

注目ディープ・ラーニング AIメカニズム実験室

第1章 メカニズム実験①…対話する人工知能

▶ 第2章 メカニズム実験②…勝手に成長する人工知能

第3章 メカニズム実験③…それらしい写真を自動生成する人工知能

山本 大輝, 速川 徹



図1 ブロック崩しを例にAIが勝手に成長していく強化学習を解説する

● あのAlpha Goと同様の学習方法を試せる

人工知能の活用例というと、画像を分類したり、会話を学習したり、数値の流れから未来を予測したりとさまざまですが、一般に知られた話題ならAlphaGoによる囲碁のプログラムが有名だと思います。

2016年、DeepMind社(Googleの1部門)の開発したAlphaGoは、韓国で「囲碁界の魔王」と呼ばれるイ・セドル九段と対戦し4勝1敗で勝利を収めました。そのニュースをご覧になった方も多いのではないのでしょうか。AlphaGoは、「強化学習」と呼ばれる方式でニューラル・ネットワークを学習させていました。

ここでは強化学習の一種であるDQN(Deep Q-Network)を用いて、ブロック崩しのゲームを学習させてみましょう。このプログラムはゲームの環境を与えると、どのようにボールをはじきとばせば高得点を得られるのか、勝手にゲームをして勝手に学習してくれます(図1)。

学習した後の動画をYouTubeにアップしました。URLは下記です。

<https://www.youtube.com/watch?v=D0SAFYTmIe0>

どのようなことをすればそうなるのでしょうか。順に追ってみましょう。

● 広がる世界

強化学習は教師あり学習や教師なし学習に並ぶ学習方式の1つであり、さまざまな応用ができます。

ゲームは得意分野の1つです。

- さまざまな種類のボード・ゲーム
- リアルタイムに状況が進むテレビ・ゲーム

などの学習が期待できます。また、ゲーム・プログラムのよう、何かの行動を入力するとその反応を計算してくれるシミュレータがあれば、その反応を見た学習が可能になります。例えば、

- 動物の歩行のシミュレーション
- 工場のロボット・アームの最適化

などができるでしょう。ただし入力できる行動の種類や、反応として返る画面のバリエーションが豊富なほど、学習に必要な計算資源は大きくなりますので注意が必要です。いかに学習対象を単純なものにできるかが重要と言えるでしょう。

プログラムの全体像

● 処理の概要

今回の処理の概要は次の通りです。

- ステップ1…学習元となるゲームの実行環境 OpenAI Gymを準備します。
- ステップ2…プログラミングに欠かせない知識：強化学習&Q学習を理解します。
- ステップ3…Deep Q-Networkを理解します。
- ステップ4…実際にDeep Q-Networkのネットワークを構築します。
- ステップ5…作成したModelで実際にゲームをプレイさせます。