

学位論文審査の結果及び試験，試問の結果報告書

(論文博士用)

山口大学大学院理工学研究科

報告番号	理工博乙 第 138 号	氏名	内山 祥吾
最終試験担当者	主査委員	大林 正直	江 鐘偉
	審査委員	田中 幹也	浜本 義彦
	審査委員	石川 昌明	
<p>【論文題目】 強化学習と小脳モデルを利用した適応型制御システムの設計法 (A Design Method for An Adaptive Control System Based on Reinforcement Learning and Cerebellar Model)</p>			
<p>【論文審査の結果及び最終試験の結果】 現代制御理論は制御対象の数式モデルを利用する設計を前提とし，システムの安定性を保証した制御方式に関するものである。しかし，現代社会に存在する多くのシステムは非線形システムであり，その時間的変化は通常，非線形の微分方程式で表わすことになる。その変化は複雑で数式(微分方程式)での表現(モデル化)が困難となる場合も多く，そのようなシステムを制御するためには，制御対象の数式モデルを必要とせずに制御器の設計が可能なモデルフリーな制御系設計法が必要となる。その一つが小脳モデル等のニューラルネットワーク(NN)やファジィ等を用いた計算手法，ソフトコンピューティング，を利用した知的制御系設計法である。ソフトコンピューティングを利用した知的適応制御は，目標を達成するように制御パラメータをオンラインで調整していく制御方式であり，「対象システムに関する事前知識を必要としない」ことが本方式の大きな強みである。但し，制御系の安定性は保証されない。一方，人間の脳機能に関する脳科学に関する研究も進展し，特に脳の学習機能を模倣した強化学習を制御システムに展開する研究も盛んに行われてきている。強化学習も制御対象との相互作用による学習の過程で制御対象の性質を獲得する学習方式であり，制御対象の数式モデルを必要とせず，ソフトコンピューティングに含まれると言える。その学習の性質上，強化学習を利用して設計された制御系も安定性は保証されない。 申請者は，本論文で，制御対象のモデル化が困難なシステムを対象として，現代制御理論とソフトコンピューティングの両者の利点を取り込み，安定性を保証し，より高い制御性能を持つ制御系の実現を目指した設計法の開発過程をまとめている。以下，各章を要約する。 第1章では，本論文の背景と目的について述べている。 第2章では，ソフトコンピューティングと現代制御理論の両者の利点を取り込み，強化学習制御とロバスト性が強い適応H_∞制御の協働により性質の異なる状態を同時に制御する制御系設計法を提案している。本方法により，通常では制御が困難な制御対象を容易に制御可能とし，頭上走行クレーンを制御対象システムとした計算機シミュレーションにより，提案法の有効性を示している。 オフライン学習による設計法は2章のみで，3章以降は全てオンライン学習による制御の開発となっている。 第3章では，第2章での設計方法がオフライン学習を利用しているのに対し，システムの不確定性に頑健なロバスト制御と強化学習をオンライン的に利用した，両者の融合型である，オンライン強化学習制御システム「リアルタイム強化学習制御システム(Real-time Reinforcement Learning Control System, RRLCS)」を提案し，加えて，リアプノフ関数による安定性解析により，システムの安定性を証明している。提案システムの有効性を台車付き倒立振子の制御対象とした計算機シミュレーションにて示している。 第4章では，第3章で提案したRRLCSが持つ確率推論による性能の不安定さを解決するため，小脳モ</p>			

デルを利用した制御系設計法について提案している。小脳モデルに関する先行研究として、Albus が提案した CMAC(Cerebellar Model Articulation Controller)と呼ばれる小脳パーセプトロンモデルがある。本論文では CMAC の短所を解消する、新たな小脳パーセプトロン改良モデル(Cerebellar Perceptron improved model, CP)を提案し、それを基にした「小脳パーセプトロン改良モデル利用型ロバスト制御システム(Cerebellar Perceptron Robust Control System, CPRCS)」を提案している。リアプノフ関数による安定性解析により、システムの安定性を証明し、台車付き倒立振り子による計算機シミュレーションにより、本提案法が 3 章で提案した RRLCS より高い追従性能を持つことを示している。

第 5 章では、フィードバック制御である CPRCS を、応答が高速であるフィードフォワード制御に拡張するために CP を改良し、改良型 CP を基に構築した「自己融合小脳パーセプトロン改良モデル利用型ロバスト制御システム(Auto-Fusion Cerebellar Perceptron Robust Control System, AFCPRCS)を提案している。フィードバック誤差学習による学習則により最適な CP を構築するシステムであり、AFCPRCS が第 4 章で提案した CPRCS よりさらに高い追従性能を持つことを、台車付き倒立振り子を制御対象とした計算機シミュレーションにより示している。

第 6 章では、第 5 章で提案した、申請者が提案した制御系設計法である、AFCPRCS をマルチエージェントシステムの合意問題に適用可能なように制御システムを拡張している。多入力多出力システムで構成される 6 体のエージェントによる合意問題の計算機シミュレーションにより、高い合意性能と CP の汎化性能を示している。

第 7 章では、結論であり本論文を総括し、今後の展望を述べている。

公聴会には学内外から多数の参加があり、活発な質疑応答がなされた。その主な内容として、

1. 小脳パーセプトロンの基本的な考え方について。
2. フォーマーションの合意問題において、提案法と従来法の基本的なメカニズムの違いについて。
3. 古典的な PID のみの制御系との比較について。
4. 制御系設計における、より注意すべき重要な部分について。

等の質問があり、いずれの質問に対しても申請者からの的確な回答がなされた。

以上より、本研究は、独創性、信頼性、有効性、実用性ともに優れており、博士(工学)の論文に十分に値するものと判断した。

試験および諮問として、以下の内容の課題が課された。

1. 小脳演算モデルと実際の小脳との対応関係について。
2. 従来法であるロバスト強化学習の詳細について

本諮問に対し、丁寧なレポート形式での回答がなされた。語学力については申請者を第 1 著者とする英論文の発表、開発過程時の数多くの英論文の深い理解等から判断して、十分な外国語能力を有すると判断した。また、論文内容及び審査会、公聴会での質問に対する説明内容などから総合的に判断して、最終試験は合格とした。

尚、主要な関連論文の発表状況は以下の通りである。(関連論文 5 編、参考論文 1 編)

- (1) 大林正直, 内山祥吾, 呉本堯, 小林邦和: 論文題目: 強化学習制御と適応 H_∞ 制御の協働型制御方式, 電気学会論文誌 C, Vol.131, No.8, pp.1467-1474 (2011)
- (2) 内山祥吾, 大林正直, 呉本堯, 小林邦和: H_∞ 追従性能補償器を備えたリアルタイム強化学習制御システム, 電気学会論文誌 C, Vol.132, No.6, pp.1008-1015 (2012)
- (3) 内山祥吾, 大林正直, 呉本堯, 小林邦和: 小脳パーセプトロン改良モデル利用型ロバスト制御システム, 電気学会論文誌 C, Vol.133, No.6, pp.1251-1258 (2013)
- (4) 内山祥吾, 大林正直, 呉本堯, 小林邦和, 間普真吾: 自己融合小脳パーセプトロン改良モデル利用型制御システムとその合意問題への適用, 電気学会論文誌 C, Vol.134, No.7, pp.990-998 (2014)
- (5) Shogo Uchiyama, Masanao Obayashi, Takashi Kuremoto, Kunikazu Kobayashi: Robust Reinforcement Learning Control System with H_∞ Tracking Performance Compensator, Proc. of 11th International Conference on Control, Automation and Systems, pp. 248 - 253 (2011)