

# マルチエージェント環境におけるロボットの構造と行動の共進化

## Coevolution of Morphology and Behavior of Robots in a Multiagent Environment

名古屋大学大学院人間情報学研究科 朝井 勇次 有田 隆也

Yuji Asai Takaya Arita

Graduate School of Human Informatics, Nagoya University

**Abstract:** Design of robots becomes far more difficult as the complexity increases in the required tasks and task environments increase. We have constructed an evolutionary model in which control of behaviors and morphology including selection and arrangement of sensors/motors in artificial creatures coevolve based on ontogenesis and biological evolution, so as to investigate a new effective method of designing robots. This paper extends this model in the context of a multi-agent environment in which cooperation of robots will be required, and analyzes the evolutionary process based on the simulation experiment. The results have shown the effectiveness of the scheme by generating various robots that behave intricately. We have also observed an interesting phenomenon which is relevant to coexistence of species in nature.

### 1 はじめに

ロボット開発の現場においては、より多様なタスクへの対応や性能の向上が求められ、設計はますます困難になりつつある。そのような事情に対応して、最近、進化ロボティクスの研究が活発である。進化ロボティクスは、生物進化のメカニズムをロボット設計の自動化に応用する試みであり、その初期の頃から多様な研究の動機や背景を持つ研究者を集めてきた。現在、その方法論に関して、多様化しつつもほぼ確立された段階に至ったと言える面もある [1]。ただし、多くの研究では（ロボット製作の場面全般において基本的にいえることであるが）ロボットの構造を人間がまず確定した上で、その構造においてセンサからの入力をどのようにモータへの出力に結びつけるかという部分に進化的計算手法を適用するものであった。

一方、自然界における生物の進化は、体の構造の進化と体の制御法（行動）の進化の相互作用という側面を持つとみなすことができる。課せられた制約の下で、生息する環境に応じて、両者が相互作用しながら選択圧を受けてきたという見解である。生物のような自律性を持つロボットにおいても、構造と行動、そして、環境の間の高度に洗練化された相互作用は本質的であろう [2]。そのようなアプローチから、いくつかの研究が行動だけでなく、構造との共進化を指向している [3]。ただし、基本的に、これらの研究は、ロボットの形状、特に、各種センサの選択や配置までも進化的に設計させるものではなく [2]、環境との相互作用こそ知能の存立を決定付けているという立場 [4] は現段階ではまだ必ずしも重視されていないとすることができる。

我々は、生物の発生と進化のメカニズムを応用したロボットの構造と行動の共進化モデルについて研究している [5][6]。提案モデルでは、個体レベルでは、構造（どのようなセンサを採用し形状はどうするか）と行動（センサ入力をどのようにモータ出力へつなげるか）を、生物の個体発生を模した発生モデルにより統一的に決定し、集団レベルでは、個体レベルの発生を規定する情報を進化モデルにより探索する。構造と行動、そして環境が密接に相互作用しながら現実的なロボットを進化的に設計する試みである。構造を固定せず探索の対象とするので、同一のハードウェア資源を前提とすると、構造を決定してスタートする場合に比べて、自由度が高くなり、探索空間が広がることを意味する。しかし、環境やタスクが複雑化すると、人間による構造の決定の困難化やコスト増加、さらに、人間の決定した構造がすでに解を持ちえない可能性も想定される。それらの意味で、全体として効率的なロボット設計の方法論となりうると考える。これまで、本モデルを基にしたシミュレーション実験を通して、センサを使い分ける多様なロボットが出現することを確認し、モデルの有効性を示した。また、シミュレーション結果を実ロボットに反映させた結果、シミュレーション中で獲得された知的な行動が実ロボットによって再現された。

本論文では、同モデルをさらに複雑な環境に適応させるため、マルチエージェント環境に応用する。具体的には、マルチエージェント環境に対応した5種類のモデルを構築し、それらに基づいたシミュレーション実験を行う。実験結果から1つの進化系列に着目し、進化過程を分析する。そして、各モデルに基づいた実験の結果を比較し、各モデルの特徴や、最良の結果が得

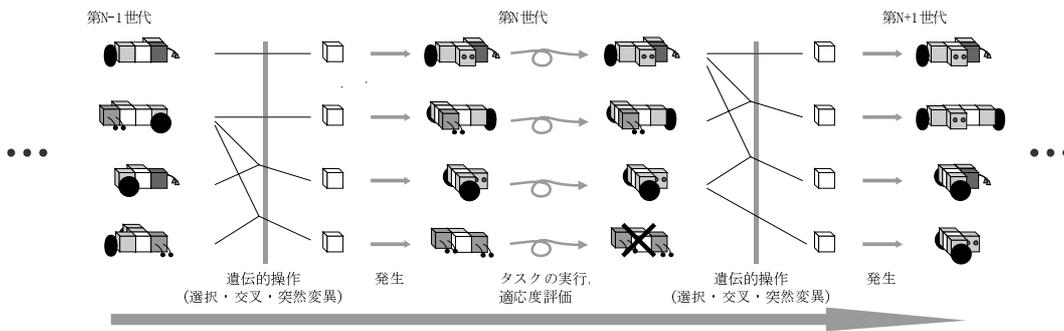


図 1: ロボットの発生と進化の概念

られるモデルの原理について考察する。

## 2 モデル

### 2.1 構造と行動の決定 [5][6]

ロボットの構造と行動の共進化は、生物の発生と進化の仕組みを模擬することで実現する(図1)。生物の形質が遺伝子に従って決定されることと同様に、ロボット(個体)の構造と行動も、遺伝子に従って決定される。生存に有利な生物が、より多くの子孫を次世代に残すことができるという、自然界と同様の仕組みによりロボットは進化する。

生物は1つの受精卵の卵割にはじまり、多数の細胞により構成される成体となる。体を構成する細胞は同一の染色体を持っている。それぞれの細胞は周囲の状態によって発現する遺伝子が異なる。その仕組みにより、細胞は異なる役割を果たし、生物は複雑な構造をとることができる。本モデルにおいても、各ロボットは遺伝子情報として記述され、その情報(文法)に基づいて発生と進化を行う。遺伝子情報の表現方式は、細胞分化に関する Nakano らの自己組織モデル [7] で採用された表記方式をロボット設計の方法論の確立を念頭において変更したものである(図2)。

発生は、染色体が与えられた1つのブロックだけが存在する状態から始まる。ブロックには、遺伝子情報に従い、様々なブロックが接続されていく。遺伝子情報は、接続された新たなブロックにもコピーされ、各ブロックにおいて同様に解釈される。全ブロックにおいて遺伝子の発現が終わると発生の段階は終了する。

1つの遺伝子は、「条件」、「実行」の各部分からなる。「条件」は、「実行」部分を実行するための条件であり、さらに2つに分けられる。1つは、上下左右(二次元平面上のブロックを真上から見る)のブロックの接続状態を調べる「接続」条件である。「接続」条件は上下左右それぞれに対して「接続」、「非接続」、「どちらでもよい」のいずれかを示す。もう1つは上下左右のブロックの発火状態を調べる「発火」条件である。「発火」条件は、上下左右それぞれに「発火」、「非発火」、「どちらでもよい」のいずれかを示す。「実行」の内容は、2

種類あり、上下左右のいずれかの1つの位置と、そこに新たに接続するブロックの情報(種類と方向)を表す「ブロック接続」(種類や方向の情報については同図では省略)、又は、ブロックの発火に関する情報を表す「発火設定」のどちらかである。発火の設定としては、「発火1」、「発火2」、「発火しない」の3種類がある。モータブロックでは、「発火1」は車輪の時計回り、「発火2」は車輪の反時計回りの回転を意味する。「実行」の内容が「ブロック接続」である遺伝子を「構造」遺伝子とし、「発火設定」である遺伝子を「発火」遺伝子とする。「構造」遺伝子は、発生の段階で作用し、「発火」遺伝子はロボットが行動する段階で作用する。遺伝子は縦に一列に並んでいるものとし、上から優先順位が与えられる。

タスク実行時には、構成する各ブロックは遺伝子に従い、発火状態が動的に変化する。センサを持つブロックだけは、センサが反応したときに発火する。環境からの入力に応じて各ブロックの発火状態が伝播し、発火したブロックが相互に影響し合い、モータなどのアクチュエータブロックの状態変化によって再び環境への出力となる。

図2に、構造の生成と発火状態の変化の例を示す。AからEの各行が遺伝子を表しており、網掛けをしている部分は使用されない。C、Eは「実行」部分が「ブロック接続」である。従って、C、Eは「構造」遺伝子であり、発生段階でのみで使用される。また、A、B、Dは「発火」遺伝子であり、タスク実行時に使用される。同図の下部は、ロボットの発生時における構造の生成過程と、タスク実行時の状態変化である。ブロックに網掛けがしてあるものは発火状態を示す。まず、発生はブロック1が1つだけ存在する状態から始まる。この周りにブロックが存在しないので、Cだけが「接続」条件を満たし、その「実行」部分が実行され、下にブロック2が接続される。次に、ブロック1は、遺伝子Eに従い、左にブロック3が接続される。タスク実行の段階では、各ブロックは「発火」遺伝子に従って発火状態と非発火状態をとる。ブロック3はDの条件、つまり、「接続」条件で右と下が「接続」、「発火」条件で右と下が「非発火」を満たすので、発火状態になる。次にブロック3が発火することにより、ブロック1がAの条件を満たし発火する。

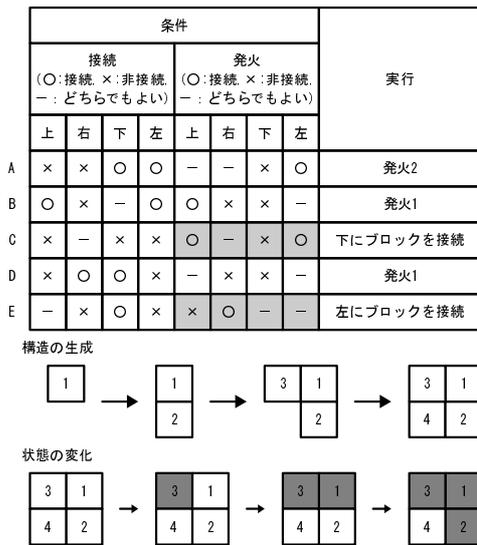


図 2: 遺伝子情報の利用の例

## 2.2 マルチエージェント環境での進化の方法

2.1 で述べた方法によって規定される個体の構造と行動をマルチエージェント環境において進化させる方法を考える。複数ロボットが指定された場所に置かれて場所を共有しながら 1 つのタスクを実行する状況を想定しており、ロボット集団の個数、適応度評価の単位、グループ選択を導入するかどうかの観点から 5 つの進化モデルを提案する。

最初に、全モデルに共通した部分を説明する。各ロボットは、発生した後、タスクを実行する。全ブロックが遺伝子情報やセンサの入力に従って状態を更新するまでを 1 ステップとする。設定されたステップ数を終えると、タスクは終了し、適応度が評価される。全てのロボットが適応度評価を行った後、遺伝的操作を行い、次世代集団を構成する。遺伝的操作では、親はルール選択により決定し、親 2 個体に対して子が 1 個体できる。2 つの親においてそれぞれ交叉点がランダムに決まり、一点交叉する（遺伝子数は増減する）。一方の親の染色体の前半部分ともう一方の親の後半部分が繋げられるが、その際、前半部分と後半部分の位置を入れ替えが 0.5 の確率で起こる。遺伝子の位置によって優先順位を割り当てているため、この入れ替えは隠れた遺伝子情報を発現することを可能とする。突然変異は、全項目、つまり「条件」部分の 8 箇所と「実行」部分の 9 箇所において等確率で起こる。

遺伝的操作を行う集団（生物種に相当する）としてのロボット集団の個数として、複数集団とするモデルと 1 集団とするモデルの 2 種類を考える（図 3）。複数集団が存在するモデルは、タスク実行時に各集団から 1 台ずつロボットをランダムに選択し、タスクを実行させる。遺伝的操作は集団毎に行う。集団毎にタスク実行時の初期配置が決定している。そのため、各ロボット集団は初期配置に特化していくと考えられる。1 集団のモデルは、集団内から複数のロボットをランダ

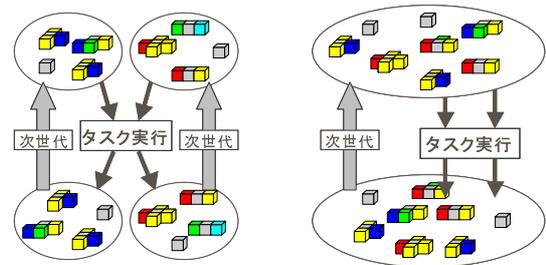


図 3: 集団構成に関する分類（左：複数集団，右：1 集団）

ムに選択し、同時にタスクを実行させる。集団内のロボットは、タスク実行時の初期位置が定まっていないため、様々な環境に対応できるロボットに進化すると考えられる。

適応度評価の方法として、各ロボットの得点を個別に評価して与える個別得点評価と、同時にタスクを実行した全てのロボットの合計得点を各ロボットに等しく評価値として与える合計得点評価（運命共同体的な状況にある複数生物個体に相当する）に 2 種類を考える。両者とも合計得点の増加を目的としているが、個別得点評価は、個々のロボットが得点を増加させていくことで、結果的に合計得点の上昇を図る間接的な方法であり、合計得点評価では、協調的な行動などを期待し、直接的に合計得点の上昇を図る方法である。

グループ選択は、複数のロボットをグループとして形成し、そのグループを単位としてタスク実行や遺伝的操作を行うことである。グループ内の各ロボットは、最初からタスク実行時の初期位置が決定している。遺伝的操作では、グループが親として選択され、交叉は各グループの対応するロボット同士で行う。各ロボットの役割分担を目指し、グループとしての得点向上を図る。細胞に寄生したミトコンドリアに関する仮説に近い、深い関係にある生物同士の関係などに例えられよう。

以上の分類をもとに 5 つのモデルを構築する（表 1）。複数集団であり個別得点評価するものを「複数集団個別得点評価モデル」、1 集団であり個別得点評価するものを「1 集団個別得点評価モデル」、複数集団であり合計得点評価するものを「複数集団合計得点評価モデル」、1 集団であり合計得点評価するものを「1 集団合計得点評価モデル」とする。グループ選択を行うモデルは、集団が 1 つで合計得点で評価する方法が想定可能であり、「組み合わせ固定モデル」と呼ぶ。

表 1: モデルの分類（上段は個体選択，下段はグループ選択）

	複数集団	1 集団
個別得点評価	複数集団個別得点評価 —	1 集団個別得点評価 —
合計得点評価	複数集団合計得点評価 —	1 集団合計得点評価 組み合わせ固定

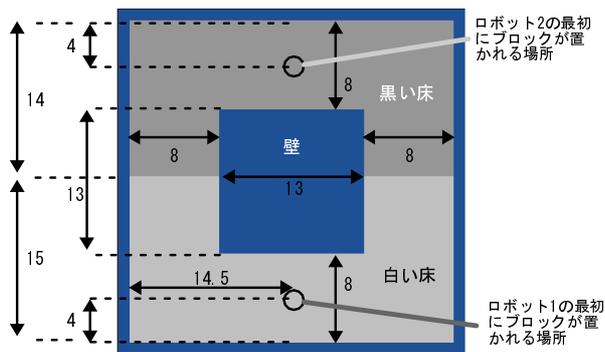


図 4: フィールド

### 3 シミュレーション

#### 3.1 設定

各種機能を持った立方体の部品を結合してロボットを構成するブロック型可変構造ロボット ROBOCUBE<sup>1</sup>の使用を想定している。ロボットを構成するブロックは、モータブロック、タッチセンサブロック、超音波センサブロック、光センサブロック、機能のないノーマルブロックの 5 種類とする。各ブロックの使用最大個数はノーマルブロックでは 5 つ、それ以外では 2 つとする。ロボットは、モータブロックが 2 つないと移動することができず、また、車輪の速度は一定で、10 ステップでブロック 1 つ分進むことができるものとする。発生段階はノーマルブロックが 1 つ存在する状態から開始する。

本論文では、フィールド (図 4) を壁に衝突せずに 2 台のロボットが動き回るというタスクを設定した。図中の数字はブロック一辺の長さを 1 とした距離を表している。同図中で、下方の位置よりタスクを開始するロボットをロボット 1 とし、上方の位置よりタスクを開始するロボットをロボット 2 とする。発生段階で最初に置かれたノーマルブロックが通過した面積を得点とする。一度通過した床は、再び通過しても得点にはならず、他のロボットがその床を通過しても得点にならない。取得可能な最高得点は 672 点である。ロボットは壁や相手ロボットに接触すると移動不可能になる。3000 ステップ経過するとタスクは終了する。黒い床は光センサ (下方を向いている) の使用を意識したものである。

パラメータとして、個体数は 200 (集団数 1 のモデル: 1 集団に 200 個体, 集団数 2 のモデル: 1 集団に 100 個体), 初期世代の染色体長 (遺伝子数) を 100, 突然変異率を  $1/300$  とした。初期集団の遺伝子はランダムである。遺伝的操作により 2000 より大きい染色体長を持つ個体が交叉によって生ずる場合は、親個体の選択からやり直す。

<sup>1</sup><http://www.watt.co.jp/>

#### 3.2 結果

各モデル毎に 5 試行の実験を行った。図 5 に、2 集団個別得点評価モデルによる実験結果の典型例を示す。同図では、最高合計得点、最高合計得点時のロボット 1 の得点、最高合計得点時のロボット 2 の得点及び、90 世代, 200 世代, 500 世代, 1000 世代のロボットの軌跡 (白線: ロボット 1, 黒線: ロボット 2) を示している。

90 世代では、ロボット 2 が壁伝いに動くようになり、200 世代では、フィールド左上でロボット 2 がロボット 1 に超音波センサを反応させることで軌道を変えて多く得点するようになっている。500 世代では、ロボット 1 の得点上昇がロボット 2 の得点減少に繋がっており、1000 世代では各ロボットが別の軌道を通ることによって全体的に高い点数を獲得している役割分担 (棲み分け) の様子が分かる。これは、ロボット 2 が使用するセンサを超音波センサからタッチセンサに変更したことで実現されている。

他の試行においても、片方のロボットがフィールドを回ることが可能になった後、もう片方のロボットも同じ方向に回るという行動に進化することが確認されている。ロボット同士の競合する部分を少なくする役割分担の様子は、他のモデルに基づいた実験から多少なからず観察されている。このことから、ロボットの構造と行動の共進化は、他のロボットの行動に対応した柔軟な進化を可能にしているといえる。

多くの試行において、最終的に進化したロボットの動きは、回転移動型と壁沿い型の 2 種類に分類できる (図 6)。回転移動型は、回転しながらセンサを利用し壁に反応して位置をずらしていく動きである。移動速度は遅いが、ロボットが回転するため、全方向にセンサを向けることができ、あらゆる方向の壁や相手ロボットに対応することができる。狭い範囲で確実に高い得点を得られ、地形や相手ロボットによる影響を受けにくい。壁沿い型は、ロボットが回転することなくセンサを利用し、壁に沿って進む動きである。移動速度が速く、広い範囲を動き、フィールドを一周できるものもある。地形に特化するものが多く、別の場所からタスクを実行させた場合、多く得点できないことが多い。地形や相手ロボットによる影響を受けやすいといえる。

さらに、得点の上昇を狙うために、5 つのモデルのそれぞれにおいて、エリート選択を導入して実験した。合計得点が最高であるロボットを 1 つのグループにして必ず次世代に残す操作である。図 7 に、各モデルに関する実験の 3000 世代における最高合計得点の 5 試行分の平均を比較した結果を示す。エリート選択を導入することで、全ての平均得点が上昇することが分かる。平均得点が最高であったのは、エリート選択を導入した 1 集団合計得点評価モデルであり、その次がエリート選択を導入した 2 集団単独得点評価モデルである。この両者は、エリート選択を行わない場合においても、他モデルと比べて高い点数を獲得している。表 1 において、正反対の位置にあるふたつのモデルが高い点数を獲得していることになる。

このことは、ロボット集団における多様性の度合い

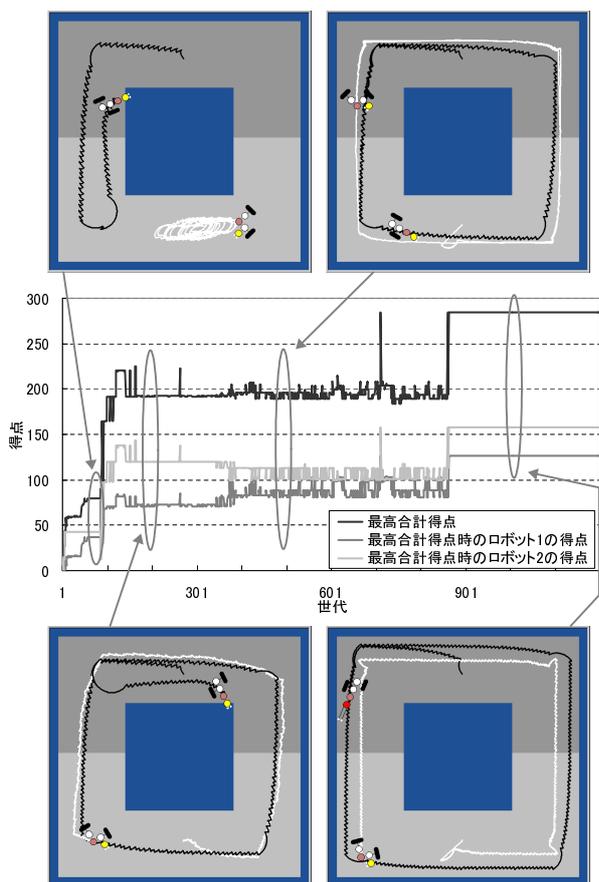


図 5: 2 集団個別得点評価モデルによる実験結果例



図 6: 2 種類の移動型

が関係していると考えられる。遺伝的アルゴリズムなどのメタヒューリスティックでは、探索と活用のトレードオフを適切に調整することが重要である [8][9]。適度な多様性が局所解収束と収束困難の両者の間の適当な状態をもたらさうからである。2 集団方式は 1 集団方式より高い多様性が期待できる。また、合計得点評価は個別得点評価に比べ、多様性を高める方向に働く。合計得点評価では、高い点数を獲得できないロボットでも、相手のロボットによって次世代に残ることができるためである。従って、多様性の高さが高すぎず低すぎないといえる集団合計得点評価モデルと、2 集団個別得点評価モデルが平均して高得点を得たと推測できる。

表 2 に、全ての試行における 3000 世代の合計得点が最高であったロボット 1 とロボット 2 の動きの組み合わせを分類したものを示す。同表において、「その他」に分類されている動きは、センサを使用した動きがで

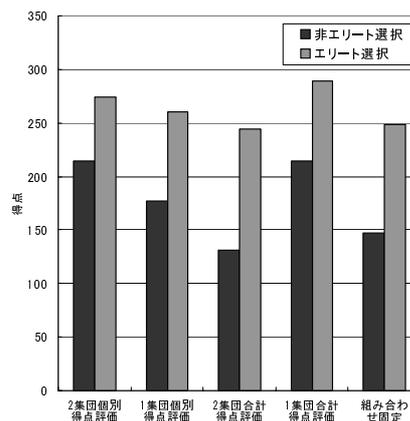


図 7: 5 試行分 3000 世代の最高合計得点の平均

きないものである。

エリート選択を行わない場合、2 集団合計得点評価モデルと組み合わせ固定モデルは、片方のロボットが「その他」になる場合が多い。これらのモデルは、片方のロボットが高い点数を獲得することで、残りの一台は得点する必要がなくなるため、片方のロボット集団だけが進化したと考えられる。さらに、片方のロボットが進化した後では、もう一台の小さな進化はほとんど影響を与えなくなってしまうことも考えられる。また、ロボット集団が 1 つのモデルは、他のモデルに比べて回転移動型になりやすい。回転移動型の特徴は、全方向にセンサが向くため、壁や相手ロボットに接触しにくく、環境の変動に強いことである。ロボット集団が 1 つのモデルでは、各ロボットはロボット 1 とロボット 2 のどちらかの位置に割り振られるか決定していないため、どちらの位置にも対応できる回転移動型が多くなったと考えられる。また、2 集団個別得点評価モデルでは、両方のロボットが壁沿い型に進化する場合が多いことも分かる。この理由としては、1) 2 集団個別得点評価モデルの場合、タスク実行時の初期位置が集団毎に決まっているため、タスク開始位置に特化した進化が可能なこと、2) 個別得点評価であるため、自分自身で高い点数を獲得する方向に進化が働くこと、3) 相手ロボットによっては、相手ロボットが接近してくることによる衝突が起こる可能性があるため、早期に点数を獲得できるほうが有利であることが考えられる。

エリート選択を導入すると、全てのモデルにおいて、

表 2: 実験結果の分類 (回 = 「回転移動型」、壁 = 「壁沿い型」、他 = 「その他」、セル内において、左側は非エリート選択、右側はエリート選択)

	2 集団 個別得点	1 集団 個別得点	2 集団 合計得点	1 集団 合計得点	組み合わ せ固定
回+回	0/2	3/1	0/0	2/3	0/0
壁+壁	3/3	2/3	0/3	1/2	0/2
回+壁	1/0	0/1	0/0	0/0	0/1
回+他	1/0	0/0	4/2	1/0	2/1
壁+他	0/0	0/0	1/0	1/0	3/1

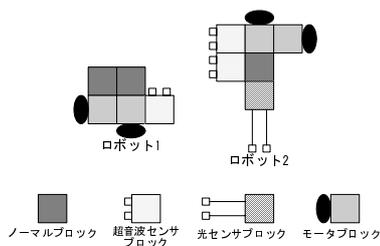


図 8: 合計得点が最高値を示したロボット

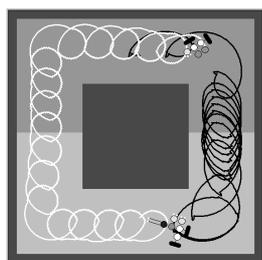


図 9: 合計得点が最高値を示したロボットの軌跡

2 台のロボットとも壁沿い型という組み合わせが増加した。壁沿い型は、移動速度が速いが、回転移動型に比べてセンサをあまり使わないので、初期配置や相手ロボットの行動に依存しやすい性質を持つ。従って、高得点と低得点のばらつきが大きい。この影響で、エリートとして選択される頻度が大きく、壁沿い型の組み合わせの増加を招いたと思われる。

図 8 に、全ての試行を通して合計得点が最大であったロボットの組を示す。このロボットの組はエリート選択を導入した 2 集団個別得点評価モデルに基づく実験で出現した。これらのロボットは回転移動型であり、合計得点は 370 点であった。図 9 に各ロボットの軌跡を示す（白線：ロボット 1，黒線：ロボット 2）。同図より、2 台のロボットがフィールド内部の広い範囲を動いた様子が分かる。両ロボットは、構成するブロックを全て使用していた。ロボット 1 では、使用しているセンサは 1 つの超音波センサのみであるが、2 つのノーマルブロックの発火状態の組み合わせにより、複雑なモータブロックの制御を実現していた。ロボット 2 では、3 つのセンサ（2 つの超音波センサと光センサ）を使用し、複雑な動きを実現していた。この結果から、各ブロックの取りうる状態数や機能の制約の中で無駄のないハードウェア設計の模索の跡が感じられる。

## 4 おわりに

我々は、複数のロボットによる協調作業を行う新しいロボットの自動設計手法に役立つ知見を得るために、ロボットの構造と行動の共進化モデルをマルチエージェント環境に応用した 5 種類の進化モデルを比較評価し

た。構築したモデルをもとに、壁に囲まれたフィールド内部を動き回るというタスクを設定し、シミュレーション実験を行った。その結果、ロボット集団の適度な多様性を保つモデル（集団数 1 で合計得点で適応度評価する方法、及び、集団数 2 で個別得点で適応度評価する方法）が効率的な進化を実現することが示された。また、構造と行動の共進化の結果、ロボットが回転しながら移動する回転移動型、壁に沿って移動する壁沿い型の 2 種類の移動方法が創発した。回転移動型は環境の変化に強く、壁沿い型は素早く点数を取ることができるなどの特徴があり、集団数 1 のモデルでは回転移動型が多く、2 集団個別得点評価モデルでは壁沿い型が多く出現した。

今後の研究の方向性として、ロボットへの実装を通して工学的実用性を一層追究するアプローチ、ロボット同士の相互作用が及ぼす構造と行動の共進化の様子から生物進化に関する知見を得るといった生物学的アプローチが考えられる。

## 参考文献

- [1] S. Nolfi and D. Floreano. *Evolutionary Robotics*. The MIT Press, 2000.
- [2] J. B. Pollack, H. Lipson, and P. Funes. Three generations of automatically designed robots. *Artificial Life*, Vol. 7, pp. 215–223, 2001.
- [3] T. Taylor and C. Massey. Recent developments in the evolution of morphologies and controllers for physically simulated creatures. *Artificial Life*, Vol. 7, pp. 77–87, 2001.
- [4] R. Pfeifer and C. Scheier. *Understanding Intelligence*. The MIT Press, 1999 (石黒他監訳：知の創成，共立出版，2001).
- [5] 朝井勇次, 有田隆也. ブロック型ロボットを用いた構造と行動の共進化の試み. 第 9 回 MPS シンポジウム予稿集, pp. 131–138, 2001.
- [6] 朝井勇次, 有田隆也. ブロック型ロボットを用いた構造と行動の共進化の試み. 情報処理学会論文誌：数理モデル化と応用, Vol. 43, No. SIG10(TOM7), pp. 110–118, 2002.
- [7] K.Nakano, K.Konishi, R.Ishiyama, and S.Ikeda. A self-organizing system with cell-specialization. In *Proceedings of 1997 IEEE International Conference on Evolutionary Computation*, pp. 279–284, 1997.
- [8] 柳浦睦憲, 茨木俊秀. 組み合わせ最適化–メタ戦略を中心として. 朝倉書店, 2001.
- [9] 中道義之, 有田隆也. ACO におけるランダム選択に基づく多様性調節の効果. 情報処理学会論文誌, Vol. 43, No. 9, pp. 2939–2947, 2002.