

学習機能を持つ自律エージェントによる

4N-3

協調的集団行動の創発*

坂本 忠昭†

三菱電機(株) 中央研究所‡

1 はじめに

環境とのインタラクションを通して自分自身の行動規則を生成し洗練させていくという適応的な学習方法は、予め行動規則を与える必要がない、環境の変化に柔軟に対応できる、といった特徴を持ち、特に人工生命の分野で多くの研究が行われている。本研究の目的は、このような適応学習の枠組を用いて、複数の自律エージェントに協調的な集団行動パターンを自発的に生成（創発）させることである。

2 問題設定

協調的な集団行動を生成させるための問題設定として、分散人工知能の小問題の一つである追跡問題[1]の変形版を考える。まず、無限に広がる2次元平面上を自由に動き回る2種類のエージェントを仮定する。各エージェントは自分の持つ視覚機能によって、自分周辺の局所的な情報を得ることができるものとする。一方のエージェントは捕食型のエージェントであり、他方の被食型のエージェントを捕まえることを目的としている。ただし、全てのエージェントの速度を一定としているため、単に捕食エージェントが被食エージェントの後ろを追いかけたのでは捕まえることはできない。そこで、例えば4つの捕食エージェントが1つの被食エージェントを囲い込み、捕まえるといった協調的な集団行動が必要になってくる。

3 エージェントの設計

全てのエージェントは任意の角度と距離で決定される扇型の視界を持ち、視界の中の情報だけから次の行動を決定する。エージェントが可能な行動は、自分の速度ベクトルの方向を変化させることである。以下、被食型、捕食型の各エージェントについて説明する。

被食エージェント

被食エージェントは、視界の中にいる捕食エージェントの位置と速度ベクトルからその捕食エージェントの次の時刻の位置を予測し、その位置から遠ざかるよ

うに次の方向を決定する。複数の捕食エージェントが視界にある場合には個々の捕食エージェントに対して同様の計算をし、得られた方向を合成する。ただし、近くの捕食エージェントから遠ざかる方向が、遠くの捕食エージェントから遠ざかる方向よりも強く影響するように距離に応じた重みを加えている。このように、被食エージェントには学習機能を持たせておらず、その行動パターンは一定規則に従っている。なお、視界の半径は捕食エージェントのそれよりも短く設定されている。

捕食エージェント

捕食エージェントは、行動規則を学習する機能を備えている。図1に示すようなメッシュ状の視界を持ち、視界に入ってきた他のエージェントの位置と種類から視覚データを生成する。視覚データはそのまま行動規則の条件部として用いられる。図1の視覚データは、メッシュの5の位置に被食エージェント（種類1）、6の位置に仲間のエージェント（種類2）が見えていることを表す。また、エージェントが選択できる移動方向は図1に示されるように番号付けされており、これが行動規則の行動部として用いられる。

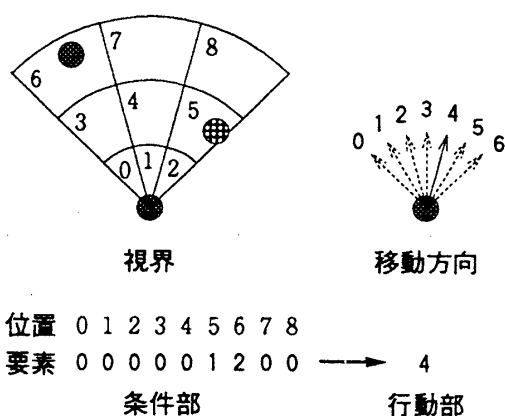


図1: 行動規則の表現

*Emergence of Cooperative Group Behaviors by Autonomous Agents with Learning Ability

†Tadaaki SAKAMOTO

‡MITSUBISHI Electric Corporation

次に捕食エージェントの内部構成を図2に示す。行動決定部は視覚データから行動データを決定する部分

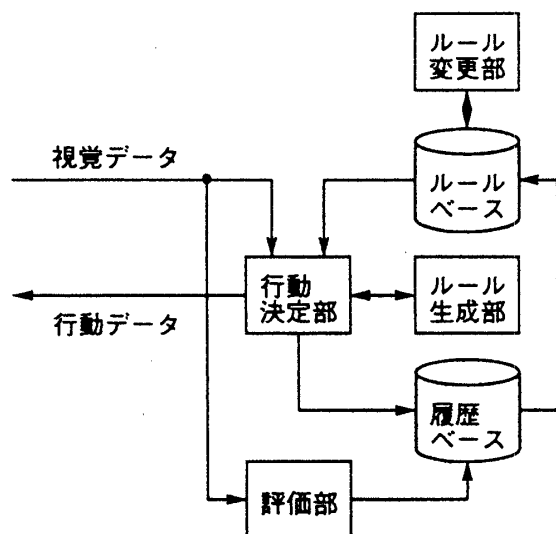


図 2: エージェントの内部構成

である。視覚データとマッチする条件部を持つ行動規則がルールベースにあれば、その行動部が行動データとなる。マッチする行動規則がなければ、ルール生成部で視覚データとランダムに生成された行動データから新たな行動規則が生成される。新たに生成されたものも含めて、用いられた行動規則は履歴ベースに蓄積される。履歴は、視界に被食エージェントが入った時点から被食エージェントを捕まえたか逃げられた時点まで記録される。なお、履歴ベースは有限長であり、それを超えた場合には古い履歴から順に削除される。

評価には、各行動規則の適用ごとに自分が被食エージェントに接近したかどうかによって与えられる短期評価と、被食エージェントを結果として逃がしたか捕まえたかによって与えられる長期評価がある。短期評価は直前に適用された行動規則に対してのみ与えられ、長期評価は過去に遡って複数の行動規則に対して与えられる。履歴ベースで評価を与えられた行動規則は再びルールベースに戻される。新しく生成された行動規則でも、この時点で評価値の低いものはルールベースに蓄積されず破棄されることがある。なお、現時点では長期評価はまだ実装されていない。ルール変更部は、ルールベースが一杯になった時に起動され、行動規則の一般化を行う。

4 実装と実験

現在、上記の仕様に基き、C++を用いてSPARC station2上に実装中である。短期評価だけを実装した段階で、試験的に4つの捕食エージェントと1つの被食エージェントによる追跡のシミュレーションを行っ

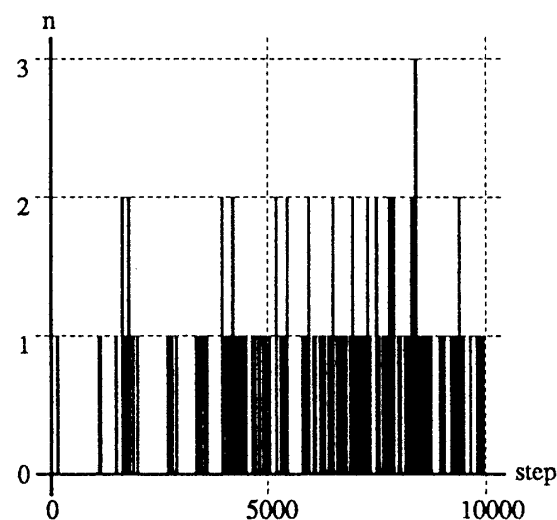


図 3: 追跡実験結果

た。その結果を図3に示す。図の横軸は時間、縦軸は被食エージェントの視界（360度）に入った捕食エージェントの数を表す。図より、時間が経つにつれて被食エージェントの視界に頻りに入ってくる、つまり、追いかける方がうまくなっていることが分かる。ただし、自分が相手に接近するような行動規則ほど評価が高いため、とにかく目の前の被食エージェントに向かって進むという行動ばかりが目につき、まだ協調的な行動が現われるまでには至っていない。そのため、ほとんどの場合が被食エージェントに逃げ切られてしまっている。これらが長期評価の実装によってどのように改善されるかという点が今後の課題の一つである。

5 おわりに

本稿では、適応的な学習機能を持つ自律エージェントにおいて、追跡問題を用いて協調的な集団行動を自発的に生成するという試みについて述べた。引続き実装を行い、上記の長期評価導入の効果を確認していく予定である。今後の課題としては、他のエージェントの速度ベクトルを認識したり、視界のメッシュを細かくしたりする視覚情報の詳細化があげられる。ただし、視覚情報の詳細化は探索空間の拡大につながるため注意する必要がある。さらに、捕食エージェント間に合図のような通信機能を導入することについても検討していきたい。

参考文献

- [1] 大沢 英一, 沼岡 千里, 石田 亨: 分散人工知能における標準的小問題, コンピュータソフトウェア, Vol.10, No.3, 1993.