

機関番号：23101

研究種目：若手研究 (B)

研究期間：2009～2010

課題番号：21700258

研究課題名 (和文) 乳幼児期における知覚・運動の相互発達模倣型強化学習システム

研究課題名 (英文) Reinforcement Learning to Mimic Mutually Developments of Infants' Perceptual and Motor Abilities

研究代表者

永吉 雅人 (NAGAYOSHI MASATO)

新潟県立看護大学・看護学部・助教

研究者番号：70426542

研究成果の概要 (和文)：本研究では、システム設計者の状態・行動空間を構成する負担を減少するためのより適用性の高い強化学習システムの実現を目的として、まず、乳幼児期の運動発達をヒントに、Switching 強化学習モデルと学習器の切替法の提案を行った。さらに、乳幼児期の知覚発達を模倣した機能を加えた発達型強化学習モデルを提案し、状態・行動空間共構成法の提案を行った。以上により、システム設計者の負担軽減を実現し、強化学習のもつ実社会への適用可能性をより高いものとした。

研究成果の概要 (英文)：This research aims to realize Reinforcement Learning (RL) system that has more applicability in order to reduce designer's burden when he constructs RL agent's state and action spaces. First, I have proposed "Switching RL model" to mimic an infants' motor development and a method to switch learning modules. Further, I have proposed "Developments RL model" to mimic infants' perceptual and motor developments and a method to co-construct state and action spaces. Therefore, proposed methods have realized to reduce designer's burden and increase an applicability of RL to the real world.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2009年度	600,000	180,000	780,000
2010年度	600,000	180,000	780,000
年度			
年度			
年度			
総計	1,200,000	360,000	1,560,000

研究分野： 計算機知能学

科研費の分科・細目： 情報学・ 感性情報学・ ソフトコンピューティング

キーワード： 機械学習 強化学習 メタヒューリスティクス
ソフトコンピューティング 人工知能

1. 研究開始当初の背景

本研究は、意思表示の困難な高齢者・障害者のために、本人の微小な運動と本人意思の関係を学習することによるコミュニケーション支援や、使用機器に対する一人ひとりの「癖」にあわせたインターフェースの自律的最適化の実現を目指し、その前段階として、

それらへの応用が期待できる強化学習システムについての研究を行う。

強化学習は、対象システムのモデルを必要とせず、明示的な教師なしに報酬（スカラー値）のみを拠り所とした試行錯誤により、自律的に適切な制御規則を獲得していく汎用性の高い学習の枠組みである。しかしながら、

制御規則は通常 if-then ルールによって構成されるため、その学習能力は if-then ルールの前件部（状態空間）と後件部（行動空間）の構成に大きく影響される。そのため、強化学習の適用の際には、システム設計者が強化学習エージェントの状態空間と行動空間を事前にかつ適切に構成しておく必要がある。しかしながら、システム設計者にとって状態空間・行動空間の事前設計は大きな負担となっている。

そこで、状態空間を適応的に構成するために、モジュール型や階層型の強化学習の枠組みが提案され、多くの研究が行われている。しかしながら、これらはモジュールの配置やタスクの分割作業、サブゴールの設定を必要としており、それらはシステム設計者の手に委ねられる場合が多く、これらもまた、システム設計者にとって大きな負担となる。さらに、行動空間を適応的に構成するための研究はこれまで行われていない。

2. 研究の目的

本研究では、乳幼児期における知覚・運動の相互発達を模倣した強化学習システムを開発することにより、強化学習エージェントの状態空間と行動空間の適応的な共構成法を提案し、システム設計者の負担減少のためのより適用性の高い強化学習システムの実現を目的とする。

3. 研究の方法

乳幼児の知覚・運動発達の中でも以下に注目し、それぞれ強化学習システムの状態・行動空間に対応させて模倣を行うための計算モデルの提案を行う。

- 1) 可能な運動が微細化していくこと（運動発達）,
- 2) それぞれの感覚器に関して、感覚能力が高まっていくこと（知覚発達）.

さらに、それぞれの実現法の提案を行い、計算機実験により有効性を評価する。

4. 研究成果

(1) Switching 強化学習

時間経過に伴って行動を細かくしていく見本として乳幼児期における運動発達過程に注目し、その模倣を行う Switching 強化学習モデル(図 1)を提案した。このモデルは図 2のように2つの学習器と1つの切替器から構成され、乳幼児期における運動発達の特徴である「粗大運動から微細運動へ進む」ということを簡単に模倣して学習を行う。

切り替え法としては、ボルツマン選択法を用いた行動選択確率のエントロピーを用いて行う。

計算機実験として 2次元連続状態空間・1次元連続行動空間を有する経路計画問題に

適用した。その結果を図 2に示す。これにより提案手法が他より有効であることが確認できた。よって、この問題において設計者の行動空間の設計負担を軽減できるだけでなく、学習の高速化を達成できることが示された。

さらに、3次元連続状態空間・2次元連続行動空間を有するロボットナビゲーション問題を用いた計算機実験に適用した。その結果を図 3に示す。これにより提案手法が他より有効であることが確認できた。よって、この問題においても設計者の行動空間の設計負担を軽減できるだけでなく、学習の高速化を達成できることが示された。

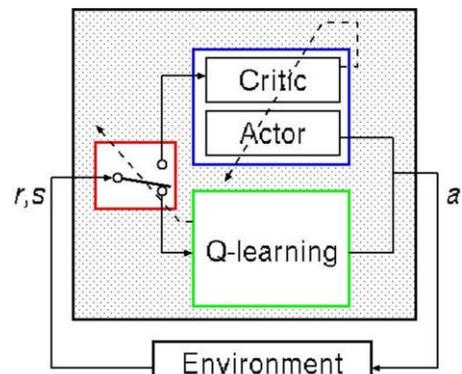


図 1 Switching 強化学習

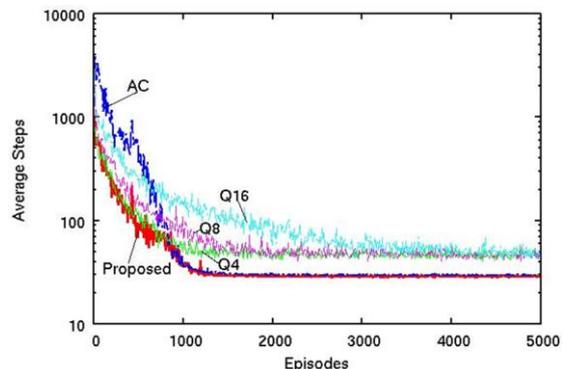


図 2 経路計画問題への適用結果

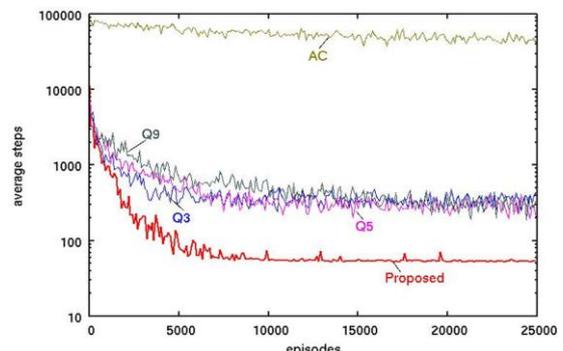


図 3 ロボットナビゲーション問題への適用結果

(2) 発達型強化学習

知覚・運動発達の模倣を行う発達型強化学習モデル(図 4)を提案した. このモデルは, 知覚発達過程の原理(未分化から分化されていくこと)を模倣するための状態フィルタ(入力状態空間(一般的な観測空間)から内部状態空間(一般的な状態空間)への写像 f を担う部分のこと)と運動発達過程の一つの特徴である粗大運動から微細運動へ進むことを簡単に模倣するための Switching 強化学習器から構成される.

切り替え法としては, Switching 強化学習と同様にボルツマン選択法を用いた行動選択確率のエントロピーを用いて行う.

計算機実験として 2 次元連続状態空間・1 次元連続行動空間を有する経路計画問題に適用した. その結果を図 5 に示す. これにより提案手法(図 5 の“FS”)が他より有効であることが確認できた. よって, この問題において設計者の行動空間の設計負担を軽減できることが示された.

さらに, 3 次元連続状態空間・2 次元連続行動空間を有するロボットナビゲーション問題を用いた計算機実験に適用した. その結果を図 6 に示す. これにより提案手法(図 6 の“FS”)が他より有効であることが確認できた. よって, この問題においても設計者の行動空間の設計負担を軽減できることが示された.

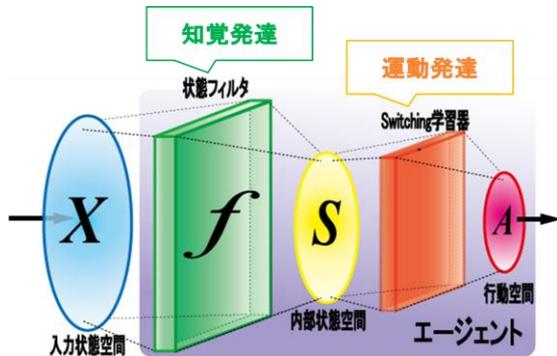


図 4 発達型強化学習

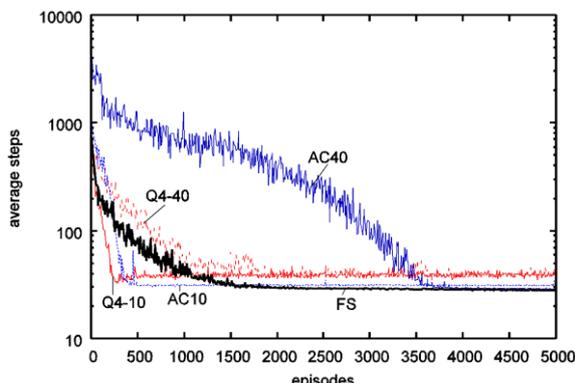


図 5 経路計画問題への適用結果

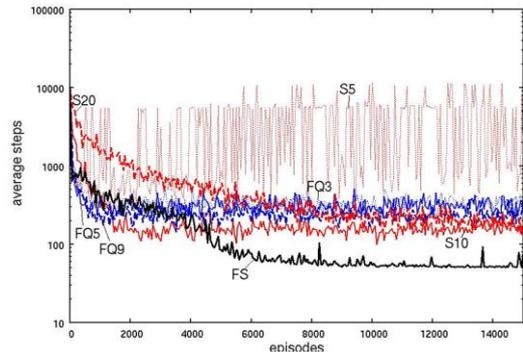


図 6 ロボットナビゲーション問題への適用結果

以上により, 強化学習のもつ実社会への適用可能性をより高いものとした. またこれにより, 内部構成的に一定でない不安定なシステムであるほうが, 一定で安定なシステムより有効なシステムとなりうるということを示す一例となりえた.

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 3 件)

- ① Masato Nagayoshi, Hajime Muraio and Hisashi Tamaki, Adaptive Co-construction of State and Action Spaces in Reinforcement Learning, Artificial Life and Robotics, Vol.16, (accepted), 2011.
- ② 永吉雅人, 村尾元, 玉置久, 連続行動空間への適用を考慮した Switching 強化学習, 電気学会論文誌 C, Vol. 131, No. 5, pp. 976-982, 2011.
- ③ Masato Nagayoshi, Hajime Muraio and Hisashi Tamaki, A Reinforcement Learning with Switching Controllers for Continuous Action Space, Artificial Life and Robotics, Vol.15, pp.97-100, 2010.

[学会発表] (計 9 件)

- ① 永吉雅人, 強化学習における状態空間と行動空間の適応的共構成法: 二次元連続行動空間への適用, 第 55 回システム制御情報学会研究発表講演会, 2011/5/19, 吹田市.
- ② Masato Nagayoshi, Adaptive Co-construction of State and Action Spaces in Reinforcement Learning, The 16th International Symposium on Artificial Life and Robotics 2011, 2011/1/29, 別府市.

- ③ 永吉雅人, 強化学習における状態空間と行動空間の適応的共構成に関する試み, システム・情報部門学術講演会 2010, 2010/11/25, 京都市.
- ④ Masato Nagayoshi, Switching Reinforcement Learning to Mimic an Infant's Motor Development - Application to Two-dimensional Continuous Action Space -, SICE Annual Conference 2010, 2010/8/19, 台北(台湾).
- ⑤ Masato Nagayoshi, A Reinforcement Learning with Switching Controllers for Continuous Action Space, The 15th International Symposium on Artificial Life and Robotics 2010, 2010/2/4, 別府市.
- ⑥ 永吉雅人, 強化学習における行動空間の適応的構成法の提案, 平成 21 年度電気関係学会北陸支部連合大会, 2009/9/12, 能美市.

[その他]

ホームページ等

<http://members.niigata-cn.ac.jp/nagayoshi/work-list-top.html>

6. 研究組織

(1) 研究代表者

永吉 雅人 (NAGAYOSHI MASATO)
新潟県立看護大学・看護学部・助教
研究者番号：70426542

(2) 研究分担者

()

研究者番号：

(3) 連携研究者

()

研究者番号：