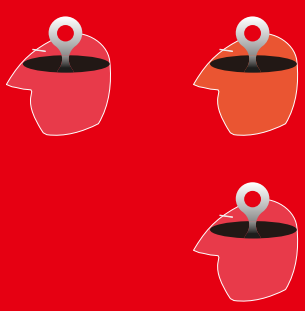
PB プレミアムブックス

強くなるロボティック・ゲームプレイヤーの作り方

実践で学ぶ強化学習

八谷大岳、杉山 将 著

機械学習の一種である「強化学習」について、数学的背景から実装、最新の動向まで幅広く解説。
「強化学習」の数学的な理論を紹介するだけでなく、ゲームやロボット制御に応用する例を、実際にプログラミングできる形で提供します。
プログラムコードと平行して理論を学ぶことで、強化学習理論の有用性を効率よく学べる構成となっています。

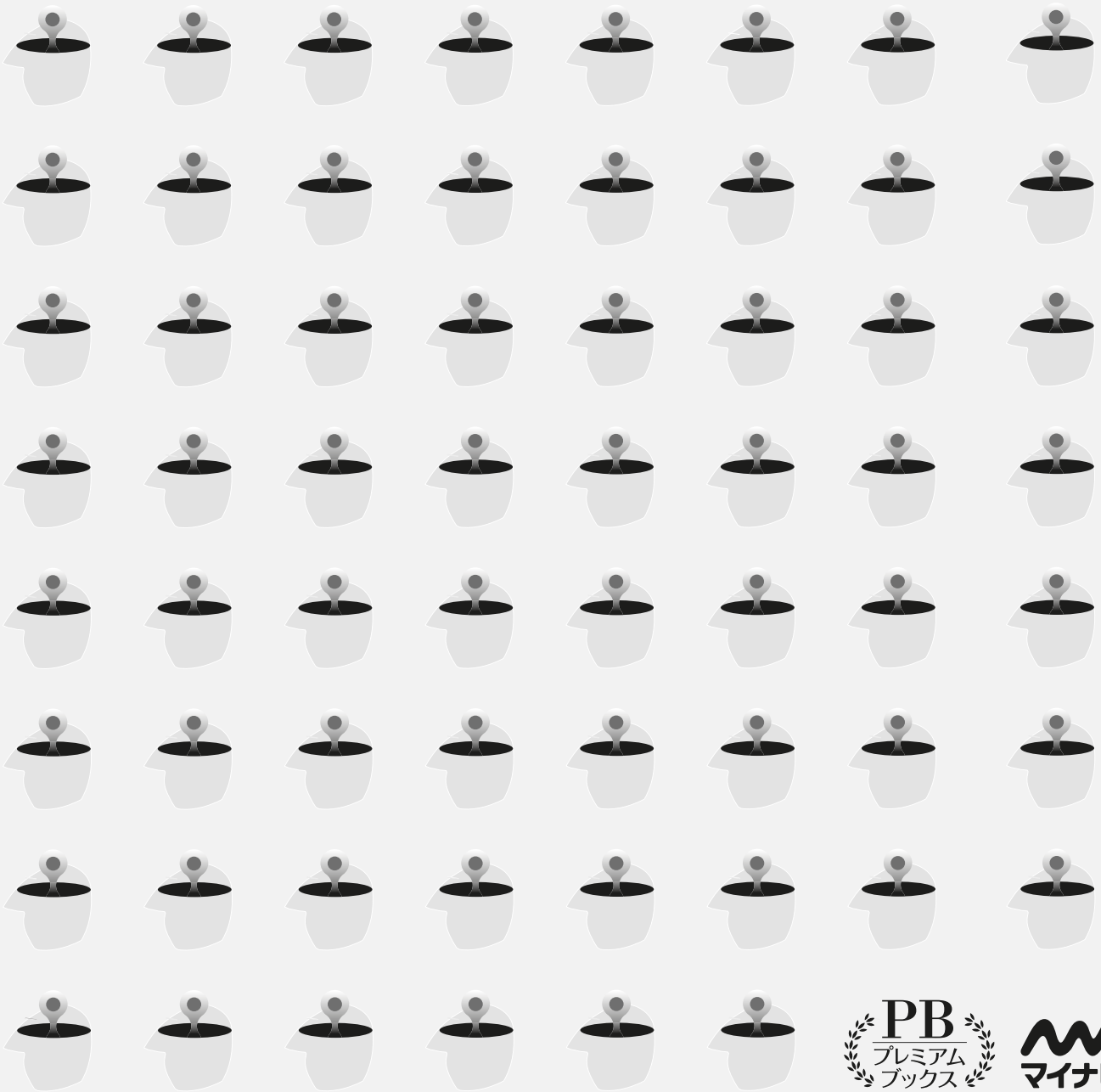


マイナビ

強くなるロボティック・ゲームプレイヤーの作り方

実践で学ぶ強化学習

八谷大岳、杉山 将 著



PB プレミアムブックス
マイナビ

●注意
本書は『強くなるロボティック・ゲームプレイヤーの作り方　～実践で学ぶ強化学習～』（2008 年 8 月弊社刊）のプレミアムブックス版です。
内容は元版から変更されておりません。
情報は、基本的に元版執筆時のものとなっております。
ご了承ください。

●サポートサイト
本書サポートサイトは以下となっています。
https://book.mynavi.jp/supportsite/detail/9784839956738.html

まえがき

階段を登ったり踊ったりするエンターテイメント用ロボットから，工場内で目にも止まらぬ速さで製品を作り上げる産業用ロボットまで，様々なロボットがテレビなどで毎日のように紹介されています．このように，日本のロボット技術は世界トップレベルであることは間違いありません．しかし未だに，人間のようなロボットが私たちの日常生活の中で活躍する日が来ないのはなぜでしょうか？ もちろん，コスト，安全性，法律の整備など社会的な要因があるのは事実ですが，何よりも，ロボットの頭脳にあたるソフトウェアの技術が十分でないことが一番の原因ではないかと著者らは考えています．

ここ数十年で，ハードウェアの技術は大きな発展を遂げてきました．実際，最近のコンピュータのハードウェアの性能は，人間をも凌駕するようになりつつあります．例えば，現在のコンピュータに含まれている演算素子の数は，人間の脳細胞の数をはるかに上回っており，計算速度や記憶容量は年々飛躍的に向上し続けています．しかし，これらの技術を統合したロボットはどうでしょうか？ 人間の代わりに食器を洗ったり掃除や洗濯をしてくれるロボットは，まだまだ実現が困難な状況です．これはロボットを制御するコンピュータのソフトウェアに，人間のような高度な認知能力や学習能力がうまく実装できていないからです．

強化学習（Reinforcement Learning）とは，私たち人間や動物が持つ学習機能の一部をコンピュータで実現することを目指すコンピュータサイエンスの研究分野の一つです．強化学習には見通しの良い統一的な理論的枠組みが存在すると共に，ゲームプレイヤーの行動戦略，ロボットの動作，マーケティング戦略をコンピュータに自動的に学習させるなど，強化学習の応用先は多岐に渡ります．このようなことから，強化学習は近年非常に活発に研究されて，すさまじい勢いで発展しています．本書では，強化学習の基礎理論から，アルゴリズム，プログラミングによる実装，さらには応用事例，最新の研究のトピックまで幅広い内容を扱います．

本書の最大の特徴は，強化学習の数学的な理論を紹介するだけでなく，強化学習をゲームやロボット制御に応用する例を，実際にプログラミングできる形で提供しているところです．数学は得意だけどプログラミングによる実装は苦手な理論派の読者の方は，まずプログラムをダウンロードして実行してみて下さい．強化学習が秘める潜在能力を十分に堪能して頂けると思います．一方，プログラミングには興味はあるけど数学的な理論はよくわからないという応用志向の強い読者の方は，プログラムコー

ドと平行して理論を学ぶことにより，強化学習理論の有用性を効率よく学んで頂けると思います．

筆者らは，機械学習（Machine Learning）とよばれる分野において研究を行なっています．機械学習の研究は，人間のような学習能力をコンピュータに持たせることが目標で，強化学習を含むより一般的なコンピュータサイエンスの研究分野の一つです．近い将来，私たちの日常生活でロボットが活躍することを夢見て，日々機械学習の基礎研究に取り組んでいます．人間のように賢いロボットを作るためには，高度な数学を自由自在に使いこなす理論的な能力と，コンピュータを自由自在に操るプログラミングの能力の両方が必要です．読者の皆さんには，本書を通して，数学的な理論の素晴らしさとプログラミングの楽しさ両方を味わって頂ければと思っています．本書が，近年発展の著しい機械学習・強化学習の分野に対する皆様の興味を高める一助になれば幸いです．

本書を執筆するにあたり，東京工業大学の杉山研究室のメンバーとの日々の議論が大変参考になりました．またソフトウェア開発に携わった方々，編集者の方々にも大変お世話になりました．この場を借りて感謝の意を述べさせて頂きます．

2008年8月

東京工業大学
八谷 大岳
杉山　 将

Contents

1章

“強くなる”ロボティック・ゲームプレイヤーを作るには

1.1 “学習機能”の必要性

コンピュータやロボットは，近年急速な発展を遂げています．工場の組み立てラインで活躍する産業用ロボットは，与えられた作業を速く，正確に，長い時間こなすことができます．また，汎用コンピュータは，高速に計算し，多くのデータを記憶し，与えられたルールに従い正確に判断をすることが出来ます．これらの能力においては，コンピュータとロボットは，人間よりも優れていると言っても過言でありません．しかし，人間が当然のように持っている機能で，ロボットなどが苦手とすることはまだあります．それは，学習を通して環境や状況の変化に柔軟に適応することです．学習というと，学校の授業などで覚えた知識を連想する人が多いかもしれませんが，普段私たちが何気なくしている動作や反応の多くも学習を通して獲得したものです．生まれたばかりの赤ちゃんができることといえば，泣くこと，母乳を吸うこと，必要最低限の反射（把握反射[脚注1.1]，モロー反射[脚注1.2]など）くらいのものです。柔軟な身体の動作，繊細な指先の動作，バランスの取れた歩行，走行，跳躍などは，私達が，生まれてから後天的に学習を通して得た能力です．もしも脳に学習能力が無ければ，どうなるかというと，いつまでたっても生まれたばかりの赤ちゃんの能力ままで，それ以上のことは何も身につけられないのです．

筆者らはコンピュータやロボットに学習をさせるという研究を行っていますが，このような人間の優れた学習能力には脱帽させられます．今のところ，ほとんどのコンピュータやロボットは，環境の変化に適応したり，より効率的な運動・動作やルール

脚注1.1 新生児の反射的行動のひとつで，手のひらに何かが触れると，意識せずに掴もうとする行動．生後2ヶ月ほどで行わなくなる．ダーウィン反射とも呼ばれる．

脚注1.2 頚が座るまでの6ヶ月ほどの間，新生児に見られる反射行動で，驚いたときに手のひらを開き，ひじを延ばして両手を挙げる動作．

を獲得するような能力を持っていません．しかし，もしもコンピュータとロボットに人間が持つような学習能力を与えることができる方法が開発されればどうなるでしょうか？　ロボットやコンピュータの持つ高い計算能力と記憶力，そして稼働力を活かせば，人間には解決できないような問題を解決したり，新しい知識や知恵を得ることができるかもしれません．また，多大な時間とコストを要するソフトウェア開発を，ある程度自動化することができるかもしれません．さて，そんな夢のような技術は実現できるのでしょうか？

本書ではロボットやコンピュータに学習機能を持たせる研究の一つである強化学習を使って，環境や状況の変化に適応できるロボットやコンピュータのための「頭脳」を作るための手法を解説します。まずは難しいことは抜きにして，現在コンピュータやロボットに学習機能を実現するためのどのような試みがあるのかを紹介しましょう．

1.2　NPC（Non-Player Character）の行動戦略の学習

NPCとは，コンピュータゲームにおいて人間が操作しないキャラクター，つまりコンピュータにより自動的に操作されるキャラクターのことです．NPCの例としては，格闘技やレース，オセロなどの対戦型ゲームにおける「敵キャラクター」や，ロールプレイングゲームにおける「サポートキャラクター」，セカンドライフを代表とする3Dインターネット仮想社会における「常駐キャラクター」などがあります．言わば，NPCはコンピュータゲームを楽しむために欠かせない存在なのです．一般に，NPCを作るためには，有限状態マシン（Finite State Machine）と呼ばれる「状態」・「行動」・「遷移」に関してキャラクターのモデルを設計し，それに基づいて行動決定ルールを

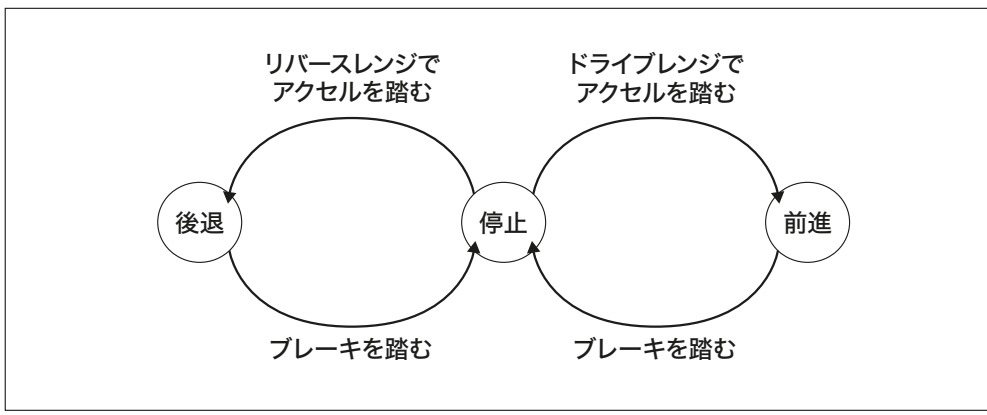


図1.1：有限状態マシン（オートマチック車の3状態マシンによる設計例）

図1.2：膨大な行動決定ルールを必要とする「Tao Feng™: First of the Lotus™（タオフェン）」

図1.3：XBoxのカーレースゲーム「PGR® 3-プロジェクト ゴッサム レーシング3-」

生成し，コーディングをします．

しかし，ゲームの規模が大きい（つまりキャラクターの数やキャラクターの状態数，行動数などが多い場合）と，有限状態マシンの設計が困難になってしまい，ルールの数も膨大になるという問題があります．たとえば，XBoxの格闘ゲーム「Tao Feng™」では，NPCが12種類あり，キャラクターの行動数は100以上あります．そして，そのNPCの行動決定に用いるコード数は，なんと約6万4,000行もあると言われています[1.1]．このような膨大な手間を解消するために，NPCに学習機能を実装し，行動戦略を自動的に学習させる研究が行われています．

マイクロソフト研究所のT. Graepelらは，強化学習法を「Tao Feng™」のキャラクターに実装し，従来のルールベースのキャラクターとの対戦を通して，人間が設計した行動決定ルールでは見られないような面白い行動パターンや，従来のキャラクターの動作の欠点を狙うような行動戦略が学習によって得られることを示しました[1.1]．

また，同研究所は，XBoxのカーレースゲーム「プロジェクト ゴッサム レーシング 3」のキャラクター（運転手）にも強化学習法を実装し，人間の操作と同程度か，もしくはそれ以上のタイムが出せるようになることを示しました[1.2]．

ほかにも、シドニー大学のK. E. Merrickらが，セカンドライフのキャラクターに強化学習法を実装し，単調な行動パターンになりがちなルールベースのキャラクターと比べ，学習により多様で複雑な行動を獲得できることを示しました[1.3]．

図1.4：インターネットの仮想社会セカンドライフ[1.3]

1.3　ロボットの動作の学習

工場の組み立てラインで作業をする産業用ロボットの多くは，「ティーチングプレイバック」や「フレキシブルロボット」と呼ばれる方法によって動作を実現しています．ティーチングプレイバックというのは，人間がロボットの動作の軌道を，座標や関節の角度によって登録（ティーチング）し，それをロボットが順番通り再生（プレイバック）する，という方法です．フレキシブルロボットは，登録された座標や関節角度をそのまま再生するだけでなく，センサーを用いて対象物の変化を調べ，その変化に合わせて動作を微調整する方法です．この２つの方法は，産業用ロボットのように，ロボットの動作環境が明確にわかっていて，かつ変化が少ない場合や，人間がロボットの動作の制御則を十分に把握し，設計することが出来る範囲においては有効です．しかし，人間の生活を支援するサービスロボットのように，環境が多様性と変化に富んでいて不確定な場合や，多自由度のロボットのように人間が制御則を設計するのが困難な場合には，利用するのは簡単ではありません．そこで，学習を通して，ロボットに自律的に動作の制御則を獲得させる研究が盛んに行われています．

マックス・プランク研究所のJ. Petersらは，7自由度のアーム型ロボットやヒューマノイド型ロボットに強化学習法を実装し，「野球のバッティング」[1.4]や到達運動(リーチング)[1.5]などの，人間が設計するのが困難な運動の制御則を，学習を通して獲得できることを示しました．

図1.5：
「Sarcos Arm」によるバッティング動作の学習[脚注1.3]

図1.6：
学習によって様々な運動を獲得している「DB」（ATR脳情報研究所）[脚注1.3]

また，スタンフォード大学のA. Y. Ngらは，ヘリコプター型ロボット[1.6]や，犬型ロボット[1.7]に強化学習法や「指導学習法（Apprenticeship Learning）」を実装し，ヘリコプターの飛行や，犬型ロボットの起伏のある道での歩行など，複雑な動作の制御則が学習によって獲得できることを示しました．

このほかにも，ソフトウェアエージェント，手書き文字認識，顔認識，音声認識，株の予測，天気予報，バイオデータマイニング，ブレインマシンインタフェースなど，様々な分野でコンピュータやロボットに学習機能を持たせる研究が行われています．近い将来，これらの研究の実用化が進み，私達の生活の中で活躍する日が来ることを期待しています．

図1.7：スタンフォード大学の自律飛行ヘリコプター

図1.8：犬型ロボット「Little Dog」

脚注1.3 Pictures are courtesy to the ATR Computational Neuroscience Laboratories, the Kawato Dynamics Brain Project funded by the Japanese Science and Technology Corporation, and the Computational Learning and Motor Control Laboratory at the University of Southern California.

2章

学習とは？

2.1　学習の定義

私達は普段何気なく「学習」という言葉を「知識を学ぶこと」,「物事を理解すること」というような意味で使っています．しかし，コンピュータやロボットで学習機能を実現するには，どのように学習が行われるかを，具体的に知る必要があります．ここでは，心理学，認知科学，脳科学などの分野における様々な学習メカニズムに関する解析を紹介します．

2.1.1　心理学における学習

心理学には「行動主義」と呼ばれる派閥があります．そこでは，学習を「経験による比較的永続的な行動変容の過程，あるいはその結果[2.1]」と定義し,「行動や反応の変化として表れ，外部から観察できる現象」として考えています．観察できる学習の代表的な例として,「パブロフ犬」で有名な「古典的条件付け」,「猫の問題箱」による「試行錯誤学習」,「スキナーの箱」で知られる「報酬学習」などがあります．

パブロフの犬［2.2］

Ivan Petrovich Pavlov（1849年–1936年）による犬を用いた学習の実験．犬に餌を与える前にベルを鳴らすことを何度か繰り返すと，犬はベルが鳴るという刺激に対して唾液を分泌をするという反応を見せるようになるというものです．これは，食べる時に“唾液を分泌する”という犬が生まれつき持つ無条件反応が，学習の過程を経て，“ベルが鳴る”という刺激に対して反応するように変化したことを意味します．このような学習は「古典的条件付け」や，刺激と反応の連合を強めることから「連合

学習」とも呼ばれます．

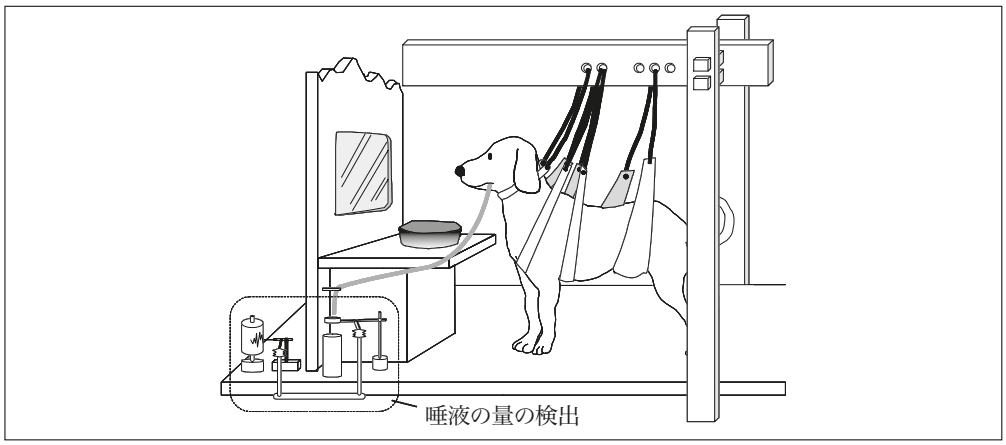


図2.1：パブロフによる犬を用いた「連合学習」の実験

猫の問題箱［2.2］

Edward Thorndike（1874年–1949年）による猫を用いた学習の実験．迷路のような箱の中で，猫は試行錯誤的に様々な反応を示し，偶然にでも外に出る行動を取り，それを何度か繰り返すと，やがて同じ行動が出現する頻度が高くなるというものです．パブロフの犬の場合と違い刺激は存在しませんが，行動を取った後の「満足」または「不快」の度合いに応じて，行動の出現頻度が学習の過程を経て変化することを意味します．このような学習は「試行錯誤学習」と呼ばれています．また，満足や不快によって行動の出現頻度が変わることを「効果の法則」といいます．

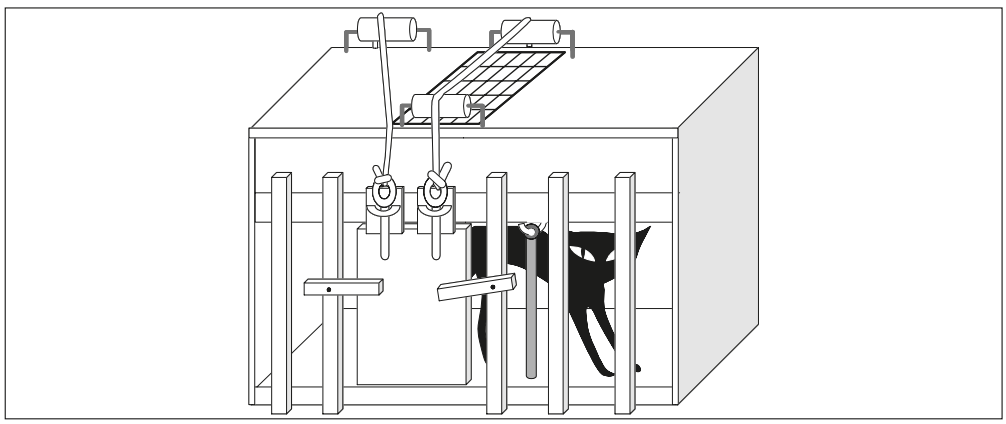


図2.2：ソーンダイクによる猫を用いた「試行錯誤学習」の実験

スキナーの箱［2.2］

Burrhus Frederic Skinner（1904年–1990年）による，ラットを用いた学習の実験．レバーを押すと餌が出る仕組みになっている箱の中で，ラットが偶然にでもレバーを押し，餌を得ることを何度か繰り返すと，ラットはレバーの近くにいることが多くなり，やがてレバーを押す行動を取る頻度が高くなるというものです．ここでも試行錯誤学習が行われていますが，「満足」，「不快」を，餌，つまり「報酬」という形で明確にしています．そして，行動を取った直後に得られる報酬に応じて，行動の自発頻度が学習の過程を経て変化することを意味します．このような学習は「オペランド条件づけ」または「報酬学習」と呼ばれます．また，状況と行動の結合を強めるような報酬を与えることを「好子」，報酬そのものを「正の強化子」，逆に状況と行動の結合を弱めるような報酬を与えることを「嫌子」，報酬そのものを「負の強化子」といいます．

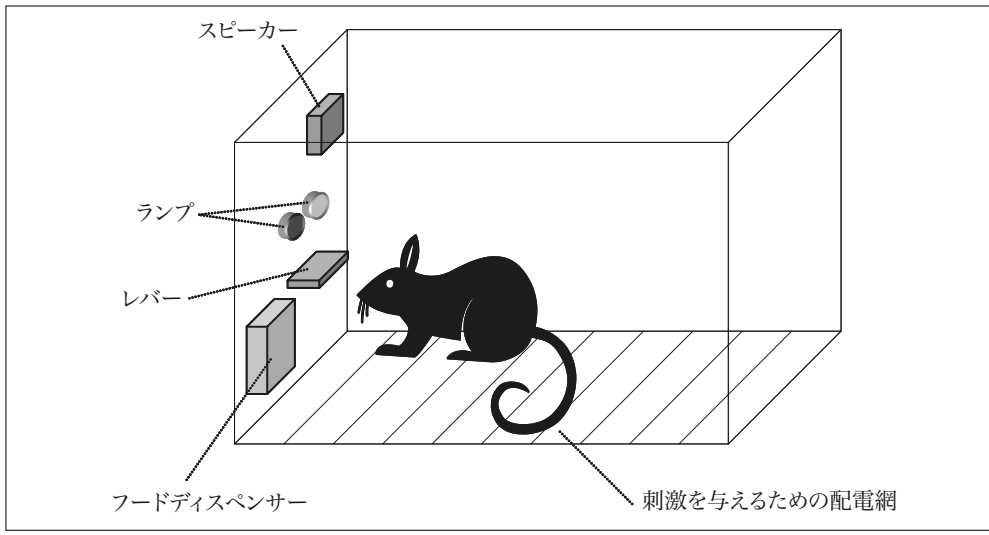


図2.3：スキナーによるラットを用いた「報酬学習」の実験

このように行動主義派の研究者は，動物がどのように行動を学習をしているのか観察を通して解析しました．私達の身近な所でも，試行錯誤学習，報酬学習を見る機会があると思います．例えば，人間の赤ちゃんが「ハイハイ」や「あんよ」などの動作を学習する姿は，まさに試行錯誤学習そのものです．また，動物のペットに餌を使って「お手」や「ふせ」などの行動を教える過程は，報酬学習に対応しています．

2.1.2　認知心理学における学習

認知心理学には，「認知主義」と呼ばれる派閥があります．そこでは，学習を次のように定義しています．

> *経験によって新しい行動傾向を獲得したり，既有の行動パターンに熟達したり，あるいはそのような行動の変化を可能にするような内的過程を獲得したり組織化，再組織化したりすること［2.1］*

学習を外部から観察できる現象と捉える行動主義と大きく異なる点は，客観的には観察できないような内面的な認知などの変化も考慮するところです．簡単な例を挙げて行動主義と認知主義の違いを考えてみましょう．

●例：Aさんが初めて世界三代珍味の一つであるキャビアを食べて美味しいと感じました．

➡〈認知主義〉：この時Aさんのキャビアに関する認知は変化しているので，Aさんは学習したと考えます．

➡〈行動主義〉：“キャビアに関する認知”の変化が行動に現れていない場合，Aさんは学習したと考えません．その後，Aさんがキャビアを食べる頻度が高くなるなど行動の変化が観察された場合に，Aさんは学習したと考えます．

さらに，学習の動機付けにおいても，行動主義と認知主義の間には違いがあります．行動主義では，「パブロフの犬」や「スキナーの箱」の例のように外部から与えられる刺激や報酬が学習を誘導するという「外発的動機付け」が考えられていますが，認知主義では，学習者の学習そのものへの興味・関心が学習を誘導するという「内発的動機付け」が考えられています．

2.1.3　脳科学における学習

行動主義では，学習を行動の変化と考えているため容易に観測ができました．しかし，認知主義では，学習を内部的な過程の変化と考えているため，実際の学習の様子を観測することは簡単には出来ません．そこで，電気生理学や解剖学に基づき脳を直

接観測することにより，学習メカニズムを分子レベルで分析する脳科学の役割が重要になります．脳科学では，学習は「脳におけるシナプスの可塑性（かそせい）により実現されている」と考えられています[2.3]．1つの神経細胞（または，ニューロン）はシナプスと呼ばれる他の神経細胞と繋がるためのコネクタを沢山持っています．そして，各シナプスは他の神経細胞のシナプスとの結合を生成，消滅，増強，抑制する可塑性と呼ばれる性質を持ちます．例えば，「記憶」は，2つの神経細胞が同時に興奮した際にシナプス結合を増強する可塑性により実現していると知られています[2.3]．また学習の動機付けは，脳内で分泌される神経伝達物質ドーパミンにより行われていると考えられています[2.3]．このドーパミンは報酬物質や快楽物質とも呼ばれ，分泌すると，私たちはやる気や満足を感じます．そのためより多くのドーパミンを分泌する行動を取るようになります．ここでもドーパミンの分泌を調整するためシナプスの可塑性が重要な役割を果たしています[2.3]．

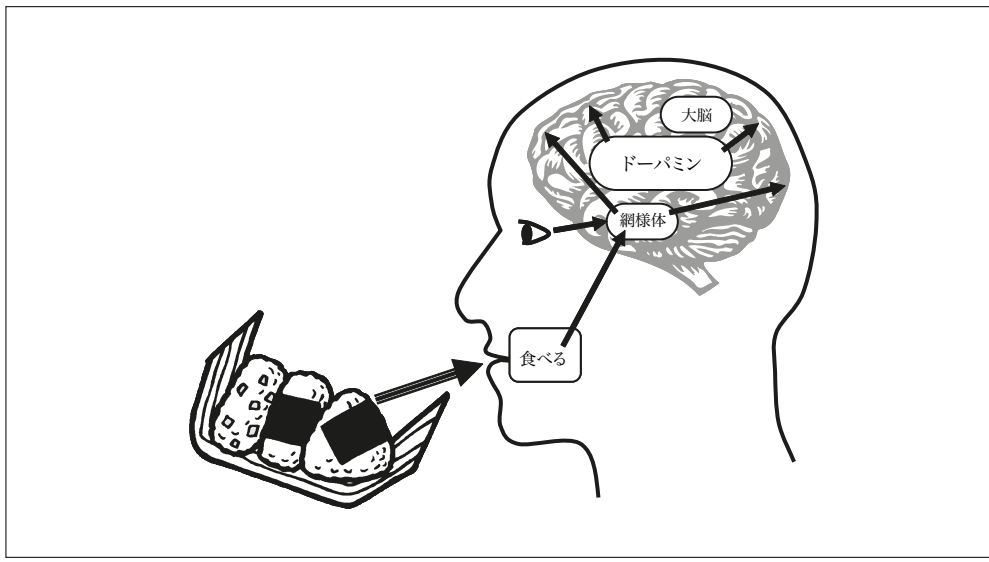


図2.4：報酬（この図では“食事”）を得ると，シナプスの可逆性によりドーパミンの分泌量が調整されると考えられている．

2.2　コンピュータの学習

人間や動物の学習の働きに関する解析が行われる中，コンピュータやロボットにて学習機能を実現するための研究も盛んに進められています．最近注目を集めている統計学を基にした「機械学習」と呼ばれる分野では，コンピュータの学習問題を大きく

次の３つに分類します．

教師付き学習問題（Supervised Learning）

入力（質問）と出力（答え）の組からなる「訓練データ」が与えられる状況で，入出力の関係（関数）を学習する問題．できるだけ少ない訓練データから，未知の入力に対しても正しい出力を予測できるような高い「汎化能力」を持つ関数を学習することが重要なポイントとなります．一般的に，教師付き学習問題は，標本（サンプル）から関数を近似する問題として定式化されます．例としては，株価予測，天気予報，フィードフォワード制御モデル（入力：センサーの値，出力：電圧，電流などの制御信号）の学習，データを分類する「分類器」の学習などがあります．

教師なし学習問題（Unsupervised Learning）

入力（質問）のみの訓練データが与えられ，文字通り教師，つまり出力（答え）の訓練データが与えられない状況で，入力データの特性を学習する問題．例としては，入力データの似たもの同士をグループ化する「クラスタリング」や，入力データの特徴を抽出し低い次元でデータを表現する「次元削減」などがあります．

強化学習問題（Reinforcement Learning）

入力（質問）と，出力（答え）に対する評価（報酬）の組からなる訓練データが与えられる状況で，政策関数と呼ばれる行動を選択する関数を学習する問題．政策関数の入出力関係を学習するという点では，教師あり学習と同じ目的を持ちますが，出力（答え）が与えられない点において異なります．例としては，1章で紹介したNPCの行動戦略の学習やロボットの動作の学習などがあります．

ここで学習したい問題がどの学習問題に分類されるかは，問題の性質または問題をどのように設計するかによって決まります．例として，アンケートの回答を分析する問題を考えてみましょう．まず，アンケートの回答から「年齢」と「収入」の関係を知りたい場合，「年齢を入力」，「収入を出力」に対応するように問題を設計すると，これは教師付き学習問題に分類されます．一方，年齢と収入から似ている回答者をグループ化したい場合，「年齢，収入を入力」に対応させたとしても，出力に対応するデー

タが無いため，これは教師なし学習問題に分類されます．さらに，年齢と収入のデータに基づき回答者に広告メールを送るための最適な戦略を選びたい場合，「年齢と収入を入力」に対応させたとしても，出力に対応する「最適な戦略」のデータは存在しません．しかし，代わりに戦略の結果として得られる売り上げを報酬に対応させることにより，強化学習の問題に分類することが出来ます．

2.3 答えを知らないロボット

さて，ロボットの動作やNPCの行動戦略を学習する問題は，どのような性質を持つのでしょうか？　また，どの学習問題に分類されるのでしょうか？　ここでは，ロボットの動作の学習に焦点を当てて考えてみたいと思います．

ロボットの動作は，一般的に制御器の制御則により実現されます．制御則とは，「制御対象の状態」に応じて「制御命令」を決定するためのルールです．制御対象の状態は，モータのエンコーダ値，障害物の距離などのセンサーの値に対応しており，制御命令は，モータの速度，トルクなどに対応しています．つまり，制御則を獲得する問題は，「制御対象の状態」と「制御命令」の関係，つまり，「制御対象の状態」を入力とし，「制御命令」を出力する関数を学習する問題として考えることが出来ます．ここでは便宜上，この関数を「制御関数」と呼ぶことにします．

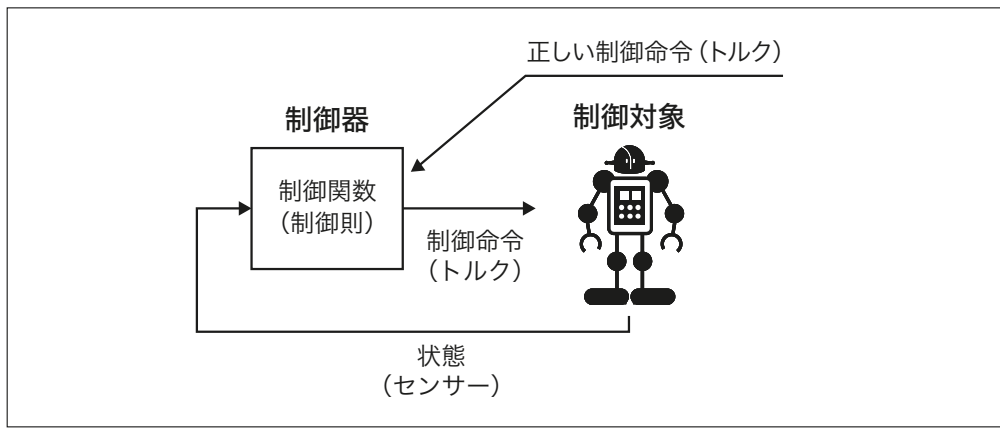


図2.5：制御器と制御対象からなるロボット制御の構成．制御器の制御関数の学習を教師付き学習問題に分類する場合は，「答え」に相当する正しい制御命令が必要となる．

ここで，制御対象の状態（センサー値など）に対する「正しい」制御命令（トルク）のデータが与えられれば，この問題は教師付き学習問題に分類することができます．

試し読みはお楽しみ
いただけましたか？

ここからはManatee
おすすめの商品を
ご紹介します。

Tech Book Zone Manatee

3.15
2017

おすすめ電子書籍

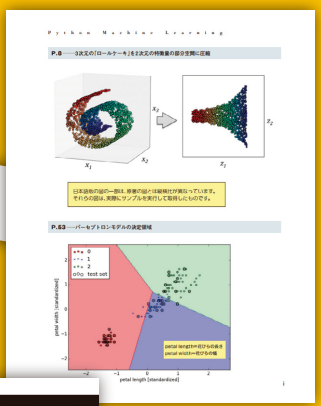
Tech Book Zone Manatee

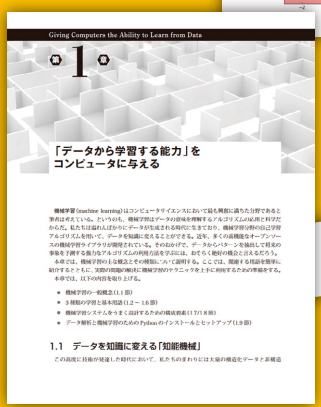
1

AI技術の基礎を習得!
分類/回帰問題や深層学習を知る

機械学習の理論とPython実践法を網羅的に解説した技術書です。機械学習とは、データから学習した結果をもとに判定や予測を行うことです。すでにさまざまな機械学習の方法が開発されています。本書では、それらの方法について背景にある理論や特徴を解説した上で、Pythonによる実装法を説明。初期の機械学習アルゴリズムから取り上げ、前処理や次元削減、Webへの展開のほか、終盤ではディープラーニングについても見ていきます。

機械学習の理論とPythonでの実践法を網羅的に解説





AIプログラミングの第一歩を踏み出すための格好の一冊

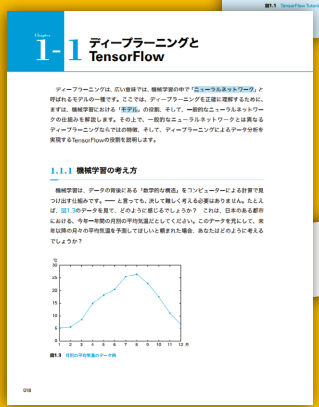
「人工知能・機械学習」

2

TensorFlowを動かしながら
ディープラーニングを理解しよう

機械学習やデータ分析が専門ではない、一般の方が対象の解説書。ディープラーニングの代表とも言える「畳み込みニューラルネットワーク」を例に、その仕組みを根本から理解すること、そしてTensorFlowを用いて実際に動作するコードを作ることが本書の目標です。ディープラーニングの根底にあるのは、古くからある機械学習の仕組みそのものです。簡単な行列計算と微分の基礎がわかっていれば、その仕組みも理解しやすいでしょう。

ディープラーニングの根本原理やTensorFlowのコードの書き方を学習できる



畳み込みニューラルネットワークを構成する1つひとつのパーツの役割を丁寧に解説

人工知能・機械学習

Python
機械学習プログラミング
達人データサイエンティストによる理論と実践

インプレス
SebastianRaschka（著者）・
株式会社クイープ（翻訳）・
福島真太朗（監訳）464 ページ
価格：4,320 円（PDF）

人工知能・機械学習

TensorFlow で学ぶ
ディープラーニング入門
畳み込みニューラルネットワーク徹底解説

マイナビ出版
中井悦司（著者） 264 ページ
価格：2,905 円（PDF）