未知環境中における移動ロボットの 知能的動作

——階層構造確率オートマトンからの接近——

馬 場 則 夫*

An Intelligent Behavior of a Mobile Robot in an Unknown Environment

—An Approach Using the Hierarchical Structure Stochastic Automaton—

Norio BABA*

This paper is concerned with an intelligent behavior of mobile robots going through a maze having numbers of gates that close with unknown rejecting probabilities.

After the brief introduction concerning learning behaviors of the hierarchical structure stochastic automata, the absolutely expedient learning algorithm under the nonstationary multiteacher environment is proposed. It is shown that the hierarchical structure stochastic automata with this algorithm can be successfully applied in order to let mobile robots gradually find the best way in the maze.

Key Words: mobile robot, hierarchical structure stochastic automata, maze, learning algorithm

1. はじめに

人間が立ち入って作業するにはあまりに危険が伴うようなところでは、できるだけロボットにその作業の肩代わりをさせる必要がある。特に、原子力関係その他の人体に直接危害が加わる可能性のあるような作業場では、ロボットの導入が必要不可欠である。しかしながら、上述のような場合には、しばしば、人間の監視が十分に行き届かないようなところでロボットに作業させねばならない。それゆえ、ロボットに自分自身で判断し行動をする能力すなわち周囲環境を適切に判断する知能をもたせることが特に重要となってくる。ところで、ロボットの知能化に対する研究は、最近よ

うやく始まったばかりであり、今後の研究成果の積み重ねに期待されるところが大きいのが現状である。

本論文では、複雑に入り組んだ迷路を通過してゴールへ到着することを目的とする移動ロボットの知能的動作に関する研究を報告する。特に、迷路内には多くのある未知確率でもって開閉するゲートがあり、移動ロボットがいかにすれば最も通過しやすい経路を首尾よく見い出すことができるかという問題を考察する。そして、階層構造確率オートマトンの概念を利用することにより、移動ロボット自身が自動的に最適な経路を見い出すことができるようになることを示す。

本論文の構成について簡単に触れよう。まず、前半(2章・3章)では、階層構造確率オートマトンが移動ロボットの知能化にいかに関与するかを示すための準備として、階層構造確率オートマトンの概念を紹介するとともに、優れた学習効果を保障する学習アルゴリズムを提案する。本論文の後半では、まず(4章)、入り組んだ迷路の通過を目標とする移動ロボットの学習的振舞いについて考察し、3章で提案された学習アルゴリズムに従って経路を決定する場合、首尾よく最も通過しやすい経路を自動的に見い出すことができることを示す。最後の章では、4章までに得られた結果の検討と問題点の指摘を行い、更に今後の見通しについて述べる。

2. 階層構造確率オートマトン

Fig. 1 はレベル数3の非定常複数教師環境中で動作する階層構造確率オートマトンを示している(これは Thathachar & Ramakrishnan⁸⁾によって考えられた学習モデルを複数教師環境の場合に拡張したもので

* 徳島大学工学部 徳島市南常三島町 2-1

* Faculty of Engineering, Tokushima University, Tokushima

(Received November 26, 1986)

(Revised March 18, 1987)

Fig. 1 Learning mechanism of the hierarchical structure stochastic automata

ある). (注 1)

図から明らかなように、第 1 レベルは確率オートマトン $\bar{A}$ 、第 2 レベルは r 個の確率オートマトン $\bar{A}_1, \bar{A}_2, \dots, \bar{A}_r$ 、第 S レベルは r^{s-1} 個の確率オートマトン $\bar{A}_{k_1 \dots k_{s-1}}$ ($k_i = 1, \dots, r; i = 1, \dots, s-1$) から階層構造は構成されている。そして、階層構造内の各オートマトンは r 個の出力をとりうる。

階層構造確率オートマトンの動作について述べよう。初期時刻 ($t=0$) における出力確率は、すべてのレベルの確率オートマトンに対して等しく $1/r$ とする。さて、時刻 t において第 1 レベルの確率オートマトン $\bar{A}$ が出力確率分布 $(p_1(t), \dots, p_r(t))'$ (注 2) より j_1 番目の出力 α_{j_1} を出したとする。 α_{j_1} に対応して、第 1 レベルの r_1 人教師環境は、 r_1 個の返答 $\beta_{j_1 i_1}(t, \omega)$ ($k_1 = 1, \dots, r_1$) (注 3) を返す。第 2 レベルでは、 $\bar{A}$ の出力 α_{j_1} に対応して確率オートマトン $\bar{A}_{j_1}$ が起動する。 $\bar{A}_{j_1}$ は t 時刻における出力確率ベクトル $(p_{j_1 1}(t), \dots, p_{j_1 r}(t))'$ に従って r 個の出力のうちのどれかを出す。以上のような動作が、最下位レベルまで続けられるわけであるが、各レベルにおいて起動された確率オートマトンは自分が出した出力とそれに対する同じレベルの複数教師環境からの応答ならびに下位レベルの複数教師環境からの応答に基づいて出力確率ベクトルを変更

(注 1) 本論文の 4 章で与えられる例では、教師環境は複数である必要はない。また、教師環境からの応答も 0 (成功) と 1 (失敗) の 2 種類で十分である。しかしながら、迷路から得られる種々の情報をいかにうまく利用するかということが今後重要になってくるものと思われる。その意味においても、(より一般化された形の) 複数教師環境における階層構造確率オートマトンの学習の性能を調べておくことが重要であろう。本論文の 3 章では、応答が $[0, 1]$ 区間内の任意の値をとりうるような非定常複数教師環境中における階層構造確率オートマトンの学習の振舞いを考察している。

(注 2) $p_i(t)$ は $\bar{A}$ が t 時刻において i 番目の出力 α_i を出す確率を示している。

(注 3) $\beta_{j_1 i_1}(t, \omega)$ は失敗度と呼ばれ、0 と 1 の間の値をとる (ω は確率測度空間の点を表わす)。

する (次章の学習アルゴリズムを参照されたい)。ただし、起動されなかった確率オートマトンに関しては、出力確率ベクトルの変更は行われない。

さて、 t 時刻において s レベルまでの出力列が $\alpha_{j_1}, \alpha_{j_1 i_2}, \dots, \alpha_{j_1 i_2 \dots i_s}$ であったとして、この事実を示す表記法として $\phi(t) = \phi_{j_1 i_2 \dots i_s}$ なる式を導入する。更に、パス確率として、

$$\pi_{j_1 \dots i_s}(t) = p_{j_1}(t) p_{j_1 i_2}(t) \dots p_{j_1 i_2 \dots i_s}(t) \quad (1)$$

を定義する。上式は、第 1 レベルから $\alpha_{j_1}, \dots$ 、第 s レベルから $\alpha_{j_1 \dots i_s}$ なる出力が t 時刻に出される確率を示している (注 4)。

3. 非定常複数教師環境中における階層構造確率オートマトンの学習の振舞

つぎのような性質をもつ非定常複数教師環境 NMT を考える。

NMT: 以下の条件を満たすパス $\phi_{j_1^* j_2^* \dots j_s^*}$ がただ一つ存在する。

条件: あらゆる時刻 t 、あらゆる ω (注 5)、あらゆる $s(s=1, \dots, N)$ 、あらゆるパス (ただし、 $\phi(t) = \phi_{j_1 i_2 \dots i_s}$ をパスとしたとき、 $\max\{|j_1^* - i_1|, \dots, |j_s^* - i_s|\} > 0$ ($s=1, \dots, N$)) に対してつぎの (2) 式が成立する。

$$\begin{aligned} E\{\beta_{j_1^* j_2^* \dots j_s^*}(t, \omega) + \dots + \beta_{j_s^* i_1^* \dots i_s^*}(t, \omega) | \phi(t) \\ = \phi_{j_1^* \dots j_s^*}\} \\ < E\{\beta_{j_1^* i_1^* \dots i_s^*}(t, \omega) + \dots + \beta_{j_s^* i_1^* \dots i_s^*}(t, \omega) | \phi(t) \\ = \phi_{i_1^* \dots i_s^*}\} \end{aligned} \quad (2)$$

上の条件式 (2) は、階層構造確率オートマトンから $\alpha_{j_1^*}, \alpha_{j_1^* j_2^*}, \dots, \alpha_{j_1^* j_2^* \dots j_s^*}$ なる出力列が出されたとき、条件付期待値の意味で最も小さい失敗度の和を複数教師環境から受けることを意味している。したがって、階層構造確率オートマトンの学習の性能に関して、“絶対的良好” という以下に示されるような定義を与えることができる。

定義: 階層構造確率オートマトンは、つぎの条件を満たすとき、非定常複数教師環境 NMT において絶対的良好である。

条件: あらゆる時刻 t およびあらゆる出力確率ベ

(注 4) 本章では、階層構造確率オートマトンの学習メカニズムについて簡単な説明を行った。学習オートマトンの基本モデルについては、Lakshmivarahan¹⁰⁾ または筆者¹¹⁾ の本を参考にされたい。

(注 5) t 時刻までに階層構造確率オートマトンから出された出力ならびに t 時刻までの複数教師環境からの応答によって構成される確率測度空間を (Q, B, μ) としたとき、 ω は Q の点を表わす。

トル $\mathbf{P}(t)$ (注 6) に対してつぎの (3) 式が成立する。

$$\Delta \pi_{j_1^* \dots j_n^*}(t) \triangleq E\{\pi_{j_1^* \dots j_n^*}(t+1) - \pi_{j_1^* \dots j_n^*}(t) | \mathbf{P}(t)\} \\ > 0 \quad (3)$$

さて、以下では、Thathachar & Ramakrishnan⁸⁾によって提案された強化法を複数教師環境の場合に拡張することによって、NMT において絶対的的良好となる学習アルゴリズムが得られることを示そう。

学習アルゴリズム ANM: (注 7)

第 1 レベル: 第 1 レベルの確率オートマトン $\bar{A}$ から t 時刻において出力 α_{j_1} が出力され、複数教師環境からの返答が $\beta_{i_1}(t, \omega)$ ($i_1=1, \dots, r_1$) であったとする。

このとき、

$$p_{j_1}(t+1) = p_{j_1}(t) + L_1(t)(1 - p_{j_1}(t)) \\ p_{j_1}(t+1) = p_{j_1}(t)(1 - L_1(t)) \quad (i_1 \neq j_1; i_1=1, \dots, r) \quad (4)$$

第 s レベル ($s=2, \dots, N$): t 時刻において、確率オートマトン $\bar{A}, \dots, \bar{A}_{j_1, \dots, j_{s-1}}$ によって出力列 $\alpha_{j_1}, \alpha_{j_1, j_2}, \dots, \alpha_{j_1, \dots, j_s}$ が出力され、 s レベルの複数教師環境からの返答が $\beta_{i_1, \dots, i_s}(t, \omega)$ ($i_1=1, \dots, r_s$) であったとする。このとき、起動された確率オートマトンは、

$$p_{j_1, \dots, j_s}(t+1) = p_{j_1, \dots, j_s}(t) + L_s(t)(1 - p_{j_1, \dots, j_s}(t)) \\ p_{j_1, \dots, j_{s-1}, i_s}(t+1) = p_{j_1, \dots, j_{s-1}, i_s}(t)(1 - L_s(t)) \quad (i_s \neq j_s; i_s=1, \dots, r) \quad (5)$$

のように出力確率ベクトルを変更する。ただし、起動されなかった確率オートマトンの出力確率ベクトルの変更はまったく行わない。

ここで、 $L_i(t)$ ($i=1, \dots, N$) は以下のように構成する。

$$\bar{\eta}_s(t) \triangleq 1 - \frac{\beta_{i_1, \dots, i_s}(t, \omega) + \dots + \beta_{r_s, i_1, \dots, i_s}(t, \omega)}{r_s} \quad (s=1, \dots, N) \quad (6)$$

とおき、

$$L_1(t) = \frac{\sum_{s=1}^N \bar{\eta}_s(t) \lambda_s, \dots, L_n(t)}{\sum_{s=1}^N \bar{\eta}_s(t) \lambda_s} \\ = \frac{p_{j_1}(t+1) \dots p_{j_1, \dots, j_{n-1}}(t+1)}{\sum_{s=1}^N \bar{\eta}_s(t) \lambda_s} \quad (7)$$

とする。ただし、

$$0 < \sum_{s=1}^N \lambda_s < 1 \quad (8)$$

とする。

上で提案された学習アルゴリズム ANM は、(2) 式

(注 6) $\mathbf{P}(t)$ は、成分として、 t 時刻におけるすべてのレベルの出力確率を含むような確率ベクトルである。

(注 7) ANM は、Absolutely Expedient Learning Algorithm in the Nonstationary Multiteacher Environment の略記である。

を満たす非定常複数教師環境のもとで絶対的的良好性を保障する。

《定理》 学習アルゴリズム ANM をもつ階層構造確率オートマトンは、非定常複数教師環境 NMT において絶対的的良好である。

(証明) 少々長くなるので、ここでは省略し、付録にその略証を与える。

4. 未知環境中で動作する知能ロボット

1 章でも述べたように、人の監視 (制御) の及ばない領域でいかにロボットに知能的な振舞いをさせるかという問題は、原子力を始めとするさまざまな分野で最近特に重要なものとなってきている。しかしながら、この問題は簡単に解決できるものではなく、将来のセンサ技術の進歩と計算機を媒体とした人工知能関係の研究の発展に依存するところが大きい。センサ技術の進歩は、ロボットの知能化のために欠くことのできないものであることは誰の目にも明らかであるが、ではなぜ人工知能関係の研究の発展が切望されるのであろうか。それは、「各種センサなどを通じてロボットが周囲の (未知) 環境から情報 (知識) を得たとき、これらをどのように扱い処理するかそしてつぎの自分の行動にいかに関与していくか」ということがきわめて重要なまた難しい問題であるということによる (注 8)。この意味において、「学習」という言葉が“ロボットの知能化”を考えると、置きざりにすることのできない一つの重要な概念として浮びあがってくるのである。

筆者は、ロボットが各種センサなどを通じて周囲環境から得た情報 (知識) を以後の行動に役立てる際、確率オートマトンも一つの有力な方法になりうると考

Fig. 2 Robot manipulators going through the maze

(注 8) T. Oren²⁴⁾は、これを、“knowledge processing knowledge” という語句を用いて表現している。

えている。それは、確率オートマトンが学習的性能をもっており、その性能により、周囲の未知環境において次第に知的に振舞うようになることが可能だからである。

以下では、複雑な迷路を通過しようとしている移動ロボットに対して、階層構造確率オートマトンの学習アルゴリズムを利用することにより、ロボットが自動的に最も通過しやすい道を見出すことができるようになることを示す。

Fig. 2 は、迷路通過問題の 1 例を示している。ロボットマニピュレータ R_i

は“START”点を出発し“GOAL”点に至ることを目標としている。ただし、途中には 24 個のゲートがあり、これらはある未知確率(注 9)をもって閉じる($i=1, \dots$)。ロボット R_i は、もし無事迷路を通過できたなら“C.C.”(Control Center の略)に成功したという情報を送る。反対に、迷路内のあるゲートで通行不可能になった場合は、どのゲートでそうなったかを“C.C.”に知らせる。

“C.C.”に課せられているのは、“START”地点にいるつぎのロボット R_{i+1} にどのルートを通ればよいかを、いままでに得られた情報に基づいて指令することである。そのために、C.C.に階層構造確率オートマトンの学習アルゴリズムをもたせ、2.3 節で述べた階層構造確率オートマトンの学習的性能を利用することにより、最も通過しやすいルートを探索する。

階層構造確率オートマトンの適用と計算機シミュレーションの結果: Fig. 2において、ロボット・マニピュレータは右または下方のどちらかに進むことができるものと仮定する。この仮定のもとで、最適ルート探索を行うための階層構造確率オートマトンの構成は以下のようになされる。

まず、第 1 レベルには、2 出力をもつ確率オートマトンを配置し、A, B のどちらかのゲートを選ぶということに対応させる。第 2 レベルには、2 出力をもつ二つの確率オートマトンを配置し、それぞれ、C, D と E, F のゲートを選ぶということにその出力を対応させる。以上のようにしてつぎつぎに階層構造が作られていき、全体として 5 レベルをもつ階層構造確率オートマトンが構成される。

階層構造確率オートマトンが環境から受ける応答について説明しよう。迷路内のゲートの通過に成功した

Fig. 3 The other type of learning mechanism of the hierarchical structure stochastic automata

かどうかという情報だけを利用する場合には、各レベルにおける教師はただ一人ということになる。そして、無事“GOAL”までロボットが到着したなら、各レベルにおける教師環境からの応答をすべて 0 (成功) とする。また、あるゲートで通行不能になった場合は、それ以降のレベルにおける教師環境からの応答を 1 (失敗) とし、それまでの上位レベルにおける教師環境からの応答を 0 (成功) とする。つまり、この問題は、環境からの応答が 0 または 1 ($\beta_{i1}^{k1}, \dots, \beta_{ir}^{kr}(t, \omega) = 0$ or 1; P-model と呼ぶ。)で、かつ各レベルは単一教師環境をもつような特別の場合に相当する(ところで、どのゲートでロボットが通行不能になったかの情報をまったく利用できない場合には、Fig. 1 の形の階層構造確率オートマトンは利用できない。代わりに、教師環境からの応答が階層構造確率オートマトン全体に及んでくるような Fig. 3 の形のものを考える必要がある)。

われわれは、Fig. 2 の迷路の各ゲートが Table 1 で

Table 1 Closing probabilities of the gates in the maze

Gate	Closing probability	Gate	Closing probability
A	$0.31 + \eta_1(t, \omega)$	M	$0.1 \sin t + 0.8 + \xi_{11}(t, \omega)$
B	$0.75 + \eta_2(t, \omega)$	N	$0.1 \cos 5t + 0.65 + \xi_{12}(t, \omega)$
C	$0.1 \sin t + 0.7 + \xi_1(t, \omega)$	O	$0.03 \cos 3t + 0.67 + \xi_{13}(t, \omega)$
D	$0.1 \sin 5t + 0.2 + \xi_2(t, \omega)$	P	$0.02 \cos t + 0.75 + \xi_{14}(t, \omega)$
E	$0.3 \sin 3t + 0.8 + \xi_3(t, \omega)$	Q	$0.03 \cos t + 0.28 + \xi_{15}(t, \omega)$
F	$0.2 \sin 4t + 0.5 + \xi_4(t, \omega)$	R	$0.5 \cos 4t + 0.75 + \xi_{16}(t, \omega)$
G	$0.1 \cos 3t + 0.63 + \xi_5(t, \omega)$	S	$0.88 + \eta_3(t, \omega)$
H	$0.1 \sin 3t + 0.75 + \xi_6(t, \omega)$	T	$0.78 + \eta_4(t, \omega)$
I	$0.1 \sin t + 0.65 + \xi_7(t, \omega)$	U	$0.63 + \eta_5(t, \omega)$
J	$0.1 \cos t + 0.2 + \xi_8(t, \omega)$	V	$0.23 + \eta_6(t, \omega)$
K	$0.1 \sin t + 0.7 + \xi_9(t, \omega)$	W	$0.85 + \eta_7(t, \omega)$
L	$0.1 \sin t + 0.62 + \xi_{10}(t, \omega)$	X	$0.27 + \eta_8(t, \omega)$

(注 9) 時刻 t ならびにある測度空間の点の関数であると
する。

$|\xi_i(t, \omega)| \leq 0.05, i=1, \dots, 16$
 $|\eta_i(t, \omega)| \leq 0.07, i=1, \dots, 8$

Fig. 4 A graphic display on the microcomputer screen

Fig. 5 A microcomputer simulation result of the intelligent behavior of the robot manipulators going through the maze

示されるような確率で閉じるような場合のロボット・マニピュレータの迷路通過問題を考え、PC-9801 パソコンによるシミュレーションを行った。Fig. 4 は探索途中におけるパソコン画面の 1 例を示している。Table 1 より、ロボット・マニピュレータにとって最も通過しやすいルートは、 $A \rightarrow D \rightarrow J \rightarrow Q \rightarrow V \rightarrow X$ というゲートを通っていくものであることが明らかである。階層構造確率オートマトンの学習アルゴリズムとして、3 章で提案した“ANM”を用いたシミュレーション結果を Fig. 5 に示す (Fig. 5 は、 $A \rightarrow D \rightarrow J \rightarrow Q \rightarrow V \rightarrow X$ というルートを階層構造確率オートマトンが選ぶ出力確率の時間変化を示している)。

“C.C.”に学習アルゴリズム ANM をもつ階層構造確率オートマトンを配置しておけば、つぎつぎに“START”地点にやってくるロボット・マニピュレータに対して徐々に最適なルートを探し出させることが可能となることは、3 章における定理より明らかであるが(注 10)、Fig. 5 で示されるシミュレーション結果もこの理論を支持している。

本章では、階層構造確率オートマトンの学習的性能が移動ロボットの知能化に応用できうという一例を示した。将来、センサ技術や人工知能の研究などの発

(注 10) Table 1 で示される非定常環境は、3 章の条件式(2)を満たす。

展と結びつけることにより、またその他のいろいろな学習方式を加味することによって、より優れた知能ロボットが構成されるであろう。

5. おわりに

本論文では、未知の確率で開閉するゲートを多数有するような迷路をロボット・マニピュレータがつぎつぎと通過しようとしている問題を考察し、階層構造確率オートマトンの学習アルゴリズムを適用することにより、ロボットが徐々に最も通過しやすいルートを自動的に見出すことができるようになることを示した。

非定常教師環境として(2)式が満たされているというかなり制約の強い条件を付加したが、今後より一般的な非定常環境での階層構造確率オートマトンの学習的振舞いを研究する必要がある。また、ロボット・マニピュレータの種々のセンサ情報を取り入れたような学習の仕方を考え、より高度な知能ロボットの作成と取り組んでいかねばならない(そうすることによって、非定常複数教師環境中における階層構造確率オートマトンの学習的性能に関する研究も生きてくるものと思われる)。

最後になったが、日頃ご指導いただく京都大学名誉教授樫木義一先生、徳島大学学長添田喬教授、徳島大学正満敏雄教授に心からの謝意を表したい。また、計算機シミュレーションなどでご助力いただいた徳島大学工学研究科大学院生三宅隆浩君に深く感謝したい。

参考文献

- 1) M.L. Tsetlin: On Behavior of Finite Automata in Random Media, Automation and Remote Control, **22**, 1345/1354 (1961)
- 2) V. I. Varshavskii and I. P. Vorontsova: On the Behavior of Stochastic Automata with Variable Structure, Automation and Remote Control, **24**, 327/333 (1963)
- 3) M. F. Norman: On Linear Models with Two Absorbing Barriers, J. Math. Psych., **5**, 225/241 (1968)
- 4) S. Lakshmivarahan and M. A. L. Thathachar: Absolutely Expedient Learning Algorithms for Stochastic Automata, IEEE Trans. Systems, Man, and Cybernetics, **SMC-3-2**, 281/286 (1973)
- 5) K. S. Narendra and M. A. L. Thathachar: Learning Automata-A Survey, IEEE Trans. Systems, Man, and Cybernetics, **SMC-4-3**, 323/334 (1974)
- 6) N. Baba and Y. Sawaragi: On the Learning Behavior of Stochastic Automata under a Nonstationary Random Environment, IEEE Trans. Systems, Man, and Cybernetics, **SMC-5-2**, 273/275 (1975)
- 7) D. E. Koditschek and K. S. Narendra: Fixed Structure Automata in a Multi-Teacher Environment, IEEE Trans. Systems, Man, and Cybernetics, **SMC-**

- 7-6, 616/624 (1977)
- 8) M.A.L. Thathachar and K.R. Ramakrishnan: A Hierarchical System of Learning Automata, IEEE Trans. Systems, Man, and Cybernetics, **SMC-11-2**, 236/241 (1981)
- 9) N. Baba: The Absolutely Expedient Nonlinear Reinforcement Schemes under the Unknown Multi-Teacher Environment, IEEE Trans. Systems, Man, and Cybernetics, **SMC-13-1**, 100/108 (1983)
- 10) N. Baba: On the Learning Behaviors of Variable-Structure Stochastic Automaton in the General N-Teacher Environment, IEEE Trans. Systems, Man, and Cybernetics, **SMC-13-2**, 224/231 (1983)
- 11) N. Baba: An Application of the Hierarchical Structure Automata to the Cooperative Game with Incomplete Information, International J. of Sys. Sci., **17**, 121/127 (1986)
- 12) J.L. Doob: Stochastic Processes, Wiley (1953)
- 13) M.F. Norman: Markov Processes and Learning Models, Academic Press (1972)
- 14) S. Lakshmivarahan: Learning Algorithms Theory and Applications, Springer-Verlag (1981)
- 15) K. R. Ramakrishnan: Hierarchical Systems and Cooperative Games of Learning Automata, PH.D. Indian Institute of Science, Bangalore, India (1982)
- 16) N. Baba: New Topics in Learning Automata Theory and Applications, Springer-Verlag (1985)
- 17) S.M. Udupa: Collision Detection and Avoidance in Computer Controlled Manipulators, 5th Int. Joint Conf. on Artificial Intelligence, 737/748 (1977)
- 18) T. Lozano-Pérez and M.A. Wesley: An Algorithm for Planning Collision-Free Paths among Polyhedral Obstacles, Commun. ACM, **22-10**, 560/570 (1979)
- 19) R.A. Brooks: Solving the Find-Path Problem by Good Representation of Free Space, IEEE Trans. Systems, Man, and Cybernetics, **SMC-13-3**, 190/197 (1983)
- 20) 尾崎弘明: マニピュレータの障害物回避, 日本ロボット学会誌, **2-6**, 76/82 (1984)
- 21) S. Tangwongsan and K.S. Fu: An Application of Learning to Robotic Planning, Tutorial on Robotics (C.S.G. Lee et al., Editors), IEEE Computer Society Press, 471/501 (1983)
- 22) R.E. Fikes, P.E. Hart and N.J. Nilsson: Some New Directions in Robot Problem Solving, Machine Intelligence, **3**, 405/430 (1972)
- 23) R.P. Paul: Robot Manipulators, The MIT Press (1981)
(吉川訳, ロボットマニピュレータ, コロナ社 (1984))
- 24) T.I. Oren: Cognizant Simulation Systems, 4th Int. Symp., Knowledge-Based Simulation and Medelling, Tucson, Arizona, U.S.A. (1987)

《付 録》

3章の定理の略証

つぎのような関数 $h_s(t)$ ($s=1, \dots, N$) を導入する¹⁵⁾.

$$h_1(t) \triangleq \delta_{q_1, j_1} \sum_{n=1}^N \bar{\eta}_n(t) \lambda_n - p_{q_1}(t) \sum_{n=1}^N \bar{\eta}_n(t) \lambda_n \quad (9)$$

$$h_s(t) \triangleq \delta_{q_1, j_1, \dots, q_s, j_s} \sum_{n=s}^N \bar{\eta}_n(t) \lambda_n$$

$$-p_{q_1, \dots, q_s}(t) \delta_{q_1, j_1, \dots, q_{s-1}, j_{s-1}} \sum_{n=s}^N \bar{\eta}_n(t) \lambda_n \quad (s=2, 3, \dots, N) \quad (10)$$

ここに,

$$\delta_{q_1, j_1, \dots, q_n, j_n} = \begin{cases} 1 & (q_i = j_i, i=1, 2, \dots, n) \\ 0 & (q_i \neq j_i \text{ (ある } i \text{ に対して)}) \end{cases} \quad (11)$$

$h_s(t)$ ($s=1, \dots, N$) を用いることにより, 3章で提案した学習アルゴリズム ANM はつぎのように書ける.

第1レベル:

$$p_{q_1}(t+1) = p_{q_1}(t) + h_1(t) \quad (12)$$

第 s レベル ($s=2, \dots, N$):

$$p_{q_1, \dots, q_s}(t+1) = p_{q_1, \dots, q_s}(t) + \frac{h_s(t)}{p_{q_1}(t+1)p_{q_1, q_2}(t+1) \dots p_{q_1, \dots, q_{s-1}}(t+1)} \quad (13)$$

(12), (13) を (1) に代入することにより,

$$\pi_{q_1, \dots, q_N}(t+1) = \pi_{q_1, \dots, q_N}(t) + \pi_{q_1, \dots, q_N}(t) \sum_{k=1}^N \frac{h_k(t)}{\pi_{q_1, \dots, q_k}(t)} \quad (14)$$

が得られる. さて,

$$\Delta \pi_{q_1, \dots, q_N}(t) \triangleq E \{ \pi_{q_1, \dots, q_N}(t+1) - \pi_{q_1, \dots, q_N}(t) | \mathbf{P}(t) \} \quad (15)$$

とおこう. すると,

$$\Delta \pi_{q_1, \dots, q_N}(t) = \pi_{q_1, \dots, q_N}(t) \sum_{k=1}^N \frac{E \{ h_k(t) | \mathbf{P}(t) \}}{\pi_{q_1, \dots, q_k}(t)} \quad (16)$$

となる.

$h_k(t)$ が λ_n ($n=1, \dots, N$) に関して線形であるから ((9), (10)), (16) 式の右辺第2項はつぎのように書かれる.

$$\sum_{k=1}^N \frac{E \{ h_k(t) | \mathbf{P}(t) \}}{\pi_{q_1, \dots, q_k}(t)} = \sum_{l=1}^N a_l \lambda_l \quad (17)$$

(9), (10) から明らかのように, $h_k(t)$ が含む λ_l は $l \geq k$ に対してである. したがって, a_l はつぎの(18)式の λ_l の係数に等しいことがわかる.

$$\begin{aligned} & \sum_{k=1}^l \frac{E \{ h_k(t) | \mathbf{P}(t) \}}{\pi_{q_1, \dots, q_k}(t)} \\ &= E \left\{ \frac{h_1(t)}{\pi_{q_1}(t)} + \frac{h_2(t)}{\pi_{q_1, q_2}(t)} + \dots + \frac{h_l(t)}{\pi_{q_1, \dots, q_l}(t)} \middle| \mathbf{P}(t) \right\} \end{aligned} \quad (18)$$

Ramakrishnan¹⁵⁾ と同様の方法を用いることにより,

a_l は

$$a_l = E \left\{ \frac{\delta_{q_1, j_1, \dots, q_l, j_l} \bar{\eta}_l(t)}{\pi_{q_1, \dots, q_l}(t)} \middle| \mathbf{P}(t) \right\} - E \{ \bar{\eta}_l(t) | \mathbf{P}(t) \} \quad (19)$$

と変形できる.

さて, つぎのような関数を考えよう.

$$\bar{d}_{j_1 \dots j_l}(t) \triangleq E\{\bar{\eta}_l(t) | \phi(t) = \phi_{j_1 \dots j_l}\} \quad (20)$$

ここで $\phi(t) = \phi_{j_1 \dots j_l}$ は階層構造確率オートマトンが時刻 t において (第 1 レベルから順に) $\alpha_{j_1}, \alpha_{j_1 j_2}, \dots, \alpha_{j_1 \dots j_l}$ なる出力を出すことを意味する。

明らかに,

$$\begin{aligned} E\{\delta_{q_1 j_1 \dots q_l j_l} \bar{\eta}_l(t) | \mathbf{P}(t)\} \\ = \sum_{j_1 \dots j_l} \pi_{j_1 \dots j_l}(t) E\{\delta_{q_1 j_1 \dots q_l j_l} \bar{\eta}_l(t) | \phi(t) = \phi_{j_1 \dots j_l}\} \\ = \pi_{q_1 \dots q_l}(t) E\{\bar{\eta}_l(t) | \phi(t) = \phi_{q_1 \dots q_l}\} \\ = \pi_{q_1 \dots q_l}(t) \bar{d}_{q_1 \dots q_l}(t) \end{aligned} \quad (21)$$

したがって,

$$E\left\{\frac{\delta_{q_1 j_1 \dots q_l j_l} \bar{\eta}_l(t)}{\pi_{q_1 \dots q_l}(t)} \middle| \mathbf{P}(t)\right\} = \bar{d}_{q_1 \dots q_l}(t) \quad (22)$$

つぎに, (19) 式の第 2 項はつぎのように変形できることがわかる。

$$\begin{aligned} E\{\bar{\eta}_l(t) | \mathbf{P}(t)\} &= \sum_{j_1} \dots \sum_{j_l} \pi_{j_1 \dots j_l}(t) E\{\bar{\eta}_l(t) | \phi(t) \\ &= \phi_{j_1 \dots j_l}\} \\ &= \sum_{j_1} \dots \sum_{j_l} \pi_{j_1 \dots j_l}(t) \bar{d}_{j_1 \dots j_l}(t) \\ &= \sum_{j_1} \dots \sum_{j_N} \pi_{j_1 \dots j_N}(t) \bar{d}_{j_1 \dots j_l}(t) \end{aligned} \quad (23)$$

(19), (22), (23) より,

$$a_l = \sum_{j_1} \dots \sum_{j_N} \pi_{j_1 \dots j_N}(t) [\bar{d}_{q_1 \dots q_l}(t) - \bar{d}_{j_1 \dots j_l}(t)] \quad (24)$$

したがって (16), (17), (24) より,

$$\begin{aligned} \Delta \pi_{q_1 \dots q_N}(t) &= \pi_{q_1 \dots q_N}(t) \sum_{j_1} \dots \sum_{j_N} \pi_{j_1 \dots j_N}(t) \\ &\quad \times \left(\sum_{l=1}^N \lambda_l \right) [\bar{D}_{q_1 \dots q_N}(t) - \bar{D}_{j_1 \dots j_N}(t)] \end{aligned} \quad (25)$$

ただし,

$$\bar{D}_{q_1 \dots q_N}(t) = \left(\sum_{l=1}^N \lambda_l \bar{d}_{q_1 \dots q_l}(t) \right) / \left(\sum_{l=1}^N \lambda_l \right) \quad (26)$$

である。

(2), (6), (20) より, あらゆる時刻 t および $l(l=1, \dots, N)$ に対してつぎの (27) 式が成立する。

$$\bar{d}_{j_1^* \dots j_l^*}(t) > \bar{d}_{j_1 \dots j_l}(t) \quad (27)$$

したがって, あらゆる時刻 t に対して,

$$\bar{D}_{j_1^* \dots j_N^*}(t) > \bar{D}_{j_1 \dots j_N}(t) \quad (28)$$

ゆえに, (25) より,

$$\Delta \pi_{j_1^* \dots j_N^*}(t) > 0$$

Q. E. D.