

[説明資料(提出ファイル)] 発明・工夫作品コンテスト 製作の動機または目的, 利用方法, 作品自体やその製作過程で工夫したことを, 文章, 写真, 図などで説明。この用紙1枚に記入し, PDFに変換した後, web提出フォームにて提出する。

個人・グループ名	遠藤直弥	大学名	静岡大学
作品名	学習する倒立振り子	人数	1名

### 制作目的

本制作は, AI 技術を学ぶことができる教材として倒立振り子に着目し, 右の視点(1),(2)を重視した中学校技術科教材を開発することを目的として行った。

- (1). AI 技術による学習過程が**実感的に理解**できること
- (2). AI 技術による学習の**最適化を体験的に学習**できること

### 制作背景

1). 研究開発の俯瞰(報告書: システム情報科学技術分野(2019))

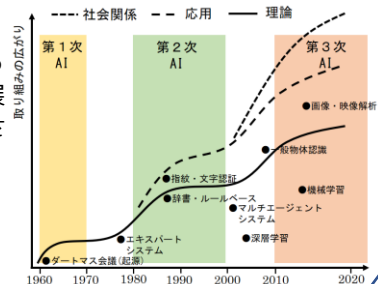
#### ○ 初等中等教育の目標

予測困難な未来社会を自立的に生き, 社会参画する資質・能力の育成  
**社会の形成者としての力をつける必要がある**

#### ○ 人工知能 (AI) の発展<sup>1)</sup>

簡単な問題しか解けなかった AI は, 計算機性能の発展により, 様々な分野で利用されるようになった。

**AIの急速な発展 (社会変化)**



### 制作の位置づけ

2). Pololu 社: <https://www.switch-science.com/catalog/3243/>

#### 人工知能と情報技術の連携した学びが必要

機械学習の中でも, 試行錯誤しながら徐々に最適な答えを探っていくアルゴリズムである強化学習を用いて, 倒立振り子に倒立維持制御を学習させるプログラムを制作することで, 機械による学習の過程を実感的に理解できる教材を開発しようと考えた。



#### 倒立振り子 × 強化学習

強化学習のプログラムを倒立振り子キットに実装することで自ら学習する倒立振り子をつくる。

### プログラム制作

#### ① 倒立振り子制御/プログラム制作環境

倒立振り子キットは Arduino 互換機として Arduino IDE でプログラム制作が可能であるが, メモリ量やライブラリーの関係から強化学習を行うのは困難である。よって, PC 側に Python をインストールし, Python による強化学習のプログラムと Arduino IDE の制御プログラムをそれぞれ制作し, シリアル通信することで学習を可能にした。

強化学習プログラム(PC)

倒立振り子プログラム



PyScripter (ver3.6)  $\longleftrightarrow$  シリアル通信  $\longleftrightarrow$  Arduino IDE (ver1.8.9)

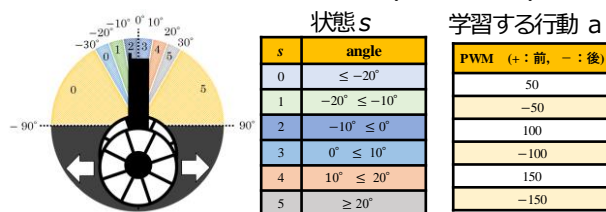
#### ② 強化学習のアルゴリズム

最初はランダムで行動を選択し, 上手くいったら報酬を受け取っていくことで状態に対する行動を学習させる。

Q学習 (報酬を最大化する学習)	探索方法 ( $\epsilon$ -Greedy法)
更新式 $Q(s_t, a_t) = Q(s_t, a_t) + a[r_t + \gamma \max_{a'} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)]$	更新式 $\epsilon = \epsilon / ep + 1$ 判別式 $\epsilon \leq \text{random}(0 \sim 1)$
$\alpha$ : 学習率 $\gamma$ : 割引率 $r$ : 報酬	$\epsilon$ : 貪欲性 $ep$ : 学習回数
状況に対する最適な行動を学習 $\rightarrow$ 徐々に最適行動を取る	

#### ③ 状態と行動の設定

状態を角度として, ある角度の時の行動 (モータ出力) を学習させるための条件を設定 (下図及び下表) した。



### 教材利用を考えたアプリケーションの制作

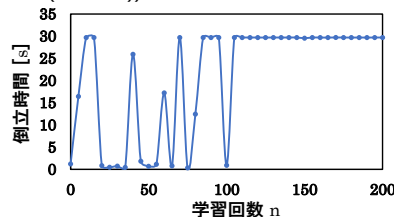
#### 学習の最適化の体験が可能なアプリケーションの開発

学習のパラメータを設定して, 学習ボタンを押すと1試行2秒(200step)で学習を行い, 学習成果ボタンで学習した結果から最適な制御を行うことができる。学習の様子は目で見て実感することもできるが, 学習回数に対する倒立時間の変化をグラフで確認することもできる。

**学習のさせ方 (学習変数)**  
を人間が決定するプログラミング

### 学習の成果

学習回数に対する倒立時間(上限 30 秒)の維持  
( $\alpha = 0.5, \gamma = 0.99, r$ : 倒れたら-200, 1step 維持できる毎に+1(最大 200))



- 学習回数に応じて倒立維持が可能になる
- 学習回数が約 100 回を超えると安定する

### まとめ

Arduino と Python とのシリアル通信により, 倒立振り子 (Balboa32U4) を制御し, 角度 (状態  $s$ ) と PWM 出力 (行動  $a$ ) をパラメータとした強化学習を可能にした。また, 計測・制御と強化学習によるプログラムの比較が可能なアプリケーションを開発し, AI 技術を体験的・実感的に理解できる教材としての可能性を示唆した。