

博士論文概要

論文題目

自律型ロボットにおける制御回路構造の学習手法
～オンライン・リアルタイムな
ネットワーク構造の強化学習～

Learning Method for Control Circuit in
Autonomous Robot
～ Online Real-Time Structural
Reinforcement Learning～

申請者

金 天海

Chyon Hae Kim

機械工学専攻 精密工学研究

2007年 12月

現在、学術的価値・産業的価値の両面から自律型ロボットの応用への期待が高まっている。一方で、自律型ロボットにおける効果的な制御系の構築法は未だ完成されたいは言い難く、ロボット開発の障害となっている。ロボットが設計者の想定できない環境へ対処する場合、設計段階における効果的な制御則の実装が原理的に不可能か、またはできたとしても非常に困難であることが多い。そして、このような場面に対して効果的な制御則を提供するには学習制御を用いることが有効であるとされている。

学習制御を搭載するロボットの要求機能としては、未知環境や新しいタスクに対して行動創発を行うための効果的な振る舞いの創発力（創発性）、環境・タスク毎の調整を省くための多様な環境への適応力（適応性）、さらには動的な変化へ対応するための即時的な対応力（即時性）等が挙げられる。ロボットシステムにおいては、これらの機能の両立が必要となるが、現在そのために使用できる学習制御システムは存在しない。上記創発性、適応性、即時性から学習制御器側の要求仕様を導くと、学習制御システムには行動創発、汎化・抽象化、柔軟性、オンライン性、漸次性といった要素が必要であり、従来型の学習システムではこれらの両立は難しい。

本論文では以上の観点から従来型の学習制御システムを分類した。現在、強化学習、ニューラルネットワーク、遺伝的アルゴリズムといった領域を軸に多くの学習制御システムが提案されている。例えば、**Direct Vision Based – Reinforcement Learning** はニューラルネットワークと強化学習を融合することで、ニューラルネットワークの苦手とする行動創発能力、強化学習の苦手とする汎化・抽象化能力を補っている。しかしながら、このシステムを設計する際にはタスクや環境に応じたネットワーク構造・パラメータの設定が必要であり、柔軟性を欠いている。また、ニューロジェネティククラニングではこの問題を遺伝的アルゴリズムとニューラルネットワークの融合によって解決しているものの、遺伝的アルゴリズムの計算を行う際の評価時間の問題からオンライン性の確保が困難である。

そこで本論文では、学習制御システムに必要な機能項目（行動創発、汎化・抽象化、柔軟性、オンライン性、漸次性）を両立するための枠組みとして、自己組織化回路素子 **Self-Organizing Network Elements (SONE)**を提案する。SONEでは、ネットワーク型の制御回路構造を学習によって獲得することで、汎化・抽象化、柔軟性を確保することができる。また、ネットワーク内に強化信号を伝播して学習する強化信号伝播規則を導入することで、行動創発を実現することができる。そして、伝播した強化信号をネットワーク内の各素子へ蓄積することで、漸次性とオンライン性の高い学習を実現することができる。よって、SONE は上記五項目を両立する枠組みであるといえる。以下に本論文各章の概要を示す。

本論文第二章では以上の基本構想に従い、自律型ロボットへの応用に関する仕様を満たす学習制御器の構成法として、自己組織化回路素子 **Self-Organizing Network Elements (SONE)**を提案する。またここでは、提案する学習制御器構成法の具体的な実装例として、自己組織化論理回路の構成法について述べ、自己組織化論理回路の学習に関する数学的な証明について述べる。

本論文で提案する SONE では、ネットワークを構成するために必要となる素子のそれぞれを独立したシステムとして記述する。また、各素子が持つ機能として、次の機能を含めるものとする。(1) 強化信号伝播 (2) 自己解体 (3) 新しい素子の生成 各素子は強化信号伝播機能によって、強化信号（報酬値）をより多く受け取るための学習である、強化学習を行う。この強化信号伝播機能によってネットワーク上の素子は互いに評価値を算出し合い、互いの重要度を決定できる。さらに、強化信号をより多

く受け取れるようにネットワークの構造を変更する機能として、自己解体機能、新しい素子の生成機能を持たせている。これらの機能は、先の強化信号伝播規則によって決定した評価値をもとに実行され、素子単位でのネットワーク構造の生成と淘汰を実現するための機能である。筆者らは、これらの機能によって、ネットワーク構造の自己組織化をオンライン・リアルタイムに行う強化学習の実現ができると考えた。また、この枠組みによるネットワークの設計は報酬量を安定的に増大させる回路素子の設計に帰着し、その設計がタスク・環境と独立したものとして扱えるため、汎用的な学習システムの構築も期待できる。

SONE へ用いる素子の具体的な実装について述べる。SONE の枠組みでは、多様な素子への実装が考えられる。著者らは、その中でも比較的容易な対象である二値による論理回路素子 (AND ノード, OR ノード, 反転リンク, 非反転リンク等) へ SONE を実装した。ここではそのうち OR ノードへの実装を例に SONE の実装法を説明する。SONE の各ノードの演算は、出力フェイズ, 伝播フェイズ, 構造変更フェイズに分類できる。出力フェイズではネットワークの順方向演算を行い、OR ノードは出力 (OR 演算) を計算する。伝播フェイズでは、各素子に割り当てられた規則に従って強化信号の伝播を行う。構造変更フェイズでは、新しい素子の導入に関するテストと不要な素子の削除に関するテストを行い、その結果に従ったネットワーク構造の変更を行う。OR ノードの例では、ノードの入力側にあるリンクのテストを行う。強化信号の伝播規則は次のように求めたものを使用した。まず、素子に対して学習効果を与えるために次のような伝播規則設定指針を構成した。(1) 各素子が、自らが受け取った強化信号を他の素子へ伝播する際に、受け取った強化信号以上の信号を周囲の素子へ与えることを禁止する。(2) 受け取る強化信号の期待値が 0 を上回るテスト用素子を実用化する。(3) 受け取る強化信号の期待値が 0 を下回る素子は削除する。以上の指針は報酬の発散や学習効果の発生を考慮した指針である。以上の指針によって伝播規則を構成した後、この伝播規則に対して①学習効果に関する証明、②回路の冗長性に関する証明が正しく行えることを確かめた。他の素子に関しても同様に実装することができる。

第三章では、SONE の教師あり学習への応用と、SONE の基本特性解析試験について述べる。ここでは、軌道学習に関する試験によって、SONE の耐ノイズ特性、追加学習特性、時系列学習特性を解析すると共に、二重螺旋問題を用いた試験によって、汎化能力に関する試験も行う。これらの試験ではリカレントニューラルネットワーク等を用いた学習法との比較を適宜行い、SONE を用いた場合の利点についてまとめる。前章で作成した SONE に関してここでは、行動創発、汎化・抽象化、柔軟性、オンライン性、漸次性の 5 項目に関する試験のうち汎化・抽象化、柔軟性、オンライン性、漸次性に関する検証を行うため、二重螺旋問題と軌道学習に関するいくつかの実験を行う。

まず、二重螺旋問題について説明する。二重螺旋問題は主に汎化・抽象化に関する能力を問うための一般的な手法であり、主にニューラルネットワークの分野で用いられている。この問題によって、二次元平面上に螺旋状に配置した点データを学習制御器に学習させることにより、データ間の補完具合を調べることができる。Fahlman がニューラルネットワークで得たデータとの比較から、SONE でも汎化・抽象化が可能であることがわかった。

次に軌道学習に関する実験について説明する。二次元平面上での軌道を学習させる実験によって、柔軟性、オンライン性、漸次性に関する評価を行った。軌道学習実験では、学習対象となる軌道は予め与えられ、ネットワークの入力データはその軌道上

の一点の座標である。ネットワークは次に移動すべき点の座標を出力し、その座標が学習データとずれている場合は誤差に対する学習を行う。本論文では、この実験をさらに拡張し、二つの問題を切り替えながら学習を行う追加学習に関する実験を導入し、オンラインな状況下における漸次的学習の成否を調べた。実験の結果、SONEによる学習はリカレント・ニューラルネットワーク (RNN) よりも安定しており、収束速度も速いことがわかった。これにより、SONE がオンラインなデータ系列に対し漸次性を持った学習を行えることがわかった。また、様々な形状の軌道を与えて学習した結果、データに対する柔軟性も確認することができた。

第四章ではシミュレーション上の移動ロボットを用いた衝突回避実験について述べる。この衝突回避実験によって、SONE が行動創発、オンライン性を両立できることを確かめた。この実験で用いた移動ロボットは八つの赤外線センサ (図中では 8 本の線分で示してある)、二つの車輪を回すモータを備えており、フィールド上の壁を探知しながら移動することができる。ただしこの実験ではロボットを制御するプログラムは何も用意せず、ロボットの入出力に対して SONE 回路素子を配置するのみである。よってロボットは、全ての必要な回路構造を SONE の自己組織化によって獲得する必要がある。ロボットには直進行動に対する報酬値と壁付近における罰 (負の報酬値) を設定し、フィールドを探索させる。この実験の結果、実験開始時にはランダムな行動が多かったロボットの動作が約 5 分程度で改善されることがわかった。ロボットはランダムな行動の経験から、状況に応じた動作を学習・獲得できたものと考えられる。ここでは、ロボットの行動創発に関する学習がオンラインに行えることを確かめた。

第五章では、SONE の内部に発生するノイズ除去に関する実験を通じて、SONE による教師あり学習と強化学習の関連性について考察する。この実験ではノイズ対策を施した SONE による移動ロボットの衝突回避を扱う。

SONE の強化学習方式では、教師あり学習の場合とは異なり、強化信号の与え方がネットワーク内部の報酬値に対しノイズを発生させることがわかった。また、従来の自己組織化論理回路へ実装している素子生成・淘汰の判定方法はこのノイズを軽減させる働きがあることもわかった。しかしながら、ノイズを効果的に除去するためには素子生成・淘汰に関する判定方法は動的に調整されなければならないことが明らかとなった。

そこで、本論文ではその動的な調整方法を対数関数の性質を応用した計算式で各素子に対して実装した。これを実装するにあたって必要となる関数形状の選定には単純な 3-bit 演算回路を構成する試験を、人為的にノイズを与えた環境下で用いた。また、選定はその正答率に従って行った。ここで得られた耐ノイズ性を向上させた SONE 素子を用いて第四章と同様の実験を行い、ロボットの振る舞いがより精度良く改善できることを確かめた。

第六章では本論文の全体に関する考察とまとめ、さらには今後の展望について述べる。自律型ロボットにおける学習制御器の要求機能である、行動創発、汎化・抽象化、柔軟性、オンライン性、漸次性の五項目を両立する学習制御器として、自己組織化回路素子 **Self-Organizing Network Elements (SONE)** を提案した。SONE を用いた基礎実験によって、汎化・抽象化、柔軟性、オンライン性、漸次性が確認できた。また、ロボットシミュレーションによって、行動創発、オンライン性が確認できた。よって、SONE による上記五項目の実現が確認できた。本手法をさらに発展させるために、入出力の連続化、遅延報酬への対応、大規模ネットワークの実現などを行っていきたい。

早稲田大学 博士（工学） 学位申請 研究業績書

氏名 金 天海 印

(2007年11月 現在)

種 類 別	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む）
論文	<ul style="list-style-type: none"> ○ 自己組織化論理回路におけるノイズの抑制, 日本ロボット学会誌 Vol.25 No.6,(2007), <u>金天海</u>, 出澤純一, 尾形哲也, 菅野重樹 ○ Enhancement of Self Organizing Network Elements for Supervised Learning, IEEE International Conference on Robotics and Automation (ICRA), (2007), <u>Chyon Hae Kim</u>, Tetsuya Ogata, and Shigeki Sugano ○ Efficient Organization of Network Topology based on Reinforcement Signals, IEEE International Conference on Intelligent Robots and Systems (IROS), (2006), <u>Chyon Hae Kim</u>, Tetsuya Ogata, and Shigeki Sugano ○ Improvement against Noises in Self-Organizing Logic Circuit, IEEE International Conference on Information Acquisition (ICIA), (2006), <u>Chyon Hae Kim</u>, Jyun-ichi Idesawa, Tetsuya Ogata, and Shigeki Sugano ○ ローカルルールに基いた論理回路の自己組織化アルゴリズム, 計測自動制御学会論文誌 Vol.42 No.4 pp.334-341, (2006), <u>金天海</u>, 尾形哲也, 菅野重樹 ○ Self-Organizing Algorithm for Logic Circuit based on Local Rules, IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), (2005), <u>Chyon Hae Kim</u>, Tetsuya Ogata, Shigeki Sugano
講演	<p>自己組織化回路素子へのフリップフロップ素子導入による時系列学習, 人工知能学会 第21回全国大会, (2007), <u>金天海</u>, 出澤純一, 菅野重樹</p> <p>自己組織化回路素子(SONE)への教師あり学習の付与, 情報処理学会第69回全国大会, (2007), <u>金天海</u>, 尾形哲也, 菅野重樹</p> <p>自己組織化回路素子 (SONE) の開発, 早稲田大学 21世紀COE 第4回 若手研究者によるWorkshop, (2006), <u>金天海</u></p> <p>自己組織化回路素子SONEへの教師あり学習機能の付与, 第7回計測自動制御学会システムインテグレーション部門講演会, (2006), <u>金天海</u>, 尾形哲也, 菅野重樹</p>

早稲田大学 博士（工学） 学位申請 研究業績書

種 類 別	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む）
講演	<p>自己組織化論理回路における対ノイズ性能の向上, 第24回日本ロボット学会学術講演会, (2006), <u>金天海</u>, 出澤純一, 尾形哲也, 菅野重樹</p> <p>An Algorithm for Self-Organizing Logic Circuit, The 3rd COE-CIR Joint Workshop, (2006), <u>Chyon Hae Kim</u>, Tetsuya Ogata, Shigeki Sugano</p> <p>自己組織化論理回路における学習アルゴリズムの解析, 第6回計測自動制御学会システムインテグレーション部門講演会, (2005), <u>金天海</u>, 尾形哲也, 菅野重樹</p> <p>Self-Organizing Logic Circuit based on Local Rules, The 2nd COE-CIR Joint Workshop, (2005), <u>Chyon Hae Kim</u>, Tetsuya Ogata, Shigeki Sugano</p> <p>ローカルルールに基づいた論理回路の自己組織化アルゴリズム, 第5回計測自動制御学会システムインテグレーション部門講演会, (2004), <u>金天海</u>, 尾形哲也, 菅野重樹</p>
特許	<p>名称：情報処理システムおよび情報処理方法, 並びにプログラム 発明者：菅野重樹, <u>金天海</u>, 尾形哲也</p> <ul style="list-style-type: none"> ・特願 2004-363742 ・特開 2006-172141 ・PCT / JP2005 / 21062
出展	<p>情報処理システムおよび情報処理方法, 並びにプログラムの特許, 計測展 2005 TOKYO, (2005), <u>金天海</u>, 尾形哲也, 菅野重樹</p>

早稲田大学 博士（工学） 学位申請 研究業績書

種 類 別	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む）
その他 講演	<p>自己組織化回路素子SONEにおけるフリップフロップ素子導入によるシーケンスの分節化と統合，第7回計測自動制御学会システムインテグレーション部門講演会，(2006)，出澤純一，<u>金天海</u>，尾形哲也，菅野重樹</p> <p>自己組織化ネットワーク素子群における対ノイズ性能向上，ロボティクス・メカトロニクス講演会，(2006)，出澤純一，<u>金天海</u>，尾形哲也，菅野重樹</p>