

模倣と自己主張に基づく行動の学習 — 集団における協調行動の生成 —

三浦正宏 山口智浩 谷内田正彦

大阪大学基礎工学部システム工学科

自律分散システムでは、モジュール間の相互作用を考慮した設計が必要となるが、一般にその解析は困難である。そこで、近年、人工生命分野で進化のメカニズムを利用したシミュレーションによる相互作用の解析が注目されている。本論文では、モジュール群が環境中での相互作用を通して行動を学習する遺伝に基づく機械学習：GBMLの枠組みの下で、模倣と自己主張に基づく行動学習を提案する。自己主張により、新規性のある多様な行動の発見を行い、模倣により、他の個体の観察から行動を獲得し、集団全体の学習を速める。今回は、複数ロボットによる荷物の搬送問題でシミュレーションを行い、模倣による学習の有効性を検証する。

Learning Cooperative Behaviors Based on Mimetism and Insistence

Masahiro MIURA, Tomohiro YAMAGUCHI and Masahiko YACHIDA

Department of System Engineering, Faculty of Engineering Science, Osaka University

1-3 Machikaneyama-cho, Toyonaka, Osaka 560, JAPAN

E-mail: {miura, tomo, yachida}@sys.es.osaka-u.ac.jp

To design the autonomous decentralized system, it is not easy to analyze interactions between modules. But, recently on Artificial Life, the mechanism of evolution is applied to develop the system. This paper proposes a new method that multiple modules in the environment learn and discover cooperative behaviors by balancing Mimetism and Insistence based on the Genetic Based Machine Learning (GBML). Once, an agent finds a better behavior by Insistence, others learn it by Mimetism. We report the result of the simulation on the transportation problem by learning cooperative behaviors with Mimetism, in which multiple robots transport materials in the factory.

1 はじめに

近年、自律分散システムに関する研究が盛んに行われている[1]。自律分散システムの利点としては、システムの柔軟性、多様性、耐故障性等が考えられる。これらは、機能と情報の分散化、システムの複数モジュールへの分割、各モジュールの自律化によって得られる利点である。そのため、自律分散システムは、集中制御が困難かつ複雑で大規模なシステムに対し有効である。

しかし、自律分散システムでは、システム中の各モジュールをどのように設計し統合化を行うかが難しい。これは、各モジュールについて、他との相互作用を考慮した行動計画を行う必要があるからである。このため、各モジュール間の協調は必要不可欠なものである。しかし、モジュール間の協調や相互作用の解析が困難であるため、システムの設計については、期待するシステムの挙動に即した全体的な制約を設けることで、モジュールの設計を行うことが多かった。

これに対し、最近、人工生命分野での進化のメカニズムをモジュール等の設計に応用する手法が注目されている。この手法では、モジュールを人工の生物個体と見立てている。そして、生物の進化の仕組みを模倣して、世代交代や遺伝操作により、個体群（集団）の協調行動を獲得しようとする試みがなされている。全体的な制約を用いる従来手法に対し、人工生命的手法では、自ら行動を学習、進化する事のできる個体群が、ローカルな制約の基で組織化され協調行動を発現していくことを期待している点が特徴である。このような学習を行う個体群が進化するためには、

(1) 新しい行動を発生させる仕組み

(2) 発生した行動を他の個体に広める仕組み
の二つのメカニズムが必要とされる。これらの基本的アイデアは、GAの枠組みで実現されているが十分ではない。そこで、通常のGAに組み込むハイブリッドの手法として、(1)に関しては、発見的学習[2]や、ニューラルネット等による強化学習[3]が考えられる。(2)に関しては、模倣による学習[4]という方法が提案されている。

後者の模倣による学習[4]では、良い個体の行

動を悪い個体に広めることで、良い行動を持つ個体の割合を高め、集団全体の学習速度を向上させることができ、また、集団の挙動の安定化が期待される。しかし模倣だけでは、各個体のとりうる行動が均質化されて、初期収束や準最適解への収束が起こりやすく、最適行動の獲得が困難になる。そのため、従来は、手本となる個体を与えて進化を導いてやるといった工夫[4]が必要であった。

そこで、我々は、個体の自律的な進化という立場から、Classifier Systemによるルールの発見をベースとした行動学習システムを用いる。その枠組みで、まず手本となる個体を与えずに模倣に基づく学習による実験を試み、模倣の効果を検証する。さらに自律的な進化を引き起こし、新しい行動を積極的に学習する仕組みとして、自己主張という概念を学習システムに導入する。

自己主張とは今までの記述にない新たな行動や、他の個体が行っていない行動を発見し、学習しようとする姿勢のことである。自己主張の効果としては、多様な行動の獲得が予想され、様々な状態に対する学習の効果が期待されるが、その反面、集団全体の秩序が失われることが予想される。

このように、自己主張と模倣とは相反する概念である。したがって、これら2つをバランスよく組み合わせることで、互いの欠点を補った、よい学習結果が期待できる。以上の考えに基づき、我々は、模倣と自己主張のバランスに基づくモデルを提案する。今回は、模倣による学習の部分について、複数ロボットによる荷物の搬送のシミュレーションを行う。これは、工場内にいくつかの機械があり、自律的なロボットが機械間の荷物の輸送を行うという問題設定である。

以下2章では、模倣と自己主張による学習における我々の基本的なアイデアについて述べる。3章では、システムの概要について、4章以降で実際のインプリメント方法及び、シミュレーション結果について述べる。

2 協調行動の生成

我々は、自律的な個体が、単一、または複数個、同一環境下で行動することを考えている。

このような環境下で個体群を進化させる手段として、各個体にGBML(Genetic Based Machine Learning, 進化型学習)[5]を用いるという方法がよくとられている。

GBMLは、GAを用いてプロダクションルールを自動生成する方法のことで、とくにClassifier System (分類システム) [5][6]が有名である。この方法では、GAによってルールの獲得及び、そのルールの評価を行う。分類システムでは、各個体が独自に行動の学習を行っていき、個体間での学習結果の共有は考えていない。しかし、個体群が、協調して進化していくには、ある個体が学習した結果を他の個体に伝播するメカニズムが必要であると考えられる。

このような各個体が獲得した行動を、他の個体が利用可能とするためのメカニズムとして、模倣による学習[4]が、提案されている。これは人間の幼児が大人の行動を真似ながら学習していくことや、多数の人がしている行動を、自分も行うことで調和をとるといったことに通じている。

Nicolas Meuleau は、車の交通のシミュレーション[4]を行い、学習速度の向上と個体群の持つ行動の均質化が、模倣を用いた学習の効果として現れることを紹介している。この研究では、手本となる良い個体をシステムが検出し、良い個体の行動を悪い個体に強制的に与えるという方法で、模倣を実現している。しかし、この方法は個体間での模倣の流れを大局的にシステムがコントロールする必要があり、各個体が自律的に学習しているとはいえない。そこで、我々は、各個体には模倣の指示を直接与えず、各個体が自発的に模倣を行い、集団の中で協調的な挙動が創発することを目指す方法をとっている。

次に、模倣による効果の1つである、個体群の行動の均質化について考える。ある環境で個体群の行動が均質化した場合、集団全体の安定性は向上することが予想される。しかしその反面、その環境に対してさらに最適な行動があったとしても、その行動を個体群が獲得するのは困難である。なぜなら、ある個体がさらに良い行動を見つけたとしても、周りの模倣により既存の行動が強化されることで、新たな行動の実行の機会が失われるからである。これは人間の世界で、古くからの慣習に縛られ、新しい行動がなかなか広まらないことに似ている。したがっ

て、模倣の効果には、集団の均質化、安定化と、進化の停滞という2つの面がある。

これに対し、我々は、新しい行動を積極的に獲得するための仕組みを導入することで、模倣を用いたときの進化の停滞を打破できると考えている。この仕組みとして、自己主張という概念を提案する。自己主張とは、「今までにない新たな行動を行う」、「他の個体とは異なった行動をとる」、「他の個体より優れた行動を行う」、といった行動の選択に対する個体の意志や自立性を表す。

この自己主張の効果としては、多様な行動の獲得が予想されるが、その反面、各個体の挙動が不安定となって、集団の秩序が失われ、集団全体の協調行動の効率が落ちる可能性がある。これは自分勝手な人々が意見を言い合っても、なかなか意見がまとまらないのと同じである。

このように模倣と自己主張とは正反対の性質を表す。そこで我々は、以上で述べたGBMLをベースとして模倣による学習と自己主張とを組み合わせた、模倣と自己主張に基づく学習のモデルを提案する。模倣と自己主張とは相反するものであるが、両者をバランスよく組み合わせることで、それぞれの欠点を補い合った、新しい進化の実現を目指す。例えば、挙動が安定しても進化が停滞せず、より良い行動のある個体が発見すると、それが集団に広まり、さらに新たな進化を目指すといったことが期待できる。

我々の提案するモデルは、個体の学習結果を模倣を用いて集団全体に伝播させることで、集団における個体の学習と集団の進化の相互作用の双方を実現している。また、自己主張を導入することで、各個体の自律的な進化を目指す。

3 システムの概要

3.1 シミュレーションの概要

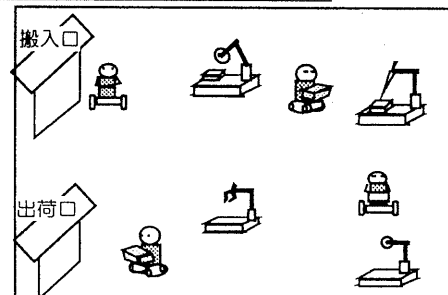


図 3.1 シミュレーション概要図

まず、問題設定とシミュレーションの概要について説明する。

我々は、資材を加工する機械が複数存在する仮定の工場を、個体群の進化する環境として設定する。ここで個体は、資材を搬送するロボットに対応している。

この工場では、

- ・ 工場の外から運ばれてきた資材は複数の工程を経て商品となり、工場の外へ出荷される。
- ・ 資材を加工する機械は複数あり、それぞれは異なる工程を担当している。
- ・ これらの機械は離れたところにあり、機械間で資材の搬送を行う必要がある。
- ・ 資材を加工する順序は決まっている。

という制約がある。ロボットは上記の制約の範囲内で、移動したり、資材を運んだりできる。

各ロボットには、工場全体として効率よく商品を生産せよ、という大まかなタスクを与える。各ロボットは、最初取るべき行動が分からないが、自律的に行動を学習する能力を持っているため、徐々に意味のある行動を学習し、最終的には、各ロボットが協調して、効率よく機械間で資材を搬送する行動を獲得する。

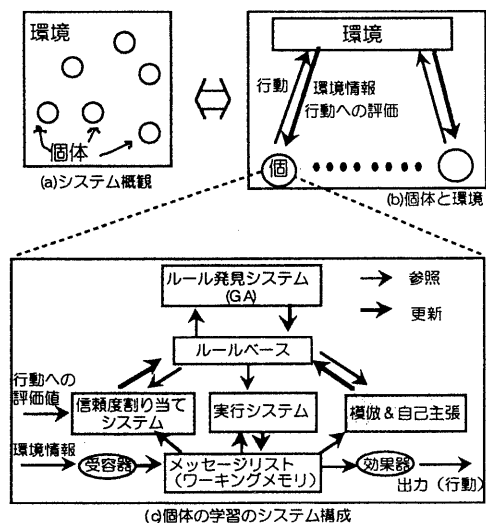


図 3.2 システム構成

図3.2に、システムの構成図を示す。図3.2(a)は、システムの概観である。複数の個体が環境中に存在し、互いに影響しあいながら行動をしている。図3.2(b)は、個体と環境の関係を表す。個体は、行動という形で環境に影響を

及ぼし、逆に、環境情報や行動に対する評価を環境から受け取る。図3.2(c)は、個体の学習に關する部分の構成である。これは、標準的な Classifier System に、模倣と自己主張によるルールの更新部分を付加した形で構成されている。

3.2 個体間での相互作用

個体間での相互作用は、環境を介して行われる。個体間での通信等はいっさい行わない。ただし、近くにいる個体が取った行動だけは、認識できる (→ 模倣、自己主張)。

個体間の相互作用としては、次の2つのことが考えられる。

(1) 個体間の競合

(2) 作業分担による個体群での協調

(1)は、複数の個体が独立して環境を更新していくため、ある個体が今まで独立に成功してきた行動が、他の個体の行動により、実行できなくなる現象。(2)は、各個体が、他の個体の行動可能な環境を互いに作りあう現象である。

3.3 個体と環境間の相互作用

個体は、行動という形で環境に対する働きかけをする。また、環境からは、個体の周りの状況および、行動に対する評価値が返る。

個体は、周りの状況から取るべき行動を判断し、実行する。そして、これに対し環境から評価値が返ってくるので、これを基に個体は、選んだ行動が良いか悪いかを判断する。

但し、環境が行う行動の評価は、反射的かつ局所的のものであり、行動のシーケンスに対する評価は行わないことにする。

3.4 Classifier System による個体の学習

本研究で用いた Classifier System と個体の学習の過程について簡単に説明する。

各個体は、環境から知覚した状態をメッセージという形式で表現する。ここでメッセージとは $\{0, 1\}$ からなる固定長の文字列である。環境からの入力、受容器の働きで、入力メッセージの形に変換後、メッセージリストに書き込まれる。メッセージリストは、メッセージを記録するワーキングメモリである。ルールベースには、プロダクションルールが記録されている。ここでのルールは条件部と帰結部からなる文字列

ルール=条件部：帰結部
 であり、条件部は {0,1,#}、帰結部は {0,1} からなる文字列である。(# は、Wild Characterを表す。)

実行システムでは、実行ルールの選択を行う。まず、メッセージリスト中の入力メッセージと、ルールベース中のルールの条件部とのマッチングを調べる。そして、マッチしたルールの間で競りを行い、競りに勝ったルールの帰結部を出力メッセージとしてメッセージリストに書き込む。

出力メッセージは個体の取る行動を表している。出力メッセージは、効果器によってメッセージリストから取り出された後、環境への出力として生成される。ただし、入力からメッセージ、メッセージから出力への変換は、あらかじめ一意に決めておく。

信頼度割り当てシステムでは、メッセージにマッチしたルールから、直前に使用したルールへの競り値の引き渡し操作を行う。また、環境からの評価値(報酬)は、直前に使用したルールへと加算される。この方法として、BBA (Bucket Brigade Algorithm)[6]という標準手法を用いている。

ルール発見システムでは、ルールベース中の各ルールを個体、ルール集合を個体群と見なし、GAを起動し、各ルールにGA operatorを適用することによりルールベースを更新する。

最後に、模倣&自己主張部について説明する。模倣では、他の個体の行動を観察して、ルールの形で認識し、ルールベースに書き込む。また、自己主張の概念に基づき、ルールベースにないルールの発見などを行い、ルールベースを更新していく。模倣及び自己主張の詳細については、4章及び6章で説明を行う。

4 模倣による学習

我々は、次の2つの方針に従って模倣のインプリメントを行う。

(方針1) : 個体が自発的に模倣を行う。

(方針2) : 模倣するときの手本は、個体の近傍から見つける。

(方針1)の意図は、「良い個体と悪い個体

をシステムが検出し、良い個体から行動ルールを悪い個体へと押しつける」[4]のではなく、「個体が自らお手本となる相手を見つけ、模倣という行動を選択する」ということである。我々は、個体が選択可能な行動の一つとして、模倣を扱うことにより、これを実現する。

(方針2)は、「模倣をするには、手本である個体の周りの状況及び行動を詳しく観察する必要があり、この観察は手本がごく近くにいるときしかできない。」ということである。

模倣を含めた、個体の行動手続きを図4.1に示す。

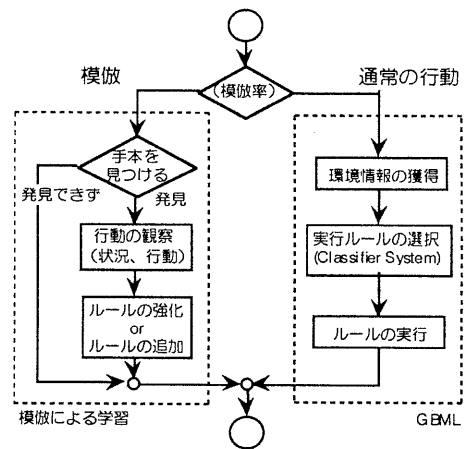


図4.1 個体の行動の手続き

行動手続きは大きく二つの部分に分かれており、一つはGBMLによる行動学習の部分、もう一つが模倣による学習部分である。以下では、模倣による学習の部分に関して説明する。

- (1) 個体が行動をするとき、まず通常の行動をするか、模倣をするかの選択を行う。この選択のために「模倣率」を設定する。模倣率とは、1回の行動選択において模倣を選ぶ確率を表す。この確率に従い、模倣を行う。例えば、模倣率20%の個体は、5回に1回の確率で模倣を試みる。
- (2) 模倣が選ばれると、まず手本となる個体を探す。もし、自分の近くに手本となる個体が全くいなければ、模倣は失敗し、その時点で一つの行動を終わる。
- (3) 手本となる個体を発見したら、その個体の状況(周りの環境)と取った行動を観察する。

この2種の情報により、手本の個体が実行したルールを再構成を行う。

(4)再構成したルールと、ルールベース中のルールを比較する。すでに存在するルールならそのルールの信頼度を強化する。存在しないのなら、そのルールをルールベースに取り込む。

以上で模倣の処理が完了し、一つの行動を終える。

5章で説明するシミュレーションの場合では、ロボットが加工機械の近くに止まる度に、上記の個体の行動手続きが実行される。

手続き中で、通常の行動が選択されたときは他の機械への移動や資材の受け取りを行う。

模倣が選択されたときは、その場に立ち止まった状態で、手本となる個体を探すことになる。

我々の行なったシミュレーションでは、いったん模倣が選択されると、一定時間、手本となる個体が現れるのを立ち止まって待つ。この時間中に自分のいる機械に他の個体が移動してきて行動をすると、手本を発見できたことになり、模倣を行なった後、次の行動選択に移る。手本を見つけられずに一定時間中に経過すると、模倣に失敗し、行動選択に移り次の行動手続きが実行される。したがって、模倣には時間コストがかかることになる。

5 模倣による行動学習の実験

5.1 シミュレーション設定

3章で説明した仮想の工場を用いて、模倣による行動学習のシミュレーション実験を行なった。それぞれのシミュレーション設定を図5.1、図5.2に示す。個体が学習するルールは、図5.1の環境の認識を条件部、図5.1の行動を帰結部として、{0,1}の文字列にコード化することで、生成される。ルール空間の大きさは、条件部が256通り、帰結部が8通り、全体では2048通りである。各個体は、50個のルールを持つことができる。

仮想の工場	
◎加工用機械の台数	4台
◎搬送ロボットの台数	1台(実験1)、2台(実験1)、3台(実験2)

加工用機械	
◎在庫できる資材の最大数	15個
◎在庫できる加工済み資材の最大数	15個
◎資材一箇の加工に要する時間	5 Step

搬送ロボット	
◎持てる荷物の数	1個
◎環境の認識	<ul style="list-style-type: none"> ・現在地(2bits) ・現在いる機械の資材の在庫数(2bits) ・現在いる機械の加工済み資材の在庫数(2bits) ・運んでいる荷物の種類(2bits)
◎行動	<ul style="list-style-type: none"> ・どの機械への移動するか(2bits) ・荷物を受け取るかどうか(1bits)
◎Classifier Systemについて	<ul style="list-style-type: none"> ・交叉率 30% (1点交叉) ・突然変異率 20% ・ルールの最大数 50個
◎模倣率	0%, 15%, 20%, 25%, 30% にて実験
◎環境からの評価	行動する度に点数を与える。
◎手本を探す時間	5 Step

図 5.1 シミュレーションの設定 その1

環境からロボットへの評価(点数)	
+は良い評価、-は悪い評価	
評価(点数) = 行動への点数 - 移動コスト	
◎行動への点数	
正しい場所への荷物移動	100点
元の場所への荷物移動	0点
間違った場所への荷物移動(受け取り拒否)	-200点
荷物が無いのに、荷物を受け取るうとした場合	-200点
◎移動コスト	
機械間を移動するのにかった時間 x 5点	

システムの評価	
◎集団全体(工場としての)のパフォーマンス	-> 出荷率
	100 Stepの間に出荷した商品の個数
◎個体(ロボット)のパフォーマンス	
	環境からの評価値の、100 Stepの間の累積

図 5.2 シミュレーションの設定 その2

5.2 実験

実験1 (手本なしの模倣学習)

5.1節で示した設定に従い、個体群の行動学習のシミュレーションを行なった。実験1では、個体(搬送ロボット)が1台の場合、及び個体が2台(ただし、模倣なし、模倣率15%、20%、25%、30%の5通り)の場合の計6通りについて、集団全体のパフォーマンス及び各個体のパフォーマンスの時間的変化を調べた。この実験は以下のようにして行なった。

(1) 無学習の個体群を環境の中に置き、行動

学習を行う。

(2) 30000 time step経過したところで、シミュレーションを終了する。

実験2 (手本ありの模倣学習)

シミュレーション設定は実験1と同じである。実験2では、安定挙動を示す学習済みの個体群の中に、無学習の新規個体を入れた場合の集団全体のパフォーマンス及び、個体のパフォーマンスの変化を調べた。この実験は以下のようにして行なった。

(1) ロボット2台からなる模倣なしと模倣率20%の個体群で、実験1と同様の行動学習を行う。

(2) 20000 time step経過して、全体挙動が安定した時点で、新たに無学習のロボット1台を個体群の中に入れる。

(3) 50000 time step経過したところで、シミュレーションを終了する。

5.3 実験結果

実験1 (手本なしの模倣学習)

実験1の結果を図5.3及び図5.4に示す。図5.3は、集団パフォーマンスの変化で、縦軸は工場の出荷率(個/100time)、横軸は時間である。図5.4は、縦軸は、個体のパフォーマンスとなっている。また、これらのグラフはシミュレーションを5回行った結果の平均である。

シミュレーション中でのロボットの挙動を観察したところ、最も出荷率が高くなる行動は、2個体両方が工場内の4つの機械を加工の順番に周回するというものであった。模倣ありの場合で獲得された行動は、ほとんど周回行動であった。模倣なしの場合では、周回行動のほかに、分担作業を獲得するものが多かった。例えば、一方が搬出口への輸送を行い、他方がその他の機械間の輸送を担当するというものである。

図5.3(a)で、集団全体のパフォーマンスの変化を見ると、個体が複数いる場合の方が、1個体のみの場合に比較して、不安定であることがわかる。また、2個体の場合では、模倣ありの変化の方が、模倣なしの変化より長期的に変動が少なく、安定している。これは、模倣の効果の一つである、挙動と効率の安定性を示してい

る。

図5.3(b)を見ると、模倣率を変えた場合では、模倣率を上げるにつれて、集団パフォーマンスの上昇の仕方、つまり集団としての行動学習のスピードが、遅くなるのがわかる。これは、「模倣を選ぶと、立ち止まって待つ」ため模倣に時間的コストがかかるためであり、また、より良い行動に対する強化学習を行う回数が減少するためであると考えられる。学習速度の厳密な比較には、これらの要因の考慮が必要とである。

実験1：ロボット1台、
2台(模倣なし、模倣率20%)

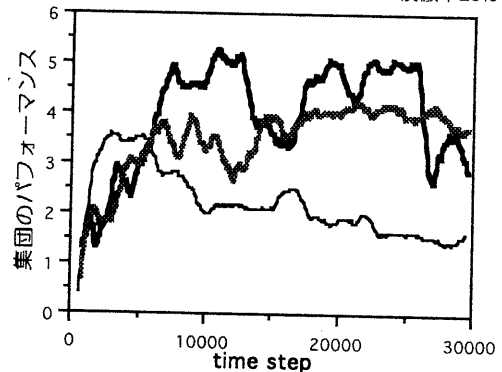


図5.3(a) 集団パフォーマンスの変化 その1

実験1：ロボット2台
(模倣率15%, 20%, 25%, 30%)

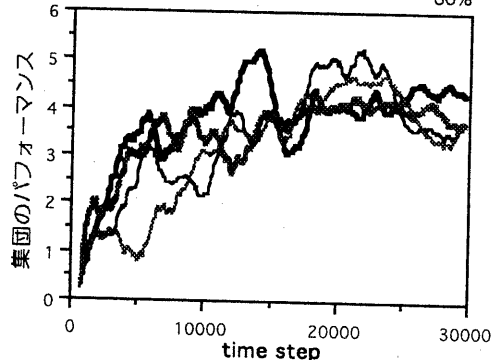


図5.3(b) 集団パフォーマンスの変化 その2

次に図5.4で、個体ごとの学習効果の比較を行う。模倣ありの場合では、模倣なしの場合に

比べると、学習速度は遅いが、スムーズにより行動を獲得していることが、個体のパフォーマンスの10000stepまでの変化から分かる。これは学習結果の他の個体への伝搬のためと考えられ、模倣の効果の1つであるといえる。

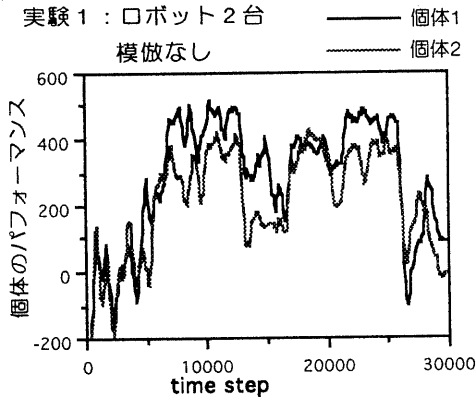


図5.4(a) 個体のパフォーマンスの変化 その1

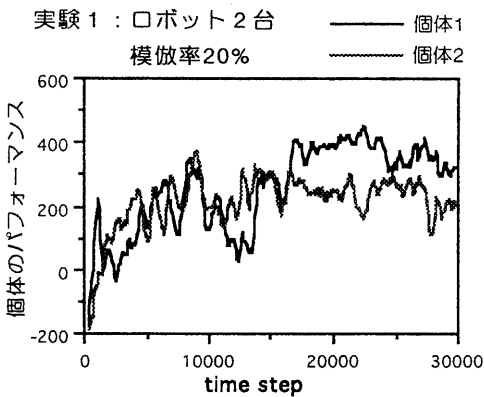


図5.4(b) 個体のパフォーマンスの変化 その2

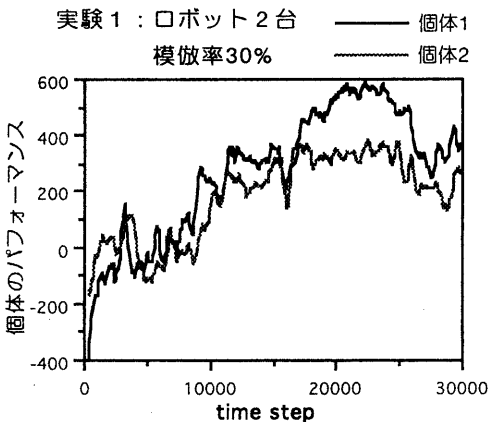


図5.4(c) 個体のパフォーマンスの変化 その3

実験2：ロボット2台+1台 —— 模倣なし
模倣なし、模倣率20% —— 模倣率20%

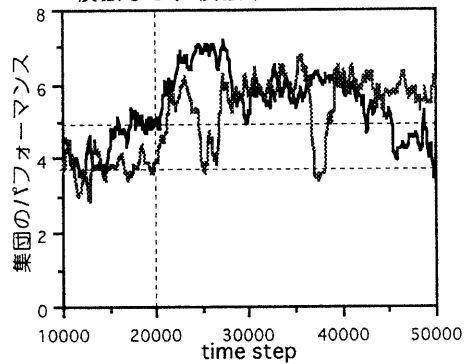


図5.5 集団パフォーマンスの変化

実験2：
新規個体のパフォーマンス —— 模倣なし
模倣なし、模倣率20% —— 模倣率20%

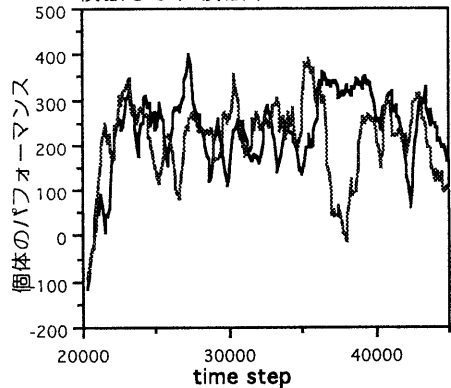


図5.6 新規個体のパフォーマンスの変化

実験2（手本ありの模倣学習）

実験2の結果を、図5.5及び図5.6に示す。

図5.5の集団パフォーマンスの変化では、模倣なし、模倣あり共に、新規個体の追加による集団パフォーマンスの向上が見られた。また、模倣の影響により、模倣ありの方が長期的に安定しているといえる。模倣ありの方で、追加後10000stepの間と、約38000stepのところで、大きな変動が見られる。

次に図5.6について新規個体のパフォーマンスの変化を見ると、初期の学習過程では、模倣なしと模倣ありがほとんど同じ速度で上昇している。模倣による時間コストを考慮すると、模

倣ありの方が学習速度は速いといえる。これは、実験1の場合と違い、手本となる個体がいるためだと考えられる。

5.3 考察

実験1、2から、倣によって、集団全体の挙動の安定性が向上することが示された。学習速度の点では、手本となる個体がいる時では、倣を用いない場合より優れているが、手本なしの場合は、倣ありの方が効率の上昇が遅い。しかし、集団で学習結果を共有できるために、スムーズに学習が行われることが分かった。

個体群の行動をみると、やはり、倣ありでは、行動の均質化の傾向が強いことが分かった。今回の問題設定では、同一の行動でも、全体のパフォーマンスを上げられるようになっていたが、複雑なタスクを与えるようになると作業の分担が、重要になってくるであろう。複雑なタスクに対応できるためには、行動の均質化を抑える仕組みが必要である。

以下で、倣の方法の問題点と改善案について述べる。第一に、倣に要するコストの設定がある。現在の倣には、時間コストがかかっており、実験1での学習速度の比較をする事ができなかった。この時間コストは、「他のロボットを詳しく観察するためには、立ち止まる必要がある」という設定のためであるが、移動しながらでも観察が出来るように変更すれば、グラフ上での学習速度の比較が出来るはずである。

また、倣可能な状況の学習の問題もある。現在では、倣行動の選択を行動選択の場面で確率的に行っている。そのため、手本となる個体が近くにいない場合がでは、倣の失敗が多く生じる。これを、どのような状況で倣が出来るかを暗黙のヒューリスティックで与えたり、倣可能な状況を学習させるなどして、確実に倣を成功させるための改善が考えられる。

6 自己主張による学習

自己主張の方法としては、その捉え方により、いろいろなものが考えられる。自己主張というと、周りに影響されずに、自らの信念を曲げずに、理想を追求する”Going my way”という解

釈もあるが、我々は自己主張を、倣による個体の均質化の防止、「新たな行動の積極的な獲得」の手段として考えている。そこで、「今までにない新しい行動を獲得する。」「周りの個体と違う行動をとる。」という2つの考えに基づき、自己主張を実現する。

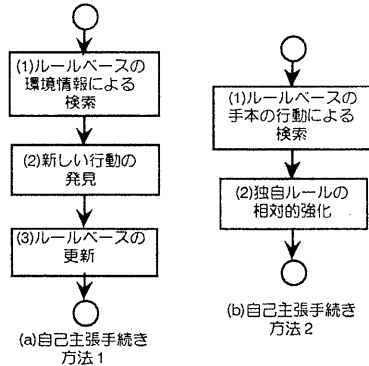


図6.1 自己主張の手続き

自己主張による学習の手続きとして、現在、以下の2つの方法を考えている。図5.1に、自己主張の手続きを示す。

(方法1)：今までにない新たな行動の獲得
(図6.1(a)の手続き)

現在の状況に適用できるルールを、ルールベースから検索する。次に、検索されたルールでは出来ない新たな行動を発見する。発見した行動をルールの形に変換し、ルールベースに加える。

(方法2)：周りの個体と異なる行動を取る
(図6.1(b)の手続き)

倣の手続き中において、手本となる個体の行動を観察できたとする。その個体の状況に適用できるルールをルールベースから検索する。検索されたルールの内、手本が用いたルール以外を独自のルールであるとして、強化を行う。

上記の自己主張手続きを、図4.1の個体の行動手続きに付加すると、図6.2のようになる。

(方法1)は通常行動の手続きの中に、(方法2)は倣の手続きの中に、それぞれ組み込まれている。

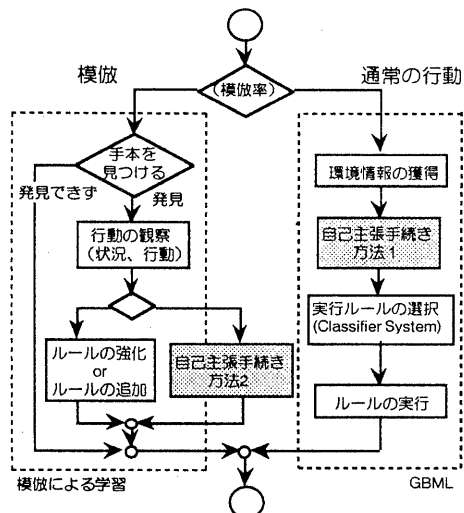


図 6.2 個体の行動の手続き (自己主張含む)

以上のように、自己主張の手続きについて述べたが、現在ではまだ、実際のシステムには組み込まれてはいない。今後、システムに組み込みシミュレーション実験を行う予定である。

7 おわりに

本稿では、環境中で個体群が最適な行動を学習するという枠組みのもとで、模倣と自己主張に基づく行動学習のモデルを提案した。模倣とは、他の個体を観察することによる行動の獲得、自己主張とは、新規性のある行動の積極的な獲得の概念である。我々のモデルでは、模倣によって、集団における個体の学習と集団の進化の相互作用を、自己主張の導入によって、各個体の自律的進化を実現している。

このモデルのうち、模倣による学習の部分構築し、複数ロボットによる荷物の搬送問題で、その効果を検証した。この結果、模倣による安定性の向上と、個体群の均質化という結果を得た。また、手本のある場合での、個体の学習速度の向上という結果も得られた。現在の模倣の方法では、学習速度についての検討を行えない、模倣の失敗による学習速度の低下がある、などの問題があり、改善をおこなう必要がある。また、自己主張の実現法について述べたが、実際にシステムに組み込み、システム全体としての

評価もおこなわなければならない、これらは今後の課題である。

参考文献

- [1] 計測と制御 (特集 群知能ロボット), Vol. 31 No.11 (1992)
- [2] Lenat, D.B., "The Role of Heuristics in Learning by Discovery, Three Case Studies", in Machine Learning (Michalski, R.S., Carbonell, J.G. and Mitchell T.M., eds.) (1983), (邦訳: "発見的学習におけるヒューリスティックの役割: 三つのケーススタディ", 知識獲得と学習シリーズ2, 学習と問題解決, 共立出版)
- [3] David Arckley and Michael Littman, "Interaction Between Learning and Evolution", Artificial Life II, pp.487-509.
- [4] Nicolas Meuleau, "Simulating Co-Evolution with Mimetism", Proceedings of the First European Conference on Artificial Life, pp.179-184.
- [5] 竹内 勝, "遺伝的アルゴリズムによる機械学習", 計測と制御 Vol.32 No.1 (1993)
- [6] Goldberg, D.E., Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, 1989