

論文内容の要旨

博士論文題目： Highly Modularized Learning System for Behavior
Acquisition of Functional Robots
(高機能ロボットの行動獲得のための高度にモジュール化
された学習システム)

氏名： 山口 明彦

(論文内容の要旨)

ヒューマノイドに代表される高機能ロボットに不可欠な機能のひとつとして、ロボット自身が行動を獲得することがあげられる。このような機能は、エンドユーザがロボットに要求を伝えることを容易にする。本論文は、目的に基づくタスク設計システムの開発を目的とする。

具体的には、本論文は「高度にモジュール化された学習システム」を提案する。このシステムでは、強化学習手法及び関連手法群がモジュール化され、学習システムを構成する。提案するシステムは次の2つの特徴を持つ。(1) 強化学習を改善する学習戦略群もモジュール化される。(2) 単一のロボットの各々のタスクに対し、複数の学習戦略群が複数回適用される。学習戦略とは、例えば次元縮約、階層化、転移学習である。学習戦略群を用いて、システムはタスクの方策を学習する行動モジュール群を自動的に生成する。転移学習モジュール群は既に学習したモジュールから新たなモジュールを生成する。このように、各行動モジュールは複数の学習戦略を通して生成される。中心となるアルゴリズムを学習戦略フュージョンと呼ぶ。

本論文では、学習戦略を定義するための要素技術として、離散行動集合 DCOB 及びダイナミクスモデルの分離手法を提案する。離散行動集合 DCOB は規模の大きいタスクを扱うことができる。DCOB は、離散集合の探索能力を保ったまま、連続行動を学習するように拡張される。この手法は WF-DCOB と名付けられる。一部の転移学習戦略群は、WF-DCOB を用いて定義される。ダイナミクスモデルの分離手法は、タスク不変の要素の抽出を可能にする。これによって、あるタスクで学習したダイナミクスモデルをほかのタスクのモデルに転移する

ことが可能になる。加えて、提案手法を実装するために、高度にモジュール化された強化学習ライブラリ SkyAI を開発し、オープンソースで公開する。

提案手法群を検証するため、多数の実験を行った。DCOB と WF-DCOB は、特にシミュレーションのヒューマノイドの運動学習において、従来手法を上回った。DCOB は実機蜘蛛型ロボットの匍匐タスクにおいても、優れたパフォーマンスを示した。ダイナミクスモデルの分離手法は、迷路タスクにおいて学習速度を改善した。WF-DCOB は優れたパフォーマンスを示したが、学習戦略フュージョンはさらにパフォーマンスを向上させた。学習戦略フュージョンの拡張性は、シミュレーションのヒューマノイドの迷路タスクにおいて検証された。このタスクでは、ロボットは、過去に学習した匍匐や旋回を用いて、ゴールまでの経路を学習した。

これらの実験を通して、本論文は、高度にモジュール化された学習システムが、目的に基づくタスク設計を実現する現実的なアプローチのひとつであると結論する。

(論文審査結果の要旨)

本論文は、目的に基づくタスク設計システムの開発を目的として、「高度にモジュール化された学習システム」を提案している。このシステムを実現するための要素技術を開発し、シミュレーション及び実機実験による検証を行っている。本論文の成果は以下に要約される。

1. **SkyAI**: 提案システムを実装するため、オープンソースで公開される強化学習のソフトウェアライブラリ。モジュール構造が導入され、実機のロボットシステムで運用できるレベルの高い実行速度と学習システムを構成するための柔軟さをあわせ持つ。**SkyAI** は実行時にモジュール構造が変更可能なように設計されている。

2. **DCOB**: 価値関数の近似のために与えられた基底関数の集合から、離散行動空間を生成する手法。規模の大きい問題（高自由度ロボットの運動学習など）に適用可能である。**DCOB** の行動の周辺を連続的に探索し、パフォーマンスの向上を図る手法 **WF-DCOB** も合わせて提案される。**WF-DCOB** は、転移学習戦略群の定義に適するようにも設計されている。

3. モデルの分離学習: ダイナミクスモデルをタスク依存の要素とタスク不変の要素に分解する手法。この方法によって、あるタスクで学習したダイナミクスモデルをほかのタスクのモデルに転移することが可能となる。

4. 学習戦略フュージョン: 提案する高度にモジュール化された学習システムの中心となる手法。さらに、**WF-DCOB** やモデルの分離学習手法を用いて、学習戦略群が定義される。

5. ヒューマノイドの歩行学習: 提案システムを等身大ヒューマノイドロボットの歩行学習へ応用。バランス制御器を駆動しながら学習することにより、転倒コストを避けながら効率的に学習している。

以上のように、本論文は従来個別に提案されていた学習戦略をシームレスに融合し、エンドユーザが容易に利用できるシステムを提案するものである。このように、強化学習を応用しロボットの行動獲得を自動化する研究分野に大きく貢献するものである。よって本論文は、博士（工学）の学位論文として十分な価値があるものと認める。

氏 名	山口 明彦
-----	-------

(最終試験結果の要旨)

平成23年2月14日、全審査委員により、学位申請者に対して、論文内容および、それに関連する専門分野についての試問を行い、十分な学力と専門知識を有していることを確認したので、合格と判定する。