

異種エージェント集団の進化的学習による 協調行動の獲得

1 T-7

藤永 貴之 犬塚 信博 伊藤 英則
名古屋工業大学

1 はじめに

現在、マルチエージェントシステムの分野において集団間の協調行動が問題とされている。このエージェント集団による協調には2つの方法がある。1つは、同一の集団内で1つの仕事を個々が分担して行なう役割分担であり、もう1つは異なる集団内で個々が自分の特徴にあった仕事をこなす適材適所である。

本稿では、異種のエージェント集団内での協調の学習のため、それぞれに異なる能力を持つエージェント集団を用いる。本稿の目的は、この異種集団による協調行動の進化を観察することである。

我々は、集団に対してある環境と目的を与え、その環境において目的を達成する上で適切な行動をとるよう進化していくことを期待する。この進化は、エージェントの能力と行動は直接関係していないので明らかではない。エージェントは自分の能力を進化の過程のみ認識することができる。我々が与えた環境において、集団は協調なしで目標を達成することができるが、協調することでより効率を良くすることが可能である。本稿では、シミュレータを用いて集団が経験を通じて行動を進化させていくことを観察する。

2 荷物運搬問題

本稿の環境は、後に定義する2種類のエージェントと荷物の山、基地とがある。エージェントは荷物の山を発見し、そこから荷物を取り、そして基地を発見し、荷物をそこに置く。

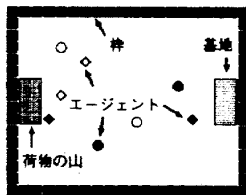


図1: フィールド

エージェントは、フィールド(図1)上で行動し、これを囲んでいる枠からは出ることができない(フィールドサイズ20×16)。図では、2種類のエージェントを丸形と菱形とで表している。

このエージェントの違いは次の章で記述する。色が反転しているのは荷物を持っている状態を示している。

我々は荷物を効率良く運搬するために、各エージェントが自分の役割を認識し行動していくことを期待している。

3 エージェントの定義

3.1 エージェントの入力と出力

エージェントは環境から得られる入力(視界は自分を中心に15×15の正方形)をもとに出力となる行動を決定する。エージェントは7種類の知覚情報をビット列として得ることができる(表1)。そして、その入力に対して決定される出力は16種類存在する(表2)。

入力ビット	ビットが1となる状態
X0	荷物を持っている
X1	荷物の山が視界にある
X2	基地が視界にある
X3	同種エージェントが視界にいる
X4	異種エージェントが視界にいる
X5	視界内の同種エージェントが荷物を持っている
X6	視界内の異種エージェントが荷物を持っている

表1: 入力ビット列の割り当て

行動の種類 / 行動	行動の種類 / 行動
A/ 荷物の山に近づく	K/ 同種エージェントから荷物を受け取る
B/ 荷物の山から離れる	L/ 同種エージェントに荷物を渡す
C/ 基地に近づく	M/ 異種エージェントから荷物を受け取る
D/ 基地から離れる	N/ 異種エージェントに荷物を渡す
E/ 同種エージェントに近づく	O/ ランダム移動
F/ 同種エージェントから離れる	P/ 何もしない
G/ 異種エージェントに近づく	
H/ 異種エージェントから離れる	
I/ 荷物を取る	
J/ 荷物を置く	

表2: エージェントの行動

3.2 行動決定の戦略

各エージェントはそれぞれ行動戦略を1つ持っている。それぞれの行動は表1の入力をもとにして表2の出力により決定される。我々は行動決定戦略としてn出力二分決定グラフ(n-BDD)を用いる。これを遺伝子表現として用い進化させていく。n-BDDはエージェントの行動戦略の進化的学習に対して用いられている[1]。この概要は図2で示している。

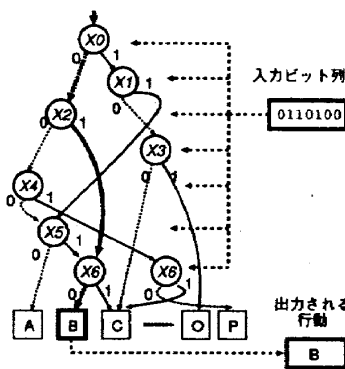


図2: n-BDDによる行動の決定

3.3 エージェントの能力の違い

本稿ではエージェントの能力の違いを与えるため、得意な行動と不得意な行動を定義した。得意な行動には、エネルギーの消費が他と比べて少ない行動や短い時間で達成できる行動などが考えられる。また、不得意な行動は、この反対と考えられる。本稿では、行動に対する時間で得意・不得意を定義した。我々は2種類の能力の異なるエージェントを定義するため、互いに自分の得意な行動は異種では不得意になるように設定した。

タイプ A (Truck) 我々はタイプ A のエージェントとしてトラックのような特徴を持たせた。これは、

荷物の運搬（行動 A~H, K~P ただし, X0 の入力ビットが1の時）を得意としており、荷物の上げ下ろし（荷物を取る行動 I と荷物を置く行動 J）を不得意としている。

タイプ B (Lift) タイプ B はリフトのような特徴を持たせた。これは、タイプ A とは反対の特徴を持っている。

4 システム概要

タイプ A とタイプ B の各 4 体を 1 グループとして 50 グループ用意した。各試行で 1 グループに対して一定時間試行を行ない、次のような適応度を得る。

$$\text{適応度} = (\text{荷物の山から取られた荷物の合計}) + 2(\text{基地へ置かれた荷物の合計})$$

この得られた適応度に従って遺伝的操作を適用することで次世代の n-BDD を生成する。ただし、適応度の高かった何グループかは遺伝的操作を加えずに残す。n-BDD に対する遺伝的操作として、節点の変更・追加・削除の 3 つを用いた。

5 実験とその評価

世代数を 3000 として実験を行なった。適応度の推移は図 3 に示すように進化するにつれて増加しているのが確認できる。

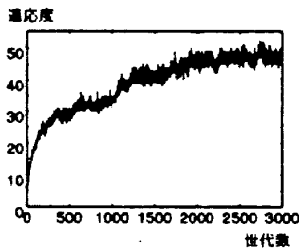
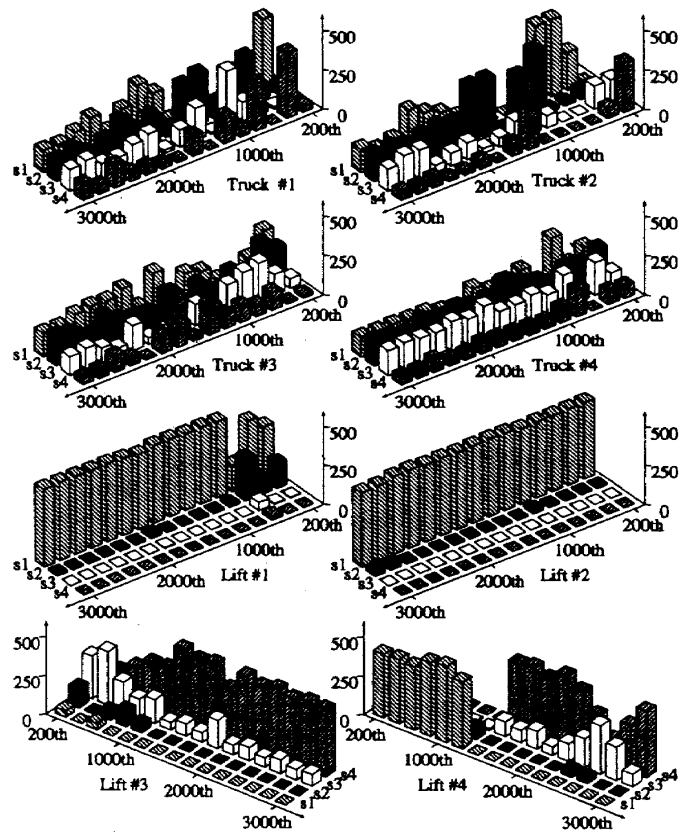


図 3: 適応度の世代推移

我々はフィールドを 4 つの範囲（荷物の山から基地までを s1, s2, s3, s4）に分割し、各世代においてどの範囲で多く行動したかを観察した。各エージェント（Truck #1~Truck #4, Lift #1~Lift #4）の行動範囲を調べてみると、図 4 のようにエージェントのタイプによって違いがみえる。この図より、タイプ A（Truck）はすべての範囲で行動しているのに対して、タイプ B（Lift）は荷物の山、基地周辺で行動していることが確認できる。これらの傾向は世代を重ねるごとにはっきりとしてくる。また、タイプ B の 4 体は 2 つのグループに分かれていくことが確認できる。

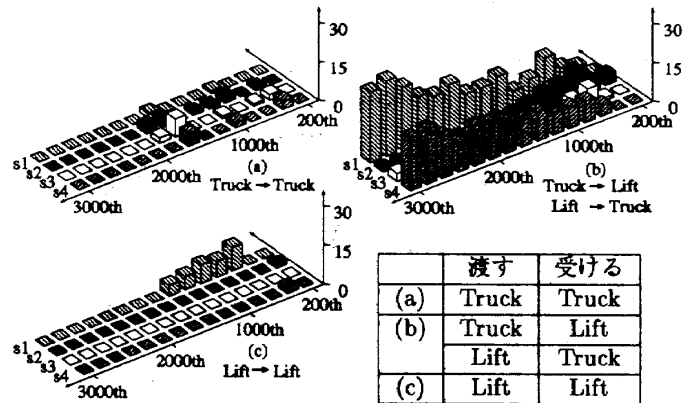
次に、我々は期待される協調が行なわれているかを確認するため、エージェント間の荷物の受渡しについて注目してみた。図 5 のように渡す側と受ける側とに分類し、世代を通して観察した。この図より、世代を重ねるにつれて (b) の異種エージェント間による受渡し行動のみ現れることが確認できる。また、この図は荷物の山・基地周辺で受渡しが行なわれるようになったことを示している。

実験結果より、各エージェントが進化・学習することで自分の役割を認識し、効果的に目的を達成するため自分の得意な行動を取ることによって適材適所が発生したと考えられる。



200th ~ 3000th : 世代数, s1 ~ s4 : フィールドの範囲, 0 ~ 500 : フィールドの範囲での行動回数

図 4: エージェントの行動範囲の分布



	渡す	受ける
(a)	Truck	Truck
(b)	Truck	Lift
	Lift	Truck
(c)	Lift	Lift

200th ~ 3000th : 世代数, s1 ~ s4 : フィールドの範囲, 0 ~ 30 : フィールドの範囲での荷物の受渡し回数

図 5: 荷物の受渡し範囲の分布

6 おわりに

本稿では、2 種類の能力の異なる進化的学習機能を持つエージェント集団の協調行動の獲得について述べ、適材適所の発生を実験により確認した。

参考文献

[1] K. Moriwaki, N. Inuzuka, M. Yamada, K. Itoh, H. Seki and H. Itoh, "Self Adaptation of Agent's Behavior using GA with n-BDD", Proc. 5th IEEE Int'l Workshop on Robot and Human Communication, pp.96-101, 1996.