

生物的適応システム

Biologically Inspired Adaptive Systems

(研究プロジェクト番号：JSPS-RFTF 96I00105)

プロジェクトリーダー

小林 重信 東京工業大学大学院総合理工学研究科・教授

コアメンバー

伊藤 宏司 東京工業大学大学院総合理工学研究科・教授

新田 克己 東京工業大学大学院総合理工学研究科・教授

中村 清彦 東京工業大学大学院総合理工学研究科・教授

喜多 一 大学評価・学位授与機構・教授

山村 雅幸 東京工業大学大学院総合理工学研究科・助教授



1. 研究目的

生命の設計と生存の原理である「進化による新しい機能の獲得」と「環境とのインタラクションによる適応」を工学的に実現することを目的として、以下の3つの観点からの研究を推進した。

- (1) 進化計算を基盤技術とする創発的設計論
- (2) 強化学習を基盤技術とする創発的制御論
- (3) マルチエージェントシステムの設計と制御

2. 研究成果概要

2.1 進化計算を基盤技術とする創発的設計論

設計問題は最適化問題として定式化されるが、多峰性・非線形性・NP-困難性などの性質を持つ場合が多く、従来の最適化手法の接近を困難にしてきた。本研究では、進化計算、特にGA(Genetic Algorithms)に基づく創発的設計論の確立に向けて研究を推進した。

GAの理想的な挙動を与える機能分担仮説を提唱し、これに基づいて形質遺伝性に優れた交叉(EAX、UNDXなど)と多様性維持に優れた世代交代モデル(MGG、DDAなど)を開発した。

EAXはTSP(Traveling Salesman Problem)を対象とする形質遺伝性と生成子個体の多さに優れた交叉で、5,000都市規模問題の最適解を現実的な時間で発見することに成功している。

実数値GAについては、UNDX、UNDX-m、SPX、ENDXなど、多峰性・非線形性・悪スケール性に対してロバストな交叉を数多く開発した。

実数値GAの設計問題への適用可能性を検証するために、レンズ系設計に適用した。その結果、専門家や領域知識に依存することなく、レンズ系を創発的に設計することに成功した。これは、光学分野での設計パラダイム変革につながる技術的ブレークスルーを与えるとの評価を得ている。

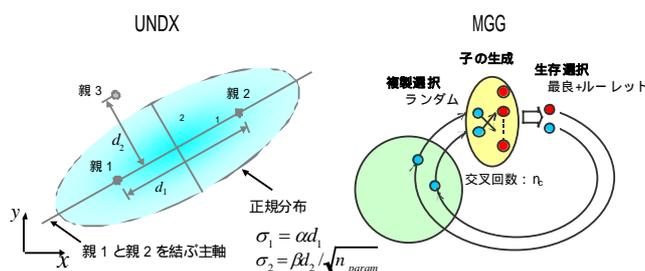


図1 交叉 UNDX と世代交代モデル MGG

2.2 強化学習を基盤とする創発的制御論

不確実性や報酬の遅れ、時間的変化を伴う実世界で動作するロボットをつくる際、環境同定の困難さが従来の制御系設計論による接近を妨げてきた。本研究では、適応制御の新しい枠組みとして優れた特徴を持つ強化学習に着目して、創発的制御論の確立に向けて研究を推進した。

強化学習の実用化に当たっては、状態観測の不完全性に対してロバストであること、連続状態・連続行動出力を陽に扱えることが求められることを考慮して、学習の基本的枠組みとして Actor-Critic を採用した。Actor には確率的政策に従い行動を選択する確率的傾斜法に加えて、適性度の履歴 (Eligibility Trace) を導入することにより、不完全な状態観測下でも安定的かつ高速に学習が行えることを理論解析で示すとともに、実験により解析結果を確認した。

上記の提案手法の有用性を確認するために、4足歩行ロボット(8自由度)への適用を試みた。環境のモデルを与えないで、実時間で約50分後に適切な歩行動作を獲得することに成功した。摩擦の小さい絨毯上ではボディを引きずる歩行、摩擦の大きいゴムマットの上ではボディを持ち上げての歩行が観察され、環境に応じて適切な政策を学習できることを確認した。

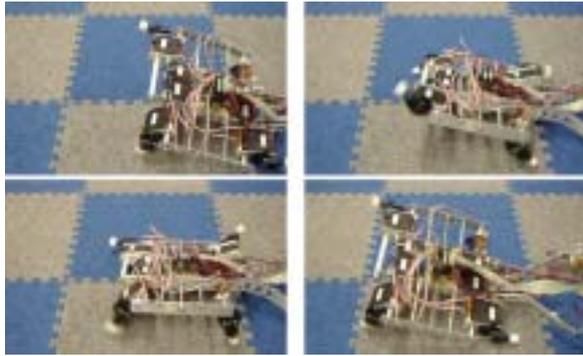


図2 学習で得られた4足ロボットの歩行動作

2.3 マルチエージェントシステムの設計と制御

マルチエージェントシステムへの取組は多様であり、本研究でも様々な観点からさまざまなアプローチを試みたが、ここではマルチエージェント強化学習によるアプローチを紹介する。

マルチエージェント強化学習では、各エージェントが他のエージェントと協調を取りながら学習を進めることが必要であり、そのためには報酬配分と情報共有を適切に行うことが求められる。Profit Sharing を用いたマルチエージェント強化学習の下で、報酬配分の理論的考察を行った結果、間接報酬（直接報酬を得たエージェントから他のエージェントに再配分される報酬）には、比率により協調が促進される場合と抑制される場合があることを明らかにした。

マルチエージェント強化学習の実用性を検討するために、上下水道系制御などへの適用を行った。下水道系制御系はひとつの処理場と複数の中継ポンプ場からなり、各施設は貯水池を持つ。処理場では汚水を浄化するため生化学処理を施すために揚水量をできるだけ一定に保つことが、各ポンプ場では貯水量の上下制限を満たすよう揚水量を柔軟に制御することが求められる。本研究では、各施設を学習エージェントとして、ツリー状の接続構造を考慮して、報酬配分と報酬共有を適切に行うこととし、シミュレータ上で Q-learning による学習を行わせた結果、処理場での切り替え回数が1日平均 1.5 回程度の極めて高い水準の制御性能を達成することに成功した。

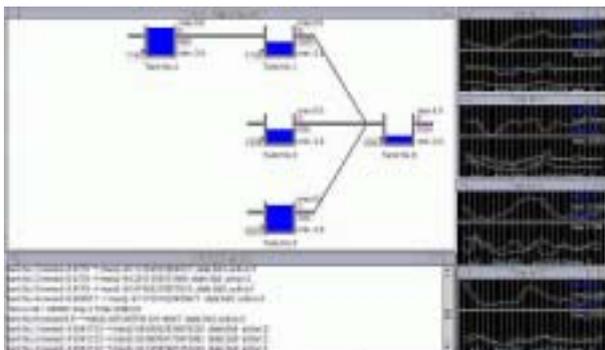


図3 強化学習による下水道系の適応制御

3 . 結論

伝統的な設計論や制御論では、領域専門家の存在が前提とされ、専門家は設計仕様や制御仕様を決めるだけでなく、それらをいかに実現するかについて試行錯誤的な役割が期待されてきた。「生物的適応システム」の研究成果は、専門家が介在することなく、設計や制御を創発的に実現できることを強く示唆しており、工学的シミュレーションの分野にパラダイム変革を生み出す可能性が大きくなったといえる。

主な発表論文

- (1) Y. Nagata and S. Kobayashi: Edge Assembly Crossover: A High-power Genetic Algorithm for the Traveling Salesman Problem, Proc. of 7th Int. Conf. on Genetic Algorithms, pp.450-457 (1997).
- (2) I. Ono and S. Kobayashi: A Real-coded Genetic Algorithm for Function Optimization Using Unimodal Normal Distribution Crossover, Proc. of 7th Int. Conf. on Genetic Algorithms, pp.246-253 (1997).
- (3) H. Kimura and S. Kobayashi: An Analysis of Actor/Critic Algorithms using Eligibility Traces: Reinforcement Learning with Imperfect Value Function, 15th Int. Conf. on Machine Learning, pp.278-286 (1998).
- (4) K. Miyazaki and S. Kobayashi: Rationality of Reward Sharing in Multi-agent Reinforcement Learning, Second Pacific Rim International Workshop on Multi-Agents, pp.111-125 (1999).
- (5) I. Ono, S. Kobayashi and K. Yoshida: Optimal Lens Design by Real-Coded Genetic Algorithms Using UNDX, Computer Methods in Applied Mechanics and Engineering, 186, pp.483-497 (2000).