

論文内容の要旨

博士論文題目

Practical Reinforcement Learning Design Methods for Home Appliances
(家電製品のための実用的な強化学習設計手法)

氏名

清水 達也

(論文内容の要旨)

家電製品に省エネ化や静音化などの所望の性能を与えるためには、その挙動の最適な制御が必要とされる。しかし、家電製品では、対象が衣類などを含み、事前にモデル化して挙動を正確に予測することが困難であるため、これまでの家電製品の制御は開発者の地道な作り込みを中心としていた。一方で、近年の機械学習技術の発展に伴い、そういったモデル化が困難な制御対象を、試行錯誤を通じて収集されたデータから最適な制御を可能とする強化学習技術が注目されており、本論文でもこれに焦点を当てている。

しかし、強化学習を家電製品など現実のシステムに適用するにあたって、幾つかの課題が残されている。特に、強化学習はシステム設計および学習段階に多くの試行錯誤が必要とされるが、家電製品では十分な試行を終える前に製品寿命を迎えるなど、現実的に実行困難である。この試行錯誤を軽減するには、システムの要求を満たすためのドメイン知識をその設計に適切に反映することが効果的だが、その反映方針は定まっておらず、やはり試行錯誤的に調整、あるいは熟練者のノウハウを頼りにすることが多かった。つまり、ある一つの家電製品に対して強化学習を適用できたとしても、他製品へと容易に転用できるような、一般的な設計論が十分に議論されていないのが現状である。

そこで本論文では、家電製品に強化学習を適用する際に、適切にドメイン知識を反映してシステム要求を満たす実用的な設計手法を提案している。具体的には、強化学習において基礎となる状態・行動空間を簡易的なモデル化を頼りに効率良く定義するとともに、その空間の粒度を計算資源に応じて調整する。その際には、粒度の粗さ次第で十分な制御性能を発揮できない事態を想定して、

氏名	清水 達也
----	-------

(論文審査結果の要旨)

強化学習を家電製品など現実のシステムに適用するにあたって、強化学習はシステム設計および学習段階に多くの試行錯誤が必要とされるが、家電製品では十分な試行を終える前に製品寿命を迎えるなど、現実的には厳しい面が存在する。この試行錯誤を軽減するために、システムの要求を満たすためのドメイン知識をその設計にシステムティックに反映する研究方針は、これまでのシステム開発で得られた多くの知見を最大限に有効活用する上で重要なものといえる。

この基本理念に基づき、本論文では強化学習の学習アルゴリズムよりも、その基盤の設計方法に注目している。つまり、強化学習では対象システムがマルコフ決定過程を満たすように、状態・行動空間を定義しなければならないが、これが満たされず不足する情報があると学習が困難となり、冗長であると計算資源の浪費や学習時間の増大が懸念される。本論文では、対象システムの挙動の簡易モデルを通じ、先んじて分析することで、必要な状態空間を明らかにするとともに、満たすべきシステム要求に応じて状態・行動空間の粒度を適切に設計する指針を提供している。また、実際のシステムでは満たすべき要求を強化学習の報酬として扱う際に、多目的化されることが避けがたいが、それらを同時に最適化しようとする学習コストが嵩むことが知られている。これを緩和すべく、満たすべき要求の内、優先度(必ず達成した制約と付加価値をもたらす要求)を設けて問題を分割・階層化する方法を提案している。これらの設計方法を、実際に3つの家電製品に適用した。その内の一つである洗濯機では、一般的な強化学習より大幅に計算資源を削減した実機での低騒音化・低振動化を達成しており、残りの二つである冷蔵庫・給湯器では、シミュレーションながらもデマンドレスポンス環境における電気料金の削減に成功している。

このように、本論文は実用性を最優先とした強化学習技術の応用方法について堅実に議論しており、実例よりその一定の社会的意義が評価できる。よって、本論文は博士(工学)の学位論文として価値あるものと認める。