



Title	Deep Reinforcement Learning Based Optimal Control of Nonlinear Systems
Author(s)	池本, 隼也
Citation	大阪大学, 2023, 博士論文
Version Type	VoR
URL	<a href="https://doi.org/10.18910/92211">https://doi.org/10.18910/92211</a>
rights	
Note	

***Osaka University Knowledge Archive : OUKA***

<https://ir.library.osaka-u.ac.jp/>

Osaka University

## 論文内容の要旨

氏 名 ( 池 本 隼 也 )	
論文題名	Deep Reinforcement Learning Based Optimal Control of Nonlinear Systems (深層強化学習に基づく非線形システムの最適制御)
論文内容の要旨	
<p>深層強化学習は、深層ニューラルネットワークを用いた強化学習法であり、従来の強化学習では対処できなかった複雑な問題に対する方策の学習を可能にした。システムのモデルを使わず方策を設計できるため、深層強化学習はシステム制御への応用も期待されている。その一方で、深層強化学習は非常に多くの相互作用を必要とし、加えて、問題に適した環境の状態や報酬を事前に定めておく必要があるため、その応用は制限されている。深層強化学習の応用範囲を拡大するため、本論文では、4つの制御問題に対する学習法を提案する。</p> <p>まず、深層強化学習には大量の学習データが必要であるが、実システムから取得するデータ量の削減法について考える。このような場合、シミュレータ内の仮想システムを利用する学習法が有用である。一方で、実システムと仮想システムのパラメータに誤差があった場合、適切な方策を学習できない可能性がある。そこで、これらのシステムのパラメータ誤差を考慮した2段階の学習法を提案する。</p> <p>次に、ネットワーク化制御によるシステムの安定化を考える。ネットワーク化制御では、制御器とシステムの情報のやり取りに通信遅延が生じる。そこで、制御入力受信時の状態変化を予測するために十分な情報を含んだ環境の状態空間を定義し、深層強化学習に基づく方策の設計法を提案する。加えて、センサがシステムの一部の情報を観測できない場合の設計法も提案する。</p> <p>次に、Signal Temporal Logic (STL) 式で記述された時間的な制御仕様が与えられたネットワーク化制御を考える。STL式で記述された仕様を満たすための行動決定にはシステムの過去の状態列が必要である。加えて、通信遅延の影響を考慮する必要がある。そこで、仕様の充足と通信遅延の両方を考慮可能な環境の状態空間を定義し、深層強化学習に基づく方策の設計法を提案する。</p> <p>最後に、STL 式によって記述された制御仕様の充足を制約とした上で、与えられた指標に関して最適な方策を設計することを考える。この制御問題を制約付き強化学習問題とみなして、制約付きマルコフ決定過程を使って定式化し、ラグランジュ緩和を用いた深層強化学習に基づく方策の設計法を提案する。</p>	

## 論文審査の結果の要旨及び担当者

氏 名 ( 池 本 隼 也 )		
	(職)	氏 名
論文審査担当者	主 査	教 授 潮 俊 光
	副 査	教 授 乾 口 雅 弘
	副 査	教 授 飯 國 洋 二

## 論文審査の結果の要旨

本論文では、深層強化学習に基づいた、モデルフリーな非線形システムの最適制御法を提案している。深層強化学習をシステム制御へ応用する際に問題となる、学習効率の悪さへの対処方法、適切な環境の状態表現や報酬関数の設定方法について検討している。第1章の緒論と第6章の結論を含め、以下の各章から構成されている。

第2章では、強化学習の基本的な枠組みを述べ、本論文で用いられる、Q学習から派生したいくつかの深層強化学習アルゴリズムについて知見をまとめている。

第3章では、実システムから取得するデータ量を削減するため、シミュレータを用いた深層強化学習アルゴリズムを考えている。シミュレータ内でモデル化された仮想システムと実システムのパラメータに誤差がある場合に対処するため、2段階の学習アルゴリズムを提案している。

第4章では、深層強化学習をネットワーク化制御によるシステムの安定化に応用している。通信遅延の影響を考慮した上で制御入力を決定するために十分な情報を含んだ状態表現を定義し、深層強化学習によって制御方策を学習する方法を提案している。さらに、システムに関する一部の情報が観測できない場合に拡張している。

第5章では、Signal Temporal Logic (STL) 式で記述された制御仕様が与えられたネットワーク化制御に深層強化学習を応用している。各時刻で制御仕様の充足状況と通信遅延の両方を考慮して制御入力を決定できるような状態表現を提案し、深層強化学習によって制御方策を学習する方法を提案している。

第6章では、STL 式によって記述された制約条件の下での最適制御を考えている。制約付きマルコフ決定過程を用いて問題を定式化し、ラグランジュ緩和を用いた深層強化学習によって方策の設計をおこなっている。また、学習初期において、制約を満たす経験を得やすくするため、2段階の学習アルゴリズムを提案している。

以上のように本論文では、4つの深層強化学習に基づいたモデルフリーな最適制御法を提案し、深層強化学習のシステム制御への応用の拡大に貢献した。よって、博士（工学）の学位論文として価値あるものと認める。