を適切に変更することで、システムの評価を最大化する。

制約

- ・エージェントの環境情報は制限されている。
- ・位置関係がシステムの評価に直接結び付かない。

2.2 SoccerServer における動的な位置決定問題

本研究では、動的な位置決定問題を提供するマルチエージェントシステムの一例である SoccerServer を対象とする。

SoccerServer における目的と制約は以下の様に見える。

目的 SoccerServer における両チームは得点及び失点による評価の最大化を目的とする。

制約

- ・視界は制限され、視界内の物体との相対位置情報と自身の位置情報を次の移動位置決定に利用可能である。
- ・システムは、得失点によって評価される。

この制約下で各々のチームの目的を達成するための有効な行動決定を実現するための方法論が必要となる。

2.3 性質

動的な位置決定問題共通の性質として、以下の2つが挙げられる。

- (1) 各エージェントに与えられる環境情報の局所性。
- (2) 個々のエージェントの評価の総和とシステムの評価の不一致

*Dynamic positioning method based on Recurrent neural network

†T.Tamashiro S.Endo K.Yamada

‡Faculty of Engineering, University of the Ryukyus

3 マルチエージェントの動的な位置決定

図1に動的な位置決定法の概念図を示す。動的な位置決定法では、各エージェントは相互結合型ニューラルネットワークによって行動決定を行なう。相互結合型ニューラルネットワークは時系列データ処理によって過去の情報の活用が可能である。相互結合型ニューラルネットワークの学習に競合共進化を適用することで、全体の試行を通じたシステムの評価を行なう。

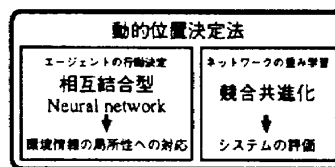


図1: 動的な位置決定法

3.1 相互結合型ニューラルネットワークの適用

動的な位置決定問題では、環境情報の局所性によりエージェントは環境を正しく認識できない。ここで、制限された環境情報から適切に行動を行なうために過去に得られたデータを活用することは有用である。

相互結合型ニューラルネットワークでは、フィードバックループを用いることで隠れ層に保持された過去の入力情報の利用が可能であり時系列モデルを表現することができる(図2)。

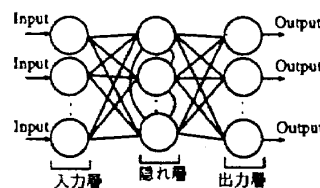


図2: 相互結合型ニューラルネットワーク

3.2 競合共進化によるニューラルネットワークの重み学習

3.2.1 動的な位置決定問題における競合共進化

動的な位置決定問題では、評価対象の違いによってシステムの評価が大きく異なる。また、評価対象を固定にすると学習が局所解に陥り易くなるためその評価が困難である。このような問題解決のために本研究では、競合共進化を用いる。競合共進化アルゴリズムは遺伝的アルゴリズム (GA)[3] を用いた進化学習法である。

動的な環境では得られた環境情報に対する正しい行動の同定は困難であるため、バックプロパゲーションなどの教師付学習形式で必要となる教師データの作成が困難である。GA では、全体の試行を通して評価を行なうので教師データの作成の必要がない。また、個々のエージェントを評価するのではなくマルチエージェントシステムの全体の評価が行なえるため本問題に対して有効である。

3.2.2 競合共進化アルゴリズム

競合共進化アルゴリズムを以下に示す。

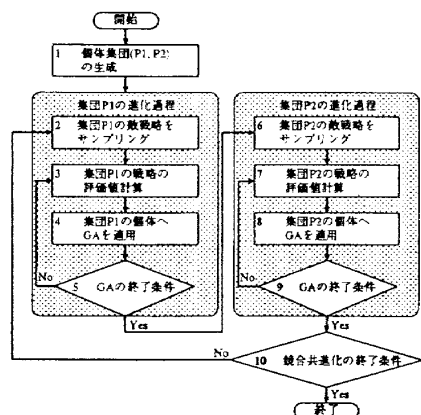


図 3: 競合共進化アルゴリズム

step1では、個体集団(P1,P2)を生成する。アルゴリズムは、集団P1の進化過程と集団P2の進化過程を交互に実行する。集団P1の進化過程において、集団P1を進化個体集団、集団P2を評価個体集団とする。step2で集団P1の比較相手として集団P2から個体をサンプリングする。step3において、集団P1の全個体は、サンプリング個体と優劣比較を行ない、その結果から評価値を計算する。その評価値を適応度として、step4で集団P1の個体に対してGAを適用する。step5で、終了条件が満たされていなければ再び集団P1を評価個体群と進化個体群とし、step2~step5と同様の進化過程をstep6~step9で実行する。step2~step9までを競合共進化1世代とし、競合共進化の終了条件を満たすまで繰り返す。

4 計算機実験

4.1 目的

プレイヤー間の位置関係の学習を目的として SoccerServer を用いてプレイヤーを攻撃側、防御側に分け、3対3として試合を行なわせる。各プレイヤーはチームのボールの支配率を最大化することを目的とする。

4.2 エージェントの設計

エージェントの出力には、次の移動位置、ボールが自らの影響範囲内にある場合にはボールに対する行動が出力される。

ボールを保持するプレイヤーの行動は以下のように定義する。

case1: 敵ゴールが視野に存在する場合

ゴールに向かって蹴る。

case2: 味方が視野に存在する場合

味方に向かってボールを蹴る。

otherwise:

次の移動位置に向かってボールを蹴る。

エージェントの移動位置は距離と角度の2パラメータでニューラルネットによって出力される。

エージェントの入力には環境情報として視野の範囲内にある物体の情報及び現在位置が与えられる。物体の種類はボール、味方エージェント*2、敵エージェント*3、近接したエージェント*1の7種類である。また視覚情報は物体との距離、敵ゴールの方向を0度とし右向き逆時計回りを正方向としたときの絶対角度が与えられる。例えば図4の状態ではエージェントは、視界内のボール、味方

エージェント敵エージェントの位置が(距離, 角度)で与えられる。また、現在位置はフィールドの中心からの位置として与えられる。

本研究で用いる相互結合型ニューラルネットは入力層、隠れ層、出力層それぞれ16,16,2のユニットから成る。

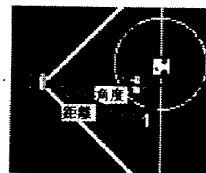
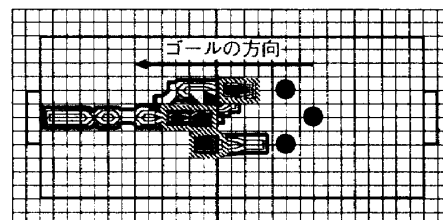


図 4: 視界

5 実験結果及び考察

以下に学習実験終了後任意のチームを対戦させた結果を示す。図5は、試合の中盤(開始10-20秒)における各プレイヤーのフィールド上の分布図である。



□: 攻撃側プレイヤーの分布 ○: 防御側プレイヤーの分布 ●: 攻撃側プレイヤーの初期位置

図 5: プレイヤーの分布

図より攻撃側プレイヤーは、ゴールに向かって進む行動を、防御側プレイヤーは、攻撃側の進行を妨げるポジションを学習していることが分かる。

以上の結果から、競合するチームに対応した位置関係を学習していると考えられる。

6 おわりに

本稿では、マルチエージェントシステムにおけるエージェントの位置関係と協調行動の関係について検討を行なうため、SoccerServerにおける動的決定問題に対し、相互結合型ニューラルネットと競合共進化アルゴリズムを用いた配置法を提案し、計算機シミュレーションにより獲得された位置関係について検討を行なった。

謝辞

本研究の一部は、財団法人テレコム先端技術研究支援センターの支援により実施した。

参考文献

- [1] I.Noda and H.Matsubara: "Soccer Server and Researches on Multi-Agent Systems", Proc. of IROS-96 Workshop on RoboCup, pp. 1-7, (1996)
- [2] 村田哲哉, 鈴木恵二, 大内東, "GAによるサッカーエージェントの動的配置探索問題に関する基礎研究", 第7回 インテリジェント・システム・シンポジウム, (1997)
- [3] 北野宏明, "遺伝的アルゴリズム", 産業出版, (1993)
- [4] 高林純示, 西田律夫, 山岡亮平, "共進化の謎に迫る-化学で見る生態系-", 平凡社, (1995)
- [5] Hugo de Garis, "Genetic Programming: Evolutionary Approaches to Multistrategy Learning", animal to animat, pp. 549-577