

シリーズ システム・制御のニューフロンティア C-2

自律的行動創発システムと身体性

—— 機械獣の構成論 ——

吉田 尚人 著

コロナ社

「シリーズ システム・制御のニューフロンティア」編集委員会

委員長 永原正章 (広島大学)
委員 河島茂生 (青山学院大学)
(五十音順) 谷口忠大 (京都大学)
望山 洋 (筑波大学)

(2025年8月現在)

シリーズ刊行のことは

21世紀も四半世紀を終え、人工知能やロボティクス、通信ネットワークなどの科学技術はかつてない速度で進化を続けています。これらの情報技術の爆発的な発展とともに、地球規模での自然環境の激変や社会構造の多様化と複雑化は、我々が直面する課題を大きく変容させました。このような状況下において、システム・制御の学問は、新たなフロンティアを開拓し続けています。

前世紀に発展した古典的なシステム・制御理論は、おもに物理法則に基づくモデル化と解析・設計を中心として発展してきました。しかし、現代社会におけるシステムは、人間、社会、経済、環境など、多様な要素が複雑に絡み合い、物理モデルのみに立脚した古典的手法では捉えきれない側面を多く持ちます。

本「シリーズ システム・制御のニューフロンティア」は、このような認識に基づき、既存の枠組みを超えた新しいシステム・制御の学問体系を構築することを目的としています。具体的には、以下の四つのカテゴリにおいて、最新の研究成果と将来展望を提示します。

【カテゴリ A：新しいシステム・制御の基礎理論】

サイバーフィジカルシステムや人間社会の中でのシステムなど、現代的なシステムを扱うための新たな基礎理論の枠組みを展開します。

【カテゴリ B：システムの知能】

人工知能とシステム制御理論が融合した新しいシステム設計論を探求します。

【カテゴリ C：システムの身体性】

身体と環境のダイナミクスとそれに基づく知能の構成を目指し、物理世界と相互作用するシステムの制御メカニズムを探求します。

【カテゴリ D：人間社会におけるシステム・制御】

社会科学や経済学，心理学など人文社会系の視点を取り入れた，人間社会におけるシステムの解析や設計について最新の理論を紹介します。

本シリーズは，大学院生，研究者，そしてシステム・制御に関心を持つすべての皆様を対象としています。各巻は，第一線の研究者によって執筆され，最新の研究成果，未解決問題，将来展望を網羅的に解説しています。本シリーズが，システム・制御の新たな地平を切り拓く一助となることを願ってやみません。

それではともに，システム・制御のニューフロンティアへと踏み出しましょう！

2025年8月

「シリーズ システム・制御のニューフロンティア」編集委員会

まえがき

草むらにいる鈴虫を拾って、虫かごに入れる。虫かごにキュウリを入れて適切な環境を与えてやると、虫は虫なりに適応して、その中での生活をわれわれに見せてくれる。ペットショップで一目惚れした柴犬ひとめぼを家に連れてくると、徐々に餌や散歩のルールを覚えて家になじみ、家の中で生きる愛らしい姿を見せてくれる。彼らは、生きているかのように動きを演じているのではなく、確かに生きていて、生きるために必要な動きを生み出すことで環境や自らの身体に適応し生活を営んでいる。彼らは、それぞれの環境で生きていく上で必要な行動を創発するための要件を満たしている点で同等であって、違いはその身体的制約と生活する環境、そして認知能力であり、それらは相互に依存し合っている。

動物が見せるこのような姿のように、ロボットもまたあたかも生きているかのように動かすのではなく、何らかの基準で「ロボットが生きている」ことを定義して、そこから生きるために必要な動きをロボットに自律的に獲得させることは可能だろうか？ また、そのようなエージェントは逆に、実際の動物について何らかの類似性を見せるのだろうか？

本書は、このような問いに対して**ロボットにおける恒常性**という観点に着目し、これに基づく行動創発を構成論的アプローチで追求するものである。ここで**恒常性**とは、身体の内側の重要な値をできるだけ所定の範囲に保とうとするプロセスである。**構成論**とは、現象を「作ることによって理解する (understanding by building)」アプローチとも呼ばれ^{251)†}、着目した現象を発生し得る

† 肩付き数字は、巻末の引用・参考文献番号を示す。

機構を仮説的に構成し、実際に動作させることを通して適切な機構を探求する。

本書は「エージェントが身体の内部状態を保つための行動最適化」という、ただ一つの技術を扱う。この技術を成立させる上で、機械学習・計算論的神経科学・生理学・伝熱工学・ロボティクス・栄養生物学・温度生物学・動物行動学・理論生物学の知見を統合し、統合された自律ロボットを構成する。また本書は、工学的に明確なタスク解決を目的としたロボットとは概念的に区別された、外在的に定められた意味や目的を持たない生活体としての自律機械を**機械獣**と呼んで位置づけ、基盤技術を構築するとともに、関連する研究の歴史と重要性、そしてその将来的な可能性について議論する。

本書の執筆のきっかけをいただいた谷口忠大氏（京都大学）と大黒達也氏（東京大学）に厚く御礼を申し上げる。強化学習の理論について最終的な確認をしていただき、全体の文章に対して建設的な意見をいただいた大塚誠氏（LiLz 株式会社）に感謝を伝えたい。また本書の草稿を確認していただき、研究会等々にて建設的な数々の意見・感想をいただいた堀部和也氏（理化学研究所）、鈴木啓介氏（北海道大学）に深く感謝する。本文については谷口忠大氏と大黒達也氏にも助言をいただき、序論の無意味性・無目的性の可能性に関しては、美術家の鮫島ゆい氏との議論がたいへん参考になった。謹んで御礼を申し上げます。本書の基になったのは筆者の博士論文「恒常性に基づく自律エージェントの行動創発に関する構成論的研究」である。東京大学大学院情報理工学系研究科での博士課程では國吉康夫氏（東京大学）に全体のご指導を、金沢星慶氏（東京大学）には日頃から相談に乗っていただき、たいへん充実した博士課程を終えられた。この機会に改めて感謝の意を表したい。また、本書の編集をご担当いただいたコロナ社の各位に深く感謝する。

最後に、博士課程を通して筆者の研究をいつも暖かく見守ってくれた両親（倫之、光子）に心から感謝する。

2025 年 秋の京都にて

吉田 尚人

目 次

1. はじめに

1.1 無目的なロボット	2
1.2 恒常性に基づく自律ロボットの行動統合とその生成	4
1.3 サイバネティクスから計算論的神経科学へ	7

2. 人工システムにおける生存

2.1 恒常性の概念	10
2.2 Ashby の超安定システム	13
2.3 先行研究におけるシミュレーション・エージェントの生存・恒常性	16
2.3.1 動物行動学に基づくアプローチ	18
2.3.2 人工生命・神経モデリング・進化計算に基づくアプローチ	19
2.3.3 認知アーキテクチャ・汎用人工知能のアプローチ	22
2.3.4 計算論的神経科学におけるアプローチ	23
2.4 先行研究における自律型ロボットの生存・恒常性	26
2.4.1 WAMOEB A シリーズ	27
2.4.2 動物行動学からのアプローチと Animat アプローチ	28
2.4.3 身体性認知科学・自律エージェント設計からのアプローチ	30
2.4.4 強化学習に基づくアプローチ	32
2.4.5 ヒューマンロボットインタラクションからのアプローチ	33

2.5	背景研究を踏まえた機械獣研究の立場	35
2.6	オートポイエシスの自律性の議論と機械獣研究の関係, および関心領域	39

3. 強化学習

3.1	エージェント・環境相互作用モデル	43
3.2	MDP のダイナミクスと期待報酬	45
3.3	価値関数・行動価値関数と最適方策	46
3.4	MDP において Markov 方策を用いる正当性	47
3.5	方策勾配に基づく強化学習	49
3.6	アクター・クリティックアーキテクチャ	51
3.7	価値関数・アドバンテージ関数の推定	53
3.8	一般化アドバンテージ推定	53
3.9	深層強化学習	55
3.9.1	深層学習の概要	55
3.9.2	深層強化学習：深層学習＋強化学習	58
3.9.3	近接方策最適化法	59

4. 恒常性強化学習

4.1	既存の恒常性に基づく動機づけ行動メカニズムに関する問題点	64
4.2	恒常性強化学習：恒常性＋強化学習	65
4.3	恒常性強化学習の定式化	69
4.4	関連するアプローチ	71
4.5	内受容感覚の模倣学習としての恒常性強化学習	73
4.6	認知発達ロボティクスにおける動機づけシステムの観点からの位置づけ	74

4.7	強化学習におけるタスク統合問題との関係	79
4.8	恒常性強化学習の応用における既存の問題	80

5. 深層恒常性強化学習

5.1	ニューラルホメオスタット	83
5.1.1	低次元入力ニューラルホメオスタットの構成	85
5.1.2	画像入力ニューラルホメオスタットの構成	85
5.1.3	最適化の詳細とハイパーパラメータ	86
5.1.4	深層恒常性強化学習における終端状態の扱い	87
5.1.5	恒常性報酬と報酬の基本設定	88
5.2	先行研究における恒常性報酬	89
5.2.1	基礎恒常性報酬	89
5.2.2	バイアス恒常性報酬	89
5.2.3	カートポール報酬	90
5.3	報酬補正と補正恒常性報酬の関係	90
5.4	実験の概要	91
5.5	実験1：連続モータ制御による2資源問題 (TRP)	92
5.5.1	連続モータ制御条件下での恒常性による採餌行動の獲得	94
5.5.2	餌取り報酬との比較	96
5.5.3	内受容固定法による行動解析	97
5.6	実験2：体温調節環境 (thermal) でのエネルギーと体温の同時制御	100
5.6.1	連続モータ制御条件下での恒常性による体温制御・採餌行動の同時獲得	101
5.6.2	内受容固定法による体温制御エージェントの行動解析	103
5.7	実験3：カメラ画像入力を用いた TRP 環境	105
5.7.1	深層恒常性強化学習の高次元環境でのスケール性の実証	107
5.7.2	内受容感覚に依存した視覚顕著性	108
5.7.3	ニューラルホメオスタットの内部表現	111
5.8	個性性の発現：個別のエージェントの運動傾向	112
5.9	探索行動の創発可能性	115

6. 機械獣の構成

6.1	なぜロボットを作るのか？	118
6.2	機械獣の構成へのアプローチ	119
6.3	機械獣のハードウェア構成	120
6.4	機械獣の充電条件と実験手順	121
6.5	出力の構成	123
6.6	観測情報の構成	124
6.6.1	内受容感覚の構成 I：平均モータ温度の定義	124
6.6.2	内受容感覚の構成 II：正規化エネルギー	125
6.7	シミュレータの構成	126
6.7.1	機械獣・外部環境の動力学シミュレーション	126
6.7.2	シミュレーションに対する実ロボットの遅延の評価	127
6.7.3	内部ダイナミクスのモデリング I：モータ温度	127
6.7.4	パラメータフィッティング	129
6.7.5	内部ダイナミクスのモデリング II：バッテリー・正規化エネルギー	130
6.8	深層恒常性強化学習による行動最適化	132
6.8.1	PPOによる行動最適化	134
6.8.2	ドメインランダム化と訓練環境のリセット	134
6.8.3	シミュレーションでの行動最適化結果の概要	135
6.9	機械獣の成立	135
6.9.1	ハードリセットに伴う温度降下の影響	137
6.9.2	温度恒常性を含まない場合の体温変化	138
6.10	内受容感覚を固定することで確認された行動特性	139
6.10.1	ナビゲーション行動の創発	139
6.10.2	モータ温度に応じた活動制御	140
6.10.3	全体的な行動戦略の可視化	141

7. 機械獣研究の展開

7.1 恒常性強化学習特有の性質に着目したネットワーク構造	144
7.1.1 内受容行動切替え機構・内受容 Mixture of Experts	145
7.1.2 ベンチマーク環境	146
7.1.3 結 果	148
7.2 恒常性を原理とした動物行動のモデリング	150
7.2.1 栄養の幾何学的アプローチと長期的な栄養摂取パターン	150
7.2.2 対応する代謝ダイナミクスの想定	153
7.2.3 GFN 解析に基づく行動解析	154
7.2.4 動因の重み付けに応じた採餌傾向の変化	155
7.3 予測制御としてのアロスタシス	157
7.3.1 最小アロスタシスモデルの構成	158
7.3.2 最小アロスタシスモデルの応答	160
7.4 恒常性に基づく動機づけが持つ多様な行動の創発能力	164
7.4.1 Homeostatic Crafter	165
7.4.2 Homeostatic Crafter での深層恒常性強化学習の概要	168
7.4.3 恒常性から創発した複雑な手続きを持つふるまい	169

8. お わ り に

8.1 本書の限界と展望	174
8.2 超長期で活動可能な完全自律認知発達ロボットに向けて	176
8.3 パーソナルロボットの多様性と機械獣	178
8.4 機械獣の幸福	179

付録 定理の証明	182
引用・参考文献	184
索 引	213

用語集

創発 コントロールされていない変数が根本にあるあらゆる現象 (Clark の定義⁷⁷⁾)。

コントロールされていない変数 (システムとその外部が持つ) 複数のパラメータの相互作用から生じているために、直接単純な操作がしにくいふるまいや性質の変化をたどっている、システム内部および外部の変数 (Clark の定義⁷⁷⁾)。

機械獣 特定の外的、明示的な目的やタスクを与えられない自律型ロボット (図 1, 図 2 参照)。

恒常性 (ホメオスタシス, homeostasis) 生体・人工を問わず、身体の内部状態に対して全体あるいはその一部の重要な要素の状態を数値化したとき、それが取り得る値域の中で狭い領域に値を保持し続ける、システムやプロセスの性質。

(システムが) **恒常性を保つ** (システムが) 恒常性の性質を有する状態を持続すること。

(ネガティブ) **フィードバック制御** 制御量に対してシステムが外界からの影響を受けることで発生した目標値との誤差に対し、現在の制御量を用いることで目標値への訂正をする制御手法。

アロスタシス (allostasis) 恒常性においてフィードバック制御の観点ではなく、実際の制定では将来の予測に基づいた応答によってフィードバック制御よりもより高効率・低誤差な制御が行われているとする考え方。

本質的変数 (essential variable) Ashby の超安定システムにおける概念で、システムにおいて適当な領域内にあるように制御される、システム内部の変数。

決定 何らかの基準・手法で複数の候補から一つを選択すること。連続値・連続ベクトルの場合は同様に、それが取り得る値域・空間から一つを選択すること。

(意思) **決定システム** 決定を行う具体的なシステムのこと。

サンプリング 確率分布からそれに従う標本を得ること。

行動 エージェント-環境の相互作用においてエージェントが決定する、自らの身体を含む環境に影響を与え得る信号のこと。行動が持続する時間はそのシステムや前提に依存する。

行動決定 決定において、特に行動を扱うもの。行動選択も同様。

運動 行動によって身体が影響を受けることで起こる、身体のおもに形状に対す

る) 物理的な変化および変化の時系列。

ふるまい エージェントを外部から観察した際に見られる、(おもに数秒以上の) 時間的に幅を持つ運動のまとまり、ないし環境との相互作用を含む運動のまとまり。

ビヘイビア (ビヘイビアモジュール) 人工エージェント・ロボットにおいて、それらエージェントが決定し、実行するふるまいの基本単位のこと。

外部環境 エージェントの外部にあり、身体をインタフェースとする、相互作用対象。

身体 運動体と内部環境からなり、それ自体にダイナミクスを持つ外部環境と強化学習エージェントのインタフェース。

内部環境 (特にエージェントの生存に関わる) 身体内部の状態とそのダイナミクス。

体内・身体の内部状態 (特にエージェントの生存に関わる) 身体内部の状態。

例: 栄養状態, エネルギー状態, バッテリー残量, 体温など。

エージェントの内部状態 エージェント内部に存在する、時間的に保持される状態のことを一般的に指す。身体の内部状態を含む概念。

例: 身体の内部状態, 短期記憶, メモリなど。

内部環境の恒常性 (特にエージェントの生存に関わる) 身体の内部状態に対して、エージェントのシステムが恒常性を保つこと。

運動体 外部環境に影響を与える身体の部分・機能。

環境 一般的な強化学習エージェントの意味での環境。身体と外部環境を含む。

強化学習エージェント 通常の強化学習理論で前提とする、環境と相互作用する行動主体。

エージェント 身体と強化学習エージェントを含む、外部環境と分離可能な行動主体。

深層恒常性強化学習 (エージェント) 強化学習エージェントの構成において、特に恒常性報酬と内受容感覚を持ち、深層恒常性強化学習の概念に基づいて構成されるもの (図 3 参照)。

ニューラルホメオスタット 深層恒常性強化学習エージェントと身体で構成され、外部環境と相互作用するエージェント。

MLP (multi-layer perceptron) 多層パーセプトロン。本書では feedforward neural network (FNN) と同義。構造に再帰的な結合を持たないニューラルネットワークのこと。

RNN (recurrent neural network) 再帰型ニューラルネットワーク。ニューラルネットワークにおいて特に再帰的な構造を含むもの。

隠れ層 MLP や RNN において入力と出力として用いられる層以外に対応する層の

こと。

MDP (Markov decision process) Markov 決定過程。強化学習理論で前提となる、エージェントと環境との相互作用ダイナミクスを表現する確率モデル。

POMDP (partially observable MDP) 部分観測 MDP。MDP における「状態」をエージェントが直接知ることができず、代わりに状態の部分的な情報を含む「観測」が与えられる相互作用ダイナミクスを考える確率モデル。

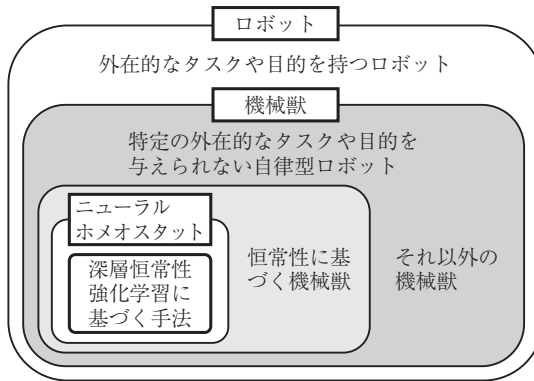
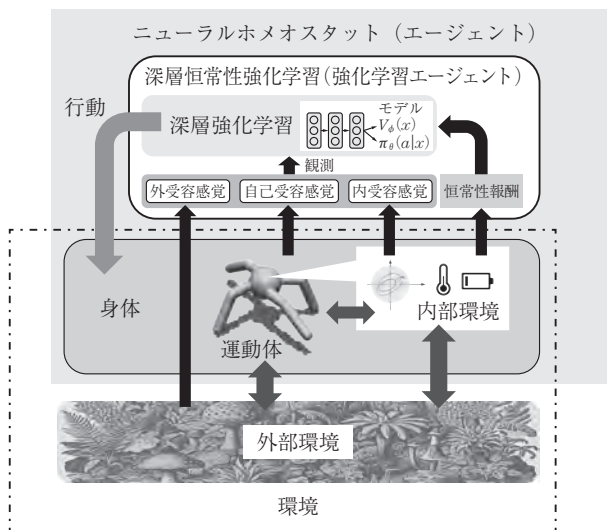


図 1 機械獣-ニューラルホメオスタット-深層恒常性強化学習の概念の関係図



図 2 実際に構成した、ニューラルホメオスタットに基づく機械獣



矢印は相互作用の方向を表す。自己受容感覚は固有感覚とも呼ばれる。本書ではエージェントを身体（外部環境とのインタフェース）と強化学習エージェントを含む行動主体として定義する。特に断らない限り、身体・内部状態を持たないことが明らかな場合は「エージェント」を通常の強化学習の意味で用いる。

図3 ニューラルホメオスタットと深層恒常性強化学習における用語の関係図

1 | はじめに

「人間の多数のしかも時間とともに変る目標は、人間の生存という大目的に向かって統一され、それが一つの価値のシステムとなっております。ここで価値というのは、客体化されたものではなく、個々の人間を動かす意志と情熱につながっている。ロボットには生命がなく、したがって、ここにいう価値も意志も情熱もありません。」（渡辺 慧, 「認識とパターン」, 1978）

私たちが普段、「生きている」と素朴に考える動物が共通に持つ性質とは何だろうか？ 自然界の動物はみな、それぞれ**身体を持っている**ことがまず共通点として考えられる。さらにもう一步踏み込んで、それが実際に物理的な存在として長期的に維持が可能となるためには、その内部の温度や栄養量、水分量などの物理的、化学的な状態が特定の範囲の値にきちんと制御されていて、その領域から大きく逸脱しないように全体的なシステムが構成されていると解釈できる（**図 1.1** 参照）。

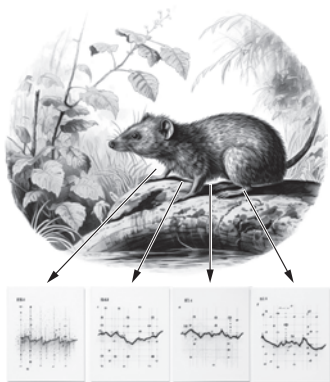


図 1.1 恒常性のイメージ図。動物の身体内部の状態を計測すると、図の下部のプロットのように血中糖度・水分量・体温などの一部重要な身体の状態が長期的に狭い範囲にとどめられており、生存が続けられる適切な状態に保たれている。

このような生体の性質は生理学の分野では**恒常性**（ホメオスタシス, homeostasis）という概念で知られている。私たちは体温が下がると震えたり、毛布を取りに行く。暑いと汗をかき、クーラーのスイッチを入れたいくなる。これらは私たちの身体温度に対する、ふるまいを通しての恒常性の保持である。私たちが直感的に「生きている」と考える動物は、何らかの形でこの恒常性を保持しており、恒常性が失われると動物はもはや自分を維持することができず、身体は崩壊し、死んでしまう。このように、われわれが日常的な意味で「生きている」と考える存在において、恒常性は最も基本的な性質である。

この恒常性のプロセスは自然物であり、本質的には意味や目的を持たない。あるのは身体の状態を適切に制御できているという事実のみであって、人間によって外的に導入された何らかの——より良い生活を送りたい、より多くの利益を得たいといった——与えられた意味や目的があってそのようになっているわけではない。これは自然物が、ハサミなどの道具や家電を始めとした何らかの存在意義があらかじめ与えられる人工物とは本質的な点で異なっていることを示唆している[†]。大学などで研究されているロボットも、多くの場合は洗濯物を畳んだり部屋を片付けたりといったロボットにとって外在的な意味や目的が人間によって与えられ、便利な道具や家電の延長として位置づけられる。

1.1 無目的なロボット

本書ではこれらの外在的な意味や目的が与えられたロボットとは対照的に、身体の前存在を前提としつつ、生命システムが持つ本質的な無意味性・無目的性に着目したロボットとその人工知能について議論を展開する。2章で詳しく紹介するように、この本質的には無意味なプロセス——生存や恒常性——を原理とした人工エージェントの研究は1950年代から脈々とシミュレーションと

[†] 動物は生殖（reproduction）を目的としているのだといわれるかもしれないが、生殖もまた本質的には無目的である。また、ラバやライガーのような不妊性の動物においても生存・恒常性は同様に議論できる。

ロボティクスの両面で行われ、一つの潮流をなしている。

本書では、このような生存や恒常性、加えて人工的な好奇心で駆動されるような「特定の外在的・明示的な目的やタスクを与えられない生活体としての自律型ロボット」を広く**機械獣** (animal machine) と呼ぶことで、一般的な外在的に目的やタスクが定められているロボットと区別し、その可能性を展開する。この「機械獣」の呼称は René Descartes の「動物機械論」にちなみ¹²⁾、Anil Seth と Manos Tsakiris の *beast machine*²⁸¹⁾ に影響を受けている。また、マジンガー Z を始めとしたポップカルチャー作品にも影響を受けた呼称である。Descartes や Seth の議論では機械・動物における哲学的、神経科学的な意識の議論が対象であったが、本書の議論は工学的な立場での自律性を持つ機械における行動生成に視点を絞り、機械獣における意識の存在やその他、哲学的問題を議論の対象とはしない。したがって、本書で論ずるところの機械獣はあくまでもわれわれの把握できる範囲において通常の自動機械を超えるものではなく、機械獣にヒトや一部の動物に認められるとされるような意識・精神的活動を認めるかについては、本書ではいかなる立場も取らないことにする。その意味で、機械獣は Descartes の動物機械論が論ずるところの動物を模したオートマタであり、本書はそれを越えた解釈を保留する。

本書で着目する恒常性で駆動される機械獣という観点では、ロボットシステム (人工知能 + 身体) の設計者は恒常性を保つことをいったん目的として学習システムを設計する一方、恒常性それ自体に外在的な目的性はなく、それが実際に恒常性をいかにして保ち、環境中でいかなる意味を持つ存在となるかについては環境に接続されたロボット側とそれを見る観察者に委ねられている。したがって、ロボットと環境を含む全体を自律的なダイナミクスとして外部から見たときに、自律ロボットのふるまい全体は設計者によって直接コントロールされていない。その意味で本書の機械獣は、Andy Clark の定義⁵⁾の意味においての**創発**的な特性からその全体的なふるまいが生じることを志向している。この原理的な理解に基づき、知的な全体的ふるまいを根こそぎ実現する方法論は、東京大学の國吉の提唱する「創発実体主義」と重なるものである。

さらに本書では機械獣の特性として、エージェントを構成する身体の内部で完結する情報処理のみで、身体内部の恒常性に対して合理的なふるまいや行動がモータ制御レベルで創発することを目指す。少し回りくどい表現だが、これは自然界の動物が生存するためのいっさいの処理が基本的には脳を含む身体の内部のメカニズムで行われているという、自律的な存在として自然に想定される状況での成立を目指すことにほかならない。別の言葉で言い換えると、これは想定するロボットの外に学習を動機づける外部システムの存在を仮定せずにすべての行動の学習が行われることを意味している[†]。

記号創発ロボティクスを提唱する京都大学の谷口によれば、このように外部に神託的な機構を想定しない発達の行動主体の性質は「認知的に閉じている」と呼ばれ¹¹⁾、本書でもこの立場、すなわち外的な機構を想定しないエージェントの構成による機械獣の構成を目指す。

1.2 恒常性に基づく自律ロボットの行動統合とその生成

このようなわかりやすい目的を与えられない、一見無意味にも見える機械獣にはどのような応用があり得るのだろうか。ここでは、パーソナルロボットにおける生活体としての統合的なふるまいの生成という観点から紹介しよう。超高齢化社会が叫ばれて久しいわが国において、独居世帯が近年ますます増加している。総務省統計局の調査によると、すでに高齢者の5人に1人が独居状態にあり、その数は年々増加傾向にあることが示唆されている⁹⁾。このような社会的状況において、ヒトとロボットが共生し、ヒトに寄り添う存在としてのペットやパートナーとなるパーソナルロボットは、今後さらに重要性を増していくことだろう。

[†] このような状況に入らない例として、例えばマニピュレータを持つ自律ロボットの手先にマーカーがつけられており、マーカーと目標位置との誤差を外部システムがモニタリングしてその誤差をロボットの運動学習の動機づけに用いるなどが考えられる。このような場合、自律ロボットの外部システム（この場合、トラッキングシステム）が学習につねに必須となってしまう、自律的な動機づけとはなっていない。

これらパーソナルロボットは単一のタスクを実行する多くの産業用ロボットとは異なり、さまざまな異なるタスクを単一のロボットで同時にこなすことが求められる。例えば、移動ロボットでは自らのオーナーとのコミュニケーションが求められつつ、自らの電源を管理し、時には自発的に充電ステーションに向かい、同時に関節を駆動するモータの温度を管理して適切な運動を決定することが必要かもしれない。このようにパーソナルロボットは必然的に、マルチタスクとその統合を志向するシステムである（**図 1.2** 参照）。

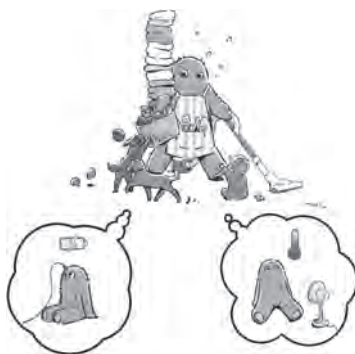


図 1.2 パーソナルロボットの置かれる状況
(イラスト：江本正弘)

このようなマルチタスク性とそのタスク統合を扱う上で、先述の恒常性に着想した自律的な行動統合システムが古くから提案されている。Blumberg らの先駆的な研究では動物行動学での知見^{195), 208)}に基づき、恒常性に着想した動機づけシステムをエージェントの行動選択に応用することで、エージェントと人間の動的なコミュニケーションの実現を試みている⁵⁴⁾。また同様に、Breazeal の Kismet⁶¹⁾や AIBO を始めとするエンターテインメントロボットにおいても同様の動機づけシステムが採用され、行動選択に用いられていることで統合的なふるまいを実現している^{25), 118)}。

これら先行研究には共通した以下の特徴がある。まず、あらかじめ個別の要求を満たすような複数の行動モジュールを想定する。それらを動機づけシステムによって実行をそれぞれ評価する。これらの評価値に応じて実行する行動モ

索引

【あ行】	完全自律的認知発達ロボット	最大事後確率推定	58
	177	最大占有原理	77
アクター・クリティックアー	観測	最短距離	151
キテクチャ	52	機械獣	3
アチーブメント	166	基礎恒常性報酬	89, 91
アドバンテージ関数	52	期待報酬	45
アロスタシス	12, 157	期待割引累積訪問数	49
安全強化学習	72	軌道	68
一次動因	12	キノコ食いロボット	16
一般化アドバンテージ推定	54	基盤モデル	176
		強化学習	22, 43
栄養の幾何学的アプローチ	150	強化学習理論	8
		近接方策最適化法	59
栄養レール	151	クエリ・キー・バリュー	145
エージェントの内部状態	70	訓練	57
エンパワーメント	77	計算論	69
オートボイエーシス	39	決定	75
オートボイエーシス論	20	決定論的方策勾配法	49
オートボイエティックマシン	39	ゲート付き再帰ユニット	168
		恒常性	2, 5, 10
オープンエンド性	166	恒常性強化学習	8, 25
重み行列	56	恒常性制御	64
		恒常性制御理論	8
【か行】		恒常性報酬	68
外受容感覚	66, 70	拘束付き MDP	72
外発的動機づけ	76	高体温	104
外部環境	14	行動	43, 75
確率的勾配降下法	58	行動価値関数	46
隠れ層	56	行動決定	7, 18
隠れマルコフモデル	25	行動選択手法	29
価値関数	46	勾配法・勾配降下法	57
活性化度	56	誤差逆伝播法	57
カートボール報酬	90		
感覚	75	【さ行】	
完全自律エージェント	16	再帰型ニューラルネット	
		ワーク	20, 168
		最大事後確率推定	58
		最大占有原理	77
		最短距離	151
		最適価値関数	47
		最適方策	47
		作動上の閉じ	41
		シグモイド関数	56
		自己教師あり強化学習	77
		自己主体感	24
		自己受容感覚	66, 70
		自己動機づけ	76
		指数加重移動分散	136
		自由エネルギー原理	23
		主成分分析	111
		状態	45, 75
		状態遷移確率	45
		自律エージェントの設計原理	30
		深層恒常性強化学習	9
		正規化エネルギー	125
		生殖	2
		生存可能領域	18
		生物規範型ロボティクス	17
		生理的変数	18
		世界モデル	176
		摂取目標量	151
		全結合 MLP	55, 145
		選好事前分布	8, 28
		層	55
		双曲線関数	56
		相互作用なし	151
		層正規化	86
		創発	3
		【た行】	
		体温調節環境	92

— 著者略歴 —

2010年 東北大学工学部機械知能・航空工学科卒業
2011年
～13年 沖縄科学技術大学院大学神経計算ユニット準研究員
2012年 奈良先端科学技術大学院大学情報科学研究科修士課程修了（情報生命科学専攻）
2016年
～20年 GROOVE X 株式会社
2022年 独立行政法人日本学術振興会特別研究員
2024年 東京大学大学院情報理工学系研究科博士課程修了（知能機械情報学専攻）
博士（情報理工学）
京都大学大学院情報学研究科特定研究員
現在に至る

自律的行動創発システムと身体性 — 機械獣の構成論 —

Autonomous Behavior Emergent Systems and Embodiment

— Construction Theory of Mechanical Beasts —

©Naoto Yoshida 2026

2026年1月16日 初版第1刷発行

検印省略

著者 吉田 尚人
発行者 株式会社 コロナ社
代表者 牛来真也
印刷所 新日本印刷株式会社
製本所 有限会社 愛千製本所

112-0011 東京都文京区千石 4-46-10

発行所 株式会社 コロナ社

CORONA PUBLISHING CO., LTD.

Tokyo Japan

振替00140-8-14844・電話(03)3941-3131(代)

ホームページ <https://www.coronasha.co.jp>

ISBN 978-4-339-03402-8 C3353 Printed in Japan

(大井)



JCOPY <出版者著作権管理機構 委託出版物>

本書の無断複製は著作権法上での例外を除き禁じられています。複製される場合は、そのつど事前に、出版者著作権管理機構(電話 03-5244-5088, FAX 03-5244-5089, e-mail: info@jcopy.or.jp)の許諾を得てください。

本書のコピー、スキャン、デジタル化等の無断複製・転載は著作権法上での例外を除き禁じられています。購入者以外の第三者による本書の電子データ化及び電子書籍化は、いかなる場合も認めていません。落丁・乱丁はお取替えいたします。