

◆創発システム

強化学習における報酬の分配に関する研究

植村 渉

wataru@rins.ryukoku.ac.jp

大阪市立大学大学院 工学研究科

指導教員：辰巳 昭治

博士 (工学), 2005年9月 取得



Keywords: 強化学習, 報酬分配, 信頼度割当問題, profit sharing

概要: 本論文では, 報酬獲得型の強化学習である Profit Sharing の特徴を理論的に分析し, 従来の強化学習法では学習が難しい領域においても, 適切に学習を進めるための報酬分配方法について議論する。

まず, 問題環境の状態遷移が理想的なマルコフ決定過程 (MDPs) 環境において, ループを構築する系列への学習の抑制を実現するための条件を検討する。ループ系列の分岐となる状態において, ループ系列へ至るルールの強化量よりも, ループ系列から脱出するルールの強化量が大きくなるための条件を提案し, 合理的な学習を進めることを確認する。

次に, 現実的な環境の条件として, エージェントの知覚能力に制限がある場合として部分観測可能マルコフ決定過程 (POMDPs) 環境を検討する。報酬の累積による学習が有効に働くための POMDPs 環境の範囲を明らかにし, どのような報酬分配が必要となるか導く。

そして, 評価値に着目し, 強化関数と評価値の関係について議論する。強化関数による分配量に対して, 評価値の初期値が適切でない場合に生じる問題を検討する。適切な初期値は, 状態ごとに異なることを明らかにし, 状態ごとに独立して初期値を設定する方法を提案する。

以上の応用例として, 複数のエージェントが協調して目的を達成するマルチエージェント環境での報酬分配を検討する。各エージェントは知覚能力が制限されるため POMDPs 環境下での学習が求められる。学習が有効に働く条件下で, どのような報酬分配が必要かを明らかにする。

公表論文: 植村 渉, 上野敦志, 辰巳昭治: POMDPs 環境のためのエピソード強化型強化学習法, 電子情報通信学会論文誌, Vol. J88-A, pp. 761-774 (2005)

現職: 龍谷大学理工学部 (電子情報学科) 助手

論文入手先: <http://friede.elec.ryukoku.ac.jp/~wataru>

抱負: 「Roma was build in a day」最初から不可能だと思ふと, できることも失敗します。固定観念を捨てて挑戦することで, 壁を乗り越えていきたいです。

◆創発システム

環境との相互作用に基づく自律適応系の構成論的研究

谷口 忠大

tanichu@groove.mbox.media.

kyoto-u.ac.jp

京都大学大学院工学系研究科

指導教員: 榎木 哲夫

博士 (工学), 2006年3月 取得



Keywords: 双シマモデル, 記号創発, 自律適応系, 強化学習, モジュール型学習, コミュニケーション, STDP 学習則, 複数内部モデル

概要: 本研究では人間と機械との“意思疎通”という問題に対し, 記号論や発生的認識論, 脳神経科学, ロボティクスなどの諸領域の知見に示唆を求めつつ構成論的な接近を行った。近年, 人間社会への自律ロボットの進出が謳われているが, その実現のためには人間環境下における人間-機械間の適応的なコミュニケーションの実現が不可欠である。しかし, 未だ“意思疎通”を数理的に扱う理論は構築されていない。その実現のためには, 現在すでに数理的に理解されている C. Shannon の意味での情報量ではなく, 自律適応系が情報を自らの身体や経験に基づいて解釈するプロセス, つまり C. S. Peirce の意味での記号過程 (semiosis) を計算論的に構成し, 理解する必要がある。

本研究では, 自律適応系が環境との相互作用を通して環境適応を行う中で, 自らの内部モデル群, 強化学習器群を組織化する過程を, 双シマモデルとして構成した。さらに, 脳神経科学の分野において発見されている STDP 学習則をこれに結合させることにより, 記号過程を教師なし学習のプロセスを通してボトムアップに創発させられることを構成論的に示した。本研究を通して得られた結果は, 人間との相互作用を通して新たな振舞いや言葉を獲得していく自律ロボットの設計論基礎として重要であるだけでなく, 人間同士のコミュニケーションの原理的理解においても重要な意味をもつと考えられる。

公表論文: 谷口忠大, 榎木哲夫: 報酬設計を通じた社会的相互作用による行為概念群の構築: シマ理論に基づいた累積的強化学習知能と情報, 知能情報ファジィ学会誌, Vol. 18, No. 4, pp. 629-640 (2006)

現職: 京都大学大学院工学系研究科機械理工学専攻日本学術振興会特別研究員 PD

論文入手先: <http://tanichu.com/>

抱負: 本研究では自律エージェントが環境と相互作用する中で自らにとっての記号をいかに組織化し得るかという議論を強化学習やモジュール型学習のスキームを援用しつつ行った。今後は集団内での記号系の組織化などを視野に入れつつ, 工学的な展開が可能な新たな記号論を構築していきたい。