

知識情報工学専攻	学籍番号	991081
申請者氏名	野口 拓也	

指導教員氏名	村越 一支
--------	-------

論文要旨(修士)

論文題目	強化信号の増減により効率的な学習を行う強化学習のパラメータ制御法
------	----------------------------------

強化学習(以下、RL;Reinforcement Learning)とは、報酬や罰という結果の善し悪しを表す強化信号を手掛かりとして、環境に適応していく学習である。人工ニューラルネットワークで学習されているような、教師あり学習とは異なり、RLでは明示的な教師を持たない。そのため、学習に時間はかかるが、試行錯誤による環境に適応した柔軟な学習を自律的に行うことができる。

RLの手法の一つとして、報酬と罰を分けた二次元評価強化学習が岡田ら(2001)によって提案されている。問題点として、環境変化によって、再学習が不得手である場合が見られたので二次元評価強化学習にMurakoshi and Mizuno(2004)の学習パラメータ制御法を二次元に拡張し導入したが、それだけでは、報酬と罰が確率的に与えられる動物実験のシミュレーションで結果を再現することができなかった。さらに、Murakoshi and Noguchi(2005)は強化信号の出現確率を考慮したアルゴリズムを提案し、前述のシミュレーションにおいて円滑な再学習を行い、再現することができた。

しかし、Murakoshi and Noguchi(2005)のアルゴリズムでは、減報酬、増罰といった負の環境変化しか、考えられてない。実際に起こり得る環境変化はそれだけではなく、増報酬、減罰といった正の環境変化も考えられる。本研究では、正負の環境変化を考慮した上で学習方策や行動を決定するように改良を行い、提案アルゴリズムの有効性を検証するために、工学的な応用を想定した問題を考えた。

取扱う問題は変量生産を想定した単一品種多工程ライン業務の自動制御とし、単一品種を生産する多工程生産ラインを対象にした変量(需要の変化)に自動で追従することを目指す。変量生産において、考えられる問題点は生産速度(単位時間当たりの生産量)と精度(良品の生産確率)には負の相関があり、状況(要求量)によって最適な生産は異なる。変量毎に計画を立てることは煩わしく、手間であると言える。提案アルゴリズムによって、問題解決を目指し有効性を検証する。

強化学習の他手法と提案アルゴリズムを用いて、上述の問題解決に取り組んだ。需要が増加する場合において速度重視の生産にシフトするように再学習し、規定時間内に十分に要求量を満たし、性能値が最も大きい手法は提案アルゴリズムであった。また、需要が減少する場合にはあまり差がみられなかったが、罰つまりコストの合計が最も少なく、性能値が最も大きい手法は提案アルゴリズムであった。これらの結果から、時々刻々と変化する需要の変動に対して、提案アルゴリズムを用いることで追従できたといえ、提案アルゴリズムの有効性を主張する。