

[別紙2]

審査の結果の要旨

論文提出者氏名 小林 祐一

[別紙2]

審査の結果の要旨

論文提出者氏名 小林 祐一

小林 祐一（こばやし ゆういち）提出の本論文は「強化学習のための自律分散型関数近似法」と題し、全6章よりなり、強化学習において問題に適した分解能の自律的な獲得の可能な、自律分散型のアルゴリズムに基づく関数近似法を提案している。

第1章では、研究の背景を説明し、研究の目的と論文の構成を述べている。強化学習の核をなす価値関数の近似において、学習にとって重要な領域のみに高い分解能を持つような表現、すなわち適応的な分解能を獲得する関数近似手法が重要な意味を持っている。適応的な分解能を自律的に獲得するための関連研究として、関数近似手法における逐次的な基底関数の追加方法や状態空間の自律的な分割手法などをあげることができるが、人手による調整が必要な設計要素が多いという問題が存在する。それに対し、本研究では自律的な情報表現獲得を目指し、設計自由度を低減した適応的分解能獲得型の関数近似手法を提案する。系(関数近似器)全体の秩序との関係を記述した形での局所的な関数近似要素の挙動を設計する手法としてグラフ上の反応拡散方程式を用いた自律分散型の関数近似法を提案した。

第2章では、強化学習の基本的なアルゴリズムを説明し、強化学習の核である価値関数近似への要求とそれに基づく関連研究の分類を行い、本研究の強化学習研究における位置付けを述べている。非定常関数近似、必要データの制約などの要求から、分解能を獲得する価値関数近似は既存の関数近似手法では容易ではない。特に、ガウス基底関数の重ねあわせで連続値関数を表現する代表的関連研究では、基底追加判別のためのしきい値や基底形状などに関して試行錯誤による設計を要するという問題が存在する。これに対し、本研究の提案関数近似法は、近似誤差でなく近似関数の形状に着目した統一的尺度（複雑度）を用いた関数近似要素（ノード）の再配置を行うため、設計自由度を低減することができる。また、自律分散型の手法であることから、ノードの動的な追加や削除が容易であるという利点も有している。

第3章では、本研究で提案する関数近似法のアルゴリズムを述べている。ノードによる超平面の組み合わせによって関数近似面を構成し、勾配変化の大きい領域にノードを密に、小さい領域に疎に分布するように挙動を設計する。そのための設計方法としてグラフ上の反応拡散方程式を用いた。各ノードの局所近傍における近似関数の複雑度を勾配変化に基づいて定義する。複雑度をノード上関数とし、これを均一にするようなノードの挙動をグラフ上の反応拡散方程

式に基づいて設計した。ノードの挙動は、系全体の状態を表すポテンシャル汎関数を極小値に導く勾配法によって設計される。ノード間のアークを構成する方法として、ユークリッドノルムに基づいた近傍を構成する TRN(Topology Representing Networks)アルゴリズムを用いた。

第4章では、提案関数近似手法の強化学習への適用法および適用時の適性と問題について述べている。提案近似手法の強化学習への適用方法として、Value Gradient 法および Q-learning 法の状態価値関数および行動価値関数の近似を行う方法を示した。Value Gradient 法では、状態価値関数を TD 誤差学習により推定し、行動決定は近似した状態価値関数の勾配情報を用いて行う。ノードの逐次的追加には、複雑度を用いる。複雑度最大のノードの近傍に新規ノードを生成させ、提案アルゴリズムにしたがって移動を行う。

第5章では、提案する関数近似手法の性質の検証および強化学習例題を用いた性能向上の確認を行っている。定常関数近似問題においては、ノード移動を適応的に行うことにより近似誤差を低減できることを確認した。また、複雑度を表す関数値の大きい領域に逐次的にノードを追加していく方法が複雑な形状の関数を近似する上で有効に機能することを確認した。強化学習への適用の評価に関しては、入力空間に格子状にノードを固定したままの場合と、ノード移動や逐次的なノード追加を行った場合との比較を行った。強化学習例題には水たまり問題と 1 自由度振子振り上げ問題を用い、各場合において、ノードの移動および逐次的追加による学習効率の変化を検証した。その結果、ノード移動により、学習効率が改善できることを示した。この改善の割合は水たまり問題よりも問題設定を変更した 2 次元環境探索問題においてより顕著に表れた。このことから、ノード移動による適応的な分解能の変更が強化学習において有効であること、および勾配変化の表現効率が報酬値に影響を与えやすい問題ほどノード移動の効果が高くなることが確認された。また、ノードの逐次追加によって、学習効率が改善できることを示した。逐次追加を行う場所の判別に複雑度を用いることで、効率の良いノードの追加が可能であることが確認された。

第6章では、結論を述べている。本論文において提案した適応的な分解能獲得型の関数近似法が、適切な関数近似によって強化学習における学習性能を向上させられることが、強化学習の標準問題を用いたシミュレーション実験によって示された。

以上を要約するに、本研究により、複雑な領域に密に分布するようにノードを移動させるという考え方が自律分散型のアルゴリズムに基づいて提案され、強化学習における情報表現方法として重要な意味を持つ分解能適応型関数近似法が確立された。このことにより、精密機械工学のみならず工学全体の発展に寄与するところが大きい。

よって本論文は博士（工学）学位請求論文として合格と認められる。