

# 行動選択ネットワークを用いたマルチエージェント系 における協調行動の獲得に関する考察\*

与那覇 賢 遠藤 聡志 山田 孝治†

琉球大学 工学部 情報工学科‡

## 1 はじめに

実世界のような巨大で複雑な環境において、単一のエージェントが自身の目的を達成することは困難である。マルチエージェントシステムは、このような問題を複数のエージェントの相互作用から創発される協調行動によって効率良く解決するためのフレームワークである。近年、サッカーゲームがマルチエージェントシステムの研究モデルとして取り上げられている [Noda and Matsubara 96],[Noda 96a]。サッカーでは複数のエージェント (プレイヤー) は、試合に勝つというチーム全体としての目標を達成しようと試みるが、この際、協調行動すなわちチームプレーが勝敗の大きな要因となる。本研究の目的は、エージェントに他のエージェントとの協調を考慮した基本的な行動選択能力を与えることである。

## 2 局所的状況の学習における問題点

本研究における基礎実験として、サッカーを例題に、局所的な状況でのパスとシュートの行動選択の学習実験をニューラルネットワークを用いて行なった。その結果、局所的なパスとシュートの成功率はある程度上昇したが、サッカーゲームによる評価では、ランダムにパスとシュートを選択するプレイヤーと比較して、性能に大きな差が見られなかった。これは、局所的な状況での学習で得られた行動選択ルールが、サッカーゲームという大きな問題空間ではうまく働かないためと考えられる。また、ニューラルネットワークによる学習でより多くの行動選択ルールを獲得することは、ネットワークの構造やサイズなどの点から困難になると考えられる。この問題を改善するため、エージェントの行動選択機構として、行動ネットワークを適用する。

\*Acquisition of Cooperative Action Based on Behavior Network in Multi-Agent System

†Satoru Yonaha, Satoshi Endo, Koji Yamada

‡Faculty of Engineering, University of the Ryukyus

## 3 Agent Network Architecture

行動選択ネットワークでは、一つの行動が一つのノードとして定義され、各ノード間は重み付き結合で接続されている。このネットワークにおいて、行動選択は刺激反応系に基づいて行なわれる。本研究では、Pattie Maesによって提案された Agent Network Architecture (ANA) [Maes 91] を行動選択ネットワークとして採用する。

ANA では各ノード (本稿では行動モジュールと呼ぶ) は、図 1 に示すように活性度、前提条件、追加リスト、削除リストで定義される。

複数存在する行動モジュールのうち、前提条件を満たし、しきい値を越えた活性度の最も高いモジュールが動作する。このとき、動作は状態 (事実の集合) に対する書き換え (事実の追加・削除) によって行なわれる。

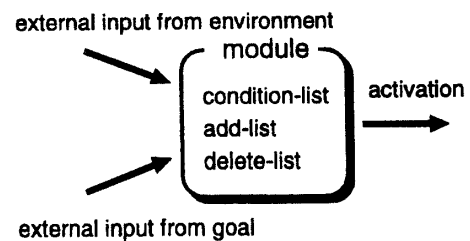


図 1: 行動モジュールの定義

行動モジュールに対する活性度の入力は、状態と目標から与えられる外部入力と、モジュール間の活性度伝搬によって得られる入力の2種類が存在する。状態からの入力は、状態と前提条件が一致するモジュールに対し、その一致する度合に比例して入力される。また、目標からの入力は目標の状態とモジュールの追加リストの状態が一致し、その目標がまだ達成されていない場合に、そのモジュールに対して入力される。

時刻  $t$  でのモジュールの活性度は次の式で表される。

$$\begin{aligned}
 \text{activation}(\text{module}, t) \approx & \\
 & D \sum_i \frac{1}{d(\text{module}, \text{sensor\_datum}_i)} \\
 & + G \sum_j \frac{1}{d(\text{module}, \text{goal}_j)} \\
 & - C \sum_k \frac{1}{d(\text{module}, \text{protected\_goal}_k)} \\
 & + M \cdot \text{activation}(\text{module}, t-1)
 \end{aligned}$$

パラメータ  $D, G$  によって、モジュールをセンサデータ指向的あるいは目標指向的に動作させることができる。また、パラメータ  $C$  は活性度を減少させる割合であり、 $M$  は減衰定数である。このように、各モジュールがネットワークを構成し、互いに活性度伝搬を行なうことによりネットワークは動作する。

ANA の特徴として、

- 現在の環境や目標からプランが生成される
- プランの推論や探索を行なわない
- 行動選択の速さ
- 拡張性

などが挙げられる。これらの特徴から、ANA は環境に対して即応性を持った行動選択機構であることが分かる。

本研究は、エージェントの基本的な行動モジュールに加え、他のエージェントとの協調を考慮したモジュールを追加することによって、マルチエージェント環境において協調行動を獲得することを試みるものである。

## 4 サッカーエージェントの設計

行動選択ネットワークの評価例題として、サッカーゲームをとりあげる。サッカーにおけるエージェント（プレイヤー）の基本的な行動要素を以下のように定義する。

- ドリブル (dribble)
- シュート (shoot)
- パス (pass)
- ボールを追いかける (chase ball)
- 適切なポジションへの移動 (move)

他のプレイヤーとの協調を考慮した行動要素として、パスと適切なポジションへの移動があり、これらはともに試合に勝つという共有された目標を円滑に達成するために必要な行動要素であると考えられる。

これら5種類の行動要素をそれぞれ行動モジュールとみなし、行動選択ネットワークを図2のように構築する。この図ではそれぞれの行動モジュールに対して successor

link のみを示し、predecessor link および conflictor link は省略されている。

また、現段階では各エージェントに対し、機能差などによる差別化は行なわない。すなわち、各エージェントが状態に応じて適切な（活性化された）行動が選択することで、協調行動が創発されることを期待する。

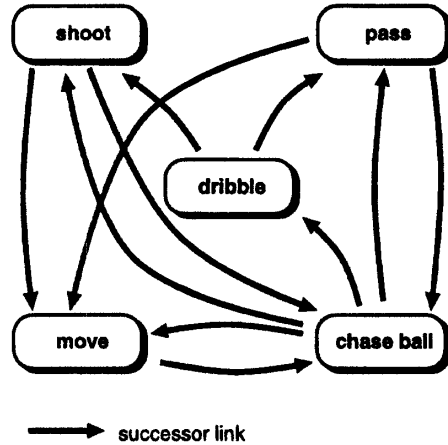


図2: サッカーエージェントの行動選択ネットワーク

設計したサッカーエージェントを評価するための実験環境として、Soccer Server [Noda and Matsubara 96] を用いる。Soccer Server は野田らによって提供されているサッカーゲームのシミュレータであり、Soccer Server 上で試合を行なうことでマルチエージェント系としての評価を行なうことができる。

## 5 おわりに

本稿では、エージェントの行動選択機構として、Agent Network Architecture に焦点を当て、その特徴について述べた。また、サッカーを例題とし、他のエージェントとの協調を考慮した行動要素を含めたエージェントの設計についても述べた。現在、計算機環境へのエージェントの実装と基礎的な評価実験の設計を行っている。

## 参考文献

- [Noda and Matsubara 96] I.Noda and H.Matsubara: "Soccer Server and Researches on Multi-Agent Systems", Proc. of IROS-96 Workshop on RoboCup, pp. 1-7, 1996.
- [Noda 96a] I.Noda, H.Matsubara and K.Hiraki: "Learning Cooperative Behavior in Multi-agent Environment", Proc. of PRICAI'96, pp. 570-579, 1996.
- [Maes 91] P.Maes: "The Agent Network Architecture (ANA)", SIGART Bulletin, Vol. 2, No. 4, pp. 115-120, 1991.