

フォーカルポイントに基づいた マルチエージェント環境における協調の実現*

与那覇賢 遠藤 聡志 山田 孝治†

琉球大学工学部‡

1 はじめに

複数のエージェントによって協調的にタスクを処理するマルチエージェントシステムでは、エージェントらの合意形成による最適な意志決定が重要である。しかし、環境情報に不確実性が含まれている場合や、明示的な通信が困難な環境では、合意形成による協調作業の実現は容易ではない。そのような場合、状況や場の必然からエージェントらの選択が自然に一致する点であるフォーカルポイントを用いた合意形成手法が有効である [1, 2]。マルチエージェントの研究分野における標準問題として認知されているサッカー [3] では、例えばパスを成功させるために「どこにパスを出すのか」あるいは「どこで受けるのか」というエージェント間での合意形成の重要性を示唆する問題が含まれている。

本稿では、フォーカルポイントの学習に強化学習の手法を用いる枠組を提案し、協調行動が要求されるマルチエージェント環境において、学習したフォーカルポイントを用いた合意形成によりタスクの達成が可能であることを示す。

2 フォーカルポイント

エージェントが取りうる選択の集合 A を

$$A = \{a_1, a_2, \dots, a_n\} \quad (1)$$

とする。エージェント A とエージェント B がそれぞれ選択 a^A, a^B を行ない、その選択を一致させる、すなわち

$$a^A = a^B \quad (2)$$

である場合を合意の形成と呼ぶ [2]。このとき一致した選択がフォーカルポイントとなる。

環境から得られる情報がノイズなどによって不確実であるような場合や、エージェントの視野が限定されている場合は、各エージェントに対する入力がそれぞれ異なるため決定的な選択アルゴリズムでは合意の形成は困難である。

本研究では、ある状態に対する行動の集合においてエージェントが選択すべき対象 $a \in A$ を確率的に選択することで不完全情報の影響を吸収し、合意の形成を行う手法を提案する。

3 強化学習を用いたフォーカルポイントの学習

3.1 強化学習

強化学習は環境内において試行錯誤的に行動し、種々の行動に対しての報酬信号を得て、最も高い報酬信号値

を得た行動を生成するように環境情報から行動への写像を内部に形成する学習手法である。

一方、フォーカルポイントは社会的な場の状況、構図から規定されるものとして、社会的側面から研究が行なわれている。これらの理由から、強化学習の生物の経験的、試行錯誤的な行動様式に従った学習メカニズムが有効であると考えられる。

標準的な強化学習のモデルでは、エージェントは環境、視覚、行動によって関連づけられており、試行によって増加していく報酬信号の総和によって行動を決定し、機械的に学習することが可能である。

3.2 定式化

本研究では基本的な強化学習の枠組を用いる。以下でその定式化を行なう。

エージェントが観測した環境 s_i に対して、エージェントが選択し得る行動の集合を (1) 式で定義する。このとき集合 A から 1 つの行動が選択されるが、この選択は確率的に行う。すなわち、エージェントが状態 s_i において行動 a_j を選択する確率を p_j^i とすれば、状態 s_i において

$$P^i(s_i) = \{p_1^i, p_2^i, \dots, p_j^i, \dots, p_n^i\}, \left(\sum_{j=1}^n p_j^i = 1.0 \right) \quad (3)$$

となる。

また、エージェントが行動 a_j を選択した結果、環境から与えられる報酬信号 R は a_j の選択される確率、すなわち p_j^i を強化するものとする。ここで、報酬信号によって $p_j^i + R$ だけ強化されると、他の確率パラメータ p_k^i は

$$p_k^i = \frac{1 - (p_j^i + R)}{1 - p_j^i} \times p_k^i \quad (k \neq j) \quad (4)$$

で更新される。状態遷移の不確実性や不完全知覚の問題を内包するマルチエージェント環境においては、行動選択の指針として確率的な接近が有効であることが報告されており、不完全知覚状態や局所解等からの脱出が期待できる [4]。

そこで、複数エージェントの存在によって状態遷移の不確実性などが生ずる環境としてサッカーを取り上げ、協調的にタスクを達成するために必要なフォーカルポイントの学習に強化学習を適用する。

4 サッカーにおける合意形成問題

4.1 問題設定

サッカーゲームにおいて、合意形成による問題解決が要求される場面として、パス問題を取り上げる。

図 1 において、Passer は Receiver へパスを出すことを目的とし、Receiver はパスを受け、相手ゴールへ

*An Agreement in Multi-agent System Based on Focal Point Technique

†S.Yonaha, S.Endo and K.Yamada

‡Faculty of Engineering, University of the Ryukyus

シュートすることを目的とする。敵味方の位置関係によって、味方プレイヤーへの直接のパスが有効な場合や、敵プレイヤーの存在しない場所へのパス（スルーパス）が有効な場合がある。どちらの場合も Passer と Receiver の間で「誰に（どこに）パスを出すのか」という合意の形成が行なわれていなければ、パスは成功しないと考えられる。すなわち、環境に存在する複数のフォーカルポイントの候補（パスを出す場所の候補）から最もパスが成功すると思われる場所を Passer と Receiver のフォーカルポイントとして学習しなければならない。本問題設定において、フォーカルポイントの候補は {Receiver1, Receiver2, Flag1, Flag2, Flag3, Goal} であるものとする。

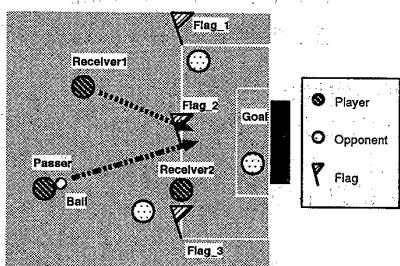


図 1: パス問題

4.2 計算機実験

実験環境として SoccerServer[3] を用い、図 2 に示す内部構造をもつエージェントを設計する。SoccerServer では物体の位置情報は連続値であるため、前処理としてこれを離散値へ置き換える。この情報はニューラルネットワークによって物体の位置関係をいくつかのパターンへ分類することで強化学習への入力となる状態空間を単純化する (図 3)。

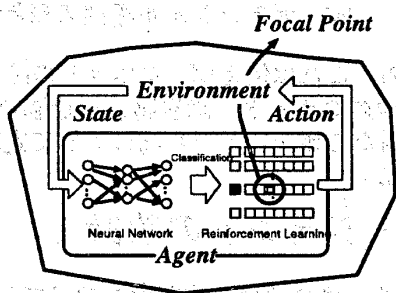


図 2: エージェントの内部構造

エージェントは、前述の学習済みのニューラルネットの出力によって得られる味方プレイヤー、敵プレイヤー、旗の位置情報の組合せを強化学習システムへの入力とする。また、状態に対するルールはパスを出す場所の集合 {Receiver1, Receiver2, Flag1, Flag2, Flag3, Goal} とする。

計算機実験の手順は以下の通りである。

- (1) ボールを持つ攻撃側のプレイヤー (Passer) は、強化学習の出力に従ってパスを出す。
- (2) 他の攻撃側プレイヤー (Receiver) も同様に出力に従ってその位置へ移動し、シュートを試みる。
- (3) 守備側のプレイヤーは近くにあるボールを追いかけ、相手陣地へ蹴り返す。
- (4) 一定時間 (約 10 秒) で 1 試行は終了する。

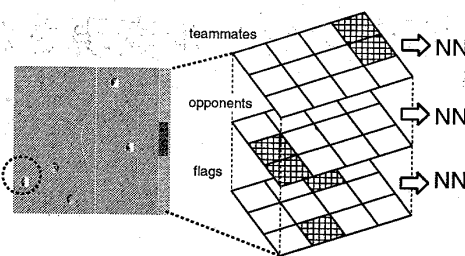


図 3: 状態空間の単純化

以上の手順をプレイヤーの配置を変えて 4500 回繰り返す。試行の結果に従って、攻撃側の各プレイヤーには報酬が与えられる。

4.3 実験結果および考察

新たに用意した事例数 2000 の 3 対 3 の状況における学習前と学習後のプレイヤー群のセットプレーの成功率の比較を表 1 に示す。この結果から、学習後のプレイヤー群の成功率が約 27% 程向上したことが分かる。学習後において成功率が約 5 割程度に留まった理由として、フォーカルポイントの候補が事前に与えられており、敵プレイヤーの存在を考慮しないものであったためと考えられる。例えば、Flag はあくまでもフィールド内の絶対的な場所を示す物体であり、必ずしもセットプレーの成功へ寄与する存在とは限らない。フォーカルポイントを用いて円滑に協調作業を達成するためには、エージェントが環境内から自律的にフォーカルポイントを獲得することで合意形成を図る枠組が必要になると考えられる。

表 1: セットプレー成功率の比較 (事例数:2000)

	success	fail
学習前	22%(439)	78%(1561)
学習後	48.8%(976)	51.2%(1024)

5 おわりに

本稿では明示的な通信を用いることができない環境において、通信を用いない合意形成手法であるフォーカルポイントに注目し、強化学習の枠組を適用するとその学習を行うことについて述べた。また、マルチエージェント環境で協調行動が要求される例としてサッカーの一場面を取り上げ、フォーカルポイントを用いることで作業が達成できることを述べた。

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] Maier Fenster, Sarit Kraus, Jeffrey S. Rosenschein: Coordination without Communication Experimental Validation of Focal Point Techniques, ICMAS95, (1995).
- [2] 本村陽一: 確率モデルの学習によるエージェント間の合意形成, MACC'97, (1997).
- [3] I.Noda and H.Matsubara: Soccer Server and Researches on Multi-Agent Systems, Proc. of IROS-96 Workshop on RoboCup, pp.1-7, (1996).
- [4] 荒井幸代, 宮崎和光, 小林重信: マルチエージェント強化学習の方法論-Q-Learning と Profit Sharing による接近-, 人工知能学会誌, Vol.13, No.4, pp.609-617, (1998).