環境にナビゲートされた自律移動体の創発的行動形成と教示戦略†

塩 瀬 隆 之*・榎 木 哲 夫*・片 井 修**

Emergent Behavior Formation and Its Guiding Strategies for an Autonomous Mobile Robot Navigated by an Environment

Takayuki SHIOSE*, Tetsuo SAWARAGI* and Osamu KATAI**

This paper presents an evolutionary algorithm for the development of a behavior formation of simulated autonomous mobile robot, that is assumed to be equipped with some range sensors. We discuss that the key for the emergence of skillful behaviors exists in a open-loop cycle consisting of a robot's perception, action, and behavior intervened by an environment. We model this cycle by introducing an evolutionary computing method, which is applied to organize a robot's behavioral rules associating between a robot's perception and actions. Especially, for promoting a flexible behavior formation we stress on the importance of maintaining diversity of behavioral rules during the evolution. We consider about the preferred instruction strategies needed when an external human designer intervenes in a robot's behavior formation and makes it be guided in some way by indirect instruction.

Key Words: affordance, autonomous robot, evolutionary computation

1. 緒 言

近年、我々が対面するシステムの大規模化に伴い、処理過程に纏わる操作が複雑化し、システムを操作するユーザに多大な負担を強いる結果をもたらしている。このようなユーザへの負担を軽減するべく、ユーザの習熟度に応じて支援内容を適応的に調整する相互適応インターフェースに関する研究も行われており⁹⁾、今後はさらに能動的に、人間の習熟や熟練の過程そのものを支援出来る人工物システムの開発が要請されている。しかし人間の熟練過程については未だ一般的なモデル化はなされておらず、熟達を促進する影響因子についても技能遂行者自身の能力に見出す立場のみならず、これを取りまく作業環境や過去の経験など、広義の文脈(コンテキスト)にその因子を求める考え方も提示されている。本研究では、人間の敏速な熟練行動の説明記述をGibsonの生態学的認識論に求める⁵⁾。元々、知覚理論として提唱されたこの概念は、熟練行動のような人間の行為一般を説明記述する可能性を有しているとして近年再び注目され、その再解釈に基づいて工学的な実現を試みる研究も行われている⁸⁾⁴⁾。

Gibsonが強調するように学習過程における「相互作用」の重要性を踏まえれば、人間が何らかの作業を繰り返し経験する際の相互作用の累積がどのようにその後の学習に対して影響を与えるかについての一般的な原理を解明する必要があると考えられる。

そこで本研究では、超音波センサを有する自律移動体を想定し、恣意的な埋め込みを仮定せずとも、人間が熟練的に獲得するような注視点の遷移が見られることを計算機上で模擬し、その過程で見られる熟練の進化为「偶発的」ではなく、多様な環境との相互作用の累積に内在するある種の「秩序」による「必然的」な進化であり、それが行動主体内部における知覚-行為概念の形成過程に多大な影響を与えることを示す。

さらにこのような機械システム概念形成過程を方向づける外部環境との相互作用の影響を積極的に活用することで、環境に対する正確な内部モデルを有しない学習主体に対して、環境教示という形態での間接的な介入によりその学習過程を所望の方向に誘導することが出来る可能性を示し、それが人間と人工物との新たな協調関係の創出のための方策となり得ることについて考察を加える。

2. 生態学的認識論の示唆

2.1 関連研究

Gibsonにより提唱されたアフォーダンス概念は、そもそも知覚理論を発端としていたため、これに関連する研究は主に画像処理やロボットビジョンの領域に限られていた。ここでは、単に特定の行為の選択可能性を有するある環境の形状特徴をしてアフォーダンスとする狭義の解釈に留まっていた。

† 自律分散システムシンポジウムで発表(1997・1)

* 京都大学大学院工学研究科 京都市左京区吉田本町

** 京都大学大学院情報学研究科 京都市左京区吉田本町

* Graduate School of Engineering, Kyoto University, Sakyo-ku, Kyoto

** Graduate School of Informatics, Kyoto University, Sakyo-ku, Kyoto

(Received September 29, 1997)

(Revised May 13, 1998)

しかし本研究では Gibson が提唱したアフォーダンス概念には、相互作用の累積に伴う行為者-環境間の関係性の変遷そのものをも方向づけるような有用な情報をしてアフォーダンスとするより深遠な意味が含まれていると解釈し、これを狭義のアフォーダンスと対比して広義のアフォーダンスとして定義する。このようなアフォーダンス概念の再解釈に基づき、「環境との相互作用」の重要性を踏まえた研究も見られる。羽倉ら⁸⁾はロボットの学習を方向づけるものとして、岡田ら⁴⁾は日常的に見られるボトムアップな発話機構のモデルとして、それぞれアフォーダンスの工学的実現を試みている。すなわち環境から切り離して生成された従来の内部モデルの限界を指摘し、羽倉らはアフォーダンス概念が示唆する「環境の主観的な知覚」を踏まえ、環境との相互作用を通じて内部モデルを獲得することでその限界を克服することを提案している。一方岡田らは、計算機上で人工物同士の対話のプロセスを模擬することで、秩序だった行為や日常的な発話が、必ずしも事前に用意されたプランによるのではなく、むしろある瞬間ごとに振る舞いがボトムアップ的に組織化されることによって成り立つ点を強調している。このように振る舞いがボトムアップに組織化される際には、膨大な自由度の破綻が懸念される。しかし人間の場合は、相互作用を通じてアフォーダンスのような行為者と環境との間に予め潜在する情報をピックアップすることによって自身の振る舞いの自由度を低減させ、その時々脈絡に沿った会話へと収斂させている。すなわち、相互作用の累積が日常会話のような何気ない発話の繰り返しの方向性を誘引するのであって、その過程を事後的に意味付けて初めて合目的な行為であると事後的に解釈されることを指摘している。いずれの研究においても「(学習主体を取り巻く) 環境との相互作用」の側面について学習主体自身の学習能力同様に重視している。

本研究においても、単に知覚理論としての狭義のアフォーダンス解釈ではなく、むしろ相互作用関係の変遷そのものまで方向づけるような広義のアフォーダンスとしての解釈を工学的に応用するための方法論について考察することに主眼を置くものである。また「環境との相互作用」の多様性がその後の行動学習に与える影響について考察した後、それを教示戦略という形態で積極的に活用する方法について考察する。

2.2 熟練行動の発現

人間の技能形成・運動生成のモデルとしては、従来からの内部システム(主体)と外部環境(客体)の独立性、いわゆる主客分離を前提とする情報処理的アプローチに対して、近年、Gibson による「アフォーダンス」⁵⁾に象徴されるような生態学的アプローチが注目を集めている。我々はこの考えに同調し、環境の中から有用な情報を抽出するプロセスそのものも、環境の中に「情報の構造」として内在していることを示唆してきた⁶⁾。そこでは認知的資源配分問題の一例として注視点配置問題を対象とした。そして環境に満ちた情報の中から無限でない認知的資源の制約下で有限の時間内に、環境の側が行為者に対して行為の可能性についての予見的情報

を提供するような特徴を行動主体にとっての注視点として定義し、それらの巧みな配し方は環境との相互作用の繰り返しを経験とした学習により進化するものと解釈している。

しかしここで我々が着目すべきは、熟練行動の事後的な説明記述ではなく、熟練行動を獲得する過程そのものである。通常人間の熟練過程においては、新規に遭遇した環境に対しては、不足した情報を補うためにあらゆる方向へ注意が注がれ、客観的な評価ではその注視先を同定することが困難であるものの、その環境での経験を重ねると注視点の乱立は減り、注視点は明らかに特定の場所にのみ注がれ、その注視先も客観的に意味付けが行い易いものとなる。このように注視点の乱立が、その環境での経験を重ねることで、対象の特定箇所に注視先が限定されていく過程を行動主体の学習として捉え、行動意図と整合性を有するようになる行為を熟練行為と呼ぶことにする。ここで注視点を洗練させていく熟練過程は、知覚-行為の体系としての行動規則をより適応的な振る舞いを産出出来るものに洗練化していくプロセスであり、その体系を行動規則に反映出来るまでは行動主体の振る舞いは稚拙で、ぎこちないものとなるものの、ここで培われる多様性の蓄積はその後の学習を方向づける上で重要な意味を持つ。そしてそのためには、外界からやってくる刺激の受動的な受け手ではなく極めて具体的な意味で環境とインタラクションを重ねながら自身の世界をイメージとして自己の内部につくりあげ、それがさらにその後のインタラクションを方向づけるような機構、いわば環境との相互作用を陽に含む開かれたループとしての学習機構を構築する必要がある。

2.3 知覚-行為体系の獲得と進化型計算

行動主体は知覚-行為の体系を行動規則として持ち、この規則に基づく環境への働きかけが挙動として産出される。学習の進行の過程では、ある行動規則に従った挙動は行動主体にとっての、環境の“見え”，すなわち環境内のいずれの箇所に注視先を配するかを変化させ、行動主体は新たに得た知覚と行為との間の関係づけを行なわなければならない。そうして得られた新たな行動規則はまた異なる挙動を導く。学習の初期にはぎこちない挙動を導いた散漫な注視先は、試行錯誤的な学習の進行に従って、洗練された挙動を導く妥当な箇所へ向けられるようになる。これら環境に開かれたループ内で、挙動と知覚-行為の体系とが相補的、段階的に洗練されていく過程こそが熟練過程であり、この過程そのものが相互作用を通じて環境に内在する「秩序」(体系付けの段階的な遷移そのものを方向づけるバイアス)により方向づけられる。このようなサイクルは動的に変化しつつ、環境からのバイアスを受けながら適応的な行動規則を獲得するまで繰り返される。このように段階的に形成される知覚-行為の体系の遷移の様子は人間の熟練的な学習の過程にも類似するものである。

本研究では、簡単な知覚-行為の機能を有する自律移動体の行動形成の問題を取り上げ、上述の行為→挙動の関係を能動的な経験(試行)収集のフェーズとして、またその挙動が評価されることによって間接的に環境に対する知覚の変化が

もたらされるようなサイクルを進化型計算を導入することにより構成する。特に、行動主体にとっての行動規則の集合の生成を、環境中でうまく振る舞うように行動形成を進める上で必要になる環境概念の一般化と特殊化⁷⁾のような互いにトレードオフの関係にある両圧力に対応する遺伝的オペレータを導入し、環境との相互作用の中から、適切な抽象度を有する行動規則の集合に最適化し獲得するための進化型計算を導入する。

3. 行動からの自律移動体の注視点獲得

3.1 能動的行動獲得アルゴリズム

本研究では行動を繰り返すことにより、環境に対する知覚が推移していくプロセスを進化型計算を用いてモデル化する。そこでさまざまな問題に対して高い一般性を持つ遺伝的アルゴリズムを逐次意思決定作業という特殊な領域に適用した Grefenstette による SAMUEL^{2) 3)} のシステムに倣って、シミュレーション上での試走を繰り返し、評価、世代更新のサイクルの中で自律移動体が環境の中で注視すべき点を見出すことを目指す。以下本文では、このアプローチを能動的行動獲得アルゴリズムと呼ぶ。対象となるのは簡単なオフィス環境を想定した所与の交差点を通過する自律移動体である。自律移動体の構成は、本体のサイズを幅 36cm、長さ 56cm、高さ 68cm とし、障害物までの距離を検出する 7 個の超音波センサ (S1~S7) を搭載するものとする。各センサは自律移動体の正面方向に対して 0 度、± 30 度、± 60 度および ± 90 度の方向に向けて固定されている。センサが測定可能な距離は、自律移動体の中心から見て 40cm から最大 500cm の間に限られる。とりうる行動としては、move Forward (20cm 直進する)、turn Right (右に 15 度方向を変え 20cm 進む)、turn Left (左に 15 度方向を変え 20cm 進む) の 3 種類に限定し、1 ステップ時間に 20cm 進行するものとする。本システムにおいて、遺伝的アルゴリズムの個体に相当するものは超音波センサによる知覚と操舵方向に関連づける行動規則の集合である。進化型計算の初期化のフェーズにおいてのみ 8~16 個の間で生成するという制限を設けているものの、進化過程では個体内のルール数に制限は設けていない (実験では 30~70 個の範囲で変動していた)。自律移動体の各時間ステップにおける各センサ方向の距離情報を便宜上 4 つの離散領域に分け、i 番目のセンサの反射波による計測距離がどの領域 Layer n ($n = 1, 2, 3, 4$; 1: 最近接, 4: 最遠方) に属するかを (Si:Ln) と記述し、それらセンサ情報の連言を前件部とし、自律移動体が行う 3 つの行動 (move Forward, turn Right, turn Left) のいずれかを後件部として行動規則を表現する。行動規則の表記例を以下に示す。

Rule1 :if (and ((S2:L3) (S4:L4) (S7:L2)) then (move Forward)

また行動規則の前件部を構成する (Si:Ln) の各々を一つの注視点として定義する。これら行動規則の集合を個体として個体数を 100 に固定して進化型計算を行う。ここでの試走

は、L 字路や十字路など日常多く見られるような交差点の通過事象を想定して、シミュレーションを行う。そして各世代で各々の個体を持つ行動規則を発火させながら仮想環境内で移動体を走行させ、壁に衝突することなく交差点を通過することができるか否かの結果から各試走における個体の評価を行う。一方世代更新は以下に述べる遺伝的オペレータにより実行する。ある行動規則に対して、より少数のセンサ観測を前件部に含むような行動規則に変更することは、そのセンサ観測値以外はどうような観測であっても発火され得ることを意味する (don't care)。すなわち、行動規則の条件部が照合する状況の外延集合を拡大することに相当し、これを行動規則の一般化と呼ぶ。その逆、すなわち条件部に新たなセンサ属性に関する制約を付加して状況の外延集合を絞り込むことを特殊化と呼ぶ。本研究で用いた一般化オペレータは、

- 条件属性の脱落：着目するセンサの数を減らすこと
- 条件属性値域の併合：ある一つのセンサについて着目するレイヤの数を増やすこと

の 2 種類で、その具体例を以下に示す。

Rule1 : if (and ((S2:L3) (S4:L4) (S7:L2)) then (move Forward)

↓<条件属性の脱落>

Rule2 : if (and ((S2:L3) (S4:L4)) then (move Forward)

↓<条件属性値域の併合>

Rule3 : if (and ((S2:L3) (S4:L3,L4)) then (move Forward)

以上の逆の操作を行うのが特殊化オペレータである。さらに行動規則の後件部の行動をランダムに変更する操作を突然変異オペレータとして用意している。本進化アルゴリズムにおいては、ランダムに生成した 100 個体の初期集団から初めて、これら試走→評価→世代更新のサイクルを 500 回繰り返したところで学習の終了とし、得られた行動規則の集合が持つ行動規則の前件部に含まれる記述を熟練行動の結果獲得される注視点として解釈する。世代更新のフェーズは以下の 2 つのステップからなる。1) 試走後、すべての個体を評価の順に並べ、まず評価の悪い下位 10% の個体を無条件に棄却する。次に残り 90% の個体からルーレット選択により親個体を選定し、その交叉により得られる子個体で損失した 10% の個体を補填する。2) 集団のエリートとして 10% の個体と、生成されたばかりの 10% の子個体とを除いた残り 80% の個体について上述の遺伝的オペレータ (一般化、特殊化、突然変異) を適用する。遺伝的オペレータの適用は各個体を構成する各々の行動規則に対して、一般化、特殊化、突然変異のオペレータをそれぞれ 40%、40%、20% の確率で適用する。特に一般化、特殊化のオペレータについては等確率にランダムで適用されるため、設計者が恣意的に集団に対して一般化や特殊化の方向づけを行うことはないものとする。

また試走のフェーズにおいて個体の中の複数の行動規則が

同時に発火する（競合状態）ことが懸念されるが、本進化アルゴリズムではランダムな選択により競合を解消する。発火の際に恣意的な競合解消のアルゴリズムを設けないのは、競合解消の過程に複雑なアルゴリズムを適用することでルール選択の即応性が損なわれることを危惧したことと、世代更新のフェーズで行われる選択淘汰との二重の淘汰圧による冗長な処理を回避することの両方を考慮してである。

3.2 行為の分化と環境知覚

先述の能動的行動獲得アルゴリズムを用い、Fig.1に示すような道幅 300cm の通路環境での通過を目標とする実験を行った。それぞれランダムシードが異なる実験を 20 回行い、平均値による評価を行った。ここで超音波センサは 400cm までに存在する障害物を検知可能であり、障害物までの距離を 100cm ごとに区切った 4 つのレイヤの離散値で表す。図中に示される線分は、この環境を実際に走行した際に用いられた行動規則の前件部、すなわち超音波センサの最小単位として規定した個々の注視点を示す。ここで注視点の配置先について詳細に述べることは節の論点ではないことから、Fig.1に示すのは成功事例の中の一例を無作為に選出した。

3.3 環境からの創発的注視点生成

前節の実験においてあるコーナを通過する挙動に着目したところ、学習の進行に伴って環境に配られる注視点の総数が減少し、さらには自身の足元（近接空間）に配られていた注視点自身が自身の進行方向の遠方（遠隔空間）へと遷移していくのが観測された⁶⁾。この様子を顕著に表している一例を Fig.2 に示す。

本実験においては、前件部の収束方向や注視点の総数の増減に関して設計者としては何ら方向づけを行っていないにもかかわらず、人間の熟練行為に類似とも思える注視点の遷移が見られたことは、自身の内部に予めインプリメントされたものではなく、環境との相互作用によりそのようなバイアスがかけられた結果であると言える。また注視点の総数が減少していくことは、行動規則の前件部が照合する外延空間を拡大することに相当し、その行動規則が適用可能な環境状況の総数は逆に増える。すなわち“環境の見え”については一般化が行われたと言える。その一方でそれぞれの“環境の見え”については異なる行為が特定されることとなり、行為については分化が起こっていると考えられる。この行為についての分化を確認するため、Fig.1の環境における学習の過程にて獲得された行動規則集合の中から後件部に同一の行為を持つ行

Fig. 1 The given environment

動規則毎にその前件部のいずれの方向に最も注視点を配しているかの注視頻度を各世代ごとに累積したグラフをレイヤ 4 についてのみ Fig.3 に示す。ここではシードの異なる 20 回の実験で得た成功個体のうち、それぞれ上位 10 個体のエリート個体全てについての平均をとった。

グラフは横軸がセンサ方向 (S1~S7)、縦軸が世代 (0~500) を表す。各々のセンサについて、この方向での知覚が行為として L,F,R (turn Left, move Forward, turn Right) のいずれを規定しているかについてその頻度を示している。縦軸は世代方向、垂直軸は注視頻度を示す。左から順に灰、白、黒のグラフがそれぞれ L,F,R に相当する。明らかにそれぞれの行為については明確に前件部が生成し分けられていることが確認される。次に距離方向の遷移としては、自身の足元（近傍）から、進行方向の遠方（遠隔）へと注視点が遷移することが確認されるが、それを確かめるべく、ここでは左曲がりや後件部を持つ行動規則のみについて注視頻度を調べた。Fig.3と同様、20 回の実験からそれぞれ 10 個体ずつのエリート個体全てについて平均をとったグラフを Fig.4 に示す。

Fig.4から明らかなように学習の初期 10~40 世代では、自

Fig. 2 The robot's gazes at each generations ((a)10th generations and (b)200th generations)

Fig. 3 Action distinguish as the robot learning (Layer4)

Fig. 4 Evolutional changes of gazes attended to four different distances

身右側のセンサ (S6, S7) の近接空間 (第1, 2レイヤ) で壁を注視していた注視点が学習に伴って減少し, 自身左30度前方 (S3) すなわち自身の進行方向の遠隔空間 (第3, 4レイヤ) で広がりを見つめるような注視点が増加していることが分かる (Fig. 4(c)). Fig. 3や Fig. 4で行った平均値による評価との類似性から, Fig. 2での議論を一般化出来ることが分かる. 今回適用したアルゴリズムでは行動規則の前件部の変化については一切のバイアスをかけていないにもかかわらず, 注視点が学習に伴って自律移動体自身の足元から前方向遠方へシフトし, かつ頻繁に注視されている方向が限定されていく様子は人間の無駄のない熟練行動との類似点として確認できる.

3.4 必然としての注視点遷移

ここで我々は上述の注視点形成が行為の分化の過程で必然的に生じていることを検証するために, 身体的制約としてセンサの可到達距離を400cmから200cmに短縮した自律移動体を想定して実験を行う. これにより先ほどの実験結果のように自身の遠方 (300~400cm先) に注視点を配することは不可能となる. センサ長さ200cmにて Fig. 1の環境を学習する実験を20回ずつ行ったところ, 失敗個体の80%が第3コーナ (図中○部分) に行き詰まっていたことから, 自律移動体にとって第3コーナを通過するためのルール生成が困難であることが分かる. これを顕著に表した事例を Fig. 5に示すが, ここでも注視点の配置先について詳細に議論することは焦点ではないことから無作為に選出した.

行き詰まっている第3コーナでは左へ舵をとる行動規則が発火することが失敗の原因と考えられる. センサレンジの短

Fig. 5 The failure of the robot with a narrow reach sensors

Fig. 6 The failure of distinguish with a narrow reach sensors

い移動体にとっては第3コーナ右下の凹角へ配した第4レイヤの注視点が, その方向へ広がりを持つかのこどく錯覚を生み, その方向へ舵をとった結果失敗が生じている. センサレンジが短いからといって右曲がりが不可能なはずはなく, 左曲がりと右曲がりを同時に経験する上で左曲がりと右曲がりの行為を明確に区別する前件部の生成がこのセンサ範囲では出来なかったのが失敗の原因である. ここで第4レイヤにおいて生成される前件部について各行為別の注視頻度のグラフを Fig. 6に示す. 先程とは異なり学習が進行しても行為による分化が見られず, 各注視点は平均的に配される. 明確な分化がなされなかったことが, 失敗の原因であると考えられる.

従って同一のアルゴリズムであるにもかかわらず異なる行為を弁別可能な前件部を生成出来るか否かの差異は, 自身の200~400cm先の情報を取得できるか否かの差であり, センサ長さが400cmまで取得可能であるならばこの領域まで注視点が遷移することは, 行為を明確に区別する上での必然であると解釈できる. 以上のことから, 環境の一般化概念を獲得する過程 (前件部の不要な注視点を捨象する過程) は行為については分化を行なっている過程 (異なる行為を明確に弁別可能な箇所へ注視点を配する過程) であるとも言える. 行為を決するための情報を自身のより前方へと求めるための動機は自身の内部に特別な埋め込みを行った結果ではなく, 人間の知的な熟練行為同様, 自らの身体を介して環境世界にインタラクションを求めた結果として, 環境との関わり方が学習を方向づける形で事後創発的に行為を弁別するに十分な注視点を獲得したものと考えることができる. 実際このような技能創発において身体性の関与の重要性を指摘している研究者も多い. これら環境との関わり方を意識することで, より頑健性の高い学習機構を実現できるものと考えられる.

Fig. 7 Three types of environment as teaching materials

4. 学習を方向づける教示戦略に関する考察

4.1 身体性に応じた学習

生涯に経験するであろう環境の数は未知であり、個体のセンサ能力や学習アルゴリズムといった自己の身体性や学習能力の限界が、そのまま環境への適応性の境界を決めてしまわないように行動系列を獲得することは有用である。ここではセンサ長さ 200cm の移動体を、前章では通過できなかった Fig.1 で示される環境において学習アルゴリズムの変更なく行動系列の獲得を改善するための手法について考察する。前章で議論したのは環境との相互作用が学習に与える影響、その方向性についてであり、学習の方向づけを行なうものが必ずしも自身の内部に存在する必要があることを示唆した。そこで環境との関わり方そのものに変化をつけることで、異なる行動系列の獲得が可能になると考えられる。すなわち目的の環境を経験する前に、前学習として別の環境を経験させる事で学習の収束の方向性が変化し、行動系列の獲得過程に別の方向性をつけることが可能になる。そこで自律移動体に対して目的の Fig.1 を経験する前に、別の環境を前学習の教材として提示する事で学習に方向づけを行うことを考える。以下では前学習としてどのような環境を提示すればよいか 3 つの環境を教材として考察を行うが、前学習の効果を評価する環境は Fig.1 同様、左右のコーナを連続して経験するような環境を想定し目標環境とする。

4.2 前学習の教材としての環境

前学習の候補として、道幅は Fig.1 と同じく 300cm であるが、形状の異なる Fig.7 のような 3 つの環境を考える。予め前件部を右方向へ方向づけることを目的とした (a) 単純な右曲がりを要求するもの。次に、カーブの通過を易しくして、左右両方向に曲がれる方向づけを行うことを目的とした (b) 緩やかなカーブ (135 度の鈍角) を経験するもの。最後に空間の広がり方が連続的に変化するようなより多様性を内包する (c) 交差点間の距離を最小 (0cm) にしたものの。これらの環境において 500 世代まで学習が進んだところで、目標環境下での学習を始める。(a) (b) については 500 世代まで学習したところで集団の 90% が成功しているのに対して、(c) で

Fig. 8 The rate of successful passage after pre-learning

Fig. 9 Action distinguish after pre-learning in environment (b)

Fig. 10 Action distinguish after pre-learning in environment (c)

は 40~50% の成功個体率であったが、そのまま転用することにした。Fig.7 に示したのは、それぞれ 500 世代まで学習した成功事例の中から無作為に選出した一例である。ここでも注視点の配置先について詳細に議論することは節の論点ではない。500 世代まで前学習を行い、その後目標環境である Fig.1 において 500 世代の再学習を行う実験を、それぞれ異なるランダムシードで 20 回ずつ行った。20 回の実験で得られた全個体についての成功個体率の平均を Fig.8 に示す。前学習の効果が見られたものは環境 (c) のみであり、目標環境に移した当初は集団の 1~5% しか成功しなかったものが、20% の成功率はおさめるようになっている。

4.3 行為の分化と多様性の維持

ここで前学習が前件部の生成においてどのようなバイアスをかけたかを検証するために、前学習から再学習にかけて変化した注視頻度のグラフをそれぞれの環境毎に生成する。ここでは特に環境 (b), (c) をそれぞれ前学習の教材として選択した場合の第 4 レイヤについてのみ Fig.9, Fig.10 に示す。ここでは Fig.3, Fig.4 と同様、20 回行った実験で得られた成

功事例の中で上位10個体ずつのエリート個体の全てについて平均をとった。

前学習を行っていない場合 (Fig.6) と、行った場合 (Fig.9, Fig.10) との差異は行為による分化が明確に行われているか否かの差異を生んでいる。Fig.6と Fig.10の比較からは適度に平易な環境を前学習の教材として経験させることで、前件部の分化をある程度進めてから転移させることが出来たことが有効であったと考えられる。一方、Fig.9と Fig.10の比較からは、前述のような前学習時での行為による分化が行われていても、必ずしもこれが後学習時に貢献をもたらすものでもないことも認められる。すなわち Fig.9においては、前学習終了時 (500 世代) にすでに行為による分化が十分に進行しており、このことがその後の異なる環境下での後学習に対して新環境に適合した行動規則の生成を不可能なものとしているのに対し、Fig.10では、前学習中期から終了時に維持されている知覚-行為の体系付けにおける多様性の存在が、その後の柔軟な行動規則生成に寄与したところが大きいことを示唆している。環境 (c) の適度な難易度が前学習における不確定性を維持したため、環境 (b) で見られたような過学習を回避することが出来たと解釈出来る。ただし、環境 (b) と (c) とで、パフォーマンスの改善に差異が生じたのは、環境 (c) の幾何的な形状そのものが目標の環境に対して寄与を持っていたからか否か、さらなる詳細な解析を要する。

4.4 創発的行動形成に向けて

以上本論文では、原始的な知覚能力を有するロボットのような人工物が、環境との相互作用の繰り返しから熟練行動を獲得するプロセスについて進化型計算を導入したモデル化を行った。とくに熟練行動を、ある特定環境下での相互作用から必然的に適応行動が漸次的に獲得されるという短期的・近視眼的進化 (microevolution) に加え、遭遇環境の変動に伴う獲得行動の進化途上で形質保存性に着目した実験結果から、能力的・身体的に限界を有する人工物が固定環境に特化されないより長期的な安定した進化 (macroevolution) を経る熟練行動を獲得する上で、ロボットが遭遇することのできる教示環境の多様性や順序性の差異という「偶発性」が影響因子として作用することを実験的に確かめた。

近年、行動パターン形成に関する研究として、群ロボットの行動創発に関する研究が盛んに進められている。そこでの技術課題の焦点は、個々のロボットが自らの空間的近傍に関する情報処理と近傍の他者との相互作用から、いかにして巨視的行動パターンが形成され、それが個人の微視的な情報処理や相互作用を規定するに至るかについての創発規範 (emergent norm) 説の工学的実現である。本論文で扱ったのは単一の行動主体としての自律移動体の熟練行動の形成であり、群ロボットとのそれとは目的を異にするものではあるが、空間的な近傍性の代わりにスパンの異なる時間的経過をたどって形成される微視的・巨視的進化の両プロセスの相互作用による創発原理という点からみた両者の共通点は興味深い。すなわち、巨視的行動パターンの形成・定着過程には、

ある特定の微視的相互作用の偶発的な発生による巨視的行動パターンの「核」の生起の過程と、この「核」が生じた後の、個々の微視的パターンの形成がその方向性に追従し、巨視的行動パターンに引き込まれていくという過程、いわば「偶然」の過程と「必然」の過程が並存することが明らかになっている。本節での実験結果でも、前学習と再学習での教示環境の適切な選択という偶然性によりロボットのライフスパンでの進化の文脈が方向づけられ、その制約下で例え同一環境のもとでも適応した行動形成の成否が影響を受けることを示した。本論文で提唱した進化型計算によるモデル、ならびに本論文の冒頭部で考察した広義のアフォーダンスのもとに展開されている種々の関連研究では、未だ上述の「必然」のプロセスを解明するにとどまるものではあるが、今後はいま一つの「偶発性」の生起の過程をいかに設計できるかについての研究の展開が求められる。すなわち環境が変動した際にそこででの行動形成が進まなくなるような膠着状態に陥ることを回避できるように、より早い段階での「核」形成を促す初期微動をどのように与え得るかについて、ロボットへの教示環境を適切に選択することで達成可能であると考えられる。この部分の一般化については現在考察を進めている。

5. 結 言

本研究では超音波センサを有する自律移動体を想定し、「創発」現象の発現の可能性を有する進化型計算を適用することで、環境の直接知覚から獲得した行動戦略の事後的な意味付けを行い、注視点生成のプロセスが自身の内的処理や設計者の埋め込みではなく環境との相互作用に内在する必然性であるとの解釈を与えた。環境の側から学習を方向づけるバイアスの存在を我々設計者が意識していない故に、その必然としての進化のプロセスは「創発」的な進化と解釈出来ることを示唆した。またセンサ能力の乏しい自律移動体の行動系列の生成を効率よく誘導するためには、それら環境のバイアスを踏まえた教示戦略が必要であることについて考察を行なった。さらにこの一連の研究を通じて、「必然性」と「偶発性」の関与という立場から、創発的行動形成に関して考察を行った。今後は、系の中での「必然性」と「偶発性」をより明確に定義し、いかに「偶発性」の生成過程を設計出来るかをより一般的に考察することが望まれる。なお本研究は、文部省科学研究費補助金重点領域研究 264「創発的機能形成のシステム理論」の援助によるものである。

参 考 文 献

- 1) R. Brooks: A robust layered control system for a mobile robot, *IEEE Journal of Robotics and Automation*, RA-2(1), Mar. (1986)
- 2) John J. Grefenstette and H. G. Cobb: Learning the Persistence of Actions in Reactive Control Rules, *Proc. of the Eighth International Workshop on Machine Learning*, 293/297 (1991)
- 3) John J. Grefenstette and D. F. Bordon: Explanations of

Empirically Derived Reactive Plans, *Proc. of the Seventh International Conference on Machine Learning*, 198/203 (1990)

- 4) 岡田, 鈴木, 猪口: ボトムアップな情報戦略, 第10回人工知能学会全国大会講演論文集, 441/444 (1996)
- 5) 佐々木 正人: アフォーダンス, 岩波書店 (1994)
- 6) 塩瀬, 榎木, 山下, 片井: 進化型計算によるアフォーダンス概念の獲得と非言語コミュニケーションに関する考察, 計測自動制御学会第23回知能システムシンポジウム講演論文集, 53/58 (1996)
- 7) 塩瀬, 榎木, 堀内, 片井: 進化型行動形成における一般化概念の獲得と教示戦略に関する考察, 第23回システム/情報合同シンポジウム'96講演論文集, 41/46 (1996)
- 8) 羽倉, 横井, 嘉数: アフォーダンス理論に基づく自律エージェントの環境認識機構, 第5回マルチエージェントと協調計算ワークショップ, On-Line Proceedings (<http://www.csl.sony.co.jp/person/nagao/macc95>) (1995)
- 9) 福田, はか: 再帰形ファジィ推論を用いた相互適応インターフェースの概念と実現, 日本機械学会論文集, 60-571, C, 970/977 (1994)

[著者紹介]

塩瀬 隆之 (学生会員)

1973年生.'96年京都大学工学部精密工学科卒業.'98年同大学院工学研究科精密工学専攻修士課程修了.現在,同大学院博士課程在学中.'98年より日本学術振興会特別研究員.主として人間-機械共生系におけるコミュニケーション支援に関する研究に従事.計測自動制御学会の会員.

榎木 哲夫 (正会員)

1957年生.'81年京都大学工学部機械系学科卒業.同大学院工学研究科精密工学専攻修士課程・博士課程を経て,'86年より同大学工学部助手(精密工学科).'94年同助教授,'96年同大学院工学研究科助教授(精密工学専攻),現在に至る.工学博士.'91~'92年米国スタンフォード大学客員研究員.主として人間-機械共存環境下での協調システムの設計と知的支援に関する研究に従事.日本機械学会,システム制御情報学会,日本ファジィ学会,日本人工知能学会,航空運航システム研究会,IEEEなどの会員.

片井 修 (正会員)

1947年生.1969年京都大学工学部機械工学科卒業.同大学院機械工学第二専攻修士・博士課程を経て,1974年同大学工学部助手(精密工学科).1983年同助教授,1994年同教授,1996年同大学院工学研究科教授(精密工学専攻),1998年同大学院情報学研究科教授(システム科学専攻),現在に至る.工学博士.その間,1980-1981年フランス国INRIA(国立情報処理自動化研究所)客員研究員.主としてシステムの知能化に関する研究と教育に従事.計測自動制御学会,日本ファジィ学会,人工知能学会,システム制御情報学会,日本機械学会,精密工学会,進化経済学会などの会員.