

Q 学習によるフルートレス生命体の合体による形態形成の研究

北海道工業大学 ○木暮 直記, 川上 敬, 中川 嘉宏, 木下 正博

要旨

本研究では、フルートレス遺伝子の特性を要する複数のエージェントが結合、あるいは接近する際にどのように形態を形成しながら行動決定を行うかについて考察する。具体的には、タイルワールドにおいて Q 学習による移動行動を行わせ、フルートレス遺伝子によってエージェントの性行動が変化することによる結合基準が形態に及ぼす影響について議論する

1. はじめに

自然界では、生物が群れを成し適切な形態を形成することで問題を解決するという現象が多く存在する。生物が群れを成すという行動は生得的な本能行動でありながらも非常に複雑で分析が困難である。近年では、群行動の原理を解析し、複数ロボット間の協調行動に応用しようという動きも見られる。また、動物における群体の一つに雄と雌によるものがあるが、性行動を決定すると言われるフルートレス遺伝子の変異により、例えば雄同士がまるで雄と雌のような関係性のまま群を成すことも起こり得るように、個体の特性変化により、群体の振る舞いにも変化が起こる可能性がある。

そこで本研究では、群行動における形態形成を分析する一つの手法として学習機能とフルートレス遺伝子を搭載した複数エージェントによる振る舞いの実験及び考察を行う。

2. 個体の結合と形態

現在、群体の制御方法がいくつか考案されている。その一つは集中型・階層型制御でトップダウンの構造であり、トップが行動命令を群体の各個体に伝達することで全個体を協調させるものである。もう一つは非集中型・局所型制御と呼ばれるもので、群体はそれぞれが自律的な個体で形成されている。この制御では、互いの個体が協調しているように見えるが、実際にはそれぞれの個体と環境の相互作用による創発された行動の集合である。また、これらを組み合わせ合わせたハイブリッド型制御が存在する。

グループ全体であるグローバルな目標を達成するのであれば、どの制御方法においても各個体間は通信を行う必要がある。通信手段には3つあり、一つは P2P 型通信と呼ばれるもので、個体が一対一で通信を行うものである。二つめはブロードバンド型通信と呼ばれるもので、一つの個体が全個体に対して通信を行う一対多のものである。三つめは環境を介した通信手段である。情報を直接的にやり取りせずに、環境に込めることで通信を行う手段である。例として、蟻の採餌行動におけるフェロモンの散布が挙げられる。これは明示的には通信を行っていないため、通信なしの協調と呼ばれることもある。また、明示的な通信を行わないものには、個体に内蔵されたセンサにより互いを感知する方法も存在する。

3. フルートレス遺伝子

フルートレスとは、突然変異で生まれた遺伝子で、生物の性行動を管理する存在である。フルートレスは元々、雄の性行動を劇的に変化させる突然変異の遺伝子として「フルーティ」という名前で登場した。フルーティの雄は、雌だけでなく雄にも求愛行動をし、しかも雌と交尾しないと

いうものであった。

キイロシヨウジョウバエの性行動を作り上げる遺伝子を解明するため、性行動に変化を引き起こす突然変異体を人為的に誘発する実験が行われた。その結果大変元気で見たところ他と変わらない系統ながら、雄は雌に関心を示さず、まるで求愛の様子を見せないという変わった突然変異系統が発見された。この系統の雄は性欲を失ったものと思われ「サトリ」(悟り)と名付けられた。その後の研究から、サトリ変異体は雄の同性愛者である、サトリとフルートレスは同じ遺伝子の違う場所に突然変異が起きているものであるということが分かった。

4. エージェント個体の振る舞い

今回の実験で用いるエージェントは全て自律的な個体であり、通信には1対1の P2P 通信を用いている。エージェントは Q 学習機能を搭載しており、各エージェントはそれぞれ独自の Q テーブルを持っている。各エージェントは個別に試行錯誤を繰り返しながら問題を解決する方法を学習する。

エージェントは1ステップ中に1回の行動が可能となっており、エージェントの選択できる行動は、上下左右のいずれかにグリッド1マス分移動するというものである。

エージェントには性別の概念が組み込まれており、異性同士が衝突することによって結合が可能である。結合した際には互いの Q テーブルに平均化や互いの最大値を代入するなどの操作を行い変化させることで知識の共有を行う。また、結合している個体は移動する際に結合している相手を牽引しながら移動を行う。さらに、個体を発生させる際に、雄個体にフルートレス遺伝子の突然変異を起こしたものとして雄同士で結合するものがある確率で発生させる。エージェントは単体ならば1ステップ中に一回しか行動できないが、結合している場合、結合している個体がそれぞれ行動するので1ステップ中に二回の行動が可能である。また、結合している個体が障害物に5回衝突したとき、その結合は解除されるものとなっている。

図1で結合の一例を示す。ステップ1では互いのエージェントが接近し、ステップ2で衝突し結合される。ステップ3、ステップ4は結合した個体が同時に移動する様子を示している。

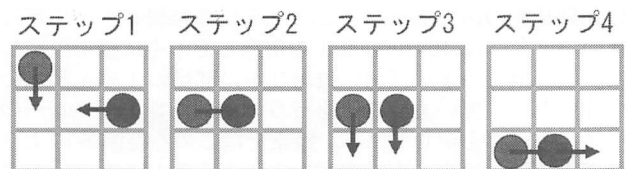


図1. 結合のパターン

5. 問題設定

本研究では、群体形成のための初期実験として、群行動の最小単位である雄と雌の結合による二体結合が可能なエージェント群によるシミュレーションを行う。以下に実験環境の概要図とその詳細を示す。

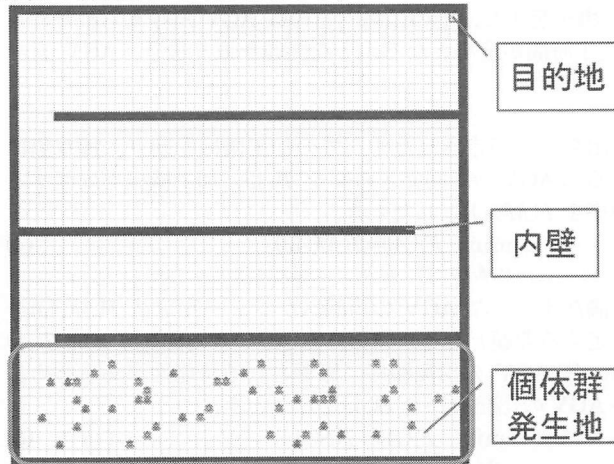


図2. 環境の概要図

50×50のグリッドフィールドを1エピソード10000ステップ以内に内壁を避け右上にある目的地へと向かうという問題を与え、これを1000エピソード繰り返す。また、各エージェントの結合時におけるQテーブルの操作は、互いのQ値を平均化するものとする。雄個体の変異確率は1/10と1/2の二通りで実験を行う。

6. 実験と考察

今回の実験では、フィールドの形状は据え置きでエージェントの個体数を雄25体、雌25体の計50体、雄50体、雌50体の計100体、雄100体、雌100体の計200体という三通りの環境下におけるシミュレーションを行う。Q学習における方策はε-グリーディ方策とし、Q値の更新式は以下のとおりである。

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha(r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)) \dots (式.1)$$

以下にその実験結果とそのグラフを示す。図3は個体数が計100体、変異確率が1/2の環境で実験を行い、いずれかの個体が最初に目標を達成したときの状態を抽出したものである。エピソード0、50、150、1000付近の状態を抽出した。グラフ上ではほとんど変化はないが、エピソード150では曲がり角付近で個体が壁際に集まっておりエピソード1000では全個体が一本の線のような形を作っている。

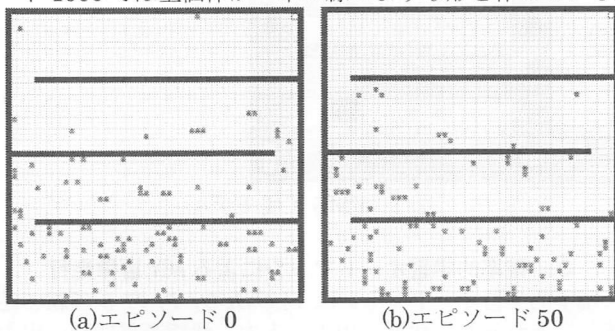


図3.1. エピソード毎の学習状況

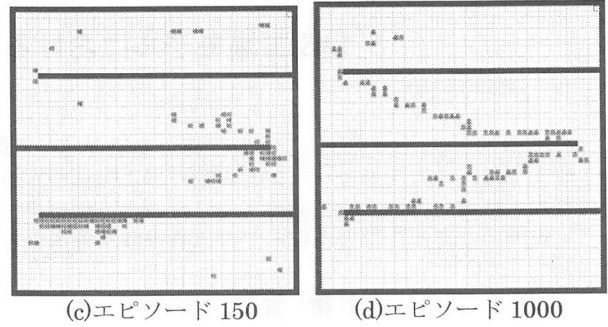


図3.2. エピソード毎の学習状況

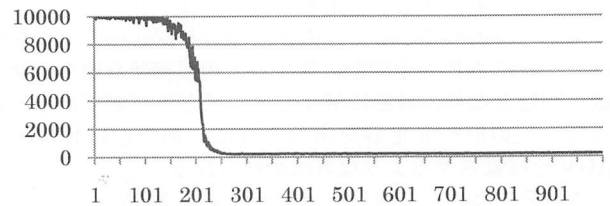


図4. 計50体の目標達成ステップ数平均

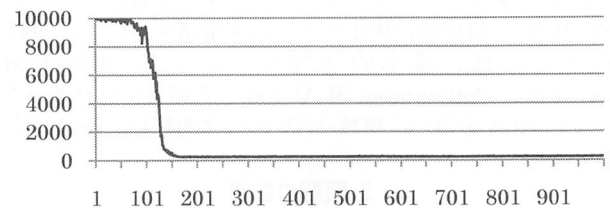


図5. 計100体の目標達成ステップ数平均

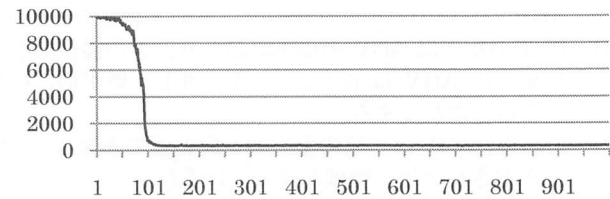


図6. 計200体の目標達成ステップ数平均

上記のグラフは計50体、100体、200体の目的達成までの所要ステップ数の全体平均を比較したものである。このグラフより、平均ステップ数が収束するまでに50体の場合は約250ステップ、100体の場合は約150ステップ、200体の場合は約110ステップというように、個体の数が多いほど平均ステップ数がより早く収束することが分かる。

7. おわりに

今回の実験では、群れを形成する最小単位である雄と雌という2体の結合を扱うことにより形態形成のための初期実験とした。結合した際に通信による知識の共有を機能として搭載した結果、形態を形成する個体数が多いほど、伝達の速度が速いことが分かった。今後は多数のフルートレス遺伝子を持つエージェント群のグルーピングに関する考察を行う。

参考文献

- “行動はどこまで遺伝するか” 山元 大輔
- “自律ロボット概論” ジョージ・A. ベーキ