

氏 名	馬 場 則 夫
学位の種類	工 学 博 士
学位記番号	工 博 第 396 号
学位授与の日付	昭 和 50 年 1 月 23 日
学位授与の要件	学 位 規 則 第 5 条 第 1 項 該 当
研究科・専攻	工 学 研 究 科 数 理 工 学 専 攻
学位論文題目	STUDIES ON THE LEARNING BEHAVIORS OF VARIABLE-STRUCTURE STOCHASTIC AUTOMATA (内部構造可変確率オートマトンの学習的性能についての研究)

論文調査委員 (主 査) 教授 榎木義一 教授 得丸英勝 教授 明石 一 教授 三根 久

論 文 内 容 の 要 旨

本論文は未知ランダム媒体のもとで動作する可変構造確率オートマトンの学習的な働きに関する基礎的な研究であって、7章より構成されている。

第1章は序論であって、ランダム媒体のもとで動作する確率オートマトンの学習的な働きに関する研究の意義およびその歴史的背景についてふれるとともに、本論文の概要と構成について述べている。

第2章では、定常ランダム媒体のもとでの線形 Lr_{-1} 強化法の学習的な性能を調べ、時間の経過とともに確率オートマトンが最適な状態に近づくかどうかについて理論的な考察を行い、あらゆる定常ランダム媒体の下で ϵ -最適性が Lr_{-1} 強化法によって保障されるということを証明している。

第3章では、失敗の時にも状態確率を変更する非線形 TNP 強化法を新しく提案し、 Lr_{-1} 強化法と同様、あらゆる定常ランダム媒体のもとで ϵ -最適性が保障されるということを証明している。この章において提案されている TNP 強化法は失敗のときにも状態確率を変更し、さらにあらゆる定常ランダム媒体のもとで ϵ -最適性を保障するので、学習速度および収束性の両面において優れている。

第4章では、学習の最終段階における確率オートマトンの振舞いについて論じている。すなわち、Kushner の確率安定論をこの問題に適用することにより、十分学習が進んだ段階から確率オートマトンが望ましくない状態へ逆もどりする確率の上界が与えられることを示している。 Lr_{-1} 強化法および TNP 強化法の両方について具体例が与えられ、両者ともこの確率は非常に小さくなることが示されている。

第5章では、ある非定常ランダム媒体のもとでの確率オートマトンの学習的な振舞いを議論している。特に確率オートマトンからの出力に対するランダム媒体からの失敗の返答の確率が時間だけではなくある確率測度空間の点の関数でもあるような非定常ランダム媒体について考察を行い、この媒体のもとで Lr_{-1} 強化法および TNP 強化法の両者によって ϵ -最適性が保障されることを証明している。

第6章では、確率オートマトンによる最適パラメータ探索問題について論じている。最適パラメータ探索問題に確率オートマトンを適用した研究はこれまでも発表されているが、数学的な議論が困難な故

に、収束の保障が与えられていないという問題があった。本論文では、この問題を理論的に考察することにより、5章で論じられた非定常ランダム媒体のもとでの確率オートマトンの学習的振舞いの問題に帰着されることを示している。

第7章では、確率オートマトンを用いたゲームについて論じている。まずゲームに参加しているプレイヤーが協定を行わない場合について考察し、学習的な性能をもつ確率オートマトンをこのゲームに利用することによって、そのプレイヤーがよりよい利益をあげることが指摘されている。さらに、そのゲームに参加している全プレイヤーが協定を行って利益をあげようとする場合、一つの確率オートマトンからの指令によって全プレイヤーが行動するような方法が提案されている。そしてその場合の確率オートマトンの学習的な働きが第5章で考察された非定常ランダム媒体のもとにおける確率オートマトンの学習的な振舞いの問題に帰着されることが示される。

論文審査の結果の要旨

本論文は、未知ランダム媒体のもとで働らく媒変構造確率オートマトンの学習的な振舞いを理論的に考察している。

この方面の研究は、1963年に Varshavskii 及び Vorontsova によって始められ、現在まで理論的な考察ならびに実的な問題への適用が極めて活発に行われてきた。しかしながら、まだこの方面の研究は十分に進んでいるとはいえず、確率オートマトンが最適な状態に近づくかどうかという問題に対してさえ十分な解答が与えられてはいない現状である。

著者はこのような問題に対し、最適な状態をとる確率が時間と共にどのように変わるかを確率過程の問題としてとらえ、その収束性を議論し、さらにランダム媒体のもとで働らく確率オートマトンの具体例をあげ、それについて考察を行なったものであって、その研究成果は次のとおりである。

(1) 線形 Lr_{-1} 強化法を持つ確率オートマトンにより、あらゆる定常ランダム媒体のもとで ϵ -最適性すなわち強化法に関するパラメータが十分小さいとき時間が十分経過すると最適な状態をとる確率が1に近づくということが保障されることを証明している。

(2) この Lr_{-1} 強化法はランダム媒体からの返答が失敗の時には状態確率を変更しないので学習速度の点であまり優れているとはいえない。本論文では、失敗の時にも状態確率を変更する非線形 TNP 強化法を提案し、この強化法によりあらゆる定常ランダム媒体のもとで ϵ -最適性が保障されることを証明している。

(3) 学習の最終段階における確率オートマトンの学習的振舞いを考察し、十分学習が進んだ状態から望ましくない状態へ逆もどりする確率の上界を与えている。

(4) 未知ランダム媒体からの失敗の返答の確率が時間だけではなく、ある確率測度空間の点の関数でもあるような非定常ランダム媒体を提案している。そして、その媒体のもとでの Lr_{-1} 強化法及び TNP 強化法の学習的な性能を議論し、両者により ϵ -最適性が保障されることを証明している。

(5) 未知最適パラメータ探索問題やゲームの問題といった実際的な問題に確率オートマトンを適用した場合の確率オートマトンの学習的な振舞いについて理論的な考察を加え、本論文で考察した理論的成果

が果す実際の効果について興味ある幾つかの結果を得ている。

以上、要するに、本論文は、確率オートマトンの学習効果に対して理論的な保障を与えたものであり、学術上、工業上寄与するところが少なくない。よって、本論文は工学博士の学位論文として価値あるものと認める。